Plasmids of Clostridioides difficile

Smits, W.K.; Roseboom, A.M.; Corver, J.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

Plasmids of *Clostridioides difficile*

Wiep Klaas Smits[1,2,3], Anna Maria Roseboom[1] and Jeroen Corver[1,2,3]

Plasmids are ubiquitous in the bacterial world. In many microorganisms, plasmids have been implicated in important aspects of bacterial physiology and contribute to horizontal gene transfer. In contrast, knowledge on plasmids of the enteropathogen *Clostridioides difficile* is limited, and there appears to be no phenotypic consequence to carriage of many of the identified plasmids. Emerging evidence suggests, however, that plasmids are common in *C. difficile* and may encode functions relevant to pathogenesis, such as antimicrobial resistance and toxin production. Here, we review our current knowledge about the abundance, functions and clinical relevance of plasmids in *C. difficile*.

Addresses
[1] Experimental Bacteriology Group, Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands
[2] Centre for Microbial Cell Biology, Leiden, The Netherlands
[3] Leiden University Center for Infectious Diseases (LU-CID), Leiden, The Netherlands

Corresponding author: Smits, Wiep Klaas (w.k.smits@lumc.nl)

Current Opinion in Microbiology 2022, **65**:87–94

This review comes from a themed issue on **Cell regulation**

Edited by **Àimee Shen** and **Rita Tamayo**

For complete overview of the section, please refer to the article collection, "Cell Regulation 2022"

Available online 11th November 2021

https://doi.org/10.1016/j.mib.2021.10.016

Introduction

Clostridioides difficile is a Gram-positive spore-forming enteropathogen and a major cause of infectious diarrhea in healthcare and community settings [1]. *C. difficile* infection (CDI) depends on the production of one or more toxins, which induce symptoms that range from mild diarrhea to potentially fatal colitis [2]. Different strains of *C. difficile* can carry different toxin genes, but toxin gene content by-and-large is conserved within phylogenetic groups [3]. Different methods for typing of *C. difficile* are used: whereas PCR ribotyping is currently the most common, whole genome sequencing (WGS)-based methods such as multilocus sequence typing are becoming more common [4]. In the 1980's, plasmid typing was explored but it was quickly abandoned when it was realized that not all *C. difficile* strains carry plasmids, which are defined here as self-replicating extrachromosomal elements made up of double stranded DNA [5–7]. Nevertheless, it underscores that the presence of plasmids in *C. difficile* has been known for over 30 years. In 2006, the first fully sequenced genome (of strain 630) showed the presence of a single ~7 kb plasmid in this strain, pCD630 [8]. With a worldwide increase in *C. difficile* infections from 2005 and the availability of the whole-genome sequence of strain 630, the need for advanced molecular biological tools for the organism increased. Instrumental in advancing the field was the development of shuttle vectors that employed (i) the replicon from ~7 kb plasmid pCD6 of strain *C. difficile* CD6, (ii) a ColE1 replicon to allow for replication in *Escherichia coli*, (iii) transfer requirements (*traJ* gene and the origin of transfer, *oriT*) and a (iv) selectable marker [9••,10]. For many years, pCD6 and pCD630 remained the only *C. difficile* plasmids that had been fully sequenced and/or characterized (Table 1). Recently, however, the interest in *C. difficile* plasmids has increased, with the demonstration of plasmids that can be relevant to disease development and treatment [11••,12••,13••]. Here, we review our current knowledge about the abundance, functions and clinical relevance of plasmids in *C. difficile*.

Plasmids are common in *C. difficile*

Though the presence of plasmids in *C. difficile* has been known for decades, there has been no unbiased investigation of plasmid prevalence. Studies have either been observational, had a strong sampling bias, or relied on techniques that lead to the preferential detection of specific types of plasmid. Moreover, there is no consensus method to determine plasmid content (see Box 1). Nevertheless, the overall picture emerging from these studies is that plasmids are common in *C. difficile*.

Early evidence from pulse-field gel electrophoresis (PFGE) suggests that 18–31% of *C. difficile* contained plasmids, and outbreak isolates showed similar plasmid-patterns [14,15]. Several years later, Clabots and coworkers reported even higher percentages, with 28–67% of strains carrying plasmids [5–7]. In 2018, two studies came out that determined the abundance of specific plasmid-families in *C. difficile*. Using a PCR-based approach targeting a conserved region of the modular pCD630 plasmid family, which also includes the ~8 kb pCD-ISS1 and ~12 kb pCD-WTSI subfamilies, it was shown that ~26% of strains from human and animal sources carry a plasmid from this family [16•]. Similarly, a significant portion (~5%) of human and animal strains carry cryptic 42–47 kb plasmids of the pDLL3026 family, which also includes pCDBI1 [17•]. In contrast to these

Table 1

Features and accession numbers for selected *C. difficile* plasmids

Plasmid	Size	Description	Accession number (GenBank)	References
pCD6	6830 bp	Part of the pCD-ECE6 family of plasmids; the replicon from this plasmid (*repA*/ORF B) forms the basis of many genetic tools. Copy number: 4–10.	AY350745.1	[9••,10,18••,28•]
pCD630	7881 bp	Plasmid from strain 630 and certain derivatives thereof. pCD630 is part of a modular family of plasmids that includes pCD-ISS1 and the pCD-WTSI subfamilies. Conserved region encompasses *cdp07-cdp10*, and includes a gene encoding a putative helicase.	AM180356.2	[8,16•,47]
pCD-METRO	7056 bp	Plasmid capable of raising metronidazole MIC >2 mg/L in diverse *C. difficile* strains. Copy number: 25–38. Replicon (ORF5 and flanking regions) not sufficient to confer resistance.	Contained in CAADHH010000000	[11••]
pDLL3026	46 192 bp	Likely non-conjugative cryptic plasmid, representing a family of 42–47 kb plasmids that includes pCDBI1. Contains a RepB replication protein, and a putative *parMRC* partitioning locus as well as several phage related functions.	Contained in GCA_004684655.1	[17•,20]
pHSJD-312	145 122 bp	Putative conjugative toxin (*tcdB*, CDT)-carrying plasmid harboring a T4SS and a quorum sensing system. Members have been identified in clade C-I, 2 and 4. Can be carried in strains harboring a chromosomal PaLoc as well.	MG973074.1	[13••,31]
pX18-498	31 985 bp	Plasmid associated with reduced vancomycin susceptibility; also replicates in *E. coli*. Encodes cell wall related functions. Carriage increases fitness of *C. difficile* in a vancomycin-treated mouse model of CDI.	Contained in SAMN14824866	[12••]

relatively abundant plasmids, a plasmid conferring metronidazole resistance (pCD-METRO) was only detected in <0.14% of all strains tested [11••]. Taken together, these studies suggest that different plasmid-families may show differential prevalence.

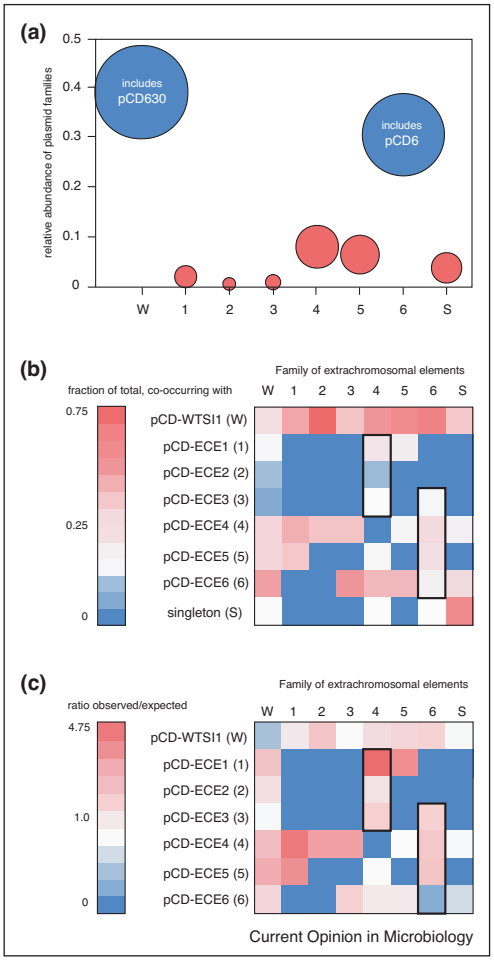
Sequence-based approaches have also been used to identify plasmids in *C. difficile*. In the largest study on plasmid prevalence and diversity to date, ~5400 short-read whole genome sequences from GenBank were analyzed; of these, 13% were predicted to harbor extrachromosomal elements that are likely plasmids [18••]. The study identified at least 6 putative families of plasmids (named pCD-ECE1 to pCD-ECE6) and suggested that pCD630- and pCD6-like plasmids account for ~50% of plasmids present in this analysis (Figure 1a). Interestingly, it was noted that plasmids are more sparse in epidemic isolates, like those belonging to PCR ribotype 027 (ST1/BI/NAP01) [18••]. Finally, long read sequencing of a collection of 419 clinical *C. difficile* isolates indicated that 36% of strains contained at least 1 plasmid [12••]. Of note, 50% of the identified plasmids were said not to fall into previously described families, but whether they are single identifications or belong to a novel group of plasmids is unclear.

Differences in the reported prevalence of plasmids may be attributable to differences in methodology, sampling bias and geographic distribution.

Box 1 Challenges in the identification of plasmids

Traditionally, plasmids are identified by gel electrophoresis. Small plasmids may be separated by regular gel electrophoresis or pulse-field gel electrophoresis; though the latter technique can reliably separate both large and small plasmids from high molecular weight chromosomal DNA and definitively establishes the extrachromosomal nature of the plasmid, it is labor intensive and technically challenging. It has been used to demonstrate the presence of plasmids in *C. difficile* total DNA [5,17•]. The extrachromosomal nature of plasmids can also be demonstrated by performing a polymerase chain reaction on total DNA. Plasmids as defined here, due to their closed circular dsDNA nature, are resistant to exonucleases such as PlasmidSafe DNase, whereas chromosomal DNA is sheared when isolated using common methodologies involving spin columns. The technique has been used to demonstrate the extrachromosomal nature of the pCD-WTSI plasmid family and pCD-METRO in recent experiments [11••,16•]. The technique is limited by the fact that larger plasmids show shearing similar to *C. difficile* chromosomes. With the advent of next generation sequencing, whole genome sequencing has become commonplace. The identification of plasmids from sequencing data is challenging. First, homology based methods like PlasMapper [41] are biased towards plasmids from Gram-negatives and perform poorly on *C. difficile* data due to limited availability of plasmids from Gram-positives in the reference database. Several graph-based computational methods have been developed [42,43]. These reference-independent methods using short-read sequence data generally perform better on higher copy number plasmids (as they are in part based on sequence coverage) and struggle to generate complete sequences for plasmids with repetitive sequences. The technique was used to identify novel plasmid families of *C. difficile* [18••]. Long read sequencing methods are expected to facilitate the identification of complex and low copy number extrachromosomal elements when size-fractionation of the chromosomal DNA and library preparation is taken into account, as routine methods may deplete smaller plasmids common to *C. difficile*. Combining short and long-read sequence data has led to a significant expansion of putative plasmids in GenBank (BioProject PRJNA524299 and Supplemental Table 1) [19]. Irrespective of method, there may be significant overlap in gene content between (pro)phages, other mobile elements and plasmids [44–46], making a definitive classification subject to experimental validation.

Figure 1



Abundance and co-occurrence of plasmid families in *C. difficile*. Data for this figure are from a previously published *in silico* analysis of next generation sequencing data [18••]. **(a)** Relative abundance of plasmid families. W = pCD-WTSI, 1–6 refers to pCD-ECE1 to pCD-ECE6, respectively. S = singletons. The most abundant families, pCD-WTSI1 and pCD6, include the plasmids pCD630 and pCD6, respectively. **(b)** Heatmap representing the relative abundance of plasmids from specific families (on the y-axis) when they co-occur with the family indicated on the x-axis. To generate the figure, data was analyzed for strains containing >1 extrachromosomal element and co-occurrence was scored. The resulting matrix was normalized to the total number of plasmids in each column. A large fraction of co-occurring plasmids due to a single family is indicated in red. For example: of the plasmids co-occurring with pCD-ECE2 (column labeled '2'), ~75% falls within the pCD-WTSI1 family (red square, top row) and the remained in the pCD-ECE4 family (light red square, 5th row). **(c)** Heatmap in which the relative contribution of each plasmid-family in (b) is corrected for the relative abundance of the particular plasmid family in the complete dataset. Because of the heterogeneity of the group that likely belong to multiple compatibility groups, singletons were omitted in this panel. In some cases (boxed regions) it becomes clear that even though that the number of plasmid co-occurrences are relative low (white or blue in panel (b)), due to the low number of plasmids in these dataset, they are in fact overrepresented (red colors in panel (c)).

C. difficile plasmids range in size from ~2 kb to possibly 300 kb

The studies above suggest that most plasmids of *C. difficile* are smaller than 50 kb, with a strong skew towards plasmids with a size <15 kb. Indeed, this is largely mirrored when the size distribution of 85 annotated *C. difficile* plasmids is plotted (GenBank nucleotide database, taxid 1496, accessed on July 15, 2021) (Figure 2 and **Table S1**).

The smallest annotated plasmid present in the database is pCD-ECE2 [18••], at 1979 bp, and contains only 3 open reading frames (ORFs), one of which is likely to be involved in replication (see further below).

Seven of the larger plasmids (>100 kb) present in the database are derived from a single study that employed long-read sequencing of *C. difficile* isolates from a pediatric inflammatory bowel disease cohort [19]. This study greatly contributed to the number of annotated plasmids, with also multiple plasmids <100 kb (**Table S1**).

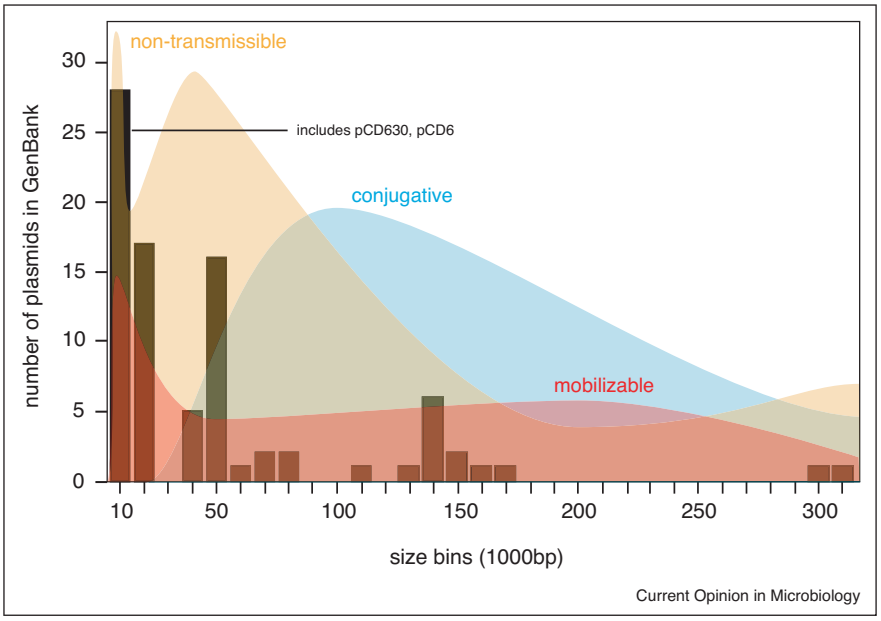
The largest element annotated as a plasmid comes from a BI1 strain and exceeds 300 kb [20]. Plasmids of this size are rare (Figure 2). It should be noted though that, although it contains the 'PLASMID' identifier, the authors explicitly indicate the possibility that this is in fact a phage. Several other plasmids, including the pDLL3026-family, also contain significant numbers of gene encoding phage-like genes that may be remnants of integrated phages [18••,21]. As a result, discriminating phage from plasmid is not trivial (see also Box 1).

It is informative to compare the size distribution of *C. difficile* plasmids with an *in silico* analysis of plasmid sizes across different phyla [22]. Though the number of plasmids for *C. difficile* is much lower, the overall size distribution is highly similar (Figure 2).

Replicons of *C. difficile* plasmids

Plasmids replicate via rolling circle replication (RCR), theta replication or strand displacement. These replication modes have been extensively reviewed here [23–25]. For clarity, we summarize a few key aspects. RCR is unidirectional, asymmetrical (i.e. leading and lagging strand synthesis are uncoupled) and is characterized by the fact that the newly synthesized strand is covalently attached to the parental strand during replication [25]. Theta replication derives its name from the shape of the replication intermediate, involves melting of dsDNA followed by coupled leading and lagging strand synthesis and can be either unidirectional or bidirectional [25]. Strand displacement replication requires three plasmid-encoded proteins for initiation and proceeds bidirectionally [25]. Though these mechanisms have not been experimentally addressed in *C. difficile*, evidence suggest that all mechanisms may occur. For instance, plasmid families pCD-ECE2 and pCD-ECE3 are believed to

Figure 2



Size distribution of plasmids of *C. difficile*.
All GenBank entries annotated with '*C. difficile*' and 'plasmid' with filter Genetic compartment 'Plasmid' ($n = 164$ as per July 8, 2021) were identified, redundant entries were merged on the basis of their descriptor and/or size (in base pairs) and non-*C. difficile* species were removed (**Table S1**). The resulting list of 85 plasmids was binned into 10 000 bp bins of which the upper limit is indicated in the graph. Size distributions for non-transmissible, conjugative and mobilizable plasmids (derived from Ref. [22]) were overlaid. Well-known *C. difficile* plasmids pCD630 and pCD6 fall into the first bin.

replicate via RCR, based on homology of encoded proteins to RCR proteins from other Firmicutes such as *Staphylococcus*, *Geobacillus* and *Streptococcus* [18••,24]. pCD-ECE5 might employ a theta-type mechanism, similar to the *E. coli* plasmid R6K, and pCD-ECE6 is expected to replicate via an IncQ/strand displacement mechanism, similar to the broad-host range plasmid RSF1010 [18••,23].

Despite these plausible mechanisms, little is known about replicon function in *C. difficile*. There is no structural information or characterization of mechanisms that determine replication initiation (including initiator proteins and origins), control copy number or determine the stability of plasmids carrying the replicon [26,27].

The best characterized replicon of *C. difficile* is the one from plasmid pCD6 that is an integral part of a suite of shuttle plasmids used for the manipulation of the species [9••,10]. The 6.8-kb pCD6 plasmid contains a 545-amino acid replication initiation (Rep) protein, RepA, encoded by ORF A. This protein shares similarity with RepA of the *Clostridium perfringens* plasmid pIP404, but is significantly larger [9••]. Approximately 200 bp downstream of the *repA* ORF is a region with 35-bp repeats (so called iterons). A region encompassing *repA* and the iterons not only sustains plasmid replication in *C. difficile*, but also in *Clostridium beijerinckii*. A second open reading frame that is generally present in the pCD6-derived region present in shuttle plasmids, ORF B, is not essential for replication [9••,10]. Plasmids containing the pCD6 replicon are efficiently maintained, with 92% of colonies retaining a resistance associated with plasmid carriage after >32 generations in strain CD3 [9••]. The copy number of plasmids with a pCD6 replicon has been estimated at 4–10 copies per chromosome equivalent, on the basis of a quantitative PCR results [11••,28•]. In contrast, the replicon from pCD-METRO sustains a copy number of 25–38 per chromosome equivalent. The 2-kb region of this plasmid that allows for replication in *C. difficile* encompasses the gene encoding the 465-amino acid putative replication factor, ORF5 (RepA), and its upstream and downstream regions. In other replicons, Rep-adjacent regions often contain the plasmid origin of replication [23]. It is expected that investigation of uncharacterized plasmid families will reveal novel replication functions, as has been the case for the *C. perfringens* pCW3 plasmid, which encodes a Rep protein without clear homology to other Rep proteins [29].

Well-characterized plasmids such as pCD630, pCD6 and pCD-METRO, appear to be efficiently maintained in the

absence of selection, consistent with their small size (<10 kb) and high copy number [9••,11••,16•,22]. Larger plasmids, which may have copy numbers close to 1 chromosome equivalent, rely on dedicated functions for stable maintenance [17•,26]. Such plasmids frequently encode post-segregational killing systems and/or partitioning systems, for example, which ensure that each daughter cell inherits at least one copy of the plasmid and is therefore not lost from the population [30]. Indeed, the 145-kb pHSJD-312 plasmid appears to encode a partitioning system [31]. Also the pDLL3026 family of plasmids appears to encode a partitioning system, homologous to the type II partitioning locus of R1 [17•]. ORF5 encodes a putative ParM partitioning protein and ORF6 encodes a protein that matches the characteristics of ParR proteins from pCW3-like plasmids of *C. perfringens*. Finally, the upstream of ORF5 contains 5 direct repeats that might constitute a centromere like *parC* locus, that acts as a recognition site for the ParR protein. Of note, no similar *parMRC* systems were identified in closely related Clostridia, suggesting that aspects of plasmid maintenance systems may be unique to *C. difficile*.

Some *C. difficile* plasmids can co-occur in the same cell

Plasmid incompatibility refers to the inability of two coresident plasmids to be stably inherited by daughter cells in the absence of specific selection and is generally dependent on features of the replicon, such as the origin and/or partitioning system [32]. For *C. perfringens*, there is limited evidence for plasmid incompatibility [33], but at the same time there is also evidence that the same or highly related replicons can coexist [34]. For *C. difficile*, a similar situation may apply, where some replicons may be compatible whereas others may not be.

PFGE experiments suggest that up to 6 plasmids can co-exist in the same strain [7,14,15] and similar numbers are reported on the basis of sequence analyses [18••]. Indeed, in laboratory experiments pCD6-replicon and pCD630-replicon plasmids can co-exist in the same cell [9••,16•]. Similarly, it was observed that pDLL3026-family plasmids are compatible with pCD6-family plasmids [17•]. Nevertheless, transfer efficiency of a pCD6-replicon plasmid into strain CD6 is lower than to a CD6-derivative cured of the native pCD6 plasmid, suggesting that the presence of pCD6 negatively effects transfer or maintenance of a vector with a similar replicon [9••].

On the basis of an *in silico* analysis, it can be predicted which family plasmids may be compatible (Figure 1b and c; see legend for details). For instance, pCD-WTSI family plasmids (including pCD630) co-occur with all other plasmid families investigated (Figure 1b). When corrected for plasmid prevalence, it is also evident that some plasmid families show clear incompatibility (Figure 1c). For instance, pCD-ECE4 co-occurs with most other families, but not with other pCD-ECE4 family plasmids; and pCD-ECE1 does not co-occur with plasmids from the pCD-ECE1, −2 or −3 families. Compatibility of these and other plasmid families awaits experimental confirmation.

C. difficile plasmids are involved in antimicrobial resistance

Many *C. difficile* plasmids are cryptic, that is, there are no phenotypic consequences to plasmid carriage [16•,17•], but in other Gram-positive pathogens plasmids are frequently associated with antimicrobial resistance (AMR) [34,35]. Well known examples include the aminoglycoside resistance plasmid pUB110 of *Staphylococcus aureus* and plasmid pCW3, that confers tetracycline resistance in *C. perfringens* [34,35]. Recent evidence suggests that plasmids are also involved in AMR in *C. difficile*. Sequence analysis has revealed the presence of putative AMR determinants such as 23S rRNA methyltransferase *cfrC* (conferring resistance to phenicols and linezolid), the 3′-aminoglycoside phosphotransferase *aphA3*, the aminoglycoside-2″-adenylyltransferase *aad(2″)*, and genes encoding a transporter from the small multidrug resistance (SMR) family on plasmids, but their functional relevance has not been demonstrated [18••,36]. In contrast, a clear role was demonstrated for *C. difficile* plasmids in tolerance and resistance to two clinically relevant antimicrobials: the nitroimidazole prodrug metronidazole and the cell-wall synthesis inhibitor vancomycin [11••,12••].

The 7-kb plasmid pCD-METRO was identified as the main difference between longitudinally collected strains from a single patient that showed a change in metronidazole resistance [11••]. When introduced in a laboratory strain, the plasmid leads to stable metronidazole resistance (minimal inhibitory concentration (MIC) >2 mg/L). It has been shown that detection of the plasmid by PCR can reliably identify metronidazole-resistant strains of diverse ribotypes [11••]. The mechanism by which the plasmid confers resistance is unknown, but it requires other regions than the replicon, as a shuttle plasmid containing just the pCD-METRO replicon does not confer resistance [11••]. It has been reported that carriage of pCD-METRO in PCR ribotype 010 (ST15) strains is associated with a single nucleotide polymorphism in the heme responsive gene *hsmA* [37].

In another study, the authors showed that plasmids are more common in *C. difficile* strains isolated from patients that failed vancomycin treatment for CDI [12••]. A 32-kb plasmid, pX18-498, was isolated and shown to raise the MIC for vancomycin 4−8-fold through a mechanism that may involve an *N*-acetylmuramoyl-L-alanine amidase. Though the MICs remained below the cut-off for resistance (<2 mg/L), it was shown that a strain carrying the plasmid showed increased virulence in a preclinical vancomycin-treated mouse model compared to an isogenic

plasmid-less comparator, as judged by for instance colonization, toxin production and inflammation of host tissue [12••].

Together, these studies provide compelling evidence for a role of plasmids in *C. difficile* AMR.

C. difficile plasmids can contain toxin genes

Several Gram-positive bacteria contain virulence plasmids, that are conjugative and carry toxin genes [34]. In *C. difficile*, the genes encoding the large clostridial toxins (TcdA and TcdB) and binary toxin (CDT) are generally carried on the chromosome [1,2,38]. Nevertheless, in 2018, sequence analysis suggested the existence of toxin-carrying plasmids, such as the 145-kb plasmid pHSJD-312, in the C-I clade of *C. difficile* strains [31]. The C-I clade (that includes ST181) predates the emergence of the typical healthcare-associated infection clades of *C. difficile* by millions of years, and is noted for divergent toxin gene architecture and unique gene clusters [39]. Plasmid pHSJD-312 carries a *tcdB*+ monotoxin pathogenicity locus and a complete binary toxin locus. Of note, the TcdB protein encoded by this plasmid clusters with variant TcdB sequences from clade 2 (including the epidemic PCR ribotype 027 strain R20291), but shows substantial differences in its glycosyltransferase domain [31]. Later, the finding of putative plasmid-located toxin genes was extended to strains of clade 2 and 4 [13••]. The presence of *tcdB* on a putatively transmissible plasmid does not preclude chromosomal carriage of the large clostridial toxins and may have significant implications, as the plasmid-encoded toxins escape routine diagnostic procedures [31].

The large size of these plasmids, and the fact that they carry additional putative virulence factors as well as putative conjugation genes including a type IV secretion system (T4SS; see further below), suggests that they should be considered legitimate virulence plasmids [26,27,34].

Transfer of *C. difficile* plasmids

Plasmids can be conjugative, mobilizable or non-mobile; the majority of plasmids in Firmicutes are believed to be non-mobile (Figure 2) [22,40]. Plasmid transfer allows cells to acquire new genetic information, which may provide a fitness advantage. Multiple lines of evidence suggest that transfer of plasmids to and from *C. difficile* can occur.

Most conjugative plasmids of Gram-positive bacteria are characterized by the presence of a T4SS, that forms the channel through which the DNA is transferred from donor to recipient cell [22,40]. To date, the only *C. difficile* plasmids postulated to be conjugative are the ~145-kb plasmids carrying *tcdB* and binary toxin [13••,31], as these contain a T4SS and several other proteins associated with conjugation, but this awaits experimental verification.

Mobilizable plasmids lack a T4SS, but most encode other functions related to mobilization, such as relaxases. The only cis-acting element required for mobilization is an origin of transfer (*oriT*) as long as the other functions are provided in trans by a compatible mobile genetic element [24]. For *C. difficile*, mobilization genes (including relaxases) have been identified on some plasmids (e.g. pCD-METRO and pCD-SMR) [11••,18••], but not others (e.g. pDLL3026) [17•].

The majority of plasmids from Firmicutes is believed to be non-mobilizable [22], and this may hold true for *C. difficile* where many plasmids lack obvious transfer-related genes. Yet, high sequence conservation has been observed between plasmids isolated from phylogenetically distinct strains and acquisition of plasmids has been documented [11••,17•]. This suggests that horizontal gene transfer might either occur via an alternative mechanism (e.g. competence) or through an uncharacterized conjugation/transfer system.

Notably, the host range and natural reservoirs of *C. difficile* plasmids are unknown. Considering their prevalence among *C. difficile* strains but paucity outside this species, it is likely that pCD6-like and pCD630-like plasmids are native to *C. difficile*. Some plasmids may originate from other Firmicutes, based on homology of the encoded proteins [18••]. It is interesting that 7 out of 25 plasmids isolated from strains of vancomycin non-responders were capable of replicating in the Gram-negative *E. coli* [12••]. Though this does not mean that this organism is a natural host for plasmids such as pX18-498, it clearly shows that some of the plasmids are broad-host range and reservoirs might include gut-resident Enterobacteriaceae. Plasmids that do not appear to be part of a larger family, so called singletons [18••] or uniquely identified plasmids [12••], may represent 'accidental' transfers. Rare transfer events may be instigated by, for instance, treatment with antimicrobials; acquisition of pCD-METRO and pX18-498 has been linked to metronidazole and vancomycin treatment, respectively [11••].

Overall, the observation of both broad and narrow host-range plasmids in *C. difficile* is consistent with observations in other Firmicutes [35].

Conclusions

Current evidence suggests that plasmids are common in *C. difficile* and may encode functions relevant to pathogenesis and treatment. Nevertheless, plasmids are understudied and frequently not annotated in whole genome sequences, partially due to lack of relevant reference sequences. The increased implementation of long-read sequencing technologies is expected to enhance the

identification of extrachromosomal elements, including plasmids, and will facilitate the functional characterization of plasmids of *C. difficile*. In particular, characterization of the mechanisms of replication and a systematic investigation of the function of plasmid-located genes using *in vivo* and *in vitro* approaches is necessary. In order to understand the role of *C. difficile* plasmids in horizontal gene transfer, interspecies and intraspecies transfer of plasmids will need to be investigated.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors thank Bastian Hornung for helpful discussions.

Funding: this work was supported, in part, by the Netherlands Organisation for Scientific Research [Vidi Fellowship 864.13.003]; and intramural funds from the Leiden University Medical Center.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.mib.2021.10.016.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Smits WK, Lyras D, Lacy DB, Wilcox MH, Kuijper EJ: **_Clostridium difficile_ infection**. *Nat Rev Dis Primers* 2016, **2**:16020.

2. Aktories K, Schwan C, Jank T: **_Clostridium difficile_ toxin biology**. *Annu Rev Microbiol* 2017, **71**:281-307.

3. Dingle KE, Griffiths D, Didelot X, Evans J, Vaughan A, Kachrimanidou M, Stoesser N, Jolley KA, Golubchik T, Harding RM *et al.*: **Clinical _Clostridium difficile_: clonality and pathogenicity locus diversity**. *PLoS One* 2011, **6**:e19993.

4. Knetsch CW, Lawley TD, Hensgens MP, Corver J, Wilcox MW, Kuijper EJ: **Current application and future perspectives of molecular typing methods to study _Clostridium difficile_ infections**. *Euro Surveill* 2013, **18**:20381.

5. Clabots C, Lee S, Gerding D, Mulligan M, Kwok R, Schaberg D, Fekety R, Peterson L: **_Clostridium difficile_ plasmid isolation as an epidemiologic tool**. *Eur J Clin Microbiol Infect Dis* 1988, **7**:312-315.

6. Clabots CR, Peterson LR, Gerding DN: **Characterization of a nosocomial _Clostridium difficile_ outbreak by using plasmid profile typing and clindamycin susceptibility testing**. *J Infect Dis* 1988, **158**:731-736.

7. Mulligan ME, Peterson LR, Kwok RY, Clabots CR, Gerding DN: **Immunoblots and plasmid fingerprints compared with serotyping and polyacrylamide gel electrophoresis for typing _Clostridium difficile_**. *J Clin Microbiol* 1988, **26**:41-46.

8. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeno-Tarraga AM, Wang H *et al.*: **The multidrug-resistant human pathogen _Clostridium difficile_ has a highly mobile, mosaic genome**. *Nat Genet* 2006, **38**:779-786.

9. •• Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP: **Conjugative transfer of clostridial shuttle vectors from _Escherichia coli_ to _Clostridium difficile_ through circumvention of the restriction barrier**. *Mol Microbiol* 2002, **46**:439-452

This study has been instrumental in establishing genetic tools for *C. difficile* by describing and characterizing a functional replicon from a *C. difficile* plasmid.

10. Minton N, Carter G, Herbert M, O'Keeffe T, Purdy D, Elmore M, Ostrowski A, Pennington O, Davis I: **The development of _Clostridium difficile_ genetic systems**. *Anaerobe* 2004, **10**:75-84.

11. •• Boekhoud IM, Hornung BVH, Sevilla E, Harmanus C, Bos-Sanders I, Terveer EM, Bolea R, Corver J, Kuijper EJ, Smits WK: **Plasmid-mediated metronidazole resistance in _Clostridioides difficile_**. *Nat Commun* 2020, **11**:598

This paper reports a clinically relevant phenotype associated with plasmid carriage, wherein the high-copy number plasmid pCD-METRO confers resistance to the antimicrobial metronidazole.

12. •• Pu M, Cho JM, Cunningham SA, Behera GK, Becker S, Amjad T, Greenwood-Quaintance KE, Mendes-Soares H, Jones-Hall Y, Jeraldo PR *et al.*: **Plasmid acquisition alters vancomycin susceptibility in _Clostridioides difficile_**. *Gastroenterology* 2021, **160**:941-945 e8

The authors demonstrate in this manuscript that vancomycin treatment failure may be associated with the carriage of plasmids, part of which can also replicate in a Gram-negative host.

13. •• Ramirez-Vargas G, Rodriguez C: **Putative conjugative plasmids with tcdB and cdtAB genes in _Clostridioides difficile_**. *Emerg Infect Dis* 2020, **26**:2287-2290

In this manuscript the authors extend the findings of Ref. [31], the authors describe that putative conjugative plasmids encoding large clostridial toxins and binary toxin can occur in Clade C-I, as well as clades 2 and 4.

14. Wust J, Sullivan NM, Hardegger U, Wilkins TD: **Investigation of an outbreak of antibiotic-associated colitis by various typing methods**. *J Clin Microbiol* 1982, **16**:1096-1101.

15. Muldrow LL, Archibold ER, Nunez-Montiel OL, Sheehy RJ: **Survey of the extrachromosomal gene pool of _Clostridium difficile_**. *J Clin Microbiol* 1982, **16**:637-640.

16. • Smits WK, Weese JS, Roberts AP, Harmanus C, Hornung B: **A helicase-containing module defines a family of pCD630-like plasmids in _Clostridium difficile_**. *Anaerobe* 2018, **49**:78-84

This paper suggests that a family of plasmids that includes pCD630 is common among human and animal derived *C. difficile* strains and is characterized by a conserved module containing a gene encoding a helicase.

17. • Amy J, Bulach D, Knight D, Riley T, Johanesen P, Lyras D: **Identification of large cryptic plasmids in Clostridioides (_Clostridium_) _difficile_**. *Plasmid* 2018, **96-97**:25-38

The authors of this paper provide a thorough analysis of a family of ~45-kb plasmids that harbor a putative *C. difficile*-specific partitioning system.

18. •• Hornung BVH, Kuijper EJ, Smits WK: **An in silico survey of _Clostridioides difficile_ extrachromosomal elements**. *Microb Genom* 2019, **5**:e000926

Authors performed a comprehensive plasmid-prediction, based on *C. difficile* whole genome sequences, identifying several new plasmid families.

19. Bushman FD, Conrad M, Ren Y, Zhao C, Gu C, Petucci C, Kim MS, Abbas A, Downes KJ, Devas N *et al.*: **Multi-omic analysis of the interaction between _Clostridioides difficile_ infection and pediatric inflammatory bowel disease**. *Cell Host Microbe* 2020, **28**:422-433 e7.

20. He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, Holt KE, Seth-Smith HM, Quail MA, Rance R *et al.*: **Evolutionary dynamics of _Clostridium difficile_ over short and long time scales**. *Proc Natl Acad Sci U S A* 2010, **107**:7527-7532.

21. Amy J, Johanesen P, Lyras D: **Extrachromosomal and integrated genetic elements in _Clostridium difficile_**. *Plasmid* 2015, **80**:97-110.

22. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EP, de la Cruz F: **Mobility of plasmids**. *Microbiol Mol Biol Rev* 2010, **74**:434-452.

23. Lilly J, Camps M: **Mechanisms of theta plasmid replication**. *Microbiol Spectr* 2015, **3**.

24. Fernandez-Lopez C, Bravo A, Ruiz-Cruz S, Solano-Collado V, Garsin DA, Lorenzo-Diaz F, Espinosa M: **Mobilizable rolling-circle replicating plasmids from Gram-positive bacteria: a low-cost conjugative transfer**. *Microbiol Spectr* 2014, **2**:8.

25. del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R: **Replication and control of circular bacterial plasmids**. *Microbiol Mol Biol Rev* 1998, **62**:434-464.

26. Sengupta M, Austin S: **Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria**. *Infect Immun* 2011, **79**:2502-2509.

27. Pilla G, Tang CM: **Going around in circles: virulence plasmids in enteric pathogens**. *Nat Rev Microbiol* 2018, **16**:484-495.

28. • Ransom EM, Ellermeier CD, Weiss DS: **Use of mCherry red fluorescent protein for studies of protein localization and gene expression in *Clostridium difficile***. *Appl Environ Microbiol* 2015, **81**:1652-1660
In this manuscript, authors describe a quantitative PCR approach to establish copy number of engineered plasmids carrying novel replicons for *C. difficile*.

29. Bannam TL, Teng WL, Bulach D, Lyras D, Rood JI: **Functional identification of conjugation and replication regions of the tetracycline resistance plasmid pCW3 from *Clostridium perfringens***. *J Bacteriol* 2006, **188**:4942-4951.

30. Baxter JC, Funnell BE: **Plasmid partition mechanisms**. *Microbiol Spectr* 2014, **2**.

31. Ramirez-Vargas G, Lopez-Urena D, Badilla A, Orozco-Aguilar J, Murillo T, Rojas P, Riedel T, Overmann J, Gonzalez G, Chaves-Olarte E *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.

32. Novick RP: **Plasmid incompatibility**. *Microbiol Rev* 1987, **51**:381-395.

33. Watts TD, Johanesen PA, Lyras D, Rood JI, Adams V: **Evidence that compatibility of closely related replicons in *Clostridium perfringens* depends on linkage to parMRC-like partitioning systems of different subfamilies**. *Plasmid* 2017, **91**:68-75.

34. Adams V, Li J, Wisniewski JA, Uzal FA, Moore RJ, McClane BA, Rood JI: **Virulence plasmids of spore-forming bacteria**. *Microbiol Spectr* 2014, **2**.

35. Schwarz S, Shen J, Wendlandt S, Fessler AT, Wang Y, Kadlec K, Wu CM: **Plasmid-mediated antimicrobial resistance in Staphylococci and other firmicutes**. *Microbiol Spectr* 2014, **2**.

36. Chatedaki C, Voulgaridi I, Kachrimanidou M, Hrabak J, Papagiannitsis CC, Petinaki E: **Antimicrobial susceptibility and mechanisms of resistance of Greek *Clostridium difficile* clinical isolates**. *J Glob Antimicrob Resist* 2019, **16**:53-58.

37. Boekhoud IM, Sidorov I, Nooij S, Harmanus C, Bos-Sanders I, Viprey V, Spittal W, Clark E, Davies K, Freeman J *et al.*: **Haem is crucial for medium-dependent metronidazole resistance in clinical isolates of *Clostridioides difficile***. *J Antimicrob Chemother* 2021, **76**:1731-1740.

38. Monot M, Eckert C, Lemire A, Hamiot A, Dubois T, Tessier C, Dumoulard B, Hamel B, Petit A, Lalande V *et al.*: ***Clostridium difficile*: new insights into the evolution of the pathogenicity locus**. *Sci Rep* 2015, **5**:15023.

39. Knight DR, Imwattana K, Kullin B, Guerrero-Araya E, Paredes-Sabja D, Didelot X, Dingle KE, Eyre DW, Rodriguez C, Riley TV: **Major genetic discontinuity and novel toxigenic species in *Clostridioides difficile* taxonomy**. *eLife* 2021, **10**.

40. Goessweiner-Mohr N, Arends K, Keller W, Grohmann E: **Conjugation in Gram-positive bacteria**. *Microbiol Spectr* 2014, **2** PLAS-0004-2013.

41. Dong X, Stothard P, Forsythe IJ, Wishart DS: **PlasMapper: a web server for drawing and auto-annotating plasmid maps**. *Nucleic Acids Res* 2004, **32**:W660-W664.

42. Arredondo-Alonso S, Willems RJ, van Schaik W, Schurch AC: **On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data**. *Microb Genom* 2017, **3**: e000128.

43. Vielva L, de Toro M, Lanza VF, de la Cruz F: **PLACNETw: a web-based tool for plasmid reconstruction from bacterial genomes**. *Bioinformatics* 2017, **33**:3796-3798 http://dx.doi.org/10.1093/bioinformatics/btx462.

44. Pfeifer E, Moura de Sousa JA, Touchon M, Rocha EPC: **Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires**. *Nucleic Acids Res* 2021, **49**:2655-2673.

45. Kazlauskas D, Varsani A, Koonin EV, Krupovic M: **Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids**. *Nat Commun* 2019, **10**:3425.

46. Krupovic M: **Recombination between RNA viruses and plasmids might have played a central role in the origin and evolution of small DNA viruses**. *Bioessays* 2012, **34**:867-870.

47. Roberts AP, Smits WK: **The evolving epidemic of *Clostridium difficile* 630**. *Anaerobe* 2018, **53**:2-4.