CHTyper, a Web Tool for Subtyping of Extraintestinal Pathogenic Escherichia coli based on the fumC and fimH Alleles

Roer, Louise; Johannesen, Thor Bech; Hansen, Frank; Stegger, Marc; Tchesnokova, Veronika; Sokurenko, Evgeni; Garibay, Natasha; Allesøe, Rosa Lundbye; Thomsen, Martin Christen Frølund; Lund, Ole; Hasman, Henrik; Hammerum, Anette M.

Published in:
Journal of Clinical Microbiology

Link to article, DOI:
10.1128/JCM.00063-18

Publication date:
2018

Document Version
Peer reviewed version

JCM Accepted Manuscript Posted Online 7 February 2018
J. Clin. Microbiol. doi:10.1128/JCM.00063-18

CHTyper, a Web Tool for Subtyping of Extraintestinal Pathogenic *Escherichia coli* based on the *fumC* and *fimH* Alleles

Louise ROER[a], Thor Bech JOHANNESEN[a] Frank HANSEN[a], Marc STEGGER[a], Veronika TCHESNOKOVA[b*], Evgeni SOKURENKO[b], Natasha GARIBAY[a], Rosa ALLESØE[c], Martin C. F. THOMSEN[c], Ole LUND[c], Henrik HASMAN[a] and Anette M. HAMMERUM[a#]

Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Copenhagen, Denmark[a];
Department of Microbiology, University of Washington, Seattle, USA[b];
Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark[c].

#Address correspondence to: Anette M. Hammerum, Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Artillerivej 5 (47/219), DK-2300 Copenhagen S, Denmark, Phone: +45 3268 3399, Fax: +45 3268 3231, E-mail: ama@ssi.dk

Sir,

Escherichia coli can cause a variety of extra-intestinal infections, such as urinary tract infection, meningitis, peritonitis and septicemia.

In 2012, Weissman *et al*. developed CH-typing, a two-locus, sequenced-based typing scheme, for a fast determination of sequence types (STs) and sub-ST clonal groups of extraintestinal pathogenic *E. coli* according to the multi-locus sequence typing (MLST) scheme (1). CH-typing is based on *fumC*, one of the household genes used in the seven-loci based MLST scheme (2), and an internal fragment of the type 1 fimbrial adhesin encoding gene, *fimH*. In May 2017, we published a web tool for subtyping of *E. coli* based on the *fimH* sequence (3). Here, we present a new web tool for CH-typing (https://cge.cbs.dtu.dk/services/chtyper/) based on both *fumC* and *fimH*, which allows users to obtain a CH-type from Sanger-generated sequences, fastq files, as well as assembled WGS data.

In the paper by Weissman *et al*., MLST and CH-typing were compared using 191 commensal and pathogenic *E. coli* isolates and 853 clinical *E. coli* isolates (2). Here, CH-Types and MLSTs were compared using assembled WGS data obtained from the Enterobase database on July 3rd 2017 (http://enterobase.warwick.ac.uk). Only *E. coli* genomes meeting the criteria of known MLSTs, according to the MLST scheme (1), and known *fimH allele* or *fimH-null* (isolates without *fimH*) were included in the analysis, resulting in 35,704 *E. coli* genomes from the Enterobase database. Discriminatory power was analyzed using the Simpsons index of diversity (D) (4).

The individual MLST loci exhibited between 240 to 428 alleles based on the available *E. coli* genomes obtained from EnteroBase, which resulted in 2,362 MLSTs, whereas the combination of *fumC* and *fimH* resulted in 1,187 unique CH-types (Table 1). The combination of *fumC* and *fimH* had a slightly higher discriminatory power (D = 0.9717 (CI;

0.9711-0.9723)) than the discriminatory power for MLST (D = 0.9606 (CI; 0.9596-0.9616). Similar observations were seen in the paper by Weissman *et al*. for the 191 commensal and pathogenic *E. coli* isolates (2).

To determine the resolution of CH-typing for clinical field application, CHTyper was used to analyze genomic data from 243 third-generation cephalosporin-resistant *E. coli* isolates obtained from patients with bloodstream infection (5). Here, 48 different STs were obtained. ST131 was most common (n=122) and 18 STs had more than one isolate. Using CHTyper, 70 CH-Types were obtained for the 243 *E. coli* isolates (Table 2). The CH-typing further subdivided 12 of the 18 STs, with more than one isolate, e.g. ST131 was subdivided into five CH-Types (Table 2).

Weissman *et al.* showed that specific CH-Types corresponded to specific STs and ST complexes with 95% accuracy, allowing good prediction of the MLST-based profile. Furthermore, CH-typing can detect the ST131 clonal subgroup *H*30, responsible for the current pandemic of fluoroquinolone and multi-drug resistant *E. coli* infections around the globe (6). Therefore, CH-typing can be used to study sub-ST clonal diversity or as a rapid screening test prior to selection for WGS.

In summary, CHTyper provides a highly suitable tool that can act as a rapid alternative to conventional MLST surveillance and for outbreak detection.

ACKNOWLEDGEMENTS

Karin Sixhøj Pedersen is thanked for her excellent technical assistance. Part of this work was supported by the Danish Ministry of Health as part of The Integrated Surveillance of ESBL/AmpC-producing *E. coli* and Carbapenemase-Producing Bacteria.

1. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M. 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol 60:1136–1151.
2. Weissman SJ, Johnson JR, Tchesnokova V, Billig M, Dykhuizen D, Riddell K, Rogers P, Qin X, Butler-Wu S, Cookson BT, Fang FC, Scholes D, Chattopadhyay S, Sokurenko E. 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic Escherichia coli. Appl Environ Microbiol 78:1353–60.
3. Roer L, Tchesnokova V, Allesoe R, Muradova M, Chattopadhyay S, Ahrenfeldt J, Thomsen MCF, Lund O, Hansen F, Hammerum AM, Sokurenko E, Hasman H. 2017. Development of a web tool for Escherichia coli subtyping based on fimh alleles. J Clin Microbiol 55.
4. Hunter PR, Gaston MA. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol 26:2465–2466.
5. 2015. DANMAP 2014 - Use of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from food animals, food and humans in Denmark.
6. Johnson JR, Tchesnokova V, Johnston B, Clabots C, Roberts PL, Billig M, Riddell K, Rogers P, Qin X, Butler-Wu S, Price LB, Aziz M, Nicolas-Chanoine M-H, Debroy C, Robicsek A, Hansen G, Urban C, Platell J, Trott DJ, Zhanel G, Weissman SJ, Cookson BT, Fang FC, Limaye AP, Scholes D, Chattopadhyay S, Hooper DC, Sokurenko E V. 2013. Abrupt emergence of a single dominant multidrug-resistant strain of Escherichia coli. J Infect Dis 207:919–28.

97 **TABLE 1.** Numbers of types found and D values of individual and combined loci of 35,704
98 *E. coli* isolates from EnteroBase

Typing method	No. of types found	D (95% CI)
Single loci and MLST		
adk	311	0.8762 (0.8740-0.8783)
fumC	428	0.8882 (0.8863-0.8900)
gyrB	318	0.9205 (0.9193-0.9217)
icd	356	0.9107 (0.9095-0.9119)
mdh	275	0.9096 (0.9085-0.9106)
purA	266	0.8646 (0.8627-0.8665)
recA	240	0.8449 (0.8425-0.8474)
ST	2362	0.9606 (0.9596-0.9616)
fimH +fimH0	300	0.9495 (0.9488-0.9502)
Loci paired with *fimH*		
adk + *fimH*	985	0.9704 (0.9698-0.9709)
fumC + *fimH*	1187	0.9717 (0.9711-0.9723)
gyrB + *fimH*	1110	0.9720 (0.9714-0.9726)
icd + *fimH*	1082	0.9714 (0.9707-0.9720)
mdh + *fimH*	984	0.9711 (0.9705-0.9717)
purA + *fimH*	925	0.9705 (0.9699-0.9711)
recA + *fimH*	891	0.9702 (0.9696-0.9708)
ST + *fimH*	3167	0.9768 (0.9762-0.9774)

99

100

101

102 **TABLE 2.** STs and CH-types for 243 third-generation cephalosporin-resistant *E. coli* isolates
103 obtained from patients with bloodstream infection

ST	CH-type(s) (n)
12	13-41 (1), 13-106 (4)
23	4-35 (1)
38	26-0 (2), 26-5 (14), 26-54 (1), 26-65 (1)
44	11-54 (2)
58	4-27 (1), 4-30 (2), 4-32 (1)
69	35-27 (10)
73	24-10 (1), 24-30 (1), 24-103 (1)
88	4-39 (1), 4-43 (1)
90	4-142 (1)
93	11-41 (1)
95	38-15 (1), 38-27 (1), 38-41 (2), 38-483 (1)
117	45-97 (1)
127	14-2 (2)
131	40-22 (1), 40-27 (14), 40-30 (95), 40-35 (1), 40-41 (11)
135	39-2 (1)
141	52-5 (1)
167	11-0 (3), 11-215 (1)
205	23-54 (1)
209	11-54 (1)
345	4-31 (1)
349	36-54 (1)
354	88-58 (1)
393	106-54 (1)
405	37- 27 (10), 37-29 (3)

410	4-24 (4)
421	38-0 (1)
443	19-24 (1)
450	11-34 (1), 11-54 (2)
453	6-31 (1)
550	14-54 (1)
603	4-517 (1)
617	11-0 (1), 11-29 (1)
624	4-27 (1)
636	108-0 (1)
648	4-0 (4), 4-27 (4)
977	188-25 (1)
1163	45-63 (1)
1177	26-65 (1)
1193	14-64 (2)
1248	29-31 (1)
1706	29-38 (1)
2509	95-60 (1)
2522	29-38 (1)
3014	41-34 (1)
3057	54-445 (1)
3285	6-35 (1)
3666	26-5 (3)
3995	4-27 (1)
5824	11-0 (1)

104

105