Klebsiella pneumoniae Population Genomics and Antimicrobial Resistant Clones

Kelly L. Wyres[1,2] and Kathryn E. Holt[1,2]

[1]Centre for Systems Genomics, University of Melbourne, Parkville, Victoria 3010, Australia
[2]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia

*Correspondence: kholt@unimelb.edu.au (K.E. Holt).

Keywords
Klebsiella pneumoniae, genomics, antimicrobial resistance, population structure

Abstract

Antimicrobial resistant *Klebsiella pneumoniae* (*Kp*) has emerged as a major global public health problem. While resistance can occur across a broad range of *Kp* clones, a small number have become globally distributed and commonly cause outbreaks in hospital settings. Here we describe recent comparative genomics investigations that have shed light on *Kp* population structure and the evolution of antimicrobial resistant clones. These studies provide the basic framework within which genomic epidemiology and evolution can be understood, but have merely scratched the surface of what can and should be explored. We assert that further large-scale comparative and functional genomics studies are urgently needed to better understand the biology of this clinically important bacterium.

Klebsiella pneumoniae Is a Major Public Health Threat

K. pneumoniae (*Kp*) is a Gram-negative bacteria belonging to the family Enterobacteriaceae. Closely related to the well-known pathogens *Salmonella enterica* and *Escherichia coli*, *Kp* can colonise a similarly wide range of animal hosts, but can also be found in association with plants; in soil, water and drains; and colonising a diversity of body sites including the respiratory tract, gut, nasopharynx, oropharynx and skin [1,2]. *Kp* is considered an opportunistic pathogen, with the majority of infections occurring in neonates, the elderly and the immunocompromised [2]. Urinary tract infection (UTI), pneumonia and wound or soft tissue infections are the most common disease syndromes. *Kp* has been amongst the most frequent agents causing hospital-acquired infections in all settings for many decades [2,3]. It is the 'K' in the ESKAPE pathogens, the six most significant and dangerous causes of drug resistant hospital infections identified by the Infectious Diseases Society of America [4]. More recently, *Kp* has been recognised by the World Health Organization, Centers for Disease Control and Prevention, European Union and other organisations as a significant threat to global public health due to its high rates of antimicrobial resistance (AMR) (see [5] and www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf). This increased attention is largely due to the increasing occurrence of high-profile hospital outbreaks and deaths associated with a particular AMR clone producing the *Kp* carbapenemase (KPC). However it is also associated with the role of *Kp* as the 'canary in the coalmine' – the organism in which most new AMR genes to be discovered in the last two decades were first detected, before becoming widespread in Gram-negative bacterial pathogens [including the extended spectrum beta-lactamase (ESBL) forms of SHV [6] and CTX-M [7]; the carbapenemases KPC [8] and NDM [9]; and most recently MCR-1 [10], the first plasmid-borne gene to be associated with colistin resistance]

The emergence of AMR *Kp* as a major global health problem has coincided with the establishment of whole genome sequencing as a viable tool for investigating and tracking bacterial pathogens, thanks to the development of cost-effective high throughput sequencing. Genomic comparisons can offer a high-resolution view of genetic variation at whole-genome scale and can be applied to explore the diversity of pathogen populations, the evolution of clinically important traits such as AMR, and

patterns of disease transmission and dissemination. Here we review recent insights into the population structure of *Kp* and the evolution of AMR clones gleaned from genomic studies; outline current tools available for genomic investigation of *Kp*; and identify outstanding questions concerning the problem of AMR *Kp* that would benefit from further application of genomics.

Population Structure and Genome Variation

The population structure of *Kp* has been elucidated using various DNA sequencing approaches. A *Kp* multi-locus sequence typing (MLST) scheme, targeting seven chromosomally encoded housekeeping genes, was established in 2005 [11,12]. MLST provides a standardised reproducible system for strain identification and nomenclature for a given species [13]. The *Kp* MLST scheme has been widely adopted and has been centrally important to the identification and investigation of clinically important phylogenetic lineages, which are typically referenced by their sequence type (ST; e.g. ST258). The availability of high throughput whole genome sequencing has since afforded much deeper resolution of the *Kp* population. In 2014, the MLST approach was extended to a core gene MLST (cgMLST) scheme targeting 694 core genes, which can be used to define high-resolution STs and their aggregation into clonal groups (CGs) [14]. The publicly available cgMLST database for *Kp* is hosted at the Institut Pasteur using the BIGSdb platform [15]. It now includes the seven-locus MLST scheme, which still forms the basis for the nomenclature of clinically important *Kp* CGs (e.g. CG258 designates the clonal group that includes ST258). *Kp* genome data can also be interrogated using phylogenetic analysis of single nucleotide polymorphisms (SNPs) across the whole genome [16,17]. In addition to identifying phylogenetic lineages or CGs, this approach can provide a very high-resolution view of recent evolution within CGs, which can be particularly useful for investigating local *Kp* outbreaks and global dissemination patterns [14,17–24].

Isolates identified as *K. pneumoniae* using standard biochemical or proteomics tests typically include three phylogenetically distinct groups or phylogroups that were originally designated KpI, KpII and KpIII but have now been designated as distinct species *K. pneumoniae*, *Klebsiella quasipneumoniae* and *Klebsiella variicola*, respectively [16,25,26]. All three are covered by the same MLST and cgMLST

schemes, which can be used to differentiate the species [11,12]. Whole genome sequence comparison has shown that these groups are distinguished by 3-4% average nucleotide divergence across the core genome, hardly ever recombine, and can be differentiated on the basis of gene content, indicating that they represent distinct independently-evolving populations and supporting their recognition as distinct species [16]. For the remainder of this review, the term *K. pneumoniae* (*Kp*) will be used to refer strictly to *K. pneumoniae* (i.e. the KpI phylogroup).

The *Kp* population is comprised of numerous deep-rooted phylogenetic lineages radiating from a single common ancestor (**Figure 1a**), with approximately 0.5% average nucleotide divergence between lineages [12,16]. These lineages show evidence of occasional homologous recombination [11,12,16,27,28] but estimates of r/m (the relative probability that a nucleotide change resulted from recombination vs point mutation) based on limited MLST data have yielded conflicting results [12,29]. Further investigation of recombination dynamics based on whole genome data is warranted, however the overall population structure appears to be relatively clonal.

A total of 157 lineages were reported based on whole genome analysis of a diverse collection of 289 *Kp* genomes [16] and 155 CGs are currently defined in the public cgMLST database [14], however the rate of discovery of new lineages suggests that the total number in existence far exceeds this, likely reaching the thousands (**Figure 1b**). The long-term persistence of so many distinct *Kp* lineages has yet to be explained. *Kp* occupies a wide range of ecological niches including many non-host associated environments [1,2,16,26]. Extensive exopolysaccharide diversity has been described, but this is not generally associated with phylogenetic lineage. Only 12 O antigen serotypes have been identified in *Kp*, each of which are shared by diverse lineages [30]. *Kp* capsular variation is more extensive: 77 phenotypically defined capsular serotypes are recognised [31–33], and genetic studies of capsule biosynthesis (K) loci indicate the existence of twice this number [18,27,28,30,34,35]. A single capsular serotype can be found in numerous distinct *Kp* lineages and extensive capsular diversity has been identified within lineages, resulting from horizontal transfer and recombination of K locus genes [12,14,16,28,30].

The average *Kp* genome is 5.5 Mbp in size and encodes ~5,500 genes. Whole genome comparisons of hundreds of isolates indicate that the core genome, that is the set of genes that are common to all *Kp*, includes fewer than 2,000 genes [14,16]. The additional 3,500 'accessory' genes in each genome are drawn from a pool of more than 30,000 protein-coding genes (using a cut-off of >30% amino acid divergence to define a new gene; or >70,000 using a cut-off of >10% amino acid divergence) [16]. The rate of accumulation of *Kp* accessory genes with increasing genome sequences indicates the *Kp* population has an open pan genome [36], meaning that *Kp* has access to a vast gene pool (**Figure 2a**). Assignment of *Kp* accessory genes to functional groups identified common functions including carbohydrate metabolism (19%), other metabolic pathways (18%), membrane transport (13%), exopolysaccharide capsule (11%), iron resistance and metabolism (2%) and resistance to antibiotics, heavy metals and stress (1%); a third of protein-coding genes found in *Kp* have as-yet unknown functions [16]. Although there is evidence that individual accessory genes can be distributed across multiple phylogenetic lineages, each lineage is associated with a distinct complement of genes that differs from that of other lineages (see **Figure 2b**) [16]. It is therefore likely that different *Kp* strains vary substantially in their metabolic capacity, which may account for the wide array of ecological niches in which *Kp* is found and also the persistence of distinct chromosomal lineages, which could potentially differ quite substantially from one another in terms of the range of niches that they can readily inhabit. Furthermore, there is evidence that the circulation of highly mobile accessory genes within the *Kp* population, via plasmids and other conjugative elements, may contribute to survival of *Kp* in different niches [16,37–39]. A recent genomic analysis found the presence of a plasmid-encoded *lac* (lactose utilisation) operon, identified in ~50% of sequenced *Kp* isolates, was significantly associated with *Kp* isolated from dairy cows with mastitis, while the presence of plasmid-encoded aerobactin, a siderophore that promotes growth in blood by removing iron from high affinity sites on human transferrin [40], was associated with *Kp* isolated from bacteraemia and other invasive infections in humans [16].

AMR Determinants

Kp is intrinsically resistant to ampicillin due to the presence of the SHV beta-lactamase in the core genome (note *K. quasipneumoniae* and *K. variicola* carry highly divergent forms of this beta-lactamase known as OKP and LEN [16]). Comparative

genomic analysis indicates that *fosA* and the efflux pump *oqxAB*, which confer low-level resistance to fosfomycin and the quinolone nalidixic acid, are also core genes in *K. pneumoniae*, *K. quasipneumoniae* and *K. variicola* [16]. However the majority of AMR in *Kp* results from the acquisition of AMR genes via horizontal transfer, mainly carried by plasmids [41]. More than 100 distinct acquired AMR genes have been identified in *Kp* [16] (**Table 1**), and hundreds of AMR-associated plasmids belonging to dozens of distinct *rep* types (plasmid replication machinery types) have been reported [16,37,41]. It is not uncommon for individual *Kp* strains to carry multiple plasmids, and for several of these to contain distinct sets of AMR genes, resulting in resistance to nearly all available antimicrobials [21,23,37,42]. Direct transfer of AMR plasmids between distinct *Kp* strains, and between *Kp* and other Enterobacteriaceae, has been detected in whole genome sequencing studies of hospitalised patients and in hospital environments, presumably driven by selection from exposure to a range of antimicrobials [42–44].

Of particular clinical concern are the dissemination of carbapenemase genes KPC, OXA-48 and NDM-1, and the ESBL gene CTX-M-15. Each of these genes is associated with a specific transposon that mobilises it between different plasmid backbones (which can then spread to other strains and species) and sometimes into the *Kp* chromosome itself [45–47]. All four genes have been reported in diverse *Kp* lineages. KPC is associated with a broad range of plasmids and is mobilised by Tn*4401*, a 10 kbp Tn*3*-like transposon, for which there are five known isoforms [48,49]. KPC was intimately linked with the emergence of ST258 and its derivative ST512 (see below), but has become more widely disseminated [45,50,51] . OXA-48 is mobilised by Tn*1999* and is most commonly, but not exclusively, associated with IncL/M plasmids [52–55]. NDM-1 is found in a broad range of plasmids of distinct *rep* types but its mechanism of mobilisation is less certain [9]. Complete or truncated IS*Aba1* is often found upstream of NDM-1, suggesting at least an historical role for this insertion sequence (IS) [9,54]. However, there is also evidence of alternative mobilisation e.g. via IS*26* or IS*CR1* [56,57]. CTX-M-15 is mobilised by IS*Ecp1* and in *Kp* is most commonly associated with IncFII plasmids that simultaneously carry other AMR genes [20,21,58–60].

Mutational resistance can also occur in *Kp*. Induced expression of intrinsic efflux pumps such as those encoded by *acrAB* and *oqxAB* have been associated with reduced susceptibility to tigecycline, fluoroquinolones and other antimicrobials [61,62]. Reduced permeability of the outer membrane via functional loss of the outer membrane porins encoded by *ompK35* and *ompK36* can cause resistance to extended spectrum cephalosporins and reduced susceptibility to carbapanems and fluoroquinolones [63]. Fluoroquinolone resistance is often conferred by a combination of substitutions in the genes encoding the topoisomerase targets, GyrA and ParC [64,65]. The presence of these mutations and of acquired AMR plasmids do not necessarily reduce fitness in terms of competitive growth or efficiency of transmission between patients [39,66,67], consequently both are often encountered on first isolation rather than evolving *in vivo* during treatment. In areas where fluoroquinolone and carbapenem resistance is common, treatment of *Kp* infections generally relies on tigecycline or colistin [68]. Colistin resistance is rare upon first isolation but often arises during treatment via mutations that upregulate the PhoQ/PhoP system and *pmrHFIJKLM* operon, most commonly by inactivation of *mgrB* via IS insertions, but also occasionally by deletions or nonsense mutations in this gene or others involved in the same pathway [69–71]. Additional mechanisms of colistin resistance have recently been reported, including mutations in the chromosomal *crrB* gene [72] and acquisition of the plasmid-borne genes *mcr-1* or *mcr-1.2* [10,73]. It was initially hoped that *mgrB* inactivation would compromise the ability of *Kp* to transmit and cause infections in new hosts. However studies to date have found no fitness cost during *in vitro* competitive growth [74] or animal models [75] and sustained outbreaks of *mgrB*-mutant colistin resistant strains have been reported [76]. Tigecycline resistance in *Kp* is usually caused by increased activity of the AcrAB efflux pump via interruption of the regulators *ramA*, *ramR* or *acrR* [77–79]. A non-synonymous substitution in the *rpsJ* gene (encoding the S10 30S ribosomal subunit) has also been implicated in tigecycline resistance [80].

Genomic Insights Into the Emergence of Antibiotic Resistant Clones

AMR has emerged within many distinct *Kp* and some *K. variicola* CGs [14,16,19,81], however a small number have become widely disseminated and commonly cause infections in a range of settings, despite the fact that they are not generally associated with any of the known *Klebsiella* virulence determinants [14,16]. **Figure 3** shows the

geographical distribution of *Kp* outbreaks reported in the literature and associated with a CG identified by MLST, as of 24th June 2016. These represent just the tip of the iceberg of the global burden of *Kp* outbreaks, since most outbreaks are not reported in the literature and MLST data are not ubiquitously generated. Notably, of all reported outbreaks where MLST was performed, 72% identified one of five common CGs (CG258, CG14/15, CG17/20, CG43, CG147, **Figure 3**). Twenty-two of the remaining 24 outbreaks were associated with *Kp* STs, one was associated with *K. variicola* (ST48 and its single locus variant, ST1236) and one was associated with *K. quasipneumoniae* (ST334). Genomic investigations of some of these common CGs, or 'clones' are beginning to provide specific insights into their evolution.

CG258

Undoubtedly the most widely recognised and globally distributed clone is CG258 (ST258, ST11, their single locus variants and other close relatives, e.g ST340, ST512, ST437, ST833, ST855 and ST1199). ST258 is widely acknowledged as the major cause of carbapenem-resistant *Kp* infections [48,82,83] and is predominantly associated with the KPC-2 and KPC-3 carbapenemases. In contrast, other members of this CG have been associated with a more diverse selection of carbapenemases and ESBLs, including NDM-1, OXA-48 and CTX-M-15 [19,81,84–86]. The epidemiology of CG258 has been well reviewed previously [48,49,82,83] so here we focus on the most recent evolutionary insights from comparative genomic studies.

An analysis of 319 *Kp* genomes, including 203 CG258 (predominantly ST258 and ST11) suggested that a large genomic recombination event of ~1.3 Mbp length distinguishes CG258 from its closest relatives [81] **(Figure 4)**. This event was dated to ~1985, suggesting that the most-recent common ancestor of CG258 was circulating in the population at that time. ST258, ST340 and ST437 each form a single monophyletic sub-clade within CG258, while ST11 is a paraphyletic group [19,28]. ST258 arose from an ST11-like ancestor following a second large-scale genomic recombination event, in which a ~1.1 Mbp genomic region was acquired from an ST442 *Kp* [27,28]. The recombinant region included the K locus, which was distinct from the ST11-like ancestor and presumably associated with a change of capsule phenotype **(Figure 4)**. Subsequently ST258 also acquired an integrative conjugative element known as ICE258.2, which encoded a type IV pilus and a type III restriction modification system [23,27]. It was speculated that the former may facilitate improved adherence, while the latter may play a role in determining which plasmids can be maintained within ST258 [23].

Early studies had suggested that ST258 was further divided into two distinct sub-lineages (I and II), distinguished by a third large-scale genomic recombination event of ~215 kbp [23,87] **(Figure 4)**. Again the recombinant region, which was acquired from an ST42 *Kp,* included a distinct K locus [23,28]. Subsequently, Bowers and colleagues showed that sub-lineages I and II actually form a monophyletic sub-clade within ST258, and the remainder of the clade is paraphyletic [19]. Isolates from the United States were distributed throughout; supporting the hypothesis that ST258 arose

in that country, where it was first identified and remains highly prevalent [19,88]. Further molecular dating analyses suggested the origin of ST258 circa 1995-1997 [19,81], just a few years before the first clinical reports [88,89].

A total of 22 distinct K loci have now been associated with CG258, each of which presumably imported by an independent recombination event [19,28]. The extensive variability of this locus suggests that it is subject to strong diversifying selection, although the drivers are as yet unclear. CG258 is also highly diverse in terms of acquired AMR genes and chromosomal AMR-conferring variants, suggesting that AMR has arisen independently multiple times, largely driven by the acquisition of a diverse array of plasmids [19,22,23,42]. ST258 isolates typically harbour between two and five plasmids of 10.9 kbp to 142.7 kbp [23,42]. The majority, although not all [19,90], ST258 harbour at least one plasmid containing either KPC-2 or KPC-3. pKpQIL is one such plasmid that is common among sub-lineages I and II [19], but rare among the rest of the clade [22,23,42]. In fact, sub-lineages I and II are generally associated with greater conservation of plasmids compared to the rest of the CG, which is highly diverse [19]. Taken together, these genomic studies unravel a story of a rapidly evolving, highly adaptive epidemic clone.

CG14/15

CG14/15 is another globally distributed MDR clone [18,20,91–93]. Similar to CG258, it has also been associated with a diverse array of AMR genes, including those encoding ESBLs (in particular CTX-M-15 [18,20,94]) and carbapenemases such as KPC [95], NDM-1 [18], OXA-48 [91], OXA-181 [93] and VIM-1 [92]. Colistin resistance has been reported both with and without concomitant ESBL and/or carbapenemase production [70,96].

Genomic analyses of ST15 isolates from The Netherlands and Nepal showed that they can be divided into at least two sub-lineages, each associated with a distinct K locus [18,20]. All of the Nepalese isolates harboured CTX-M-15, while 42 also harboured NDM-1. The latter isolates were part of an outbreak from which nine NDM-1 negative isolates were also identified [18,21]. Long read SMRT sequencing of a representative outbreak isolate identified four distinct plasmid replicons ranging from 69 kbp to 305 kbp. Three of the four plasmids contained AMR genes and/or heavy

metal resistance genes. The fourth plasmid contained a tellurite resistance cassette. The largest plasmid, pMK1-NDM, harboured NDM-1 in combination with CTX-M-15, OXA-1, *aac(6')-Ib-cr*, *aadA2*, *folP*, *catA1*, *dfrA12* and *armA* [21]. Short read Illumina sequencing data suggested that all of the outbreak isolates harboured pMK1-NDM-like plasmids, including those that were NDM-1 negative due to deletion of the NDM-1 region [18,21].

Other Clonal Groups

Several other globally distributed MDR clones including CG17/20, CG43 and CG147 have been associated with a number of disease outbreaks (**Figure 3**). All were first recognised in the mid-late 2000s and are associated with a range of different AMR genes. Of note, ST101 from CG43 seems to be widely distributed in Europe and is commonly associated with CTX-M-15, largely through plasmid acquisition [46,70,97–100]. However, a genome sequence from a representative isolate of an ST101 outbreak in Germany showed that this strain harboured a chromosomal copy of the IS*Ecp1*-CTX-M-15 transposon [46]. Isolates from this outbreak were resistant to extended spectrum beta-lactams, gentamicin, tetracycline, ciprofloxacin and sulphamethoxazole/trimethoprim and harboured CTX-M-15, TEM-1, and plasmid replicons FIA and FIB. Aside from CTX-M-15, the location of the remaining AMR genes was unclear [46]. This finding is potentially of concern given that the fitness cost of chromosomal CTX-M-15 is likely much reduced compared to the cost of maintenance of an entire CTX-M-15 plasmid. Consequently, it is more likely that the host will retain the gene even in the absence of antimicrobial selective pressure. Unfortunately, CG43 is not the only *Kp* AMR clone within which chromosomal CTX-M-15 has been reported. More worryingly, the genome of an ST147 isolate from the United Arab Emirates contained a chromosomal IS*Ecp1*-CTX-M-15 plus three chromosomal copies of IS*Ecp1*-OXA-181, which conferred resistance to the carbapenems [47]. The situation was worsened by the fact that one of the IS*Ecp1*-OXA-181 transposons had interrupted the *mgrB* gene, resulting in colistin resistance and generating a truly pan-resistant strain [47].

Concluding Remarks and Future Perspectives

There is now widespread recognition of the immense potential for genomics to enhance surveillance and tracking of specific pathogens and of AMR more generally, and to aid infection control and outbreak investigations. Several studies have reported the use of genomics to aid investigations of AMR *Kp* outbreaks in hospitals, with emerging themes being the detection of persistent polyclonal outbreaks resulting from transmission of AMR plasmids as well as AMR clones; asymptomatic colonisation of healthcare workers and patients with AMR clones; and sinks, taps and drains as persistent reservoirs of infection [17,22,42,43]. We contend that analysis and interpretation of genome data generated in such studies will be greatly assisted in the future by the emerging genomic framework for *Kp*, which helps investigators to readily extract the most useful information and place it in the context of the existing knowledge base. Currently the key elements of the *Kp* genomic framework are identification of CGs; AMR determinants including acquired genes and common mutations; known virulence genes and alleles; plasmids; and capsular and O antigen loci. Details of current data sources and tools for extracting these elements from *Kp* genome data are given in **Box 1**.

While the availability of thousands of *Kp* genomes may sound ample to some, we believe there is a pressing need to dramatically expand our current understanding of the *Kp* population through further functional, clinical and ecological genomics studies. Understanding of *Kp* disease, transmission and evolution is arguably decades behind that of other human pathogens, but genomics can help scientists and clinicians to rapidly advance our knowledge of this important threat to global health. Studies to date show population structure of *Kp* is complex and intriguing, and raises important questions about the functional and ecological differences between lineages, which are highly relevant to understanding why certain *Kp* lineages appear to pose greater clinical problems than others (see **Outstanding Questions**). Functional genomics studies are needed to identify factors involved in environmental persistence of *Kp*, as well as transmission, colonisation, and pathogenicity in humans [101]. Functional genomics can also be used to search for lineage-specific factors that might explain why certain AMR determinants appear to be maintained in some CGs but transient in others [67,102], which could be novel targets for inhibition of the seemingly never-ending accumulation of AMR in the problem clones. Analysis of the available

361 genome data indicates that the *Kp* sequenced so far represent the tip of the iceberg of
362 a much larger *Kp* population (**Figure 1b, 2a**). Much deeper sampling will be required
363 in order to begin to understand the ecology of *Kp*, which could identify important
364 reservoirs of bacterial diversity and help to understand why *Kp* appears to have so
365 often been the first step in the trafficking of AMR genes from environmental bacteria
366 into human-associated bacterial populations.
367
368 After simmering away for decades, the problem of AMR *Kp* has become too
369 important to ignore and the international medical, public health and scientific
370 communities now need to play catch-up. Genomics has played a key role in the past
371 few years and has plenty more to offer in tackling the global threat of AMR *Kp*.
372 Given the scale of the challenge, it will be important to continue to build a deeper
373 understanding of the underlying population out of which problem clones emerge and
374 to share genomic data together with associated source and phenotypic data, in order to
375 maximize the potential benefits of genomic approaches.

Figure Legends

Figure 1. Lineage Diversity in *Klebsiella pneumoniae*. (a) Core gene phylogeny for *K. pneumoniae*. Unrooted maximum likelihood phylogenetic tree for 283 isolates sampled from diverse sources and locations, tips are coloured by country as indicated in panel b. (b) Discovery of novel *K. pneumoniae* lineages with increasing sampling of isolates in different locations. Curves show the discovery rate for new *K. pneumoniae* lineages as more isolates were sampled for whole genome sequencing; Simpson's diversity index is shown in parentheses. Plots are reproduced from [16]; tree and source information are available for interactive viewing at https://microreact.org/project/BJClQz9H.

Figure 2. Gene Content Diversity in *Klebsiella pneumoniae*. (a) *K. pneumoniae* pan genome. Curves show the discovery rate for new *K. pneumoniae* protein-coding genes as more isolates were sampled for whole genome sequencing (mean and 95% confidence interval for each sample size). Different absolute numbers are obtained depending on the level of amino acid (aa) identity used to define a new protein-coding gene, however both curves show that the *K. pneumoniae* population has an open pan genome, indicating there is no upper limit to the number of accessory genes that the population can sustain. (b) Differences in gene content within and between *K. pneumoniae* lineages. Boxplots show the distribution of gene content distances (measured using Jaccard distance) for pairs of *K. pneumoniae* genomes that belong to the same (blue) or different (green) lineages. Plots are reproduced from data in [16].

Figure 3. Distribution of *Klebsialla pneumoniae* Outbreaks by Clonal Group (CG) and Region. Outbreak reports as of June 2016 were identified in the literature by PubMED search using the following search terms; "*Klebsiella pneumoniae*" AND "outbreak" AND (one of "MLST" OR "multilocus sequence typing"); "*Klebsiella pneumoniae*" AND "outbreak" AND (one of "ST1*" … "ST9*" OR "CG1*" ... "CG9*" OR "CC1*" … "CC9*"). Pie graph areas are proportional to the total number of outbreaks reported in each World Health Organization region (each region is indicated by a different shade of grey), slices indicate frequency of each CG. CG 258 is divided into two categories; ST258 and its derivative ST512; and the remaining sequence types (STs) identified in the literature search (ST11, ST340 and ST437). CG 14/15 includes ST14 and ST15; CG 17/20 includes ST16, ST17 and ST20; CG 43 includes ST101; CG 147 includes ST147 and ST273; other indicates outbreaks caused by 22 different *Kp* STs that are not part of any named CG, one *K. variicola* ST and its derivative (ST48 and ST1236, respectively) and one *K. quasipneumoniae* ST334. Red stars indicate the locations of the earliest recorded ST258 outbreaks in the United States and Israel, for which MLST was not applied. Blue star indicates the location of the Nepalese ST15 outbreak, which did not meet the search criteria but is described in the main text.

Figure 4. Genomic Evolution of *Klebsiella pneumoniae* Clonal Group (CG) 258. A schematic cladogram of the relationships within CG258 is shown alongside colour bars that represent the bacterial chromosome. Coloured blocks represent regions of the genome acquired through horizontal transfer from a *K. pneumoniae* that is not part of CG258, as indicated by the arrows. The relative positions of the seven *K. pneumoniae* multi-locus sequence typing loci are indicated by grey pointers. The position of the K locus is indicated by an orange pointer. ST258 lineage I and II are labelled ST258-I and ST258-II, respectively.

428 **Table 1. Genetic Determinants of AMR in *Klebsiella pneumoniae* Genomes.**

429

Beta-lactamases	***bla* genes conferring resistance (*intrinsic)**	
Class A	CARB-3, PSE-1, SCO-1, SHV-1*, TEM-1	
- ESBL	CTX-M, SHV-5, TEM-10, VEB	
- Carbapenemase	KPC, GES-5	
Class B (Metallo-beta-lactamase)	CphA, IMP, NDM, SIM, VIM	
Class C (Cephalosporinase)	AmpC, CMY, DHA, FOX, MIR	
Class D	OXA-1, OXA-2, OXA-7, OXA-9, OXA-10, OXA-12	
- ESBL	OXA-11, OXA-15	
- Carbapenemase	OXA-48, OXA-51, OXA-181, OXA-237	
Other AMR	**Genes conferring resistance (*intrinsic)**	**Mutations**
Aminoglycosides	*aac, aadA, aadB, aph, armA, rmt, strAB*	-
Carbapenems	(see carbapenemase *bla* genes, class A & D above)	Mutations in *ompK35*, *ompK36*
Colistin	*mcr-1, mcr1.2*	Inactivation of *pmrB, mgrB*; mutations in *crrB*
Fluoroquinolones	*qepA, qnrA, qnrB, qnrD, qnrS, qepA*	SNPs in *gyrA, parC;* Upregulation of *acrAB* or *oqxAB* efflux
Macrolides	*ereA, ereB, ermB, mef, mph, msrE*	-
Phenicols	*catA, catB, cml, floR*	-
Rifampin	*arr*	
Sulfonamides	*folP, sul1, sul2, sul3*	-
Tetracycline	*tet* genes	-
Tigecycline	-	Upregulation of *acrAB* or *oqxAB* efflux; mutation in *rpsJ*
Trimethoprim	*dfr* genes	-

430

Box 1. Tools and Databases for *K. pneumoniae* Genomic Analyses.

- ***Klebsiella pneumoniae* BIGSdb:** An online database and integrated set of tools for analysis of genome assemblies [14]. The *K. pneumoniae* MLST database, cgMLST, virulence and resistance gene databases are available through this single resource, which also hosts a searchable repository of *K. pneumoniae, K. variicola and K. quasipneumoniae* genomes. As of June 2016 the database includes 2328 distinct STs. Available at bigsdb.pasteur.fr/klebsiella.

- **Centre for Genomic Epidemiology:** A suite of online tools for analysis of genome assemblies or short read data. *K. pneumoniae* MLST analysis [103], virulence and AMR gene screening [104,105], and plasmid screening [106] are all available. The AMR screening protocol uses the ResFinder database [105]. Available at www.genomicepidemiology.org.

- **SRST2:** An offline tool for allelic typing from short read sequence data [107]. MLST, virulence and resistance gene screening can be achieved directly from sequence reads. In fact, SRST2 can be used in conjunction with any appropriately formatted gene or allelic database. Available at github.com/katholt/srst2.

- **ISmapper:** An offline tool for determination of insertion sequences (IS), copy number and insertion sites within genomes [108]. ISmapper takes as input paired-end short read sequence data, a genome assembly or reference genome and a set of IS references. Available at github.com/jhawkey/IS_mapper.

- **ISfinder:** A searchable online database of bacterial IS. Users can access and/or download IS nucleotide sequences and relevant information, including general features, direct and inverted repeat sequences and predicted protein sequences. There is a BLASTn query function, a list of IS annotated bacterial genomes and a browser for visualisation of IS within genomes. Available at www-is.biotoul.fr.

- **Kaptive:** A database of complete sequences of *Klebsiella* capsule loci and accompanying tool for identification and typing of capsule loci from genome assemblies. Available at github.com/katholt/kaptive.

- **NCBI Pathogen Detection resources**: Curated databases of AMR genes and genomes of antimicrobial resistant bacterial pathogens. As at June 2016 the databases include 3,275 AMR gene nucleotide sequences and 2,391 annotated genomes drawn from Genbank. Genome-wide phylogenetic analyses, pre-

computed at the species level, can also be accessed. Available at www.ncbi.nlm.nih.gov/pathogens.

- **PATRIC database:** An integrated resource for analysis and exploration of pathogen genomes including Klebsiella. Users can access and download hundreds of Klebsiella genome assemblies with accompanying annotation and source information. Protein sorting and metabolic pathway comparison tools are also included. Available at www.patricbrc.org.

References

1 Bagley, S.T. (1985) Habitat association of *Klebsiella* species. *Infect Contr.* 6, 52–58

2 Podschun, R. and Ullmann, U. (1998) *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev.* 11, 589–603

3 Jones, R.N. (2010) Microbial etiologies of hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. *Clin Infect Dis.* 51, S81–S87

4 Pendleton, J.N. *et al.* (2013) Clinical relevance of the ESKAPE pathogens. *Exp Rev Anti Infect Ther.* 11, 297–308

5 World Health Organization (2014) *Antimicrobial resistance: global report on surveillance*,

6 Chaves, J. *et al.* (2001) SHV-1 β-lactamase is mainly a chromosomally encoded species-specific enzyme in *Klebsiella pneumoniae*. *Antimicrob Agents Chemother.* 45, 2856–2861

7 Sirot, J. *et al.* (1988) *Klebsiella pneumoniae* and other Enterobacteriaceae producing novel plasmid-mediated beta-lactamases markedly active against third-generation cephalosporins: Epidemiologic studies. *Clin Infect Dis.* 10, 850–859

8 Nordmann, P. *et al.* (2009) The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. *Lancet Infect Dis.* 9, 228–236

9 Nordmann, P. *et al.* (2011) The emerging NDM carbapenemases. *Trends Microbiol.* 19, 588–595

10 Liu, Y.Y. *et al.* (2016) 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.* 16, 161–168

11 Diancourt, L. *et al.* (2005) Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *J Clin Microbiol.* 43, 4178–4182

12 Brisse, S. *et al.* (2009) Virulent clones of *Klebsiella pneumoniae*: Identification and evolutionary scenario based on genomic and phenotypic characterization. *PLoS One* 4, e4982

13 Maiden, M.C.J. (2006) Multilocus sequence typing of bacteria. *Annu Rev*

Microbiol. 60, 561–588

14 Bialek-Davenet, S. *et al.* (2014) Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis.* 20, 1812–1820

15 Jolley, K.A. and Maiden, M.C. (2010) BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics.* 10, 595

16 Holt, K.E. *et al.* (2015) Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci U S A.* 112, E3574-81

17 Snitkin, E.S. *et al.* (2012) Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med.* 4, 148ra116

18 The, H.C. *et al.* (2015) A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of *Klebsiella pneumoniae*. *EMBO Molec Med.* 7, 227–239

19 Bowers, J.R. *et al.* (2015) Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS One* 10, e0133727

20 Zhou, K. *et al.* (2016) Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum β-lactamase producing ST15 *Klebsiella pneumoniae*. *Sci Rep.* 6, 20840

21 Stoesser, N. *et al.* (2014) 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

22 Marsh, J.W. *et al.* (2015) Genomic epidemiology of an endoscope-associated outbreak of *Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae*. *PLoS One* 10, e0144310

23 Deleo, F.R. *et al.* (2014) Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A.* 111, 4988–4993

24 Struve, C. *et al.* (2015) Mapping the evolution of hypervirulent *Klebsiella pneumoniae*. *MBio.* 6, 1–12

25 Brisse, S. *et al.* (2014) Description of *Klebsiella quasipneumoniae* sp., isolated from human infections, with two subspecies, *Klebsiella quasipneumoniae subsp. quasipneumoniae* subsp. nov. and *Klebsiella quasipneumoniae subsp. similipneumoniae* subsp. nov., and. *Int J Syst Evol Microbiol.* DOI: 10.1099/ijs.0.062737-0

26 Rosenblueth, M. *et al.* (2004) Klebsiella variicola , a novel species with clinical and plant-associated isolates. *Syst Appl Microbiol.* 27, 27–35

27 Chen, L. *et al.* (2014) Epidemic *Klebsiella pneumoniae* ST258 Is a hybrid strain. *MBio* 5, e01355-14

28 Wyres, K.L. *et al.* (2015) Extensive capsule locus variation and large-scale genomic recombination within the *Klebsiella pneumoniae* clonal group 258. *Genome Biol Evol.* 7, 1267–1279

29 Vos, M. and Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. *ISME J* 3, 199–208

30 Follador, R. *et al.* (2016) The diversity of *Klebsiella pneumoniae* surface polysaccharides. *MGen.* 2, DOI: 10.1099/mgen.0.000073

31 Edwards, P.R. and Fife, M.A. (1952) Capsule types of *Klebsiella*. *J Infect Dis.* 91, 92–104

32 Edmunds, P.N. (1954) Further *Klebsiella* capsule types. *J Infect Dis.* 94, 65–71

33 Ørskov, I.D.A. and Fife-Asbury, M.A. (1977) New *Klebsiella* capsular antigen, K82, and the deletion of five of those previously assigned. *Int J Syst Bacteriol.* 27, 386–387

34 Pan, Y.-J. *et al.* (2015) Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of *Klebsiella* spp. *Nat Sci Rep.* 5, 15573

35 Wyres, K.L. *et al.* (2016) Identification of *Klebsiella* capsule synthesis loci from whole genome data. *bioRxiv Prepr.* DOI: http://dx.doi.org/10.1101/071415

36 Medini, D. *et al.* (2005) The microbial pan-genome. *Curr Opin Genet Dev.* 15, 589–594

37 Ramirez, M.S. *et al.* (2014) Plasmid-mediated antibiotic resistance and virulence in gram-negatives: the *Klebsiella pneumoniae* paradigm. *Microbiol Spectr.* 2, 1–15

38 Bojer, M.S. *et al.* (2013) ClpP-dependent and -independent activities encoded by the polycistronic ClpK-encoding locus contribute to heat shock survival in

Klebsiella pneumoniae. *Res Microbiol.* 164, 205–210

39 Löhr, I.H. *et al.* (2015) Persistence of a pKPN3-like CTX-M-15-encoding IncFIIK plasmid in a *Klebsiella pneumoniae* ST17 host during two years of intestinal colonization. *PLoS One* 10, e0116516

40 Paczosa, M.K. and Mecsas, J. (2016) *Klebsiella pneumoniae*: Going on the offense with a strong defense. *Microbiol Mol Biol Rev.* 80, 629–661

41 Iredell, J. *et al.* (2016) Antibiotic resistance in Enterobacteriaceae: mechanisms and clinical implications. *BMJ* 352, h6420

42 Conlan, S. *et al.* (2014) Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Sci Transl Med.* 6, 254ra126

43 Mathers, A.J. *et al.* (2015) *Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae* at a single institution: Insights into endemicity from whole-genome sequencing. *Antimicrob Agents Chemother.* 59, 1656–1663

44 Mathers, A.J. *et al.* (2011) Molecular dissection of an outbreak of carbapenem-resistant Enterobacteriaceae reveals intergenus KPC carbapenemase transmission through a promiscuous plasmid. *MBio.* 2, e00204–e00211

45 Sheppard, A.E. *et al.* (2016) Nested Russian doll-like genetic mobility drives rapid dissemination of the carbapenem resistance gene *bla*KPC. *Antimicrob Agents Chemother.* 60, 3767–3778

46 Mshana, S.E. *et al.* (2015) Molecular epidemiology and characterization of an outbreak causing *Klebsiella pneumoniae* clone carrying chromosomally located blaCTX-M-15 at a German University-Hospital. *BMC Microbiol.* 15, 122

47 Zowawi, H.M. *et al.* (2015) Stepwise evolution of pandrug-resistance in *Klebsiella pneumoniae*. *Sci Rep.* 5, 15082

48 Chen, L. *et al.* (2014) Carbapenemase-producing *Klebsiella pneumoniae*: Molecular and genetic decoding. *Trends Microbiol.* 22, 686–696

49 Lee, C.-R. *et al.* (2016) Global dissemination of carbapenemase-producing *Klebsiella pneumoniae*: epidemiology, genetic context, treatment options, and detection methods. *Front Microbiol.* 7, 1–30

50 Conlan, S. *et al.* (2016) Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. 7, e00742-16

51 Adler, A. *et al.* (2016) Dissemination of the *bla*KPC gene by clonal spread and horizontal gene transfer: comparative study of incidence and molecular

mechanisms. *J Antimicrob Chemother.* 71, 2143–2146

52 Poirel, L. *et al.* (2012) Genetic features of the widespread plasmid coding for the carbapenemase OXA-48. *Antimicrob Agents Chemother.* 56, 559–562

53 Potron, A. *et al.* (2013) Intercontinental spread of OXA-48 beta-lactamase-producing Enterobacteriaceae over a 11-year period, 2001 to 2011. *Euro Surveill.* 18, 20549

54 Czobor, I. *et al.* (2016) Efficient transmission of IncFIIY and IncL plasmids and *Klebsiella pneumoniae* ST101 clone producing OXA-48, NDM-1 or OXA-181 in Bucharest hospitals. *Int J Antimicrob Agents.* 48, 223–224

55 Pérez-Vázquez, M. *et al.* (2016) Phylogeny, resistome and mobile genetic elements of emergent OXA-48 and OXA-245 Klebsiella pneumoniae clones circulating in Spain. *J Antimicrob Chemother.* 71, 887–896

56 Doi, Y. *et al.* (2014) Whole-genome assembly of *Klebsiella pneumoniae* coproducing NDM-1 and OXA-232 carbapenemases using single-molecule, real-time sequencing. *Antimicrob Agents Chemother.* 58, 5947–5953

57 Chen, C.J. *et al.* (2014) Closely related NDM-1-encoding plasmids from *Escherichia coli* and *Klebsiella pneumoniae* in Taiwan. *PLoS One* 9, 1–6

58 Coelho, A. *et al.* (2010) Characterisation of the CTX-M-15-encoding gene in *Klebsiella pneumoniae* strains from the Barcelona metropolitan area: Plasmid diversity and chromosomal integration. *Int J Antimicrob Agents.* 36, 73–78

59 Markovska, R. *et al.* (2014) Predominance of IncL/M and IncF plasmid types among CTX-M-ESBL-producing *Escherichia coli* and *Klebsiella pneumoniae* in Bulgarian hospitals. *Apmis* 122, 608–615

60 Dolejska, M. *et al.* (2013) Plasmid content of a clinically relevant *Klebsiella pneumoniae* clone from the Czech Republic producing CTX-M-15 and QnrB1. *Antimicrob Agents Chemother.* 57, 1073–1076

61 Padilla, E. *et al.* (2010) *Klebsiella pneumoniae* AcrAB efflux pump contributes to antimicrobial resistance and virulence. *Antimicrob Agents Chemother.* 54, 177–183

62 Bialek-Davenet, S. *et al.* (2015) Differential contribution of AcrAB and OqxAB efflux pumps to multidrug resistance and virulence in *Klebsiella pneumoniae*. *J Antimicrob Chemother.* 70, 81–88

63 Martinez-Martinez, L. (2008) Extended-spectrum beta-lactamases and the permeability barrier. *Clin Microbiol Infect.* 14 Suppl 1, 82–89

64 Weigel, L.M. *et al.* (1998) *gyrA* mutations associated with fluoroquinolone resistance in eight species of Enterobacteriaceae. *Antimicrob Agents Chemother.* 42, 2661–2667

65 Chen, F. *et al.* (2003) The roles of mutations in *gyrA, parC*, and *ompK35* in fluoroquinolone resistance in *Klebsiella pneumoniae*. *Microb. Drug Res.* 9, 265–271

66 Tsai, Y.K. *et al.* (2011) *Klebsiella pneumoniae* outer membrane porins OmpK35 and OmpK36 play roles in both antimicrobial resistance and virulence. *Antimicrob Agents Chemother.* 55, 1485–1493

67 Tóth, A. *et al.* (2014) Fitness cost associated with resistance to fluoroquinolones is diverse across clones of *Klebsiella pneumoniae* and may select for CTX-M-15 type extended-spectrum β-lactamase. *Eur J Clin Microbiol Infect Dis.* 33, 837–843

68 Doi, Y. and Paterson, D.L. (2015) Carbapenemase-producing Enterobacteriaceae. *Semin Respir Crit Care Med.* 36, 74–84

69 Cannatelli, A. *et al.* (2013) In vivo emergence of colistin resistance in *Klebsiella pneumoniae* producing KPC-type carbapenemases mediated by insertional inactivation of the PhoQ/PhoP mgrB regulator. *Antimicrob Agents Chemother.* 57, 5521–5526

70 Jayol, A. *et al.* (2014) Resistance to colistin associated with a single amino acid change in protein PmrB among *Klebsiella pneumoniae* isolates of worldwide origin. *Antimicrob Agents Chemother.* 58, 4762–4766

71 Poirel, L. *et al.* (2015) The *mgrB* gene as a key target for acquired resistance to colistin in *Klebsiella pneumoniae*. *J Antimicrob Chemother.* 70, 75–80

72 Cheng, Y.-H. *et al.* (2016) Amino acid substitutions of CrrB responsible for resistance to colistin through CrrC in *Klebsiella pneumoniae*. *Antimicrob Agents Chemother.* 60, 3709–3716

73 Di Pilato, V. *et al.* (2016) MCR-1.2: a new MCR variant encoded by a transferable plasmid from a colistin-resistant KPC carbapenemase-producing Klebsiella pneumoniae of sequence type 512. *Antimicrob Agents Chemother.* 60, 5612–5615

74 Cannatelli, A. *et al.* (2015) Polymyxin resistance caused by *mgrB* inactivation is not associated with significant biological cost in *Klebsiella pneumoniae*. *Antimicrob Agents Chemother.* 59, 2898–28900

75 Arena, F. *et al.* (2016) Colistin resistance caused by inactivation of the MgrB regulator is not associated with decreased virulence of sequence type 258 KPC carbapenemase-producing *Klebsiella pneumoniae*. *Antimicrob Agents Chemother.* 60, 2509–2512

76 Giani, T. *et al.* (2015) Large nosocomial outbreak of colistin-resistant, carbapenemase-producing *Klebsiella pneumoniae* traced to clonal expansion of an *mgrB* deletion mutant. *J Clin Microbiol.* 53, 3341–3344

77 Hentschke, M. *et al.* (2010) *ramR* mutations in clinical isolates of *Klebsiella pneumoniae* with reduced susceptibility to tigecycline. *Antimicrob Agents Chemother.* 54, 2720–2723

78 Roy, S. *et al.* (2013) Tigecycline susceptibility in *Klebsiella pneumoniae* and *Escherichia coli* causing neonatal septicaemia (2007-10) and role of an efflux pump in tigecycline non-susceptibility. *J Antimicrob Chemother.* 68, 1036–1042

79 Ruzin, A. *et al.* (2005) Influence of transcriptional activator RamA on expression of multidrug efflux pump AcrAB and tigecycline susceptibility in *Klebsiella pneumoniae*. *Antimicrob Agents Chemother.* 49, 1017–1022

80 Villa, L. *et al.* (2014) Genomics of KPC-producing *Klebsiella pneumoniae* sequence type 512 clone highlights the role of RamR and ribosomal S10 protein mutations in conferring tigecycline resistance. *Antimicrob Agents Chemother.* 58, 1707–1712

81 Gaiarsa, S. *et al.* (2015) Genomic epidemiology of *Klebsiella pneumoniae* in Italy and novel insights into the origin and global evolution of its resistance to carbapenem antibiotics. *Antimicrob Agents Chemother.* 59, 389–396

82 Munoz-Price, L.S. *et al.* (2013) Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis.* 13, 785–796

83 Pitout, J.D.D. *et al.* (2015) Carbapenemase-producing *Klebsiella pneumoniae*, a key pathogen set for global nosocomial dominance. *Antimicrob Agents Chemother.* 59, 5873–5884

84 Lian-Hui, W. *et al.* (2016) Diversity of the genetic environment of the *bla*KPC-2 gene among *Klebsiella pneumoniae* clinical isolates in a Chinese hospital. *Microb. Drug Res.* 22, 15–21

85 Pereira, P.S. *et al.* (2013) Update of the molecular epidemiology of KPC-2-producing *Klebsiella pneumoniae* in Brazil: Spread of clonal complex 11

(ST11, ST437 and ST340). *J Antimicrob Chemother.* 68, 312–316

86 Baraniak, A. *et al.* (2016) NDM-producing Enterobacteriaceae in Poland, 2012-14: Inter-regional outbreak of *Klebsiella pneumoniae* ST11 and sporadic cases. *J Antimicrob Chemother.* 71, 85–91

87 Wright, M.S. *et al.* (2014) Population Structure of KPC-producing *Klebsiella pneumoniae* from Midwestern US hospitals. *Antimicrob Agents Chemother* 58, 4961–4965

88 Kitchel, B. *et al.* (2009) Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* isolates in the United States: clonal expansion of multilocus sequence type 258. *Antimicrob Agents Chemother.* 53, 3365–3370

89 Deshpande, L.M. *et al.* (2006) Occurrence and characterization of carbapenemase-producing Enterobacteriaceae: report from the SENTRY Antimicrobial Surveillance Program (2000-2004). *Microb Drug Res.* 12, 223–230

90 Adler, A. *et al.* (2012) A swordless knight: Epidemiology and molecular characteristics of the *bla*KPC-negative sequence type 258 Klebsiella pneumoniae clone. *J Clin Microbiol.* 50, 3180–3185

91 Thomas, C.P. *et al.* (2013) Early (2008-2010) hospital outbreak of *Klebsiella pneumoniae* producing OXA-48 carbapenemase in the UK. *Int J Antimicrob Agents.* 42, 531–536

92 Sánchez-Romero, I. *et al.* (2012) Nosocomial outbreak of VIM-1-producing *Klebsiella pneumoniae* isolates of multilocus sequence type 15: molecular basis, clinical risk factors, and outcome. *Antimicrob Agents Chemother.* 56, 420–427

93 Balm, M.N.D. *et al.* (2013) OXA-181-producing *Klebsiella pneumoniae* establishing in Singapore. *BMC Infect Dis.* 13, 58

94 Novais, A. *et al.* (2012) Spread of an OmpK36-modified ST15 *Klebsiella pneumoniae* variant during an outbreak involving multiple carbapenem-resistant Enterobacteriaceae species and clones. *Eur J Clin Microbiol Infect Dis.* 31, 3057–3063

95 Stillwell, T. *et al.* (2015) Outbreak of KPC-3 producing carbapenem-resistant *Klebsiella pneumoniae* in a US pediatric hospital. *J Pediatr Infect Dis.* 4, 330–338

96 Mammina, C. *et al.* (2012) Ongoing spread of colistin-resistant *Klebsiella*

pneumoniae in different wards of an acute general hospital, Italy, June to December 2011. *Euro Surveill.* 17, 1–6

97 Cubero, M. *et al.* (2015) Carbapenem-resistant and carbapenem-susceptible isogenic isolates of *Klebsiella pneumoniae* ST101 causing infection in a tertiary hospital. *BMC Microbiol.* 15, 177

98 Hrabák, J. *et al.* (2009) International clones of *Klebsiella pneumoniae* and *Escherichia coli* with extended-spectrum β-lactamases in a Czech hospital. *J Clin Microbiol.* 47, 3353–3357

99 Österblad, M. *et al.* (2012) Carbapenemase-producing enterobacteriaceae in Finland: The first years (2008-11). *J Antimicrob Chemother.* 67, 2860–2864

100 Marcade, G. *et al.* (2012) The emergence of multidrug-resistant *Klebsiella pneumoniae* of international clones ST13, ST16, ST35, ST48 and ST101 in a teaching hospital in the Paris region. *Epidemiol Infect.* DOI: 10.1017/S0950268812002099

101 Bachman, M.A. *et al.* (2015) Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung infection. *MBio* 6, e00775

102 Bruchmann, S. *et al.* (2015) Deep transcriptome profiling of clinical *Klebsiella pneumoniae* isolates reveals strain and sequence type-specific adaptation. *Env. Microbiol.* 17, 4690–4710

103 Larsen, M. V *et al.* (2012) Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol.* 50, 1355–1361

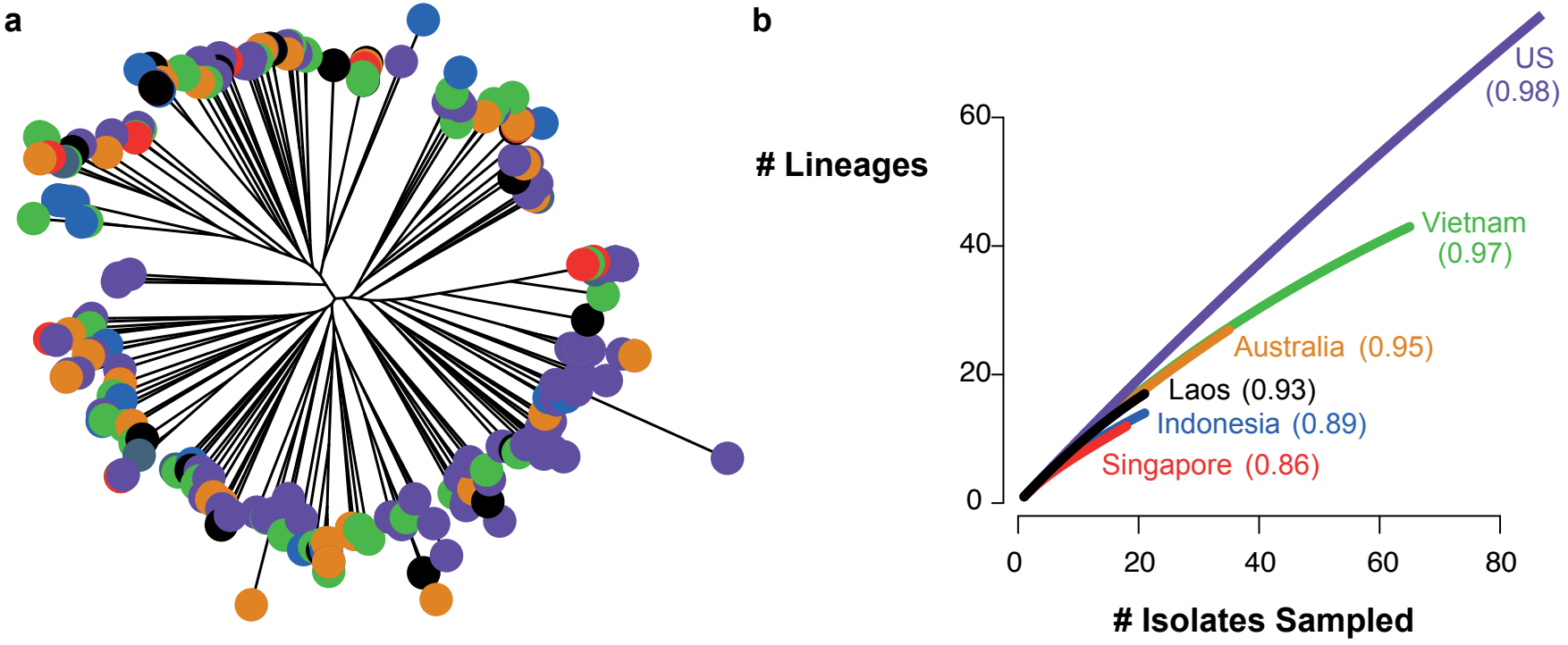
104 Joensen, K.G. *et al.* (2014) Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol.* 52, 1501–1510

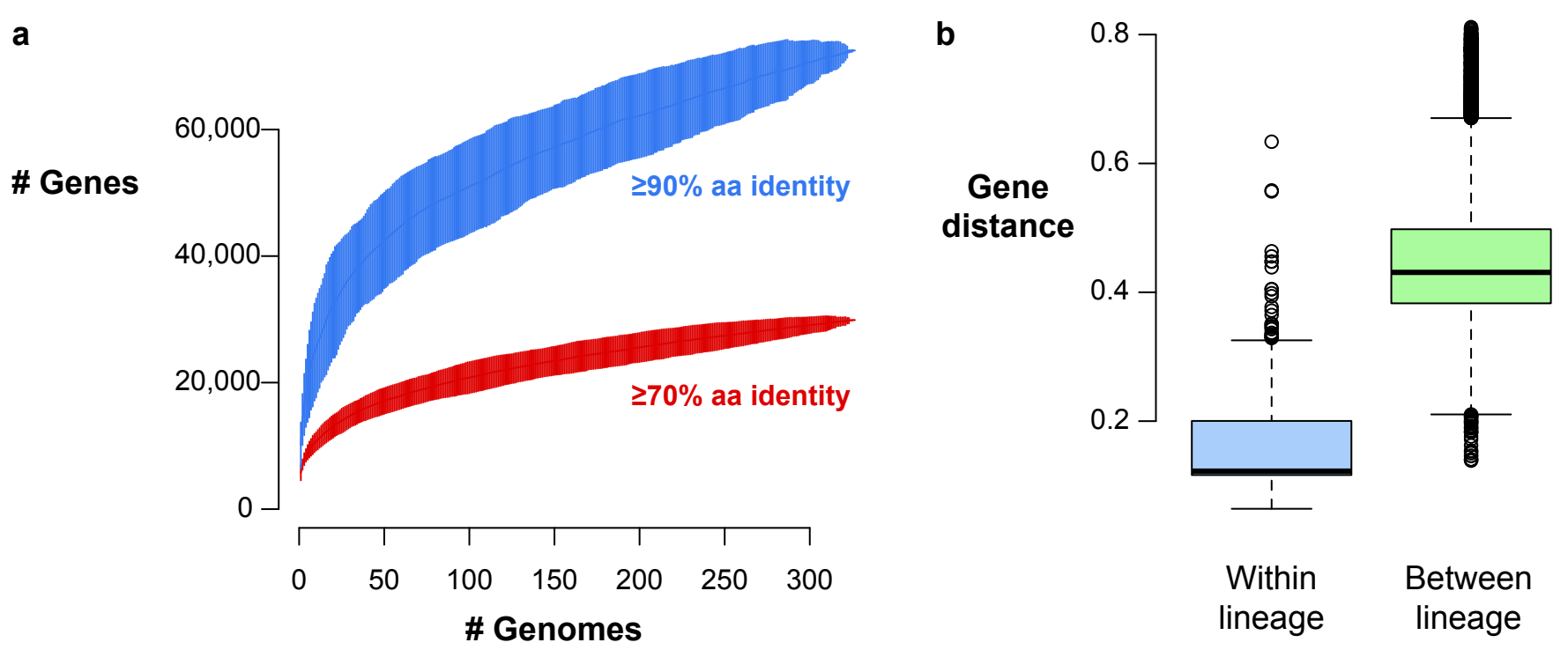
105 Zankari, E. *et al.* (2012) Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother.* 67, 2640–2644

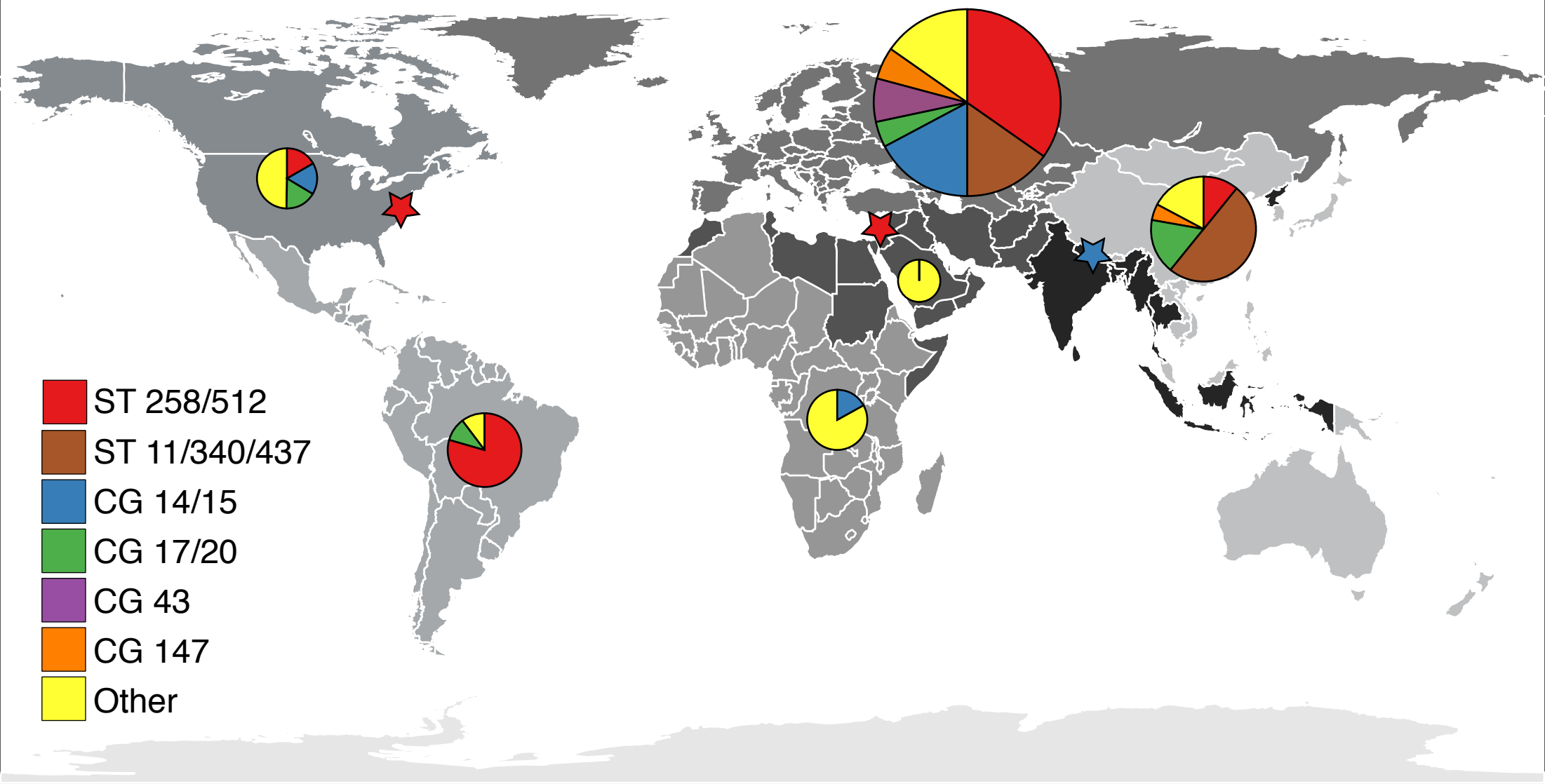
106 Carattoli, A. *et al.* (2014) PlasmidFinder and pMLST: *in silico* detection and typing of plasmids. *Antimicrob Agents Chemother.* 58, 3895–3903

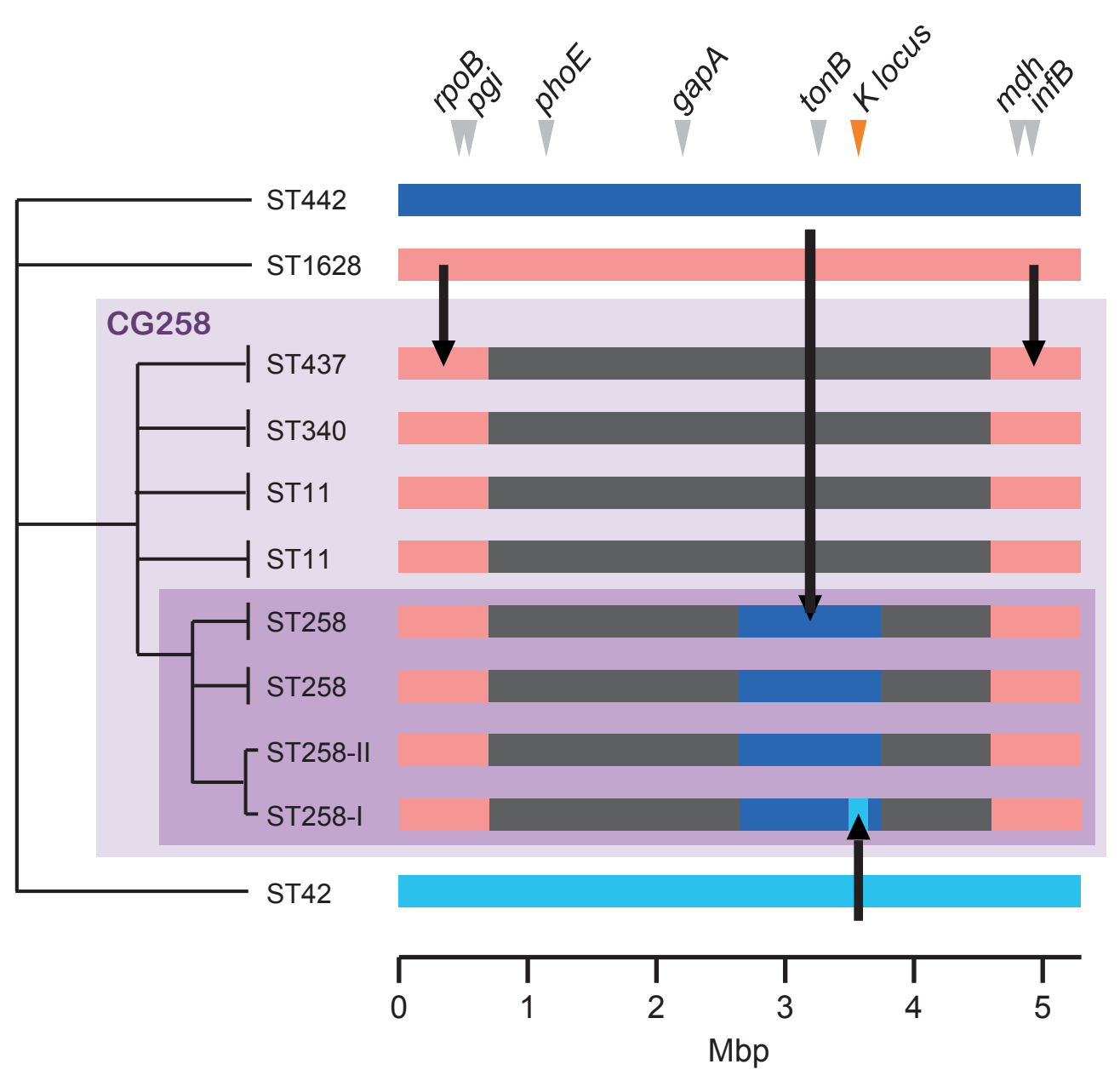
107 Inouye, M. *et al.* (2014) SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 6, 90

108 Hawkey, J. *et al.* (2015) ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. *BMC Genomics* 16, 667


a
b
Lineages
60
40
20
0
0
20
40
60
80
Isolates Sampled
US
(0.98)
Vietnam
(0.97)
Australia (0.95)
Laos (0.93)
Indonesia (0.89)
Singapore (0.86)


a
Genes
60,000
40,000
20,000
0
≥90% aa identity
≥70% aa identity
0
50
100
150
200
250
300
Genomes
b
Gene distance
0.8
0.6
0.4
0.2
Within lineage
Between lineage


ST 258/512
ST 11/340/437
CG 14/15
CG 17/20
CG 43
CG 147
Other


rpoB
pgi
phoE
gapA
tonB
K locus
mdh
infB
ST442
ST1628
CG258
ST437
ST340
ST11
ST11
ST258
ST258
ST258-II
ST258-I
ST42
0
1
2
3
4
5
Mbp