HAL
open science

Genomics of Clostridium tetani

Holger Brüggemann, Elzbieta Brzuszkiewicz, Diana Chapetón-Montes, Lucile Plourde, Denis Speck, Michel R. Popoff

To cite this version:

Holger Brüggemann, Elzbieta Brzuszkiewicz, Diana Chapetón-Montes, Lucile Plourde, Denis Speck, et al.. Genomics of Clostridium tetani. Research in Microbiology, 2015, 166 (4), pp.326 - 331. 10.1016/j.resmic.2015.01.002 . pasteur-01770121

HAL Id: pasteur-01770121

https://pasteur.hal.science/pasteur-01770121

Submitted on 20 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Institut Pasteur
Research in Microbiology 166 (2015) 326–331
www.elsevier.com/locate/resmic

Genomics of *Clostridium tetani*

Holger Brüggemann [a], Elzbieta Brzuszkiewicz [b], Diana Chapeton-Montes [c], Lucile Plourde [d], Denis Speck [d], Michel R. Popoff [c,*]

[a] *Department of Biomedicine, Aarhus University, Denmark*
[b] *Department of Genomics and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August-University of Göttingen, Germany*
[c] *Institut Pasteur, Bactéries anaérobies et Toxines, 28 rue du Dr Roux, F-75724 Paris Cedex 15, France*
[d] *Sanofi-Pasteur, Marcy l'Etoile, France*

Received 22 December 2014; accepted 13 January 2015
Available online 29 January 2015

Abstract

Genomic information about *Clostridium tetani*, the causative agent of the tetanus disease, is scarce. The genome of strain E88, a strain used in vaccine production, was sequenced about 10 years ago. One additional genome (strain 12124569) has recently been released. Here we report three new genomes of *C. tetani* and describe major differences among all five *C. tetani* genomes. They all harbor tetanus-toxin-encoding plasmids that contain highly conserved genes for TeNT (tetanus toxin), TetR (transcriptional regulator of TeNT) and ColT (collagenase), but substantially differ in other plasmid regions. The chromosomes share a large core genome that contains about 85% of all genes of a given chromosome. The non-core chromosome comprises mainly prophage-like genomic regions and genes encoding environmental interaction and defense functions (e.g. surface proteins, restriction-modification systems, toxin-antitoxin systems, CRISPR/Cas systems) and other fitness functions (e.g. transport systems, metabolic activities). This new genome information will help to assess the level of genome plasticity of the species *C. tetani* and provide the basis for detailed comparative studies.
© 2015 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Keywords: *Clostridium tetani*; Tetanus neurotoxin; Genomics; Plasmid

1. Introduction

Clostridium tetani is a Gram-positive, spore-forming and anaerobic rod-shaped bacterium which produces a potent neurotoxin, the tetanus toxin (TeNT). *C. tetani* forms spores which appear as translucent terminal enlargements having the general appearance of drumsticks. Spores are resistant to desiccation, heat and exposure to oxygen, and thus *C. tetani* can survive in environments such as soil for long periods of time [1]. More rarely, *C. tetani* can be found in the intestinal flora of animals and humans. Usually, *C. tetani* enters an organism through an open wound. Spores germinate and *C. tetani* grow in anaerobic conditions in necrotic tissues. TeNT is synthesized at the end of exponential growth and is released in the surrounding tissues. TeNT enters motorneurons and is retrogradely transported to the central nervous system. Then, TeNT enters inhibitory interneurons and blocks the release of neurotransmitters (glycine, GABA) upon proteolytic cleavage of the SNARE protein VAMP2 [2,3]. TeNT is responsible for all the symptoms of tetanus, characterized by spastic paralysis (muscle rigidity and spasms) [4]. In contrast to *botulinum* neurotoxins (BoNT), which show wide genetic diversity (see companion articles in this issue), only one TeNT type has been reported until now. It is presumed that *C. tetani* strains exhibit limited genetic variability. A previous study conducted via DNA/DNA hybridization showed that 10 neurotoxigenic and 3 non-toxigenic *C. tetani* strains were homogeneous [5]. At least two genetic profiles have been evidenced

* Corresponding author.
E-mail address: mpopoff@pasteur.fr (M.R. Popoff).

http://dx.doi.org/10.1016/j.resmic.2015.01.002
0923-2508/© 2015 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

in five *C. tetani* strains, as tested by pulse field gel electrophoresis [6]. It was first demonstrated that the *tent* gene is located on a large size plasmid and that the loss of this plasmid yields non-toxigenic strains [7]. The first complete genome sequence of a toxigenic *C. tetani* strain (E88) was determined in 2003 [8]. The genome contained a chromosome of approximately 2.8 Mb and the tetanus-toxin-encoding plasmid of 74 kb. One additional *C. tetani* genome was determined in 2014 [9]. Here we report the genome sequences of three additional *C. tetani* strains. We performed comparative genome studies in order to define the core and flexible genome content of *C. tetani*.

2. Materials and methods

2.1. Bacterial strains

The three *C. tetani* strains used for this study were CN655, 184.08 and strain A. Strain CN655 is a collection strain from Prevot's collection (Institut Pasteur) that originated from the Wellcome Institute (1947). Strain 184.08 was isolated from a suppurative arthritis case subsequent to a traumatism in a 75-year-old man in the south of France in 2008. Strain A is a variant of the Harvard strain.

C. tetani strains were grown in TGY broth (pH 7.5) containing trypticase (trypticase-glucose yeast peptone BBL, BD Biosciences; 30 g/L), yeast extract (Bacto Yeast Extract, BD Biosciences; 20 g/L), glucose (5 g/L) and cysteine, HCl (0.5 g/L) under anaerobic conditions.

2.2. DNA extraction and genome sequencing

Genomic DNA from all strains of *C. tetani* was extracted and purified as previously described [10]. Draft sequences of the genomes were obtained by using Illumina/Solexa GAIIx machines at the Beijing Genomics Institute (BGI) (Shenzhen, China). The assembly of sequence reads was done using SOAPdenovo (version 1.05).

2.3. Sequence analysis

Automatic annotation of all contigs was done using RAST (http://rast.nmpdr.org/). To identify homologs in different strains, comparisons of all coding DNA sequences (CDS) were made using a protein sequence-based bidirectional BLAST approach (blastP version 2.2.18). Sequence homologies were only mentioned in this study for proteins with an amino acid identity of >25% and an overlap of the query and subject sequence of >75%. Genome representations were created by a DNA plotter (Sanger Institute).

2.4. Sequence accession

GenBank accession numbers for the draft genome sequences are JSWC00000000 for strain CN655, JSWD00000000 for strain 184.08 and JWIX00000000 for strain A.

3. Results and discussion

Genomes of neurotoxigenic clostridia have been sequenced. Whereas many genome sequences of *C. botulinum* have been determined in the last decade (over 10 completed genomes and more than 30 draft genomes), only two *C. tetani* genomes are available thus far. The first genome sequence of *C. tetani* was determined in 2003 [8]. Recently, the genome of strain 12124569 was sequenced, an isolate that caused a rare case of chronic osteitis [9]. We decided to obtain more information on the genomic diversity of *C. tetani* by sequencing three additional genomes, those of strains CN655, 184.08 and A. Strain 184.08 was isolated from a suppurative arthritis of a 75-year-old man in the south of France in 2008. CN655 is a collection strain from Prevot's collection (Institut Pasteur) that originated from the Wellcome Institute (1947). Strain A is a variant of a Harvard strain. All these strains produce the tetanus toxin as tested in a mouse bioassay (data not shown).

3.1. Tetanus-toxin-encoding plasmids

All five strains possess a tetanus-toxin-encoding plasmid. G + C content, size and number of coding sequences (CDS) vary slightly, with the exception of strain 12124569 that carries a considerably smaller plasmid (Table 1). Sequence comparison revealed that plasmids of strains E88, CN665 and A (hereafter named pE88, pCN665 and pA) are highly similar to each other (Fig. 1); in contrast, plasmids of strains 184.08 and 12124569 (named p184.08 and p12124569, respectively) strongly differ (Fig. 2). Bidirectional blast (BIBlast) comparison revealed the extent of sequence divergence (Supplemental Table 1). Several genes of pE88, pCN655 and pA are not present on p184.08 and p12124569, including a two-component system and an ABC transport system. In contrast, p184.08 contains additional genes, including a type III restriction-modification system, a phage integrase and a putative bacitracin resistance transport system (Supplemental table 1B).

The sequences of TeNT and collagenase ColT are highly conserved: TeNT of pE88, pCN655 and pA are 100% identical at the amino acid level and are 98% identical (1294/1315) to TeNT from strains 184.08 and 12124569. Similarly, ColT of pE88, pCN655 and pA are 100% identical and they are 97% (963/991) identical to ColT of p184.08. Other highly conserved proteins encoded by all five plasmids are the two

Table 1
Features of tetanus-toxin-encoding plasmids.

Plasmid name	Size	G + C (%)	CDS[b]
pE88	74.0 kb	24.5	98
p184.08	ca. 73.0 kb[a]	24.9	93
pA	ca. 71.7 kb[a]	24.5	90
pCN655	ca. 69.6 kb[a]	24.3	90
p12124569	58.4 kb	24.0	77

[a] Multiple contigs.
[b] Based on RAST annotation (http://rast.nmpdr.org/).

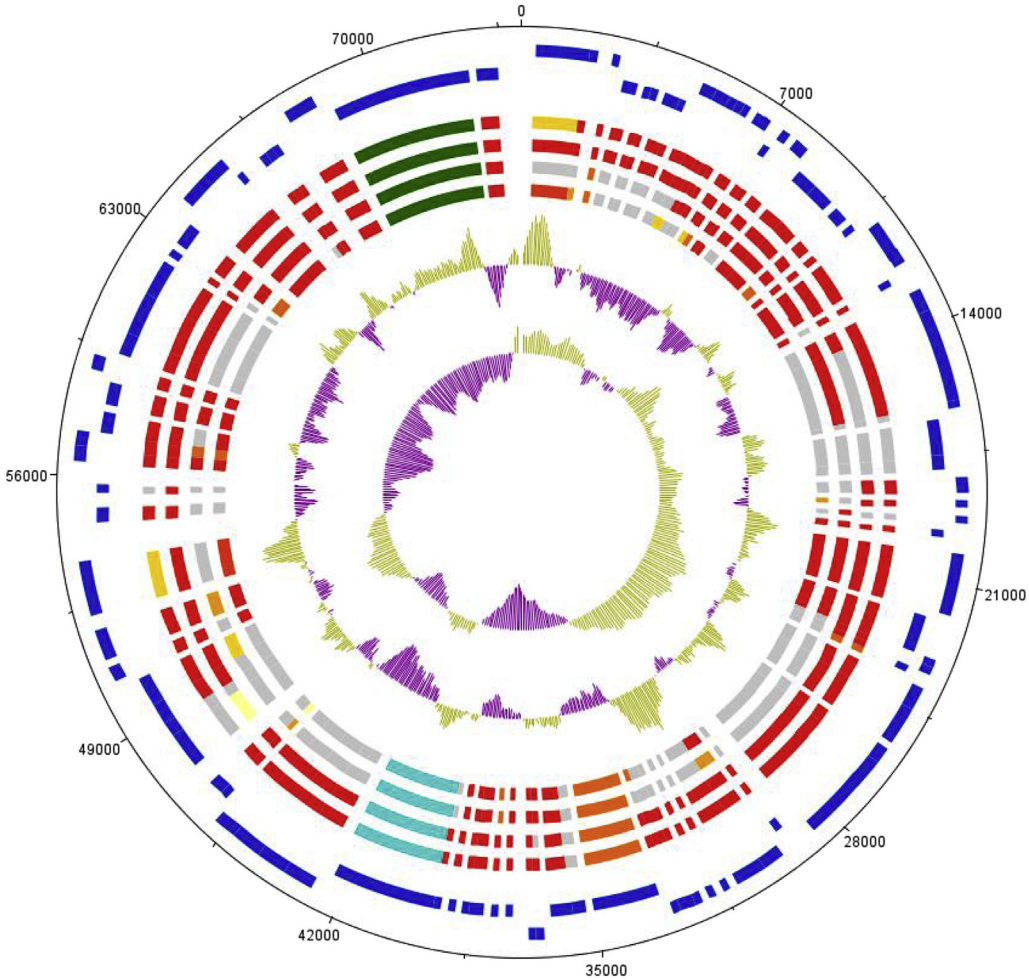

Fig. 1. A BiBlast-based comparison of five tetanus-toxin-encoding plasmids of *C. tetani* with pE88 as reference plasmid. The blue outer rings depict the CDS of pE88 (plus and minus strand); the four inner rings depict bidirectional Blast results regarding pCN655, pA, p12124569 and p184.08, respectively. Red: high similarity; orange: medium similarity; yellow: low similarity; gray: absence. The highly conserved tetanus toxin and collagenase genes are depicted in dark green and cyan, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) two inner rings depict the G + C content and the G + C skew.

putative RNA polymerase sigma factors (CTP10 and CTP11 in pE88; 94% and 99% protein identity, respectively), an ABC transport system (CTP30/CTP31) and two putative amino-peptidases (CTP20 and CTP56 in pE88; 97% and 98% protein identity, respectively).

3.2. *Chromosome of* C. tetani

The chromosomes of the five *C. tetani* strains have a similar size (ca. 2.8 Mb) and G + C content (28.6–28.8%) (Table 2). They are highly conserved: the core genome comprises 85% and 83% of all CDS in strains E88 and 184.08, respectively, when applying a cutoff of 30% protein identity (Fig. 3). Even with a cutoff of 50% protein identity, 84% and 81.5% of all CDS of strains E88 and 184.08, respectively, belong to the core genome (Supplemental Table 2).

The majority of strain-specific functions are encoded in prophage regions. For example, the chromosome of strain E88 contains three strain-specific prophage regions (Supplemental Table 2). Other strain-specific functions include transposases (in particular, in the genome of E88), cell-surface/adhesion/S-layer proteins, restriction-modification-systems, ABC transport systems and efflux pumps, two-component systems, resistance determinants and defense/repair functions as well as specific metabolic functions. For instance, strains 184.08 and 12124569 contain peptide methionine sulfoxide reductase MsrB. Strain E88 specifically contains a plasmid-like locus comprising an adenine-specific methyltransferase and a toxin-antitoxin system. Regarding metabolic functions, for example, strain 184.08 possesses the genes for an allophanate hydrolase that converts allophanate to ammonium and carbon dioxide, an essential system for using urea as a nitrogen source.

Another difference between the strains concerns the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas loci and the diversity of CRISPR spacers. The CRISPR/Cas system constitutes an adaptive immune system that protects the bacterium from bacteriophage infections and from mobile genetic elements. Strain E88 contains two loci:

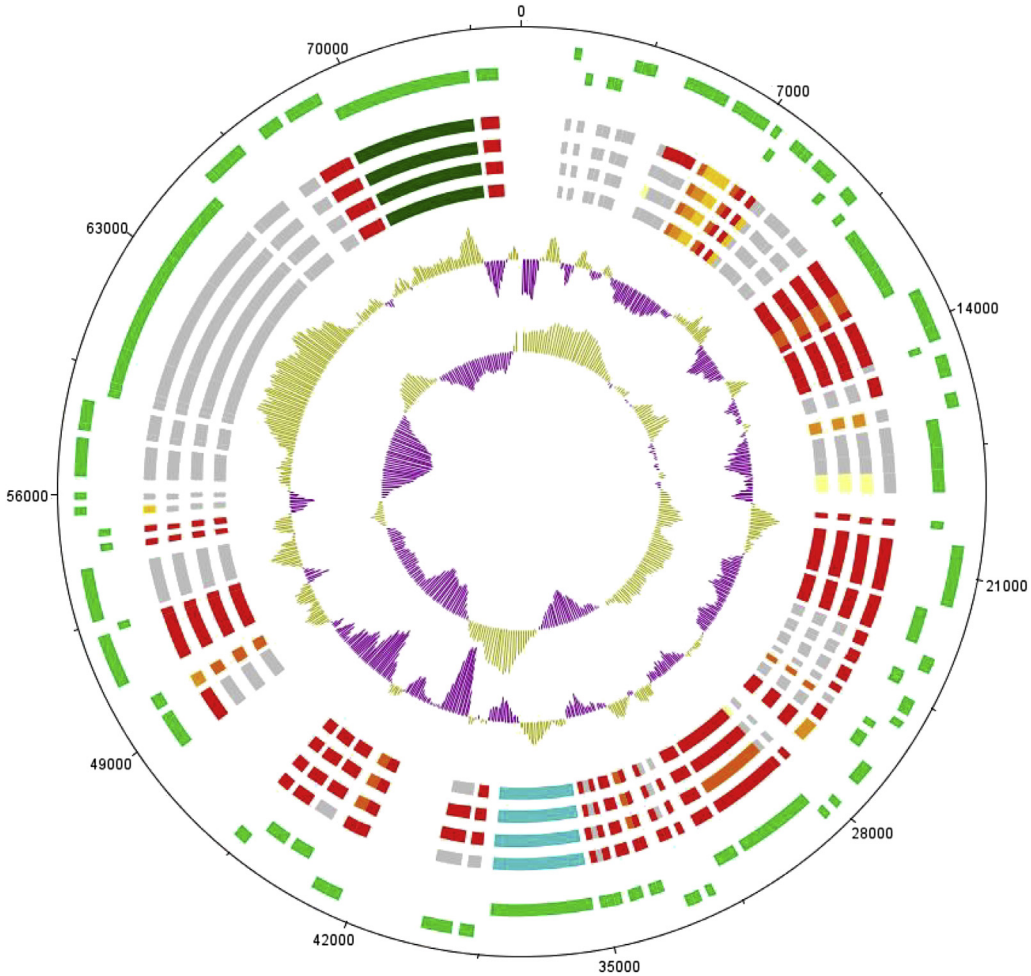

Fig. 2. A BiBlast-based comparison of five tetanus-toxin-encoding plasmids of *C. tetani* with p184.08 as reference plasmid. The green outer rings depict the CDS of p184.08 (plus and minus strand); the four inner rings depict BiBlast results regarding p12124569, pE88, pCN655 and pA, respectively. Red: high similarity; orange: medium similarity; yellow: low similarity; gray: absence. The highly conserved tetanus toxin and collagenase genes are depicted in dark green and cyan, respectively. The two inner rings depict the G + C content and the G + C skew. Please note that the map of p184.08 was reconstructed from 9 contigs aligned to pE88; thus, sequence gaps in p184.08 exist. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one locus (CTC1142-CTC1148) is composed of the *cas* genes *cas6, csx8 (cas8a1), devR (cas7), cas5, cas3, cas4 and cas1* (type I-A-like system) with a repeat consensus 5′-GTATTAGTAGCACCATATTGGAATGTAAAT-3′ and the other one (CTC1463-CTC1469) is composed of *cas6, csh1 (cas8b), csh2 (cas7), cas5, cas3, cas4, cas1* and *cas2* (type I–B system) with a repeat consensus 5′-ATTTAAATACAACTCTTGTTATTGTTCAAC-3′. Most of the spacers (66 in strain E88) show no similarity to sequences deposited in databases. Some spacers match strain-specific prophage regions in other *C. tetani* and/or *C. botulinum* genomes. Interestingly, a number of spacers also match regions within the same genome: for example, spacers in E88 match E88 genomic regions encoding stage IV sporulation protein (CTC1140), sporulation sigma-E-factor-processing peptidase (CTC1125), vancomycin b-type resistance protein VanW (CTC0360) and S-layer protein/N-acetylmuramoyl-L-alanine amidase (CTC0462). This indicates that such regions are located on mobile genetic elements that have been acquired by horizontal gene transfer events. The type IA-like CRISPR/Cas system in strain 12124569 is associated with a different repeat (5′-AATTTACATTCCAATATGGTGCTACTAATAC-3′) and contains 22 spacers. Some of them match prophage-like genomic regions in *C. tetani* and *C. botulinum*. The genome of strain 184.08 does not contain any CRISPR/Cas region. Interestingly, this genome also contains more prophage-like genomic regions that account for its larger genome size

Table 2
Features of 5 chromosomes of *C. tetani* strains.

	Size	Contigs	G + C (%)	CDS[a]
E88	2799251	1	28.7	2736
CN655	ca. 2780622	111	28.6	2734
A	ca. 2754351	90	28.6	2704
12,124,569	2807472	1	28.8	2809
184.08	ca. 2841472	135	28.6	2826

[a] Based on RAST annotation (http://rast.nmpdr.org/).

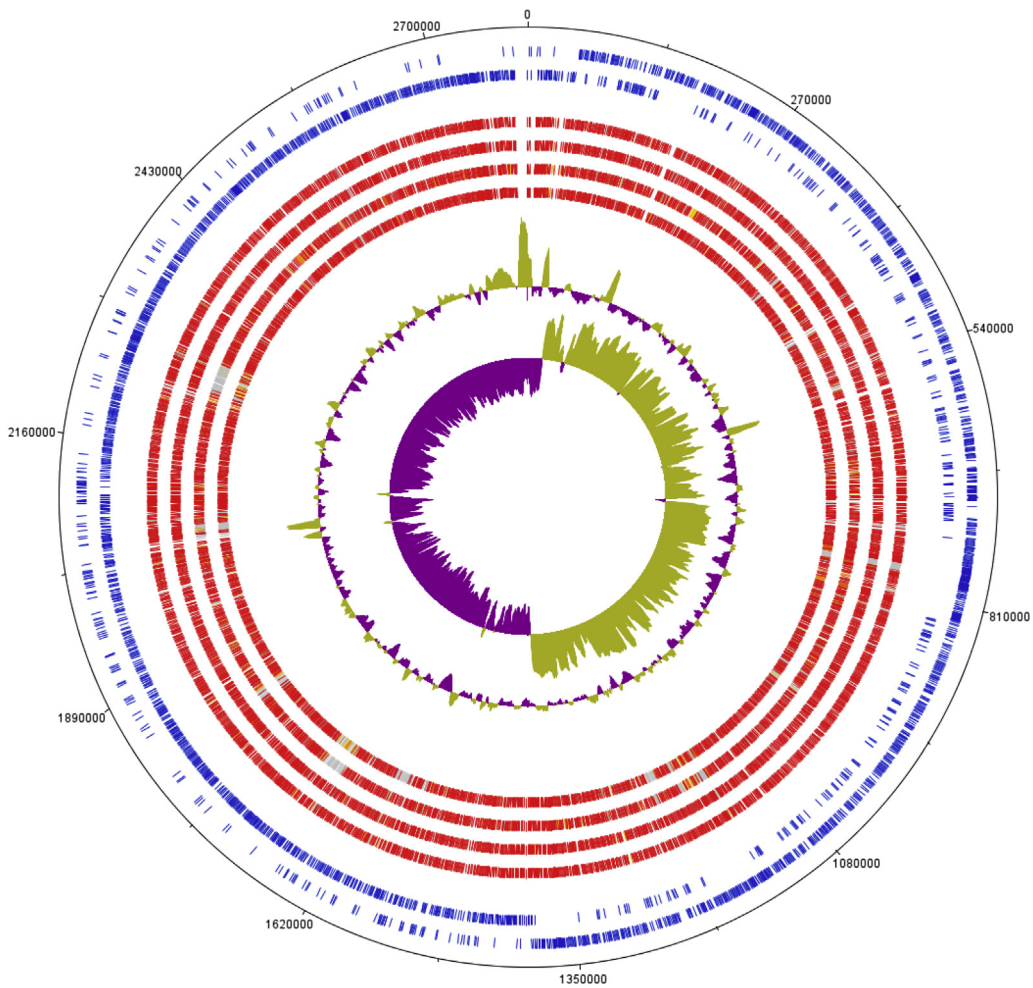

Fig. 3. A bidirectional blast comparison of the chromosomes of five *C. tetani* strains. The outer blue ring depict the CDS of reference strain E88; the inner four rings depict BiBlast results in the genomes of CN655, A, 12124569 and 184.08, respectively. Red: high similarity, orange: medium similarity, yellow: low similarity, gray: absence. The two inner rings depict G + C content and the G + C skew. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

compared to the other four *C. tetani* genomes (Table 2; Supplemental Table 2B). This indicates that the lack of CRISPR/Cas systems renders strain 184.08 more vulnerable to phage infections. The divergence of the CRISPR/Cas regions of *C. tetani* reflects the distinct bacteriophage infection history of each individual *C. tetani* strain.

In conclusion, genome comparison of *C. tetani* highlights flexible, non-core genome content. Larger islands within individual genomes often constitute prophage-like regions. Smaller genomic regions of divergence encode a set of fitness functions, increase competitiveness and protection and guarantee survival under stress conditions and attack, e.g. against bacteriophage infections.

The tetanus-toxin-encoding plasmids of the five isolates have several differences, underlying their highly flexible nature compared to the chromosome. In particular, plasmids of strains 184.08 and 12124569 differ significantly from pE88, the plasmid from the vaccine-producing strains E88 and A.

Conflict of interest

The authors have no conflict of interest.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.resmic.2015.01.002.

References

[1] Popoff MR, Mazuet C, Poulain B. Botulism and tetanus. The prokaryotes: human microbiology. 4th edn. Berlin Heidelberg: Springer-Verlag; 2013. p. 247–90.

[2] Bohnert S, Schiavo G. Tetanus toxin is transported in a novel neuronal compartment characterized by a specialized pH regulation. J Biol Chem 2005;280:42336–44.

[3] Rossetto O, Scorzeto M, Megighian A, Montecucco C. Tetanus neurotoxin. Toxicon 2013;66:59–63. http://dx.doi.org/10.1016/j.toxicon.2012.1012.1027 [Epub 2013 Feb 1016].

[4] Hassel B. Tetanus: pathophysiology, treatment and the possibility of using botulinum toxin against tetanus-induced rigidity and spasms. Toxins Basel 2013;5:73−83. http://dx.doi.org/10.3390/toxins5010073.

[5] Nakamura S, Okado I, Abe T, Nishida S. Taxonomy of *Clostridium tetani* and related species. J Gen Microbiol 1979;113:29−35.

[6] Plourde-Owobi L, Seguin D, Baudin MA, Moste C, Rokbi B. Molecular characterization of *Clostridium tetani* strains by pulsed-field gel electrophoresis and colony PCR. Appl Environ Microbiol 2005;71:5604−6.

[7] Finn CW, Silver RP, Habig WH, Hardegree MC, Zon G, Garon CF. The structural gene for tetanus neurotoxin is on a plasmid. Science 1984;224:881−4.

[8] Brüggemann H, Bäumer S, Fricke WF, Wiezr A, Liesagang H, Decker I, et al. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. Proc Ntl Acad Sci USA 2003;100:1316−21.

[9] Levy PY, Fournier PE, Lotte L, Million M, Brouqui P, Raoult D. *Clostridium tetani* osteitis without tetanus. Emerg Infect Dis 2014;20:1571−3. http://dx.doi.org/10.3201/eid2009.131579.

[10] Popoff MR, Guillou JP, Carlier JP. Taxonomic position of lecithinase-negative strains of *Clostridium sordellii*. J Gen Microbiol 1985;131:1697−703.