Analysis of the *Clostridium difficile* binary toxin locus

Metcalf, D., Weese, J.S.*
Dept. of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, ON, N1G 2W1

Introduction

The function and regulation of the *Clostridium difficile* binary toxin (CDT) is currently poorly understood and the role of this toxin in pathogenesis of infection has yet to be defined. The objective of this study was to compare full binary toxin loci sequences from different *C. difficile* ribotypes and toxinotypes in an effort to further the understanding of the regulation of the binary toxin and its putative upstream regulator (CdtR). We used pyrosequencing to screen isolates for a truncating single nucleotide *cdtR* mutation and we also sought to perform sequence and phylogenetic analysis to determine concordance between ribotype and binary toxin sequence relatedness.

Materials and Methods

Isolate Selection and Culture Conditions. All *C. difficile* isolates were obtained from a collection of clinical isolates from human patients in Ontario, Canada, collected between 2004-2006 (1) except one porcine isolate collected in 2002. Ten isolates were chosen representing 6 different binary toxin positive ribotypes belonging to 9 different toxinotypes. All *C. difficile* strains were cultured on Colombia Blood Agar plates (Oxoid), and incubated anaerobically for 24 hours at 37°C. Genomic DNA was extracted using a commercial kit following the manufacturer's instructions (Instagene Matrix, Biorad, Richmond, CA). Twenty-four primers were used to sequence the binary toxin locus including the *cdtR* promoter, *cdtR*, *cdtAB* promoter, *cdtA*, and *cdtB* genes.

Binary Toxin Locus Analysis. Binary toxin locus sequence processing and assembly was performed using the Geneious bioinformatics software (Biomatters, Auckland, NZ). Signal sequences and their respective cleavage sites were predicted using the online SignalP 3.0 Server (2). Multiple sequence alignments were performed using ClustalW2 tool (3) and phylogenetic trees were created using ClustalX version 2.0. Secondary structure prediction was performed using the Advanced Protein Secondary Structure Prediction (APSSP) Server (4).

qPCR. Quantitative real-time PCR was used to assess expression of *cdtA* and *cdtR* in some *C. difficile* isolates. *C.difficile* isolates were grown to early stationary phase (OD_{600} 1.0) and RNA was isolated using a Maxwell®16, a robotic, magnetic bead-based system (Promega). Samples were treated with the Turbo DNA-free DNase system (Ambion) and subjected to a reverse transcription reaction using the Omniscript Reverse Transcription Kit (Qiagen) and gene specific primers (as described by Carter et al. (5) for *cdtA*, and for *cdtR*, c*dtR*-F 5' cgcaaaatataagataagtaatg 3' and *cdtR*-R 5'gcgataatttttttctttaatg 3') following the manufacturer's instructions. Expression was normalized using the *rpoA* gene.

Pyrosequencing. A pyrosequencing protocol was developed to identify a single nucleotide polymorphism (SNP) in the *cdtR* gene. Forward, sequencing, and biotinylated reverse primers were designed using the Pyromark Assay Design Software 2.0 (Qiagen). A 100 bp region covering the SNP was PCR amplified in 54 ribotype 078 or toxinotype V strains, 22 ribotype 027 and 6 strains belonging to other ribotypes. Pyrosequencing was performed using Pyro Gold reagents on a PyroMark ID instrument following the manufacturer's instructions (Qiagen).

Results

•Analysis of the full binary toxin loci of 10 *C. difficile* isolates yielded information on putative promoter elements of the *cdtR* and *cdtAB* genes not previously identified (Figure 1).

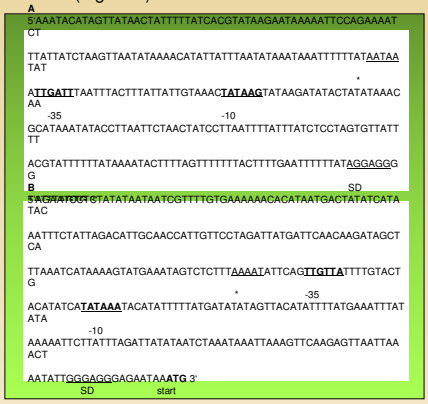

Figure 1. Model of the putative *cdtR* **(A)** and *cdtA* **(B)** promoter regions in *C. difficile*. The start codon is labelled as start and is in bold. The putative Shine-Dalgarno site (ribosomal binding site) is underlined with dashes. The putative -10 and -35 promoter elements are in bold, underlined and labelled. The underlined sequence labelled with a * indicates the putative -45 Gram positive consensus sequence commonly found upstream of the -35 promoter element.

•The ribosomal binding sites (RBS) for both promoters strongly resemble the RBS consensus sequence. Putative -10 and -35 promoter elements were also identified along with a -45 consensus sequence specific to Gram positive organisms.

•Overall sequence analysis revealed numerous polymorphisms ranging from SNPs to a nucleotide insertion in an 078 strain and the toxinotype IV strain.

•Most polymorphisms resulted in conserved amino acids and the few polymorphisms that encoded non-conserved amino acids likely wouldn't affect secondary structure as predicted by APSSP (4).

•A truncating mutation in the *cdtR* gene in a ribotype 078 strain was identified during sequencing of the total loci. This SNP reduces CdtR from 248 amino acids to 107 amino acids.

•qPCR expression analysis of *cdtA* and *cdtR* in the ribotype 078 strain with the *cdtR* SNP revealed no expression of *cdtR* but expression of *cdtA* was still detected.

Table 1. Screening of 82 *C. difficile* isolates for the presence of the *cdtR* truncating mutation identified by pyro-sequencing.

	Source	Truncating mutation	No mutation
Ribotype 078	Human	**13**	**11**
	Animal	**9**	**2**
	Other	**14**	0
ToxV/non-078	Human	0	**2**
	Animal	**3**	0
Ribotype 027	Human	0	**15**
	Animal	0	**4**
	Other	0	**3**
Other Ribotypes	Human	0	**6**
Total		39	43

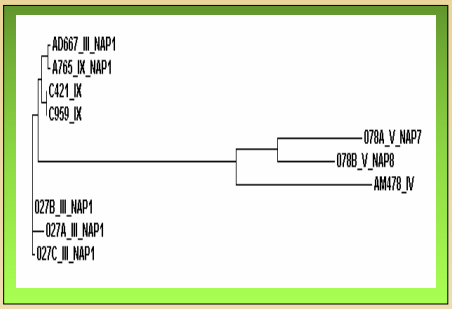

Figure 2. Phylogram generated from complete binary toxin loci sequences of the 10 *C. difficile* isolates (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Roman numerals indicate toxinotype and the North American pulsotype (NAP) is indicated if known.

- Pyrosequencing identified the *cdtR* SNP in 39/82 *C. difficile* isolates from animal, human, and other sources including food and the environment (Table 1).
- The SNP was only found in ribotype 078 or non-078 toxinotype V isolates.
- The SNP was significantly more common in *C. difficile* strains isolated in or after 2008 compared to before 2008 (P=0.0001).
- Phylogenetic analysis of the binary toxin loci revealed an overall conserved relationship between isolates of the same ribotype or toxinotype with one exception. AD667 is a toxinotype III isolate and has a binary toxin locus sequence more closely resembling toxinotype IX isolates.
- The toxinotype IV isolate appears to be more closely related to toxinotype V isolates than any other type.

Discussion

- Analysis of the promoter sequences of the *cdtR* and *cdtA* gene indicates consensus sequences for several promoter elements are present in some *C. difficile* promoters.
- Despite no expression of *cdtR* in a 078 strain and the presence of a mutation that would truncate CdtR even if it was expressed, *cdtA* expression was still detected by qPCR. This suggests CdtR may not be the only positive regulator of *cdtA* and other regulatory mechanisms are likely involved in regulating expression of the binary toxin.
- The significance of the *cdtR* SNP only being found in toxinotype V isolates remains unknown. In general, the SNP was more common in recent isolates suggesting this mutation has recently emerged.
- Bouvet and Popoff (6) found the same mutation in the *cdtR* gene and also reported the SNP to be restricted to ribotype 078 or toxinotype V strains.
- Phylogenetic analysis of the whole binary toxin loci tended to reflect the relationship indicated by ribotyping and toxinotyping.

References
1) Martin et al. 2008. J. Clin. Microbiol. 46:2999-3004. 2) SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). 3) ClustalW2 tool (http://www.ebi.ac.uk/Tools/clustalw2/index.html) 4) Advanced Protein Secondary Structure Prediction (APSSP) Server (http://imtech.res.in/raghava/apssp/). 5) Carter et al. 2007. J. Bacteriol. 189:7290-7301. 6) Bouvet and Popoff. 2008. J. Clin. Microbiol. 46:3703-3713.