

SCIENTIFIC REPORTS



OPEN

Resistance mechanisms and population structure of highly drug resistant *Klebsiella* in Pakistan during the introduction of the carbapenemase NDM-1

Eva Heinz¹, Hasan Ejaz^{2,3,4}, Josefin Bartholdson Scott⁵, Nancy Wang³, Shruti Gujran³, Derek Pickard¹, Jonathan Wilksch³, Hanwei Cao³, Ikram-ul Haq⁶, Gordon Dougan^{1,5} & Richard A. Strugnell³

Klebsiella pneumoniae is a major threat to public health with the emergence of isolates resistant to most, if not all, useful antibiotics. We present an in-depth analysis of 178 extended-spectrum beta-lactamase (ESBL)-producing *K. pneumoniae* collected from patients resident in a region of Pakistan, during the period 2010–2012, when the now globally-distributed carbapenemase *bla*-NDM-1 was being acquired by *Klebsiella*. We observed two dominant lineages, but neither the overall resistance profile nor virulence-associated factors, explain their evolutionary success. Phenotypic analysis of resistance shows few differences between the acquisition of resistance genes and the phenotypic resistance profile, including beta-lactam antibiotics that were used to treat ESBL-positive strains. Resistance against these drugs could be explained by inhibitor-resistant beta-lactamase enzymes, carbapenemases or *ampC* type beta-lactamases, at least one of which was detected in most, but not all relevant strains analysed. Complete genomes for six selected strains are reported, these provide detailed insights into the mobile elements present in these isolates during the initial spread of NDM-1. The unexplained success of some lineages within this pool of highly resistant strains, and the discontinuity between phenotypic resistance and genotype at the macro level, indicate that intrinsic mechanisms contribute to competitive advantage and/or resistance.

Coincident with the development of antibiotic therapy, there has been a steady increase in the numbers and types of bacterial pathogens acquiring resistance to antimicrobials, with some clades showing a propensity to spread globally^{1,2}. *Klebsiella pneumoniae* has been particularly successful in this regard and is a member of the ESKAPE pathogen group that are acknowledged as causes of serious infections associated with multidrug resistance^{3,4}. Before this rapid increase in drug resistance, *K. pneumoniae* was primarily known as a major cause of infections in neonates, especially in Low- and Middle-Income Countries (LMICs)^{5–8} and community-acquired and nosocomial infections in immunocompromised patients^{9–11}.

The population structure of *K. pneumoniae* was redefined recently as a species ‘complex’. This complex includes *K. pneumoniae sensu stricto*, which poses the main burden of hospital-acquired infections; *K. variicola*, which can be plant-associated¹² but also causes hospital infections^{13–15}; and *K. quasipneumoniae*, which were initially thought to be asymptomatic carriage isolates, until more recent reports highlighted their potential virulence and increased drug resistance^{16–18}. DNA sequence-based phylogenetic typing analysis of the *K. pneumoniae*

¹Parasites and Microbes, Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK. ²Department of Clinical Laboratory Sciences, CAMS, Jouf University, Al-Jouf, Saudi Arabia. ³Department of Microbiology and Immunology, The University of Melbourne, at Peter Doherty Institute for Infection and Immunity, Melbourne, Australia.

⁴Department of Microbiology, The Children’s Hospital & The Institute of Child Health, Lahore, Pakistan. ⁵Department of Medicine, University of Cambridge, Cambridge, UK. ⁶Institute of Industrial Biotechnology, GC University, Lahore, Pakistan. Correspondence and requests for materials should be addressed to E.H. (email: eva.heinz@sanger.ac.uk) or R.A.S. (email: rastru@unimelb.edu.au)

species has revealed complex, deeply branched lineages, and these data can be used to inform the epidemiological analysis of different high-risk, globally distributed clones¹⁹.

The acquisition of extended-spectrum beta-lactamases (ESBLs) rapidly increased in *K. pneumoniae* from the 1990s, particularly in hospital isolates^{20,21}, and was driven by mobile elements (usually plasmids) often encoding several other resistance genes. The year 2009 saw the first description of NDM-1, a metallo-beta-lactamase. NDM-1 hydrolyzes carbapenems and is not targeted by beta-lactamase inhibitors²², thus allowing bacteria expressing NDM-1 to bypass the two main treatment options for ESBL-positive strains. The other known enzymes in *K. pneumoniae* conferring resistance against beta-lactamase inhibitors are *ampC*-type beta-lactamases and inhibitor-resistant beta-lactamases.

There is currently little understanding of why some *K. pneumoniae* sequence types are more successful than others. High-risk lineages in related pathogens, e.g. *Escherichia coli* ST131 and *Salmonella enterica* serovar Typhimurium ST313, acquired multiple antibiotic resistance determinants prior to their clonal expansions. Here, we describe the population structure and resistance profiles of *K. pneumoniae* isolated from a large hospital in Pakistan during routine sampling in 2010–2012. Our analysis reveals that there are many lineages that were prevalent at the time which have subsequently not spread more globally¹¹, and that dominant lineages which are now recognised as high-risk clones did not carry NDM-1. We combined short-read and long-read sequencing and phenotypic resistance profiles for selected isolates, and observed NDM-1 to be unstable in some of these *K. pneumoniae* lineages. Our study again strongly emphasises the relevance of the genetic background and intrinsic resistance mechanisms to provide some strains with a competitive advantage within a pool of highly resistant *K. pneumoniae* population.

Results

Conserved ESBL gene repertoire vs. high diversity in additional beta-lactamase genes. The isolates were collected between 2010–2012 through the routine microbiological screening of bacterial infections in The Children's Hospital, Lahore, Pakistan, and were pre-selected for ESBL expression through the E-test²³. At this time, ESBL-resistant *K. pneumoniae* were responsible for a significant clinical burden at Lahore hospital. The patients ranged in age between neonates (<29 d) and 15 years, and all received at least one invasive procedure during their hospital stay (97% intravenous lines) as described in detail in Ejaz *et al.*²³. ESBL-positive *K. pneumoniae* were responsible for a high number of negative outcomes ('left against medical advice' or death) in the hospital during the study period (87/214). Whilst the patient symptoms and the overall epidemiology of the strain collection was discussed in detail in Ejaz *et al.*²³, we now perform a detailed genomic analysis to identify resistance and virulence genes, and explore whether any of these genes conferred a competitive advantage within this highly drug-resistant isolate collection.

Phenotypic screening of antibiotic susceptibility (Table S1) identified isolates that were negative for AmpC, but resistant to cefoxitin, a second-generation cephalosporin that is, contrary to other cephalosporins, resistant to the activity of ESBL enzymes²⁴ (Fig. 1). To identify any lineages responsible for this resistance, and to assess whether a new resistance mechanism was spreading, we conducted a detailed analysis of the resistance profiles. We combined phenotypic analysis using the Vitek system, which facilitates measuring susceptibility to a wide range of antimicrobials at high throughput, with genomic analysis of candidate acquired resistance genes.

Klebsiella spp. normally harbour two chromosomally integrated enzymes related to low-level beta-lactam resistance. AmpH is an AmpC-related enzyme functioning as penicillin-binding protein, whereas beta-lactamases of the SHV, OKP or LEN family are usually present as one chromosomal copy in *K. pneumoniae*, *quasipneumoniae* and *variicola*, respectively²⁵ (Figs 2 and S1). In addition to these well established resistance genes, we detected a high number of acquired resistance genes, where the extended-spectrum beta-lactamase CTX-M-15 was ubiquitously present. We detected three OXA and three TEM beta-lactamase sequence variants (the TEM DNA sequence variants translate into identical amino acid sequences), and observed that a small number of isolates carried the rare VEB-5 beta lactamase. We furthermore detect several isolates with an additional copy of the SHV beta-lactamase family, most likely acquired on a plasmid. We detected other beta-lactam resistance genes with extended spectrums including the the AmpC enzymes CMY-2 and CMY-6, and the NDM-1 carbapenemases (Figs 2 and S1). The measured MIC values for widely-used antimicrobials of this isolate collection generally matched the genomically-predicted resistances (Table S1, Fig. S1). Intrinsic resistance mechanisms such as deactivation of porins, upregulation of efflux pumps, or differential expression of resistance genes, might account for resistance in the non-matching isolates^{26,27}.

Sensitivity was observed among a high number of isolates for cefoxitin and piperacillin-tazobactam. This is expected as cefoxitin is generally insensitive to ESBL^{24,28} (e.g. *bla*-CTX-M-15, which was present in all but one of our isolates), whereas resistance against piperacillin is widespread; in contrast, the beta-lactamase inhibitor tazobactam was still usually effective^{24,29}. In most of our isolates, resistance against cefoxitin was explained either by the presence of the carbapenemase *bla*-NDM-1 which hydrolyses cefoxitin³⁰, the *bla*-CMY AmpC³¹, or both (Fig. S1). Importantly, these data reveal that there was not simply a single acquisition of an additional resistance gene in the population leading to the extended resistance phenotype. Three isolates (HE021, HE205, HE206; Fig. S1) expressed phenotypic resistance that could not be explained by acquired genes. A pan-genome analysis of the *K. quasipneumoniae* lineage reveals a number of differences which might affect drug sensitivity such as transposases and hypothetical proteins (likely phage-derived following comparisons to public databases), however the only clear genetic difference between HE205 and HE206, and closely related sensitive isolates, is a predicted iron uptake system (*fec*) that was absent from the resistant strains. However, it is unclear whether *fec* has any effect on the observed resistance, or whether other changes like differential expression of genes, for example upregulation of efflux pumps or downregulation of porins, leads to cefoxitin resistance in these strains.

In considering other alternative drug treatments for ESBL-positive organisms, we noted that sensitivity against piperacillin/tazobactam (TZA) often coincided with lower MIC values against amoxicillin/clavulanate (Fig. S1; Table S1).

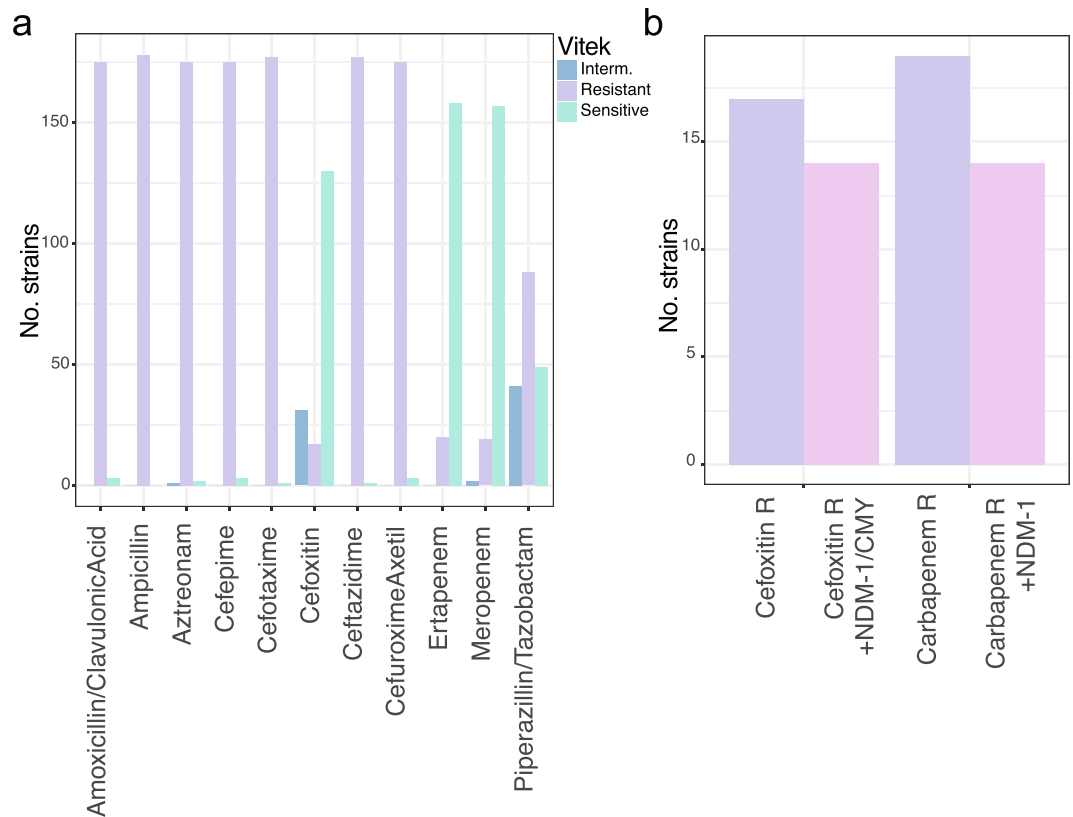


Figure 1. Resistance profiles. **(a)** Vitek data showing the number of strains read as sensitive, intermediate or resistant against the respective antimicrobial. **(b)** Isolates resistant against cefoxitin or carbapenems were assessed for the presence of genes expected to confer this resistance, highlighting a small gap of strains with non-predicted resistance mechanisms.

This is despite the fact that, based on a phenotypic sensitive/resistant evaluation, these isolates grew at drug levels above the defined resistance threshold, and thus were assessed as amoxicillin/clavulanate resistant. More detailed analysis of the encoded SHV/OKP *bla* genes revealed that several *K. quasipneumoniae* isolates (HE031, HE029, HE025, HE034, HE023; ST334) harboured a *bla*-SHV gene in addition to the intrinsic *bla*-OKP. The hypothesis that this mismatch in genotype and phenotypic resistance was plasmid-derived is supported by the plasmid replicon profiles of these isolates, which differ from closely related strains lacking the additional *bla*-SHV (Fig. 2). These putatively plasmid-borne *bla*SHV genes encode the mutations G238S and E240K, which correlate with TZA resistance or TZA-intermediate sensitivity (Fig. S1)^{32–34}. We also observed several sequences encoding the L35Q mutation. However, the L35Q mutation only provides a subtle increase in TZA MICs^{32,34,35}, and no mutation was found at S130³⁶. In summary, we found four different mechanisms that confer resistance against cefoxitin and/or inhibitor combinations in addition to the endogenous ESBLs. These include NDM-1, CMY-2, CMY-6, and additional SHV copies and their combinations, as well as three strains with unexplained profiles, highlighting the diversity of resistance mechanisms present in our collection.

No single factor gives competitive advantage. To test whether the observed resistance patterns resulted from expansion of single lineages that had earlier acquired the respective mechanism, or multiple introductions into the same or distinct lineages, we analysed the data in context of their phylogenetic history. Our dataset included 21 different sequence types (STs)²³ (Fig. 3a), with 56 isolates of *K. quasipneumoniae*, a species that was until recently thought to be less virulent than its related species, *K. pneumoniae sensu stricto* and *K. variicola* (Fig. 3)²³. The prevalence for each ST ranged from a single to 52 isolates (ST 334), showing that the STs that acquired ESBL genes were not equally competitive in the hospital. These observations suggest that unresolved factors in the genetic background lead to higher prevalence of some lineages over others when all the isolates with a similar resistance profile are evaluated in the same setting.

The capsule (K) is an essential virulence factor for *K. pneumoniae* and, together with lipopolysaccharide O-antigen (O) types, used for typing. K- and O-type can be identified and assigned based on their respective operons in the genomes. Recent work^{37,38} highlighted the considerable diversity found among different *K. pneumoniae* capsule types³⁹. This phenomenon is also reflected in our collection, where we observed a different capsule type for almost each sequence type, and no single type correlating with more dominant lineages (Fig. 3b, Table S2). This argues against any one capsule type itself conferring a competitive advantage in this setting, but rather that either multiple capsule types are equally advantageous, and/or that some combinations of capsule type and genotype are more competitive in a clinical setting than others.

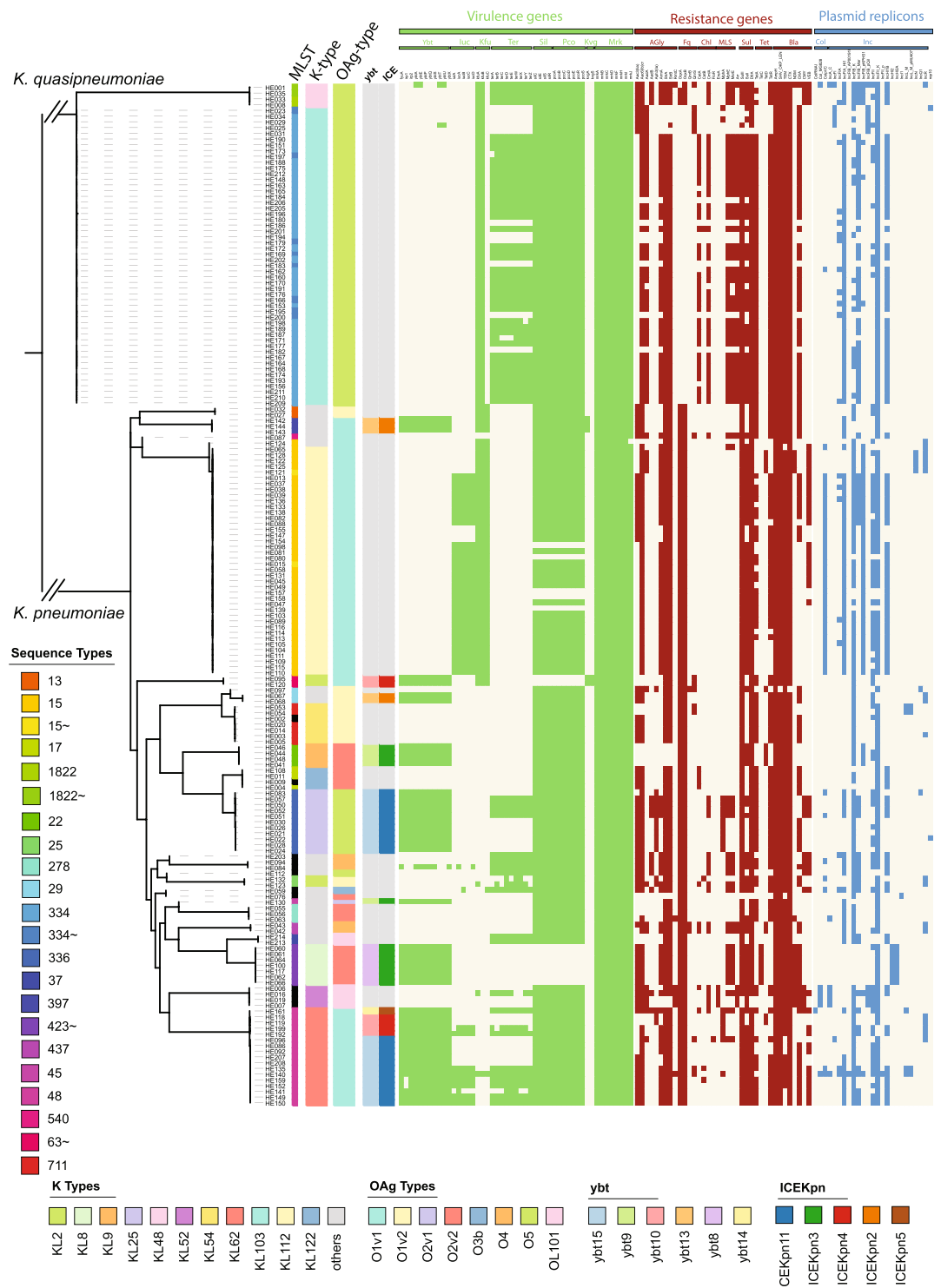


Figure 2. Antimicrobial resistance (AMR), plasmid replicons and virulence genes, and K- and O-type. The guide tree is based on roary, the branches leading to *K. pneumoniae* and *K. quasipneumoniae* are not to scale to facilitate visualisation; and shows the diversity of capsule and O-antigen type, which correlate with sequence types (one sequence type usually shares the capsule- and O-antigen combination). The predictions of AMR genes, virulence genes and plasmid replicons was performed using ariba⁴³, a mapping-based approach independent of assemblies. The guide tree is based on roary as in Fig. 1 where the branches leading to *K. pneumoniae* and *K. quasipneumoniae* are not to scale to facilitate visualisation.

In contrast to capsule loci, there was less diversity of LPS O-antigens with two main types in these ESBL-producing isolates (Fig. 3c). While other global studies^{37,40} observed a combination of O1 (52%/26%; from Follador *et al.*, and Pennini *et al.*, respectively), O2 (16%/36%) and O3 (15%/13%) as main LPS types^{37,40,41}, we observed

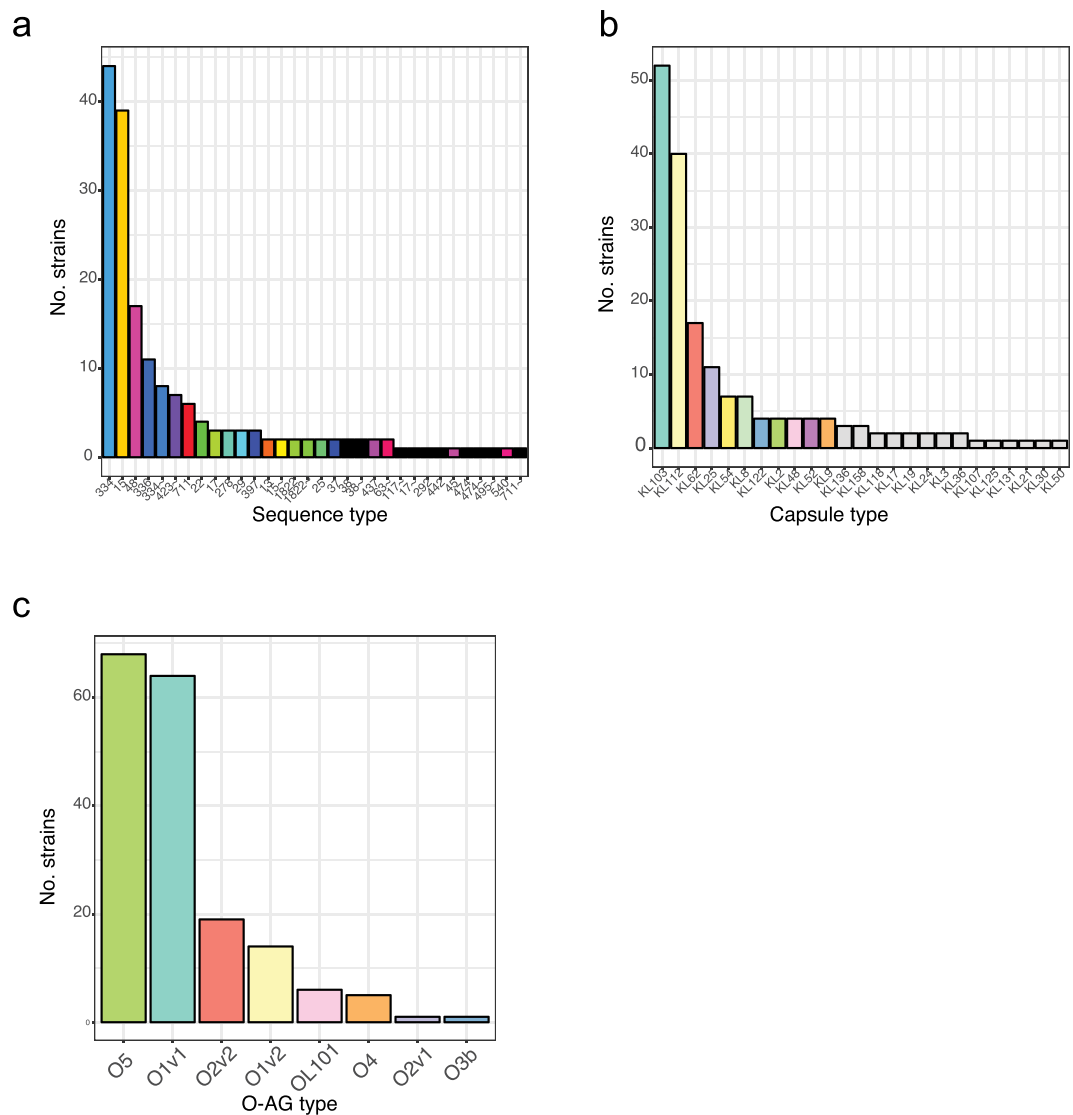


Figure 3. LPS O-antigen, capsule and sequence types in this study. **(a)** Two main (ST15, ST334) sequence types dominate the collection, followed by several medium-abundance (e.g. ST45, ST336) types and a high diversity of low-abundance sequence types. **(b)** The capsule types almost share almost exactly the same level as diversity as the sequence types, with a different capsule type associated with sequence types; whereas the O-antigen types **(c)** are as expected less diverse, but dominated through the commonly as less diverse perceived type O5.

O1 (44%) and O5 (38%; vs. 6%/1% in the global studies) as the most common. The high prevalence of O5 was due to a single lineage expansion, one of the two very dominant lineages was O5 (ST334), whereas the other one was O1v1 (ST15). The differences in O-antigen abundance of different studies are highly relevant as O-antigens are often considered vaccine targets in *K. pneumoniae*.

Virulence factor distribution. The genes encoding the siderophores aerolysin and yersiniabactin were only detected in *K. pneumoniae* isolates, in accordance with a previous large-scale study investigating 2499 isolates⁴² which also reports their absence in *K. quasipneumoniae*. Forty-one isolates, representing 33.6% *K. pneumoniae*, carried the yersiniabactin determinant, and 42 *K. pneumoniae* (34.4%) carried aerobactin. Only 7 of these isolates (5.7%) encoded both siderophores⁴² (Fig. 2), and no salmochelin or colibactin encoding genes were identified.

The profiles present in our strains include six different types of yersiniabactin using the classification of yersiniabactin (ybt) and their associated mobile elements (Kp integrative conjugative element ICEKp) by Lam *et al.*⁴². These ybt loci were present in eight sequence types. Our collection also included a less common combination of ybt8 with ICEKp3 (ST423). In addition, we found one sequence type (ST48) with three different ybt-ICE variants (Fig. 2; ICEKpn11 + ybt15, ICEKpn4 + ybt10, ICEKpn5 + ybt14), suggesting independent acquisition of different siderophore types even within a confined setting.

We also noted a high number of genes predicted as pseudogenes by ariba⁴³, especially multiple occurrences of the virulence gene *rmpA2* as pseudogene, which is one of the capsule upregulators and diagnostic for

Strain	MIC (Meropenem)	NDM-1 coverage
HE006 (14893_8#74)	64	75.0
HE007 (14893_8#75)	32	83.8
HE021 (14893_8#89)	2	Not Detected
HE016 (14893_8#84)	64	59.3
HE019 (14893_8#87)	64	79.1
HE121 (14936_2#2)	32	105.4
HE122 (14936_2#3)	16	49.4
HE065 (14936_3#41)	<1.25	Not Detected
HE125 (14936_2#6)	32	73.3
HE128 (14936_2#9)	32	43.4
HE001 (14893_8#69)	16	8.1
HE035 (14936_3#11)	4	Not Detected
HE063 (14936_3#39)	16	6.2
HE055 (14936_3#31)	16	Not Detected
HE060 (14936_3#36)	64	44.7
HE061 (14936_3#37)	64	47.1
HE062 (14936_3#38)	64	53.8
HE213 (14936_2#94)	<1.25	Not Detected
HE064 (14936_3#40)	64	52.9
HE066 (14936_3#42)	64	56.3
HE100 (14936_3#76)	64	43.5
HE117 (14936_3#93)	64	53.9

Table 1. Detection of carbapenemase using minimum inhibitory concentration and gene prediction (ariba read coverage).

hypervirulent *K. pneumoniae* (Fig. S2a). The *rmpA2* genes were all disrupted through a single-nucleotide addition or deletion in either a poly-G or poly-A region in this gene causing a frameshift as described previously⁴⁴ (Fig. S2a), and the encoded potential for random slip-strand misrepair to bring *rmpA2* into frame and increase capsule production. RmpA and aerobactin are indicative of the presence of the *K. pneumoniae* virulence plasmid found in hypervirulent strains^{45–48}, and we therefore compared the typical *K. pneumoniae* virulence plasmids, indicative in particular of the hypervirulent clonal group 23 (CG23), with our data. We performed comparisons against the plasmid pSGH10, which is representative of the hypervirulent clade CG23⁴⁸. This comparison showed that sequences similar to the pSGH10 plasmid are present in two main STs in our collection but that both STs lack salmochelin and other elements of the plasmid pSGH10 (Fig. S2b)^{49,50}. There was also variability of which regions of the plasmid were conserved within the STs in our study, again highlighting the highly dynamic and plastic nature of *K. pneumoniae* genomes.

The plasmid background of NDM-1. To identify the number of different plasmids that were present in our dataset, and to assess how many are still in circulation (e.g. reported in the recent literature), we performed long-read sequencing of representative isolates, using Pacific Biosciences (PacBio) technology (Tables 1, S2). Some NDM-1 genes detected in Illumina reads were not confirmed by PacBio sequencing, and were only detected with low read coverage (Table 1 and Fig. S3). Hence, we performed additional MIC tests for meropenem (E-tests, see methods; Table 1) using cultures generated after DNA extraction for sequencing. This phenotypic analysis confirmed that, where only lower read coverage for NDM-1 was observed in the earlier Illumina data, the bacteria were meropenem-sensitive indicating either loss of the plasmid carrying NDM-1 or of a highly mobile cassette carrying NDM-1. Even though we cannot exclude contamination during the multiplexed short-read sequencing run or the earlier stages of the sample preparation, the unstable nature of NDM-1 has been observed previously in several independent reports, where loss can occur after two generations of laboratory subculture, and even under meropenem selection^{51–53}. The putative loss of the entire plasmid carrying NDM-1 in part of our samples, or at least a larger transposable cassette, is further supported by the observation that several resistance genes that were predicted at equally low abundance in Illumina sequencing, were absent from the PacBio assemblies, indicating an unstable plasmid carrying these resistances, which included NDM-1 (Fig. S3).

Incompatibility group F and H plasmids were fixed in the population across all sequence types, with IncN and IncR only rarely observed; this is in agreement with a large comparative study describing IncFII, IncA/C, IncL/M, and IncI1 as the main resistance-carrying Inc types in *Klebsiella*⁵⁴ (Figs 2 and 4a). Two of the plasmids resolved by long-read sequencing (strains HE016 and HE066) show the original, more complete cassette including the GroEL/S from the *Acinetobacter* construct⁵⁵ (Fig. 4b,c), whilst in one case (HE125) a smaller portion of this gene region is conserved (Fig. 4c); this variation in the different genes surrounding NDM-1 has been described in other studies^{55–58}.

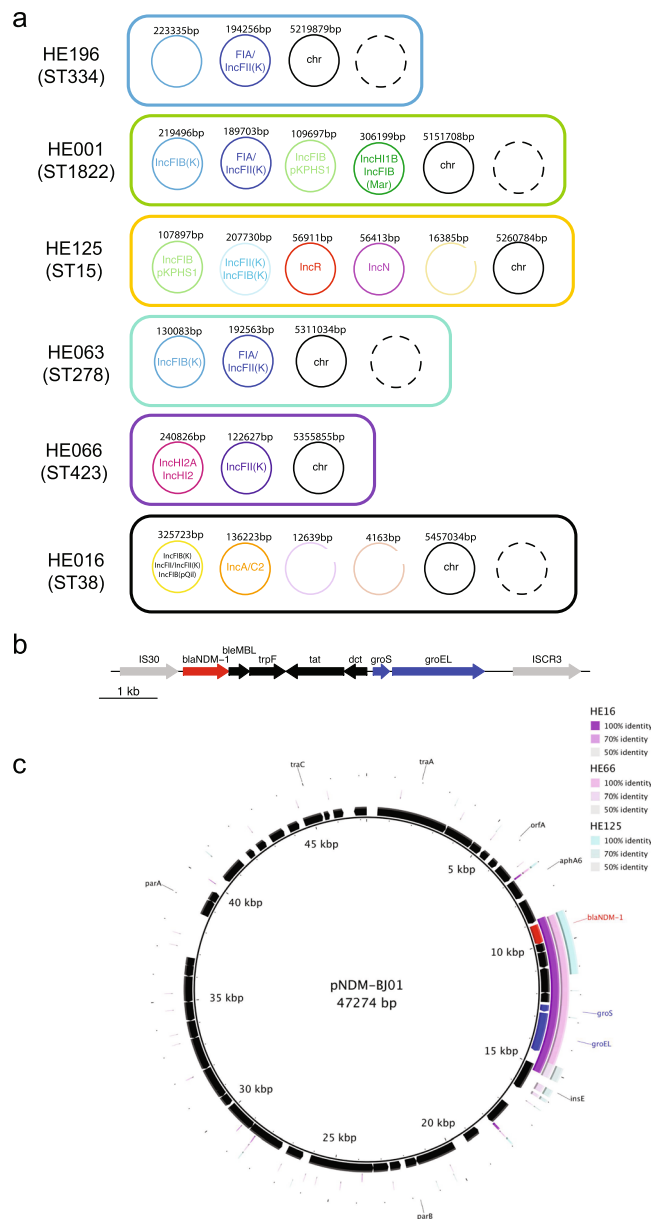


Figure 4. Complete plasmid repertoire of selected isolates. **(a)** The plasmids, their replicons and sizes as predicted after pacbio sequence analysis are shown for the respective isolates. **(b)** A close-up of the original NDM-1 cassette³⁶. **(c)** The three plasmid contigs from the PacBio assemblies of HE016, HE066 and HE125 which contain the NDM-1 gene, were compared against a plasmid encoding the original cassette as described for *Acinetobacter*⁵⁵. This shows the different levels of conservation of the cassette, whilst nothing else of the original plasmid can be found.

Comparison of ST15 isolates from Pakistan with a reported ST15 outbreak in Nepal. ST15 is an example of a widespread problematic *K. pneumoniae* clade. Given the relative geographic proximity between Pakistan and Nepal, we analysed the relationships between the two ST15 lineages reported in Nepal^{59,60} with those isolates in our collection (Table S3). We assessed the presence or absence of genes via a pan-genome analysis of a subset of isolates, as well as mapping of the short read data for the same isolates against a PacBio-sequenced reference sequence (accession number CP008929; Figs 5 and S4) from the Nepal outbreak⁵⁹. The phylogeny indicated that the ST15 population from Lahore was comprised of several subgroups, and differences indicated by longer branches correlated with changes seen in the mapping results (Fig. S4). Even within the tight group of the isolates from Lahore, at least three different sublineages had emerged. Each of these sublineages lack one, two or three regions found in the reference strain from Nepal (Fig. S4). A detailed overview of resistance, virulence gene profiles, plasmid replicons and surface determinants (capsule and O-antigen type) is provided in Fig. S4, and includes part of the Type VI secretion system as well as cellulose biosynthesis. Analysing the capsule operons, we observed the same novel capsule type as described in the Nepal isolates⁶⁰ (Fig. 5), further highlighting that ST15 appears to be spreading throughout the South Asian subcontinent.

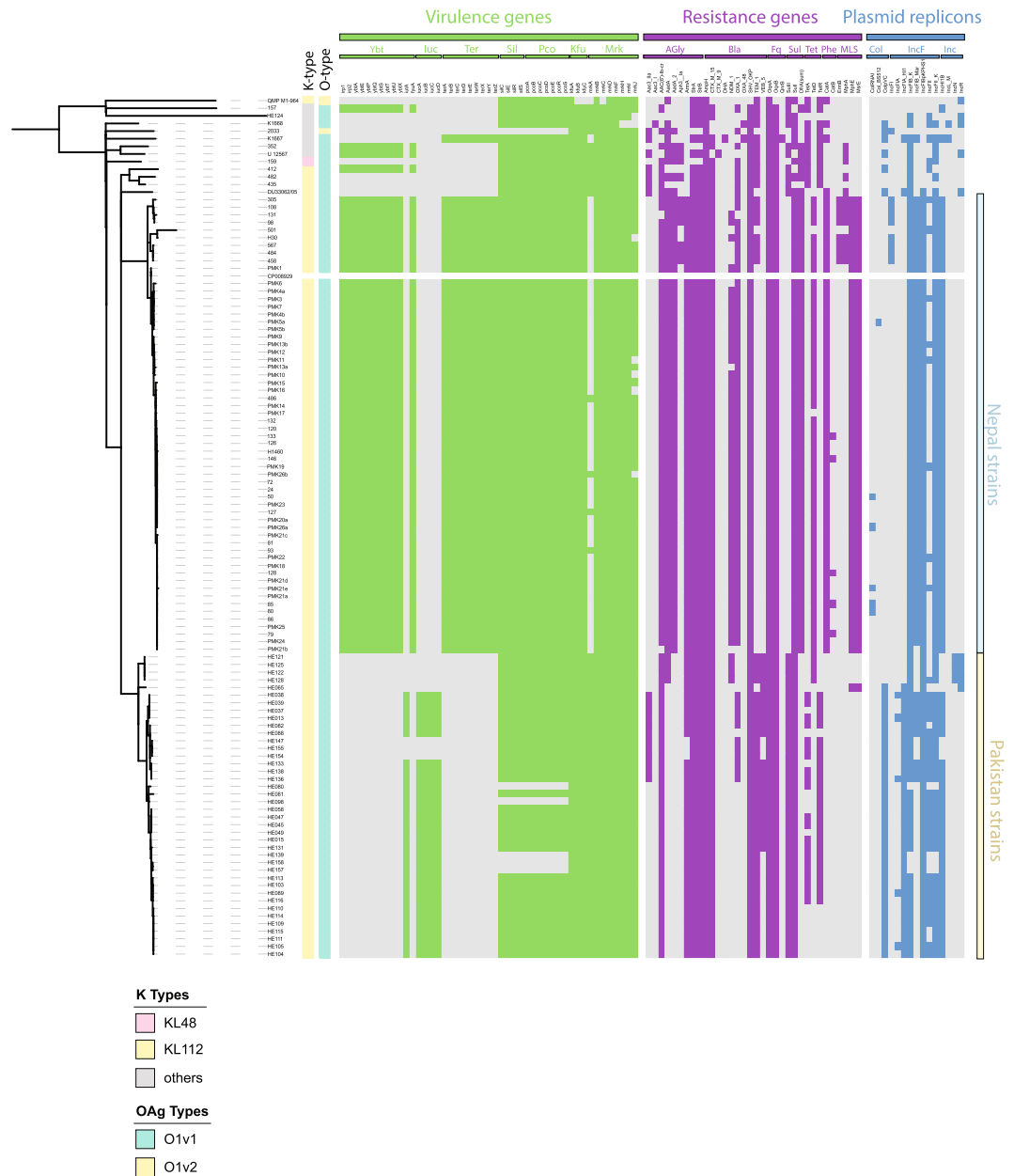


Figure 5. Comparison of the ST15 isolates from an outbreak in Nepal and Pakistan. Plasmid replicons, resistance genes and virulence genes were predicted by ariba; the tree is a whole-genome tree after removing recombination with gubbins against the reference strain from Nepal (CP008929.1).

In contrast, a comparison of the plasmid content showed almost entirely different compositions of plasmids in Nepal and Lahore, Pakistan (Figs 5 and S5). Only one of the Nepal plasmids (plasmid A) is partially conserved in the isolates we examined (Fig. S5), in accordance with the absence of NDM-1 (Fig. S5). Several gene clusters from the Nepal outbreak plasmids are conserved and likely to be encoded on different plasmids, including (Nepal plasmid A) *lacIZY*, phosphate transport (*phnD/E ptxD*); arsenic (*ars*), copper (*pco*), silver (*sil*) and cation (*cus*) efflux systems, as well as plasmid maintenance proteins from this plasmid. Only sparse similarity can be found with other plasmids (e.g. Nepal plasmid C streptomycin resistance), in accordance with the different virulence and resistance profiles of the two hospital lineages. A subgroup of the ST15 isolates from Pakistan (HE080, HE098, HE139, HE157, HE158) did not share the genes from plasmid A in Nepal, which can also be seen by their different virulence factor profile (Figs 5, S4 and S5).

Discussion

The *K. pneumoniae* isolates described in this study represent routinely collected samples in a Pakistan hospital of the *K. pneumoniae* species complex over an extended period²³. The main burden of global clinical isolates in hospital strains is usually associated with *K. pneumoniae sensu stricto*. There were two highly prevalent STs, and one of these was *K. quasipneumoniae* subsp. *similipneumoniae*, a subspecies that was previously thought to be less

Siderophore	Outcome		Total
	Death	Discharged	
Yersiniabactin+	16	26	42
Yersiniabactin–	34	75	109
Total	50	101	151
p-value ^a	$P_{chi} = 0.5388$		
Aerobactin+	13	22	35
Aerobactin–	37	79	116
Total	50	101	151
p-value ^a	$P_{chi} = 0.709$		

Table 2. Testing outcome versus presence of yersiniabactin or aerobactin. ^aP-value determined via a chi-square test of independence using the R function `chisq.test(input matrix)`.

clinically important (Fig. 2)²³. The major *K. quasipneumoniae* sequence type (ST334) has been rarely observed in other studies (e.g. 1/250⁶¹; 1/167⁶²; 1/328⁶³; 2/198⁶⁴), but was common in Lahore – we found 52 isolates (44 if excluding ST variants) all originating from bloodstream infections, representing 29.2% (24.7% of samples investigated). There is however very little data on *K. pneumoniae* diversity from the region, and it is unclear whether the expansion of ST334 was an unusual occurrence, or whether this ST is still dominant in the area. Given its similarity to *K. pneumoniae* in standard diagnostics, *K. quasipneumoniae* numbers are very likely underestimated in routine testing⁶⁵.

With the continuous loss of effective antimicrobials and the lag in the development of new ones being released, antibody-based treatments should be considered and have support through governments and global health programs^{66,67}. There are challenges associated with immunoprophylaxis against *Klebsiella* infections as there is no single prevalent capsule- or O-antigen type. Two large-scale analyses of global datasets including all three species identified serotypes O1, O2 and O3 in 80%/75% of hospital infections^{37,40}. We observed a high number of non-O1/O2 types in our collection, and the increase of genomic and functional studies of O-types in the past years led to an increase in recognised LPS diversity. The diversity and distribution is of high importance as O-antigen has often been investigated as a potential target for an anti-*Klebsiella* vaccine, as only few O-antigen serotypes are known, in contrast to the considerable diversity in capsule types^{37,38}. However, the lack of apparent correlation to disease severity³⁷ and LPS type indicate that if an O-antigen targeted vaccine were introduced, switches to non-vaccine targeted LPS serotypes or the emergence of new serotypes might occur once immune pressure is applied^{68,69}.

Given the reported association of the siderophore yersiniabactin with disease, we tested whether a correlation existed between siderophores and patient outcome, but neither yersiniabactin nor aerobactin were significantly associated with negative or positive events (death or discharged), in contrast to a recent description of high mortality with a reduced virulence factor repertoire of ST11⁷⁰ (details Table 2). We observed a similar prevalence of aerobactin and yersiniabactin in our population, and neither of these was associated with higher number of isolates which could indicate competitive advantage. Both were present only in *K. pneumoniae*, and relatively few strains (7; 5.7%) encoded both siderophores. This analysis, showing parts of the virulence plasmid as well as high levels of resistance in the same strains, suggests that the evolutionary convergence between virulent and resistant strains was already occurring within these historical isolates⁶³. Other studies have also suggested that the accumulation of virulence factors is an ongoing process in *K. pneumoniae*, and more work needs to be done in longitudinal studies to monitor the changes of the virulence potential of sequence types over time⁷¹. This is particularly relevant given recent observations of high mortality associated with strains with reduced virulence potential⁷⁰.

The collection is derived from a time (2010 to 2012) where carbapenem resistance was still at comparatively low levels and NDM-1 had only recently been introduced into the *K. pneumoniae* population, whereas now we see very high prevalence in LMIC as well as high-income countries^{2,26,27}. Even though the KPC carbapenemase enzyme was described first in 1996 in the USA⁷², it spread comparatively slowly but has become highly problematic in the American continent and southeast Asia (China, Taiwan, Singapore)⁵⁶. NDM-1 was first reported in 2008²², but spread rapidly within the Indian subcontinent, and across all continents by 2010^{73,74}. Our dataset contains several lineages with NDM-1, however these did not become globally spread high-risk lineages in the subsequent years. Acquisition of an NDM-1 plasmid *per se* is neither a sign that the plasmid will spread stably with the population, nor that this lineage has a major advantage over all carbapenemase-negative lineages, which likely reflects intrinsic factors that need to be present to provide an NDM-1 positive lineage a competitive advantage over other ESBL-positive lineages as observed in Nepal. The apparent instability of the NDM-1 locus in the isolates we examined, most likely through plasmid instability, suggests that laboratory analysis of this important resistance must be conducted carefully.

Monitoring resistance to inhibitor-based treatments is crucial as this combination treatment is a common clinical response to ESBL infections²⁹, and there are ongoing efforts to develop inhibitors against beta-lactamases and more recently also carbapenemases^{75–77}. There is considerable subtlety in the evolution of drug resistance, and especially for beta-lactams a variety of activity spectra across resistance enzymes in one strain is likely to be of importance, as enzymes with lower activity spectra are often more resistant to inhibitors than broad-spectrum ESBL enzymes⁷⁸. Detailed analyses, including the fitness cost of resistance genes as indicated by the unstable nature of NDM-1 or, as contrasting example, the rapid spread and fixation of CTX-M-15, are crucial to our understanding of the different mechanisms of resistance and their fixation in a population.

Materials and Methods

Strain collection. The strains were derived from routine diagnostic bacterial cultures in the Microbiology Department of The Children's Hospital & The ICH Lahore, Lahore, Pakistan, during 2010–2012, where 13% of all laboratory-positive samples were *K. pneumoniae*. The strains were pre-selected for ESBL resistance, and a detailed description of the infections in the hospital during the time as well as specific data on patient symptoms and other metadata regarding the strains analysed here can be found in Ejaz *et al.*²³. The study was ethically approved by the institutional review board of The Children's Hospital, Lahore under No. 1625/PH&I. The studies were compliant with the ethics approval granted by the Institutional Review Board of the Lahore Children's Hospital, all the methods were performed in accordance with the standard ethical guidelines and do not include any human or animal trials.

PacBio DNA preparation and sequencing. DNA from overnight cultures was isolated by Phenol:Chloroform:Isoamyl Alcohol (25:24:1) and Chloroform:IAA (24:1) extractions using Phase-lock tubes (Qiagen) and re-dissolved in 10 mM Tris pH 7.4 buffer. Sequencing was performed on the PacBio RSII using P6/C4 sequencing chemistry, the library was prepared using the SMRTbell Template Prep Kit 1.0. Filtered sub-reads were generated with the pacbi-smrt software, and assembled with canu v1.1⁷⁹. The assemblies were then circularised using circulator v1.5.3⁸⁰ with canu as assembler, and polished with unicycler-polish v0.4.0⁸¹ and the Illumina reads of the respective sample. HE016 gave better results in assembling with the unicycler-hybrid assembler v0.4.0 combining the PacBio data with the Illumina reads⁸¹, and was subsequently circularised and polished as above. All assemblies were then annotated with PROKKA v1.11⁸².

Bacteria Mapping and Variant Detection. Mapping was performed against the chromosome of *Klebsiella pneumoniae* NTUH-K2044 (AP006725), a published whole genome from an outbreak in Nepal⁵⁹ and HE196 (this study). Sequence reads were mapped against the reference genome as indicated using SMALT⁸³ (v0.7.4) to produce a BAM file. Variation detection was performed using samtools mpileup v0.1.19 bcftools v0.1.19 to produce a BCF file of all variant sites.

Pan-genome analysis. The samples as indicated in the respective experiments were assembled and annotated as described above, and the GFF3 files generated by PROKKA v1.11⁸² were used as input for the pan-genome pipeline Roary⁵⁰ (v3.7.0) with a BLASTp percentage identity of 90% and a core definition of 99%. This gives a core gene alignment of 3486 genes for all strains from this study (Figs 1, 2). The core gene alignment was generated with mafft v7.205⁸⁴, SNPs were first extracted using snp-sites v2.3.2⁸⁵, and then a maximum likelihood tree with RAxML v8.2.8⁸⁶ was calculated. For the ST15 comparison, the pan-genome was calculated as described above, and gene presence/absence analysed further using R v3.4.2 and the resulting presence/absence matrix and phandango⁸⁷.

Phylogenetic analyses. Trees were calculated using RAxML⁸⁶ (v8.2.8) with the time-reversible GTR model and 100 bootstrap repeats. Tree demonstrations were prepared in itol⁸⁸, coverage plots were generated using bedtools genome coverage⁸⁹ and displayed using the R package ggplot v2.2.1⁹⁰. For whole-genome mapping trees, recombinant regions were removed using Gubbins⁹¹ (v1.4.9) and a maximum likelihood tree calculated with RAxML to obtain bootstrap support values, as described above, and sites with more than 5% N were not considered in the tree calculation.

Gene content analysis. For the determination of the antibiotic resistance profile, the virulence factor profile and plasmid types, we used ariba. The resistance, virulence and plasmid profiles were matched against the modified version of ARG-ANNOT⁹² available at the SRST2 website (<https://github.com/katholt/srst2/tree/master/data>; download date 02.10.2016), a dataset of virulence factors obtained from the *Klebsiella*-specific BIGSDB (<http://bigsdb.pasteur.fr/klebsiella/klebsiella.html>; download date 22.02.2016), and the PlasmidFinder database as implemented in ariba v2.10.0^{93,43}. Sequence gene profiles and types were determined using MLST check⁹⁴ comparing assembled genomes against the MLST database for *Klebsiella pneumoniae* (pubmlst.org/Klebsiellapneumoniae/). For further resistance determinants (SNP-based or porin inactivation), a database with the genes of interest was created as outlined in the ariba manual pages. The assignment of *bla*-SHV alleles was controlled based on amino acids using the beta-lactamase database <http://www.laced.uni-stuttgart.de>. Yersiniabactin and ICE alleles were annotated using kleborate⁴² (<https://github.com/katholt/Kleborate>). Capsule and LPS O-antigen was annotated using kleborate and a custom database of LPS O-antigens³⁷. The assignment to O3 subgroups was updated based on recent nomenclature (O3l/s; now a/b/c) using the online tool Kaptive⁴¹ (<http://kaptive.holtlab.net/>).

Minimum inhibitory concentrations (MIC) measurements. All isolates were analysed using the VITEK 2 system (bioMérieux, UK). In brief, suspensions of colonies were made in 0.45% saline solution from growth on iso-sensitest agar (Thermo Scientific Oxiod, UK), adjusted to a turbidity equivalent to that of a 0.5 McFarland standard and used to load the test cards, using manufacturer's instructions. The *Enterobacteriaceae* AST-N350 cards was automatically filled, sealed and inserted into the VITEK 2 reader-incubator module (incubation temperature 37 °C), and fluorescence measurements were performed every 15 min for up to 18 h. For detailed analysis of several NDM-1 containing strains, the MICs for meropenem were also assessed manually following the protocol of Wiegand 2008⁹⁵ (Table 1).

Data Availability

Details on strains and accession numbers can be found in Table S1 (Vitek resistance data), Table S2 (accessions and further information on isolates from Pakistan) and Table S3 (additional published strains used for the analyses). Data provided for the Pakistan samples includes accessions for raw sequence data as well as assemblies of both Illumina and PacBio sequenced genomes and patient metadata. All relevant data is in the manuscript and supporting material, alignment and tree files are available for download at figshare <https://figshare.com/s/cdadd-b659a6e178102df>.

References

1. WHO. Antimicrobial resistance: global report on surveillance 2014. (2014).
2. WHO. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics (2017).
3. Wyres, K. L. & Holt, K. E. *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Curr. Opin. Microbiol.* **45**, 131–139 (2018).
4. Karlowsky, J. A. *et al.* In Vitro Activity of Imipenem–Relebactam against Clinical Isolates of Gram–Negative Bacilli Isolated in Hospital Laboratories in the United States as Part of the SMART 2016 Program. *Antimicrob. Agents Chemother.* **62** (2018).
5. Investigators of the Delhi Neonatal Infection Study (DeNIS) collaboration. Characterisation and antimicrobial resistance of sepsis pathogens in neonates born in tertiary care centres in Delhi, India: a cohort study. *Lancet Glob. Health* **4**, e752–760 (2016).
6. Kabwe, M. *et al.* Etiology, Antibiotic Resistance and Risk Factors for Neonatal Sepsis in a Large Referral Center in Zambia. *Pediatr. Infect. Dis. J.* **35**, e191–198 (2016).
7. Shehab El-Din, E. M. R., El-Sokkary, M. M. A., Bassiouny, M. R. & Hassan, R. Epidemiology of Neonatal Sepsis and Implicated Pathogens: A Study from Egypt. *BioMed Res. Int.* **2015**, 509484 (2015).
8. Boulos, A., Rand, K., Johnson, J. A., Gautier, J. & Koster, M. Neonatal Sepsis in Haiti. *J. Trop. Pediatr.* **63**, 70–73 (2017).
9. Ghotaslou, R., Ghorashi, Z. & Nahaei, M.-R. *Klebsiella pneumoniae* in neonatal sepsis: a 3-year-study in the pediatric hospital of Tabriz, Iran. *Jpn. J. Infect. Dis.* **60**, 126–128 (2007).
10. Lee, C.-R. *et al.* Global Dissemination of Carbapenemase-Producing *Klebsiella pneumoniae*: Epidemiology, Genetic Context, Treatment Options, and Detection Methods. *Front. Microbiol.* **7**, 895 (2016).
11. Wyres, K. L. & Holt, K. E. *Klebsiella pneumoniae* Population Genomics and Antimicrobial-Resistant Clones. *Trends Microbiol.* **24**, 944–956 (2016).
12. Martínez-Romero, E. *et al.* Draft Genome Sequences of *Klebsiella variicola* Plant Isolates. *Genome Announc.* **3** (2015).
13. Maatallah, M. *et al.* *Klebsiella variicola* is a frequent cause of bloodstream infection in the stockholm area, and associated with higher mortality compared to *K. pneumoniae*. *PLoS One* **9**, e113539 (2014).
14. Di, D. Y. W., Jang, J., Unno, T. & Hur, H.-G. Emergence of *Klebsiella variicola* positive for NDM-9, a variant of New Delhi metallo- β -lactamase, in an urban river in South Korea. *J. Antimicrob. Chemother.* **72**, 1063–1067 (2017).
15. Andrade, B. G. N., de Veiga Ramos, N., Marin, M. F. A., Fonseca, E. L. & Vicente, A. C. P. The genome of a clinical *Klebsiella variicola* strain reveals virulence-associated traits and a p19-like plasmid. *FEMS Microbiol. Lett.* **360**, 13–16 (2014).
16. Garza-Ramos, U. *et al.* Draft Genome Sequence of a Hypermucoviscous Extended-Spectrum- β -Lactamase-Producing *Klebsiella quasipneumoniae* subsp. *similipneumoniae* Clinical Isolate. *Genome Announc.* **4** (2016).
17. Shankar, C. *et al.* Draft genome of a hypervirulent *Klebsiella quasipneumoniae* subsp. *similipneumoniae* with novel sequence type ST2320 isolated from a chronic liver disease patient. *J. Glob. Antimicrob. Resist.* **9**, 30–31 (2017).
18. Breurec, S. *et al.* Liver Abscess Caused by Infection with Community-Acquired *Klebsiella quasipneumoniae* subsp. *quasipneumoniae*. *Emerg. Infect. Dis.* **22**, 529–531 (2016).
19. Woodford, N., Turton, J. F. & Livermore, D. M. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiol. Rev.* **35**, 736–755 (2011).
20. Doi, Y. *et al.* Gram-Negative Bacterial Infections: Research Priorities, Accomplishments, and Future Directions of the Antibacterial Resistance Leadership Group. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **64**, S30–S35 (2017).
21. Pitout, J. D. D. & Laupland, K. B. Extended-spectrum beta-lactamase-producing *Enterobacteriaceae*: an emerging public-health concern. *Lancet Infect. Dis.* **8**, 159–166 (2008).
22. Yong, D. *et al.* 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 (2009).
23. Ejaz, H. *et al.* Phylogenetic Analysis of *Klebsiella pneumoniae* from Hospitalized Children, Pakistan. *Emerg. Infect. Dis.* **23**, 1872–1875 (2017).
24. Bouxom, H., Fournier, D., Bouiller, K., Hocquet, D. & Bertrand, X. Which non-carbapenem antibiotics are active against extended-spectrum β -lactamase-producing *Enterobacteriaceae*? *Int. J. Antimicrob. Agents* **52**, 100–103 (2018).
25. Haeggman, S., Löfdahl, S., Paaup, A., Verhoef, J. & Brisse, S. Diversity and evolution of the class A chromosomal beta-lactamase gene in *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* **48**, 2400–2408 (2004).
26. Navon-Venezia, S., Kondratyeva, K. & Carattoli, A. *Klebsiella pneumoniae*: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol. Rev.* **41**, 252–275 (2017).
27. Grundmann, H. *et al.* Occurrence of carbapenemase-producing *Klebsiella pneumoniae* and *Escherichia coli* in the European survey of carbapenemase-producing *Enterobacteriaceae* (EuSCAPE): a prospective, multinational study. *Lancet Infect. Dis.* **17**, 153–163 (2017).
28. Kernéis, S. *et al.* Cefoxitin as a carbapenem-sparing antibiotic for infections caused by extended-spectrum beta-lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae*. *Infect. Dis. Lond. Engl.* **47**, 789–795 (2015).
29. Schuetz, A. N., Reyes, S. & Tamma, P. D. Point-Counterpoint: Piperacillin–Tazobactam Should Be Used To Treat Infections with Extended-Spectrum-Beta-Lactamase-Positive Organisms. *J. Clin. Microbiol.* **56** (2018).
30. Poirel, L., Dortet, L., Bernabeu, S. & Nordmann, P. Genetic features of blaNDM-1-positive. *Enterobacteriaceae*. *Antimicrob. Agents Chemother.* **55**, 5403–5407 (2011).
31. Polsfuss, S. *et al.* Practical approach for reliable detection of AmpC beta-lactamase-producing *Enterobacteriaceae*. *J. Clin. Microbiol.* **49**, 2798–2803 (2011).
32. Dubois, V. *et al.* Molecular and biochemical characterization of SHV-56, a novel inhibitor-resistant beta-lactamase from *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* **52**, 3792–3794 (2008).
33. Prinarakis, E. E., Miriagou, V., Tzelepi, E., Gazouli, M. & Tzouveleki, L. S. Emergence of an inhibitor-resistant beta-lactamase (SHV-10) derived from an SHV-5 variant. *Antimicrob. Agents Chemother.* **41**, 838–840 (1997).
34. Manageiro, V. *et al.* Characterization of the inhibitor-resistant SHV β -lactamase SHV-107 in a clinical *Klebsiella pneumoniae* strain coproducing GES-7 enzyme. *Antimicrob. Agents Chemother.* **56**, 1042–1046 (2012).
35. Mendonça, N., Ferreira, E. & Louro, D. ARSIP Participants & Caniça, M. Molecular epidemiology and antimicrobial susceptibility of extended- and broad-spectrum beta-lactamase-producing *Klebsiella pneumoniae* isolated in Portugal. *Int. J. Antimicrob. Agents* **34**, 29–37 (2009).

36. Winkler, M. L., Papp-Wallace, K. M., Taracila, M. A. & Bonomo, R. A. Avibactam and inhibitor-resistant SHV β -lactamases. *Antimicrob. Agents Chemother.* **59**, 3700–3709 (2015).
37. Follador, R. *et al.* The diversity of *Klebsiella pneumoniae* surface polysaccharides. *Microb. Genomics* **2**, e000073 (2016).
38. Wyres, K. L. *et al.* Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb. Genomics* **2**, e000102 (2016).
39. Jenney, A. W. *et al.* Seroepidemiology of *Klebsiella pneumoniae* in an Australian Tertiary Hospital and its implications for vaccine development. *J. Clin. Microbiol.* **44**, 102–107 (2006).
40. Pennini, M. E. *et al.* Immune stealth-driven O2 serotype prevalence and potential for therapeutic antibodies against multidrug resistant *Klebsiella pneumoniae*. *Nat. Commun.* **8**, 1991 (2017).
41. Wick, R. R., Heinz, E., Holt, K. E. & Wyres, K. L. Kaptive Web: User-Friendly Capsule and Lipopolysaccharide Serotype Prediction for *Klebsiella* Genomes. *J. Clin. Microbiol.* **56** (2018).
42. Lam, M. M. C. *et al.* Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in *Klebsiella pneumoniae* populations. *Microb. Genomics*, <https://doi.org/10.1099/mgen.0.000196> (2018).
43. Hunt, M. *et al.* ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb. Genomics* **3**, e000131 (2017).
44. Yu, W.-L., Lee, M.-F., Tang, H.-J., Chang, M.-C. & Chuang, Y.-C. Low prevalence of *rmpA* and high tendency of *rmpA* mutation correspond to low virulence of extended spectrum β -lactamase-producing *Klebsiella pneumoniae* isolates. *Virulence* **6**, 162–172 (2015).
45. Wu, K.-M. *et al.* Genome sequencing and comparative analysis of *Klebsiella pneumoniae* NTUH-K2044, a strain causing liver abscess and meningitis. *J. Bacteriol.* **191**, 4492–4501 (2009).
46. Shon, A. S., Bajwa, R. P. S. & Russo, T. A. Hypervirulent (hyper-mucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. *Virulence* **4**, 107–118 (2013).
47. Chen, Y.-T. *et al.* Sequencing and analysis of the large virulence plasmid pLVPK of *Klebsiella pneumoniae* CG43. *Gene* **337**, 189–198 (2004).
48. M C Lam, M. *et al.* Population genomics of hypervirulent *Klebsiella pneumoniae* clonal-group 23 reveals early emergence and rapid global dissemination. *Nat. Commun.* **9**, 2703 (2018).
49. 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).
50. Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinforma. Oxf. Engl.* **31**, 3691–3693 (2015).
51. González, L. J. *et al.* Membrane anchoring stabilizes and favors secretion of New Delhi metallo- β -lactamase. *Nat. Chem. Biol.* **12**, 516–522 (2016).
52. Huang, T.-W. *et al.* Copy Number Change of the NDM-1 sequence in a multidrug-resistant *Klebsiella pneumoniae* clinical isolate. *PLoS One* **8**, e62774 (2013).
53. Wang, X. *et al.* High rate of New Delhi metallo- β -lactamase 1-producing bacterial infection in China. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **56**, 161–162 (2013).
54. Carattoli, A. Resistance plasmid families in *Enterobacteriaceae*. *Antimicrob. Agents Chemother.* **53**, 2227–2238 (2009).
55. Hu, H. *et al.* Novel plasmid and its variant harboring both a bla(NDM-1) gene and type IV secretion system in clinical isolates of *Acinetobacter lwoffii*. *Antimicrob. Agents Chemother.* **56**, 1698–1702 (2012).
56. Nordmann, P., Cuzon, G. & Naas, T. The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. *Lancet Infect. Dis.* **9**, 228–236 (2009).
57. Wailan, A. M. *et al.* Mechanisms Involved in Acquisition of blaNDM Genes by IncA/C2 and IncFIIY Plasmids. *Antimicrob. Agents Chemother.* **60**, 4082–4088 (2016).
58. Wailan, A. M. *et al.* Genomic Characteristics of NDM-Producing *Enterobacteriaceae* Isolates in Australia and Their blaNDM Genetic Contexts. *Antimicrob. Agents Chemother.* **60**, 136–141 (2016).
59. Stoesser, N. *et al.* 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. *Antimicrob. Agents Chemother.* **58**, 7347–7357 (2014).
60. Chung The, H. *et al.* A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of *Klebsiella pneumoniae*. *EMBO Mol. Med.* **7**, 227–239 (2015).
61. Moradigaravand, D., Martin, V., Peacock, S. J. & Parkhill, J. Evolution and Epidemiology of Multidrug-Resistant *Klebsiella pneumoniae* in the United Kingdom and Ireland. *mBio* **8** (2017).
62. Bowers, J. R. *et al.* Genomic Analysis of the Emergence and Rapid Global Dissemination of the Clonal Group 258 *Klebsiella pneumoniae* Pandemic. *PLoS One* **10**, e0133727 (2015).
63. Holt, K. E. *et al.* Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc. Natl. Acad. Sci. USA* **112**, E3574–3581 (2015).
64. Henson, S. P. *et al.* Molecular epidemiology of *Klebsiella pneumoniae* invasive infections over a decade at Kilifi County Hospital in Kenya. *Int. J. Med. Microbiol. IJMM* **307**, 422–429 (2017).
65. Long, S. W. *et al.* Whole-Genome Sequencing of Human Clinical *Klebsiella pneumoniae* Isolates Reveals Misidentification and Misunderstandings of *Klebsiella pneumoniae*, *Klebsiella variicola*, and *Klebsiella quasipneumoniae*. *mSphere* **2** (2017).
66. *The vaccine book* (Elsevier, AP, 2016).
67. Lob-Levyt, J. Vaccine coverage and the GAVI Alliance Immunization Services Support initiative. *Lancet Lond. Engl.* **373**, 209 (2009).
68. Mostowy, R. J. *et al.* Pneumococcal Capsule Synthesis Locus cps as Evolutionary Hotspot with Potential to Generate Novel Serotypes by Recombination. *Mol. Biol. Evol.* **34**, 2537–2554 (2017).
69. Corander, J. *et al.* Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nat. Ecol. Evol.* **1**, 1950–1960 (2017).
70. Gu, D. *et al.* A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. *Lancet Infect. Dis.* **18**, 37–46 (2018).
71. Turton, J. F. *et al.* Virulence genes in isolates of *Klebsiella pneumoniae* from the UK during 2016, including among carbapenemase gene-positive hypervirulent K1-ST23 and ‘non-hypervirulent’ types ST147, ST15 and ST383. *J. Med. Microbiol.* **67**, 118–128 (2018).
72. Yigit, H. *et al.* Novel carbapenem-hydrolyzing beta-lactamase, KPC-1, from a carbapenem-resistant strain of *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* **45**, 1151–1161 (2001).
73. Rolain, J. M., Parola, P. & Cornaglia, G. New Delhi metallo-beta-lactamase (NDM-1): towards a new pandemic? *Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis.* **16**, 1699–1701 (2010).
74. Kumarasamy, K. K. *et al.* Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect. Dis.* **10**, 597–602 (2010).
75. Drawz, S. M. & Bonomo, R. A. Three decades of beta-lactamase inhibitors. *Clin. Microbiol. Rev.* **23**, 160–201 (2010).
76. Papp-Wallace, K. M. *et al.* Relebactam Is a Potent Inhibitor of the KPC-2 β -Lactamase and Restores Imipenem Susceptibility in KPC-Producing *Enterobacteriaceae*. *Antimicrob. Agents Chemother.* **62** (2018).
77. Jorgensen, S. C. J. & Rybak, M. J. Meropenem and Vaborbactam: Stepping up the Battle against Carbapenem-resistant *Enterobacteriaceae*. *Pharmacotherapy* **38**, 444–461 (2018).
78. Bradford, P. A. Extended-spectrum beta-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clin. Microbiol. Rev.* **14**, 933–951, table of contents (2001).

79. Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
80. Hunt, M. *et al.* Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* **16**, 294 (2015).
81. 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).
82. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinforma. Oxf. Engl.* **30**, 2068–2069 (2014).
83. Ponstingl, H. *SMALT: A mapper for DNA sequencing reads* (2014).
84. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
85. Page, A. J. *et al.* SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genomics* **2**, e000056 (2016).
86. Stamatakis, A. Using RAxML to Infer Phylogenies. *Curr. Protoc. Bioinforma.* **51**(6), 14.1–14 (2015).
87. Hadfield, J. *et al.* Phandango: an interactive viewer for bacterial population genomics. *Bioinforma. Oxf. Engl.*, <https://doi.org/10.1093/bioinformatics/btx610> (2017).
88. Letunic, I. & Bork, P. Interactive tree of life (iTOL)v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–245 (2016).
89. Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinforma.* **47**(11), 12.1–34 (2014).
90. Wickham, H. *ggplot2 Elegant Graphics for Data Analysis* (Springer, 2016).
91. 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).
92. 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).
93. 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).
94. Page, A. J., Taylor, B. & Keane, J. A. Multilocus sequence typing by blast from de novo assemblies against PubMLST. *The Journal of Open Source Software*, <https://doi.org/10.21105/joss.00118> (2016).
95. Wiegand, I., Hilpert, K. & Hancock, R. E. W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175 (2008).
96. Toleman, M. A., Spencer, J., Jones, L. & Walsh, T. R. bla_{NDM-1} is a chimera likely constructed in *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **56**, 2773–2776 (2012).

Acknowledgements

The authors thank Karen Oliver for help with the PacBio sequencing, and Andrew Page for helpful discussions on PacBio assembly strategies, Simon Harris for helpful discussions on the manuscript, as well as the Sanger Institute Pathogens Informatics group. We thank Dr. Badr Alzahrani (Jouf University) for comments on the manuscript. This work was supported by the NHMRC (Program Grant 1092262), the Wellcome Trust (206194), and the Higher Education Commission of Pakistan and The Children's Hospital & The ICH, Lahore, Pakistan. H.E. was supported by a scholarship from Higher Education Commission (HEC) Pakistan under the International Research Support Initiative Program (IRSIP). GD is supported by the NIMR Cambridge BRC AMR theme.

Author Contributions

E.H., H.E., I.u.H. and R.A.S. designed the study. J.B.S., N.W., S.G., D.P., J.W., H.C. performed laboratory experiments. E.H. performed all bioinformatics and analysed the data. E.H. and R.A.S. wrote and G.D. contributed to, the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-38943-7>.

Competing Interests: The authors declare no competing interests.

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) 2019