Leiden University Medical Center

Ana M. Oliveira Paiva[1,2], Erika van Eijk[1], Annemieke H. Friggen[1], Christoph Weigel[3], and W.K. Smits[1,2] (w.k.smits@lumc.nl)

[1] Leiden University Medical Center, Leiden, The Netherlands [2] Centre for Microbial Cell Biology, Leiden, The Netherlands
[3] Insititute of Biotechnology, Technische Universität Berlin, Berlin, Germany

Identification of the unwinding region in the *C. difficile* chromosomal origin of replication

Introduction

Faithful DNA replication is crucial for viability of cells across all kingdoms. Targeting DNA replication is a viable strategy for the inhibition of bacterial pathogens (Van Eijk *et al*, 2017).
Knowledge about DNA replication in *C. difficile* is limited, but replication inhibitors are being developed as a novel therapeutic (Van Eijk *et al*, 2019).

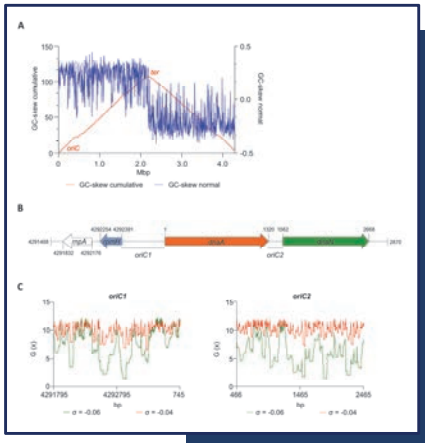
To date, no information is available on the very first steps of DNA replication in *C. difficile*. In most bacteria, replication of the chromosome starts with the assembly of the replication machinery at the origin of replication (*oriC*) and proceeds bidirectionally.
In the majority of bacteria, replication is initiated by the DnaA protein, that binds to specific sequences in *oriC* and induces origin unwinding (melting) (Chodavarapu and Kaguni, 2016).

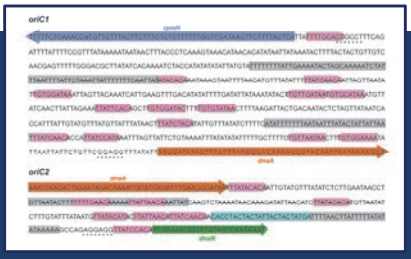
Conclusions

We used a combination of in silico predictions and in vitro experiments to demonstrate that:

1. *C. difficile* employs a bipartite origin of replication, wherein both *oriC1* and *oriC2* are necessary for unwinding *in vitro*
2. DnaA-dependent origin melting occurs at *oriC2* , located between the *dnaA* and *dnaN* genes
3. Origin architecture is conserved, suggesting that other Clostridia employ a similar mechanism

Scan the QR code below to go to the manuscript

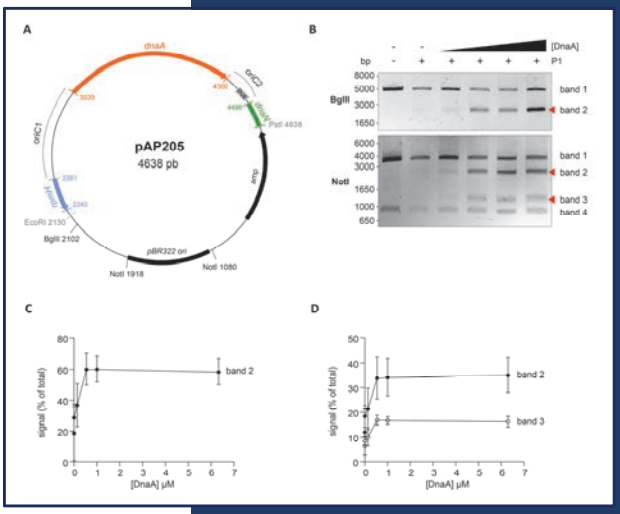




Prediction of origin features places the *C. difficile oriC* in the *dnaA-dnaN* region

[**A**] GG-skew was calculated using the Genskew Java application [**B**] Chromosomal region around the inflection point [**C**] Helically unstable AT-rich DNA stretches, putative DNA unwinding elements (DUEs) were predicted using SIST (Zhabinskaya *et al*, 2015).

[**above**] DnaA binding sites were predicted using Pattern Locator (Mrazek and Xie, 2006) using the consensus DnaA box TTWTNCACA, unstable regions using SIST and DnaA-trio (Richardson *et al*, 2016) and ribosome binding sites were manually identified. The *rpmH*, *dnaA* and *dnaN* ORFs are indicated with colored arrows.

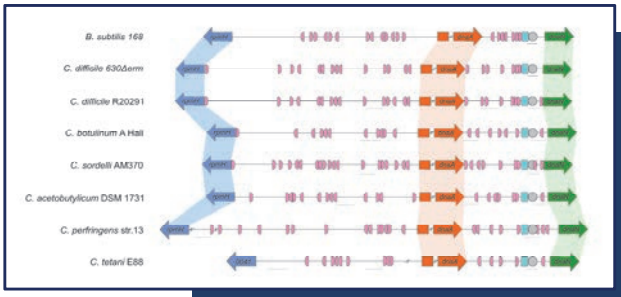


DnaA-dependent melting occurs at *oriC2*, at the DUE proximal to the DnaA-trio motif

[**A**] Plasmid containing the bipartite *C. difficile oriC* (including *dnaA*, *dnaN*), used in P1 nuclease assays [**B**] P1 nuclease assay using indicated restriction enzymes [**C**] Quantification of the BglII P1 nuclease assay (n=3) [**D**] Quantification of the NotI P1 nuclease assay (n=3).

Localized melting resulting from DnaA activity exposes ssDNA to the action of the ssDNA-specific P1 nuclease (Donczew *et al*, 2012). After incubation of a vector containing the *oriC* fragment with DnaA protein and cleavage by the P1 nuclease, the vector is purified and digested with different endonucleases to map the location of the unwound region. Signal in a band relative to total signal in the lane can be quantified.

We only observed DnaA-dependent unwinding ◀ that saturates around 1 uM DnaA for vector pAP205 that contains both *oriC1* and *oriC2*, but not for vectors containing either of these regions alone (not shown).



Origin features are conserved in Clostridia

Putative origin regions of the indicated Clostridium spp. were obtained from the DoriC 10.0 database (Luo and Gao, 2019). DnaA binding sites were predicted using Pattern Locator using the consensus DnaA box TTWTNCACA, unstable regions using SIST and DnaA-trio and ribosome binding sites were manually identified. The *rpmH*, *dnaA* and *dnaN* ORFs are indicated with colored arrows. Origins were aligned on the DnaA-trio motif.

References

GenSkew: http://genskew.csb.univie.ac.at/
Chodavarapu and Kaguni (2016) doi: 10.1016/bs.enz.2016.03.001
Donczew et al (2012) doi: 10.1093/nar/gks742
Luo and Gao (2019) doi: 10.1093/nar/gky1014
Mrazek and Xie (2006) doi: 10.1093/bioinformatics/btl551
Richardson et al (2016) doi: 10.1038/nature17962
Van Eijk et al (2017) doi: 10.1093/jac/dkw548
Van Eijk et al (2019) doi: 10.1128/AAC.01363-18
Zhabinskaya et al (2015) doi: 10.1093/bioinformatics/btu657

Acknowledgements

We thank Alan Grossman for kindly providing the pAV13 vector and *E. coli* strain CYB1002, Anna Zawilak-Pawlik for kindly providing the pori1ori2 vector and expert help in setting up the P1 assays and Luís Sousa for help with the SIDD and Pattern Locator coding files.