VIRULENCE
2025, VOL. 16, NO. 1, 2502554
https://doi.org/10.1080/21505594.2025.2502554

Taylor & Francis
Taylor & Francis Group

Differences in virulence and drug resistance between *Clostridioides difficile* ST37 and ST1 isolates

Zirou Ouyang[a*], Jing Yang[a*], Huimin Zhang[b*], Min Zhao[a], Huimin Yang[a], Jiafeng Zhao[a], Yaxuan Yang[c], Cuixin Qiang[a], Zhirong Li[a], Pu Qin[a], Weigang Wang[a], Yanan Niu[a], and Jianhong Zhao [a]

[a]Hebei Provincial Center for Clinical Laboratories, The Second Hospital of Hebei Medical University, Shijiazhuang, Hebei, China; [b]Department of Clinical Laboratory, The Fourth Hospital of Hebei Medical University, Shijiazhuang, Hebei, China; [c]Science and Education Department, Handan First Hospital, Handan, Hebei, China

ABSTRACT

One of the most common hospital-acquired infections is caused by toxigenic Clostridioides difficile. Although C. difficile ST37 only produces a functional toxin B, it causes disease as severe as that caused by hypervirulent ST1. We aim to compare the differences in virulence and drug resistance between ST37 and ST1 isolates. We conducted whole-genome sequencing on ST37 and ST1 isolates, analyzing their type-specific genes, and the distribution and mutation of genes related to virulence and antibiotic resistance. We compared the in vitro virulence-related phenotypes of ST37 and ST1 isolates, including: TcdB concentration, number of spores formed, aggregation rate, biofilm formation, swimming diameter in semi-solid medium, motility diameter on the surface of solid medium, and their resistance to 14 CDI-related antibiotics. We detected 4 ST37-specific genes related to adherence, including lytC, cbpA, CD3246, and srtB. We detected 97 virulence-related genes in ST37 isolates that exhibit genomic differences compared to ST1. ST37 isolates showed increased aggregation, biofilm formation, and surface motility compared to ST1 in vitro. Chloramphenicol resistance gene catQ and tetracycline resistance gene tetM are present in ST37 but absent in ST1 strains. The resistance rates of ST37 to chloramphenicol and tetracycline were 45.4% and 81.8%, respectively, whereas ST1 isolates were sensitive to both antibiotics. ST1 was more resistant to rifaximin than ST37. ST37 isolates showed stronger aggregation, biofilm formation and surface motility, and had higher resistance rates to chloramphenicol and tetracycline. ST1 isolates showed stronger ability to produce toxin and sporulation, and was highly resistant to rifaximin.

ARTICLE HISTORY
Received 10 December 2024
Revised 3 March 2025
Accepted 21 April 2025

KEYWORDS
Clostridioides difficile; ST37; ST1; virulence; antibiotic resistance

Introduction

Clostridioides difficile is a Gram-positive, anaerobic, spore-forming intestinal opportunistic pathogen [1]. *C. difficile* infection (CDI) caused by toxigenic strains is one of the most common hospital-acquired infections and the leading cause of antibiotic-associated diarrhea [1].

The occurrence and development of CDI is influenced by both host intestinal homeostasis and *C. difficile* virulence factors including spores, toxins, adhesions and biofilms [2]. *C. difficile* can survive for several months after sporulation in an environment that is not conducive to its growth and then infect the host via the fecal-oral route [3]. Prolonged use of antibiotics or proton pump inhibitors leads to disruption of intestinal homeostasis and decreased resistance to *C. difficile* colonization, and hypnozoite spores hidden in the intestine can take the opportunity to germinate into vegetative cells [4]. Subsequently, *C. difficile* colonizes the intestine via adhesion factors and produces toxin A (TcdA) and toxin B (TcdB), which in turn cause diarrhea, pseudomembranous colitis, toxic megacolon, and even death [5]. Furthermore, some *C. difficile* strains use adhesions such as cell wall proteins, flagella and type IV pili to form biofilm in the intestine, and biofilm contributes to the occurrence of recurrent CDI [2]. Hence, the fluctuation of each virulence factor can affect the pathogenicity of *C. difficile*.

Since the beginning of the twenty-first century, the epidemic BI/NAP1/027 strain of *C. difficile*, which was characterized as ST1 by multilocus sequencing typing (MLST), has emerged in Europe and North America [6]. Park and Yeo [7] reported that more than half of hospital-acquired CDI and community-acquired CDI

CONTACT Jianhong Zhao zhaojh_2002@hebmu.edu.cn
*These authors have contributed equally to this work and share first authorship. Author order is determined after consultation.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/21505594.2025.2502554

in Europe and the United States is related to ST1. ST1 is typically characterized by hypervirulence via production of high levels of TcdA and TcdB as well as binary toxin [5]. Increased toxin production in ST1 is related to mutations in *tcdC*. *tcdC* is located downstream of *tcdA* within the pathogenicity locus (PaLoc) and functions as a negative regulator of toxin production during the exponential phase of growth [7]. Moreover, Burns et al. [8] reported that ST1 has a stronger sporulation ability than non-ST1 strains and is highly resistant to fluoroquinolone and rifampicin, resulting in high morbidity, recurrence, and mortality of CDI.

In Asia, the dominant strain of *C. difficile* is ST37 rather than ST1 [9]. ST37 produces only a functional TcdB. A 1.8 kb deletion in the 3' end of the *tcdA* gene and a point mutation in the 5' end result in the production of a truncated, non-functional TcdA, and ST37 does not produce binary toxin [9]. Despite lacking a functional TcdA, ST37 has spread widely, caused CDI outbreaks globally, and resulted in CDI as severe as that caused by other toxigenic strains that produce two toxins and even hypervirulent ST1 strains [9,10]. Studies have suggested that ST37 outbreaks may be associated with antimicrobial resistance (AMR), as ST37 has higher rates of resistance to clindamycin, fluoroquinolone, tetracycline, and rifaximin compared to other strains [11,12]. Although toxin production and AMR of ST37 are well documented, other virulence factors and phenotypes that may affect its pathogenicity and prevalence, such as sporulation, adhesions, aggregation and biofilm formation, have rarely been reported.

To explore the characteristics of virulence factors of ST37, and compare the differences in virulence and antibiotic resistance between ST37 and ST1, we used whole-genome sequencing to analyze the genomic features, virulence-related genes, and antibiotic resistance-related genes of ST37 and ST1 isolates. Additionally, we compared the *in vitro* pathogenicity-related phenotypes of ST37 and ST1 strains, including toxin production, sporulation, aggregation, biofilm formation and motility, as well as their resistance to antibiotics closely associated with CDI. We hope that our results can be applied to help control the global prevalence and transmission of ST37 and ST1.

Materials and methods

Collection and identification of C. difficile *strains*

We obtained all standard *C. difficile* strains (ATCC BAA-1870, ATCC43598, ATCC BAA-1382, ATCC700057) from the American Type Culture Collection (ATCC; Manassas, VA, USA). All *C. difficile* clinical isolates were isolated from patients with suspected CDI with diarrhea in China as previously described [13,14]. The stool samples inclusion criteria were soft or liquid stools, one specimen per patient, fresh or frozen stool within 48 h. Stool samples were treated with alcohol, and the mixture was inoculated onto *C. difficile* moxalactam-norfloxacin (CDMN, Oxoid, Cambridge, UK) agar plates. After incubation for 48 h at 37°C in an anaerobic chamber, presumptive *C. difficile* colonies were selected for Gram staining and observed with a microscope based on colony morphology and odor. Finally, colonies were identified by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (QuanTOF, IntelliBio, Qingdao, China) or Vitek 2 ANC card (bioMérieux, France) as previously described [13–15]. We amplified the toxin genes *tcdA*, *tcdB*, *cdtA*, and *cdtB* of the isolates and performed multilocus sequencing typing (MLST) using previously described methods [15,16]. All *C. difficile* strains were stored at −80°C until use. The information of all strains used in this study is shown in supplementary Table S1.

Whole genome sequencing and analysis

For economic reasons, only 6 of the 11 ST37 clinical isolates were sequenced, and the 12 ST1 isolates were all sequenced as part of other studies (data to be published). We extracted the genomic DNA from the samples using the Cetyltrimethyl Ammonium Bromide method with minor modification [17]. We assessed DNA concentration and quality using a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing libraries were generated using the TruSeq DNA Sample Preparation Kit (Illumina, San Diego, CA, USA) and the Template Prep Kit (Pacific Biosciences, Menlo Park, CA, USA). Subsequently, Personal Biotechnology Company (Shanghai, China) sequenced the genomes using the Illumina Novaseq platform. The filtered reads were assembled using SPAdes and A5-miseq after removal of raw reads with adapter contamination, low quality, < 50 base pairs, or 10% ambiguous bases [18,19]. The genome sequences were acquired after rectification using pilon software [20].

We used GeneMarkS v4.32 [21] to predict the coding sequences and tRNAscan-SE, Barrnap v0.9, and Rfam to identify transfer RNA, ribosomal RNA, and other non-coding RNA, respectively [22,23]. The Virulence Factors of Pathogenic Bacteria (VFDB) and Comprehensive Antibiotic Resistance (CARD) databases were used to retrieve the virulence loci and

antibiotic resistance-associated genes, respectively [24,25]. We performed function annotation by conducting blast searches against different databases, including the Non-Redundant Protein (NR) Database, Gene Ontology, Kyoto Encyclopedia of Gene and Genomes (KEGG), and Cluster of Orthologous Groups of proteins (COG) [26–29]. The orthologous genes of different *C. difficile* isolates were identified using the online tool OrthoMCL v2.0.8 [30].

Single nucleotide polymorphisms (SNPs) analyses

We used *C. difficile* ST1 epidemic strain R20291 (accession number ERR047167) as the reference strain for alignment. We identified the SNP sites that differed between the ST37 and ST1 isolates and R20291 in virulence-related genes and antibiotic resistance-associated genes using GATK software and annotated them using ANNOVAR software, respectively [31,32]. The maximum likelihood trees based on SNPs were constructed using FastTree software [33].

In vitro assays

Using simple random sampling combined with phylogenetic tree analysis, 6 ST1 isolates were selected and compared with 6 sequenced ST37 isolates for *in vitro* virulence phenotypes (supplementary Table S1). The standard *C. difficile* strain ATCC BAA-1382 was included in each test as the control.

Quantification of TcdB concentration

The quantitative test of *in vitro* TcdB production was performed as previously described [34]. Briefly, *C. difficile* overnight cultures grown in brain heart infusion broth supplemented with 5% yeast extract and 0.1% L-cysteine (BHIS) were subcultured 1:100 into BHIS broth and grown to an optical density at 600 nm (OD_{600}) of 0.5. *C. difficile* TcdB enzyme-linked immunosorbent assay (ELISA) quantitative kits (Jingmei, Jiangsu, China) were used to determine the amount of TcdB in the culture supernatants at 12 h of inoculation. The OD_{450} was measured and compared with the linear range of the standard curve to calculate the total toxin concentration.

Sporulation

We conducted sporulation assays as previously described with some modification [34]. Briefly, *C. difficile* overnight cultures grown in BHIS were subcultured 1:100 into BHIS broth and grown to an optical density at OD_{600} of 0.5. We collected 1 mL of the culture every day for the next 5 days to quantify the spores. We centrifuged the 1 mL cultures for 10 min at 3200 × *g*, removed the supernatant, washed the pellet with 1 mL sterile phosphate-buffered saline (PBS), centrifuged the sample again, and resuspended the pellet in 1 mL sterile PBS prior to heat shock for 30 min at 65°C. We then serially diluted the heat shock-treated samples, plated them on Brucella agar plates with 0.1% taurocholate, and incubated them at 37°C for 48 h.

Self-aggregation

Aggregation assays were performed as previously described [35]. Briefly, we added 1 mL of overnight cultures with OD_{620} = 0.5 into glass tubes containing 10 mL of BHIS broth and incubated the samples at 37°C. We assessed aggregation by measuring the cell density in the tubes every 2 h before (OD_{620} pre) and after (OD_{620} post) vortex resuspension of the cell aggregates. The aggregation rate was calculated using the following formula: $[(OD_{620}\ \text{post} - OD_{620}\ \text{pre})/OD_{620}\ \text{post}] \times 100$.

Biofilm formation

We conducted the biofilm assays as previously described with some modification [34]. Briefly, overnight cultures with OD_{620} of 0.5 were diluted 1:100 in BHISG medium (BHIS supplemented with 1.8% glucose) and then inoculated into 24-well plates. They were incubated anaerobically at 37°C for 1, 3, 5, and 7 days. After incubation, supernatants were removed and the wells were washed twice with sterile PBS. We air dried the biofilms, stained them with Crystal Violet (0.2% w/v) for 30 min, washed them twice with sterile PBS, and air dried them. The dye was solubilized in 75% ethanol and quantified at 570 nm. Assays were repeated three times, and the BHISG broth was included as a blank.

Motility

Motility assays were performed as previously described with some modification [36,37]. We cultured the isolates in BHIS at 37°C in an anaerobic chamber. For swimming motility assays, we added 2 µL of overnight cultures to BHIS-0.3% agar. For surface motility assays, we added 2 µL of overnight cultures to BHIS-1.8% agar supplemented with 1% glucose. After spots were dry, we incubated the plates at 37°C in an anaerobic chamber for 7 days and measured the spots at 24-h intervals. The diameter of each motility spot (widest part of the colony and perpendicular to the widest part) was measured twice and averaged.

Antimicrobial susceptibility testing (AST)

We performed AST using the agar dilution method recommended by the Clinical and Laboratory Standards Institute (CLSI, Malvern, PA, USA). Eleven ST37 and 12 ST1 isolates were tested for susceptibility to metronidazole, vancomycin, meropenem, fidaxomicin, chloramphenicol, rifaximin, tetracycline, ceftiaxone, ceftazidime, ciprofloxacin, levofloxacin, tobramycin, erythromycin, and clindamycin. The standard *C. difficile* strain ATCC 700,057 was included in each test as the control. We identified minimum inhibitory concentrations (MICs) according to the criteria in the CLSI guidelines and previously reported references [38,39].

Detection of R505K mutation in RpoB

The presence of *rpoB* gene in 5 unsequenced ST37 isolates were detected by conventional PCR. Primers specific for the *rpoB* gene were designed using Primer BLAST (sequence [5'to 3']: ATGGAAGCTATAACGCCTCAA; ACAGCACCATTTACAGTTCTA). The PCR reactions were conducted in a reaction volume of 50 µl that consisted of 25 µL of 2× Taq MasterMix, 19 µL of double-distilled H_2O, 2 µL of DNA, and 4 µL of primers. The cycling parameters of conventional PCR were as follows: 95°C for 5 min; 35 cycles of 95°C for 1 min, 50°C for 1 min, and 72°C for 1 min; and a final extension at 72°C for 5 min. The PCR products were analyzed by electrophoresis on a 1% agarose gel. Then, the amplified products were sequenced at Sangon Biotech (Shanghai, China). All *rpoB* gene sequences were performed multiple sequence alignment using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and Meinverse (https://meinverse.cn/home) software.

Statistical analysis

Data analysis was performed using SPSS Version 21.0 (IBM Corp., Armonk, NY, USA). The differences in TcdB concentration, sporulation, biofilm formation and motility capacity were analyzed using a *t*-test. The Mann-Whitney test was used to analyze the self-aggregation assay data. A p value < 0.05 was considered to be statistically significant.

Results

Characteristics of C. difficile ST37 and ST1 strains

The strains used in this study are shown in supplementary Table S1. Some strains were isolated from different cities, and the strains from the same city were from different hospitals or different departments of the same hospital. All ST37 strains were *tcdA-tcdB+cdtA-cdtB-*, all ST1 strains were *tcdA+tcdB+cdtA+cdtB+*.

Genomic features and type-specific genes of C. difficile ST37 and ST1

Supplementary Table S2 lists the general information and details about the genomic sequencing results for the 6 ST37 and 12 ST1 strains. All *C. difficile* isolates were similar in GC content (28.39–28.61%).The genome size of 6 ST37 isolates (4.16–4.26 Mb) was similar, and the genome size of 12 ST1 strains (4.07–4.09 Mb) was smaller than that of ST37 strains. All strains shared 3,398 core genes (Supplementary Figure S1). Strain ST37-B2 and ST37-B7 had the most total ($n = 3954$) and unique ($n = 43$) coding sequences, respectively. The phylogenetic tree based on SNPs showed that ST37 strains were tightly clustered together (Supplementary Figure S2).

Then, we divided the two types of *C. difficile* into two groups, and found that 3,395 genes were shared between ST37 and ST1 groups, and 263 genes and 213 genes were unique to ST37 and ST1, respectively (Figure 1(a)). We assigned the unique genes of the two *C. difficile* strains to COG functional categories and found that the largest number of them were involved in genetic information storage and processing (Figure 1(b,c)). We found 18 and 36 unique genes related to metabolism in ST37 and ST1 strains, respectively. Among them, ST37 contain more genes related to carbohydrate transport and metabolism (44.4%, 8/18), while only five genes (13.9%, 5/36) related to carbohydrate transport and metabolism were detected in ST1strains (Figure 1(b,c)). Among the carbohydrate transport and metabolism-related genes unique to ST37 strains, 3 genes encode glycoside hydrolases, including 6-phosphate-β-glucosidase (GH1), glycoside hydrolase 38 (GH38) and sucrose hydrolase (GH13_18) (Supplementary Table S3). The five unique carbohydrate transport and metabolism-related genes identified in ST1 strains all encode phosphate transferase system (PTS) transporter subunits (Supplementary Table S3). We also detected the *srlD* gene in ST1 strains, which encodes sorbitol-6-phosphate dehydrogenase. Although the *srlD* gene is not classified as a carbohydrate transport and metabolism-related gene in the COG annotation, it has been reported that *C. difficile* 630 (ST54) containing the *srlD* gene can metabolize sorbitol, while the ST81 lacking the *srlD* gene cannot utilize sorbitol.

The ST37-specific genes also included four genes related to *C. difficile* adherence, including *lytC* (*cwp84*), *cbpA*, *srtB* and CD3246 genes

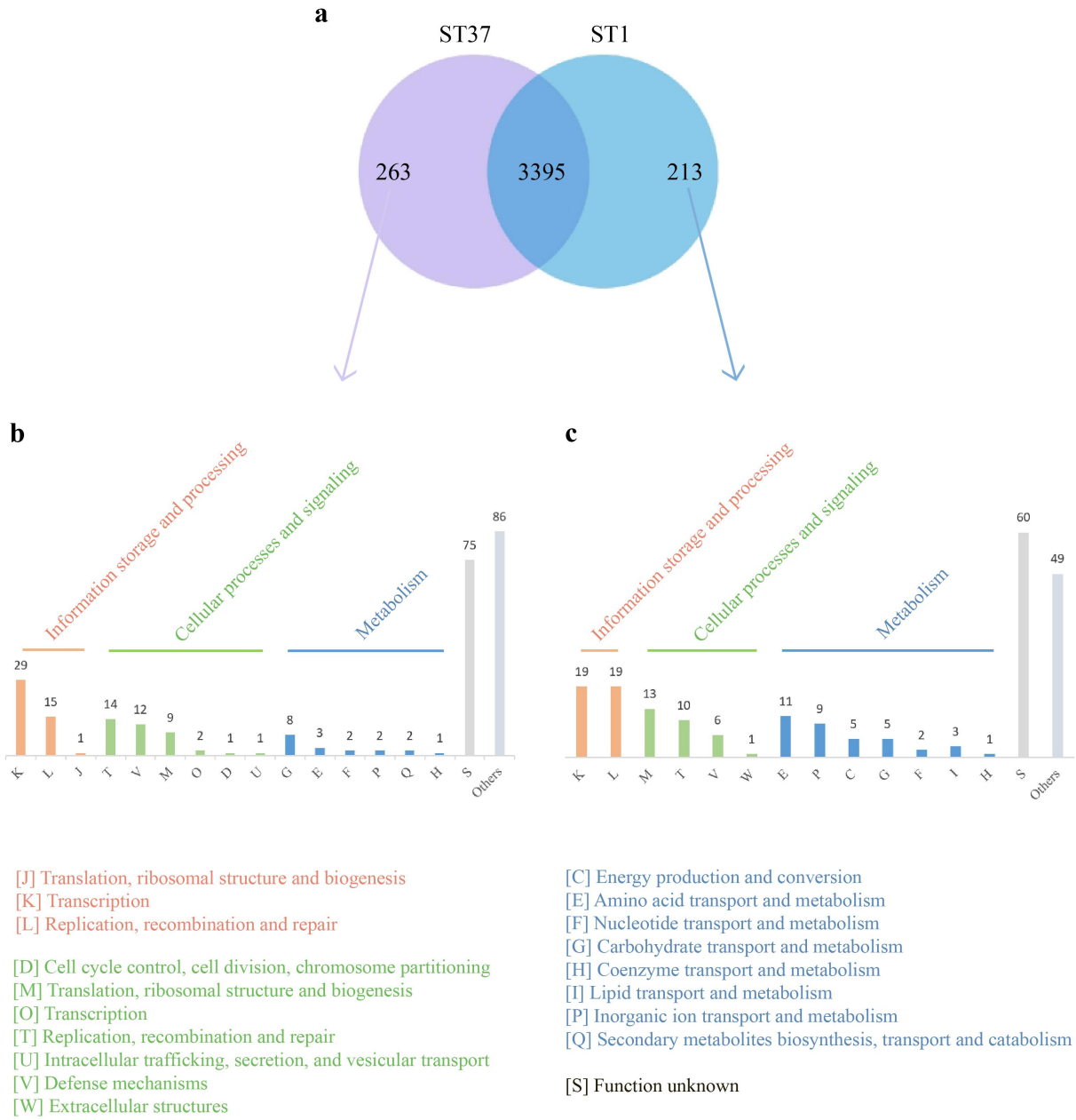

Figure 1. Distribution of orthologous genes in ST37 and ST1 strains. (a) Venn diagram showing the number of shared and unique genes for the ST37 and ST1 strains. (b) and (c) show the enrichment of cluster of orthologous groups of proteins (COG) annotation of unique genes in ST37 and ST1, respectively. Other groups represent genes not annotated in the COG database.

(Supplementary Table S3). CWP84 is a cysteine protease that plays an important role in maturation of SlpA, the principal component of surface layer (S-layer). Class B sortases (SrtB), a cysteine transpeptidase enzyme that catalyzes covalent attachment of collagen-binding protein (CbpA) to the meso-diaminopimelic acid of the peptidoglycan, thereby promoting the biofilm formation of *C. difficile*. The type-specific genes in ST37 strains also include the tetracycline resistance gene *tetM*, and three genes encoding GGDEF and/or EAL domain-containing proteins (Supplementary Table S3). These proteins are predicted to have diguanylate cyclases (DGCs) and/or phosphodiesterases (PDEs) activity that regulate the level of the second messenger cyclic diguanylate (c-di-GMP) in *C. difficile*. DGCs catalyze the synthesis of c-di-GMP, and the high intracellular levels of c-di-GMP can promote the aggregation and biofilm formation of *C. difficile*.

Among the ST1-specific genes, in addition to the toxin genes *tcdA*, *cdtA* and *cdtB*, more genes related to vancomycin resistance are worthy of attention.

Taken together, we found more genes related to carbohydrate transport and metabolism and bacterial adhesion in ST37 strains.

SNP analyses of virulence-related genes in ST37 and ST1 strains

Given that multiple virulence factors play different important roles in the pathogenesis of *C. difficile*, we then focused on the genomic differences of virulence-related genes between ST37 and ST1 strains. Based on KEGG annotation and existing studies, in addition to toxin genes *tcdA*, *tcdB*, *cdtA* and *cdtB*, genes related to sporulation, adherence, biofilm formation, and motility, which are involved in the pathogenic process of *C. difficile*, were included in the category of virulence-related genes.

We first analyzed the PaLoc region of ST37 strains. As shown in Figure 2, the PaLoc region of the 6 ST37 strains were highly similar. They all contained *tcdR*, *tcdB*, *tcdE*, *tcdC*, and incomplete and discontinuous truncated *tcdA* genes. In addition to the mutant *tcdA* gene, the *tcdB*, *tcdR*, and *tcdC* genes all contained non-synonymous single nucleotide variants (SNVs). The *tcdB* gene contained the highest number of SNVs, whereas the *tcdE* gene was relatively conserved, with only one synonymous SNP (Figure 2). Mutations in genes within the PaLoc region were not detected in any of the 12 ST1 strains sequenced.

As shown in Figure 3(a), in addition to the toxin genes *tcdA* and *tcdB*, we detected 95 virulence-related genes containing non-synonymous SNPs in ST37 strains, and most were motility-related genes ($n = 40$), followed by 29 sporulation-, 22 adherence- and 4 biofilm-related genes (Figure 3(a)). Notably, flagellar genes accounted for the vast majority of all mutated motility-related genes, and only five mutated type IV pili genes were detected. The filament cap gene *fliD* had the most mutation sites, followed by the flagellar hook gene *fliK*, the filament gene *fliC*, and hook-filament junction genes *flgL* and *flgK*. Flagellin proteins FliD, FliC, FlgL, FlgK, and FliK together form the extracellular part of flagella, which is involved in the swimming motility, adherence and biofilm formation of *C. difficile*.

Among the mutated adherence-related genes in the ST37 strains, we identified 18 genes encoding cell-wall proteins (CWPs) (Figure 3(a)). CWPs are components of the surface layer (S-layer), and more than 30 CWPs and the S-layer protein SlpA together form S-layer proteins, which are involved in *C. difficile* adhesion and colonization and in evoking a strong immune response. We found that gene CD2831, which encodes a surface protein that binds collagen and promotes adhesion and biofilm formation in *C. difficile*, harbored the most non-synonymous SNP sites among the mutated adherence-related genes. We also identified the gene *zmpI* (CD2830), which encodes zinc metalloprotease (ZmpI). It contains five non-synonymous SNPs, which can cleave the surface protein CD2831 from peptidoglycan. The expressions of two other identified genes, CD2831 and *zmpI*, are affected by the bacterial second messenger cyclic diguanylate (c-di-GMP): under high levels of c-di-GMP, expression of CD2831 is upregulated, whereas *zmpI* expression is repressed and *C. difficile* adhesion and early biofilm formation are promoted. We also identified three genes encoding c-di-GMP metabolizing enzymes, including *adrA*, *yegE*, and *bifA*. The first two genes encode diguanylate cyclase and the last one encodes c-di-GMP phosphodiesterase (Figure 3(a)).

Figure 3(b) shows all genes with non-synonymous mutations in the ST1 strains. Clearly, we detected fewer

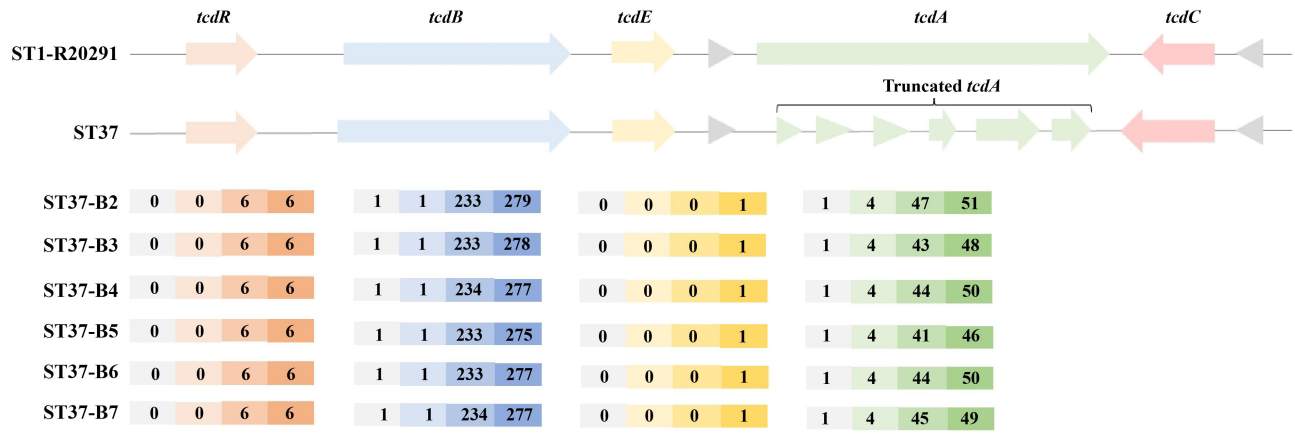

Figure 2. Comparative analysis of the PaLoc region of ST37 strains and R20291. The heat map below the PaLoc region corresponds to mutations in genes of the same color within the PaLoc region. The color changes from light to dark represent frameshift, stop gain, non-synonymous single nucleotide variants (SNV), and synonymous SNV, respectively. Numbers within the heat map represent the number of mutation sites.

a

	Gene	ST37 B2	B3	B4	B5	B6	B7
Toxin	tcdB	233	233	234	233	233	234
	tcdA	47	43	44	41	44	45
Sporulation (n=29)	spoIVCA	46	44	45	46	43	44
	cotA	22	24	22	30	27	22
	cotH	18	18	18	17	18	18
	spoIID	5	5	5	5	5	5
	spoIIIAF	3	3	3	3	3	3
	spoVFA	3	3	3	3	3	3
	sleC	2	2	2	2	2	2
	spoVAD	2	2	2	2	2	2
	oppD	2	2	2	2	2	2
	spoIIIAA	2	2	2	2	2	2
	spoIIIAB	2	2	2	2	2	2
	spoIIM	2	2	2	2	2	2
	whiA	2	2	2	2	2	2
	spoIIE	2	2	2	2	2	2
	spoIIID	1	1	1	1	1	1
	cotJC	1	1	1	1	1	1
	oppB	1	1	1	1	1	1
	spoVB	1	1	1	1	1	1
	spoIIIAE	1	1	1	1	1	1
	spoIIIAG	1	1	1	1	1	1
	sigK	1	1	1	1	1	1
	codY	1	1	1	1	1	1
	cotJB	1	1	1	1	1	1
	yqfD	1	1	1	1	1	1
	spoIIP	1	1	1	1	1	1
	spoIVA	1	1	1	1	1	1
	sigH	1	1	1	1	1	1
	spmB	1	1	1	1	1	1
	jag	1	1	1	1	1	1

	Gene	ST37 B2	B3	B4	B5	B6	B7
Adherence (n=22)	CD2831	52	53	51	55	53	53
	slpA	40	40	39	39	42	40
	cwp12	27	27	27	27	27	27
	cwp21	23	21	23	21	23	23
	cwp20	19	19	19	19	19	19
	cwp2	17	17	17	17	17	17
	cwp26	11	12	15	12	12	13
	cwp10	11	11	11	11	11	11
	cwp28	9	9	9	9	9	9
	cwp25	8	8	8	8	8	8
	cwp23	8	8	8	8	8	8
	cwp22	6	6	6	6	6	6
	cwp11	6	6	6	6	6	6
	zmp1	5	5	5	5	5	5
	cwp27	5	5	5	5	5	5
	cwp13	5	5	5	5	5	5
	cwp5	5	5	5	5	5	5
	cwp84	3	3	3	3	3	3
	cwp18	3	3	3	3	3	3
	cwp8	3	3	3	3	3	3
	groEL	2	2	2	2	2	2
	cwp7	2	2	2	2	2	2
Biofilm	adrA	49	49	49	49	49	49
	yegE	16	16	16	16	16	16
	bifA	14	14	14	14	14	14
	srtB	5	5	5	5	5	5

	Gene	ST37 B2	B3	B4	B5	B6	B7
Motility (n=40)	fliD	58	60	60	65	67	61
	fliK	57	58	55	54	55	57
	fliC	35	33	33	35	35	35
	flgL	33	32	28	30	33	27
	flgK	24	24	24	24	24	24
	fliW	24	24	23	22	24	24
	motA	20	20	20	20	20	20
	flgM	18	19	19	17	18	19
	fliG	15	15	13	15	15	15
	flgD	18	15	18	17	14	18
	fliF	13	14	14	14	14	14
	fliR	13	14	14	14	14	14
	flgE	10	11	9	11	11	11
	fliS	10	10	10	10	10	10
	pilC	10	10	10	10	10	10
	fliJ	9	10	11	9	9	11
	flhG	7	7	7	7	7	7
	fliH	6	6	6	6	6	6
	flhF	5	5	5	5	5	5
	fliI	5	5	5	5	5	5
	fliO	5	5	5	5	5	5
	motB	5	5	5	5	5	5
	pilD	5	5	5	5	5	5
	flgG	4	4	4	4	4	4
	flgJ	4	4	4	4	4	4
	fliM	4	4	4	4	4	4
	pilT	4	4	4	4	4	4
	pilV	4	4	4	4	4	4
	cheC	3	3	3	3	3	3
	flhA	3	3	3	3	3	2
	flbD	2	2	2	2	2	2
	flgB	2	2	2	2	2	2
	flgF	2	2	2	2	2	2
	fliA	2	2	2	2	2	2
	fliE	2	2	2	2	2	2
	fliL	2	2	2	2	2	2
	fliQ	2	2	2	2	2	2
	pilB	2	2	2	2	2	2
	flgC	1	1	1	1	1	1
	fliNY	1	1	1	1	1	1

b

Gene	ST1 A2	A3	A4	A5	A7	A8	A9	A10	A11	A12	A13	A14
cdeA	2	2	2	2	2	2	2	2	2	2	2	2
nimA	2	2	2	2	2	2	2	2	2	2	2	2
rpoB	2	2	2	2	2	2	2	2	2	2	2	2
hisZ	1	1	1	1	1	1	1	1	1	1	1	1
mntH	1	1	1	1	1	1	1	1	1	1	1	1
mutS	1	1	1	1	1	1	1	1	1	1	1	1
nadA	1	1	1	1	1	1	1	1	1	1	1	1
vanR	1	1	1	1	1	1	1	1	1	1	1	1
yceD	1	1	1	1	1	1	1	1	1	1	1	1
slpA	1	1	1	1	1	1	1	1	1	0	1	1
carB	1	1	1	1	1	1	0	0	0	0	0	0
atpD	0	0	0	0	0	0	1	1	0	1	1	1
rsmB	0	0	0	0	0	0	1	1	0	1	1	1
tsaD	0	0	0	0	0	0	1	1	0	1	1	1
gyrB	0	0	0	0	0	0	1	0	0	0	0	1
pcp	0	0	0	0	0	0	0	0	0	0	0	1
prfB	0	0	0	0	0	0	0	0	0	0	0	1
gyrA	1	0	0	0	0	0	0	1	0	1	1	0
marR	0	0	0	1	0	1	0	0	1	0	0	0

Figure 3. Heat map of non-synonymous mutations of genes in ST37 and ST1 strains. (a) Heat map of non-synonymous mutations of virulence-related genes in ST37 strains. (b) Heat map of all genes with non-synonymous mutations in ST1 strains. Numbers within the heat map represent the number of mutation sites.

non-synonymous mutated genes and mutation sites in ST1 strains compared to ST37 strains. Only one mutated virulence-related gene, *slpA*, was detected in the ST1 isolates, which is the only mutant virulence-related gene shared between ST37 and ST1 isolates (Figure 3).

The above results showed that the virulence-related genes of ST37 were genomically distinct from those of ST1, but whether this indicates that there are differences in pathogenesis-related phenotypes between the two strains remains to be investigated.

ST37 showed decreased sporulation but increased aggregation, biofilm formation, and surface motility compared to ST1

Since the virulence-related genes of ST37 were genomically distinct from those of ST1, we accordingly compared *in vitro* pathogenesis-related phenotypes of ST37 and ST1, including TcdB production, sporulation, aggregation, biofilm formation and motility.

We first determined the growth of *C. difficile* in BHIS medium. As shown in Supplementary Figure S3, both ST37 and ST1 strains entered the stabilization phase after approximately 6 h of growth. Therefore, we tested the concentration of TcdB protein in the culture supernatant after 12 h of growth. As shown in Figure 4(a), the accumulated TcdB protein concentration of ST1 strain after 12 h of growth was approximately 1.87 pg/mL, which was slightly higher than that of ST37 strain (1.80 pg/mL, $p > 0.05$).

Subsequently, we performed spore counts for five consecutive days to compare the sporulation capacity of ST37 and ST1. As shown in Figure 4(b), at all five experimental time points, the number of spores of ST1 strains was more than that of ST37 strains, and the difference on day 1 and day 2 was statistically significant ($p < 0.05$).

The adherence of *C. difficile* includes not only the adhesion to host intestinal epithelial cells, but also the adhesion between individual bacteria. Here, we compared the adhesion ability between bacteria of ST37 and ST1 isolates by self-aggregation assay. We calculated the self-aggregation rate by continuously measuring the absorbance changes of *C. difficile* growing in tubes over a 12 h period. As shown in Figure 4(c), the aggregation rate of ST37 strains increased over time, but that of ST1 strains changed little, and the aggregation rate of ST1 strains was significantly lower than that of ST37 strains at 6 h, 8 h and 10 h. At 24 h of culture, the aggregation rate of ST1 was only about 20%, whereas more than half of the ST37 strains were aggregated at the bottom of the glass tubes (Figure 4(c)).

We further compared the biofilm formation ability of the two strains on days 1, 3, 5, and 7. As shown in Figure 4(d), the biofilm formation of ST1 was slightly higher than that of ST37 in the early stages (day 1) ($p > 0.05$). However, with the prolongation of culture time, the biofilm formation of ST1 strains increased slightly, while the biofilm formation of ST37 strains increased significantly. On days 5 and 7, the biofilm formation of ST37 was significantly higher than that of ST1 (Figure 4(d)).

Finally, we compared the swimming motility of ST37 and ST1 strains on BHIS-0.3% agar and surface motility on BHIS-1.8% agar plates. The results showed that ST1 was better at locomotion in semi-solid agar (Figure 4(e)), while ST37 had significantly stronger motility on the solid agar surface than ST1 (Figure 4(f)).

In summary, the virulence-related genes of ST37 were genomically distinct from those of ST1, and the *in vitro* pathogenesis-related phenotypes in the two types of strains were also different. ST37 isolates showed stronger self-aggregation and biofilm formation ability and surface motility, while ST1 was superior in toxin production, sporulation and swimming motility.

Antibiotic resistance of the ST37 and ST1 strains

Given the significance of AMR in the epidemic and pathogenesis of *C. difficile*, we then analyzed the presence of antibiotic resistance-related genes and their SNP profiles in ST37 and ST1 isolates. The retrieval of antibiotic resistance-related genes combines the CARD databases, the KEGG annotation, and available studies. As shown in Table 1, We detected 33 and 32 antibiotic resistance-related genes in ST37 and ST1, respectively. *catQ* gene and *tetM* gene were present in ST37 strains but absent in ST1 strains, which are associated with chloramphenicol resistance and tetracycline resistance, respectively. In contrast, the ST1 strains contained more genes associated with vancomycin resistance, including *vanB*, *vanRG2*, *vanSG* and *vanTG* genes (Table 1).

Table 2 shows non-synonymous mutation sites in common antibiotic resistance-related genes in ST37 and ST1 isolates. The number of non-synonymous SNP sites detected in ST37 isolates was also higher than in ST1 isolates. The XdhD and XdhA related to metronidazole resistance in ST37 isolates contained the most non-synonymous SNPs. There were only two identical mutations in ST37 and ST1 isolates, including the I155L mutation in NimA and R505K mutation in RpoB.

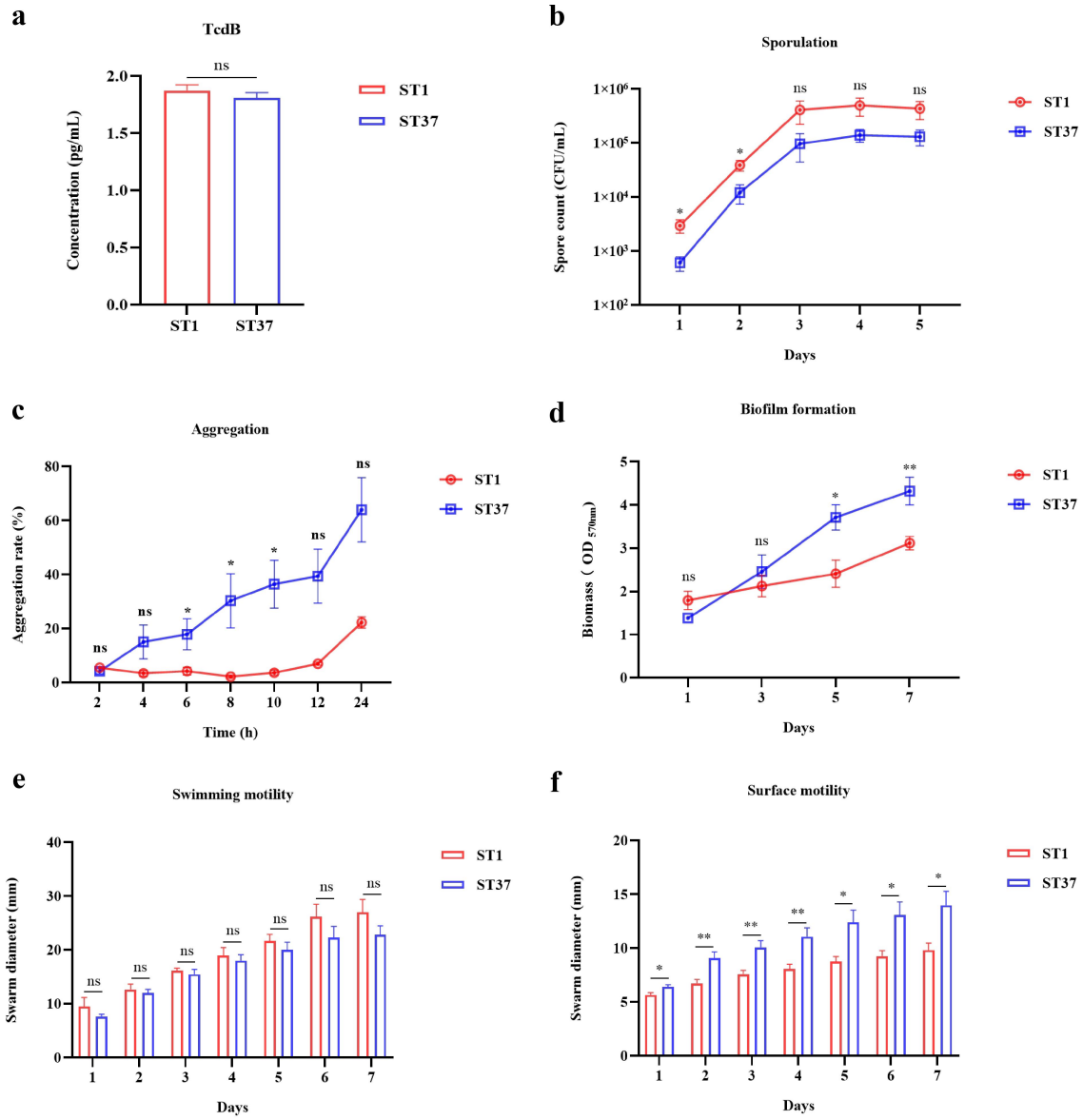

Figure 4. Comparison of pathogenesis-related phenotypes between ST37 and ST1 strains. (a) Comparison of TcdB production between ST37 and ST1 strains. (b) Comparison of the number of spores measured for five consecutive days. (c) Comparison of self-aggregation rates. (d) Comparison of biofilm formation capacity. (e) and (f) show comparison of swimming motility and surface motility ability measured for seven consecutive days, respectively. Error bars represent standard error of the mean. ns, no significance, *$p < 0.05$; **$p < 0.01$.

To investigate the relationship between the detected antibiotic resistance-related genes and the drug resistance of ST37 and ST1 isolates, we conducted *in vitro* AST. In addition to the 6 sequenced ST37 isolates, we also included 5 ST37 clinical isolates from different sources for the AST to obtain more comprehensive results. We selected 14 antibiotics according to the above antibiotic resistance-related genes and common antibiotics associated with the pathogenesis and treatment of CDI. As shown in Figure 5, both ST37 and ST1 were 100% resistant to ceftiaxone and ceftazidime (beta-Lactam), ciprofloxacin and levofloxacin (fluoroquinolones), tobramycin (aminoglycoside), erythromycin and clindamycin (macrolide-lincosamide-streptogramin B, MLS_B). Compared to ST1 strains, ST37 strains had higher resistance rates to chloramphenicol and tetracycline, which were 45.4% and 81.8%, respectively, whereas ST1 strains were sensitive to both antibiotics (Figure 5). The resistance rates of ST1 strains to rifaximin is as high as 100%, while the resistance rate of ST37 to rifaximin is lower than that of ST1, at 63.6 %. MICs of 14 antimicrobial agents against 11 ST37 and 12 ST1 *C. difficile* clinical isolates were shown in supplementary Tables S4 and S5, respectively.

Table 1. Description and function of antibiotic resistance-related genes in ST37 and ST1 strains.

Gene symbol	Description of the encoded protein	Putative function	ST37 (n = 33)	ST1 (n = 32)
aac(6″) -Ie-aph(2″″) -Ia	AAC(6″)-APH(2″) bifunctional resistance protein	Aminoglycoside resistance	√[a]	√
ant(6)-Ib	Aminoglycoside adenylyltransferase	Aminoglycoside resistance	√	×[b]
aph(3″)-Ib (strA)	aminoglycoside O-phosphotransferase APH(3″)-Ib	Aminoglycoside resistance	√	√
dfrD_2	Dihydrofolate reductase	Antifolate resistance	√	×
bcrA	bacitracin resistance ABC transporter ATP-binding subunit	Bacitracin resistance	√	√
mecRI	BlaR1 family beta-lactam sensor/signal transducer	beta-Lactam resistance	√	√
rtcB	RNA ligase RtcB family protein	beta-Lactam resistance	√	√
SaurJH1_0031	Penicillinase repressor	beta-Lactam resistance	√	×
catQ	Chloramphenicol acetyltransferase	Chloramphenicol resistance	√	×
gyrA	DNA gyrase subunit A	Fluoroquinolone resistance	√	√
gyrB	DNA gyrase subunit B	Fluoroquinolone resistance	√	√
arlR	Member of the two-component regulatory system ArlS/ArlR	Involved in the regulation of adhesion, autolysis, multidrug resistance and virulence	√	√
ermB	23S rRNA (adenine(2058)-N(6))-methyltransferase Erm(B)	Macrolide-lincosamide-streptogramin B(MLSB) resistance	√	√
mecI (blaI)	Penicillinase repressor	Methicillin resistance	×	√
feoB1	Ferredoxin oxidoreductase	Metronidazole resistance	√	√
nimA	5-nitroimidazole reductase	Metronidazole resistance	√	√
xdhD	Xanthine dehydrogenase	Metronidazole resistance	√	√
xdhA	Xanthine dehydrogenase	Metronidazole resistance	√	√
mepA	MATE family efflux transporter	Multidrug resistance	√	√
rpoB	DNA-directed RNA polymerase subunit beta	Rifampicin/fidaxomicin resistance	√	√
rpoC	DNA-directed RNA polymerase subunit beta	Rifampicin/fidaxomicin/vancomycin resistance	√	√
rpsL	30S ribosomal protein S12	Streptomycin resistance	√	√
tetM	Tetracycline resistance protein	Tetracycline resistance	√	×
gdpP	DHH family phosphoesterase	Vancomycin resistance	√	√
murG	Undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	Vancomycin resistance	√	√
vanB	D-Ala:D-Ser ligase	Vancomycin resistance	×	√
vanR	Vancomycin resistance response regulator transcription factor VanR	Vancomycin resistance	√	√
vanRF	Two-component response regulator	Vancomycin resistance	√	√
vanRG	Response regulator VanRG	Vancomycin resistance	√	√
vanRG2	D,D-peptidase	Vancomycin resistance	×	√
vanS	Vancomycin resistance histidine kinase VanS	Vancomycin resistance	√	√
vanSG	Sensor histidine kinase	Vancomycin resistance	×	√
vanTG	Serine racemase	Vancomycin resistance	×	√
vanU	Glycopeptide resistance transcriptional regulator VanU	Vancomycin resistance	√	√
vanZF	Glycopeptide antibiotics resistance protein	Vancomycin resistance	√	×
SAV2088	Cardiolipin synthetase	–	√	√
tufA	Elongation factor Tu	–	√	√
vatB	Acetyltransferase	–	√	√

[a]Presence; [b]Absence, [c]Function unknown.

Finally, we detected the R505K mutation in RpoB in 5 unsequenced ST37 isolates by Sanger sequencing. As shown in Supplementary Figure S4, the R505K (AGA>AAA) mutation in RpoB was detected in all four rifaximin-resistant ST37 strains.

Discussion

The pathogenicity of ST37 was not appreciated until 1999, when an outbreak of TcdA-TcdB+ *C. difficile*-associated diarrhea occurred in a Canadian tertiary care hospital [40,41]. Since then, a global outbreak of ST37 has occurred, especially in Asia, but also in many non-Asian countries such as Germany, Canada, Portugal, South Africa, and the United States [10]. As a toxigenic strain that produces only functional TcdB, ST37 is not only capable of causing disease as severe as that caused by other toxigenic strains that produce both TcdA and TcdB, but its pathogenicity is also comparable to the most to that of the "notorious" epidemic hypervirulent *C. difficile* ST1 [42,43]. Moreover, researchers have reported that ST37 is usually associated with severe CDI or even lethal CDI [9,44]. Hence, *C. difficile* ST37 is considered to be one of the most successful epidemic strains in terms of both pathogenicity and prevalence.

The pathogenicity of *C. difficile* is affected by both the host intestinal environment and its own virulence factors. Imbalance of intestinal homeostasis and reduced colonization resistance create the prerequisites for the occurrence of CDI, and the synergistic toxic effect of many virulence factors of *C. difficile* is the decisive factor for the occurrence and development of CDI [45]. In this study, we explored the differences in virulence and drug resistance between ST37 and ST1 isolates through comparative genomics and *in vitro*

Table 2. Nonsynonymous mutations in common antibiotic resistance-related genes of ST37 and ST1 strains.

Gene symbol	Antibiotic resistance	ST37	ST1
gyrA	Fluoroquinolones resistance	I406L, N468D	D71A[a], R296G[b]
gyrB	Fluoroquinolones resistance	S366A	D426N[c]
ermB	Macrolide-lincosamide-streptogramin B(MLS_B) resistance	×[d]	×
feoB1	Metronidazole resistance	×	×
nimA	Metronidazole resistance	I155L	I155L, Y130S
xdhD	Metronidazole resistance	E561D, G559D, T413I, M212I, M212T, M144V, T143I, T143A, L129I, V58I, N31D, Q767E, L747I, I678K, A598S, D548E, R420K, N351D, A311V, V299I, I231L, A186V, D264G, D559N, K233R, V95I, T51I	×
xdhA	Metronidazole resistance	S146N, S694G, L601R, D567N, T550I, V541L, M417I, R399H, I726M, I640V, D355E	×
rpoB	Rifampicin/fidaxomicin resistance	H502N[e,f], R505K [e,f], I750M, Q1037E, V1205A, A1207D[f], A1207T[f], T1208A[f], E1232D	D492E, R505K
rpoC	Rifampicin/fidaxomicin resistance	L833I	×
gdpP	Vancomycin resistance	×	×
murG	Vancomycin resistance	G236E, D158E	×
vanB	Vancomycin resistance	-[g]	×
vanR	Vancomycin resistance	S2N, I117L	T115A
vanRF	Vancomycin resistance	×	×
vanRG	Vancomycin resistance	D45N	×
vanRG2	Vancomycin resistance	–	×
vanS	Vancomycin resistance	K17E, N179D	×
vanSG	Vancomycin resistance	–	×
vanTG	Vancomycin resistance	–	×
vanU	Vancomycin resistance	×	×

[a]The mutation was only detected in ST1-A2; [b]The mutations were only detected in ST1-A10, [c]The mutations were only detected in ST1-A9, ST1-A14, [d]No mutations were detected, [e]The mutations were only detected in ST37-B2, ST37-B4 and ST37-B7; [f]The mutation was not detected in ST37-B5; ST1-A12 and ST1-A13, [g]The gene is absent in ST37.

phenotypic analysis. Genomic analysis revealed four ST37-specific genes known to be associated with *C. difficile* adhesion, as well as significant differences in virulence-related genes between ST37 and ST1 isolates. We also compared the *in vitro* pathogenicity-related phenotypes of ST37 and ST1 isolates and found that ST37 isolates exhibited stronger capabilities in aggregation, biofilm formation, and surface motility, while ST1 isolates exhibited superior TcdB production and spore formation capabilities.

In this study, we observed that ST37 isolates exhibit a notable characteristic of aggregative growth, suggesting that ST37 isolates may possess a strong capacity for inter-bacterial adhesion. Adhesion is a critical factor influencing the colonization and survival of *C. difficile* in the host's intestinal [2,45]. Adhesion encompasses not only the attachment to host intestinal epithelial cells but also the adherence between individual bacterial cells [45]. Adhesion to intestinal epithelial cells facilitates successful colonization within the gut, making it less likely for the bacteria to be expelled with the intestinal contents [46]. Adhesion between bacterial cells aids in the rapid establishment of communities, enabling them to collectively resist adverse environmental influences [46]. *C. difficile* face various survival pressures in the host intestine, such as antibiotic exposure and nutrient and niche competition from other intestinal bacteria, which makes it difficult for a single *C. difficile* to survive [47]. In this case, the same or a variety of bacteria use extracellular polymers as the matrix, and adhere to and cross-link under the mediation of extracellular accessory structures such as capsule, flagella and type IV pili, forming the phenomenon of aggregation growth [48]. Aggregation is not only beneficial for *C. difficile* to obtain more nutrients, more importantly, cross-linking between aggregated bacteria facilitates the rapid transmission of virulence genes and drug resistance genes in the colony, further endowing *C. difficile* with stronger resistance and pathogenicity [35,36].

Greater biofilm formation ability is another survival and pathogenic advantage of ST37 [48]. Biofilms are infectious complexes formed by the adhesion of bacteria to hydrated extracellular matrices such as proteins, polysaccharides, and extracellular DNA [49]. Biofilm formation not only provides nutrition and a suitable growth environment for *C. difficile*, but it also protects it from drug toxicity and host immune system attack [49]. Biofilms are extremely difficult to clear via drugs, and biofilm-related CDI is prone to developing into chronic and recurrent infections [50]. Most CDI patients are immunocompromised older patients with underlying disease, and if they are infected with a strong biofilm-forming *C. difficile* such as ST37, they are likely to have

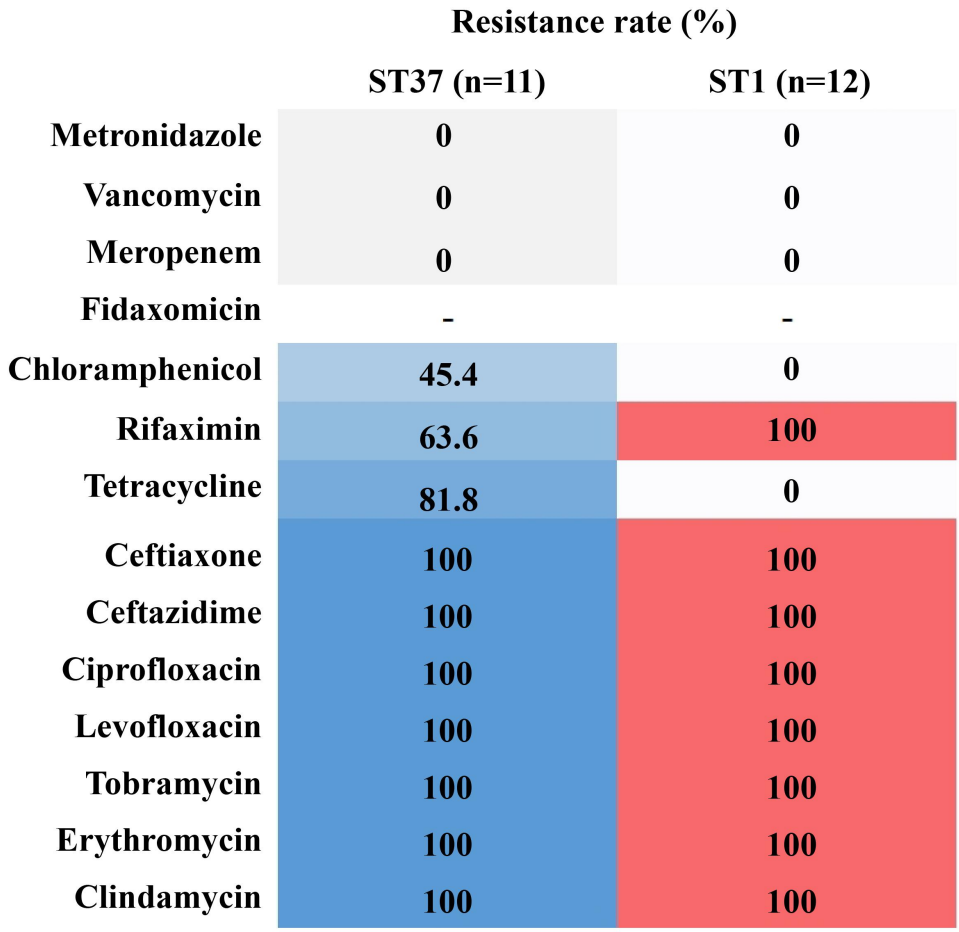

Figure 5. Comparison of antibiotic resistance rates between ST37 and ST1 strains.

more severe infection symptoms, worse prognosis, and higher recurrence rate. The strong biofilm formation ability of ST37 has been reported. Rahmoun et al. [51] compared the biofilm production capacity of isolates from different STs such as ST1, ST2, ST37 and ST42, and found that most (44.4%) isolates of ST37 were strong biofilm producers, while most (42.9%) isolates of ST1 were weak/moderate biofilm producers.

The strong adhesion and biofilm formation ability of ST37 isolates may be related to its four specific adhesion proteins, including LytC (CWP84), CbpA (CD2831), CD3246 and SrtB, as well as proteins containing GGDEF domain. Dawson et al. [52] have reported that cell surface proteins CD2831 (CbpA) and CD3246 can promote biofilm formation in *C. difficile*. SrtB and LytC (CWP84) catalyze the binding of collagen binding protein CbpA to cell wall peptidoglycan and the maturation of surface protein SlpA, respectively [52]. Interestingly, we also detected three putative DGCs containing GGDEF domain that can catalyze the synthesis of c-di-GMP, and the expression of CD2831 (CbpA) is increased in the presence of high levels of c-di-GMP, and *C. difficile* adhesion and biofilm formation are also promoted [48,52,53]. Unfortunately, we did not test the expression levels of the above four proteins and the concentration of c-di-GMP in ST37 and ST1 aggregates and biofilms, so the specific virulence factors and mechanisms responsible for the strong aggregation and biofilm formation of ST37 need to be further investigated. Nonetheless, our research supports the notion that prioritizing drugs or strategies that inhibit bacterial aggregation and biofilm formation may be more effective in treating CDI caused by ST37.

Surface motion on BHIS-1.8% agar plates reflects the motility of *C. difficile* on the surface of intestinal epithelial cells [54]. The elastic modulus of healthy intestinal tissue is about 10 kPa, corresponding to 1% ~ 2% agar percentages. Inflammation caused by CDI thickened the intestinal tract from 6–8 mm to 22–32 mm, and the elastic modulus increased to about 25 kPa [54]. Inflammatory intestinal tissues may facilitate the rapid migration of ST37 to nutrient-rich niche and then multiply, facilitating early dissemination of infection. Therefore, stronger surface motility may also play an important role in the pathogenesis of some ST37 strains.

Sporulation capacity is an important factor affecting the survival and environmental transmission of *C. difficile in vitro* and *in vivo*, and hospital-acquired CDIs are associated with transmission of *C. difficile* spores in the hospital setting [55]. The 29 sporulation-related genes with non-synonymous mutations that we identified in ST37 strains did not appear to confer greater sporulation ability, as ST37 strains formed significantly fewer spores than ST1 strains *in vitro*. Previous studies have also reported that ST1 has stronger sporulation ability than non-ST1 strains [8,56]. Hence, as the most successful epidemic strains, it is not surprising that ST1 has a strong sporulation ability.

The evolution of AMR plays a crucial role in the pathogenesis and spread of *C. difficile*. AMR not only can lead to the failure of CDI treatment, but also facilitates the colonization of resistant strains in the gut, which, through the formation of spores and/or biofilms, can easily cause recurrent CDI [57]. The persistent dissemination of resistant spores in healthcare and/or community settings significantly contributes to the increased prevalence of CDI. Multidrug resistance (MDR) (resistant to at least three or more antibiotics) is a common feature of ST37 and ST1 strains. Previous studies suggested that many outbreaks of ST37 are related to clindamycin and fluoroquinolone resistance and that AMR and selective stress may be responsible for the successful epidemic of ST37 [11,12]. In the present study, we found the differences between ST37 and ST1 strains in resistance to chloramphenicol, tetracycline and rifaximin. The resistance of ST37 isolates to chloramphenicol and tetracycline may be associated with the *catQ* and *tetM* genes, respectively, as both genes were detected in 6 ST37 isolates, whereas they were absent in all 12 susceptible ST1 isolates. Notably, the presence of *catQ* did not completely correspond to the resistance of ST37 to chloramphenicol, because 4 of the 6 sequenced ST37 strains were sensitive to chloramphenicol, and we also did not detect mutation in the *catQ* gene. Chloramphenicol acetyltransferase, encoded by the *cat* gene family, can catalyze the transfer of the acetyl group of acetyl-CoA to the 3-hydroxyl group of chloramphenicol, resulting in the loss of antibacterial activity of chloramphenicol. Study reported that the resistance of *C. difficile* to chloramphenicol is mostly related to the *catP* gene, but we did not detect any other *cat* genes except *catQ* [58].

The high prevalence of ST37 in non-developed regions such as Asia and the high resistance rate to chloramphenicol and tetracycline may be related to the extensive use of these two antibiotics in these regions. In some developed countries, the use of chloramphenicol is severely restricted due to side effects and drug resistance [59]. However, chloramphenicol is still widely used in some developing countries due to its high cost-effectiveness. Especially in some areas where medical resources are relatively scarce, chloramphenicol is an important choice for the treatment of infectious diseases [60].

The high resistance of *C. difficile* to rifaximin is another point of concern, and previous studies have reported the high resistance of ST1 to rifaximin [61]. Rifaximin exerts its antibacterial effect by inhibiting bacterial RNA polymerase, with the β-subunit encoded by the *rpoB* gene being one of the primary structures of RNA polymerase [62]. Mutations in the rifamycin resistance-determining region of RpoB (such as Ser488Tyr, Asp492Tyr, His502Asn/Tyr/Asp, and Arg505Lys) disrupt the interaction between rifaximin and RpoB, thereby reducing the affinity of rifaximin for its target protein and leading to decreased bacterial susceptibility to rifaximin [62]. In this study, we detected the R505K (Arg505Lys) mutation in RpoB in all 12 rifaximin-resistant ST1 strains. Among the 6 sequenced ST37 strains, this mutation was only detected in 3 rifaximin-resistant strains. We also detected the R505K mutation in RpoB in four rifaximin-resistant unsequenced ST37 isolates by Sanger sequencing. Overall, our study further confirms the importance of the R505K mutation in RpoB in *C. difficile* resistance to rifaximin.

Metronidazole and vancomycin have been the first-line therapies for CDI for many years. In this study, we noted that the sensitivity of ST1 strain to metronidazole and vancomycin was lower than that of ST37 strain. The decreased sensitivity of ST1 strains to metronidazole may be related to the Y130S mutation in NimA, and this mutation was not detected in the sequenced ST37 strains. Arcay et al. [63] detected this mutation in two metronidazole-resistant *C. difficile* strains. The I155L point mutation in NimA was detected in both ST37 and ST1 strains, but the relationship between this mutation and metronidazole resistance has not been reported. The reduced sensitivity of ST1 strains to vancomycin may be related to the T115A mutation in VanR, which is consistent with the mutation site reported by Shen et al. [64]. The ST1 strains had four more genes related to vancomycin resistance than the ST37 strains, which may also be one of the reasons for its lower sensitivity to vancomycin. Taken together, although the ST37 and ST1 strains in this study were sensitive to metronidazole and vancomycin, the presence of antibiotic resistance genes and the detection of mutant sites warned us of the importance of continuously monitoring the prevalence and resistance spectrum of *C. difficile*.

In summary, our results further confirmed the MDR of ST37 and ST1. Therefore, we need to accelerate the

standardization of antibiotic use and management to prevent and control the prevalence and spread of ST37 on a wider scale and in non-developed regions, and especially in Asia where ST37 is the dominant strain.

Conclusions

In conclusion, we found genomic differences in type-specific genes, virulence-related genes, and drug resistance-related genes between ST37 and ST1 isolate. ST37 isolates showed stronger aggregation, biofilm formation and surface motility than ST1 isolates, and had higher resistance rates to chloramphenicol and tetracycline. ST1 isolate showed stronger ability to produce toxin and sporulation, and was highly resistant to rifaximin.

Ethics approval and consent to participate

The protocol has been reviewed by the Ethics Committee (IRB) of The Second Hospital of Hebei Medical University. All C. difficile strains are from strains cultured from residual samples used in clinical diagnosis, it involves the confidentiality of patient data and compliance with the Declaration of Helsinki. Since the data did not affect patient care, after consulting the IRB of The Second Hospital of Hebei Medical University, a formal ethical review was approved, and written informed consent from the patient was not required (ethical approval number 2021-R521).

Acknowledgements

We would like to thank all members of the study team the patients who provided stool samples.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Hebei Provincial Government Subsidized Clinical Medicine Excellence Training Program [ZF2025084], Hebei Provincial Natural Science Foundation [H2022206358 and H2022206549], Hebei County General Hospital Appropriate Health Technology Promotion Project [20200018].

Authors' contributions

Zirou Ouyang, investigation, methodology, visualization, writing – original draft, writing – review and editing. Jing Yang, methodology, writing – review and editing. Huimin Zhang, visualization, writing – review and editing. Yaxuan Yang and Min Zhao, investigation. Huimin Yang and Jiafeng Zhao, formal analysis. Cuixin Qiang, Zhirong Li, Pu Qin, Weigang Wang, and Yanan Niu, software. Jianhong Zhao, funding acquisition, supervision. Jianhong Zhao final approval of the version to be published and offered professional guidance. All authors were involved in editing the manuscript. All authors have read and approved the final work.

Data availability statement

The dataset generated from this study has been deposited in the NCBI database under the BioProject number PRJNA1160995. The genomes of *C. difficile* R20291 (accession number ERR047167) can be found in the NCBI database. The data can also be obtained from the Science Data Bank database (https://doi.org/10.57760/sciencedb.18394). The supplementary tables and figures can also be obtained from the Science Data Bank database (https://doi.org/10.57760/sciencedb.24044).

ORCID

Jianhong Zhao http://orcid.org/0000-0002-7368-7961

References

[1] Park SO, Yeo I. Trends in *Clostridioides difficile* prevalence, mortality, severity, and age composition during 2003–2014, the national inpatient sample database in the US. Ann Med. 2022;54(1):1851–1858. doi: 10.1080/07853890.2022.2092893

[2] Buddle JE, Fagan RP. Pathogenicity and virulence of *Clostridioides difficile*. Virulence. 2023;14(1):2150452. doi: 10.1080/21505594.2022.2150452

[3] Yuille S, Mackay WG, Morrison DJ, et al. Drivers of *Clostridioides difficile* hypervirulent ribotype 027 spore germination, vegetative cell growth and toxin production in vitro. Clin Microbiol Infect. 2020;26(7):. e941.1–.e.7. doi: 10.1016/j.cmi.2019.11.004

[4] Kochan TJ, Shoshiev MS, Hastie JL, et al. Germinant synergy facilitates *Clostridium difficile* spore germination under physiological conditions. mSphere. 2018;3(5). doi: 10.1128/mSphere.00335-18

[5] Chandrasekaran R, Lacy DB. The role of toxins in *Clostridium difficile* infection. FEMS Microbiol Rev. 2017;41(6):723–750. doi: 10.1093/femsre/fux048

[6] Stabler RA, He M, Dawson L, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102. doi: 10.1186/gb-2009-10-9-r102

[7] Dupuy B, Govind R, Antunes A, et al. *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. J Med Microbiol. 2008;57(6):685–689. doi: 10.1099/jmm.0.47775-0

[8] Burns DA, Heeg D, Cartman ST, et al. Reconsidering the sporulation characteristics of hypervirulent *Clostridium difficile* BI/NAP1/027. PLOS ONE. 2011;6(9):e24894. doi: 10.1371/journal.pone.0024894

[9] Xu X, Luo Y, Chen H, et al. Genomic evolution and virulence association of *Clostridioides difficile* sequence type 37 (ribotype 017) in China. Emerg Microbes Infect. 2021;10(1):1331–1345. doi: 10.1080/22221751.2021.1943538

[10] Imwattana K, Knight DR, Kullin B, et al. *Clostridium difficile* ribotype 017 – characterization, evolution and epidemiology of the dominant strain in Asia. Emerg Microbes Infect. 2019;8(1):796–807. doi: 10.1080/22221751.2019.1621670

[11] Imwattana K, Knight DR, Kullin B, et al. Antimicrobial resistance in *Clostridium difficile* ribotype 017. Expert Rev Anti Infect Ther. 2020;18(1):17–25. doi: 10.1080/14787210.2020.1701436

[12] Imwattana K, Putsathit P, Collins DA, et al. Global evolutionary dynamics and resistome analysis of *Clostridioides difficile* ribotype 017. Microb Genom. 2022;8(3). doi: 10.1099/mgen.0.000792

[13] Li R, Xiao D, Yang J, et al. Identification and characterization of *clostridium difficile* sequence type 37 genotype by matrix-assisted laser desorption ionization – time of flight mass spectrometry. J Clin Microbiol. 2018;56(5). doi: 10.1128/JCM.01990-17

[14] Li Z, Ouyang Z, Zhang H, et al. Novel target and PCR assay for identification of hypervirulent ST1 (BI/NAP1/027) Clostridioides difficile and detection of toxigenic C. Difficile. Clinica (Rome) Acta. 2024;559:119728. doi: 10.1016/j.cca.2024.119728

[15] Cui QQ, Yang J, Niu YN, et al. Epidemiological investigation of *Clostridioides difficile* colonization in Chinese community infants. Anaerobe. 2019;56:116–123. doi: 10.1016/j.anaerobe.2019.01.008

[16] Griffiths D, Fawley W, Kachrimanidou M, et al. Multilocus sequence typing of *Clostridium difficile*. J Clin Microbiol. 2010;48(3):770–778. doi: 10.1128/JCM.01796-09

[17] Lienhard A, Schäffer S. Extracting the invisible: obtaining high quality DNA is a challenging task in small arthropods. PeerJ. 2019;7:e6753. doi: 10.7717/peerj.6753

[18] Bankevich A, Nurk S, Antipov D, et al. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–477. doi: 10.1089/cmb.2012.0021

[19] Coil D, Jospin G, Darling AE. A5-miseq: an updated pipeline to assemble microbial genomes from illumina MiSeq data. Bioinformatics. 2015;31(4):587–589. doi: 10.1093/bioinformatics/btu661

[20] Walker BJ, Abeel T, Shea T, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLOS ONE. 2014;9(11):e112963. doi: 10.1371/journal.pone.0112963

[21] Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 2001;29(12):2607–2618. doi: 10.1093/nar/29.12.2607

[22] Lowe TM, Eddy SR. tRnascan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25(5):955–964. doi: 10.1093/nar/25.5.955

[23] Kalvari I, Argasinska J, Quinones-Olvera N, et al. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res. 2018;46(D1):D335–D342. doi: 10.1093/nar/gkx1038

[24] Chen L, Yang J, Yu J, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 2005;33(Database issue):D325–8. doi: 10.1093/nar/gki008

[25] McArthur AG, Waglechner N, Nizam F, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother. 2013;57(7):3348–3357. doi: 10.1128/AAC.00419-13

[26] Blake JD, Cohen FE. Pairwise sequence alignment below the twilight zone. J Mol Biol. 2001;307(2):721–735. doi: 10.1006/jmbi.2001.4495

[27] Moriya Y, Itoh M, Okuda S, et al. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 2007;35(Web Server):W182–5. doi: 10.1093/nar/gkm321

[28] Conesa A, Götz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. 2008;2008:1–12. doi: 10.1155/2008/619832

[29] Powell S, Forslund K, Szklarczyk D, et al. eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res. 2014;42(D1):D231–9. doi: 10.1093/nar/gkt1253

[30] Chen F, Mackey AJ, Stoeckert CJ, et al. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res. 2006;34(90001):D363–8. doi: 10.1093/nar/gkj123

[31] McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–1303. doi: 10.1101/gr.107524.110

[32] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164. doi: 10.1093/nar/gkq603

[33] Price MN, Dehal PS, Arkin AP, et al. FastTree 2–approximately maximum-likelihood trees for large alignments. PLOS ONE. 2010;5(3):e9490. doi: 10.1371/journal.pone.0009490

[34] Wang W, Cao J, Yang J, et al. Antimicrobial activity of tannic acid in vitro and its protective effect on mice against *Clostridioides difficile*. Microbiol Spectr. 2023;11(1):e0261822. doi: 10.1128/spectrum.02618-22

[35] Bordeleau E, Purcell EB, Lafontaine DA, et al. Cyclic di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. J Bacteriol. 2015;197(5):819–832. doi: 10.1128/JB.02340-14

[36] Purcell EB, McKee RW, McBride SM, et al. Cyclic diguanylate inversely regulates motility and aggregation in *clostridium difficile*. J Bacteriol. 2012;194(13):3307–3316. doi: 10.1128/JB.00100-12

[37] Garrett EM, Sekulovic O, Wetzel D, et al. Phase variation of a signal transduction system controls *Clostridioides difficile* colony morphology, motility, and virulence. PLOS Biol. 2019;17(10):e3000379. doi: 10.1371/journal.pbio.3000379

[38] Tian TT, Zhao JH, Yang J, et al. Molecular characterization of *Clostridium difficile* isolates from human subjects and the environment. PLOS ONE. 2016;11(3):e0151964. doi: 10.1371/journal.pone.0151964

[39] Cheng JW, Liu C, Kudinha T, et al. The tcdA-negative and tcdB-positive *Clostridium difficile* ST81 clone

exhibits a high level of resistance to fluoroquinolones: a multi-centre study in Beijing, China. Int J Antimicrob Agents. 2020;56(1):105981. doi: 10.1016/j.ijantimicag.2020.105981

[40] Bartlett JG, Chang TW, Onderdonk AB. Will the real Clostridium species responsible for antibiotic-associated colitis please step forward? Lancet. 1978;1(8059):338. doi: 10.1016/S0140-6736(78)90118-6

[41] Al-Barrak A, Embil J, Dyck B, et al. An outbreak of toxin a negative, toxin B positive *Clostridium difficile*-associated diarrhea in a Canadian tertiary-care hospital. Can Commun Dis Rep. 1999;25(7):65–69.

[42] Goorhuis A, Debast SB, Dutilh JC, et al. Type-specific risk factors and outcome in an outbreak with 2 different *Clostridium difficile* types simultaneously in 1 hospital. Clin Infect Dis. 2011;53(9):860–869. doi: 10.1093/cid/cir549

[43] Kim J, Kim Y, Pai H, et al. Clinical characteristics and treatment outcomes of *Clostridium difficile* infections by PCR ribotype 017 and 018 strains. PLOS ONE. 2016;11(12):e0168849. doi: 10.1371/journal.pone.0168849

[44] Arvand M, Hauri AM, Zaiss NH, et al. *Clostridium difficile* ribotypes 001, 017, and 027 are associated with lethal *C. difficile* infection in Hesse, Germany. Euro Surveill. 2009;14(45). doi: 10.2807/ese.14.45.19403-en

[45] Stewart D, Anwar F, Vedantam G. Anti-virulence strategies for *Clostridioides difficile* infection: advances and roadblocks. Gut Microbes. 2020;12(1):1802865. doi: 10.1080/19490976.2020.1802865

[46] McKee RW, Aleksanyan N, Garrett EM, et al. Type IV pili promote *clostridium difficile* adherence and persistence in a mouse model of infection. Infect Immun. 2018;86(5):e00943. doi: 10.1128/IAI.00943-17

[47] Marshall A, McGrath JW, Graham R, et al. Food for thought—the link between *Clostridioides difficile* metabolism and pathogenesis. PLOS Pathog. 2023;19(1):e1011034. doi: 10.1371/journal.ppat.1011034

[48] Dawson LF, Peltier J, Hall CL, et al. Extracellular DNA, cell surface proteins and c-di-GMP promote biofilm formation in Clostridioides difficile. Sci Rep. 2021;11(1):3244. doi: 10.1038/s41598-020-78437-5

[49] Taggart MG, Snelling WJ, Naughton PJ, et al. Biofilm regulation in *Clostridioides difficile*: novel systems linked to hypervirulence. PLOS Pathog. 2021;17(9):e1009817. doi: 10.1371/journal.ppat.1009817

[50] Tijerina-Rodríguez L, Villarreal-Treviño L, Baines SD, et al. High sporulation and overexpression of virulence factors in biofilms and reduced susceptibility to vancomycin and linezolid in recurrent Clostridium [Clostridioides] difficile infection isolates. PLOS ONE. 2019;14(7):e0220671. doi: 10.1371/journal.pone.0220671

[51] Rahmoun LA, Azrad M, Peretz A. Antibiotic resistance and biofilm production capacity in *Clostridioides difficile*. Front Cell Infect Microbiol. 2021;11:683464. doi: 10.3389/fcimb.2021.683464

[52] Ouyang Z, Zhao M, Li J, et al. Cyclic diguanylate differentially regulates the expression of virulence factors and pathogenesis-related phenotypes in *Clostridioides difficile*. Microbiol Res. 2024;286:127811. doi: 10.1016/j.micres.2024.127811

[53] Arato V, Gasperini G, Giusti F, et al. Dual role of the colonization factor CD2831 in Clostridium difficile pathogenesis. Sci Rep. 2019;9(1):5554. doi: 10.1038/s41598-019-42000-8

[54] Purcell EB, McKee RW, Bordeleau E, et al. Regulation of type IV pili contributes to surface behaviors of historical and epidemic strains of clostridium difficile. J Bacteriol. 2015;198(3):565–577. doi: 10.1128/JB.00816-15

[55] Srikhanta YN, Hutton ML, Awad MM, et al. Cephamycins inhibit pathogen sporulation and effectively treat recurrent *Clostridioides difficile* infection. Nat Microbiol. 2019;4(12):2237–2245. doi: 10.1038/s41564-019-0519-1

[56] Kerlund T, Persson I, Unemo M, et al. Increased sporulation rate of epidemic *Clostridium difficile* type 027/NAP1. J Clin Microbiol. 2008;46(4):1530–1533. doi: 10.1128/JCM.01964-07

[57] Sholeh M, Krutova M, Forouzesh M, et al. Antimicrobial resistance in *Clostridioides* (*Clostridium*) *difficile* derived from humans: a systematic review and meta-analysis. Antimicrob Resist Infect Control. 2020;9(1):158. doi: 10.1186/s13756-020-00815-5

[58] Imwattana K, Rodríguez C, Riley TV, et al. A species-wide genetic atlas of antimicrobial resistance in *Clostridioides difficile*. Microb Genom. 2021;7(11):000696. doi: 10.1099/mgen.0.000696

[59] Qu X, Yin C, Sun X, et al. Consumption of antibiotics in Chinese public general tertiary hospitals (2011–-2014): trends, pattern changes and regional differences. PLOS ONE. 2018;13(5):e0196668. doi: 10.1371/journal.pone.0196668

[60] Browne AJ, Chipeta MG, Haines-Woodhouse G, et al. Global antibiotic consumption and usage in humans, 2000–18: a spatial modelling study. Lancet Planetary Health. 2021;5(12):e893–e904. doi: 10.1016/S2542-5196(21)00280-1

[61] Kabała M, Gofron Z, Aptekorz M, et al. *Clostridioides difficile* ribotype 027 (RT027) outbreak investigation due to the emergence of rifampicin resistance using multilocus variable-number tandem repeat analysis (MLVA). Infect Drug Resist. 2021;14:3247–3254. doi: 10.2147/IDR.S324745

[62] O'Connor JR, Galang MA, Sambol SP, et al. Rifampin and rifaximin resistance in clinical isolates of *Clostridium difficile*. Antimicrob Agents Chemother. 2008;52(8):2813–2817. doi: 10.1128/AAC.00342-08

[63] Arcay RM, Suárez-Bode L, López-Causapé C, et al. Emergence of high-level and stable metronidazole resistance in Clostridioides difficile. Int J Antimicrob Agents. 2020;55(1):105830. doi: 10.1016/j.ijantimicag.2019.10.011

[64] Shen WJ, Deshpande A, Hevener KE, et al. Constitutive expression of the cryptic vanGcd operon promotes vancomycin resistance in Clostridioides difficile clinical isolates. J Antimicrob Chemother. 2020;75(4):859–867. doi: 10.1093/jac/dkz513