Supplementary note 1 - Analysing the data via the microreact project at:

https://microreact.org/project/Wastewater

Columns B - CX include output from Kleborate v0.4.0-beta, including species, MLST, virulence loci and antimicrobial resistance genes. Column CY lists the number of resistance genes identified in the isolates using Abricate v0.9.8 with the ResFinder database (downloaded 29 April 2020) using the threshold of >80 % nucleotide identity and coverage. The locations of these genes are detailed in columns CZ - FE, using output from MOB-suite and mlplasmids v1.0.0; ‘chromosomal’, ‘ambiguous’ (MOB-suite and mlplasmids predictions were discordant) or plasmid (when MOB-suite and mlplasmids both called ‘plasmid’, the accession number returned by MOB-suite is shown, or unknown in the case of novel plasmids). Columns FF - GE show replicon types identified using Abricate with the PlasmidFinder database (downloaded 13 Jan 2020) with a threshold of >80 % nucleotide identity and coverage. Columns GF - JG are presence (‘1’)/ absence (‘-‘) of plasmids identified by MOB-suite identified by accession number (as above). Columns JH and JI show the number of virulence factors identified using Abricate with the virulence factors database (downloaded 19 April 2020) with thresholds of both >40 and >80 % nucleotide identity and coverage respectively. Columns JJ - NG show the presence of these virulence factors in the isolates as identified using thresholds of >40 % (‘40’) or >80 % (‘80’) nucleotide identity and coverage. Detailed instructions on how to use microreact are available at https://microreact.org/instructions.

Supplementary note 2 - linked resistance genes on multiple plasmids in Kpne ST983

In all nine ST983 isolates, the resistance genes \( bla_{CTX-M-15}, bla_{TEM-1b}, aph(3'')-Ib, sul2 \) and \( aph(6)-Id \) genes are on homologous contigs of 9 - 14 kb, which also harbour genes for IS1380 family transposase \( IS\text{E}cp1 \) and transposon \( Tn3 \) resolvase. In the longer contigs these are flanked by an \( IS110 \) family transposase \( IS5075 \) and \( Tn3 \) family transposase \( Tn2 \). The resistance genes \( aac(6')-Ib-cr \) and \( bla_{OXA-1} \) are on another homologous contig with a gene encoding chloramphenicol acetyltransferase in all isolates, and also carries \( dfrA14 \) in one isolate (5Rd). Resistance gene \( aac(3)-llb \) is on a homologous contig in eight of the isolates, with an \( IS3 \) family transposase \( ISKpn11 \) and \( tmrB \) which encodes a tunicamycin resistance protein. Similarly, \( dfrA14 \) is on a homologous contig in these eight isolates, alongside an \( IS6 \) family transposase \( IS600 \) and flanked in the longer contigs by \( IS6 \) family transposase \( IS26 \). When aligned to CP021953 and JX424423, two of the resistance plasmids predicted by MOB-suite to harbour all of these resistance genes in four of these isolates, all of the contigs from all of the ST983 isolates align to the plasmid over an approximately 30 kb region, interspersed with insertion sequences (data not shown). These arrangements likely explain the mobility of these resistance genes. The ST983 isolates in a previous study from South Africa [1] show a similar profile of resistance genes with those in the present study (\( aph(6')-Id, aph(3'')-Ib, bla_{TEM-1b}, bla_{SHV-168}, \) \( bla_{SHV-38}, bla_{CTX-M-15}, qnrB6, oqxA, oqxB, fosA, sul2, tet(A), tet(C), dfrA14, qnrB1 \). Phylogenetic analysis confirmed the close relatedness between the ST983 reported from South Africa and those in the current study (Figure S14).

1. Founou RC, Founou LL, Allam M, Ismail A, Essack SY. Whole Genome Sequencing of Extended Spectrum beta-lactamase (ESBL)-producing Klebsiella pneumoniae Isolated from Hospitalized Patients in KwaZulu-Natal, South Africa. Sci Rep. 2019;9(1):6266. doi: 10.1038/s41598-019-42672-2
Supplementary note 3 – Other *Kpne* lineages

*Kpne* ST35 is a global multidrug-resistant clone [1]. Two isolates of ST35 were recovered from the WWTP (7Sl, 8Sd) and one from the river (River_C). These two isolates have identical plasmid and resistance gene profiles, and differ by 80 core SNPs over the alignment of 5,100,720 nucleotides. Both isolates harbour a single NC_021231_00058-like plasmid, which contains 6 resistance genes (*aac(3′)-Ia*, *aac(6′)-Ib-cr*, *bla*OXA-1, *bla*CTX-M-15, *dfrA14* and *qnrB1*; Figure 3), possess the same two replicon types (IncFII_1_pKP91 and IncFIB(K)_Kpn3; Figure S12), the same four chromosomal resistance genes *oqxA*, *oqxB*, *fosA6* and *bla*SHV-33 (Figure 3) and the virulence factor yersiniabactin (type ybt9; ICEKp3; as discussed below). The high identity between these two isolates is striking, given that one isolate was from the WWTP influent, and the other sampled directly from the river approximately 8 km upstream from the WWTP and 3 months earlier. This suggests that this multidrug-resistant clone is stable and relatively abundant in the local aquatic environment. The third ST35 isolate, 8Sd from the WWTP, is slightly divergent on the tree and exhibits a different plasmid and resistance gene profile.

*Kpne* ST13 is also a globally disseminated clone of clinical importance [2, 3, 4, 5]. This ST was represented by 3 hospital wastewater isolates, each of which contains two plasmids, a CP000823-like plasmid (which does not carry any resistance genes) and a CP021856(tig00000001_pilon)-like plasmid, which harbours *ant(3′)-Ia*, *dfrA1* and *sul1*. All three ST13 isolates possess the chromosomal resistance genes *oqxA*, *oqxB*, *fosA6* and *bla*SHV-101. Similarly, ST17 has been detected from global sources and is known to harbour multiple resistance genes [6, 7]. ST17 was also represented by 3 hospital isolates which show more heterogeneous plasmid and resistance profiles; a total of 5 plasmids were detected within these 3 isolates, but 4 of these were only present in only one isolate. Only a single plasmid-borne resistance gene was detected in ST17, *tet(D)*, but 7 chromosomal genes were noted. Finally, ST584 was assigned to three isolates from the WWTP sample. One of these isolates, SSf, harbours 5 plasmids including a JX843238(pTOR_02)-like plasmid [8] that carries *catA1* encoding resistance to chloramphenicol (the only unambiguous example of the presence of this gene on a plasmid in our data) and *sul1* genes. ST584 is not recognised as clinically significant, and is most notable for being recovered at high prevalence from wild boar and barbary macaques in Algeria [9].

1. Shen Z, Gao Q, Qin J, Liu Y, Li M. Emergence of an NDM-5-Producing Hypervirulent Klebsiella pneumoniae Sequence Type 35 Strain with Chromosomal Integration of an Integrative and Conjugative Element, ICEKp1. Antimicrob Agents Chemother. 2019;64(1). doi: 10.1128/AAC.01675-19

2. Marcade G, Brisse S, Bialek S, Marcon E, Leflon-Guibout V, Passet V, et al. The emergence of multidrug-resistant Klebsiella pneumoniae of international clones ST13, ST16, ST35, ST48 and ST101 in a teaching hospital in the Paris region. Epidemiol Infect. 2013;141(8):1705-12. doi: 10.1017/S0950268812002099

3. Cejas D, Elena A, Guevara Nunez D, Sevillano Platero P, De Paulis A, Magarinos F, et al. Changing epidemiology of KPC-producing Klebsiella pneumoniae in Argentina: Emergence of hypermucoviscous ST25 and high-risk clone ST307. J Glob Antimicrob Resist. 2019;18:238-42. doi: 10.1016/j.jgar.2019.06.005

4. Mairi A, Pantel A, Ousalem F, Sotto A, Touati A, Lavigne JP. OXA-48-producing Enterobacterales in different ecological niches in Algeria: clonal expansion, plasmid characteristics and virulence traits. J Antimicrob Chemother. 2019;74(7):1848-55. doi: 10.1093/jac/dkz146
5. Fursova NK, Astashkin EI, Gabrielyan NI, Novikova TS, Fedyukina GN, Kubanova MK, et al. Emergence of Five Genetic Lines ST395(NDM-1), ST13(OXA-48), ST3346(OXA-48), ST39(CTX-M-14), and Novel ST3551(OXA-48) of Multidrug-Resistant Clinical Klebsiella pneumoniae in Russia. Microb Drug Resist. 2020;26(8):924-33. doi: 10.1089/mdr.2019.0289

6. Ding Y, Wang Y, Hsia Y, Sharland M, Heath PT. Systematic review of carbapenem-resistant Enterobacteriaceae causing neonatal sepsis in China. Ann Clin Microbiol Antimicrob. 2019;18(1):36. doi: 10.1186/s12941-019-0334-9

7. Strydom KA, Chen L, Kock MM, Stoltz AC, Peirano G, Nobrega DB, et al. Klebsiella pneumoniae ST307 with OXA-181: threat of a high-risk clone and promiscuous plasmid in a resource-constrained healthcare setting. J Antimicrob Chemother. 2020;75(4):896-902. doi: 10.1093/jac/dkz550

8. Rahube TO, Viana LS, Koraimann G, Yost CK. Characterization and comparative analysis of antibiotic resistance plasmids isolated from a wastewater treatment plant. Front Microbiol. 2014;5:558. doi: 10.3389/fmicb.2014.00558

9. Bachiri T, Bakour S, Ladjouzi R, Thongpan L, Rolain JM, Touati A. High rates of CTX-M-15-producing Escherichia coli and Klebsiella pneumoniae in wild boars and Barbary macaques in Algeria. J Glob Antimicrob Resist. 2017;8:35-40. doi: 10.1016/j.jgar.2016.10.005
Figure S1. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 95 isolates analysed in this study constructed using an alignment of 184,671 core SNPs. Species and source of each isolate are indicated, with the location of resistance genes identified using Abricate with the Resfinder database. Contigs harbouring the resistance genes were classified as chromosomal or plasmid using MOB-suite and mlplasmids; contigs classified as plasmid by one method and chromosomal by another are marked as ‘ambiguous’. Where both methods assigned a contig to plasmid origin, the accession number of the plasmid as reported by MOB-suite is shown. Contigs assigned as plasmids but without a match to the database are assigned as “unknown”.
Figure S2. Distribution of resistance genes identified using Abricate with the Resfinder database between *Korn* and *Kpne*; not significantly different by a Wilcoxon test ($p = 0.059$).
Figure S3. Complete hierarchical clustering analysis (ward.D2) of predicted plasmids, resistance genes and replicon types in 95 isolates. Green refers to absence and pink to presence.
Figure S4. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the pOXA-48-like plasmids recovered in our study and 21 publicly available sequences (listed in Table S2). Two Korn isolates (6Rbi and 7Sd) separated by a single SNP indicated with a red circle.
Figure S5. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 95 isolates analysed in this study constructed using an alignment of 184,671 core SNPs. Species, source of each isolate and presence of plasmids identified using MOB-suite are indicated. A red asterisk indicates plasmids associated with resistance genes.
**Figure S6.** Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 95 isolates analysed in this study constructed using an alignment of 184,671 core SNPs. Species, source of each isolate and presence of replicon types identified using Abricate with the PlasmidFinder database are indicated.
Figure S7. Comparison of isolates from hospital drain and WWTP influent. (A) Number of plasmids as determined using MOB-suite; those that also not confirmed as plasmid by mlplasmids were classed as ambiguous, (B) number of replicon types identified by Abricate using the PlasmidFinder database. Wilcoxon test shows a significant difference in both cases.
Figure S8. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 25 Korn isolates in this study isolated from the hospital drain and WWTP influent showing the presence of replicon types identified using Abricate with the PlasmidFinder database.
Figure S9. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 25 Korn isolates from this study with isolates from WWTP in the East of England (n=2; Ludden et al., 2017) and unpublished sequences of isolates from Italy (n=25) and Pakistan (n=5). Isolates were aligned to SPARK_1625_C1 from Italy; the tree was constructed using an alignment of 6,004,523 core SNPs. The source of each isolate is indicated.
Figure S10. Alignment of the GU595196 and KY798506 reference plasmids with the contigs identified in our data as KY798506-like. In red are the $\text{bla}_{\text{KPC}}$ genes, in grey the $\text{bla}_{\text{TEM}}$ and in blue the $\text{bla}_{\text{SHV}}$. 
**Figure S11.** Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 59 *Kpne* isolates in this study isolated from the hospital wastewater, WWTP influent and river showing ST and the presence of plasmids as identified by MOB-suite. Contigs were also analysed using mlplasmids and any that were not identified as plasmid by both methods were omitted from this data. Plasmids associated with resistance genes are marked with a red asterisk.
Figure S12. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 59 Kpne isolates in this study isolated from the hospital drain, WWTP influent and river showing ST and the presence of replicon types identified using Abricate with the PlasmidFinder database.
Figure S13. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 9 *Kpne* ST983 isolates in this study isolated from the hospital drain and WWTP influent. The presence of plasmids identified using MOB-suite and mlplasmids, listed by accession number, is shown. The location of resistance genes (plasmid accession number / chromosomal / ambiguous) is shown; ambiguous was assigned to contigs when results from MOB-suite and mlplasmids did not agree. Plasmids associated with resistance genes are marked with an asterisk.
Figure S14. Approximate maximum likelihood phylogenetic tree (midpoint rooted) of the 59 *Kpne* isolates in this study with ED01500733 (ST983) from South Africa (accession number NZ_POWS00000000.1). The South African ST983 isolate differed from the ST983 isolates in the current study by between 180 and 208 core SNPs.
Figure S15. Phylogenetic reconstruction of the ybt locus identified in our samples (n=33). Blue leaves correspond to *Kpne* isolates and pink to *Korn*. 
Figure S16. Tanglegram linking the phylogenetic trees constructed using SNPs in the core genome (Left) and the ybt locus (Right). Both trees are midpoint rooted and include 33 isolates. The purple branches include *Kpne* ybt-positive isolates, while the green branches include *Korn* ybt-positive isolates. Lines have been drawn between tips in the trees representing the same isolate, while the tree branches were sorted to minimize the number of overlapping lines required. The lines are coloured by the leaves that are common in subtrees of both trees. The lines of the leaves that are not common in subtrees of both trees were left in grey. The leaves are coloured according to the maximum number of matching clusters the trees contain (k=5).