Klebsiella pneumoniae has been a leading health care-acquired infection (HAI) agent for decades (1, 2). The emergence of multidrug-resistant K. pneumoniae, especially extended-spectrum β-lactamase (ESBL) producers and carbapenemase producers, has elevated the morbidity and mortality rates and health care costs associated with K. pneumoniae to highly significant levels (3–6). Health care- and outbreak-associated strain types of K. pneumoniae that appear highly transmissible and have a propensity for antimicrobial resistance (AMR) or virulence gene acquisition are a growing proportion of the K. pneumoniae species (7–18). Sequence type 258 (ST258), the crux of the worldwide carbapenemase-producing Enterobacteriaceae (CPE) threat, has disseminated rapidly around the world’s health care systems despite its recent emergence (17). Its progenitor strains in clonal group 258 (CG258) also cause outbreaks and carry many important ESBL- and carbapenemase-encoding genes (9, 19–21). Several other strain types, such as those in CG14, CG20, and CG37, also frequently appear as multidrug resistant and in outbreak situations (7, 10, 12, 15).

Host colonization is likely an important reservoir driving the transmission of these strains. In the health care environment, intestinal colonization of K. pneumoniae is a risk factor for infection (22–24), and carriers of CPE are at high risk for invasive disease (25). Rates of CPE and ESBL-producing K. pneumoniae colonization are rising in patient and health care worker populations, increasing the size of the reservoir and increasing chances of transmission (26, 27). Asymptomatic transmission of multidrug-resistant strains is rapid (16, 28), and transmission events that lead to outbreaks often go undetected (29, 30). Early detection of K. pneumoniae colonization of patients, especially multidrug-resistant K. pneumoniae or epidemic strain type colonization, is now considered critical to infection control (24, 30–33).

Infection control programs that include the detection and isolation of carriers have repeatedly been successful in markedly decreasing multidrug-resistant or epidemic strain infections (31, 34–37), but this practice is uncommon for several reasons. Many of these programs use culture-based methods for detecting CPE or ESBL producers, which have several limitations, including turnaround time, narrow application, fair sensitivity and specificity, and extensive labor for high-throughput screening (31, 38). PCR-based assays are rapid but often use DNA from culture, and a limited number of tests can be run simultaneously, potentially missing important AMR genes not previously known to circulate in a given locale (31, 39).

Next-generation sequencing has gained a foothold in health care with whole-genome sequencing (WGS) for outbreak detection, transmission mapping, and source tracing (40, 41), microbiome sequencing (e.g., targeted 16S rRNA gene sequencing) to understand microbial population structure (42, 43) and with metagenomic sequencing to attempt to determine all of the genetic factors present (44). Although metagenomic sequencing does not require an a priori understanding of the genetic targets in a clinical sample, it does have significant drawbacks, limiting its translation to the clinical microbiology laboratory. Chief among these are that (i) the required amount of sequencing space increases the cost and
time, (ii) limited coverage across targets lessens the confidence in diagnostic calls, and (iii) the necessary computing power and highly complex analysis limit the ability for local analysis. Targeted amplicon sequencing, on the other hand, allows for rapid, cost-effective, highly multiplexed, and accurate detection of numerous clinically important targets directly in clinical samples (45). Such assays have recently been approved by the FDA for clinical diagnostics (46).

In this report, we describe a new amplicon sequencing tool, KlebSeq, for screening and surveillance that detects and characterizes Klebsiella bacteria in complex samples such as wound and nasal swabs or fecal samples without culturing. KlebSeq includes a sizeable panel of assays for species identification, strain identification, and important AMR and virulence gene targets designed to generate information for health care epidemiology and infection prevention. KlebSeq also includes an analysis pipeline for instant interpretation of the data. Results from the screening of a patient population with this system would rule in or rule out the possibilities of particular transmission events and identify patients carrying high-risk strains like ST258 or other multidrug-resistant Klebsiella strains. The highly multiplexed nature of KlebSeq greatly expands the capacity of a single sequencing run, minimizing costs, and allows for high-throughput patient sample testing. This innovation can also serve as a foundation system on which to build in other HAI agents such as Escherichia coli or Staphylococcus aureus and their multiple AMR mechanisms or as a model for many other applications.

**MATERIALS AND METHODS**

**Samples.** Isolates for target identification and assay validation and DNA extracted from clinical specimens were acquired through collaborations with a large hospital reference laboratory that receives specimens from 10 system-wide medical centers in Arizona and from a high-volume private reference laboratory that receives specimens from regional inpatient, long-term care, and outpatient facilities. Isolates were identified with Vitrek 2 (bioMérieux). Clinical specimen types included various respiratory specimens (nasal, ear, and throat swabs; sputum samples; tracheal aspirates; and bronchial alveolar lavage samples), urine, and wound swabs or tissue. DNA was extracted from isolates with the Qiagen DNAeasy Blood and Tissue kit with additional lytic enzymes when appropriate. DNA was extracted from clinical specimens by NucliSens easyMAG (bioMérieux, Durham, NC). DNA from healthy donor fecal samples was acquired from a family microbiome study; samples had been collected from members of seven families over multiple time points. DNA was extracted in accordance with the Earth Microbiome Project protocol (47). All of the samples were obtained from studies approved by the institutional review boards of the participating institutions.

**Assay target identification and assay design.** Figure 1 illustrates the methodologies and resources, also described below, utilized to assay a target library and develop several types of amplicon sequencing assays.

**WGS, single nucleotide polymorphism (SNP) detection, and phylogenetic analysis.** In-house genome libraries were prepared from 31 Klebsiella isolates and 6 non-Klebsiella isolates (to validate KlebSeq assays) with a 500-bp insert size with the KAPA Library Preparation kit and Standard PCR Library Amplification (Kapa Biosystems, Wilmington, MA) and sequenced on Illumina’s GAIIx or MiSeq. Additional in-house genomes that we have previously described were also included and comprised 111 K. pneumoniae, 1 K. quasipneumoniae, and 5 K. variicola genomes. Public genome sequence data from 256 K. pneumoniae, 18 K. quasipneumoniae, and 13 K. variicola isolates were downloaded from the SRA database (http://www.ncbi.nlm.nih.gov/Traces/sra/), and genome sequence data from 177 K. pneumoniae, 4 K. quasipneumoniae, and 11 K. variicola isolates were downloaded from the Assembly database (http://www.ncbi.nlm.nih.gov/assembly), and all passed filters for high quality; i.e., assemblies and SRA data aligned with ≥80% of MGH 78578 or ≥88% of the strict core genome multilocus sequence typing (scgMLST) refer-

---

**FIG 1** Workflow of the amplicon sequencing target library and assay development pipeline.
quences (described below), SRA data at a ≥10× read depth. Accession numbers are listed in Tables S1 and S2 in the supplemental material.

NASP (48), developed for microbial genome analysis, was used to detect SNPs among genomes (see Table S1 in the supplemental material). In brief, reads were aligned with a reference genome, either one concatenated from scgMLST alleles (10) or MGH 78578 (GenBank accession no. CP000647) with Novoalign V3.04.04 (Novacraft Technologies, Selangor, Malaysia) and SNPs were called with GATK version 2.7-2 (49). Data filtered out included SNP loci with <10% coverage or with <90% consensus in any one sample, regions duplicated in the reference genome as identified by Nucler, and SNP loci that were not present in all of the genomes in the data set. In NASP, results were output in a SNP matrix from a core genome common to all of the isolates in the analysis. Phylogenetic trees were generated from the NASP SNP matrices with MEGA 6.0 (50) and subsequently plotted by means of ITOL v2 or v3 (51).

**Genomic target identification.** To find whole gene targets for assay design, selected genomes were assembled with UGAP (https://github.com/jasonsahl/UGAP), which uses the SPAdes genome assembler, version 3.6. For this work (52), Assembled genomes were then run through LS-BSR (53), which generates a list of open reading frames (ORFs) that have high identity among target species genomes and that have low identity or are not present in nontarget genomes. Alleles of the candidate target ORFs were collected by BLAST, including alleles from nontarget genomes, if present. Lastly, alleles of candidate ORFs were aligned for assay design. Canonical SNPs (canSNPs) were identified from the SNP matrix generated by NASP. Sequence flanking each SNP was collected from the NASP reference genome.

**AMR and virulence gene target collection.** AMR and virulence gene sequences were identified and collected in several ways, including from http://www.lahey.org/studies/other.asp#table1, http://www.lahey.org/qnrstudies/, the Klebsiella BGIRepository at bgi.org, public literature, and the NCBI at http://www.ncbi.nlm.nih.gov. Public literature included a paper by Holt et al. (54) in which a species-wide analysis of K. pneumoniae genomes revealed several siderophore systems and other virulence factors associated more with infections than with colonizing strains. AMR genes included the major ESBL and carbapenemase genes and plasmid-mediated quinolone resistance determinants, as well as the gyrA and parC chromosome genes, several aminoglycoside resistance genes, trimethoprim-sulfamethoxazole, tetracycline, streptomycin, chloramphenicol, and fosfomycin resistance genes, and the recently discovered plasmid-mediated colistin resistance gene mcr-1. Virulence targets included several siderophore systems, for which multiple genes from each were used as assay targets; the regulator of the mucoid phenotype (an indicator of hypervirulence); the wzi gene for capsule typing, for which we used the previously published assay (55); and two genes highly associated with invasive infection, pK2044_00025 and pK2044_00325 (54). For genes that consist of highly diverse alleles, for example, blaCTXMA, gyrB, or drfA, phylogenetic trees based on nucleotide sequences were generated in order to group similar alleles for assay design.

**Assay design and validation.** Gene-based target alleles were aligned in SeqMan (DNAStar, Madison, WI) to identify conserved regions for primer design, and assays were designed with guidance from RealTimeDesign (Biosearch Technologies, Petaluma, CA), or gene-based assays were generated with AlleleID (Premier Biosoft, Palo Alto, CA), which designs assays to capture alleles in an alignment rather than individual sequences. SNP assay primers were designed with RealTimeDesign, and primer sequences were checked for conservation in the NASP SNP matrix. Lastly, assays were run through BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to check for cross-reactivity with other relevant targets or species, including human. Universal tails were added to each primer sequence for library preparation as described below. The assays and their primer sequence are listed in Table S3 in the supplemental material.

Individual assays were screened across positive controls when they were accessible and screened across several isolate genomic DNAs (gDNAs) to test robustness. Additionally, multiplex PCR was validated by initial gene-specific PCR (described below), followed by PCR product dilution and then screening of individual assays by Sybr green-based quantitative PCR (qPCR). For this, 10-μl reaction mixtures of 1× Platinum SYBR green qPCR SuperMix (ThermoFisher Scientific, Waltham, MA), 200 nM forward and reverse primers of one assay, and 1 μl of diluted multiplex PCR product were run at 95°C for initial denaturation for 4 min and then 40 cycles of 95°C for 15 s and 60°C for 1 min. Lastly, several panels of known isolate DNAs were screened by the amplicon sequencing method to test the sensitivity and specificity of the species and strain identification assays. AMR and virulence gene assays were validated by comparing amplicon sequencing results with WGS data.

**Amplicon library preparation and sequencing.** Amplicon library preparation with universal tails was described in detail previously (56). Here, assays were combined into three assay pools for multiplex PCR (see Table S3 in the supplemental material), requiring three initial PCRs for each sample. The initial gene-specific PCR mixture comprised 12.5 μl of Kapa Multiplex PCR Mastermix (Kapa Biosystems, Wilmington, MA), 10 μl of primer mix (final concentration of 200 nM each), and 2.5 μl of template DNA from each sample and was denatured at 95°C for 3 min; cycled 25 times at 95°C for 15 s, 60°C for 30 s, and 72°C for 1 min 30 s and subjected to a final extension at 72°C for 1 min. The three multiplex PCR products from the same sample were combined, and PCR products were cleaned with 1× Agencourt AMPure XP beads (Beckman Coulter, Indianapolis, IN). A second PCR with the universal tail-specific primers added Illumina’s sample-specific index and sequencing adapters. This PCR mixture comprised 12.5 μl of 2× Kapa HiFi HotStart Ready Mix (Kapa Biosystems), 400 nM each primer, and 1 to 10 μl of cleaned gene-specific PCR product and was denatured at 98°C for 2 min; cycled 6 to 12 times at 98°C for 30 s, 65°C for 20 s, and 72°C for 30 s; and subjected to a final extension at 72°C for 30 s. Final PCR products were cleaned with 0.8× Agencourt AMPure XP beads (Beckman Coulter). Amplicon libraries from individual samples were quantified by qPCR with the Kapa Library Quantification kit (Kapa Biosystems). Samples were then pooled in equimolar concentrations for sequencing on the Illumina MiSeq platform with the 2×250bp version 2 kit.

**Analysis.** Amplicon sequencing results were automatically analyzed with a newly developed amplicon sequencing analysis pipeline (ASAP) (45) that uses a JavaScript Object Notation (JSON) file customized to describe all of the assays in a multiplex. The information in the JSON file includes (i) a category for each assay (presence/absence, SNP, gene variant, or region of interest) that dictates how ASAP will report results and (ii) reference sequences for read mapping. In ASAP, amplicon sequence reads are first trimmed of adapter or readthrough sequences with Trimomatic (57) and then mapped to the reference sequences with an aligner of choice. BAM alignment files are analyzed alongside the JSON file assay descriptions to determine the presence, percent identity, and breadth and depth of coverage of the reference and proportions of nucleotide polymorphisms for each amplicon. User-defined parameters for KlebSeq-prepared samples included the bowtie2 aligner (58) for all of the assays except for wzi, for which bwa (59) was chosen (because the reference sequence is shorter than the expected amplicon and reads need to be clipped to align [55]), and thresholds for determining results of screening included percent identities listed in Table S3 in the supplemental material, 80% breadth at 100× depth of coverage for isolate DNA, 80% at 20× (clinical specimens) or 10× (fecal specimens) depth, and a ≥10% proportion of polymorphism for informative SNP loci for complex-specimen DNA (meaning that at least 10% of the reads had to share a SNP state at a given locus for it to be reported). For WGS-prepared data, the parameters were bwa aligner (for clipping) and 80% breadth at 5× depth. The ASAP output includes an XML file containing details of the analysis of each assay target for each sample, which can be converted into a webpage interface by XSLT transformations. An example of a KlebSeq ASAP output for one sample is shown in Fig. 2. SeqMan NGEN (DNAStar, Madison, WI) and Tablet (60) were used to verify results.
KlebSeq ASAP Results for Sample: F4-Mom-TP2

| Alignment statistics | Analysis parameters |
|-----------------------|----------------------|
| Total reads: 132184   | Depth filter: 10%    |
| Mapped reads: 70104   | Breadth filter: 80% |
| Unmapped reads: 62080 | Proportion filter: 10% |
| Aligner used: bowtie2 |                      |

Summary for Sample: F4-Mom-TP2

| Klebsiella species and strain types detected | Antimicrobial resistance determinants detected | Other phenotype determinants detected |
|---------------------------------------------|---------------------------------------------|-------------------------------------|
| **Klebsiella pneumoniae**                   | nmpA (aminoglycoside resistance)            | Yersiniabactin siderophore          |
|                                            | sul2 (TMP-SMX resistance)                  |                                     |
|                                            | aph3-I (aminoglycoside resistance)         |                                     |
|                                            | strA (streptomycin resistance)             |                                     |
|                                            | strB (streptomycin resistance)             |                                     |
|                                            | fosA (fosfomycin resistance)               |                                     |

Species and strain identification assays for sample: F4-Mom-TP2

| Assay Name | Average Read Depth | Coverage Breadth | Significance | SNPs found |
|------------|--------------------|------------------|--------------|------------|
| Kp_M1_UT1  | 814.01             | 100%             | Klebsiella pneumoniae present | details... |
| Kp_M2_UT1  | 925.17             | 100%             | Klebsiella pneumoniae present | details... |
| Kcani_UT   | 571.19             | 98.95%           | Not Klebsiella variicola      | details... |

Antimicrobial resistance assays for sample: F4-Mom-TP2

| Assay Name | Average Read Depth | Coverage Breadth | Significance | SNPs found |
|------------|--------------------|------------------|--------------|------------|
| npmA_UT    | 71.98              | 100%             | Resistant to all aminoglycosides | details... |
| sul2_UT1   | 135.43             | 100%             | Resistant to trimethoprim-sulfamethoxazole | details... |
| aph3-I_UT1 | 2046.08            | 100%             | Resistant to kanamycin, neomycin | details... |

FIG 2 Partial sample output of KlebSeq ASAP report; some AMR gene assays and the virulence gene assays are hidden from view to limit the size of the image. The top box shows a summary of what was detected according to selected ASAP filters. Details of each assay appear below that. If additional SNPs are detected in comparison to the assay reference, hovering over “details...” expands a list of the SNPs. Clicking on an assay name pops up a graph of coverage depth across the reference sequence.

**KlebSeq validation.** Figure 3 and the following text outline the processes used to validate KlebSeq, and Table S3 in the supplemental material shows the extent to which each assay was validated in multiplex. First, WGS data from 73 K. pneumoniae samples were analyzed for AMR genes, subjected to MLST via SRST2 (61) and species identification confirmation via phylogenetic analysis, and also analyzed by ASAP. gDNA from these same 73 samples plus gDNA from 149 other species was screened with KlebSeq. To validate KlebSeq’s K. pneumoniae strain identification and AMR gene profiles in specimens, six isolates that had been cultured and identified in six of the specimens were sequenced and analyzed. Additionally, a PCR for MLST was performed with selected specimen DNAs by the protocol from the Klebsiella BigSdb. DNA libraries from the PCR products were prepared for sequencing by the same protocol as for whole gDNA. Sequence data were run through SRST2 to determine the ST of the K. pneumoniae present in the specimen.

**RESULTS**

**Phylogenetic analysis and canSNP identification.** With the Klebsiella scgMLST (10) assembly as a reference, SNPs among a diverse set of genomes from K. pneumoniae and genomes from newly defined K. quasipneumoniae (22 from the public databases and 1 from in-house isolates) and K. variicola (24 from the public databases and 5 from in-house isolates) were identified with NASP. canSNPs that differentiate K. quasipneumoniae and K. variicola from K. pneumoniae were selected for assay development.

With the reference genome MGH 78578 and 547 diverse K. pneumoniae genomes, NASP generated a SNP matrix from which canSNPs for each of the major clonal groups were selected for assay development. Clonal groups and locations of canSNPs, and AMR and virulence genes and subsequent assay design resulted in two assays specific to K. pneumoniae (Kp-M1 and Kp-M2), one each for K. oxytoca (Koxy_UT), K. variicola (Kvari_UT), and K. quasipneumoniae (Kquasi_UT), 37 assays to identify clonal groups or lineages within clonal groups of K. pneumoniae, 76 AMR gene assays, and 15 virulence gene assays (see Table S3 in the supplemental material). The canSNP states for each strain identification assay are specific to that clonal group of K. pneumoniae, except in the case of CG35, where the amplicon must match the reference sequence 98%, allowing up to four additional SNPs, in order to be called CG35. Otherwise, identity thresholds for each strain identification assay are optional; they
merely make the assays completely Klebsiella specific, regardless of the canSNP state.

**KlebSeq validation on isolate DNA.** To validate the species and clonal group identification assays, gDNA from 73 *K. pneumoniae* isolates whose whole genomes were sequenced (4 of which were later identified as *K. quasipneumoniae* and *K. variicola* [see below]), 22 *K. oxytoca* isolates, and 157 other enteric opportunistic pathogen isolates, which included *E. coli*, *Enterobacter aerogenes*, *E. amnigenus*, *E. cloacae*, *E. hormaechei*, *Enterococcus faecalis*, *E. faecium*, *Providencia stuartii*, and *Serratia marcescens*, and 1 *Achromobacter baumannii* isolate, were screened with KlebSeq. Sensitivity and specificity results of the species identification assays compared with clinical microbiological identification (Vitek 2) are in Table 1. With the redundancy built into the multiplex by including two assays, Kp-M1 and Kp-M2, that target two different *K. pneumoniae* species-specific genes (M1 and M2), 100% sensitivity is achieved. One isolate previously identified as *K. pneumoniae* was identified as *K. quasipneumoniae*, and two were identified as *K. variicola*. These isolates’ whole genomes were added to the phylogenetic analysis of these three species that was previously run to find the species-specific canSNPs (see Materials and Methods). The *K. quasipneumoniae* and *K. variicola* genomes identified by our assay clustered with their respective species in the phylogeny (Fig. 5). Clinical methods do not currently distinguish among all three of these species, so assay sensitivity and specificity were not calculated for *K. quasipneumoniae* and *K. variicola* (Table 1).

Table 2 shows the KlebSeq results of the *K. pneumoniae* clonal group identification and capsule typing assays of isolate DNA. Each isolate’s strain type was correctly captured by the appropriate assays or not captured in cases where no assay was designed for that clonal group. Included in Table 2 are results from partial sequencing of the *wzi* gene for capsule typing. This gave surprisingly clear results, given that approximately 75% of the informative region are missing from our sequence output, as the PCR amplicon is approximately 580 bp (55), which is too long to cover with the Illumina version 2 sequencing chemistry. However, full capsule typing by *wzi* sequencing would be possible with longer-read chemistry (i.e., Illumina version 3 chemistry, for 600-bp reads). Results from screening of nontarget organisms showed that several of the *K. pneumoniae* clonal group assays amplified DNA from other organisms, as expected. An identity threshold can be applied (see Table S3 in the supplemental material); however, all of the SNP states that define a particular clonal group are specific to that clonal group, except for CG35, so the identity threshold is optional except for this assay. Sequence analysis by ASAP reports when a clonal group is present only if the defining canSNP state is present and reports nothing if it is not.

AMR gene detection by amplicon sequencing was validated by comparing ASAP results with AMR gene screening of WGS with SRST2 (61) and with ASAP. Results showed an almost perfect correlation between KlebSeq ASAP and WGS ASAP, indicating that the KlebSeq PCRs are performing well. There were a few discrepancies with SRST2, which reported uncertainty (SNPs or low-coverage indicators) for most of the discrepancies. Some discrepancies were in the presence of the *dfrA* gene. This group of genes is very diverse, so it may be that KlebSeq does not capture the full repertoire of *dfrA* genes. Virulence gene detection was validated by comparing ASAP results from WGS data with those from amplicon sequence data, and results showed concordance. In addition, by targeting multiple genes that are part of the same virulence factors (i.e., siderophore systems), sensitivity and confidence in results were increased.

These results also confirm that KlebSeq is applicable to pure isolates as well as complex specimens. Screening of isolate DNA has the added benefit of traceability of the AMR and virulence genes, which are often carried on mobile genetic elements, to their host. Isolate screening could be used for surveillance and other purposes for identifying or characterizing Klebsiella.

**Specimen sample results.** KlebSeq was run on DNA from 87 respiratory specimens, 46 urine specimens, 40 wound specimens, and 89 fecal samples from healthy individuals (see Table S4 in the supplemental material). Sensitivity and specificity results of the species identification assays compared with those of clinical mi-
Microbiological methods are shown in Table 1. In most cases, sensitivity was high, except in the wound specimens, where the one sample clinically identified as *K. pneumoniae* was identified as *K. variicola*. Some sensitivity and specificity calculations may be misleadingly low, as amplicon sequencing identified some samples as containing *K. pneumoniae* that actually contained *K. quasipneumoniae* or *K. variicola*, and several as containing *Klebsiella* that went undetected by clinical microbiological methods, including several in which clonal groups were also detected. In the healthy donor specimens, *K. oxytoca* (n/H11005:13) was more prevalent than *K. pneumoniae* (n/H11005:9). Sequencing read depth was low in some samples (see Table S4 in the supplemental material); this may have been due to dilution of the DNA samples before screening, as each was diluted 1:10 in water.

Important *Klebsiella* clonal groups were detected in multiple specimens (see Table S4 in the supplemental material). In the 17 urine samples positive for *K. pneumoniae*, clonal identifications included CG34, ST20, CG45, CG392 (which includes the NDM producer ST147 [62]), though this sample was negative for *bla*NDM, ST133, and CG111. In wounds, the only sample *K. pneumoniae* positive by KlebSeq was CG29. From respiratory specimens, groups CG37 (n/H11005:2), ST134 (n/H11005:1), ST258 (n/H11005:2), CG36 (n/H11005:3), and inner ST14 (n/H11005:1) were identified. Interestingly, several clonal groups were identified in the healthy donor fecal specimens as well. In the nine *K. pneumoniae*-positive samples, the groups included ST20, CG37, and CG76, which are all members of multidrug-resistant outbreak strain types (11, 12, 15), along with ST133 and ST380. ST380 is associated with a K2 capsule type and hypervirulence and causes pyogenic liver abscesses in healthy people, especially those of Asian ethnicity (63). Many Asians are colonized by hypervirulent K1 or K2 capsule strain types; however, the level of risk of subsequent liver infection is unknown (63). For

---

**FIG 4** Maximum-parsimony tree with 100 bootstraps of the SNPs among 548 *K. pneumoniae* genomes. Major clonal groups are colored, and locations of canSNPs for strain identification assays are marked with stars. All of the branches labeled with canSNPs had >99% bootstrap support, except for the three branches indicated.
TABLE 1 Results of KlebSeq species identification assays of genomic DNA from isolates whose whole genomes were also sequenced, DNA from specimens for which clinical culture results are known, and DNA from specimens with unknown content

| DNA type (no. of samples) and species identified by clinical methods or parameter<sup>a</sup> | Total no. screened | No. of isolates identified by amplicon sequencing assay<sup>b</sup> | Kp-M1 | Kp-M2 | Kp-M1 + Kp-M2 | Kquasi_UT | Kyari_UT | Koxy_UT |
|---|---|---|---|---|---|---|---|---|
| Isolate DNA (252) | | | | | | | | |
| K. pneumoniae | 69 | 68 | 67 | 69 | 0 | 0 | 0 | 0 |
| K. quasipneumoniae | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| K. variicola | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 0 |
| K. oxytoca | 14 | 0 | 0 | 0 | 0 | 0 | 14 | 0 |
| Nontarget species | 149 | 2/88 | 0/88 | 2/88 | 0/155 | 0/155 | 0/135 | 0/135 |
| % Sensitivity | 99 | 97 | 100 | 100 | 100 | 100 | 100 | 100 |
| % Specificity | 98 | 100 | 98 | 100 | 100 | 100 | 100 | 100 |
| Urine DNA (46) | | | | | | | | |
| K. pneumoniae | 16 | 14 | 15 | 16 | 2 (1 mix)<sup>c</sup> | 1 (mix)<sup>c</sup> | 0 | 0 |
| K. oxytoca | 6 | 1 | 1 | 1 | (CG34) | 0 | 0 | 6 |
| Other species | 24 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Unknown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| % Sensitivity | 88 | 94 | 100 | 100 | 100 | 100 | 100 | 100 |
| % Specificity | 90 | 93 | 93 | 100 | 100 | 100 | 100 | 100 |
| Wound DNA (40) | | | | | | | | |
| K. pneumoniae | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| K. oxytoca | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Other species | 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Unknown | 7 | 1 | 1 | 1 | (CG29) | 0 | 0 | 1 |
| % Sensitivity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| % Specificity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Respiratory specimen DNA (87) | | | | | | | | |
| K. pneumoniae | 6 | 6 | 6 | 6 | 0 | 0 | 0 | 0 |
| K. oxytoca | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| Other species | 77 | 7 (1 ST258) | 5 | 7 (2 CG36, 1 CG37) | 0 | 0 | 1 | 1 |
| Unknown | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| % Sensitivity | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| % Specificity | 91 | 94 | 91 | 99 | 99 | 99 | 99 | 99 |
| Fecal specimen DNA (89) | | | | | | | | |
| K. pneumoniae | 89 | 9 | 5 | 9 | 1 (mix)<sup>c</sup> | 3 (2 mix)<sup>c</sup> | 13 | 0 |
| All specimens (isolates not included) (262) | | | | | | | | |
| K. pneumoniae | 23 | 20 | 21 | 22 | 2 | 2 | 0 | 0 |
| K. oxytoca | 8 | 1 | 1 | 1 | 0 | 0 | 8 | 0 |
| Other species | 132 | 8 | 6 | 8 | 0 | 0 | 1 | 1 |
| Unknown | 99 | 10 | 6 | 10 | 1 | 3 | 14 | 0 |
| % Sensitivity | 87 | 91 | 96 | 100 | 100 | 100 | 100 | 100 |
| % Specificity | 94 | 95 | 94 | 99 | 99 | 99 | 99 | 99 |

<sup>a</sup> K. quasipneumoniae is not distinguished from K. pneumoniae by the clinical identification method used (Vitek 2).

<sup>b</sup> Kp-M1 and Kp-M2 are K. pneumoniae species identification assays that detect targets M1 and M2 in the K. pneumoniae genome. Kquasi_UT, Kyari_UT, and Koxy_UT are K. quasipneumoniae-, K. variicola-, and K. oxytoca-specific assays, respectively.

<sup>c</sup> These species were found as mixtures with K. pneumoniae on the basis of a proportion (≥10%) of the sequencing reads containing the species-defining SNP.
this sample, no wzi gene sequence was obtained; thus, the capsule type is unknown. A majority of the *K. pneumoniae* isolates in our samples did not fall into the major clonal groups targeted by KlebSeq. These strains probably all belong to lesser-known clonal groups, as more studies are showing that many *K. pneumoniae* infections are caused by nonepidemic, sporadic strains (64, 65).

Numerous and variable AMR genes were detected in the specimens, including different variants of the same gene that confer different phenotypes (see Table S4 in the supplemental material). With sequence-based information, we demonstrate that seven of the *K. pneumoniae* had key mutations in the *gyrA* gene known to confer resistance to fluoroquinolones. Additionally, several samples contained the *aac(6')-Ib* gene for aminoglycoside resistance, and five of those contained the sequence variant *aac(6')-Ib-cr* for fluoroquinolone resistance; mixtures of these two genes were also detected. Many of the infection specimens (nonhealthy donor specimens), both positive and negative for *K. pneumoniae*, were positive for other aminoglycoside resistance genes, as well as tetracycline, trimethoprim-sulfamethoxazole, streptomycin, fosfomycin, and chloramphenicol resistance genes. A few contained plasmid-mediated quinolone resistance genes. Several samples, especially the respiratory specimens, were also positive for KPC and CTX-M group 1 and 9 genes. Most of the healthy donor specimens were positive for trimethoprim-sulfamethoxazole resistance genes, and many were positive for streptomycin, aminoglycoside, tetracycline, and fosfomycin resistance genes. Interestingly, 39 of the 89 were positive for * npmA*, a relatively recently described pan-aminoglycoside resistance gene (66). These specimens were the only samples positive for this gene. Three specimens contained plasmid-mediated quinolone resistance genes. Fortunately, none were found to contain ESBL or carbapenemase genes. No complex specimens in this study were positive for genes encoding the important carbapenemases OXA-48, VIM, and NDM, and none were positive for the plasmid-mediated colistin resistance gene *mcr-1* (see Table S4 in the supplemental material).

These sets of samples did not appear to contain especially virulent strains of *K. pneumoniae*. The yersiniabactin siderophore genes were, by far, the most prevalent of the virulence genes tested, although positive samples made up less than half of the *K. pneumoniae*-positive samples. No specimens were positive for *rmpA*, the regulator of mucoid phenotype gene, including the ST380-containing sample, and few were positive for the salmochelin siderophore genes, which are associated with invasive *K. pneumoniae*. 

FIG 5 Neighbor-joining tree with 100 bootstraps of the SNPs in the diverse set of *K. pneumoniae*, *K. variicola*, and *K. quasipneumoniae* genomes used in this study. Unknown isolates that were identified as *K. variicola* and *K. quasipneumoniae* are boxed.
infection (54). One respiratory specimen that contained an ST14 strain was positive for a K2 capsule type by partial \textit{wzi} sequencing. K2 strains of \textit{K. pneumoniae} are associated with hypermucoviscosity and hypervirulence, as previously mentioned. However, this respiratory sample was not positive for \textit{rmpA}, and a recent study proposed that the presence of multiple siderophore system genes (linked to K1 or K2 capsule genes) explains hypervirulence rather than capsule type (54). In our data, \textit{K. pneumoniae}-containing samples were positive for multiple siderophores or other virulence-associated genes only 15% of the time. Sequencing of \textit{wzi} revealed a variety of capsule types and cases in which the same clonal groups had different \textit{wzi} genotypes and in which they had the same genotype. This character would help identify or rule out a transmission event when patients carrying the same strain are found.

On an interesting note, in the healthy donor fecal samples collected from members of the same families over time, out of the nine \textit{K. pneumoniae}-positive samples, only two came from the

![Table 2: Isolates used for assay validation and results of strain typing by amplicon sequencing](image)
same person over time. The characterization assays suggest that the same strain of *K. pneumoniae* was present at both time points. *K. pneumoniae*-positive samples were found in multiple members of two of the seven families. In one of these families, the positive members carried strains different from one another, and in the other, it appears that two members had CG37 isolates with the same capsule type. The sample set is too small to draw conclusions from these data; however, the data raise interesting questions about community *K. pneumoniae* carriage.

**Validation of KlebSeq strain identification in specimens.** Table 3 shows MLST results from WGS data of six isolates cultured from specimens run on KlebSeq and from MLST PCR and sequencing of 11 specimens run on KlebSeq. In each case, KlebSeq appears to have identified the correct strain. Two isolates for which no strain type was identified by KlebSeq have novel STs. Sample TG75900 was identified as *K. quasipneumoniae* by KlebSeq and typed as ST196 on the basis of whole-genome data. This genome was added to the phylogeny of the three species and clustered with *K. quasipneumoniae* (Fig. 5). MLST of the specimen DNA did not yield results for all of the MLST loci of all 11 samples, which is to be expected given the complexity of the specimen DNA sample. In cases where only partial data were retrieved, at least three alleles from each match the strain identified by KlebSeq.

**DISCUSSION**

In the United States, HAI is estimated to affect 1 in 25 hospital patients, totaling hundreds of thousands of patients, with significant mortality (2). HAIs have a significant impact on health care costs; a 2009 CDC report estimated upwards of $45 billion in annual additional cost (67). Infections with AMR organisms cause significantly higher mortality rates, significantly more intensive care unit (ICU) admissions, and significant excess costs, including hospitalization, medical care, and antimicrobial therapy, than do infections with susceptible strains (5, 68). HAI prevention measures, although costly in and of themselves (69), have the potential to save thousands of lives and billions of dollars (67). Periodic patient screening and isolation of AMR organism carriers have proven successful in controlling transmission and outbreaks in several hospitals (31, 34–37). Use of a highly informative screening and surveillance tool such as KlebSeq has cost-effective and life-saving potential.

Early detection of colonization of health care patients by *K. pneumoniae*, especially multidrug-resistant *K. pneumoniae*, would allow health care staff to make more informed patient management decisions. In outbreak situations, rapid identification of transmissions before subsequent infections would allow for proactive measures to curb an outbreak. In nonoutbreak situations, identification of particular strains and AMR genes would help to assess the risk of *K. pneumoniae* carriage to the host patient, as well as to other patients, as some strains are more closely associated with adverse outcomes (e.g., outbreaks, HAI, AMR, and treatment failure) than others (7–13, 70). Although our understanding of many Klebsiella virulence factors is limited, identification of virulence genes gains us understanding of the correlations between particular virulence factors and the risk of disease (54). Additionally, many *K. pneumoniae* infections, including HAIs and non-multidrug-resistant infections, are caused by nonepidemic, lesser-known strain types (64, 65). Classifying the *K. pneumoniae* isolate in each patient sample would help an institution to decide when and which intervention procedures should be enacted and also to understand more about transmission dynamics and local strain type circulation.

KlebSeq has several characteristics that make it attractive as a health care screening approach. With a single assay, enough infor-

---

**Table 3** Results of KlebSeq strain identification validation by MLST of isolates cultured from specimens tested by KlebSeq and MLST of specimens tested by KlebSeq

| Sample   | Type          | KlebSeq identification of original specimen | No. of loci retrieved from sequence data | ST by MLST     |
|----------|---------------|--------------------------------------------|-----------------------------------------|----------------|
| TG69923  | Isolate       | CG29                                       | 7                                       | Novel; DLV<sup>a</sup> of ST29  |
| TG75899  | Isolate       | CG392                                      | 7                                       | ST392          |
| TG75900  | Isolate       | *K. quasipneumoniae*                       | 7                                       | ST196          |
| TG75901  | Isolate       | ST133                                      | 7                                       | Novel; SLV<sup>b</sup> of ST133 |
| TG75902  | Isolate       | No strain ID                               | 7                                       | Novel; DLV of ST248 |
| TG75911  | Isolate       | Mixture of *K. pneumoniae* with no strain ID and *K. variicola* | 7 | Novel; TLV<sup>c</sup> of ST633 |
| TG69737  | Urine         | CG34                                       | 7                                       | Novel; 4 alleles match ST34 |
| TG69766  | Urine         | CG45                                       | 6                                       | 6 alleles match ST45  |
| TG69776  | Urine         | CG111                                      | 7                                       | Novel; DLV of ST111 |
| TG69861  | Respiratory   | No strain ID                               | 7                                       | Novel; SLV of ST393 |
| TG69865  | Respiratory   | ST134                                      | 6                                       | 5 alleles match ST134 |
| TG69871  | Respiratory   | CG37                                       | 3                                       | 3 alleles match ST37 |
| TG69883  | Respiratory   | ST258                                      | 7                                       | Novel; 3 alleles match ST258 |
| TG73865  | Respiratory   | CG36                                       | 7                                       | Novel; 3 alleles match ST36 |
| TG73911  | Respiratory   | Mixture of *K. pneumoniae* with no strain ID and CG36 | 6 | 4 alleles match ST36 |
| TG73916  | Respiratory   | Inner ST14                                 | 7                                       | Novel; 4 alleles match ST14 |
| TG74003  | Respiratory   | Mixture of *K. pneumoniae* no strain ID and CG36 | 7 | Novel; DLV of ST461 |

<sup>a</sup> DLV, double-locus variant.

<sup>b</sup> SLV, single-locus variant.

<sup>c</sup> TLV, triple-locus variant.
mation is garnered about a patient’s Klebsiella carriage status to contribute greatly to patient management and to infection control decisions. Indexing samples by means of the universal tail during sample preparation allows the characterization of a large number of specimens in one run, minimizing sequencing costs per specimen and allowing for high-throughput screening of hundreds of patient samples simultaneously at a cost of tens of dollars per patient. KlebSeq uses DNA extracted directly from a specimen, so targets from entire populations of a species are analyzed to capture different strains in the same sample, which can be numerous (71, 72). If culture-based methods are used for screening, different strains are missed when one genotype (i.e., colony) is chosen for characterization, and the resulting information is limited. Additionally, culture-based methods can miss “silent” multidrug-resistant K. pneumoniae strains that test negative for carbapenemases in vitro (16), and if used for high-throughput screening, they can be laborious, time-consuming, costly, and subjective (31, 38). If screening of large numbers of patients by amplicon sequencing is cost-prohibitive, it can be limited to the highest-risk groups of patients, i.e., long-term care facility patients (31, 73), travelers returning from regions where Klebsiella carriage is endemic (74, 75), ICU patients (28), patients with previous K. pneumoniae carriage (75–77), patients who have shared a room with a known carrier (78) or case contacts of carriers (79), those who have recently taken antibiotics (80, 81), or patients on mechanical ventilation or enteral feeds or who have had prior Clostridium difficile infections (82). Additionally, using ASAP makes the analysis in KlebSeq streamlined and results are easily interpretable. Lastly, the amplicon sequencing and ASAP package is customizable and updateable (45). Individual assays can be added or removed, adding only the cost of new primers.

The results we present here show that KlebSeq is effective with DNA from numerous sample types, including pure organism culture, or complex, multiorganism samples and swab samples with low-level microbial DNA in a presumably high human DNA background without culture methods. In addition to clinically important clonal lineages of K. pneumoniae, KlebSeq also reliably distinguishes among the Klebsiella species, two of which, K. quasipneumoniae and K. variicola, are continuously misidentified as K. pneumoniae (as exemplified in our study) and cause invasive disease (83, 84). Additionally, we highlight several instances where culture methods failed to produce a positive K. pneumoniae result, including one sample that contained the critical ST258 strain. The sensitivity of KlebSeq is superior to that of culture-based methods for complex specimens, lowering the risk of false negatives in patient screening. We identified dozens of AMR and virulence genes within individual samples, demonstrating the additional function of profiling for clinically important characteristics, and were able to distinguish minor genotype differences that confer different phenotypes, such as the gyrA gene, aac(6’)-Ib versus aac(6’)-Ib-cr, and the wzi gene.

Notably, our data show that healthy individuals may carry clinically important strains of K. pneumoniae and frequently K. oxytoca, as well as many AMR genes and siderophore virulence systems. For our purposes, these healthy donor fecal DNA samples were used to validate the use of our amplicon sequencing approach with highly complex fecal metagenome samples. Much more study is needed to elucidate the implications of healthy host carriage of known pathogenic strains of K. pneumoniae and their virulence factors. Furthermore, the fact that we observed carriage of the hypervirulence-associated ST380 strain from a healthy person and the hypervirulence-associated K2 capsule type in an ST14 strain from a respiratory infection lends credence to the idea that we need much more information about K. pneumoniae carriage strains to be able to draw conclusions about these associations. Aside from these important observations, other observations from these data raise questions about the dynamics of K. pneumoniae carriage and microbiome sharing. The fact that there were not more cases of positive results from the same person at multiple rather than single time points is interesting. This could be due to intermittent shedding of K. pneumoniae in feces, intermittent colonization by K. pneumoniae, or heterogeneity in the sample itself, underrepresenting the full microbial community when a small sample is taken; this observation further warrants periodic rather than one-time screening of patients at risk. Of the two instances where two family members carried K. pneumoniae, each tells a different story about microbiome sharing. An amplicon sequencing-based diagnostic approach would facilitate longitudinal patient screening because of ease of use and limited costs.

Among the many pathogens encountered in health care institutions, we focus only on Klebsiella because of its high priority in public health and the high risk of CPE establishment in a facility. Additionally, although many other HAI agents cause devastating and costly infections, directing a complex assay such as KlebSeq at a subset of those agents greatly simplifies the validation process and speeds the availability of assay results (especially in the case of validation for FDA approval [85]). KlebSeq is an important step toward a comprehensive yet accessible tool for all pathogen identification and characterization. Metagenomic analysis is attractive in its breadth of coverage capabilities, but its current costs and complexities are prohibitive. Amplicon assays targeting the other ESKAPE (Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter species) pathogens and more AMR determinants could relatively easily be added to the tool described here, with a step-by-step validation process. We have not found the limit of the multiplexing capacity. (We run three multiplexed PCRs merely because the PCR volume limits the number of primers that can be added; future assays will utilize primers at a higher concentration to test the multiplexing limit.) The multiplexing limit is also presumably dependent on the number and sizes of targets that are present in a sample, potentially exhausting the polymerase or sequencing space. These pieces are unpredictable for specimens of unknown content. Lastly, our report serves as evidence in favor of the concept that highly multiplexed amplicon sequencing is one good answer to the call for early detection tools in infection control.

Overall, the KlebSeq method was able to accurately and consistently identify and characterize Klebsiella from complex specimens. A limitation of our study is that clonal group identification in the complex specimens was not confirmed by isolation and WGS of the Klebsiella isolate from the specimen. Additionally, profiling of complex specimens directly for AMR and virulence genes, most of which are on mobile elements, can be confounding, as it cannot be known which organism carries the genes of interest. Third, our validation efforts focused on the detection of targets but did not characterize the limits of that detection, and we saw evidence that KlebSeq may not find all of the variants of all of the AMR genes. However, KlebSeq is designed for screening and periodic surveillance in high-risk situations with a rule-in/rule-out determination of the possibility of transmission events and
through identification of high-risk multidrug-resistant or epidemic strains of \textit{Klebsiella}. For these purposes, KlebSeq is ideal. The specimen types used for validation could be considered a limitation, as we did not test rectal swabs, a specimen type commonly used for CPE surveillance. However, we show that KlebSeq works with different swab types and fecal specimens, which addresses the challenges of detection in rectal swabs. The turnaround time from sample collection to result is dependent only on current technology (not organism culture), and we recently conducted a proof-of-concept study of a 24-h sample-to-answer test with different targets (data not shown). This test was done on an Illumina MiSeq with only 60 cycles. Other platforms and upcoming technology may allow this turnaround time to be decreased even further.

Rapid amplicon sequencing with automated analysis and reporting is a promising response to the need for constant surveillance for highly transmissible or highly drug-resistant pathogens. Our model system, directed at \textit{Klebsiella}, can easily be adapted to multiple other pathogens and to different purposes, such as environmental sampling and community host screening and, as smaller, more on-demand next-generation systems become available, to diagnostics and individual patient monitoring. The targeted, highly multiplexed nature of amplicon sequencing and the ability to interpret the data instantly make it an applicable tool for health care facility surveillance. As these technologies are adopted, considerable coordination within the health care facility is paramount to the success of infection and outbreak prevention, with the integration of isolation and barrier precautions, excellent communication, and good stewardship. Nevertheless, several institutions have shown that the combination of surveillance and systematic response reduces outbreaks and multidrug-resistant infections (31, 33–37).

**ACKNOWLEDGMENTS**

The Division of Healthcare Quality Promotion at the Centers for Disease Control and Prevention kindly contributed \textit{K. pneumoniae} DNA used in this study. Many thanks to Kristyal Sheridan, Mike Valentine, and Veronica Harrison for technical assistance and Tricia O’Reilly for operational support.

**FUNDING INFORMATION**

This work, including the efforts of Jolene R. Bowers, Darrin Lemmer, Jason W. Sahl, Talima Pearson, Elizabeth Driebe, David M. Engeltalher, and Paul Stephen Keim, was funded by HHS | National Institutes of Health (NIH) (R01AI090782).

**REFERENCES**

1. Montgomery JZ. 1979. Epidemiology of Klebsiella and hospital-associated infections. Rev Infect Dis. 1:736–753. http://dx.doi.org/10.1093/clinids/1.5.736.
2. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, Lyncfield R, Maloney M, McAlistier-Hollod L, Nadle J, Ray SM, Thompson DL, Wilson LE, Fridkin SK. 2014. Emerging Infections Program Healthcare-Associated Infections, Antimicrobial Use Prevalence Survey Team. 2014. Multistate point-prevalence survey of health care-associated infections. N Engl J Med. 370:1198–1208. http://dx.doi.org/10.1056/NEJMoA1308801.
3. Centers for Disease Control and Prevention. 2013. Antibiotic resistance threats in the United States, 2013. Centers for Disease Control and Prevention, Atlanta, GA. http://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf.
4. Yu Y, Gu B, Huang M, Liu H, Xu T, Xia W, Wang T. 2015. Epidemiology of carbapenem-resistant Enterobacteriaceae (CRE) during 2000–2012 in Asia. J Thorac Dis. 7:376–385. http://dx.doi.org/10.3978/j.issn.2072-1439.2014.12.33.
5. World Health Organization. 2014. Antimicrobial resistance global report on surveillance. World Health Organization, Geneva, Switzerland. http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf.
6. Ossen L, Martino MD, Siqueira I, Pasternak J, Gales AG, Silva CV, Camargo TZ, Scherer PF, Marra AR. 2013. A hospital-based matched case-control study to identify clinical outcome and risk factors associated with carbapenem-resistant Klebsiella pneumoniae infection. BMC Infect Dis 13:80. http://dx.doi.org/10.1186/1471-2334-13-80.
7. Kitchel B, Sundin DR, Patel JB. 2009. Regional dissemination of KPC-producing \textit{Klebsiella pneumoniae}. Antimicrob Agents Chemother 53: 4511–4513. http://dx.doi.org/10.1128/AAC.00784-09.
8. Frasson I, Lavezzo E, Franchin E, Toppo S, Barzon L, Cavallaro A, Richter SN, Palu G. 2012. Antimicrobial treatment and containment measures for an extremely drug-resistant \textit{Klebsiella pneumoniae} ST101 isolate carrying blKP1101-IT, a novel fully sequenced bla(KPC-2) plasmid. J Clin Microbiol 50:3768–3772. http://dx.doi.org/10.1128/JCM .01892-12.
9. Lascols C, Peirano G, Hackel M, Laupland KB, Pitout JD. 2013. Surveillance and molecular epidemiology of \textit{Klebsiella pneumoniae} isolates that produce carbapenemases: first report of OXA-48-like enzymes in North America. Antimicrob Agents Chemother 57:130–136. http://dx.doi.org/10.1128/AAC.01686-12.
10. Bialek-Davenet S, Criscuolo A, Allioud F, Passet V, Jones L, Delannoy-Vieillard AS, Garin B, Le Hullo S, Arlet G, Nicolas-Chanoine MH, Dautry-Varsacq F, Brisse S. 2014. Genomic definition of hypervirulent and multidrug-resistant \textit{Klebsiella pneumoniae} clonal groups. Emerg Infect Dis 20:1812–1820. http://dx.doi.org/10.3201/eid2011.140206.
11. Nagasaka Y, Kimura K, Yamada K, Wachino J, Iin W, Notake S, Yanagisawa H, Arakawa Y. 2015. Genetic profiles of fluoroquinolone-nonsusceptible \textit{Klebsiella pneumoniae} among cephalosporin-resistant \textit{K. pneumoniae}. Microb Drug Resist 21:224–233. http://dx.doi.org/10.1089/mdr.2014.0180.
12. Liu Y, Wan LG, Deng Q, Cao XW, Yu Y, Xu QF. 2015. First description of NDM-1-, KPC-2-, VIM-2- and IMP-4–producing \textit{Klebsiella pneumoniae} strains in a single Chinese teaching hospital. Epidemiol Infect 143:376–384. http://dx.doi.org/10.1017/S0950268814000995.
13. Cubero M, Grau I, Tubau F, Pallares R, Dominguez MA, Linares J, Ardany G. 2016. Hypervirulent \textit{Klebsiella pneumoniae} clones causing bacteremia in adults in a teaching hospital in Barcelona, Spain (2007–2013). Clin Microbiol Infect 22:154–160. http://dx.doi.org/10.1016/j.cmi.2015.09.025.
14. Jacobson RK, Manenes MR, Moodley C, Smith M, Williams SG, Nicol MP, Bamford CM. 2015. Molecular characterisation and epidemiological investigation of an outbreak of blAOXA-181 carbapenemase–producing isolates of \textit{Klebsiella pneumoniae} in South Africa. S Afr Med J 105:1030–1035. http://dx.doi.org/10.7196/SAMJ.2015.v105i12.9926.
15. Zhu J, Sun L, Ding B, Yang Y, Xu X, Liu W, Zhu D, Yang F, Zhang H, Hu F. 2016. Outbreak of NDM–1–producing \textit{Klebsiella pneumoniae} ST76 and ST37 isolates in neonates. Eur J Clin Microbiol Infect Dis http://dx.doi.org/10.1007/s10096-016-2578-z.
16. Vila RA, Hujer AM, Marshall SH, Perez F, Hujer KM, Brisceno DF, Duk M, Jacobs MR, Grossberg R, Toltzis P, Bonomo RA. 2012. “Silent” dissemination of \textit{Klebsiella pneumoniae} isolates bearing \textit{K. pneumoniae} carbapenemases in a long-term care facility for children and young adults in northeast Ohio. Clin Infect Dis 54:1314–1321. http://dx.doi.org/10.1093/cid/cis036.
17. Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D, de Man T, Rasheed JK, Engelthaler DM, Keim P, Limbago BM. 2015. Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 Klebsiella pneumoniae pandemic. PLoS One 10: e0133727. http://dx.doi.org/10.1371/journal.pone.0133727.
18. Zheng R, Zhang Q, Guo Y, Feng Y, Liu L, Zhang A, Zhao Y, Yang X, Xia X. 2016. Outbreak of plasmid-mediated NDM–1–producing \textit{Klebsiella pneumoniae} ST105 among neonatal patients in Yunnan, China. Ann Microbiol Antimicrob 15:10. http://dx.doi.org/10.1186/s12941-016-0124-6.
19. Pena I, Picazo JJ, Rodriguez-Avial C, Rodriguez-Avial I. 2014. Carbapenemase-producing Enterobacteriaceae in a tertiary hospital in Madrid, Spain: high percentage of colistin resistance among VIM–1–producing \textit{Klebsiella pneumoniae} ST11 isolates. Int J Antimicrob Agents 43:660–464. http://dx.doi.org/10.1016/j.ijantimicag.2014.01.021.
20. Voulgari E, Gartzonika C, Vrioni G, Politi L, Priavali E, Levidiotou-Stefanou S, Tsakris A. 2014. The Balkan region: ND1-producing Klebsiella pneumoniae ST11 clonal strain causing outbreaks in Greece. J. Antimicrob Chemother 69:2091–2097. http://dx.doi.org/10.1093/jac/dku105.

21. Woodford N, Turton JF, Livermore DM. 2011. Multiresistant Gram-negative Hospital Enterobacteriaceae: The role of high-risk clones in the dissemination of antibiotic resistance. FEMS Microbiol Rev 35:76–75. http://dx.doi.org/10.1111/j.1574-6966.2011.02028.x.

22. Swaminathan M, Sharma S, Poliainsky Blash S, Patel G, Banach DB, Phillips M, LaBombardi V, Anderson KF, Kitchel B, Srinivasan A, Calfee DP. 2013. Prevalence and risk factors for acquisition of carbapenem-resistant Enterobacteriaceae in the setting of endoscopy. Infect Control Hosp Epidemiol 34:809–817. http://dx.doi.org/10.1086/671270.

23. Pena C, Pujol M, Ardanyun C, Ricart A, Pallares R, Linares J, Ariza J, Gudiol F. 1998. Epidemiology and successful control of a large outbreak due to Klebsiella pneumoniae producing extended-spectrum beta-lactamases. Antimicrob Agents Chemother 42:53–58.

24. Tischendorf J, de Avila RA, Safdar N. 2016. Risk of infection following colonization with carbapenem-resistant Enterobacteriaceae [sic] a systematic review. Am J Infect Control 44:539–543. http://dx.doi.org/10.1016/j.ajic.2015.12.005.

25. Amit S, Mishali H, Kotlovsky T, Schwaber MJ, Carmeli Y. 2015. Bloodstream infections among carriers of carbapenem-resistant Klebsiella pneumoniae: etiology, incidence and predictors. Clin Microbiol Infect 21:30–34. http://dx.doi.org/10.1016/j.cmi.2014.08.001.

26. Thaden JT, Lewis SS, Hazen KC, Hudlage K, Fowler VG Jr, Moehring RW, Chen LL, Lindsey CD, Moore ZS, Sexton DJ, Anderson DJ. 2014. Rising rates of carbapenem-resistant enterobacteriaceae in community hospitals: a mixed-methods review of epidemiology and microbiology practices in a network of community hospitals in the southeastern United States. Infect Control Hosp Epidemiol 35:978–983. http://dx.doi.org/10.1086/677157.

27. Basyouni RH, Gaber SN, Weggan AA. 2015. Fecal carriage of extended-spectrum beta-lactamase- and AmpC-producing Escherichia coli among healthcare workers. J Infect Dev Ctries 9:304–308. http://dx.doi.org/10.3855/jidc.5633.

28. Papadimitriou-Olivieris M, Marangos M, Fligou F, Christofidou M, Sklavou C, Vamvakopoulos S, Anastassiou ED, Filos KS. 2013. KPC-producing Klebsiella pneumoniae enteric colonization acquired during intensive care unit stay: the significance of risk factors for its development and its impact on mortality. Diagn Microbiol Infect Dis 77:169–173. http://dx.doi.org/10.1016/j.diagmicrobio.2013.06.007.

29. Monari C, Merlini L, Nardelli E, Cacioni M, Repetto A, Mencacci A, Vecchiarelli A. 2016. Carbapenem-resistant Klebsiella pneumoniae: results of a laboratory surveillance program in an Italian general hospital (August 2014-January 2015): surveillance of carbapenem-resistant Klebsiella pneumoniae. J Antimicrob Chemother 71:1371–1379. http://dx.doi.org/10.1093/jac/dku117.

30. Vooi in ‘t Holt AF, Severin JA, Goessens WH, te Witt R, Vos MC. 1998. Epidemiology and successful control of a large outbreak of carbapenem-resistant Klebsiella pneumoniae due to high-risk clones in the dissemination of antibiotic resistance. Clin Microbiol Infect 4:1111–1117. http://dx.doi.org/10.1016/S1198-743X(98)00058-0.

31._via whole genome sequencing. Am J Transplant 14:2633–2639. http://dx.doi.org/10.1111/ajt.12898.

32. Snitkin ES, Zelazny AM, Thomas PJ, Stock F, NISC Comparative Sequencing Program Group, Henderson DK, Palmore TN, Segre JA. 2012. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med 4:184ra811. http://dx.doi.org/10.1126/scitranslmed.3001529.

33. Iqbal S, Quigley EM. 2016. Progress in our understanding of the gut microbiome: implications for the clinician. Curr Gastroenterol Rep 18:49. http://dx.doi.org/10.1007/s11894-016-0524-y.

34. Cope EK, Lynch SV. 2015. Novel microbiome-based therapeutics for chronic rhinosinusitis. Curr Allergy Asthma Rep 15:504–515. http://dx.doi.org/10.1007/s11882-014-0504-y.

35. Millares L, Perez-Brocal V, Ferrari R, Gallego M, Pomeres X, Garcia-Nunez M, Monton C, Capilla S, Monso E, Moya A. 2015. Functional metagenomics of the bronchial microbiome in COPD. PLoS One 10:e0136135. http://dx.doi.org/10.1371/journal.pone.0136135.

36. Laine S. 2013. FDA allows marketing of four “next generation” gene sequencing devices. U.S. Food and Drug Administration, Silver Spring, MD. http://www.fda.gov/Newsroom/NewsroomAnnouncements/ucm375742.htm. Accessed 31 July 2016.

37. Gilbert JA, Meyer F, Janssens J, Gordon J, Pace N, Tiedje J, Ley R, Fierer N, Field D, Kyripides N, Glocker FO, Klenk HP, Wommack KE, Glass E, Doherty K, Gallery R, Stevens R, Knight R. 2010. The Earth Microbiome Project: meeting report of the “1 EMF Meeting on Sample Selection and Acquisition” at Argonne National Laboratory October 6 2010. Stand Genomic Sci 3:249–253. http://dx.doi.org/10.4056/ags.1443528.

38. Sahl JW, Lemmer D, Travis J, Schupp J, Gilleece J, Aziz M, Driebe E, Drees K, Hicks N, Williamsson C, Hepp C, Smith D, Roe C, Engelthaler D, Wagner D, Keim P. 2016. The northern Arizona SNP pipeline: accurate, flexible, rapid identification of SNPs in WGS data sets. bioRxiv http://dx.doi.org/10.1101/037267.

39. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. http://dx.doi.org/10.1101/gr.107119.110.
50. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725–2729. http://dx.doi.org/10.1093/molbev/mst197.

51. Letunic I, Bork P. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 39:W475–478. http://dx.doi.org/10.1093/nar/gkw270.

52. Buter G, Pancholi A, Neves GR, Antipov G, Gurevich AA, Dvorok M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pijibelski AD, Pyshkin AV, Sirokvin AV, Vyuhhi N, Tesler G, Alexseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

53. Sahi JW, Caporaso JG, Rasko DA, Keim P. 2014. The large-scale blast score format (LS-BSF) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ 2:e332. http://dx.doi.org/10.7717/peerj.332.

54. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, Brisse S, Cao H, Wilksch J, Decre D. 2015. Examination of the enterotoxic Escherichia coli population structure during human infection. mBio 6(1):e00501-15. http://dx.doi.org/10.1128/mBio.00501-15.

55. van Duin D, Cober E, Richter SS, Perez F, Kalayjian RC, Salzata RA, Evans S, Fowler VG, Jr, Kaye KS, Bonomo RA. 2015. Impact of therapy and strain type on outcomes in urinary tract infections caused by carbapenem-resistant Klebsiella pneumoniae. J Antimicrob Chemother 70:1203–1211. http://dx.doi.org/10.1093/acin/nku169.

56. Wachino K, Nagel JL. 2015. Clinical outcomes of Enterobacteriaceae infections stratified by carbapenem MICs. J Clin Microbiol 53:201–205. http://dx.doi.org/10.1128/JCM.03057-14.

57. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. http://dx.doi.org/10.1093/bioinformatics/btu170.

58. Kim HJ, Choi HS, Kim MN, Choi HR, Jeong JS, Choi SH. 2015. Examination of the enterotoxic Escherichia coli population structure during human infection. mBio 6(1):e00501-15. http://dx.doi.org/10.1128/mBio.00501-15.

59. Allyn C, Cormican M, Vellinga A, Johnson JR, Austin B, Morris RE, Schupp JM, Hicks ND, Smith DE, Buchhagen JL, Valafar RD, II. 2015. Factors associated to prevalence and incidence of car-

60. Antimicrob Agents Chemother 51:4401–4409. http://dx.doi.org/10.1128/AAC.00926-07.

61. Scott RD, II. 2009. The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. Centers for Disease Control and Prevention, Atlanta, GA. https://www.cdc.gov/hai/pdfs/hai/Scott_CostPaper.pdf.

62. Patel TS, Sahl HW, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, Brisse S, Cao H, Wilksch J, Decre D. 2015. Examination of the enterotoxic Escherichia coli population structure during human infection. mBio 6(1):e00501-15. http://dx.doi.org/10.1128/mBio.00501-15.

63. Wachino K, Nagel JL. 2015. Clinical outcomes of Enterobacteriaceae infections stratified by carbapenem MICs. J Clin Microbiol 53:201–205. http://dx.doi.org/10.1128/JCM.03057-14.

64. Daroukh A, Delauana C, Bigot S, Ceci JM, Siddihoun N, Bukuieya I, Raisin J, Porcheret H, Maisonneuve I, Boulouyre MA. 2014. Characteristics and costs of carbapenemase-producing enterobacteria carriers (2012/2013), Med Mal Infect 44:321–326. http://dx.doi.org/10.1016/j.medmal.2014.06.004.

65. Lin MY, Lyles-Banks RD, Lolans K, Hines DW, Spear JB, Petrak R, Trick WE, Weinstein RA, Hayden MK, Centers for Disease Control and Prevention Epicenters Program. 2013. The importance of long-term acute care hospitals in the regional epidemiology of Klebsiella pneumoniae carbapenemase-producing Enterobacteriaceae. Clin Infect Dis 57:1246–1252. http://dx.doi.org/10.1093/cid/cit500.

66. Janvier F, Delacour H, Tese S, Larreche S, Sammartini N, Ollat D, Rapp C, Merens A. 2014. Facetal carriage of extended-spectrum beta-lactamase-producing enterobacteria among soldiers at admission in a French military hospital after aeromedical evacuation from overseas. Eur J Clin Microbiol Infect Dis 33:1719–1723. http://dx.doi.org/10.1007/s10096-014-2141-8.

67. Allyn J, Angue M, Belmonte O, Lugagne N, Traversier N, Vandroux D, Lefort Y, Allou N. 2015. Delayed diagnosis of high drug-resistant microorganisms carriage in repatriated patients: three cases in a French inten-

68. Bart Y, Paul M, Eluk O, Geffen Y, Rahino G, Hussein K. 2015. Risk factors for recurrence of carbapenem-resistant Enterobacteriaceae carriage: case-control study. Infect Control Hosp Epidemiol 36:936–941. http://dx.doi.org/10.1017/ice.2015.82.

69. Kim KR, Lee JY, Park HY, Kwak SH, Lim YJ, Hong MJ, Kim SK, Park SY, Kim HJ, Choi HS, Kim MN, Choi HR, Jeong SH, Choi SH. 2015. Clearance rate of carbapenemase-producing Enterobacteriaceae carriage among hospitalized patients. Infect Control Hosp Epidemiol 36:1361–1362. http://dx.doi.org/10.1017/ice.2015.169.

70. Torres-Gonzalez P, Cervera-Hernandez ME, Niembro-Ortega MD, Leal-Vega F, Cruz-Hervert LP, Garcia-Garcia L, Galindo-Fraga A, Mar-tinez-Gamboa A, Bobadilla-Del Valle M, Sifuentes-Osornio J, Ponce-de-Leon A. 2015. Factors associated to prevalence and incidence of car-

71. Gaglioti C, Ciccarese V, Sarti M, Giordani S, Barozzi A, Braglia C, Cormican M, Vellinga A, Johnson JR, Austin B, Morris RE, Schupp JM, Hicks ND, Smith DE, Buchhagen JL, Valafar RD, II. 2015. Factors associated to prevalence and incidence of car-

72. Antimicrob Agents Chemother 51:4401–4409. http://dx.doi.org/10.1128/AAC.00926-07.
81. Armand-Lefèvre I, Angebault C, Barbier F, Hamelet E, Defrance G, Ruppe E, Bronchard R, Lepeule R, Lucet JC, El Mniai A, Wolff M, Montravers P, Plesiat P, Andremont A. 2013. Emergence of imipenem-resistant Gram-negative bacilli in intestinal flora of intensive care patients. Antimicrob Agents Chemother 57:1488–1495. http://dx.doi.org/10.1128/AAC.01823-12.

82. Prasad N, Labaze G, Kopacz J, Chwa S, Platis D, Pan CX, Russo D, LaBombardi VJ, Osorio G, Pollack S, Kreiswirth BN, Chen L, Urban C, Segal-Maurer S. 2016. Asymptomatic rectal colonization with carbapenem-resistant Enterobacteriaceae and Clostridium difficile among residents of a long-term care facility in New York City. Am J Infect Control 44:525–532. http://dx.doi.org/10.1016/j.ajic.2015.11.021.

83. Breurec S, Melot B, Hoen B, Passet V, Schepers K, Bastian S, Brisse S. 2016. Liver abscess caused by infection with community-acquired Klebsiella quasipneumoniae subsp. quasipneumoniae. Emerg Infect Dis 22:529–531. http://dx.doi.org/10.3201/eid2203.151466.

84. Maatallah M, Vading M, Kabir MH, Bakhrouf A, Kalin M, Nauclet P, Brisse S, Giske CG. 2014. Klebsiella variicola is a frequent cause of bloodstream infection in the Stockholm area, and associated with higher mortality compared to K. pneumoniae. PLoS One 9:e113539. http://dx.doi.org/10.1371/journal.pone.0113539.

85. Sichtig H. 2016. Infectious disease next generation sequencing based diagnostic devices: microbial identification and detection of antimicrobial resistance and virulence markers, draft guidance for industry and Food and Drug Administration staff. Division of Microbiology Devices, U.S. Food and Drug Administration, Rockville, MD. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM500441.pdf.