Emergence of carbapenem, beta-lactamase inhibitor and cefoxitin resistant lineages from a background of ESBL-producing *Klebsiella pneumoniae* and *K. quasipneumoniae* highlights different evolutionary mechanisms

Running title: Resistance to ESBL alternatives in *Klebsiella* spp.

Eva Heinz[1#]

Hasan Ejaz[2,3,4]

Josefin Bartholdson Scott[5]

Nancy Wang[3]

Shruti Guanjaran[3]

Derek Pickard[1]

Jonathan Wilksch[3]

Hanwei Cao[3]

Ikram-ul-Haq[6]

Gordon Dougan[1,5]

Richard A Strugnell[3#]

1: Infection Genomics Program, Wellcome Trust Sanger Institute, Hinxton, UK

2: Department of Clinical Laboratory Sciences, CAMS, Jouf University, Al-Jouf, Kingdom of Saudi Arabia

3: Department of Microbiology and Immunology, The University of Melbourne, at Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

4: Department of Microbiology, The Children's Hospital & The Institute of Child Health, Lahore, Pakistan

5: Department of Medicine, University of Cambridge, Cambridge, UK

6: Institute of Industrial Biotechnology, GC University, Lahore, Pakistan

#Correspondence:

R.A. Strugnell, Department of Microbiology and Immunology, The University of Melbourne, at Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; rastru@unimelb.edu.au

Eva Heinz, Infection Genomics Program, Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK; eva.heinz@sanger.ac.uk

Word count abstract: 247

Word count manuscript: 4773

Keywords

Klebsiella, pneumoniae, quasipneumoniae, cefoxitin, ESBL, NDM-1, ST15

Importance

Carbapenem-resistant and extended-spectrum beta-lactamase (ESBL) carrying *Enterobacteriaceae* were recently highlighted as critical priority fo the development of new treatments by the WHO. *Klebsiella pneumoniae* is a member of the *Enterobacteriaceae* and has seen a dramatic rise in clinical relevance due to its uncanny ability to accumulate multidrug-resistance plasmids. We present a detailed analysis of a set of ESBL-resistant *K. pneumoniae* clinical isolates, and our high-resolution whole-genome sequence analyses highlight that acquisition of drug resistances is not a one-way street in *K. pneumoniae*, but a highly dynamic process of gain and loss, and that the most successful lineages in the clinic are not necessarily the most resistant or most virulent ones. Analysis of the virulence potential also shows that these strains harbour some, but not all, hallmarks of hypervirulent strains, emphasizing that it is not a clear distinction between hypervirulent and other strains, but equally in flux.

Abstract

Klebsiella pneumoniae is recognised as a major threat to public health, with increasing emergence of multidrug-resistant lineages including strains resistant to all available antibiotics. We present an in-depth analysis of 178 extended-spectrum beta-lactamase (ESBL)-producing *Klebsiella* strains, with a high background diversity and two dominant lineages, as well as several equally resistant lineages with less prevalence. Neither the overall resistance profile nor the virulence factors explain the prevalence of some lineages; we observe several putative hypervirulence factors across the population, including a reduced virulence plasmid, but this does not correlate with expansion of one or few highly virulent and resistant lineages. Phenotypic analysis of the profiles of resistance traits shows that the vast majority of the phenotypic resistance profiles can be explained by detailed genetic analyses. The main discrepancies are observed for beta-lactams combined with beta-lactamase inhibitors, where most, but not all, resistant strains carry a carbapenemase or *ampC*. Complete genomes for six selected strains, including three of the 21 carbapenem-resistant ones, are reported, which give detailed insights into the early evolution of the *bla-NDM-1* enzyme, a carbapenemase that was first reported in 2009 and is now globally distributed. Whole-genome based high-resolution analyses of the dominant lineages suggests a very dynamic picture of gene transfer and selection, with phenotypic changes due to plasmid acquisition and chromosomal changes, and emphasize the need to monitor the bacteria at high resolution to understand the rise of high-risk clones, which cannot be explained by obvious differences in resistance profiles or virulence factors.

Introduction

The past four decades have seen a continuous escalation of bacterial pathogens acquiring resistance mechanisms against antimicrobials, and especially antimicrobial resistance determinants associated with mobile elements have spread with exponentially increasing speed across the globe (1, 2). A particularly successful pathogen in this group is *Klebsiella pneumoniae*, which was formerly known as a major cause of infections in neonates, especially in developing countries (3-6) and community-acquired and nosocomial infections in immunocompromised patients (7-9). The acquisition of extended-spectrum beta-lactamases (ESBLs) rapidly increased in *Klebsiella* spp. from the 1990s, particularly in hospital isolates (10, 11). In 2009 there was first description of NDM-1, a metallo-betalactamase which hydrolyzes carbapenems and can escape beta-lactamase inhibitors, which were until then the main alternative treatments for third-generation cephalosporin-resistant bacteria (12). NDM-1 appears to have originated as a chromosomal gene fusion in *Acinetobacter* (13), but rapidly disseminated around the globe as well as across several species (e.g. *E. coli*, *Provatella;* 14).

We report a high-resolution analysis of a set of ESBL-carrying *K. pneumoniae* isolates from a tertiary care children's hospital in Lahore, Pakistan, collected between 2010 and 2012 (15). The selected strains were all ESBL producers, and our population study shows that whilst we observe two dominant lineages, no clear advantage for these lineages can be seen either in the resistance profile, or the virulence factor repertoire, despite some

111 strains carrying markers of hypervirulent *Klebsiella;* these markers were not
112 widely disseminated within the dataset, but occurred intermittently.
113
114 The main treatment alternatives for ESBL-producing
115 *Enterobacteriaceae* are i) cefoxitin, which is resistant to ESBL enzymes but
116 sensitive to AmpC; ii) beta-lactams in combination with beta-lactamase
117 inhibitors, with two widely used combinations piperacillin/tazobactam and
118 amoxicillin/clavulanic acid; and iii) carbapenems. Comparing phenotypic with
119 predicted resistance profiles showed that resistance to carbapenems
120 correlated with the presence of carbapenemase. Resistance against either or
121 both of the alternative treatments (cefoxitin or the use of beta-lactamase
122 inhibitors) could be explained in some cases through the presence of NDM-1,
123 whereas for several strains, the resistance mechanisms are still unclear.
124
125 We furthermore have added complete genome sequences of six
126 selected strains including one representative for each lineage harbouring the
127 *bla-NDM-1* gene; this gives us important insights into the early spread of
128 carbapenemases and the associated plasmid diversity. We note no
129 differences in the *bla-CTX* or *bla-TEM* genes, but high variation in allelic
130 distribution of *bla-OXA* and *bla-SHV*, including several ESBL variants and
131 additional plasmid-encoded copies of *bla-SHV*. Our study highlights the
132 mobile and plastic nature of *Klebsiella* resistance determinants, which vary
133 widely, even within strains selected for similar resistance traits from a single
134 hospital, and within the same sequence type. This dynamic relationship
135 between a now key opportunistic pathogen and a very large pool of mobile

and readily transferrable resistance mechanisms is very sobering and is in stark contrast to observations of other resistant pathogens, where acquisition of an advantageous resistance/mobile element leads to rapid, clonal spread of a distinct lineage linked to the resistance element (16-18).

Results

Strain collection shows high variety of sequence types and virulence determinants

The strains were collected between 2010-2012 through the routine microbiological screening in The Children's Hospital, Lahore, Pakistan, and were pre-selected for ESBL presence through E-test (15). The population structure of our dataset includes a high diversity of sequence types (15; Fig. 1A, 1B), as well as a high number of isolates from *K. quasipneumoniae*, a subspecies that was previously thought to be less virulent (Fig. 1; 15). In contrast, no *K. variicola* was identified, a closely related species which is also routinely identified as *K. pneumoniae* with non-distinguishable clinical symptoms. To analyse the molecular epidemiology in finer detail, we identified the capsule (K) and lipopolysaccharide O-antigen types based on their respective operons in the genomes, which are important markers to monitor the epidemiology of *Klebsiella* spp. Recent work (19, 20) highlighted the considerable diversity found among different *Klebsiella* capsule types (21), and this is also reflected in the collection reported here, where the diversity of sequence types and capsule types is equivalent, with a different capsule type for almost each sequence type (Fig. 1C, Table S1). In contrast to capsule loci,

there is less diversity of LPS O-antigens with two main types in our collection; however these are not the usually reported main types (i.e. not O1, O2 and O3; 20, 22) but include less common types. LPS serotype 05 was the second most common, present in 68 of our isolates across 3 sequence types (Fig. 1).

With the collection of strains there is an almost ubiquitous presence of heavy metal resistance (silver, *sil*; copper, *pco*) as well as the type 3 fimbriae (i.e. *mrk*; Fig. S1). The genes encoding the siderophores aerolysin and yersiniabactin are distributed only partially throughout the collection (Fig. S1; 23), and no salmochelin or colibactin encoding genes were identified. Our collection encodes for six different types of yersiniabactin in eight sequence types (23), including the less common combination of ybt8 with ICEKp3 (ST423), and one sequence type (ST48) with three different variants (Fig. S1), highlighting the wide distribution of the different siderophore types even within a confined setting already in these historical isolates.

We identified a high number of genes predicted as pseudogenes by ariba (Fig. S2A), including the virulence gene *rmpA2*, the capsule upregulator (Fig. S2B), as well as the fluoroquinolone resistance gene *qnrB* (Fig. S2C). RmpA is inactivated through a frameshift in either the poly-G and poly-A regions (Fig. S2B), as described previously (24). RmpA and aerobactin are indicative of the presence of the *Klebsiella* virulence plasmid (25-28), and we therefore compared several virulence plasmids found with our data, and we performed comparisons against the plasmid pSGH10, which is present in a representative model strain of the hypervirulent clade CG23 (28). The plasmid

is not evenly spread across the sequence types from Lahore and lacks salmochelin; both when compared mapping (Fig. 2A) or assemblies (Fig. 2B). The presence or absence of these putative key virulence factors (15) could not be correlated with clinical outcome (details Table 2).

Most multidrug-resistance in *Klebsiella* is caused by acquired resistances on mobile elements, which is clearly reflected in our dataset (Fig. S1), with the expected high number of resistances against possible first-line drug classes aminoglycosides, fluoroquinolones, sulphonamides and beta-lactams, noting that the isolates were selected for ESBL expression, as well as other antimicrobials (Fig. S1).

Comparison of predicted to measured MICs highlights importance of genetic background and potential unknown resistance determinants

Measured MIC values of widely-used antimicrobials of our strain collection generally matched the genomics-predicted resistances (Table S3, Fig. S3); intrinsic resistance mechanisms such as deactivation of porins, upregulation of efflux pumps or resistance genes, might account for some of the non-matching results (29-31). Resistance against aminoglycosides is conferred almost exclusively by AAC3 or AAC6 (found in 87% of all isolates); however, it is noteworthy that despite some of the widespread key integrons in *Klebsiella* spp. usually including aminoglycoside resistance (32, 33) there is still sensitivity against this class in our group of clinical isolates. Thus, this

antibiotic class might still provide a treatment option for some ESBL-positive strains.

Closer investigation of the potential loss of fluoroquinolone resistance via the predicted *qnrB* pseudogenes (Fig. S2A) showed that it might reflect a functional gene with an alternative start codon (Fig. S2C), which however is recognised as truncated due to its slightly shorter sequence at the N-terminus. The full-length sequence found in the CARD and ARGANNOT database (accession ABC86904.1) contains 12 amino acids at the N-terminus. A frameshift mutation present in the *qnrB* coding sequences of our strain set (Fig. S2A) does not affect the sequence after the alternative start codon. Comparison to the structural information of the functional QnrB (34) shows that the N-terminal structural elements are conserved despite the shorter N-terminus (Fig. S2C; coloured boxes indicate structural elements as in Figure A1 from 34); the remainder of the predicted proteins in our data are 100% identical with the reference sequences. We also note that the sequence in Genbank has been updated (ABC86904.2), which renders it the same length as the more recently described qnrB2 (35; APU91821.1, 5 amino acids difference), and we therefore assume the putatively shortened QnrB to be functional and to confer quinolone resistance in the strains reported here. Only few lineages show additional resistance to fluoroquinolones through mutations in *gyrA* and *parC*.

Alternative beta-lactam treatments for ESBL-positive organisms

Klebsiella spp. harbour two intrinsic enzymes related to low-level beta-lactam resistance; AmpH, an AmpC-related enzyme functioning as penicillin-

235 binding protein; and a chromosomally integrated beta-lactamase of the SHV,
236 OKP or LEN family for *K. pneumoniae*, *quasipneumoniae* and *variicola*,
237 respectively (36). In addition to these chromosomal enzymes, we detected
238 CTX-M-15, three OXA and three TEM variants which however result in the
239 identical amino acid sequence, VEB-5; the AmpC enzymes CMY-6 and CMY-
240 18, and NDM-1 carbapenemases as well as additional copies of SHV (Fig. 3).
241 Sensitivity was observed among a high number of isolates for cefoxitin
242 and piperacillin-tazobactam; this is expected as cefoxitin is generally
243 insensitive to ESBL (e.g. *bla-CTX-M-15*, which is present in all but one of our
244 isolates), whereas resistances against piperacillin are widespread, but the
245 beta-lactamase inhibitor tazobactam is still usually effective. Resistance
246 against cefoxitin can be explained in most cases by the presence of the
247 carbapenemase *bla-NDM-1* which hydrolyses cefoxitin, the *bla-CMY* AmpC,
248 or both; only three strains (HE021, HE205, HE206; Fig. 3) seem to confer
249 resistance due to other factors. A pan-genome analysis of the *K.*
250 *quasipneumoniae* lineage reveals a number of differences which might
251 tangentially affect drug sensitivity such as transposases and hypothetical
252 proteins (likely phage-derived following comparisons to public databases),
253 however the only clear genetic difference between HE205 and HE206 is an
254 iron uptake system (*fec*) that is absent from the two former strains (Fig. S4).
255
256 It is recognised that sensitivity against piperacillin/tazobactam (TZA)
257 often coincides with lower MIC values against amoxicillin/clavulanate (Fig. S4;
258 Table S2), even though, based on a phenotypic sensitive/resistant evaluation,
259 these strains are assessed as resistant. More detailed analysis of the

encoded SHV/OKP *bla* genes revealed that several *K. quasipneumoniae* strains (HE031, HE029, HE025, HE034, HE023; ST334) gained a *bla-SHV* gene in addition to the intrinsic *bla-OKP*, and the contigs encoding the *bla-SHV* copies are similar to plasmid sequences based on Blast data from public websites, and a plasmid location is also supported by the unique plasmid replicon profile of these strains (Fig. S1). These putatively plasmid-borne *blaSHV* genes encode for a sequence with the mutations G238S and E240K, which correlate with TZA resistance, or TZA-intermediate sensitivity (Fig. 3; 37-39). We also observed several sequences in our dataset encoding the L35Q mutation; however this only provides a subtle increase in TZA MICs (37, 40, 41), and no mutation was found at S130 (42).

The emergence of NDM-1 on mobile elements

The MIC data highlights that the same gene (NDM-1) might cause different resistance phenotypes in different genetic backgrounds (Table 1). This could be due to both the genetic background provided by the bacterial chromosome, as well as the plasmid harbouring the gene. To identify the number of different plasmids that were present in our dataset, and to test how many of them are still circulating as reported in recent literature, we performed PacBio sequencing of representative isolates (Table 1, S1).

Since some NDM-1 genes detected in Illumina reads could not be confirmed by PacBio sequencing, and were only present in low coverage (Table 1), we performed additional MIC tests for meropenem (E-tests, see methods; Table 1). This confirms that the low-coverage enzymes only resulted in lower MICs read as sensitive, which can indicate either loss of the

plasmid carrying NDM-1, or of a highly mobile cassette carrying NDM-1. The unstable nature of NDM-1 has been observed previously in several independent reports, where loss can occur already after two generations of re-growth, and even under meropenem selection (43-45). The 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 presence of several resistance genes predicted at low abundance and absent in the PacBio assemblies, potentially indicating the unstable nature of the plasmid carrying the resistance (Fig. 4).

In addition, replicon analysis and mapping of the resistance and virulence genes back to the plasmids across all strains where the genomes were resolved with PacBio indicates the involvement of several different plasmids, showing highly dynamic profiles between strains even within closely related isolates (Fig. 4). Almost fixed in the population across all sequence types are as expected the incompatibility groups F and H; with only few isolates showing rarer types such as N and R (Fig. S1, Fig. 4).

Two of the plasmids show the original, more complete cassette including the GroEL/S from the original *Acinetobacter* construct (Fig. S5), whilst in one case a smaller portion of this gene region is conserved; this variation in the different forms of NDM-1 has been described in other studies (46, 47).

Comparison ST15 isolates from Lahore with outbreak in Nepal

To gain further insights into the diversity within dominant lineages in Fig. 1, we analysed the differences between the two ST15 lineages reported

in Nepal (48, 49), and in this study (Table S3). This included assessing 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 isolate against a PacBio-sequenced reference sequence (accession number CP008929; Fig. S6, Fig. 5), derived from the Nepal outbreak which has the complete genome (49). A comparison of the plasmid content showed entirely different compositions, with only one plasmid that was commonly conserved (Fig. S7). Analysing the capsule operons, we observed the same novel capsule type as described for the Nepal isolates (48; Fig. S6); further highlighting the likely close relationship of this virulent lineage that seems to be spreading throughout this area of the world.

The phylogeny indicated additional structure in the isolates from Lahore, which was confirmed by analysing the mapping results in more detail (Fig. 5). The mapping-based approach shows that, even within the tight cluster of the isolates sequenced in this analysis, at least three different patterns were identified. These altered patterns lack three regions found in the reference strain from Nepal (Fig. S7) and a detailed overview of resistance, virulence gene profiles, plasmid replicons and surface determinants (capsule and O-antigen type) is given in Fig. S6.

The first region is a gain in the Nepal lineage as it is also absent from more distantly related members of ST15, and covers the ICE element carrying yersiniabactin and other virulence determinants as described originally (Fig. S6, Fig. 5; 48, 49). The other two regions were selectively lost in parts of the ST15 lineage from the Lahore strain dataset (Fig. 5). Region 2, lost in 35

strains, includes a putative chitinase and cellulose synthase *acsAB*. In addition, this lineage has lost genes encoding type VI secretion systems, which are important for virulence as well as between-bacteria competition (50). We also note an additional type VI secretion locus located immediately downstream of the island is conserved. This region is recognised as an intact prophage by the phage prediction program phaster (51). Twenty of these strains have lost a third region that encoded genes involved in biofilm formation (*acsC*, *acsD*; the second copy of *bcsA* and *bcsB* required for cellulose biosynthesis; Fig. S6).

Only one of the Nepalese plasmids (plasmid A) is partially conserved in the isolates we examined (Fig. S7), in accordance with the absence of the genes for NDM-1 and yersiniabactin (Fig. S6). 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 strains does not share the genes from plasmid A in Nepal, which can also be seen by their different virulence factor profile (Fig. S7).

Discussion

The *Klebsiella* isolates represent routinely collected samples over a protracted period. We observed sporadic single-isolate lineages; smaller, related clusters of 5-10 isolates per lineage; in addition to two larger clusters of strains. One of the latter two represents members of *K. quasipneumoniae* subsp. *similipneumoniae*, a subspecies that was previously thought to be less virulent (Fig. 1; 15). The major *K. quasipneumoniae* sequence type (334) was so far observed in other studies as occasional isolate (e.g. 1/250 in 52; 1/167 in 53; 1/328 in 54; 2/198 in 55, whereas we report 52 (44 if excluding ST variants) isolates all originating from bloodstream infections, representing 29.2% (24.7%) of samples investigated. Given its similarity to *K. pneumoniae* in standard diagnostics, *K. quasipneumoniae* numbers could be potentially underestimated (56). *K. variicola*, which is equally difficult to distinguish from *K. pneumoniae*, occurs more frequently in clinical samples (e.g. 28/134 samples in 57; 9/198 in 55; 14/250 in 52; 20/328 in 54) and was suggested to be highly virulent (58), but we did not observe any strain belonging to this species.

With the continuous decline of effective antimicrobials and the lag in new ones being released, antibody-based treatments are becoming increasingly important and are being strongly supported by governments and global health programs (59, 60). There are challenges associated with immunoprophylaxis against *Klebsiella* infections as there is no single prevalent capsule- or O-antigen type; on the contrary, our data reveals a very high diversity in these potential vaccine targets especially with respect to O-Ag, that is generally far less variable than the capsule. A recent large-scale

385 analysis of several *Klebsiella* datasets spanning a global collection identified
386 serotypes O1, O2 and O3 in 80% of hospital infections (20); contrary to this
387 previous finding, we observed a high number of non-O1/O2 types in our
388 collection of *Klebsiella*; these are mainly, but not exclusively, contributed by
389 the large number of *K. quasipneumoniae* (Fig. 1D, Table S1). This diversity,
390 and the lack of apparent correlation to disease severity (20), indicate that
391 switches to non-vaccine targeted serotypes might be easily done once
392 immune pressure for certain types is applied and has been investigated in
393 detail in other organisms (61, 62); a better understanding of the dynamics and
394 diversity is crucial for the development of O-antigen based vaccines, which is
395 ongoing.
396
397 Screening for the resistance and virulence genes revealed no single
398 factors that characterise the most successful lineages or clinical outcome,
399 which is different from other pathogens (e.g. 16, 18). Putative virulence
400 factors in *Klebsiella* spp. are not well-defined beyond the polysaccharide
401 capsule and typically include genes that are thought to mediate attachment,
402 such as Mrk, or survival in the human host, e.g. siderophores (63). Despite
403 spread and chromosomal integration of yersiniabactin however, aerolysin
404 seem to be the main siderophore for *Klebsiella* with highest affinity for iron in
405 *in vitro* experiments despite similar gene numbers expressed and its presence
406 has a significant impact on virulence (64). The considerable variation in
407 yersiniabactin systems suggests that there may be other selective pressures
408 on iron acquisition systems such as antibodies or bacteriophage. The unique
409 clinical manifestation of CG23 however might strongly depend on the unique

feature of colibactin (i.e. enhanced gut colonisation and dissemination to the liver and brain; 65-67). Even though there are discussions whether hypervirulent and hypermucoid should be considered equally and the key virulence factor(s) is still not clear (25, 68, 69), the hallmarks are the virulence plasmid encoding the capsule regulator *rmpA/rmpA2*, aerobactin and metal resistances; and chromosomal integration of yersiniabactin and colibactin, microcin E942 and the sialic acid-containing serum-resistant capsule K1 (70). We report a strain set that is a hybrid type; either encoding the virulence plasmid but with *rmpA2* inactivated, or the yersiniabactin, but no strains with both features even though there are several variants circulating through the population. All strains with the virulence plasmid encode *rmpA2* as a pseudogene, and no strain in the collection encodes colibactin or the functional *rmpA*.

A similar observation was made on a set of recent strains, indicating that the accumulation of virulence factors is an ongoing process, and more work needs to be done of longitudinal studies to monitor the changes of the virulence potential of sequence types over time (71). Findings of these mixed strains, harbouring some but not all virulence factors known, emphasises the importance to understand if there is a negative fitness effect of the combination of all virulence factors or a mutual exclusivity of the mobile elements involved that are only made possible through the genetic background of CG23, or if virulence mechanisms simply spread somewhat slower, and we are only seeing a delay until they appear co-distributed with antimicrobial resistance plasmids.

435

436 The collection is derived from a time (2010 to 2012) where
437 carbapenem resistance was still at comparatively low levels and only starting
438 to spread, whilst it is at a very high prevalence today in *Klebsiella* in LMIC as
439 well as high-income countries (1, 29, 31). The origin of NDM-1 is likely a gene
440 fusion promoted by *Acinetobacter* (13), which was first reported in an IncA/C2
441 plasmid with the original cassette. Our dataset contains several lineages with
442 NDM-1, however these did not become the major problems of the region in
443 the subsequent years. We find three different plasmid constructs at least, with
444 the classical IncA/C2 (similar to NZ_CP006661.1 = pNDM-US; HE016_3
445 found in *Klebsiella* 72), the IncHI2A/IncHI2 (NC_009838.1 = pAPEC-O1-R;
446 H066_2 from *E. coli* 73), and a non-circularised (HE125) plasmid fragment,
447 that showed highest similarity to a plasmid from *Klebsiella* and *E. coli*
448 (KJ440075.1 = pLK78 74). Whilst the IncA/C2 and the MAR plasmid are
449 nowadays spread widely through the *Klebsiella* population, the IncR plasmid
450 is more diagnostic of MDR *E. coli*, and might have used *Klebsiella* rather as
451 initial vehicle than a successful means of spread. The samples are from a
452 time where ESBL strains were already widespread, and carbapenems used
453 as optional treatment; an NDM-1 carrying plasmid would therefore provide a
454 clear advantage at face value. However, the patterns are more complex and
455 much remains to be understood about the interaction between plasmids and
456 bacteria, but it seems clear that acquisition of an NDM-1 plasmid alone is
457 neither a sign that the plasmid will now spread stably with the population, nor
458 that this lineage has an undisputable advantage over all carbapenemase-
459 negative lineages.

460

461 The genomics generally predicted the phenotype with respect to drug
462 resistance. This is important as technology is being developed to identify
463 resistance genes at the point-of-care. Subtle differences in the sensitivity that
464 are overlooked by evaluating strains only below or above a specified drug
465 concentration, and might indicate an additional mechanism to deactivate beta-
466 lactamase inhibitors, especially since we could not detect any genomic
467 differences between the beta-lactamases of closely related strains. Such an
468 analysis does not preclude expression differences due to promoter
469 differences that might impact phenotypic resistance. Monitoring this
470 resistance in detail is important as piperacillin/tazobactam treatment is a
471 common clinical response to ESBL infections (75). A recent report on
472 stepwise resistance against an inhibitor also highlighted changes in the major
473 porins (Omp35 and Omp36) as well as the LPS O-antigen locus (*rfb*); this
474 could explain some resistances (e.g. the porin disruption in ST15) (76).
475 AmpH is a penicillin-binding protein, but exhibits only extremely low beta-
476 lactamase activity against nitrocefin (77, 78). None of the OXA alleles
477 detected in our dataset conferred extended-spectrum beta-lactamase activity,
478 but OXA enzymes are usually poorly inhibited by clavulanic acid, which might
479 provide an important advantage (79).

480

481 Understanding the rapid spread of ESBL- and carbapenemase –
482 carrying plasmids in combination with the spread of high-risk lineages arising
483 from the diverse background population of *Klebsiella* is of crucial importance
484 to the early identification of highly resistant and virulent clones, and

485 preventing their spread. There is considerable subtlety in the evolution of drug
486 resistance where having a variety of activity spectra across resistance
487 enzymes is likely to be of importance, as enzymes with lower activity spectra
488 are often more resistant to inhibitors than highly resistant ESBL enzymes (79).
489 Detailed analyses are crucial to our understanding of the different
490 mechanisms of resistance against alternative treatments such as beta-
491 lactamase inhibitors. Inhibitors could become a highly effective treatment if
492 resistance can be avoided, or more likely recognised earlier.

Methods

PacBio DNA preparation and sequencing

DNA was further purified by Phenol:Chloroform:Isoamyl Alcohol (25:24:1) and Chloroform:IAA (24:1) extractions using Phase-lock tubes (Qiagen) and re-dissolved in 10mM Tris pH7.4 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 v1.1 (80). The assemblies were then circularised using circulator (81) with canu as assembler, and polished with unicycler-polish (82) and the Illumina reads of the respective sample. HE016 gave better results in assembling with the unicycler-hybrid assembler combining the PacBio data with the Illumina reads (15), and was subsequently circularised and polished as above. All assemblies were then annotated with prokka (83).

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 Nepal (49) and HE196 (this study). Sequence reads were mapped against the reference genome as indicated using SMALT (84; 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 v1.11 (83) were used as input for the pan-genome pipeline Roary (v3.7.0; 85) 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 (Figure 1, 2). The core gene alignment was generated with mafft (86); SNPs were first extracted using snp-sites v2.3.2 (87), and then a maximum likelihood tree with RAxML (88) was calculated. For the ST15 comparison, the pan-genome was calculated as described above, and gene presence/absence analysed further using R and the resulting presence/absence matrix and phandango (89).

Phylogenetic analyses

Trees were calculated using RAxML (v8.2.8; 88) with the time-reversible GTR model and 100 bootstrap repeats. Tree demonstrations were prepared in itol (90) as well as in house python and R scripts. For whole-genome mapping trees, recombinant regions were removed using Gubbins (v1.4.9; 91) 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-ANNOT

(92) 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, 94). Sequence gene profiles and types were determined using MLST check (95) 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 kleborate (23; https://github.com/katholt/Kleborate). Capsule and LPS O-antigen was annotated using kleborate and a custom database of LPS O-antigens (20). The assignment to O3 subgroups was updated based on recent nomenclature (O3l/s; now a/b/c) using the online tool captive (http://kaptive.holtlab.net/, 22).

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

568 reader–incubator module (incubation temperature 37°C), and fluorescence
569 measurements were performed every 15 min for up to 18 h. For detailed
570 analysis of several NDM-1 containing strains, the MICs for meropenem were
571 also assessed manually following the protocol of Wiegand 2008 (96; Table 1).
572
573 **Data availability:**
574 Details on strains and accession numbers can be found in Table S1. All
575 relevant data is in the manuscript and supporting material, alignment and tree
576 files are available for download at figshare
577 https://figshare.com/s/cdaddb659a6e178102df.
578
579
580
581
582
583
584

585 **Table 1: Detection of carbapenemase using minimum inhibitory**

586 **concentration and gene prediction (ariba read coverage).**

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

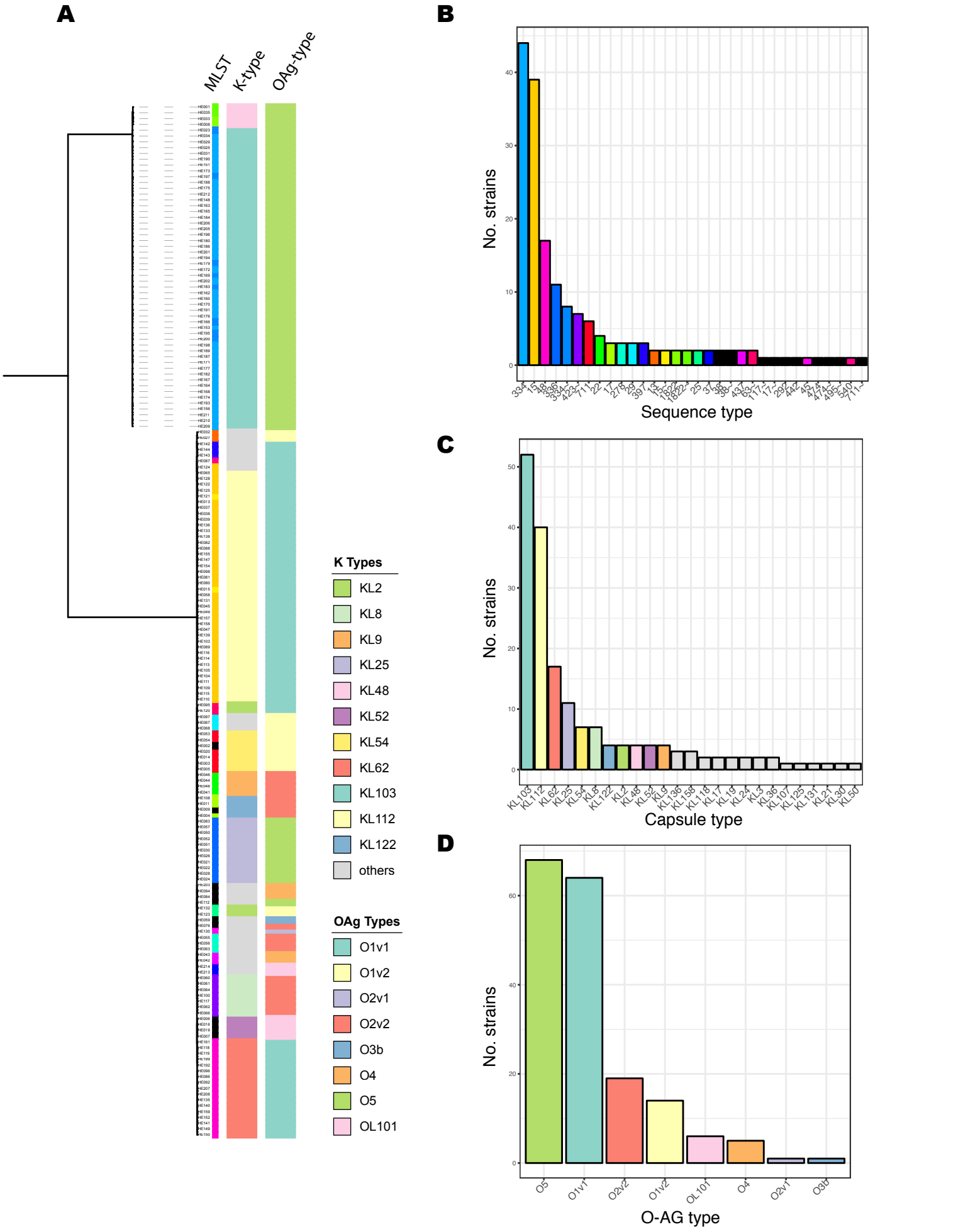
587

588

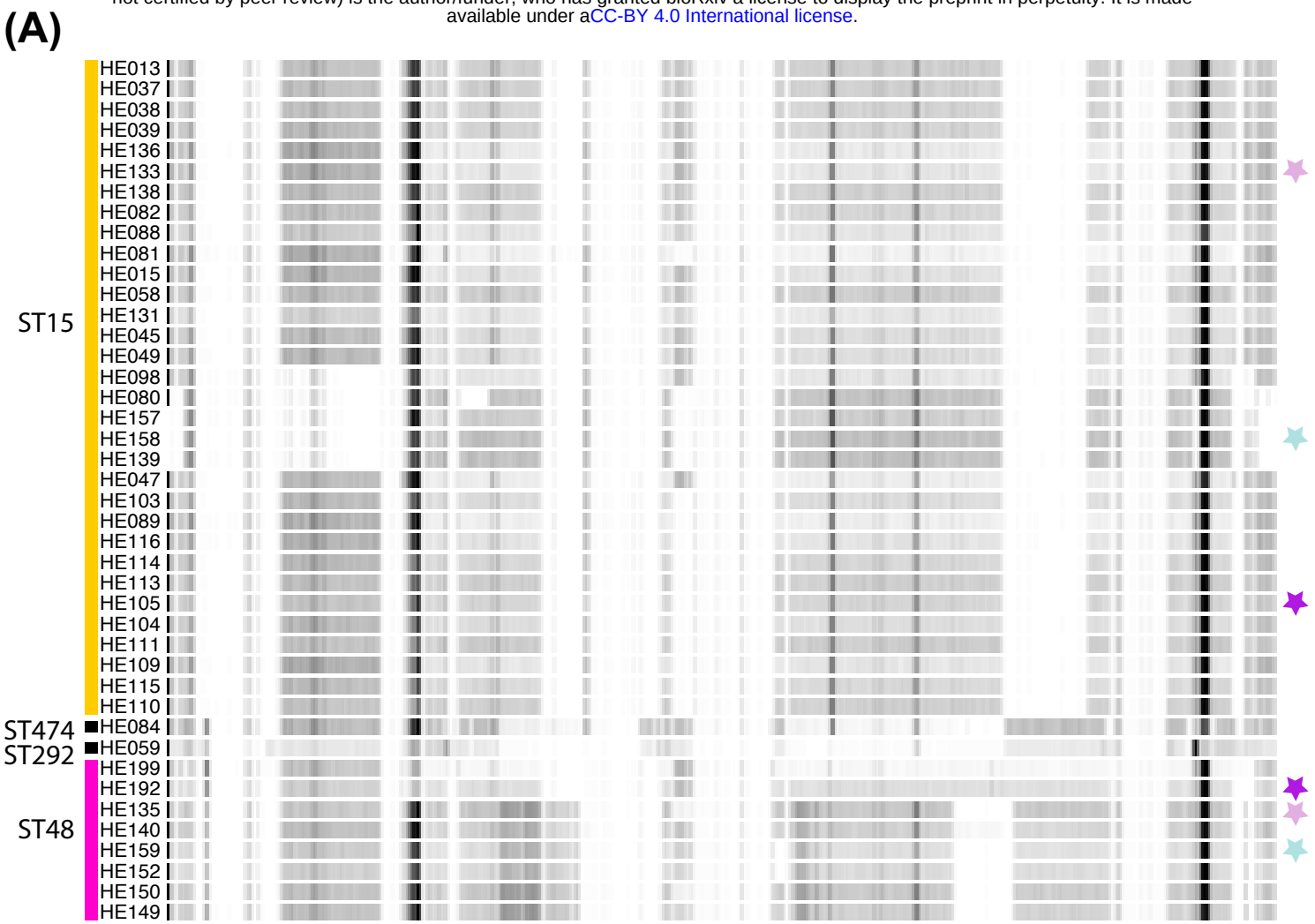
589 **Table 2: Testing outcome versus presence of yersiniabactin or**
590 **aerobactin.**

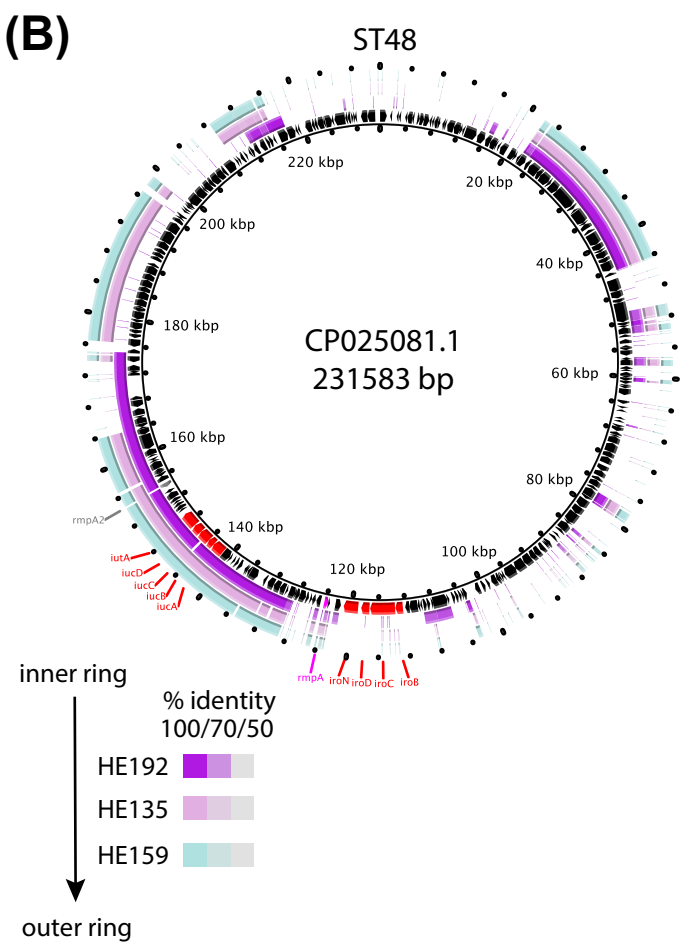
	outcome		
Siderophore	Death	Discharged	Total
Yersiniabactin +	16	26	42
Yersiniabactin -	34	75	109
Total	50	101	151
p-value[1]	p_{chi}=0.5388; p_{fisher}=0.4444		
Aerobactin +	13	22	35
Aerobactin -	37	79	116
Total	50	101	151
p-value[1]	p_{chi}=0.709; p_{fisher}=0.6823		

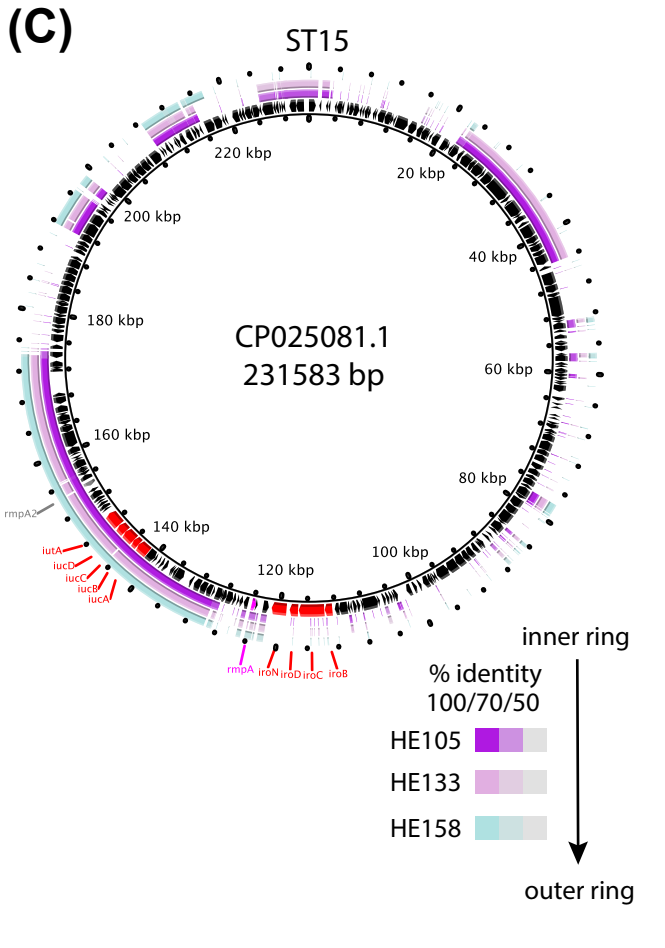
591
592 [1]: P-value determined via a chi-square test of independence using the R
593 function chisq.test(input matrix) annotated as p_{chi}; and a fisher's exact test
594 using the R function fisher.test(input matrix) annotated as p_{fisher}.
595



596 **Fig. 1: LPS O-antigen, capsule and sequence types in this study. (A)** The
597 guide tree is based on roary as in Fig. 1; and shows the diversity of capsule
598 and O-antigen type, which correlate with sequence types (one sequence type
599 usually shares the capsule- and O-antigen combination). **(B)** Two main (ST15,
600 ST334) sequence types dominate the collection, followed by several medium-
601 abundance (e.g. ST45, ST336) types and a high diversity of low-abundance
602 sequence types. **(C)** The capsule types almost share almost exactly the same
603 level as diversity as the sequence types, with a different capsule type
604 associated with sequence types; whereas the O-antigen types **(D)** are as
605 expected less diverse, but dominated through the commonly as less diverse
606 perceived type O5.
607







608 **Fig. 2: Comparison of strains with *rmpA2* pseudogenes to Klebsiella**
609 **virulence plasmids. (A)** shows a heatmap from mapping the reads against
610 the recently published reference plasmid for hypervirulent strains, pSGH10
611 (28). Representative strains were chosen as indicated by star icons in **(A)** for
612 ST48 **(B)** and ST15 **(C)**, the relevant Illumina contigs extracted using blastn
613 and abacas, and mapped against the reference plasmid using BRIG (97) to
614 further illustrate the partial conversation of the virulence plasmid also in the
615 assembled data.
616

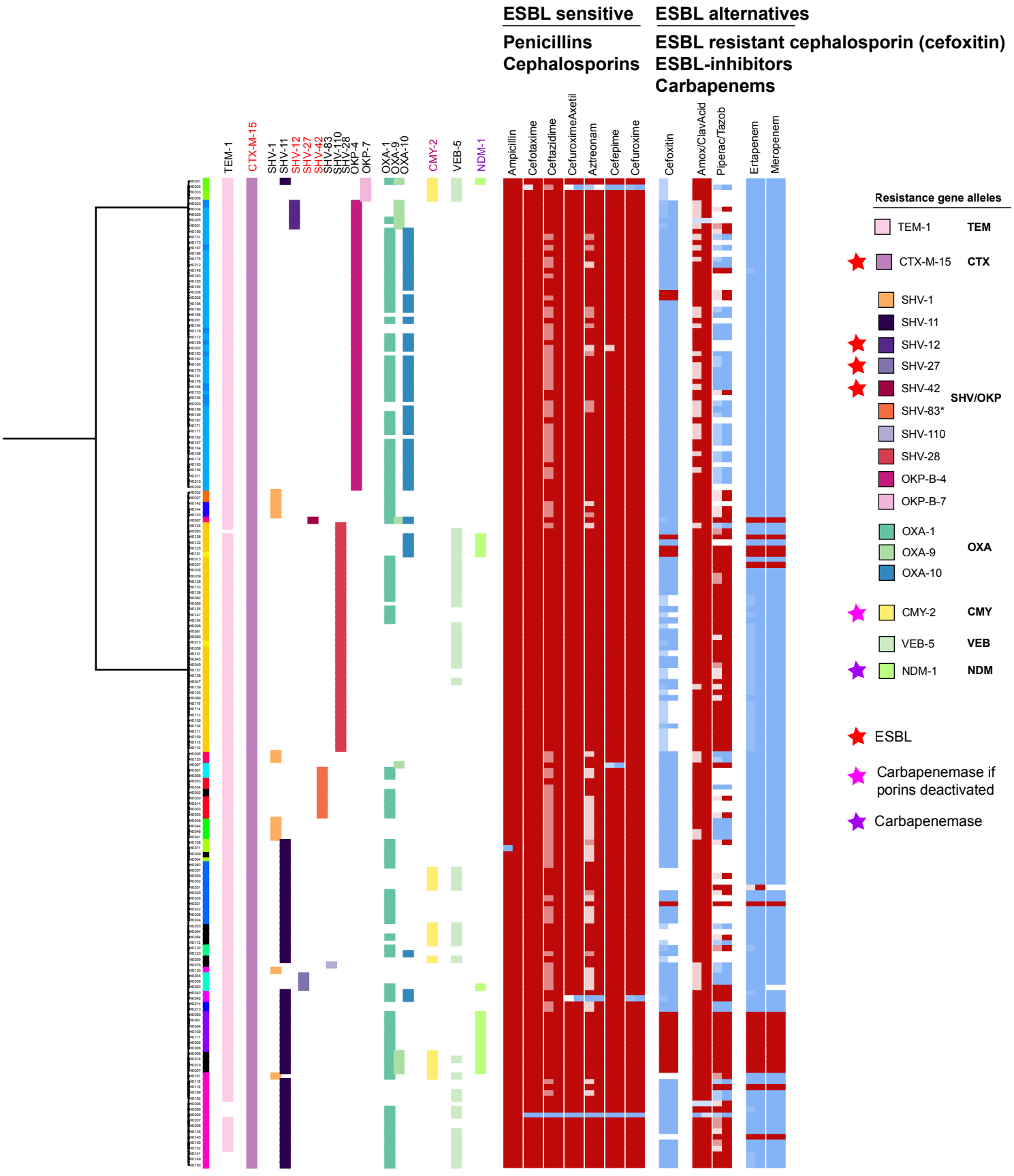
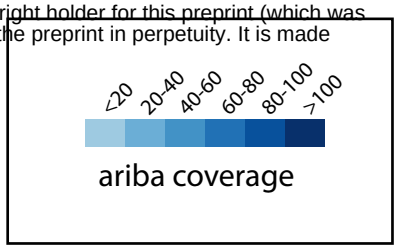
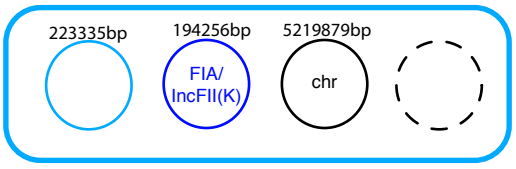
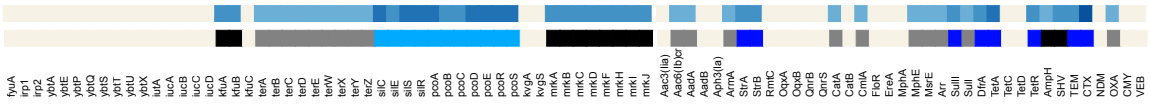


Fig. 3: Details of predicted beta-lactamase enzymes and resistance profiles to treatment alternatives for ESBL organisms. (**A)** The guidance tree is as in Fig. 1. All strains with TEM or CTX genes encode for the broad-spectrum beta-lactamase TEM-1 and the ESBL enzyme CTX-M-15, respectively, and only the VEB-5 and NDM-1 allele, and two alleles for the AmpC CMY (6 and 18) could be identified. On the contrary, we see (amino acid) variation in the different alleles for SHV/OKP and OXA; colours are according to **(B)** for SHV/OKP and **(C)** for OXA. *Bla*-SHV can have different activity spectrums; the red stars indicate extended-spectrum beta-lactamase activity. Pink indicates weak carbapenemase activity which needs additional mutations (e.g. porin deactivation) to confer resistance, dark violet denotes carbapenemase activity. The allele assignment was controlled at the beta-lactamase online database (http://www.laced.uni-stuttgart.de); activity assignments are according to (98).

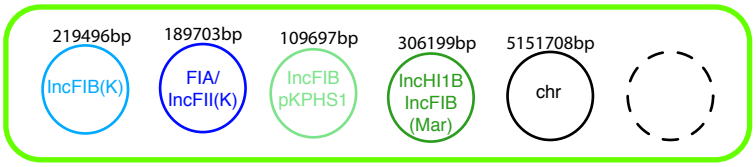


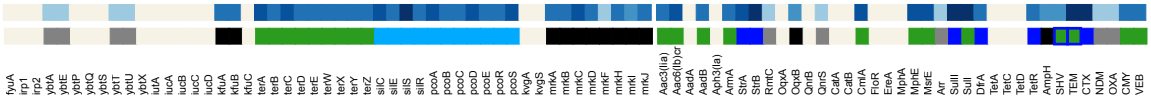
HE196 (ST334)



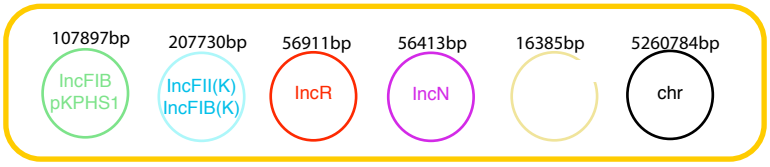


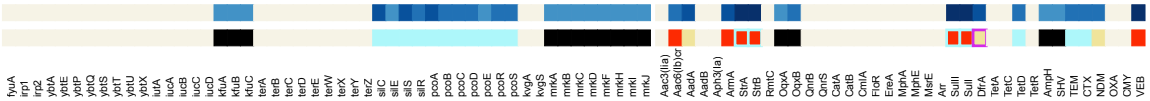
HE001 (ST1822)



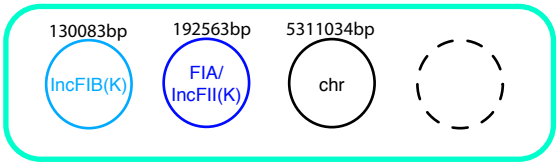


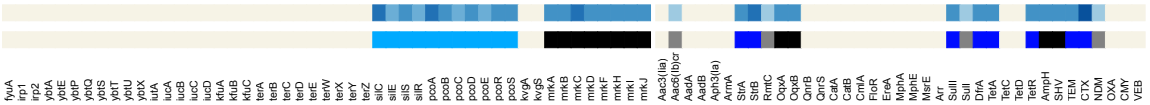
HE125 (ST15)



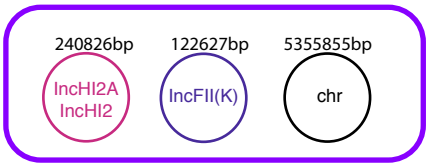


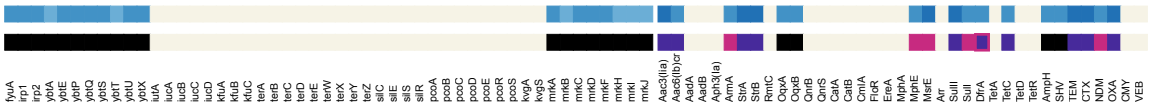
HE063 (ST278)



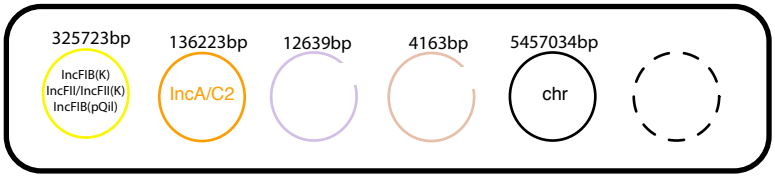


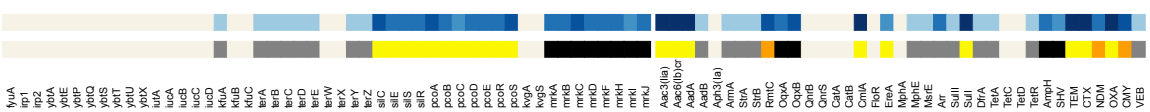
HE066 (ST423)



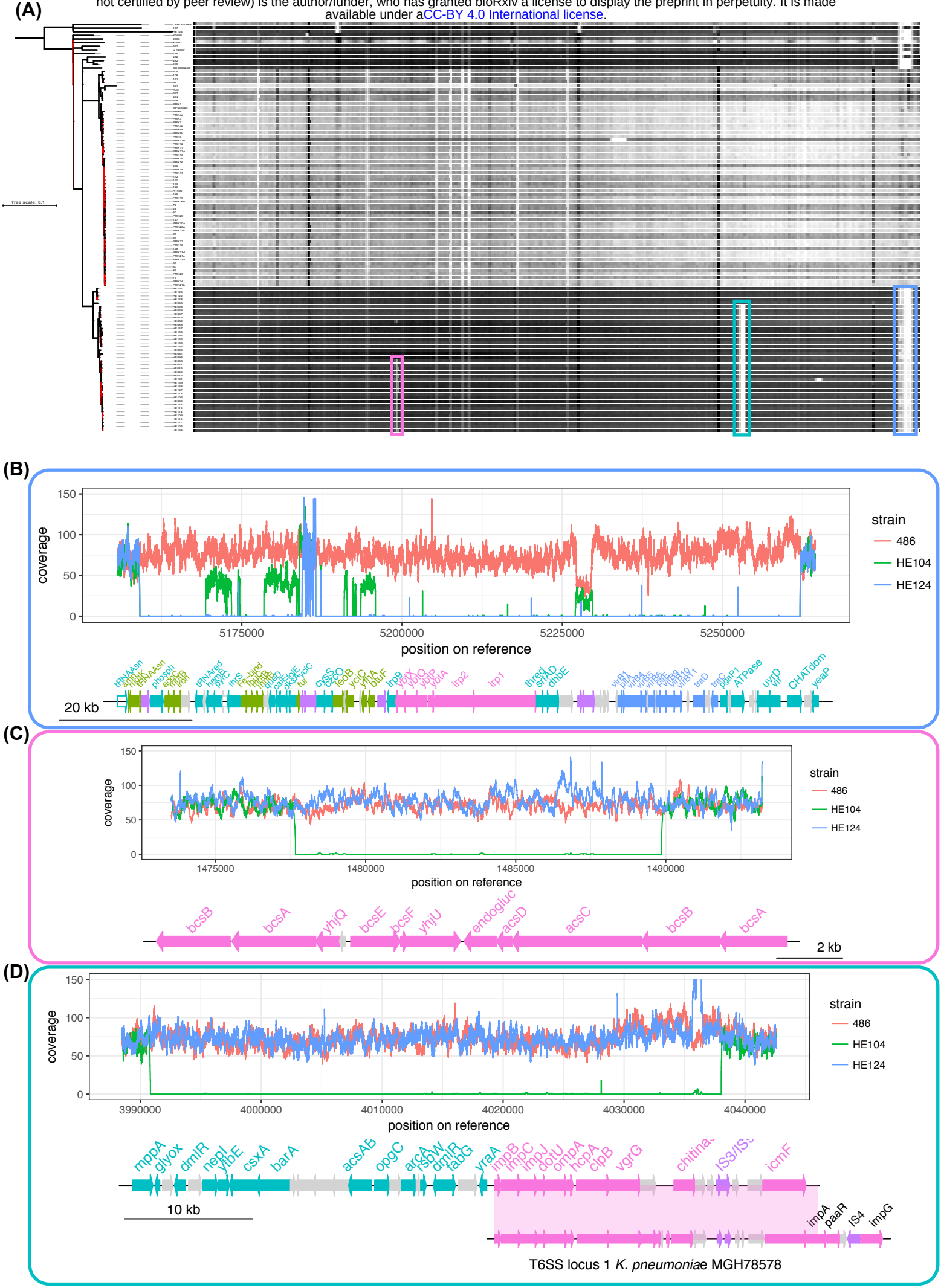


HE016 (ST38)





632 **Fig. 4: The plasmid diversity for representative isolates.** The contigs
633 obtained from PacBio for selected strains are shown; plasmid replicons as
634 predicted by the PlasmidFinder web server
635 (https://cge.cbs.dtu.dk/services/PlasmidFinder/, 94). Below each strain are the
636 resistance genes as predicted by ariba on a colour gradient displaying the
637 read coverage of the hit; and the row below these indicate on which plasmid
638 the respective genes were found (colours according to the respective cell
639 overview). Black indicates chromosomal, dark grey indicates gene only found
640 in Illumina reads but not in the PacBio data. The presence of genes was
641 investigated by blastn, only complete perfect (100% identity) hits were
642 considered.
643



644 **Fig. 5: In-depth comparison of expanded lineages. (A)** The strains
645 belonging to ST15 of this study was compared to the ST15 outbreak lineage
646 from Nepal (48, 49) by mapping against the completed genome of an isolate
647 from the Nepal outbreak (49), which also carried four plasmids. The heatmap
648 shows the mapping coverage across the chromosome before removal of
649 recombination from white (low coverage) to black (high coverage),
650 emphasizing the large deletions and acquisitions in the Pakistan and Nepal
651 line, respectively. Even coverage of the genome could be observed, with
652 three genetic islands missing in different subsets of our strains, indicating the
653 dynamic within this closely related lineage. The tree was constructed after
654 removing recombination by gubbins, bootstrap values are depicted as colour
655 gradient; from red (0) to black (100). **(B)** A detailed view of two regions in the
656 genome absent in groups of strains within the isolates from this study (plotted
657 areas: 1473529-1493215, region 1; 3988478-4042608, region 2; coordinates
658 from CP008929.1), but present in the lineage from Nepal, as well as other
659 isolates from the same ST but not part of the clonal lineages as outgroups. In
660 addition, the Nepal lineage acquired the chromosomal locus including
661 yersiniabactin (23; plotted region 5155581-5264456); these three regions
662 were also removed by gubbins before the tree construction.
663
664
665 **Supplement: Figures S1-S7 (with Figure legends) and Tables S1-S3:**
666 **Table S1: Details on the strains from Pakistan.**
667 **Table S2: Vitek measurements.**
668 **Table S3: Additional strains included in the detailed analysis of ST15.**

669

670 **Acknowledgments**

671 The authors thank Karen Oliver for help with the PacBio sequencing, and
672 Andrew Page for helpful discussions on PacBio assembly strategies, as well
673 as the Pathogens Informatics group. We thank Dr. Badr Alzahrani (Jouf
674 University) for comments on the manuscript. This work was supported by the
675 NHMRC (Program Grant 1092262), the Wellcome Trust (206194), and the
676 Higher Education Commission of Pakistan and The Children's Hospital & The
677 ICH, Lahore, Pakistan. H.E. was supported by a scholarship from Higher
678 Education Commission (HEC) Pakistan under the International Research
679 Support Initiative Program (IRSIP).

680

References

1. WHO. GLOBAL PRIORITY LIST OF ANTIBIOTIC-RESISTANT BACTERIA TO GUIDE RESEARCH, DISCOVERY, AND DEVELOPMENT OF NEW ANTIBIOTICS 2017 [Available from: http://www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en/.
2. WHO. Antimicrobial resistance: global report on surveillance 2014. WHO; 2014.
3. Boulos A, Rand K, Johnson JA, Gautier J, Koster M. Neonatal Sepsis in Haiti. J Trop Pediatr. 2017;63(1):70-3.
4. Shehab El-Din EM, El-Sokkary MM, Bassiouny MR, Hassan R. Epidemiology of Neonatal Sepsis and Implicated Pathogens: A Study from Egypt. Biomed Res Int. 2015;2015:509484.
5. Kabwe M, Tembo J, Chilukutu L, Chilufya M, Ngulube F, Lukwesa C, et al. Etiology, Antibiotic Resistance and Risk Factors for Neonatal Sepsis in a Large Referral Center in Zambia. Pediatr Infect Dis J. 2016;35(7):e191-8.
6. Investigators of the Delhi Neonatal Infection Study c. Characterisation and antimicrobial resistance of sepsis pathogens in neonates born in tertiary care centres in Delhi, India: a cohort study. Lancet Glob Health. 2016;4(10):e752-60.
7. Ghotaslou R, Ghorashi Z, Nahaei MR. Klebsiella pneumoniae in neonatal sepsis: a 3-year-study in the pediatric hospital of Tabriz, Iran. Jpn J Infect Dis. 2007;60(2-3):126-8.
8. Wyres KL, Holt KE. Klebsiella pneumoniae Population Genomics and Antimicrobial-Resistant Clones. Trends Microbiol. 2016;24(12):944-56.
9. Lee CR, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global Dissemination of Carbapenemase-Producing Klebsiella pneumoniae: Epidemiology, Genetic Context, Treatment Options, and Detection Methods. Front Microbiol. 2016;7:895.
10. Pitout JD, Laupland KB. Extended-spectrum beta-lactamase-producing Enterobacteriaceae: an emerging public-health concern. Lancet Infect Dis. 2008;8(3):159-66.
11. Doi Y, Bonomo RA, Hooper DC, Kaye KS, Johnson JR, Clancy CJ, et al. Gram-Negative Bacterial Infections: Research Priorities, Accomplishments, and Future Directions of the Antibacterial Resistance Leadership Group. Clin Infect Dis. 2017;64(suppl_1):S30-S5.
12. Yong D, Toleman MA, Giske CG, Cho HS, Sundman K, Lee K, 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. 2009;53(12):5046-54.
13. Toleman MA, Spencer J, Jones L, Walsh TR. blaNDM-1 is a chimera likely constructed in Acinetobacter baumannii. Antimicrob Agents Chemother. 2012;56(5):2773-6.
14. Kumarasamy KK, Toleman MA, Walsh TR, Bagaria J, Butt F, Balakrishnan R, et al. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. Lancet Infect Dis. 2010;10(9):597-602.
15. Ejaz H, Wang N, Wilksch JJ, Page AJ, Cao H, Gujaran S, et al. Phylogenetic Analysis of Klebsiella pneumoniae from Hospitalized Children, Pakistan. Emerg Infect Dis. 2017;23(11):1872-5.
16. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events. Nat Genet. 2015;47(6):632-9.

17. Jamrozy DM, Harris SR, Mohamed N, Peacock SJ, Tan CY, Parkhill J, et al. Pan-genomic perspective on the evolution of the Staphylococcus aureus USA300 epidemic. Microb Genom. 2016;2(5):e000058.
18. McNally A, Oren Y, Kelly D, Pascoe B, Dunn S, Sreecharan T, et al. Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. PLoS Genet. 2016;12(9):e1006280.
19. Wyres KL, Wick RR, Gorrie C, Jenney A, Follador R, Thomson NR, et al. Identification of Klebsiella capsule synthesis loci from whole genome data. bioRxiv preprint. 2016;doi: http://dx.doi.org/10.1101/071415.
20. Follador R, Heinz E, Wyres KL, Ellington MJ, Kowarik M, Holt KE, et al. The diversity of Klebsiella pneumoniae surface polysaccharides. Microbial Genomics. 2016;2(8):doi:10.1099/mgen.0.000073.
21. Jenney AW, Clements A, Farn JL, Wijburg OL, McGlinchey A, Spelman DW, et al. Seroepidemiology of Klebsiella pneumoniae in an Australian Tertiary Hospital and its implications for vaccine development. J Clin Microbiol. 2006;44(1):102-7.
22. Wick RR, Heinz E, Holt KE, Wyres KL. Kaptive Web: user-friendly capsule and lipopolysaccharide serotype prediction for Klebsiella genomes. 2018.
23. Lam MMC, Wick RR, Wyres KL, Gorrie C, Judd LM, Brisse S, et al. Frequent emergence of pathogenic lineages of Klebsiella pneumoniae via mobilisation of yersiniabactin and colibactin. bioRxiv preprint. 2017.
24. Yu WL, Lee MF, Tang HJ, Chang MC, Chuang YC. Low prevalence of rmpA and high tendency of rmpA mutation correspond to low virulence of extended spectrum beta-lactamase-producing Klebsiella pneumoniae isolates. Virulence. 2015;6(2):162-72.
25. Shon AS, Bajwa RP, Russo TA. Hypervirulent (hypermucoviscous) Klebsiella pneumoniae: a new and dangerous breed. Virulence. 2013;4(2):107-18.
26. Wu KM, Li LH, Yan JJ, Tsao N, Liao TL, Tsai HC, et al. Genome sequencing and comparative analysis of Klebsiella pneumoniae NTUH-K2044, a strain causing liver abscess and meningitis. J Bacteriol. 2009;191(14):4492-501.
27. Chen YT, Chang HY, Lai YC, Pan CC, Tsai SF, Peng HL. Sequencing and analysis of the large virulence plasmid pLVPK of Klebsiella pneumoniae CG43. Gene. 2004;337:189-98.
28. Lam MMC, Wyres KL, Duchene S, Wick RR, Judd LM, Gan Y-H, et al. Population genomics of hypervirulent Klebsiella pneumoniae clonal group 23 reveals early emergence and rapid global dissemination.
29. Grundmann H, Glasner C, Albiger B, Aanensen DM, Tomlinson CT, Andrasevic AT, 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. 2017;17(2):153-63.
30. Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, et al. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis. 2016;16(2):161-8.
31. Navon-Venezia S, Kondratyeva K, Carattoli A. Klebsiella pneumoniae: a major worldwide source and shuttle for antibiotic resistance. FEMS Microbiol Rev. 2017;41(3):252-75.
32. Hannecart-Pokorni E, Depuydt F, de wit L, van Bossuyt E, Content J, Vanhoof R. Characterization of the 6'-N-aminoglycoside acetyltransferase gene aac(6')-Im [corrected] associated with a sull-type integron. Antimicrob Agents Chemother. 1997;41(2):314-8.
33. Kaushik M, Kumar S, Kapoor RK, Virdi JS, Gulati P. Integrons in Enterobacteriaceae: diversity, distribution and epidemiology. Int J Antimicrob Agents. 2018;51(2):167-76.

789 34. Vetting MW, Hegde SS, Wang M, Jacoby GA, Hooper DC, Blanchard JS.
790 Structure of QnrB1, a plasmid-mediated fluoroquinolone resistance factor. J Biol
791 Chem. 2011;286(28):25265-73.
792 35. Wen Y, Pu X, Zheng W, Hu G. High Prevalence of Plasmid-Mediated
793 Quinolone Resistance and IncQ Plasmids Carrying qnrS2 Gene in Bacteria from
794 Rivers near Hospitals and Aquaculture in China. PLoS One. 2016;11(7):e0159418.
795 36. Haeggman S, Lofdahl S, Paauw A, Verhoef J, Brisse S. Diversity and
796 evolution of the class A chromosomal beta-lactamase gene in Klebsiella
797 pneumoniae. Antimicrob Agents Chemother. 2004;48(7):2400-8.
798 37. Manageiro V, Ferreira E, Cougnoux A, Albuquerque L, Canica M, Bonnet R.
799 Characterization of the inhibitor-resistant SHV beta-lactamase SHV-107 in a clinical
800 Klebsiella pneumoniae strain coproducing GES-7 enzyme. Antimicrob Agents
801 Chemother. 2012;56(2):1042-6.
802 38. Prinarakis EE, Miriagou V, Tzelepi E, Gazouli M, Tzouvelekis LS. Emergence
803 of an inhibitor-resistant beta-lactamase (SHV-10) derived from an SHV-5 variant.
804 Antimicrob Agents Chemother. 1997;41(4):838-40.
805 39. Pitton JS. Mechanisms of bacterial resistance to antibiotics. Ergeb Physiol.
806 1972;65:15-93.
807 40. Dubois V, Poirel L, Demarthe F, Arpin C, Coulange L, Minarini LA, et al.
808 Molecular and biochemical characterization of SHV-56, a novel inhibitor-resistant
809 beta-lactamase from Klebsiella pneumoniae. Antimicrob Agents Chemother.
810 2008;52(10):3792-4.
811 41. Mendonca N, Ferreira E, Louro D, Participants A, Canica M. Molecular
812 epidemiology and antimicrobial susceptibility of extended- and broad-spectrum beta-
813 lactamase-producing Klebsiella pneumoniae isolated in Portugal. Int J Antimicrob
814 Agents. 2009;34(1):29-37.
815 42. Winkler ML, Papp-Wallace KM, Taracila MA, Bonomo RA. Avibactam and
816 inhibitor-resistant SHV beta-lactamases. Antimicrob Agents Chemother.
817 2015;59(7):3700-9.
818 43. Wang X, Liu W, Zou D, Li X, Wei X, Shang W, et al. High rate of New Delhi
819 metallo-beta-lactamase 1-producing bacterial infection in China. Clin Infect Dis.
820 2013;56(1):161-2.
821 44. Huang TW, Chen TL, Chen YT, Lauderdale TL, Liao TL, Lee YT, et al. Copy
822 Number Change of the NDM-1 sequence in a multidrug-resistant Klebsiella
823 pneumoniae clinical isolate. PLoS One. 2013;8(4):e62774.
824 45. Gonzalez LJ, Bahr G, Nakashige TG, Nolan EM, Bonomo RA, Vila AJ.
825 Membrane anchoring stabilizes and favors secretion of New Delhi metallo-beta-
826 lactamase. Nat Chem Biol. 2016;12(7):516-22.
827 46. Wailan AM, Paterson DL, Kennedy K, Ingram PR, Bursle E, Sidjabat HE.
828 Genomic Characteristics of NDM-Producing Enterobacteriaceae Isolates in Australia
829 and Their blaNDM Genetic Contexts. Antimicrob Agents Chemother. 2015;60(1):136-
830 41.
831 47. Wailan AM, Sidjabat HE, Yam WK, Alikhan NF, Petty NK, Sartor AL, et al.
832 Mechanisms Involved in Acquisition of blaNDM Genes by IncA/C2 and IncFIIY
833 Plasmids. Antimicrob Agents Chemother. 2016;60(7):4082-8.
834 48. Chung The H, Karkey A, Pham Thanh D, Boinett CJ, Cain AK, Ellington M, et
835 al. A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of
836 Klebsiella pneumoniae. EMBO Mol Med. 2015;7(3):227-39.
837 49. Stoesser N, Giess A, Batty EM, Sheppard AE, Walker AS, Wilson DJ, et al.
838 Genome sequencing of an extended series of NDM-producing Klebsiella
839 pneumoniae isolates from neonatal infections in a Nepali hospital characterizes the
840 extent of community- versus hospital-associated transmission in an endemic setting.
841 Antimicrob Agents Chemother. 2014;58(12):7347-57.

842 50. Sarris PF, Zoumadakis C, Panopoulos NJ, Scoulica EV. Distribution of the
843 putative type VI secretion system core genes in Klebsiella spp. Infect Genet Evol.
844 2011;11(1):157-66.
845 51. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a
846 better, faster version of the PHAST phage search tool. Nucleic Acids Res.
847 2016;44(W1):W16-21.
848 52. Moradigaravand D, Martin V, Peacock SJ, Parkhill J. Evolution and
849 Epidemiology of Multidrug-Resistant Klebsiella pneumoniae in the United Kingdom
850 and Ireland. MBio. 2017;8(1).
851 53. Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D, et al.
852 Genomic Analysis of the Emergence and Rapid Global Dissemination of the Clonal
853 Group 258 Klebsiella pneumoniae Pandemic. PLoS One. 2015;10(7):e0133727.
854 54. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al.
855 Genomic analysis of diversity, population structure, virulence, and antimicrobial
856 resistance in Klebsiella pneumoniae, an urgent threat to public health. Proc Natl
857 Acad Sci U S A. 2015;112(27):E3574-81.
858 55. Henson SP, Boinett CJ, Ellington MJ, Kagia N, Mwarumba S, Nyongesa S, et
859 al. Molecular epidemiology of Klebsiella pneumoniae invasive infections over a
860 decade at Kilifi County Hospital in Kenya. Int J Med Microbiol. 2017;307(7):422-9.
861 56. Long SW, Linson SE, Ojeda Saavedra M, Cantu C, Davis JJ, Brettin T, et al.
862 Whole-Genome Sequencing of Human Clinical Klebsiella pneumoniae Isolates
863 Reveals Misidentification and Misunderstandings of Klebsiella pneumoniae,
864 Klebsiella variicola, and Klebsiella quasipneumoniae. mSphere. 2017;2(4).
865 57. Gorrie CL, Mirceta M, Wick RR, Edwards DJ, Thomson NR, Strugnell RA, et
866 al. Gastrointestinal Carriage Is a Major Reservoir of Klebsiella pneumoniae Infection
867 in Intensive Care Patients. Clin Infect Dis. 2017;65(2):208-15.
868 58. Maatallah M, Vading M, Kabir MH, Bakhrouf A, Kalin M, Naucler P, et al.
869 Klebsiella variicola is a frequent cause of bloodstream infection in the stockholm
870 area, and associated with higher mortality compared to K. pneumoniae. PLoS One.
871 2014;9(11):e113539.
872 59. Lob-Levyt J. Vaccine coverage and the GAVI Alliance Immunization Services
873 Support initiative. Lancet. 2009;373(9659):209.
874 60. Ahmed R, Anderson RM, Bakken D, Bethony JM, Bloom BR, Blomke CJ, et
875 al. The Vaccine Book. Bloom BR, Lambert P-H, editors: Elsevier Inc.; 2016.
876 61. Corander J, Fraser C, Gutmann MU, Arnold B, Hanage WP, Bentley SD, et
877 al. Frequency-dependent selection in vaccine-associated pneumococcal population
878 dynamics. Nat Ecol Evol. 2017;1(12):1950-60.
879 62. Mostowy RJ, Croucher NJ, De Maio N, Chewapreecha C, Salter SJ, Turner
880 P, et al. Pneumococcal Capsule Synthesis Locus cps as Evolutionary Hotspot with
881 Potential to Generate Novel Serotypes by Recombination. Mol Biol Evol.
882 2017;34(10):2537-54.
883 63. Paczosa MK, Mecsas J. Klebsiella pneumoniae: Going on the Offense with a
884 Strong Defense. Microbiol Mol Biol Rev. 2016;80(3):629-61.
885 64. Russo TA, Olson R, Macdonald U, Metzger D, Maltese LM, Drake EJ, et al.
886 Aerobactin mediates virulence and accounts for increased siderophore production
887 under iron-limiting conditions by hypervirulent (hypermucoviscous) Klebsiella
888 pneumoniae. Infect Immun. 2014;82(6):2356-67.
889 65. Nougayrede JP, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk
890 G, et al. Escherichia coli induces DNA double-strand breaks in eukaryotic cells.
891 Science. 2006;313(5788):848-51.
892 66. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrede JP.
893 Escherichia coli induces DNA damage in vivo and triggers genomic instability in
894 mammalian cells. Proc Natl Acad Sci U S A. 2010;107(25):11537-42.

67. Guerra L, Guidi R, Frisan T. Do bacterial genotoxins contribute to chronic inflammation, genomic instability and tumor progression? FEBS J. 2011;278(23):4577-88.
68. Russo TA, Shon AS, Beanan JM, Olson R, MacDonald U, Pomakov AO, et al. Hypervirulent K. pneumoniae secretes more and more active iron-acquisition molecules than "classical" K. pneumoniae thereby enhancing its virulence. PLoS One. 2011;6(10):e26734.
69. Catalan-Najera JC, Garza-Ramos U, Barrios-Camacho H. Hypervirulence and hypermucoviscosity: Two different but complementary Klebsiella spp. phenotypes? Virulence. 2017;8(7):1111-23.
70. Struve C, Roe CC, Stegger M, Stahlhut SG, Hansen DS, Engelthaler DM, et al. Mapping the Evolution of Hypervirulent Klebsiella pneumoniae. MBio. 2015;6(4):e00630.
71. Turton JF, Payne Z, Coward A, Hopkins KL, Turton JA, Doumith M, 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. 2018;67(1):118-28.
72. Hudson CM, Bent ZW, Meagher RJ, Williams KP. Resistance determinants and mobile genetic elements of an NDM-1-encoding Klebsiella pneumoniae strain. PLoS One. 2014;9(6):e99209.
73. Johnson TJ, Wannemeuhler YM, Scaccianoce JA, Johnson SJ, Nolan LK. Complete DNA sequence, comparative genomics, and prevalence of an IncHI2 plasmid occurring among extraintestinal pathogenic Escherichia coli isolates. Antimicrob Agents Chemother. 2006;50(11):3929-33.
74. Chen CJ, Wu TL, Lu PL, Chen YT, Fung CP, Chuang YC, et al. Closely related NDM-1-encoding plasmids from Escherichia coli and Klebsiella pneumoniae in Taiwan. PLoS One. 2014;9(8):e104899.
75. Schuetz A, Salcedo SR, Tamma PD. Point-Counterpoint: Piperacillin-tazobactam should be used to treat infections with ESBL-positive organisms. J Clin Microbiol. 2017.
76. Giddins MJ, Macesic N, Annavajhala MK, Stump S, Khan S, McConville TH, et al. Successive emergence of ceftazidime-avibactam resistance through distinct genomic adaptations in blaKPC-2-harboring Klebsiella pneumoniae ST307. Antimicrob Agents Chemother. 2017.
77. Gonzalez-Leiza SM, de Pedro MA, Ayala JA. AmpH, a bifunctional DD-endopeptidase and DD-carboxypeptidase of Escherichia coli. J Bacteriol. 2011;193(24):6887-94.
78. Henderson TA, Young KD, Denome SA, Elf PK. AmpC and AmpH, proteins related to the class C beta-lactamases, bind penicillin and contribute to the normal morphology of Escherichia coli. J Bacteriol. 1997;179(19):6112-21.
79. Bradford PA. Extended-spectrum beta-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. Clin Microbiol Rev. 2001;14(4):933-51, table of contents.
80. Koran S. WBP, Berlin K., Miller J.R., Philippy A.M. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. bioRxiv preprint. 2016.
81. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol. 2015;16:294.
82. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595.
83. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.

84. WTSI. SMALT: A mapper for DNA sequencing reads [Available from: https://sourceforge.net/projects/smalt/.
85. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.
86. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80.
87. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microbial Genomics. 2016.
88. Stamatakis A. Using RAxML to Infer Phylogenies. Curr Protoc Bioinformatics. 2015;51:6 14 1-.
89. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics. 2017.
90. 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. 2016;44(W1):W242-5.
91. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15.
92. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother. 2014;58(1):212-20.
93. Hunt M, Mather AE, Sanchez-Buso L, Page AJ, Parkhill J, Keane JA, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequence reads. bioRxiv preprint. 2017.
94. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895-903.
95. Pathogen Informatics W. MLST Check: For determining the sequence type from an assembly [Available from: https://github.com/sanger-pathogens/mlst_check.
96. Wiegand I, Hilpert K, Hancock RE. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Nat Protoc. 2008;3(2):163-75.
97. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 2011;12:402.
98. Liakopoulos A, Mevius D, Ceccarelli D. A Review of SHV Extended-Spectrum beta-Lactamases: Neglected Yet Ubiquitous. Front Microbiol. 2016;7:1374.
99. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189-91.
100. Hu H, Hu Y, Pan Y, Liang H, Wang H, Wang X, 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. 2012;56(4):1698-702.