

# Integrating whole-genome sequencing within the National Antimicrobial Resistance Surveillance Program in the Philippines

Silvia Argimón<sup>1</sup>, Melissa A. L. Masim<sup>2</sup>, June M. Gayeta<sup>2</sup>, Marietta L. Lagrada<sup>2</sup>, Polle K. V. Macaranas<sup>2</sup>, Victoria Cohen<sup>1</sup>, Marilyn T. Limas<sup>2</sup>, Holly O. Espiritu<sup>2</sup>, Janziel C. Palarca<sup>2</sup>, Jeremiah Chilam<sup>2</sup>, Manuel C. Jamoralin Jr.<sup>2</sup>, Alfred S. Villamin<sup>2</sup>, Janice B. Borlasa<sup>2</sup>, Agnetta M. Olorosa<sup>2</sup>, Lara F. T. Hernandez<sup>2</sup>, Karis D. Boehme<sup>2</sup>, Benjamin Jeffrey<sup>1</sup>, Khalil Abudahab<sup>1</sup>, Charmian M. Hufano<sup>2,6</sup>, Sonia B. Sia<sup>2</sup>, John Stelling<sup>3</sup>, Matthew T. G. Holden<sup>4</sup>, David M. Aanensen<sup>1,5,7</sup>✉ & Celia C. Carlos<sup>2,7</sup>✉

National networks of laboratory-based surveillance of antimicrobial resistance (AMR) monitor resistance trends and disseminate these data to AMR stakeholders. Whole-genome sequencing (WGS) can support surveillance by pinpointing resistance mechanisms and uncovering transmission patterns. However, genomic surveillance is rare in low- and middle-income countries. Here, we implement WGS within the established Antimicrobial Resistance Surveillance Program of the Philippines via a binational collaboration. In parallel, we characterize bacterial populations of key bug-drug combinations via a retrospective sequencing survey. By linking the resistance phenotypes to genomic data, we reveal the interplay of genetic lineages (strains), AMR mechanisms, and AMR vehicles underlying the expansion of specific resistance phenotypes that coincide with the growing carbapenem resistance rates observed since 2010. Our results enhance our understanding of the drivers of carbapenem resistance in the Philippines, while also serving as the genetic background to contextualize ongoing local prospective surveillance.

<sup>1</sup>Centre for Genomic Pathogen Surveillance, Wellcome Genome Campus, Hinxton, UK. <sup>2</sup>Antimicrobial Resistance Surveillance Reference Laboratory, Research Institute for Tropical Medicine, Muntinlupa, The Philippines. <sup>3</sup>Brigham and Women's Hospital, Boston, MA, USA. <sup>4</sup>University of St Andrews School of Medicine, St Andrews, Scotland, UK. <sup>5</sup>Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. <sup>6</sup>Present address: St. Luke's Medical Center—Global City, Taguig, Metro Manila, The Philippines. <sup>7</sup>These authors contributed equally: David M. Aanensen, Celia C. Carlos. ✉email: [david.aanensen@bdi.ox.ac.uk](mailto:david.aanensen@bdi.ox.ac.uk); [ccarlosph@yahoo.com](mailto:ccarlosph@yahoo.com)

Antimicrobial resistance (AMR) is an increasingly serious threat to global public health and the economy that requires concerted action across countries, government sectors and non-government organizations<sup>1</sup>. Without AMR containment, an adverse impact on medical costs, global gross domestic product, livestock production and international trade is expected by 2050, and the sustainable development goals for 2030 are less likely to be attained<sup>2</sup>. The Global Action Plan developed by the World Health Organization (WHO) to tackle AMR highlights the need to strengthen our understanding of how resistance develops and spreads, and the underlying resistance mechanisms<sup>3</sup>. One of the pillars of this objective is the national surveillance system for AMR.

An advanced example of a national surveillance system is the Antimicrobial Resistance Surveillance Program (ARSP), the national laboratory-based surveillance of the Philippine Department of Health (DOH). The ARSP began in 1988 with eight hospital-based laboratories (sentinel sites) and the Antimicrobial Resistance Surveillance Reference Laboratory (ARSRL) in Metro Manila and has since expanded to 24 sentinel sites and two gonorrhoea surveillance sites in 16 regions of the country (Supplementary Fig. 1). Surveillance encompasses common pathogens of public health importance in the Philippines, where infectious diseases represented nine out of the ten leading causes of morbidity<sup>4</sup>, and four out of the ten leading causes of infant mortality<sup>5</sup> in 2013–2014. Results of culture-based identification and antimicrobial susceptibility are stored centrally at the ARSRL within the clinical microbiology management software WHONET (Fig. 1a<sup>6</sup>). A report summarizing the resistance trends is published annually and disseminated to local, national and international surveillance stakeholders<sup>7</sup>. For policy change, surveillance data generated by ARSP are used by the DOH to develop clinical practice guidelines and determine the panel of antibiotics to be included in the national formulary. Furthermore, the ARSP has been contributing data to international AMR surveillance programmes since the 1980s, including the Global Antimicrobial Surveillance System (GLASS)<sup>8</sup> since 2015. Importantly, the ARSP data has informed the Philippine National Action Plan to combat AMR, one of the key requirements for alignment with the WHO Global Action Plan.

The Philippines has seen a steady increase in resistance rates for several key pathogen–antibiotic combinations in the last 10 years, including carbapenem-resistant organisms (Fig. 1b). The genetic mechanisms underlying carbapenem resistance, one of the biggest therapeutic challenges in the treatment of antimicrobial-resistant infections, include increased upregulation of efflux pumps, decreased uptake by altered expression/loss of porin function, and acquisition of hydrolytic enzymes—carbapenemases<sup>9</sup>. Whole-genome sequencing (WGS) of bacterial pathogens can identify distinct clonal lineages (strains) on phylogenetic trees, known AMR mechanisms (genes or mutations) and the vehicles for the acquired AMR genes (mobile genetic elements, MGEs) which, in turn, enables enhanced detection and characterization of high-risk clones (Fig. 1a). WGS is routinely used for infectious disease epidemiology in several high-income countries around the world, where it has improved outbreak investigations and epidemiological surveillance<sup>10,11</sup> and enhanced our knowledge of the spread of antimicrobial-resistant strains and their resistance mechanisms<sup>12,13</sup>. Elucidation of resistance mechanisms and their context can be critical for effective infection control. For example, upregulation of efflux pumps and loss of porin function by mutation are vertically transmitted, while acquired carbapenemases carried in transmissible plasmids or integrative conjugative elements have the potential for horizontal dissemination between strains and species, thus necessitating enhanced infection control measures<sup>14</sup>. Moreover, diverse

carbapenemase classes, variants, flanking MGEs and associated plasmid incompatibility groups have been described, requiring multiple biochemical and molecular tests for identification. For example, the class B New Delhi metallo-beta-lactamase (NDM) is found worldwide, represented by over ten different variants that are usually upstream of intact or truncated *ISAbal25* elements, within plasmid backbones of different incompatibility groups, across multiple lineages of *E. coli*—as well as of other carbapenem-resistant organisms<sup>15,16</sup>. WGS can determine the interplay between these different components, i.e. the gene, the vehicle and the strain, thus maximizing the epidemiological benefit derived from its cost.

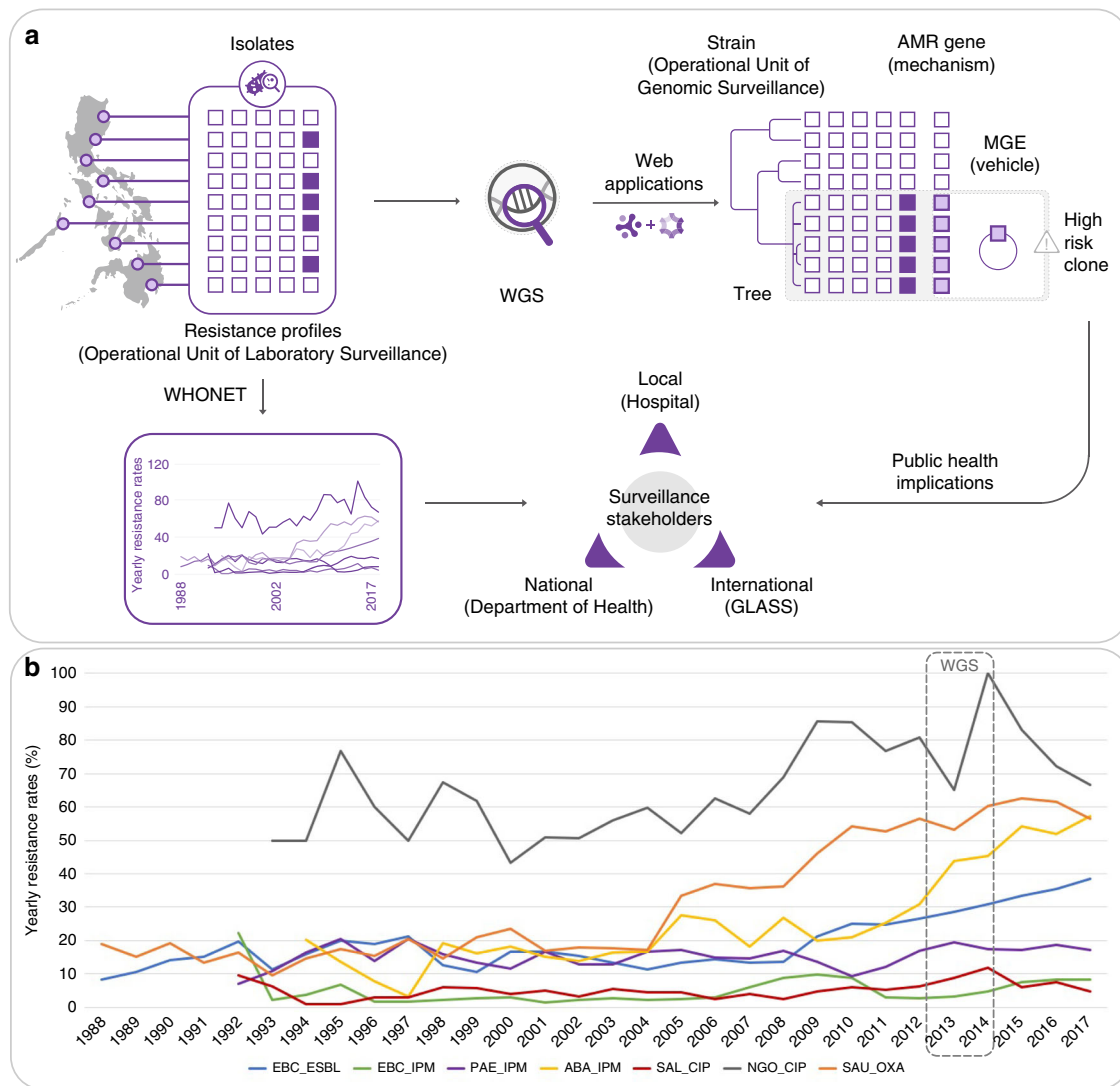
Implementation of WGS within existing national surveillance systems in low- and middle-income countries (LMICs) has the potential to enhance local prevention and control of resistant infections in a sustainable and equitable manner. Integration of WGS into routine surveillance of AMR can be facilitated by international collaboration focused on transfer of expertise and ownership. Our collaboration (See and Sequence, Fig. 1a) aimed to implement WGS within the ARSP via a multi-faceted approach that included a large initial retrospective sequencing survey, technology transfer, utilization of user-friendly web applications and capacity building in laboratory and bioinformatic procedures for local prospective sequencing. We linked what has traditionally been used as the operational unit of laboratory surveillance, the resistance profile (RP), with the operational unit of genomic epidemiology now provided by WGS, genetic relatedness. Here we provide exemplars from the retrospective sequencing survey that highlight how the granular view of strain-gene-vehicle in carbapenem-resistant populations at the local, national and global operational scales can be leveraged for surveillance of AMR and public health.

## Results

**Integration of laboratory and genomic surveillance.** The Philippine ARSP collects bacterial isolates and stores the associated clinical and epidemiological data in the WHONET software<sup>6</sup>. Antibiograms are stored as RPs, which represent the diversity of AMR phenotypes in the country. We implemented WGS within the ARSP via ongoing technology transfer, capacity building in and binational collaboration. In parallel, we conducted a large retrospective sequencing survey to characterize bacterial populations, dissect resistance phenotypes of key bug–drug combinations and provide genomic context for local prospective surveillance. Here, we focus on the carbapenem-resistant organisms, which have been classified as of critical priority for the development of new antibiotics by the WHO.

### Operational unit of laboratory surveillance—increasing the burden of specific resistance phenotypes.

Carbapenem-resistant *P. aeruginosa*, *A. baumannii*, *K. pneumoniae* and *E. coli* were all first isolated by the ARSP between 1992 and 1994. Since the year 2000, resistance rates to imipenem and meropenem have remained below 30% for all organisms except for *A. baumannii*, which has seen a steady rise in resistance rates between 2009 and 2017, reaching 56% (Fig. 2a). This coincides with the expansion of two RPs, first a striking expansion of true extremely drug-resistant (XDR) and possible pan drug-resistant (PDR) RP-1, and later of true-XDR RP-2 (full profiles can be found in Fig. 2b). Yearly resistance rates for *P. aeruginosa* have oscillated between 10 and 25% since the year 2000, and have doubled between 2010 and 2013 (Fig. 2a). Two RPs have seen the largest combined expansion between 2011 and 2013, possible-PDR RP-3 and true-XDR RP-4 (Fig. 2b). The yearly resistance rates to imipenem for Enterobacteriaceae were ~3% in 2011 but have since steadily



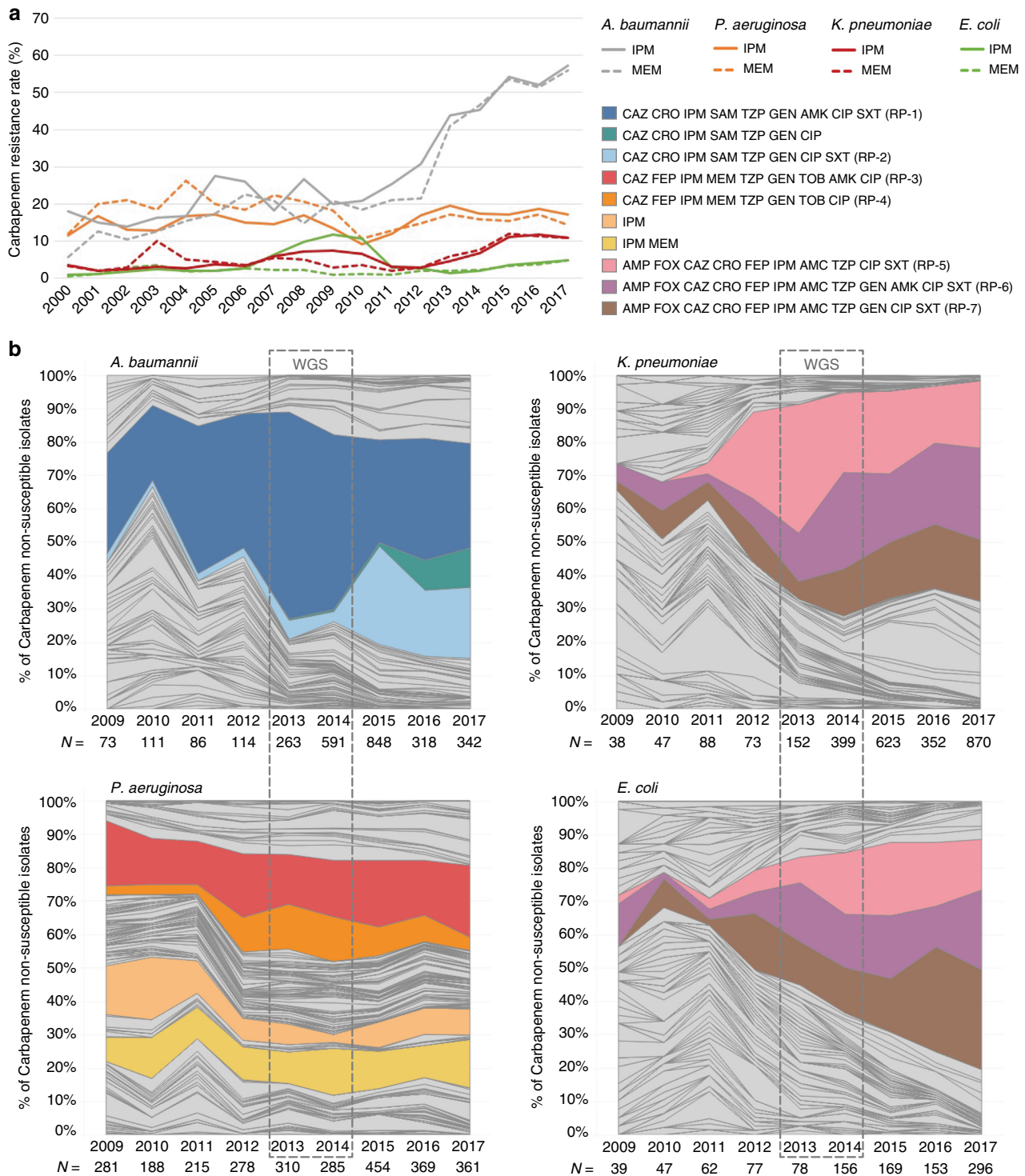
**Fig. 1** Implementing whole-genome sequencing (WGS) for AMR surveillance in the Philippines. **a** ARSP workflow and enhanced detection of high-risk clones by WGS. Isolates collected by sentinel sites are tested for susceptibility to antibiotics (open squares: susceptible, solid squares: resistant). The data are stored as resistance profiles in WHONET and summaries of resistance trends are shared yearly with surveillance stakeholders. WGS of bacterial isolates and interpretation with web applications like Microreact and Pathogenwatch provide information on genetic relatedness (strains), known AMR determinants (mechanisms), and the mobile genetic elements (MGE, vehicles) for their dissemination, thus allowing us to detect high-risk clones. **b** Detail of trends in antimicrobial resistance in the Philippines. Yearly resistance rates for key bug–drug combinations based on phenotypic data collected by sentinel sites. EBC Enterobacteriaceae (*K. pneumoniae*, *E. coli*, *Salmonella enterica*), PAE *P. aeruginosa*, ABA *A. baumannii* (*Acinetobacter* spp. before the year 2000), SAL *Salmonella enterica*, SAU *S. aureus*, ESBL extended-spectrum beta-lactamase production suspected, or non-susceptible to the following antibiotics IPM imipenem, CIP ciprofloxacin, OXA oxacillin, WGS box: period covered by the retrospective sequencing survey.

increased to 5% for *E. coli* and 11% for *K. pneumoniae*, respectively (Fig. 2a). This coincides with the expansion of three possible-XDR RPs, RP-5, RP-6 and RP-7 (Fig. 2b). The observed expansion of specific RPs could be driven by the rapid dissemination of a discrete genetic lineage (strain) carrying AMR determinants or, alternatively, of transmissible MGEs, which shuttle resistance determinants across different genetic lineages.

**Operational unit of genomic surveillance—genetic diversity and mechanisms underpinning carbapenem resistance phenotypes.** The sequences of 805 genomes linked to WHONET data were obtained from isolates belonging to the four bacterial pathogens in the WHO critical list (Table 1), which were collected in 2013 and 2014 by between 13 and 18 ARSP sites. The isolates sequenced were non-susceptible to carbapenems and/or to

extended-spectrum cephalosporins (Table 1), with the exception of *P. aeruginosa*, for which susceptible isolates were available and also sequenced.

The distribution of the carbapenem RPs highlighted in Fig. 2 was not concordant with the major clades observed on the phylogenetic trees (Fig. 3). Instead, the same RP was usually observed in multiple genetic lineages across the tree. Yet, the distribution of the multi-locus sequence types (STs), a molecular genotyping method inferred from the sequences of seven house-keeping genes, was largely concordant with the major clades observed on the trees inferred from whole-genome single-nucleotide polymorphisms (SNPs) (Fig. 3), confirming that the homoplasic distribution of RPs was not an artefact of the tree topology. The distribution of the same carbapenem RPs across multiple genetic lineages was observed both at the national and local levels (Supplementary Fig. 2), as the number of ST



**Fig. 2 Temporal dynamics of carbapenem resistance profiles.** **a** Yearly carbapenem resistance rates (IPM imipenem, MEM meropenem) for *A. baumannii*, *P. aeruginosa*, *K. pneumoniae* and *E. coli*. **b** Relative abundance of carbapenem resistance profiles (non-susceptible to imipenem and/or meropenem). The three-letter antibiotic codes (as in Table 1) indicate that the isolate is non-susceptible (resistant or intermediate). Only carbapenem non-susceptible isolates with complete susceptibility data were included, as indicated by the numbers under the x-axis (N). WGS box: period covered by the retrospective whole-genome sequencing survey.

assignments increased with the number of isolates representing a RP. Thus, the genomic data suggest that the expansion of specific carbapenem RPs in the Philippines is driven by the horizontal dissemination of resistance determinants across diverse genetic backgrounds, rather than by the sole expansion of single resistant genetic clones, and that resistance to multiple antibiotics is

acquired en bloc via MGEs, as previously observed for carbapenem-resistant organisms<sup>16</sup>.

Carbapenemases are a diverse collection of hydrolytic enzymes<sup>16</sup>, and we identified representatives of classes A, B and C carbapenemases in the Philippines (Supplementary Fig. 3). Class B NDM-1 and Verona integron-borne metallo-beta-



**Table 1** Organisms from the retrospective sequencing survey described in this study.

Organism (code)	Key resistance(s)	Antibiotics tested by ARSP	Genomes	Sentinel sites
<i>K. pneumoniae</i> (Kpn)	Carbapenems, ESBL-producing	AMP, FOX, CAZ, CRO, FEP, IPM, AMC, TZP, GEN, AMK, CIP, SXT	340	18
<i>E. coli</i> (Eco)	Carbapenems, ESBL-producing	AMP, FOX, CAZ, CRO, FEP, IPM, AMC, TZP, GEN, AMK, CIP, SXT	181	17
<i>P. aeruginosa</i> (Pae)	Carbapenems	CAZ, FEP, IPM, MEM, TZP, GEN, TOB, AMK, CIP	176	17
<i>A. baumannii</i> (Aba)	Carbapenems	CAZ, CRO, IPM, SAM, TZP, GEN, AMK, CIP, SXT	108	13

Key resistances were used to select the isolates for WGS. The number of draft genomes obtained and the number of ARSP sentinel sites they represent are indicated. Antibiotics tested by ARSP for each organism are abbreviated according to their WHONET codes: AMC amoxicillin/clavulanic acid, AMK amikacin, AMP ampicillin, CAZ ceftazidime, CIP ciprofloxacin, CRO ceftriaxone, GEN gentamicin, FEP cefepime, FOX ceftoxitin, IPM imipenem, MEM meropenem, SAM ampicillin/sulbactam, SXT trimethoprim/sulfamethoxazole, TOB tobramycin, TZP piperacillin/tazobactam.

lactamase VIM-2 were the most prevalent and geographically disseminated in the Enterobacteriaceae and *P. aeruginosa*, respectively, while the same was true for class D OXA-23 in *A. baumannii*. Importantly, different carbapenemase genes/variants (or combinations thereof) were found underlying the same RP in all four organisms (Fig. 3).

**Local scale of surveillance—plasmid-driven hospital outbreak of carbapenem-resistant *K. pneumoniae*.** Fifty-seven percent ( $N = 33$ ) of the *K. pneumoniae* genomes with the possible-XDR resistance phenotype AMP FOX CAZ CRO FEP IPM AMC TZP GEN AMK CIP SXT (RP-6) were isolated from a single hospital (MMH) but belonged to 12 different STs, with almost half of them ( $N = 15$ ) placed within ST340 (Supplementary Fig. 4a). The phylogenetic analysis of these 15 genomes in the wider context of ST340 showed three main lineages, with the 15 possible-XDR isolates from MMH forming a tight cluster (clade III, Fig. 4a). The shorter branch lengths in the imipenem-resistant clade III ( $6.5 \pm 3.7$  pairwise SNP differences) compared with the imipenem-susceptible clade II ( $26.8 \pm 6.9$  pairwise SNP differences, Mann–Whitney  $U$  test  $z$ -score  $-6.265$ ,  $p$  value  $= 3.71 \times 10^{-10}$ , Supplementary Fig. 4b), and the isolation dates spanning 12 months, suggest a rapid expansion of this clone, which coincides with the acquisition of the *bla*<sub>NDM-1</sub> gene (Fig. 4a).

Epidemiological data showed that the samples were mostly from hospital-acquired infections (HAI,  $N = 14$ ) and from neonates ( $N = 12$ ) (Fig. 4a). These observations triggered a retrospective epidemiological investigation that revealed that ten of the isolates originated from patients of the neonatal intensive care unit (NICU), all of which had umbilical catheters and were on mechanical ventilators. The remaining cases were either from paediatric wards ( $N = 4$ ) or the male ward ( $N = 1$ ), suggesting wider dissemination of this high-risk clone within the hospital environment.

The genetic diversity underlying the 33 possible-XDR isolates from MMH (12 STs) prompted us to investigate the hypothesis of dissemination of carbapenem resistance in the hospital by a plasmid carrying *bla*<sub>NDM-1</sub>. We identified a 101,540 bp IncFII plasmid with *bla*<sub>NDM-1</sub>, *rmtC* and *sul1* (p13ARS\_MMH0112-3, Table 2, Supplementary Note 1 and Supplementary Fig. 5a, b) in a representative isolate from ST340 clade III. We found the entire p13ARS\_MMH0112-3 plasmid sequence (i.e.,  $\geq 95\%$  of the length) represented in 27 genomes from 9 different STs, including 20 samples from 6 different STs isolated from patients under 1 year old (Fig. 4b). A second plasmid identified in the representative isolate carried 13 AMR genes, including the extended-spectrum beta-lactamase gene *bla*<sub>CTX-M-15</sub> (p13ARS\_MMH0112-2, Table 2, Supplementary Note 2 and Supplementary Fig. 5a, c).

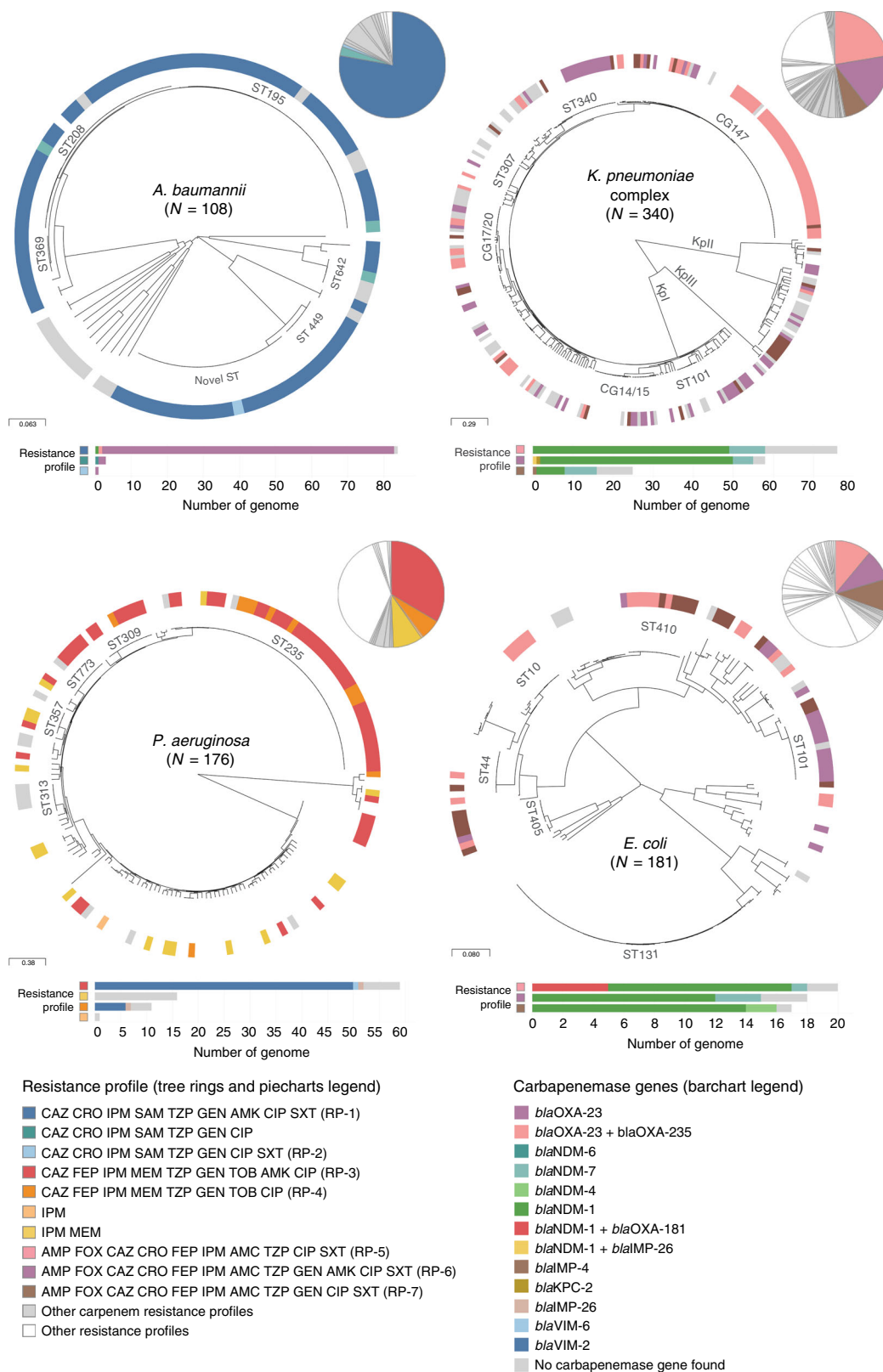
Altogether, our findings suggest that the burden of carbapenem-resistant *K. pneumoniae* infections in hospital MMH was largely linked to plasmid p13ARS\_MMH0112-3 with *bla*<sub>NDM-1</sub>, circulating within diverse genetic lineages, and leading to outbreaks in

high-risk patient populations. Hospital authorities were informed of these findings, and measures for infection control were implemented, including designating a separate multi-drug resistance organism room for cohorting, active surveillance upon identification of any new carbapenem-resistant *K. pneumoniae* from the NICU, and referral of any new carbapenem-resistant *K. pneumoniae* from the NICU to ARSRL for sequencing.

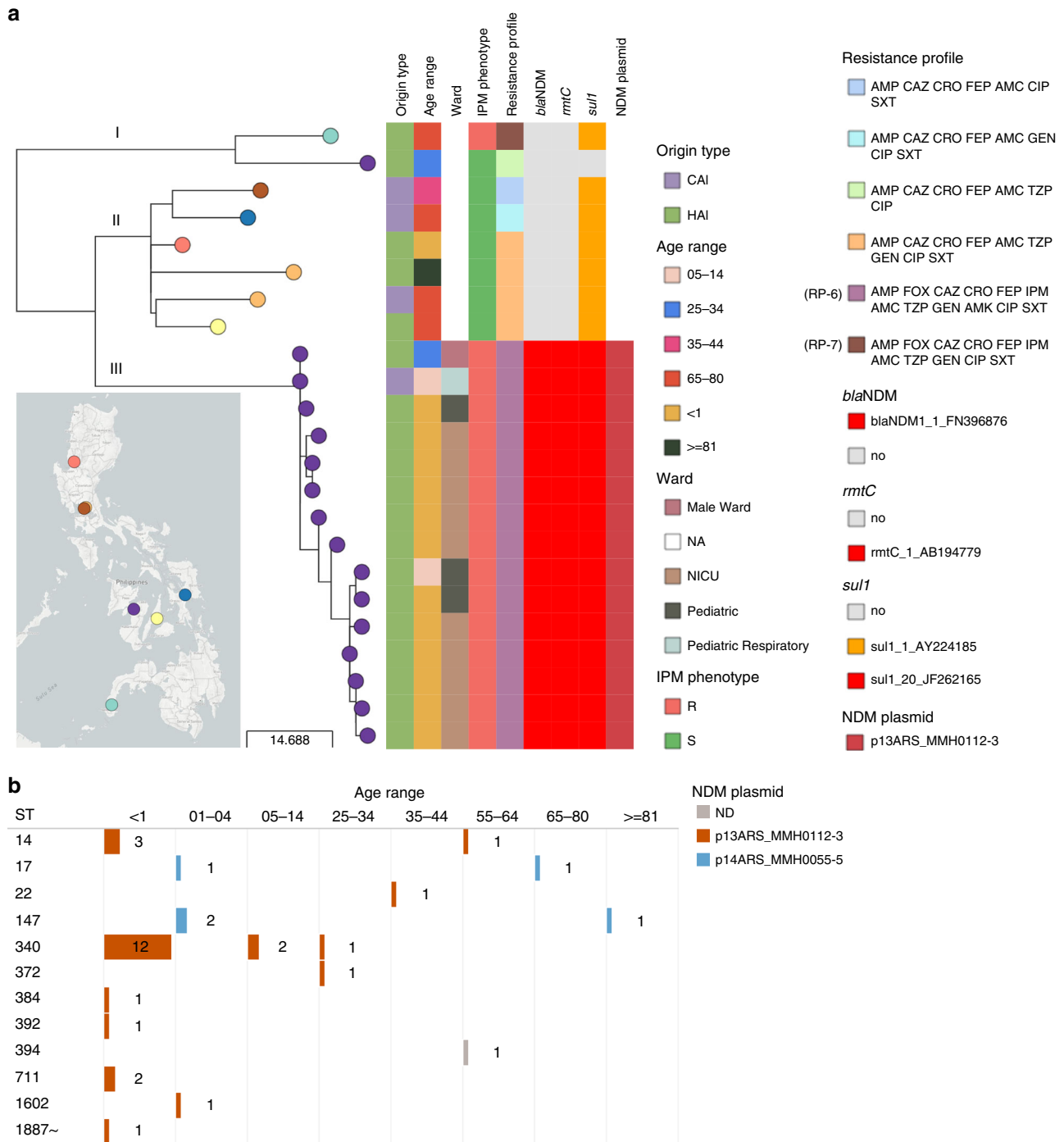
**National scale of surveillance—regional circulation of a successful *K. pneumoniae* lineage.** The possible-XDR resistance phenotype AMP FOX CAZ CRO FEP IPM AMC TZP CIP SXT (RP-5) was represented by 76 *K. pneumoniae* isolates from 14 different STs. Seventy-one percent ( $N = 54$ ) belonged to ST147, an international epidemic clone that was found in 11 sentinel sites in all three island groups in the Philippines. The phylogeny of the ST147 genomes showed that carbapenem non-susceptible isolates (78.8%,  $N = 63$ ) were found in clades arising from three out of four deep branches of the tree (Fig. 5), which represent separate groups of the global population (Supplementary Note 3 and Supplementary Fig. 6). Carbapenem resistance was largely explained by the presence of *bla*<sub>NDM-1</sub> in clade IV genomes, while clade III-B showed geographically distinct clusters with either *bla*<sub>NDM-1</sub> or *bla*<sub>NDM-7</sub> (Fig. 5, Supplementary Note 4 and Supplementary Fig. 7).

Five plasmid sequences obtained from five different isolates representing carbapenem-resistant clusters in clades III and IV showed that the NDM genes were carried within different variants of the insertion sequence IS*Aba125* (Supplementary Fig. 8a) on different plasmid backbones, two of which showed high sequence similarity to international plasmids (Table 3 and Supplementary Figs. 5b and 8b). The distribution of plasmids harbouring *bla*<sub>NDM-1</sub> and *bla*<sub>NDM-7</sub> matched the strong phylogeographic signal in the terminal branches of the tree (Fig. 5). Within clade III-B, a cluster of genomes from sentinel sites MMH and STU was characterized by plasmid p14ARS\_MMH0055-5 carrying *bla*<sub>NDM-7</sub>, while another cluster of genomes from VSM and NMC was distinguished by plasmid p13ARS\_VSM0593-1 with *bla*<sub>NDM-1</sub>. Plasmids p13ARS\_GMH0099 and p14ARS\_VSM0843-1, both carrying *bla*<sub>NDM-1</sub>, were found in different subclades of clade IV, one of local circulation in hospital VSM, and another one showing regional expansion across five sentinel sites (Fig. 5). Taken together, the phylogeographic signal (strains) in combination with the distribution of *bla*<sub>NDM</sub> variants (mechanism) and plasmids (vehicle) revealed local and regional patterns of circulation of carbapenem non-susceptible ST147 within the Philippines, which could not be identified based solely on the RPs and MLST information. Furthermore, comparison with global ST147 genomes indicated that clade IV may be unique to the Philippines (Supplementary Fig. 6).

**International scale of surveillance—high-risk clone of *E. coli* ST410 with NDM-1 and OXA-181.** Recent reports of *E. coli*



**Fig. 3 Carbapenem resistance profiles are not associated with specific genetic lineages.** Phylogenetic trees showing the major lineages of *A. baumannii*, *P. aeruginosa*, *K. pneumoniae* species complex and *E. coli*, indicating the position of select sequence types (ST) and clonal groups (CG). KpI *K. pneumoniae* sensu stricto, KpII *K. quasipneumoniae*, KpIII *K. variicola*. Tree ring: select carbapenem resistance profiles from Fig. 2 are shown in colour. The remaining carbapenem resistance profiles are shown in light grey. Other resistance profiles (susceptible to carbapenems) are shown in white for simplicity. The three-letter antibiotic codes are as in Table 1. The pie charts show the relative abundance of the resistance profiles in the retrospective collection of sequenced genomes. The bar charts show the distribution of carbapenemase genes across the key resistance profiles. The scale bars show the number of SNPs per variable site. The data are available at [https://microreact.org/project/ARSP\\_ABA\\_2013-14](https://microreact.org/project/ARSP_ABA_2013-14) (Aba), [https://microreact.org/project/ARSP\\_PAE\\_2013-14](https://microreact.org/project/ARSP_PAE_2013-14) (Pae), [https://microreact.org/project/ARSP\\_KPN\\_2013-14](https://microreact.org/project/ARSP_KPN_2013-14) (Kpn) and [https://microreact.org/project/ARSP\\_ECO\\_2013-14](https://microreact.org/project/ARSP_ECO_2013-14) (Eco).



**Fig. 4** Detection of a plasmid-driven outbreak of *K. pneumoniae* ST340. **a** Phylogenetic tree and linked epidemiological and genotypic data of 24 retrospective ST340 genomes. Origin type defined as either community-acquired infection (CAI) or hospital-acquired infection (HAI). NA ward information not available. The imipenem (IPM) phenotype was either resistant (R) or susceptible (S). The three-letter antibiotic codes in the resistance profiles are as in Table 1. This interactive view is available at [https://microreact.org/project/ARSP\\_KPN\\_ST340\\_2013-14/ac2a0920](https://microreact.org/project/ARSP_KPN_ST340_2013-14/ac2a0920). The maximum-likelihood tree was inferred from 196 SNP positions identified by mapping the genomes to reference CAV1217 (CP018676.1 [<https://www.ncbi.nlm.nih.gov/nucleotide/CP018676.1>]), and masking regions corresponding to mobile genetic elements and recombination. The scale bar shows the number of SNPs per variable site. The full data are available at [https://microreact.org/project/ARSP\\_KPN\\_ST340\\_2013-14](https://microreact.org/project/ARSP_KPN_ST340_2013-14). **b** Distribution of 33 isolates from hospital MMH with resistance profile RP-6 by patient age group, sequence types (ST) and NDM plasmid. Short reads of the 33 isolates were mapped to the plasmid sequences and a match was counted when the reads covered at least 95% of the sequence length with at least 5× depth of coverage. ND not detected.

ST410 international high-risk clones carrying class D and class B carbapenemases<sup>17,18</sup> are particularly alarming in light of their wide geographic distribution and the broad variety of niches this clone can occupy<sup>19,20</sup>. ST410 was the second most prevalent ST in the retrospective collection of *E. coli* (13.2%, *N* = 24),

encompassing a large proportion of imipenem-resistant isolates (*N* = 15), and the three expanding possible-XDR RP-5,6,7 (Supplementary Fig. 9). The phylogenetic tree of 24 ST410 genomes showed that isolates with the possible-XDR profiles clustered within one clade (Fig. 6a), which can be further delineated by the

**Table 2 Plasmids from *K. pneumoniae* ST340 strain 13ARS\_MMH012.**

Plasmid	Size (bp)	Replicons	AMR genes	Match accession (size)	Accession
p13ARS_MMH012-2	243,650	IncFIB(K), IncFII(K)	<i>bla</i> <sub>CTX-M-15</sub> , <i>bla</i> <sub>TEM-1B</sub> , <i>bla</i> <sub>OXA-1</sub> , <i>aac</i> (3)-IIa, <i>aac</i> (6')-Ib-cr, <i>aph</i> (6)-Ia, <i>aph</i> (3'')-Ib, <i>aph</i> (3')-Ia, <i>aac</i> (6')-Ib-cr, <i>qnrB1</i> , <i>mph</i> (A), <i>sul2</i> , <i>tet</i> (A)	NA	LR697123
p13ARS_MMH012-3	101,540	IncFII(Yp)	<i>bla</i> <sub>NDM-1</sub> , <i>rmtC</i> , <i>sul1</i>	MG462729.1 (110,787 bp)	LR697124

The replicons and AMR genes were identified with the PlasmidFinder and ResFinder databases, respectively. The match accession indicates the accession code when at least one match of more than 90% query sequence coverage and 99% identity was found in the NCBI nucleotide database. NA not available.

distribution of carbapenemase genes and variants into three different clones carrying *bla*<sub>NDM-4</sub> ( $N = 1$ ), *bla*<sub>NDM-1</sub> ( $N = 8$ ) or both *bla*<sub>NDM-1</sub> and *bla*<sub>OXA-181</sub> ( $N = 5$ ). The phylogenetic analysis with global *E. coli* ST410 genomes confirmed that these are independent clones, as they were found interspersed with international genomes within a major high-risk lineage (Fig. 6b and Supplementary Note 5).

*K. pneumoniae* carrying *bla*<sub>NDM-1</sub> and *bla*<sub>OXA-181</sub> have been previously reported<sup>21,22</sup>, as well as *E. coli* ST410 harbouring *bla*<sub>NDM-5</sub> and *bla*<sub>OXA-181</sub> from Egypt<sup>23</sup>, Denmark and the UK<sup>18</sup>, but, to our knowledge, this is the first report of *E. coli* carrying both *bla*<sub>NDM-1</sub> and *bla*<sub>OXA-181</sub>, which is likely to have disseminated between two sentinel sites (NMC and VSM). We identified five plasmids in the representative strain 14ARS\_NMC0074 (Table 4 and Supplementary Fig. 10). The IncX3 plasmid carrying *bla*<sub>OXA-181</sub> (p14ARS\_NMC0074-5) was identical to plasmid pAMA1167-OXA-181 isolated previously from an *E. coli* ST410 strain<sup>24</sup> (Supplementary Fig. 10). Mapping short reads to p14ARS\_NMC0074-5 showed that this plasmid is the main vehicle of *bla*<sub>OXA-181</sub> in the Philippines, as in the international genomes (Fig. 6b<sup>18</sup>). However, we did not identify any known plasmids similar to the IncA/C2 plasmid with *bla*<sub>NDM-1</sub> (p14ARS\_NMC0074-2), though it shared ~90% of its backbone with the IncA/C2 plasmid described above from *K. pneumoniae* ST147 strain 13ARS\_VSM0593 (Supplementary Fig. 10). This plasmid backbone was found in *E. coli* ST410 genomes from the Philippines, but not in international genomes (Fig. 6b).

Altogether, our results show that the Philippine *E. coli* ST410 genomes represent diverse lineages of the global circulating population, with evidence of frequent international introductions. These lineages are characterized by a diverse repertoire of carbapenemase genes and variants, amassed via a combination of conserved international plasmids and locally circulating plasmids.

## Discussion

National networks of laboratory-based surveillance are a key pillar within the Global Action Plan to combat AMR, with the reference laboratory playing a central role in the dissemination of surveillance data to local, national and international stakeholders. The Philippines ARSP surveillance data have shown increasing resistance trends for key bug–drug combinations (Figs. 1b and 2a). In this study, the analysis of the dynamics of RPs showed the concomitant expansion of specific resistance phenotypes (Fig. 2b). By complementing laboratory data with WGS and linking the operational units of laboratory and genomic surveillance, we revealed a diversity of genetic lineages, AMR mechanisms and vehicles underlying the expansion of carbapenem resistance phenotypes (Figs. 1a and 3). The combination of these three components vastly improved our understanding of the drivers of carbapenem resistance in the Philippines at different geographical scales, from an plasmid-driven local outbreak of *K. pneumoniae* ST340 (Fig. 4), to the expansion of a *K. pneumoniae*

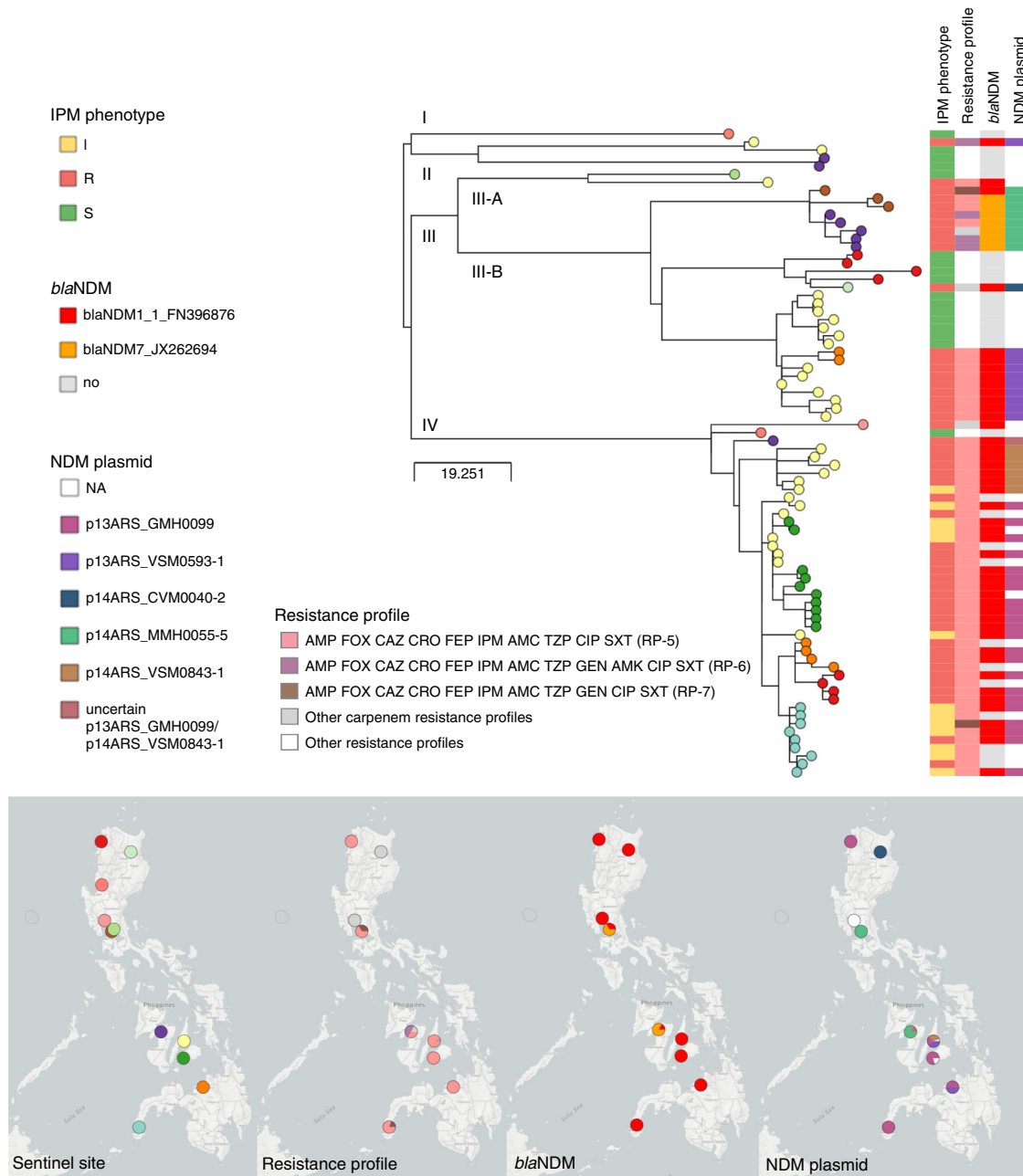
ST147 clone carrying NDM-1 in an IncA/C2 plasmid without any close match in international genomes (Fig. 5), to the independent introductions of international epidemic clone *E. coli* ST410 that can acquire plasmids of local circulation (Fig. 6).

A crucial aspect of AMR surveillance is the detection of the emergence and spread of high-risk clones. In traditional laboratory surveillance, RPs can be analyzed with spatiotemporal algorithms to detect hospital outbreaks<sup>25</sup> or dissemination of phenotypic subpopulations between hospitals<sup>26</sup>. Cluster detection based on RPs depends on consistent and complete antibiograms within and across sentinel sites. During the course of our project it became apparent that there were inconsistencies in the panels of antibiotics tested by the sentinel sites (missing data or different antibiotics tested), which led to the exclusion of samples from the analysis of the dynamics of RPs (Fig. 2). These inconsistencies would also hamper cluster detection based on RPs, or the early detection of novel emerging RPs that could act as indicators of novel high-risk clones. The joint discussions of these observations by the two collaborating teams led to an effort to reinforce the standardization of antibiotic testing across the ARSP sentinel sites, which was coordinated by the ARSRL.

Even with complete and comprehensive susceptibility testing, RPs can only provide a coarse view of the spread of AMR, and sometimes lead to clusters of disparate genetic relatedness (Fig. 3 and Supplementary Fig. 2). Integration of WGS into laboratory-based AMR surveillance can substantially improve the detection of high-risk clones by providing a high-resolution picture of genetic lineages (strains), AMR mechanisms (genes/mutations) and vehicles (MGEs, Fig. 1a). We identified high-risk clones within known international epidemic clones at the local, national and international scales by linking clonal relatedness, geographic clustering, epidemiological data and gene and plasmid content with interactive web tools<sup>27</sup>.

At the local scale, we identified a plasmid-driven NICU outbreak of carbapenem-resistant *K. pneumoniae*. The epidemiological data captured in WHONET was key to support the interpretation of the genomic findings, triggering a retrospective investigation. This previously undetected outbreak was traced to an IncFII plasmid carrying *bla*<sub>NDM-1</sub> in the genetic context of ST340 (Fig. 4a). Yet, the endemic IncFII plasmid was found across multiple wards, and genetic backgrounds (STs, Fig. 4b), indicating a major role in the burden of carbapenem resistance *K. pneumoniae* in this hospital. A second IncFIB–IncFII plasmid found only within the ST340 lineage might have contributed to the persistence and transmission of this clone in the NICU, by carrying several genomic islands with a role in survival in the host and in the environment (Supplementary Fig. 5). ST340 is a member of the drug-resistant clonal complex 258 and has been reported to cause outbreaks worldwide<sup>28</sup>. Our findings were disseminated to the local stakeholders (i.e., hospital) via a forum with NICU staff, paediatricians and ARSRL representatives. Infection control strategies were implemented, which ultimately bolstered the infection control team of this hospital. Control of healthcare-associated infections is crucial to containing





**Fig. 5 WGS dissects the circulation of *K. pneumoniae* ST147 in the Philippines.** Phylogenetic tree and linked epidemiological and genotypic data of 80 ST147 genomes. The imipenem (IPM) phenotype was either resistant (R), intermediate (I) or susceptible (S). The three-letter antibiotic codes in the resistance profiles are as in Table 1. This interactive view is available at [https://microreact.org/project/ARSP\\_KPN\\_ST147\\_2013-14/4c33ace7](https://microreact.org/project/ARSP_KPN_ST147_2013-14/4c33ace7). The maximum-likelihood tree was inferred from 809 SNP positions identified by mapping the genomes to reference MS6671 (LN824133.1 [<https://www.ncbi.nlm.nih.gov/nucore/LN824133.1>]) and masking regions corresponding to mobile genetic elements and recombination. The scale bar shows the number of SNPs per variable site. The distribution of plasmids with NDM genes was inferred by mapping the short reads of the genomes to the complete plasmid sequences, and a match was counted when the reads covered at least 95% of the sequence length with at least 5x depth of coverage. The full data are available at [https://microreact.org/project/ARSP\\_KPN\\_ST147\\_2013-14](https://microreact.org/project/ARSP_KPN_ST147_2013-14).

the spread of AMR<sup>29</sup>. The roadmap from sequence data to actionable data for infection control and prevention within a hospital has been mapped by multiple use cases<sup>30,31</sup> and supported by recent studies on implementation and cost-effectiveness in high-resource settings<sup>32,33</sup>. This study makes a case for extending this roadmap to LMICs that have infection control capacity in place.

At the national scale, the combined information on strain, carbapenemase gene/variant and vehicle shows a granular picture of carbapenem-resistant *K. pneumoniae* ST147 that uncovers the

geographic distribution of high-risk clones. ST147 is an international epidemic clone that causes carbapenem-resistant infections mediated by NDM, OXA-48-like and VIM carbapenemases<sup>28,34</sup>. Notably, ST147 was not reported in a recent study of *K. pneumoniae* from seven healthcare facilities across South and South East Asia<sup>35</sup>, which did not include the Philippines. Within the seemingly well-established genetic background of clade IV (Fig. 5), a high-risk clone characterized by the presence of bla<sub>NDM-1</sub> within plasmid p14ARS\_VSM0843-1 displayed local clonal expansion in one sentinel site, while another high-risk

**Table 3 Plasmids carrying NDM genes in *K. pneumoniae* ST147 representative isolates.**

Plasmid	Size (bp)	Replicons	AMR genes	Match accession (size)	Accession
p14ARS_CVM0040-2	110,787	IncFII(Yp)	<i>bla</i> <sub>NDM-1</sub> , <i>rmtC</i> , <i>sul1</i>	MG462729.1 (110,787 bp)	LR697128
p13ARS_GMH0099	273,158	IncA/C2, IncR, IncFIA(HI1)	<i>bla</i> <sub>NDM-1</sub> , <i>bla</i> <sub>CTX-M-15</sub> , <i>bla</i> <sub>OXA-1</sub> , <i>bla</i> <sub>OXA-10</sub> , <i>bla</i> <sub>TEM-1B</sub> , <i>aph</i> (3')-VI, <i>aph</i> (3')-Ia, <i>aadA1</i> (x2), <i>aac</i> (6')-Ib-cr, <i>aph</i> (6)-Ia, <i>strA</i> , <i>strB</i> , <i>qnrS1</i> , <i>mph</i> (A), <i>erm</i> (42), <i>ere</i> (A), <i>floR</i> , <i>cmlA1</i> (x2), <i>arr-2</i> (x2), <i>sul1</i> (x3), <i>sul2</i> , <i>dfrA14</i>	NA	LR697099
p13ARS_VSM0593-1	181,414	IncA/C2	<i>bla</i> <sub>NDM-1</sub> , <i>aph</i> (3')-Ia, <i>mph</i> (A), <i>cmlA1</i> , <i>floR</i> , <i>sul1</i> , <i>sul2</i> , <i>qnrS1</i> , <i>arr-3</i> ,	NA	LR697125.1
p14ARS_MMH0055-5	44,885	IncX3	<i>bla</i> <sub>NDM-7</sub>	KP776609.1 (45,122 bp)	LR697126
p14ARS_VSM0843-1	95,284	IncR, IncFIA(HI1)	<i>bla</i> <sub>NDM-1</sub> , <i>bla</i> <sub>CTX-M-15</sub> , <i>bla</i> <sub>TEM-1B</sub> , <i>aph</i> (3')-VI, <i>aph</i> (6)-Ia, <i>aph</i> (3')-Ib, <i>aph</i> (3')-VI, <i>aadA1</i> , <i>cmlA1</i> , <i>sul1</i> , <i>sul2</i> , <i>arr-3</i>	NA	LR697132

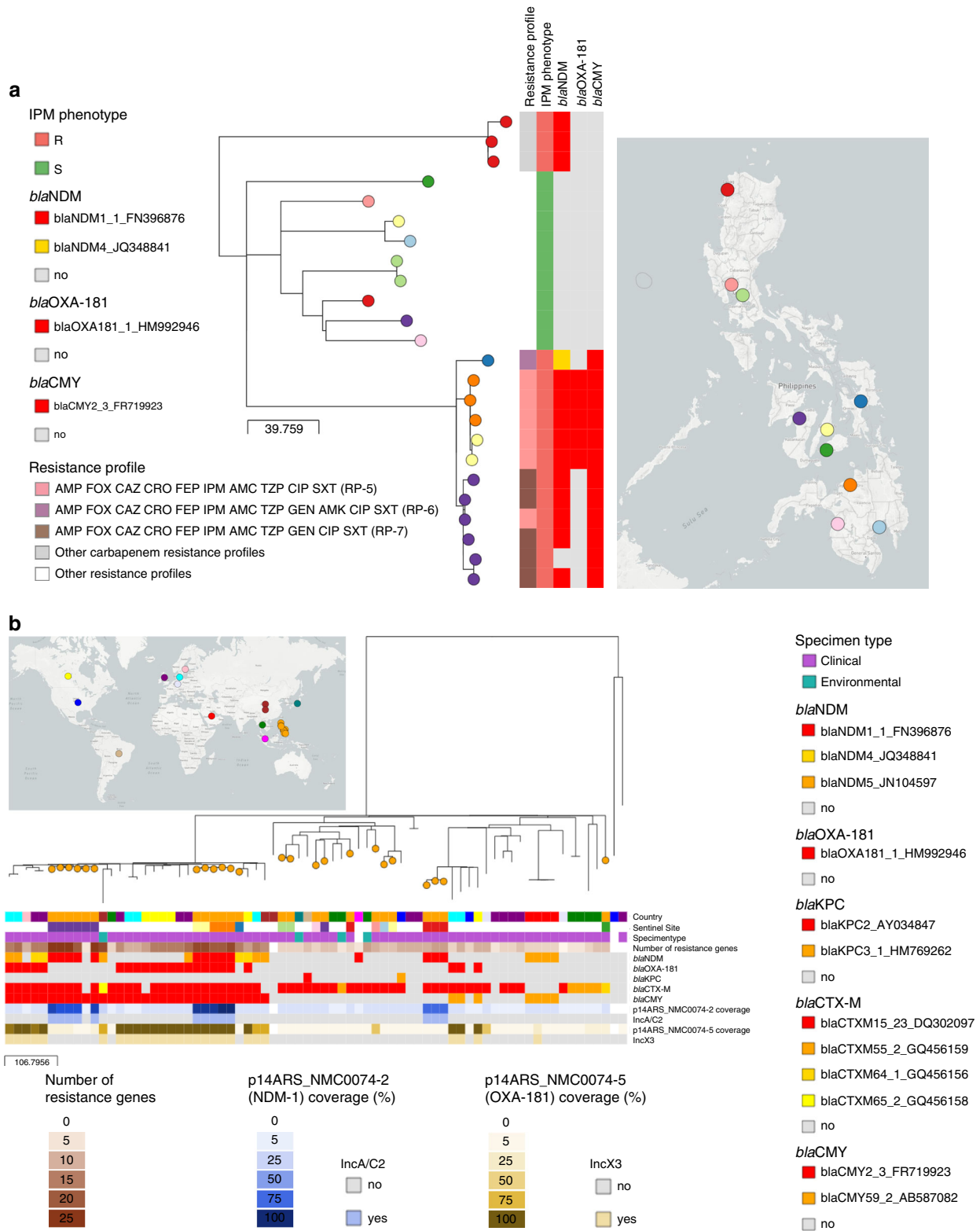
The replicons and AMR genes were identified with the PlasmidFinder and ResFinder databases, respectively. The match accession indicates the accession code when at least one match of more than 90% query sequence coverage and 99% identity was found in the NCBI nucleotide database. NA not available.

clone characterized by the presence of *bla*<sub>NDM-1</sub> within plasmid p13ARS\_GMH0099 had also disseminated geographically across five different sites (Fig. 5). Clonal expansion followed by geographical dissemination is usually attributed to the acquisition of an AMR determinant<sup>36</sup>. The extent of the geographical dissemination of the different high-risk clones could be attributed to the different plasmid backbones (vehicle), but differences in the genetic background (strain) could also be at play. The presence of a robust AMR surveillance network such as the ARSP is key for the detection of geographical dissemination of high-risk clones at a regional/national level, and for the dissemination of the information to national stakeholders. In this respect, the ARSRL is currently evaluating formats for the inclusion of genomic data into their annual surveillance reports. The reference laboratory directors can also alert hospitals across the country to establish concerted infection control measures, as well as relay this information to the national DOH for the formulation of evidence-based guidelines.

At the global scale, we identified several high-risk clones of *E. coli* ST410 circulating in the Philippines, with the evidence of frequent international transmission, in agreement with a previous report<sup>18</sup>. Previous reports of carbapenem-resistant *E. coli* ST410 from the Western Pacific and South East Asia regions described the presence of *bla*<sub>OXA-181</sub> in China<sup>17,37</sup>, *bla*<sub>NDM-1</sub> in Singapore<sup>38</sup>, *bla*<sub>NDM-5</sub> in South Korea<sup>39</sup>, or the combination of *bla*<sub>NDM-5</sub> and *bla*<sub>OXA-181</sub> in Myanmar<sup>40</sup> and South Korea<sup>39</sup>. The repertoire of carbapenemases in the retrospective collection of ST410 from the Philippines seemed to be in tune with the genes/variants circulating locally, as most isolates carried the prevalent variant *bla*<sub>NDM-1</sub>. Of note was the high-risk clone carrying both *bla*<sub>NDM-1</sub> and *bla*<sub>OXA-181</sub>, a combination hitherto unreported in *E. coli* ST410, and acquired through the combination of a plasmid of local circulation and a conserved international plasmid, respectively. This suggests that *E. coli* ST410, a successful pandemic lineage, can not only easily disseminate, but it can also adjust the complement of carbapenemases by acquiring endemic plasmids along the way. Global AMR surveillance networks, such as the WHO GLASS<sup>8</sup>, are paramount to detect the emergence and monitor the spread of resistance at the international level and inform the implementation of targeted prevention and control programmes. The ARSRL has been supplying laboratory surveillance data to GLASS since 2015, and presented their experience with WGS at a WHO technical consultation on the application of WGS for national, regional and global surveillance in 2020, thus becoming an international contributor in both fields.

Collecting information on AMR that can be rapidly transformed into action requires harmonized standards, especially at the national and international levels<sup>8</sup>. For laboratory data, WHONET serves this purpose in the ARSP and in over 2000 clinical, public health, veterinary and food laboratories in over 120 countries worldwide<sup>26</sup>, while GLASS serves this purpose at the global level. Genomic data are amenable to standardization<sup>29</sup>, and national public health agencies<sup>41</sup> and international surveillance networks<sup>42</sup> are adopting diverse schema to identify and name genetic lineages. However, a global standard system for defining a cluster has not been implemented beyond the level of discrimination of MLST. Likewise, different databases of AMR mechanisms<sup>43–45</sup> may differ in content and nomenclature. Thus, the standardization of genomic data, as well as the provision of platforms, for the uptake of genomic information is crucially moving forward, and for ongoing AMR surveillance.

Containment of AMR at a global level requires an international concerted effort. High-risk clones have the propensity to disseminate rapidly, and genomics can improve the detection of their emergence and spread. This highlights the increasing need to build equitable partnerships to facilitate ownership transfer of genomic epidemiology capacity (operational, analytical and interpretational) to enhance national AMR surveillance programmes within low-resource settings. The binational partnership of this project led to a large retrospective survey of bacterial pathogens that has provided the first in-depth genomic view of the AMR landscape in the Philippines and established contextual data for ongoing local prospective sequencing. In parallel, through training and transfer of expertise in laboratory procedures and bioinformatics, the open exchange of data, and collective interpretation of results, we have expanded the existing capacity of a national reference laboratory with WGS focused on action for public health. WGS commenced at the ARSRL in 2018 with the Illumina MiSeq equipment available locally. A new dedicated bioinformatics server was installed at ARSRL for sequence data storage and analysis. Collective data interpretation was aided by data sharing via interactive web tools such as Microreact ([www.microreact.org](http://www.microreact.org))<sup>27</sup> and Pathogenwatch ([www.pathogen.watch](http://www.pathogen.watch)) to leverage the expertise from both CGPS and ARSRL. The results of the retrospective survey were presented to the representatives of the sentinel sites during the ARSP annual meetings and at international conferences, at first by the members of CGPS, followed by presentations by the members of the ARSRL in subsequent years<sup>46,47</sup>. The ARSRL staff conducted its first outbreak investigation using WGS in July 2019, with



**Fig. 6 Phylogeographic analysis of *E. coli* ST410 from the Philippines.** **a** Phylogenetic tree and linked epidemiological and genotypic data of 24 retrospective ST410 genomes. The imipenem (IPM) phenotype was either resistant (R) or susceptible (S). The three-letter antibiotic codes in the resistance profiles are as in Table 1. This interactive view is available at [https://microreact.org/project/ARSP\\_ECO\\_ST410/088ba65b](https://microreact.org/project/ARSP_ECO_ST410/088ba65b). **b** Philippine isolates (orange nodes) in global context. This interactive view is available at [https://microreact.org/project/ARSP\\_ECO\\_ST410\\_GLOBAL/c701506e](https://microreact.org/project/ARSP_ECO_ST410_GLOBAL/c701506e). The maximum-likelihood trees were inferred from 703 (A) and 2851 (B) SNP positions, respectively, identified by mapping the genomes to reference AMA1167 (CP024801.1) and masking regions corresponding to mobile genetic elements and recombination. The scale bars show the number of SNPs per variable site. The distribution of plasmids with carbapenemases genes in (b) was inferred by mapping the short reads of the genomes to the complete plasmid sequences, and a match was counted when the reads covered at least 95% of the sequence length with at least 5× depth of coverage. The full data are available at [https://microreact.org/project/ARSP\\_ECO\\_ST410](https://microreact.org/project/ARSP_ECO_ST410) (a) and [https://microreact.org/project/ARSP\\_ECO\\_ST410\\_GLOBAL](https://microreact.org/project/ARSP_ECO_ST410_GLOBAL) (b).

**Table 4 Plasmid repertoire of carbapenem-resistant *E. coli* strain 14ARS\_NMC0074.**

Plasmid	Size (bp)	Replicons	AMR genes	Match accession (size)	Accession
p14ARS_NMC0074-2	146,817	IncA/C2	<i>bla</i> <sub>NDM-1</sub> , <i>bla</i> <sub>OXA-10</sub> , <i>aadA1</i> , <i>aph(3')-VI</i> , <i>aph(3')-Ia</i> , <i>mph(A)</i> , <i>cmlA1</i> , <i>flor</i> , <i>sul1</i> , <i>sul2</i> , <i>dfrA14</i> , <i>qnrS1</i> , <i>arr-2</i>	NA	LR697127
p14ARS_NMC0074-5	51,479	IncX3	<i>bla</i> <sub>OXA-181</sub> , <i>qnrS1</i>	CP024806.1 (51,479 bp)	LR697131
p14ARS_NMC0074-4	81,712	IncFII/IncFIA	<i>bla</i> <sub>CTX-M-15</sub> , <i>bla</i> <sub>TEM-1B</sub> , <i>bla</i> <sub>OXA-1</sub> , <i>aadA5</i> , <i>aac(6')-Ib-cr</i> , <i>sul1</i> , <i>tet(B)</i> , <i>dfrA17</i>	CP024805.1 (111,310 bp)	LR697131
p14ARS_NMC0074-3	82,418	IncFII	<i>bla</i> <sub>TEM-1</sub> , <i>mph(A)</i> , <i>erm(B)</i>	NA	LR697130
p14ARS_NMC0074-6	2088	Col(BS512)	-	CP024803.1 (2088 bp)	LR697133

The replicons and AMR genes were identified with the PlasmidFinder and ResFinder databases, respectively. The match accession indicates the accession code when at least one match of more than 90% query sequence coverage and 99% identity was found in the NCBI nucleotide database. NA not available.

sequencing, bioinformatics, and reporting to the DOH conducted locally, with a turnaround time of 8 days from sample receipt to report preparation.

While progress has been made, significant challenges remain to bring genomics into routine surveillance in low-resource settings. These include, but are not limited to, challenges in the supply chain and procurement of WGS reagents and equipment, vastly differing costs between high- and low-income settings, shortage of skilled local bioinformaticians due to both limited access to training and difficulties in staff retention, and the lack of platforms to feedback the actionable genomic data to sentinel sites. In addition, it was expected that the timeline from implementation of a disruptive technology such as WGS to full adoption by all the relevant stakeholders was to exceed the period of one academically funded project. The continued collaboration with the ARSRL (via an NIHR-funded project) is building on the foundations laid by this study to tackle some of the challenges mentioned above, and to improve on the translation of genomic data into public-health action and policy. We laid the foundations for the sustainable implementation of WGS and genomic epidemiology within national surveillance networks in LMICs, which can be extended to other locations to tackle the global challenge of AMR.

## Methods

**Ethics statement.** Ethical approval is not applicable as per the Research Institute for Tropical Medicine—Institutional Review Board. This study uses archived bacterial samples processed by ARSP. No identifiable data were used in this study.

**ARSP workflow and data.** The ARSP implements AMR surveillance on aerobic bacteria from clinical specimens. The programme collects culture and antimicrobial susceptibility data from its 24 sentinel sites and 2 gonorrhoeae surveillance sites. The sentinel sites participate in an external quality assessment scheme conducted by the reference laboratory to ensure quality of laboratory results. Case finding is based on priority specimens sent routinely to the hospitals' laboratories for clinical purposes, and is thus based on the diagnostic practices of the clinicians, as well as the resources available to the patient to cover the culture and susceptibility testing. Sentinel sites submit monthly results of all culture and susceptibility tests to the ARSRL via WHONET, a free Windows-based database software for the management and analysis of microbiology laboratory data with a special focus on the analysis of antimicrobial susceptibility test results<sup>6</sup>. Multi-drug-resistant phenotypes are classified according to recommended standard definitions<sup>48</sup>. Select isolates are referred to the ARSRL for confirmatory testing if they fall in one of the three following categories: (1) organisms of public health importance regardless of their resistance phenotype, such as *Salmonella enterica* ser. Typhi, *Streptococcus pneumoniae* and *Vibrio cholerae*; (2) organisms with emerging resistance, such as carbapenem resistance; (3) organisms with rare susceptibility patterns, such as vancomycin-resistant *Staphylococcus aureus* or ceftriaxone-resistant *Neisseria gonorrhoeae*. At the ARSRL, all referred isolates are re-identified using standardized culture, and susceptibility tested by disc diffusion or minimum inhibitory concentration (MIC) determined using the most current guidance for antibiotic test panels and breakpoints recommended by the Clinical and Laboratory Standards Institute. Serotyping for *S. pneumoniae*, *H. influenzae*, *Salmonellae* and *V. cholerae* is conducted to further characterize the isolates.

**ARSP data and bacterial isolates included in this study.** Phenotypic data collected since 1988 from antimicrobial susceptibility tests were summarized with WHONET v5.6 to compute yearly resistance rates for key pathogen–antibiotic combinations. Only the first bacterial isolate per patient per species per year was included in the calculations.

RP derived from the antibiograms were summarized per organism with WHONET v5.6<sup>6</sup>, and their relative abundance visualized with Tableau Desktop. Data collected between 2009 and 2017 were included in the analysis as the number of sentinel sites collecting data remained relatively stable during this period (22–24 sites). Isolates with missing data for any antibiotic on the panels listed on Table 1 were excluded.

The bacterial pathogens included in the See and Sequence study were those responsible for the majority of antimicrobial-resistant infections in the Philippines. These included carbapenemase-producing *Pseudomonas aeruginosa* and *Acinetobacter baumannii*, carbapenemase-producing and ESBL-suspect *Escherichia coli* and *Klebsiella pneumoniae*, methicillin-resistant *Staphylococcus aureus*, *Neisseria gonorrhoeae*, non-typhoidal *Salmonella*, and *Salmonella enterica* ser. Typhi. Linked epidemiological data included location and date of specimen collection, type of specimen, type of patient (in or outpatient) and sex and age of the patient. We utilized a proxy definition for infection origin whereby patient first isolates collected in the community or on either of the first 2 days of hospitalization were categorized as community infection isolates, while isolates collected on hospital day 3 or later were categorized as HAI isolates<sup>8</sup>.

In this article, we focus on carbapenemase-producing organisms. All carbapenemase-producing *K. pneumoniae* and *E. coli* isolates referred to, confirmed, and banked by the ARSRL in 2013–2014 were selected for the retrospective sequencing survey. Approximately 100 isolates of carbapenemase-producing *P. aeruginosa* and *A. baumannii* each were selected according to the following criteria: (1) referred to ARSRL in 2013–2014; (2) complete RP (i.e., no missing data); (3) overall prevalence of the RP in the ARSP data (including both referred and non-referred isolates); (4) geographical representation of different sentinel sites. The number of isolates included from each sentinel site was proportional to their relative abundance and estimated from  $(n/N) \times 100$  (rounded up), where  $n$  is the total number of isolates from one site, and  $N$  is grand total of isolates; (5) when both invasive and non-invasive isolates representing a combination of RP, sentinel site and year of collection were available, invasive isolates (i.e., from blood, or cerebrospinal, joint, pleural and pericardial fluids) were given priority. In addition, ~100 isolates of ESBL-producing *E. coli* and *K. pneumoniae* each were included as per the criteria above. Pan-susceptible isolates were only available for *P. aeruginosa* isolates and were also included in the retrospective sequencing survey.

Lyophilisates of the selected bacterial isolates were re-suspended in 0.5 mL of tryptic soy broth, plated onto MacConkey agar plates and incubated at 35–37 °C for 18–24 h. The species identity of the revived bacterial isolates was confirmed by characterization of the colony morphology and conventional biochemical tests (e.g., oxidase, motility and sugar fermentation using triple sugar iron agar test). Phenotypic re-testing was performed on the resuscitated isolates based on the phenotype stored in WHONET. Carbapenemase production in resuscitated isolates of Enterobacteriaceae was tested with the modified Hodge test, and confirmed using the imipenem MIC method with imipenem Etest strips (BioMerieux). *P. aeruginosa* and *A. baumannii* were screened for metallo-beta-lactamase production using the EDTA synergy test between imipenem disk (10 µg) and EDTA (750 µg/mL), which was confirmed by imipenem/imipenem inhibitor Etest strips. *E. coli* and *K. pneumoniae* isolates were screened for ESBL with the double-disk synergy test between aztreonam (30 µg) and amoxicillin/clavulanic acid (20/10 µg) and confirmed by using the gradient diffusion method with cefotaxime/cefotaxime with clavulanic acid or ceftazidime/ceftazidime with clavulanic acid Etest strips.

**Whole-genome sequencing, assembly and annotation.** *A. baumannii*, *E. coli*, *K. pneumoniae* and *P. aeruginosa* bacterial strains were shipped to the Wellcome Sanger Institute streaked onto Nutrient Agar butts contained in screw-cap cryovials, where DNA was extracted from a single colony of each isolate with the



QIAamp 96 DNA QIAcube HT kit and a QIAcube HT (Qiagen; Hilden, Germany). DNA extracts were multiplexed and sequenced on the Illumina HiSeq platform (Illumina, CA, USA) with 100-bp paired-end reads.

The quality of the sequence data was assessed for median base quality >30, mapping coverage >80% of the length of a reference genome sequence of the same species (*A. baumannii* ATCC 17978, *E. coli* UPEC\_ST131, *K. pneumoniae* subsp. *pneumoniae* Ecl8 and *P. aeruginosa* LESB58), and >80% of the sequence reads assigned to the corresponding species with Kraken v0.10.6-a2d113dc8<sup>49</sup> and a reference a database of all viruses, archaea and bacteria genomes in RefSeq<sup>50</sup> and the mouse and human reference.

Annotated assemblies were produced using the pipeline described in<sup>51</sup>. For each sample, sequence reads were used to create multiple assemblies using VelvetOptimiser v2.2.5<sup>52</sup> and Velvet v1.2<sup>53</sup>. An assembly improvement step was applied to the assembly with the best N50 and contigs were scaffolded using SSPACE v2.0<sup>54</sup> and sequence gaps were filled using GapFiller v1.11<sup>55</sup>. Automated annotation was performed using PROKKA v1.5<sup>56</sup> and a genus-specific database from RefSeq<sup>50</sup>. The quality of the assemblies was assessed for total size, number of contigs, N50 and GC content appropriate for each species, with all genomes passing quality control characterized by assemblies of <150 contigs and N50 >60,000.

**Phylogenetic analysis.** Evolutionary relationships between isolates were inferred from SNPs by mapping the paired-end reads to the reference genomes of *A. baumannii* A1 (CP010781), *E. coli* EC958 (HG941718), *K. pneumoniae* subsp. *pneumoniae* (AP006725) or *P. aeruginosa* LESB58 (FM209186), using the Burrows Wheeler aligner v0.7.12<sup>57</sup> to produce a BAM file. PCR duplicate reads were identified using Picard v1.92<sup>58</sup> and flagged as duplicates in the BAM file.

Variation detection was performed using samtools mpileup v0.1.19<sup>59</sup> with parameters -d 1000 -DSugBf and bcftools v0.1.19<sup>60</sup> to produce a BCF file of all variant sites. The option to call genotypes at variant sites was passed to the bcftools call. All bases were filtered to remove those with uncertainty in the base call. The bcftools variant quality score was required to be >50 (quality < 50) and mapping quality >30 (map\_quality < 30). If not all reads gave the same base call, the allele frequency, as calculated by bcftools, was required to be either 0 for bases called the same as the reference, or 1 for bases called as an SNP (af1 < 0.95). The majority base call was required to be present in at least 75% of reads mapping at the base (ratio < 0.75), and the minimum mapping depth required was 4 reads, at least two of which had to map to each strand (depth < 4, depth\_strand < 2). Finally, strand\_bias was required to be <0.001, map\_bias < 0.001 and tail\_bias < 0.001. If any of these filters were not met, the base was called as uncertain.

A pseudogenome was constructed by substituting the base call at each site (variant and non-variant) in the BCF file into the reference genome, and any site called as uncertain was substituted with an N. Insertions with respect to the reference genome were ignored and deletions with respect to the reference genome were filled with Ns in the pseudogenome to keep it aligned and the same length as the reference genome used for read mapping. For sequence-type-specific phylogenies, MGEs and recombination regions detected with Gubbins v1.4.10<sup>61</sup> were masked from the alignment of pseudogenomes produced by mapping against the genomes of *K. pneumoniae* strain CAV1217 (ST340, CP018676.1), *K. pneumoniae* strain MS6671 (ST147, LN824133.1) or *E. coli* strain AMA1167 (ST410, CP024801.1). Alignments of only variant positions were inferred with snp-sites v2.4.1<sup>62</sup> and were also used to compute pairwise SNP differences between isolates ([https://github.com/simonrharris/pairwise\\_difference\\_count](https://github.com/simonrharris/pairwise_difference_count)).

Maximum-likelihood phylogenetic trees were generated with RAXML v8.2.8<sup>63</sup> based on the generalized time reversible model with GAMMA method of correction for among-site rate variation and 100 bootstrap replications. Trees were midpoint rooted.

The phylogenetic trees, genomic predictions of AMR and genotyping results were visualized together with the ARSP epidemiological data using Microreact<sup>27</sup>.

**In silico predictions of AMR determinants and plasmid replicons.** Known antibiotic resistance determinants were identified using Antibiotic Resistance Identification By Assembly v.2.6.1 (ARIBA)<sup>64</sup>. For acquired genes an in-house curated database comprising 2316 known resistance gene variants was used (<https://figshare.com/s/94437a301288969109c2>)<sup>65</sup>. Full-length assembled matches with identity >90% and coverage >5× were considered as evidence of the presence of the gene in the query genome. Further manual inspection of the results excluded 13 matches that passed the above filters but that were suspected low-level contamination. Mutations in *ompK35* and *ompK36* genes were detected with ARIBA and the Comprehensive Antimicrobial Resistance Database<sup>44</sup> and the output was inspected for truncations, interruptions and frameshifts.

Plasmid replicons were identified with ARIBA and the PlasmidFinder database<sup>66</sup>. Full-length assembled matches with 100% identity were considered as evidence of the presence of the replicon gene.

**In silico multi-locus sequence typing.** Multi-locus STs were determined from assemblies with MLSTcheck v1.007001<sup>67</sup> or from sequence reads with ARIBA<sup>64</sup>, using species-specific databases hosted at PubMLST (<https://pubmlst.org>)<sup>68</sup> for *E. coli*<sup>69</sup>, *A. baumannii*<sup>70</sup>, *P. aeruginosa*<sup>71</sup> or at BIGSdb (<https://bigsdbs.readthedocs.io/en/latest/#>)<sup>68</sup> for *K. pneumoniae*<sup>72</sup>.

MLST calls from *P. aeruginosa* genome sequences were manually curated as we noted the presence of two genes exhibiting 100% sequence identity to different *acsA* alleles in the pubmlst database (accessed April 2019), which confounded the in silico predictions. This was not privy to the genomes in this study, since the genome sequence of reference strain NCGM2\_S1 (accession AP012280), known to belong to ST235, showed two genes with locus tags NCGM2\_1665 (nt position 1808730–1810685) and NCGM2\_0806 (nt positions c879796–881733) that are annotated as *acsA* and *acsB*, respectively, and are 100% identical to alleles *acsA\_38* and *acsA\_225*, respectively (<https://pubmlst.org/data/alleles/paeruginosa/acsA.tfa>, accessed Apr 2019). Furthermore, the same was only observed for the *acsA* allele both from assemblies (MLST\_check) and raw reads (ARIBA-MLST), and therefore unlikely the result of contamination with other strains. Allele *acsA\_38* defines ST235, but allele *acsA\_225* defines a novel ST. Thus, when two different *acsA* alleles were detected in a *P. aeruginosa* genome we selected the one associated with a known ST.

**Characterization of carbapenemase plasmids.** Select *K. pneumoniae* (14ARS\_CVM0040, 14ARS\_MMH0055, 13ARS\_GMH0099, 14ARS\_VSM0843, 13ARS-VSM0593 and 13ARS\_MMH0112) and *E. coli* (14ARS\_NMC0074) isolates were also sequenced with PacBio RS II or Sequel (Pacific Biosciences, CA, USA) or Oxford Nanopore Gridion (Oxford Nanopore, Oxford, UK) platforms. Hybrid assemblies from short and long reads were obtained with Unicycler v0.4.0<sup>73</sup> with default parameters, except for 13ARS\_GMH0099 for which the bold assembly mode was used. The sequence of the p13ARS\_GMH0099 plasmid was manually curated with Bandage v0.8.1<sup>74</sup> circularized with Circlator v1.5.3<sup>75</sup>, and further polished with Quiver (Pacific Biosciences, CA, USA).

The presence of AMR genes and plasmid replicons in the plasmid sequences was detected with ResFinder<sup>45</sup> and PlasmidFinder<sup>66</sup>, respectively. Matches with identity larger than 98% were reported. Detailed annotations of plasmids were obtained with MARA<sup>76</sup>. Visual comparisons between plasmid sequences were created with BRIG v1.0<sup>77</sup> with default BLAST v2.7.0 parameters.

The distribution of plasmids across draft genomes was inferred by mapping short reads to the plasmid sequences with smalt v0.7.4 with identity threshold  $y = 96\%$  and random alignment of reads with multiple mapping positions of equal score ( $-r 1$ ). The BAM file was sorted and indexed with samtools v0.1.19–44428cd<sup>59</sup>, which was also used to generate a pileup (mpileup -DSugBf) of the alignment records in BCF format. A pseudo-plasmid sequence in fastq format was constructed with the bcftools view and vcfutils.pl vcf2fq programmes, and the sequence length coverage (%) of the reference plasmid was computed.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Raw sequence data generated during this study have been deposited in the European Nucleotide Archive (ENA) under the study accession PRJEB17615. Run accessions are provided in the Microreact projects linked throughout the paper and below. The plasmid sequence assemblies generated during this study are available from Tables 2–4. The resistance database used in this study is available at: <https://figshare.com/s/94437a301288969109c2>. The multi-locus sequence-type databases used in this study are available at PubMLST (<https://pubmlst.org>) or at BIGSdb (<https://bigsdbs.readthedocs.io/en/latest/#>).

The data presented in this study are available from the following Microreact projects: [https://microreact.org/project/ARSP\\_ABA\\_2013-14](https://microreact.org/project/ARSP_ABA_2013-14) [https://microreact.org/project/ARSP\\_PAE\\_2013-14](https://microreact.org/project/ARSP_PAE_2013-14) [https://microreact.org/project/ARSP\\_KPN\\_2013-14](https://microreact.org/project/ARSP_KPN_2013-14) [https://microreact.org/project/ARSP\\_ECO\\_2013-14](https://microreact.org/project/ARSP_ECO_2013-14) [https://microreact.org/project/ARSP\\_KPN\\_ST340\\_2013-14](https://microreact.org/project/ARSP_KPN_ST340_2013-14) [https://microreact.org/project/ARSP\\_KPN\\_ST147\\_2013-14](https://microreact.org/project/ARSP_KPN_ST147_2013-14) [https://microreact.org/project/ARSP\\_KPN\\_ST147\\_GLOBAL](https://microreact.org/project/ARSP_KPN_ST147_GLOBAL) [https://microreact.org/project/ARSP\\_ECO\\_ST410](https://microreact.org/project/ARSP_ECO_ST410) [https://microreact.org/project/ARSP\\_ECO\\_ST410\\_GLOBAL](https://microreact.org/project/ARSP_ECO_ST410_GLOBAL)

Received: 6 March 2020; Accepted: 25 April 2020;

Published online: 01 June 2020

## References

- World Health Organization. Antimicrobial resistance: global report on surveillance <http://www.who.int/iris/handle/10665/112642> (2014).
- World Bank. Drug-resistant infections: a threat to our economic future <http://documents.worldbank.org/curated/en/323311493396993758/pdf/114679-REVISED-v2-Drug-Resistant-Infections-Final-Report.pdf> (2017).
- World Health Organization. Global Action Plan on antimicrobial resistance [http://www.wpro.who.int/entity/drug\\_resistance/resources/global\\_action\\_plan\\_eng.pdf](http://www.wpro.who.int/entity/drug_resistance/resources/global_action_plan_eng.pdf) (2015).
- Department of Health, Republic of the Philippines. Morbidity <https://www.doh.gov.ph/morbidity> (2014).

5. Department of Health, Republic of the Philippines. Mortality <https://www.doh.gov.ph/mortality> (2013).
6. O'Brien, T. F. & Stelling, J. M. WHONET: an information system for monitoring antimicrobial resistance. *Emerg. Infect. Dis.* **1**, 66 (1995).
7. Antimicrobial Resistance Surveillance Reference Laboratory, Research Institute for Tropical Medicine. Annual reports <https://arsp.com.ph/publications/> (2019).
8. World Health Organization. Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2016–2017 <https://www.who.int/glass/resources/publications/early-implementation-report/en/> (2017).
9. Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems: past, present, and future. *Antimicrob. Agents Chemother.* **55**, 4943–4960 (2011).
10. Ashton, P. M. et al. Identification of *Salmonella* for public health surveillance using whole genome sequencing. *PeerJ* **4**, e1752 (2016).
11. Deng, X., den Bakker, H. C. & Hendriksen, R. S. Genomic epidemiology: whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annu. Rev. Food Sci. Technol.* **7**, 353–374 (2016).
12. Doumith, M. et al. Detection of the plasmid-mediated *mcr-1* gene conferring colistin resistance in human and food isolates of *Salmonella enterica* and *Escherichia coli* in England and Wales. *J. Antimicrob. Chemother.* **71**, 2300–2305 (2016).
13. Wong, V. K. et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events. *Nat. Genet.* **47**, 632–639 (2015).
14. Koser, C. U., Ellington, M. J. & Peacock, S. J. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet.* **30**, 401–407 (2014).
15. Dadashi, M. et al. Frequency distribution, genotypes and the most prevalent sequence types of New Delhi metallo-beta-lactamase-producing *Escherichia coli* among clinical isolates around the world: a review. *J. Glob. Antimicrob. Resist.* **19**, 284–293 (2019).
16. Diene, S. M. & Rolain, J. M. Carbapenemase genes and genetic platforms in gram-negative bacilli: Enterobacteriaceae, *Pseudomonas* and *Acinetobacter* species. *Clin. Microbiol. Infect.* **20**, 831–838 (2014).
17. Liu, Y. et al. First report of OXA-181-producing *Escherichia coli* in China and characterization of the isolate using whole-genome sequencing. *Antimicrob. Agents Chemother.* **59**, 5022–5025 (2015).
18. Roer, L. et al. *Escherichia coli* sequence type 410 is causing new international high-risk clones. *mSphere* **3**, e00337–e00318 (2018).
19. Falgenhauer, L. et al. Circulation of clonal populations of fluoroquinolone-resistant CTX-M-15-producing *Escherichia coli* ST410 in humans and animals in Germany. *Int. J. Antimicrob. Agents* **47**, 457–465 (2016).
20. Schauler, K. et al. Clonal spread and interspecies transmission of clinically relevant ESBL-producing *Escherichia coli* of ST410—another successful pandemic clone? *FEMS Microbiol. Ecol.* **92**, fiv155 (2016).
21. Dortet, L., Poirel, L., Al Yaqoubi, F. & Nordmann, P. NDM-1, OXA-48 and OXA-181 carbapenemase-producing Enterobacteriaceae in Sultanate of Oman. *Clin. Microbiol. Infect.* **18**, E144–E148 (2012).
22. Szekeley, E. et al. First description of *bla*(NDM-1), *bla*(OXA-48), *bla*(OXA-181) producing Enterobacteriaceae strains in Romania. *Int. J. Med. Microbiol.* **303**, 697–700 (2013).
23. Gamal, D., Fernandez-Martinez, M., El-Defrawy, I., Ocampo-Sosa, A. A. & Martinez-Martinez, L. First identification of NDM-5 associated with OXA-181 in *Escherichia coli* from Egypt. *Emerg. Microbes. Infect.* **5**, e30 (2016).
24. Overballe-Petersen, S. et al. Complete nucleotide sequence of an *Escherichia coli* sequence type 410 strain carrying *bla*<sub>NDM-5</sub> on an IncF multidrug resistance plasmid and *bla*<sub>OXA-181</sub> on an IncX3 plasmid. *Genome Announc.* **6**, e01542–e01517 (2018).
25. Abboud, C. S. et al. A space-time model for carbapenemase-producing *Klebsiella pneumoniae* (KPC) cluster quantification in a high-complexity hospital. *Epidemiol. Infect.* **143**, 2648–2652 (2015).
26. Park, R. et al. Statistical detection of geographic clusters of resistant *Escherichia coli* in a regional network with WHONET and SaTScan. *Expert Rev. Anti Infect. Ther.* **14**, 1097–1107 (2016).
27. Argimon, S. et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Micro. Genom.* **2**, e000093 (2016).
28. Wyres, K. L. et al. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Micro. Genom.* **2**, e000102 (2016).
29. Gwinn, M., MacCannell, D. R. & Khabbaz, R. F. Integrating advanced molecular technologies into public health. *J. Clin. Microbiol.* **55**, 703–714 (2017).
30. Epton, E. E. et al. Carbapenem-resistant *Klebsiella pneumoniae* producing New Delhi metallo-beta-lactamase at an acute care hospital, Colorado, 2012. *Infect. Control Hosp. Epidemiol.* **35**, 390–397 (2014).
31. Harris, S. R. et al. Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect. Dis.* **13**, 130–136 (2013).
32. Peacock, S. J., Parkhill, J. & Brown, N. M. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. *Microbiology* **164**, 1213–1219 (2018).
33. Dymond, A. et al. Genomic surveillance of methicillin-resistant *Staphylococcus aureus*: a mathematical early modelling study of cost effectiveness. *Clin. Infect. Dis.* **ciz480** <https://doi.org/10.1093/cid/ciz480> (2019).
34. Lascols, C., Peirano, G., Hackel, M., Laupland, K. B. & Pitout, J. D. Surveillance and molecular epidemiology of *Klebsiella pneumoniae* isolates that produce carbapenemases: first report of OXA-48-like enzymes in North America. *Antimicrob. Agents Chemother.* **57**, 130–136 (2013).
35. Wyres, K. L. et al. Genomic surveillance for hypervirulence and multi-drug resistance in invasive *Klebsiella pneumoniae* from South and Southeast Asia. *Genome Med* **12**, 11 (2020).
36. Baker, S., Thomson, N., Weill, F. X. & Holt, K. E. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. *Science* **360**, 733–738 (2018).
37. Qin, S., Cheng, J., Wang, P., Feng, X. & Liu, H. M. Early emergence of OXA-181-producing *Escherichia coli* ST410 in China. *J. Glob. Antimicrob. Resist.* **15**, 215–218 (2018).
38. Khong, W. X. et al. Tracking inter-institutional spread of NDM and identification of a novel NDM-positive plasmid, pSg1-NDM, using next-generation sequencing approaches. *J. Antimicrob. Chemother.* **71**, 3081–3089 (2016).
39. Baek, J. Y. et al. Plasmid analysis of *Escherichia coli* isolates from South Korea co-producing NDM-5 and OXA-181 carbapenemases. *Plasmid* **104**, 102417 (2019).
40. Aung, M. S. et al. Prevalence of extended-spectrum beta-lactamase and carbapenemase genes in clinical isolates of *Escherichia coli* in Myanmar: dominance of *bla*<sub>NDM-5</sub> and Emergence of *bla*<sub>OXA-181</sub>. *Micro. Drug Resist.* **24**, 1333–1344 (2018).
41. Dallman, T. J. et al. Use of whole-genome sequencing for the public health surveillance of *Shigella sonnei* in England and Wales, 2015. *J. Med. Microbiol.* **65**, 882–884 (2016).
42. Nadon, C. et al. PulseNet International: vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. *Eur. Surveill.* **22**, 23.30544 (2017).
43. Gupta, S. K. et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* **58**, 212–220 (2014).
44. Jia, B. et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* **45**, D566–D573 (2017).
45. Zankari, E. et al. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).
46. Gayeta, J. et al. Establishment of whole genome sequencing in the antimicrobial resistance surveillance program in the Philippines: assessment of whole genome sequence quality. In *Applied Bioinformatics and Public Health Microbiology*. (Advanced Courses and Scientific Conferences, Hinxtion, Cambridge, UK, 2019).
47. Masim, M. A. et al. Genomic surveillance report for methicillin-resistant *Staphylococcus aureus* in the Philippines. In *Applied Bioinformatics and Public Health Microbiology*. (Advanced Courses and Scientific Conferences, Hinxtion, Cambridge, UK, 2019).
48. Magiorakos, A. P. et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**, 268–281 (2012).
49. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).
50. Pruitt, K. D., Tatusova, T., Brown, G. R. & Maglott, D. R. NCBI reference sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**, D130–D135 (2012).
51. Page, A. J. et al. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Micro. Genom.* **2**, e000083 (2016).
52. Gladman, S. & Seemann, T. Velvet optimiser: for automatically optimising the primary parameter options for the Velvet de novo sequence assembler. Victorian Bioinformatics Consortium <https://github.com/Victorian-Bioinformatics-Consortium/VelvetOptimiser> (2008).
53. Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
54. Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. & Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011).
55. Boetzer, M. & Pirovano, W. Toward almost closed genomes with GapFiller. *Genome Biol.* **13**, R56 (2012).
56. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
57. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

58. Broad Institute. Picard: a set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data <https://broadinstitute.github.io/picard/> (2020).
59. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. Samtools. Bcftools: utilities for variant calling and manipulating VCFs and BCFs <http://samtools.github.io/bcftools/> (2020).
61. Croucher, N. J. et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15 (2015).
62. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Micro. Genom.* **2**, e000056 (2016).
63. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
64. Hunt, M. et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Micro. Genom.* **3**, e000131 (2017).
65. David, S. et al. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat. Microbiol.* **4**, 1919–1929 (2019).
66. Carattoli, A. et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).
67. Page, A. J., Taylor, B. & Keane, J. A. Multilocus sequence typing by blast from de novo assemblies against PubMLST. *J. Open Source Softw.* **8**, 2 (2016).
68. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* **3**, 124 (2018).
69. Wirth, T. et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.* **60**, 1136–1151 (2006).
70. Bartual, S. G. et al. Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii*. *J. Clin. Microbiol.* **43**, 4382–4390 (2005).
71. Curran, B., Jonas, D., Grundmann, H., Pitt, T. & Dowson, C. G. Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. *J. Clin. Microbiol.* **42**, 5644–5649 (2004).
72. Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. & Brisse, S. Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *J. Clin. Microbiol.* **43**, 4178–4182 (2005).
73. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).
74. Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352 (2015).
75. Hunt, M. et al. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* **16**, 294 (2015).
76. Partridge, S. R. & Tsafnat, G. Automated annotation of mobile antibiotic resistance in gram-negative bacteria: the multiple antibiotic resistance annotator (MARA) and database. *J. Antimicrob. Chemother.* **73**, 883–890 (2018).
77. Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402 (2011).

## Acknowledgements

We are grateful to the members of the Antimicrobial Resistance Surveillance Program that collected bacterial isolates and linked epidemiological data. The full list can be found in the Supplementary Information (Supplementary Note 6). This work was funded by the Newton Fund, Medical Research Council (UK) grant MR/N019296/1, Philippine Council for Health Research and Development project number FP160007. J.S. was partially supported by research grants RR025040 and U01CA207167 from the National Institutes of Health (NIH). S.A. and D.M.A. were additionally supported by the National Institute for Health Research (UK) Global Health Research Unit on genomic Surveillance of AMR (16\_136\_111) and by the Centre for Genomic Pathogen Surveillance (<http://pathogensurveillance.net>).

## Author contributions

C.C.C. and D.M.A. conceived the study. S.A., M.A.L.M., J.M.G., M.L.L., P.K.V.M., V.C., M.T.L., H.O.E., J.C.P., J.C., M.C.J., A.S.V., J.B.B., A.M.O., L.F.T.H., K.D.B., B.J., K.A., C.M.H., S.B.S., J.S., M.T.G.H., C.C.C. and D.M.A. performed the data analysis. S.A., M.A.L.M., M.T.G.H., C.C.C. and D.M.A. wrote the paper. All authors read and approved the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-16322-5>.

Correspondence and requests for materials should be addressed to D.M.A. or C.C.C.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020