Clostridioides difficile infection with isolates of cryptic clade C-II: a genomic analysis of polymerase chain reaction ribotype 151

Ducarmon, Q.R.; Bruggen, T. van der; Harmanus, C.; Sanders, I.M.J.G.; Daenen, L.G.M.; Fluit, A.C.; ... ; Smits, W.K.

Clinical Microbiology and Infection 29 (2023) 538.e1–538.e6

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Original article

Clostridioides difficile infection with isolates of cryptic clade C-II: a genomic analysis of polymerase chain reaction ribotype 151

Quinten R. Ducarmon [1, 2], Tjomme van der Bruggen [3], Céline Harmanus [1], Ingrid M.J.G. Sanders [1], Laura G.M. Daenen [4], Ad C. Fluit [3], Rolf H.A.M. Vossen [5], Susan L. Kloet [5], Ed J. Kuijper [1, 2, 6], Wiep Klaas Smits [1, 2, 7, *]

[1] *Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands*
[2] *Center for Microbiome Analyses and Therapeutics, Leiden University Medical Center, Leiden, the Netherlands*
[3] *Department of Medical Microbiology, University Medical Centre Utrecht, Utrecht, the Netherlands*
[4] *Department of Haematology, University Medical Centre Utrecht, Utrecht, the Netherlands*
[5] *Leiden Genome Technology Center, Leiden University Medical Center, Leiden, the Netherlands*
[6] *Center for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands*
[7] *Centre for Microbial Cell Biology, Leiden, the Netherlands*

ARTICLE INFO

Article history:
Received 23 August 2022
Received in revised form
28 November 2022
Accepted 1 December 2022
Available online 9 December 2022

Editor: L. Gerard

Keywords:
Clostridioides difficile
Cryptic clade
PaLoc
Polyphyletic
ribotype 151

ABSTRACT

Objectives: We report a patient case of pseudomembranous colitis associated with a monotoxin-producing *Clostridioides difficile* belonging to the very rarely diagnosed polymerase chain reaction (PCR) ribotype (RT) 151. To understand why this isolate was not identified using a routine commercial test, we performed a genomic analysis of RT151.

Methods: Illumina short-read sequencing was performed on n = 11 RT151s from various geographical regions to study their genomic characteristics and relatedness. Subsequently, we used PacBio circular consensus sequencing to determine the complete genome sequence of isolates belonging to cryptic clades C–I and C-II, which includes the patient isolate.

Results: We found that 1) RT151s are polyphyletic with isolates falling into clades 1 and cryptic clades C–I and C-II; 2) RT151 contains both nontoxigenic and toxigenic isolates and 3) RT151 C-II isolates contained monotoxin pathogenicity loci. The isolate from our patient case report contains a novel-pathogenicity loci insertion site, lacked *tcdA* and had a divergent *tcdB* sequence that might explain the failure of the diagnostic test.

Discussion: This study shows that RT151 encompasses both typical and cryptic clades and provides conclusive evidence for *C. difficile* infection due to clade C-II isolates that was hitherto lacking. Vigilance towards *C. difficile* infection as a result of cryptic clade isolates is warranted. **Quinten R. Ducarmon, Clin Microbiol Infect 2023;29:538.e1–538.e6**

Introduction

Clostridioides difficile is a gram-positive, spore-forming anaerobic bacterium and the leading cause of antibiotic-associated diarrhoea, both in healthcare facilities and the community [1]. *C. difficile* can produce several toxins (toxin A or TcdA, toxin B or TcdB and the binary toxin or CDT) and these are responsible for *C. difficile* infection (CDI), symptoms of which can range from mild and self-limiting diarrhoea to pseudomembranous colitis, toxic megacolon and ultimately death [1]. Phylogenetic analyses have revealed that most *C. difficile* isolates associated with human and animal disease fall within five clades [2]. In recent years, reports have appeared on (mainly environmental) isolates that are phylogenetically distinct from these five clades, but rather fall within at least three so-called cryptic clades (C–I, C-II and C-III) that can be considered a separate genomospecies [2–5]. Though cryptic clade isolates can contain toxin genes [2,4–6], knowledge about the occurrence and clinical significance of these isolates is limited.

* Corresponding author. Wiep Klaas Smits, Department of Medical Microbiology, Postzone E4-P, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, The Netherlands.
E-mail address: w.k.smits@lumc.nl (W.K. Smits).

https://doi.org/10.1016/j.cmi.2022.12.003

Within clades 1–5, the pathogenicity locus harbouring the *tcdA* and/or *tcdB* genes is generally located at a consistent location in the *C. difficile* genome (*cdu1*/*cdd1*) and further includes several other genes (*tcdR*, *tcdC* and *tcdE*) [6,7]. Importantly, isolates belonging to the cryptic clades C–I and C-III can have divergent pathogenicity locus (PaLoc) sequences and atypical PaLoc insertion sites that can impact diagnostic tests [2,4,6,8].

Here, we report a case of human CDI caused by an RT151 clade C-II isolate for which initial diagnostic tests using Cepheid Xpert *C. difficile* BT (XCBT) assay were negative, even though *C. difficile* could be cultured from patient faeces and clinical symptoms corresponded to a CDI case. We show that the cryptic C-II isolate cultured from this patient contains a monotoxin B PaLoc at a novel location and that other RT151 isolates fall in phylogenetically divergent clades 1, C–I and C-II.

Methods

RT151 isolates analysed as part of this study were derived from various national and international studies in which the Dutch National Expertise Center for *C. difficile* at Leiden University Medical Center (LUMC) participated and had isolates available. The isolates originated from Greece, Malta, The Netherlands and Spain between 2011 and 2021. Detailed information on all ribotype (RT) 151 isolates, including relevant accession numbers, sequence types (STs), clade and metadata is provided in Table S1

Isolates were cultured on TSS plates (Tryptic Soy Agar with 5% sheep blood; bioMérieux, The Netherlands) or CLO plates (selective *C. difficile* medium containing cefoxitin, amphotericin B and cycloserine; bioMérieux, The Netherlands). Capillary electrophoresis PCR ribotyping was performed at the LUMC, according to standard procedures [9]. Faeces samples and *C. difficile* isolates were analysed using the XCBT assay (Cepheid), which targets *tcdB*, *cdtA* and a variant *tcdC* gene associated with epidemic strains, according to the manufacturer's instructions. They were further characterized using an in-house multiplex PCR targeting the 16S rRNA gene, *gluD* and the toxin genes [10] or a PCR targeting *tcdB* alone [11]. Toxin status was additionally assessed using a cell cytotoxicity neutralization assay (CCNA) using 0.4 μM filter-sterilized culture supernatant from a 48-hour culture in brain-heart infusion broth on VeroE6 cells, using purified toxins (SML1153-2UG, Lot No SLBT4085, Sigma) and neutralizing anti-TcdA/B antitoxin (T1000, Lot No 1015235, Techlab) [12]. Supernatants from 48-hour brain-heart infusion cultures were also analysed using the VIDAS *C. difficile* Toxin A & B (bioMérieux), according to the manufacturer's instructions. For detailed methods, please see supplementary material.

Results

Case description and identification of an RT151 isolate from a CDI patient

A patient at the University Medical Center Utrecht developed diarrheal complaints after an allogeneic stem cell transplantation complicated by graft failure and prolonged neutropenia. Faeces initially tested positive for norovirus and sapovirus, but diarrhoea persisted despite negative diagnostic tests for a panel of gastrointestinal pathogens, including the XCBT assay. Macroscopic findings during colonoscopy and colon biopsies were suggestive of pseudomembranous colitis and showed no signs of graft-versus-host disease. Faecal samples were culture-positive for *C. difficile*, despite a negative XCBT assay. After the initial resolution of symptoms on vancomycin treatment, the patient experienced a relapse that spontaneously resolved. Faecal material from the relapse was again culture-positive for *C. difficile* as well as an in-house developed PCR targeting the toxin B gene [11], but negative in the XCBT assay. Both isolates were then sent to the Dutch National Expertise Centre for *C. difficile* at the LUMC (LUMCMM21 0001 and LUMCMM21 0003) and were found to be RT151 with the help of capillary PCR ribotyping (Fig. S1). A more detailed case description is provided as supplementary material. The patient provided written informed consent.

RT151 isolates have different toxin profiles

We retrieved all isolates (n = 7) typed as RT151 using either agarose gel-based ribotyping or capillary electrophoresis-based PCR ribotyping from the collection of the Dutch National Expertise Center for *C. difficile*, which encompasses >22 000 isolates collected since 2005. This shows that RT151 is a rarely diagnosed PCR ribotype (<0.035%).

Five of the isolates were nontoxigenic *C. difficile* and two were toxigenic based on a multiplex PCR. The isolates obtained from the case described above (LUMCMM21 0001/00 03) were positive for the gene encoding toxin B, but not the gene encoding toxin A (A^-B^+), whereas an isolate obtained from a patient in Malta in 2019 (LUMCMM19 2333) had a monotoxin A (A^+B^-) PaLoc; neither of the isolates contained the binary toxin locus, according to the PCR results.

We also assessed toxigenicity using a diagnostic enzyme-linked fluorescent assay (VIDAS, bioMérieux) and a VeroE6-based CCNA and found that only LUMCMM21 0001/0003 tested positive for toxin production under the conditions tested (Table S1).

Core genome phylogeny places some RT151 isolates in cryptic clade C–I and C-II

Next, we performed Illumina short-read whole genome sequencing (WGS). In addition to the seven RT151 isolates obtained from our in-house collection, the United Kingdom Anaerobic Reference Unit provided WGS data from four additional RT151 isolates. Isolates LUMCMM21 0001 and LUMCMM21 0003 were indistinguishable in a detailed analysis (data not shown).

By placing the eleven RT151 isolates into a phylogenetic tree with representative isolates from cryptic clades I–III and all "classical" clades, we assigned the RT151 isolates to their respective clades (Fig. 1). The majority (n = 8) of our RT151 isolates belonged to clade 1 and multilocus sequence type ST205 (or a highly related ST type; ST205-like) (Table S1 and Fig. 1). The n = 4 RT151 isolates obtained from the United Kingdom Anaerobic Reference Unit fell into ST205 (or ST205-like) groups and were, therefore, not further analysed. Three of the isolates in our analyses were classified into cryptic clades, with two independent isolates belonging to clade C-II (LUMCMM21 0001/0003 and LUMCMM19 2333) and one isolate to clade C–I (LUMCMM16 0013). An analysis of the average nucleotide identities (ANI) of the cryptic C-II isolates versus all other genomes shows that these isolates have ANIs of >98.7% against other C-II isolates, but that ANI values to other clades fall well below the 96% that is considered a cut-off for the same species (Fig. 1(B) and (C)) [13].

A clustering analysis based on accessory genome was consistent with the core genome phylogeny as it clearly shows clustering of the ST205(-like) isolates, with a divergent accessory genome of the cryptic clade isolates (Fig. 1(D)).

These data demonstrate that *C. difficile* ribotype 151 is polyphyletic with isolates from the same ribotype falling into two different cryptic clades as well as a "classical" clade, which has not been reported before.

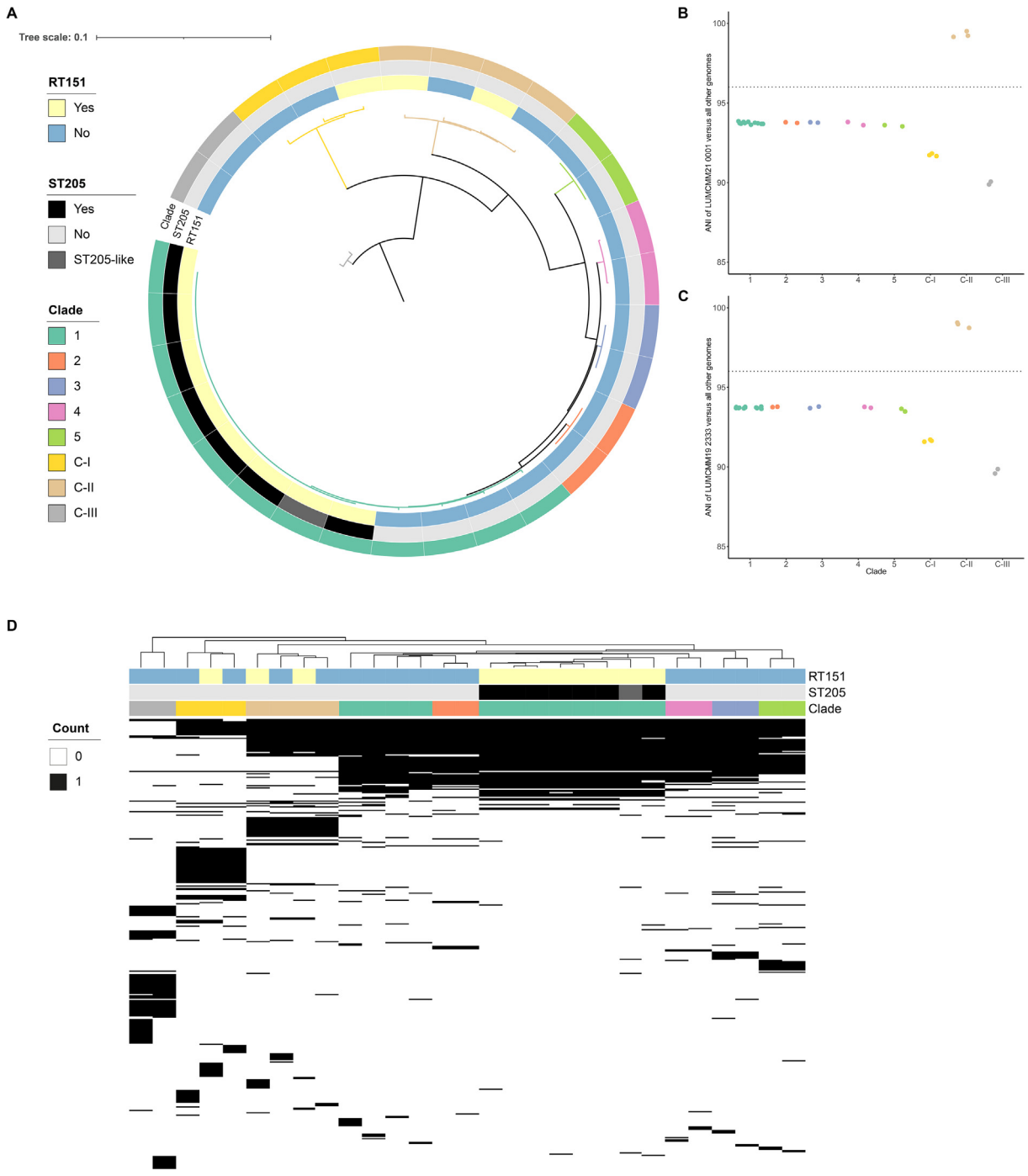

Fig. 1. RT151 isolates are part of clade 1, cryptic clade C–I and C-II and belong to various STs. **(A)** The tree was rooted at the midpoint in iTOL version 6. Detailed information, including the depicted classifications, are available in Table S1. **(B)** ANI calculations of LUMCMM21 0001/0003 versus all other included isolate genomes. **(C)** ANI calculations of LUMCMM19 2333 versus all other included isolate genomes. **(D)** Clustering using accessory genome content recapitulates core genome phylogeny. Black denotes the presence of a gene, white denotes the absence of a gene. The clustering of accessory genes was performed using the complete linkage method with Euclidian distances in the ComplexHeatmap package [24]. Detailed information, including the depicted classifications, are available in Table S1 and supplemental material. ANI, average nucleotide identities; RT, ribotype; ST, sequence types.

Cryptic clade C-II RT151 isolates can have a monotoxin pathogenicity locus at a novel insertion site

Based on our initial WGS, we were unable to recover PaLoc sequences on a single contig for all isolates and, importantly, genomic neighbourhoods around the PaLoc could not be fully resolved. In particular, for LUMCMM21 0003, *tcdB* was contained on an 18-kb contig lacking flanking chromosomal regions. Additionally, we noted limited homology to the pHSJD-312 plasmid (Fig. S2(a)), which carries a similar PaLoc [8]. However, in contrast to the PaLoc on this plasmid, the LUMCMM21 0003 PaLoc lacked the binary toxin genes *cdtA* and *cdtB*.

To identify the PaLoc insertion site and to establish whether it is carried on a plasmid, we generated complete genomes for the three

cryptic clade RT151 isolates using PacBio circular consensus sequencing (Table 1). The nontoxigenic C–I isolate (LUMCMM16 0013) was found to contain extrachromosomal elements of 132 kb and 60 kb, which are classified in a PHASTER prediction [14] as possible and putative phage, respectively. The *tcdB*-positive C-II isolate (LUMCMM21 0003) consisted of a single contig, demonstrating that the toxin gene is not carried on a plasmid. The *tcdA*-positive C-II isolate (LUMCMM19 2333) harbours a 13-kb extrachromosomal element, but this putative plasmid does not contain the toxin gene. Thus, in contrast to reports from C–I strains that carry toxins on a plasmid (as well as on the chromosome at an unresolved location) [4,8], the C-II RT151 isolate appears to carry its toxin solely on the chromosome. DiffBase [15] classifies the LUMCMM21 0003 TcdB as subtype B12, and the LUMCMM19 2333 TcdA as subtype A7.

PacBio sequencing also allows us to identify motifs associated with m4C and m6A methylation events (Table S2). Methylation has been reported to affect the virulence characteristics of *C. difficile* [16]. Notably, all strains contain the CamA-dependent modification of CAAAA**A** sequences (modified residue in bold) [16,17], indicating that the action of this methyltransferase is conserved in the RT151 C–I and C-II strains. We also observed m6A methylation on a subset of CAAA**A**AASNV motifs in the C-II strains, but not the C–I strain, suggesting it may be specific for clade C-II. Finally, we observed m4C modification on a subset of CTATTATCW motifs in the C-II strains that may be conserved in C–I strains (CWATTATCW). This motif has not previously been identified by others [16].

Next, we investigated the genomic location of the PaLocs of the RT151 isolates by aligning the observed insertion sites to those known ones from the literature (Fig. 2(A)) [4,6,18]. This revealed that the PaLoc in the Dutch C-II isolate (Fig. 2(B)) is inserted at a location that has not previously been described, between genes homologous to *cd0628* and *cd0629* of strain 630 that encode a putative membrane protein and transcriptional regulator, respectively (Fig. 2(C)).

Discussion

Here, we show that the rare PCR ribotype RT151 of *C. difficile* is polyphyletic and includes pathogenic isolates from cryptic clade C-II with monotoxin pathogenicity loci that can escape routine diagnostics. We also describe a novel insertion site for the PaLoc in cryptic clade C-II isolates.

The prevalence of RT151 reported here is in line with findings of the Leeds Teaching Hospitals NHS Trust and the University of Leeds that report a prevalence for RT151 of <0.03% with varying toxin status in a CCNA (personal communication K.A. Davies and J. Freeman). An analysis of a global collection of n = 12 098 *C. difficile* genome sequences yields a prevalence of 0.033% for ST205 [2], which constitutes the majority of RT151 isolates.

We find that RT151 strains fall into distinct phylogenetic groups (clade 1, clade C–I and clade C-II) based on MLST. Though there generally is good concordance between PCR ribotype and MLST-based approaches [19–21], this finding underscores that PCR ribotyping is insufficiently discriminatory for phylogenetic placement.

We assert that C-II strains may previously have been misclassified or not identified as such. The toxin A-negative toxin B-positive toxin type XXXII strain 173 070 [5] contains a multigene insertion interrupted by an IS256 insertion element at its *cdu1-cdd1* locus and its PaLoc has a *tcdR-tcdB-tcdE* architecture, lacks the binary toxin genes and has a divergently transcribed transcriptional regulator gene upstream from *tcdR* (presumably *cdtR*), similar to LUMCMM21 0003. By downloading raw reads (SRR1514909) and performing ANI analyses on the assembled genome, we confirmed that 173 070 is a cryptic C-II isolate (data not shown). This is in line with the SLO148 ribotype [5] being highly similar to RT151 (data not shown) and the reported ST (ST200) being classified as C-II in other studies [2,22]. Strain RA09-70 [6] has its PaLoc inserted between *spoVAE* and *cd0776*, similar to LUMCMM19 2333 (Figs. 1(a), Fig. S2). RA09-70 was classified as clade 5 [6] but the phylogenetic analysis by the authors of the article included only strains from clade 1–5 and C–I. Using the deposited genome assembly (GCF_001299495.1), we confirmed through ANI analyses that RA09-70 is, in fact, a clade C-II isolate (data not shown).

To the best of our knowledge, this study is the first to conclusively link a clade C-II isolate to human disease, although the data above indicate that clade C-II isolates could be more common than hitherto assumed. Diagnosis of CDI due to LUMCMM21 0001/0003 was hampered by the fact that faecal samples as well as *C. difficile* isolates tested negative in the XCBT assay used by many laboratories worldwide. Toxin gene sequences of cryptic clade isolates differ significantly from those of regular clade 1–5 isolates leading to false negative results [3–5]. Owing to the proprietary nature of the XCBT assay, we could not confirm that this failure is due to mismatches between the primers used in the assay and the C-II *tcdB* sequence but we have informed the company and provided the WGS data to investigate this further. We do note that positive identifications were made using the European Centre for Disease Prevention and Control-endorsed multiplex PCR (relating to little or no mismatches in the primer sequence compared to the genome).

Toxin-based assays (CCNA and VIDAS) demonstrated toxin production for the monotoxin B isolate (LUMCMM21 0001/0003), but not the monotoxin A isolate (LUMCMM19 2333). Though the latter was isolated from a hospitalized patient, lack of metadata prevented us to establish whether the patient had symptoms consistent with CDI; it remains to be established if this isolate is pathogenic. Recently, it has been reported that toxin A-positive toxin B-negative C-III strains cannot be identified on ChromID agar owing to their esculin-negative nature [23]. We found, however, that all RT151 strains, including those belonging to clades C–I and C-II, were esculin-positive (Table S1).

To date, reports of cryptic clade isolates that cause human CDI are very rare. However, our study adds to the increasing body of evidence that such isolates can cause CDI. Increased vigilance is

Table 1
Genome characteristics of the cryptic clade ribotype 151 isolates

Isolate	Clade	Size (bp)	%GC	#ORFs	GenBank
LUMCMM16 0013	C–I	4169505 (chromosome)	28.8%	3747	GCA_945909635
		132414 (possible phage)	26.3%	174	
		59572 (putative phage)	26.6%	79	
LUMCMM21 0003	C-II	4108918 (chromosome)	28.7%	3688	GCA_945861085
LUMCMM19 2333	C-II	4424808 (chromosome)	28.7%	4020	GCA_945909465
		12730 (putative plasmid)	25.9%	17	

bp, base pair; %GC, guanine-cytosine content (in per cent); ORF, open reading frame; #, number of.

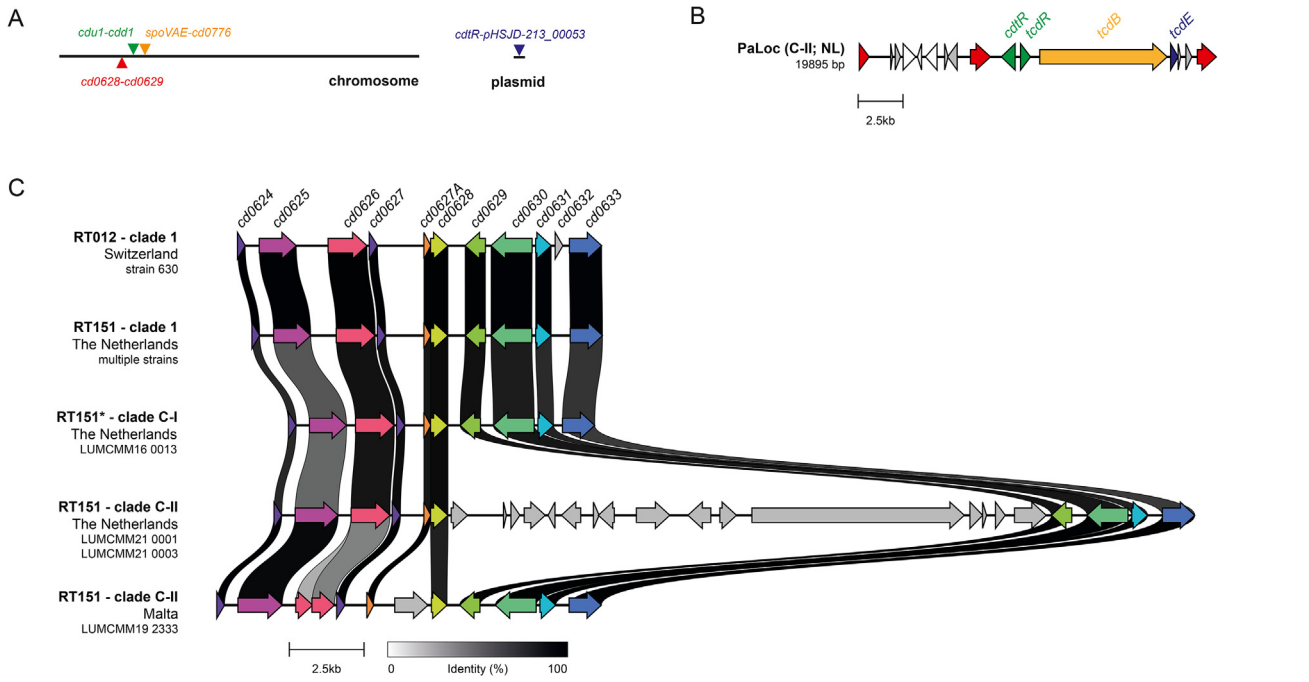

Fig. 2. Structure of the monotoxin pathogenicity locus and PaLoc insertion sites of two cryptic clade-II RT151 strains. **(A)** Overview of known PaLoc insertion sites (above the lines) [4,6] and the novel PaLoc insertion site (below the line) identified in the present study. Three sites are located on the chromosome (nomenclature from RT012 reference strain 630) [18] and one is located on a plasmid (nomenclature from clade C–I strain HSJD-312) [4,8]. **(B)** Structure of the PaLoc in the $A^{-}B^{+}$ RT151 clade C-II strain (LUMCMM21 0001/0003) showing all open reading frames predicted using Prokka [25]. Genes involved in transposition or recombination are indicated in red, hypothetical genes with homology to the PaLoc on pHSJD-312 are indicated in grey, genes encoding toxin regulators are indicated in green, toxin gene is in orange and the holin gene in purple. **(C)** Alignment of the chromosomal region incorporating the *cd0628-cd0629* insertion site for RT151 isolates in comparison with the RT012 reference strain 630 [18]. RT151 clade 1 is LUMCMM11 0004, RT151 clade C–I is LUMCMM16 0013, RT151 clade C-II is LUMCMM21 0001/0003 (The Netherlands) and LUMCMM19 2333 (Malta). Figure was generated using clinker [26]and centred on *cd0628*. Coloured arrows indicate similar genes; links are drawn between similar genes on neighbouring clusters and are shaded on the basis of sequence identity (0% white, 100% black, identity threshold for visualization 0.35). PaLoc, pathogenicity loci; RT, ribotype.

warranted towards CDI as a result of cryptic clade isolates as these isolates can escape routine diagnostics owing to divergent toxin sequences and PaLoc insertion sites.

Author contributions

Q.R.D. contributed to software, validation, formal analysis, data curation, writing the original draft and visualization; T.v.d.B. contributed to writing the original draft; C.H., I.M.J.G.S. and A.C.F. contributed to investigation; L.D. provided resources; R.H.A.M.V. contributed to investigation and formal analysis; S.L.K. provided resources and supervised; E.J.K. contributed to conceptualization and writing the original draft; W.K.S. contributed to conceptualization, formal analysis, writing the original draft, visualization and supervision. All authors reviewed and edited the manuscript and approved the final version.

Transparency declaration

E.J.K. holds an unrestricted research grant from Vedanta Biosciences. All other authors have no conflict of interest to declare. No external funding was received for this study.

Access to data

Raw sequence data is available at the European Nucleotide Archive under BioProject PRJEB52887. Associated metadata can be found in Table S1. All bioinformatic tools used for the analyses are freely available through the references provided. Complete chromosomes for PacBio-sequenced isolates can be found as accession numbers GCA_945861085 (LUMCMM21 0003), GCA_945909465 (LUMCMM19 2333) and GCA_945909635 (LUMCMM16 0013). For PaLoc alignment and insertion analyses, the corresponding regions were extracted from the complete genome sequence of the reference strain *C. difficile* 630 (AM180355) [18]. The pHSJD-312 sequence was obtained from accession MG973074 [4]. The patient provided written informed consent for the anonymized materials and results to be used for publication.

Acknowledgements

The authors thank K. Galea (L-Università ta' Malta); J. Freeman and K. Davies (University of Leeds) for sharing unpublished data; Michael Perry and Trefor Morris (Anaerobic Reference Unit, Cardiff) for providing RT151 whole genome sequence information; I. Sidorov for implementation of clinker and members of the Experimental Bacteriology Group at the LUMC for helpful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cmi.2022.12.003.

References

[1] Smits WK, Lyras D, Lacy DB, Wilcox MH, Kuijper EJ. Clostridium difficile infection. Nat Rev Dis Primers 2016;2:16020. https://doi.org/10.1038/nrdp.2016.20.

[2] Knight DR, Imwattana K, Kullin B, Guerrero-Araya E, Paredes-Sabja D, Didelot X, et al. Major genetic discontinuity and novel toxigenic species in *Clostridioides difficile* taxonomy. Elife 2021;10:e64325. https://doi.org/10.7554/eLife.64325.

[3] Williamson CHD, Stone NE, Nunnally AE, Roe CC, Vazquez AJ, Lucero SA, et al. Identification of novel, cryptic *Clostridioides species* isolates from environmental samples collected from diverse geographical locations. Microb Genom 2022;8:000742. https://doi.org/10.1099/mgen.0.000742.
[4] Ramírez-Vargas G, López-Ureña D, Badilla A, Orozco-Aguilar J, Murillo T, Rojas P, et al. Novel Clade C-I *Clostridium difficile* strains escape diagnostic tests, differ in pathogenicity potential and carry toxins on extrachromosomal elements. Sci Rep 2018;8:13951. https://doi.org/10.1038/s41598-018-32390-6.
[5] Janezic S, Marín M, Martín A, Rupnik M. A new type of toxin A-negative, toxin B-positive *Clostridium difficile* strain lacking a complete tcdA gene. J Clin Microbiol 2015;53:692–5. https://doi.org/10.1128/JCM.02211-14.
[6] Monot M, Eckert C, Lemire A, Hamiot A, Dubois T, Tessier C, et al. *Clostridium difficile*: new insights into the evolution of the pathogenicity locus. Sci Rep 2015;5:15023. https://doi.org/10.1128/JCM.02211-14.
[7] Braun V, Hundsberger T, Leukel P, Sauerborn M, von Eichel-Streiber C. Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. Gene 1996;181:29–38. https://doi.org/10.1016/S0378-1119(96)00398-8.
[8] Ramírez-Vargas G, Rodríguez C. Putative conjugative plasmids with tcdB and cdtAB Genes in *Clostridioides difficile*. Emerg Infect Dis 2020;26:2287–90. https://doi.org/10.3201/eid2609.191447.
[9] Fawley WN, Knetsch CW, MacCannell DR, Harmanus C, Du T, Mulvey MR, et al. Development and validation of an internationally-standardized, high-resolution capillary gel-based electrophoresis PCR-ribotyping protocol for *Clostridium difficile*. PLOS ONE 2015;10:e0118150. https://doi.org/10.1371/journal.pone.0118150.
[10] European Centre for Disease Prevention and Control. Laboratory procedures for diagnosis and typing of human *Clostridium difficile* infection. Stockholm: ECDC; 2018.
[11] Wolfhagen MJ, Fluit AC, Torensma R, Poppelier MJ, Verhoef J. Rapid detection of toxigenic *Clostridium difficile* in fecal samples by magnetic immuno PCR assay. J Clin Microbiol 1994;32:1629–33. https://doi.org/10.1128/jcm.32.7.1629-1633.1994.
[12] van den Berg RJ, Bruijnesteijn van Coppenraet LS, Gerritsen HJ, Endtz HP, van der Vorm ER, Kuijper EJ. Prospective multicenter evaluation of a new immunoassay and real-time PCR for rapid diagnosis of *Clostridium difficile*-associated diarrhea in hospitalized patients. J Clin Microbiol 2005;43:5338–40. https://doi.org/10.1128/JCM.43.10.5338-5340.2005.
[13] Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 2009;106:19126–31. https://doi.org/10.1073/pnas.0906412106.
[14] Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 2016;44:W16–21. https://doi.org/10.1093/nar/gkw387.
[15] Mansfield MJ, Tremblay BJ, Zeng J, Wei X, Hodgins H, Worley J, et al. Phylogenomics of 8,839 *Clostridioides difficile* genomes reveals recombination-driven evolution and diversification of toxin A and B. PLOS Pathog 2020;16:e1009181. https://doi.org/10.1371/journal.ppat.1009181.
[16] Oliveira PH, Ribis JW, Garrett EM, Trzilova D, Kim A, Sekulovic O, et al. Epigenomic characterization of *Clostridioides difficile* finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. Nat Microbiol 2020;5:166–80. https://doi.org/10.1038/s41564-019-0613-4.
[17] van Eijk E, Anvar SY, Browne HP, Leung WY, Frank J, Schmitz AM, et al. Complete genome sequence of the *Clostridium difficile* laboratory strain 630Δerm reveals differences from strain 630, including translocation of the mobile element CTn5. BMC Genomics 2015;16:31. https://doi.org/10.1186/s12864-015-1252-7.
[18] Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, et al. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet 2006;38:779–86. https://doi.org/10.1038/ng1830.
[19] Zhao H, Nickle DC, Zeng Z, Law PYT, Wilcox MH, Chen L, et al. Global landscape of *Clostridioides difficile* phylogeography, antibiotic susceptibility, and toxin polymorphisms by post-hoc whole-genome sequencing from the MODIFY I/II studies. Infect Dis Ther 2021;10:853–70. https://doi.org/10.1007/s40121-021-00426-6.
[20] Baktash A, Corver J, Harmanus C, Smits WK, Fawley W, Wilcox MH, et al. Comparison of whole-genome sequence-based methods and PCR ribotyping for subtyping of *Clostridioides difficile*. J Clin Microbiol 2022;60:e0173721. https://doi.org/10.1128/jcm.01737-21.
[21] Seth-Smith HMB, Biggel M, Roloff T, Hinic V, Bodmer T, Risch M, et al. Transition from PCR-ribotyping to whole genome sequencing based typing of *Clostridioides difficile*. Front Cell Infect Microbiol 2021;11:681518. https://doi.org/10.3389/fcimb.2021.681518.
[22] Imwattana K, Knight DR, Kullin B, Collins DA, Putsathit P, Kiratisin P, et al. *Clostridium difficile* ribotype 017 - characterization, evolution and epidemiology of the dominant strain in Asia. Emerg Microbe. Infect 2019;8:796–807. https://doi.org/10.1080/22221751.2019.1621670.
[23] Shivaperumal N, Knight DR, Imwattana K, Androga GO, Chang BJ, Riley TV. Esculin hydrolysis negative and TcdA-only producing strains of *clostridium (Clostridioides) difficile* from the environment in Western Australia. J Appl Microbiol 2022;133:1183–96. https://doi.org/10.1111/jam.15500.
[24] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847–9. https://doi.org/10.1093/bioinformatics/btw313.
[25] Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068. https://doi.org/10.1093/bioinformatics/btu153. –9.
[26] Gilchrist CLM, Chooi Y-H. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 2021;37:2473–5. https://doi.org/10.1093/bioinformatics/btab007.