Molecular epidemiology of Clostridium difficile in a tertiary hospital of China

Yun-Bo Chen,1,2 Si-Lan Gu,1,2 Ze-Qing Wei,1,2 Ping Shen,1,2 Hai-Sheng Kong,1,2 Qing Yang1,2 and Lan-Juan Li1,2

1State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, PR China
2Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, PR China

Clostridium difficile infection (CDI) is caused by toxin-producing strains. It accounts for 20–30% of antibiotic-associated diarrhoea and particularly accounts for 90% of pseudomembranous colitis. The epidemiological study of C. difficile is thus important. In this study, we report the molecular epidemiology and ward distribution of C. difficile in a tertiary hospital of China. A total of 161 toxigenic strains were isolated from 1845 patients originating from different wards and the strains were characterized based on toxin profile and multilocus sequence typing. Variable isolation rates were observed in different wards and the occurrence was higher in intensive care unit and geriatric wards. Toxin gene profiling revealed that, out of the 161 isolates, 134 (83.2%) were positive for both toxin A (tcdA) and toxin B (tcdB) (A+B+) followed by toxin A-negative and B-positive (A−B+) (16.8%) isolates. However, only three of the toxigenic strains (1.9%) were positive for both the cdtA and cdtB genes. Based on the molecular epidemiology study, a total of 30 different sequence types (STs), including one new ST (ST-220), were distinguishable. ST-54 was the most prevalent (23.0%), followed by ST-35 (19.3%) and ST-37 (10.0%). None of the isolates belonged to ST-1 (ribotype 027) or ST-11 (ribotype 078). Taken together, the toxin profile and the molecular epidemiological data showed that all the ST-37 clades were of toxin type A−B+, which accounted for 59.3% of all type A−B+ isolates. Meanwhile the clade 1 genotype, ST-54, was widely distributed among the geriatric, infection and haematology wards. There was no outbreak of C. difficile infection during our study; however the possibility of prolonged outbreaks cannot be completely ignored.

INTRODUCTION

Clostridium difficile infection (CDI), which is caused by toxin-producing C. difficile, accounts for 20–30% of antibiotic-associated diarrhoea and is the causative agent of about 90% of pseudomembranous colitis (Bartlett, 2002). Over the last decade, CDI has increased markedly in both healthcare and community settings (Rupnik et al., 2009) and is regarded as the leading cause of infectious nosocomial diarrhoea among adults in developed countries. Its symptoms include a range of clinical presentations from asymptomatic carriage to diarrhoea or more severe manifestations, such as pseudomembranous colitis and toxic megacolon and about 1–5% of CDI patients require colectomy or intensive care, or eventually die (Bartlett, 2008).

Two structurally similar toxins, toxin A (TcdA, an enterotoxin) and toxin B (TcdB, a cytotoxin), are the main virulence determinants associated with CDI, and the role of these toxins in the pathogenesis of CDI has been well described (Pothoulakis, 1996). The genes for these toxins, together with three regulatory genes, are chromosomally located on the 19.6 kb pathogenicity locus (PaLoc) (Rupnik, 2008). Most pathogenic strains of C. difficile produce both toxins (A+B+). However many recent reports demonstrate the clinical importance of tcdA-negative, tcdB-positive (A−B+) isolates, which have been reported to cause sporadic cases, outbreaks and cases of pseudomembranous colitis (Kim et al., 2012). Two reports from mainland China showed that A−B+ strains accounted for 30% of all isolates (Huang et al., 2009; Yan et al., 2013), which was similar to the levels in South Korea but somewhat higher than that found in western countries.

Genotyping is a useful epidemiological tool and plays an important role in the identification of epidemic and hypervirulent genotypes. It also significantly adds to the
investigations related to the relationships between types and enteropathogenicity. Walker et al. (2013) reported that C. difficile genotype predicts mortality and outcome after CDI. Furthermore, it helps in the analysis of evolutionary paths and comparisons of lineages on a global context (He et al., 2013). Although many different genotyping methods have been developed for C. difficile (Killgore et al., 2008), none is broadly accepted. As a microbial genotyping method, multilocus sequence typing (MLST), which is based on nucleotide sequence, has been widely used for population genetics studies and global epidemiological investigations at various levels. Lemee et al. (2004) described the MLST method, and later Griffiths et al. (2010) developed a new robust MLST scheme for C. difficile, and also set up an internet-accessible database to allow simple deposition, retrieval and comparison of data.

Since the 1980s, outbreaks and serious cases of CDI reported in North America and Europe have emphasized the importance of the epidemiological surveillance of CDI. Many researchers have examined CDI isolates from different parts of the world using different typing methods to determine the molecular epidemiology of C. difficile in different regions or settings. All of these studies will greatly help us in understanding and taking some measures to control CDI. Unlike western countries, data on CDI in China are limited because of the poor awareness of CDI among general healthcare workers and the absence of clinical laboratories that can detect C. difficile under routine diagnosis for hospitalized patients presenting with diarrhoea. With emergent strains and changes in incidences, persistent surveillance in specific and targeted populations is essential. The current study was conducted to analyse the molecular epidemiology of C. difficile in our setting in order to initiate active local surveillance, identify epidemic strains and compare the results with the other epidemiological data.

**METHODS**

**Collection of C. difficile isolates.** The study was conducted at the First Affiliated Hospital of Zhejiang University (Hangzhou, China), which is a teaching hospital with 2300 beds spread across 30 wards. Stool samples (semi-formed, unformed or liquid) submitted to the clinical microbiology laboratory between 1 September 2009 and 31 December 2011 were subjected to isolation of C. difficile.

Stool specimens were treated with alcohol to a final concentration of 75% for spore selection before the anaerobic isolation of C. difficile using the selective medium cycloserine-cefoxitin-taurocholate agar (CCFA-TA; Oxoid) supplemented with 7% sheep blood at 35 °C for 48 h. The number of C. difficile colonies isolated from samples was semiquantitatively determined as: L, large number of colonies (more than 100); M, medium number of colonies (between 50 and 100); and S, small number of colonies (less than 50). Strains were identified on the basis of morphological features (flat, yellow, ground-glass appearance) and were confirmed by matrix-assisted laser desorption/ionization (MALDI)-time of flight mass spectrometry using the Bruker Daltonics microflex LT system (Bruker Daltonik) according to the operating manual. Briefly, for pre-treatment, one colony was suspended in 300 μl distilled RNase-free water. Absolute ethanol (900 μl) was added and mixed carefully, and then the sample was centrifuged (14,000 g for 2 min). The supernatant was discarded and the pellet was air dried. Formic acid (70%; 25 μl) and acetonitrile (25 μl) were added to the pellet, mixed thoroughly, and then the sample was centrifuged again (14,000 g for 2 min). The supernatant (1 μl) was deposited on the MALDI plate. Pre-treated samples were overlaid with 1 μl matrix solution (saturated solution of 2-cyano-4-hydroxybenzilic acid in 50% acetonitrile and 2.5% trifluoroacetic acid). The matrix sample was co-crystallized by air drying at room temperature. Measurements were performed with a Microflex mass spectrometer (Bruker Daltonik) using FlexControl software (version 3.0). For each spectrum, 240 shots in 40-shot steps from different positions of the target spot (automatic mode) were collected and analysed. The spectra were imported into the BioTyper software (version 2.0; Bruker) and analysed by standard pattern matching with default settings. For each isolate, the highest score of a match against a spectrum in the database was used for identification. A score of ≥2.0 was used for species identification. A commercial bacterial bead preservation system (Microbank) was used for long-term storage of the isolates at −80 °C.

**Clinical data.** Clinical data were collected by an electronic medical record system to obtain information regarding the patient’s gender, age, date of admission, ward and sample collection date. Diarrhoea was defined as two or more loose stools per day for at least 2 days. A case of CDI was defined as diarrhoeal when the patient was positive for C. difficile culture and toxin tests.

**Detection of toxin genes by PCR.** Bacteria obtained after 48 h anaerobic blood agar culture were suspended in 1 ml distilled water in a microcentrifuge tube and boiled for 15 min. After the removal of cellular debris by centrifugation (15,000 g for 2 min), the supernatant containing the genomic DNA was used for PCR amplification.

All strains were tested for the presence of tcdA and tcdB genes by PCR as described previously (Kato et al., 1998). The primer pairs were NK9/NK11 for the repetitive domain of tcdA and NK104/NK105 for tcdB. Toxin A−, toxin B− (A−B−) strains were negative with the NK9/NK11 and NK104/NK105 primer sets in PCR. A−B+ strains were positive by PCR with these two primer sets and generated an approximately 1200 bp product with the NK11/NK9 primer set. A−B+ strains were positive by PCR with the NK104/NK105 primer sets and yielded an approximately 700 bp product with the NK11/NK9 primer set. The presence of the binary toxin genes cdtA and cdtB was detected as described by Stubbs (Stubbs et al., 2000).

**MLST and analysis.** MLST with seven housekeeping genes (adh, atpA, ddx, gyaA, recA, sodA and tpi) was performed on all isolates as described previously by Griffiths et al. (2010). DNA sequences were submitted to the MLST database homepage (http://pubmlst.org/cdifficile/) to obtain the sequence type (ST). The allele of each ST was aligned from the MLST database homepage (http://pubmlst.org/cdifficile/), Molecular Evolutionary Genetics Analysis (MEGA) software (version 5.2) (http://www.megasoftware.net/) was used to reconstruct phylogenetic tree by the neighbour-joining method. The phylogenetic trees were reconstructed based on the absolute number of nucleotide differences between STs. Bootstrapting was performed with 1000 replicates.

**Statistical analysis.** SPSS version 18.0 for Windows (SPSS) was used for statistical analysis. Categorical variables were analysed using the χ² test. A P value <0.05 was considered statistically significant.

**RESULTS**

**Distribution of C. difficile isolates in wards**

Over 2 years, a total of 181 (9.8%) non-duplicate C. difficile strains were isolated from 1845 patients. To
determine the first episode of infection only the first clinical samples of stools were used for the isolation of C. difficile from each patient. Among the isolates, 161 (89 %) were toxigenic strains and the remaining 20 (11.0 %) were nontoxicigen. Out of the 161 toxigenic strains, 134 (83.2 %) were positive for both tcdA and tcdB genes (A+B+), and 27 (16.8 %) contained only the tcdB gene (A−B+). Surprisingly, only three of the toxigenic strains (1.9 %) were positive for both the cdtA and cdtB genes (A+B+). Altogether, 25 (15.5 %) strains were isolated from outpatients and 136 (84.5 %) from inpatients.

Among the isolates from inpatients, altogether 33 (24.3 %) strains were isolated from the infection ward, followed by 28 (20.6 %) strains from the gastroenterology department, 19 (14.0 %) strains from the haematology department, 19 (14.0 %) strains from the geriatric ward and 14 (10.3 %) strains from the intensive care unit (ICU) (Fig. 1). However, the isolation rates from the ICU, geriatric ward, haematology department, gastroenterology department and infection ward were 24.1 % (14/58), 22.6 % (19/84), 17.0 % (28/165), 10.3 % (19/185) and 9.7 % (33/339), respectively.

### Molecular epidemiology of the isolates

Toxigenic C. difficile strains were analysed by MLST and divided into 30 different STs, including one new ST, ST-220. The most prevalent type was ST-54 (23.0 %), followed by ST-35 (19.3 %) and ST-37 (10.0 %), while none of the isolates was identified as ST-1 (BI/NAP1/027) or ST-11 (ribotype 078). However, 13 STs were represented by only one strain. The relative abundance of the STs is summarized in Table 1. When combined with the toxin profile, MLST revealed that there was a correlation between STs and toxin genotypes. All of the ST-37 strains were toxin type A2B+, which accounted for 59.3 % of all type A2B+ strains. All of the strains of ST-26, ST-39, ST-48, ST-81, ST-100 and ST-109 were toxin type A−B+. In comparison with other STs of A−B+ type, ST-37 yielded an approximately 700 bp product by PCR with the NK11/
NK9 primer set for tcdA, whilst others were negative by PCR. The three binary toxin isolates belonged to ST-5, ST-22 and the ST-220, which were type A+B+ (Table 1).

Some correlations were observed between the genotype and the wards. ST-54 was the major type in the geriatric ward and accounted for 57.9 % (11/19). In the infection ward, ST-54 and ST-35 were the most frequent STs and both were equally represented at 19.4 %. In the haematology department, ST-54 and ST-35 accounted for 41.1 and 17.6 %, of the isolates respectively. Unlike the above three wards, ST-139 was the main type in the gastroenterology department and accounted for 17.9 %. Surprisingly, 14 strains isolated from patients in the ICU belonged to eight different STs (Fig. 1). Similarly, 25 isolates from outpatients were found distributed among 10 STs, in which ST-35 was the most frequent type and accounted for 24 %.

The monthly distribution of the top three frequent STs (ST-54, ST-35 and ST-37) is shown in Fig. 2. The highest number of CDI cases occurred in November and December of 2011 with 13 and 15 strains, respectively. In November, two CDI cases were from outpatients and 11 cases were distributed in eight different wards. In December, four cases were from outpatients and 11 cases were from seven different wards. A total of six cases were reported from the gastroenterology department in these two months, which belonged to five different STs. According to the combined data of hospital location of patients (room and bed), CDI episode date and genotype data, no evidence of outbreaks was observed (according to the definition of CDI outbreak) during the studied period in different departments with CDI patients.

ST-54 and ST-35 were unevenly distributed among males and females with male/female ratios of 27/11 and 21/10, respectively ($P<0.05$); however, there was no significant gender difference in the distribution of ST-37 (male/female=8/10).

**Phylogenetic analysis**

Phylogenetic analysis was based on the 3.5 kb sequence of housekeeping genes. We reconstructed the phylogenetic tree using the neighbour-joining method (Fig. 3) as described previously (Griffiths et al., 2010). Most genotypes were found to be from clades 1 and 4, and only two were from clade 3 (ST-5 and ST-22). Genotypes from clades 2 and 5 were not observed in our study. Three clades...
(2, 3 and 5) were binary toxin positive, and a new ST (ST-220) was found that was positive for the \textit{tcdA}, \textit{tcdB}, \textit{cdtA} and \textit{cdtB} genes. Compared with ST-54, the \textit{glyA} allele in the newly identified ST-220 was changed from allele 1 to allele 8.

**DISCUSSION**

\textit{C. difficile} is a toxin-producing pathogen that accounts for antibiotic-associated diarrhoea and is the causative agent of majority of pseudomembranous colitis (Bartlett, 2002). \textit{C. difficile} infection has increased markedly in the last decade and is regarded as the leading cause of nosocomial diarrhoea among adults in developed countries (Rupnik \textit{et al.}, 2009). The molecular epidemiology of \textit{C. difficile} is thus important, and several laboratories have reported the molecular epidemiology of this pathogen (Bauer \textit{et al.}, 2011; Dong \textit{et al.}, 2013). There is a lack of epidemiological studies of \textit{C. difficile} from China. Here, we report the molecular characterization of strains isolated from the stool specimens of patients admitted to a tertiary care centre in China. In our study, 8.7% of the samples were positive for toxigenic \textit{C. difficile}. Unfortunately, this positive rate only represents the tip of the iceberg due to the lack of clinical suspicion of \textit{C. difficile} infection in diarrhoea. Clinicians in mainland China, such as those in our hospital with 2500 beds, often fail to request tests for \textit{C. difficile} toxins in the case of unexplained diarrhoea and, thus, a significant percentage of CDI cases are missed.

In contrast to the toxigenic \textit{C. difficile} strains positive for binary toxin in European countries (Soes \textit{et al.}, 2012), we found a much lower number of binary toxin-positive strains in our study, corresponding to only three strains (1.9%). These strains revealed three STs, which were different from that of the PCR ribotype 027 strain (ST-001). Similarly, Rupnik \textit{et al.} (2003) found an occurrence prevalence of only 1.6% for the binary toxin among 310 isolates from various hospitals in Japan and Korea. We hypothesize that this may be due to different genotypes of strains prevalent in Asia. However, more data are needed to support our hypothesis.

In the present study, A\textnormal{B}+ strains accounted for 16.8% of the total isolates and 59.3% of all A\textnormal{B}+ strains were ST-37. Other A\textnormal{B}+ strains included ST-26, ST-39, ST-48, ST-81, ST-100, ST-109 and ST-219. Compared with other A\textnormal{B}+ strains, ST-37 possessed a variant of the \textit{tcdA} gene. Phylogenetic analysis showed that ST-37, ST-39, ST-81, ST-109 and ST-219 belonged to clade 4, whilst the remaining strains belonged to clade 1. A\textnormal{B}+ strains in our study were from different STs, and we therefore propose a high diversity of these strains in mainland China. Lemée \textit{et al.} (2005) reported that A\textnormal{B}+ strains were clustered in the same ST (ST-1) and proposed a low genetic diversity of A\textnormal{B}+ strains. This difference in the occurrence of A\textnormal{B}+ strains reported by Lemée \textit{et al.} (2005) and in the current study may be due to different geographical locations (France and China, respectively) where the study was performed. However, it appears that most of the A\textnormal{B}+ strains belonging to clade 4 are evolutionarily close. Furthermore, the A\textnormal{B}+ strains from clades 1 and 4 may be an example of convergent evolution. Conclusions on the potential evolutionary relationship between different A\textnormal{B}+ isolates can be drawn when all A\textnormal{B}+ isolates from different clades and sequence types are simultaneously studied and compared at the genetic level.

Although A\textnormal{B}+ strains play a major role in CDI, an increasing number of reports mention infections due to A\textnormal{B}+ strains in recent years, especially in East Asia (Huang \textit{et al.}, 2009; Shin \textit{et al.}, 2008). It is still unknown why A\textnormal{B}+ strains, especially ST-37 (ribotype 017), cause widespread disease in Asia, even though these strains do not produce binary toxin.

Genotyping by MLST identified 30 different STs including one new ST (ST-220). The hypervirulent \textit{C. difficile} ST-1 (ribotype 027) and ST-11 (ribotype 078) strains were not observed in our study. Among the strains, ST-54 (23.0%), ST-35 (19.3 %) and ST-37 (9.9 %) were the dominant types, and the remaining strains (47.8 %) belonged to 27 STs in our study. This diversity may be attributed to our hospital, which is a university teaching hospital with 2500...
beds and has patients from different communities or who are transferred from other hospitals. These data are in accordance with data published by Yan et al. (2013) who found that ST-37, ST-35 and ST-54 were the three most frequent types, accounting for 24.0, 15.4 and 11.5%, respectively. Unlike reports from a European survey that reported PCR ribotypes 014/020 (ST-2), 001 (ST-3) and 078 (ST-11) as the most prevalent strains (Bauer et al., 2011), ST-37 (PCR ribotype 017) is the most common ribotype in the region of East Asia (Kim et al., 2010; Dong et al., 2013). The reason for the widespread occurrence of ST-37 strains in East Asia is still not understood and needs dedicated research.

The novel ST-220 was isolated from a male patient in the geriatric ward. Compared with ST-54, the glyA allele in ST-220 was changed from allele 1 to 8, but ST-220 was positive for the tcdA, tcdB, cdtA and cdtB genes, and its antimicrobial resistance profiles were different from ST-54 strains isolated from the same ward (unpublished data). Although phylogenetic analysis suggests that ST-220 may have evolved from ST-54, further study is needed to verify our hypothesis.

In the present study, variable rates of isolation from different wards were observed and the rate of isolation was higher in the ICU and geriatric wards (Fig. 2). These results are similar to a report by Chung et al. (2010), who found that the incidence of CDI was highest in the ICU followed by the infection ward at a Taiwan hospital. This phenomenon may be attributed to the use of antibiotics, proton-pump inhibitors, older age of patients and underlying illness (Gorschützer et al., 2001), all of which have been found to be significantly associated with C. difficile infection (Loo et al., 2011).

Finally, in this study, we described the molecular epidemiology of the C. difficile strains isolated from different wards of the First Affiliated Hospital of Zhejiang University.

Fig. 3. Phylogenetic tree reconstructed using the neighbour-joining method based on composite sequences of seven housekeeping gene fragments. Numbers on respective branches are bootstrap values for the cluster supported by that branch. Bootstraps were generated using 1000 replicates. Values in the bracket after each ST indicate the clade and toxin types present in the strain.
University (Hangzhou, China), which is a teaching hospital with 2500 beds spread across 30 wards. There were some interesting findings that were in contrast to the study reported from the Europe and North America, like the prevalence of A−B+ strain of C. difficile, which accounted for 16.8% of the total isolates. Furthermore, 59.3% of all type A−B+ strains were of the ST-37 type. Similarly, fewer binary toxin positive strains were observed in our study, corresponding to only three strains (1.9%). The absence of binary toxin-positive strains reflects the low prevalence of clade 2, 3 and 5 isolates in the study.

Our study reported that ST-54 (23.0%), ST-35 (19.3%) and ST-37 (9.9%) were the dominant types and the remaining 27 types accounted for the remaining 47.8% of the strains. This was in contrast to reports from a European survey that described PCR ribotypes 014/020 (ST-2), 001 (ST-3), and 078 (ST-11) as the most prevalent strains (Bauer et al., 2011). Large-scale epidemiological studies at molecular level are thus warranted to establish the true load of this pathogen among the Chinese population.

ACKNOWLEDGEMENTS

The authors wish to thank Dr K. Y. Yuen and W. C. Yam of the Department of Microbiology Queen Mary Hospital Faculty of Medicine, University of Hong Kong, who provided the PCR ribotype 027 strain of C. difficile. The authors also wish to thank Medjaden Bioscience Limited for editing and proofreading the manuscript. This work was supported by a grant from the National Program on Key Basic Research Project (973 Program) (grant no. 2013CB531401), and by grants from the National Science and Technology Major Project (grant no. 2012ZX10004215) and the Education Bureau of Zhejiang Province Project (grant no. Y201223856).

REFERENCES


