Genome Sequencing of an Extended Series of NDM-Producing *Klebsiella pneumoniae* Isolates from Neonatal Infections in a Nepali Hospital Characterizes the Extent of Community- versus Hospital-Associated Transmission in an Endemic Setting

N. Stoesser,[a] A. Giess,[a] E. M. Batty,[b] A. E. Sheppard,[a] A. S. Walker,[c] D. J. Wilson,[a,b] X. Didelot,[d] A. Bashir,[e] R. Sebra,[e] A. Kasarskis,[e] B. Sthapit,[f] M. Shakya,[f] D. Kelly,[c,g] A. J. Pollard,[c,g] T. E. A. Peto,[a] D. W. Crook,[a] P. Donnelly,[b] S. Thorson,[f] P. Amatya,[f] S. Joshi[f]

Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom[a]; Wellcome Trust Center for Human Genetics, Oxford, United Kingdom[b]; NIHR Biomedical Research Center, University of Oxford/Oxford University Hospitals NHS Trust, Oxford, United Kingdom[c]; School of Public Health, Imperial College London, London, United Kingdom[d]; Icahn School of Medicine at Mount Sinai, New York, New York, USA[e]; Department of Pediatrics, Patan Hospital, Kathmandu, Nepal[f]; Oxford Vaccine Group, Center for Clinical Vaccinology and Tropical Medicine, University of Oxford, Oxford, United Kingdom[g]

NDM-producing *Klebsiella pneumoniae* strains represent major clinical and infection control challenges, particularly in resource-limited settings with high rates of antimicrobial resistance. Determining whether transmission occurs at a gene, plasmid, or bacterial strain level and within hospital and/or the community has implications for monitoring and controlling spread. Whole-genome sequencing (WGS) is the highest-resolution typing method available for transmission epidemiology. We sequenced carbapenem-resistant *K. pneumoniae* isolates from 26 individuals involved in several infection case clusters in a Nepali neonatal unit and 68 other clinical Gram-negative isolates from a similar time frame, using Illumina and PacBio technologies. Within-outbreak chromosomal and closed-plasmid structures were generated and used as data set-specific references. Three temporally separated case clusters were caused by a single NDM *K. pneumoniae* strain with a conserved set of four plasmids, one being a 304,526-bp plasmid carrying bla_{NDM-1}. The plasmids contained a large number of antimicrobial/heavy metal resistance and plasmid maintenance genes, which may have explained their persistence. No obvious environmental/human reservoir was found. There was no evidence of transmission of outbreak plasmids to other Gram-negative clinical isolates, although bla_{NDM} variants were present in other isolates in different genetic contexts. WGS can effectively define complex antimicrobial resistance epidemiology. Wider sampling frames are required to contextualize outbreaks. Infection control may be effective in terminating outbreaks caused by particular strains, even in areas with widespread resistance, although this study could not demonstrate evidence supporting specific interventions. Larger, detailed studies are needed to characterize resistance genes, vectors, and host strains involved in disease, to enable effective intervention.

*K*lebsiella pneumoniae is a major pathogen, especially in neonatal critical care (1), and in association with extended-spectrum beta-lactamases (ESBLs) and/or carbapenemases, it results in increased hospital costs, longer stays, and high rates of mortality (2, 3). Furthermore, the success of controlling outbreaks of carbapenemase-producing *Klebsiella* spp. is mixed (4, 5). In resource-limited settings, such as South Asia, multidrug-resistant *Enterobacteriaceae* are widespread (1, 6), and therefore, delineating the relative contributions of within-hospital transmission of resistant strains and mobile genetic elements versus recurrent importation of these from the community is challenging.

Carbapenem resistance in *Klebsiella* spp. is typically caused by hydrolytic enzymes or nonexpression of porin genes/increased efflux pump activity in the presence of ESBLs (7). Of particular concern is the recent emergence of transmissible carbapenemases, such as *Klebsiella pneumoniae* carbapenemase (bla_{KPC}) (8) and New Delhi metallo-beta-lactamase (bla_{NDM}) (9). These genes are widely transmitted intra- and interspecies on mobile genetic elements. For example, bla_{NDM} has been observed in at least 40 countries, on different plasmids, in several species of *Enterobacteriaceae* (10), and in environmental and animal samples (11). The reported prevalence likely underrepresents the problem, as screening frequently relies on insensitive phenotypic methods; molecular methods are more robust (12).

Whole-genome sequencing (WGS) (13–15) can markedly alter classical interpretations of transmission (14, 15). Several studies have used lower-resolution typing methods, such as pulsed-field gel electrophoresis (PFGE), to characterize outbreaks caused by carbapenem-resistant *Klebsiella* spp., and two published studies have successfully utilized WGS, describing KPC- and NDM-1-producing strains, respectively (16, 17). The former study described a 6-month outbreak (22 isolates) using genetic and epidemiological data, supporting both patient-to-patient and ventilator-to-patient transmission. The latter study was smaller (8 isolates) and revealed

the likely transmission of NDM *K. pneumoniae* at a U.S. hospital over 4 months.

To extend the results of previously published work, we used Illumina and PacBio SMRT (single-molecule real-time sequencing) technology (Pacific Biosciences) to recover nearly complete genetic information, including that of episomal structures, of a series of disease-causing, multidrug-resistant *K. pneumoniae* isolates collected in a hospital neonatal service in Nepal. In addition, we made genetic comparisons with other contemporaneously collected clinical isolates (both *K. pneumoniae* and other Gram-negative bacilli), thereby investigating the hypothesis that these drug-resistant cases were occurring as a result of independent introductions of strains/plasmids from the wider community.

MATERIALS AND METHODS

Patan hospital is a 450-bed hospital in Kathmandu, Nepal, with a pediatric unit handling ~9,500 live births, 35,000 outpatient visits, and 2,600 admissions/year. Neonatal care is managed in two nurseries, one "clean," for non-sepsis-related supportive care, and one "septic," for patients who have major risk factors for or a diagnosis of sepsis. The neonatal and pediatric intensive care units (NICU and PICU) are located close to the nurseries.

The suspected first infected neonatal case was a low-birth-weight premature baby born in the hospital on 9 August 2011. Postnatally, he was admitted to the NICU for respiratory distress. Early blood cultures were negative, but on day 8, *K. pneumoniae* was cultured, initially susceptible only to carbapenems. The child died 4 days later, despite treatment.

Cases of sepsis associated with similar drug-resistant *K. pneumoniae*-positive cultures susceptible only to one or more of the carbapenems, colistin, or tigecycline were defined as part of an epidemiological cluster, with the last observed case on 30 June 2012. To contextualize genetic variation in cluster isolates and identify wider spread of resistance genes/elements, we randomly sampled a subset of stored *Enterobacteriaceae* and other Gram-negative bacilli, stratified by susceptibility profile, age (adults/children), and hospital-associated (sampled ≥48 h after admission) versus community-associated (sampled <48 h after admission, with no previous admission to Patan hospital and delivery outside a health care facility) infections (see Section 1 in the supplemental material).

Infection control was enhanced from December 2011, including changes in cleaning, equipment, and surveillance protocols and closure of the original clean nursery (see Section 2 in the supplemental material).

Laboratory and sequencing methods. All samples were originally processed locally. Species identification was based on biochemical profiling. Antimicrobial susceptibility was assayed using the disk diffusion method (18) and was reassessed using broth microdilution in Oxford (BD Phoenix automated microbiology system; Oxford, United Kingdom).

Environmental sampling is performed routinely as part of infection control surveillance, and this included the study period. Typically, paired swab samples are collected at 2-week intervals before and after intensive cleaning from five different randomly selected sites in the NICU and the established clean and septic nurseries, including laryngoscope blades, ventilators, stethoscopes, equipment trolleys, incubators, central lines, endotracheal/suction tubes, tap water, floors, and door handles. If a patient was considered part of a *K. pneumoniae* case cluster, a set of these surveillance swabs was taken from the patient's incubator or bed. Rectal swab samples were taken from all nursing and medical staff in the neonatal nurseries, NICU, and PICU (between 9 and 11 July 2012). Samples were also collected from purified water used for clinical purposes (26 September 2011), disinfectants (13 December 2011), and air conditioners in the nurseries and ICUs (12 December 2011). All samples were plated on MacConkey agar and incubated aerobically; *Klebsiella* isolates were identified as described above.

DNA was extracted from subcultured frozen isolate stocks (Quick-Gene; Fujifilm, Tokyo, Japan) with a mechanical lysis step (FastPrep; MP

Biomedicals, Santa Ana, CA). All extracts were sequenced on the HiSeq 2000 platform (Illumina, San Diego, CA), generating 100-base paired-end reads. For the isolate cultured from the first infected neonatal case (PMK1), PacBio SMRT sequencing was also undertaken.

Sequence data analysis. Species identification was confirmed with Kraken (19), and Illumina reads were mapped to species-specific reference genomes (see Sections 3 and 4 in the supplemental material). Mapping/variant-calling methodologies were performed as previously described (14). Reads were also *de novo* assembled using Velvet and VelvetOptimizer (20, 21); BLASTn was used to identify the presence/absence of resistance gene variants in assemblies (22) and infer multilocus sequence types (STs) (23).

A maximum-likelihood phylogeny of the first *K. pneumoniae* isolate per case based on single-nucleotide variants (SNVs) distributed over the core mapped genome was constructed in PhyML (see Section 5 in the supplemental material) (24).

The PacBio reads for PMK1 were assembled with HGAP (hierarchical genome assembly process) and Quiver (25). Chromosomal and plasmid contigs were identified using BLASTn (26) against the NCBI nucleotide sequence database. Circular overlaps in plasmid contigs were identified with nucmer (27). Reference plasmid sequences were closed and corrected with visual inspection of BWA (Burrows-Wheeler aligner)-generated Illumina and PacBio read alignments to plasmid contigs (28) and annotated with Prokka (29).

All strains were mapped to the PacBio-generated reference sequence (chromosome plus plasmids) with BWA. Structural variation in plasmids was identified by plotting the mean read coverage for each 1,000 bp of the reference sequence divided by the mean read coverage across the whole reference sequence, capped to a maximum of one. Precise breakpoints were identified from the inspection of mapped reads. The presumptive plasmid structures generated are denoted pPMK[isolate number]-[reference plasmid suffix], e.g., pPMK17-NDM. The NDM gene copy number was estimated by read counts of the NDM regions divided by the total number of reads, scaled to the count for pPMK1-NDM, and then rounded to the nearest whole number.

The genetic outbreak was then defined as any cases from the epidemiological cluster and closely genetically related isolates forming a distinct group on the maximum-likelihood tree. Comparing these to the PacBio-generated reference chromosome, a mutation rate for the outbreak strains was estimated using a time-scaled analysis in BEAST (see Section 6 in the supplemental material) (30), including longitudinal isolates from individuals. The most likely transmission chain was inferred using the R Outbreaker package (31) (see Section 7 in the supplemental material).

Nucleotide sequence accession numbers. Illumina-generated sequence data for the whole data set have been deposited at NCBI (project accession number PRJNA253300) (see Section 1 in the supplemental material). Sequence data for the PacBio/Illumina-generated PMK1 reference chromosome and plasmids have been deposited in GenBank with accession numbers as follows: (i) PMK1 (chromosome), CP008929; (ii) pPMK1-A, CP008930; (iii) pPMK1-B, CP008931; (iv) pPMK1-C, CP008932; and (v) pPMK-NDM, CP008933. Sequence data were also deposited in the Sequence Read Archive under accession number SRP043586.

RESULTS

Of 102 strains sequenced, 55 were confirmed as *K. pneumoniae*, 43 as other *Enterobacteriaceae*, and 4 as other Gram-negative bacilli. Of 55 *K. pneumoniae* isolates sequenced from 47 individuals, 34 isolates sampled from 26 individuals were part of the genetically defined outbreak. All outbreak isolates were cultured from blood, except for PMK9, which was cultured from cerebrospinal fluid. Figure 1 shows the timeline of cases in the epidemiologically defined cluster and/or the genetically defined outbreak. The prolonged time intervals between some cases (PMK1 to PMK3, 96 days; PMK9 to PMK10, 57 days; and PMK12 to PMK15, 34 days)
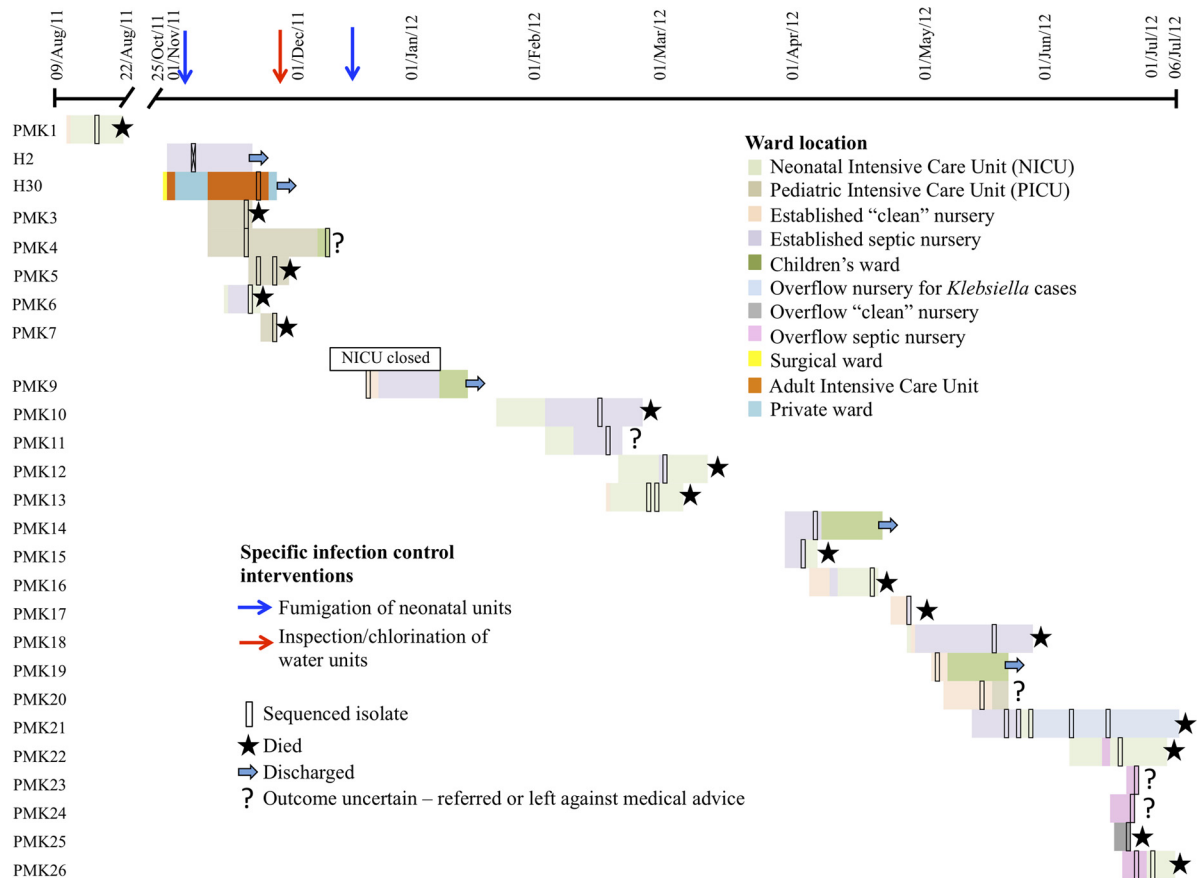
FIG 1 Timeline of *Klebsiella pneumoniae* cases, including individuals who were both part of epidemiologically defined case clusters and had genetically linked outbreak strains. H2 was found to be genetically unrelated to the other outbreak isolates; H30 was thought to share a relatively recent common ancestor but was not involved in the immediate transmission network. There were no clinical details available for H1460, which is therefore not shown.

demonstrate the challenge of distinguishing between multiple importations and ongoing spread. The outbreak affected all units, including the overflow nurseries established toward the end of the outbreak in a different hospital building.

The outbreak-associated mortality was high, with 16 (64%) inpatient deaths in 25 neonates (6 had an unknown outcome, e.g., referred elsewhere or left against medical advice), in contrast with a hospital-wide contemporaneous neonatal death rate of 46/6,908 (0.7%). The neonatal critical care mortality rates were 46% (45/98 cases) during the outbreak period and 27% (32/117) in the year following the final case ($P = 0.007$, Fisher's exact test).

Approximately 30 sites per month were sampled from the three neonatal units surveyed; the heaviest environmental contamination with *Klebsiella* spp. was observed at the onset of the outbreak, between August 2011 and January 2012 (mean of 2 environmental swabs positive [range, 0 to 3]), as opposed to the preceding and subsequent 6-month periods (no sites positive at any time point; mean of 0 sites positive [range, 0 to 1], respectively) (see Section 8 in the supplemental material). *Klebsiella* spp. bacteria were isolated from a number of environmental cultures, including samples obtained from laryngoscope blades, suctioning equipment, purified water containers, tap water, soap dispensers, a burette set surface swab, and a health care worker's hands; none of these underwent susceptibility testing. Eighteen of 69 (26%) rectal swab

samples cultured *Klebsiella* spp., but none were ESBL or carbapenemase producers. None of these isolates were stored, and therefore none were sequenced.

The sequenced *K. pneumoniae* isolates represented 14 STs; six were novel (Fig. 2). The strains exhibited considerable diversity (56,947 core variable sites as mapped to reference strain MGH78578; mean call rate of 87%). H2, part of the epidemiologically defined cluster, was distantly related to the outbreak cases (16,557 SNVs from PMK1). Conversely, H1460, not part of the cluster, was part of the genetically defined outbreak (1 SNV from PMK26a), as was a hospital-associated strain from an adult, H30 (21 SNVs from PMK1). All outbreak isolates were sequence type 15 (ST15). Three additional smaller, closely genetically related clusters were observed; two neonatal hospital-associated clusters and a pair of community-associated infections within 3 weeks of each other (Fig. 2; see also Section 1 in the supplemental material).

We also sequenced 47 other clinical isolates representing different species (Fig. 3A and B); the earliest isolate was from 5 January 2008 and the latest from 1 August 2012. Within the wider hospital, several isolates causing bloodstream infections in neonates remained unsequenced, including five *K. pneumoniae* isolates, four isolates defined as *Klebsiella* spp., two *Escherichia coli* isolates, seven *Acinetobacter* spp. isolates, and two *Klebsiella oxytoca* isolates. Genotypic and phenotypic multidrug resistance was very common in both hospital- and community-associated iso-
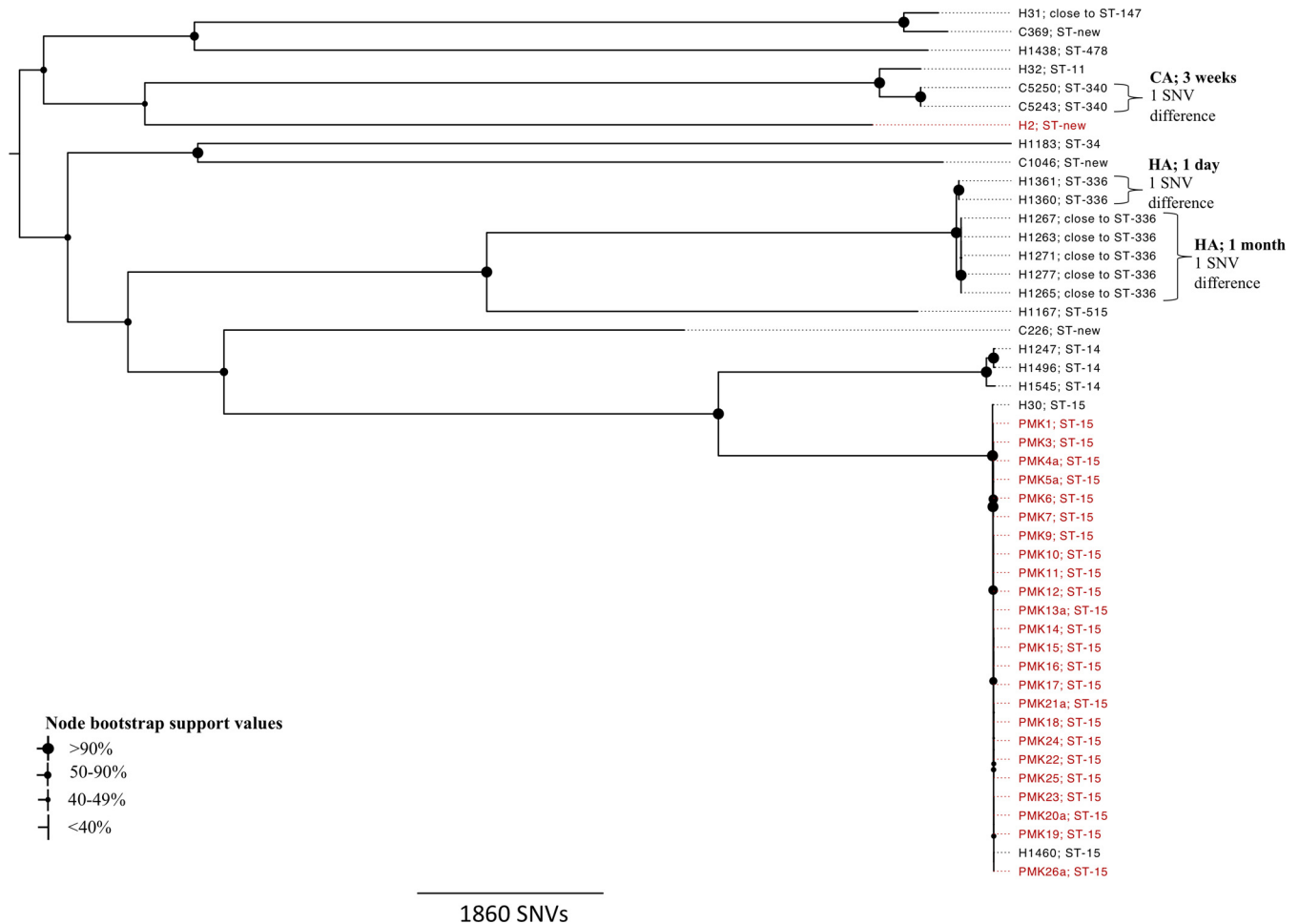
FIG 2 Maximum-likelihood phylogeny of all sequenced *Klebsiella pneumoniae* study isolates (deduplicated by individual). Isolates in red represent those originally considered part of case clusters on the basis of clinical suspicion and susceptibility testing. Additional closely genetically related clusters of nonoutbreak strains are indicated with curly brackets. "CA" denotes community-associated strains, and "HA" hospital-associated strains; the time interval listed is that spanning the isolation dates of clustered isolates. Symbols indicating bootstrap support at nodes are defined in the key.

lates: 13/21 (62%) nonoutbreak *K. pneumoniae* isolates, 12/14 (86%) *Enterobacter cloacae* isolates, and 6/8 (75%) *Klebsiella oxytoca* isolates contained at least one variant of each of *aac*, bla_{TEM}, bla_{OXA}, bla_{CTX-M}, and *qnr* in combination. Three *E. coli* isolates and one *E. cloacae* isolate contained bla_{NDM-1}, and a further *E. coli* isolate contained bla_{NDM-6}.

Detailed outbreak strain analysis. Fifty-three high-confidence chromosomal SNVs were identified during the outbreak (Fig. 4A and B), 21 uniquely in the adult H30 strain. Sixteen SNVs emerged and persisted in more than one isolate, 14 in coding sequences. Of these, eight resulted in nonsynonymous mutations (Fig. 4A). There was a large, 121,366-bp deletion in PMK13b (reference bp positions, 3047928 to 3169294).

From the time-scaled phylogenetic tree (see Section 6 in the supplemental material), the outbreak strain's mutation rate was estimated at 3.65×10^{-6} (95% confidence interval [CI], 2.45×10^{-6} to 4.89×10^{-6}) mutations per called site per year, equating to 18.4 (95% CI, 12.3 to 24.6) mutations per genome per year. The time to most recent common ancestor of the adult-associated H30 strain and the PMK1 strain cultured from the first infected neonatal case was estimated at between 1 and 7 months before the neonatal outbreak was observed.

The transmission network inferred by Outbreaker (Fig. 5) demonstrates uncertainty around the specific transmission links for early strains (PMK3 to PMK9). Four individuals harboring these strains shared ward space and time (PMK3 to PMK5 and PMK7), and the network is consistent with direct transmissions. PMK9 may have contaminated the ward or equipment in the established septic nursery or colonized an unsampled asymptomatic host, leading to the infection represented by PMK10 and, possibly, also to PMK11 (although this case was also consistent with a direct transmission from PMK10) and later spreading to the individual harboring PMK14. The link between PMK11 and PMK12 again most likely represents an indirect transmission through either an environmental or asymptomatic source, and the link between PMK11 and PMK13 represents an event across wards. Based on the available epidemiologic data (Fig. 1), the spread from PMK14 most likely occurred within the established septic nursery (possibly directly to PMK15, otherwise indirectly via equipment or a colonized asymptomatic contact). The links between PMK18, PMK21a, PMK22, PMK24, and PMK25 could potentially have been established through sequential transfer from the established septic nursery to the NICU and then to the overflow septic nursery. PMK19 to PMK20 may represent a direct transmission event,
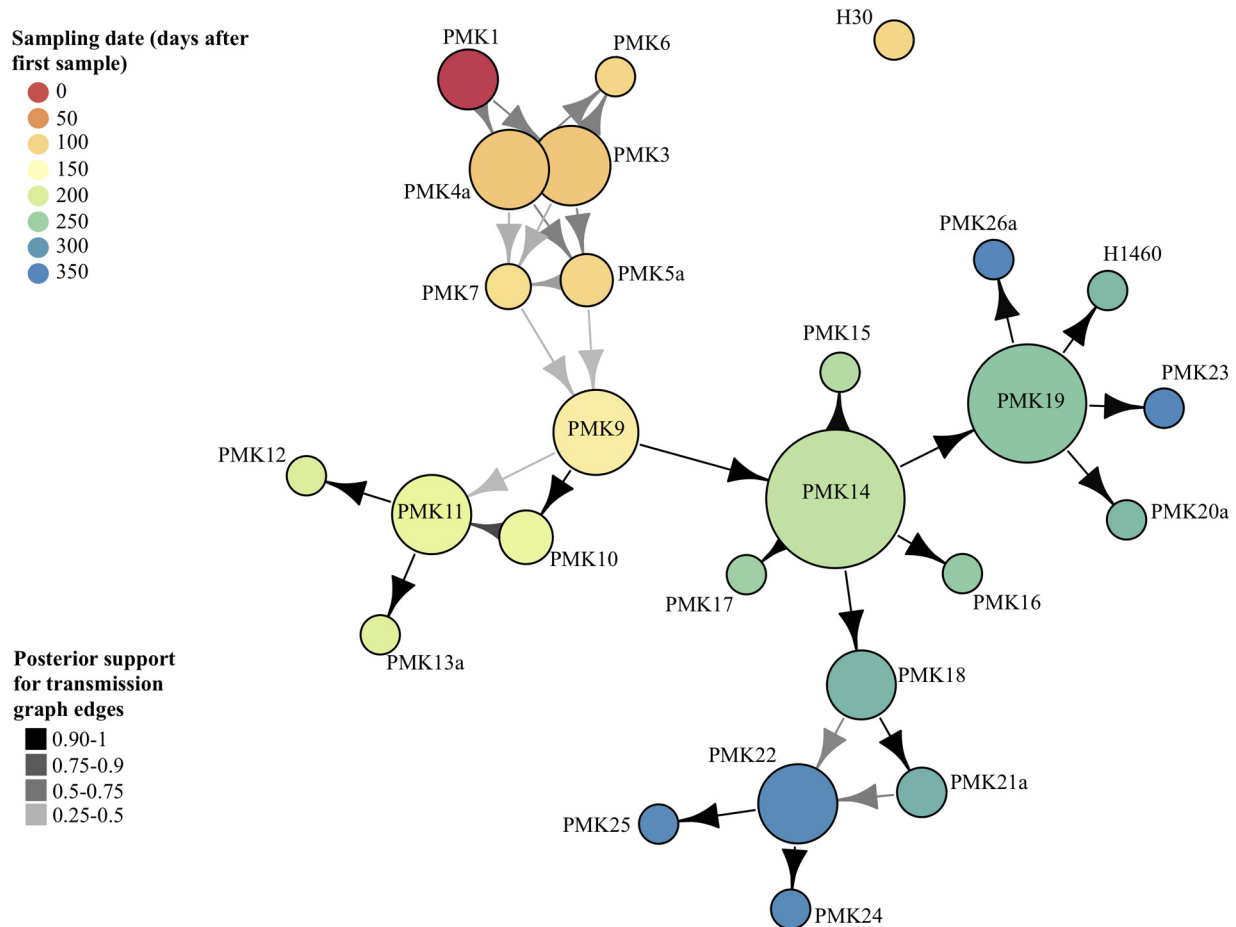
FIG 3 Resistance genotypes and susceptibility phenotypes as determined by the Phoenix automated system are shown for all sequenced *Klebsiella pneumoniae* isolates (A) and for all sequenced non-*Klebsiella pneumoniae* isolates (B). Copy numbers calculated for bla_{NDM}, where present, are also shown. Susceptibility categories were determined by Phoenix in accordance with EUCAST breakpoints. S, susceptible; I, intermediate; R, resistant; NA, no result available; AMP, ampicillin-amoxicillin; AMC, amoxicillin-clavulanate; TZP, piperacillin-tazobactam; CXM, cefuroxime; CRO, ceftriaxone; CAZ, ceftazidime; CIP, ciprofloxacin; ERT, ertapenem; MEM, meropenem; ATM, aztreonam; GEN, gentamicin; AMK, amikacin; SXT, trimethoprim-sulfamethoxazole; CST, colistin.

A

B

FIG 5 Transmission network inferred using Outbreaker. The sizes of the nodes reflect the numbers of secondary cases per infected individual. The colors of nodes reflect the sampling date as indicated in the key at the upper left. Shading of arrows represents the degree of posterior support for the transmission link, as indicated in the key at the bottom left.

but the nature of the link between PMK19 and cases PMK23 and PMK26a is less evident. The model identifies PMK14 and PMK19 as contributing to the largest number of secondary outbreak cases, with H30 excluded from the inferred transmission network.

Four complete plasmid sequences were identified in PMK1. The *bla*$_{NDM-1}$-containing plasmid, pPMK1-NDM (304,526 bp), was a multireplicon (IncHI1B/IncFIB) plasmid with antibiotic resistance determinants that included *aac(6′)-Ib-cr*, *aadA2*, *bla*$_{CTX-M-15}$, *bla*$_{OXA-1}$, *folP*, *catA1*, *dfrA12*, *armA*, and a large conjugative transfer module. A number of mercury resistance (*mer*) genes were present (see Section 9 in the supplemental material).

pPMK1-A (187,571 bp), also a multireplicon (IncFII$_K$/IncFIB$_K$) and likely a conjugative plasmid, contained tetracycline resistance genes *tetA* and *tetR* and iron (*fec*), arsenic (*ars*), copper (*cop*), tellurite, and silver (*sil*) resistance gene cassettes. pPMK1-B (111,693 bp) was a colE1-type, IncFIB plasmid, containing a tellurite resistance gene but lacking obvious conjugative transfer genes. pPMK1-C (69,947 bp) contained *aph(6)-Id* and *aph(3′)-Ib*-like resistance genes encoding streptomycin and kanamycin/neomycin resistance, respectively (see Section 9 in the supplemental material).

All four plasmids were highly conserved across the outbreak isolates (Fig. 6). There were five SNVs in the NDM-containing plasmid and two in pPMK-B, with three and one SNV, respectively, in these plasmids only in H30. No SNV-level variation was observed for pPMK-A or pPMK-C (Fig. 4A).

FIG 4 (A) Nucleotide and amino acid level variations in outbreak isolates with respect to the sequences of the PMK1 reference chromosome and pPMK1-B and pPMK1-NDM. del, deletion; NA, unannotated hypothetical protein; -, stop codon. The order of the isolates approximates the outbreak time frame but also accommodates the grouping of isolates that share identical genetic sequences based on mapping to the PMK1 reference chromosome (i.e., 0 SNV differences; shown in boxes with black borders). Brown-shaded chromosomal positions represent noncoding positions, gray-shaded positions represent positions in plasmids. Nonsynonymous, persistent mutations occurred in *ppx* (an exopolyphosphatase), *azlC* (an azaleucine resistance protein), *adaA* (a methyltransferase), *gcd* (a quinoprotein glucose dehydrogenase), *bmR* (a transcriptional repressor), *oprD* (an outer membrane porin), *ssuC* (an alkane sulfonate transporter subunit), and *virB4* (a type IV secretion system). (B) Phylogenetic tree summarizing chromosomal genetic relationships between all outbreak isolates. Colored nodes represent sampled isolates and black nodes unsampled intermediates; node colors are defined in the key. Each solid branch represents a single SNV, with branch colors indicating mutation types as follows: orange, nonsynonymous; green, synonymous; gray, intergenic; gray dashed line, 121,366-bp deletion; black dashed line, 21 single-nucleotide variants (SNVs).
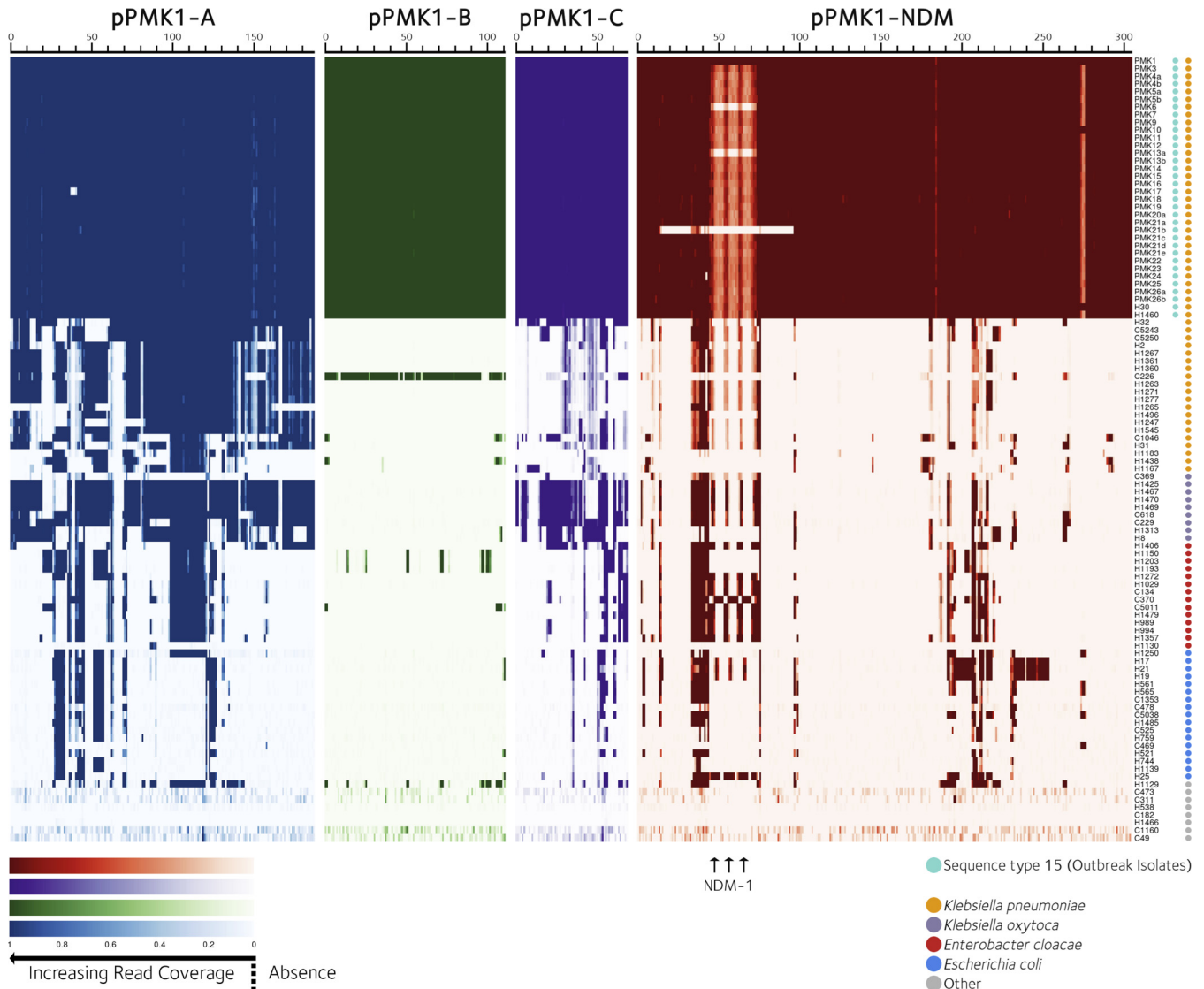
FIG. 6 Mapping of coverage of reference plasmids in outbreak and nonoutbreak strains. Reference plasmid coordinates are displayed in kilobases. Heatmap values represent the number of bases mapped, scaled by the number of bases mapped to the whole reference plasmid. A value of zero represents absence. A value of 1 represents presence at the average coverage or greater. Values between zero and 1 represent presence at a lower-than-average coverage.

Plasmid pPMK1-NDM contained two regions not present in other outbreak strains. The first was a tandem duplication of the bla_{NDM-1} region (kb positions 56 to 76), resulting in three bla_{NDM-1} copies, and the second an acquisition of two transposases (bp positions 272471 to 275170) in the plasmid structure. pPMK6-NDM and pPMK13a-NDM shared a 26-kb deletion (bp positions 454501 to 70264), including bla_{NDM-1}. pPMK21b-NDM contained a larger, 83-kb deletion (bp positions 13135 to 96114) involving bla_{NDM-1} and other antibiotic resistance genes, including $aac(6')$-Ib-cr, bla_{OXA-1}, and catB3. bla_{NDM-1} gene deletions correlated with reversion to ertapenem susceptibility. pPMK24-NDM contained a small, 2,401-bp deletion representing three phage-related open reading frames. The only other variation in plasmid gene presence was a deletion in bp positions 36552 to 40801 of pPMK17-A, involving tetracycline resistance genes (tetA and tetR).

Comparison of plasmid structures across nonoutbreak Kleb-siella strains and non-K. pneumoniae bacteria found outbreak plasmids to be restricted to the outbreak Klebsiella strains (Fig. 6). Partial exceptions were (i) pPMK1-B, large tracts of which were also found in a community-associated K. pneumoniae strain (C226), and (ii) pPMK1-A, regions of which were found in several nosocomial and community-associated K. pneumoniae and K. oxytoca strains. pPMK1-NDM, however, was not observed in any nonoutbreak strains, and in the five other NDM-positive isolates, bla_{NDM} was located in non-pPMK1-NDM genetic backgrounds (four were E. coli isolates H17, H19, H21 [closely genetically related; data not shown], and H25, and one was the E. cloacae isolate C370).

DISCUSSION

Despite the time lapses between isolates and initial uncertainty as to whether all cases were linked, WGS clearly demonstrated that the outbreak was caused by a single, clonal strain of ST15 K. pneu-

moniae in association with a conserved population of four plasmids, including a bla_{NDM-1}-containing plasmid. The intervals between case clusters suggest persistence of the NDM-1 *K. pneumoniae* strain in the unit environment or in asymptomatic carriage, potentially supported by the isolation of *Klebsiella* spp. from environmental samples and rectal swabs taken from staff. Both environmental contamination and asymptomatic colonization are likely contributors to the transmission of drug-resistant *K. pneumoniae* in the nosocomial setting (17), although the relative contribution of each route is unknown and could feasibly vary among lineages. The assessment of combined epidemiologic data and the transmission network inferred from sampling dates and the genomic data in this study strongly supports the view that even-wider sampling frames are needed to fully understand the dynamics of these outbreaks, given that both direct and indirect human and environmental transmissions are likely to be occurring. The time-scaled analysis of the outbreak isolates and H30 suggests that an ancestral strain predating PMK1 and H30 was present somewhere between 1 and 7 months before the identification of the first infected neonatal case—this ancestral strain may have been present in the parents, other patients, hospital staff, or the hospital environment.

While it is impossible to exclude repeated introductions of either the strain or the bla_{NDM-1}-containing plasmid into the pediatric critical care setting from the community or elsewhere in the hospital, the extraordinary degree of similarity between outbreak strains, in contrast to other contemporaneous strains from both locations, makes this less likely. Regarding the selection of the wider set of isolates for sequencing, we specifically avoided characterizing only phenotypically carbapenem-resistant organisms, given the known lack of sensitivity of phenotypic methods in the presence of carbapenemases (12), aiming to determine whether there was any evidence for wider dissemination of the outbreak plasmids or the outbreak *K. pneumoniae* strain in the absence of bla_{NDM}. The wider sampling and detailed plasmid analysis are major strengths of our study and expand on the two previous WGS outbreak investigations (16, 17); without such data, uncertainty about transmission versus repeated importations of strains or dissemination of drug resistance plasmids remains, particularly in high-prevalence contexts.

This study exemplifies the potential of using WGS to benefit both outbreak management and antimicrobial treatment. In particular, long-read (30-kb) PacBio sequencing enabled us to produce reference assemblies of the isolate cultured from the first infected neonatal case, including both chromosome and plasmids, which were then used as a comparator for Illumina short-read data sets for the other isolates, allowing fine-scale definition of genetic differences between strains and probabilistic interpretation of likely transmission pathways. As well as resolving the temporally distinct clusters of NDM-1 *K. pneumoniae* isolates into a single year-long outbreak, phylogenetic analysis identified several other unrecognized clusters of antimicrobial-resistant isolates, indicating both nosocomial (two separate CTX-M-15 *K. pneumoniae* outbreaks, one NDM-1 *E. coli* outbreak, and two *K. oxytoca* and three *E. cloacae* clusters [data not shown]) and community (CTX-M-15 *K. pneumoniae*) transmission.

Laboratory susceptibility phenotyping for susceptible, intermediate, and resistant (SIR) categories were inconsistent in some cases between isolates with the same complement of resistance genes, even for the highly genetically related outbreak strains,

highlighting the challenges of relying on this for cluster identification (observed for meropenem in the presence of NDM-1 and gentamicin in the presence of *armA*) (Fig. 4A). Although sequence-based susceptibility prediction has yet to be correlated with patient-level clinical outcomes, it appears to be sensitive and specific (22, 32) and indicates the potential future value of WGS in managing patients when current routine laboratory turnaround times are matched. Single-strand sequencing platforms and new, fast genome assemblers (33) are likely to make resistance prediction from clinical specimens, such as blood cultures, possible within hours of sampling.

This NDM *K. pneumoniae* outbreak terminated after a year, and hospital-wide neonatal critical care deaths almost halved after it ended, consistent with its suppression or eradication. This is of interest for two reasons: First, bla_{NDM} is known to be locally prevalent and was found in 5 of 68 (7%) sequenced, nonoutbreak clinical isolates in this study, and second, the hospital is of older construction and potentially less amenable to infection control. Carbapenemase-associated *K. pneumoniae* outbreaks have been shown to be difficult to control in some settings, even those where resources are less restricted (5). More recently, a separate NDM-1 *Enterobacter cloacae* outbreak was observed in the hospital, albeit in association with a different plasmid vector (data not shown), and was terminated following fumigation of a number of the clinical units, demonstrating the need for ongoing surveillance and comprehensive infection control.

ST15 *K. pneumoniae* is one of the dominant global clones, associated with a range of beta-lactamases, including NDM and CTX-M-15 (34, 35). Its success may partly relate to the accumulation of resistance without fitness costs (36). The outbreak strain stably supported several plasmids totaling nearly 700 kb, 13% of the chromosome size. The outbreak plasmids contained resistance genes, plasmid addiction modules, and genes encoding systems protecting host bacteria against bacteriophages or plasmids, which may have contributed to the strain's dominance and stable persistence.

pPMK1-NDM showed substantial homology to pNDM-MAR (JN420336.1), first identified from ST15 *K. pneumoniae* isolates in Morocco, but it differed from other NDM-containing plasmids (37). The major differences were the presence of additional NDM-1 copies and further resistance genes in pPMK1-NDM (see Section 10 in the supplemental material). Carbapenem exposure affects NDM-1 copy number, and lack of selection pressure can lead to complete gene deletion (38), providing further evidence to support minimizing unnecessary carbapenem therapy.

Our mutation rate estimate for the outbreak *K. pneumoniae* strain is higher than the mutation rates of other *Enterobacteriaceae*, such as *E. coli* (~1.1/genome/year) (39). There are no published data on *K. pneumoniae* mutation rates to our knowledge; it is therefore impossible to ascertain whether this strain was adaptable because of hypermutability or whether this is a species-level phenomenon. Mutation rates define the plausible time frame of acquisition events in bacteria and are critical now that genomic data are increasingly being relied upon to refine transmission epidemiology. Some SNVs were intergenic, highlighting the potential additional resolution achieved using the complete mapped genome rather than extended multilocus sequence typing for transmission analyses (40).

There are several limitations to our study. The NDM-1 copy number variation may have been due to variable selection pres-

sure in strain culture and storage. Selective culture would have aided environmental detection of NDM-containing isolates, and the identification of an environmental source or the presence of a susceptible variant of the outbreak clone was unfortunately impossible given the lack of further susceptibility testing on and/or storage of environmental/fecal carriage isolates for sequencing. Routine surveillance of patients admitted to at-risk units would have been ideal but was not possible given the local resource constraints. The wider set of clinical isolates did not contain other ST15 *K. pneumoniae* bacteria; these were either present at low frequencies and thus unsampled or were not present at all. Despite our wider sampling, we were still unable to identify the source of the outbreak, a limitation which is likely to be overcome when it is possible to implement WGS as a high-resolution, real-time typing method and use it to refine and extend ongoing outbreak investigations.

In summary, this study highlights challenges in managing clusters of resistant *K. pneumoniae* infections and demonstrates that, despite a lack of obvious transmission pathways, the same strain can persist in hospital environments for months, sporadically causing disease. Whether this is a consequence of particular combinations of host bacterial strains and NDM-containing and/or other plasmids, representing clinically successful entities with a significant impact on patient outcome, is unclear. However, WGS provides a high-resolution mechanism by which the contribution of these different elements can start to be unraveled.

ACKNOWLEDGMENTS

REFERENCES

1. Downie L, Armiento R, Subhi R, Kelly J, Clifford V, Duke T. 2013. Community-acquired neonatal and infant sepsis in developing countries: efficacy of WHO's currently recommended antibiotics—systematic review and meta-analysis. Arch. Dis. Child. **98:**146–154. http://dx.doi.org/10.1136/archdischild-2012-302033.
2. Giske CG, Monnet DL, Cars O, Carmeli Y, ReAct-Action on Antibiotic Resistance. 2008. Clinical and economic impact of common multidrug-resistant gram-negative bacilli. Antimicrob. Agents Chemother. **52:**813–821. http://dx.doi.org/10.1128/AAC.01169-07.
3. Gupta N, Limbago BM, Patel JB, Kallen AJ. 2011. Carbapenem-resistant Enterobacteriaceae: epidemiology and prevention. Clin. Infect. Dis. **53:**60–67. http://dx.doi.org/10.1093/cid/cir202.
4. Munoz-Price LS, Quinn JP. 2013. Deconstructing the infection control

5. Tofteland S, Naseer U, Lislevand JH, Sundsfjord A, Samuelsen O. 2013. A long-term low-frequency hospital outbreak of KPC-producing Klebsiella pneumoniae involving intergenus plasmid diffusion and a persisting environmental reservoir. PLoS One **8:**e59015. http://dx.doi.org/10.1371/journal.pone.0059015.
6. Jean SS, Hsueh PR. 2011. High burden of antimicrobial resistance in Asia. Int. J. Antimicrob. Agents **37:**291–295. http://dx.doi.org/10.1016/j.ijantimicag.2011.01.009.
7. Pfeifer Y, Cullik A, Witte W. 2010. Resistance to cephalosporins and carbapenems in Gram-negative bacterial pathogens. Int. J. Med. Microbiol. **300:**371–379. http://dx.doi.org/10.1016/j.ijmm.2010.04.005.
8. Yigit H, Queenan AM, Anderson GJ, Domenech-Sanchez A, Biddle JW, Steward CD, Alberti S, Bush K, Tenover FC. 2001. Novel carbapenem-hydrolyzing beta-lactamase, KPC-1, from a carbapenem-resistant strain of Klebsiella pneumoniae. Antimicrob. Agents Chemother. **45:**1151–1161. http://dx.doi.org/10.1128/AAC.45.4.1151-1161.2001.
9. Yong D, Toleman MA, Giske CG, Cho HS, Sundman K, Lee K, Walsh TR. 2009. Characterization of a new metallo-beta-lactamase gene, bla(NDM-1), and a novel erythromycin esterase gene carried on a unique genetic structure in Klebsiella pneumoniae sequence type 14 from India. Antimicrob. Agents Chemother. **53:**5046–5054. http://dx.doi.org/10.1128/AAC.00774-09.
10. Johnson AP, Woodford N. 2013. Global spread of antibiotic resistance: the example of New Delhi metallo-beta-lactamase (NDM)-mediated carbapenem resistance. J. Med. Microbiol. **62:**499–513. http://dx.doi.org/10.1099/jmm.0.052555-0.
11. Woodford N, Wareham DW, Guerra B, Teale C. 2014. Carbapenemase-producing Enterobacteriaceae and non-Enterobacteriaceae from animals and the environment: an emerging public health risk of our own making? J. Antimicrob. Chemother. **69:**287–291. http://dx.doi.org/10.1093/jac/dkt392.
12. Doyle D, Peirano G, Lascols C, Lloyd T, Church DL, Pitout JD. 2012. Laboratory detection of Enterobacteriaceae that produce carbapenemases. J. Clin. Microbiol. **50:**3877–3880. http://dx.doi.org/10.1128/JCM.02117-12.
13. Didelot X, Bowden R, Wilson DJ, Peto TE, Crook DW. 2012. Transforming clinical microbiology with bacterial genome sequencing. Nat. Rev. Genet. **13:**601–612. http://dx.doi.org/10.1038/nrg3226.
14. Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, Ip CL, Golubchik T, Batty EM, Finney JM, Wyllie DH, Didelot X, Piazza P, Bowden R, Dingle KE, Harding RM, Crook DW, Wilcox MH, Peto TE, Walker AS. 2013. Diverse sources of C. difficile infection identified on whole-genome sequencing. N. Engl. J. Med. **369:**1195–1205. http://dx.doi.org/10.1056/NEJMoa1216064.
15. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TE. 2013. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. Lancet Infect. Dis. **13:**137–146. http://dx.doi.org/10.1016/S1473-3099(12)70277-3.
16. Epson EE, Pisney LM, Wendt JM, MacCannell DR, Janelle SJ, Kitchel B, Rasheed JK, Limbago BM, Gould CV, Kallen AJ, Barron MA, Bamberg WM. 2014. Carbapenem-resistant Klebsiella pneumoniae producing New Delhi metallo-beta-lactamase at an acute care hospital, Colorado, 2012. Infect. Control Hosp. Epidemiol. **35:**390–397. http://dx.doi.org/10.1086/675607.
17. 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**(148)**:**148ra116. http://dx.doi.org/10.1126/scitranslmed.3004129.
18. NCCLS. 2011. Performance standards for antimicrobial susceptibility testing; 21st informational supplement, vol 31, no 1. M100-S21. National Committee for Clinical Laboratory Standards, Wayne, PA.
19. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. **15:**R46. http://dx.doi.org/10.1186/gb-2014-15-3-r46.
20. Gladman S, Seemann T. 2008. VelvetOptimiser, v2.1.7. Victorian Bioinformatics Consortium, Monash University, Clayton, Australia.
21. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read

assembly using de Bruijn graphs. Genome Res. **18**:821–829. http://dx.doi.org/10.1101/gr.074492.107.

22. **Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, Johnson JR, Walker AS, Peto TE, Crook DW.** 2013. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J. Antimicrob. Chemother. **68**:2234–2244. http://dx.doi.org/10.1093/jac/dkt180.

23. **Diancourt L, Passet V, Verhoef J, Grimont PA, Brisse S.** 2005. Multilocus sequence typing of Klebsiella pneumoniae nosocomial isolates. J. Clin. Microbiol. **43**:4178–4182. http://dx.doi.org/10.1128/JCM.43.8.4178-4182.2005.

24. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. **59**:307–321. http://dx.doi.org/10.1093/sysbio/syq010.

25. **Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.

26. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

27. **Delcher AL, Phillippy A, Carlton J, Salzberg SL.** 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. **30**:2478–2483. http://dx.doi.org/10.1093/nar/30.11.2478.

28. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754–1760. http://dx.doi.org/10.1093/bioinformatics/btp324.

29. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**:2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.

30. **Drummond AJ, Rambaut A.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. **7**:214. http://dx.doi.org/10.1186/1471-2148-7-214.

31. **Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N.** 2014. Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. PLoS Comput. Biol. **10**:e1003457. http://dx.doi.org/10.1371/journal.pcbi.1003457.

32. **Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, Kearns AM, Pichon B, Young B, Wilson DJ, Llewelyn MJ, Paul J, Peto TE, Crook DW, Walker AS, Golubchik T.** 2014. Prediction of Staphylococcus aureus antimicrobial resistance by whole-genome sequencing. J. Clin. Microbiol. **52**:1182–1191. http://dx.doi.org/10.1128/JCM.03117-13.

33. **Iqbal Z, Turner I, McVean G.** 2013. High-throughput microbial population genomics using the Cortex variation assembler. Bioinformatics **29**:275–276. http://dx.doi.org/10.1093/bioinformatics/bts673.

34. **Peirano G, Ahmed-Bentley J, Fuller J, Rubin JE, Pitout JD.** 2014. Travel-related carbapenemase-producing Gram-negative bacteria in Alberta, Canada: the first 3 years. J. Clin. Microbiol. **52**:1575–1581. http://dx.doi.org/10.1128/JCM.00162-14.

35. **Quinones D, Valverde A, Rodriguez-Banos M, Kobayashi N, Zayaz A, Abreu M, Canton R, del Campo R.** 2014. High clonal diversity in a non-outbreak situation of clinical ESBL-producing Klebsiella pneumoniae isolates in the first national surveillance program in Cuba. Microb. Drug Resist. **20**:45–51. http://dx.doi.org/10.1089/mdr.2013.0021.

36. **Toth A, Kocsis B, Damjanova I, Kristof K, Janvari L, Paszti J, Csercsik R, Topf J, Szabo D, Hamar P, Nagy K, Fuzi M.** 2014. Fitness cost associated with resistance to fluoroquinolones is diverse across clones of Klebsiella pneumoniae and may select for CTX-M-15 type extended-spectrum beta-lactamase. Eur. J. Clin. Microbiol. Infect. Dis. **33**:837–843. http://dx.doi.org/10.1007/s10096-013-2022-6.

37. **Villa L, Poirel L, Nordmann P, Carta C, Carattoli A.** 2012. Complete sequencing of an IncH plasmid carrying the blaNDM-1, blaCTX-M-15 and qnrB1 genes. J. Antimicrob. Chemother. **67**:1645–1650. http://dx.doi.org/10.1093/jac/dks114.

38. **Huang TW, Chen TL, Chen YT, Lauderdale TL, Liao TL, Lee YT, Chen CP, Liu YM, Lin AC, Chang YH, Wu KM, Kirby R, Lai JF, Tan MC, Siu LK, Chang CM, Fung CP, Tsai SF.** 2013. Copy number change of the NDM-1 sequence in a multidrug-resistant Klebsiella pneumoniae clinical isolate. PLoS One **8**:e62774. http://dx.doi.org/10.1371/journal.pone.0062774.

39. **Reeves PR, Liu B, Zhou Z, Li D, Guo D, Ren Y, Clabots C, Lan R, Johnson JR, Wang L.** 2011. Rates of mutation and host transmission for an Escherichia coli clone over 3 years. PLoS One **6**:e26907. http://dx.doi.org/10.1371/journal.pone.0026907.

40. **Maiden MC, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND.** 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat. Rev. Microbiol. **11**:728–736. http://dx.doi.org/10.1038/nrmicro3093.