



UNIVERSITY OF LEEDS

This is a repository copy of *Clostridium difficile: Investigating Transmission Patterns Between Infected and Colonized Patients Using Whole Genome Sequencing*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/134934/>

Version: Accepted Version

Article:

Kong, LY, Eyre, DW, Corbeil, J et al. (18 more authors) (2019) *Clostridium difficile: Investigating Transmission Patterns Between Infected and Colonized Patients Using Whole Genome Sequencing*. *Clinical Infectious Diseases*, 68 (2). pp. 204-209. ISSN 1058-4838

<https://doi.org/10.1093/cid/ciy457>

© 2018, The Author(s). Published by Oxford University Press for the Infectious Diseases Society of America. This is an author produced version of a paper published in *Clinical Infectious Diseases*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

1 **Clostridium difficile: investigating transmission patterns between infected and**
2 **colonized patients using whole genome sequencing**

3

4 L. Y. Kong MD¹, D. W. Eyre DPhil^{2,3}, J. Corbeil MD⁴, F. Raymond PhD⁴, A. S. Walker
5 PhD³, M.H. Wilcox MD⁵, D. W. Crook MD^{2,6}, S. Michaud MD⁷, B. Toye MD⁸, E. Frost
6 MD⁷, N. Dendukuri PhD⁹, I. Schiller MSc¹⁰, A.M. Bourgault MD^{1,11}, A. Dascal MD¹²,
7 M. Oughton MD¹², Y. Longtin MD¹², L. Poirier MD¹³, P. Brassard MD¹, N. Turgeon
8 MD¹⁴, R. Gilca MD¹⁵, V. G. Loo MD¹

9

10 ¹Division of Infectious Diseases and Department of Medical Microbiology, McGill
11 University Health Centre, Montréal, Canada

12 ²Nuffield Department of Medicine, John Radcliffe Hospital, Oxford, UK

13 ³NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK

14 ⁴Centre de recherche CHUQ, Université Laval, Québec, Canada

15 ⁵Department of Microbiology, Leeds Teaching Hospitals and University of Leeds, Leeds,
16 UK

17 ⁶National Infection Service, Public Health England, London, UK

18 ⁷Department of Microbiology, Centre Hospitalier Universitaire de Sherbrooke,
19 Sherbrooke, Canada

20 ⁸Division of Microbiology, Ottawa Hospital, University of Ottawa, Ottawa, Canada

21 ⁹Technology Assessment Unit, McGill University Health Centre, Montréal, Canada

22 ¹⁰Centre for Outcomes Research, McGill University Health Centre – Research Institute,
23 Montréal, Canada

1 ¹¹Department of Microbiology, Centre Hospitalier de l'Université de Montréal, Montréal,
2 Canada

3 ¹²Division of Infectious Diseases, Jewish General Hospital, Montréal, Canada

4 ¹³Department of Microbiology, Hôpital Maisonneuve-Rosemont, Montréal, Canada

5 ¹⁴Department of Microbiology, Centre Hospitalier Universitaire de Québec – Hôtel-Dieu
6 de Québec, Québec, Canada

7 ¹⁵ Quebec Institute of Public Health, Québec, Canada

8

9 Key words: Clostridium difficile; whole genome sequencing; colonization

10 Running title: C. difficile colonization & transmission

11

12 Corresponding author:

13 Ling Yuan Kong, MD

14 Division of Infectious Diseases and Department of Medical Microbiology

15 McGill University Health Centre

16 1001 Decarie Blvd

17 Montreal, QC H4A 3J1

18 Email : ling.kong@mail.mcgill.ca

19 Telephone : (514) 835-7093

20 Fax : (514) 412-4354

21

22 Alternate Corresponding Author:

23 Vivian Loo, MD

24 Division of Infectious Diseases and Department of Medical Microbiology

1 McGill University Health Centre
2 1001 Decarie Blvd
3 Montreal, QC H4A 3J1
4 Email: vivian.loo@mcgill.ca
5 Telephone: (514) 934-1934 ext. 42818
6 Fax : (514) 412-4354
7
8 Word count: Text 2352 Abstract 229
9 Figures: 1
10 Tables: 2

1 **Summary of main point**

2 Using whole genome sequencing of isolates from a cohort of patients with Clostridium
3 difficile infection (CDI) and colonization, we found that incident CDI cases were more
4 likely to be linked to an infected than colonized donor, in this setting with high rates of
5 the NAP1/027/ST1 strain.

1 **Abstract**

2

3 **Background**

4 Whole genome sequencing (WGS) studies can enhance our understanding of the role of
5 patients with asymptomatic *Clostridium difficile* colonization in transmission.

6

7 **Methods**

8 Isolates obtained from patients with *Clostridium difficile* infection (CDI) and colonization
9 identified in a study conducted during 2006 - 2007 at six Canadian hospitals underwent
10 typing by pulsed-field gel electrophoresis, multilocus sequence typing, and WGS.

11 Isolates from incident CDI cases not in the initial study were also sequenced where
12 possible. Ward movement and typing data were combined to identify plausible donors for
13 each CDI case, as defined by shared time and space within predefined limits. Proportions
14 of plausible donors for CDI cases that were colonized, infected, or both were examined.

15

16 **Results**

17 Five hundred and fifty-four isolates were sequenced successfully, 353 from colonized and
18 201 from CDI cases. The NAP1/027/ST1 strain was the most common strain, found in
19 124 (62%) of infected and 92 (26%) of colonized patients. A donor with a plausible ward
20 link was found for 81 CDI cases (40%) using WGS with a threshold of ≤ 2 single
21 nucleotide variants to determine relatedness. Sixty-five (32%) CDI cases could be linked
22 to both infected and colonized donors. Exclusive linkages to infected and colonized
23 donors were found for 28 (14%) and 12 (6%) CDI cases, respectively.

1

2 **Conclusion**

3 Colonized patients contribute to transmission, but CDI cases are more likely linked to
4 other infected patients than colonized patients in this cohort with high rates of
5 NAP1/027/ST1 strain, highlighting the importance of local prevalence of virulent strains
6 in determining transmission dynamics.

1 **Background**

2

3 Clostridium difficile is a leading cause of healthcare-associated diarrhea and a major
4 cause of morbidity and mortality for hospitalized patients[1]. Patients with symptomatic
5 infection and asymptomatic colonization are both known to shed spores into the
6 environment[2]. Currently recommended infection control measures focus on the
7 detection and isolation of symptomatic patients, believed to be responsible for most
8 healthcare-associated transmission events[3]. However, recent molecular studies using
9 whole genome sequencing (WGS) have found that most new cases of C. difficile
10 infection (CDI) in endemic settings could not be explained by transmission from
11 symptomatic cases[4], raising interest in the role of colonized patients in transmission of
12 C. difficile.

13

14 Typing methods used to identify transmission leading to CDI include pulsed-field gel
15 electrophoresis (PFGE), PCR ribotyping, and multilocus sequence typing (MLST),
16 among others[2]. With the advent of high-throughput sequencing technologies, WGS is
17 increasingly being adopted as a preferred typing/fingerprinting method with high
18 discriminatory power, and so has been used in multiple molecular epidemiology studies
19 on C. difficile transmission[4-7]. In this study, using WGS of isolates and
20 epidemiological data from a prospective cohort study, we aimed to elucidate the role of
21 patients colonized with C. difficile in onward transmission of infection.

22

23 **Methods**

1

2 **Study population and definitions**

3 A multicenter prospective study was conducted between March 6, 2006 and June 25,
4 2007 to determine host and pathogen factors for health care-associated *C. difficile*
5 infection and colonization, with results previously published[8]. Briefly, data were
6 collected in six Canadian, university-affiliated hospitals, on 15 study units (seven surgical
7 units and eight medical units). The selected units were those with a historically high or
8 low incidence of CDI. All patients 18 years or older admitted to these hospital units were
9 eligible for participation. Exclusion criteria included hemodynamic instability, palliative
10 status, neutropenia (absolute neutrophil count ≤ 1000 per cubic millimeter), or inability to
11 participate in the informed-consent process.

12

13 Patients were followed daily until ward discharge, death, or withdrawal from the study.
14 Rectal swabs or stool samples were obtained for culture on admission, weekly during
15 hospitalization, and at onset of diarrhea (if applicable). Toxigenic *C. difficile* culture was
16 performed on stool samples or rectal swabs using standard methods[9]. The cell cytotoxin
17 neutralization assay was the diagnostic assay used in routine clinical care during the
18 study period. Isolates were tested for presence of *tcdA* and *tcdB* using nucleic acid
19 amplification methods[10, 11].

20

21 CDI was defined as the presence of diarrhea without an alternative explanation and a
22 positive *C. difficile* cytotoxin assay or toxigenic culture, an endoscopic diagnosis of
23 pseudomembranes, or a pathological diagnosis of CDI. Diarrhea was defined as at least

1 three loose stools within at least one 24-hour period. Asymptomatic *C. difficile*
2 colonization was defined as a positive stool *C. difficile* culture in the absence of diarrhea.
3 Non-toxigenic strains of *C. difficile* were defined as culture positive and tcdB negative.
4
5 In order to capture a more comprehensive picture of transmission, we also reviewed
6 infection control data to determine the incidence of CDI cases in non-participants
7 occurring on the study units during the study period. For one of the six participating
8 hospitals, isolates were conserved for the purpose of infection control surveillance and
9 were available for non-study incident CDI cases on study units; all incident CDI cases
10 participated in the study for one other hospital. These isolates were included in the
11 current analysis. Hospital and study unit admission and discharge dates were collected for
12 every participant admitted to study units.

14 **PFGE**

15 Each isolate underwent PFGE using standard methods[12] at the time of the study. Strain
16 relatedness was determined using the criteria of Tenover et al using BioNumerics
17 (Applied Maths)[13]. The Dice coefficient was used to measure similarity between
18 patterns.

20 **DNA preparation, sequencing, mapping and single nucleotide polymorphism (SNP)** 21 **detection**

22 DNA was extracted using Purelink viral RNA/DNA minikit (Invitrogen, Burlington, ON,
23 Canada) on a sub-cultured colony from frozen isolates. DNA was quantified using the

1 QuantiFluor dye (Promega). Sequencing libraries were prepared using the Nextera XT
2 Sample Preparation Kit (Illumina, San Diego, CA, USA) with 1 ng of purified DNA per
3 sample. Dual indices were added during library preparation. Library concentrations were
4 normalized using bead normalization as described by the manufacturer. Ninety-six
5 libraries were pooled per HiSeq lane. Sequencing was performed on the HiSeq 2500
6 sequencer (Illumina) using v3 chemistry, generating paired-end 101 bp reads. Reads and
7 assemblies have been deposited in the European Nucleotide Archive database in project
8 PRJEB11776.

9

10 Sequence reads were analyzed and assembled using a previously described pipeline
11 developed specifically for bacterial genomes[4]. The set of reads from each isolate was
12 mapped using Stampy v. 1.0.11 (without Burrows-Wheeler Aligner pre-mapping, using
13 an expected substitution rate of 0.01)[14] to the *C. difficile* 630 reference genome
14 (Genbank: AM180355.1)[15]. Base-pair calls were identified across all mapped non-
15 repetitive core genome sites using SAMtools (version 0.1.19) mpileup with the extended
16 base-alignment quality flag, using parameters based on bacterial sequences[4]. A
17 consensus of $\geq 75\%$ was required to support a nucleotide call, and calls were required to
18 be homozygous under a diploid model. Only calls supported by ≥ 5 reads, including one
19 in each direction were accepted.

20

21 Sequences were compared using single nucleotide polymorphisms (SNPs), obtaining
22 differences between sequences from maximum likelihood phylogenies constructed using
23 PhyML[16] with generalized time-reversible substitution model and “BEST” tree

1 topology search algorithm, corrected for the effect of recombination using
2 ClonalFrameML[17] (with default settings). Sequence reads were also assembled de novo
3 with Velvet[18] and MLSTs and toxigenic strains identified using BLAST searches of de
4 novo assemblies (≥ 1000 nucleotide identities with *tcdA* or *tcdB* genes).

6 **Transmission analysis**

7 Isolates' PFGE, MLST and toxigenic status were first examined according to colonized
8 or infected status. Ward movement and WGS data were then combined to identify
9 plausible donors for each CDI case. Proportions of plausible donors that were colonized
10 or infected were calculated. A donor was identified for an isolate when they were
11 determined to be clonal (differed by ≤ 2 SNPs by WGS), and a plausible epidemiological
12 link could be identified between the pair based on a previously described model[19],
13 namely the pair shared a ward after the donor tested positive and before the recipient
14 tested positive, shared a ward before either tested positive, or if the recipient occupied a
15 ward after the donor tested positive and was discharged. Maximum infectious period of 8
16 weeks, incubation period of 12 weeks and ward contamination period of 26 weeks were
17 allowed[20].

18
19 The analyses were first done for all available isolates, then restricted to two hospitals
20 where 80% or more of all incident CDI cases occurring on study units during the study
21 period were sequenced, whether or not part of the prospective study.

23 **Results**

1 Five hundred and thirteen of 568 isolates from the cohort study were available for
2 sequencing. An additional 52 isolates from 77 incident CDI cases from one of the
3 participating hospitals were included for a total of 565 isolates. The participation rate in
4 the initial prospective cohort study was 57.1% of eligible patients admitted to the study
5 units. For one hospital contributing 9.6% of isolates, all incident CDI cases on study units
6 were captured in the study. Figure 1 provides a breakdown of sample sources and patient
7 statuses.

8

9 Overall, 554 (98%) samples were sequenced successfully, from 550 patients (4 patients
10 contributed 2 samples). There were 353 samples from colonized patients and 201 from
11 infected patients. Two isolates did not have a PFGE pattern available, and 17 isolates
12 could not be assigned to a known MLST.

13

14 The epidemic NAP1/ST1(ribotype 027) strain was the most commonly occurring strain
15 among both infected and colonized patients, found in 124 (62%) and 92 (26%) patients,
16 respectively. However, the majority of colonized patients carried strains from a variety of
17 different sequence types (Figure 2). Strains from 27 different sequence types were found
18 among infected patients, whereas a greater variety with 41 sequence types was found
19 among colonized patients. The majority (74%) of colonized patients carried toxigenic
20 strains.

21

22 By comparing all samples from infected patients with prior samples from within the
23 cohort, using a threshold of ≤ 2 SNPs to determine relatedness, overall 105 (52%) cases

1 could be linked genetically to a prior sample (Table 1); 65 patients (32%) could be linked
2 to both infected and colonized donors. More cases were found to be related to isolates
3 only from infected patients than isolates only from colonized patients, 28 cases (14%)
4 and 12 cases (6%) respectively. Within all 105 cases related to a previous infected or
5 colonized donor using WGS, a donor with a plausible ward link could be found for 81
6 patients (77%; 40% of all 201 cases). Nearly all the identified donors were of the
7 epidemic NAP1/ST1 strain. Only 7 patients with genetic and ward links were found to
8 have non-NAP1/ST1 donors, including 3 linked to colonized donors only, 3 linked to
9 infected donors only and one to both infected and colonized donors.

10

11 Restricting analyses to the 2 hospitals with most complete data (Table 2), overall similar
12 patterns were observed, including for those cases substantiated with ward links. Thirty
13 out of 117 cases (26%) could be linked to isolates from both infected and colonized
14 patients and 26 (22%) to isolates from only infected patients, whereas only 4 (3%) were
15 linked to samples from only colonized patients. Of 46 cases with a ward link, 30 (26% of
16 all 117 cases) had an exclusive link to an infected donor, and only 2 (2% of all 117 cases)
17 had an exclusive link to a colonized donor.

18

19 **Discussion**

20 The role of colonized patients in transmission of CDI has been subject of several previous
21 molecular epidemiology studies [7, 19, 21]. Curry et al. used multilocus variable number
22 tandem repeats analysis genotyping and concluded that 29% of 56 incident CDI cases

1 could be linked to colonized patients [21]. Using WGS, Eyre et al. did not find evidence
2 of any onward transmission from 18 asymptomatic colonized patients to CDI cases [19].
3
4 Using WGS, we investigated the contribution of colonized and infected patients in
5 onward transmission toward incident CDI cases. In our larger cohort, 52% of cases could
6 be linked to a previous patient. This is higher than previously reported rates [4], in part
7 because our study includes both infected and colonized patients as sources, although
8 higher linkage rates to symptomatic patients, 93/201 (46%) of cases, were also found.
9 This difference may be explained in part by the diagnostic laboratory methods used. In
10 the study by Eyre et al, the laboratory method used was immunoassay whereas in our
11 study, the laboratory method was toxigenic culture which has a higher sensitivity than
12 enzyme immunoassay for detecting *C. difficile*. Therefore, more patients would have
13 been classified as CDI and a higher linkage would be made with CDI patients. However,
14 patients met the case definition for CDI and did not have an alternative explanation for
15 diarrhea. In addition, the high incidence of CDI of 28.1 cases per 10,000 patient-days in
16 our cohort reflected the epidemic setting of the study, with a large pool of symptomatic
17 patients, and a higher infection-to-colonization ratio compared to other cohorts[22]. The
18 high proportion of infected patients is likely explained by the predominance of the
19 NAP1/ST1 strain, which is more virulent and likely to cause infection[8].
20
21 Examining data from all units, an incident CDI case was 2.3 times more likely to be
22 linked to an infected patient only than to a colonized patient only, whereas in the subset
23 of hospitals with most complete data, this was 6 times more likely. Within the hospitals

1 where data were most complete, exclusive linkage to colonized donors was less common;
2 however, in these hospitals the proportion of infected cases sequenced (77-86%) was
3 substantially higher than on the other units (26-27%) due to availability of additional
4 isolates. In both analyses, many cases could be linked to both infected and colonized
5 patients, reflecting the outbreak setting in which the cohort study took place and the
6 relatively slow rate of *C. difficile* evolution relative to the time between transmitted cases,
7 enabling additional potential transmission links to be identified.

8

9 Our analyses suggest that colonized patients may be a source of onward transmission to
10 incident CDI cases, but that spread from infected donors is likely more frequent. This
11 could plausibly be explained by lower levels of shedding seen in colonized patients
12 (without diarrhea) as compared with infected patients [23]. Onward transmission events
13 from colonized individuals to infected patients in our cohort frequently carried the
14 epidemic NAP1/ST1 strain, possibly reflecting strain-specific characteristics, such as
15 higher transmissibility [24] (increasing the chance of acquisition) and higher propensity
16 to cause symptomatic infection and thereby increasing detection. For example,
17 NAP1/ST1 may be shed more frequently/persist more effectively in the environment. A
18 study using WGS to track transmission similar to ours, but examining only ribotype-027
19 (NAP1/ST1) strains within one UK hospital, found that 60% of their genetically-related
20 strains were circulated by ward-based contamination [7]. However, another possibility for
21 the greater degree of linkage is the relatively recent emergence of this fluoroquinolone-
22 resistant NAP1/ST1, resulting in less population-wide genetic diversity, and thus
23 increasing the chance of observing genetic linkage without direct transmission.

1

2 The limitations in our study include the incomplete sampling in the participating
3 hospitals. Overall, we only obtained fecal samples from 57% of eligible participants, and
4 did not capture all CDI cases on all study units. Incomplete sampling leads to the
5 proportion of linked cases being under-estimated as some potential transmission donors
6 are missed. Patients who were ineligible in the initial cohort study represent another pool
7 of potential missed linkages, since previously determined eligibility criteria (e.g.
8 neutropenia) for the prospective study do not necessarily translate to a ward-based
9 transmission analysis study. Ideally, studies focused on ward-based transmission would
10 be less restrictive, given the very low risk posed to patients of undergoing rectal swabs.
11 Increased participation could have been achieved by waiving written informed consent
12 and obtaining verbal consent and implementation of ward-based communication tools
13 explaining the option to opt-out.

14

15 When limiting the analyses to two hospitals with more than 80% incident cases
16 contributing isolates for sequencing, rates of linkage to infected patients increased, but
17 this could represent sampling bias given more infected donors were available. Finally,
18 although all transmission events were inferred from the genetic data, other sources, such
19 as patients not included in analyses, including ineligible patients, and the environment
20 were not sampled and may be other reservoirs of *C. difficile* leading to CDI.

21

22 Our study provides new insight into the epidemiology of transmission between colonized
23 and infected patients, by deriving data from the largest cohort to date of colonized and

1 infected patients along with geographic ward information. We also confirm the utility of
2 WGS in conjunction with epidemiological data to track transmission, which is
3 increasingly studied including in healthcare epidemiological models.

4

5 **Conclusion**

6 Patients colonized with *C. difficile* without diarrhea contribute to the transmission of
7 infection, but more transmission events appear to originate from infected patients with
8 diarrhea. Certain strains, such as the epidemic NAP1/ST1 strain, may be more
9 transmissible and virulent, and hence more likely to cause more symptomatic infection
10 following contact with infected and asymptotically colonized patients. Thus, the
11 relative contribution of colonized and infected patients toward onward transmission is
12 likely dependent on the local prevalence of virulent strains.

13

14 **Funding**

15 This work was supported by the funding source of the prospective cohort study, the
16 Consortium de Recherche sur le Clostridium difficile. The partners of the Consortium
17 include: Fonds de la recherche en santé du Québec; Canadian Institutes of Health
18 Research; Ministère de la Santé et des Services sociaux du Québec; Institut national de
19 santé publique du Québec; Health Canada; Centre hospitalier de l'Université de
20 Montréal; McGill University Health Centre; Centre hospitalier universitaire de Québec;
21 and Centre hospitalier universitaire de Sherbrooke. As the current study uses existing
22 information collected during the study, the funding sources did not have any role in the
23 design of the current study, its analysis, interpretation of data, and writing of the report.

References

- 1
2 1. Evans CT, Safdar N. Current Trends in the Epidemiology and Outcomes of
3 Clostridium difficile Infection. Clin Infect Dis **2015**; 60 Suppl 2: S66-71.
- 4 2. Martin JS, Monaghan TM, Wilcox MH. Clostridium difficile infection:
5 epidemiology, diagnosis and understanding transmission. Nat Rev Gastroenterol
6 Hepatol **2016**; 13(4): 206-16.
- 7 3. Cohen SH, Gerding DN, Johnson S, et al. Clinical practice guidelines for
8 Clostridium difficile infection in adults: 2010 update by the society for healthcare
9 epidemiology of America (SHEA) and the infectious diseases society of America
10 (IDSA). Infect Control Hosp Epidemiol **2010**; 31(5): 431-55.
- 11 4. Eyre DW, Cule ML, Wilson DJ, et al. Diverse sources of C. difficile infection
12 identified on whole-genome sequencing. N Engl J Med **2013**; 369(13): 1195-205.
- 13 5. Didelot X, Eyre DW, Cule M, et al. Microevolutionary analysis of Clostridium
14 difficile genomes to investigate transmission. Genome Biol **2012**; 13(12): R118.
- 15 6. Eyre DW, Golubchik T, Gordon NC, et al. A pilot study of rapid benchtop
16 sequencing of Staphylococcus aureus and Clostridium difficile for outbreak
17 detection and surveillance. BMJ Open **2012**; 2(3).
- 18 7. Kumar N, Miyajima F, He M, et al. Genome-Based Infection Tracking Reveals
19 Dynamics of Clostridium difficile Transmission and Disease Recurrence. Clin
20 Infect Dis **2016**; 62(6): 746-52.
- 21 8. Loo VG, Bourgault AM, Poirier L, et al. Host and pathogen factors for
22 Clostridium difficile infection and colonization. N Engl J Med **2011**; 365(18):
23 1693-703.

- 1 9. Clabots CR, Gerding SJ, Olson MM, Peterson LR, Gerding DN. Detection of
2 asymptomatic *Clostridium difficile* carriage by an alcohol shock procedure. *J Clin*
3 *Microbiol* **1989**; 27(10): 2386-7.
- 4 10. Spigaglia P, Mastrantonio P. Molecular analysis of the pathogenicity locus and
5 polymorphism in the putative negative regulator of toxin production (TcdC)
6 among *Clostridium difficile* clinical isolates. *J Clin Microbiol* **2002**; 40(9): 3470-
7 5.
- 8 11. Goncalves C, Decre D, Barbut F, Burghoffer B, Petit JC. Prevalence and
9 characterization of a binary toxin (actin-specific ADP-ribosyltransferase) from
10 *Clostridium difficile*. *J Clin Microbiol* **2004**; 42(5): 1933-9.
- 11 12. Fawley WN, Wilcox MH. Pulsed-field gel electrophoresis can yield DNA
12 fingerprints of degradation-susceptible *Clostridium difficile* strains. *J Clin*
13 *Microbiol* **2002**; 40(9): 3546-7; author reply 7.
- 14 13. Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA
15 restriction patterns produced by pulsed-field gel electrophoresis: criteria for
16 bacterial strain typing. *J Clin Microbiol* **1995**; 33(9): 2233-9.
- 17 14. Lunter G, Goodson M. Stampy: a statistical algorithm for sensitive and fast
18 mapping of Illumina sequence reads. *Genome Res* **2011**; 21(6): 936-9.
- 19 15. Sebahia M, Wren BW, Mullany P, et al. The multidrug-resistant human pathogen
20 *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet* **2006**; 38(7):
21 779-86.
- 22 16. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large
23 phylogenies by maximum likelihood. *Syst Biol* **2003**; 52(5): 696-704.

- 1 17. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in
2 whole bacterial genomes. *PLoS Comput Biol* **2015**; 11(2): e1004041.
- 3 18. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using
4 de Bruijn graphs. *Genome Res* **2008**; 18(5): 821-9.
- 5 19. Eyre DW, Griffiths D, Vaughan A, et al. Asymptomatic *Clostridium difficile*
6 colonisation and onward transmission. *PLoS One* **2013**; 8(11): e78445.
- 7 20. Walker AS, Eyre DW, Wyllie DH, et al. Characterisation of *Clostridium difficile*
8 hospital ward-based transmission using extensive epidemiological data and
9 molecular typing. *PLoS Med* **2012**; 9(2): e1001172.
- 10 21. Curry SR, Muto CA, Schlackman JL, et al. Use of multilocus variable number of
11 tandem repeats analysis genotyping to determine the role of asymptomatic carriers
12 in *Clostridium difficile* transmission. *Clin Infect Dis* **2013**; 57(8): 1094-102.
- 13 22. Longtin Y, Paquet-Bolduc B, Gilca R, et al. Effect of Detecting and Isolating
14 *Clostridium difficile* Carriers at Hospital Admission on the Incidence of *C*
15 *difficile* Infections: A Quasi-Experimental Controlled Study. *JAMA Intern Med*
16 **2016**; 176(6): 796-804.
- 17 23. Donskey CJ, Kundrapu S, Deshpande A. Colonization versus carriage of
18 *Clostridium difficile*. *Infect Dis Clin North Am* **2015**; 29(1): 13-28.
- 19 24. Eyre DW, Fawley WN, Rajgopal A, et al. Comparison of Control of *Clostridium*
20 *difficile* Infection in Six English Hospitals Using Whole-Genome Sequencing.
21 *Clin Infect Dis* **2017**.

22
23

- 1 Table 1. Proportions of CDI cases genetically and epidemiologically linked to prior
- 2 infected and colonized donors using WGS – all hospitals (201 cases)

	Genetically linked, n (%)	NAP1/027/ST1 among genetically linked donors, n (%)	Genetic and ward link, n (%)	NAP1/027/ST1 among genetically and ward linked donors, n (%)
Linked to prior case	105 (52)	95 (91)	81 (40)	74 (91)
Linked to infected patients only	28 (14)	23 (82)	34 (17)	31 (91)
Linked to colonized patients only	12 (6)	8 (67)	19 (10)	16 (84)
Linked to both infected and colonized patients	65 (32)	64 (99)	28 (14)	27 (96)

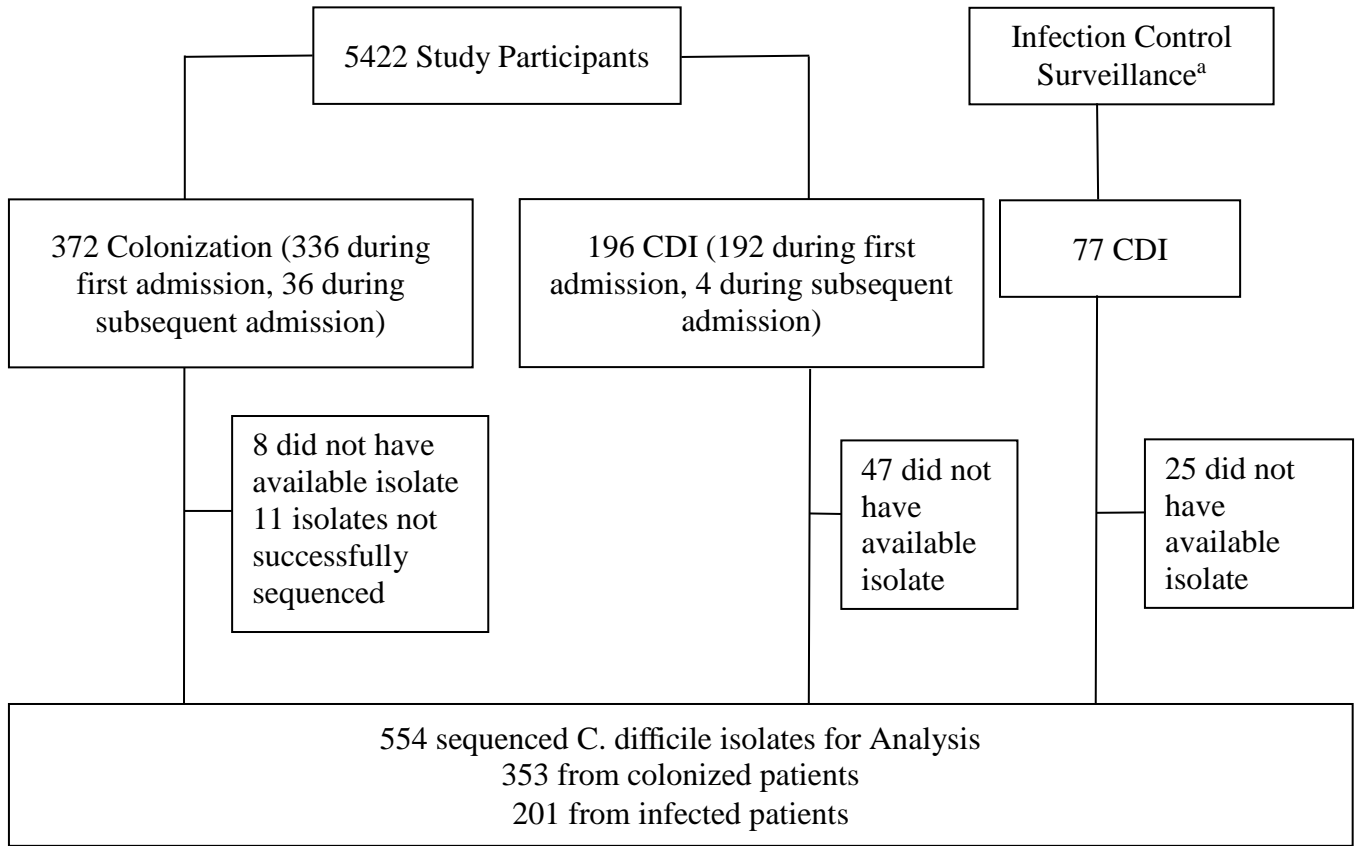
3

- 1 Table 2. Proportions of CDI cases genetically and epidemiologically linked to prior
- 2 infected and colonized donors using WGS – 2 hospitals (117 cases)

Possible source	Genetically linked, n (%)	NAP1/027/ST1 among genetically linked donors, n (%)	Genetic and ward link, n (%)	NAP1/027/ST1 among genetically and ward linked donors, n (%)
Linked to prior case	60 (51)	53 (88)	46 (39)	42 (91)
Linked to infected patients only	26 (22)	21 (81)	30 (26)	27 (90)
Linked to colonized patients only	4 (3)	3 (75)	2 (2)	2 (100)
Linked to both infected and colonized patients	30 (26)	29 (97)	14 (12)	13 (93)

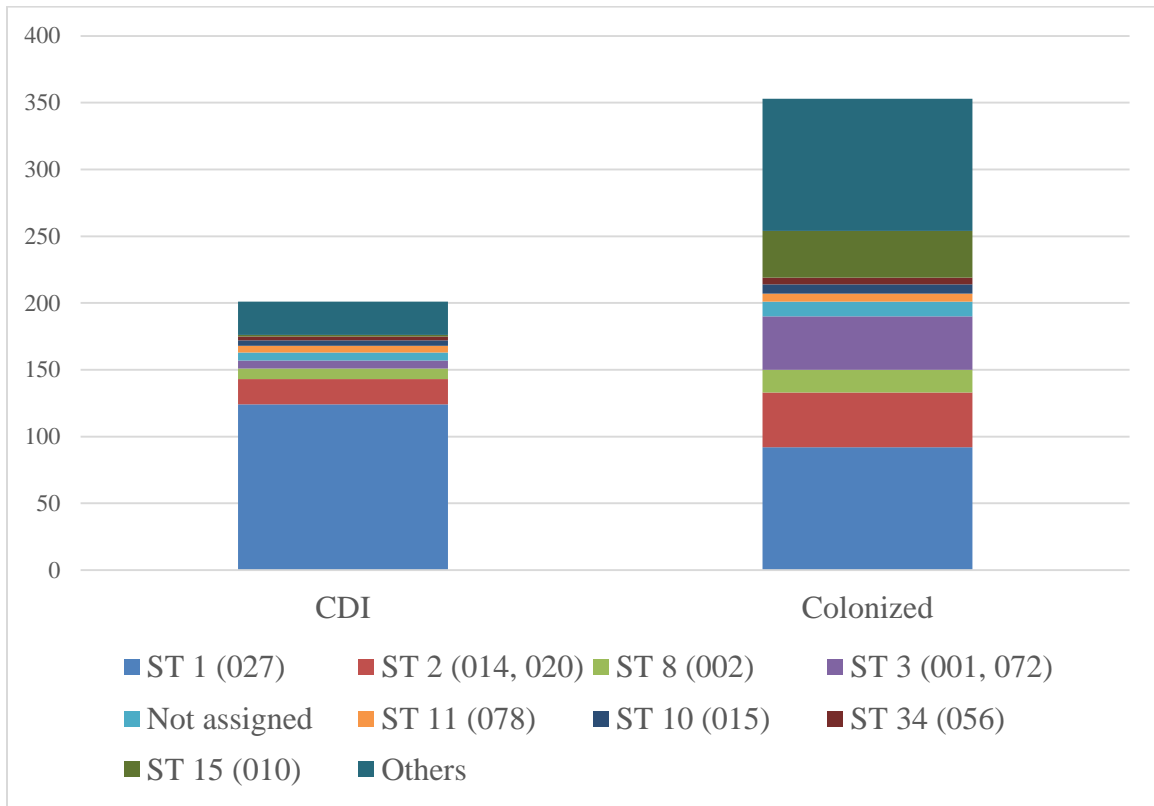
3

1 Figure 1. Flowchart of patients and isolates included in analysis



2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32 ^aInfection control surveillance isolates were available for one site
33
34
35
36

1 Figure 2. Multilocus sequence types by infected or colonized status



2

3 ST: Sequence type

4 PCR ribotype in parentheses