Clinical Infectious Diseases

MAJOR ARTICLE

Clostridioides difficile Whole-genome Sequencing Differentiates Relapse With the Same Strain From Reinfection With a New Strain

Janice Cho,[1] Scott Cunningham,[2] Meng Pu,[1] Ryan J. Lennon,[3] Jennifer Dens Higano,[4] Patricio Jeraldo,[5] Priya Sampathkumar,[6] Samantha Shannon,[2] Purna C. Kashyap,[1,a] and Robin Patel[2,6,a]

[1]Division of Gastroenterology, Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA, [2]Division of Clinical Microbiology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA, [3]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA, [4]Mayo Clinic Alix School of Medicine, Rochester, Minnesota, USA, [5]Division of Surgical Research, Department of Surgery, Mayo Clinic, Rochester, Minnesota, USA, and [6]Division of Infectious Diseases, Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA

Background. Current approaches in tracking *Clostridioides difficile* infection (CDI) and individualizing patient management are incompletely defined.

Methods. We recruited 468 subjects with CDI at Mayo Clinic Rochester between May and December 2016 and performed whole-genome sequencing (WGS) on *C. difficile* isolates from 397. WGS was also performed on isolates from a subset of the subjects at the time of a recurrence of infection. The sequence data were analyzed by determining core genome multilocus sequence type (cgMLST), with isolates grouped by allelic differences and the predicted ribotype.

Results. There were no correlations between *C. difficile* isolates based either on cgMLST or ribotype groupings and CDI outcome. An epidemiologic assessment of hospitalized subjects harboring *C. difficile* isolates with ≤2 allelic differences, based on standard infection prevention and control assessment, revealed no evidence of person-to-person transmission. Interestingly, community-acquired CDI subjects in 40% of groups with ≤2 allelic differences resided within the same zip code. Among 18 subjects clinically classified as having recurrent CDI, WGS revealed 14 with initial and subsequent isolates differing by ≤2 allelic differences, suggesting a relapse of infection with the same initial strain, and 4 with isolates differing by >50 allelic differences, suggesting reinfection. Among the 5 subjects classified as having a reinfection based on the timing of recurrence, 3 had isolates with ≤2 allelic differences between them, suggesting a relapse, and 2 had isolates differing by >50 allelic differences, suggesting reinfection.

Conclusions. Our findings point to potential transmission of *C. difficile* in the community. WGS better differentiates relapse from reinfection than do definitions based on the timing of recurrence.

Keywords. *Clostridioides difficile*; whole-genome sequencing; *Clostridium difficile*; ribotype; clinical outcomes.

Clostridioides difficile infection (CDI) is the leading cause of health care–associated diarrhea. Health care–associated CDI increases the cost of hospitalization, in some cases up to 4-fold [1, 2].

Clostridioides difficile exhibits genetic diversity that can possibly be leveraged to track its spread [3, 4]. With the advent of next-generation sequencing, whole-genome sequencing (WGS) is increasingly used as a fingerprinting method, allowing precise tracking of transmission and providing a tool to identify and help control outbreaks [5–7]. WGS may also detect genes that might contribute to particular outcomes, alongside strains that are more likely to spread in health-care settings. Specific *C. difficile* strains may be associated with clinical severity [4, 8–12] and/or antibiotic resistance [9, 13]. That being said, specific genotypes are not always clearly associated with disease severity; for example, by a lack of correlation between disease severity and levels of toxin production [14].

CDI can be challenging to treat, given its propensity for recurrence following treatment. But another episode of CDI is not always a recurrence, as infection can be acquired *de novo*. Current definitions of recurrence leverage differential timing of symptom onset and do not include strain information. Likewise, strain information is not part of treatment strategies, in part due to a lack of availability and to the cost and turnaround time, but also as a result of a lack of data supporting clinical utility.

The aim of this study was to determine the utility of WGS in evaluating transmission patterns of *C. difficile*, differentiating relapses versus reinfections in subjects with recurrent CDI and defining those strains correlated with specific clinical outcomes. We evaluated 468 subjects with *C. difficile*–positive tests by polymerase chain reaction (PCR) who were seen at Mayo Clinic

Received 22 August 2019; editorial decision 6 February 2020; accepted 14 February 2020; published online February 17, 2020.

[a]P. C. K. and R. P. contributed equally to this work.

Correspondence: R. Patel, Division of Clinical Microbiology, Division of Infectious Diseases, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. (patel.robin@mayo.edu).

Clinical Infectious Diseases® 2021;72(5):806–13

DOI: 10.1093/cid/ciaa159

Rochester from May to December 2016. We analyzed the host factors contributing to outcomes. *Clostridioides difficile* was isolated and subjected to WGS. Strain type, determined by a core genome MLST (cgMLST) analysis, was used to assess potential person-to-person transmission and was correlated with outcomes. A subset of subjects had more than 1 detection of *C. difficile* over time; their initial and subsequent isolates were analyzed by WGS. The data represent 1 of the largest cohorts of genetically analyzed *C. difficile* from a single center.

METHODS

Subject Recruitment

This study was approved by the Mayo Clinic Institutional Review Board. Subjects 18 years and older who had CDI, as evidenced by a positive *C. difficile* toxin PCR and diarrhea, between May and December 2016 at Mayo Clinic Rochester were included. Subjects who provided no research authorization or from whom stool samples were unavailable were excluded.

Isolation of *Clostridioides difficile* strains

Samples positive for *C. difficile* by BioFire FilmArray Gastrointestinal Panel (BioFire, Salt Lake City, UT) or a laboratory-developed PCR assay targeting *tcdC* [15] were reflexively cultured to CHROMagar *C. difficile* media and incubated for 24 hours at 35–37°C anaerobically. Culture plates were examined with a Wood's ultraviolet lamp (365 nm), and colonies showing fluorescence were Gram stained and subcultured to Centers for Disease Control and Prevention sheep blood agar and CHROMagar *C. difficile* media. The identification of *C. difficile* was based on growth on CHROMagar *C. difficile* medium, colony and Gram stain morphology, and a positive quick indole reaction. There were 428 samples that grew isolates identified as *C. difficile*.

DNA was extracted using a Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research, Irvine, CA) following the manufacturer's protocol, except that material (approximately 10 µg loop-full) from the subculture plate was directly inoculated into a lysis tube and elution buffer preheated to 60°C prior to the elution step. Eluted DNA was quantified with a Quantus fluorometer using the Quantifluor ONE dsDNA system (Promega, Madison, WI).

Whole-genome Sequencing and Ribotyping of *Clostridioides difficile* Strains

Paired-end sequencing libraries were created with a NEBNext Ultra DNA Library Prep Kit (New England Biolabs, Ipswich, MA) with the following specifications: a fragmentation target size of 500 bp, 1:5 adapter dilution, NEBNext Ultra II Q5 PCR mix, 8 PCR cycles, and dual 0.6X AMPure XP bead (Beckman Coulter, Indianapolis, IN) clean-up. Libraries were pooled to a maximum of 80 samples per lane (150X target coverage) on a HiSeq 2500 (Illumina, San Diego, CA) using V2 PE250 chemistry in rapid run mode. The manufacturer's minimum quality metrics for the HiSeq Rapid SBS Kit V2 500 cycle kit were applied.

Raw FASTQ read files were adapter- and quality-trimmed with Atropos version 1.1.16 [16, 17] and fed into SeqSphere+ version 5.1 (Ridom, Muenster, DE) for *de novo* assembly, cgMLST analysis, and the construction of a circular cladogram with the nearest neighbor-joining method in Geneious Prime (2019.0.4) [18, 19]. We evaluated 4 relatedness thresholds (ie, number of allelic differences between isolates; 0, ≤2, ≤7 and ≤50 allelic differences). The PCR ribotype was predicted by comparing reference strains (Supplementary Table 1) with clinical isolates. Clusters and isolates in close allelic proximity (≤150 allelic differences) to PCR ribotype reference strains were predicted to match the reference strain's ribotype.

Clinical Data Abstraction and Statistical Analysis

The frequency of previously described risk factors for susceptibility to CDI was studied, including current hospitalization, recent hospitalization (within 12 weeks), recent antibiotic use (within 28 days), prior CDI history, and immunosuppressed state [20, 21]. Severe CDI was defined as a white blood cell count ≥15 000 cells/mm^3, creatinine >1.5 mg/dL, or albumin <3 g/dL [22]. Severe CDI with complications was defined as being in an intensive care setting, hypotension, white blood cell count ≥35 000 or <2000 cells/mm^3, serum lactate >2.2 mmol/L, end organ failure, megacolon, or the need for a colectomy. Colonization was defined as a positive *C. difficile* toxin PCR and no symptoms [22]. Community-acquired CDI was defined as a positive *C. difficile* test within 72 hours of admission or as an outpatient with no health-care exposure in the antecedent 12 weeks. Long-term care facility–acquired CDI was defined as a positive *C. difficile* test within 72 hours of admission in a subject who resided in a long-term care facility prior to admission. Hospital-acquired CDI was defined as a positive *C. difficile* test after 72 hours of admission.

Continuous variables were summarized with means (standard deviation) and medians (Q1, Q3), unless otherwise indicated. Discrete variables were summarized as frequencies (percentages). We used 2-sample t-tests and Pearson chi-squared tests to assess differences between subjects with community-acquired versus hospital-acquired CDI. Logistic regression was used to estimate the odds ratios for associations between subject characteristics and binary disease and treatment outcomes. All hypothesis tests were 2-sided with .05 significance levels. Analyses were conducted using R software version 3.4.2 (Vienna, Austria).

A subset of patients underwent repeat testing for *C. difficile* within the study period. If subjects had diarrhea recur within 8 weeks of their initial test, they were classified as having a recurrence. If subjects had diarrhea beyond 8 weeks of their initial test, they were classified as having reinfection [22]. Clinical data correlating to the sample collected from the subject's initial infection are shown in Tables 1–4. In subjects who underwent repeat testing, *C. difficile* was isolated and

WGS performed to determine whether the repeat detection reflected a relapse with the same strain (defined as isolates differing by ≤2 allelic differences) or infection with a new strain (defined as isolates differing by >50 allelic differences). Some samples did not grow an isolate and, thus, those subjects were excluded. Also, for samples collected within 24 hours of the initial sample collection, the subjects were excluded for redundancy.

RESULTS

Subject Characteristics

We analyzed 468 subjects with *C. difficile* to identify the host factors contributing to outcomes and performed a cgMLST analysis on isolates from 397 subjects. We also performed a subset analysis on 26 subjects at the time of recurrence or reinfection (Figure 1). Among our cohort, 45.9% were male, with the median age being 58 years (interquartile range [IQR], 35–72 years) and 93.3% identifying as Caucasian. The median body mass index was 25.9 (IQR, 21.7–30.6). See Table 1 for detailed characteristics.

Subject Characteristics and *Clostridioides difficile* Infection Outcomes

The frequencies of previously described risk factors for CDI are outlined in Table 2 and clinical outcomes are shown in Table 3. By univariate analysis, severe CDI was significantly associated with being aged >65 years, antibiotic use within 4 weeks, current hospitalization, prior hospitalization within 12 weeks, and an immunosuppressed state, while female sex, a history of inflammatory bowel disease, and community-acquired CDI were negatively correlated with severe CDI. Current and recent hospitalizations were associated with severe complications. Age >65 years and current hospitalization were associated with the composite outcome of severe complications or death.

A prior history of CDI had a significant association with the composite outcome of treatment failure or recurrence. Age >65 years, female sex, proton pump inhibitor use, recent antibiotics, inflammatory bowel disease, body mass index, recent hospitalization, immunosuppression, and mode of acquisition were not associated with treatment failure or recurrence. Relative to subjects not hospitalized, those hospitalized for non-CDI reasons were less likely to have recurrence or the composite of treatment failure and recurrence (Table 4).

Table 1. Characteristics of the Study Subjects

	Overall, n = 468[a]	Community, n = 306	Hospital, n = 139	*P* Value
Sex	…	…	…	.005
Male	215 (45.9%)	130 (42.5%)	79 (56.8%)	
Female	253 (54.1%)	176 (57.5%)	60 (43.2%)	
Age, years	…	…	…	< .001
Mean (standard deviation)	53.2 (24.0)	47.8 (24.4)	62.9 (18.4)	
Median (1st, 3rd quartile)	58 (35, 72)	52.0 (29.0, 66.0)	66.0 (56.0, 75.0)	
Race	…	…	…	.36
Caucasian	435 (92.9%)	285 (93.1%)	130 (93.5%)	
African American/Black	6 (1.3%)	4 (1.3%)	2 (1.4%)	
Asian	5 (1.1%)	4 (1.3%)	1 (.7%)	
Native American/Alaskan	2 (.4%)	2 (.7%)	0 (.0%)	
Other	14 (3.0%)	6 (2.0%)	6 (4.3%)	
Unknown	6 (1.3%)	5 (1.6%)	0 (.0%)	
Body mass index	…	…	…	.006
Mean (standard deviation)	27.0 (7.9)	26.3 (7.6)	28.6 (8.5)	
Median (1st, 3rd quartile)	25.9 (21.7, 30.6)	25.8 (21.3, 29.9)	26.2 (22.9, 32.9)	
Body mass index group	…	…	…	.003
<18.5	54 (11.5%)	44 (14.4%)	7 (5.0%)	
18.5–24.9	150 (32.1%)	94 (30.7%)	49 (35.3%)	
25–29.9	133 (28.4%)	93 (30.4%)	33 (23.7%)	
30+	131 (28.0%)	75 (24.5%)	50 (36.0%)	
Proton pump inhibitor use	…	…	…	< .001
No	292 (62.4%)	216 (70.6%)	62 (44.6%)	
Yes	176 (37.6%)	90 (29.4%)	77 (55.4%)	
Gut graft-versus-host disease	…	…	…	.24
No	465 (99.4%)	303 (99.0%)	139 (100.0%)	
Yes	3 (.6%)	3 (1.0%)	0 (.0%)	
Inflammatory bowel disease	…	…	…	< .001
None	401 (85.7%)	248 (81.0%)	130 (93.5%)	
Crohn disease	33 (7.1%)	25 (8.2%)	8 (5.8%)	
Ulcerative colitis	34 (7.3%)	33 (10.8%)	1 (.7%)	

Table 2. Previously Studied Risk Factors for *Clostridioides difficile* Infection Acquisition

	Overall, n = 468[a]	Community, n = 306	Hospital, n = 139	*P* Value
Reason for hospitalization[b]	…	…	…	< .001
Not hospitalized	277 (59.3%)	234 (76.5%)	36 (25.9%)	
Hospitalization for CDI	67 (14.3%)	36 (11.8%)	23 (16.5%)	
Already in hospital	123 (26.3%)	36 (11.8%)	80 (57.6%)	
Recent hospitalization[b]	…	…	…	< .001
Discharge within 0–4 weeks of symptom onset	105 (22.5%)	13 (4.2%)	80 (57.6%)	
Discharge within 4–12 weeks of symptom onset	42 (9.0%)	23 (7.5%)	17 (12.2%)	
None within 12 weeks of testing	320 (68.5%)	270 (88.2%)	42 (30.2%)	
Recent antibiotic use (<4 weeks)	…	…	…	< .001
No	273 (58.3%)	220 (71.9%)	44 (31.7%)	
Yes	195 (41.7%)	86 (28.1%)	95 (68.3%)	
Prior CDI episode(s)[b]	…	…	…	.003
No	361 (77.5%)	228 (74.8%)	121 (87.1%)	
Yes	105 (22.5%)	77 (25.2%)	18 (12.9%)	
Immunodeficiency/immunosuppression				
No	267 (57.1%)	187 (61.1%)	67 (48.2%)	.011
Yes	201 (42.9%)	119 (38.9%)	72 (51.8%)	

Abbreviation: CDI, *Clostridioides difficile* infection.

[a]23 subjects had onset of symptoms in a long-term care facility or it was unclear where the onset of symptoms were.
[b]Percentages based on those with data available. Data were missing for hospitalization reason from 1 subject, recent hospitalization from 1 subject, and prior CDI from 2 subjects.

Characteristics of Subjects with Community- and Hospital-acquired *Clostridioides difficile* Infection

Community-acquired CDI was more common in women than men (57.5% vs 42.5%, respectively; $P = .005$), with hospital-acquired CDI being more common in men than women (56.8% vs 43.2%, respectively; $P = .003$). Subjects with community-acquired CDI were younger (median age, 52; IQR, 29–66; 74.2% ≤65 years) than those with hospital-acquired CDI (median age, 66; IQR, 56–75; 48.9% ≤65 years; $P < .001$). A prior history of CDI was more common in subjects with community-acquired CDI (25.2%), compared to those with hospital-acquired CDI (12.9%; $P = .003$). In comparison, proton pump inhibitor use was more common in hospital-acquired CDI subjects (55.4%), compared to those with a community-acquired CDI (29.4%, $P < .001$). Ribotype information on community- versus hospital-acquired CDI is presented in Supplementary Table 2.

Table 3. Clinical Outcomes of *Clostridioides difficile* Infection

	Overall, n = 468[a]	Community, n = 306	Hospital, n = 139	*P* Value
CDI severity[b]	…	…	…	<.001
Mild/moderate	360 (77.8%)	256 (83.9%)	91 (65.9%)	
Severe	86 (18.6%)	42 (13.8%)	41 (29.7%)	
Severe and complicated	17 (3.7%)	7 (2.3%)	6 (4.3%)	
Response to treatment[b]	…	…	…	.32
Response	350 (78.5%)	225 (76.8%)	113 (83.1%)	
Failure	24 (5.4%)	16 (5.5%)	6 (4.4%)	
Recurrence after response	72 (16.1%)	52 (17.7%)	17 (12.5%)	
Number of recurrences[b]	…	…	…	.15
0	389 (84.4%)	251 (82.8%)	122 (87.8%)	
1	56 (12.1%)	40 (13.2%)	14 (10.1%)	
2	15 (3.3%)	11 (3.6%)	3 (2.2%)	
6	1 (.2%)	1 (.3%)	0 (.0%)	
Severe complications[b]	…	…	…	.15
No	448 (96.8%)	298 (97.7%)	131 (94.9%)	
Yes	6 (1.3%)	4 (1.3%)	2 (1.4%)	
Death	9 (1.9%)	3 (1.0%)	5 (3.6%)	

Abbreviation: CDI, *Clostridioides difficile* infection.

[a]23 subjects had a different onset class than community or hospital (typically long-term care facility).
[b]Percentages based on those with data available. Data were missing for CDI severity for 5 subjects, response to treatment for 22, number of recurrences for 7, and severe complications for 5.

Table 4. Subject Characteristics and Association With *Clostridioides difficile* Infection Outcomes

	Severe CDI		Severe Complications or Death		Treatment Failure or Recurrence	
	Odds Ratio (95% CI)	*P* Value	Odds Ratio (95% CI)	*P* Value	Odds Ratio (95% CI)	*P* Value
Age >65	2.38(1.52–3.72)	<.001	7.87 (2.46–34.9)	.002	1.06 (.66–1.69)	.81
Female sex	.62 (.40–.97)	.036	.73 (.25–2.07)	.55	1.26 (.80–2.00)	.32
PPI use	1.50 (.96–2.34)	.072	1.92 (.68–5.58)	.21	.70 (.43–1.12)	.14
Recent antibiotic use (<4 weeks)	2.51 (1.61–3.95)	<.001	1.64 (.58–4.75)	.35	.87 (.55–1.38)	.57
IBD history	.37 (.15–.77)	.015	.41 (.02–2.11)	.40	.94 (.47–1.77)	.85
BMI	1.00 (.97–1.03)	.94	.94 (.86–1.01)	.15	1.00 (.97–1.03)	.99
Hospitalization	…	<.001	…	.006	…	.072
Not hospitalized	1.00 (reference)		1.00 (reference)		1.00 (reference)	
Hospitalized for CDI	12.3 (6.51–23.7)	<.001	7.31 (1.75–36.4)	.007	.89 (.46–1.65)	.72
Hospitalized for non-CDI	6.05 (3.48–10.8)	<.001	5.57 (1.52–26.2)	.014	.52 (.28–.91)	.028
Recent hospitalization	…	<.001	…	.85	…	.93
None within 12 weeks prior	1.00 (reference)		1.00 (reference)		1.00 (reference)	
Discharged <4 weeks prior	2.78 (1.67–4.62)	<.001	.90 (.20–3.01)	.88	.90 (.51–1.54)	.70
Discharged 4-12 weeks prior	4.60 (2.31–9.09)	<.001	1.53 (.23–6.07)	.59	.97 (.42–2.05)	.93
Prior history of CDI	.76 (.43–1.30)	.34	.52 (.08–1.92)	.40	2.01 (1.21–3.31)	.007
Immunosuppressed state	1.62 (1.04–2.51)	.033	1.16 (.40–3.27)	.78	.70 (.43–1.11)	.13
Acquisition of CDI	…	<.001	…	.30	…	.24
Community-onset	1.00 (reference)		1.00 (reference)		1.00 (reference)	
Hospital-onset	2.70 (1.69–4.31)	<.001	2.28 (.76–6.77)	.13	.67 (.39–1.12)	.14
Other	2.81 (1.01–7.24)	.036	2.24 (.12–13.5)	.46	1.38 (.43–3.86)	.56

Abbreviation: BMI, body mass index; CDI, *Clostridioides difficile* infection; CI, confidence interval; IBD, inflammatory bowel disease; PPI, proton pump inhibitor.

Whole-genome Sequencing Analysis

In analyzing genetic relatedness, we considered all isolates, including those obtained from the same subject over time. The most common ribotype was RT-14–20. We evaluated 4 thresholds to determine relatedness based on cgMLST. There were 386 unique isolates, as well as 21 groupings of genetically identical isolates, using the most stringent cutoff of no allelic differences (Figure 2). There were 36 groupings of genetically similar isolates using a cutoff of ≤2 allelic differences, with the largest group comprising 13 isolates; 313 isolates did not cluster with any of the others at this allelic threshold cutoff. There were 51 groupings of genetically similar isolates using a cutoff of ≤7 allelic differences, with the largest group making up 29 isolates; 199 isolates did not cluster with any of the others at this allelic threshold cutoff. Finally, the application of a more generous cutoff of ≤50 allelic differences yielded 48 groups of genetically similar isolates, with the largest group comprising 42 isolates. Supplementary Figures 1 and 2 highlight the relationships among isolates from immunosuppressed subjects and subjects who had recurrent CDI.

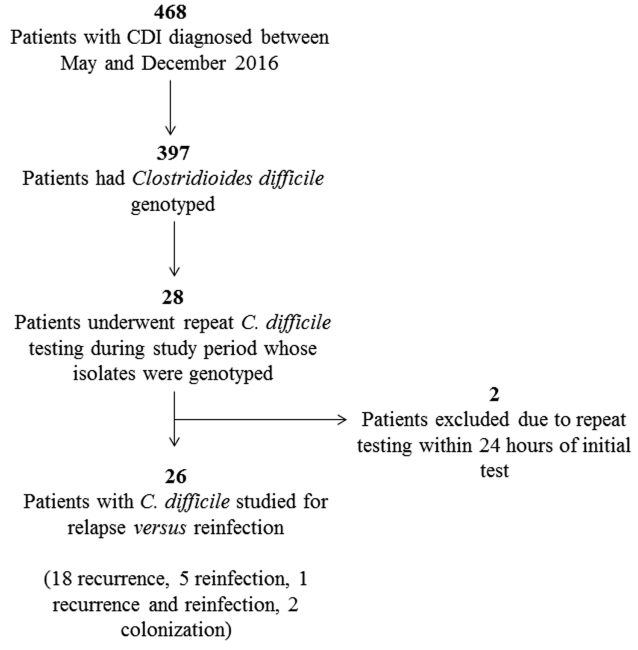

Figure 1. Flowchart of sample analysis. Clinical data was obtained in association with the first *Clostridioides difficile*–positive sample from each subject. A separate analysis of the 28 subjects who underwent repeat testing was performed to determine whether the isolates reflected recurrence or reinfection (by examining genotype and ribotype; Table 5). There was 1 subject who had 2 episodes of recurrent CDI, with 1 reflecting relapse and the other reflecting reinfection. Abbreviation: CDI, *Clostridioides difficile* infection.

Whole-genome Sequencing to Define Relapse Versus Reinfection

During the study period, 28 subjects underwent repeat testing for *C. difficile* (Table 5), 2 of whom had repeated testing done within 24 hours and were excluded. Among the remaining 26 subjects for whom WGS data was available at the time of subsequent detection of *C. difficile*, 18 were clinically classified as having recurrent CDI (based on symptom onset ≤8 weeks after primary infection). Of these 18 subjects, 14 had isolates with ≤2 allelic differences between them, suggesting relapse with the same strain, while 4 had isolates with >50 allelic differences between them, suggesting reinfection with a new strain. Among the 5 subjects clinically classified as having a reinfection (based

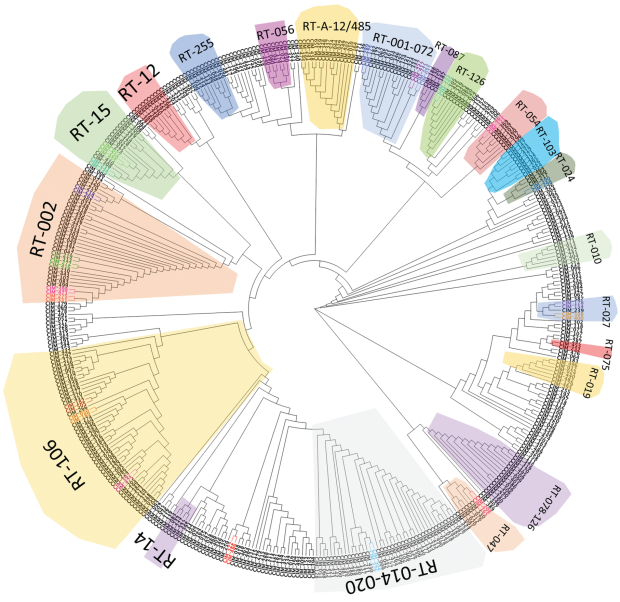

Figure 2. Genomic analysis of *Clostridioides difficile* isolates. The tree was built using the nearest neighbor-joining method. The circular cladogram shows relationships between *C. difficile* genomic sequences. Isolates are labeled with CIM and a number. Isolates with no allelic differences are labeled with the same color. There were 36 clusters of genetically similar isolates derived using a cutoff of ≤2 allelic differences. Isolates of the same predicted ribotype are shaded the same color. Abbreviations: CIM, Center for Individualized Medicine; RT, ribotype.

on symptom onset >8 weeks after primary infection), 3 had isolates with ≤2 allelic differences between them, suggesting relapse of the same strain , while 2 had isolates with >50 allelic differences between them, suggesting true reinfection. There was 1 subject who had a recurrence based on a positive test within 8 weeks of the initial infection and was found to have an isolate with no allelic differences compared to the primary isolate, and then subsequently had a reinfection based on another positive test 8 weeks after the second episode, at which time they had an isolate with >50 allelic differences from the other 2, consistent with reinfection with a new strain. There were 2 subjects who were retested in the absence of diarrhea, 1 of whom had an indistinguishable isolate (suggesting persistent colonization), and the other of whom had an isolate differing from the first by ≤50 allelic differences, suggesting either that the subject was initially infected with more than 1 strain or that they acquired a second strain after the initial diagnosis without developing symptoms.

Table 5. Subgroup Analysis of Subjects Who Underwent Repeat *Clostridioides difficile* Testing

Number of Sets of Samples from the Same Subject[a]	Number of Allelic Differences	Matching Ribotype[b]	Classification[c]
10	0	Yes	Recurrence
1	0	Yes	Reinfection
1	0	Yes	Colonization
2	≤2	Yes	Recurrence
3	≤2	Unknown	Recurrence
2	≤2	Yes	Reinfection
4	>50	Unknown	Recurrence
3	>50	Unknown	Reinfection
1	≤50	Yes	Colonization

Abbreviation: CDI, *Clostridioides difficile* infection.

[a]Each set of samples represents the initial sample and a subsequent sample obtained during the study period.

[b]Ribotypes from the initial sample and subsequent sample were compared and, if they matched, then a label of "Yes" was applied. Some samples were not able to be ribotyped and thus are labeled as "Unknown" in this column.

[c]CDI was classified as a recurrence if symptoms occurred within 8 weeks of an initial *Clostridioides difficile* test positivity, reinfection if symptoms occurred beyond 8 weeks of *C. difficile* test positivity, or colonization if a *C. difficile* test was positive but there were no associated symptoms.

Whole-genome Sequencing–Guided Tracking of Infection

There were 36 clusters of genetically similar isolates using a cutoff of ≤2 allelic differences. Subjects within each cluster were investigated using standard infection prevention and control approaches, including an assessment of hospital room location, paths crossed when undergoing diagnostic imaging, and home addresses to identify potential epidemiologic linkages. Of 94 isolates reviewed, there were 17 pairs that were from the same subject at different points in time. The medical records of the other 60 subjects were reviewed by an infection prevention and control practitioner using standard infection prevention and control approaches, to identify factors that might suggest transmission between these subjects. Within only 1 of these clusters was there a clear epidemiologic link found. That link was between 2 subjects, 1 of whom was a 5-month-old with diarrhea who tested positive for *C. difficile*. The infant's mother and primary caregiver developed watery diarrhea 2 weeks later and was diagnosed with *C. difficile* infection. In another group, there was possible transmission between 2 subjects: 1 was hospitalized on a medical unit with community-onset CDI, with the second developing hospital-onset CDI while hospitalized on the same patient care unit. The 2 cases were cared for by the same nursing staff, though their hospital stays were separated by 3 months and they were not housed in the same hospital room.

In order to assess for possible community transmission, we studied clusters of more than 1 subject with ≤2 allelic differences between isolates associated with community-acquired *C. difficile*. We found that 8 of 20 such clusters included subjects residing within the same zip code. (Supplementary Figure 3). In looking further at clusters in the same geographic region, 1 of the groups included 3 subjects who had community-acquired CDI within 3 weeks of each other. Of these, 2 were hospitalized after onset of CDI symptoms, with the subjects roomed on different floors of the hospital and cared for by different nursing and provider teams. There were no obvious commonalities noted, such as outpatient visits to the same clinic, diagnostic imaging, or wound care; it is possible that they had links in the community that were not obvious from a review of their medical records.

Infection Outcomes Are Not Associated with Whole-genome Sequencing–Based Classification of Strains

There was no association between *C. difficile* isolates grouped based on allelic difference cutoffs of 0, ≤2, ≤7, and ≤50 allelic differences, or by ribotype and CDI severity, response to treatment, or recurrence after successful treatment.

DISCUSSION

Our findings highlight the relevance of WGS in determining a relapse with the same strain versus reinfection with a new strain of *C. difficile* in patients with repeated episodes of CDI, providing different classifications than those provided by current definitions based on the time since the primary infection. This has potential implications in informing individualized therapeutic strategies for subsequent CDI. A relapse with the same strain based on WGS may indicate the suppression of a primary infection or a persistent reservoir in the gastrointestinal tract, leading to recurrent disease. It may also suggest that a different therapy be considered at the time of relapse to prevent multiple recurrent CDI. Alternately, if the strain is different from the strain causing the primary infection, repeat therapy with the initial regimen may be a consideration.

We evaluated different thresholds of relatedness, since there are no definitive criteria for a definition of relatedness. The cutoffs for our analysis are in line with Kociolek et al [23], who defined isolates collected <124 days apart as isogenic if they differed by ≤2 single nucleotide variants and those collected 124–364 days apart as isogenic if they differed by ≤3 single nucleotide variants. We found that 40% of groups with community-acquired CDI and ≤2 allelic differences reside in the same zip code, suggesting a potential common community reservoir. There are several possible routes of transmission of *C. difficile* in the community, including food, water, pets, soil, plants, sod or other landscaping materials, wind, gatherings, schools, workplaces, and restrooms, to name a few. The identification of a community reservoir was beyond the scope of this work. We found associations of host features and outcomes of CDI based on a univariate analysis. Our findings align with those of prior studies showing that older age is associated with severe CDI and the development of severe complications [24, 25], while differing from previous findings with regard to the association of recent antibiotic use and immunosuppression with severe CDI [26]. We also found a risk profile for community-onset CDI, allowing for potential early consideration of testing for *C. difficile* in certain outpatients.

It is not surprising that we found no correlation between isolate groupings and outcomes of CDI, as disease is a result of interactions of the host, environmental, and pathogen features, rather than being related to a singular factor.

There are several limitations to our study. This was a single-center study with limited demographic and racial diversity; hence, our findings may not be broadly applicable. Further, only a small subset of subjects had their *C. difficile* isolate sequenced at the time of subsequent CDI. Also, we did not sequence multiple colonies from the same subject, so it is possible that we missed the presence of more than 1 strain at the time of primary infection; our finding of the same strain at the time of subsequent CDI in some subjects allays some of this concern. Our study did not capture asymptomatic carriers, who may be a source of *C. difficile* in the hospital [27], though our previous study found a low rate of apparent nosocomial transmission [28]. These limitations aside, our findings may be relevant for identifying strategies for risk reduction in CDI, as well as treatment stratification in recurrent CDI. They also add to the literature suggesting that CDI may be acquired in health-care facilities through occult routes [29–32], and that there is a need to identify and curb the community spread of CDI, which appears to be on the rise.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. The authors thank the staff of the Mayo Clinic Anaerobic Bacteriology Laboratory for culture and isolation of the study isolates; Bruce Eckloff and Julie Lau from the Mayo Clinic Medical Genome Facility for performing the sequencing; and Lyndsay Busby for secretarial assistance.

Financial support. This work was supported by the National Institutes of Health (grant number National Institutes of Health DK114007 to P. C. K.) and the Center for Individualized Medicine at the Mayo Clinic.

Potential Conflicts of Interest. S. C. has received speaker honoraria, paid to the Mayo Clinic, from the American Society for Microbiology, outside the submitted work. P. C. K. has received advisory board fees from Novome Biotechnologies and consulting fees from Otsuka Pharmaceuticals and Pendulum Therapeutics, outside the submitted work. R. P. has received grants from CD Diagnostics, Merck, Hutchison Biofilm Medical Solutions, Accelerate Diagnostics, ContraFect, TenNor Therapeutics Limited, and Shionogi; has a patent on *Bordetella pertussis/parapertussis* polymerase chain reaction issued, a patent on a device/method for sonication with royalties paid by Samsung to Mayo Clinic, and a patent on an antibiofilm substance issued; and receives travel reimbursements from the American Society for Microbiology and the Infectious Diseases Society of America (IDSA), an editor's stipend from IDSA, and honoraria from the National Board of Medical Examiners, Up-to-Date, and the Infectious Diseases Board Review Course. R. P. is a consultant to Curetis, Specific Technologies, Next Gen Diagnostics, PathoQuest, Selux Diagnostics, 1928 Diagnostics, and Qvella and consultant fees are paid to the Mayo Clinic. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Zimlichman E, Henderson D, Tamir O, et al. Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system. JAMA Intern Med **2013**; 173:2039–46.

2. Lessa FC, Winston LG, McDonald LC; Emerging Infections Program *Clostridioides difficile* Surveillance Team. Burden of *Clostridium difficile* infection in the United States. N Engl J Med **2015**; 372:2369–70.
3. Krutova M, Nyc O, Matejkova J, Kuijper EJ, Jalava J, Mentula S. The recognition and characterisation of Finnish *Clostridium difficile* isolates resembling PCR-ribotype 027. J Microbiol Immunol Infect **2018**; 51:344–51.
4. Zhou Y, Burnham CA, Hink T, et al. Phenotypic and genotypic analysis of *Clostridium difficile* isolates: a single-center study. J Clin Microbiol **2014**; 52:4260–6.
5. Kumar N, Miyajima F, He M, et al. Genome-based infection tracking reveals dynamics of *Clostridium difficile* transmission and disease recurrence. Clin Infect Dis **2016**; 62:746–52.
6. Eyre DW, Walker AS. *Clostridium difficile* surveillance: harnessing new technologies to control transmission. Expert Rev Anti Infect Ther **2013**; 11:1193–205.
7. Krutova M, Nyc O, Matejkova J, Allerberger F, Wilcox MH, Kuijper EJ. Molecular characterisation of Czech *Clostridium difficile* isolates collected in 2013–2015. Int J Med Microbiol **2016**; 306:479–85.
8. Sim JH, Truong C, Minot SS, et al. Determining the cause of recurrent *Clostridium difficile* infection using whole genome sequencing. Diagn Microbiol Infect Dis **2017**; 87:11–6.
9. Jin D, Luo Y, Huang C, et al. Molecular epidemiology of *Clostridium difficile* infection in hospitalized patients in Eastern China. J Clin Microbiol **2017**; 55:801–10.
10. Walker AS, Eyre DW, Wyllie DH, et al; Infections in Oxfordshire Research Database. Relationship between bacterial strain type, host biomarkers, and mortality in *Clostridium difficile* infection. Clin Infect Dis **2013**; 56:1589–600.
11. Corbellini S, Piccinelli G, De Francesco MA, Ravizzola G, Bonfanti C. Molecular epidemiology of *Clostridium difficile* strains from nosocomial-acquired infections. Folia Microbiol **2014**; 59:173–9.
12. Forgetta V, Oughton MT, Marquis P, et al. Fourteen-genome comparison identifies DNA markers for severe-disease-associated strains of *Clostridium difficile*. J Clin Microbiol **2011**; 49:2230–8.
13. Dong D, Peng Y, Zhang L, Jiang C, Wang X, Mao E. Clinical and microbiological characterization of *Clostridium difficile* infection in a tertiary care hospital in Shanghai, China. Chin Med J **2014**; 127:1601–7.
14. Sirard S, Valiquette L, Fortier LC. Lack of association between clinical outcome of *Clostridium difficile* infections, strain type, and virulence-associated phenotypes. J Clin Microbiol **2011**; 49:4040–6.
15. Sloan LM, Duresko BJ, Gustafson DR, Rosenblatt JE. Comparison of real-time PCR for detection of the tcdC gene with four toxin immunoassays and culture in diagnosis of *Clostridium difficile* infection. J Clin Microbiol **2008**; 46:1996–2001.
16. Didion JP, Martin M, Collins FS. Atropos: specific, sensitive, and speedy trimming of sequencing reads. Peer J **2017**; 5:e3720.
17. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet **2011**; 17:10–2.
18. Jünemann S, Sedlazeck FJ, Prior K, et al. Updating benchtop sequencing performance comparison. Nat Biotechnol **2013**; 31:294–6.
19. Bletz S, Janezic S, Harmsen D, Rupnik M, Mellmann A. Defining and evaluating a core genome multilocus sequence typing scheme for genome-wide typing of *Clostridium difficile*. J Clin Microbiol **2018**; 56:e01987–17.
20. Loo VG, Bourgault AM, Poirier L, et al. Host and pathogen factors for *Clostridium difficile* infection and colonization. N Engl J Med **2011**; 365:1693–703.
21. Battaglioli EJ, Hale VL, Chen J, et al. *Clostridioides difficile* uses amino acids associated with gut microbial dysbiosis in a subset of patients with diarrhea. Sci Transl Med **2018**; 10:eeam7019.
22. McDonald LC, Gerding DN, Johnson S, et al. Clinical practice guidelines for *Clostridium difficile* infection in adults and children: 2017 update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clin Infect Dis **2018**; 66:987–94.
23. Kociolek LK, Gerding DN, Espinosa RO, Patel SJ, Shulman ST, Ozer EA. *Clostridium difficile* whole genome sequencing reveals limited transmission among symptomatic children: a single-center analysis. Clin Infect Dis **2018**; 67:229–34.
24. Abou Chakra CN, McGeer A, Labbé AC, et al. Factors associated with complications of *Clostridium difficile* infection in a multicenter prospective cohort. Clin Infect Dis **2015**; 61:1781–8.
25. Henrich TJ, Krakower D, Bitton A, Yokoe DS. Clinical risk factors for severe *Clostridium difficile*-associated disease. Emerg Infect Dis **2009**; 15:415–22.
26. Rodríguez-Pardo D, Almirante B, Bartolomé RM, et al; Barcelona *Clostridium difficile* Study Group. Epidemiology of *Clostridium difficile* infection and risk factors for unfavorable clinical outcomes: results of a hospital-based study in Barcelona, Spain. J Clin Microbiol **2013**; 51:1465–73.
27. Kong LY, Dendukuri N, Schiller I, et al. Predictors of asymptomatic *Clostridium difficile* colonization on hospital admission. Am J Infect Control **2015**; 43:248–53.
28. Cho J, Seville MT, Khanna S, Pardi DS, Sampathkumar P, Kashyap PC. Screening for *Clostridium difficile* colonization on admission to a hematopoietic stem cell transplant unit may reduce hospital-acquired *C difficile* infection. Am J Infect Control **2018**; 46:459–61.
29. Gonzalez-Orta M, Saldana C, Ng-Wong Y, et al. Are many patients diagnosed with healthcare-associated *Clostridioides difficile* infections colonized with the infecting strain on admission? Clin Infect Dis **2019**; 69:1801–4.
30. Durovic A, Widmer AF, Tschudin-Sutter S. New insights into transmission of *Clostridium difficile* infection-narrative review. Clin Microbiol Infect **2018**; 24:483–92.
31. Tschudin-Sutter S. *Clostridioides difficile* infection in outpatient settings - the need for studies on clinical impact. Clin Microbiol Infect **2019**; 25:534–5.
32. Muñoz-Price LS, Hanson R, Singh S, et al. Association between environmental factors and toxigenic *Clostridioides difficile* carriage at hospital admission. JAMA Netw Open **2020**; 3:e1919132.