The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/136388

Please be advised that this information was generated on 2018-11-08 and may be subject to change.
Next-Generation Sequencing for Typing and Detection of Resistance Genes: Performance of a New Commercial Method during an Outbreak of Extended-Spectrum-Beta-Lactamase-Producing *Escherichia coli*

J. Veenemans,† I. T. Overdevest,† E. Snelders,‡ I. Willemsen,§ Y. Hendriks,§ A. Adesokan,∥ G. Doran,∥ S. Bruso,∥ A. Rolfe,∥ A. Pettersson,∥ J. A. J. W. Kluytmans∥

Laboratory for Microbiology and Infection Control, Amphia Hospital, Breda, The Netherlands; Department for Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands; Pathogenica, Cambridge, Massachusetts, USA

Next-generation sequencing (NGS) has the potential to provide typing results and detect resistance genes in a single assay, thus guiding timely treatment decisions and allowing rapid tracking of transmission of resistant clones. We evaluated the performance of a new NGS assay (Hospital Acquired Infection BioDetection System; Pathogenica) during an outbreak of sequence type 131 (ST131) *Escherichia coli* infections in a nursing home in The Netherlands. The assay was performed on 56 extended-spectrum-beta-lactamase (ESBL) *E. coli* isolates collected during 2 prevalence surveys (March and May 2013). Typing results were compared to those of amplified fragment length polymorphism (AFLP), whereby we visually assessed the agreement of the BioDetection phylogenetic tree with clusters defined by AFLP. A microarray was considered the gold standard for detection of resistance genes. AFLP identified a large cluster of 31 indistinguishable isolates on adjacent departments, indicating clonal spread. The BioDetection phylogenetic tree showed that all isolates of this outbreak cluster were strongly related, while the further arrangement of the tree also largely agreed with other clusters defined by AFLP. The BioDetection assay detected ESBL genes in all but 1 isolate (sensitivity, 98%) but was unable to discriminate between ESBL and non-ESBL TEM and SHV beta-lactamases or to specify CTX-M genes by group. The performance of the hospital-acquired infection (HAI) BioDetection System for typing of *E. coli* isolates compared well with the results of AFLP. Its performance with larger collections from different locations, and for typing of other species, was not evaluated and needs further study.

Genes encoding extended-spectrum beta-lactamas (ESBL) are rapidly spreading among Gram-negative bacteria. Due to growing reservoirs of resistant *Enterobacteriaceae* in humans and farm animals and the widespread presence of these bacteria in the food chain, the frequency of outbreaks in hospitals and long-term-care facilities is expected to increase (1–6). Nosocomial spread of resistant bacteria poses a serious threat to patient safety all around the world.

When an outbreak is detected, rapid assessment of its extent and the identification of transmission routes and sources are essential to prevent further transmission among susceptible patients. Several molecular typing methods are available to investigate the degree of relatedness of strains recovered from patients during outbreaks. Due to differences in performance (e.g., discriminatory power and reproducibility), none of these typing methods can be universally applied for all pathogens in all situations. Ideally, typing results should be reproducible so that all data generated by different laboratories can be compared. Costs, time to result, and the required level of technical skills are further determinants of whether a typing method is suitable to be widely implemented in microbiology laboratories (7, 8).

Amplified fragment length polymorphism (AFLP) is often used for typing of *Enterobacteriaceae*. This technique relies on amplification of genomic fragments generated by one or two restriction enzymes and uses the amplified products to create digitized profiles. It has good discriminatory power but is relatively expensive and time-consuming and is technically demanding. Whole-genome sequencing (WGS) can provide information about bacterial relatedness with across-genome single-nucleotide resolution, but rapid translation of these data into epidemiologically useful information remains difficult. Data derived from whole-genome sequencing have nonetheless facilitated the search for suitable targets that allow bacterial typing based on sequencing only a part of the genome. In addition to providing information on phylogenetic relatedness, sequence data can be used to detect resistance genes, enabling clinicians to rapidly track transmission of resistant clones or mobile genetic elements and to initiate timely appropriate treatment during outbreak situations. Thus, it is likely that sequence-based typing methods will eventually replace fragment-based methods such as AFLP.

Next-generation sequencing (NGS) is a new high-throughput sequencing method that can rapidly generate millions of reads (ranging between 35 and 700 bp in length) at relatively low per-base sequencing costs. The Hospital Acquired Infection BioDetection System (Pathogenica) is a new commercial NGS-based assay that targets a selection of short DNA regions and is able to detect
the presence of several different pathogens and resistance genes simultaneously in a single sample while sequencing only a few kilobases of genome. In addition, the kit includes software that can be used to type the strains and provide a phylogenetic tree for sets of isolates.

In June to July 2012, a survey in 9 nursing homes in the southern region of The Netherlands identified several facilities with a high prevalence of fecal ESBL carriage. The results of multilocus sequence typing (MLST) and AFLP analysis of isolates collected in the two facilities with the highest prevalences (19% and 21% rectal carriage, respectively) indicated clonal spread in both facilities. Widespread dissemination of E. coli ST131 (a multiresistant clone that spreads extensively in elderly populations in regions of hyperendemicity all around the world) was found in one of the facilities. In 2013, two follow-up surveys were performed to assess the effect of the implemented interventions. In this paper, we compare the typing performance of the new NGS assay to that of AFLP for the isolates collected during these follow-up surveys.

MATERIALS AND METHODS

Isolate collection and epidemiological setting. Two consecutive cross-sectional surveys were performed in March and May 2013. During these surveys, feces or rectal swabs (Eswab medium; Copan Diagnostics) were collected from all admitted patients, and hand cultures were taken from health care workers (HCW). Every first ESBL-positive E. coli isolate per patient or health care worker (collected either in March or in May) was included in the isolate collection for the current study. The isolate collection thus included 56 E. coli isolates from 48 patients and 1 HCW. Forty-nine isolates (from 42 patients and 1 HCW) were collected in the first nursing home (NH1), and the remaining isolates were collected from the second location (NH2) located at a 45-km distance from NH1. In NH1, patients were residing in 3 semisepare buildings (A, B, and C); each of these buildings was divided into 2 or 3 separate departments (A1 to A3, B1 to B3, and C1 to C3). The study included 15, 20, and 14 isolates from buildings A, B, and C, respectively. Staff members were not exchanged to B3, and C1 to C3. The study included 15, 20, and 14 isolates from buildings A, B, and C, respectively. Staff members were not exchanged between these buildings, and transfer of patients between departments seldom occurred. From NH2, 7 isolates (from 6 patients residing in departments D1 and D2) were included in the collection.

Identification and detection of resistance genes. Identification of Gram-negative bacteria growing on a selective medium (EBSA; AlphaOpartments D1 and D2) were included in the collection.

Identification of hospital-acquired infection (HAI) BioDetection System. The pathogens that can be detected by the HAI BioDetection System (Pathogenica, Boston, MA) are Acinetobacter baumannii, Acinetobacter pittii, Acinetobacter calcoaceticus, Clostridium difficile, Escherichia coli, Enterococcus fecalis, Enterococcus faecium, Enterobacter aerogenes, Enterobacter cloacae, Klebsiella oxytoca, Klebsiella pneumoniae, Proteus mirabilis, Pseudomonas aeruginosa, Staphylococcus aureus, Staphylococcus epidermidis, Staphylococcus saprophyticus, and Staphylococcus haemolyticus. The drug resistance genes that can be detected by the HAI BioDetection System are NDM, OXA, KPC, IMP, CARB, CMY, CTX-M, TEM, SHV, PER, VEB, and VIM (beta-lactamases). meca (mecillin resistance), vanA (glycopeptide resistance), ermA (macrolide/lincosamide resistance), and mecA (multidrug resistance efflux operon). In addition to the genes mentioned here, which are all listed in the product brochure, the assay can also detect several aminoglycoside resistance genes (aac, aad, and ant) and sulfonamide resistance genes (sulI) (Table 1). The probe set may not detect all forms of these genes, however, and the detection of these genes has not yet been validated.

The assay is based on approximately 300 probes that capture and amplify short DNA regions of interest for each gene that are selected by the manufacturer based on previously performed whole-genome analyses. The kit simultaneously sequences all probe targets from any organism in a sample, minimizing the hands-on time needed to identify and type bacterial species, while detecting resistance genes at the same time. By targeting variable regions, the selected probe set can identify species and distinguish between strains while sequencing only a few kilobases of a genome.

Isolates were thawed, DNA was extracted, and each sample was diluted to a working DNA concentration of 0.5 ng/μl and processed further, according to the instructions of the manufacturer. Individually barcoded isolates were combined into a clean microcentrifuge tube, and this pooled PCR product was purified using 0.8× Agencourt AMPure XP Reagent (Beckman Coulter). An Ion Torrent PG M sequencing platform (Life Technologies) was used to sequence the DNA fragments. Template preparation was carried out with an Ion PG M Template OT2 200 kit (Life Technologies), according to the instructions of the manufacturer. Sequencing data were analyzed using an internal version of Pathogenica’s HAI BioDetection software v1.2. This software uses 10^6 to 10^8 sequencing reads to determine the genomic sequence of all the loci targeted by the assay for each organism present in the sample. RAxML was used to generate a phylogenetic tree for the samples (12). A shell-scripting interface to the software (rather than the Torrent Server interface) was used to create the tree, whereby the probe weight and partition values were shown to be

Next-Generation Sequencing for E. coli Typing

jcm.asm.org 2455

July 2014 Volume 52 Number 7

Downloaded from http://jcm.asm.org on May 1, 2017 by RADBOUD UNIVERSITEIT NIJMEGEN
<table>
<thead>
<tr>
<th>Detection method</th>
<th>CTX-M-1 (n = 8)</th>
<th>CTX-M-3 (n = 1)</th>
<th>CTX-M-15 (n = 32)</th>
<th>CTX-M-9 (n = 6)</th>
<th>TEM (n = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-ESBL TEM positive</td>
<td>4</td>
<td>1</td>
<td>31</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Non-ESBL TEM negative</td>
<td>4</td>
<td>0</td>
<td>31</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>HAI BioDetection system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta-lactamase genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTX-M</td>
<td>7</td>
<td>1</td>
<td>32</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>TEM</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Other resistance genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aac(6')-Ib</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ant</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>ant(3')-Ia</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>sulI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| Resistance genes detected by microarray (Check-MDR CT103, Checkpoints) and by the HAI BioDetection System |
|------------------------------------------------------|------------------------------------------------------|
| Detection method | CTX-M-1 (n = 8) | CTX-M-3 (n = 1) | CTX-M-15 (n = 32) | CTX-M-9 (n = 6) | TEM (n = 7) |
| Non-ESBL TEM positive | 4 | 1 | 31 | 7 |             |
| Non-ESBL TEM negative | 4 | 0 | 31 | 3 | 7 |
| HAI BioDetection system | | | | | |
| Beta-lactamase genes | | | | | |
| CTX-M | 7 | 1 | 32 | 6 | 0 |
| TEM | 4 | 1 | 1 | 2 | 7 |
| Other resistance genes | | | | | |
| aac(6')-Ib | 0 | 0 | 0 | 0 | 0 |
| ant | 0 | 1 | 5 | 5 | 0 |
| ant(3')-Ia | 1 | 1 | 0 | 3 | 1 |
| sulI | 0 | 0 | 0 | 0 | 0 |

**Results**

- **Comparison of the typing results.** First, we visually assessed the agreement of the phylogenetic tree generated by the HAI BioDetection kit software with the results generated by AFLP. We also assessed to what extent the arrangement of this tree was concordant with the available epidemiological data.

- **RESULTS**

**Typing.** Among the 56 *E. coli* isolates, five clusters of isolates were identified by AFLP, and 12 isolates were considered unique. The largest cluster consisted of 31 indistinguishable isolates (5 representatives in green), whereas the other 4 clusters included only 3 or 4 indistinguishable isolates (yellow, blue, pink, and red; Fig. 1). Isolates from the green cluster were exclusively found in departments in buildings B and C in nursing home 1. Isolates unrelated to this cluster were not found in these departments. The smaller clusters were found in department A1 (red cluster), in department A2 (pink and blue clusters) in nursing home 1, and in department D1 (yellow cluster) in nursing home 2. Results related to the AFLP typing, ST131 PCR, and department of residence at the time of sample collection are indicated in Fig. 1. All isolates from the green cluster belonged to phylogroup B2 and were found to be ST131 by PCR. For one isolate representative of this cluster, sequence type results were confirmed by MLST.

The HAI BioDetection kit correctly identified 54 isolates as *E. coli*. Due to technical problems with the sample, sequence results

**TABLE 1 Resistance genes detected by microarray (Check-MDR CT103, Checkpoints) and by the HAI BioDetection System**
could not be generated for 2 isolates. Figure 2 shows the phylo-
genetic tree produced by the HAI BioDetection v1.2 software. For each isolate, the AFLP cluster is indicated by its designated color in the figure, followed by the department designation. As shown in the figure, the BioDetection phylogenetic tree shows strong relatedness for all isolates belonging to the green AFLP cluster, whereby 5 strains in the cluster, 4 from department B2 and 1 from department C1, appeared in a slightly separate location in the tree. The arrangement of other parts of the phylogenetic tree also largely agreed with the clusters defined by AFLP and with the epidemiological data.

For several isolates, the results of the two typing methods yielded discordant interpretations. Specifically, there were 2 patients (P1 and P3; see Fig. 2) who carried an ESBL E. coli strain during both consecutive surveys and for whom the isolates collected in March belonged to the green AFLP cluster, whereby 5 strains in the cluster, 4 from department B2 and 1 from department C1, appeared in a slightly separate location in the tree. The arrangement of other parts of the phylogenetic tree also largely agreed with the clusters defined by AFLP and with the epidemiological data.

Whole-genome sequencing (WGS) was performed on a subset of 11 isolates, indicated by arrows in Fig. 2. This subset included the 4 isolates collected from the 2 patients (P1 and P3) described in the previous paragraph and 2 isolates from another patient (P2; isolates 27 and 52, both from the yellow cluster, collected in March and May [see Fig. 2]). WGS was also performed on 5 isolates with various positions in the HAI tree to determine the genetic diversity in the isolate collection: 1 isolate from the blue AFLP cluster (isolate 51) and a set of 4 isolates from the green AFLP cluster, of which 2 isolates (12 and 25) appeared in a slightly separate “group” in the HAI tree (see Fig. 2).

WGS results are shown in Fig. 3. As predicted by the placement in the HAI BioDetection tree and AFLP results and shown in panel A of the figure, WGS data from isolates 44 and 46 (P3) and isolate 51 (blue cluster) show substantial divergence from those from the other 8 isolates. To better examine the relationships among those 8 isolates, we generated a new WGS tree with higher resolution using this subset; these results are shown in panel B. In this new tree, the ordering of the 4 isolates from the green cluster (53 plus 55 versus 12 plus 25) was consistent with the HAI phylogenetic tree, and the 2 pairs of samples of patient P1 (41 plus 11) and P2 (52 plus 27) were presented as expected.
Lastly, we examined the association between the number of SNPs observed between sample pairs (55 pairs for 11 samples with WGS data) by the HAI BioDetection data and the (log-transformed) number of SNPs between the same sample pairs in the WGS data. This confirmed the correlation between the number of SNP differences detected by the BioDetection System and the true genomic distance (Pearson’s correlation coefficient, 0.69; \( P < 0.01 \)).

Detection of resistance genes. An overview of the resistance genes detected by the 2 assays is provided in Table 1. CTX-M and TEM ESBL beta-lactamases were detected by microarray in 47/54 and 7/54 isolates, respectively. SHV genes were not detected. The HAI BioDetection System did not detect the CTX-M beta-lactamase gene for 1 isolate (CTX-M-15; see Table 1 [sensitivity, 98%]) and was not able to discriminate between the genes from the different CTX-M groups. TEM beta-lactamase genes were detected by microarray in 16 isolates (TEM wild type [TEM-WT], 9 isolates; TEM-ESBL, 7 isolates), while the HAI BioDetection kit found TEM beta-lactamase gene in 15 isolates (8 TEM-WT and 7 TEM-ESBL), whereby no discrimination was provided between ESBL and non-ESBL TEM beta-lactamase genes. An overview of other resistance genes that were detected by the kit is provided in Table 1.

\[ \text{aac(6')-Ib} \] (encoding aminoglycoside N-acetyltransferase, which modifies aminoglycosides by acetylation) was found in 32 isolates and was found more frequently among isolates carrying CTX-M-15 (28/32) than among isolates with other CTX-M genes (4/22) (\( P < 0.01 \); Fisher’s exact text). Of the isolates carrying the \[ \text{aac(6')-Ib} \] gene, 21 (66%) were resistant to tobramycin, while
the 11 remaining isolates (34%) showed intermediate susceptibility.

DISCUSSION

In this collection of 56 *E. coli* isolates, the BioDetection phylogenetic tree correctly reflected the strong relatedness of all isolates belonging to the large outbreak cluster identified by AFLP, as well as the relatedness of isolates belonging to the smaller AFLP clusters. It should be noted that the discordant results for several isolates may be partly due to the fact that the assays were not performed simultaneously on the same culture. For the isolates that were analyzed by whole-genome sequencing, the results indicate that, although the BioDetection system sequences only a small part of the genome, the number of SNPs reported is predictive of the true genomic divergence as determined by WGS.

A limitation of the version of the HAI BioDetection software that was used is that it did not translate the typing results into a set of distinct clusters. To facilitate the interpretation of the typing results, further work on the definition of species-specific cutoffs is currently under way to improve the software so that clusters can be defined in a standardized way. Finally, although the test collection included a considerable number of *E. coli* isolates and consisted of several separate AFLP clusters with distinct epidemiological backgrounds, an evaluation of the typing performance with a larger collection (that reflects the diversity present among the members of the species as a whole) is justified.

The BioDetection assay accurately detected resistance genes in all but 1 CTX-M-positive isolate (sensitivity 98%), but it was unable to discriminate between ESBL and non-ESBL TEM and SHV beta-lactamases or to specify CTX-M genes by group. When the assay is used to screen direct patient material for the presence of resistance genes, an additional test is thus required in case of a positive test result. This is particularly true for TEM and SHV genes but may also be useful when CTX-M genes are detected, as discriminating between different groups of resistance genes is important for epidemiological purposes. We have no data about the prevalence of non-ESBL TEM genes in direct fecal material of the patients in this collection, as a selective culture step preceded the DNA isolation. With regard to the failure of the kit to detect an ESBL CTX-M gene for 1 isolate, we cannot exclude the possibility that plasmids carrying the resistance gene were lost during storage, a phenomenon that has been described previously (18). As reported by others (19), the *aac(6’)-Ib* gene coding for the aminoglycoside-modifying enzyme was found in almost all ST131 isolates.

In conclusion, the performance of the HAI BioDetection System for typing of *E. coli* isolates compared well with the results of AFLP. Its performance in larger collections from different locations and for typing of other species was not evaluated and needs further study. A recent study showed accurate discrimination of clonal lineages in a collection of 32 carbapenem-resistant *Klebsiella pneumoniae* isolates (20). Its ease of use, the ability to simultaneously detect resistance genes, and the potential to create an electronic database for surveillance are important advantages. Although the assay has good sensitivity to detect resistance genes and can be used to screen for their presence, TEM and SHV genes would need to be characterized further by an additional test.

ACKNOWLEDGMENTS

We thank the laboratory technicians of the VU University Medical Centre for their work on the AFLP.

A. Adesokan, G. Doran, S. Bruso, and A. Rolfe are members of Pathogenica and were responsible for and/or participated in the development of the HAI BioDetection System.

REFERENCES


