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
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
CHTyper, a Web Tool for Subtyping of Extraintestinal Pathogenic Escherichia coli Based on the fumC and fimH Alleles

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

Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Copenhagen, Denmark
Department of Microbiology, University of Washington, Seattle, Washington, USA
Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark

KEYWORDS Escherichia coli, WGS

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

In 2012, Weissman et al. developed fumC fimH (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 strains according to the multilocus sequence typing (MLST) scheme (1). CH typing is based on fumC, one of the household genes used in the seven-locus-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 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 sequencing-generated sequences and fastq files, as well as assembled whole-genome sequencing (WGS) data.

In the paper by Weissman et al., the results of MLST and CH typing were compared using 191 commensal and pathogenic E. coli isolates and 853 clinical E. coli isolates (2).

TABLE 1

<table>
<thead>
<tr>
<th>Typing method</th>
<th>No. of types found</th>
<th>D (95% confidence interval)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Single loci or MLST</th>
</tr>
</thead>
<tbody>
<tr>
<td>adk</td>
</tr>
<tr>
<td>fumC</td>
</tr>
<tr>
<td>gyrB</td>
</tr>
<tr>
<td>icd</td>
</tr>
<tr>
<td>mdh</td>
</tr>
<tr>
<td>purA</td>
</tr>
<tr>
<td>recA</td>
</tr>
<tr>
<td>ST</td>
</tr>
<tr>
<td>fimH + fimH0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loci or ST paired with fimH</th>
</tr>
</thead>
<tbody>
<tr>
<td>adk + fimH</td>
</tr>
<tr>
<td>fumC + fimH</td>
</tr>
<tr>
<td>gyrB + fimH</td>
</tr>
<tr>
<td>icd + fimH</td>
</tr>
<tr>
<td>mdh + fimH</td>
</tr>
<tr>
<td>purA + fimH</td>
</tr>
<tr>
<td>recA + fimH</td>
</tr>
<tr>
<td>ST + fimH</td>
</tr>
</tbody>
</table>
Here, CH types and MLSTs were compared using assembled WGS data obtained from the EnteroBase database on 3 July 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 (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 and 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 (\(D = 0.9717\) [confidence interval, 0.9711 to 0.9723]) than the discriminatory power of MLST

<table>
<thead>
<tr>
<th>ST</th>
<th>CH type(s) (no. of isolates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>13-41 (1), 13-106 (4)</td>
</tr>
<tr>
<td>23</td>
<td>4-35 (1)</td>
</tr>
<tr>
<td>38</td>
<td>26-0 (2), 26-5 (14), 26-54 (1), 26-65 (1)</td>
</tr>
<tr>
<td>44</td>
<td>11-54 (2)</td>
</tr>
<tr>
<td>58</td>
<td>4-27 (1), 4-30 (2), 4-32 (1)</td>
</tr>
<tr>
<td>69</td>
<td>35-27 (10)</td>
</tr>
<tr>
<td>73</td>
<td>24-10 (1), 24-30 (1), 24-103 (1)</td>
</tr>
<tr>
<td>88</td>
<td>4-39 (1), 4-43 (1)</td>
</tr>
<tr>
<td>90</td>
<td>4-142 (1)</td>
</tr>
<tr>
<td>93</td>
<td>11-41 (1)</td>
</tr>
<tr>
<td>95</td>
<td>38-15 (1), 38-27 (1), 38-41 (2), 38-483 (1)</td>
</tr>
<tr>
<td>117</td>
<td>45-97 (1)</td>
</tr>
<tr>
<td>127</td>
<td>14-2 (2)</td>
</tr>
<tr>
<td>131</td>
<td>40-22 (1), 40-27 (14), 40-30 (95), 40-35 (1), 40-41 (11)</td>
</tr>
<tr>
<td>135</td>
<td>39-2 (1)</td>
</tr>
<tr>
<td>141</td>
<td>52-5 (1)</td>
</tr>
<tr>
<td>167</td>
<td>11-0 (3), 11-215 (1)</td>
</tr>
<tr>
<td>205</td>
<td>23-54 (1)</td>
</tr>
<tr>
<td>209</td>
<td>11-54 (1)</td>
</tr>
<tr>
<td>345</td>
<td>4-31 (1)</td>
</tr>
<tr>
<td>349</td>
<td>36-54 (1)</td>
</tr>
<tr>
<td>354</td>
<td>88-58 (1)</td>
</tr>
<tr>
<td>393</td>
<td>106-54 (1)</td>
</tr>
<tr>
<td>405</td>
<td>37-27 (10), 37-29 (3)</td>
</tr>
<tr>
<td>410</td>
<td>4-24 (4)</td>
</tr>
<tr>
<td>421</td>
<td>38-0 (1)</td>
</tr>
<tr>
<td>443</td>
<td>19-24 (1)</td>
</tr>
<tr>
<td>450</td>
<td>11-34 (1), 11-54 (2)</td>
</tr>
<tr>
<td>453</td>
<td>6-31 (1)</td>
</tr>
<tr>
<td>550</td>
<td>14-54 (1)</td>
</tr>
<tr>
<td>603</td>
<td>4-517 (1)</td>
</tr>
<tr>
<td>617</td>
<td>11-0 (1), 11-29 (1)</td>
</tr>
<tr>
<td>624</td>
<td>4-27 (1)</td>
</tr>
<tr>
<td>636</td>
<td>108-0 (1)</td>
</tr>
<tr>
<td>648</td>
<td>4-0 (4), 4-27 (4)</td>
</tr>
<tr>
<td>977</td>
<td>188-25 (1)</td>
</tr>
<tr>
<td>1163</td>
<td>45-63 (1)</td>
</tr>
<tr>
<td>1177</td>
<td>26-65 (1)</td>
</tr>
<tr>
<td>1193</td>
<td>14-64 (2)</td>
</tr>
<tr>
<td>1248</td>
<td>29-31 (1)</td>
</tr>
<tr>
<td>1706</td>
<td>29-38 (1)</td>
</tr>
<tr>
<td>2509</td>
<td>95-60 (1)</td>
</tr>
<tr>
<td>2522</td>
<td>29-38 (1)</td>
</tr>
<tr>
<td>3014</td>
<td>41-34 (1)</td>
</tr>
<tr>
<td>3057</td>
<td>54-445 (1)</td>
</tr>
<tr>
<td>3285</td>
<td>6-35 (1)</td>
</tr>
<tr>
<td>3666</td>
<td>26-5 (3)</td>
</tr>
<tr>
<td>3995</td>
<td>4-27 (1)</td>
</tr>
<tr>
<td>5824</td>
<td>11-0 (1)</td>
</tr>
</tbody>
</table>
(D = 0.9606) (confidence interval, 0.9596 to 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 E. coli isolates that were resistant to third-generation cephalosporins and obtained from patients with bloodstream infection (5). Here, 48 different STs were obtained. ST131 was the most common (n = 122), and 18 STs were represented by more than one isolate. Using CHTyper, 70 CH types were obtained for the 243 E. coli isolates (Table 2). CH typing further subdivided 12 of the 18 STs represented by more than one isolate; e.g., ST131 was subdivided into 5 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 H30, responsible for the current pandemic of fluoroquinolone- and multidrug-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 is a highly suitable tool that can act as a rapid alternative to conventional MLST surveillance and for outbreak detection.

ACKNOWLEDGMENTS
Karina 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.

REFERENCES