

Supporting materials

Longitudinal analysis within one hospital in sub-Saharan Africa over 20 years reveals repeated replacements of dominant clones of *Klebsiella pneumoniae* and stresses the importance to include temporal patterns for vaccine design considerations

*Eva Heinz^{1,2}, Oliver Pearse^{2,3}, Allan Zuza³, Sithembile Bilima³, Chisomo Msefula⁴, Patrick Musicha^{2,3}, Patriciah Siyabu⁵, Edith Tewesa⁵, Fabrice E Graf², Rebecca Lester^{3,6}, Samantha Lissauer^{3,7}, Jennifer Cornick^{3,7}, Joseph M Lewis^{2,3,7}, Kondwani Kawaza^{3,4}, Nicholas R Thomson^{8,9}, *Nicholas A Feasey^{2,3,10}

1. Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK
2. Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK
3. Malawi Liverpool Wellcome Programme, Kamuzu University of Health Sciences, Blantyre, Malawi
4. Kamuzu University of Health Sciences, Blantyre, Malawi
5. Queen Elizabeth Central Hospital, Blantyre, Malawi
6. Division of Infection & Immunity, University College London, UK
7. Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, Liverpool, UK
8. Parasites and Microbes Program, Wellcome Sanger Institute, Hinxton, UK
9. London School of Hygiene and Tropical Medicine, London, UK
10. School of Medicine, St Andrews University, UK

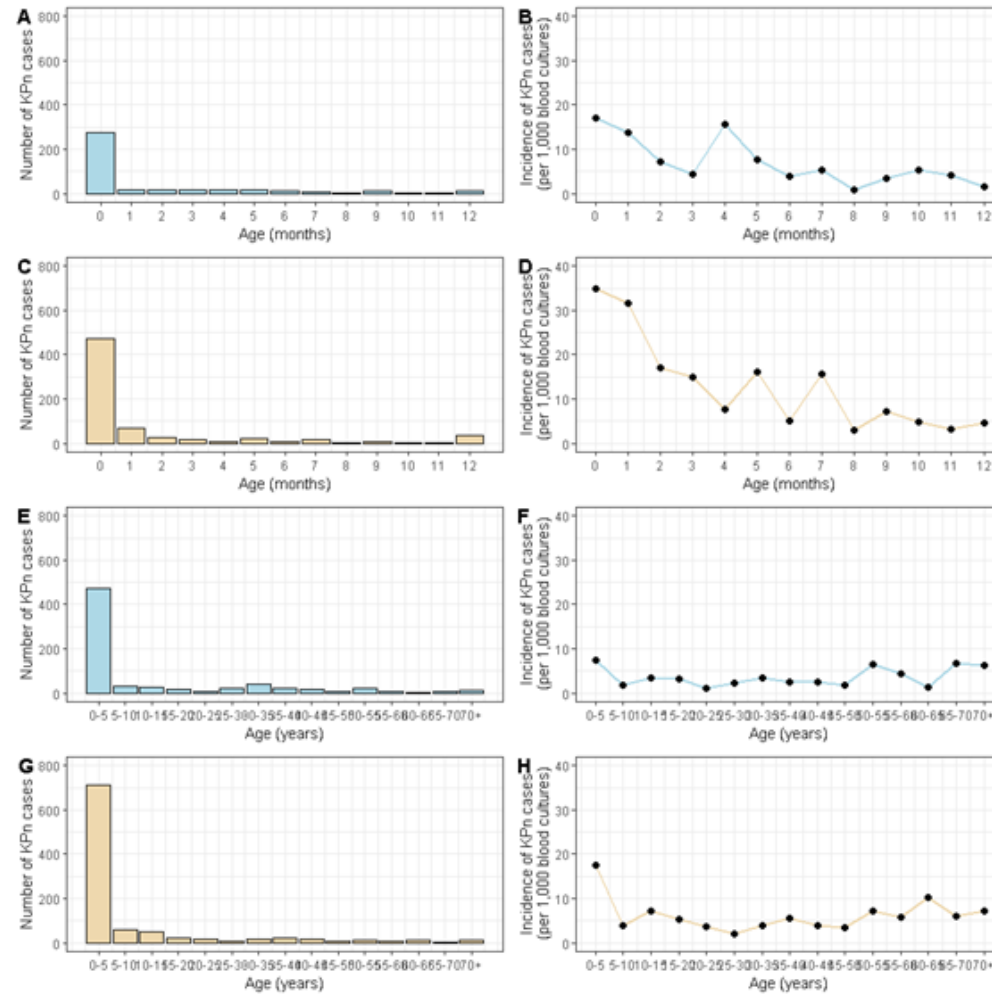


Fig S1. Age specific *Kpn* crude frequency and blood culture positivity rate. (A) Number of *Kpn* cases up to age 12 months (1998-2015). (B) Blood culture positivity rate (per 1,000 blood cultures) of *Kpn* cases up to age 12 months (1998-2015). (C) Number of *Kpn* cases up to age 12 months (2016-2021). (D) Blood culture positivity rate (per 1,000 blood cultures) of *Kpn* cases up to age 12 months (2016-2021). (E) Number of *Kpn* cases up to 100 years (1998-2015). (F) Blood culture positivity rate (per 1,000 blood cultures) of *Kpn* cases up to 100 years (1998-2015). (G) Number of *Kpn* cases up to 100 years (2016-2021). (H) Blood culture positivity rate (per 1,000 blood cultures) of *Kpn* cases up to 100 years (2016-2021).

Figure S2

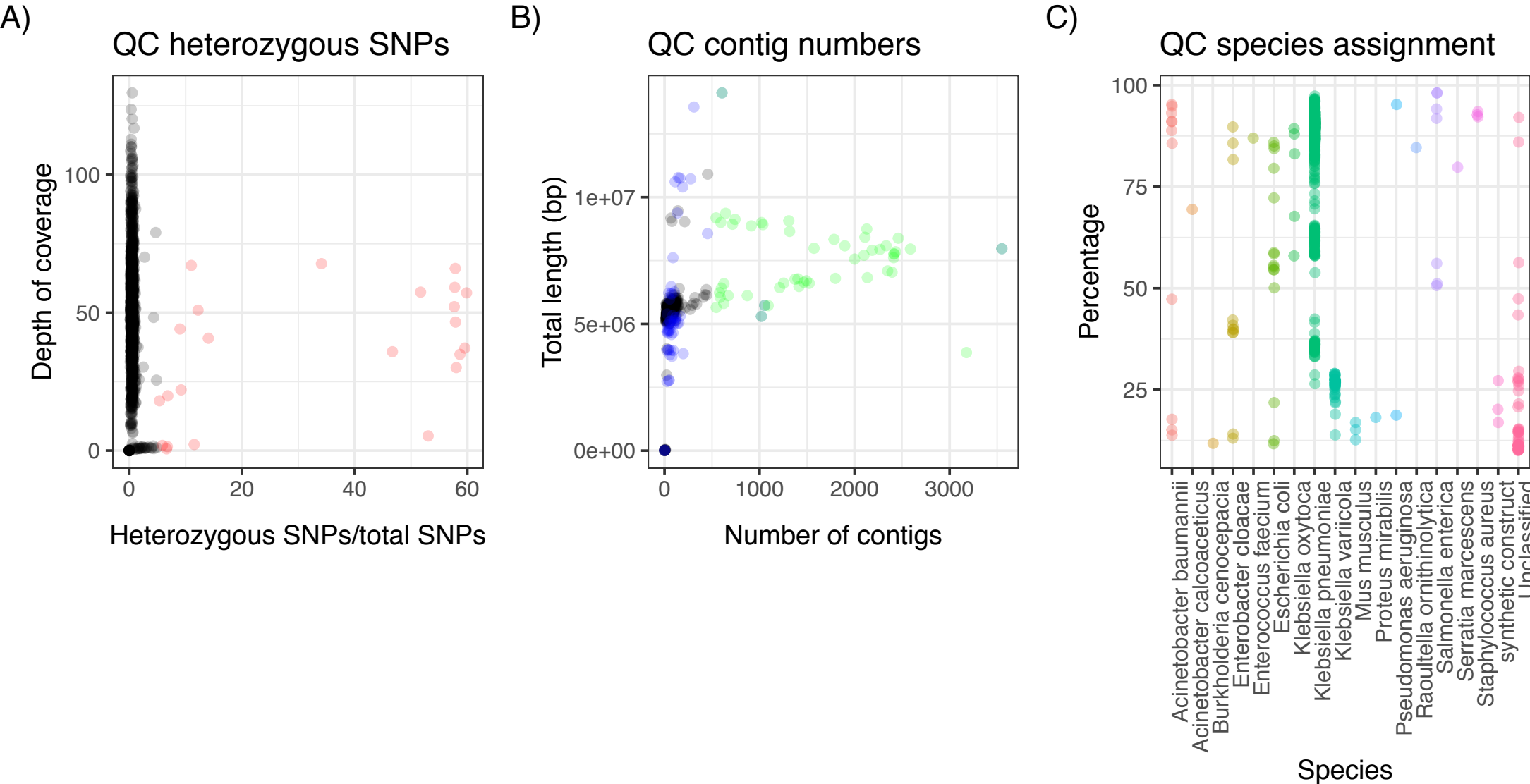


Fig S2. Quality control of sequencing data. The species assignment of the sequencing reads according to kraken are shown in panel **(A)**, where all isolates with more than 10% of a different species were removed. **(B)** The plot shows the distribution of homozygous vs heterozygous SNPs, isolates represented by red points (>5% heterozygous SNPs) were removed as likely consisting of two related strains. **(C)** Number of contigs vs total assembly length is shown; with isolates that passed QC in black, isolates that failed the species composition QC in blue, and isolates with >500 contigs in green, which were also removed.

Major STs per wards



Major wards

- 3B MALE MEDICAL WARD
- 4A FEMALE MEDICAL WARD
- A ETC
- BURNS UNITS
- CHATINKHA NURSERY
- INTENSIVE CARE UNIT
- PAEDIATRIC A&E
- PAEDIATRIC MEDICAL BAY/WARD
- PAEDIATRIC MOYO WARD
- PAEDIATRIC NURSERY
- PAEDIATRIC ONCOLOGY
- PAEDIATRIC SPECIAL CARE WARD
- PAEDIATRIC SURGICAL WARD
- PICU/HDU

Fig S3. Main ST isolates per wards over time. Each panel shows the number of isolates for the relevant ST as in the legend. The number of isolates per year are stratified by the major wards as in the figure legend.

A)

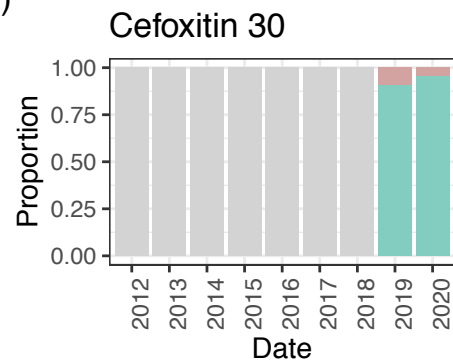
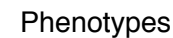
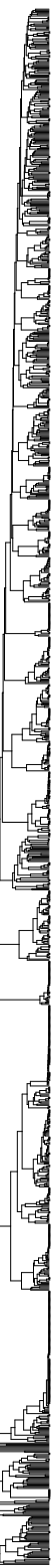


Fig S4. Phenotypic resistance profiles of all sequenced isolates. This highlights very high resistance against widely used first-line treatments like **(A)** cefpodoxime (ceftriaxone; 3GC), **(B)** augmentin (BLBLI combination), **(C)** gentamicin (AGly), **(D)** cotrimoxazole (TMT) but still a large proportion of sensitive isolates against less commonly used antimicrobials; **(E)** ciprofloxacin (Fq), **(F)** amikacin (AGly), and a pattern of increasing sensitivity against **(G)** chloramphenicol. The currently only available last-line treatments are **(H)** piperazillin-tazobactam (BLBLI combination), which show very high resistance levels; and **(I)** meropenem (carbapenem), the currently only viable alternative tested. **(J)** Cefoxitin also shows a lot of sensitive isolates; this antimicrobial is routinely used in diagnostic laboratories world-wide to test for *ampC* production as it was rapidly replaced with third-generation cephalosporins after its initial characterisation. The resistance profile is however of high interest as clinical trials have been started recently to assess it as a potential treatment alternative.



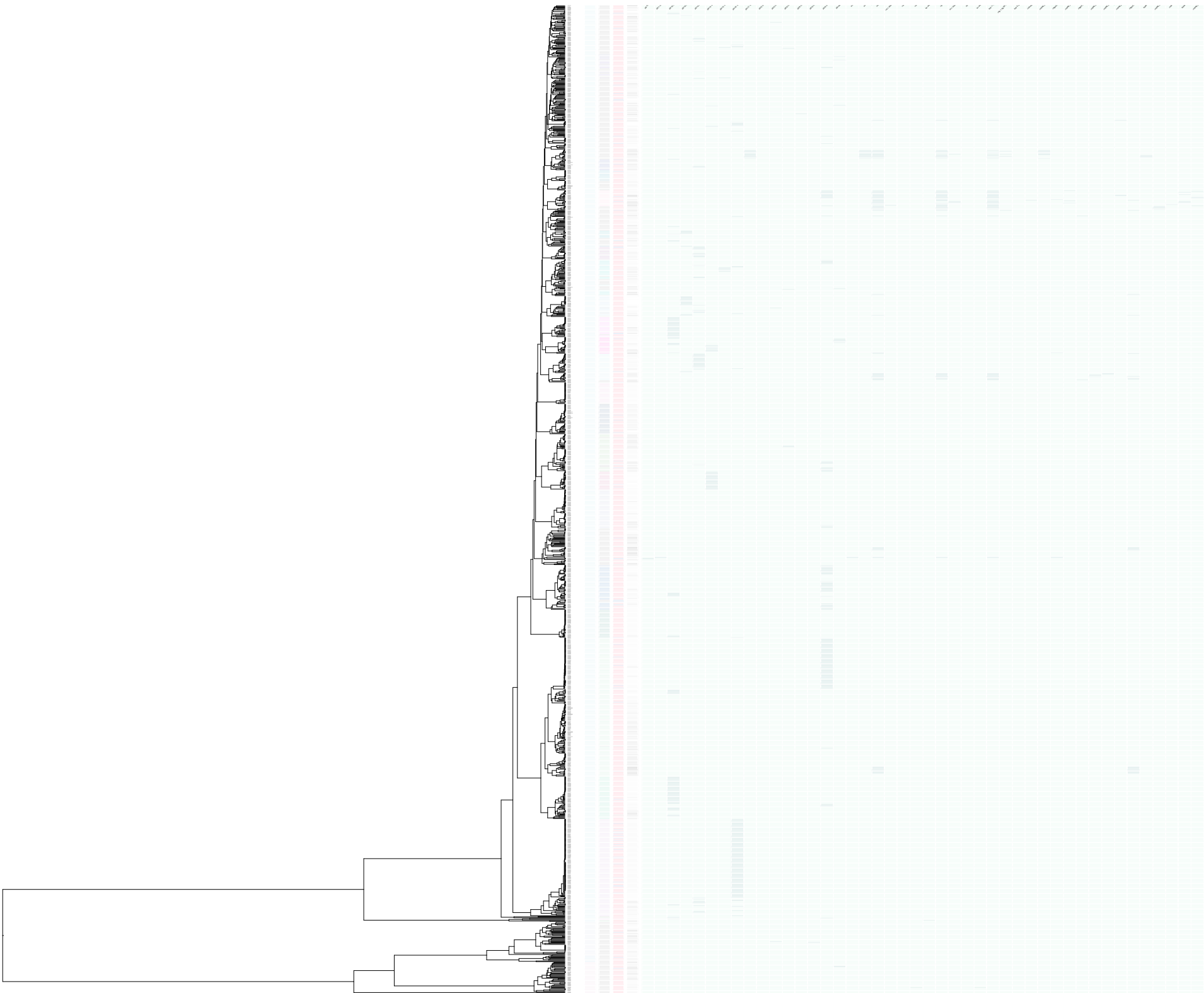
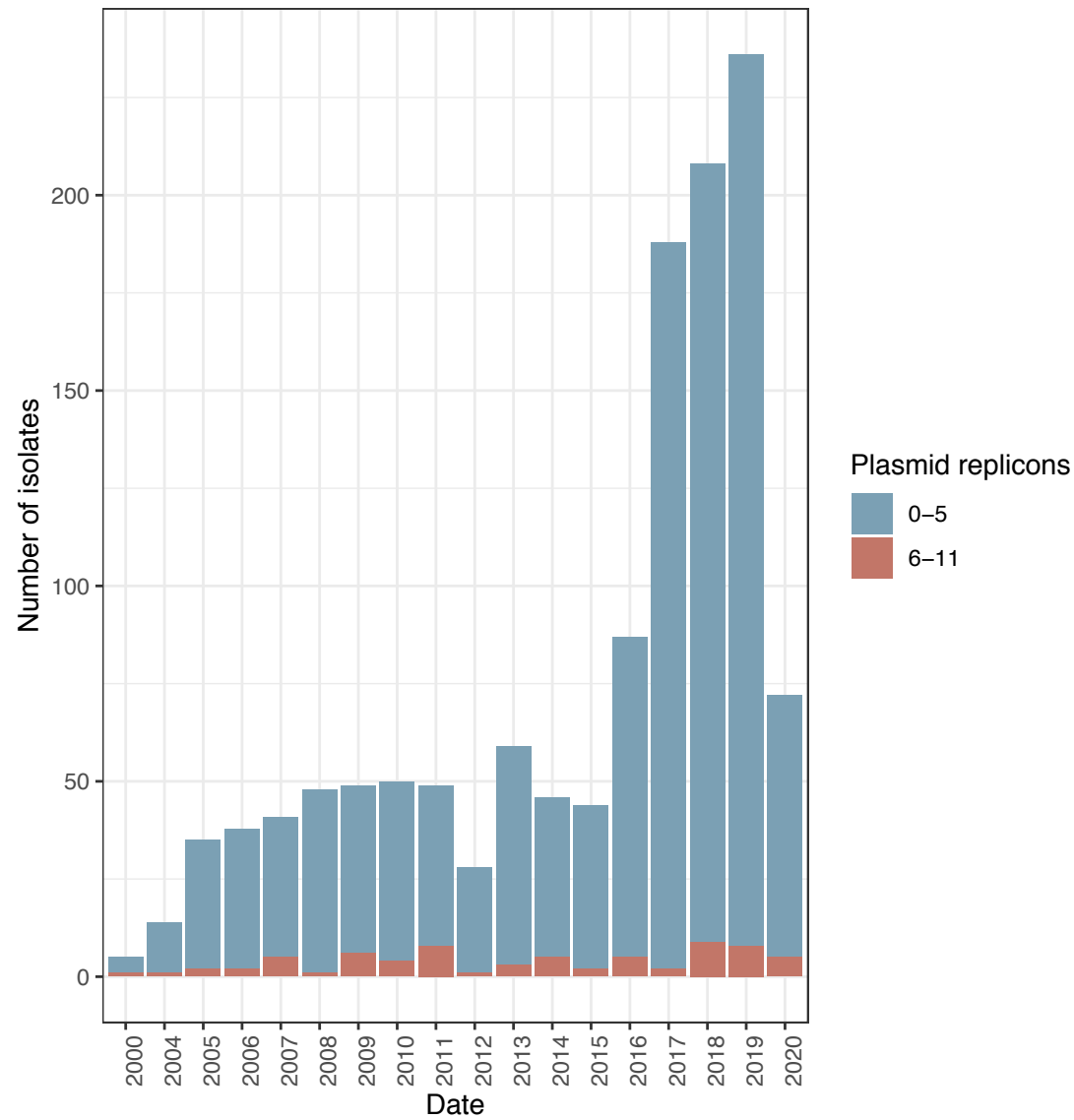




Fig S5. Predicted resistance and virulence genes and plasmid replicons. The guidance tree was calculated with iqtree based on the core gene alignment derived from panaroo. The metadata columns displayed are *Kpn* species, main ST, sample type and year of isolation; as well as heatmaps showing the presence (dark shade) or absence (light shade) of (A) resistance genes, (B) virulence genes and (C) plasmid replicons.

Figure S6

A) Plasmid replicons per isolate over time



B) AMR gene numbers per isolate over time

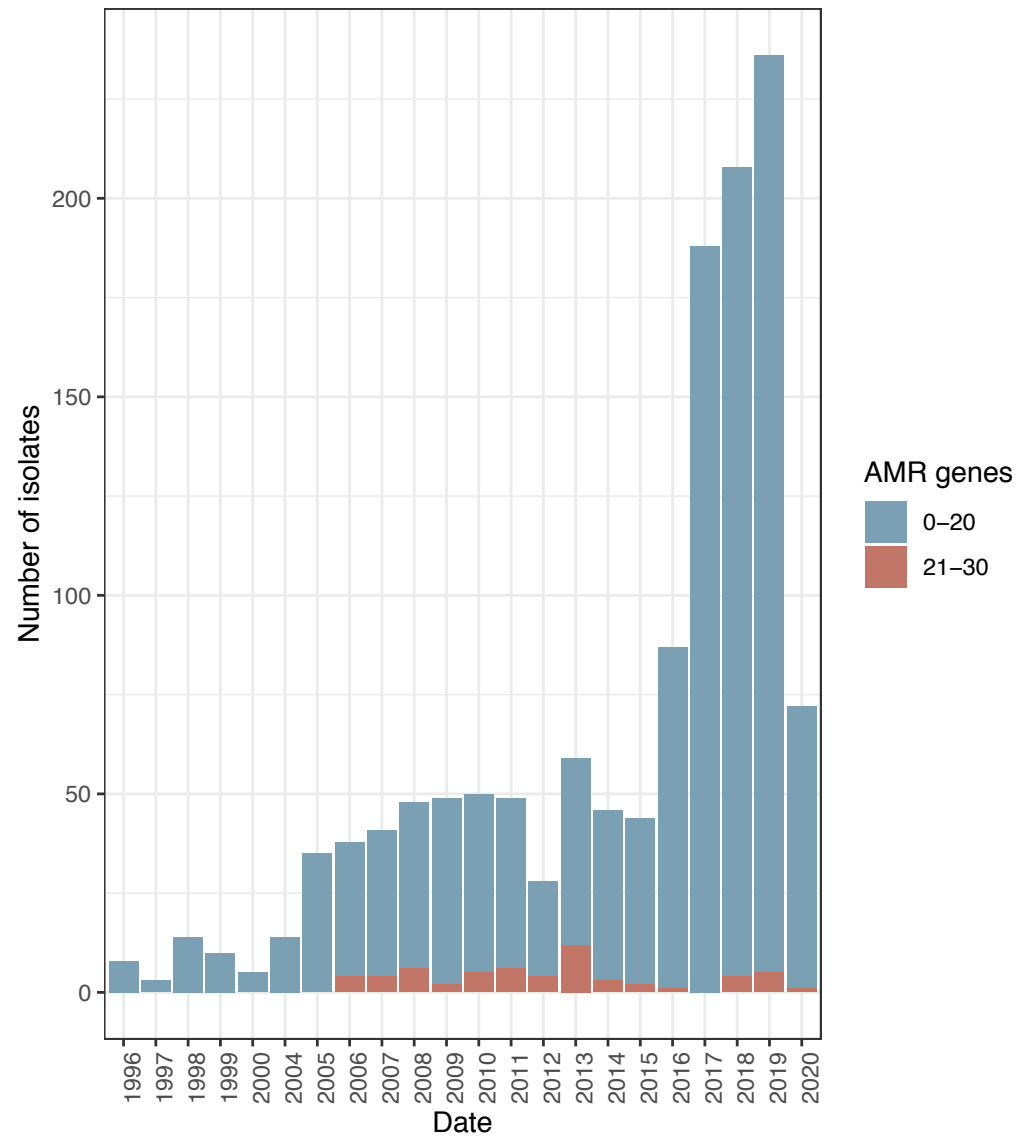


Fig S6. Plasmid replicons and resistance gene numbers. (A) *Kpn* isolates over time showing the number of plasmid replicons encoded, and (B) the number of isolates for numbers of resistance genes present over time.

Figure S7

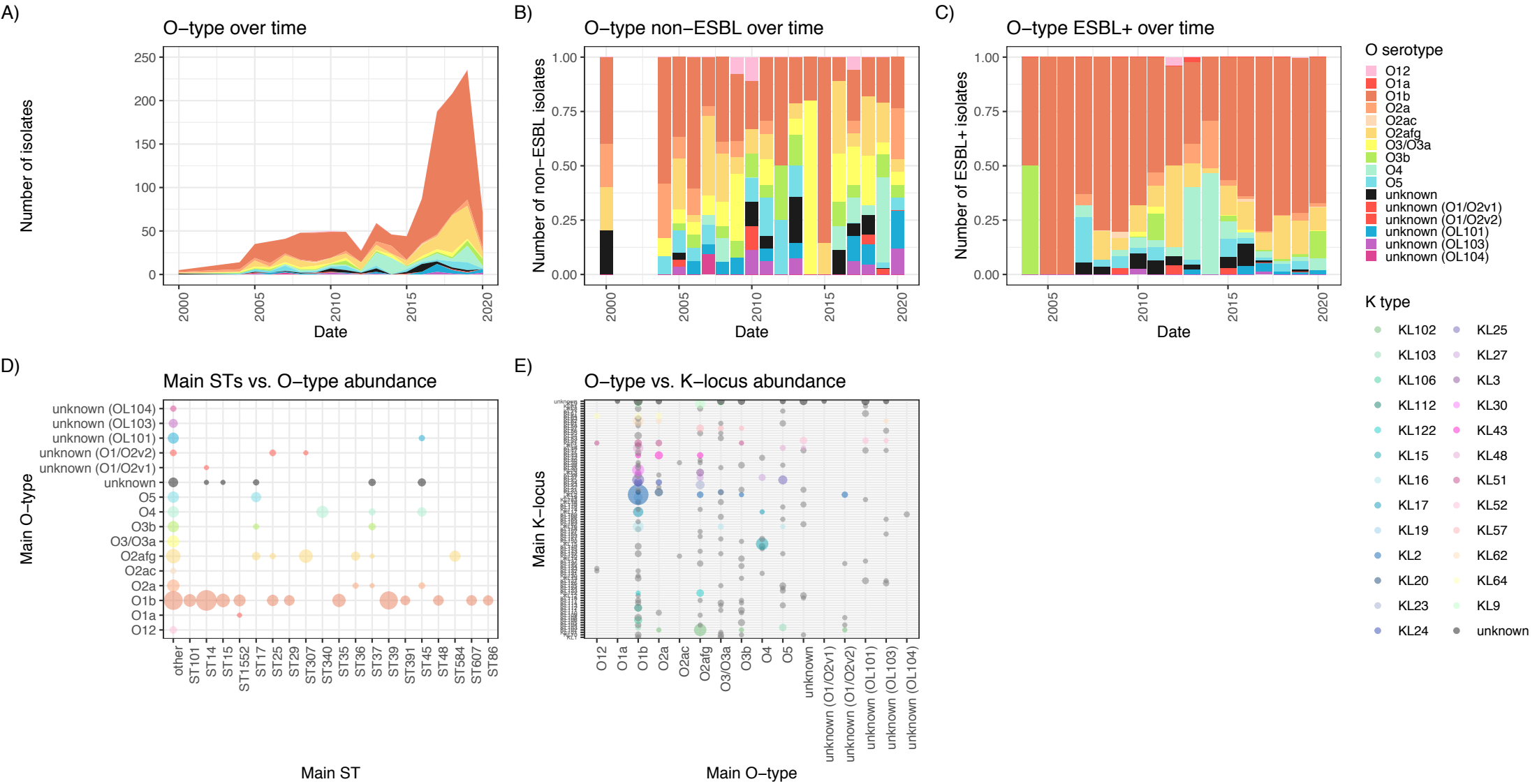
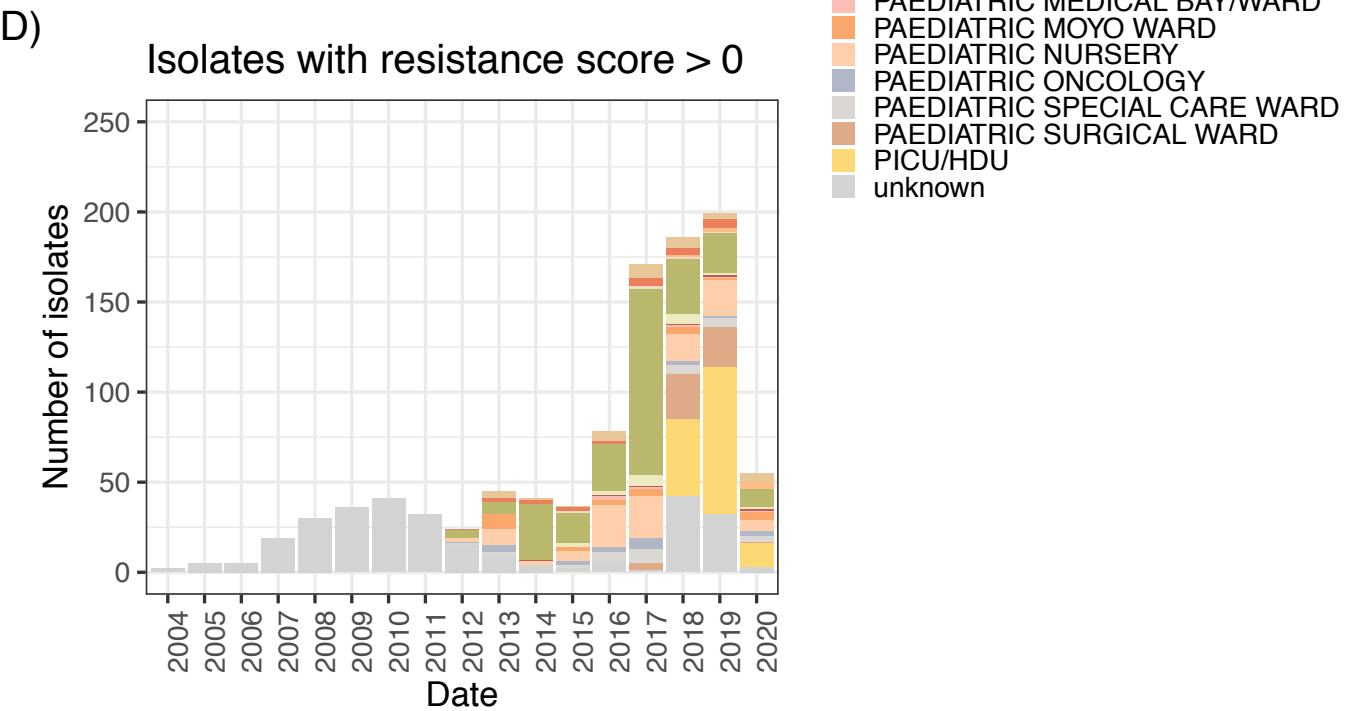
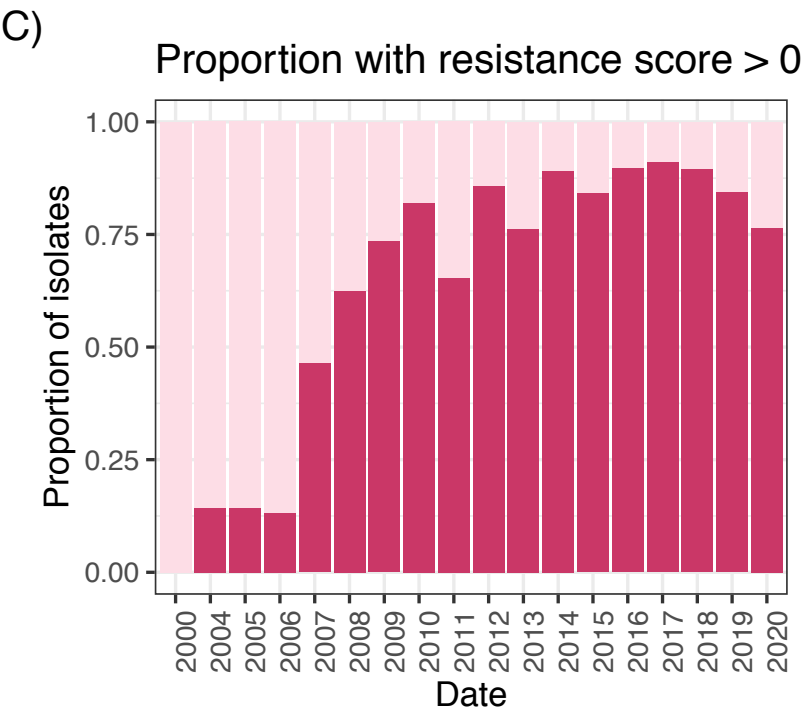
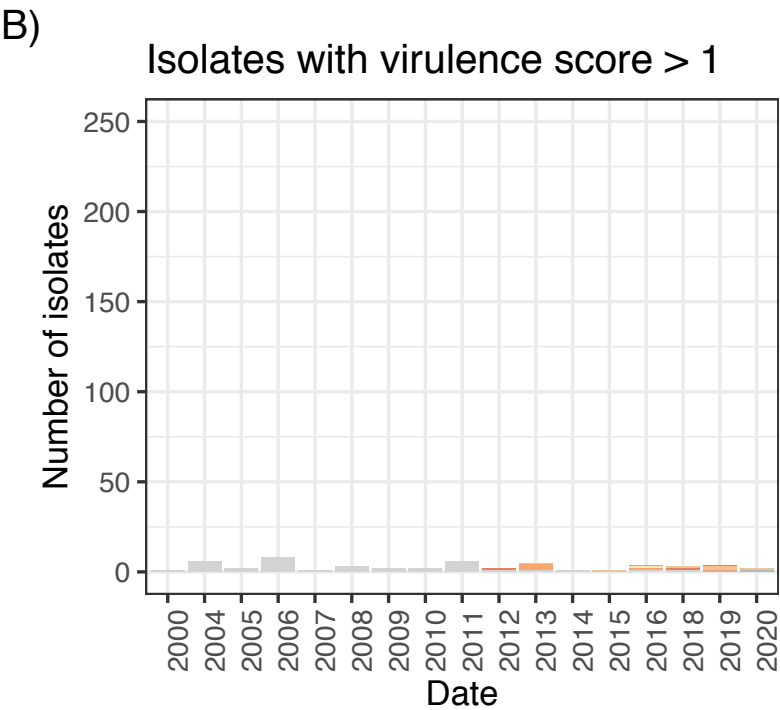
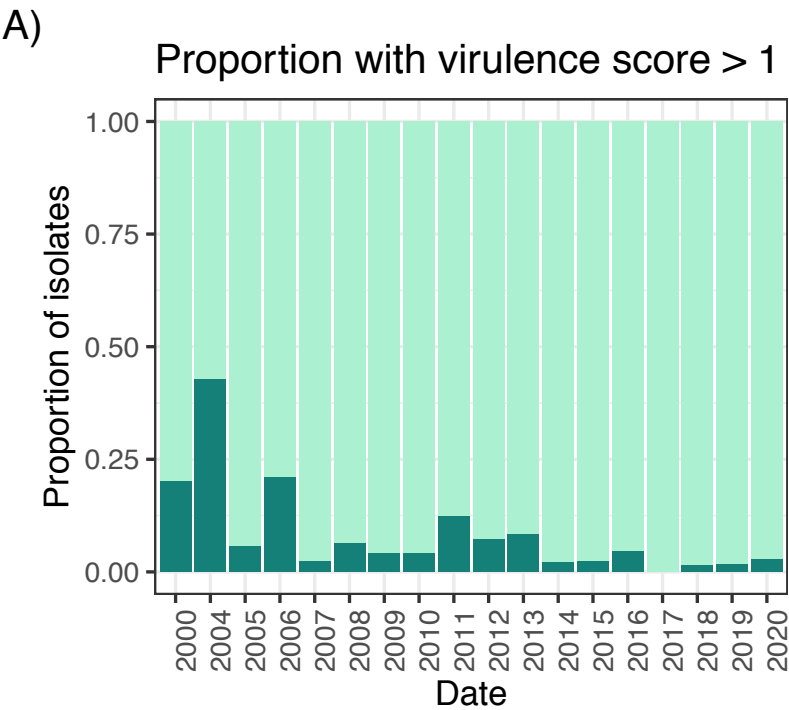


Fig S7. O-Ag serotypes over time and resistance profiles. The distribution of predicted O-Ag serotypes based on their genomic loci(24,27) over time for (A) all isolates, (B) non-ESBL and (C) ESBL-encoding isolates, and their distribution across (D) main STs and (E) K-locus types.

Figure S8



- Major wards
- 3B MALE MEDICAL WARD
 - 4A FEMALE MEDICAL WARD
 - A ETC
 - BURNS UNITS
 - CHATINKHA NURSERY
 - INTENSIVE CARE UNIT
 - PAEDIATRIC A&E
 - PAEDIATRIC MEDICAL BAY/WARD
 - PAEDIATRIC MOYO WARD
 - PAEDIATRIC NURSERY
 - PAEDIATRIC ONCOLOGY
 - PAEDIATRIC SPECIAL CARE WARD
 - PAEDIATRIC SURGICAL WARD
 - PICU/HDU
 - unknown

Fig S8. Proportions of virulent isolates and resistant isolates over time have markedly different pattern. (A) The number of isolates with a kleborate-predicted virulence score > 1 (1 usually indicates yersiniabactin which is chromosomally fixed in several STs) which remains very low and stable over time and found sporadically across various wards, whereas (B) isolates with resistance scores 1 or over (indicating ESBL or, in the case of score of 2, carbapenemase present) rapidly increase over time and biased towards several wards. This is reflected in the proportions; whilst isolates encoding several virulence factors remain at very low proportion over time (C), we can see the rapid increase of resistant isolates which remains at a very high level since the initial increase (D).

Table S1: Accessions, metadata and kleborate predictions per strain.

Table S2: Antimicrobial gene predictions using ariba.

Table S3: Plasmid replicon predictions using ariba.

Table S4: Phenotypic resistance profiles.

Table S5: O- and K-antigen predictions using updated definitions in kaptive.