Population genomics of *Klebsiella pneumoniae*

Kelly L. Wyres[1], *Margaret M. C. Lam*[1] *and Kathryn E. Holt*[1,2]✉

Abstract | *Klebsiella pneumoniae* is a common cause of antimicrobial-resistant opportunistic infections in hospitalized patients. The species is naturally resistant to penicillins, and members of the population often carry acquired resistance to multiple antimicrobials. However, knowledge of *K. pneumoniae* ecology, population structure or pathogenicity is relatively limited. Over the past decade, *K. pneumoniae* has emerged as a major clinical and public health threat owing to increasing prevalence of healthcare-associated infections caused by multidrug-resistant strains producing extended-spectrum β-lactamases and/or carbapenemases. A parallel phenomenon of severe community-acquired infections caused by 'hypervirulent' *K. pneumoniae* has also emerged, associated with strains expressing acquired virulence factors. These distinct clinical concerns have stimulated renewed interest in *K. pneumoniae* research and particularly the application of genomics. In this Review, we discuss how genomics approaches have advanced our understanding of *K. pneumoniae* taxonomy, ecology and evolution as well as the diversity and distribution of clinically relevant determinants of pathogenicity and antimicrobial resistance. A deeper understanding of *K. pneumoniae* population structure and diversity will be important for the proper design and interpretation of experimental studies, for interpreting clinical and public health surveillance data and for the design and implementation of novel control strategies against this important pathogen.

Disability-adjusted life years
A measure of disease burden estimated as the number of years lost to ill-health, disability and/or death.

Problem clones
Klebsiella pneumoniae clones that are over-represented among human infection isolates.

[1]*Department of Infectious Diseases, Monash University, Melbourne, Victoria, Australia.*

[2]*Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, UK.*

✉*e-mail: kathryn.holt@monash.edu*

https://doi.org/10.1038/s41579-019-0315-1

Klebsiella pneumoniae belongs to the family Enterobacteriaceae, which includes the well-known genera *Salmonella* and *Escherichia*[1]. *K. pneumoniae* has long been recognized as an agent of disease (first described as a cause of pneumonia by Carl Friedländer in 1882), and remains among the world's most common nosocomial pathogens[2]. It is also a major cause of neonatal sepsis, ranking in the top three causative agents in most settings[3,4]. The World Health Organization recognizes extended-spectrum β-lactam (ESBL)-producing and carbapenem-resistant *K. pneumoniae* (CRKp) as a critical public health threat[5]. In Europe alone, such strains reportedly account for >90,000 infections, >7,000 deaths annually and 25% of the total disability-adjusted life years lost to multidrug-resistant (MDR) bacterial infections[6]. Precise burden estimates are lacking in other regions but MDR rates are trending upwards globally; for example, >75% of *K. pneumoniae* bloodstream infections in Malawi are now MDR[7].

In addition to its importance as a nosocomial pathogen, *K. pneumoniae* can be found living in a wide range of host-associated and environmental niches[8] and exhibits extensive phenotypic and genetic diversity. However, relatively little is known about how this genetic diversity is structured within the bacterial population (that is, the species population structure) and how this relates to the organism's ecology or its capacity to cause different types of disease. For instance, although a virulence plasmid has been described, it is absent from the vast majority of isolates from clinical infections. High-throughput genomic analyses allow us to interrogate and compare the entire genetic complement of hundreds or thousands of individual *K. pneumoniae* and have revealed key insights into this pathogen's population structure that can help us to better understand how it evolves, spreads and causes disease. For example, current data suggest that genes associated with antimicrobial resistance (AMR) and virulence are each concentrated in distinct subpopulations of *K. pneumoniae*[9,10]. However, there are signs that this structuring may be breaking down, resulting in potentially dangerous strains that are both highly pathogenic and resistant to all or most available antibiotics[11,12].

In this Review, we explore recent advances in *K. pneumoniae* population genomics, focusing on aspects related to ecology and the epidemiology of human infections. We discuss the overall population structure of the species and the epidemiology of problem clones, as well as the diversity and distribution of key AMR and virulence determinants. We highlight the importance of the

emerging *K. pneumoniae* population genomic framework and how it can be used to inform and interpret the surveillance and experimental studies that are sorely needed to advance our understanding and management of this important pathogen.

Taxonomy

Whole-genome sequencing (WGS) has clarified that a substantial proportion of isolates identified as *K. pneumoniae* by biochemical or proteomics assays in clinical and research laboratories actually belong to closely related species that share 95–96% average nucleotide identity with *K. pneumoniae*[9,13–15]. This group has no formal taxonomic designation but we refer to it here as the *K. pneumoniae* species complex (KpSC) (BOX 1), noting that its members share only 90% average nucleotide identity with other *Klebsiella* species (FIG. 1). Among the KpSC, *K. pneumoniae sensu stricto* is highly prevalent in clinical collections, typically comprising ~85% of the isolates identified as *K. pneumoniae*[9,16–19]. The remainder of this Review therefore focuses on *K. pneumoniae* as this is the species of clearest clinical importance and on which the majority of available data are concentrated.

Ecology and lifestyle

K. pneumoniae can survive in a multitude of ecological niches, both free-living and host-associated. These niches include soil, water, a range of plant species, insects, birds, reptiles and many different mammals in which this bacterium can be either a commensal organism or a potential pathogen[20]. However, there has been a lack of large-scale systematic sampling efforts, which limits our knowledge about the prevalence, abundance and features of *K. pneumoniae* in different niches. Moreover, most studies have relied on standard identification techniques without sequence data and have not distinguished *K. pneumoniae* from other KpSC members. Nevertheless, culture-based estimates suggest that *K. pneumoniae* (and/or the broader KpSC) is a common colonizer of the mammalian intestinal tract (39% and 44% prevalence in dogs[21] and dairy cattle[22], respectively), whereas 16S rRNA sequence-based estimates indicate frequent carriage by birds and various insect species[20].

From commensal to pathogen

The interaction between *K. pneumoniae* and human hosts is complex and changeable, and *K. pneumoniae* can play the role of either commensal, opportunistic pathogen or pathogen. Commensal colonization in the gut and respiratory tract is common but prevalence estimates vary by age group, geographical location and recent health-care contact. Studies from the USA and Australia have estimated gut colonization at ~4–6% prevalence in the community[14,23] but up to ~25% among individuals with recent health-care exposure in the USA, Australia and England[14,24,25]. Notably, higher community carriage rates for KpSC (between 18 and 87%) have been estimated for healthy adults in Korea, Japan, Singapore, Taiwan and Malaysia[26,27]. Duration of gut colonization is not well understood but can exceed 12 months[28,29].

The majority of *K. pneumoniae* infections globally are opportunistic health-care-associated infections (HAIs), sometimes referred to as 'classical' *K. pneumoniae* infections. The most common manifestations are pneumonia, urinary tract and wound infections, any of which can progress to bacteraemia[30]. Most at risk are vulnerable patient groups such as neonates, the elderly, those with inserted medical devices and the immunocompromised, in which infections are thought to result from overgrowth and lack of immunological control of commensal *K. pneumoniae* strains. As such, intestinal carriage is a key risk factor for *K. pneumoniae* HAI, associated with a fourfold increased risk of infection among intensive care and oncology patients[14,24]. Further supporting this, genomic comparisons indicate that gut-colonizing strains are the most common source of *K. pneumoniae* infections in these settings[14,24], and an elevated *K. pneumoniae* load in faecal samples has been associated with an increased risk of bacteraemia caused by CRKp in a long-term care facility[31]. Of greatest clinical concern is the rise of MDR and particularly CRKp: a recent meta-analysis estimated mortality associated with *K. pneumoniae* HAI to be 42% for CRKp, compared with 21% for carbapenem-susceptible strains[32]. Limited antimicrobial therapy options are available for MDR and CRKp infections, leading to a resurgence in the use of colistin (to which resistance is also now rising) and a renewed interest in vaccines and other preventatives against *K. pneumoniae*[33,34].

Outside the hospital setting, *K. pneumoniae* can act as a 'true' pathogen — that is, it can cause severe community-acquired infections (CAIs), which are not considered

Box 1 | *Klebsiella pneumoniae* species complex

Members of the *Klebsiella pneumoniae* species complex (KpSC) were first distinguished on the basis of *gyrA* sequences and designated as phylogroups of *K. pneumoniae*[148]. Whole-genome sequencing (WGS) later confirmed genome-wide average nucleotide identity ≥3%, sufficient to designate new species[9,40], and has facilitated the identification of additional member species (FIG. 1):

- *Klebsiella pneumoniae* (Kp1)
- *Klebsiella quasipneumoniae* subsp. *quasipneumoniae* (Kp2)[149]
- *Klebsiella quasipneumoniae* subsp. *similipneumoniae* (Kp4)[149]
- *Klebsiella variicola* subsp. *variicola* (Kp3)[150]
- *Klebsiella variicola* subsp. *tropica* (Kp5)[74]
- *Klebsiella* quasivariicola (Kp6)[151]
- *Klebsiella africana* (Kp7)[74]

The niche specificity of these species is not yet well understood, although there is clear evidence that *K. variicola* is plant-associated and typically carries a nitrogen-fixing operon (*nif*) and cellulases that are lacking from the other species[9,150,152]. All species have been isolated from the human gut, and all but *K. variicola* subsp. *tropica* have been isolated from human infections[74]. *K. variicola* and *K. quasipneumoniae* are relatively common agents of nosocomial infections (10–20% as frequent as *K. pneumoniae*[9,16,153]), can acquire antimicrobial resistance genes and plasmids from *K. pneumoniae*[154,155] and have been reported to cause nosocomial outbreaks[154,156]. Both species have been reported as the cause of community-acquired liver abscess[156,157] in humans and have also been isolated from animals[158].

Many species designations in culture collections and sequence databases require updating to reflect the new WGS-informed taxonomy[159]. Hence, species identification methods that rely on comparison with reference strains or genomes should be treated carefully, including sequence-based classifiers or mass spectrometry (MALDI-TOF) platforms commonly used in clinical laboratories[13–15]. These methods can be improved using custom databases populated by isolates whose species have been accurately defined using WGS[15]. The *K. pneumoniae* multilocus sequence typing (MLST) scheme covers the entire species complex and can be used to accurately assign isolates to the correct species[40].

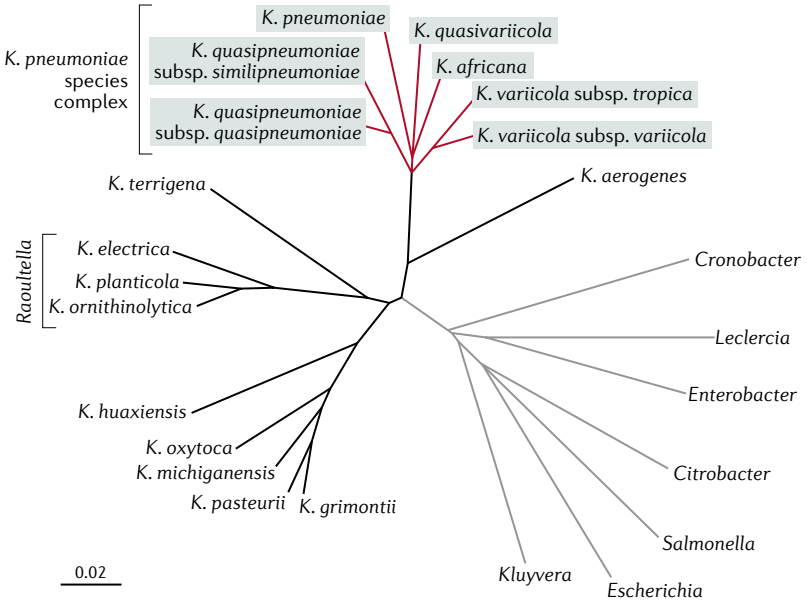


Fig. 1 | **Taxonomic position of *Klebsiella pneumoniae*.** Whole-genome-based tree showing the phylogenetic relationships between *K. pneumoniae*, its close relatives in the *K. pneumoniae* species complex (red branches), other select members of the *Klebsiella* genus (black branches) and family Enterobacteriaceae (grey branches). The tree was inferred from mash distances of representative whole genome sequences (Supplementary methods and Supplementary Table 1). Scale bar is the estimated average nucleotide divergence. Note *Klebsiella terrigena*, *Klebsiella planticola* and *Klebsiella ornithinolytica* have been assigned the genus name *Raoultella* based on *gyrB* sequences[169], but this is debated as it renders *Klebsiella* non-monophyletic.

Endophthalmitis
Inflammation of the eye, usually caused by bacterial or fungal infection.

Necrotizing fasciitis
An infection resulting in rapid death of the skin and soft tissues.

Hypervirulent
A term used to describe a clinical phenomenon of severe community-acquired *Klebsiella pneumoniae* disease, which is typically associated with the presence of a combination of multiple acquired virulence loci.

Core genes
Genes that are present in all members of a given species (or nearly all, typically ≥95%).

Accessory genes
Genes that are present in some members of a given species but not all (typically <95%).

Core-genome multilocus sequence typing
(cgMLST). A classification scheme and nomenclature based on nucleotide sequence variation in core genes (usually hundreds of genes).

opportunistic, in otherwise healthy patients who do not share the risk factors for HAIs[35]. Common CAIs include endophthalmitis, pneumonia, necrotizing fasciitis, non-hepatic abscess, meningitis and pyogenic liver abscess in the absence of biliary tract disease[36]. The ability to cause infections in unusual and/or multiple sites is considered characteristic of hypervirulent *K. pneumoniae* infections, which are often accompanied by bacteraemia and/or metastatic spread[36] (discussed below). Host risk factors for *K. pneumoniae* CAI include alcoholism (pneumonia)[37] and diabetes (pyogenic liver abscess)[38]. Numerous pathogen risk factors are also indicated (discussed below).

Population genomic framework

Genomic studies show that the *K. pneumoniae* population is diverse but highly structured, and this natural structure provides a useful framework for understanding the epidemiology and evolution of clinically relevant genetic variation. Such understanding is key for the effective design and interpretation of experimental studies aimed at elucidating mechanisms of AMR, pathogenicity and/or virulence in *K. pneumoniae*, and for the design of effective control strategies.

Population structure

Typical *K. pneumoniae* genomes are ~5–6 Mbp in size, encoding ~5,000–6,000 genes. Approximately 1,700 genes are conserved in all members of the species (core genes), whereas the remainder are variably present (accessory genes)[9,17]. The total pan-genome (the sum of all core and accessory genes)[39] is extremely diverse and likely exceeds 100,000 protein coding sequences[9]. The majority of accessory genes are rare in the population, that is, they are present in <10% of genomes[9]. Taxonomic and GC content analyses suggest that these genes are shared with a wide range of other bacterial species, most commonly other *Klebsiella* species followed by other Enterobacteriales but also including more distant orders[9].

Phylogenetic analyses based on 1,000–2,000 core or common chromosomal genes show that the *K. pneumoniae* population comprises hundreds of deep-branching lineages that differ from each other by ~0.5% nucleotide divergence[9,40]. These lineages correspond closely to the clonal groups (CGs) defined by core-genome multilocus sequence typing (cgMLST) as subsets of isolates that each share ≥594 of 694 cgMLST alleles with at least one other member of the group[40]. Whether defined on the basis of core-genome phylogeny (lineages) or cgMLST (CGs), the resulting groups are typically referred to as clones (FIG. 2a) and are identified and labelled based on the dominant seven-gene multilocus sequence type, allowing backwards comparison with the original seven-gene MLST scheme[41] (which covers the entire KpSC). Note that the seven-gene MLST scheme can also be used alone to define CGs, but sometimes fails to correctly distinguish groups whose recent ancestry is affected by chromosomal recombination[42,43].

K. pneumoniae clones can be distinguished from one another on the basis of accessory gene content[9]. This may be explained by clone-specific niche adaptation through horizontal gene transfer (HGT), so long as migration between niches occurs[44], and there is evidence that it does (discussed below). However, there is also ample evidence of between-clone HGT, largely driven by chromosomal recombination and plasmid-mediated conjugation[10,45], although phage-mediated transduction and integrative conjugative elements (ICEs) also play a role[46]. Homologous recombination between chromosomes is dominated by exchange of capsule biosynthesis loci[10] (a key pathogenicity determinant) and can result in the acquisition of regions of DNA exceeding 1 Mbp in length[46,47]. Many diverse plasmids have been sequenced from *K. pneumoniae*, ranging considerably in terms of length and incompatibility types, although $IncFII_K$ and $IncFIB_K$ are the most prevalent[9,48] (BOX 2). It seems that some *K. pneumoniae* strains may be particularly permissive for plasmid uptake and/or maintenance, resulting in plasmid loads that are typically greater than those reported for *Escherichia coli* and other Gram-negative ESKAPE pathogens[20]. For example, it is not uncommon for a *K. pneumoniae* isolate to carry between four and six different plasmids, with up to 10 reported[20,49]. Most complete *K. pneumoniae* genomes carry multiple prophages[46], and multiple distinct *K. pneumoniae* phages have been isolated and sequenced, including for potential use as therapeutic agents[34]. However, there have not yet been any systematic studies of prophage diversity within the population. CRISPR–Cas9 systems and restriction-modification systems are variably present in the *K. pneumoniae* population, however, how these relate to plasmid and phage diversity is not yet clear[10,50].

Clones
A generic term used to describe subpopulations of *Klebsiella pneumoniae* strains with a recent common ancestor identified through allelic variation in core genes, either by multilocus sequence typing or core-genome multilocus sequence typing (specifically called 'clonal groups' or 'CGs') or by core-genome phylogenetics (specifically called 'lineages').

Global problem clones

K. pneumoniae infections are caused by diverse clones that are widely geographically distributed. However, a subset of these contribute disproportionately to global disease burden, which we refer to as 'global problem clones' (FIG. 2).

Multidrug-resistant clones. MDR, which is defined as resistance to ≥3 antimicrobial classes in addition to ampicillin to which all *K. pneumoniae* infections are intrinsically resistant, has evolved many times in hundreds of distinct *K. pneumoniae* lineages (FIG. 2a,b). Some of these lineages emerge to cause localized problems, spreading within a single hospital or health-care network, for example, MDR clones sequence type 70 (ST70) and ST323 that caused outbreaks in Kilifi, Kenya and Melbourne, Australia, respectively[17,19]. Such events likely represent chance emergence in a given time and place, and the factors influencing their frequency, likelihood and duration of persistence are not known[51]. Many will remain localized problems, causing no or limited infections elsewhere, but a subset of the most highly resistant lineages (for example, those resistant to third-generation cephalosporins and/or carbapenems) have become global problems. These include the well-studied CG258 alongside CG15, CG20 (CG17), CG29, CG37, CG147, CG101 (CG43) and CG307 (FIG. 2c), which are not related to one another (FIG. 2a) but are each widely geographically distributed and common causes of MDR HAIs and/or outbreaks[48,52,53]. Here, we refer to these collectively as the 'global MDR clones', although there

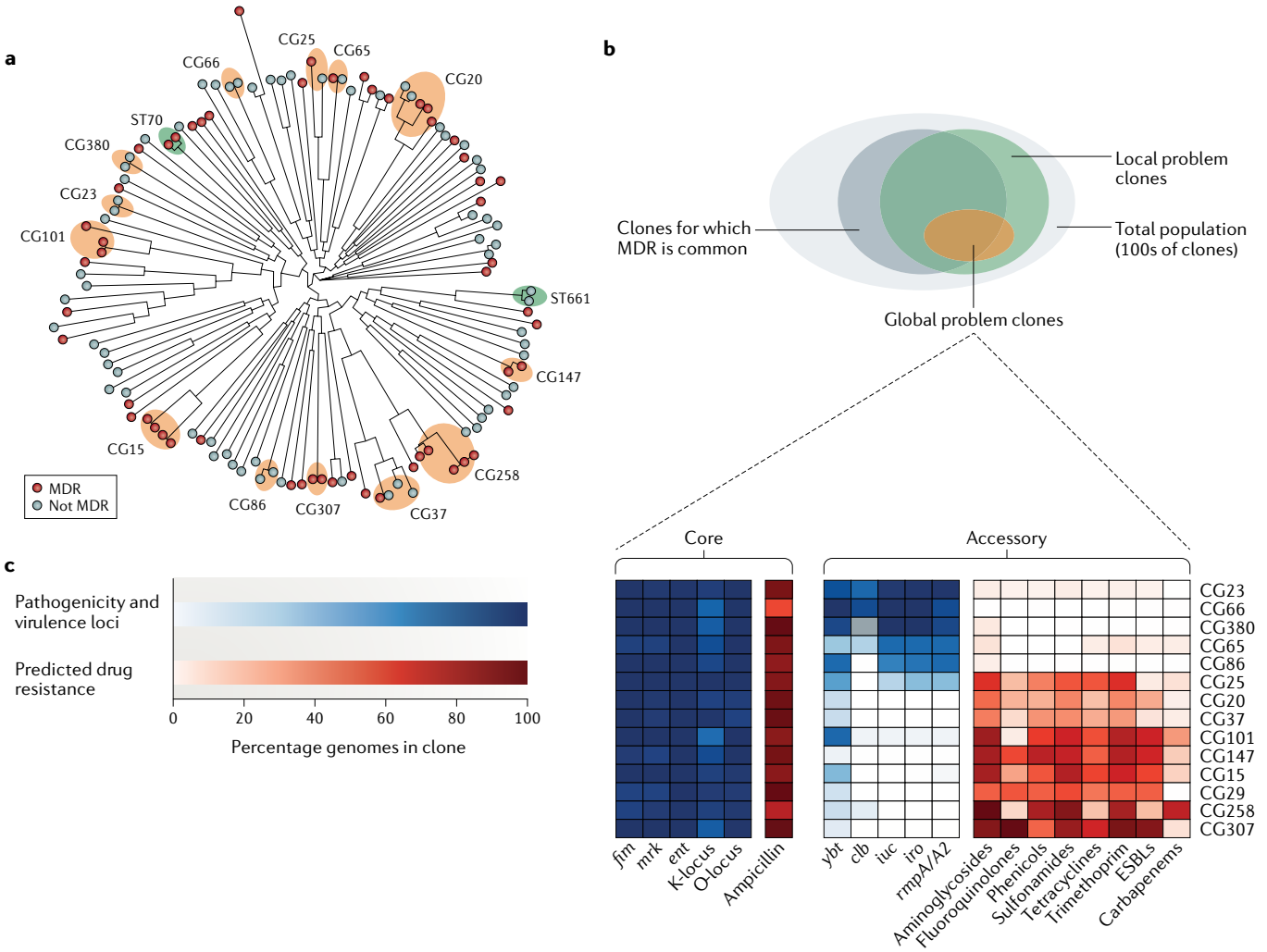


Fig. 2 | ***Klebsiella pneumoniae* population structure and global problem clones.** There are hundreds of distinct phylogenetic lineages or clones within the *K. pneumoniae* population; a subset of these are shown in the whole-genome phylogeny (part **a**), with tips coloured to indicate individual isolates that are multidrug resistant (MDR), as predicted from genome data. Global problem clones are highlighted in orange, including eight MDR clones (clonal group 15 (CG15), CG20, CG29, CG37, CG147, CG101, CG258 and CG307) and six hypervirulent clones (CG23, CG25, CG65, CG66, CG86 and CG380). Two examples of local problem clones are highlighted in green. The maximum-likelihood tree was inferred from a core-genome alignment constructed from 83 genomes randomly selected from a global diversity study[9] as well as 41 genomes representing problem clones (Supplementary Table 2 and Supplementary methods). Note that MDR clones may be identified only sporadically and not spread to become local or global problems; and some global problem clones are not MDR (part **b**). The prevalence of pathogenicity and virulence factors (blue) and drug resistance genes (red) among each of the global problem clones is shown (part **c**), based on published data from REF.[10] (geographically and temporally diverse set of 1,092 genomes from 28 clones, $n \geq 10$ genomes per clone) and REF.[53] ($n = 95$ CG307) (Supplementary Tables 3, 4 and Supplementary methods). *clb*, colibactin locus; *iro*, salmochelin locus; *iuc*, aerobactin locus; *rmpA/A2*, regulators of mucoid phenotype genes; ST, sequence type; *ybt*, yersiniabactin locus.

exist important geographical differences in their contributions to MDR infection burden (we refer readers to previous systematic reviews[48,52] and summarize the latest trends in FIG. 3). Importantly, although many studies do attribute the majority of third-generation cephalosporin-resistant and/or CRKp infections to a small number of clones (57% of CRKp infections in the recent Europe-wide study EuSCAPE were ST11, ST15, ST101 or ST258/ST512)[54], in many regions the burden attributed to sporadic MDR strains or local problem clones remains substantial (for example, >33% of strains in a recent national survey of CRKp in the UK[55]).

Hypervirulent clones. In contrast to MDR infections, hypervirulent *K. pneumoniae* infections in all regions are dominated by the same subset of lineages. By far the most common is CG23, followed by CG65 (including ST65 and ST375) and CG86, whereas CG25, CG66 and CG380 are implicated to a lesser extent[56–59]. Consistent with the high prevalence of hypervirulent disease in the Asia-Pacific rim, CG23 was recently identified among the top two most common clones associated with *K. pneumoniae* bloodstream infections in China, Vietnam and Laos, accounting for >10% of isolates[60,61]. Limited data are available about the prevalence of ST23 outside the Asia-Pacific rim because the majority of studies focus solely on MDR strains whereas ST23 isolates are usually drug susceptible[10,46] (FIG. 2c). However, the available evidence suggests that ST23 accounts for ≤2% of clinical isolates in other regions[17,61–63], despite circulating among humans since its emergence >100 years ago[46], far longer than the most well-known MDR clones (ST258 and ST307 both emerged in the mid-1990s)[47,53].

Clones found in human versus other sources

There are limited data on lineage distributions in non-human niches, however, there is some overlap between lineages isolated from clinical and other sources. Global problem clones have been isolated from a range of animals, for example, ST11 from poultry[64], ST15 from companion animals[65], ST23 from non-human primates and horses[46,66] and ST25 from pigs[67]. Despite the ubiquity of *K. pneumoniae* outside the human host, and growing interest to identify the routes of transmission into the human population, just four studies have so far reported WGS surveys of *K. pneumoniae* collected using non-selective isolation from non-human sources and compared the data with contemporaneous clinical isolate samples. These studies uncovered substantial diversity among clinical isolates (Simpson diversity ≥0.96) and also among non-human samples (Simpson diversity ≥0.94): 45 sequence types (STs) in 49 isolates from dairy farms in New York state, USA[9]; 16 STs from 29 canal or farm isolates in Thailand[68]; 24 STs from 28 wastewater isolates and four STs in cattle or turkey isolates from six farms in the North of England[25]; and 39 STs from 44 retail meat isolates in Arizona, USA[69]. In each study, the proportion of STs from *K. pneumoniae* clinical isolates that were also detected in local non-human samples was low (5%, 15%, 8% and 15%, respectively). However, given the small sample sizes, high diversity in each source and the lack of specific spatiotemporal models for expected transmission between sources to inform sampling efforts, this overlap cannot be dismissed as insignificant.

Notably, these overlap fractions are similar in scale to the proportion of STs detected in multiple patients,

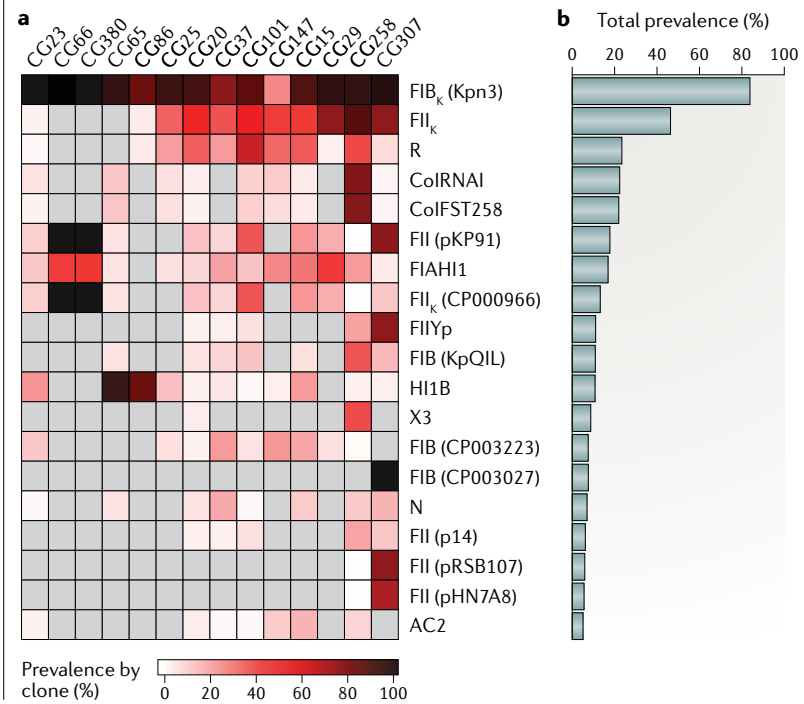
Box 2 | **Plasmid diversity in *Klebsiella pneumoniae***

Plasmids are important vehicles for the transfer of antimicrobial resistance (AMR), virulence and other accessory genes between bacterial cells. In *K. pneumoniae*, the majority of horizontally acquired AMR genes are carried on large conjugative (self-transmissible) plasmids belonging to a small number of incompatibility groups, which have been recently reviewed[48] (IncFII, IncN, IncR and IncX3), although small (mobilizable but not self-transmissible) plasmids can also harbour AMR genes in *K. pneumoniae*[160].

Most available data on plasmid diversity and distribution in the *K. pneumoniae* population come from studies utilizing short-read whole-genome sequencing to survey populations, yielding information on replicon markers[161] or *mob* (relaxase) gene variants[162,163] rather than whole plasmids[164] (although this is changing with increased use of long-read sequencing)[165]. Using these techniques, plasmid load and diversity have been shown to be significantly higher in multidrug-resistant clones than hypervirulent clones or other clonal groups[10].

The figure shows a heatmap of the distribution of 19 common plasmid replicon markers (*y* axis) across 14 *K. pneumoniae* problem clones (part **a**) based on screening 1,187 genomes from 29 clones[10,53] against the PlasmidFinder database[161] using BLASTn. Bars show the prevalence of each replicon across the total genome collection (part **b**). Replicon markers shown are those with prevalence ≥5%, named as per the PlasmidFinder database. Strain-level data for the full set of 69 replicons detected are available in Supplementary Table 5. The FIB_K replicon is associated with both multidrug resistant and virulence plasmids, conjugative and non-conjugative, and is common in all clones. Also common are incompatibility type FII_K and R replicons (associated with large conjugative plasmids) and small (Col) plasmids, followed by other F plasmid variants (FII, FIA and FIB) and incompatibility types X3, N, HI1B and AC/2. A further 50 replicon markers were detected in this genome set (at <5% prevalence), with median total of three per genome (range 0–5; see Supplementary Table 5).

Notably, although there can be significant plasmid variation within a single clone[10,47,48], phylogenomic evolutionary analyses suggest that some problem clones have maintained specific plasmids throughout decades of clonal expansion, punctuated by occasional instances of rearrangement and/or gene deletions[46,166]. For example, sequence type 258 (ST258) has carried FIB_K bla_{KPC} plasmid pKpQIL[47] and ST307 has carried a FII_K $bla_{CTX\text{-}M\text{-}15}$ plasmid[53] since each emerged in the mid-1990s; and the FIB_K *K. pneumoniae* virulence plasmid (KpVP-1) has been present in ST23 dating back to the 1870s[46].



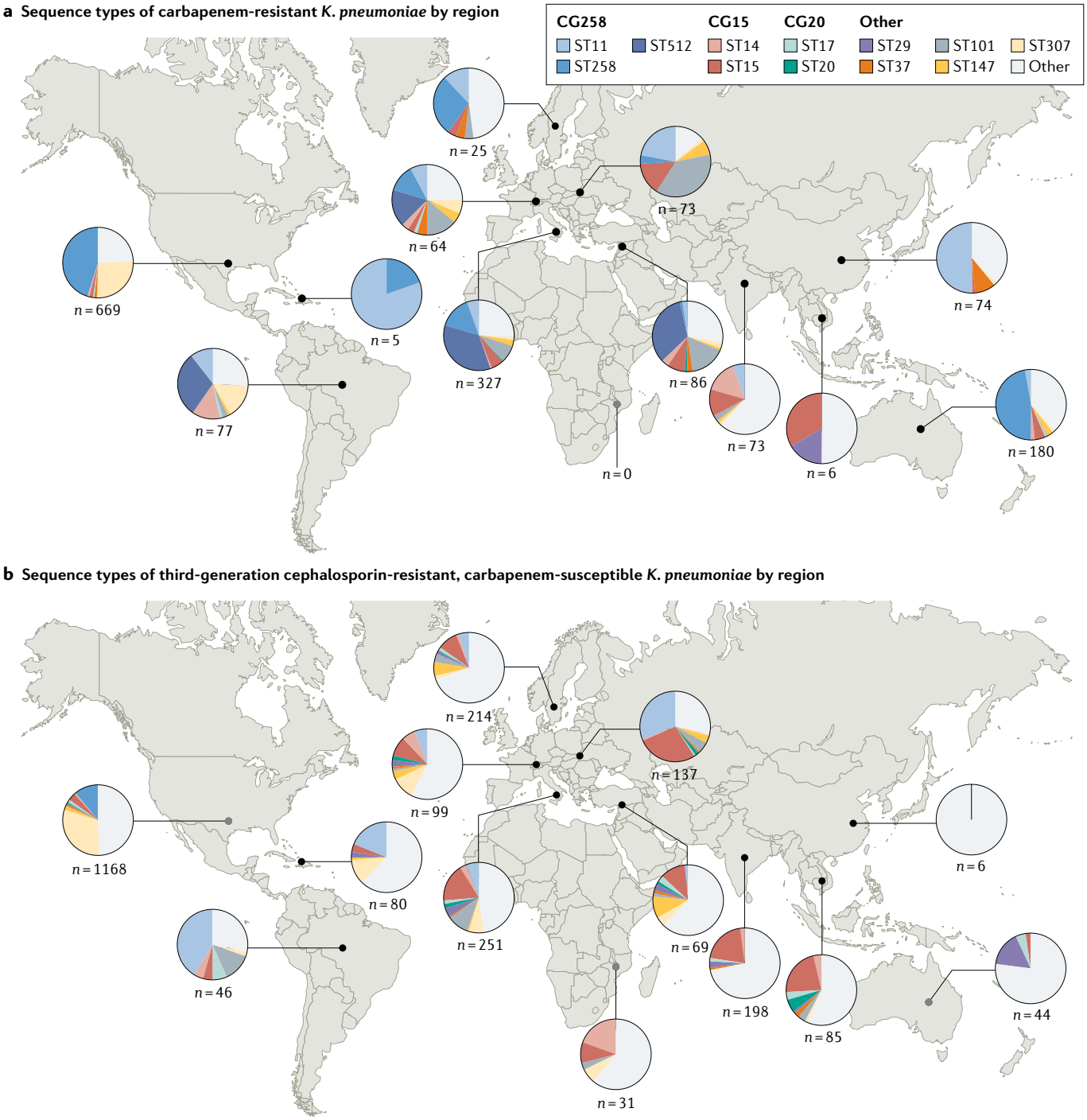


Fig. 3 | **Geographical distribution of *Klebsiella pneumoniae* clones harbouring resistance to carbapenems and third-generation cephalosporins.** Regional frequencies of sequence types (STs) of carbapenem-resistant (part **a**) and third-generation cephalosporin-resistant, carbapenem-susceptible *K. pneumoniae* (part **b**) are shown. Data are derived from recent studies in which isolates were selected on the basis of neither ST nor local transmission events or outbreaks[14,16,19,54,61,62,170–175] and are grouped by region as defined by the United Nations Statistics Division (Supplementary Table 6 and Supplementary methods). Black and grey location markers indicate data from multiple and single health-care institutions, respectively. Although ST258 and its derivative ST512 are dominant carbapenem-resistant *K. pneumoniae* (CRKp) in the Americas and southern Europe, they are comparatively rare in other regions of the world. ST11 (a diverse clade from which ST258 was derived by recombination[45,176]) represents ~12% of CRKp across Europe and is much more broadly distributed than ST258 or ST512. ST11 is also the single dominant cause of CRKp infections in China. Clonal group 307 (CG307) has recently begun to displace ST258 and ST512 among CRKp in the Americas and southern Europe and has emerged as the dominant CRKp lineage in South Africa[177], but outside these regions it is most commonly carbapenem-susceptible and third-generation cephalosporin-resistant (due to the $bla_{CTX\text{-}M\text{-}15}$ plasmid). CG15 (ST15, ST14 and close relatives) is among the most common third-generation cephalosporin-resistant clones in diverse geographies and associated with numerous reports of disseminated carbapenem-resistant variants[18,54,178].

Integrative conjugative elements
(ICEs). Mobile pieces of DNA that encode the machinery required for their own integration and excision from the bacterial host chromosome and transfer between bacterial cells.

ESKAPE pathogens
The six most common causes of multidrug-resistant health-care-associated infection defined by the Infectious Diseases Society of America: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species.

Simpson diversity
A measure of diversity that considers the total number of distinct entities (species, sequence types and so on) as well as their relative abundance.

Minimum inhibitory concentration
(MIC). The lowest concentration of a compound (usually an antibiotic) that is able to inhibit growth of a given bacterial isolate.

when comparing clinical isolates from a given setting and time period (in studies where infections were sampled randomly and not directed towards specific AMR phenotypes or suspected outbreak investigations). In the four studies outlined above, 17%, 40%, 0% and 27% of STs, respectively, were isolated from >1 patient[9,25,68,69]. Similarly, WGS studies of all *K. pneumoniae* clinical isolates from a given intensive care or geriatrics ward over a 1-year period found that 8% or 20% of STs were isolated from >1 patient, respectively[14,19]. We conclude that simple niche-sampling studies are unlikely to directly capture transmission chains occurring between niches, rather, this would require frequent, intensive and contemporaneous sampling of isolates from self-contained ecosystems. For example, there is some evidence of strain sharing between humans and cohabiting companion animals[21]. Although several studies have shown that *K. pneumoniae* and other KpSC members can be isolated from retail meats, seafood, vegetables and ready-to-eat processed foods[69–72], the ultimate source of KpSC in foods remains unclear; however, there is evidence that contamination during slaughter or harvest and/or processing plays a role[71,73].

Antimicrobial resistance

Resistance to all drug classes used to treat *K. pneumoniae* has been observed clinically. The specific genetic mechanisms of resistance in *K. pneumoniae* have been recently reviewed elsewhere, and the vast majority of cases are associated with horizontally acquired AMR genes[48,52]. Hence, we focus here on the diversity and distribution of these genes in the *K. pneumoniae* population, following a brief discussion of the role of core genes in AMR.

Core genes and drug resistance

K. pneumoniae is intrinsically resistant to ampicillin (that is, the aminopenicillin used to treat Gram-negative infections) through production of the class A β-lactamase enzyme SHV, encoded in the *K. pneumoniae* chromosome by the core gene bla_{SHV}[9] (FIG. 2c). Orthologues of this class A β-lactamase are a conserved feature of KpSC chromosomes (designated LEN in *Klebsiella variicola* and *Klebsiella quasivariicola*, and OKP in other species[74]). The bla_{SHV} gene has been captured multiple times from the *K. pneumoniae* chromosome by the transposase IS*26* (REF.[75]), forming a mobile genetic element (MGE) that facilitated dissemination to other species through plasmid transfer[76]. Notably, the resulting mobile variants of bla_{SHV} often carry point mutations that result in ESBL activity (conferring resistance to third-generation cephalosporins), and, occasionally, even carbapenemase activity[76] (conferring resistance to carbapenems), and are under stronger promoters than the chromosomal forms[77]. Consequently, some *K. pneumoniae* now carry two copies of bla_{SHV}: a chromosomal core gene variant and an acquired (typically plasmid borne) ESBL variant under control of a strong IS*26* promoter. This can create difficulties in accurately identifying bla_{SHV} alleles present in *K. pneumoniae* based on short-read WGS data, which is necessary to accurately predict the spectrum of β-lactamase activity and associated drug resistance.

The core loci *fosA* (glutathione *S*-transferase) and *oqxAB* (efflux pump) have also been captured from the *K. pneumoniae* chromosome by MGEs and disseminated to other species. At wild-type levels of gene expression, *fosA* and *oqxAB* confer on *K. pneumoniae* reduced susceptibility to fosfomycin (minimum inhibitory concentration (MIC) 16–32 mg/l) and quinolones (ciprofloxacin MIC 0.008–0.25 mg/l), respectively, which do not meet the recognized break points for clinically relevant resistance. However, the mobile forms of these genes are often expressed at higher levels, which can confer clinically relevant resistance in *E. coli* or *K. pneumoniae*[78,79].

Numerous other core genes contribute to antimicrobial susceptibility in *K. pneumoniae*, particularly those associated with lipopolysaccharide (LPS) production, efflux or membrane permeability[80,81]. Mutations in some of these are known to increase resistance to clinically important antimicrobials (FIG. 4c). Changes in expression and/or activity of the efflux pumps OqxAB and AcrAB have been associated with resistance to multiple antibiotics, such as fluoroquinolones, nitrofurantoin, tigecycline, chloramphenicol and carbapenems[79,82–84]. Various genetic mechanisms for this have been described, including through the regulators *ramA* (affecting both efflux pumps), *soxS* (affecting AcrAB) and *rarA* (affecting OqxAB)[79]. Specific mutations in the outer membrane porin OmpK36, particularly when coupled with loss of OmpK35, contribute substantially to the carbapenem-resistance problem in *K. pneumoniae*[85,86]. These porin defects are most common in carbapenemase-associated clones[85] and can increase the MIC above the level conferred by acquired carbapenemase genes alone. In some cases (such as β-lactamase OXA-48), porin defects are necessary to reach or exceed the break points established for clinically relevant carbapenem resistance[86,87]. CRKp infections are often treated with colistin[33], however, resistance to this last-line drug can emerge during treatment as a consequence of modification of lipid A, most commonly through inactivation of the *mgrB* gene or point mutations in the two-component regulator systems *phoPQ*, *pmrAB* or *crrAB*[88].

Acquired resistance genes

Most resistance observed in *K. pneumoniae* is associated with horizontally acquired accessory AMR genes rather than mutations in chromosomal genes[48] (FIG. 2c). These are most often plasmid borne[48], but can also be integrated into the chromosome[53]. Counting acquired AMR genes in *K. pneumoniae* isolates from WGS data is not straightforward unless complete genome sequences are available, as this requires differentiating the chromosomal genes bla_{SHV}, *fosA* and *oqxAB* from MGE-encoded forms as detailed above. However, even excluding these genes entirely, hundreds of distinct acquired AMR alleles have been detected in *K. pneumoniae*: 52 were detected in a WGS screen of 195 unselected KpSC blood culture isolates sampled over 10 years at a Kenyan hospital[17]; 78 were detected in a WGS screen of 328 unselected globally representative KpSC genomes[9]; 111 were detected in a WGS screen of 1,092 *K. pneumoniae* isolates from 28 clonal groups[10] (FIG. 2c); and 410 were detected in a screen of all

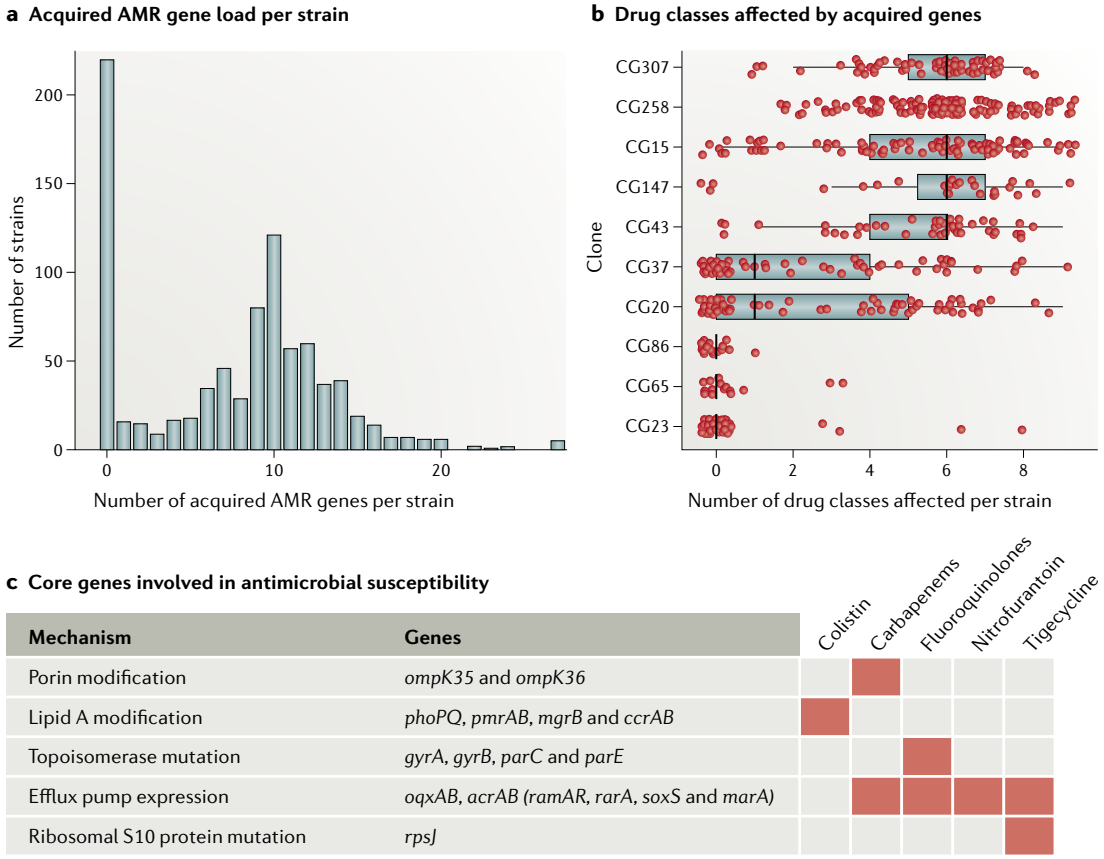


Fig. 4 | **Antimicrobial resistance in *Klebsiella pneumoniae*.** The majority of resistance to clinically relevant antibiotics is associated with horizontally acquired antimicrobial resistance (AMR) genes. These are non-randomly distributed in the population, and the number of acquired AMR genes per strain follows a bimodal distribution (part **a**), with most strains carrying either zero acquired AMR genes (that is, they are susceptible to all drugs except ampicillin, due to the core β-lactamase gene bla_{SHV}) or ~10 acquired AMR genes (encoding resistance to multiple drug classes). The number of drug classes per strain that are affected by horizontally acquired AMR genes (red data points) are summarized by box plots for each of 10 global problem clones (part **b**). The data for panels **a** and **b** are from REF.[10] and REF.[53] (Supplementary Table 3 and Supplementary methods). Acquired resistance mechanisms that are conferred by mutations in core chromosomal genes that are present in all *K. pneumoniae* (not by horizontally acquired genes) are indicated (part **c**). For these genes, wild-type alleles are associated with susceptibility whereas mutations are associated with resistance. Note that for all drugs, except nitrofurantoin, resistance in *K. pneumoniae* can also be conferred by horizontally acquired AMR genes. CG, clonal group.

198 complete genomes of *K. pneumoniae* available in the National Center for Biotechnology Information (NCBI) GenBank database in March 2018 (REF.[20]).

It has been noted among *K. pneumoniae* hospital infection isolates that resistance to a given drug class is often strongly correlated with resistance to other drug classes[89]. However, few studies have explicitly assessed the distribution of acquired AMR genes in the *K. pneumoniae* population. Doing so using WGS analysis of 1,187 isolates from 29 clones reveals a bimodal distribution, with most isolates either carrying multiple accessory AMR genes (mode = 10) or none (FIG. 4a). These accessory AMR genes are associated with acquired resistance to a wide range of drug classes (FIG. 4b), with bimodal peaks at either zero or six classes (in addition to the intrinsic ampicillin resistance). Clear differences in AMR gene load are evident between clones. The global MDR clones are characterized by consistently high numbers of accessory AMR genes encoding resistance to ≥6 drug classes, whereas the hypervirulent clones rarely carry any acquired resistance genes (FIGS 2c and 4b). A bimodal distribution is evident within CG20 and CG37, indicating MDR and non-MDR subpopulations within these clones (FIG. 4b). Similar patterns were recently reported for an independent sample of 228 *K. pneumoniae* isolates from the Caribbean that showed significant differences in acquired AMR gene load by clone, with a median of 5–10 genes in ST11, ST258, ST152, ST15, ST392, ST405, ST307 and ST39 and a median of one gene in ST23, ST86, ST1605 and ungrouped isolates[62].

Several contributing factors have been proposed for the uneven distribution of MDR within bacterial pathogen populations[89]. Among these factors, there is evidence in *K. pneumoniae* for genetic linkage between AMR genes encoded on the same MGE[19,48] and for lineage differences in recombination rates[10]. In *Streptococcus pneumoniae*, strain differences in human commensal carriage rates play a role[90]; this could potentially be relevant to *K. pneumoniae* but has yet to be explored. The genomic epidemiology of MDR in *K. pneumoniae*

Pathogenicity factors
Features encoded by loci that are present in all *Klebsiella pneumoniae* (although there may be important allelic variants) and required for 'classical' opportunistic infections.

K-locus
The capsule (K antigen) biosynthesis locus.

Rhinoscleromatis
A rare chronic infection characterized by granulomas (structure formed from a collection of white blood cells) in the upper airways.

O-loci
The outer lipopolysaccharide (O antigen) biosynthesis loci.

suggests that there may also be lineage-specific differences in the fitness costs associated with maintaining plasmids, even in the absence of selection for specific AMR phenotypes. Carriage of multiple plasmids each harbouring multiple AMR genes is not uncommon in *K. pneumoniae*[19,91]. WGS surveys of diverse isolate collections that are not pre-selected for AMR have noted that large complements of AMR genes can be detected in diverse and more rarely encountered lineages of *K. pneumoniae* or KpSC members, particularly among hospital isolates[9,55,92]. This suggests that although all lineages are exposed to AMR plasmids, some maintain these only transiently. By contrast, the global MDR clones show evidence of maintaining specific AMR genes and plasmids through decades of clonal expansion, diversification and widespread dissemination (BOX 2).

Detailed genomic investigations of *K. pneumoniae* at individual health-care facilities provide further evidence for variation in plasmid maintenance dynamics. Several WGS studies document AMR plasmid transmission between *K. pneumoniae* strains (and between *K. pneumoniae* and other species) in the hospital setting[19,49,92]. The detection of MDR plasmids in rare *K. pneumoniae* STs in hospital patients could thus be explained by occasional acquisition of plasmids from other Enterobacteriales in the gut, which are maintained in the hospital setting owing to selection from antimicrobial use but are readily lost upon withdrawal of therapy or transmission to another host. Consistent with this, hospital-based WGS studies tend to reveal an imbalanced population of strains carrying the target gene of interest (for example, bla_{KPC}), with one or two plasmid-bearing STs spreading between large numbers of patients (often global MDR clones such as CG258 or CG307) alongside multiple rare STs that acquire the same plasmid but do not spread to other patients, perhaps owing to reduced fitness[92,93]. There is some experimental evidence supporting clone-specific differences in plasmid maintenance costs in *K. pneumoniae*[94,95], but the mechanisms for this remain to be explored.

Resistance outside hospitals

Numerous studies have reported selective isolation of ESBL or carbapenemase-producing *K. pneumoniae* from animals and wastewater treatment plants. Notably, these tend to recover the same MDR clones isolated from clinical cases in neighbouring locations and are often attributed to contamination from humans rather than vice versa[96]. As noted above, there is a lack of reported attempts to non-selectively isolate *K. pneumoniae* from non-human sources, hence the frequency of AMR among non-human-associated populations and the flow of AMR genes and strains between niches remains cryptic.

Pathogenicity and virulence

Numerous genetic factors contribute to the ability of *K. pneumoniae* strains to cause disease in humans. The underlying mechanisms[97] and host immune evasion strategies[98] have been reviewed elsewhere. In this Review, we focus on current knowledge of the diversity and distribution of disease-related genes in the *K. pneumoniae* population.

Pathogenicity factors

All *K. pneumoniae* strains harbour a subset of core chromosomally encoded pathogenicity factors that form the basic requirements for establishing opportunistic infections in mammalian hosts. These include the core locus *ent* encoding biosynthesis of the siderophore enterobactin (Ent), the core *fim* and *mrk* loci encoding type 1 and type 3 fimbriae, respectively, as well as the variable capsular polysaccharide (K antigen) and LPS (O antigen) biosynthesis loci[9,40,99,100]. Ent (or an alternative acquired siderophore; see below) is required for growth in most niches[101]; the other factors mediate processes at the cell surface involved in the earlier stages of infection and/or evasion of host defence mechanisms[97,98].

Large-scale genomic comparisons have revealed substantial allelic and gene-content heterogeneity at the capsule (K antigen) and LPS (O antigen) biosynthesis loci, which comprise ~10% of the pan-genome[9]. The capsule is produced through a Wzy-dependent process for which the conserved machinery is encoded by genes that are present in most K-antigen biosynthesis loci (*wzi*, *wza*, *wzb*, *wzc*, *wzx* and *wzy*)[102,103]. However, the capsule-specific sugar synthesis machinery is encoded by a set of highly diverse genes that are variably present and frequently reassorted in the population. To date, >138 distinct combinations have been identified[61,99,100], each defining a distinct K-locus thought to encode a distinct capsule type, but only 77 of these have been distinguished by traditional serological typing[104]. Capsule types K1 and K2 are associated with invasive disease and enhanced pathogenicity in murine models[30] and are highly conserved in hypervirulent clones (K1 in CG23 and K2 in the others)[10,36,40,59]. Capsule type K5 is also associated with liver abscess in diverse strain backgrounds[59], and K3 is restricted to the rare rhinoscleromatis lineage (ST67)[42]. Little is understood about the relative virulence of other capsule types, and most non-hypervirulent clones, including the global MDR clones, exhibit substantial K-locus diversity[10,47]. A notable exception is CG307, which so far always carries the K-locus 102 (KL102) locus, as well as an additional putative capsule synthesis locus that has a distinct structure to the K-locus and is rare in the broader *K. pneumoniae* population[53].

Similar gene-content variation has been used to define 12 distinct O-loci[99,105]. Unlike the K-loci, however, the genes responsible for defining the nine recognized LPS O serotypes and >5 subtypes[106,107] include those at the O-locus and also other regions of the genome (locations are type dependent)[106]. Serotypes O1 and O2 are most common among clinical *K. pneumoniae* isolates[99] and may provide comparatively enhanced protection against phagocytosis[99,100]. There are no firm data as yet on whether K-locus or O-locus variation is related to niche or host specialization, as is the case in other bacterial species[108].

The extent and clinical impact of allelic diversity at the *fim* and *mrk* loci remain relatively unexplored; however, there is evidence that both types of fimbriae contribute to intestinal colonization, biofilm formation on catheters and the ability to cause pneumonia and urinary tract infections[109,110]. Notably, both loci also play a role in adhesion to plant cells[109].

Virulence factors
Features encoded by accessory loci which are entirely absent from most *Klebsiella pneumoniae* but whose presence increases either disease severity or propensity to cause disease.

FIB_K replicons
Plasmids of incompatibility type $IncFIB_K$.

Hypermucoidy
A phenotypic state characterized by 'sticky' growth and identified by production of a viscous filament (≥5 mm) when a colony is stretched by a culture loop.

Virulence factors

Numerous accessory gene-encoded virulence factors are known to further enhance the severity of *K. pneumoniae* infections and/or the propensity to cause disease. Here, we focus on those that have been subject to extensive experimental validation and the frequency, mobilization and diversity of which in *K. pneumoniae* has been investigated.

Acquired siderophores. Siderophore systems comprise iron-chelating molecules that can competitively scavenge iron from host proteins or other sources, and surface receptors for internalization. Four systems have been described in *K. pneumoniae*: Ent (core siderophore; encoded by *ent*) and yersiniabactin, aerobactin and salmochelin (accessory or acquired siderophores; encoded by *ybt*, *iuc* and *iro*, respectively)[9]. All three accessory systems can enhance virulence in murine models[111–114], and the presence of these loci are statistically associated with invasive CAIs in humans compared with classical HAI or asymptomatic carriage[9,115,116]. Although there is some functional redundancy between the systems, the siderophores vary in their iron-binding affinities[109] and host–immune interactions, including survival in macrophages[117–119]. For example, the core siderophore Ent is bound by host lipocalin 2 (Lcn2) secreted by neutrophils and epithelial cells, disrupting Ent's iron-scavenging function and inducing an inflammatory response[120]. However, the acquired *iro* locus results in production of glycosylated Ent (Gly-Ent; also known as salmochelin), which evades Lcn2 binding while maintaining iron-scavenging activity; aerobactin and yersiniabactin have lower affinity for iron than Ent but are not bound by Lcn2[97,101,112].

The *ybt* locus is statistically associated with infections (both classical HAIs and hypervirulent CAIs), being present in 30–40% of *K. pneumoniae* human clinical isolates and 13% of colonizing isolates[9,115]. It is widely distributed in the *K. pneumoniae* population, found at 6–80% and 78–100% frequency in the global MDR and hypervirulent clones, respectively[10,115] (FIG. 2). In *K. pneumoniae*, the *ybt* locus is usually mobilized by an ICE that integrates at an aspartate (Asn) tRNA gene (ICE*Kp*; of which numerous allelic and structural variants have been described), but also occasionally by plasmids[115,121]. ICE*Kp* appears to be quite mobile in the population, with hundreds of distinct acquisition events (defined as unique combinations of ICE*Kp* locus structure, integration site and chromosomal ST) identified across clinical, non-clinical, classical and hypervirulent *K. pneumoniae* clones[115]. Some variants are comparatively restricted in their distribution (for example, ICE*Kp10*), whereas others (for example, ICE*Kp3* and ICE*Kp4*) are widely detected across clones[115].

Iuc and *iro* are found in <10% of *K. pneumoniae*[9,10,116] (FIG. 2). With the exception of one *iro* lineage (*iro3*) co-localized with *ybt* and *rmpA* in ICE*Kp1*, and a chromosomally fixed *iuc* lineage (*iuc4*) in the rhinoscleromatis-associated ST67, the main dispersal mechanism is through plasmids, typically large (>100 kbp) FIB_K replicons known as *K. pneumoniae* virulence plasmids (KpVPs)[116]. Two of these plasmids, pK2044 (KpVP-1 carrying *iuc1* and *iro1*) and Kp52.145 pII (KpVP-2 carrying *iuc2* and *iro2*) have been extensively characterized owing to their association with hypervirulent clones: KpVP-1 in CG23 (98%), CG86 (70%) and CG65 (78%); and KpVP-2 in CG380 (100%) and CG66 (100%)[10,116]. The dominant virulence plasmids KpVP-1 and KpVP-2 are non-conjugative (that is, they are not self-transmissible). However, they can be mobilized between bacterial cells by other conjugative elements[122], and mosaic plasmids comprising regions of KpVP-1 fused with those of conjugative plasmids can result in conjugative virulence plasmids[12,123]. Several additional plasmids carrying distinct *iuc* or *iro* lineages (*iuc3*, *iuc5*, *iro4* and/or *iro5*) have also been identified, including many that carry genes predicted to encode conjugative machinery, and these appear to circulate outside the hypervirulent clones[116].

Colibactin. The genotoxic polyketide colibactin, which induces DNA damage in eukaryotic cells, was first described in *E. coli* but is also found in ~10% of *K. pneumoniae*, encoded in ICE*Kp10* by the *clb* (*pks*) locus[115,124]. Colibactin is proposed to promote mucosal and gut colonization, dissemination to the blood and other organs, and may contribute to colorectal cancer[124–126]. In *K. pneumoniae*, colibactin is primarily associated with the liver abscess clones CG66 (100%) and CG23 (80%) but is also present in subpopulations of hypervirulent clones CG65 and CG380 (FIG. 2). In CG23, it is restricted to the globally disseminated CG23-I sublineage that accounts for the majority of CG23 liver abscess[46]. Interestingly, ICE*Kp10* has also been acquired several times in CG258 and is common in the widely distributed ST258-*wzi154* (KL107) subclade, but the *clb* locus is disrupted and likely inactive, suggesting that it may be negatively selected in this clone[115].

Hypermucoidy. Hypermucoidy is perhaps the most well-known virulence factor for *K. pneumoniae* but its genetic basis and contribution to disease are not straightforward (BOX 3). The phenotype is commonly associated with capsule overproduction due to the presence of one or both of the accessory regulator genes *rmpA* or *rmpA2* (which share >80% homology)[127,128], although it has recently been demonstrated that hypermucoidy can occur without capsule overproduction[129]. *rmpA* and *rmpA2* are typically situated near *iro* and *iuc*, respectively, on the virulence plasmids and thus share the distribution of both KpVP-1 and KpVP-2; however, *rmpA* and *iro* are also co-localized in ICE*Kp1*, and *rmpA2* and *iuc* are co-localized in non-canonical virulence plasmids[12,116].

Other factors. Analysis of fimbrial operons in two genomes suggests that there are diverse accessory adhesion factors in the *K. pneumoniae* population that may also contribute to virulence[109,130]. Various other accessory loci have been identified as contributing to *K. pneumoniae* virulence, mostly on the basis of screening targeted gene knockouts or mutant libraries constructed in the background of hypervirulent strains (for example, strains B5055/Kp52.145 (K2 ST66), NTUH-K2044 (K1 ST23), CG43 (K2 ST86) and KPRR1/ATCC 43816 (K2 ST493))

Psicose sugar
A monosaccharide molecule, also known as allulose.

in murine models[97]. However, the genetic diversity and population distribution of these loci, their functional relevance in non-hypervirulent strain backgrounds and their association with clinical infections in humans remain largely unexplored. Pan-genome-wide association studies of isolates associated with different clinical manifestations, which encompass diverse human clinical isolates and adjust for potential confounding of the associations by bacterial population structure[131], may help to clarify these issues. For example, all three acquired siderophores as well as *rmpA*, *rmpA2*, *clb*, microcin and allantoinase, but not the proposed virulence factors *kfu* or *kvgAS* (BOX 3), were associated with invasive disease in a post hoc analysis of diverse human isolates[9], and comparison of prospectively-collected pneumonia and blood isolates with gut-colonizing isolates from a single patient population identified five factors including *ter* (tellurium resistance) and a novel psicose sugar utilization locus as independently associated with disease[132].

Hypervirulence

A small number of clones account for the majority of hypervirulent infections; these clones typically carry a combination of virulence-associated variants of core pathogenicity factors (K1 or K2 capsules; O1 or O2 LPS) as well as accessory virulence factors (*rmpA* and/or *rmpA2*, and acquired siderophores *ybt*, *iuc* and *iro*)[10,59] (FIG. 2). In rarer instances, strains with non-K1 or non-K2 capsules (for example, K5, K20, K54 and K57) and/or without the accessory virulence factors have also been reported to cause hypervirulent infections[59,133,134]. The strong influence of population structure (FIG. 2) and the strong genetic linkage between virulence loci co-located in ICE*Kp* and/or virulence plasmids[115,116] have complicated attempts to tease apart the relative importance of these factors to human clinical infections and thus to predict hypervirulence from genotype (BOX 3). Nevertheless, the virulence plasmid-associated *iuc*, *iro* and *rmpA* or *rmpA2* loci are recognized as important predictors for clinical hypervirulence[9,133], and the acquisition of *iuc* and *rmpA2* by classical strains has been associated with enhanced morbidity and mortality[11]. The *peg-344* transporter gene (occasionally annotated as *pagO*) has also been reported as a marker for hypervirulence[133,135]. However, the studies in question accounted for neither population structure nor the gene's strong linkage with other known hypervirulence factors (*peg-344* is located between *rmpA* and *iro* on KpVP-1 (REFS[136,137]), KpVP-2 (REF.[138]) and ICE*Kp*1 (REF.[121])), and *peg-344* knockout mutants were not attenuated in either of two independent models of hypervirulent infection[139,140].

Convergence of resistance and virulence

MDR and virulence have historically been associated with non-overlapping populations of *K. pneumoniae* (for example, see FIG. 2c), with only the occasional and sporadic report of a hypervirulent strain acquiring AMR[40,46,55,141]. However, as both functions are mobile in the population, the convergence of MDR and virulence factors in the same strain is possible and may begin to erode the current boundaries between MDR and hypervirulent clones, exacerbating the wider public health threat posed by *K. pneumoniae*.

The scenarios for convergence can be broadly divided along two axes: clone background, that is, hypervirulent clones acquiring MDR versus MDR clones acquiring virulence factors; and MGE dynamics, that is, transmission of complete MDR[55,141,142] or virulence plasmids[11] versus transmission and linkage of MGEs to create hybrid vectors of both AMR and virulence. Hybrid vectors have been reported as a result of several mechanisms, including recombination of AMR and virulence plasmid backbones[12], insertion of AMR genes into common virulence plasmids[143,144] or insertion of virulence loci into MDR plasmids[55]. The most common scenario so far appears to involve the presence of virulence plasmids or hybrid AMR–virulence plasmids in MDR clones (Supplementary Table 7). This is consistent with reported differences in HGT dynamics between hypervirulent and MDR clones, with MDR clones displaying much greater diversity of gene content associated with various HGT mechanisms (plasmids, phage and homologous recombination)[10]. It is important to note, however, that although MDR clones carrying the virulence plasmids have been reported to manifest enhanced pathogenicity in the hospital setting[11], they have not yet been shown to cause hypervirulent CAIs.

Geographically, the focal point for convergence appears to be Asia (33 of 58 documented cases in Supplementary Table 7), where both MDR and hypervirulence are common. Carbapenem-resistant or MDR ST11 strains carrying KL64 or KL47 and the virulence

Box 3 | **Revisiting putative hypervirulence determinants**

Genomics studies show how diversity is structured in the *Klebsiella pneumoniae* population and reveal how this structure has sometimes confounded apparent associations between proposed virulence genes and hypervirulence phenotypes, resulting in some common misconceptions regarding the 'hypervirulence markers'.

Capsule-specific genes

- K1 and K2 capsules are highly conserved within individual hypervirulent clones and are found in numerous non-hypervirulent clones. They are not sufficient for hypervirulence[10,42].
- Two proposed hypervirulence markers are now known to be capsule-locus genes: *magA* (K1-specific *wzy* capsule repeat unit polymerase)[167] and *wcaG*, a component of K-locus 1 (KL1) as well as KL6, KL16, KL54, KL58, KL63 and KL113 (REFS[99,102,167]).

Clone-specific markers

- *kfu* (ferric iron uptake) and *allS* (allantoin metabolism) are conserved in clonal group 23 (CG23)[46] but not in other hypervirulent clones[40].
- *kvgAS* is conserved in CG86, CG65 and CG25 (REF.[40]), but absent from CG23 (REF.[46]).

Hypermucoidy and *rmpA* genes

- Hypermucoidy is not exclusive to hypervirulent strains, and not all hypervirulent strains are hypermucoid by the string test. Concordance between the string test and clinically or animal model-assessed hypervirulence varies (51–98%)[133].
- Truncations in *rmpA* are common; half of the recorded *rmpA* or *rmpA2* alleles carry frameshift mutations arising from indels at a poly(G) sequence, causing loss of function and lack of hypermucoidy[168].

Siderophore receptors

Among the *K. pneumoniae* core genes is a TonB-dependent receptor that shares ~71% homology with the aerobactin receptor *iutA*. Detection of this core gene by PCR or sequencing is sometimes mistaken for the presence of aerobactin[68], despite the lack of aerobactin synthesis genes *iucABCD*.

Classical infections
A term used to refer to opportunistic healthcare-associated *Klebsiella pneumoniae* infections.

plasmid with *iuc* (and often *rmpA2*) are disseminating in China[11,145]; multiple convergent strains and plasmid vectors have also been reported from South and Southeast Asia[61] (Supplementary Table 7).

Genomic surveillance

WGS serves as a powerful tool to monitor and explore *K. pneumoniae* epidemiology by affording a comprehensive picture of strain populations in a single assay, allowing the simultaneous detection of species, lineage, K-loci and O-loci, core and accessory AMR and virulence determinants, and also base pair-resolved measurement of strain relatedness with which to assess evidence of transmission. Indeed, much of this can now be accomplished in a few steps using tools such as Kleborate, PathogenWatch and BIGSdb.

Genomic surveillance has already been instrumental in elucidating patterns of clonal spread at local, regional and global levels and gaining insights into the burden and spread of *K. pneumoniae* HAIs[16–19,54,61,62]. Given the genetic diversity and complexity of *K. pneumoniae* clinical threats, WGS is rapidly becoming a crucial tool for epidemiology and surveillance of this pathogen, particularly for monitoring global emergence and dissemination of problem clones, tracking the convergence of AMR and hypervirulence, and more fully exploring links between clinical infections and potential ecological reservoirs.

Importantly, the utility of WGS surveillance data extends well beyond its initial purpose, as each study provides information not only on local disease epidemiology, but also contributes to understanding the ongoing evolution and geographical spread of clinically relevant strains and features. For example, publicly deposited WGS data from a recent Europe-wide genomic surveillance study of CRKp[54] helped to resolve the likely origin of two cases of AMR *iuc*$^+$ *K. pneumoniae* infections identified in Norway[12], before the original surveillance study report was published. Even studies of local *K. pneumoniae* hospital outbreaks have provided lessons for infection control well beyond their initial sphere of concern, showing that MDR *K. pneumoniae* are more transmissible than susceptible strains in hospitals[19,54] and confirming hospital plumbing as a source of prolonged outbreaks[146], providing focal points for intervention and prevention strategies.

WGS investigations of CRKp have also led to widespread recognition of the roles of plasmid and MGE transfer in hospital outbreaks[92], with *K. pneumoniae* playing a leading role as both donor and recipient for AMR gene transfer[49,92,93,146,147]. Understanding the potential for 'plasmid outbreaks' is transforming the way that screening and surveillance of carbapenemase-producing Enterobacteriaceae is managed, shifting the focus from individual organisms to the enzymes and plasmids that make the organism a threat[147].

Conclusions and future perspectives

The application of genomics has revealed remarkable genetic diversity and has drastically accelerated understanding of *K. pneumoniae* population structure, AMR, pathogenicity and transmission in clinical environments. The emerging population genomic framework has helped clarify many points of confusion related to taxonomy, resistance and virulence; for example, by highlighting those AMR and virulence determinants that are core to all strains (FIG. 2) and revealing how clonal spread can confound association analyses, such as that between the *kfu* operon and the hypervirulent phenotype (BOX 3). This framework also provides important context to help guide the design and interpretation of future experimental and clinical studies, which are critical to confirm — and to elucidate the mechanisms underlying — observed associations between genetic variation and clinically relevant phenotypes. For example, it is becoming apparent that there is no such thing as 'typical' *K. pneumoniae*, as gene-content turnover is extensive and even the 'core' loci harbour widespread genetic variation of clinical relevance. Experimental studies could benefit from taking this into account, by including multiple diverse strains and exploring the functional dependence of specific genes on factors that vary substantially with genetic background, including both allelic and gene-content variation. Ideally, such studies would also report the genome sequences of strains, enabling others to interpret the findings and make comparisons with their own strain collections.

Increased reporting of convergent AMR-virulent isolates and the discovery that the infamous 'hypervirulent' MDR ST11 isolated in China was already widely disseminated prior to initial reporting both clearly highlight the need for enhanced global surveillance. Genomics enables identification and typing of a wide array of clinically relevant features in a single assay, facilitating rapid detection of emerging threats and informing containment strategies (see Related links for tools and resources). There is now also a considerable body of evidence supporting the use of genomics to track strain and/or plasmid transmissions within hospitals to inform infection control and help reduce the burden of disease. However, not all infections are associated with hospital transmission (for example, hypervirulent infections and sporadic classical infections), and these cannot be prevented by enhanced infection control measures. This highlights the need to better understand the interaction between *K. pneumoniae* and human hosts, including core and variable bacterial mechanisms, host factors and the role of gut colonization, which generally precedes progression to disease.

Combining comparative genomic analyses with targeted population surveys and experimental laboratory studies could prove a powerful approach to identify and functionally validate additional clinically relevant genomic markers. We envisage that this will help to populate a risk framework for *K. pneumoniae*, which captures the spectrum of strain-specific risk for colonization, persistent carriage, disease progression and nosocomial transmission. Such a risk framework could have many benefits, such as prioritizing scarce infection control resources in hospital and informing the design of novel therapeutics, control strategies and patient care practices.

An important outstanding question remains, regarding the ultimate reservoirs of *K. pneumoniae* and how they contribute to human infections. The growing international focus on a 'One Health' approach to understanding

AMR and its drivers has spurred an increased focus on the ecology of *K. pneumoniae* and other opportunistic pathogens. Ecological studies and cross-niche surveillance for *K. pneumoniae* are gaining popularity and can be expected to provide evidence regarding the major animal, food and environmental reservoirs of *K. pneumoniae*. However, given the ubiquity and diversity of the species, detection alone is unlikely to be of much public health benefit. Rather, population genomic analyses will be needed to explore which reservoirs contribute to the amplification and spread of AMR and/or virulence determinants and to identify pathways for transmission to humans that could potentially be disrupted in the name of disease control.

Published online 13 February 2020

1. Adeolu, M., Alnajar, S., Naushad, S. & Gupta, R. S. Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. *Int. J. Syst. Evol. Microbiol.* **66**, 5575–5599 (2016).
2. Pendleton, J. N., Gorman, S. P. & Gilmore, B. F. Clinical relevance of the ESKAPE pathogens. *Expert. Rev. Anti Infect. Ther.* **11**, 297–308 (2013).
3. Okomo, U. et al. Aetiology of invasive bacterial infection and antimicrobial resistance in neonates in sub-Saharan Africa: a systematic review and meta-analysis in line with the STROBE-NI reporting guidelines. *Lancet Infect. Dis.* **19**, 1219–1234 (2019).
4. Zaidi, A. K. M. et al. Hospital-acquired neonatal infections in developing countries. *Lancet* **365**, 1175–1188 (2005).
5. World Health Organization. Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics (WHO, 2017).
6. Cassini, A. et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* **19**, 56–66 (2019).
7. Musicha, P. et al. Trends in antimicrobial resistance in bloodstream infection isolates at a large urban hospital in Malawi (1998–2016): a surveillance study. *Lancet. Infect. Dis.* **17**, 1042–1052 (2017).
8. Bagley, S. T. Habitat association of *Klebsiella* species. *Infect. Contr.* **6**, 52–58 (1985).
9. Holt, K. E. et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc. Natl Acad. Sci. USA* **112**, E3574–E3581 (2015).
This large-scale comparative genomics study of diverse *K. pneumoniae* and related members of the species complex from seven countries establishes the global genomic framework and the scale and heterogeneity of the pan-genome, and identifies genetic loci that are statistically associated with invasive disease versus asymptomatic colonization in humans.
10. Wyres, K. L. et al. Distinct evolutionary dynamics of horizontal gene transfer in drug resistant and virulent clones of *Klebsiella pneumoniae*. *PLoS Genet.* **15**, e1008114 (2019).
11. Gu, D. et al. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. *Lancet Infect. Dis.* **3099**, 1–10 (2017).
This study presents an initial report of MDR-ST11 harbouring a KpVP-1 virulence plasmid variant, which displays enhanced survival in a human neutrophil assay and enhanced virulence in the *Galleria mellonella* infection model in comparison with typical MDR-ST11.
12. Lam, M. M. C. et al. Convergence of virulence and MDR in a single plasmid vector in MDR *Klebsiella pneumoniae* ST15. *J. Antimicrob. Chemother.* **74**, 1218–1222 (2019).
13. Long, S. W. et al. Whole-genome sequencing of human clinical *Klebsiella pneumoniae* isolates reveals misidentification and misunderstandings of *Klebsiella pneumoniae*, *Klebsiella variicola*, and *Klebsiella quasipneumoniae*. *mSphere* **2**, e00290–e00317 (2017).
14. Gorrie, C. L. et al. Gastrointestinal carriage is a major reservoir of *K. pneumoniae* infection in intensive care patients. *Clin. Infect. Dis.* **65**, 208–215 (2017).
15. Rodrigues, C., Passet, V., Rakotondrasoa, A. & Brisse, S. Identification of *Klebsiella pneumoniae, Klebsiella quasipneumoniae, Klebsiella variicola* and related phylogroups by MALDI-TOF mass spectrometry. *Front. Microbiol.* **9**, 1–7 (2018).
16. Long, S. W. et al. Population genomic analysis of 1,777 extended-spectrum β-lactamase-producing *Klebsiella pneumoniae* isolates, Houston, Texas: unexpected abundance of clonal group 307. *mBio* **8**, e00489–e00517 (2017).
17. Henson, S. P. et al. Molecular epidemiology of *Klebsiella pneumoniae* invasive infections over a decade at Kilifi County Hospital in Kenya. *Int. J. Med. Microbiol.* **307**, 422–429 (2017).
18. Heinz, E. et al. Resistance mechanisms and population structure of highly drug resistant *Klebsiella* in Pakistan during the introduction of the carbapenemase NDM-1. *Sci. Rep.* **9**, 2392 (2019).
19. Gorrie, C. L. et al. Antimicrobial resistant *Klebsiella pneumoniae* carriage and infection in specialized geriatric care wards linked to acquisition in the referring hospital. *Clin. Infect. Dis.* **67**, 161–170 (2018).
20. Wyres, K. L. & Holt, K. E. *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Curr. Opin. Microbiol.* **45**, 131–139 (2018).
21. Marques, C. et al. Evidence of *Klebsiella pneumoniae* sharing between healthy companion animals and co-habiting humans. *J. Clin. Microbiol.* **57**, e01537–e01618 (2019).
22. Zadoks, R. N. et al. Sources of *Klebsiella* and *Raoultella* species on dairy farms: be careful where you walk. *J. Dairy Sci.* **94**, 1045–1051 (2011).
23. Conlan, S., Kong, H. H. & Segre, J. A. Species-level analysis of DNA sequence data from the NIH human microbiome project. *PLoS One* **7**, e47075 (2012).
24. Martin, R. M. et al. Molecular epidemiology of colonizing and infecting isolates of *Klebsiella pneumoniae*. *mSphere* **1**, e00261–e00316 (2016).
25. Ludden, C. et al. A One Health study of the genetic relatedness of *Klebsiella pneumoniae* and their mobile elements in the East of England. *Clin. Infect. Dis.* **70**, 219–226 (2020).
26. Chung, D. R. et al. Fecal carriage of serotype K1 *Klebsiella pneumoniae* ST23 strains closely related to liver abscess isolates in Koreans living in Korea. *Eur. J. Clin. Microbiol. Infect. Dis.* **31**, 481–486 (2012).
27. Lin, Y.-T. et al. Seroepidemiology of *Klebsiella pneumoniae* colonizing the intestinal tract of healthy Chinese and overseas Chinese adults in Asian countries. *BMC Microbiol.* **12**, 13 (2012).
28. Löhr, I. H. et al. Long-term faecal carriage in infants and intra-household transmission of CTX-M-15-producing *Klebsiella pneumoniae* following a nosocomial outbreak. *J. Antimicrob. Chemother.* **68**, 1043–1048 (2013).
29. Mo, Y. et al. Carriage duration of carbapenemase-producing Enterobacteriaceae in a hospital cohort—implications for infection control measures. *medRxiv*. https://doi.org/10.1101/19001479 (2019).
30. Podschun, R. & Ullmann, U. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin. Microbiol. Rev.* **11**, 589–603 (1998).
31. Shimasaki, T. et al. Increased relative abundance of *Klebsiella pneumoniae* carbapenemase-producing *Klebsiella pneumoniae* within the gut microbiota is associated with risk of bloodstream infection in long-term acute care hospital patients. *Clin. Infect. Dis.* **68**, 2053–2059 (2019).
32. Xu, L., Sun, X. & Ma, X. Systematic review and meta-analysis of mortality of patients infected with carbapenem-resistant *Klebsiella pneumoniae*. *Ann. Clin. Microbiol. Antimicrob.* **16**, 18 (2017).
33. Bassetti, M., Peghin, M., Vena, A. & Giacobbe, D. R. Treatment of infections due to MDR Gram-negative bacteria. *Front. Med.* **6**, 74 (2019).
34. Manohar, P., Tamhankar, A. J., Lundborg, C. S. & Nachimuthu, R. Therapeutic characterization and efficacy of bacteriophage cocktails infecting *Escherichia coli, Klebsiella pneumoniae*, and *Enterobacter* species. *Front. Microbiol.* **10**, 574 (2019).
35. Meatherall, B. L., Gregson, D., Ross, T., Pitout, J. D. D. & Laupland, K. B. Incidence, risk factors, and outcomes of *Klebsiella pneumoniae* bacteremia. *Am. J. Med.* **122**, 866–873 (2009).
36. Russo, T. A. & Marr, C. M. Hypervirulent *Klebsiella pneumoniae*. *Clin. Microbiol. Rev.* **32**, e00001–e00019 (2019).
37. Ko, W. C. et al. Community-acquired *Klebsiella pneumoniae* bacteremia: global differences in clinical patterns. *Emerg. Infect. Dis.* **8**, 160–166 (2002).
38. Kim, J. K., Chung, D. R., Wie, S. H., Yoo, J. H. & Park, S. W. Risk factor analysis of invasive liver abscess caused by the K1 serotype *Klebsiella pneumoniae*. *Eur. J. Clin. Microbiol. Infect. Dis.* **28**, 109–111 (2009).
39. Brockhurst, M. A. et al. The ecology and evolution of pangenomes. *Curr. Biol.* **29**, R1094–R1103 (2019).
40. Bialek-Davenet, S. et al. Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg. Infect. Dis.* **20**, 1812–1820 (2014).
This work establishes the cgMLST scheme for *K. pneumoniae* and the associated species complex, which is based on 694 core genes and is available through the BIGSdb-*Kp* online database.
41. Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. & Brisse, S. Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *J. Clin. Microbiol.* **43**, 4178–4182 (2005).
42. Brisse, S. et al. Virulent clones of *Klebsiella pneumoniae*: identification and evolutionary scenario based on genomic and phenotypic characterization. *PLoS One* **4**, e4982 (2009).
43. Breurec, S. et al. *Klebsiella pneumoniae* resistant to third-generation cephalosporins in five African and two Vietnamese major towns: multiclonal population structure with two major international clonal groups, CG15 and CG258. *Clin. Microbiol. Infect.* **19**, 349–355 (2013).
44. McInerney, J. O., McNally, A. & O'Connell, M. J. Why prokaryotes have pangenomes. *Nat. Microbiol.* **2**, 17040 (2017).
45. Wyres, K. L. et al. Extensive capsule locus variation and large-scale genomic recombination within the *Klebsiella pneumoniae* clonal group 258. *Genome Biol. Evol.* **7**, 1267–1279 (2015).
46. Lam, M. M. C. et al. Population genomics of hypervirulent *Klebsiella pneumoniae* clonal group 23 reveals early emergence and rapid global dissemination. *Nat. Commun.* **9**, 2703 (2018).
47. Bowers, J. R. et al. Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS One* **10**, e0133727 (2015).
48. Navon-Venezia, S., Kondratyeva, K. & Carattoli, A. *Klebsiella pneumoniae*: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol. Rev.* **41**, 252–275 (2017).
49. Conlan, S. et al. Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. *mBio* **7**, e00742–e00816 (2016).
50. Shen, J., Lv, L., Wang, X., Xiu, Z. & Chen, G. Comparative analysis of CRISPR–Cas systems in *Klebsiella* genomes. *J. Basic. Microbiol.* **57**, 325–336 (2017).
51. Ellington, M. J. et al. Contrasting patterns of longitudinal population dynamics and antimicrobial resistance mechanisms in two priority bacterial

pathogens over 7 years in a single center. *Genome Biol.* **20**, 184 (2019).
52. Wyres, K. L. & Holt, K. E. *Klebsiella pneumoniae* population genomics and antimicrobial-resistant clones. *Trends Microbiol.* **24**, 944–956 (2016).
53. Wyres, K. L. et al. Emergence and rapid global dissemination of CTX-M-15-associated *Klebsiella pneumoniae* strain ST307. *J. Antimicrob. Chemother.* **74**, 577–581 (2019).
54. David, S. et al. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat. Microbiol.* **4**, 1919–1929 (2019).
This study is a genomic analysis of >1,700 CRKp and carbapenem-susceptible *K. pneumoniae* infections isolated from patients in 244 hospitals in 32 countries across Europe, providing the first large-scale systematic sample for comparison of clone and resistance gene distributions.
55. Turton, J. F. et al. Virulence genes in isolates of *Klebsiella pneumoniae* from the UK during 2016, including among carbapenemase gene-positive hypervirulent K1-ST23 and 'non-hypervirulent' types ST147, ST15 and ST383. *J. Med. Microbiol.* **67**, 118–128 (2017).
56. Siu, L. K. et al. Molecular typing and virulence analysis of serotype K1 *Klebsiella pneumoniae* strains isolated from liver abscess patients and stool samples from noninfectious subjects in Hong Kong, Singapore, and Taiwan. *J. Clin. Microbiol.* **49**, 3761–3765 (2011).
57. Lin, J. C. et al. Genotypes and virulence in serotype K2 *Klebsiella pneumoniae* from liver abscess and non-infectious carriers in Hong Kong, Singapore and Taiwan. *Gut Pathog.* **12**, 21 (2014).
58. Shi, Q. et al. Diversity of virulence level phenotype of hypervirulent *Klebsiella pneumoniae* from different sequence type lineage. *BMC Microbiol.* **18**, 1–6 (2018).
59. Lee, I. R. et al. Differential host susceptibility and bacterial virulence factors driving *Klebsiella* liver abscess in an ethnically diverse population. *Sci. Rep.* **13**, 29316 (2016).
60. Zhang, Y. et al. High prevalence of hypervirulent *Klebsiella pneumoniae* infection in China: geographic distribution, clinical characteristics, and antimicrobial resistance. *Antimicrob. Agents Chemother.* **60**, 6115–6120 (2016).
61. Wyres, K. L. et al. Genomic surveillance for hypervirulence and multi-drug resistance in invasive *Klebsiella pneumoniae* from South and Southeast Asia. *Genomic med.* **12**, 11 (2019).
62. Heinz, E., Brindle, R., Morgan-McCalla, A., Peters, K. & Thomson, N. R. Caribbean multi-centre study of *Klebsiella pneumoniae:* whole genome sequencing, antimicrobial resistance and virulence factors. *Microb. Genomics.* **5**, e000266 (2019).
63. Musicha, P. et al. Genomic analysis of *Klebsiella pneumoniae* isolates from Malawi reveals acquisition of multiple ESBL determinants across diverse lineages. *J. Antimicrob. Chemother.* **74**, 1223–1232 (2019).
64. Zhang, R. et al. Presence of NDM in non-*E. coli* Enterobacteriaceae in the poultry production environment. *J. Antimicrob. Chemother.* **74**, 2209–2213 (2019).
65. Marques, C. et al. *Klebsiella pneumoniae* causing urinary tract infections in companion animals and humans: population structure, antimicrobial resistance and virulence genes. *J. Antimicrob. Chemother.* **74**, 594–602 (2019).
66. Anzai, E. K. et al. First case report of non-human primates (*Alouatta clamitans*) with the hypervirulent *Klebsiella pneumoniae* serotype K1 strain ST23: a possible emerging wildlife pathogen. *J. Med. Primatol.* **46**, 337–342 (2017).
67. Bowring, B. G., Fahy, V. A., Morris, A. & Collins, A. M. An unusual culprit: *Klebsiella pneumoniae* causing septicaemia outbreaks in neonatal pigs? *Vet. Microbiol.* **203**, 267–270 (2017).
68. Runcharoen, C. et al. Whole genome sequencing reveals high-resolution epidemiological links between clinical and environmental *Klebsiella pneumoniae*. *Genome Med.* **9**, 6 (2017).
69. Davis, G. S. et al. Intermingled *Klebsiella pneumoniae* populations between retail meats and human urinary tract infections. *Clin. Infect. Dis.* **61**, 892–899 (2015).
70. Singh, A., Lekshmi, M., Prakasan, S., Nayak, B. & Kumar, S. Multiple antibiotic-resistant, extended spectrum-β-lactamase (ESBL)-producing Enterobacteria in fresh seafood. *Microorganisms* **5**, E53 (2017).
71. Zekar, F. M. et al. From farms to markets: Gram-negative bacteria resistant to third-generation cephalosporins in fruits and vegetables in a region of North Africa. *Front. Microbiol.* **8**, 1569 (2017).
72. Yaici, L. et al. Spread of ESBL/AmpC-producing *Escherichia coli* and *Klebsiella pneumoniae* in the community through ready-to-eat sandwiches in Algeria. *Int. J. Food Microbiol.* **245**, 66–72 (2017).
73. Projahn, M. et al. Contamination of chicken meat with extended-spectrum β-lactamase-producing *Klebsiella pneumoniae* and *Escherichia coli* during scalding and defeathering of broiler carcasses. *Food Microbiol.* **77**, 185–191 (2019).
74. Rodrigues, C. et al. Description of *Klebsiella africanensis* sp. nov., *Klebsiella variicola* subsp. *tropicalensis* subsp. nov. and *Klebsiella variicola* subsp. *variicola* subsp. nov. *Res. Microbiol.* **170**, 165–170 (2019).
75. Ford, P. & Avison, M. Evolutionary mapping of the SHV β-lactamase and evidence for two separate IS*26*-dependent bla_{SHV} mobilization events from the *Klebsiella pneumoniae* chromosome. *J. Antimicrob. Chemother.* **54**, 69–75 (2004).
76. Liakopoulos, A., Mevius, D. & Ceccarelli, D. A review of SHV extended-spectrum β-lactamases: neglected yet ubiquitous. *Front. Microbiol.* **5**, 1374 (2016).
77. Turner, M. et al. Plasmid-borne bla_{SHV} genes in *Klebsiella pneumoniae* are associated with strong promoters. *J. Antimicrob. Chemother.* **64**, 960–964 (2009).
78. Ito, R. et al. Widespread fosfomycin resistance in Gram-negative bacteria attributable to the chromosomal *fosA* gene. *mBio* **8**, e00749–e00817 (2017).
79. Li, J. et al. The nature and epidemiology of OqxAB, a multidrug efflux pump. *Antimicrob. Resist. Infect. Control.* **8**, 44 (2019).
80. Bernardini, A. et al. The intrinsic resistome of *Klebsiella pneumoniae*. *Int. J. Antimicrob. Agents.* **53**, 29–33 (2019).
81. Jana, B. et al. The secondary resistome of multidrug-resistant *Klebsiella pneumoniae*. *Sci. Rep.* **7**, 42483 (2017).
82. Nicolas-Chanoine, M.-H., Mayer, N., Guyot, K., Dumont, E. & Pagès, J.-M. Interplay between membrane permeability and enzymatic barrier leads to antibiotic-dependent resistance in *Klebsiella pneumoniae*. *Front. Microbiol.* **9**, 1422 (2018).
83. Xu, Q. et al. Efflux pumps AcrAB and OqxAB contribute to nitrofurantoin resistance in an uropathogenic *Klebsiella pneumoniae* isolate. *Int. J. Antimicrob. Agents* **54**, 223–227 (2019).
84. He, F. et al. Tigecycline susceptibility and the role of efflux pumps in tigecycline resistance in KPC-producing *Klebsiella pneumoniae*. *PLoS one* **10**, e0119064 (2015).
85. Fajardo-Lubián, A., Ben Zakour, N. L., Agyekum, A., Qi, Q. & Iredell, J. R. Host adaptation and convergent evolution increases antibiotic resistance without loss of virulence in a major human pathogen. *PLoS Pathog.* **15**, e1007218 (2019).
This study shows that deletions or mutations within OmpK35/36 porins are widespread and have emerged in multiple independent lineages of *K. pneumoniae*, facilitating enhanced resistance to β-lactams (including carbapenems) without significantly reducing the ability to colonize the gut or to cause pneumonia (in murine models).
86. Wong, J. L. C. et al. OmpK36-mediated carbapenem resistance attenuates ST258 *Klebsiella pneumoniae* in vivo. *Nat. Commun.* **10**, 3957 (2019).
This study experimentally confirms predictions that specific di-amino acid insertions in OmpK36 constrict its central pore, restricting diffusion of nutrients and carbapenems, and demonstrates using competition assays that these mutations do reduce fitness in the context of the murine pneumonia model.
87. Lunha, K. et al. High-level carbapenem-resistant OXA-48-producing *Klebsiella pneumoniae* with a novel OmpK36 variant and low-level, carbapenem-resistant, non-porin-deficient, OXA-181-producing *Escherichia coli* from Thailand. *Diagn. Microbiol. Infect. Dis.* **85**, 221–226 (2016).
88. Cain, A. K. et al. Morphological, genomic and transcriptomic responses of *Klebsiella pneumoniae* to the last-line antibiotic colistin. *Sci. Rep.* **8**, 9868 (2018).
89. Chang, H.-H. et al. Origin and proliferation of multiple-drug resistance in bacterial pathogens. *Microbiol. Mol. Biol. Rev.* **79**, 101–116 (2015).
90. Lehtinen, S. et al. Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proc. Natl Acad. Sci. USA* **114**, 1075–1080 (2017).
91. Conlan, S. et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Sci. Transl Med.* **6**, 254ra126 (2014).
92. Martin, J. et al. Covert dissemination of carbapenemase-producing *Klebsiella pneumoniae* (KPC) in a successfully controlled outbreak: long- and short-read whole-genome sequencing demonstrate multiple genetic modes of transmission. *J. Antimicrob. Chemother.* **72**, 3025–3034 (2017).
93. Sheppard, A. E. et al. Nested Russian doll-like genetic mobility drives rapid dissemination of the carbapenem resistance gene bla_{KPC}. *Antimicrob. Agents Chemother.* **60**, 3767–3778 (2016).
This study, using a genomic comparison of carbapenem-resistant Enterobacteriaceae isolated from a single health-care institution, highlights the importance of strain, plasmid and transposon transmission for the dissemination of bla_{KPC} carbapenemase genes.
94. Buckner, M. M. C. et al. Clinically relevant plasmid–host interactions indicate that transcriptional and not genomic modifications ameliorate fitness costs of *Klebsiella pneumoniae* carbapenemase-carrying plasmids. *mBio* **9**, e02303–e02317 (2018).
This study demonstrates that the pKpQIL plasmid of ST258 alters *K. pneumoniae* gene expression and shows differences in transfer efficiencies of pKpQIL and related plasmids depending on the genetic background of donor and recipient strains.
95. Hardiman, C. et al. Horizontal transfer of carbapenemase-encoding plasmids and comparison with hospital epidemiological data. *Antimicrob. Agents Chemother.* **60**, 4910–4919 (2016).
96. Lepuschitz, S. et al. Whole genome sequencing reveals resemblance between ESBL-producing and carbapenem resistant *Klebsiella pneumoniae* isolates from Austrian rivers and clinical isolates from hospitals. *Sci. Total. Environ.* **662**, 227–235 (2019).
97. Paczosa, M. K. & Mecsas, J. *Klebsiella pneumoniae*: going on the offense with a strong defense. *Microbiol. Mol. Biol. Rev.* **80**, 629–661 (2016).
98. Bengoechea, J. A. & Sa Pessoa, J. *Klebsiella pneumoniae* infection biology: living to counteract host defences. *FEMS Microbiol. Rev.* **43**, 123–144 (2019).
99. Follador, R. et al. The diversity of *Klebsiella pneumoniae* surface polysaccharides. *Microb. Genom.* **2**, e000073 (2016).
100. Wyres, K. L. et al. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb. Genom.* **2**, e000102 (2016).
101. Bachman, M. A., Lenio, S., Schmidt, L., Oyler, J. E. & Weiser, J. N. Interaction of lipocalin 2, transferrin, and siderophores determines the replicative niche of *Klebsiella pneumoniae* during pneumonia. *mBio* **3**, e00224-11 (2012).
102. Pan, Y.-J. et al. Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of *Klebsiella* spp. *Sci. Rep.* **5**, 15573 (2015).
103. Whitfield, C. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* **75**, 39–68 (2006).
104. Ørskov, I. D. A. & Fife-Asbury, M. A. New *Klebsiella* capsular antigen, K82, and the deletion of five of those previously assigned. *Int. J. Syst. Bacteriol.* **27**, 386–387 (1977).
105. Wick, R. R., Heinz, E., Holt, K. E. & Wyres, K. L. Kaptive Web: user-friendly capsule and lipopolysaccharide serotype prediction for *Klebsiella* genomes. *J. Clin. Microbiol.* **56**, e00197–e00218 (2018).
106. Clarke, B. R. et al. Molecular basis for the structural diversity in serogroup O2-antigen polysaccharides in *Klebsiella pneumoniae*. *J. Biol. Chem.* **293**, 4666–4679 (2018).
107. Guachalla, L. M. et al. Discovery of monoclonal antibodies cross-reactive to novel subserotypes of *K. pneumoniae* O3. *Sci. Rep.* **7**, 6635 (2017).
108. Mostowy, R. J. & Holt, K. E. Diversity-generating machines: genetics of bacterial sugar-coating. *Trends Microbiol.* **26**, 1008–1021 (2018).
109. Khater, F. et al. *In silico* analysis of usher encoding genes in *Klebsiella pneumoniae* and characterization of their role in adhesion and colonization. *PLoS One* **10**, e0116215 (2015).
110. Murphy, C. N., Mortensen, M. S., Krogfelt, K. A. & Clegg, S. Role of *Klebsiella pneumoniae* type 1 and type 3 fimbriae in colonizing silicone tubes implanted into the bladders of mice as a model of

catheter-associated urinary tract infections. *Infect. Immun.* **81**, 3009–3017 (2013).
111. Bachman, M. A. et al. *Klebsiella pneumoniae* yersiniabactin promotes respiratory tract infection through evasion of lipocalin 2. *Infect. Immun.* **79**, 3309–3316 (2011).
112. Russo, T. A., Olson, R., Macdonald, U., Beanan, J. & Davidson, B. A. Aerobactin, but not yersiniabactin, salmochelin, or enterobactin, enables the growth/survival of hypervirulent (hypermucoviscous) *Klebsiella pneumoniae ex vivo* and *in vivo*. *Infect. Immun.* **83**, 3325–3333 (2015).
113. Russo, T. A. et al. Aerobactin mediates virulence and accounts for increased siderophore production under iron-limiting conditions by hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*. *Infect. Immun.* **82**, 2356–2367 (2014).
114. Fischbach, M. A. et al. The pathogen-associated *iroA* gene cluster mediates bacterial evasion of lipocalin 2. *Proc. Natl Acad. Sci. USA* **103**, 16502–16507 (2006).
115. Lam, M. M. C. et al. Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICE *Kp* in *Klebsiella pneumoniae* populations. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000196 (2018).
116. Lam, M. C. C. et al. Tracking key virulence loci encoding aerobactin and salmochelin siderophore synthesis in *Klebsiella pneumoniae*. *Genome Med.* **10**, 77 (2018).
117. Holden, V. I., Bachman, M. A. & Holden, V. I. Diverging roles of bacterial siderophores during infection. *Metallomics* **7**, 986–995 (2015).
118. Saha, P. et al. The bacterial siderophore enterobactin confers survival advantage to *Salmonella* in macrophages. *Gut Microbes* **10**, 412–423 (2019).
119. Achard, M. E. S. et al. An antioxidant role for catecholate siderophores in *Salmonella*. *Biochem. J.* **454**, 543–549 (2013).
120. Holden, V. I., Breen, P., Houle, S., Dozois, C. M. & Bachman, M. A. *Klebsiella pneumoniae* siderophores induce inflammation, bacterial dissemination, and HIF-1α stablization during pneumonia. *mBio* **7**, e01397–e01416 (2016).
121. Lin, T.-L., Lee, C.-Z., Hsieh, P.-F., Tsai, S.-F. & Wang, J.-T. Characterization of integrative and conjugative element ICEKp1-associated genomic heterogeneity in a *Klebsiella pneumoniae* strain isolated from a primary liver abscess. *J. Bacteriol.* **190**, 515–526 (2008).
122. Nassif, X., Fournier, J., Arondel, J. & Sansonetti, P. J. Mucoid phenotype of *Klebsiella pneumoniae* is a plasmid-encoded virulence factor. *Infect. Immun.* **57**, 546–552 (1989).
123. Yang, X., Wai-Chi Chan, E., Zhang, R. & Chen, S. A conjugative plasmid that augments virulence in *Klebsiella pneumoniae*. *Nat. Microbiol.* **4**, 2039–2043 (2019).
124. Nougayrède, J. P. et al. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848–851 (2006).
125. Lai, Y. C. et al. Genotoxic *Klebsiella pneumoniae* in Taiwan. *PLoS One* **9**, e96292 (2014).
126. Lu, M.-C. et al. Colibactin contributes to the hypervirulence of *pks*$^{+}$ K1 CC23 *Klebsiella pneumoniae* in mouse meningitis infections. *Front. Cell. Infect. Microbiol.* **7**, 1–14 (2017).
127. Wacharotayankun, R. et al. Enhancement of extracapsular polysaccharide synthesis in *Klebsiella pneumoniae* by RmpA2, which shows homology to NtrC and FixJ. *Infect. Immun.* **61**, 3164–3174 (1993).
128. Arakawa, Y. et al. Biosynthesis of *Klebsiella* K2 capsular polysaccharide in *Escherichia coli* HB101 requires the functions of *rmpA* and the chromosomal *cps* gene cluster of the virulent strain *Klebsiella pneumoniae* Chedid (O1:K2). *Infect. Immun.* **59**, 2043–2050 (1991).
129. Walker, K. A. et al. A *Klebsiella pneumoniae* regulatory mutant has reduced capsule expression but retains hypermucoviscosity. *mBio* **10**, e00089–e00119 (2019).
This study teases apart the roles of *rmpA* and a newly identified gene immediately downstream of it (*rmpC*) in hypermucoidy and capsule production.
130. Wu, C. C., Huang, Y. J., Fung, C. P. & Peng, H. L. Regulation of the *Klebsiella pneumoniae* Kpc fimbriae by the site-specific recombinase KpcI. *Microbiology* **156**, 1983–1992 (2010).
131. Earle, S. G. et al. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. *Nat. Microbiol.* **1**, 16041 (2016).
132. Martin, R. M. et al. Identification of pathogenicity-associated loci in *Klebsiella pneumoniae* from hospitalized patients. *mSystems* **3**, e00015–e00018 (2018).
This genome-wide association study tests for *K. pneumoniae* genes associated with clinical isolates (bacteraemia or pneumonia) versus asymptomatic gut-colonizing isolates and demonstrates independent effects of the tellurite resistance operon *ter* and a psicose utilization locus.
133. Russo, T. A. et al. Identification of biomarkers for differentiation of hypervirulent *Klebsiella pneumoniae* from classical *K. pneumoniae*. *J. Clin. Microbiol.* **56**, e00776-18 (2018).
This study is the first systematic exploration of virulence markers and their relative effect on prediction of the hypervirulence phenotype, highlighting the importance of the aerobactin synthesis locus *iuc* and other plasmid co-located loci.
134. Liu, C. & Guo, J. Hypervirulent *Klebsiella pneumoniae* (hypermucoviscous and aerobactin positive) infection over 6 years in the elderly in China: antimicrobial resistance patterns, molecular epidemiology and risk factor. *Ann. Clin. Microbiol. Antimicrob.* **18**, 4 (2019).
135. Ye, M. et al. Clinical and genomic analysis of liver abscess-causing *Klebsiella pneumoniae* identifies new liver abscess-associated virulence genes. *Front. Cell Infect. Microbiol.* **6**, 1–12 (2016).
136. Wu, K. M. et al. Genome sequencing and comparative analysis of *Klebsiella pneumoniae* NTUH-K2044, a strain causing liver abscess and meningitis. *J. Bacteriol.* **191**, 4492–4501 (2009).
137. Chen, Y., Chang, H., Lai, Y., Pan, C. & Tsai, S. Sequencing and analysis of the large virulence plasmid pLVPK of *Klebsiella pneumoniae* CG43. *Gene* **337**, 189–198 (2004).
138. Lery, L. M. et al. Comparative analysis of *Klebsiella pneumoniae* genomes identifies a phospholipase D family protein as a novel virulence factor. *BMC Biol.* **12**, 41 (2014).
139. Tu, Y. C. et al. Genetic requirements for *Klebsiella pneumoniae*-induced liver abscess in an oral infection model. *Infect. Immun.* **77**, 2657–2671 (2009).
140. Bulger, J., MacDonald, U., Olson, R., Beanan, J. & Russo, T. A. Metabolite transporter PEG344 is required for full virulence of hypervirulent *Klebsiella pneumoniae* strain hvKP1 after pulmonary but not subcutaneous challenge. *Infect. Immun.* **85**, e00093-17 (2017).
141. Xie, Y. et al. Emergence of the third-generation cephalosporin-resistant hypervirulent *Klebsiella pneumoniae* due to the acquisition of a self-transferable bla_{DHA-1}-carrying plasmid by an ST23 strain. *Virulence* **9**, 838–844 (2018).
142. Surgers, L., Boyd, A., Girard, P. M., Arlet, G. & Decré, D. ESBL-producing strain of hypervirulent *Klebsiella pneumoniae* K2, France. *Emerg. Infect. Dis.* **22**, 1687–1688 (2016).
143. Shen, D. et al. Emergence of a multidrug-resistant hypervirulent *Klebsiella pneumoniae* of ST23 with a rare $bla_{CTX-M-24}$-harboring virulence plasmid. *Antimicrob. Agents Chemother.* **63**, e02273–e02318 (2019).
144. Dong, N., Lin, D., Zhang, R., Chan, E. W. C. & Chen, S. Carriage of bla_{KPC-2} by a virulence plasmid in hypervirulent *Klebsiella pneumoniae*. *J. Antimicrob. Chemother.* **73**, 3317–3321 (2018).
145. Dong, N. et al. Genome analysis of clinical multilocus sequence Type 11 *Klebsiella pneumoniae* from China. *Microb. Genomics* **4**, https://doi.org/10.1099/mgen.0.000149 (2018).
146. Weingarten, R. A. et al. Genomic analysis of hospital plumbing reveals diverse reservoir of bacterial plasmids conferring carbapenem resistance. *mBio* **9**, e02011–e02017 (2018).
147. Sherry, N. L. et al. Genomics for molecular epidemiology and detecting transmission of carbapenemase-producing Enterobacterales in Victoria, Australia, 2012–2016. *J. Clin. Microbiol.* **57**, e00573-19 (2019).
148. Brisse, S. & Verhoef, J. Phylogenetic diversity of *Klebsiella pneumoniae* and *Klebsiella oxytoca* clinical isolates revealed by randomly amplified polymorphic DNA, *gyrA* and *parC* genes sequencing and automated ribotyping. *Int. J. Syst. Evol. Microbiol.* **51**, 915–924 (2001).
149. Brisse, S., Passet, V. & Grimont, P. A. D. 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. demonstration that *Klebsiella singaporensis* is a junior heterotypic synonym of *Klebsiella variicola*. *Int. J. Syst. Evol. Microbiol.* **64**, 3146–3152 (2014).
150. Rosenblueth, M., Martínez, L., Silva, J. & Martínez-Romero, E. *Klebsiella variicola*, a novel species with clinical and plant-associated isolates. *Syst. Appl. Microbiol.* **27**, 27–35 (2004).
151. Long, S. W. et al. Whole-genome sequencing of a human clinical isolate of the novel species *Klebsiella quasivariicola* sp. nov. *Genome Announc.* **5**, e01057-17 (2017).
152. Blin, C., Passet, V., Touchon, M., Rocha, E. P. C. & Brisse, S. Metabolic diversity of the emerging pathogenic lineages of *Klebsiella pneumoniae*. *Env. Microbiol.* **19**, 1881–1898 (2017).
This study is the first exploration of metabolic variation among diverse *K. pneumoniae* lineages and other members of the species complex, showing that metabolic capabilities — in particular, carbon substrate utilization — can vary substantially between strains.
153. Potter, R. F. et al. Population structure, antibiotic resistance, and uropathogenicity of *Klebsiella variicola*. *mBio* **9**, e02481-18 (2018).
154. Mathers, A. J. et al. *Klebsiella quasipneumoniae* provides a window into carbapenemase gene transfer, plasmid rearrangements, and patient interactions with the hospital environment. *Antimicrob. Agents Chemother.* **63**, e02513–e02518 (2019).
155. Brinkac, L. M. et al. Emergence of New Delhi metallo-β-lactamase (NDM-5) in *Klebsiella quasipneumoniae* from neonates in a Nigerian hospital. *mSphere* **4**, e00685-18 (2019).
156. Rodríguez-Medina, N., Barrios-Camacho, H., Duran-Bedolla, J. & Garza-Ramos, U. *Klebsiella variicola*: an emerging pathogen in humans. *Emerg. Microbes Infect.* **8**, 973–988 (2019).
157. Breurec, S. et al. Liver abscess caused by infection with community-acquired *Klebsiella quasipneumoniae* subsp. *quasipneumoniae*. *Emerg. Infect. Dis.* **22**, 529–531 (2016).
158. Brisse, S. & Van Duijkeren, E. Identification and antimicrobial susceptibility of 100 *Klebsiella* animal clinical isolates. *Vet. Mirobiol.* **105**, 307–312 (2005).
159. Martínez-Romero, E. et al. Genome misclassification of *Klebsiella variicola* and *Klebsiella quasipneumoniae* isolated from plants, animals and humans. *Salud Publica. Mex.* **60**, 52–62 (2018).
160. Ramirez, M. S., Iriarte, A., Reyes-Lamothe, R., Sherratt, D. J. & Tolmasky, M. E. Small *Klebsiella pneumoniae* plasmids: neglected contributors to antibiotic resistance. *Front. Microbiol.* **10**, 2182 (2019).
161. Carattoli, A. et al. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).
162. Robertson, J. & Nash, J. H. E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb. Genom.* **4**, https://doi.org/10.1099/mgen.0.000206 (2018).
163. Orlek, A. et al. Ordering the mob: insights into replicon and MOB typing schemes from analysis of a curated dataset of publicly available plasmids. *Plasmid* **91**, 42–52 (2017).
164. Arredondo-Alonso, S., Willems, R. J., van Schaik, W. & Schürch, A. C. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb. Genom.* **3**, e000128 (2017).
165. De Maio, N. et al. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000294 (2019).
166. Villa, L. et al. Diversity, virulence and antimicrobial resistance of the KPC-producing *Klebsiella pneumoniae* ST307 clone. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000110 (2017).
167. Yeh, K.-M. et al. Revisiting the importance of virulence determinant *magA* and its surrounding genes in *Klebsiella pneumoniae* causing pyogenic liver abscesses: exact role in serotype K1 capsule formation. *J. Infect. Dis.* **201**, 1259–1267 (2010).
168. Yu, W.-L., Lee, M.-F., Tang, H.-J., Chang, M.-C. & Chuang, Y.-C. Low prevalence of *rmpA* and high

tendency of *rmpA* mutation correspond to low virulence of extended spectrum β-lactamase-producing *Klebsiella pneumoniae* isolates. *Virulence* **6**, 162–172 (2015).
169. Martinez, J., Martinez, L., Rosenblueth, M., Silva, J. & Martinez-Romero, E. How are gene sequence analyses modifying bacterial taxonomy? The case of *Klebsiella*. *Int. Microbiol.* **7**, 261–268 (2004).
170. Ejaz, H. et al. Phylogenetic analysis of *Klebsiella pneumoniae* from hospitalized children, Pakistan. *Emerg. Infect. Dis.* **23**, 1872–1875 (2017).
171. Cerqueira, G. C. et al. Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks. *Proc. Natl Acad. Sci. USA* **114**, 1135–1140 (2017).
172. Moradigaravand, D., Martin, V., Peacock, S. J. & Parkhill, J. Evolution and epidemiology of multidrug-resistant *Klebsiella pneumoniae* in the United Kingdom. *mBio* **8**, e01976-16 (2017).
173. Liu, L. et al. Carbapenem-resistant isolates of the *Klebsiella pneumoniae* complex in western China: the common ST11 and the surprising hospital-specific types. *Clin. Infect. Dis.* 67 S263–S265 (2018).
174. Ocampo, A. M. et al. A two-year surveillance in five Colombian tertiary care hospitals reveals high frequency of non-CG258 clones of carbapenem-resistant *Klebsiella pneumoniae* with distinct clinical characteristics. *Antimicrob. Agents Chemother.* **60**, 332–342 (2015).
175. Andrade, L. N. et al. Virulence genes, capsular and plasmid types of multidrug-resistant CTX-M(-2, -8, -15) and KPC-2-producing *Klebsiella pneumoniae* isolates from four major hospitals in Brazil. *Diagn. Microbiol. Infect. Dis.* **91**, 164–168 (2018).
176. Deleo, F. R. et al. Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc. Natl Acad. Sci. USA* **111**, 4988–4993 (2014).
177. Lowe, M. et al. *Klebsiella pneumoniae* ST307 with $bla_{OXA-181}$, South Africa, 2014–2016. *Emerg. Infect. Dis.* **25**, 739–747 (2019).
178. Chung The, H. et al. A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of *Klebsiella pneumoniae*. *EMBO Mol. Med.* **7**, 227–239 (2015).

Acknowledgements
The authors are supported by Monash University and the Viertel Foundation of Australia. The authors thank Ryan Wick for assistance in creating the phylogenetic tree shown in Fig. 1.

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Microbiology thanks J. Pitout, T. Russo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Supplementary information
Supplementary information is available for this paper at https://doi.org/10.1038/s41579-019-0315-1.

RELATED LINKS
BIGSdb: https://bigsdb.pasteur.fr/klebsiella/klebsiella.html
Kleborate: https://github.com/katholt/Kleborate
PathogenWatch: https://pathogen.watch/