**Genomic analysis of diversity, antimicrobial resistance and virulence of extended-spectrum beta-lactamase producing *Klebsiella pneumoniae* colonising adults in Blantrye, Malawi**

Joseph M. Lewis1,2,3,4, Madalitso Mphasa1, Rachel Banda1, Mathew Beale4, Eva Heinz2, Jane Mallewa5, Nicholas Thomson4, Nicholas A Feasey1,2

1 Malawi-Liverpool Wellcome Clinical Research Programme

2 Liverpool School of Tropcial Medicine

3 University of Liverpool

4 Wellcome Sanger Institute

5 College of Medicine, University of Malawi

6 University of Lancaster

Corresponding Author

Joseph M. Lewis

Department of Clinical Infection, Microbiology and Immunology

University of Liverpool

8 West Derby Street, Liverpool, L69 7BE

United Kingdom

[jmlewis@liverpool.ac.uk](mailto:jmlewis@liverpool.ac.uk)

+44 (0)151 795 9687

**Abstract**

**Background**

Klebsiella pneumoniae is an antimicrobial resistance (AMR) associated global pathogen but genomes from sub-Saharan Africa (sSA) are underrepresented in global sequencing efforts.

**Objectives**

We carried out a genomic investigation of extended-spectrum beta-lactamase (ESBL)-producing *K. pneumoniae* colonising adults in Blantyre, Malawi to describe diversity and AMR determinants, to place these isolates in a global genomic context, and to describe virulence determinants, comparing colonising to previously sequenced infecting isolates.

**Methods**

We carried out short-read whole-genome sequencing of 203 ESBL *K. pneumoniae* isolates. Presence of AMR genes, multilocus sequence type, K-locus, O-locus and virulence loci were determined. Diversity was compared to a global collection of 484 *K. pneumoniae* genomes with a core-gene phylogeny.

**Results**

Diversity of Malawian isolates reflects global Klebsiella diversity, and the commonest carriage STs correspond to locally and globally successful invasive and AMR-associated lineages. Recognised virulence loci are associated with invasive disease in Malawi but most invasive isolates lack these determinants. Virulence determinants are mostly associated with lack of ESBL genes but we find that lineages of ESBL-hypervirulent *K. pneumoniae* are present in Malawi.

**Conclusions**

Diversity of K. pneumoniae in Blantyre, Malawi reflects global diversity, suggesting unrestricted global spread, and the determinents of a successful pathogenic *K. pneumoniae lineage* are the same in Malawi as elsewhere. Host factors or alternate virulence loci to those sought are likely to be the major determinants of invasive disease in Malawi. Presence of AMR-hypervirulent lineages is a cause for concern and highlights a need for genomic AMR surveillance across sSA.

**Word count: 2624**

**Introduction**

*Klebsiella pneumoniae* is a human gut commensal and opportunistic pathogen which is often significantly associated with antimicrobial resistance (AMR), and has been identified by the World Health Organisation as a global priority pathogen1. It is a frequent cause of nosocomial infection2, and a common cause of neonatal sepsis worldwide3. In low resource settings such as many of the countries of sub-Saharan Africa (sSA), a lack of access to last-line antimicrobials can mean that AMR *K. pneumoniae* a significant therapeutic challenge. In Blantyre, Malawi, for example, 91% of *K.* *pneumoniae* are now resistant to third-generation cepahlosporins (3GC), largely mediated through production of extended-spectrum beta-lactamases (ESBLs) and are often untreatable with locally available antimicrobials4.

Global high-throughput sequencing has provided significant insight into the biology of *K. pneumoniae*. For example, organisms classified as *K. pneumoniae* by standard phenotypic testing are now recognised to be multiple species which form the *K. pneumoniae* species complex5; global AMR spread has been linked to clonal expansion of AMR-associated high-risk clones6; and genomic loci associated with virulence7 (including the hypermucoid phenotype8) have been identified. Historically, antimicrobial resistance and virulence were associated with different *K. pneumoniae* populations, but so-called convergence leading to AMR-hypervirulent clones is increasingly described, especially in South and South-East Asia9. However, genomes from sSA are under-represented in global collections and it may be that the genomic epidemiology of *K. pneumoniae* differs in this setting. We therefore present the results of a genomic analysis of *K. pneumoniae* from a study of colonisation with ESBL Enterobacterales in Blantyre, Malawi, with three aims: to describe the diversity and AMR determinants of *K. pneumoniae* in this setting; to place these isolates in a global genomic context; and to describe virulence determinants, comparing colonising to previously sequenced infecting isolates.

**Methods**

The isolates analysed in this study were selectively cultured from stool and rectal swabs collected from adults in Blantyre, Malawi, as part of a study of longitudinal carriage of ESBL-producing Enterobacterales, as previously described[REF carriage paper]. Briefly, 225 adults (≥ 16 years) with sepsis were recruited in the emergency department of Queen Elizabeth Central Hospital (QECH), Blantyre, Malawi, along with 100 adults who were planned to be admitted to QECH but with no plans for antimicrobial administration and 100 community dwelling adults with no antimicrobial exposure (except for longterm co-trimoxazole preventative therapy, CPT, or antituberculous chemotherapy) in the previous four weeks. Up to five stool samples (or rectal swab samples performed by trained study team members if participants were unable to provide stool) were collected over the course of six months and selectively aerobically cultured overnight at 37ºC on ChromAGAR ESBL-selective chromogenic media (ChromAGAR, France) before being speciated with the API system (Biomeriuex, France).

A subsample of isolates identified as *K. pneumoniae* underwent DNA extraction and sequencing: one *K. pneumoniae* colony pick from the first 217 samples where *K. pneumoniae* was identified. DNA was extracted from overnight nutrient broth cultures using the Qiagen DNA mini kit (Qiagen, Germany) as per the manufacturer’s instructions. DNA was shipped at ambient to the Wellcome Sanger Institute for paired-end short-read 150bp sequencing on the Illumina HiSeq X10 instrument. Species was confirmed with Kraken v0.10.6 and Braken v1.010 *De novo* assembly was undertaken with SPAdes v3.14.011 and the quality of the assemblies assessed with QUAST v5.0.212; assembly failures with a total assembled length of < 1Mb or assemblies with a QUAST-defined contamination of ≥ 10% were excluded from further analysis. Assemblies were annotated with Prokka v1.5 with a genus-specific database from RefSeq13 and the Roary v1.007 pangenome pipeline14 used to identify core genes. Genes present in ≥ 99% samples considered to be core, and were concatenated to a core gene pseudosequence. Variable sites were identified with SNP-sites v2.4.115 and used to infer a maximum-likelihood phylogenetic tree in IQ-TREE v1.6.316 with the ModelFinder module used to select the best fitting nucleotide substitution model, using ascertainment bias correction. Trees were visualized in the R *ggtree* v2.2.417package.

ARIBA v.2.12.118 was used to identify AMR-associated genes using the SRST2 curated version of the ARG-ANNOT database19 on the raw reads, and was used to call SNPs in the quinolone-resistance determining regions (QRDR) *gyrA, gyrB, parC* and *parE*, after downloading wild-type genes from the *Escherichia coli* K-12 substr. MG1655 (NC\_000913.3) reference genome. For the purposes of describing aggregate quinolone resistance, any mutation in any QRDR was considered to confer quinolone resistance. Beta lactamases were considered to be extended spectrum based on the phenotypic classifications at <https://ftp.ncbi.nlm.nih.gov/pathogen/betalactamases/Allele.tab>. ARIBA was also used to determine multilocus sequence type (ST) as defined by the 7-gene Pasteur scheme20 hosted at pubMLST (<https://pubmlst.org/>). Kleborate v2.0.121 was used to infer Klebsiella species, capsule polysaccharide (K-type) and lipopolysaccharide (O-type) serotypes, and to identify the presence of the siderophore virulence loci *ybt* (yersiniabactin), *iuc* (aerobactin)and *iro* (salmochelin)*,* the genotoxin locus *clb* (colibactin)*,* and the hypermucoidy genes *rmpA* and *rmpA2*.

To place the isolates from this study in a local and global context, two further phylogenies were constructed, using collections of Malawian and global isolates. Two studies from our centre in Blantyre provided previously published genomes: a genomic investigation into a *K. pneumoniae* outbreak on the neonatal ward at QECH22 which sequenced 100 bloodstream infection isolates from children between 2012 and 2015, and a study which sequenced 72 infecting (bacteremia and meningitis) and rectal carriage isolates from QECH from 1996-2014, which were selected to maximise diversity23. A phylogeny for Malawian isolates was inferred as described above, using the same nucleotide substitution model as was selected for the isolates from this study. To build a global phylogeny we included all Malawian isolates plus 66 genomes from a study of the genomic epidemiology of Klebsiella in Kenya24 and 288 genomes from a global description of Klebsiella population structure5, again using the same methods. ST and ESBL presence or absence was inferred using the methods described above.

All statistical analyses were carried out in R v4.0.2 (R Foundation for Statistical Computing, Vienna, Austria). Summary statistics, where presented, are medians and interquartile ranges or proportions with exact binomial confidence intervals unless otherwise stated. Comparisons of proportions use Fisher’s exact test. The clinical study which provided the isolates for this analysis was approved by the Liverpool School of Tropical Medicine (16-062) and Malawi College of Medicine (P.11/16/2063) research ethics committees. All sequenced isolates have been deposited in the European Nucleotide Archive, and accession numbers (as well as accession numbers of publicly available genomes used in this analysis) are provided as supplementary data.

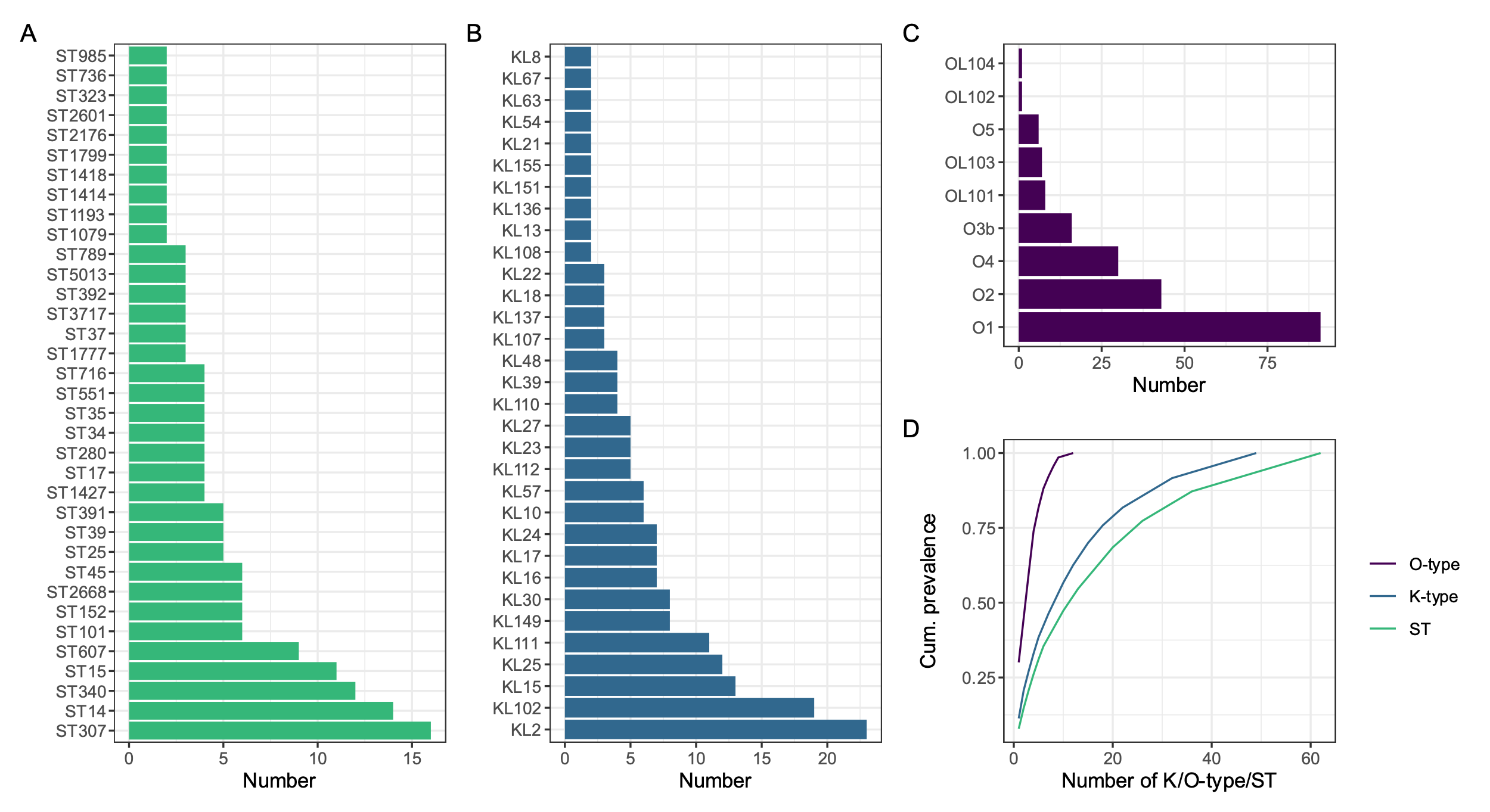
**Results**

Following quality control, 203 *K. pneumoniae* genomes form this study were available for analysis. These were largely *Klebsiella. pneumoniae sensu stricto* (n=190) but *Klebsiella quasipneumoniae subsp. Similipneumoniae* (n=7*), Klebsiella variicola subsp. Variicola* (n=3), *Klebsiella quasipneumoniae subsp. quasipneumoniae* (n=2) and *Klebsiella quasivariicola* (n=1) were also identified. We first described diversity in terms of inferred ST, K- and O-types (Figure 1). The commonest STs identified were those that show well-recognised associations with invasive disease and AMR: ST307, ST14, ST340, and ST15. There was significant diversity in ST with 61 STs identified and a median 2 (IQR 1-4) samples per ST; 26/61 STs were only represented once in the collection. There was also significant K-type diversity with 49 K-types identified. K-types were ST-associated with median 1 (IQR 1-2) STs per K-type. O-types were less diverse, with 12 identified, and they were more likely to be associated with multiple STs: each O-type was associated with a median 4.5 (IQR 1.75 – 10.5) STs. The largest three O-types accounted for 81% of samples whereas the largest three K-types and STs accounted for 27% and 22% of samples, respectively. The best fitting nucleotide substitution model to the core gene pseudosequence alignment as identified by IQ-TREE was the general time reversible model with FreeRate site heterogeneity with 8 parameters; the STs reflected the structure of the inferred phylogeny well (Supplementary Figure 1), which, as expected, had the characteristic Klebsiella deep branching morphology.

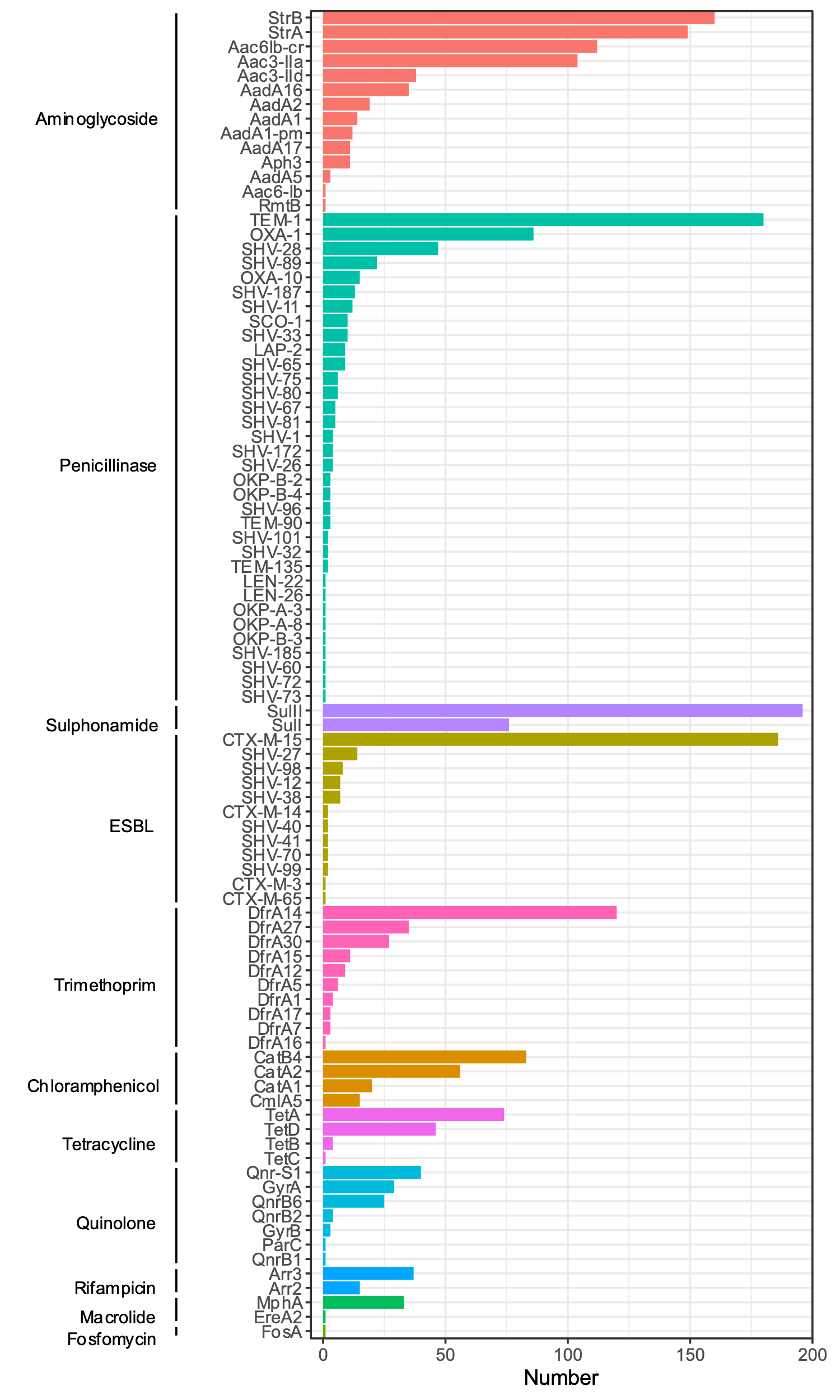
The isolates contained a median of 15 (IQR 11.5-17) antimicrobial resistance genes per genome (Figure 2). At least one ESBL-encoding gene was identified in 99% (200/203) isolates, as were genes encoding narrow spectrum beta lactamases(200/203, 99% isolates). The most commonly identified ESBL-encoding gene was *bla*CTX-M-15 in 186/203 (92%) of isolates. Genes conferring resistance to sulphonamides (201/203, 99%), trimethorprim (198/203, 98%) and aminoglycosides (198/203 98%) were near-ubiquitous; determinants of resistance to chloramphenicol (140/203, 69%) and quinolones (87/203, 43%) were less common. No genes conferring carbapenemase resistance were identified, though these have been previously identified in *E. coli* in Malawi25.

We next placed the carriage isolates in a local and global context. Following quality control, 150 genomes from other studies in Malawi, 65 from Kenya, and 264 global genomes were included, in addition to the 203 from this study: 687 genomes in total (see Supplementary Data). The Roary pan-genome pipeline identified 49,385 genes, of which 2754 were core; these formed a 0.95Mb pseudosequence with 200,622 variable sites, which were used to infer a phylogenetic tree (Figure 3). The Malawian isolates were distributed throughout the tree, showing that Malawian human carriage and infection isolates sample global Klebsiella diversity. The commonly identified Malawian STs in this study were overrepresented in the Malawian as compared to global collections, forming tabletops in the tree topography: ST14, ST340 and ST307. ST15 also formed a tabletop but this included both global and Malawian isolates.

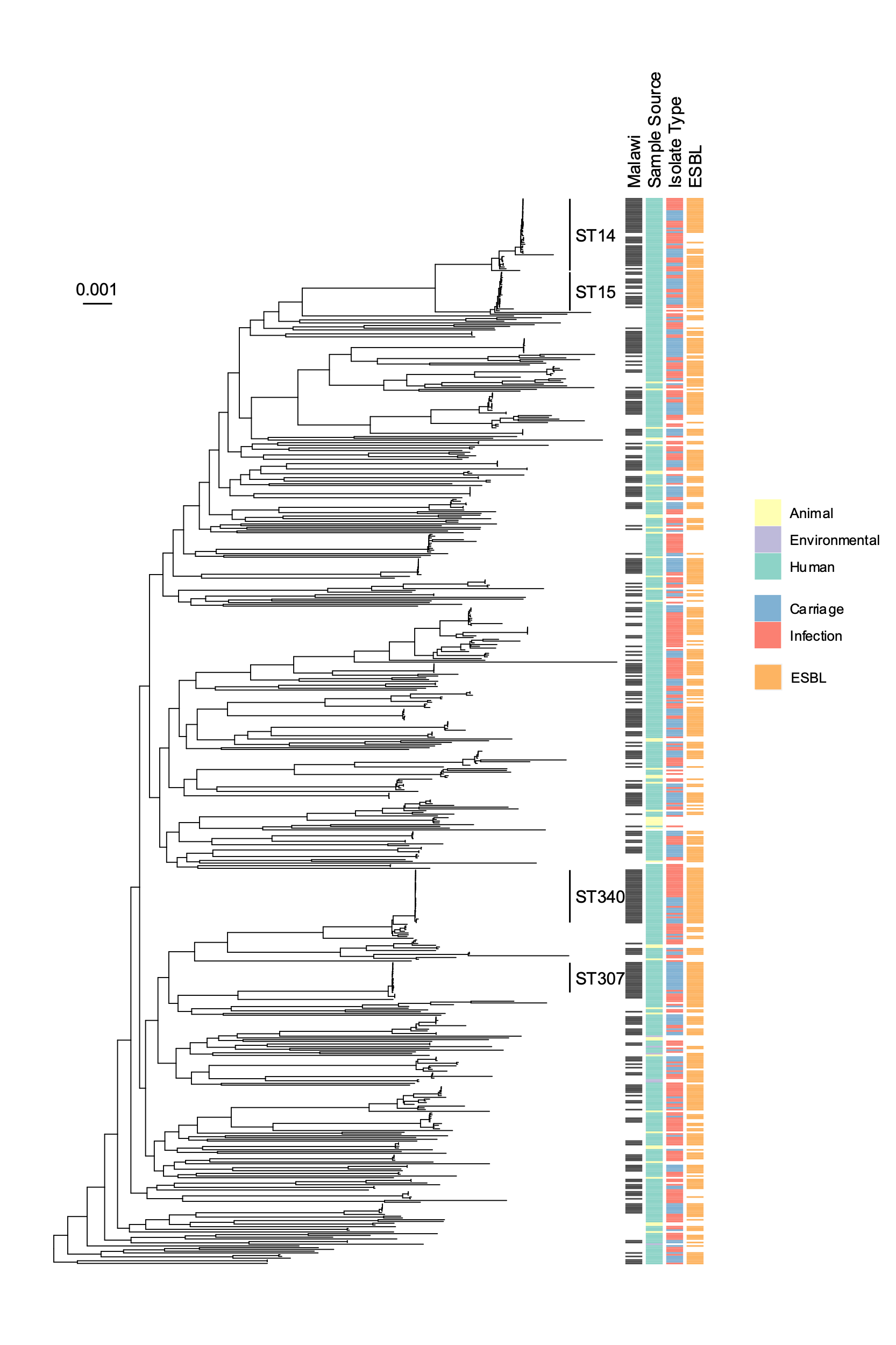
We further explored the presence and associations of a number of recognized virulence determinants in these Malawi-associated isolates (Figure 4), and their association with infection or carriage. Most commonly identified was the siderophore locus *ybt*, present in 27% (93/365) of Malawian isolates, and more commonly in carriage (68/214, 32%) compared to invasive (26/139, 19%, p = 0.006) isolates, and in ESBL isolates (89/306, 29%) compared to non-ESBL isolates (5/47, 11%, p = 0.007). All other virulence determinants were less commonly identified (the siderophore loci *iuc* in 5% and *iro* in 3%; the genotoxin loci *clb* in 3%; and the hypermucoidy genes *rmpA* and *rmpA2* in 5% and 3% respectively), but were strongly associated with invasive isolates: only 1/214 (0.5%) carriage isolate contained any of these virulence genes (Supplementary Figure 2), compared to 18/139 of invasive isolates (12.9%, p < 0.001). Generally, isolates containing these non-*ybt* virulence determinants were less likely to contain ESBL-encoding genes (11/47 [23%] non-ESBL isolates vs 8/306 [3%], p < 0.001) but two lineages contained both: ST268 (with *ybt, clb, iro,iuc* and *rmpA2* +/- *rmpA*) and ST218 (with *iuc, iro* and *rmpA2* +/- *rmpA*) also all carried ESBL-encoding genes. These ST218/268 isolates were all from blood culture, the earliest in 2004; no isolates of these STs types were identified in any carriage samples. They did not carry the classical virulence-associated K loci (K1 or K2), but rather all ST218 were identified as K-locus KL57 and ST268 as KL20.



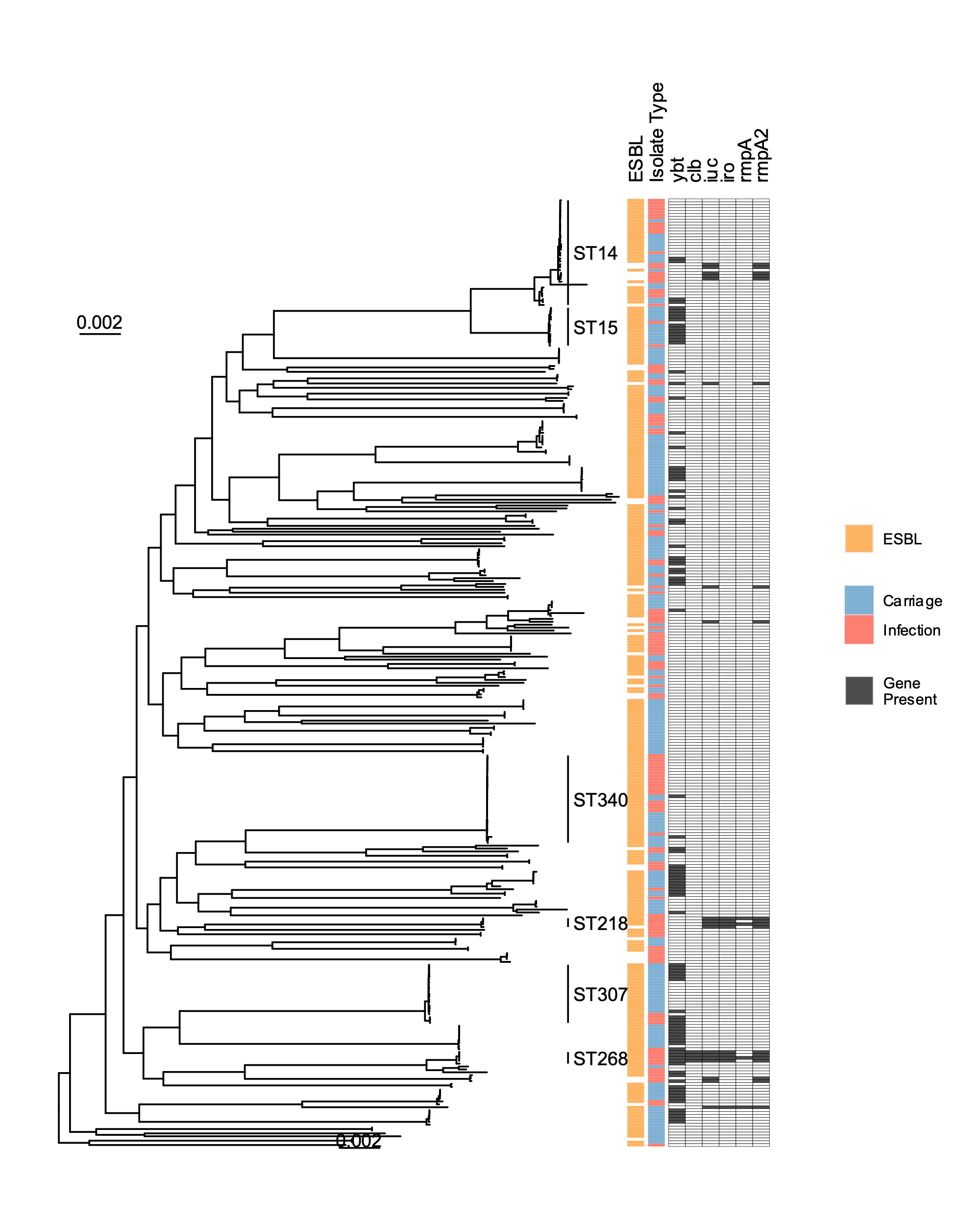
**Figure 1:** Diversity of K. pneumoniae included in the study. Distributions of (A) sequence type (ST) (B) K-type (C) O-type. (D) shows cumulative prevalence as a function of number of K/O-types or ST, where K/O-type or ST is ordered from largest to smallest.



**Figure 2:** Antimicrobial resistance determinants identified.



**Figure 3:** Midpoint rooted core-gene phylogenetic tree of Malawian and global isolates, restricted to *Klebsiella pneumoniae* sensu stricto (KpI). Heatmaps show whether isolated in Malawi, whether animal/human/environmental, whether carriage or infection, and whether ESBL genes are present. Scale bar shows nucleotide substitutions per site.



**Figure 4:** Midpoint rooted core-gene maximum-likelihood phylogenetic tree of Malawian isolates, including all genomes from this study and context genomes from Malawian studies (n = 353), and restricted to *Klebsiella pneumoniae* *sensu stricto* (KpI). The Roary pan-genome pipeline identified 20,853 genes in these genomes, of which 3391 were core, forming a 2.82Mb alignment with 378,596 variable sites which were used to infer the tree. Heatmaps show whether ESBL genes are present, whether carriage or infection, and whether the siderophore virulence loci *ybt* (yersinobactin), *iuc* (aerobactin), *iro* (salmochein), the genotoxin virulence locus *clb* (colibactin) and the hypermucoidy genes *rmpA* and *rmpA2* are present. Scale bar shows nucleotide substitutions per site.

**Discussion**

We present a genomic investigation of 203 colonising ESBL-producing *K. pneumoniae* species complex genomes from Blantyre, Malawi, an area of the world that is undersampled in global *K. pneumoniae* collections. We can draw several conclusions that increase our understanding of the dynamics of colonisation and infection in this setting.

Largely, Malawian colonising ESBL *K. pneumoniae* reflect global *K. pneumoniae* diversity, as previously described for invasive isolates,23 and suggest unrestricted transmission of strains on a global scale. Some STs are overrepresented in the Malawian as compared to the global collection, representing successful infecting and colonising isolates in the Malawian setting; largely, the common Malawian infecting STs are also common Malawian colonising STs, consistent with the hypothesis that the source of infecting *K. pneumoniae* is the host microbiota26. These successful STs are well-described global invasive and AMR-associated lineages: ST14, ST15, ST3076, and ST34027, suggesting that their overrepresentation in the Malawian collections is not evidence of geographic restriction but reflects the sampling scheme of the global collection, in which isolates were selected to represent the full diversity of the species. Their frequency of isolation in Malawi suggests that the determinants of success for pathogenic *K. pneumoniae* lineages are similar in the Malawian setting as they are globally.

There was significant diversity in the Malawian collection in K- and O-types; though K-types tended to be ST-associated, single O-types were associated with multiple ST and K-types. Given interest in the development of vaccines against multidrug-resistant *K. pneumoniae*28, these data could inform putative vaccine targets. AMR-gene content was also diverse, though *bla*CTX-M-15 was by far the most commonly identified ESBL-encoding gene, present in 92% of isolates. The near-ubiquity of trimethoprim and sulphonamide resistance determinants may be related to the WHO recommendation for widespread lifelong co-trimoxazole preventative therapy for people living with HIV29. No genes conferring resistance to carbapenemases were identified, though the carbapenemase *bla*NDM-5 has been described in *E. coli* in Blantyre25, and the carbapenemases *bla*KPC-2 and *bla*OXA-48 in ST340 K. pneumoniae and *K. variicola*, respecively in Central Malawi30. Carbapenems are at best sporadically available in QECH, and it seems very likely that increasing exposure would result in rapid expansion of carbapenemases in *K. pneumoniae*.

We found some evidence that the well described virulence loci *clb, ioc, iuc* and hypermucoidy genes *rmpA* ad *rmpA2* are, as expected, associated with invasive disease in the Malawian collection. However, though presence of these loci was associated with invasive disease, the majority of invasive isolates lacked any recognised virulence determinants, suggesting that host factors or other pathogen factors are the primary determinant of whether carriage develops into infection in the Malawian setting. The siderophore locus *ybt* was unexpectedly associated with carriage rather than invasive isolates; though this could be due to confounding brought about by the sampling scheme (carriage isolates were enriched for ESBL-producers so *ybt-*ESBL associations could introduce confounding), it highlights the importance of understanding the genomic epidemiology of *K. pneumoniae* across diverse settings. Generally, presence of virulence determinants were associated with absence of ESBL-encoding genes. However, two lineages (ST218 and ST268) demonstrated the presence of both. This phenomenon – convergence of AMR and virulence - is increasingly described in South and South East Asia9, and potentially of significant public health concern for Malawi. ST26831 and ST21832 are have been described as hypervirulent clones in China (and hypervirulent ESBL producers in the case of ST268) but the genomic epidemiology of Klebsiella virulence and AMR in sSA is poorly described. Given our findings, it is likely that hypervirulent-AMR *K. pneumoniae* are already circulating in Malawi, and across sSA, and reiterates the need to for scale up of genomic AMR surveillance across sSA. Neither ST268 nor ST218 was represented in the carriage collection. This could be related to sampling of rare lineages (only six ST218/268 in total were present in the collection) or could relate to the biology of these hypervirulent lineages, and a preponderance for invasive disease rather than colonisation.

There are limitations to our study. Most importantly, our sampling scheme is not unbiased. ESBL-producing carriage isolates were selected for, and one of the Malawian studies providing invasive context genomes was an investigation of a K. pneumoniae outbreak on the QECH neonatal unit. This is likely to have introduced bias into the collection of Malawian genomes. All Malawian genomes are from a single centre, which may limit generalisation to other settings in sSA. Multiple samples were cultured from single individuals and so were not independent, which could introduce bias.

In conclusion, we present a genomic analysis of ESBL K pneumoniae colonising adults in Blantyre, Malawi. Diversity of isolates refects global Klebsiella diversity, and the most commonly identified STs correspond to globaly successful invasive and AMR-associated lineages. This suggests that Blantyre is sampling the global Klebsiella diversity; that the determinents of successful pathogenic lineages are similar in Malawi as elsewhere; and provides support for the hypothesis that *K. pneumoniae* carriage precedes infection. We find that lineages of ESBL-hypervirulent *K. pneumoniae* are present in Malawi, a finding with public health implications which highlights the need for expanded genomic surveillance of AMR across sSA.

**References**

1. World Health Organisation. Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis. (2017).

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. Musicha, P. et al. Trends in antimicrobial resistance in bloodstream infection isolates at a large urban hospital in Malawi (1998–2016): a surveillance study. The Lancet Infectious Diseases **17**, 1042–1052 (2017).

5. Holt, K. E. et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. PNAS **112**, E3574–E3581 (2015).

6. Wyres, K. L., Lam, M. M. C. & Holt, K. E. Population genomics of Klebsiella pneumoniae. Nat Rev Microbiol **18**, 344–359 (2020).

7. Lam, M. M. C. et al. Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in Klebsiella pneumoniae populations. Microbial Genomics, **4**, e000196 (2018).

8. Walker, K. A. et al. A Klebsiella pneumoniae Regulatory Mutant Has Reduced Capsule Expression but Retains Hypermucoviscosity. mBio **10**, (2019).

9. Wyres, K. L. et al. Genomic surveillance for hypervirulence and multi-drug resistance in invasive Klebsiella pneumoniae from South and Southeast Asia. Genome Medicine **12**, 11 (2020).

10. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology **15**, R46 (2014).

11. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology : a journal of computational molecular cell biology **19**, 455–77 (2012).

12. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. Bioinformatics **29**, 1072–1075 (2013).

13. Seemann, T. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**, 2068–2069 (2014).

14. Page, A. J. et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics **31**, 3691–3693 (2015).

15. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microbial Genomics **2**, (2016).

16. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution **32**, 268–274 (2015).

17. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution **8**, 28–36 (2017).

18. Hunt, M. et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microbial genomics **3**, e000131 (2017).

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

20. Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. D. & Brisse, S. Multilocus Sequence Typing of Klebsiella pneumoniae Nosocomial Isolates. Journal of Clinical Microbiology **43**, 4178–4182 (2005).

21. Lam, M. M. C. et al. Genomic surveillance framework and global population structure for Klebsiella pneumoniae. bioRxiv 2020.12.14.422303 (2021) doi:10.1101/2020.12.14.422303.

22. Cornick, J. et al. Genomic investigation of a suspected multi-drug resistant Klebsiella pneumoniae outbreak in a neonatal care unit in sub-Saharan Africa. bioRxiv 2020.08.06.236117 (2020) doi:10.1101/2020.08.06.236117.

23. Musicha, P. et al. Genomic analysis of Klebsiella pneumoniae isolates from Malawi reveals acquisition of multiple ESBL determinants across diverse lineages. Journal of Antimicrobial Chemotherapy **74**, 1223–1232 (2019).

24. 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).

25. Lewis, J. M. et al. Emergence of carbapenemase producing Enterobacteriaceae, Malawi. J Glob Antimicrob Resist **20**, 225–227 (2019).

26. Gorrie, C. L. et al. Gastrointestinal Carriage Is a Major Reservoir of Klebsiella pneumoniae Infection in Intensive Care Patients. Clinical Infectious Diseases **65**, 208–215 (2017).

27. Safavi, M. et al. Global genotype distribution of human clinical isolates of New Delhi metallo-β-lactamase-producing Klebsiella pneumoniae; A systematic review. J Glob Antimicrob Resist **23**, 420–429 (2020).

28. Choi, M., Tennant, S. M., Simon, R. & Cross, A. S. Progress towards the development of Klebsiella vaccines. Expert Review of Vaccines **18**, 681–691 (2019).

29. World Health Organisation. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Second edition. (2016).

30. Kumwenda, G. P. et al. First Identification and genomic characterization of multidrug-resistant carbapenemase-producing Enterobacteriaceae clinical isolates in Malawi, Africa. J Med Microbiol **68**, 1707–1715 (2019).

31. Zhang, Y. et al. High Prevalence of Hypervirulent Klebsiella pneumoniae Infection in China: Geographic Distribution, Clinical Characteristics, and Antimicrobial Resistance. Antimicrobial Agents and Chemotherapy **60**, 6115–6120 (2016).

32. 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. Annals of Clinical Microbiology and Antimicrobials **18**, 4 (2019).