Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli.

Joensen, Katrine Grimstrup; Scheutz, Flemming; Lund, Ole; Hasman, Henrik; Kaas, Rolf Sommer; Nielsen, Eva M.; Aarestrup, Frank Møller

Published in:
Journal of Clinical Microbiology

Link to article, DOI:
10.1128/JCM.03617-13

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic Escherichia coli

Katrine Grimstrup Joensen, Flemming Scheutz, Ole Lund, Henrik Hasman, Rolf S. Kaas, Eva M. Nielsen and Frank M. Aarestrup

PUBLISHED AHEAD OF PRINT 26 February 2014.

Updated information and services can be found at:
http://jcm.asm.org/content/52/5/1501

**SUPPLEMENTAL MATERIAL**

These include:

Supplemental material

**REFERENCES**

This article cites 33 articles, 16 of which can be accessed free at: http://jcm.asm.org/content/52/5/1501#ref-list-1

**CONTENT ALERTS**

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Fast and accurate identification and typing of pathogens are essential for effective surveillance and outbreak detection. The current routine procedure is based on a variety of techniques, making the procedure laborious, time-consuming, and expensive. With whole-genome sequencing (WGS) becoming cheaper, it has huge potential in both diagnostics and routine surveillance. The aim of this study was to perform a real-time evaluation of WGS for routine typing and surveillance of verocytotoxin-producing *Escherichia coli* (VTEC). In Denmark, the Statens Serum Institut (SSI) routinely receives all suspected VTEC isolates. During a 7-week period in the fall of 2012, all incoming isolates were concurrently subjected to WGS using IonTorrent PGM. Real-time bioinformatics analysis was performed using web-tools (www.genomicepidemiology.org) for species determination, multilocus sequence type (MLST) typing, and determination of phylogenetic relationship, and a specific VirulenceFinder for detection of *E. coli* virulence genes was developed as part of this study. In total, 46 suspected VTEC isolates were characterized in parallel during the study. VirulenceFinder proved successful in detecting virulence genes included in routine typing, explicitly verocytotoxin 1 (*vtx1*), verocytotoxin 2 (*vtx2*), and intimin (*eae*), and also detected additional virulence genes. VirulenceFinder is also a robust method for assigning verocytotoxin (*vtx*) subtypes. A real-time clustering of isolates in agreement with the epidemiology was established from WGS, enabling discrimination between sporadic and outbreak isolates. Overall, WGS typing produced results faster and at a lower cost than the current routine. Therefore, WGS typing is a superior alternative to conventional typing strategies. This approach may also be applied to typing and surveillance of other pathogens.

Bacterial pathogens still pose a major threat to public health, and in order to limit their spread and prevent infectious disease outbreaks, accurate and rapid diagnostics and classification of isolates are of great importance. In current routine practice, isolation and identification are mostly performed at clinical microbiological laboratories, and verification and further characterization are performed for a few selected pathogens at national, or regional, reference laboratories, using a variety of species-specific methods. Typing and surveillance of bacterial pathogens rely mainly on well-established, standardized phenotypic and molecular typing methods, such as serotyping and pulsed-field gel electrophoresis (PFGE) (1, 2). However, to obtain sufficient discrimination between isolates, it is typically necessary to combine typing results from several different typing techniques, both phenotypic and genotypic. As a result, it is laborious, time-consuming, and expensive to perform proper typing for surveillance and outbreak detection.

As the cost of whole-genome sequencing (WGS) has decreased and benchtop sequencing machines enable fast turnaround, it has become increasingly attractive for use in routine diagnostics and typing, and the approach has already been found useful in retrospective outbreak investigations (3, 4).

Although WGS provides detailed information that will, in theory, enable diagnostics and typing solely on the basis of the features in the bacterial genome, it is a challenge to define and extract the appropriate information from the large amount of sequence data that is generated. Thus, to facilitate the use of WGS data for routine diagnostics, typing, and surveillance, it is important that the sequence data can be automatically and quickly converted to clinically relevant information that can be easily interpreted by physicians and public health professionals with limited bioinformatics skills. To achieve this, the Center for Genomic Epidemiology (CGE) provides public, user-friendly web-tools for rapid handling of WGS data and extraction of relevant information, useful for diagnostics, surveillance, and outbreak investigations for the global medical society (www.genomicepidemiology.org).

In this study, as a proof of concept, we demonstrate the usefulness of WGS for routine typing, surveillance, and outbreak detection of verocytotoxin-producing *Escherichia coli* (VTEC) infections in Denmark. VTEC, also known as Shiga toxin-producing *E. coli* (STEC), is a gastrointestinal pathogen, which is typically spread by ingestion of contaminated food or water or person-to-person contact. Rapid and reliable diagnostics and detection of outbreak clusters are of utmost importance for control. VTEC infections cause bloody diarrhea and in some cases hemolytic-uremic syndrome (HUS), which is characterized by kidney failure, thrombocytopenia, and microangiopathic hemolytic anemia, and can be fatal to young children. VTEC pathogenicity is facili-
tated by the Shiga toxin (Stx) and a number of other virulence factors (5, 6).

We performed real-time benchtop sequencing of VTEC isolates from patients in Denmark during a 7-week period in parallel with the current routine procedure. During the period of the study, Denmark experienced a small _E. coli_ O157:H7 outbreak with a total of 13 VTEC isolates. The VTEC outbreak strain had an infrequent toxin subtype profile (vtx, vtx1a, and vtx2a) and a high proportion of HUS cases (62%), and the toxin subtype profiling proved important in the outbreak investigation and risk assessment of the VTEC strain (7).

Here, we demonstrate that this WGS-based typing approach is a superior alternative to the current routine typing of VTEC infections, rapidly producing typing results comparable to those of current routine typing and thus of great value in surveillance and outbreak detection. This approach may also be applicable to other pathogens. In addition, we here present the VirulenceFinder, a new CGE web-tool made for automatic detection and extraction of _E. coli_ virulence genes from WGS data.

**MATERIALS AND METHODS**

**Study design and isolates.** The use of whole-genome sequencing (WGS) for routine typing purposes of verocytotoxin-producing _Escherichia coli_ (VTEC) isolates was examined by conducting a WGS study parallel with the routine typing performed at the Statens Serum Institut (SSI) to evaluate WGS for typing purposes with regard to typing results, time to final result, labor time, and total cost. Suspected VTEC isolates from infected patients were, as part of the routine surveillance in Denmark, sent from hospitals to the SSI for confirmation and for further phenotypic and molecular characterization. Simultaneously, the isolates were typed by the WGS approach at DTU Food (National Food Institute, Technical University of Denmark), using the IonTorrent PGM (Life Technologies, Carlsbad, CA) benchtop sequencer. All isolates included in the study were obtained from human fecal samples from patients with regular diarrhea, bloody diarrhea, or hemolytic-uremic syndrome (HUS), and epidemiological information was collected by the SSI as standard procedure. We defined the test period in the fall of 2012 to last until a total of at least 40 isolates were collected, but with a maximum duration of 12 weeks, in case fewer isolates were received than expected.

**Routine typing procedures.** As part of the routine typing and surveillance at the SSI, suspected VTEC isolates were subjected to several phenotypic and molecular typing methods according to the standard procedures at the SSI. In instances where the isolates received at the SSI exhibited mixed-colony morphology, both types of isolates were subjected to typing and included in the study.

All suspected VTEC isolates were serotyped by O typing and H typing, identifying the specific cell wall antigen and flagellar antigen, respectively (8, 9). In addition, the isolates were K typed detecting capsule antigens by K1/K5 bacteriophage susceptibility (8). All isolates were additionally tested for hemolysin production (10) and β-glucuronidase activity (11) and for the production of verocytotoxins by the Vero cell assay (12).

All the suspected VTEC isolates were subjected to DNA hybridization with specific DNA probes for the _E. coli_ attaching and effacing gene (aee) (13), bundle-forming pilus gene (bfpA) (14), enteropathogenic _E. coli_ (EPEC) adherence factor (EAF) (15), plasmid-encoded O157 enterohe-molysin gene (ehxA) (16), verotoxin 1 gene (vtx1) (17), verotoxin 2 gene (vtx2) (18), the verotoxin 2F variant gene (vtx2f) (19), and Shiga toxin-producing _E. coli_ (STEC) autoagglutinating adhesion gene (saa) (20). In addition, subtyping of vtx1 and vtx2 was carried out for all isolates by PCR with subtype-specific primers detecting vtx1a, vtx1c, vtx1d, vtx2a, vtx2b, vtx2c, vtx2d, vtx2e, vtx2f, and vtx2g subtypes (21). Only isolates considered to be potential outbreak isolates based on the typing methods described above were additionally typed by PFGE (22).

**Whole-genome sequencing.** Genomic DNA (gDNA) was purified from the isolates using the Easy-DNA extraction kit (Invitrogen, Carlsbad, CA), and DNA concentrations were determined using the Qubit dsDNA (double-stranded DNA) BR assay kit (Invitrogen). Subsequently, gDNA was fragmented by sonication on the Covaris S2 system. Specifically, 130 μl (200 ng) gDNA diluted in low-TE (Tris-EDTA) buffer was used as input, and conditions were set for generation of 200- to 300-bp fragments, using 3 cycles each of 1 min, 5°C bath temperature, frequency sweeping, 10% duty cycling, intensity of 5, and cycles/burst of 100. A total of 100 ng gDNA was introduced to the IonTorrent PGM 200-bp work flow. gDNA libraries were prepared according to the IonXpress Plus gDNA Fragment Library Preparation (Life Technologies) protocol, consisting of end repair, nick repair, ligation, and size selecting with the E-Gel SizeSelect agarose gel system and subsequent library amplification. Library concentration was determined employing a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and the Agilent high-sensitivity DNA kit. Template preparation was done according to the Ion OneTouch 200 template kit (Life Technologies) protocol for 200-base-read libraries on the Ion OneTouch system (Life Technologies) and subsequent quality control of template Ion Spheres using the Qubit 2.0 fluorometer (Life Technologies). Sequencing was done on the IonTorrent PGM sequencer following the Ion PGM 200 sequencing kit (Life Technologies) protocol for either Ion 316 chips or Ion 318.

**The _E. coli_ virulence gene database.** For automatic detection of virulence genes in the suspected VTEC isolates, an _E. coli_ FASTA database was constructed, as part of VirulenceFinder, which is a component of the publicly available web-based tools for WGS analysis hosted by the Center for Genomic Epidemiology (CGE) (http://www.genomicepidemiology.org/). The content of the _E. coli_ virulence database was constructed on the basis of the Identibac scheme (Alere Technologies GmbH, Jena, Germany) for genotyping detection of _E. coli_ virulence genes. All genes, and gene variants, represented in the Identibac scheme by GenBank accession numbers or identifiers (IDs) were BLASTed against the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore/), and gene variants that matched 90% on identity and size were collected. All partial genes were excluded from the database, and also gene variants belonging to genera other than _Escherichia, Klebsiella, Citrobacter, Enterobacter, or Shigella_ were excluded. Subtypes of the vtx genes were assigned according to the previously described, sequence-based nomenclature for verocytotoxins (21). The _E. coli_ virulence gene database contains 76 genes, and the gene content and number of gene variants can be seen in Table 1.

**Using VirulenceFinder.** VirulenceFinder was constructed to enable detection of virulence genes related to _E. coli_ in WGS data while being simple and user friendly. Sequence data can be submitted either as assembled genomes or raw reads from various sequencing technologies, since assembly of raw sequence reads is incorporated into the tool, as described previously for other CGE tools (23, 24). It is possible to select configurations for the organism of interest, and in addition, it is possible to select percent identity (%ID) threshold between the input and the best matching database gene.

The output consists of best-matching genes from BLAST analysis of the selected database, against the submitted genome, with genes set to cover a minimum of three-fifths of the length of the database genes (24). The output contains information on the virulence gene, the %ID, the length of query and database gene, the position of the hit in the contig, and the accession number of the hit. In addition, the vtx subtypes are outputted for typing purposes.

**Analysis of sequence data.** Sequence data were analyzed, without further processing, using the CGE web-tools for species detection using the KmerFinder, for determination of multilocus sequence types (MLSTs) (23) and by employing VirulenceFinder for detection of _E. coli_ virulence genes and vtx1 and vtx2 subtypes. For VirulenceFinder, the configuration was set for the _E. coli_ database with an 85.00%ID threshold. For the MLST tool, the configuration was set to _E. coli_ scheme 1 (23, 25).

KmerFinder is a novel program/database for rapid species identifica-
TABLE 1 Gene content of the E. coli virulence database

<table>
<thead>
<tr>
<th>Gene</th>
<th>Descriptiona</th>
<th>No. of variants in the database</th>
</tr>
</thead>
<tbody>
<tr>
<td>astA</td>
<td>Heat-stable enterotoxin 1</td>
<td>11</td>
</tr>
<tr>
<td>bfpA</td>
<td>Major subunit of bundle-forming pili</td>
<td>5</td>
</tr>
<tr>
<td>cba</td>
<td>Colicin B</td>
<td>15</td>
</tr>
<tr>
<td>cci</td>
<td>Cloacin</td>
<td>4</td>
</tr>
<tr>
<td>cdtB</td>
<td>Cytolethal distending toxin B</td>
<td>14</td>
</tr>
<tr>
<td>celB</td>
<td>Endonuclease colicin E2</td>
<td>10</td>
</tr>
<tr>
<td>cfa_c</td>
<td>Colonization factor antigen I</td>
<td>4</td>
</tr>
<tr>
<td>cif</td>
<td>Type III secreted effector</td>
<td>4</td>
</tr>
<tr>
<td>cma</td>
<td>Colicin M</td>
<td>19</td>
</tr>
<tr>
<td>cndf</td>
<td>Cytoxic necrotizing factor</td>
<td>7</td>
</tr>
<tr>
<td>cofA</td>
<td>Longus type IV pilus subunit</td>
<td>1</td>
</tr>
<tr>
<td>eae</td>
<td>Intimin</td>
<td>45</td>
</tr>
<tr>
<td>eatA</td>
<td>Serine protease autotransporters of Enterobacteriaceae (SPATE)</td>
<td>3</td>
</tr>
<tr>
<td>efaI</td>
<td>EHEC factor for adherence</td>
<td>11</td>
</tr>
<tr>
<td>ehxA</td>
<td>Enterohemolysin</td>
<td>12</td>
</tr>
<tr>
<td>epeA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>1</td>
</tr>
<tr>
<td>espA</td>
<td>Type III secretion system</td>
<td>23</td>
</tr>
<tr>
<td>espB</td>
<td>Secreted protein B</td>
<td>14</td>
</tr>
<tr>
<td>espC</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>3</td>
</tr>
<tr>
<td>espE</td>
<td>Type III secretion system</td>
<td>13</td>
</tr>
<tr>
<td>espF</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>2</td>
</tr>
<tr>
<td>espI</td>
<td>Prophage-encoded type III secretion system effector</td>
<td>2</td>
</tr>
<tr>
<td>espP</td>
<td>Putative exoprotein precursor</td>
<td>4</td>
</tr>
<tr>
<td>etpD</td>
<td>Type II secretion protein</td>
<td>3</td>
</tr>
<tr>
<td>f17A</td>
<td>Subunit A of F17 fimbral protein</td>
<td>7</td>
</tr>
<tr>
<td>f17G</td>
<td>Adhesin subunit of F17 fimbrae</td>
<td>9</td>
</tr>
<tr>
<td>fanA</td>
<td>Involved in biogenesis of K99/F5 fimbriae</td>
<td>1</td>
</tr>
<tr>
<td>fasA</td>
<td>Fimbrial 987P/F6 subunit</td>
<td>1</td>
</tr>
<tr>
<td>fedA</td>
<td>Fimbrial protein F107 subunit A</td>
<td>3</td>
</tr>
<tr>
<td>fedB</td>
<td>Fimbrial adhesin AC precursor</td>
<td>6</td>
</tr>
<tr>
<td>fimH14a</td>
<td>Mature FimH14a/F41 protein</td>
<td>2</td>
</tr>
<tr>
<td>gud</td>
<td>Glutamate decarboxylase</td>
<td>70</td>
</tr>
<tr>
<td>hlyE</td>
<td>Avian E. coli hemolysin</td>
<td>1</td>
</tr>
<tr>
<td>iha</td>
<td>Adherence protein</td>
<td>19</td>
</tr>
<tr>
<td>ipaD</td>
<td>Invasion protein Shigella flexneri</td>
<td>9</td>
</tr>
<tr>
<td>ipaH9.8</td>
<td>Invasion plasmid antigen</td>
<td>8</td>
</tr>
<tr>
<td>ircA</td>
<td>Siderophore receptor</td>
<td>4</td>
</tr>
<tr>
<td>irnA</td>
<td>Enterobactin siderophore receptor protein</td>
<td>13</td>
</tr>
<tr>
<td>iss</td>
<td>Increased serum survival</td>
<td>14</td>
</tr>
<tr>
<td>K88a/b</td>
<td>K88/F4 subunit protein</td>
<td>10</td>
</tr>
<tr>
<td>katP</td>
<td>Plasmid-encoded catalase peroxidase</td>
<td>1</td>
</tr>
<tr>
<td>lngA</td>
<td>Longus type IV pilus</td>
<td>2</td>
</tr>
<tr>
<td>lpfA</td>
<td>Long polar fimbriae</td>
<td>11</td>
</tr>
<tr>
<td>ltcA</td>
<td>Heat-labile enterotoxin A subunit</td>
<td>17</td>
</tr>
<tr>
<td>mchB</td>
<td>Microcin H47 part of colicin H</td>
<td>2</td>
</tr>
<tr>
<td>mchC</td>
<td>Microcin protein</td>
<td>6</td>
</tr>
<tr>
<td>mchF</td>
<td>ABC transporter protein MchF</td>
<td>15</td>
</tr>
<tr>
<td>mcmA</td>
<td>Microcin M part of colicin H</td>
<td>4</td>
</tr>
<tr>
<td>nfaE</td>
<td>Diffuse adherence fibrillar adhesin gene</td>
<td>5</td>
</tr>
<tr>
<td>nleA</td>
<td>Non-LEE-encoded effector A</td>
<td>18</td>
</tr>
<tr>
<td>nleB</td>
<td>Non-LEE-encoded effector B</td>
<td>14</td>
</tr>
<tr>
<td>nleC</td>
<td>Non-LEE-encoded effector C</td>
<td>6</td>
</tr>
<tr>
<td>pet</td>
<td>Enterohaemolysin</td>
<td>19</td>
</tr>
<tr>
<td>provB</td>
<td>P-related fimbrial regulatory gene</td>
<td>22</td>
</tr>
<tr>
<td>prestige</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>1</td>
</tr>
<tr>
<td>sat</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>6</td>
</tr>
<tr>
<td>senB</td>
<td>Plasmid-encoded enterotoxin</td>
<td>3</td>
</tr>
<tr>
<td>sepA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>7</td>
</tr>
<tr>
<td>sfaS</td>
<td>S-fimbrial minor subunit</td>
<td>1</td>
</tr>
<tr>
<td>sigA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>2</td>
</tr>
<tr>
<td>staI</td>
<td>Heat-stable enterotoxin ST-Ia</td>
<td>2</td>
</tr>
<tr>
<td>stk</td>
<td>Heat-stable enterotoxin II</td>
<td>3</td>
</tr>
<tr>
<td>stxlA</td>
<td>Shiga-like toxin 1 A-subunit</td>
<td>18</td>
</tr>
<tr>
<td>stxB</td>
<td>Shiga-like toxin 1 B-subunit</td>
<td>14</td>
</tr>
<tr>
<td>stx2A</td>
<td>Shiga toxin 2 subunit A</td>
<td>114</td>
</tr>
<tr>
<td>stx2B</td>
<td>Shiga toxin 2 subunit B</td>
<td>43</td>
</tr>
<tr>
<td>subA</td>
<td>Subtilase toxin subunit</td>
<td>5</td>
</tr>
<tr>
<td>saa</td>
<td>STEC autoagglutinating adhesin</td>
<td>1</td>
</tr>
<tr>
<td>tcpP</td>
<td>TIR cytoskeleton coupling protein</td>
<td>34</td>
</tr>
<tr>
<td>tir</td>
<td>Translocated intimin receptor protein</td>
<td>36</td>
</tr>
<tr>
<td>toxB</td>
<td>Toxin B</td>
<td>4</td>
</tr>
<tr>
<td>tsh</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>3</td>
</tr>
<tr>
<td>vat</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>7</td>
</tr>
<tr>
<td>vifF</td>
<td>VirF transcriptional activator</td>
<td>3</td>
</tr>
</tbody>
</table>

a LEE, locus of enterocyte effacement.

TABLE 1 (Continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Descriptiona</th>
<th>No. of variants in the database</th>
</tr>
</thead>
<tbody>
<tr>
<td>prfB</td>
<td>P-related fimbrial regulatory gene</td>
<td>22</td>
</tr>
<tr>
<td>rpeA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>1</td>
</tr>
<tr>
<td>sat</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>6</td>
</tr>
<tr>
<td>senB</td>
<td>Plasmid-encoded enterotoxin</td>
<td>3</td>
</tr>
<tr>
<td>sepA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>7</td>
</tr>
<tr>
<td>sfaS</td>
<td>S-fimbrial minor subunit</td>
<td>1</td>
</tr>
<tr>
<td>sigA</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>2</td>
</tr>
<tr>
<td>staI</td>
<td>Heat-stable enterotoxin ST-Ia</td>
<td>2</td>
</tr>
<tr>
<td>stk</td>
<td>Heat-stable enterotoxin II</td>
<td>3</td>
</tr>
<tr>
<td>stxlA</td>
<td>Shiga-like toxin 1 A-subunit</td>
<td>18</td>
</tr>
<tr>
<td>stxB</td>
<td>Shiga-like toxin 1 B-subunit</td>
<td>14</td>
</tr>
<tr>
<td>stx2A</td>
<td>Shiga toxin 2 subunit A</td>
<td>114</td>
</tr>
<tr>
<td>stx2B</td>
<td>Shiga toxin 2 subunit B</td>
<td>43</td>
</tr>
<tr>
<td>subA</td>
<td>Subtilase toxin subunit</td>
<td>5</td>
</tr>
<tr>
<td>saa</td>
<td>STEC autoagglutinating adhesin</td>
<td>1</td>
</tr>
<tr>
<td>tcpP</td>
<td>TIR cytoskeleton coupling protein</td>
<td>34</td>
</tr>
<tr>
<td>tir</td>
<td>Translocated intimin receptor protein</td>
<td>36</td>
</tr>
<tr>
<td>toxB</td>
<td>Toxin B</td>
<td>4</td>
</tr>
<tr>
<td>tsh</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>3</td>
</tr>
<tr>
<td>vat</td>
<td>Serine protease autotransporters of Enterobacteriaceae</td>
<td>7</td>
</tr>
<tr>
<td>vifF</td>
<td>VirF transcriptional activator</td>
<td>3</td>
</tr>
</tbody>
</table>
evaluating the number of reads with the most common nucleotide at the specific position, $X$, in relation to the number of reads with other nucleotides at the position, $Y$. A Z-score was calculated as $Z = (X - Y) / \sqrt{X + Y}$, and $Z = 3.29$ was used as the threshold, and additionally, nucleotide differences were considered only when the most common nucleotide was at least 10 times more abundant than other nucleotides at the position. Positions with nonsignificant variations were assigned $N$, and the same for non-mapped positions of the reference genome.

Sequences were compared in pairs, nucleotide differences in positions were counted, and the tree was constructed by the unweighted-pair group method using average linkages (UPGMA) algorithm from the neighbor program in the phylip package (http://evolution.genetics.washington.edu/phylip.html). The phylogenetic trees were rooted by midpoints.

**Evaluation of WGS for typing and surveillance of VTEC.** The WGS-derived typing data were, for each suspected VTEC isolate, compared to the typing data received from the routine procedure at the SSI. The phylogenies obtained from the two different methods were compared and related to the epidemiological information received from the SSI to evaluate the ability to discriminate correctly between isolates. In addition, the application of WGS for typing and surveillance of VTEC infections was assessed by comparing the WGS-based approach and the routine typing with respect to hands-on time, time for obtaining typing results, and estimated cost per isolate.

**RESULTS**

The study included all suspected VTEC isolates received at the SSI during this time period. An increased number of isolates were received during the period of the study due to the occurrence of a small O157:H7 outbreak (7), and thus, both sporadic isolates and outbreak isolates (C812-12, C818-12, C819-12, C849-12, C852-12, and C863-12) were available for evaluation of the WGS-based typing method.

During the 7 weeks, a total of 42 isolates were received at the SSI for further characterization and subtyping, and since 4 of these isolates exhibited mixed-colony morphology upon visual inspection of primary plates, a total of 46 different suspected VTEC isolates were included in the study. In addition, two of the isolates, i.e., C770-12 and C679-12A, were subjected to WGS twice for verification of the WGS method.

As part of the study, the *E. coli* database for VirulenceFinder was constructed for detection of important *E. coli* virulence genes and vtx subtypes from WGS data to enable comparison to the current routine typing results. Additionally, a newly developed method, NDtree, for determining nucleotide differences among related isolates was employed for establishing phylogenetic relationships.

**WGS-based typing of the suspected VTEC isolates.** All 48 whole-genome sequences, representing the 46 isolates and 2 replicates, were subjected to bioinformatics analysis using the above-mentioned web-tools. The initial WGS-based species identification led to the discovery that one isolate, C849-12, was in fact not *E. coli*, but *Morganella morganii*, and this isolate was thus not included in further analysis. The remaining isolates were all confirmed as *E. coli*. In addition, one isolate, C767-12, was, based on comparison of WGS typing results and routine typing results, a clear mix-up, and was thus excluded from the comparison of typing results. MLST types were successfully assigned for the remaining isolates, apart from two isolates, C887-12 and C893-12, for which the ST types could not be assigned due to problems with the *de novo* assembly. Table 2 shows a comparison between the most important typing results for each isolate in the study by routine typing and WGS.

For all WGS analysis, apart from the SNPtree, the replicate sequences showed identical results. The complete list of typing results can be seen in Table S1 in the supplemental material, as well as the information on sequence quality and epidemiological information. The average coverage of the *E. coli* genome sequences ranged from 18× to 110×, and VirulenceFinder succeeded in detecting virulence genes in all the *E. coli* genomes. The complete list of detected virulence genes for each isolate can also be seen in Table S1.

Overall, there was high concordance between the routine typing results and the results obtained with VirulenceFinder, as can be seen in Table 3.

The *eae* gene was detected in the same 30 isolates by routine typing and by VirulenceFinder. The routine typing detected the presence of *vtx* genes in 40 of the isolates, of which 11 isolates had *vtx1*, 16 isolates had *vtx2*, and the remaining 13 isolates possessed both *vtx1* and *vtx2*. There was exact concordance between the *vtx1* genes found using routine typing and VirulenceFinder. The *vtx2* gene was detected in 29 isolates by routine typing, and VirulenceFinder found the *vtx2* gene in 27 isolates. For one of the two isolates, C641-12B, in which VirulenceFinder did not detect *vtx2*, subsequent retyping at the SSI confirmed the lack of *vtx2*, while the presence of *vtx2* was confirmed for the other isolate, C541-12. The inability of VirulenceFinder to detect the gene in this isolate was a feature of poor *de novo* assembly in this specific region in the gene of this isolate.

For 35 of the isolates that were hemolytic, routine typing detected *ehxA*, encoding enterohemolysin A, in 32 isolates, while VirulenceFinder detected *ehxA* in only 31 of these isolates. However, *ehxA* was subsequently detected in the sequence data for the last isolate, C813-12. The *ehxA* gene did not show up in VirulenceFinder due to problems in the *de novo* assembly, leading to only around half of the gene being present in the same contig of the assembled sequence data.

It was not possible to detect the plasmid-encoded virulence factor genes *saa* and *bfpA* with VirulenceFinder in these isolates, C749-12, C862-12, and C820-12, shown by routine typing to harbor one of these genes, although the genes were included in the database. The *saa* gene was, however, detected in C749-12 by subsequent reference mapping of raw reads, and retyping at the SSI confirmed the absence of *bfpA* in the isolates. The *saa* gene did not show up in VirulenceFinder due to a low-coverage region in the middle of the gene that caused problems in the *de novo* assembly step in VirulenceFinder.

For two isolates, C893-12 and C892-12, *vtx* subtypes were initially assigned by routine typing as *vtx1a* and *vtx2d*, respectively. VirulenceFinder assigned the subtypes *vtx1c* and *vtx2g*. Retyping at the SSI confirmed the *vtx1c* and *vtx2g* subtypes detected by VirulenceFinder.

**Phylogeny of the suspected VTEC isolates.** NDtree and SNPtree were constructed from the sequence data from the 44 confirmed *E. coli* isolates, as well as the two replicate sequences (Fig. 1 and Fig. 2). For NDtree, on average, 78% of the 5,498,450 bases were called in each of the strains. The isolates clustered completely according to serotype, and there was clear concordance between serotype and MLST type. For SNPtree, a total of 118,834 SNPs were called. Most isolates clustered according to serotype, although this method did not manage to cluster all three O117:
TABLE 2 Comparison of the most important typing results for each isolate in the study by routine typing and WGS typing

<table>
<thead>
<tr>
<th>Serotype</th>
<th>Routine typing result(s)</th>
<th>VirulenceFinder result(s)</th>
<th>MLST</th>
<th>Isolate</th>
</tr>
</thead>
<tbody>
<tr>
<td>O5:H−</td>
<td>vtx1c</td>
<td>vtx1c</td>
<td>ST-447</td>
<td>C750-12</td>
</tr>
<tr>
<td>O26:H11</td>
<td>eae, vtx1a</td>
<td>eae, vtx1a</td>
<td>ST-21</td>
<td>C696-12</td>
</tr>
<tr>
<td>O27:H30</td>
<td>vtx2b</td>
<td>vtx2b</td>
<td>ST-753</td>
<td>C770-12 (replicate)</td>
</tr>
<tr>
<td>O36:H−</td>
<td>vtx1c</td>
<td>vtx1c</td>
<td>ST-101</td>
<td>C749-12</td>
</tr>
<tr>
<td>O55:H12</td>
<td>vtx1a</td>
<td>vtx1a</td>
<td>ST-675</td>
<td>C904-12A</td>
</tr>
<tr>
<td>O76:H−</td>
<td>vtx1c</td>
<td>vtx1c</td>
<td>ST-17</td>
<td>C757-12</td>
</tr>
<tr>
<td>O103:H2</td>
<td>eae, vtx1a</td>
<td>eae, vtx1a</td>
<td>ST-1967</td>
<td>C813-12</td>
</tr>
<tr>
<td>O103:H−</td>
<td>eae, vtx1a</td>
<td>eae, vtx1a</td>
<td>ST-17</td>
<td>C850-12</td>
</tr>
<tr>
<td>O rough:H2</td>
<td>eae, vtx1a</td>
<td>eae, vtx1a</td>
<td>ST-333</td>
<td>C885-12B</td>
</tr>
<tr>
<td>O115:H−</td>
<td>eae, vtx2a</td>
<td>eae, vtx2a</td>
<td>ST-333</td>
<td>C896-12A</td>
</tr>
<tr>
<td>O117:K1:H7</td>
<td>vtx1a</td>
<td>vtx1a</td>
<td>ST-504</td>
<td>C751-12</td>
</tr>
<tr>
<td></td>
<td>vtx1a</td>
<td>vtx1a</td>
<td>ST-504</td>
<td>C659-12</td>
</tr>
<tr>
<td></td>
<td>vtx1a</td>
<td>vtx1a</td>
<td>ST-504</td>
<td>C760-12</td>
</tr>
<tr>
<td>O121:H9</td>
<td>eae, vtx2a</td>
<td>eae, vtx2a</td>
<td>ST-655</td>
<td>C821-12</td>
</tr>
<tr>
<td>O128ab:H2</td>
<td>vtx2b</td>
<td>vtx2b</td>
<td>ST-25</td>
<td>C748-12</td>
</tr>
<tr>
<td>O128ac:H−</td>
<td>vtx1c, vtx2b</td>
<td>vtx1c, vtx2b</td>
<td>ST-811</td>
<td>C864-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx2f</td>
<td>eae, vtx2f</td>
<td>ST-20</td>
<td>C820-12</td>
</tr>
<tr>
<td>O145:H−</td>
<td>eae, vtx2a</td>
<td>eae, vtx2a</td>
<td>ST-32</td>
<td>C816-12</td>
</tr>
<tr>
<td>O146:H21</td>
<td>vtx1c, vtx2b</td>
<td>vtx1c, vtx2b</td>
<td>ST-442</td>
<td>C874-12</td>
</tr>
<tr>
<td></td>
<td>eae</td>
<td>eae</td>
<td>ST-442</td>
<td>C896-12B</td>
</tr>
<tr>
<td></td>
<td>eae</td>
<td>eae</td>
<td>ST-442</td>
<td>C885-12A</td>
</tr>
<tr>
<td>O157:H7  (O157:H−)</td>
<td>eae, vtx1a, vtx2c</td>
<td>eae, vtx1a, vtx2c</td>
<td>ST-11</td>
<td>C570-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2c</td>
<td>eae, vtx1a, vtx2c</td>
<td>ST-11</td>
<td>C697-12A (replicate)</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2c</td>
<td>eae, vtx1a, vtx2c</td>
<td>ST-11</td>
<td>C697-12B</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2c</td>
<td>eae, vtx1a, vtx2c</td>
<td>ST-11</td>
<td>C541-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C812-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C818-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C819-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C849-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C852-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-11</td>
<td>C863-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx2a</td>
<td>eae, vtx2a</td>
<td>ST-966</td>
<td>C894-12</td>
</tr>
<tr>
<td></td>
<td>eae, vtx2a</td>
<td>eae, vtx2a</td>
<td>ST-2966</td>
<td>C891-12</td>
</tr>
<tr>
<td>O165:H−</td>
<td>eae, vtx1a, vtx2a</td>
<td>eae, vtx1a, vtx2a</td>
<td>ST-119</td>
<td>C905-12</td>
</tr>
<tr>
<td>O180:H−</td>
<td>eae</td>
<td>eae</td>
<td>ST-301</td>
<td>C641-12A</td>
</tr>
<tr>
<td>O181:H16</td>
<td>eae</td>
<td>eae</td>
<td>ST-301</td>
<td>C641-12B</td>
</tr>
<tr>
<td>OX187:H28</td>
<td>vtx1a†</td>
<td>vtx1a†</td>
<td>ST-200</td>
<td>C892-12</td>
</tr>
<tr>
<td>O rough:−</td>
<td>vtx2f</td>
<td>vtx2f</td>
<td>ST-678</td>
<td>C895-12</td>
</tr>
</tbody>
</table>

† Retyped as vtx2g at the SSI.
‡ Retyped as vtx1c at the SSI.

K1:H7 isolates, C751-12, C760-12, and C659-12, together, and also the two replicate sequences of C770-12 were not grouped with SNPtree. This seemed to be a feature of IonTorrent data in combination with the SNPtree algorithm, causing trouble for some sequences that were very dissimilar to the employed O157 reference genome.

NDtree thus seemed to cluster the isolates more correctly for this data set, although both approaches managed to group the six outbreak isolates.

The nucleotide difference (ND) and SNP divergence between the 46 genome sequences is included in Dataset S2 and Dataset S3 in the supplemental material, respectively. In addition, for further comparison of the phylogenetic approaches, a phylogenetic tree based on the NDtree matrix, but with positions called in all of the sequences, and constructed using FastTree, is included in Dataset S4.

**Evaluating the NDtree phylogeny.** NDtree clustered the isolates in complete agreement with the epidemiological information obtained. The replicate sequences, C697-12A and C770-12, both had no nucleotide differences to their respective partner sequences, C641-12A and C641-12B, which were isolated from the same plate due to colony morphology differences, also exhibited no differences and were thus identical, as also concluded by routine typing. Also, C697-12A (replicate) and C697-12B were isolates from the same plate and had no nucleotide differences and were also found identical by routine typing. The C885-12 and
C896-12 isolates, which in both cases exhibited mixed-colony morphology upon reception at the SSI, were from the same patient received 3 days apart. They clustered together with 2 nucleotide differences pairing isolate C885-12A together with C896-12B, and C885-12B together with C896-12A, which was also evident from the routine typing data, where the C885-12A and C896-12B isolates were both serotype O115:H2/H11002, and the C885-12B and C896-12A isolates were both serotype O146:H21.

In addition, the two isolates C816-12 and C884-12, which were not known to be epidemiologically related, but both were serotyped as O145:H9252/H11002, clustered together with no nucleotide differences, and could potentially have originated from the same source.

During the 7 weeks, a VTEC outbreak occurred with high risk of HUS, with a total of 13 diagnosed cases (7). Six of these outbreak isolates were included in the study. Within the outbreak, the nucleotide differences ranged from 1 to 7, whereas the distance to the other isolates in the O157:H7 group ranged from 607 to 1,617 nucleotide differences.

### Comparison of time for typing

Typing of the isolates was initiated Monday each week during the study, both at the SSI for routine analysis and at DTU Food for WGS analysis, in parallel, starting from the exact same isolates. All work was performed to fit into an average working day in Denmark of maximum 7.4 h.

At the SSI, several different typing methods were employed, and although the majority were initiated Monday (hemolysis test, β-glucuronidase test, O typing, H typing, and Vero cell assay), others were initiated Tuesday (K1/K5 typing and PCR), Wednesday (hybridization), or Monday the following week (PFGE), for practical reasons as part of the standard work flow in the routine laboratory.

The time for obtaining results was dependent on the method used. Results on hemolytic activity and β-glucuronidase activity were obtained on Tuesday, K1/K5 results were ready on Wednesday, PCR results for vtx subtyping were ready on Thursday, and results from hybridizations, O typing, H typing, and Vero cell assay were ready Friday. The PFGE results were obtained on Wednesday of the second week, on day 10.

The WGS-based typing was different, with only one work flow and each step necessary for initiating the next. The first step was gDNA purification, which was initiated Monday, and took approximately 2 h (35-min hands-on time), followed by fragmentation that took around 1 h (1 h hands-on) and subsequently, library preparation (3.5 h [2.5-h hands-on]). On Tuesday, the library concentration was determined (1 h [15-min hands-on]) and the template was subsequently prepared (5 h [0.5 h hands-on]). On

---

**TABLE 3** Concordance of results from routine typing and VirulenceFinder

<table>
<thead>
<tr>
<th>Virulence gene</th>
<th>No. of isolates with the virulence gene found by:</th>
<th>Routine typing</th>
<th>VirulenceFinder</th>
</tr>
</thead>
<tbody>
<tr>
<td>eae</td>
<td>30</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>vtx1</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>vtx2</td>
<td>29 (28)</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>eha</td>
<td>32</td>
<td>31 (32)</td>
<td></td>
</tr>
<tr>
<td>saa</td>
<td>1</td>
<td>0 (1)</td>
<td></td>
</tr>
<tr>
<td>bfpA</td>
<td>2 (0)</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*a* When there was a discrepancy between the results for routine typing and VirulenceFinder, the value after retyping at the SSI or after reference mapping of raw reads is shown in parentheses.

---

**FIG 1** Phylogeny of the isolates by the NDtree method. Isolates known to be epidemiologically related are shown in the same color, with the red group constituting the outbreak isolates. Serotypes and MLST types are shown for the main clusters. An asterisk indicates types with slight variations within the cluster.
Wednesday, the template quality was determined (0.5 h [0.5 h hands-on]), and then sequencing was initiated, finishing after 5 h (45 min hands-on). On Thursday, the sequence data were extracted and submitted to the CGE web-tools, and the final results for all analyses were ready after 3 to 3.5 h (0.5 h hands-on).

In Fig. 3, a comparison between the time for obtaining typing results by routine typing and by the WGS approach is illustrated. The routine typing relies heavily on serotyping and vtx subtyping, and thus, sufficient results for accurate typing were not obtained before Friday. For further discrimination of suspected outbreak isolates, PFGE results were not ready before Wednesday of the following week. For the WGS approach, however, all results on MLST, vtx subtypes, and phylogenetic relationships among isolates were ready for interpretation on Thursday, saving at least half a day compared to the routine typing.

The time for WGS typing depended on the number of isolates to be tested. In the setup, four isolates could be run simultaneously. With a number of isolates between four and eight, an additional template preparation and sequencing were necessary. This added a total of 1 h and 45 min hands-on time to the procedure, but results were still obtained Thursday.

The total hands-on time each week for the typing procedures carried out at the SSI was approximately 14.5 h, while for the WGS, it was around 6.5 h or 8 h depending on the number of isolates.

DISCUSSION

The objective of this study was to evaluate in real time WGS for typing and surveillance of VTEC infections in Denmark by comparing the typing results, time, and cost to those of the routine typing procedure currently carried out. A set of 46 suspected VTEC isolates was employed for evaluation, with WGS-based typing being conducted in parallel to the routine typing on all suspected VTEC isolates received at the SSI during the study period.

Several recent studies have already proved the usefulness of rapid benchtop sequencing for investigations of various outbreaks (4, 31, 32). These studies have analyzed outbreaks retrospectively, and for useful implications in clinical microbiology, it is necessary to be able to conduct the relevant WGS analysis in real time, to be able to take action both regarding patient care and outbreak control. In this study, it was apparent that typing data could be extracted from WGS and analyzed in real time, and lead to faster conclusions, and more information, than routine typing.

For real-time WGS typing to work for routine surveillance, it is essential that clinical health personnel without bioinformatics skills be able to quickly extract and interpret the relevant information from the massive amount of sequence data. Different comprehensible web-tools for WGS analysis are gradually becoming available enabling extraction of relevant information from WGS data for typing purposes (23, 24, 26, 33). However, the well-established traditional typing procedures varies between pathogens, and thus, a lot of effort still needs to be put into development of more useful and user-friendly tools.

We conducted WGS typing and analysis in real time with the ongoing routine typing and surveillance of VTEC and developed VirulenceFinder for extraction of virulence genes for typing. In our study, the WGS-based typing using the CGE web-tools was able to compete with the current typing, both on typing results, time, and price. The use of VirulenceFinder enabled quick and accurate detection of eae, ehxA, and vtx genes and was in addition
more robust assigning correct vtx subtypes than routine typing was. VirulenceFinder was, however, not successful in detecting saa and vtx2 in isolates C749-12 and C541-12, respectively, harboring these genes according to routine typing. This was probably a feature of poor sequence quality and low average coverage. The presence of many other important virulence genes in the isolates was detected by VirulenceFinder, giving much more information on the virulence profiles of the isolates than obtained by routine typing.

WGS-based typing also proved valuable for species detection, confirming that one of the suspected VTEC isolates was not *E. coli*, but instead *Morganella morganii*, adding to the value of performing WGS and strongly emphasizing that species confirmation should be performed as the first step of WGS typing.

The WGS-based phylogeny was very efficient for discrimination between the suspected VTEC isolates. As routine typing and surveillance did not offer any direct measure of relationships between all the isolates, but were instead based on grouping according to serotype and virulence gene profiles, and only in some cases PFGE, WGS-based typing could in this discipline offer much more information on how the isolates were related and potentially lead to faster detection of outbreaks.

The NDtree method managed to cluster the suspected VTEC isolates in complete agreement with the epidemiological information, whereas the SNPtree tool failed to cluster some of the isolates known to be identical. This was probably because the IonTorrent sequence data did not work well with SNPtree, since SNPtree is optimized for Illumina sequence data. The problem with SNPtree could be that this algorithm assigns the nucleotide for the reference genome at positions where the SNP is not accepted. This could, in theory, lead to two identical sequences not being clustered if the sequence data for one of these sequences had more sequencing errors, or perhaps low average coverage compared to the other, and was thus “mistakenly” assigned bases from the reference genome. Since SNPtree for this reason required all the query sequences to be closely related to the reference genome employed, it was not suitable for routine typing where many diverse isolates were included.

However, SNPtree did manage to cluster the main part of the isolates according to serotype, and also to identify the outbreak cluster, but since it did not seem completely reliable with the IonTorrent sequence data, NDtree, which was based on a different way to assign nucleotide differences, was employed instead, as it performed better for the data set in this study. NDtree clustered all isolates in agreement with the epidemiological data and showed no SNP differences between replicate sequences, and all isolates were clustered according to serotype. The method also perfectly clustered the outbreak isolates together and made it easy to distinguish outbreak isolates from sporadic isolates even of the same serotype. Although serotyping is an essential part of routine typing of *E. coli* as well as other important food-borne pathogens, for WGS-based typing, it was not necessary, since the MLST typing alone did just as good a job with clustering the isolates. This could be coincidental and due to the limited number of strains or restriction to analyze only VTEC, because *in silico* MLST performed on 61 whole-genome sequences of various *E. coli* strains rendered many of the various strains jumbled and less well resolved (34).

For the WGS analysis, many of the conventional typing results may be accessory. WGS offers ultimate resolution, and it would be valuable to develop tools to extract additional information, such
as the serotype, thus enabling comparison to historical data generated by conventional serotyping. Although WGS offers ultimate resolution, it is important to note that epidemiological information is still necessary for any analysis in typing, outbreak investigation, and surveillance.

For this study, WGS typing could compete with routine typing, by employing automated extraction of WGS data for typing of VTEC. With this approach, all relevant information could be extracted from the sequence, making WGS typing faster than routine typing, and especially with regard to the workload, WGS was advantageous, since it required much less hands-on time.

The SSI, many different E. coli isolates passed through the routine flow, being typed to different discriminatory levels by diverse methods, and it was thus impossible to determine the exact price per isolate included in the study. However, it was estimated that the overall price per isolate for each of the two approaches, routine typing and WGS typing, was around 430 euros for typing of suspected VTEC in this study, including both salaries and materials.

For routine typing, this included serotyping, PCR subtyping, DNA hybridizations, hemolysis and β-glucuronidase tests, Vero cell assay, and PFGE.

For both approaches, time could be optimized if necessary. In routine typing, PFGE could be initiated for isolates on the first day, and thus yield results Wednesday or Thursday of the first week. Similarly, in the WGS workflow, sequencing could be initiated on Tuesday, thus enabling analysis of sequencing results on Wednesday. However, for both routine and WGS typing, the procedure for typing was done according to what was believed to be most advantageous considering time, economical aspects, accessibility to equipment, etc.

Since the WGS part of this study was performed using the IonTorrent PGM, the price of routine typing and surveillance could be further decreased by performing sequencing on the Illumina MiSeq benchtop sequencer (35), as our current in-house price per isolate could be further decreased by performing sequencing on the IonTorrent PGM, the price of routine typing and surveillance using user-friendly web-tools are already applicable for both approaches, time could be optimized if necessary. In routine typing, PFGE could be initiated for isolates on the first day, and thus yield results Wednesday or Thursday of the first week. Similarly, in the WGS workflow, sequencing could be initiated on Tuesday, thus enabling analysis of sequencing results on Wednesday. However, for both routine and WGS typing, the procedure for typing was done according to what was believed to be most advantageous considering time, economical aspects, accessibility to equipment, etc.

In conclusion, this study shows that WGS-based typing and surveillance using user-friendly web-tools are already applicable for routine purposes and that this approach can make the process even faster and cheaper. Finally, WGS delivers typing results that equal or even surpass the current typing methodologies in terms of microbiological information.

ACKNOWLEDGMENTS

This study was supported by the Center for Genomic Epidemiology (www.genomicepidemiology.org) grant 09–067103/DSF from the Danish Council for Strategic Research. We thank Susanne Jespersen, Pia Möller Hansen, Christian Vråby Pedersen, and Christina Aaby Sørensen for excellent technical assistance. We declare that we have no conflicts of interest.

REFERENCES


Downloaded from http://jcm.asm.org on April 27, 2014 by guest