MICROBIAL GENOMICS

Diversity, virulence, and antimicrobial resistance of the KPC-producing *Klebsiella pneumoniae* ST307 clone

ABSTRACT

The global spread of KPC carbapenemase-producing *Klebsiella pneumoniae* has been mainly associated with the dissemination of high-risk clones. Most of the hospital outbreaks reported in the last decade have been attributed to isolates belonging to Clonal Group 258. However, recent epidemiological analysis suggests that a new clone, sequence type (ST) 307, is emerging in different parts of the world and is a candidate to become one of the most prevalent high-risk clones in the near future. Here we show that the ST307 genome encodes features that can provide an advantage in adaptation in the hospital environment and in the human host. These include novel plasmid-located virulence clusters, such as a cluster for glycogen synthesis. Glycogen production is considered one of the possible adaptive responses to long-term survival and growth in environments outside the host. Chromosomally-encoded virulence traits, including fimbriae, an Integrative Conjugative Element carrying the yersiniabactin siderophore, and two different capsular loci were identified. Resistance to complement was verified in capsulated and uncapsulated ST307 strains. The acquired genetic features identified in the genome of this new emerging clone

may contribute to increased persistence of ST307 in the hospital environment, and shed light on its potential epidemiological success.

DATA SUMMARY

1. Whole Genome Shotgun project KP48-IT has been released at DDBJ/ENA/GenBank, accession no.: PRJNA295649 (https://www.ncbi.nlm.nih.gov/bioproject/295649)

2. Whole Genome Shotgun project PRJNA354908 has been deposited at DBJ/ENA/GenBank, accession no. PRJNA354908 (https://www.ncbi.nlm.nih.gov/bioproject/354908), individual accession numbers are listed in Table S1

3. Novel plasmid nucleotide sequences have been deposited in GenBank, accession numbers: pKpQIL_307, KY271403; pKPN3_307_typeA, KY271404; pKPN3_307_typeB, KY271405; pKPN307_TypeC, KY271406; pKPN3_307_TypeD, KY271407; IncN_typeA, KY271413; IncN_typeB, KY271414; IncN_typeC, KY271415: pTet_ 7201, KY271408

4. The complete DNA sequences of the following prophage genomes have been deposited in GenBank, accession number: Prophage1_ST307, KY271401; Prophage2_ST307, KY271396; Prophage2b_ST307, KY271395; Prophage3_ST307, KY271397; Prophage4_ST307, KY271398; Prophage5_ST307, KY271399; Prophage6_ST307, KY271400; Phage48_ST307, KY271402

5. Type A and Type B Integrative Conjugative Elements have been deposited in GenBank; accession numbers KY271411 and KY271412, respectively

6. The π-fimbria cluster sequence has been deposited in GenBank, accession number KY271409

7. The capsula_entero_ST307 cluster sequence has been deposited in GenBank, accession number KY271410.

I/We confirm all supporting data, code and protocols have been provided within the article or through supplementary data files. X

IMPACT STATEMENT

Klebsiella pneumoniae is one of the most relevant cause of healthcare-associated infections. The global spread of carbapenemase-producing *Klebsiella pneumoniae* high-risk clones is a concern. Most of the hospital outbreaks reported in the last decade have been attributed to isolates belonging to Clonal Group 258 but new clones, showing higher mortality and longer hospital stay can emerge.

Recent epidemiological evidences indicate that a new lineage, sequence type 307, has been detected in different parts of the world, in some hospitals displacing the CG258. Here we report the first description of the ST307 genome, studying isolates from different geographical origin. Resistome and mobilome were fully characterized, by complete assembly of resistance and virulence plasmids, identifying integrative conjugative elements associated with the acquisition of the yersiniabactin virulence cluster, and phage content. ST307 resistance to human sera was measured.

ST307 genome encodes plasmid-located and chromosomally-encoded features that can provide an advantage in adaptation in the hospital environment and in the human host.

Some of the genetic features described in this study are novel or very rare for *Klebsiella* spp. and may help in tracing emergence of ST307 in future surveillance studies performed

around the world.

INTRODUCTION

The worldwide spread of carbapenem-producing *Klebsiella pneumoniae* (*Kp*) has become a major threat for healthcare facilities (Wyres & Holt, 2016). This global phenomenon has been mainly associated with the clonal dissemination of high-risk clones. One of the most successful is the KPC-producing *Kp* (KPC-*Kp*) sequence type (ST) 258 clone, and its related variants belonging to Clonal Group 258 (CG258) (Chen *et al*., 2014; Bowers *et al.*, 2014). In recent years new extensively drug-resistant lineages have emerged internationally (Bialek-Davenet et al., 2014; Struve et al., 2015; Wyres & Holt, 2016; Zhou et al., 2016; Ruiz-Garbajosa *et al.* 2013; Giske *et al.* 2012). Among them, *KPC-Kp* ST307 is candidate to become one of the most relevant clones, since its emergence has been recognized in several countries in the last five years (Girlich *et al.,* 2014; Castanheira *et al.,* 2013; Gona *et al.,* 2014; Richter *et al.,* 2012). ST307 was first defined in 2008 in the MLST database (an unpublished isolate), but has since been described in 2013 in the US (Castanheira *et al.,* 2013). It can be hypothesized that ST307 was initially associated with the production of the globally disseminated extended-spectrum beta-lactamase (ESBL) CTX-M-15. The acquisition of KPC enzyme was subsequent to that of CTX-M-15, as deduced by the fact that CTX-M-15-producing *Kp* ST307 were previously reported at high frequencies (70 to 90%) in Italy, Korea, Pakistan, Morocco and in pets from Japan (Girlich *et al.,* 2014; Castanheira *et al.,* 2013; Gona *et al.,* 2014; Richter *et al.,* 2012; Habeeb *et al*., 2013; Harada *et al*., 2016; Park *et al*., 2015).

The Italian experience is particularly interesting because it illustrates what might happen elsewhere in the future, with replacement among KPC producers of CG258 by ST307 in a

short period of time. In fact, the first outbreak of KPC-3-producing *Kp* ST258 in Palermo, Sicily occurred in 2008 (Mammina *et al*., 2010), and then this clone became a key epidemiological feature of many healthcare facilities in this region until 2013 (Giuffrè *et al*., 2013). A surveillance study performed in March-August 2014 in the three largest hospitals of Palermo recognized an epidemiological change, since multifocal dissemination of KPC-3-producing *Kp* clones was observed. In particular, the predominant KPC-3 CG258 clone was identified in 38/94 (40%) patients, but ~~in~~ 27/94 (28%) ~~of them~~ were ST307 producing both KPC-3 and CTX-M-15 ~~was found~~ (Bonura *et al*., 2015; Geraci *et al*., 2015).

In Colombia, a two-year surveillance study was performed from June 2012 to June 2014 in five tertiary-care centers in Medellín, collecting 193 carbapenem-resistant *Kp* strains. Remarkably, whilst 62.2% of isolates were from STs unrelated to CG258, whereas 14.2 % of them were ST307 (Ocampo *et al*., 2016). Patients infected with KPC-*Kp* ST307 presented high mortality (over 50%) and longer hospital stay compared with other clones, supporting the evidence that this lineage is probably possessing additional factors contributing to its epidemiological success. Low prevalence is observed by the National Infection Service, Public Health England in United Kingdom with just eight KPC-*Kp* ST307 isolates identified in a collection of >3000 carbapenem-producing *Kp* isolates from 2014-2016. Interestingly, one of these strains was from a patient transferred to the UK from Italy.

In this work, we performed whole-genome sequencing (WGS) and compared the genetic structures of KPC-*Kp* ST307 isolates from Italy, Colombia, and the UK seeking to identify factors that contribute to the success and spread of the KPC-*Kp* ST307 clone.

METHODS

Clinical isolates and sequenced strains

A total of 24 ST307 KPC-*Kp* isolates were studied. Twelve on 27 ST307 KPC-3-*Kp* isolates collected during the surveillance study performed in March-August 2014 in Palermo (Bonura *et al.*, 2015).These were representatives of the isolates obtained from the three participating hospitals, and were selected on the basis of the slightly different Pulsed-Field Gel Electrophoresis (PFGE) patterns (defined as D1, D4 subtypes; Bonura *et al.*, 2015; Geraci *et al.*, 2015). Among the 12 selected strains, seven and five were CTX-M-15 positive or negative, respectively (Table 1) .

The four Colombian isolates were selected from the collection of 17 KPC-ST307- *Kp* obtained during the two-year surveillance study (Ocampo *et al.*, 2016); two strains were KPC-2 and two KPC-3 positives and came from two different hospitals (Table 1).

The complete genomes of 12 strains were obtained: isolates 48-IT and CIV4-IT were chosen for WGS from the Italian collection as representatives of the KPC-3, CTX-M-15 positives and KPC-3-positive, CTX-M-15 negative isolates, respectively. KL-49-CO and KH-43-CO were chosen for WGS as representatives of the KPC-2 positive strains from Colombia.

In addition, the short-read genomic data of eight KPC-*Kp* isolates from the UK and sequenced by the National Infection Service, Public Health England were also included in the study (Table 1). These were the only ST307 KPC-*Kp* isolates among >3000 CPE genomes, and one of these was from a patient who had been transferred from Italy (strain H155360912-IT). Of these eight KPC-*Kp* isolates, seven belonged to ST307 and one was a single-locus variant of ST307 (H154440769-UK).

Whole-Genome sequencing

~~Strain~~ Genomic DNA was purified from the 48-IT ~~was~~ strain by the Macherey Nagel DNA extraction kit. Plasmid DNA from the 48-IT was purified using the Plasmid Midi Kit

(Invitrogen). Genomic and plasmid DNA were used to prepare two different shot-gun libraries and sequenced ~~(Villa *et al.*, 2016) on~~at the 454-GS platform ~~according to~~following the standard sequencing procedure (Roche Diagnostics) Reads obtained were assembled using the GS-FLX gsAssembler software (Roche Diagnostics).

WGS of strains CIV4-IT, KH-43-CO and KL-49-CO was performed on DNA extracted using the Macherey Nagel kit. Genomic DNA paired-end libraries were generated using the Nextera XT DNA sample preparation kit (Illumina Inc, San Diego, CA, USA) and sequenced using the Illumina MiSeq next generation sequencer with 2x300PE (Illumina Inc). DNA from UK isolates was extracted with a Qiasymphony DSP (Qiagen). DNA libraries were prepared using the Nextera XT sample preparation method and sequenced with a standard 2x100 PE protocol on a HiSeq 2500 instrument (Illumina). De novo assembly was performed using the Galaxy version 20150522 of A5 pipeline through the ARIES public Galaxy server (https://w3.iss.it/site/aries/; Tritt *et al.*, 2012)

Genome annotation ~~and analysis~~

~~Draft genomes were ordered using the MAUVE comparison tool against the *K. pneumoniae* NJST258_2 (NZ_CP006918), KPNIH1 (NZ_ CP008827) and HS11286 (CP003200), reference genomes (Deleo *et al.*~~ Draft genomes were~~2014; Snitnik *et al.*, 2012; Liu *et al.*, 2012),~~ ~~respectively and~~ uploaded to the RAST server (http://rast.nmpdr.org/) and BASYs Server for functional annotation (Van Domselaar *et al.*, 2005).

Antimicrobial resistance gene and replicon content were detected using the ResFinder (Zankari *et al.*, 2012) and PlasmidFinder (Carattoli *et al.*, 2014) tools (https://cge.cbs.dtu.dk/services/).

Virulence genes were identified using the BIGSdb-Kp database, at the Institut Pasteur (http://bigsdb.pasteur.fr/klebsiella/klebsiella.html) and phage sequences using the PHAST

prediction tool (http://phast.wishartlab.com/). Capsular type was deduced by comparing the sequence of the *wzi* gene with those previously described (Brisse *et al.*, 2013).

Genome comparative analysis

The 48-IT was used as reference genome at the SEED Viewer version 2.0 (http://rast.nmpdr.org/seedviewer.cgi) for comparison with the other ST307 WGSs, to identify genetic differences occurring within the ST307 clone (Figure S1).

Using the SEED Viewer version 2.0, and the MAUVE comparison tool, the 48-IT WGS was compared with the *Kp* NJST258_2 (NZ_CP006918), KPNIH1 (NZ_CP008827) and HS11286 (CP003200) reference genomes (Deleo *et al.* 2014; Snitnik *et al.*, 2012; Liu *et al.*, 2012). Results were reconfirmed comparing three representative ST307 strains with the three reference genomes (Figures S2 and S3).

The cut off defining the major differences was: presence or absence (<70% aminoacid identity) of at least >5 consecutive CDSs (Table S2). Analysis of the major differences was manually curated and the respective DNA sequence was analysed by BlastN (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against the NCBI database.

Core genome MLST was performed using the BIGSdb tool (Jolley & Maiden, 2010) installed at Institut Pasteur (http://bigsdb.pasteur.fr) based on the 634-gene strict core-genome MLST scheme of Bialek-Davenet *et al.,* 2014. Classification of strains was performed by the neighbor-joining method using as distance matrix, the proportion of mismatches among allelic profiles of the strict core genome MLST scheme (Gouy *et al.*, 2010).

The integration site of Integrative Conjugative Elements (ICEs) was obtained comparing in silico the 48-IT and H151440671-UK sequences with the genome of the *K. pneumoniae* strain CAV1193 (CP013322; Sheppard *et al*., 2016), which does not carry the yersiniabactin

ICE. ICE groups and tRNA-ASN integration sites followed the classification proposed by Marcoleta *et al*., 2016.

Plasmid reconstruction

Plasmid ~~pKPN-307 type A and pKpQIL-307 plasmids were assembled using the~~ DNA was purified from strain 48-IT and used to transform *Escherichia coli* DH5α chemically competent cells (Invitrogen), selecting on Luria–Bertani agar plates (Sigma), containing ampicillin (50 µg/ml), obtaining bla_{KPC-3}-positive transformants.

Plasmid contigs ~~obtained~~were tested by ~~the GS FLX gsAssembler software.~~ResFinder (Zankari *et al.*, 2012) and PlasmidFinder (Carattoli *et al.*, 2014) tools (https://cge.cbs.dtu.dk/services/). The assembly of the contigs was initially done ~~*in silico* by~~ using the 454 ReadStatus output file, generated by the gsAssembler software (Roche Diagnostics), identifying reads overlapping adjacent contigs. ~~The assembly was confirmed and verified by PCR followed by Sanger DNA sequencing. pKPN-307 type A and pKpQIL-307 were used as DNA sequence reference for *in silico* identification and assembly of contigs in the other ST307 genomes. Plasmid scaffolds showing pair-end overlapping were assembled. The plasmid assembly was confirmed and verified by PCR based gap closure for plasmids identified in CIV4-IT, KL-49-CO and KH-43-CO strains~~Plasmid pKpQIL carrying bla_{KPC-3} was split in 9 contigs and the complete plasmid sequence was reconstructed by PCR-based gap closure method using the 48-IT transformant as DNA template. The pKpQIL-307 of 48-IT strain was submitted in GenBank as prototype of this plasmid type.

~~Plasmids pKPN-307 types B, C and D were obtained from H151440672-UK, H151400611-UK, and KL-49-CO genomes, respectively. Plasmids of IncN types A, B and C were assembled from the genomes of KL-49-CO, H151440671-UK and H151400611-UK, respectively.~~Thirty contigs were identified relative to a pKPN3-like plasmid. The complete sequence was

reconstructed by PCR, closing all the gaps and reconfirming the plasmid map, which was named pKPN-307 type A and submitted to GenBank as prototype of this plasmid type. Apart strain 48-IT, plasmid pKPN-307 type A was identified at 99% identity and 100% coverage in genomes H154440769-UK (43 contigs), KH-43-CO (20 contigs) and KL-49-CO (19 contigs), respectively. Predicted plasmid assembly in these strains was verified checking pair-end overlapping and confirmed by PCR-based closure method.

In the other ST307 genomes, plasmid contigs were identified by ResFinder, PlasmidFinder, and BlastN against the pKPN-307, pKpQIL-307, and R46 IncN plasmids used as reference sequences. The assembly of plasmid contigs predicted by alignment with plasmid references was checked for pair-end overlapping, and some prototypes were confirmed by PCR-based gap closure method. In detail, plasmid pKPN-307 type B was obtained from H151440672-UK (21 contigs, selected as prototype pKPN-307 type B for submission in GenBank), H151300628-UK (22 contigs), H151400611-UK (20 contigs) and H151440672-UK (21 contigs), respectively. Plasmids pKPN-307 type C and D were obtained from H150820806-UK (62 contigs) and CIV4-IT (23 contigs), respectively and submitted in GenBank. Plasmids of IncN types A, and C were fully assembled for GenBank submission from the genomes of KL-49-CO (6 contigs) and H151400611-UK (10 contigs), respectively. Plasmid IncN type B was assembled from H151440671-UK (8 contigs, selected as prototype IncN type B for submission in GenBank), H151300628-UK (12 contigs), H151400610-UK (11 contigs) and H151440672-UK (15 contigs), respectively. Plasmids pTetA and pTetA-QnrB1 were in single scaffolds with complementary pair-ends, in the genomes of H151440672-UK and KH-43-CO strains, respectively.

Manual annotation of complete plasmid sequences was done using Artemis Version 8 (Sanger Institute) in combination with a pairwise alignment, performed by BLASTP homology search (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Serum resistance

Serum resistance assays were performed using 100 µl of *K. pneumoniae* overnight bacterial LB liquid culture, diluted to the final concentration of $1.5x10^4$ cells/ml, mixed with 300 µl of fresh, non-heated human sera obtained from healthy volunteers. A pool of three sera from different volunteers was used for every experiment. The 1:3 bacteria-sera volume ratio mixture was incubated at 37 °C and aliquots of 100 µl were plated on LB agar plates at T_0 and after 30, 60 and 120 min of incubation with sera. Plates were incubated overnight at 37 °C and viable cell counts were determined. The assays were repeated three times using three different pools of sera obtained from different volunteers.

PCR analysis of ST307 specific features.

Specific features identified in the ST307 genomes were screened for by PCR in the entire collection of isolates listed in Table 1, using primer pairs listed in the ~~Supplementary~~ Table ~~2~~S3. PCR assays were performed on total DNAs extracted by Macherey Nagel kit, using the following PCR conditions: 1 cycle of denaturation at 94°C for 5 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at temperature indicated in Supplementary Table1 for 30 s and elongation at 72°C for 1 min. The amplifications were concluded with an extension program of 1 cycle at 72°C for 5 min. The *bla*$_{CTX-M}$ and *wzi* genes were screened using previously described primers and conditions (Brisse *et al.*, 2013; Carattoli *et al.*, 2008) PCR products were fully sequenced.

Plasmid typing was performed using the PCR-Based Replicon Typing Kit (PBRT-KIT, DIATHEVA, Srl, Fano, IT).

RESULTS

Whole-genome sequencing of KPC-*Kp* ST307

A total of 24 ST307 isolates from Italy, Colombia and UK were studied. Twelve isolates, representative of our collections were sequenced (Suppl. Table 1).

Cluster analysis based on 634 strict core-genome MLST genes demonstrated the clear phylogenetic distinction of the ST307 genomes from previously analyzed isolates, [25, 28] showing that they represent a unique sub-lineage (or clonal group) of *K. pneumoniae* very distant from the two ST258 clades (Fig. 1). .

WGS comparative analysis

Comparative analysis of strains from Italy, Colombia and UK evidenced that the major differences in the WGSs of ST307 were in antimicrobial resistance gene complement (resistome), plasmid and phage gene content (mobilome).

The ST307 core genome was highly conserved among strains, while different variants of plasmids and Integrative Conjugative Elements (ICEs) were detected (Figure S1). A total of 4745 common genes and 637 accessory genes, present in at least one of the 12 sequenced isolates, were identified. Among them, 202 were hypothetical proteins with unknown function and 83 were phage-associated proteins. ~~The major differences among the genomes were in antimicrobial resistance gene complement (resistome) and plasmid and phage gene content (mobilome).~~

Comparative analysis using both Seed Viewer (BlastP-based comparison, Figure S2) and MAUVE alignment tool (BlastN-based alignment, Figure S3) performed among the chromosome of 48-IT, NJST258_2 (NZ_CP006918), KPNIH1 (NZ_CP008827) and HS11286 (CP003200) genomes, identified 16 major regions of discontinuity (Table S2). Regions that

were unique of 48-IT, being absent in the other three reference genomes encoded capsules, LPS modification, fimbriae, secretion and efflux systems and were analysed in detail.

ST307 resistome

~~Eighteen~~Sixteen strains were CTX-M-15-positives (all the Colombian, 7/12 Italian and 5/8 UK isolates) and, among them ~~11~~nine were KPC-3 and seven KPC-2 producers (Table 1). The ~~resistance content showed that~~*bla*$_{KPC-3}$ gene was detected in all the Italian ~~isolates had the~~ ~~*bla*$_{KPC-3}$ gene, comprising also~~strains, in the UK ~~isolate~~strain imported from Italy~~, the UK isolates had the~~ and in two strains from Colombia. The *bla*$_{KPC-2}$ gene~~, while the~~ was detected in all strains from UK and in two Colombian isolates ~~had both types.~~. Most of the strains carried additional acquired resistance genes, such as *bla*$_{CTX-M-15}$, *bla*$_{TEM}$, *bla*$_{OXA}$, *aac(3)-IIa, aac(6')Ib-cr* , *qnrB*, *tet*(A), *strAB*, *sul2, dfrA14,* and *catB3*, but the precise complement of resistance genes differed among the isolates. The *bla*$_{SHV-28}$, *oqxAB* and *fosA,* previously described as intrinsic resistance genes in *K. pneumoniae*, were detected in all genomes (Table 2).

ST307 mobilome – plasmids

Plasmids were classified into four major groups on the basis of replicon and resistance gene content:

i) *The KPC-carrying plasmids*.

The *bla*$_{KPC-3}$ gene was always located on pKpQIL-like plasmids, 116,499 ~~bb~~bp in size, highly similar to those previously described in the CG258 clones (Leavitt *et al.*, 2010; Chen *et al.*, 2014) and characterized by the presence of two replicons (FIIk2, FIB-pKpQIL). The *bla*$_{KPC-2}$ gene was located on three different plasmids: pKpQIL-like, IncN and untypable (in the KH-43-CO strain) plasmids (Table 3). Three types of IncN plasmids were identified and named types A, B and C (Gootz

307 *et al.*, 2009). Type A was detected in a Colombian strain while types B and C were
308 found in isolates from UK.
309 IncN type A was 54,345 bp in size, showed the integration of the Tn*5403*-
310 ΔIS*Kpn6*- bla_{KPC-2} –IS*Kpn7* transposon into a class 1 integron, containing the
311 *aac(6')Ib-cr* gene cassette located close to the *uvp1* resolvase gene.
312 IncN type B was 52,848 bp in size, showed the integration of a composite bla_{KPC-2}
313 gene environment, including IS*26*, IS*Ecp1*, a portion of Tn*2*-bla_{TEM-1} transposon
314 and the deleted ΔIS*Kpn6* element of Tn*4401*. This structure was integrated within
315 the *nuc* gene of the IncN plasmid.
316 IncN type C was 55,680 bp in size and was identical to type B except that it
317 carried a Δ3'CS-class 1 integron carrying the *dfrA14* gene cassette close to the
318 *uvp1* resolvase gene.
319 *ii) pKPN-307 plasmids.*
320 Four types of pKPN-307 were identified and named type A to D (Table 1, Fig. 3).
321 Type A pKPN-307 was 227,989 bp in size, contained two replicons (FIIk7, FIB-
322 pKPN3), and a multi-drug resistance region (MDR) of 38 Kb carrying $bla_{CTX-M-15}$,
323 bla_{TEM-1B}, bla_{OXA-1}, *aac(3)-IIa, aac(6')Ib-cr* , *qnrB1*, *strAB*, *sul2, dfrA14,* and *catB3*
324 genes and arsenic, copper and silver resistance clusters. Five putative virulence-
325 encoding clusters were identified: the *lac* operon, the Fec-like iron(III) dicitrate
326 transport system, a glutathione ABC-transport system (Garcia-Fernandez *et al.*,
327 2012) a novel urea ABC-transport system and a novel cluster for glycogen
328 synthesis. The urea transport system included UrtA (urea binding protein), UrtB
329 (urea permease), UrtC (urea transporter) and AmiF (formamidase). The glycogen
330 synthesis cluster included a 4-alpha-glucan branching enzyme, glucose-1-

phosphate adenylyltransferase, glycogen synthase, glycogen phosphorylase and phosphoglucomutase enzymes.

Type B pKPN-307 was 133,069 bp in size, carried the FIB-pKPN3 replicon, arsenic and copper/silver resistance and the same MDR region as found in type A plasmid, but lacked ~~a~~ *catB3* ~~gene~~, *qnrB* and *dfrA14* genes and the FIIk7 replicon and transfer locus. This plasmid carried a urea ABC-transport system and a partial ~~glycogen synthesis cluster.~~glutathione ABC-transport system.

Type C pKPN-307 was 212,319 bp in size. In this plasmid the virulence clusters, transfer and arsenic and copper/silver loci, and replicons were as found in type A, while the MDR region lacked the $bla_{CTX\text{-}M\text{-}15}$, *aac(3)-IIa, strAB* and *sul2* resistance determinants.

Type D pKPN-307 was 116,325 bp in size. Transfer region and replicons were as in type A plasmids but the MDR region carried only the ~~*bla*~~$_{TEM\text{-}1B}$, $bla_{OXA\text{-}1}$, *aac(6')Ib-cr*, *qnrB1* and *dfrA14* genes and a partial copper/silver resistance determinant. None of the five virulence clusters identified in type A plasmids were detected on this plasmid variant.

iii) *The pTet-plasmids*. The Tn*1721*::*tet*(A) element was identified on ~5 Kb plasmids in two strains from the UK and in KL-49-CO from Colombia (Table 1). In strain KH-43-CO a *tet*(A)-*qnrB1* plasmid of 13,262 bp was detected.

iv) *Other large plasmids*.

FIB-M, HIB-M and R replicons were identified in the genomes of two UK isolates (Table 1). It was not possible to get complete assembly of plasmids carrying these replicons and to link them to any resistance gene(s) because of the short contigs

generated by short-read WGS technology and the lack of a proper reference plasmid for the assembly.

The 12 ST307 isolates that were not sequenced were screened by PCR to detect the most relevant plasmid-mediated features identified (Table 1). Overall WGS and PCR results confirmed the frequent association of pKpQIL with KPC-3 and of an IncN plasmid with KPC-2. CTX-M-15 was associated with the pKPN-307 plasmids. Type A pKPN-307 was the most diffused pKPN-307-like plasmid, present in strains from Italy, Colombia and the UK; pKPN-307 types B and C were detected in Italian and UK isolates, whereas pKPN-307 type D was present only in two Italian isolates (CIV57-IT and CIV4-IT).

ST307 mobilome - prophages, phages, and integrative conjugative elements

Six different prophages were identified in the ST307 genomes, with ϕ1 and ϕ2 the most prevalent, being detected in 10/12 genomes. One extrachromosomal phage was identified in strain 48-IT (Table 2, Fig. S1).

Integrative conjugative elements (ICEs) associated with the cluster encoding the yersiniabactin virulence trait (Schubert *et al*., 2004; Lin et al., 2008) were found in 8/12 genomes (Table 2). Two yersiniabactin cluster variants were identified, designated ICE-YB-Type A and B, respectively. Type A and B differed in *ybt*, *irp1*, *irp2*, and *fyuA* alleles (~~Suppl.~~ Table ~~3~~S4) and the ICE structure. Both ICEs were constituted of a Type IV secretion system~~,~~ (T4SS), *mobA* and *mobB* genes but showed a different assortment of associated ORFs, encoding conserved or hypothetical proteins~~. The two ICEs also showed integration into two different tRNA-Asn sites (Menard & Grossman, 2013) within their respective *K.*~~ (Lin *et al.*, 2008). Type A was 99% identity, 96% coverage with the group VI ICE, as previously described in HS11286 reference genome (Marcoleta et al., 2016). Type B showed 98% identity, 86% coverage with group VI, including a 8 Kb region, encoding restriction-methylation enzymes,

ABC transport system and hypothetical proteins. Only four matches (SKGH01-ST147, CP015500.1; CAV1016-ST45, CP017934.1; RJF293-ST374, CP014008.1; *E. coli* ED1a, CU928162.2) among hundreds of genomic sequences available in public sequence databases were identified by BLASTN showing homology with this 8 Kb terminal portion of the ICE, suggesting that this is a new and rare type of ICE element.

The two ICEs also showed integration into two different tRNA-Asn1B (Type A) and tRNA-Asn1D (Type B) sites (Menard & Grossman, 2013; Marcoleta et al., 2016) within their respective *K. pneumoniae* genomes (Fig. 4).

ST307 fimbriae

A region of ~13 Kb was ~~unique to~~identified in the ST307 clone ~~and~~ encoded a π-fimbrial chaperone/usher pathway, including the fimbrial subunit, the usher and chaperone proteins. π-fimbriae were previously described in uropathogenic, piliated *Escherichia coli* (Nuccio & Baumler, 2007). PCR analysis demonstrated that this π-fimbrial cluster was present in all ST307 isolates in our collection (Table 1). This cluster was not present in the ST258 and ST11 genomes, and by BLASTN on the entire GenBank database it was identified only in fifteen *Kp* genomes, belonging to different STs (ST147, ST273, ST392, ST86, ST278, ST37, ST941 and ST442).

Eight additional fimbriae-encoding clusters and the *mrk* cluster coding for type 3 pili were also identified in all ST307 genomes (Suppl. Table 3).

These fimbriae are not unique of the ST307 clone being described in many other *Kp* genomes (Holt *et al.*, 2015) **ST307 capsular loci and resistance to serum complement**

We identified the *wzi*-173 allele, previously associated with the KN2 capsular type (Follador *et al.*, 2016; Pan *et al.*, 2015) in 20/24 ST307 isolates (Table 1). In the genome of CIV4-IT, representative of the four *wzi*-negative strains, an IS*Kpn7* element disrupted the capsular

cluster at the *kpb6* gene. The remaining portion of the cluster was not detected within the genome, suggesting that integration of the insertion sequence was followed by a deep rearrangement causing the deletion of approximately 12 ORFs of the *cps*-cluster (ΔCp1 in Fig. ~~S2).~~

5, panel A). Beside ST307, the *wzi173* was also identified in two *Kp* isolates in the BIGSdb-Kp database (http://bigsdb.pasteur.fr/klebsiella/klebsiella.html) belonging to ST1272 (KP-11 and KP-7) isolated in North America from humans. A second complete cluster potentially encoding a different capsular type was identified in all ST307 genomes (Cp2 in Fig. ~~S2). This cluster was highly homologous to capsular clusters identified in *Enterobacter aerogenes*.~~S5) within a 14 Kb region of discontinuity with respect to the NJST258, HS11286 and KPN1H1 reference genomes. This cluster showed the best homology to capsular clusters identified in *Enterobacter aerogenes* (93% nucleotide identity with *E. aerogenes* strain CAV1320, CP011574). In all the other *Kp* genomes two hypothetical proteins are encoded at this site, while this cluster is present (93% nucleotide identity) in the genome of *Klebsiella quasipneumoniae* strain ATCC 700603 (CP014696.2).

Isolates carrying both Cp1 and Cp2 capsular clusters and those showing the deleted Cp1 cluster were analyzed for complement resistance using a pool of three different fresh human sera from healthy volunteers at a final concentration of 20%. ST258 and ST101 strains were also tested in the same experiments as internal comparators of the experiments. Results showed that ST307 isolates endowed with intact Cp1 and Cp2 clusters were more resistant to complement than ST258, but both were more susceptible than ST101. However, IS*Kpn7*-mediated disruption of the Cp1 cluster strongly affected the complement resistance of *K. pneumoniae*, despite the presence of Cp2, showing 2 Logs of

reduction of cfu in the first hour, and 4 Logs of reduction after two hours of incubation with human sera (Fig. 5).

Other features

Other discontinuity regions were detected comparing ST307 with the reference ST258 and ST11 genomes, encoding sugar transport via the phosphoenolpyruvate phosphotransferase system, antirestriction proteins, toxin-antitoxin systems, ethanolamine utilization and other functions, whose role in fitness and virulence of ST307 cannot be predicted (Table S2). Of note, a Type VI secretion system (T6SS) was identified in all ST307 and by BLASTN only in three other *Kp* genomes (CAV1016-ST45, CP017934; MGH 78578-ST38, CP000647; Kp52.145-ST66, FO834906), and corresponded to the cluster II T6SS, previously described in MGH 78578 (Sarris *et al.*, 2011). ST307 also carried the cluster for the metabolism of the 4-hydroxyproline that exists in collagen, and most bacteria cannot metabolize this hydroxyamino acid (Watanabe et al., 2012). This was detected by BlastN in other 17 *Kp* genomes in GenBank, the majority belonging to the ST147 clone.

DISCUSSION

KPC-*Kp* ST307 is a novel lineage that has potential to become an epidemic or 'high risk' clone. Our analysis revealed that ST307 represents a distinctive clonal group and demonstrates that the main carbapenemase KPC was acquired through horizontal transfer of plasmids. In each country of isolation, the most frequent KPC variant on its respective plasmid type (*i.e.,* pKpQIL-KPC-3 and IncN-KPC-2) moved into ST307 (Garcia-Fernandez et al., 2012; Findlay *et al.*, 2016; Cheng *et al.*, 2016; Snitkin *et al.*, 2012). The acquisition of KPC was probably subsequent to that of CTX-M-15 and this event occurred independently in

different countries after the spread of ST307, as deduced by the fact that strains had distinct KPC plasmids but related CTX-M-15 carrying plasmids.

We found major characteristics that ~~can~~may provide an advantage to this clone in adaptation to the hospital environment and in the human host. Plasmid pKPN-307 is likely one of the crucial players in the evolution of this clone. The largest variant of this plasmid identified in this study (type A) carried five putative virulence clusters: the *lacIZY* operon, the Fec-like iron (III) dicitrate transport and the glutathione ABC-transport system, the urea transport system and the cluster for glycogen synthesis. In *Escherichia coli*, glycogen synthesis is regulated by the stress sigma factor RpoS and is considered one of the possible adaptive responses to long-term survival and growth in environments outside the host (Somorin *et al.*, 2016). It can be hypothesized that plasmid-mediated glycogen synthesis may help ST307 isolates survive under limited nutrient availability and that the urea transport system may facilitate colonization of the urinary tract by this clone. Urinary tract colonization may also be sustained by the unusual π-fimbria identified in all of our ST307 genomes. This kind of fimbria is characteristic of uropathogenic *E. coli* (Menard & Grossman, 2013).ST307 also carries the yersiniabactin siderophore mobilized by an ICE, previously recognized as a relevant and frequent virulence factor in *Kp* (Schubert *et al.*, 2004; Holt *et al.*, ~~*coli* (Menard & Grossman, 2013).~~2015).

~~ST307 also carries a yersiniabactin siderophore mobilized by an ICE, recognized by one major virulence factor in *Kp* (Schubert *et al.*, 2004; Holt *et al.*, 2015).~~

Two different capsular loci were identified in ST307. One is ~~unique to ST307,~~ characterized by the ~~*wzi123*~~*wzi*-173 allele, the second cluster ~~was previously detected in~~is homologous to the *Enterobacter* spp. genomes and has never been described in *Kp*. Capsules are used by microbes to escape the host immune response has been associated with biofilm formation,

protection from desiccation and contributes to serum survival (Miajlovich *et al.*, 2014; Doorduijn *et al.*, 2016). ~~We~~Since there are not functional studies on *cps2*, we cannot predict the role of this addiction capsular locus in ST307 genome. However, we demonstrated that capsulated ST307 isolates endowed with the two ~~functional~~ clusters were more resistant to serum complement than ST258 isolates. Overall, some of the genetic features identified in the ~~genome~~ ST307 genome despite the lack of a formal functional validation, are interesting and rare and may contribute to ~~increased~~increase fitness, persistence and adaptation of this clone in the hospital environment and in the human host ~~and shed light on the potential epidemiological success of ST307.~~..

AUTHOR STATEMENTS

Funding information

This study was funded by the Italian National Center for Disease Prevention and Control of the Italian Minister of Health (ISS:J4B; CCM 2014), and the Swiss National Science Foundation (32003B_153377).

Conflicts of interest

Authors declare no conflicts of interest in the manuscript. However, their Institutions receive royalties from DIATHEVA S.r.l. (Fermo, Italy) and financial support for conference attendance, lectures, research projects or contracted evaluations from numerous sources, including: Accelerate Diagnostics, Achaogen Inc, Allecra Therapeutics, Amplex, AstraZeneca UK Ltd, Basilea Pharmaceutica, Becton Dickinson Diagnostics, bioMérieux, Bio-Rad Laboratories, The BSAC, Cepheid, Check-Points B.V., Cubist Pharmaceuticals, Department of Health, Enigma Diagnostics, Food Standards Agency, GlaxoSmithKline Services Ltd, Henry

Stewart Talks, IHMA Ltd, Kalidex Pharmaceuticals, Melinta Therapeutics, Merck Sharpe & Dohme Corp, Meiji Seika Pharmo Co., Ltd, Mobidiag, Momentum Biosciences Ltd, Nordic Pharma Ltd, Norgine Pharmaceuticals, Rempex Pharmaceuticals Ltd, Roche, Rokitan Ltd, Smith & Nephew UK Ltd, Trius Therapeutics, VenatoRx Pharmaceuticals and Wockhardt Ltd.

ABBREVIATIONS

KPC: *Klebsiella pneumoniae* carbapenemase; KPC-*Kp:* KPC-producing *Klebiella pneumoniae;* ST: Sequence type; CG: Clonal Group; MLST: Multi-Locus Sequence typing; PBRT: PCR-Based Replicon Typing; ESBL: Extended-Spectrum Beta-Lactamase; IS: insertion sequence; CDS: coding sequence; ICE: Integrative Conjugative Element; WGS: Whole Genome Sequence; T4SS: Type IV secretion system; T6SS: Type VI secretion system.

REFERENCES

Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Jones, L., Delannoy-Vieillard,A. S., Garin, B., Le Hello, S., Arlet, G. & other authors. (2014). Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis* **20**, 1812–1820.

Bonura, C., Giuffre, M., Aleo, A., Fasciana, T., Di Bernardo, F., Stampone, T., Giammanco, A.; MDR-GN Working Group, Palma, D.M. & Mammina, C. (2015). An Update of the Evolving Epidemic of blaKPC Carrying *Klebsiella pneumoniae* in Sicily, Italy, 2014: Emergence of Multiple Non-ST258 Clones. *PLoSOne* 10:e0132936.

Bowers, J. R., Kitchel, B., Driebe, E. M., MacCannell, D. R., Roe, C., Lemmer, D., de Man, T., Rasheed, J. K., Engelthaler, D. M. & other authors. (2015). Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoSOne* **10**, e0133727.

Brisse, S., Passet, V., Haugaard, A. B., Babosan, A., Kassis-Chikhani, N., Struve, C. & Decré, D. (2013). *wzi* gene sequencing, a rapid method for determination of capsular type for *Klebsiella* strains. *J Clin Microbiol* **51**, 4073–4078.

Carattoli, A., García-Fernández, A., Varesi, P., Fortini, D., Gerardi, S., Penni, A., Mancini, C. & Giordano, A. (2008). Molecular epidemiology of *Escherichia coli* producing extended-spectrum beta-lactamases isolated in Rome, Italy. *J Clin Microbiol* **46**:103-108.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F. & Hasman H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* **58**:3895-3903.

Castanheira, M., Farrell, S. E., Wanger, A., Rolston, K.V., Jones, R. N. & Mendes, R. E. (2013). Rapid expansion of KPC-2-producing *Klebsiella pneumoniae* isolates in two Texas hospitals due to clonal spread of ST258 and ST307 lineages. *Microb Drug Resist* **19**:295-297.

Chen, L., Chavda, K. D., Melano, R. G., Jacobs, M.R., Koll, B., Hong, T., Rojtman, A. D., Levi, M. H., Bonomo, R. A. & Kreiswirth, B. N. (2014). Comparative genomic analysis of KPC-encoding pKpQIL-like plasmids and their distribution in New Jersey and New York Hospitals. *Antimicrob Agents Chemother* **58**:2871-2877.

Chen, L., Mathema, B., Pitout, J. D., DeLeo, F. R. & Kreiswirth, B. N. (2014). Epidemic *Klebsiella pneumoniae* ST258 Is a hybrid strain. *MBio* **5**, e01355-14.

Cheng, L., Cao, X. L., Zhang, Z. F., Ning, M. Z., Xu, X. J., Zhou, W., Chen, J. H., Zhang, J. H., Shen, H. & Zhang, K. (2016). Clonal dissemination of KPC-2 producing *Klebsiella pneumoniae* ST11 clone with high prevalence of oqxAB and rmtB in a tertiary hospital in China: results from a 3-year period. *Ann Clin Microbiol Antimicrob* **15**:1.

Deleo, F. R., Chen, L., Porcella, S. F., Martens, C. A., Kobayashi, S. D., Porter, A. R., Chavda, K. D., Jacobs, M. R., Mathema, B. & other authors. **(2014).** Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A* **111**, 4988–4993.

Doorduijn, D. J., Rooijakkers, S. H. M., van Schaik, W. & Bardoel, B. W. (2016). Complement resistance mechanisms of *Klebsiella pneumoniae. Immunology* **221**: 1102-1109

Findlay, J., Hopkins, K. L., Doumith, M., Meunier, D., Wiuff, C., Hill, R., Pike, R., Loy, R., Mustafa, N., Livermore, D. M. & Woodford N. (2016). KPC enzymes in the UK: an analysis of the first 160 cases outside the North-West region. *J Antimicrob Chemother* **71**:1199-1206.

Follador, R., Heinz, E., Wyres, K. L., Ellington, M. J., Kowarik, M., Holt, K. E. & Thomson, N. R. (2016). The diversity of *Klebsiella pneumoniae* surface polysaccharides. MGen 2.

García-Fernández, A., Villa, L., Carta, C., Venditti, C., Giordano, A., Venditti, M., Mancini, C. & Carattoli, A. (2012). Klebsiella pneumoniae ST258 producing KPC-3 identified in italy carries novel plasmids and OmpK36/OmpK35 porin variants. *Antimicrob Agents Chemother* **56**:2143–2145.

Geraci, D. M., Bonura, C., Giuffrè, M., Saporito, L., Graziano, G., Aleo, A., Fasciana, T., Di Bernardo, F., Stampone, T., Palma, D. M. & Mammina, C. (2015). Is the monoclonal spread of the ST258, KPC-3-producing clone being replaced in southern Italy by the dissemination of multiple clones of carbapenem-nonsusceptible, KPC-3-producing *Klebsiella pneumoniae*? *Clin Microbiol Infect* **21**:e15-17.

Girlich, D., Bouihat, N., Poirel, L., Benouda, A. & Nordmann P. (2014). High rate of faecal carriage of extended-spectrum beta-lactamase and OXA-48 carbapenemase-producing Enterobacteriaceae at a university hospital in Morocco. *Clin Microbiol Infect* **20**:350-354.

Giske, C. G., Froding, I., Hasan, C. M., Turlej-Rogacka, A., Toleman, M., Livermore, D., Woodford, N. & Walsh T. R. (2012). Diverse sequence types of *Klebsiella pneumoniae* contribute to the dissemination of blaNDM-1 in India, Sweden, and the United Kingdom. *Antimicrob Agents Chemother* **56**:2735-2738.

Giuffrè, M., Bonura, C., Geraci, D. M., Saporito, L., Catalano, R., Di Noto, S., Nociforo, F., Corsello, G. & Mammina C. (2013). Successful control of an outbreak of colonization by *Klebsiella pneumoniae* carbapenemase-producing *K. pneumoniae* sequence type 258 in a neonatal intensive care unit, Italy. *J Hosp Infect* **85**:233-236.

Gona, F., Barbera, F., Pasquariello, A. C., Grossi, P., Gridelli, B., Mezzatesta, M. L., Caio, C., Stefani, S. & Conaldi, P. G. (2014). In vivo multiclonal transfer of bla(KPC-3) from *Klebsiella pneumoniae* to *Escherichia coli* in surgery patients. *Clin Microbiol Infect* **20**:O633-635.

Gootz, T. D., Lescoe, M. K., Dib-Hajj, F., Dougherty, B. A., He, W., Della-Latta, P. & Huard, R. C. (2009). Genetic organization of transposase regions surrounding blaKPC carbapenemase genes on plasmids from *Klebsiella* strains isolated in a New York City hospital. *Antimicrob Agents Chemother* **53**:1998-2004.

Gouy, M., Guindon, S. & Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**:221-224.

Habeeb, M. A., Haque, A., Nematzadeh, S., Iversen, A. & Giske, C. G. (2013). High prevalence of 16S rRNA methylase RmtB among CTX-M extended-spectrum beta-lactamase-producing Klebsiella pneumoniae from Islamabad, Pakistan. *Int J Antimicrob Agents* **41**:524-526.

Harada, K., Shimizu, T., Mukai, Y., Kuwajima, K., Sato, T., Usui, M., Tamura, Y., Kimura, Y., Miyamoto, T. & other authors. (2016). Phenotypic and Molecular Characterization of Antimicrobial Resistance in Klebsiella spp. Isolates from Companion Animals in Japan: Clonal Dissemination of Multidrug-Resistant Extended-Spectrum beta-Lactamase-Producing *Klebsiella pneumoniae*. *Front Microbiol* **7**:1021.

Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., Jenney, A., Connor,T. R., Hsu, L. Y. & other authors. **(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**, E3574-81.

Jolley, K.A. & Maiden, M. C. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**:595.

Leavitt, A., Chmelnitsky, I., Ofek, I., Carmeli, Y. & Navon-Venezia, S. (2010). Plasmid pKpQIL encoding KPC-3 and TEM-1 confers carbapenem resistance in an extremely drug-resistant epidemic

Klebsiella pneumoniae strain. *J Antimicrob Chemother* **65**:243-248. **Lin, T.L., Lee, C. Z., Hsieh, P. F., Tsai, S. F. & Wang, J. T. (2008).** Characterization of integrative and conjugative element ICEKp1-associated genomic heterogeneity in a *Klebsiella pneumoniae* strain isolated from a primary liver abscess. J Bacteriol **190**: 515-526.

Liu, P., Li, P., Jiang, X., Bi, D., Xie, Y., Tai, C., Deng, Z., Rajakumar, K. & Ou, H. Y. (2012). Complete Genome Sequence of *Klebsiella pneumoniae* subsp. pneumoniae HS11286, a Multidrug-Resistant Strain Isolated from Human Sputum. *J. Bacteriol* **194**:1841-1842

Mammina, C., Palma, D. M., Bonura, C., Anna Plano, M. R., Monastero, R., Sodano, C., Calà, C. & Tetamo, R. (2010). Outbreak of infection with *Klebsiella pneumoniae* sequence type 258 producing Klebsiella pneumoniae Carbapenemase 3 in an intensive care unit in Italy. *J Clin Microbiol*. **48**:1506-1507.

Marcoleta, A. E., Berríos-Pastén, C., Nuñez, G., Monasterio, O. & Lagos, R. (2016). Klebsiella pneumoniae Asparagine tDNAs Are Integration Hotspots for Different Genomic Islands Encoding Microcin E492 Production Determinants and Other Putative Virulence Factors Present in Hypervirulent Strains. Front Microbiol **7**: 1-17.

Mathers, A. J., Stoesser, N., Sheppard, A. E., Pankhurst, L., Giess, A., Yeh, A. J., Didelot, X., Turner, S. D., Sebra, R., & other authors. (2015). Klebsiella pneumoniae carbapenemase (KPC)-producing *K. pneumoniae* at a single institution: insights into endemicity from whole-genome sequencing. *Antimicrob Agents Chemother* **59**:1656-1663

Menard, K. L. & Grossman, A. D. (2013). Selective pressures to maintain attachment site specificity of integrative and conjugative elements. *PLoS Genet* **9**:e1003623.

Miajlovic, H. & Smith, S. G. (2014). Bacterial self-defence: how Escherichia coli evades serum killing. *FEMS Microbiol Lett*. **354**:1-9.

Nuccio, S. P. & Bäumler, A. J. (2007). Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev*. **71**:551-575.

Ocampo, A. M., Chen, L., Cienfuegos, A. V., Roncancio, G., Chavda, K. D., Kreiswirth, B. N. & Jiménez, J. N. (2016). A Two-Year Surveillance in Five Colombian Tertiary Care Hospitals Reveals High Frequency of Non-CG258 Clones of Carbapenem-Resistant *Klebsiella pneumoniae* with Distinct Clinical Characteristics. *Antimicrob Agents Chemother* **60**:332-342.

Pan, Y. J., Lin, T. L., Chen, C. T., Chen, Y. Y., Hsieh, P. F., Hsu, C. R., Wu, M. C. & Wang, J. T. (2015). Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of *Klebsiella* spp. *Nat Sci Rep* **5**, 15573.

Sarris, P. F., Zoumadakis, C., Panopoulos, N. J., Scoulica, E.V. (2011). Distribution of the putative type VI secretion system core genes in *Klebsiella* spp. Infect Genet Evol. **11**:157-66.

Park, D. J., Yu, J. K., Park, K. G. & Park, Y. J. (2015). Genotypes of Ciprofloxacin-Resistant Klebsiella pneumoniae in Korea and Their Characteristics According to the Genetic Lineages. *Microb Drug Resist* **21**:622-630.

Richter, S. N., Frasson, I., Franchin, E., Bergo, C., Lavezzo, E., Barzon, L., Cavallaro, A. & Palù G. (2012). KPC-mediated resistance in *Klebsiella pneumoniae* in two hospitals in Padua, Italy, June 2009-December 2011: massive spreading of a KPC-3-encoding plasmid and involvement of non-intensive care units. *Gut Pathog* **4**:7.

Ruiz-Garbajosa, P., Curiao, T., Tato, M., Gijon, D., Pintado, V., Valverde, A., Baquero, F., Morosini, M. I., Coque, T. M. & Cantón, R. (2013). Multiclonal dispersal of KPC genes following the emergence of non-ST258 KPC-producing *Klebsiella pneumoniae* clones in Madrid, Spain. *J Antimicrob Chemother* **68**:2487-2492.

Schubert, S., Dufke, S., Sorsa, J. & 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.

Sheppard, A. E., Stoesser, N., Sebra, R., Kasarskis, A., Deikus, G., Anson, L., Walker, A. S., Peto, T. E., Crook, D. W. & Mathers, A. J. (2016). Complete Genome Sequence of KPC-Producing *Klebsiella pneumoniae* Strain CAV1193. *Genome Announc* **4**, e01649-15.

Snitkin, E. S., Zelazny, A. M., Thomas, P. J., Stock, F., Program, N. C. S., Henderson, D. K., Palmore, T. N. & Segre, J. A. (2012). Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med* **4**, 148ra116.

Somorin, Y., Abram, F., Brennan, F. & O'Byrne, C. (2016). The General Stress Response Is Conserved in Long-Term Soil-Persistent Strains of *Escherichia coli*. *Appl Environ Microbiol* **82**:4628-4640.

Struve, C., Roe, C. C., Stegger, M., Stahlhut, S. G., Hansen, D. S., Engelthaler, D. M., Andersen, P. S., Driebe, E. M., Keim, P. & Krogfelt, A. (2015). Mapping the evolution of hypervirulent Klebsiella pneumoniae. MBio 6, 1-12.

Tritt, A., Eisen, J. A., Facciotti, M.T. & Darling, A. E. (2012). An integrated pipeline for de novo assembly of microbial genomes. *PLoS One* **7**:e42304.

Van Domselaar, G. H., Stothard, P., Shrivastava, S., Cruz, J. A., Guo, A., Dong, X., Lu, P., Szafron, D., Greiner, R. & Wishart, D. S. (2005). BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res.* **33**:W455-9.

Villa, L., Feudi, C., Fortini, D., Iacono, M., Bonura, C., Endimiani, A., Mammina, C. & Carattoli A. (2016). Complete Genome Sequence of KPC-3- and CTX-M-15-Producing *Klebsiella pneumoniae* Sequence Type 307. *Genome Announc* **4**:17.

Watanabe, S., Morimoto, D., Fukumori, F., Shinomiya,H., Nishiwaki,H., Kawano-Kawada,M., Sasai,Y., Tozawa,Y., and Watanabe, Y. (2012). Identification and Characterization of Δ-Hydroxyproline Dehydrogenase and Δ1-Pyrroline-4-hydroxy-2-carboxylate Deaminase Involved in Novel L-Hydroxyproline Metabolism of Bacteria. J. Biol. Chem. 287: 32674–32688.

Wyres, K. L. & Holt, K. E. (2016). *Klebsiella pneumoniae* Population Genomics and Antimicrobial-Resistant Clones. *Trends Microbiol* **24**:944-956.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F. M. & Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* **67**:2640-2644.

Zhou, K., Lokate, M., Deurenberg, R. H., Tepper, M., Arends, J. P., Raangs, E. G. C., Lo-Ten-Foe, J., Grundmann, H., Rossen, J. W. A. & Friedrich, A. W. (2016). Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum β-lactamase producing ST15 Klebsiella pneumoniae. Sci Rep **6**: 20840.

DATA BIBLIOGRAPHY

1. XXX, XXXX, XXXX et al. Complete Genome Sequence of KPC-3- and CTX-M-15-Producing Klebsiella pneumoniae Sequence Type 307. Genome Announc. 2016 Apr 7;4(2). pii: e00213-16. NCBI Bioproject PRJNA295649

2. XXXX, XXXXXX. NCBI Bioproject PRJNA354908

3. Deleo, F. R., Chen, L., Porcella, S. F., Martens, C. A., Kobayashi, S. D., Porter, A. R., Chavda, K. D., Jacobs, M. R., Mathema, B. & other authors. Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A* 111, 4988–4993. European Nucleotide PRJNA237670 (2014)

4. Snitkin, E. S., Zelazny, A. M., Thomas, P. J., Stock, F., Program, N. C. S., Henderson, D. K., Palmore, T. N. & Segre, J. A. (2012). Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med* **4**, 148ra116. NCBI Bioproject PRJNA224116 (2012)

5. Liu,P., Li,P., Jiang,X., Bi,D., Xie,Y., Tai,C., Deng,Z.,Rajakumar,K. and Ou,H.Y.Complete Genome Sequence of Klebsiella pneumoniae subsp. pneumoniae HS11286, a Multidrug-Resistant Strain Isolated from Human Sputum J. Bacteriol. 194 (7), 1841-1842 (2012). NCBI BioProject PRJNA78789 (2012)

6. Sheppard,A.E., Stoesser,N., Sebra,R., Kasarskis,A., Deikus,G.,Anson,L., Walker,A.S., Peto,T.E., Crook,D.W. and Mathers,A.J. (2016) Complete Genome Sequence of KPC-Producing Klebsiella pneumoniae Strain CAV1193. Genome Announc 4, e01649-15. NCBI BioProject PRJNA246471

7. XXXX & XXXX, Genbank KY271395-KY271415 (2016)

8. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. (2012) Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother.* 67: *2640-4.* Database: https://cge.cbs.dtu.dk//services/data.php

9. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. (2014) In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother.* 58:3895-903. Database: https://cge.cbs.dtu.dk//services/data.php

10. Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Jones, L., Delannoy-Vieillard, A. S., Garin, B., Le Hello, S., Arlet, G. & other authors. Genomic definition of hypervirulent and multidrug resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis* 20. European Nucleotide Archive PRJEB6688 (2014)

FIGURES LEGENDS

Figure 1. Unrooted neighbor-joining tree of *K. pneumoniae* core genome

Unrooted neighbor-joining tree is based on the proportion of mismatches among allelic profiles of the strict core genome MLST scheme. Numbers at the tip of branches are sequence types. The positions of ST307 and the two ST258 clades are indicated by orange and green boxes, respectively

Figure 2. KPC-positive plasmids identified in ST307

White arrows indicated plasmid scaffold genes and their direction of transcription. The locus Tra is indicated by a squared white arrow with capitol letters indicating the respective *tra* genes (i.e. J: *traG*, G; *traF*, F; *traO,* O etc..). Resistance genes are indicated by orange coloured arrows. Transposon-related genes [*tnpA, tnpR, tnpM*], class 1 integrase and insertion sequences are indicated by red arrows. Other genes are indicated by coloured boxes as follows: violet, replicase genes; grey, restriction enzyme and DNA methylase genes; green, *ccg* cluster; blue *fipA* and *nuc* genes .

Figure 3. Variant pKPN-307 plasmids identified in ST307

White arrows indicate plasmid scaffold genes and their direction of transcription. The locus Tra is indicated by a squared white arrow with capital letters indicating the respective *tra* genes (i.e. J: *traG*, G; *traF*, F; *traO*, O etc..). Resistance genes are indicated by orange coloured arrows. Transposon-related genes [*tnpA, tnpR, tnpM*], class 1 integrase and insertion sequences are indicated by red arrows. Other genes are indicated by colored boxes as follows: violet, replicase genes; green, clusters encoding putative virulence determinants.

Figure 4. Integrative Conjugative Elements mobilizing the Yersiniabactin cluster

Type A ICE (group VI as defined by Marcoleta et al., 2016) identified in strain 48-IT and type B ICE (group VI-like) identified in strain H151440671-UK are drawn indicating their integration site with respect to the tRNA genes (tRNA Asn1A, 1B, 1C and 1D as described in Marconeta et al., 2016) as detected in the complete genome sequence of strain CAV1193 that does not contain ICEs. Arrows indicate genes and their direction of transcription. Colours indicate clusters encoding the yersiniabactin system (brown), Type IV secretion system (green), hyphotetical proteins (blue) and other ICE associated genes (yellow), respectively. Abbreviations: R:Restriction, M: Methylation; RT: Reverse Transcriptase.

Figure 5. Complement resistance of ST307

Bars represent **s**erum resistance results performed using fresh, non-heated human sera obtained from healthy volunteers on ST307 strains 48-I (blue), representative of strains carrying both Cp1 and Cp2 capsular loci, and CIV4-IT (white) as representative of strains carrying ΔCp1 and Cp2 cluters. As comparators strain ST258 (green) and ST101 (orange) were also tested. Colony-forming units were measured immediately after 1:3 mixture with sera (T_0) and after 30 (T30), 60 (T60) and 120 (T120) min of incubation.

Table 1. Characteristics and features of the ST307 *K. pneumoniae* isolates from three countries

Strain[a]	Country	Carbapenemase and ESBL		Plasmids						Capsule		π-fimbria
				Replicons	Urea	Glycogen	pKPN307 type	pKpQIL KPC-3	IncN KPC-2 type	Cp1 wzi	Cp2	
48-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
CIV2-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
CIV10-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
VSC1-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
CIV13-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
CIV66-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	neg	B	pos	neg	173	pos	pos
CIV65-IT	IT	KPC-3	CTX-M-15	FIIK,FIBK	pos	neg	B	pos	neg	173	pos	pos
KH-24-CO	CO	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	neg	neg	173	pos	pos
KH-37-CO	CO	KPC-3	CTX-M-15	FIIK,FIBK	pos	pos	A	neg	neg	173	pos	pos
KH-43-CO	CO	KPC-2	CTX-M-15	N, FIIK,FIBK	pos	pos	A	neg	neg	173	pos	pos
KL-49-CO	CO	KPC-2	CTX-M-15	N, FIIK,FIBK	pos	pos	A	neg	A	173	pos	pos
H151300628-UK	UK	KPC-2	CTX-M-15	N,FIIK,FIBK	pos	neg	B	neg	B	173	pos	pos
H151400610-UK	UK	KPC-2	CTX-M-15	N, FIIK,FIBK	pos	neg	B	neg	B	173	pos	pos
H151400611-UK	UK	KPC-2	CTX-M-15	N, FIIK, FIBK, FIBM,HIBM	pos	neg	B	neg	C	173	pos	pos

H151440672-UK	UK	KPC-2	CTX-M-15	N, FIIK,FIBK	pos	neg	B	neg	B	173	pos	pos
H154440769-UK	UK	KPC-2	CTX-M-15	FIIK,FIBK	pos	pos	A	pos	neg	173	pos	pos
H150820806-UK	UK	KPC-2	neg	FIIK,FIBK	pos	pos	C	pos	neg	173	pos	pos
CIV57-IT	IT	KPC-3	neg	FIIK,FIBK	neg	neg	D	pos	neg	173	pos	pos
H155360912-IT	UK	KPC-3	neg	FIIK,FIBK,R	neg	neg	neg	pos	neg	173	pos	pos
H151440671-UK	UK	KPC-2	neg	N	neg	neg	neg	neg	B	173	pos	pos
CIV4-IT	IT	KPC-3	neg	FIIK,FIBK	neg	neg	D	pos	neg	neg	pos	pos
CIV78-IT	IT	KPC-3	neg	FIIK,FIBK	pos	pos	C	pos	neg	neg	pos	pos
21-IT	IT	KPC-3	neg	FIIK,FIBK	pos	pos	C	pos	neg	neg	pos	pos
VSC21-IT	IT	KPC-3	neg	FIIK,FIBK	pos	neg	B	pos	neg	neg	pos	pos

[a] Whole Genome Sequencing was performed for underlined strains

Table 2. Resistome, Integrative Conjugative Elements, prophages and phages in ST307 genomes

Strain	Beta-lactamases	Aminoglycosides	Quinolones	Others	ICE-YB	Prophages						Phage
48-IT	*bla*$_{KPC-3}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{OXA-9}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, catB3, fosA*	A	Φ1	Φ2c					Φ48
CIV4-IT	*bla*$_{KPC-3}$, *bla*$_{TEM-1A}$, *bla*$_{OXA-1}$, *bla*$_{OXA-9}$, *bla*$_{SHV-28}$	*aac(6')Ib-cr*	*qnrB1, oqxAB*	*dfrA14, catB3, fosA*	A	Φ1	Φ2c					
KH-43-CO	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, tetA, catB3, fosA*	neg	Φ1	Φ2b		Φ4	Φ5	Φ6	
KL-49-CO	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul1, sul2, dfrA14, tetA, catB3, fosA*	neg	Φ1	Φ2b	Φ3	Φ4			
H150820806-UK	*bla*$_{KPC-2}$, *bla*$_{TEM-1A}$, *bla*$_{OXA-1}$, *bla*$_{OXA-9}$, *bla*$_{SHV-28}$	*aac(6')Ib-cr*	*qnrB1, oqxAB*	*dfrA14, tetA, catB3, fosA*	neg	Φ1		Φ3	Φ4			
H154440769-UK	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{OXA-9}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, tetA, catB3, fosA*	neg	Φ1						
H151300628-UK	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, tetA, fosA*	B	Φ1	Φ2					
H151400610-UK	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$,	*strA, strB, aac(3)-IIa,*	*qnrB1,*	*sul2, dfrA14,*	B	Φ1	Φ2					

	bla$_{SHV-28}$	*aac(6')Ib-cr*	*oqxAB*	*tetA, fosA*								
H151400611-UK	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, tetA, fosA*	B	Φ1	Φ2					
H151440672-UK	*bla*$_{KPC-2}$, *bla*$_{CTX-M-15}$, *bla*$_{TEM-1B}$, *bla*$_{OXA-1}$, *bla*$_{SHV-28}$	*strA, strB, aac(3)-IIa, aac(6')Ib-cr*	*qnrB1, oqxAB*	*sul2, dfrA14, tetA, fosA*	B	Φ1	Φ2					
H155360912-IT	*bla*$_{KPC-3}$, *bla*$_{TEM-1A}$, *bla*$_{OXA-9}$, *bla*$_{SHV-28}$		*oqxAB*	*fosA*	B		Φ2					
H151440671-UK	*bla*$_{KPC-2}$, *bla*$_{TEM-1B}$, *bla*$_{SHV-28}$		*oqxAB*	*fosA*	B	Φ1	Φ2					