Application of Next-Generation Sequencing for Characterization of Surveillance and Clinical Trial Isolates Analysis of the Distribution of β-lactamase Resistance Genes and Lineage Background in the United States

Rodrigo E Mendes, Ronald N Jones, Leah N Woosley, Vincent Cattoir, Mariana Castanheira

IDSA
Infectious Diseases Society of America

hivma
hiv medicine association

OXFORD

Application of Next-Generation Sequencing for Characterization of Surveillance and Clinical Trial Isolates: Analysis of the Distribution of β-lactamase Resistance Genes and Lineage Background in the United States

Rodrigo E. Mendes,[1] Ronald N. Jones,[1] Leah N. Woosley,[1] Vincent Cattoir,[2,3,4] and Mariana Castanheira[1]

[1]JMI Laboratories, North Liberty, Iowa; [2]University Hospital of Rennes, Department of Clinical Microbiology, Rennes, France; [3]National Reference Center for Antimicrobial Resistance, Rennes, France; [4]University of Rennes 1, Unit Inserm U1230, Rennes, France

Background. Sequencing technologies and techniques have seen remarkable transformation and innovation that have significantly affected sequencing capability. Data analyses have replaced sequencing as the main challenge. This paper provides an overview on applying next-generation sequencing (NGS) and analysis and discusses the benefits and challenges. In addition, this document shows results from using NGS and bioinformatics tools to screen for β-lactamase genes and assess the epidemiological structure of *Escherichia coli*– and *Klebsiella pneumoniae*–causing bloodstream (BSIs) and urinary tract (UTIs) infections in patients hospitalized in the United States during the SENTRY Antimicrobial Surveillance Program for 2016.

Methods. A total of 3525 isolates (2751 *E. coli* and 774 *K. pneumoniae*) causing BSIs (n = 892) and UTIs (n = 2633) in hospitalized patients in the United States were included. Isolates were tested for susceptibility by broth microdilution, and those that met a minimum inhibitory concentration (MIC)–based screening criteria had their genomes sequenced and analyzed.

Results. A total of 11.6% and 16.1% of *E. coli*–causing UTIs and BSIs, respectively, met the MIC-based criteria, whereas 11.0% and 13.7% of *K. pneumoniae* isolates causing UTIs and BSIs, respectively, met the criteria. Among *E. coli*, bla_{CTX-M} variants (87.6% overall) prevailed (60.5% of CTX-M group 1 and 26.9% of group 9). A total of 60.3% of *K. pneumoniae* isolates carried bla_{CTX-M} variants (52.7% and 7.6% of groups 1 and 9, respectively). Two *E. coli* (0.6%) and 13 *K. pneumoniae* (12.9%) isolates harbored bla_{KPC}. Among KPC-producing *K. pneumoniae* (2 from BSIs and 11 from UTIs), 84.6% (11/13) were ST258 (CC258). Seventeen and 38 unique clonal complexes (CCs) were noted in *E. coli* that caused BSIs and UTIs, respectively, and CC131 (or ST131) was the most common CC among BSI (53.6%) and UTI (58.2%) isolates. Twenty-three and 26 CCs were noted among *K. pneumoniae*–causing BSIs and UTIs, respectively. CC258 (28.3%) prevailed in UTI pathogens, whereas CC307 (15.0%) was the most common CC among BSI isolates.

Conclusions. This study provides a benchmark for the distribution of β-lactamase genes and the population structure information for the most common *Enterobacteriaceae* species responsible for BSIs and UTIs in US medical centers during the 2016 SENTRY Program.

Keywords. genome; high-throughput; MLST; resistance.

Several surveillance studies conducted in the last few decades collected and generated massive amounts of data, including prevalence of pathogens causing hospital and community infections and susceptibility information for monitoring antimicrobial resistance. In general, the vast majority of these studies have generated phenotypic information associated with bacteria and fungi, from organism identification to antimicrobial susceptibility profiles. Although small local surveillance studies have long utilized various molecular methodologies for further characterization of organisms, large-scope investigations introduced the use of molecular techniques in a more gradual manner, mostly due to feasibility issues. Eventually, when applicable, large surveillance studies utilized molecular techniques to investigate specific phenotypes observed in subsets of isolates within a collection of surveillance strains [1–5].

In the past 2 decades, large surveillance studies used to apply singleplex and multiplex polymerase chain reaction (PCR) assays for characterizing clinically relevant pathogens of interest responsible for human infections. These reactions were used for several applications, including screening of resistance and virulence genes, typing, and serotyping of organisms [6–15]. With the availability of a "first-generation" DNA sequencing method (Sanger sequencing),

PCR amplicon sequencing was incorporated as an additional step after PCR assays to confirm and identify gene variants [5, 16]. The Sanger method comprises the process of selectively incorporating chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication; it was commercialized by Applied Biosystems (Foster City, CA). Although these techniques provide good and reliable results, they tend to be very time-consuming and labor-intensive. Currently, a number of automated methods can be utilized to further characterize organisms in a higher-throughput and friendly manner [17–22]. Despite the advantages of higher-throughput next-generation sequencing methods, Sanger sequencing is still used for specific purposes.

Since the first bacterial genome sequence was completed 2 decades ago [23], sequencing technologies and techniques have undergone remarkable transformation and innovation [23, 24]. Technological advances have significantly affected sequencing capability and, consequently, decreased the cost per megabase of sequence, becoming economically feasible for scientific research, clinical diagnostics, and surveillance programs [24]. Thus, sequencing has not been the recent limitation, but analyzing the gigabytes of data generated and providing accurate and meaningful interpretations have become limitations. In addition, sequence data and analyses need to provide translational information that can provide actionable items in real time, such as information to minimize the dissemination of clones or resistance genes, genomic surveillance, and the detection and control of hospital outbreaks [25, 26].

This paper provides an overview of the application of bacterial genome sequencing and analysis in large surveillance programs, among other settings, and discusses the benefits and challenges associated with the detection of resistance mechanisms, epidemiology typing, serotyping, and other means of molecular bacterial characterization. In addition, this document offers a study example of the application of the combination of next-generation sequencing (NGS) and bioinformatics tools to screen for β-lactamase genes and to assess the epidemiologic structure of *Escherichia coli* and *Klebsiella pneumoniae* that produce extended-spectrum β-lactamases (ESBLs), including carbapenemases, causing bloodstream (BSIs) and urinary tract (UTIs) infections in patients hospitalized in the United States during the SENTRY Antimicrobial Surveillance Program for 2016.

The Power of Genome Sequencing and Bioinformatics Tools
Phenotypic Prediction
Some studies have published promising results demonstrating that phenotypic profiles can be predicted based on the presence of certain genetic determinants. These studies have empirically demonstrated that predicting susceptibility and resistance using high-throughput genome sequencing and bioinformatics data analysis may be feasible or potentially feasible for *Staphylococcus aureus*, *Enterobacteriaceae*, *Pseudomonas*

aeruginosa, *Campylobacter* spp., *Helicobacter pylori*, *Neisseria gonorrhoeae*, and *Mycobacterium tuberculosis* [27].

The causal relationship between the presence of any given gene(s) and the respective phenotype has been well established for genes, such as β-lactamase genes, vancomycin resistance genes (eg, *vanA*), macrolide-lincosamide-streptogramin B (MLS_B), fluoroquinolones (acquired genes and mutations in the quinolone resistance–determining region [QRDR]), pleuromutilins, oxazolidinones, and tetracycline, to name a few. The pre-established knowledge of resistance genes and respective phenotypes allows for the possibility of predicting these phenotypes based on genotypic data. However, resistance phenotypes associated with glycopeptides, lipoglycopeptides, and cyclic lipopeptides, for instance, may involve multiple and/or complex resistance mechanisms that are yet not very well understood and may remain a challenge for genotypic detection and, consequently, phenotypic prediction. In addition, up- or downregulation of intrinsic genes can also present challenges for phenotype prediction.

A recent investigation utilized a machine-learning approach to predict the susceptibility phenotype in *M. tuberculosis* [28]. This approach could have the potential to better predict those resistance phenotypes associated with complex and multiple mutations, such as those linked with or close to the drug target site, alterations in porin channels (being just a minimal number of single nucleotide polymorphisms [SNPs] or a disruption due to an insertion sequence), alterations in promoter regions, or repressor genes. However, it is conceivable that phenotypic and genotypic information will need to be generated continuously, at least at a global level, to provide training data, so the "learning" of mechanisms to old and new drugs remains constant. This information would be utilized to maintain a database available to any data analysis center or clinical laboratory.

Genome Sequencing and Bioinformatics in Clinical Trials
Many studies were performed to molecularly characterize clinical trial isolates to determine the presence of resistance genes, virulence genes, and/or typing information [29–35]. The latter can be utilized to identify lineages, or the genetic background of strains, recovered from enrolled patients. When pathogens from the same species are recovered from clinical specimens collected during multiple study visits, typing methods are applied to suggest cases of persistent infection or re-infection. Proven re-infection cases have the potential to alter the efficacy data for the investigational and comparator agents [36]. Moreover, the information obtained from characterized resistance mechanisms among clinical trial isolates can be utilized to form genetic subsets of isolates, which can provide a more granular analysis of the in vivo efficacy and in vitro activity data. It is also important to mention that pathogens (baseline and follow-up isolates) that develop a resistance phenotype to the study drug during treatment can be investigated to determine the resistance mechanisms [36].

NGS and proper data analysis more recently have been applied to characterize a particular microbiome (eg, the gut microbiome) using metagenomics. Two approaches can be utilized: shotgun metagenomics or targeted-amplicon sequencing. Shotgun metagenomics attempts to sequence the entire genetic content present in a sample, and targeted-amplicon sequencing represents a more biased approach that targets a particular group of microorganisms [37]. Because antimicrobial therapy can affect normal human flora, some pharmaceutical companies have studied the impact of investigational drugs on the gut flora. DNA can be sequenced directly from human clinical samples by metagenomics sequencing. This approach can identify and quantify organisms present in human samples and then compare the effect of the investigational and comparator agents on the intestinal flora during pre- and post-treatment periods, as well as the effect of different therapeutic regimens [38, 39].

Another area of possible interest during drug development would be a metagenomics investigation of all possible pathogens present in a given human clinical sample (eg, urine, sputum). NGS-based methods can offer a relatively unbiased approach for pathogen detection and quantification. Clinical samples collected pre- and post-treatment from patients who failed therapy can be investigated by metagenomics to detect organisms. This approach may help us understand if the presence of other uncultured organisms or species not detected by current conventional microbiology laboratory methods, including those that can be refractory to the drug therapy being used, can explain poor clinical outcomes [40].

Challenges of Sequencing and Data Analysis
Powerful high-throughput genome sequencing has opened doors for massive DNA sequencing. Many current sequencing methodologies can be optimized and standardized in a protocol to be utilized for several species, facilitating implementation. In contrast, data analysis represents a challenge and, in general, requires bioinformatics knowledge to perform the various steps needed to answer any given question(s) [41, 42]. Many software programs are available online, and several options can be chosen to perform a single step of the process; each has its pros and cons [24]. Moreover, any in-depth analysis also requires scientific knowledge of genomic features and the biological background of the organism(s) under investigation to properly interpret results [42]. However, the ultimate goal would be to obtain a bioinformatics pipeline that would process the raw data and provide information in real time.

Regardless of software or the combination of software programs employed, the use of quality control (QC) criteria remains crucial for any laboratory process. Each sequencing methodology has its own QC metrics to ensure the quality of the sequencing data, and this does not enable the use of a single and uniform set of QC parameters [43]. Moreover, different downstream applications would require different levels of QC

parameters and metrics, such as SNP analysis from screening resistance/virulence genes, as the former requires greater coverage and sequencing for data analysis than the latter. In addition to the set of QC parameters related to the sequencing data, a set of QC strains with known genotypes associated with the downstream application should be used systematically to ensure that all steps from DNA extraction to data analysis are acceptable.

Using a database containing curated resistance/virulence genes is essential for proper detection, identification, and interpretation of data. Databases need to be constantly updated to minimize the possibility of not detecting genes recently discovered. A number of databases are freely available, such as ResFinder, CARD, ARDB, MEGARes, ARG-ANNOT, and SSTAR, or a combination of databases (eg, JMI bioinformatics pipeline) [44–47]. In general, genetic markers and their respective annotations have not been standardized, and similar genes can have different annotations from database to database. As alluded to earlier, the organism's biological background is crucial for properly interpreting results. These high-throughput NGS and screening tools are very sensitive in detecting gene determinants, and these systems and databases will not indicate the location (chromosomal, plasmid, or both), relevance, or association of gene(s) with any resistance phenotypes. Many resistance genes in the databases described here are chromosomally located and ubiquitous to several organisms and, when expressed in low levels, may not necessarily cause resistance phenotypes, or resistance may be limited to certain drugs within a class. Thus, the presence of an intrinsic gene may not be relevant for certain species–drug combinations, but potentially may be important in others. To complicate things further, an isolate can display a resistance phenotype when such intrinsic genes are overexpressed due to alterations in the promoter region, insertion of a promoter upstream, and/or gene mobilization into multicopy plasmids [48–57]. Therefore, certain genes detected in some species require careful interpretation.

Many organisms express resistance phenotypes due to overexpression of intrinsic systems (efflux pumps), decreased or null expression of other alterations (porin channels needed for antimicrobial cell entry), and/or drug target alterations [58–68]. The complete molecular characterization of resistance mechanisms in some species–drug class combinations requires quantifying certain cell components and/or sequencing and analyzing drug targets. In addition, drug targets need to be known, as well as whether any alterations in close or distal locations of the drug target affect binding [69–76]. Mutations causing frame-shift alterations or a premature stop codon within genes encoding for porin membranes (eg, OmpK35, OmpC) are simpler to interpret and less subjective. However, the interpretation of mutations within housekeeping genes or drug target sequences is challenging and may require complex experiments to link any given alterations with a resistance phenotype. Also, the sequencing analysis requires the use of proper reference

sequences from organisms susceptible to the drug and/or class of drugs being analyzed. Ideally, the reference strain should belong to the same lineage as the query isolate in order to minimize the detection and report of less relevant polymorphisms.

METHODS

Organism Collection

A total of 3525 isolates causing BSIs (n = 892) and UTIs (n = 2633) in hospitalized patients in the United States were collected as part of the SENTRY Antimicrobial Surveillance Program. In total, 2751 *E. coli* (601 and 2150 isolates responsible for BSIs and UTIs, respectively) and 774 *K. pneumoniae* (291 and 483 isolates responsible for BSIs and UTIs, respectively) were included. These isolates were consecutively collected (1 per patient) from 81 sites located in 36 states in 9 US census divisions and were submitted to JMI Laboratories (North Liberty, IA) as part of the 2016 SENTRY Program. Isolates were initially identified by the participating laboratory, and identifications were confirmed at JMI Laboratories using matrix-assisted laser desorption ionization-time of flight mass spectrometry (Bruker Daltonics, Bremen, Germany) and genome sequencing.

Susceptibility Testing

Isolates were tested for susceptibility by the broth microdilution method using cation-adjusted Mueller-Hinton broth. Susceptibility testing was performed at a central reference laboratory (JMI Laboratories) according to Clinical and Laboratory Standards Institute (CLSI) methods [77, 78]. Manufacturing of broth microdilution minimum inhibitory concentration (MIC) panels (JMI Laboratories) and quality control practices followed CLSI guidelines [77, 78].

Screen for β-Lactamase Genes by NGS and Bioinformatics Analysis

Isolates displaying elevated MIC results (≥2 mg/L) for ceftriaxone, aztreonam, ceftazidime, and/or imipenem/meropenem were selected. A total of 11.6% (n = 249) and 16.1% (n = 97) of *E. coli*–causing UTIs and BSIs, respectively, met the MIC-based criteria, whereas 11.0% (n = 53) and 13.7% (n = 40) of *K. pneumoniae* isolates causing UTIs and BSIs, respectively, met the criteria (data not shown). Total genomic DNA was extracted using the Thermo Scientific KingFisher Flex Magnetic Particle Processor (Cleveland, OH), which was used as input material for library construction. DNA libraries were prepared using the NexteraXT library construction protocol (Illumina, San Diego, CA) following the manufacturer's instructions and were sequenced on a MiSeq Sequencer (JMI Laboratories). FASTQ format sequencing files for each sample set were assembled independently using de novo assembler SPAdes 3.9.0, and a JMI Laboratories–designed software workflow was applied to the assembled sequences to align against a database containing known β-lactamase-encoding genes.

Epidemiology Typing

The same JMI Laboratories–designed software analysis workflow utilized for screening β-lactamase-encoding genes was employed on assembled genomes to identify and extract previously defined sets of 7 housekeeping gene fragments

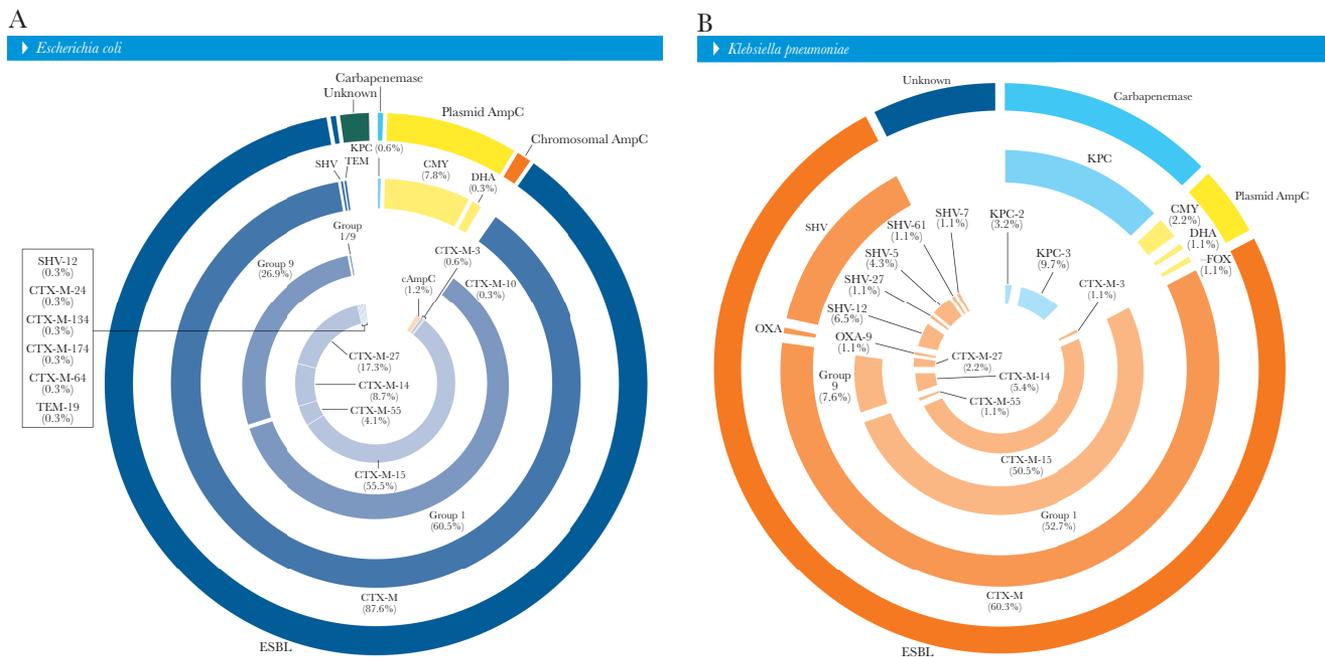**Figure 1.** Overall distribution of β-lactamase-encoding genes detected in *Escherichia coli* (A) and *Klebsiella pneumoniae* (B) causing urinary tract infections or bloodstream infections. Percentages obtained from total number of *E. coli* (346) or *K. pneumoniae* (93) as denominators. Chromosomal AmpC defined by the overexpression of AmpC; unknown defined as the lack of any β-lactamase screened during the study.

(~500 bp) for multilocus sequence typing (MLST) assignment. Briefly, once extracted, each fragment was compared with known allele variants for each locus (housekeeping gene) on the respective MLST website (http://bigsdb.pasteur.fr/klebsiella/ or http://enterobase.warwick.ac.uk/species/ecoli). An allele sharing 100% genetic identity with a known variant received its numeric designation, and a 7-number sequence (1 for each housekeeping gene) formed an allelic profile, defined as sequence types (STs). ST profiles sharing 100% genetic identity in at least 6 of 7 MLST loci were grouped into a clonal complex (CC), named after its presumed ancestral genotype. Isolates containing alleles that did not match an existing sequence in the MLST database were submitted/deposited in the respective database for allele and ST assignments.

RESULTS

In general, most *E. coli* (88.2% [305]) responsible for UTIs or BSIs that met the MIC-based screening criteria carried ESBL-encoding genes (99.3% [303] of CTX-M, 0.3% [1] of SVH-12, and 0.3% [1] of TEM-19), followed by isolates associated with AmpC (9.2% [32]), either carrying plasmid AmpC (81.3% [26] of CMY-2, 3.1% [1] of CMY-42, or 3.1% [1] of DHA-1) or showing overexpression of the chromosomal AmpC (12.5% [4]). A total of 2.0% (n = 7) of *E. coli* did not show any of the β-lactamase genes investigated and were assigned here as unknown. Two (0.6%) *E. coli* clinical isolates harbored 1 each of the *Klebsiella pneumoniae* carbapenemase (KPC)–2 and KPC-3-encoding genes. In addition, among *E. coli* isolates carrying bla_{CTX-M} variants, CTX-M group 1 (60.5% overall or 69.3% within CTX-M genes) and group 9 (26.9% overall or 30.7% within CTX-M genes) prevailed (Figure 1A).
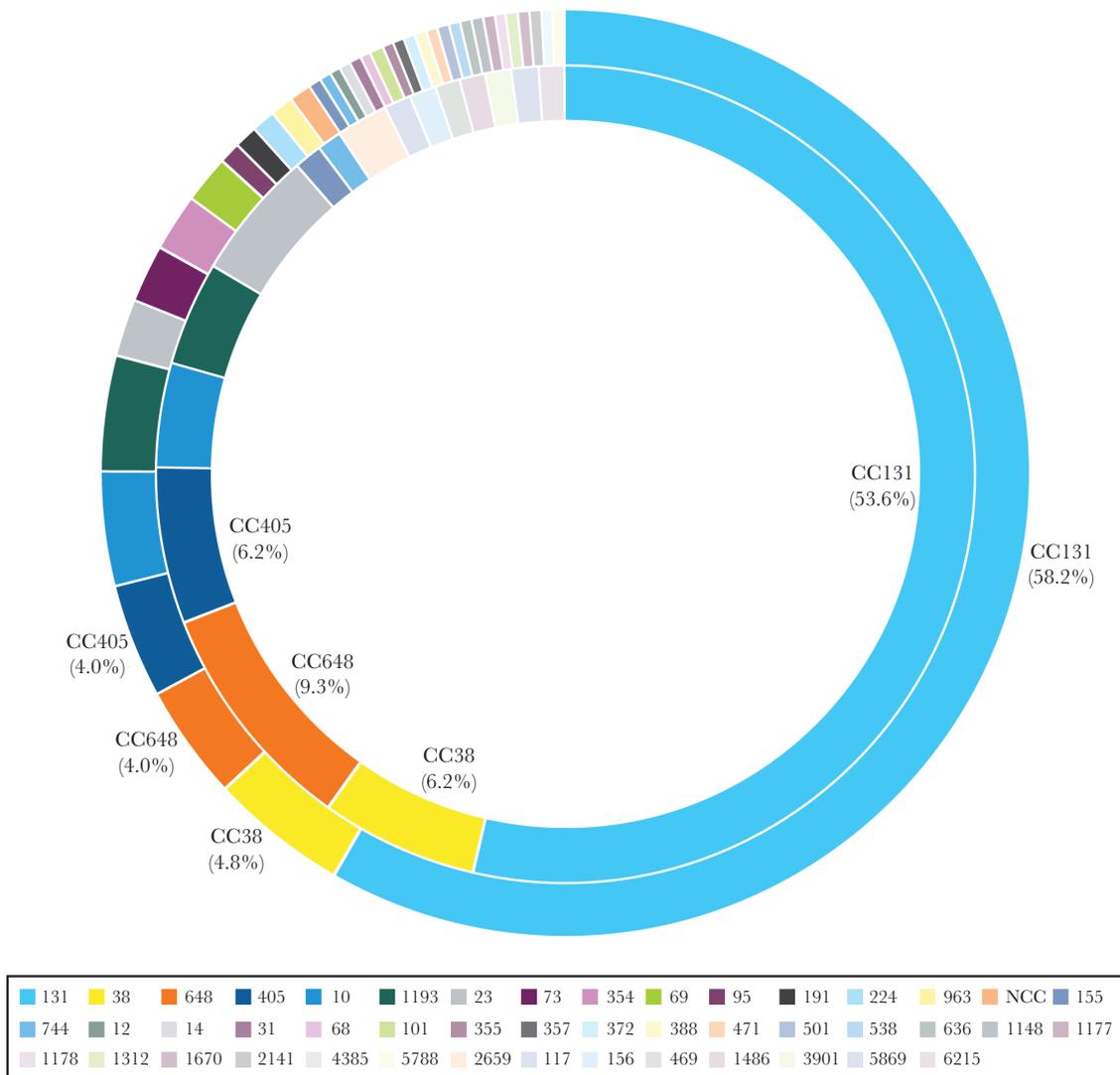

Figure 2. Distribution of clonal complexes among *Escherichia coli* isolates recovered from urinary tract infections (outer circle) and bloodstream infections (inner circle).

Among *K. pneumoniae* isolates responsible for UTIs or BSIs that met the MIC-based screening criteria, 75.3% (n = 70) carried ESBL-encoding genes (80.0% [56] of CTX-M, 18.6% [13] of SVH variants, and 1.4% [1] of OXA-9), whereas 4.4% (n = 4) had plasmid AmpC (2.2% [2] of CMY-2, 1.1% [1] of FOX-5, or 1.1% [1] of DHA-1). A total of 12.9% of *K. pneumoniae* clinical isolates harbored KPC-encoding genes (9 KPC-3 and 3 KPC-2). A total of 7.5% (n = 7) of *K. pneumoniae* isolates did not show any of the β-lactamase genes investigated. Among *K. pneumoniae* isolates, 52.7% (or 87.5% within CTX-M) and 7.6% (or 12.5% within CTX-M) of bla_{CTX-M} variants belonged to groups 1 and 9, respectively (Figure 1B).

Among KPC-producing *K. pneumoniae* (2 from BSI and 11 from UTI), 84.6% (11/13) were ST258 (CC258). Seventeen and 38 unique CCs were noted in *E. coli*–causing BSIs and UTIs, respectively, and CC131 (or ST131) was the most common CC among BSI (53.6%) and UTI (58.2%) *E. coli* isolates (Figure 2). Twenty-four and 26 CCs were noted among *K. pneumoniae*

isolates causing BSIs and UTIs, respectively. CC258 (28.3%; all but 1 ST1326 were ST258) prevailed in UTI pathogens, whereas CC307 (15.0%) was the most common CC among BSI isolates (Figure 3).

Among *E. coli* isolates, the majority of $bla_{CTX-M-15}$ (70.3%) were detected within ST131 isolates, with similar occurrences among those causing BSIs (68.4%) and UTIs (71.1%) (data not shown). Most ST131 *E. coli* isolates causing BSIs carried bla_{CTX-M} genes belonging to CTX-M group 1 (75.0%), and a small percentage of genes belonged to group 9 (17.3%) (Figure 4). ST131 *E. coli* isolates causing UTIs had a greater number of bla_{CTX-M} genes belonging to CTX-M group 9 (30.4%) when compared with isolates causing BSIs (17.3%) (Figure 4).

DISCUSSION

The study presented here illustrates the applicability of NGS and proper data analysis to evaluate the presence and distribution of β-lactamase-encoding genes and bacterial lineages among
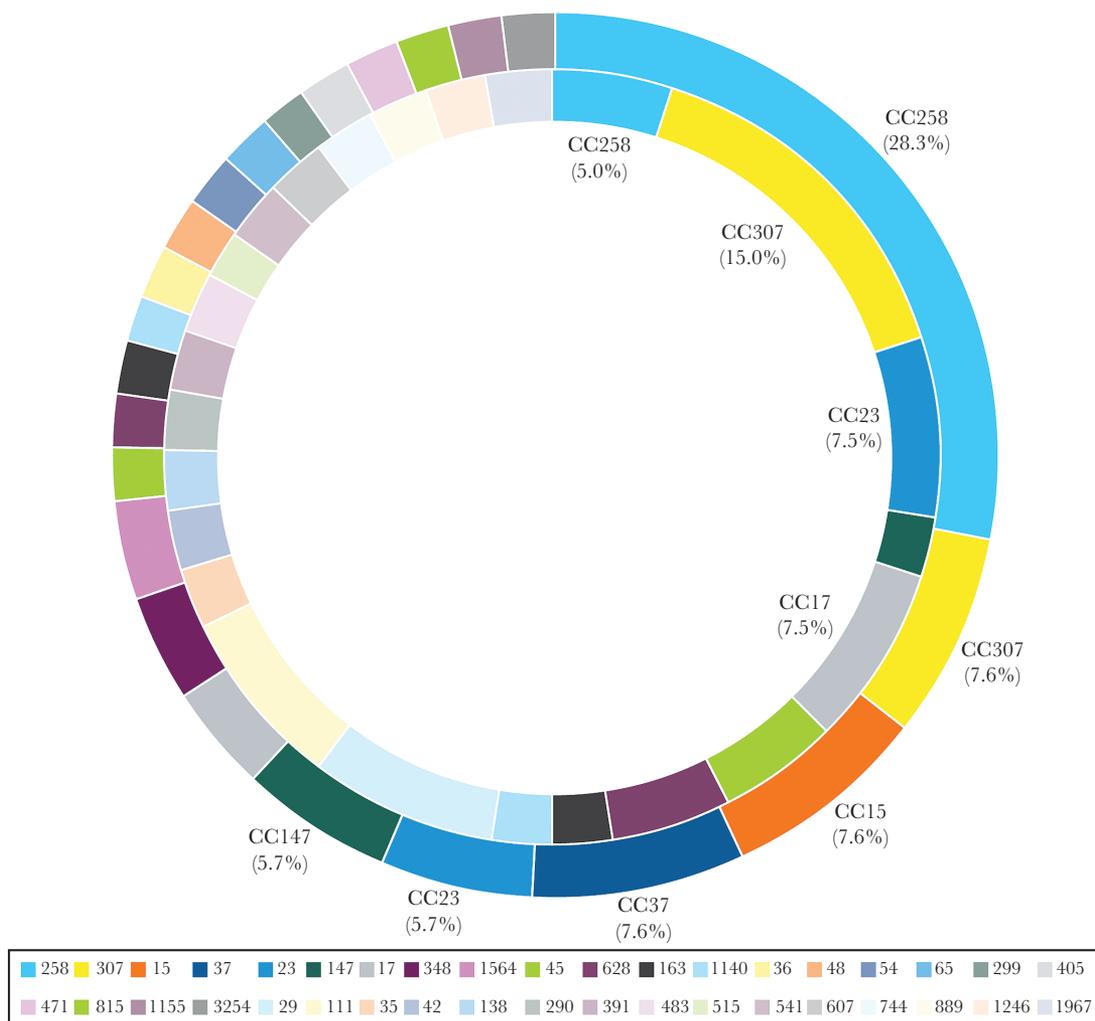

Figure 3. Distribution of clonal complexes among *Klebsiella pneumoniae* isolates recovered from urinary tract infections (outer circle) and bloodstream infections (inner circle).
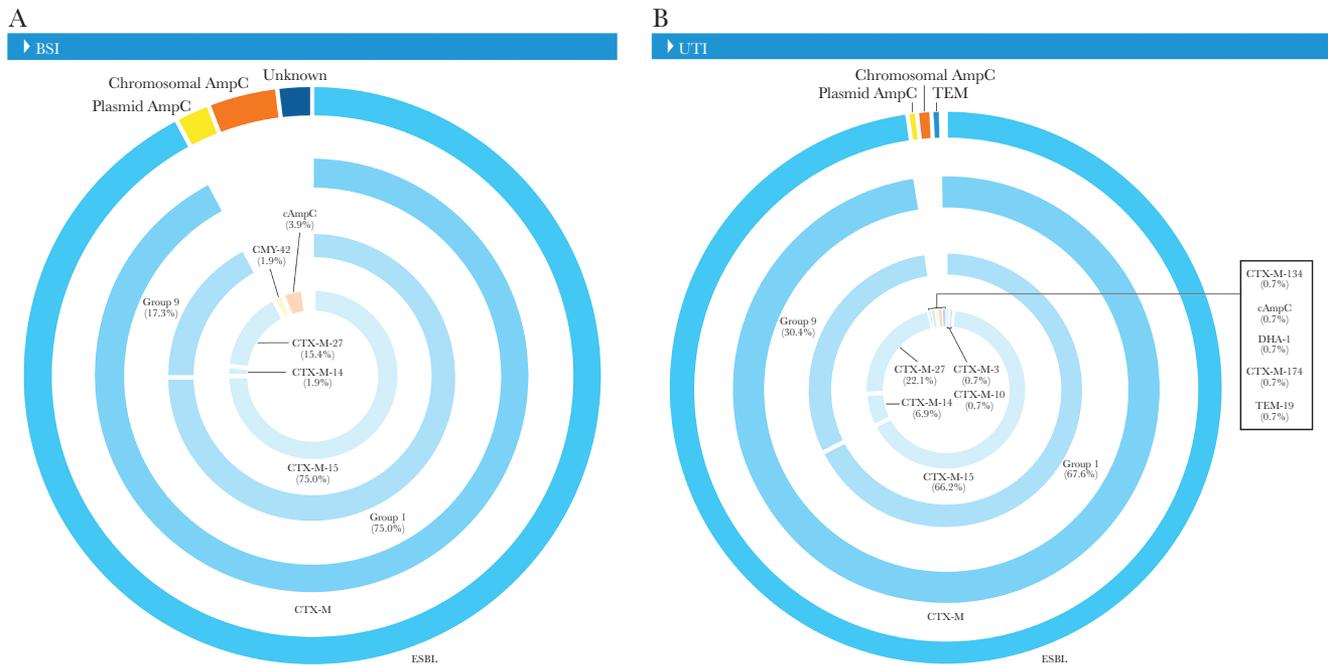
Figure 4. Distribution of β-lactamase-encoding genes detected in *Escherichia coli* ST131 causing bloodstream infection (A) and urinary tract infection (B). Chromosomal AmpC defined by the overexpression of AmpC; unknown defined as the lack of any extended-spectrum β-lactamase screened during the study. Abbreviations: BSI, bloodstream infection; UTI, urinary tract infection.

specific groups of organisms causing infections in humans. *E. coli* and *K. pneumoniae* are prevalent *Enterobacteriaceae* causing a variety of infection types in hospital and community settings. The data clearly show the dissemination of *E. coli* ST131 isolates in the United States, regardless of infection site, whereas the population structure of *K. pneumoniae* was more heterogeneous. In addition, the data confirm the dominance of group 1 CTX-M-encoding genes among *E. coli, E. coli* ST131, and *K. pneumoniae*; however, the data presented here may also suggest that genes belonging to CTX-M group 9 may be spreading among UTI *E. coli* isolates. A recent clinical trial comparing piperacillin-tazobactam and meropenem for treating BSIs caused by *E. coli* or *Klebsiella* spp. that were nonsusceptible to ceftriaxone and susceptible to piperacillin-tazobactam also reported the dominance of bla_{CTX-M} genes (83.5%), especially $bla_{CTX-M-15}$ (54.5%) [35].

Several studies have demonstrated the dominance of ST131, but often among isolates from small geographical regions or local studies and among a certain population or a convenience sample. A previous study that investigated the timing and tempo of the ST131 expansion in the United States that also used a sampling of *E. coli* isolates from the SENTRY Program reported a rate of 0.0% of ST131 among ESBL-producing *E. coli* in 2000. This rate increased to 45% in 2009 [79]. The study presented here showed a combined rate (UTI and BSI isolates) of 57% of ST131 in ESBL-producing *E. coli* in 2016 (data not shown). These data indicate that the proportion of ST131 increased significantly among ESBL-producing *E. coli* between 2000 and

2009 in the United States and that it continues to increase, but perhaps at a slower pace. In addition, the ST131 rate reported here (57%) was very similar to that (56.8%) observed in a recent clinical trial, although the geographic origins of isolates were distinct [35]. Continuous surveillance over time should provide more granular information regarding the expansion of ST131 and the proportion of genes belonging to CTX-M group 1 and group 9, as the latter group is increasingly being reported in the Asia-Pacific region [80].

Not many publications exist on the population genomic analysis of ESBL-producing *K. pneumoniae* isolates in the United States. In contrast, much attention has been devoted to the emergence and dissemination of carbapenemase-producing *K. pneumoniae*, which has been largely attributed to the clonal expansion of a dominant genetic lineage, the clonal group (CG) 258 (includes ST258 and single-locus variants ST11, ST512, ST340, and ST437) carrying carbapenemase-encoding genes [81–84]. The population structure of *K. pneumoniae* isolates included here was more heterogeneous than that observed in *E. coli* isolates, which emphasizes the opportunistic nature of these species. However, the results obtained among KPC-producing *K. pneumoniae* also reflect the well-known dominance of CC258, as 84.6% of these isolates belonged to ST258.

Interestingly, the prevalence rates of *E. coli* (11.6%–16.1%) and *K. pneumoniae* (11.0%–13.7%) isolates that met the MIC-based screening criteria for ESBL production in this study (2016) were similar. Previous studies published in the late 1990s reported rates of ESBL production among *E. coli*–causing

(0.4%–1.4%) and *K. pneumoniae*–causing UTIs (4.2%–6.6%) [85] or *K. pneumoniae*–causing (5.8%) BSIs [86] in the United States that were much lower than those observed here. Another older study reported prevalence rates of ESBL-producing *E. coli* and *K. pneumoniae* of 3.3% and 7.6%, respectively, among isolates causing various infection types in the United States (1997–1999) [87]. It appears that the prevalence of the ESBL phenotype in *K. pneumoniae* used to be higher than its prevalence in *E. coli*. The data presented here suggest that the difference in ESBL phenotypes between species no longer exists and that it may, at least in part, be explained by the successful dissemination of *E. coli* ST131. No prevalent clonal type exists among *K. pneumoniae*, except for a small predominance of CC11 (28.3%) among UTI pathogens and CC307 (15.0%) among BSI isolates, which differs from *E. coli*.

NGS use will likely increase continuously, and its numerous applications depend on a proper bioinformatics workflow to obtain the required data in a report format that would be useful, whether it relates to a single sample/pathogen or a greater collection of samples/isolates. JMI Laboratories has used NGS to characterize isolates originating from surveillance or clinical trial studies, and the vast majority of the investigations have involved characterization of resistance mechanisms associated with a particular class of agents, any resistance genes, or epidemiology typing. Some studies have also utilized NGS to investigate the genetic context of resistance genes of interest [26, 88]. Also, importantly, storing digital data has become cheaper and more accessible with cloud-based options, and permanently storing digital sequencing data makes it possible for timeless ad hoc in silico analysis.

Furthermore, this study provides a benchmark for the distribution of β-lactamase genes and the population structure information obtained by NGS among the most common *Enterobacteriaceae* species responsible for BSIs and UTIs in US medical centers during the SENTRY Program for 2016. Coupling high-throughput genome sequencing with powerful bioinformatics tools can be useful to detect and monitor resistance genes and bacterial populations in surveillance programs and other investigations in different bacterial collections. Future studies that include isolates recovered from US and non-US medical centers during multiple years can contribute to the rapidly growing knowledge base of the distribution of resistance/virulence genes and the evolution of bacterial populations within the context of clinically relevant human pathogens.

References

1. Gales AC, Gordon KA, Wilke WW, et al. Occurrence of single-point *gyr*A mutations among ciprofloxacin-susceptible *Escherichia coli* isolates causing urinary tract infections in Latin America. Diagn Microbiol Infect Dis **2000**; 36:61–4.
2. Gales AC, Biedenbach DJ, Winokur P, et al. Carbapenem-resistant *Serratia marcescens* isolates producing Bush group 2f beta-lactamase (SME-1) in the United States: results from the MYSTIC Programme. Diagn Microbiol Infect Dis **2001**; 39:125–7.
3. Kugler KC, Denys GA, Wilson ML, Jones RN. Serious streptococcal infections produced by isolates resistant to streptogramins (quinupristin/dalfopristin): case reports from the SENTRY antimicrobial surveillance program. Diagn Microbiol Infect Dis **2000**; 36:269–72.
4. Jones RN, Deshpande L, Fritsche TR, Sader HS. Determination of epidemic clonality among multidrug-resistant strains of *Acinetobacter* spp. and *Pseudomonas aeruginosa* in the MYSTIC Programme (USA, 1999-2003). Diagn Microbiol Infect Dis **2004**; 49:211–6.
5. Deshpande LM, Rhomberg PR, Sader HS, Jones RN. Emergence of serine carbapenemases (KPC and SME) among clinical strains of *Enterobacteriaceae* isolated in the United States medical centers: report from the MYSTIC Program (1999-2005). Diagn Microbiol Infect Dis **2006**; 56:367–72.
6. Goering RV. Pulsed field gel electrophoresis: a review of application and interpretation in the molecular epidemiology of infectious disease. Infect Genet Evol **2010**; 10:866–75.
7. Sader HS, Hollis RJ, Pfaller MA. The use of molecular techniques in the epidemiology and control of infectious diseases. Clin Lab Med **1995**; 15:407–31.
8. Miernyk K, Debyle C, Harker-Jones M, et al. Serotyping of *Streptococcus pneumoniae* isolates from nasopharyngeal samples: use of an algorithm combining microbiologic, serologic, and sequential multiplex PCR techniques. J Clin Microbiol **2011**; 49:3209–14.
9. Pai R, Gertz RE, Beall B. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. J Clin Microbiol **2006**; 44:124–31.
10. Wilke WW, Marshall SA, Coffman SL, et al. Vancomycin-resistant *Enterococcus raffinosus*: molecular epidemiology, species identification error, and frequency of occurrence in a national resistance surveillance program. Diagn Microbiol Infect Dis **1997**; 29:43–9.
11. Marshall SA, Wilke WW, Pfaller MA, Jones RN. *Staphylococcus aureus* and coagulase-negative staphylococci from blood stream infections: frequency of occurrence, antimicrobial susceptibility, and molecular (mecA) characterization of oxacillin resistance in the SCOPE program. Diagn Microbiol Infect Dis **1998**; 30:205–14.
12. Nelson GE, Pondo T, Toews KA, et al. Epidemiology of invasive group A streptococcal infections in the United States, 2005-2012. Clin Infect Dis **2016**; 63:478–86.
13. Limbago B, Fosheim GE, Schoonover V, et al; Active Bacterial Core surveillance MRSA Investigators. Characterization of methicillin-resistant *Staphylococcus aureus* isolates collected in 2005 and 2006 from patients with invasive disease: a population-based analysis. J Clin Microbiol **2009**; 47:1344–51.
14. Mendes RE, Bell JM, Turnidge JD, et al. Emergence and widespread dissemination of OXA-23, -24/40 and -58 carbapenemases among *Acinetobacter* spp. in Asia-Pacific nations: report from the SENTRY surveillance program. J Antimicrob Chemother **2009**; 63:55–9.
15. Mendes RE, Kiyota KA, Monteiro J, et al. Rapid detection and identification of metallo-beta-lactamase-encoding genes by multiplex real-time PCR assay and melt curve analysis. J Clin Microbiol **2007**; 45:544–7.
16. Pfaller MA, Acar J, Jones RN, et al. Integration of molecular characterization of microorganisms in a global antimicrobial resistance surveillance program. Clin Infect Dis **2001**; 32(Suppl 2):S156–67.
17. Jones RN, Marshall SA, Pfaller MA, et al. Nosocomial enterococcal blood stream infections in the SCOPE Program: antimicrobial resistance, species occurrence, molecular testing results, and laboratory testing accuracy. SCOPE Hospital Study Group. Diagn Microbiol Infect Dis **1997**; 29:95–102.
18. Diekema DJ, Pfaller MA, Turnidge J, et al; Sentry Participants Group. Genetic relatedness of multidrug-resistant, methicillin (oxacillin)-resistant *Staphylococcus aureus* bloodstream isolates from SENTRY antimicrobial resistance surveillance centers worldwide, 1998. Microb Drug Resist **2000**; 6:213–21.
19. Castanheira M, Farrell SE, Deshpande LM, et al. Prevalence of β-lactamase-encoding genes among *Enterobacteriaceae* bacteremia isolates collected in 26 U.S. hospitals: report from the SENTRY antimicrobial surveillance program (2010). Antimicrob Agents Chemother **2013**; 57:3012–20.
20. Karlowsky JA, Lob SH, Kazmierczak KM, et al. In vitro activity of imipenem against carbapenemase-positive *Enterobacteriaceae* isolates collected by the SMART global surveillance program from 2008 to 2014. J Clin Microbiol **2017**; 55:1638–49.

21. Nitschke H, Slickers P, Müller E, et al. DNA microarray-based typing of *Streptococcus agalactiae* isolates. J Clin Microbiol **2014**; 52:3933–43.

22. Sanchini A, Campanile F, Monaco M, et al. DNA microarray-based characterisation of Panton-Valentine leukocidin-positive community-acquired methicillin-resistant *Staphylococcus aureus* from Italy. Eur J Clin Microbiol Infect Dis **2011**; 30:1399–408.

23. Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. Nat Rev Microbiol **2015**; 13:787–94.

24. Dark MJ. Whole-genome sequencing in bacteriology: state of the art. Infect Drug Resist **2013**; 6:115–23.

25. Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. Microbiology **2018**; 164:1213–19.

26. Deshpande LM, Castanheira M, Flamm RK, Mendes RE. Evolving oxazolidinone resistance mechanisms in a worldwide collection of enterococcal clinical isolates: results from the SENTRY antimicrobial surveillance program. J Antimicrob Chemother **2018**; 73:2314–22.

27. Rhoads DD. Commentary: lowering the barriers to routine whole genome sequencing of bacteria in the clinical microbiology laboratory. J Clin Microbiol **2018**; 56:e00813–18.

28. Yang Y, Niehaus KE, Walker TM, et al. Machine learning for classifying tuberculosis drug-resistance from DNA sequencing data. Bioinformatics **2018**; 34:1666–71.

29. Mendes RE, Castanheira M, Woosley LN, et al. Molecular β-lactamase characterization of Gram-negative pathogens recovered from patients enrolled in the ceftazidime-avibactam phase 3 trials (RECAPTURE 1 and 2) for complicated urinary tract infections: efficacies analysed against susceptible and resistant subsets. Int J Antimicrob Agents **2018**; 52:287–92.

30. Mendes RE, Castanheira M, Woosley LN, et al. Molecular β-lactamase characterization of aerobic gram-negative pathogens recovered from patients enrolled in the ceftazidime-avibactam phase 3 trials for complicated intra-abdominal infections, with efficacies analyzed against susceptible and resistant subsets. Antimicrob Agents Chemother **2017**; 61:e02447.

31. Mendes RE, Castanheira M, Gasink L, et al. β-Lactamase characterization of gram-negative pathogens recovered from patients enrolled in the phase 2 trials for ceftazidime-avibactam: clinical efficacies analyzed against subsets of molecularly characterized isolates. Antimicrob Agents Chemother **2015**; 60:1328–35.

32. Mendes RE, Sader HS, Deshpande LM, et al. Characterization of baseline methicillin-resistant *Staphylococcus aureus* isolates recovered from phase IV clinical trial for linezolid. J Clin Microbiol **2010**; 48:568–74.

33. Mendes RE, Deshpande LM, Smyth DS, et al. Characterization of methicillin-resistant *Staphylococcus aureus* strains recovered from a phase IV clinical trial for linezolid versus vancomycin for treatment of nosocomial pneumonia. J Clin Microbiol **2012**; 50:3694–702.

34. Campbell SJ, Deshmukh HS, Nelson CL, et al. Genotypic characteristics of *Staphylococcus aureus* isolates from a multinational trial of complicated skin and skin structure infections. J Clin Microbiol **2008**; 46:678–84.

35. Harris PNA, Tambyah PA, Lye DC, et al; MERINO Trial Investigators and the Australasian Society for Infectious Disease Clinical Research Network (ASID-CRN). Effect of piperacillin-tazobactam vs meropenem on 30-day mortality for patients with *E coli* or *Klebsiella pneumoniae* bloodstream infection and ceftriaxone resistance: a randomized clinical trial. JAMA **2018**; 320:984–94.

36. Mendes RE, Shortridge D, Woosley LN, et al. Molecular characterization of clinical trial isolates exhibiting increased MIC results during fosfomycin (ZTI-01) treatment in a phase 2/3 clinical trial for complicated urinary tract infections (ZEUS). Paper presented at: ASM Microbe 2018; June 7–11 2018; Atlanta, GA.

37. Forbes JD, Knox NC, Peterson CL, Reimer AR. Highlighting clinical metagenomics for enhanced diagnostic decision-making: a step towards wider implementation. Comput Struct Biotechnol J **2018**; 16:108–20.

38. Wu K, O'Grady H, Workentine M, et al. Impact of ascending dosages of cadazolid or vancomycin on the intestinal microbiome during treatment of *Clostridium difficile* infection. Open Forum Infect Dis **2015**; 2(Suppl 1):S149.

39. Cannon K, Byrne B, Happe J, et al. Enteric microbiome profiles during a randomized phase 2 clinical trial of surotomycin versus vancomycin for the treatment of *Clostridium difficile* infection. J Antimicrob Chemother **2017**; 72:3453–61.

40. Sabat AJ, van Zanten E, Akkerboom V, et al. Targeted next-generation sequencing of the 16S-23S rRNA region for culture-independent bacterial identification – increased discrimination of closely related species. Sci Rep **2017**; 7:3434.

41. Lesho E, Clifford R, Onmus-Leone F, et al. The challenges of implementing next generation sequencing across a large healthcare system, and the molecular epidemiology and antibiotic susceptibilities of carbapenemase-producing bacteria in the healthcare system of the U.S. Department of Defense. Plos One **2016**; 11:e0155770.

42. Deurenberg RH, Bathoorn E, Chlebowicz MA, et al. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol **2017**; 243:16–24.

43. Angers-Loustau A, Petrillo M, Bengtsson-Palme J, et al. The challenges of designing a benchmark strategy for bioinformatics pipelines in the identification of antimicrobial resistance determinants using next generation sequencing technologies. F1000Res **2018**; 7:459.

44. Zankari E, Hasman H, Kaas RS, et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother **2013**; 68:771–7.

45. Lakin SM, Dean C, Noyes NR, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res **2017**; 45:D574–80.

46. Gupta SK, Padmanabhan BR, Diene SM, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother **2014**; 58:212–20.

47. de Man TJ, Limbago BM. SSTAR, a stand-alone easy-to-use antimicrobial resistance gene predictor. mSphere **2016**; 1:e00050.

48. Rice LB, Carias LL, Hujer AM, et al. High-level expression of chromosomally encoded SHV-1 beta-lactamase and an outer membrane protein change confer resistance to ceftazidime and piperacillin-tazobactam in a clinical isolate of *Klebsiella pneumoniae*. Antimicrob Agents Chemother **2000**; 44:362–7.

49. Héritier C, Poirel L, Fournier PE, et al. Characterization of the naturally occurring oxacillinase of *Acinetobacter baumannii*. Antimicrob Agents Chemother **2005**; 49:4174–9.

50. Figueiredo S, Poirel L, Papa A, et al. Overexpression of the naturally occurring blaOXA-51 gene in *Acinetobacter baumannii* mediated by novel insertion sequence ISA*ba9*. Antimicrob Agents Chemother **2009**; 53:4045–7.

51. Poirel L, Figueiredo S, Cattoir V, et al. Acinetobacter radioresistens as a silent source of carbapenem resistance for *Acinetobacter* spp. Antimicrob Agents Chemother **2008**; 52:1252–6.

52. Figueiredo S, Bonnin RA, Poirel L, et al. Identification of the naturally occurring genes encoding carbapenem-hydrolysing oxacillinases from *Acinetobacter haemolyticus*, *Acinetobacter johnsonii*, and *Acinetobacter calcoaceticus*. Clin Microbiol Infect **2012**; 18:907–13.

53. Shannon K, Williams H, King A, Phillips I. Hyperproduction of TEM-1 beta-lactamase in clinical isolates of *Escherichia coli* serotype O15. FEMS Microbiol Lett **1990**; 55:319–23.

54. Hanson ND, Sanders CC. Regulation of inducible AmpC beta-lactamase expression among *Enterobacteriaceae*. Curr Pharm Des **1999**; 5:881–94.

55. Bou G, Martínez-Beltrán J. Cloning, nucleotide sequencing, and analysis of the gene encoding an AmpC beta-lactamase in *Acinetobacter baumannii*. Antimicrob Agents Chemother **2000**; 44:428–32.

56. Hamidian M, Hall RM. Tn6168, a transposon carrying an ISAba1-activated ampC gene and conferring cephalosporin resistance in *Acinetobacter baumannii*. J Antimicrob Chemother **2014**; 69:77–80.

57. Hot C, Berthet N, Chesneau O. Characterization of sal(A), a novel gene responsible for lincosamide and streptogramin A resistance in *Staphylococcus sciuri*. Antimicrob Agents Chemother **2014**; 58:3335–41.

58. Chalhoub H, Saenz Y, Nichols WW, et al. Loss of activity of ceftazidime-avibactam due to Mex-AB-OprM efflux and overproduction of AmpC cephalosporinase in *Pseudomonas aeruginosa* isolated from patients suffering from cystic fibrosis. Int J Antimicrob Agents **2018**; 52:697–701.

59. Housseini BIK, Phan G, Broutin I. Functional mechanism of the efflux pumps transcription regulators from *Pseudomonas aeruginosa* based on 3D structures. Front Mol Biosci **2018**; 5:57.

60. Blanco P, Hernando-Amado S, Reales-Calderon JA, et al. Bacterial multidrug efflux pumps: much more than antibiotic resistance determinants. Microorganisms **2016**; 4:14.

61. Li XZ, Plésiat P, Nikaido H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. Clin Microbiol Rev **2015**; 28:337–418.

62. Schindler BD, Kaatz GW. Multidrug efflux pumps of Gram-positive bacteria. Drug Resist Updat **2016**; 27:1–13.

63. Martínez-Martínez L. Extended-spectrum beta-lactamases and the permeability barrier. Clin Microbiol Infect **2008**; 14(Suppl 1):82–9.

64. Alm RA, Johnstone MR, Lahiri SD. Characterization of *Escherichia coli* NDM isolates with decreased susceptibility to aztreonam/avibactam: role of a novel insertion in PBP3. J Antimicrob Chemother **2015**; 70:1420–8.

65. Zapun A, Contreras-Martel C, Vernet T. Penicillin-binding proteins and beta-lactam resistance. FEMS Microbiol Rev **2008**; 32:361–85.

66. Cannavino CR, Mendes RE, Sader HS, et al. Evolution of ceftaroline-resistant mrsa in a child with cystic fibrosis following repeated antibiotic exposure. Pediatr Infect Dis J **2016**; 35:813–5.

67. Sanchez EH, Mendes RE, Sader HS, Allison GM. In vivo emergence of ceftaroline resistance during therapy for MRSA vertebral osteomyelitis. J Antimicrob Chemother **2016**; 71:1736–8.

68. Mendes RE, Tsakris A, Sader HS, et al. Characterization of methicillin-resistant *Staphylococcus aureus* displaying increased MICs of ceftaroline. J Antimicrob Chemother **2012**; 67:1321–4.

69. Pringle M, Poehlsgaard J, Vester B, Long KS. Mutations in ribosomal protein L3 and 23S ribosomal RNA at the peptidyl transferase centre are associated with reduced susceptibility to tiamulin in *Brachyspira* spp. isolates. Mol Microbiol **2004**; 54:1295–306.

70. Long KS, Poehlsgaard J, Hansen LH, et al. Single 23S rRNA mutations at the ribosomal peptidyl transferase centre confer resistance to valnemulin and other antibiotics in *Mycobacterium smegmatis* by perturbation of the drug binding pocket. Mol Microbiol **2009**; 71:1218–27.

71. Locke JB, Hilgers M, Shaw KJ. Mutations in ribosomal protein L3 are associated with oxazolidinone resistance in staphylococci of clinical origin. Antimicrob Agents Chemother **2009**; 53:5275–8.

72. Klitgaard RN, Ntokou E, Nørgaard K, et al. Mutations in the bacterial ribosomal protein l3 and their association with antibiotic resistance. Antimicrob Agents Chemother **2015**; 59:3518–28.

73. Locke JB, Hilgers M, Shaw KJ. Novel ribosomal mutations in *Staphylococcus aureus* strains identified through selection with the oxazolidinones linezolid and torezolid (TR-700). Antimicrob Agents Chemother **2009**; 53:5265–74.

74. Long KS, Vester B. Resistance to linezolid caused by modifications at its binding site on the ribosome. Antimicrob Agents Chemother **2012**; 56:603–12.

75. Gregory ST, Dahlberg AE. Erythromycin resistance mutations in ribosomal proteins L22 and L4 perturb the higher order structure of 23 S ribosomal RNA. J Mol Biol **1999**; 289:827–34.

76. Mendes RE, Deshpande LM, Jones RN. Linezolid update: stable in vitro activity following more than a decade of clinical use and summary of associated resistance mechanisms. Drug Resist Updat **2014**; 17:1–12.

77. CLSI. *M07Ed11E. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard*. 11th ed. Wayne, PA: Clinical and Laboratory Standards Institute; **2018**.

78. CLSI. *M100Ed28E. Performance Standards for Antimicrobial Susceptibility Testing: 28th Informational Supplement*. Wayne, PA: Clinical and Laboratory Standards Institute; **2018**.

79. Johnson JR, Porter S, Thuras P, Castanheira M. Epidemic emergence in the United States of *Escherichia coli* sequence type 131-H30 (ST131-H30), 2000 to 2009. Antimicrob Agents Chemother **2017**; 61:e00732.

80. Harris PNA, Ben Zakour NL, Roberts LW, et al. Whole genome analysis of cephalosporin-resistant *Escherichia coli* from bloodstream infections in Australia, New Zealand and Singapore: high prevalence of CMY-2 producers and ST131 carrying *bla*CTX-M-15 and *bla*CTX-M-27. J Antimicrob Chemother **2017**; 73:634–42.

81. Bowers JR, Kitchel B, Driebe EM, et al. Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. Plos One **2015**; 10:e0133727.

82. Peirano G, Bradford PA, Kazmierczak KM, et al. Importance of clonal complex 258 and IncFK2-like plasmids among a global collection of *Klebsiella pneumoniae* with *bla*KPC. Antimicrob Agents Chemother **2017**; 61:e02610.

83. Kitchel B, Rasheed JK, Patel JB, et al. Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* isolates in the United States: clonal expansion of multilocus sequence type 258. Antimicrob Agents Chemother **2009**; 53:3365–70.

84. Cuzon G, Naas T, Truong H, et al. Worldwide diversity of *Klebsiella pneumoniae* that produce beta-lactamase bla$_{KPC-2}$ gene. Emerg Infect Dis **2010**; 16:1349–56.

85. Mathai D, Jones RN, Pfaller MA; SENTRY Participant Group North America. Epidemiology and frequency of resistance among pathogens causing urinary tract infections in 1,510 hospitalized patients: a report from the SENTRY Antimicrobial Surveillance Program (North America). Diagn Microbiol Infect Dis **2001**; 40:129–36.

86. Biedenbach DJ, Moet GJ, Jones RN. Occurrence and antimicrobial resistance pattern comparisons among bloodstream infection isolates from the SENTRY Antimicrobial Surveillance Program (1997-2002). Diagn Microbiol Infect Dis **2004**; 50:59–69.

87. Winokur PL, Canton R, Casellas JM, Legakis N. Variations in the prevalence of strains expressing an extended-spectrum beta-lactamase phenotype and characterization of isolates from Europe, the Americas, and the Western Pacific region. Clin Infect Dis **2001**; 32(Suppl 2):S94–103.

88. Mendes RE, Deshpande LM, Bonilla HF, et al. Dissemination of a pSCFS3-like cfr-carrying plasmid in *Staphylococcus aureus* and *Staphylococcus epidermidis* clinical isolates recovered from hospitals in Ohio. Antimicrob Agents Chemother **2013**; 57:2923–8.