



Diversity, Virulence, and Antimicrobial Resistance in Isolates From the Newly Emerging *Klebsiella pneumoniae* ST101 Lineage

Chandler C. Roe¹, Adam J. Vazquez¹, Eliana Pia Esposito², Raffaele Zarrilli^{2*} and Jason W. Sahl^{1*}

¹ Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ, United States, ² Department of Public Health, University of Naples Federico II, Naples, Italy

OPEN ACCESS

Edited by:

Peter Mullany,
University College London,
United Kingdom

Reviewed by:

Miklos Fuzi,
Semmelweis University, Hungary
Roberta Migliavacca,
University of Pavia, Italy

*Correspondence:

Raffaele Zarrilli
razzari@unina.it
Jason W. Sahl
Jason.Sahl@nau.edu

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 20 November 2018

Accepted: 01 March 2019

Published: 02 April 2019

Citation:

Roe CC, Vazquez AJ, Esposito EP,
Zarrilli R and Sahl JW (2019) Diversity,
Virulence, and Antimicrobial
Resistance in Isolates From the Newly
Emerging *Klebsiella pneumoniae*
ST101 Lineage.
Front. Microbiol. 10:542.
doi: 10.3389/fmicb.2019.00542

The global dissemination of *Klebsiella pneumoniae* and *Klebsiella pneumoniae* carbapenemase (KPC) has been largely attributed to a few high-risk sequence types (STs) (ST258, ST11, ST512) associated with human disease. ST101 is an emerging clone that has been identified in different parts of the world with the potential to become a global, persistent public health threat. Recent research suggests the ST101 lineage is associated with an 11% increase in mortality rate in comparison to non-ST101 infections. In this study, we generated a high-quality, near-finished genome assembly of a multidrug-resistant (MDR) isolate from Italy (isolate 4743) that is a single locus variant of ST101 (ST1685). We demonstrate that the 4743 genome contains virulence features such as an integrative conjugative element carrying the yersiniabactin siderophore (ICEKp3), the mannose-resistant *Klebsiella*-like (type III) fimbriae cluster (mrkABCDFHIJ), the ferric uptake system (kfuABC), the yersiniabactin receptor gene *fyuA*, a capsular K type K17, and an O antigen type of O1. *K. pneumoniae* 4743 carries the *bla*KPC-2 carbapenemase gene along with genes conferring resistance to aminoglycosides, beta-lactams, fluoroquinolones, fosfomycin, macrolides, lincosamides, and streptogramin B. A comparative genomics analysis of 44 ST101 genomes as well as newly sequenced isolate 4743 identified variable antimicrobial resistance (AMR) resistance profiles and incompatibility plasmid types, but similar virulence factor profiles. Using Bayesian methodologies, we estimate the common ancestor for the ST101 lineage emerged in 1990 (95% HPD: 1965 to 2007) and isolates within the lineage acquired *bla*KPC after the divergence from its parental clonal group and dissemination. The identification of virulence factors and antibiotic resistance genes acquired by this newly emerging clone provides insight into the reported increased mortality rates and highlights its potential success as a persistent nosocomial pathogen. With a combination of both colistin resistance, carbapenem resistance, and several known virulence factors, the ST101 genetic repertoire may be a “perfect storm” allowing for a newly emerging, high-risk, extensively antibiotic resistant clone. This high-risk clone appears adept at acquiring resistance and may perpetuate the dissemination of extensive antimicrobial resistance. Greater focus on the acquisition of virulence factors and antibiotic resistance genes is crucial for understanding the spread of antibiotic resistance.

Keywords: *Klebsiella*, phylogenetics, antimicrobial resistance, ST101, genomics

INTRODUCTION

The Gram-negative bacterium, *Klebsiella pneumoniae*, is abundantly distributed in the environment and has traditionally been considered an opportunistic pathogen associated with hospital-acquired and community-acquired infections (Holt et al., 2015). Also, the rapid and global dissemination of multidrug-resistant (MDR) and extensively drug resistant (XDR or pan-resistant) *Klebsiella pneumoniae* was recently recognized by the CDC as an urgent public health threat requiring immediate and aggressive action (CDC, Antibiotic resistance threats in the U.S., 2013). Currently, carbapenem and broad-spectrum antibiotics are considered last resort treatment options, but the increasing incidence of extended-spectrum beta-lactamase (ESBL) and carbapenem resistant (CRE) isolates coupled with their global distribution highlights the potential for rapid dissemination of mobile MDR genes to other highly virulent, nosocomial pathogens (Ventola, 2015). *K. pneumoniae* contains significant genome variability and large accessory genomes, which includes virulence functions associated with invasive disease in humans and antimicrobial genes associated with hospital-acquired infections (Bialek-Davenet et al., 2014a; Holt et al., 2015; Cerqueira et al., 2017). Recent genomic epidemiology studies of *K. pneumoniae* isolates identified two types of high-risk clonal groups among bacterial populations: hyper-virulent clonal complexes (CCs) responsible for community-acquired invasive infections and multidrug-resistant CCs responsible for health-care associated infections (Struve et al., 2015; Cerqueira et al., 2017). The existence of hyper-virulent clones is a key feature of *K. pneumoniae* and typically show increased pathogenesis not commonly associated with antimicrobial resistance (Struve et al., 2015).

Klebsiella pneumoniae isolates responsible for health-care associated infections are usually CRE due to production of either class B metallo- β -lactamases (MBLs) (IMP, VIM, NDM) or class A (KPC) or class D (OXA-48) serine carbapenemases (Pitout et al., 2015; Grundmann et al., 2017; Logan and Weinstein, 2017). Multidrug-resistant (MDR) *K. pneumoniae* isolates have primarily been assigned to CC258 (Gaiarsa et al., 2015) and to additional emerging genotypes ST11, ST15 and ST101 (Bialek-Davenet et al., 2014a; Bowers et al., 2016; Conte et al., 2016; Cerqueira et al., 2017; Moradigaravand et al., 2017). *K. pneumoniae* isolates associated with hospital-acquired infections have acquired antimicrobial resistance genes and are usually devoid of virulence genes (Bialek-Davenet et al., 2014a; Cerqueira et al., 2017). However, the acquisition of yersiniabactin (an iron sequestering system crucial for disease establishment) has been observed in many isolates of the epidemic KPC-producing CC258 (Holt et al., 2015). High-risk clones have a highly flexible accessory genome and are adept at both acquiring resistance as well as switching resistance profiles (Woodford et al., 2011). Their global success as a nosocomial pathogen can be correlated with excessive genome plasticity (Conlan et al., 2016). Understanding high-risk clones' ability to adapt and survive in a hospital environment is vital to assuage further spread of antibiotic resistance. While hyper-virulent and drug-resistant *K. pneumoniae* populations remain mostly non-overlapping,

combinations of these groups (isolates within CC23) have been described (Bialek-Davenet et al., 2014a). The risk of combination of these two high-risk genetic profiles highlights the importance of understanding the spread of genomic regions from the *K. pneumoniae* accessory genome.

Among the newly emerging genotypes, *K. pneumoniae* isolates belonging to ST101 are associated with hospital-acquired infections and epidemics worldwide (Giani et al., 2013; Skálová et al., 2016; Gonçalves et al., 2017). *K. pneumoniae* isolates assigned to ST101 are CRE because of the production of KPC-2 (Oteo et al., 2016) or OXA-48 (Skálová et al., 2016; Avgoulea et al., 2018). Also, the emergence of colistin resistance has been observed in KPC-2 producing (Del Franco et al., 2015) and OXA-48 producing (Giani et al., 2013; Del Franco et al., 2015; Papagiannitsis et al., 2016) *K. pneumoniae* ST101 isolates. The focus of this study was to characterize the phylogenetic and genomic diversity within the emerging nosocomial high-risk sequence type ST101. MDR *K. pneumoniae* isolates typically belong to specific high-risk sequence types (ST11, ST258, ST512) (Bialek-Davenet et al., 2014a), with extensive comparative genomics research investigating the genomic backbone of these clones, however, few studies have investigated the global genomic diversity of the newly emerging and clinically relevant ST101 clone (Gonçalves et al., 2017; Moradigaravand et al., 2017; Avgoulea et al., 2018). In this study, we investigated the lineage relatedness, resistance determinants, plasmid profiles, core and accessory genome content, and the evolutionary rate of the ST101 sequence type through whole-genome sequencing. We propose a timeline for emergence as well as provide a better understanding of the evolution and composition of an increasingly successful MDR global lineage that may aid in the identification of newly emerging novel *K. pneumoniae* lineages and help understand the proliferation of extensively antibiotic resistance pathogens.

MATERIALS AND METHODS

Isolate Information of an MDR Strain From Italy

Klebsiella pneumoniae 4743 strain was isolated from rectal swab of a patient in the intensive care unit ward of public hospital of USL Valle D'Aosta, Aosta, Italy on November 24th 2013. To compare this isolate in a phylogenetic context, a global collection of 1,723 publicly available *K. pneumoniae* isolates were downloaded and analyzed (**Supplementary Table S1**).

Antimicrobial Resistance Profiling

Antimicrobial susceptibilities were performed for 4743 using the Vitek 2 system and the AST-GN card (bioMérieux, Marcy l'Étoile, France). Values were interpreted according to breakpoint table for interpretation of MIC values and zone diameters (European Committee on Antimicrobial Susceptibility Testing, 2016). Colistin susceptibility assay was performed according to recommendations of joint CLSI-EUCAST guidelines: http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/General_documents/Recommendations_for_MIC_determination_of_colistin_March_2016.pdf.

DNA Extraction, Sequencing, and Assembly

For Illumina MiSeq sequencing of strain 4743, genomic DNA was extracted with the GenElute DNA extraction kit (Sigma-Aldrich, Milan, Italy). A sequence library was generated for paired-end sequencing using previously described methods (Stone et al., 2016) and sequenced to an average depth of 50x. For PacBio sequencing, genomic DNA of sample 4743 was extracted using a DNeasy Blood and Tissue Kit according to the manufacturer's instructions (Qiagen, Milan, Italy). Approximately 10 µg of DNA was fragmented to 10–20 kbp using the G-tube apparatus (Covaris) following the manufacturer's recommendations. A PacBio Sequencing library was constructed using the SMRTbell™ Template Prep Kit 1.0 and by following the PacBio 20 kb library protocol; the library molecules, after adapter ligation and damage repair, were size selected for 15 kb and larger using Blue Pippin instrument (Sage Sciences) by following the manufacturer's instructions. The final library was processed for sequencing by using PacBio MagBead kit v2 with the P6/C4 chemistry and following PacBio protocols. Sequencing was performed on PacBio RSII instrument in one SMRT cell (v3) for 6 h.

The 4743 genome was assembled using Unicycler v0.4.7 (Wick et al., 2017), resulting in seven contigs. The genome was polished by running six rounds of Pilon (Walker et al., 2014) and was then processed with Circlator (Hunt et al., 2015). Assembly statistics are shown in **Supplementary Table S2**. The assembly and raw reads were deposited in NCBI under BioProject PRJNA477005. Additionally, genome annotation was performed using PROKKA v1.13 (Seemann, 2014).

In silico MLST Typing

Multi-locus sequence typing (MLST) analysis was performed on all genome assemblies using an *in silico* MLST script¹ which implemented the Institut Pasteur's MLST scheme as previously described (Diancourt et al., 2005).

In silico K Typing and O Typing

Polysaccharide capsule (K typing) and lipopolysaccharide O antigen typing was performed using the Kaptive tool (Wyres et al., 2016). Briefly, the source code for the command-line version of Kaptive was downloaded from the GitHub repository². Both K and O typing of 45 ST101 isolates was accomplished running default settings within the Kaptive.py script.

Core Genome SNP Phylogenies

All external *K. pneumoniae* genomes ($n = 3,352$) were downloaded from GenBank in February 2018. Genomes were filtered from the dataset if they contained: (1) greater than 10 ambiguous nucleotides; (2) an anomalous number of contigs (>462); (3) an anomalous genome assembly size (<5,195,738, >5,992,749) and; (4) an anomalous pairwise MASH distance (>0.017) (Ondov et al., 2016); this filtering resulted in a final

dataset of 1,504 genomes. All genome assemblies were aligned against the completed reference ST101 genome GCA_001902435 using NUCmer (Delcher et al., 2002) and SNPs were identified as part of the NASP pipeline (Sahl et al., 2016). A global maximum likelihood phylogeny was generated using the TVM+ASC+G4 nucleotide substitution model in IQTREE v1.6.1 (Nguyen et al., 2015) with 1,000 bootstrap replicates. Trees were visualized in FigTree³. The Retention Index (Farris, 1989) was calculated with Phangorn (Schliep, 2011).

For the ST101 clade analysis, assembled genomes were aligned to sample Kp_Goe_33208 (GCA_001902435) using the previously described methods above. A ST101-only maximum likelihood phylogeny was inferred using the K3P+ASC substitution model with 1,000 bootstrap replicates in IQTREE. Phylogenetic trees were visualized in FigTree. The Retention Index was calculated with Phangorn. Additionally, the presence of recombination within this dataset was determined using the PHI statistic from the package PhiPack (Bruen et al., 2006).

LS-BSR Analysis

The Large-Scale Blast Score Ratio (LS-BSR) pipeline (Sahl et al., 2014) was used to identify differential coding region conservation within the ST101 clade. This pipeline predicted coding regions (CDSs) using the program Prodigal v2.60 (Hyatt et al., 2010) and clustered the putative CDS regions with 90% identity using VSEARCH v1.11.1 (Rognes et al., 2016). The resulting regions were aligned against themselves using the program BLAT v35×1 (Kent, 2002) to generate a reference bit score. Genomic regions were then aligned back to every sample in the ST101 dataset using BLAT to generate a query bit score. The query bit score was divided by the reference bit score for each region to obtain the BLAST Score Ratio (BSR) (Rasko et al., 2005). The core CDSs were identified as genomic regions that had a BSR value of > 0.80 across all genomes. Unique genomic regions were identified based on a BSR value < 0.40 in all but a single genome.

Genes Associated With Virulence, Heavy Metal Resistance, and Drug Resistance

Genes associated with virulence (120), colibactins (91), yersiniabactins (165), heavy metal resistance (203), efflux pumps and regulators (489), and non-scheme genes (278) were downloaded from BIGSdb-Kp database (Bialek-Davenet et al., 2014b). LS-BSR was run on the ST101-only lineage with the genes parameter flagged using the virulence, heavy metal resistance, and efflux pump databases as input. LS-BSR was run using nucleotides and the BLAT alignment option. A BSR value > 0.95 was considered present.

QRDR Mutations

Quinolone resistance determining regions (QRDRs) within the ST101 genomes were inspected for mutations conferring increased fitness within gyrase (*gyrA* and *gyrB*) and topoisomerase IV (*parC* and *parE*) genes. Reference genes for *parE* (AFQ64062.1), *parC* (AFQ64068.1), *gyrA* (AFQ64815.1),

¹<https://gist.github.com/jasonsahl/2eedc0ea93f90097890879e56b0c3fa3>

²<http://github.com/kelwyres/Kaptive-Web>

³<http://tree.bio.ed.ac.uk/software/figtree/>

and *gyrB* (AFQ63426.1) were downloaded from NCBI. Blastx from blast+ v2.2.29 (Camacho et al., 2009) was used to identify the composition of each gene in the ST101 genomes. Genes were extracted from each genome and alignments were performed using ClustalW (Thompson et al., 1994) in MEGA7 (Kumar et al., 2016).

Plasmid Analysis

The *Enterobacteriaceae* plasmid database from PlasmidFinder (Carattoli et al., 2014) was used to identify plasmid types within the ST101 clade. The PlasmidFinder database consists of representative genes of varying incompatibility groups. Here we screened 122 typing genes from PlasmidFinder against the ST101 genomes using LS-BSR/BLAT. BSRs were generated and visualized in Rstudio using the shinyheatmap program (Khomtchouk et al., 2017). Typing genes with a BSR value > 0.95 were considered to be present.

Antibiotic Resistance Gene Screening

The distribution of genes associated with virulence and antimicrobial resistance was identified for all ST101 genomes. Acquired genes associated with antimicrobial resistance were identified using the ResFinder database (Kleinheinz et al., 2014) with LS-BSR. The ResFinder database included genes that may confer resistance to the following antibiotics: aminoglycosides, beta-lactams, colistin, fluoroquinolone, fosfomycin, fusidic acid, glycopeptide, macrolide, lincosamide, and streptogramin B (MLS), nitroimidazole, oxazolidinone, phenicol, rifampicin, sulphonamide, tetracycline, and trimethoprim. Mutations within the *mgrB* regulator gene that are known to confer resistance to colistin were also investigated (Esposito et al., 2018). Briefly, the program BLAT was used to identify the *mgrB* gene within the ST101 samples and the sequence was extracted from each alignment. The *mgrB* gene sequences were aligned using MUSCLE (Edgar, 2004) in MEGA7 (Kumar et al., 2016) and subsequently translated into amino acids. The amino acid sequences were visually inspected for premature stop codons. Additionally, genes associated with virulence were obtained from the BIGSdb-KP database and screened for within the ST101 genomes using LS-BSR. A heatmap representing presence/absence of antimicrobial resistance genes was produced using shinyheatmap in R.

Population Structure

In order to assess the genomic relationships within the ST101 lineage, we applied Plink v1.07 (Purcell et al., 2007) and fastStructure (Raj et al., 2014), a Bayesian model-based clustering algorithm; the Plink output was used as input for fastStructure. Briefly, SNP positions with more than two allele states were removed from the dataset. Plink was implemented using default parameters. As the *K* parameter represents populations in fastStructure, we implemented $K = 2$ through $K = 9$ with default parameters. An optimal *K* value was chosen based on model complexity that maximizes marginal likelihood as well as optimum model components used to explain the structure.

Beast Timing Analysis

BEAST analysis included 36 of the 45 ST101 genomes (Supplementary Table S3). Strains were only included in this analysis if the collection date was listed under the biosample data on NCBI. A core SNP matrix was generated by aligning raw reads to sample 4743 using the NASP pipeline. The presence of recombination within a dataset can confound molecular clock analyses and was therefore identified and removed using ClonalframeML running default parameters (Didelot and Wilson, 2015). A regression analysis implementing root-to-tip genetic distance as a function of the sample collection year was conducted using the software package TempEst version 1.5.1 (Rambaut et al., 2016). A measure of clocklike behavior was assessed using the determination coefficient R^2 and the rooted ST101-only phylogeny. Additionally, a 10,000 date-randomization permutation of sampling collection dates was performed in an effort to compare our regression coefficient to that observed by random chance (Murray et al., 2016).

RESULTS

4743 Strain Details

Klebsiella pneumoniae 4743 was isolated from a rectal swab of a patient admitted to Intensive Care Unit of USL Valle D'Aosta, Aosta, Italy on November 24th 2013. *K. pneumoniae* 4743 isolate was representative of an epidemic of KPC-2 producing *K. pneumoniae* CC101 occurring from November 2013 to August 2014 (Del Franco et al., 2015). *K. pneumoniae* 4743 isolate was able to transfer resistance to carbapenems and ESBL activity along with conjugative plasmids carrying *bla*_{KPC-2} and *bla*_{CTX-M-group1} genes, respectively (Del Franco et al., 2015).

4743 Genome Assembly

The genome assembly using PacBio reads combined with Illumina MiSeq reads produced a total of 7 contigs, 6 of which are plasmids and 1 of which is chromosomal; 4 of the plasmids were determined to be circular. The total size of sample 4743 is 5,857,478 bases. PHASTER identified four intact phage within the large chromosome contig, and two more complete phage in two of the plasmid contigs (4743 plasmid unnamed 3 and 4743 plasmid unnamed 6).

Antimicrobial Resistance Profiling

Antimicrobial susceptibilities were performed for sample 4743 using the Vitek 2 system and the AST-GN card. *Klebsiella pneumoniae* 4743 isolate was resistant to imipenem, meropenem, ertapenem, beta-lactam/beta-lactamase inhibitor combinations (clavulanic acid/amoxicillin, piperacillin/tazobactam), third and fourth generation cepheems, ciprofloxacin, but susceptible to trimethoprim-sulfamethoxazole and colistin (Supplementary Table S4).

External Genomes

External genome assemblies from *K. pneumoniae* were downloaded from the assembly database in GenBank and

filtered based on several quality criteria, resulting in 1,504 genome assemblies. *in silico* MLST typing was performed on all assemblies using the previously described Pasteur system (Diancourt et al., 2005). Of all *K. pneumoniae* isolates analyzed, 41 were assigned to ST101, three were assigned to ST2017, and one sample (4743) was assigned to ST1685. Both ST2017 and ST1685 are single locus variants (locus *rpoB*) of ST101 and were included in subsequent analyses.

Phylogenetics and Comparative Genomics

WGS of one isolate (4743) was performed and compared to whole genome sequences of 44 *K. pneumoniae* isolates assigned to the ST101 lineage available in GenBank and 1,504 non-ST101 *K. pneumoniae* reference genomes (**Supplementary Table S1**). The core genome phylogeny based on 1.9 Mb of conserved sequence and 179,342 SNPs demonstrated the position of the ST101 lineage (**Figure 1**) in relation to most global lineages. The retention index (RI) of the global isolate SNP alignment was 0.958, indicating little homoplasy throughout the *K. pneumoniae* core genome. The presence of recombination introduces incongruences in the phylogenetic placement of lineages with deeply branching nodes and long branches (Schierup and Hein, 2000). The PHI statistic was used to test for evidence of recombination. The *p*-value for the PHI statistic was 0.00e+00, providing evidence, but not quantification, of recombination.

A phylogeny of the ST101 lineage only, as defined by a monophyletic clade in the global phylogeny, was based on 3,992 concatenated SNPs and demonstrated little homoplasy with an RI value of 0.95, indicating that the genomes in this lineage are closely related and their evolution was minimally driven by recombination. However, the PHI statistic revealed statistically significant recombination within this dataset (*p*-value = 0.0). ClonalframeML identified 372 recombination events that spanned 138,212 bases, both monomorphic and polymorphic, across the chromosome of the ST101 lineage. Masking of these regions removed 1,882 SNPs from the ST101 core genome. The high-quality, non-recombination core genome size of the ST101 clade was 4.85 Mb.

Population Structure in the ST101 Lineage

In order to visualize the shared genomic regions within the ST101 lineage, fastStructure was applied using the ST101-only dataset that included 3,992 SNPs. We evaluated our dataset for shared ancestry using a range of *K* values (number of expected populations) from 2 to 10. Model complexity that maximized the marginal likelihood was 2. Model components used to explain the structure in the data was 3. When *K* = 2, fastStructure revealed two subpopulations within the ST101 lineage with limited genome sharing between the two populations; only two samples (GCA 001720745 Pakistan 2013 and GCA 001720815 Pakistan 2013) displayed genome sharing (**Figure 2**). For *K* = 3, fastStructure

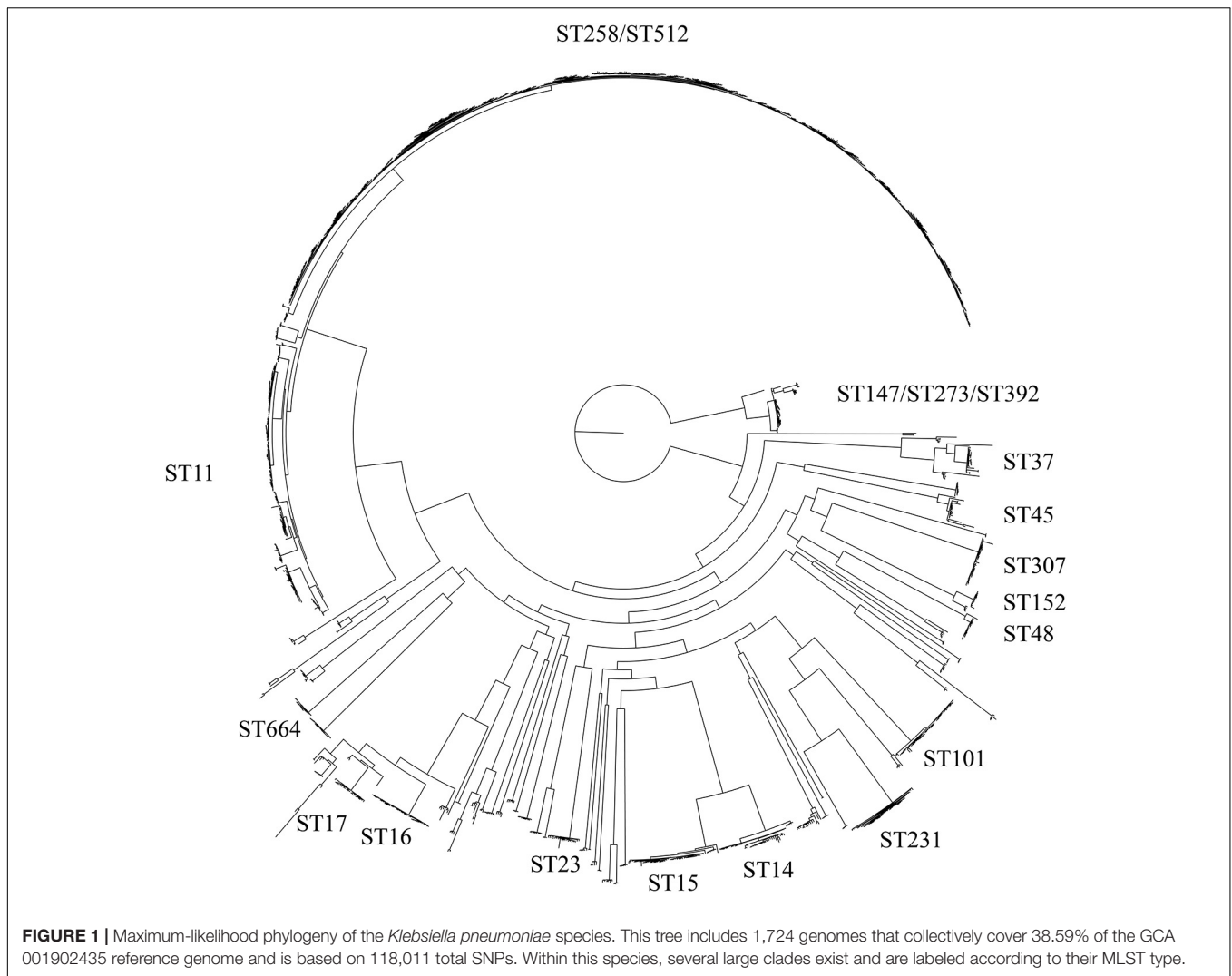
revealed limited genome sharing between two of the three of the populations; eight samples demonstrated admixture. In both models, samples GCA 002187295 (Nigeria 2014) and GCA 002247645 (Thailand 2015) make up a separate ST101 population.

In silico Antimicrobial Resistance Profiling

As a complement to the laboratory-determined antimicrobial susceptibility profile for strain 4743 and in order to compare the ST101 lineage antimicrobial susceptibility patterns, *in silico* antimicrobial susceptibility testing was performed using the ResFinder database. Genes responsible for conferring aminoglycoside resistance, beta-lactam resistance, and fluoroquinolone resistance were conserved across all ST101 clade genomes (**Figure 3** and **Supplementary Table S5**). Fluoroquinolone resistance genes were identified within every ST101 genome (*oqxA*, *oqxB*). Genes responsible for fosfomycin resistance were identified in 97.8% of ST101 genomes (*fosA*). Phenicol resistance genes (*catB4*, *catA2*, *cmlA1*, *floR*) were conserved in 73% of the ST101 genomes. In total, 26% ST101 isolates carried genes responsible for macrolide, lincosamide, and streptogramin B resistance (*mphE*, *msrE*, *ereA*, *ereB*, *mphA*). Rifampicin resistant genes (*arr-2* and *arr-3*) were present in 24% of genomes, and 71.1% of isolates carried genes responsible for sulphonamide resistance (*sul1*, *sul2*, *sul3*). Additionally, 60% of the isolates harbored genes conferring Tetracycline resistance (*tetD*, *tetX*, and *tetA*) and 86.7% carried at least one gene responsible for trimethoprim resistance (*dfrA1*, *dfrA5*, *dfrA14*, *dfrA16*, or *dfrA27*). A total of 15 different beta-lactamase genes, including ESBLs, were identified within the ST101 lineage. The most prevalent ESBL gene identified was *bla_{CTX-M-15}* (84%); *bla_{CTX-M-14}* gene was found in two ST101 genomes. The *bla_{SHV-1}* gene was identified in 87% of the ST101 samples, conferring resistance to broad spectrum beta-lactams. Sample 4743 carried the *bla_{CTX-M-15}* gene. Two genomes, sample 4743 and ST101 (Oteo et al., 2016), carried the *bla_{KPC-2}* gene (pKP048_p019, NC_014312.1). Additionally, 33.3% of the 45 genomes carried the *bla_{OXA-48}* gene. Alarmingly, 7 of 45 (15.5%) ST101 genomes contain alterations in *mgrB* gene (4 genomes harbor either a frameshift mutation or IS insertional inactivation while 3 genomes show a deletion of the *mgrB* locus) conferring colistin resistance (Esposito et al., 2018).

Plasmid Composition

Plasmid typing was analyzed in ST101 genomes using 122 previously characterized *K. pneumoniae* plasmids from the database PlasmidFinder. The results demonstrate the plasmid diversity across the ST101 lineage (**Figure 4**), although this methodology does not reveal whether the typing genes are present chromosomally or on plasmids. Plasmid replicon analysis revealed various types of plasmid incompatibility groups among the ST101 isolates, none of which were detected in all 45 ST101 samples. However, incompatibility plasmid groups FIB, FII, and R were detected in the majority of ST101 samples [IncFIB (K) in 84%, IncFII (K) in 75%, IncR in 61%, IncFIA



(HI1) in 61%, and IncL/M (pOXA-48) in 29%]. The following incompatibility groups were found in 3 out of 4 plasmids identified in sample 4743 as well as one non-circular contig: IncFIB(K) (4743 plasmid unnamed 2), IncFII(K) (4743 plasmid unnamed 2), IncR (4743 plasmid unnamed 3), IncFIA (HI1) (4743 plasmid unnamed 3), IncFII (4743 plasmid unnamed 4), and ColRNAL (4743 plasmid unnamed 5). Plasmid pKP-KPC2 corresponding to 4743 plasmid unnamed 3 (111,854 nts) carried IncFII(K) and *bla*_{KPC-2} gene, plasmid pKP-CTXM-15 corresponding to 4743 plasmid unnamed 4 (67,867 nts) carried IncFII and *bla*_{CTXM-15} gene. Interestingly, the blastn alignment of plasmid pKP-KPC2 from isolate 4743 and plasmid pKP048 from ST101 *K. pneumoniae* isolate described by Oteo et al., 2016 showed that the *bla*_{KPC-2} gene is included in Tn1721 transposon structure of 14,470 nts (residues 11,810–26,279 in plasmid pKP048).

Unique Genomic Regions

The complete gene content for 1,504 *K. pneumoniae* genomes was compared using LS-BSR. By using default values in

LS-BSR, we did not identify any coding regions that were unique to the ST101 clade and absent from all others. The core genome for the ST101 lineage consisted of 4,465 coding region sequences (CDSs). The ST101 lineage accessory genome had a total of 2,247 CDSs. There were a total of 564 CDSs found only in a single genome within ST101 and the average number of unique CDSs per genome was 12.8. While these genes are unique within the ST101 lineage, all 564 CDSs were identified throughout the global *K. pneumoniae* genomes.

Virulence and AMR Associated Genes

Known genes associated with virulence (Supplementary Table S6) from the BIGSdb-Kp database were screened against the ST101 lineage using LS-BSR and BLAT. These genes included heavy metal resistance genes, efflux pumps, colibactin and yersiniabactin genes. Additionally, we screened for the yersiniabactin-encoding mobile element ICEKp as well as capsule and lipopolysaccharide serotyping. Samples within the ST101 lineage all carry the iron-scavenging siderophore



FIGURE 2 | Ancestry proportions were inferred for the ST101 lineage using fastStructure v1.0 with the optimal choices of K (populations). Two different K -values (number of populations) were selected as most appropriate, $K = 2$ and $K = 3$. $K = 2$ graph represents two populations with little shared ancestry between them. Population one is represented by dark red. Population two is distinguished as brown. When running fastStructure with the number of expected populations set at three ($K = 3$), we observed limited shared ancestry.

yersiniabactin on the mobile genetic element ICEKp3 except for two samples, GCA 002187295 and GCA 002247645, which fall outside the main ST101 clade (Figure 1). The ST101

lineage was screened for a total of 489 efflux pump genes from the BIGSdb-KP database. We identified 115 conserved efflux pump genes within all ST101 samples including the

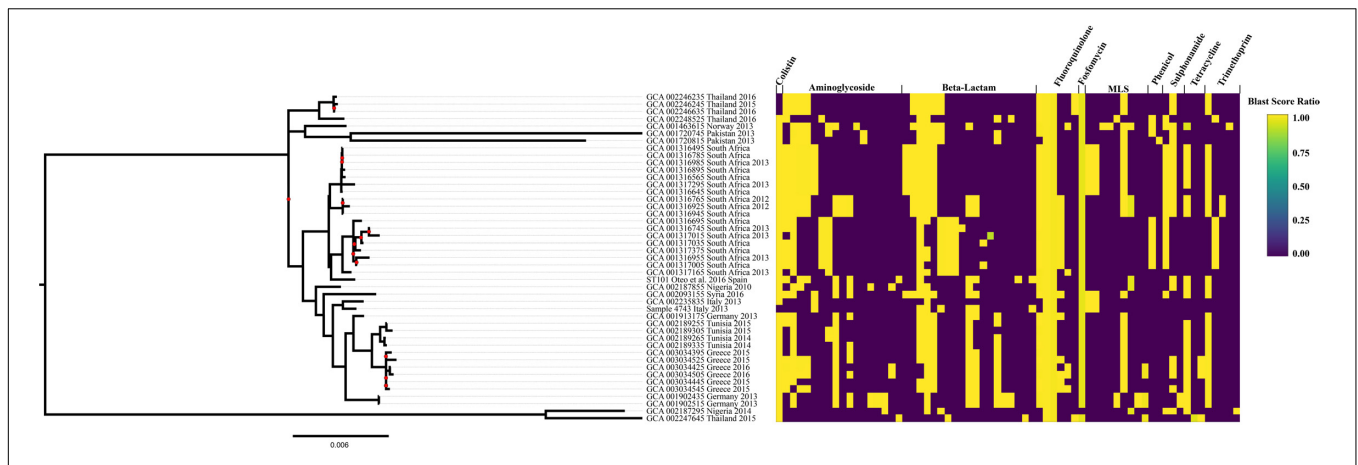


FIGURE 3 | Maximum-likelihood phylogeny of the *Klebsiella pneumoniae* ST101 lineage. The tree is based on 4,077 SNPs using the K3P+ASC+G4 nucleotide substitution model and includes 45 genomes that collectively cover 85.95% of the reference genome. The PHI test found statistically significant evidence for recombination in this dataset ($P = 0.00e+00$). All samples were screened for antibiotic resistance using ResFinder. The percent identity for each resistance gene was plotted using shinyheatmap alongside the phylogenetic tree. Gene lists are in **Supplementary Table S6**. MLS includes genes that confer resistance to macrolides-lincosamides-streptogramin B.

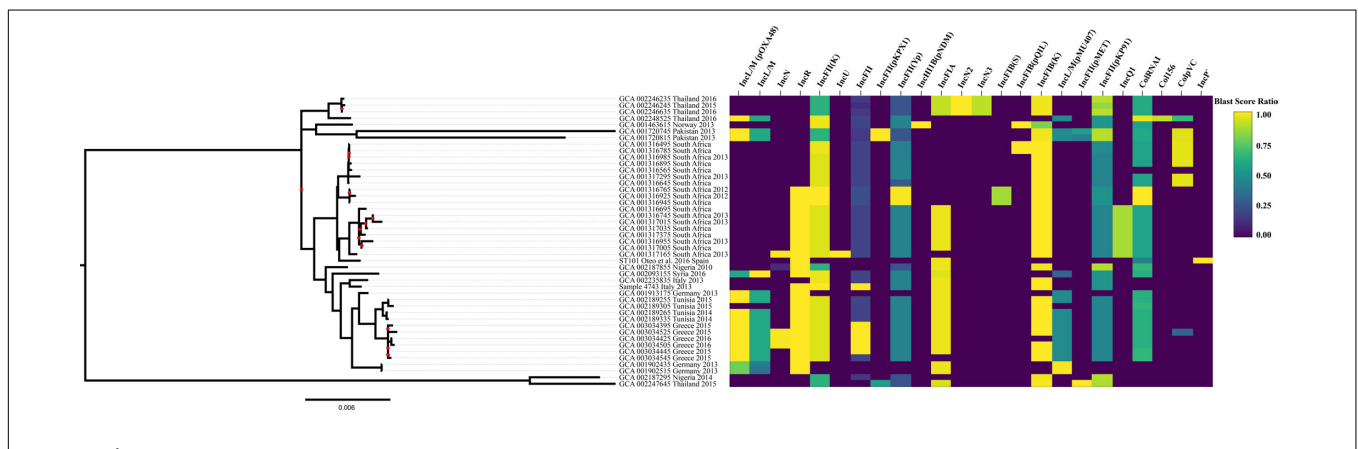


FIGURE 4 | Plasmid replicon gene screening. The ST101 lineage assemblies were screened for plasmid types using PlasmidFinder. The percent identity for each plasmid typing gene was plotted using shinyheatmap alongside the maximum-likelihood phylogenetic tree.

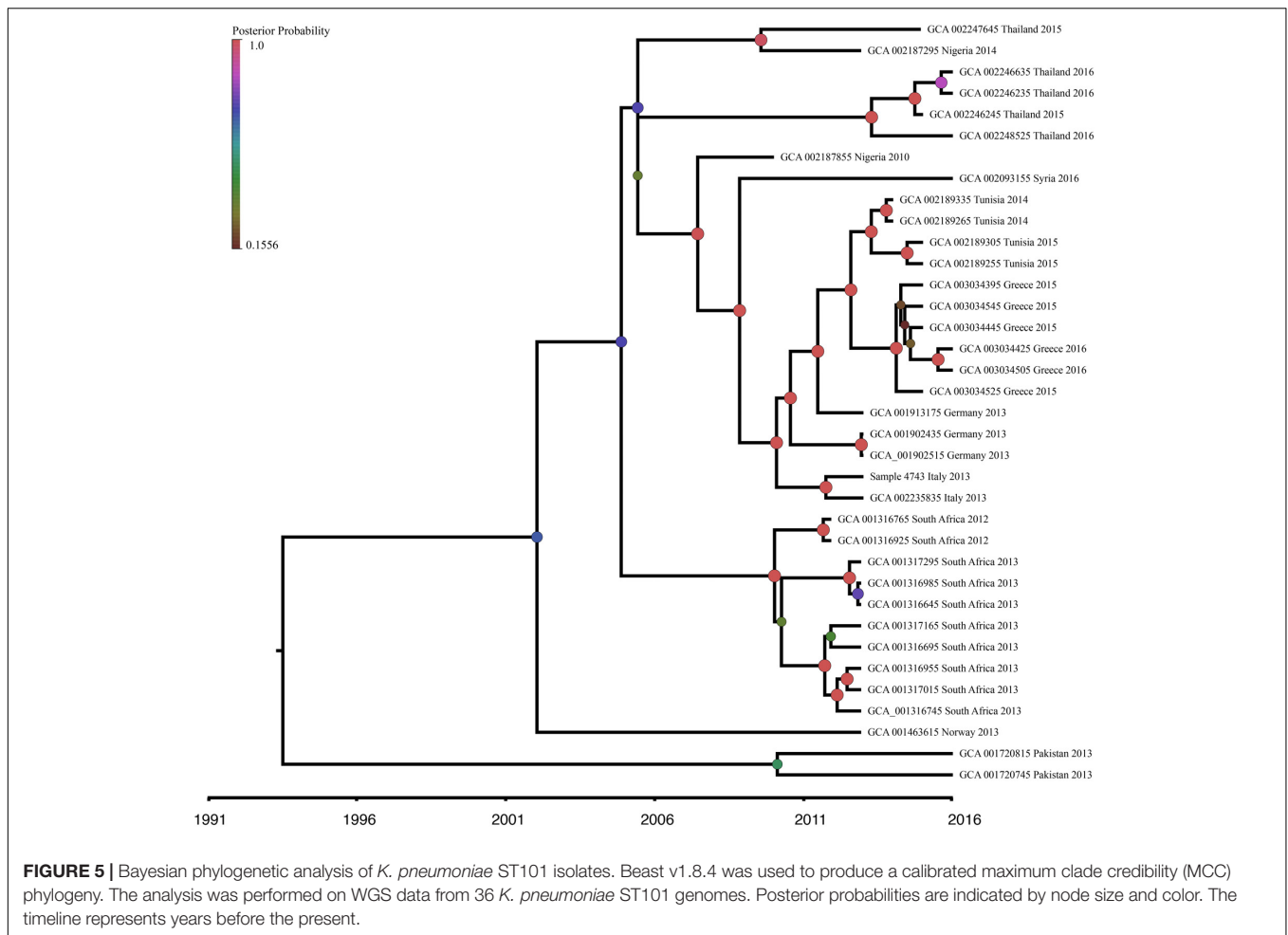
virulence and antibiotic cross-resistance associated AcrAB and OqxAB efflux pumps as well as the *rarA* regulator gene (**Supplementary Table S5**). A total of 203 heavy metal resistance genes were screened against the ST101 lineage. The complete *pcoABCDRE* operon conferring copper resistance was identified in 36 of the 44 ST101 samples while 41 of the genomes carry genes responsible for silver resistance (*silCERS*).

Both K typing and O typing have meaningful clinical and epidemiological significance. Capsular polysaccharide characterization (K typing) is widely used to define clinical *K. pneumoniae* isolates (Hansen et al., 2002; Vimont et al., 2008; Woodford et al., 2011). The polysaccharide capsule is often considered a virulence determinant as K-type variations have been linked to specific presence/absence of genes within the locus. Specific lipopolysaccharide antigens (O typing) also contribute to *K. pneumoniae* pathogenicity. In an effort to better understand the pathogenicity of the ST101 clone,

we performed both K and O typing using the program Kaptive Web. Limited capsular diversity was observed with 43 samples having the KL17 loci and 2 samples with the KL106 loci. O loci typing revealed the majority of the ST101 lineage as O1V1. Interestingly, the O1 antigen has been previously described as a major contributor to the virulence of pyogenic liver abscess causing *K. pneumoniae*. Three ST101 samples (GCA 001316955, GCA 002189265, GCA 001316985) were typed as O1/O2v1. O1/O2 signifies that either gene *wbbY* or *wbbZ* could not be found within the sample (Wick et al., 2018).

QRDR Mutations

All of the ST101 genomes but two (GCA_002187295 and GCA_002247645) showed amino acid substitutions in codon 83 (Ser83Tyr) and codon 87 (Asp87Gly, Asp87Asn, Asp87Ala) of the *gyrA* gene (**Supplementary Table S7**). Additionally, a Ser80Ile substitution in the



parC gene was detected in all ST101 genomes with the exception of two samples (GCA_002187295 and GCA_002247645). Amino acid substitutions within QRDR of the *parE* or the *gyrB* genes were not observed in any of the ST101 genomes.

Beast Analysis

The ST101-only Bayesian dataset consisted of 35 genomes with reported collection dates from NCBI and the newly sequenced sample 4743. This dataset contained 3,977 SNPs, with 1,889 SNP positions falling within genomic regions determined by ClonalframeML as recombining and subsequently removed from this analysis. The root-to-tip regression analysis identified weak clocklike behavior ($R^2 = 0.025$) within the ST101 lineage with 2.5% of the diversity explained by time. However, we believe we are capturing little of the true temporal signal due to our narrow sampling dates (2010–2016) and small sample size ($n = 36$). Given the positive regression slope and the detection of a clocklike signal, we determined molecular clock analysis was appropriate and reliable for SNP accumulation rate estimation (Duchêne et al., 2016). MEGA7 identified the best fitting nucleotide substitution model as the GTR model and was applied in the

BEAST analysis. In an effort to determine if the observed R^2 value was better than random chance, a 10,000-date permutation testing on the recombination removed ST101-only dataset. This testing produced a P value of 0.339, indicating that our observed R^2 value was not statistically different than random chance.

In order to investigate the evolution of *K. pneumoniae* ST101, BEAST analysis was performed using the recombination-free, high quality SNP dataset previously described. The constant population demographic model with a relaxed molecular clock was selected as the most appropriate model and clock combination to describe ST101 evolution. Molecular clock calibration estimated the evolutionary rate for the ST101 lineage as 2.8527×10^{-6} substitutions per site per year (95% highest posterior density [HPD], 1.0762×10^{-6} to 4.7396×10^{-6}). The mean time to most recent common ancestor (TMRCA) was estimated at 26.6 years ago (95% HPD, 8.35 to 51.24 years ago) from the time of the last sample date, which was 2016 (Figure 5). This dataset has narrow sampling dates with a limited number of samples and it is likely we are capturing very little of the temporal signal. Given a wider sampling time and additional samples, we would likely narrow our MRCA estimate.

DISCUSSION

The focus of this study was to perform a comprehensive genomics analysis of the ST101 lineage to better understand the genomic content of this newly emerging, extensively resistant lineage. Previous studies have identified this sequence type as high risk for furthering the spread of carbapenem resistance (Del Franco et al., 2015; Oteo et al., 2016). Here we demonstrate that ST101 is a distinct lineage in which one sample in this study, 4743, acquired the KPC-2 enzyme. Additionally, a recent study identified 13 ST101 samples carrying the KPC-2 gene (WGS data not yet available) (Del Franco et al., 2015; Oteo et al., 2016). The inconsistent carriage of KPC-2 within the ST101 clone suggest the KPC enzyme was acquired through horizontal transfer of plasmids after the divergence of the ST101 lineage rather than vertical gene transfer among the entire lineage. In further support of this, our data showed that ST101 *K. pneumoniae* isolate 4743 and ST101 isolate described by Oteo et al., 2016 both carry plasmids containing the *bla*_{KPC-2} gene into Tn1721 transposon.

We found multiple genomic features that may provide this clone with an advantage for adaptation within a hospital environment as well as a human host. Multiple putative virulence factors were identified throughout the ST101 lineage including the siderophore genes *irp1* and *irp2*, the yersiniabactin receptor gene *fyuA*, the ICEKp-3 element that carries the yersiniabactin siderophore cluster (ybtAEPQSTUX), the mannose-resistant *Klebsiella*-like (type III) fimbriae cluster (mrkABCDFHIJ), and the ferric uptake system (kfuABC). The *irp* and *irp2* genes encode iron-repressible high molecular weight proteins that are involved with yersiniabactin production (Schubert et al., 1998); these were identified in all but two samples within the ST101 lineage. The *fyuA* (ferric yersiniabactin uptake) gene is also involved with the iron-acquiring yersiniabactin system by coding for yersiniabactin receptors (Schubert et al., 1998). In *E. coli*, the *fyuA* gene is required for biofilm formation in urinary tracts and is important for disease establishment in iron-poor environments such as the urinary tract (Hancock et al., 2008). The ST101 lineage also carry the yersiniabactin siderophore cluster on an ICE, which has been identified as a frequent virulence factor in *K. pneumoniae* (Schubert et al., 2004; Holt et al., 2015; Lam et al., 2018). The ICEKp element increases the ST101 clone's ability to cause disease by coding for an iron scavenging system; this element also allows this clone to further spread this important virulence factor because it is located on an integrative conjugative element (Lam et al., 2018). The mannose-resistant *Klebsiella*-like (type III) fimbriae cluster (mrkABCDFHIJ) is considered a virulence factor and has been previously described in several species of *Enterobacteriaceae* as contributors to mucous adherence, tissue colonization, and biofilm formation (Monroy et al., 2005; El Fertas-Aissani et al., 2013). Finally, the ferric uptake system (kfuABC) is also associated with increased virulence (Ma et al., 2005). In pyogenic liver abscess causing *K. pneumoniae*, the kfu system is described as usually more prevalent in invasive tissue strains, however, the presence of multiple iron-acquisition systems suggests the ST101 lineage is capable of acquiring and using iron from diverse sources similar to pyogenic liver causing strains (Ma et al., 2005; Hsieh et al., 2008). Previous research

suggests iron acquisition gene diversity allows for more capable environmental iron acquisition (Lee et al., 2012). This increased efficiency of iron uptake has been hypothesized to result in increased capsule production, which is an important phenotype of hypervirulence (Russo et al., 2011; Chaturvedi et al., 2012; Shon et al., 2013; Moradigaravand et al., 2017).

Hyper-virulent and drug-resistant *K. pneumoniae* populations remain mostly non-overlapping, although combinations of these groups have been previously reported (Bialek-Davenet et al., 2014a; Gu et al., 2018). Here we demonstrate the presence of several known virulence factors within the ST101 lineage as well as highlight the extensive drug resistance gene repertoire this clone harbors (colistin, beta-lactams, aminoglycosides, fluoroquinolones, and fosfomycin). The ST101 lineage is an example of a dual-risk clone combining the genetic profiles of hypervirulent *K. pneumoniae* as well as extensive drug resistant strains. Alarming, the ST101 resistome is similar to the antibiotic resistant profile of the global ST258 lineage. The majority of both ST101 and ST258 clones possess genes that are responsible for resistance to aminoglycosides [*aac*(6'), *aph*(6'), *aadA* genes], third generation-cephalosporins and aztreonam (*bla*CTXM), fosfomycin (*fosA*), and fluoroquinolones [*aac*(6')-Ib-cr, *oqxA*, *oqxB*, *qnrB*, *qnrS* genes]. Both lineages confer resistance to colistin due to inactivation of the *mgrB* gene and to carbapenems due to the acquisition of the *bla*KPC-2 gene. Both lineages carry the mannose-resistant *Klebsiella*-like (type III) fimbriae cluster (mrkABCDFHIJ), and the ICEKp containing the yersiniabactin siderophore cluster (ybtAEPQSTUX). While drug-resistance has been shown to be associated with a large loss of fitness, researchers have demonstrated the presence of three amino acid substitutions within QRDR of *parC* and *gyrA* genes that result in no loss of vitality in multiple bacterial species, including *K. pneumoniae* (Marcusson et al., 2009; Toth et al., 2014). In addition to extensive drug resistance, genomes from the ST101 lineage harbor these three amino acid substitutions within the *gyrA* and *parC* genes. *K. pneumoniae* minor sequence types that are fluoroquinolone resistant were reported to lack these mutations entirely, carry a "non-serine" mutation, or only one of the three mutations (Fuzi et al., 2017). However, major clones, specifically ST258, were all shown to carry the *gyrA* Ser83Ile and *parC* Ser80Ile double mutations (Bowers et al., 2015). Given the global success of ST258 and similar virulence, antibiotic resistance, and QRDR mutation profiles compared to the ST101 lineage, this newly emerging clone has the potential to become a global epidemic dual-risk clone and major public health threat.

DATA AVAILABILITY

The datasets generated for this study can be found in NCBI, PRJNA477005.

ETHICS STATEMENT

The study was approved by the ethics committee of the Aosta Regional Hospital (protocol number 836/2015). All

microbiological samples were taken as part of standard care procedures. Patients included in the study were anonymized, no written informed consent was acquired because of the retrospective nature of the study.

AUTHOR CONTRIBUTIONS

CR analyzed the data and wrote the manuscript. AV and EE performed laboratory analyses. RZ designed the study and helped to write the manuscript. JS designed the study, analyzed the data, and helped to write the manuscript.

REFERENCES

- Antibiotic Resistance Threats in the U.S. (2013). *Centers for Disease Control and Prevention*. Available at: <https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf>
- Avgoulea, K., Di Pilato, V., Zarkotou, O., Sennati, S., Politi, L., Cannatelli, A., et al. (2018). Characterization of extensively drug-resistant or pandrug-resistant sequence type 147 and 101 OXA-48-producing *Klebsiella pneumoniae* causing bloodstream infections in patients in an intensive care unit. *Antimicrob. Agents Chemother.* 62:e2457-17. doi: 10.1128/AAC.02457-17
- Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Jones, L., Delannoy-Vieillard, A.-S., et al. (2014a). Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg. Infect. Dis.* 20, 1812–1820. doi: 10.3201/eid2011.140206
- Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Nicolas-Chanoine, M.-H., Decre, D., et al. (2014b). Development of a multiplex PCR assay for identification of *Klebsiella pneumoniae* hypervirulent clones of capsular serotype K2. *J. Med. Microbiol.* 63, 1608–1614. doi: 10.1099/jmm.0.081448-0
- Bowers, J. R., Kitchel, B., Driebe, E. M., Maccannell, D. R., Roe, C., Lemmer, D., et al. (2015). Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS One* 10:e0133727. doi: 10.1371/journal.pone.0133727
- Bowers, J. R., Lemmer, D., Sahl, J. W., Pearson, T., Driebe, E. M., Wojack, B., et al. (2016). KlebSeq, a diagnostic tool for surveillance, detection, and monitoring of *Klebsiella pneumoniae*. *J. Clin. Microbiol.* 54, 2582–2596. doi: 10.1128/JCM.00927-16
- Bruen, T. C., Philippe, H., and Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172, 2665–2681. doi: 10.1534/genetics.105.048975
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421
- Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). In silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14
- Carqueira, G. C., Earl, A. M., Ernst, C. M., Grad, Y. H., Dekker, J. P., Feldgarden, M., et al. (2017). Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1135–1140. doi: 10.1073/pnas.1616248114
- Chaturvedi, K. S., Hung, C. S., Crowley, J. R., Stapleton, A. E., and Henderson, J. P. (2012). The siderophore yersiniabactin binds copper to protect pathogens during infection. *Nat. Chem. Biol.* 8, 731–736. doi: 10.1038/nchembio.1020
- Conlan, S., Park, M., Deming, C., Thomas, P. J., Young, A. C., Coleman, H., et al. (2016). Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. *mBio* 7:e742-16. doi: 10.1128/mBio.00742-16
- Conte, V., Monaco, M., Giani, T., D'Ancona, F., Moro, M. L., Arena, F., et al. (2016). Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* from invasive infections in Italy: increasing diversity with predominance of the ST512

FUNDING

This work was funded under the State of Arizona Technology and Research Initiative Fund (TRIF), administered by the Arizona Board of Regents, through Northern Arizona University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.00542/full#supplementary-material>

- clade II sublineage. *J. Antimicrob. Chemother.* 71, 3386–3391. doi: 10.1093/jac/dkw337
- Del Franco, M., Paone, L., Novati, R., Giacomazzi, C. G., Bagattini, M., Galotto, C., et al. (2015). Molecular epidemiology of carbapenem resistant *Enterobacteriaceae* in Valle d'Aosta region, Italy, shows the emergence of KPC-2 producing *Klebsiella pneumoniae* clonal complex 101 (ST101 and ST1789). *BMC Microbiol.* 15:260. doi: 10.1186/s12866-015-0597-z
- Delcher, A. L., Phillippy, A., Carlton, J., and Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 30, 2478–2483. doi: 10.1093/nar/30.11.2478
- Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. D., and Brisse, S. (2005). Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. *J. Clin. Microbiol.* 43, 4178–4182. doi: 10.1128/JCM.43.8.4178-4182.2005
- Didelot, X., and Wilson, D. J. (2015). ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput. Biol.* 11:e1004041. doi: 10.1371/journal.pcbi.1004041
- Duchêne, S., Holt, K. E., Weill, F.-X., Le Hello, S., Hawkey, J., Edwards, D. J., et al. (2016). Genome-scale rates of evolutionary change in bacteria. *Microbial. Genom.* 2:e000094. doi: 10.1099/mgen.0.000094
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- El Fertat-Aissani, R., Messai, Y., Alouache, S., and Bakour, R. (2013). Virulence profiles and antibiotic susceptibility patterns of *Klebsiella pneumoniae* strains isolated from different clinical specimens. *Pathol. Biol.* 61, 209–216. doi: 10.1016/j.patbio.2012.10.004
- Esposito, E. P., Cervoni, M., Bernardo, M., Crivaro, V., Cuccurullo, S., Imperi, F., et al. (2018). Molecular epidemiology and virulence profiles of colistin-resistant *Klebsiella pneumoniae* blood isolates from the hospital agency “Ospedale dei Colli,” Naples, Italy. *Front. Microbiol.* 9:1463. doi: 10.3389/fmicb.2018.01463
- European Committee on Antimicrobial Susceptibility Testing (2016). *Breakpoint Tables for Interpretation of MICs and Zone Diameters*, Version 6.0. Available at: http://www.eucast.org/clinical_breakpoints
- Farris, J. S. (1989). THE RETENTION INDEX AND THE RESCALED CONSISTENCY INDEX. *Cladistics* 5, 417–419. doi: 10.1111/j.1096-0031.1989.tb00573.x
- Fuzi, M., Szabo, D., and Csércsik, R. (2017). Double-serine fluoroquinolone resistance mutations advance major international clones and lineages of various multi-drug resistant bacteria. *Front. Microbiol.* 8:2261. doi: 10.3389/fmicb.2017.02261
- Gaiarsa, S., Comandatore, F., Gaibani, P., Corbella, M., Dalla Valle, C., Epis, S., et al. (2015). Genomic epidemiology of *Klebsiella pneumoniae* in Italy and novel insights into the origin and global evolution of its resistance to carbapenem antibiotics. *Antimicrob. Agents Chemother.* 59, 389–396. doi: 10.1128/AAC.04224-14
- Giani, T., Pini, B., Arena, F., Conte, V., Bracco, S., Migliavacca, R., et al. (2013). Epidemic diffusion of KPC carbapenemase-producing *Klebsiella pneumoniae* in Italy: results of the first countrywide survey, 15 May to 30 June 2011. *Eur. Surveill.* 18:20489.
- Gonçalves, G. B., Furlan, J. P. R., Vespero, E. C., Pelisson, M., Stehling, E. G., and Pitondo-Silva, A. (2017). Spread of multidrug-resistant high-risk *Klebsiella*

- pneumoniae* clones in a tertiary hospital from southern Brazil. *Infect. Genet. Evol.* 56, 1–7. doi: 10.1016/j.meegid.2017.10.011
- Grundmann, H., Glasner, C., Albiger, B., Aanensen, D. M., Tomlinson, C. T., Andrasević, A. T., et al. (2017). Occurrence of carbapenemase-producing *Klebsiella pneumoniae* and *Escherichia coli* in the European survey of carbapenemase-producing *Enterobacteriaceae* (EuSCAPE): a prospective, multinational study. *Lancet Infect. Dis.* 17, 153–163. doi: 10.1016/S1473-3099(16)30257-2
- Gu, D., Dong, N., Zheng, Z., Lin, D., Huang, M., Wang, L., et al. (2018). A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. *Lancet. Infect. Dis.* 18, 37–46. doi: 10.1016/S1473-3099(17)30489-9
- Hancock, V., Ferrieres, L., and Klemm, P. (2008). The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microbiology* 154, 167–175. doi: 10.1099/mic.0.2007/011981-0
- Hansen, D. S., Skov, R., Benedi, J. V., Sperling, V., and Kolmos, H. J. (2002). *Klebsiella* typing: pulsed-field gel electrophoresis (PFGE) in comparison with O:K-serotyping. *Clin. Microbiol. Infect.* 8, 397–404. doi: 10.1046/j.1469-0691.2002.00411.x
- Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., et al. (2015). Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc. Natl. Acad. Sci. U.S.A.* 112, E3574–E3581. doi: 10.1073/pnas.1501049112
- Hsieh, P. F., Lin, T. L., Lee, C. Z., Tsai, S. F., and Wang, J. T. (2008). Serum-induced iron-acquisition systems and TonB contribute to virulence in *Klebsiella pneumoniae* causing primary pyogenic liver abscess. *J. Infect. Dis.* 197, 1717–1727. doi: 10.1086/588383
- Hunt, M., Silva, N. D., Otto, T. D., Parkhill, J., Keane, J. A., and Harris, S. R. (2015). Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 16:294. doi: 10.1186/s13059-015-0849-0
- Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Kent, W. J. (2002). BLAT—The blast-like alignment tool. *Genome Res.* 12, 656–664. doi: 10.1101/gr.229202
- Khomtchouk, B. B., Hennessy, J. R., and Wahlestedt, C. (2017). shinyheatmap: ultra fast low memory heatmap web interface for big data genomics. *PLoS One* 12:e0176334. doi: 10.1371/journal.pone.0176334
- Kleinheinz, K. A., Joensen, K. G., and Larsen, M. V. (2014). Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. *Bacteriophage* 4:e27943.
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Lam, M. M. C., Wick, R. R., Wyres, K. L., Gorrie, C. L., Judd, L. M., Jenney, A. W. J., et al. (2018). Genetic diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in *Klebsiella pneumoniae* populations. *Microb. Genom.* 4:e000196. doi: 10.1099/mgen.0.000196
- Lee, W., Van Baalen, M., and Jansen, V. A. (2012). An evolutionary mechanism for diversity in siderophore-producing bacteria. *Ecol. Lett.* 15, 119–125. doi: 10.1111/j.1461-0248.2011.01717.x
- Logan, L. K., and Weinstein, R. A. (2017). The epidemiology of carbapenem-resistant *Enterobacteriaceae*: the impact and evolution of a global menace. *J. Infect. Dis.* 215, S28–S36. doi: 10.1093/infdis/jiw282
- Ma, L. C., Fang, C. T., Lee, C. Z., Shun, C. T., and Wang, J. T. (2005). Genomic heterogeneity in *Klebsiella pneumoniae* strains is associated with primary pyogenic liver abscess and metastatic infection. *J. Infect. Dis.* 192, 117–128. doi: 10.1086/430619
- Marcusson, L. L., Fridmodt-Møller, N., and Hughes, D. (2009). Interplay in the selection of fluoroquinolone resistance and bacterial fitness. *PLoS Pathog.* 5:e1000541. doi: 10.1371/journal.ppat.1000541
- Monroy, M. A., Knöbl, T., Bottino, J. A., Ferreira, C. S., and Ferreira, A. J. (2005). Virulence characteristics of *Escherichia coli* isolates obtained from broiler breeders with salpingitis. *Comp. Immunol. Microbiol. Infect. Dis.* 28, 1–15. doi: 10.1016/j.cimid.2004.03.001
- Moradigaravand, D., Martin, V., Peacock, S. J., and Parkhill, J. (2017). Evolution and epidemiology of multidrug-resistant *Klebsiella pneumoniae* in the United Kingdom and Ireland. *mBio* 8:e01976-16. doi: 10.1128/mBio.01976-16
- Murray, G. G. R., Wang, F., Harrison, E. M., Paterson, G. K., Mather, A. E., Harris, S. R., et al. (2016). The effect of genetic structure on molecular dating and tests for temporal signal. *Methods Ecol. Evol.* 7, 80–89. doi: 10.1111/2041-210X.12466
- Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17:132. doi: 10.1186/s13059-016-0997-x
- Oteo, J., Pérez-Vázquez, M., Bautista, V., Ortega, A., Zamarrón, P., Saez, D., et al. (2016). The spread of KPC-producing *Enterobacteriaceae* in Spain: WGS analysis of the emerging high-risk clones of *Klebsiella pneumoniae* ST11/KPC-2, ST101/KPC-2 and ST512/KPC-3. *J. Antimicrob. Chemother.* 71, 3392–3399. doi: 10.1093/jac/dkw321
- Papagiannitsis, C. C., Di Pilato, V., Giani, T., Giakkoupi, P., Riccobono, E., Landini, G., et al. (2016). Characterization of KPC-encoding plasmids from two endemic settings, Greece and Italy. *J. Antimicrob. Chemother.* 71, 2824–2830. doi: 10.1093/jac/dkw227
- Pitout, J. D. D., Nordmann, P., and Poirel, L. (2015). Carbapenemase-producing *Klebsiella pneumoniae*, a key pathogen set for global nosocomial dominance. *Antimicrob. Agents Chemother.* 59, 5873–5884. doi: 10.1128/AAC.01019-15
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi: 10.1534/genetics.114.164350
- Rambaut, A., Lam, T. T., Max Carvalho, L., and Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using tempest (formerly Path-O-Gen). *Virus Evol.* 2, vew007. doi: 10.1093/ve/vew007
- Rasko, D. A., Myers, G. S., and Ravel, J. (2005). Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* 6:2. doi: 10.1186/1471-2105-6-2
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584
- Russo, T. A., Shon, A. S., Beanan, J. M., Olson, R., Macdonald, U., Pomakov, A. O., et al. (2011). Hypervirulent *K. pneumoniae* secretes more and more active iron-acquisition molecules than “Classical” *K. pneumoniae* thereby enhancing its virulence. *PLoS One* 6:e26734. doi: 10.1371/journal.pone.0026734
- Sahl, J. W., Caporaso, J. G., Rasko, D. A., and Keim, P. (2014). The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. *PeerJ* 2:e00332. doi: 10.7717/peerj.332
- Sahl, J. W., Lemmer, D., Travis, J., Schupp, J. M., Gillette, J. D., Aziz, M., et al. (2016). NASP: an accurate, rapid method for the identification of SNPs in WGS datasets that supports flexible input and output formats. *Microbi. Genom.* 2:e000074. doi: 10.1099/mgen.0.000074
- Schierup, M. H., and Hein, J. (2000). Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156, 879–891.
- Schliep, K. P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 592–593. doi: 10.1093/bioinformatics/btq706
- Schubert, S., Dufke, S., Sorsa, J., and Heesemann, J. (2004). A novel integrative and conjugative element (ICE) of *Escherichia coli*: the putative progenitor of the Yersinia high-pathogenicity island. *Mol. Microbiol.* 51, 837–848. doi: 10.1046/j.1365-2958.2003.03870.x
- Schubert, S., Rakin, A., Karch, H., Carniel, E., and Heesemann, J. (1998). Prevalence of the “high-pathogenicity island” of Yersinia species among *Escherichia coli* strains that are pathogenic to humans. *Infect. Immun.* 66, 480–485.
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shon, A. S., Bajwa, R. P. S., and Russo, T. A. (2013). Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. *Virulence* 4, 107–118. doi: 10.4161/viru.22718

- Skálová, A., Chudějová, K., Rotová, V., Medvecký, M., Študentová, V., Chudáčková, E., et al. (2016). Molecular characterization of OXA-48-like-producing *Enterobacteriaceae* in the Czech Republic: evidence for horizontal transfer of pOXA-48-like plasmids. *Antimicrob. Agents Chemother.* 61:e1889-16. doi: 10.1128/AAC.01889-16
- Stone, N. E., Sidak-Loftis, L. C., Sahl, J. W., Vazquez, A. J., Wiggins, K. B., Gillice, J. D., et al. (2016). More than 50% of *Clostridium difficile* isolates from pet dogs in flagstaff, USA, carry toxigenic genotypes. *PLoS One* 11:e0164504. doi: 10.1371/journal.pone.0164504
- Struve, C., Roe, C. C., Stegger, M., Stahlhut, S. G., Hansen, D. S., Engelthaler, D. M., et al. (2015). Mapping the evolution of hypervirulent *Klebsiella pneumoniae*. *mBio* 6:e00630. doi: 10.1128/mBio.00630-15
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Toth, A., Kocsis, B., Damjanova, I., Kristof, K., Janvari, L., Paszti, J., et al. (2014). Fitness cost associated with resistance to fluoroquinolones is diverse across clones of *Klebsiella pneumoniae* and may select for CTX-M-15 type extended-spectrum beta-lactamase. *Eur. J. Clin. Microbiol. Infect. Dis.* 33, 837–843. doi: 10.1007/s10096-013-2022-6
- Ventola, C. L. (2015). The antibiotic resistance crisis: part causes and threats. *P T* 40, 277–283.
- Vimont, S., Mnif, B., Fevre, C., and Brisse, S. (2008). Comparison of PFGE and multilocus sequence typing for analysis of *Klebsiella pneumoniae* isolates. *J. Med. Microbiol.* 57, 1308–1310. doi: 10.1099/jmm.0.2008/003798-0
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963
- Wick, R. R., Heinz, E., Holt, K. E., and Wyres, K. L. (2018). Kaptive web: user-friendly capsule and lipopolysaccharide serotype prediction for *Klebsiella* Genomes. *J. Clin. Microbiol.* 56:e197-18. doi: 10.1128/JCM.00197-18
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Woodford, N., Turton, J. F., and Livermore, D. M. (2011). Multiresistant gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiol. Rev.* 35, 736–755. doi: 10.1111/j.1574-6976.2011.00268.x
- Wyres, K. L., Wick, R. R., Gorrie, C., Jenney, A., Follador, R., Thomson, N. R., et al. (2016). Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microbiol. Genom.* 2:e000102. doi: 10.1099/mgen.0.00102

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Roe, Vazquez, Esposito, Zarrilli and Sahl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.