Editor: Stefania Stefani

Journal Pre-proof

Genome stability during serial sub-culturing in hyperepidemic multidrug-resistant Klebsiella pneumoniae and Escherichia coli

Aline I. Moser , Edgar I. Campos-Madueno , Vincent Perreten , Andrea Endimiani

PII: S2213-7165(22)00204-1
DOI: https://doi.org/10.1016/j.jgar.2022.08.014
Reference: JGAR 1940

To appear in: *Journal of Global Antimicrobial Resistance*

Received date: 10 June 2022
Revised date: 19 July 2022
Accepted date: 9 August 2022

Please cite this article as: Aline I. Moser , Edgar I. Campos-Madueno , Vincent Perreten , Andrea Endimiani , Genome stability during serial sub-culturing in hyperepidemic multidrug-resistant Klebsiella pneumoniae and Escherichia coli, *Journal of Global Antimicrobial Resistance* (2022), doi: https://doi.org/10.1016/j.jgar.2022.08.014

Highlights

- We determined the genomic changes of 4 different MDR isolates after 20 propagation steps on 2 types of agar plates
- Hybrid whole-genome sequencing was implemented to focus on the single nucleotide variants (SNVs)
- An ST101 *K. pneumoniae* lost 2 plasmidic carbapenemases (OXA-48, NDM-1) and showed 1-12 SNVs
- A CTX-M-15-producing *K. pneumoniae* of ST307 was structurally stable and showed only 1-2 SNVs
- Both OXA-181- (ST410) and NDM-5-producing (ST167) *E. coli* were stable and showed 0-9 SNVs

Genome stability during serial sub-culturing in hyperepidemic multidrug-resistant *Klebsiella pneumoniae* and *Escherichia coli*

Aline I. Moser,[1] Edgar I. Campos-Madueno,[1,2] Vincent Perreten,[3] Andrea Endimiani [1,*]

[1]Institute for Infectious Diseases, University of Bern, Bern, Switzerland; [2]Graduate School of Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland; [3]Institute of Veterinary Bacteriology, University of Bern, Bern, Switzerland

Short running title: *In vitro* genetic stability of MDR *E. coli* and *K. pneumoniae*

***Corresponding author:**

Prof. Andrea Endimiani, MD, PhD

Institute for Infectious Diseases - University of Bern

Friedbühlstrasse 51, CH-3001, Switzerland

Phone: +41-31-632-8632; Fax +41-31-632-8766

E-mails: aendimiani@gmail.com; andrea.endimiani@ifik.unibe.ch

ABSTRACT

Background: Core-genome single-nucleotide variant (cgSNV) analysis represents a powerful tool for epidemiological investigations of multidrug-resistant (MDR) bacteria. However, cgSNV thresholds to confirm whether isolates are the same clone are not formally defined.

Methods: We implemented hybrid whole-genome sequencing to study the genomic changes of 4 MDR isolates belonging to hyperepidemic sequence types (STs) during 20 propagation steps (T20) on MacConkey and CHROMID ESBL plates. The following strains were analyzed: *K. pneumoniae* AE-2247421 (OXA-48/NDM-1-producing, ST101), *K. pneumoniae* MCL-2017-2 (CTX-M-15-producing, ST307), *E. coli Ec*-042 (OXA-181-producing, ST410), and *E. coli Ec*-050 (NDM-5-producing, ST167). The genome assembly at T5 and T20 was compared to that at time point zero (T0) and to two reference genomes.

Results: At T20, AE-2247421 lost the IncL $bla_{\text{OXA-48}}$-carrying plasmid when grown on CHROMID ESBL plates, while a large fragment encompassing $bla_{\text{NDM-1}}$ was lost from its IncC plasmid when grown on both plates. In contrast, no structural changes were noted for the other 3 strains. With regard to the cgSNVs, the following results were obtained at T5 and T20 (ranges considering the different agar plates and reference genomes): AE-2247421 (1-8 and 2-12 cgSNVs), MCL-2017-2 (both 1-2 cgSNVs), *Ec*-042 (both 0 cgSNVs), and *Ec*-050 (0-6 and 0-9 cgSNVs).

Conclusions: We showed that structural changes and accumulation of cgSNVs can occur in few propagation steps under laboratory conditions. These changes might also arise in the clinical context in a short time, especially under antibiotics treatment. This phenomenon should be carefully considered because it might affect the final interpretation of epidemiological genomic analyses.

KEYWORDS: SNVs, cgSNVs, genome, plasmid, stability, *K. pneumoniae*, *E. coli*

1. INTRODUCTION

Whole genome sequencing (WGS) is increasingly being used for epidemiological studies of multidrug-resistant (MDR) bacteria [1-4]. In particular, the WGS in combination with single-nucleotide variant (SNV) analysis represents an appropriate tool to support tracking and surveillance of these pathogens in large epidemiological studies and outbreak investigations [2, 3, 5, 6]. This technique provides an increased resolution power over traditional typing methods, which mostly relied on *i*) restriction analysis of large genomic DNA fragments (pulsed-field gel electrophoresis), *ii*) sequence analysis of specific housekeeping genes (multi-locus sequence typing, MLST), and *iii*) sequence analysis of core and accessory genes (core-genome MLST, cgMLST; whole-genome MLST, wgMLST).

In the more recently developed cg/wgMLST, the allele-based analyses allow for easy and standardized laboratory use, while the core-genome SNV (cgSNV) methods may provide the highest resolution [7, 8]. However, because of the complexity of cgSNV-based approaches (e.g., expertise, various variant-calling software available, read mapping- *vs*. assembly-based) the final number of cgSNVs identified are variable and inconsistent across studies [9]. It should also be noted that, so far, there is no consensus for a cgSNV threshold to distinguish whether isolates are descendants of the same strain and can be considered as clonal. This makes the interpretation of such data challenging and may bias the overall epidemiological interpretations.

Depending on the species, cutoff values ranging from 2 up to 37 SNVs have been proposed when cgSNV analyses were implemented [9]. While these criteria were obtained from retrospective studies of clinical outbreaks to determine relatedness between the isolates, it is still unknown how fast SNVs and other genomic changes arise during the propagation of the same clone. In fact, only two studies conducted *in vitro* experiments to assess the genome stability of clinical pathogens (including *Escherichia coli*, but not *Klebsiella pneumoniae*) over the course of 20 or 100 sub-culturing steps [10, 11]. Both of these studies used liquid cultures for their serial passaging experiments providing a good baseline for the number of mutations that

accumulated in the bacterial populations over time. However, they did not determine the spontaneous mutations occurring in a unique strain. To estimate this phenomenon, a strong population bottleneck (i.e., frequently reducing the genetic diversity of the population) has to be imposed in order to minimize selection that would otherwise wipe out some variants [12]. By doing so, we can observe intrinsic mutations, since selection will not be able to favor beneficial mutations and reduce deleterious ones.

In the present study, we applied such a bottleneck by picking only one colony for each propagation step to study the cgSNVs that accumulated in four MDR *K. pneumoniae* and *E. coli* clinical isolates. All of the strains chosen for the experiment belonged to hyperepidemic sequence types (STs) circulating in both human and veterinary settings [5, 13-16]. We also used complete assembled and circularized genomes – obtained with a combined short- and long-read sequencing approach – to analyze our strains, other than the two previous studies that only used a short-read sequencing approach [10, 11]. This allowed us to not only assess cgSNVs between timepoints, but also detect structural changes in the genomes, such as loss of mobile genetic elements (MGEs).

2. MATERIALS AND METHODS

2.1 Strains. Four clinical MDR isolates producing carbapenemase and/or extended-spectrum β-lactamases (ESBLs) and belonging to hyperepidemic clones were selected from our laboratory collection for the sub-culturing experiments: *K. pneumoniae* AE-2247421 [sequence type (ST) 101, OXA-48 and NDM-1 producer], *K. pneumoniae* MCL-2017-2 (ST307, CTX-M-15 producer), *E. coli Ec*-042 (ST410, OXA-181 producer), and *E. coli Ec*-050 (ST167, NDM-5 producer) (Table 1). AE-2247421 was isolated from a human rectal swab in 2013, while MCL-2017-2 was detected in 2017 in a not specified human sample [5, 16]. *Ec*-042 and *Ec*-050 were isolated from human stools in 2019 and 2018, respectively [13].

2.2 Sub-culturing experiments. The isolates were reactivated from 20% glycerol stocks stored at -80°C by plating them on MacConkey agar II plates (BBL™, Becton-Dickinson) and subsequent overnight incubation at 36±1°C. For each strain, two single and well isolated colonies were obtained from the overnight culture. From there, one colony was streaked on MacConkey agar plates and the other one was streaked on plates selective for third-generation cephalosporin-resistant strains (CHROMID ESBL, bioMérieux), followed by an overnight incubation at 36±1°C. This procedure was repeated in parallel for both plate types for a total of 20 propagation steps. Notably, in the present work, the number of generations was estimated to be ~500 for both *K. pneumoniae* and *E. coli* strains (data not shown) [17]. For simplicity, MacConkey and CHROMID ESBL agar plates are referred hereinafter to as "MAC-plates" and "ESBL-plates", respectively.

Antimicrobial susceptibility testing was performed using the broth microdilution GNX2F Sensititre panels (Thermo Fisher Scientific). MICs for antibiotics were interpreted according to the 2021 European Committee on Antimicrobial Susceptibility Testing (EUCAST) criteria (version 11.0).

2.3 Hybrid genome assembly. DNA was extracted from the first agar culture at timepoint 0 (T0), after 5 (T5), and after 20 propagation steps (T20) using the PureLink™ Microbiome DNA

Purification Kit (Thermo Fisher Scientific). WGS was performed using both short-read (Illumina, NovaSeq 6000 platform, NEBNext Ultra II DNA library prep kit; 2 x 150-bp paired-end) and long-read (Oxford Nanopore Technologies, MinION device, SQK-RBK004 library; FLO-MIN 106D R9 flow cell) sequencing for each clone.

Raw reads were trimmed using Trimmomatic (v0.36) and Porechop (v0.2.4) for short- and long-reads, respectively [18, 19]. *De novo* hybrid assemblies were performed using Unicycler v.0.4.8 and the quality of the final assemblies was assessed by mapping the trimmed Illumina-reads to the assemblies using Bowtie2 (v2.3.4.1) followed by an analysis with Qualimap (v2.2.1) [20, 21]. Circularity of the chromosome and plasmid sequences of the final assemblies was confirmed by mapping the contigs of an independent short-read assembly (SPAdes, v3.12.0) to the hybrid assemblies [22]. Annotation was performed with the NCBI annotation pipeline. The final assemblies were analyzed with the Center for Genomic Epidemiology (CGE; www.genomicepidemiology.org/) ResFinder (v4.1) and PlasmidFinder (v2.1) (50% threshold for minimum percentage identity), and cgMLSTFinder (*E. coli*, Enterobase; v1.1) [23-25]. Assemblies were also analyzed for IS elements, MGEs, and prophages using ISFinder, Islandviewer4, and PHASTER, respectively [26-28]. Unless specified, all bioinformatic analyses were done with default parameters.

2.4 Core-genome SNV (cgSNV) analysis. The resulting hybrid genomes were used for the cgSNV analyses using Parsnp (v1.2) with recombination filter (parameter: -x) with the corresponding reference genomes described in the section below [29].

For all strains, the hybrid assembly at T0 was used as reference genome for the cgSNV analysis. In addition, the analysis was performed using *i*) a randomly selected strain from the NCBI genome database belonging to the same ST and *ii*) the NCBI genome database reference genome of the same species as the reference (i.e., genomes that were "closely-related" and "distant-related" to our genomes, respectively): *K. pneumoniae* BA33875 (GenBank: GCA_002740955.2) and *K. pneumoniae* HS11286 for strain AE-2247421, *K. pneumoniae*

Kp616 (GenBank: GCA_003076555) and *K. pneumoniae* HS11286 (GenBank: GCA_000240185.2) for strain MCL-2017-2, *E. coli* 124 (GenBank: GCA_010365465) and *E. coli* K-12 (GenBank: GCA_000005845.2) for strain *Ec*-042, and *E. coli* Ec129 (GenBank: GCA_005156265) and *E. coli* K-12 for strain *Ec*-050. The SNVs identified between the isolates at T0 and the proceeding time steps were manually inspected to determine whether they were located within coding sequences (CDS) or predicted MGEs as described above.

3. RESULTS AND DISCUSSION

In the present work, we aimed to assess structural and nucleotide changes in four hyperepidemic clinical isolates over the duration of 500 generations. Notably, the isolates were chosen due to their epidemiological impact for both human and non-human settings. To do so, we conducted serial propagation experiments using two MDR *K. pneumoniae* strains belonging to ST101 and ST307 [5, 16, 30, 31], and two carbapenemase-producing *E. coli* isolates of ST410 and ST167 [13, 15, 32-35]. We emphasize that, unlike many previous studies (e.g., [36, 37]), we assessed the stability of MDR clinical isolates and not that of laboratory strains (e.g., transconjugants).

3.1 Structural changes in the K. pneumoniae isolates. The ST101 *K. pneumoniae* AE-2247421 strain underwent major structural changes during the propagation steps (Table 1). In particular, after 20 propagation steps, strain AE-2247421 was missing a plasmid and several antimicrobial resistance genes (ARGs) when grown on ESBL-plates. Indeed, the assembly for the isolate from T20 (i.e., AE-2247421-T20-ESBL) lost the ~63 kb IncL plasmid carrying the $bla_{\text{OXA-48}}$ gene (Table 1). No Illumina reads from isolate AE-2247421-T20-ESBL mapped to the IncL plasmid of the isolate at T0 (pAE-2247421-T0_1), confirming the absence of the plasmid in AE-2247421-T20-ESBL. Since the IncL plasmid carried only the $bla_{\text{OXA-48}}$ ARG (i.e., conferring resistance to carbapenems, but not to third-generation cephalosporins), it can be speculated that this MGE was not maintained in the original AE-2247421 isolate on ESBL-plates due to the lack of carbapenem selective pressure. However, we are unable to explain why the same phenomenon was not observed using the non-selective MAC-plates. Nevertheless, these observations indicate that IncL $bla_{\text{OXA-48}}$-carrying plasmids are not stably maintained in *K. pneumoniae* independently of the selective pressure.

In addition, the ARI-A region of the 189.9 kb IncC plasmid (p2247421_T0_1) from strain AE-2247421 underwent structural changes during propagation on both MAC- and ESBL-plates (Figure 1). In particular, in plasmid p2247421-T5-MAC_1, the transposition of IS*CR1* resulted in the loss of a 15,795 bp sequence containing *sul1*, *aadA2*, *dfrA12*, *mph(E)*, *msr(E)*, *armA*,

bla$_{NDM-1}$, and *aph(3')-VI* ARGs (Figure 1B). In plasmid p2247421-T5-ESBL_1, the IS*CR1*-mediated transposition caused the loss of an 26,909 bp sequence containing the same ARGs as for p2247421-T5-MAC_1, and also of an additional *sul1* and *qnrA6* (Figure 1B). The changes observed in both p2247421-T5-MAC_1 and p2247421-T5-ESBL_1 remained stable until T20 (Figure 1). We note that the continuous rearrangement of the ARI-A island in IncC plasmids through the loss and introduction of ARGs is a known phenomenon [38]. However, to the best of our knowledge, this has never been shown in serial propagation experiments of clinical isolates.

Consistent with the loss of the *bla*$_{OXA-48}$-carrying plasmid and the sequences containing *bla*$_{NDM-1}$ in AE-2247721 at T20 on ESBL-plates, the isolate was susceptible to imipenem, meropenem, and doripenem, whereas the one at T20 on MAC-plates that only lost *bla*$_{NDM-1}$ remained non-susceptible to all carbapenems tested (data not shown). Both isolates at T20 remained resistant to third- and fourth-generation cephalosporins due to additional β-lactamases, such as the CTX-M-15 ESBL (Table 1).

The loss of a plasmid and several ARGs in *K. pneumoniae* after 20 propagation steps indicate that the strain AE-2247421 is unstable. This might be a strain specific trait, as these results differ from what previous *in vitro* studies have observed in other *K. pneumoniae* strains after 35 serial passages [39]. The chromosome of the *K. pneumoniae* isolates propagated on MAC-plates (AE-2247421-T5-MAC and AE-2247421-T20-MAC) differed from all of the other chromosomes by the insertion of an IS*1R* element. The IS*1R* seemed to originate from the IncL plasmid containing *bla*$_{OXA-48}$, as this IS element was not present anywhere else in the genome assembly of the index isolate AE-2247421-T0. Remarkably, IS*1R*-mediated transposition has been described to mediate the insertion of *bla*$_{OXA-48}$ from an IncL plasmid to the chromosome in *E. coli* [40]. However, in our study, only the IS*1R* was transferred to the chromosome of *K. pneumoniae* AE-2247421.

For *K. pneumoniae* strain MCL-2017-2 of ST307, we only observed small structural changes in the plasmids pMCL-2017-T5-ESBL_1 and pMCL-2017-T20-ESBL_1. In particular, the 246,787 bp IncFIB(K)/IncFII(K) plasmid sequences showed a 302 bp deletion in the ferric citrate transport *fecR* gene that was located at position 91,234 to 91,881 bp (data not shown; GenBank: CP086462 and CP086456).

3.2 Structural changes in the E. coli isolates. As summarized in Table 1, for the two clinical *E. coli* isolates, no changes were observed in the number of plasmids and ARGs at the different timepoints. These findings correspond to what has been found by Sabol et al. conducting 20 serial passaging experiments with an *E. coli* clinical isolate [11]. However, in this previous study one cgMLST allele difference was identified for the 15th passage in 2 out of 3 replicates and for the 20th passage in 1 out of 3 replicates, whereas for our two *E. coli* isolates the cgSTs remained the same during all propagation steps (Table 1). Overall, these results demonstrate that both OXA-181-producing ST410 and NDM-5-producing ST167 *E. coli* clones are structurally highly stable (see also section about cgSNVs), which may contribute to their strong ability to spread among humans and animals (as observed in Switzerland and over the world [13, 15, 32-35]), and which can facilitate their tracking during molecular epidemiological studies and outbreak investigations.

For *E. coli* strain *Ec*-042 of ST410, a 329 bp difference between the IncIγ plasmid from the isolate at T5 on MAC-plates (pEc-042-T5-MAC_2) compared to the plasmid from the isolate at T0 (pEc-042-T0_2) was the result of a rearrangement in the shufflon region between the *pilV* gene and the *rci* shufflon recombinase gene at position 31,595 bp to 33,558 bp (GenBank: CP086540). However, comparison of the shufflon between the plasmids from all isolates revealed that similar rearrangements have occurred in the same region in strains from both agar plate types at all timepoints (data not shown; GenBank: CP086546, CP086528, CP086534, CP086522). In fact, the shufflon of the IncI plasmid is known to be rearranged constantly, resulting in a large within-population heterogeneity of bacterial cultures [41]. This, together with

the experimental approach used in the current study (i.e., picking only one colony for further propagation) that imposed a population bottleneck, probably leads to the observed diversity among plasmids.

The chromosome of strain *Ec*-042 was slightly larger (~800 bp) in *Ec*-042-T20-MAC compared to the sequences from the other timepoints and plate types. This increase was due to the insertion of an IS*1R* element at position 2,478,980 causing the partial deletion of the N-terminus in a predicted polymer-forming cytoskeletal protein. This IS*R1* element was probably originating from a duplication and transposition event from one of the six other copies located in the chromosome (GenBank: CP086526). The transposition of IS elements already present in the chromosome is a known phenomenon for *E. coli* and has been reported before [42].

In *E. coli Ec*-050 of ST167, the isolates from MAC-plates at T5 and T20, had 1,688 bp larger IncI1-Iα plasmids (pEc-050-T5-MAC_1, GenBank: CP086501; pEc-050-T20-MAC_1, GenBank: CP086481) compared to that at T0. This was due the insertion of an IS*1294* element at position 6'673 bp to 7'872 bp that derived from the IncFII plasmid (pEc-050-T5-MAC_4, GenBank: CP086504; pEc-050-T20-MAC_4, GenBank: CP086484) (data not shown).

3.3 cgSNV identification in K. pneumoniae isolates. As shown in Table 2, the number of cgSNVs identified after 20 propagation steps when using the hybrid assembly of timepoint T0 as reference sequence varied between 2 (MCL-2017-2) to 12 (AE-2247421) when considering the chromosome together with plasmids. When using closely-related strains of the same ST the number of cgSNVs decreased to 2 to 5, while with the more distant-related reference genomes they were 1 to 4. We also note that the number of cgSNVs recorded for plasmids at T20 were overall comprised between 0 and 3 (i.e., maximum ~20-25% of the total cgSNVs recorded).

Though serial passage experiments have not been performed for clinical isolates of *K. pneumoniae* in the past, the number of cgSNVs recorded for our isolates were below the previously indicated cgSNV threshold of relatedness for epidemiologic analyses (i.e., ≤18 cgSNV) [9]. In this context, we note that for local and relatively small studies considering

outbreak-related *K. pneumoniae* isolates cgSNVs comprised between 0 and 16 were recorded [4, 43, 44].

Based on our results and the overall considerations, a cautionary cutoff value of ≤20 cgSNVs might be considered for *K. pneumoniae* when using assembly-based methods considering both chromosome and MGEs. This threshold might be applied in case of local outbreak investigations occurring in relative short time, since our *in vitro* experiments mimic this specific scenario in terms of time.

3.4 cgSNV identification in E. coli isolates. Our cgSNV analysis indicated that after 20 propagation steps the two *E. coli* isolates showed 0 to 9 cgSNVs when compared to T0, while they were 0 to 6 and 0 to 4 when compared to the closely-related or the more distant-related reference genomes, respectively (Table 2).

Overall, the number of cgSNVs recorded were below the threshold suggested for epidemiologic analyses (i.e., ≤10 cgSNV) [9]. Moreover, despite using a different SNV calling method, the changes observed for *Ec*-042 (0 cgSNVs) were consistent to what has been reported by Sabol et al. (1 cgSNV) and Petronella et al. (0-3 cgSNVs after 100 serial passages), while those for *Ec*-050 were slightly higher (2 to 9 cgSNVs, of which 2-3 involving one plasmid) [10, 11]. Furthermore, small and local investigations considering outbreak-related *E. coli* isolates reported cgSNVs comprised between 0 and 13 [3, 6, 45]. Therefore, a cautionary cutoff value of ≤15 cgSNVs might be implemented for *E. coli* isolates responsible for local outbreaks occurring in a relative short time. We finally emphasize that for *K. pneumoniae* and *E. coli* isolates collected over a longer period of time and from different geographic areas, higher cutoff values of cgSNVs may probably have to be considered.

3.5 The impact of repeat regions. The numbers of SNVs in our study were slightly lower when not considering the SNVs located in regions that occurred more than once in the genome. In particular, they were 0 to 7, 0 to 4, and 0 to 4 cgSNVs for *K. pneumoniae* when the reference was T0, the closely related, and the more distant-related genome, respectively, while they were 0

to 6, 0 to 6, and 0 to 4 cgSNVs for *E. coli* with the respective reference sequences (Table 2). Since SNVs detected in such regions are dubious, especially when only short read sequencing and draft assemblies are used, repetitive regions are often excluded from cgSNV analyses [46, 47]. The reduction in cgSNVs detected in our study when excluding repeat regions further emphasizes the importance of a harmonized analysis pipeline when comparing cgSNV thresholds from different studies [48]. It furthermore demonstrates the advantage of a combined long- and short-read sequencing approach and the resulting completely assembled genomes.

3.6 The impact of the agar type on cgSNVs. For *E. coli Ec*-050 and *K. pneumoniae* AE-2247421, the number of cgSNVs at T20 was higher on ESBL-plates compared to MAC-plates (Table 2). Moreover, several of these cgSNVs only present in isolates from ESBL-plates were located in CDS (Table 3).

It has been shown that exposure to antibiotics not only increases the selection pressure on bacterial populations, but also enhances the genome-wide mutation rate of bacterial cells and movement of MGEs [49]. The antibiotics-induced mutagenesis has not only been observed for antibiotics that directly act on DNA replication (e.g., fluoroquinolones), but also for β-lactams which interfere with peptidoglycan synthesis [50-52]. Therefore, it is possible that the cephalosporin in the ESBL-plates lead to an increased mutation-rate in some of our isolates. However, as our study is limited to one experiment per isolate, this needs to be confirmed with more replicates in future studies. In addition, the design of our analysis likely minimized selection pressure through the implementation of frequent population bottlenecks and thus, could have allowed more of these mutations to accumulate step-by-step. By applying a frequent and strong population bottleneck, genetic diversity is reduced because only one clone is allowed to proliferate. This reduces selection and enhances the impact of genetic drift which promotes the accumulation of mutations [12, 53].

4. CONCLUSIONS

Although we proposed thresholds of relatedness for cgSNV analyses, the present study emphasizes the difficulty for defining them and confirms that such cutoff values have to be regarded as recommendations rather than strict values [9]. Nevertheless, the results from our analyses can serve as a baseline for epidemiological cgSNV cutoffs in future studies.

Our data underscore the importance of a standardized protocol when conducting SNV analyses to get consistent results that are comparable across laboratories. In particular, we demonstrated that the number of cgSNVs identified is affected by the reference genome used and, to a lesser extent, by considering the plasmid(s) for the overall analysis. We also know that the number of cgSNVs can be heavily influenced by the sequencing procedure, the quality of the genome assemblies (e.g., assembly errors), SNV-calling software (e.g., the parameters used in read-based methods), and recombination removal tools (e.g., ClonalframeML or Gubbins) with up to 150 cgSNVs differences observed between different methods [9, 48, 54]. Therefore, cgSNV values from different studies can only be compared with each other to a limited extent, unless common and standardized approaches are set and implemented.

We have also shown that the hybrid WGS approach allows not only for the detection of high quality cgSNVs, but importantly the detection of structural changes in the genome over time. In fact, large structural changes such as the loss of ARGs or an entire plasmid may arise in only 20, or even less, propagation steps. Furthermore, we showed that SNVs can be accumulated in a short time-span under laboratory conditions. Overall, this has direct implications on the handling of strains analyzed in the laboratory, as already a few propagation steps can lead to major changes in the genome [11]. Although we would expect to have less bacterial generations *in vivo* during an equal period of time, our results imply that mutations and structural changes might also arise in the clinical context in a short time, especially under antibiotics treatment. This possible phenomenon should be carefully considered because it might affect the final epidemiological interpretation of genomic analyses.

ACKNOWLEDGEMENTS

We thank Mrs. Angela Vallone for technical support and Dr. Carola Maffioli (MCL Medizinische Laboratorien, Niederwangen, Switzerland) for providing the *K. pneumoniae* strain MCL-2017-2.

FUNDING

This work was supported by the Swiss National Science Foundation (SNF) grant No. 192514 (to AE). Edgar I. Campos-Madueno is a PhD student (2021-2024) supported by SNF.

ETHICAL APPROVAL

Not required.

COMPETING INTERESTS

None declared.

Table 1. Assembly characteristics for all of the sequenced clones

Passage/plate	Coverage	Genome size	Plasmids	cgST [a]	PlasmidFinder 2.0	ResFinder 4.1 [b]	NCBI accessions
***K. pneumoniae* AE-2247421 (ST101)**							
T0	223	5'960'745	7	NA	IncC, IncFIB(K), IncFIA (HI1), IncR, IncL, Col(pHAD28), Col440II	***bla*$_{NDM-1}$, *bla*$_{OXA-48}$, *bla*$_{CTX-M-15}$** *(2x)*, ***bla*$_{CMY-4}$**, *bla*$_{OXA-10}$, *bla*$_{SHV-1}$, *bla*$_{SCO-1}$, *bla*$_{TEM-1B}$, *bla*$_{TEM-1A}$, *armA, aph(6)-Id, aph(3'')-Ib, aph(3')-VI, aac(3)-IIa, aadA2, aadA1, tet(A), oqxB, oqxA, qnrA6, sul1 (4x), sul2, dfrA12, dfrA14 (2x), cmlA1, floR, fosA, msr(E), mph(E), arr-2, (2x)*	CP086447-CP086454
T5-MAC	452	5'938'471	7	NA	IncC, IncFIB(K), IncFIA (HI1), IncR, IncL, Col(pHAD28), Col440II	***bla*$_{OXA-48}$, *bla*$_{CTX-M-15}$** *(2x)*, ***bla*$_{CMY-4}$**, *bla*$_{OXA-10}$, *bla*$_{SHV-1}$, *bla*$_{SCO-1}$, *bla*$_{TEM-1B}$, *bla*$_{TEM-1A}$, *aph(6)-Id, aph(3'')-Ib, aac(3)-IIa, aadA1, tet(A), oqxB, oqxA, qnrA6, sul1 (2x), sul2, dfrA14 (2x), cmlA1, floR, fosA, arr-2*	CP086439-CP086446
T20-MAC	238	5'939'313	7	NA	IncC, IncFIB(K), IncFIA (HI1), IncR, IncL, Col(pHAD28), Col440II	***bla*$_{OXA-48}$, *bla*$_{CTX-M-15}$** *(2x)*, ***bla*$_{CMY-4}$**, *bla*$_{OXA-10}$, *bla*$_{SHV-1}$, *bla*$_{SCO-1}$, *bla*$_{TEM-1B}$, *bla*$_{TEM-1A}$, *aph(6)-Id, aph(3'')-Ib, aac(3)-IIa, aadA1, tet(A), oqxB, oqxA, qnrA6, sul1 (2x), sul2, dfrA14 (2x), cmlA1, floR, fosA, arr-2*	CP086423-CP086430
T5-ESBL	269	5'927'416	7	NA	IncC, IncFIB(K), IncFIA (HI1), IncR, IncL, Col(pHAD28), Col440II	***bla*$_{OXA-48}$, *bla*$_{CTX-M-15}$** *(2x)*, ***bla*$_{CMY-4}$**, *bla*$_{OXA-10}$, *bla*$_{SHV-1}$, *bla*$_{SCO-1}$, *bla*$_{TEM-1B}$, *bla*$_{TEM-1A}$, *aph(6)-Id, aph(3'')-Ib, aac(3)-IIa, aadA1, tet(A), oqxB, oqxA, sul1, sul2, dfrA14 (2x), cmlA1, floR, fosA, arr-2*	CP086431-CP086438
T20-ESBL	299	5'862'987	6	NA	IncC, IncFIB(K), IncFIA (HI1), IncR, Col(pHAD28), Col440II	*bla*$_{CTX-M-15}$ *(2x)*, ***bla*$_{CMY-4}$**, *bla*$_{OXA-10}$, *bla*$_{SHV-1}$, *bla*$_{SCO-1}$, *bla*$_{TEM-1B}$, *bla*$_{TEM-1A}$, *aph(6)-Id, aph(3'')-Ib, aac(3)-IIa, aadA1, tet(A), oqxB, oqxA, sul1, sul2, dfrA14 (2x), cmlA1, floR, fosA, arr-2*	CP086416-CP086422
***K. pneumoniae* MCL-2017-2 (ST307)**							
T0	278	5'703'400	2	NA	IncFIB(K), IncFII(K), Col(pHAD28)	***bla*$_{CTX-M-15}$**, *bla*$_{SHV-28}$, *bla*$_{OXA-1}$, *bla*$_{TEM-1B}$, *aph(6)-Id, aph(3'')-Ib, aac(6')-Ib-cr, aac(3)-IIa, tet(A), oqxB, oqxA, qnrB1, sul2, dfrA14, catB3, fosA*	CP086467-CP086469
T5-MAC	462	5'703'394	2	NA	IncFIB(K), IncFII(K), Col(pHAD28)	***bla*$_{CTX-M-15}$**, *bla*$_{SHV-28}$, *bla*$_{OXA-1}$, *bla*$_{TEM-1B}$, *aph(6)-Id, aph(3'')-Ib, aac(6')-Ib-cr, aac(3)-IIa, tet(A), oqxB, oqxA, qnrB1, sul2, dfrA14, catB3, fosA*	CP086464-CP086466
T20-MAC	431	5'703'389	2	NA	IncFIB(K), IncFII(K), Col(pHAD28)	***bla*$_{CTX-M-15}$**, *bla*$_{SHV-28}$, *bla*$_{OXA-1}$, *bla*$_{TEM-1B}$, *aph(6)-Id, aph(3'')-Ib, aac(6')-Ib-cr, aac(3)-IIa, tet(A), oqxB, oqxA, qnrB1, sul2, dfrA14, catB3, fosA*	CP086458-CP086460
T5-ESBL	247	5'703'095	2	NA	IncFIB(K), IncFII(K), Col(pHAD28)	***bla*$_{CTX-M-15}$**, *bla*$_{SHV-28}$, *bla*$_{OXA-1}$, *bla*$_{TEM-1B}$, *aph(6)-Id, aph(3'')-Ib, aac(6')-Ib-cr, aac(3)-IIa, tet(A), oqxB, oqxA, qnrB1, sul2, dfrA14, catB3, fosA*	CP086461-CP086463
T20-ESBL	420	5'703'097	2	NA	IncFIB(K), IncFII(K), Col(pHAD28)	***bla*$_{CTX-M-15}$**, *bla*$_{SHV-28}$, *bla*$_{OXA-1}$, *bla*$_{TEM-1B}$, *aph(6)-Id, aph(3'')-Ib, aac(6')-Ib-cr, aac(3)-IIa, tet(A), oqxB, oqxA, qnrB1, sul2, dfrA14, catB3, fosA*	CP086455-CP086457
***E. coli Ec*-042 (ST410)**							
T0	294	4'808'195	5	137220	IncX3, ColKP3, IncIγ, IncX4, Col440I	***bla*$_{OXA-181}$**, ***bla*$_{CMY-42}$**, *mdf(A), qnrS1*	CP086544-CP086549
T5-MAC	430	4'808'531	5	137220	IncX3, ColKP3, IncIγ, IncX4, Col440I	***bla*$_{OXA-181}$**, ***bla*$_{CMY-42}$**, *mdf(A), qnrS1*	CP086538-CP086543
T20-MAC	290	4'809'001	5	137220	IncX3, ColKP3, IncIγ, IncX4, Col440I	***bla*$_{OXA-181}$**, ***bla*$_{CMY-42}$**, *mdf(A), qnrS1*	CP086526-CP086531
T5-ESBL	325	4'808'193	5	137220	IncX3, ColKP3, IncIγ, IncX4, Col440I	***bla*$_{OXA-181}$**, ***bla*$_{CMY-42}$**, *mdf(A), qnrS1*	CP086532-CP086537
T20-ESBL	273	4'808'210	5	137220	IncX3, ColKP3, IncIγ, IncX4, Col440I	***bla*$_{OXA-181}$**, ***bla*$_{CMY-42}$**, *mdf(A), qnrS1*	CP086520-CP086525
***E. coli Ec*-050 (ST167)**							
T0	271	5'332'076	9	137002	IncI1-Iα, IncFIA, IncFIB (AP001918), IncFII, IncI2δ, IncX4, ColRNAI, Col(pHAD28), Col (MG828)	***bla*$_{NDM-5}$**, ***bla*$_{CMY-2}$**, ***bla*$_{TEM-30}$**, *aac(3)-IIa, aadA1, aadA2, tet(A), sul1 (2x), sul2, dfrA12, dfrA, mdfA, floR, mph(A), erm(B)*	CP086510-CP086519
T5-MAC	275	5'333'616	9	137002	IncI1-Iα, IncFIA, IncFIB (AP001918), IncFII, IncI2δ, IncX4, ColRNAI, Col(pHAD28), Col (MG828)	***bla*$_{NDM-5}$**, ***bla*$_{CMY-2}$**, ***bla*$_{TEM-30}$**, *aac(3)-IIa, aadA1, aadA2, tet(A), sul1 (2x), sul2, dfrA12, dfrA, mdfA, floR, mph(A), erm(B)*	CP086500-CP086509
T20-MAC	308	5'333'341	9	137002	IncI1-Iα, IncFIA, IncFIB (AP001918), IncFII, IncI2δ, IncX4, ColRNAI, Col(pHAD28), Col (MG828)	***bla*$_{NDM-5}$**, ***bla*$_{CMY-2}$**, ***bla*$_{TEM-30}$**, *aac(3)-IIa, aadA1, aadA2, tet(A), sul1 (2x), sul2, dfrA12, dfrA, mdfA, floR, mph(A), erm(B)*	CP086480-CP086489
T5-ESBL	225	5'330'360	9	137002	IncI1-Iα, IncFIA, IncFIB (AP001918), IncFII, IncI2δ, IncX4, ColRNAI, Col(pHAD28), Col (MG828)	***bla*$_{NDM-5}$**, ***bla*$_{CMY-2}$**, ***bla*$_{TEM-30}$**, *aac(3)-IIa, aadA1, aadA2, tet(A), sul1 (2x), sul2, dfrA12, dfrA, mdfA, floR, mph(A), erm(B)*	CP086490-CP086499
T20-ESBL	390	5'331'662	9	137002	IncI1-Iα, IncFIA; IncFIB (AP001918), IncFII, IncI2δ, IncX4, ColRNAI, Col(pHAD28), Col(MG828)	***bla*$_{NDM-5}$**, ***bla*$_{CMY-2}$**, ***bla*$_{TEM-30}$**, *aac(3)-IIa, aadA1, aadA2, tet(A), sul1 (2x), sul2, dfrA12, dfrA, mdfA, floR, mph(A), erm(B)*	CP086470-CP086479

[a] cgST, core-genome sequence type (cgST); NA, not available
[b] In some assemblies certain ARGs occurred more than once in the genome (e.g., 2x, 4x); The main ARGs are indicated in bold

Table 2. SNVs at the different timepoints compared to T0

Table 3. Single nucleotide variants for all of the three isolates where SNVs were detected among the different timepoints and plate types

Strain	Reference genome [b]	Cov. [%] [c]	Location [d]	SNVs compared to timepoint T0 [a]													
				MacConkey agar plates (MAC-plates)								CHROMID ESBL agar plates (ESBL-plates)					
				T5	CDS [%] [e]	M [%] [f]	R [%] [g]	T20	CDS [%] [e]	M [%] [f]	R [%] [g]	T5	CDS [%] [e]	M [%] [f]	R [%] [g]	T20	CDS [%] [e]	M [%] [f]	R [%] [g]
***K. pneumoniae* AE-2247421**	AE-2247421-T0 (at timepoint 0)	98.8	Chromosome	5	80	0	80	4	75	0	75	3	66.7	0	33.3	9	66.7	0	55.5
			p2247421-T0_2	0	-	-	-	1	0	0	0	0	-	-	-	0	-	-	-
			p2247421-T0_3	0	-	-	-	0	-	-	-	5	100	40	40	3	0	0	0
	BA33875 (closely-related)	92.7	Chromosome	3	66.7	0	66.7	4	75	0	75	3	66.7	0	33.3	5	40	0	20
			pBA33875_IncFIB	0	-	-	-	1	0	0	0	0	-	-	-	0	-	-	-
	HS11286 (distantly-related)	82.8	Chromosome	1	0	0	0	2	50	0	50	2	50	0	0	4	25	0	0
***K. pneumoniae* MCL-2017-2**	MCL-2017-T0 (at timepoint 0)	100	Chromosome	2	100	50	100	2	100	50	100	2	100	50	100	2	100	50.0	100
	616 (closely-related)	90.3	Chromosome	2	100	50	100	2	100	50	100	2	100	50	100	2	100	50.0	100
	HS11286 (distantly-related)	85.7	Chromosome	1	100	100	100	1	100	100	100	1	100	100	100	1	100	100	100
***E. coli* *Ec*-042**	*Ec*-042-T0 (at timepoint 0)	99.9	Chromosome	0	-	-	-	0	-	-	-	0	-	-	-	0	-	-	-
	124 (closely-related)	95.2	Chromosome	0	-	-	-	0	-	-	-	0	-	-	-	0	-	-	-
	K-12 (distantly-related)	86.9	Chromosome	0	-	-	-	0	-	-	-	0	-	-	-	0	-	-	-
***E. coli* *Ec*-050**	*Ec*-050-T0 (at timepoint 0)	99.8	Chromosome	2	100	0	100	0	-	-	-	2	100	0	50	6	83.3	33.3	0
			pEc-050-T0_3	0	-	-	-	0	-	-	-	1	100	100	100	0	-	-	-
			pEc-050-T0_5	3	100	100	100	2	100	100	100	3	100	100	100	3	100	100	100
	EC-129 (closely-related)	87.4	Chromosome	2	100	100	100	0	-	-	-	2	100	0	50	6	83.3	33.3	0
	K-12 (distantly-related)	76.1	Chromosome	0	-	-	-	0	-	-	-	1	100	0	0	4	75	0	0

[a] SNV analysis was performed using Parsnp (v1.2). SNVs were analyzed between timepoints T0 (starting culture), T5 (culture after 5 times of propagation), and T20 (culture after 20 times of propagation).

[b] Genome sequences used as reference for the Parsnp (v1.2) analysis were *Ec*-042-T0 (GenBank: CP086544-CP086549), *Ec*-050-T0 (GenBank: CP086510-CP086519), MCL-2017-T0 (GenBank: CP086467-CP086469), AE-2247421-T0 (GenBank: CP086447-CP086454), *E. coli* 124 (GenBank: GCA_010365465.1), *E. coli* K-12 (GenBank: GCA_000005845.2), *E. coli* EC-129 (GenBank: GCF_005156265.1), *K. pneumoniae* 616 (GenBank: GCA_003076555.1), *K. pneumoniae* HS11286 (GenBank: GCA_000240185.2), and *K. pneumoniae* BA33875 (GenBank: GCA_002740955.2).

[c] Total coverage of the core genome alignment used for the SNV analysis among all sequences [%]

[d] Only sequences where SNVs were detected are listed. Plasmid pBA33875_IncFIB is deposited (GenBank: CP035180).

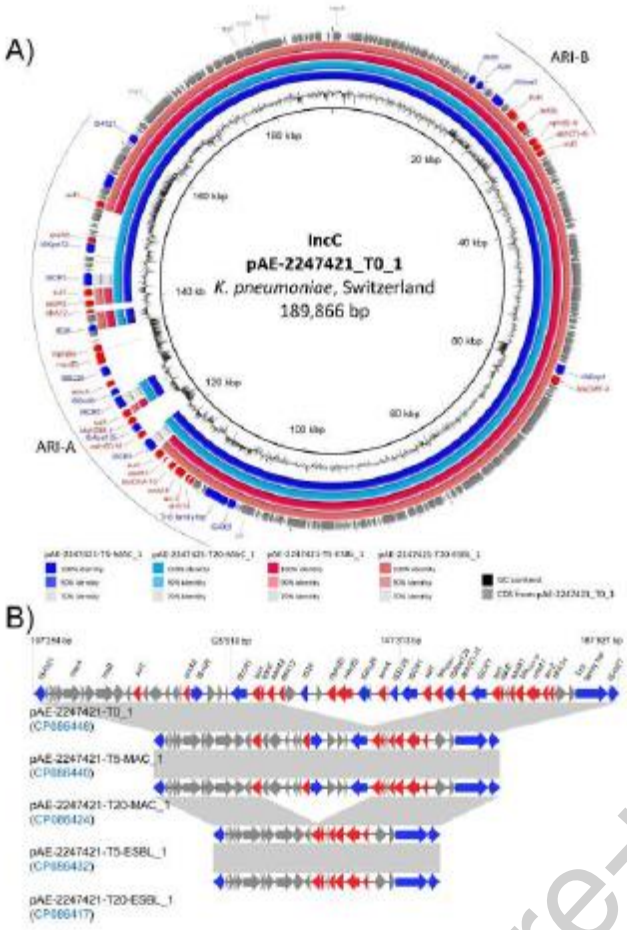
[e] CDS: percentage of SNVs located within a coding sequence

[f] M: percentage of SNVs located within a predicted mobile element. IS elements were identified with ISfinder, prophage regions were identified with PHASTER (upgrade 6), and genomic islands were predicted with IslandViewer4

[g] R: percentage of SNVs located within a CDS that occurs more than once in the genome (i.e., 23S rRNA, 16S rRNA, IS elements)

SNV	Location [a]	Position [b]	REF/ALT [c]	Annotation [d]	Predicted MGE [e]	Present/Absence [f]			
						T5 MAC	T20 MAC	T5 ESBL	T20 ESBL
***K. pneumoniae* AE-2247421 (ST101)**									
1	2247421-T0	561'581	C/A	*argG* (Argininosuccinate synthase)	no	0	0	1	1
2	2247421-T0	1'214'581	C/T	tRNA-Glu	no	1	0	0	0
3	2247421-T0	1'602'764	T/A	non-coding region	no	0	0	0	1
4	2247421-T0	1'602'766	G/A	non-coding region	no	0	0	0	1
5	2247421-T0	4'052'567	C/G	non-coding region	no	1	1	1	1
6	2247421-T0	4'426'352	G/A	23S rRNA	no	0	1	0	0
7	2247421-T0	4'783'742	C/T	23S rRNA	no	0	0	0	1
8	2247421-T0	4'783'743	A/G	23S rRNA	no	0	0	0	1
9	2247421-T0	4'783'750	T/C	23S rRNA	no	0	0	0	1
10	2247421-T0	4'783'751	G/A	23S rRNA	no	0	0	0	1
11	2247421-T0	5'170'937	A/G	tRNA-Glu	no	1	1	1	1
12	2247421-T0	5'215'153	A/G	23S rRNA	no	0	1	0	0
13	2247421-T0	5'411'346	A/G	23S rRNA	no	1	0	0	0
14	2247421-T0	5'411'738	T/C	23S rRNA	no	1	0	0	0
15	p2247421-T0_2	54'457	A/G	non-coding region	no	0	1	0	0
16	p2247421-T0_3	17'752	T/C	*merA* (Mercuric reductase)	no	0	0	1	0
17	p2247421-T0_3	31'419	G/A	*ltrA* (Group II intron-encoded protein LtrA)	no	0	0	1	0
18	p2247421-T0_3	31'953	C/T	*ltrA* (Group II intron-encoded protein LtrA)	no	0	0	1	0
19	p2247421-T0_3	46'390	A/G	non-coding region	no	0	0	0	1
20	p2247421-T0_3	46'407	T/C	non-coding region	no	0	0	0	1
21	p2247421-T0_3	46'420	T/C	non-coding region	no	0	0	0	1
22	p2247421-T0_3	62'100	G/C	class 1 integron integrase IntI1	Integrase	0	0	1	0
23	p2247421-T0_3	62'105	G/A	class 1 integron integrase IntI1	Integrase	0	0	1	0
***K. pneumoniae* MCL-2017-2 (ST307)**									
1	MCL-2017-2-T0	464'196	C/T	23S rRNA	No	0	1	1	1
2	MCL-2017-2-T0	1'304'980	T/C	16S rRNA	Incomplete prophage region	0	1	1	1
***E. coli Ec*-050 (ST167)**									
1	Ec-050-T0	107'408	A/C	*lldP* (L-lactate permease)	no	0	0	1	1
2	Ec-050-T0	382'791	A/G	*frlB* (Fructosamine deglycase FrlB)	no	0	0	0	1
3	Ec-050-T0	1'239'111	G/A	23S rRNA	no	1	0	0	0
4	Ec-050-T0	1'884'998	T/C	*rfbB* (dTDP-glucose 4,6-dehydratase 2)	incomplete prophage	0	0	0	1
5	Ec-050-T0	1'884'999	C/T	*rfbB* (dTDP-glucose 4,6-dehydratase 2)	incomplete prophage	0	0	0	1
6	Ec-050-T0	3'879'295	T/G	non-coding region	no	0	0	0	1
7	Ec-050-T0	4'665'919	G/T	*ompL* (Porin OmpL)	no	0	0	0	1
8	Ec-050-T0	4'692'403	T/C	23S rRNA	no	1	0	0	0
9	Ec-050-T0	4'784'422	T/C	23S rRNA	no	0	0	1	0
10	pEc-050-T0_1	115'570	G/A	IS*26*	IS element	1	0	1	1
11	pEc-050-T0_1	115'724	G/A	IS*26*	IS element	1	1	1	1
12	pEc-050-T0_1	115'725	G/A	IS*26*	IS element	1	1	1	1
13	pEc-050-T0_4	4'834	G/T	IS*26*	IS element	0	0	1	0

[a] Location of the cgSNV in the genome assembly at T0; [b] Nucleotide position of the cgSNV in the genome for T0; [c] Nucleotide of the reference (REF) and the variant (ALT); [d] Annotation of the CDS where the cgSNV was located according to the NCBI annotation pipeline; [e] Information whether the cgSNV was located within a predicted mobile genetic element (MGE) or not (no); [f] Indication whether the cgSNV was present (1) or absent (0) for a specific timepoint (T5 or T20) and plate type (MAC- or ESBL-plate)



LEGEND TO FIGURE 1

BLASTn comparison of the IncC plasmid in AE-2247421 at different time points. A) The IncC plasmids at time points T0, T5, and T20 on MAC-plates and ESBL-plates were compared with BRIG (BLAST Ring Image Generator) v.0.95. AE-2247421_T0 (GenBank: CP086448) was used as reference sequence. The colored rings represent similarities to the reference sequence. CDS from AE-224721-T0 are depicted as arrows in the outermost circle. Mobile genetic elements (MGEs) are depicted in blue, antimicrobial resistance genes in red, and all other CDS in grey. B) Genetic environment of the ARI-A region for the different time points. The sequences were compared with EasyFig v.2.2.5. The grey area between the sequences indicates a sequence identity ≥99%. The GenBank accession number for each sequence is indicated in blue in brackets.

REFERENCES

[1] Jamal AJ, Mataseje LF, Williams V, Leis JA, Tijet N, Zittermann S, et al. Genomic Epidemiology of Carbapenemase-Producing Enterobacterales at a Hospital System in Toronto, Ontario, Canada, 2007 to 2018. Antimicrob Agents Chemother. 2021;65:e0036021.

[2] Moser AI, Kuenzli E, Campos-Madueno EI, Budel T, Rattanavong S, Vongsouvath M, et al. Antimicrobial-Resistant *Escherichia coli* Strains and Their Plasmids in People, Poultry, and Chicken Meat in Laos. Front Microbiol. 2021;12:708182.

[3] Rumore J, Tschetter L, Kearney A, Kandar R, McCormick R, Walker M, et al. Evaluation of whole-genome sequencing for outbreak detection of Verotoxigenic *Escherichia coli* O157:H7 from the Canadian perspective. BMC Genomics. 2018;19:870.

[4] Ruppe E, Olearo F, Pires D, Baud D, Renzi G, Cherkaoui A, et al. Clonal or not clonal? Investigating hospital outbreaks of KPC-producing *Klebsiella pneumoniae* with whole-genome sequencing. Clin Microbiol Infect. 2017;23:470-5.

[5] Campos-Madueno EI, Moser AI, Jost G, Maffioli C, Bodmer T, Perreten V, et al. Carbapenemase-producing *Klebsiella pneumoniae* strains in Switzerland: Human and non-human settings may share high-risk clones. J Global Antimicrob Res. 2022;in press.

[6] Moser AI, Kuenzli E, Budel T, Campos-Madueno EI, Bernasconi OJ, DeCrom-Beer S, et al. Travellers returning from the island of Zanzibar colonized with MDR *Escherichia coli* strains: assessing the impact of local people and other sources. J Antimicrob Chemother. 2021;76:330-7.

[7] Lee K, Izumiya H, Iyoda S, Ohnishi M. Effective Surveillance Using Multilocus Variable-Number Tandem-Repeat Analysis and Whole-Genome Sequencing for Enterohemorrhagic *Escherichia coli* O157. Appl Environ Microbiol. 2019;85.

[8] Miro E, Rossen JWA, Chlebowicz MA, Harmsen D, Brisse S, Passet V, et al. Core/Whole Genome Multilocus Sequence Typing and Core Genome SNP-Based Typing of OXA-48-Producing *Klebsiella pneumoniae* Clinical Isolates From Spain. Front Microbiol. 2019;10:2961.

[9] Schurch AC, Arredondo-Alonso S, Willems RJL, Goering RV. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. Clin Microbiol Infect. 2018;24:350-4.

[10] Petronella N, Kundra P, Auclair O, Hebert K, Rao M, Kingsley K, et al. Changes detected in the genome sequences of *Escherichia coli*, *Listeria monocytogenes*, *Vibrio parahaemolyticus*, and *Salmonella enterica* after serial subculturing. Can J Microbiol. 2019;65:842-50.

[11] Sabol A, Joung YJ, VanTubbergen C, Ale J, Ribot EM, Trees E. Assessment of Genetic Stability During Serial In Vitro Passage and In Vivo Carriage. Foodborne Pathog Dis. 2021;18:894-901.

[12] Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nat Rev Genet. 2013;14:827-39.

[13] Endimiani A, Brilhante M, Bernasconi OJ, Perreten V, Schmidt JS, Dazio V, et al. Employees of Swiss veterinary clinics colonized with epidemic clones of carbapenemase-producing *Escherichia coli*. J Antimicrob Chemother. 2020;75:766-8.

[14] Nadimpalli ML, de Lauzanne A, Phe T, Borand L, Jacobs J, Fabre L, et al. *Escherichia coli* ST410 among humans and the environment in Southeast Asia. Int J Antimicrob Agents. 2019;54:228-32.

[15] Nigg A, Brilhante M, Dazio V, Clement M, Collaud A, Gobeli Brawand S, et al. Shedding of OXA-181 carbapenemase-producing *Escherichia coli* from companion animals after hospitalisation in Switzerland: an outbreak in 2018. Euro Surveill. 2019;24.

[16] Seiffert SN, Marschall J, Perreten V, Carattoli A, Furrer H, Endimiani A. Emergence of *Klebsiella pneumoniae* co-producing NDM-1, OXA-48, CTX-M-15, CMY-16, QnrA and ArmA in Switzerland. Int J Antimicrob Agents. 2014;44:260-2.

[17] Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. Proc Natl Acad Sci U S A. 2012;109:E2774-83.

[18] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-20.

[19] Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom. 2017;3:e000132.

[20] Garcia-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Gotz S, Tarazona S, et al. Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics. 2012;28:2678-9.

[21] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357-9.

[22] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455-77.

[23] Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020;75:3491-500.

[24] Carattoli A, Hasman H. PlasmidFinder and In Silico pMLST: Identification and Typing of Plasmid Replicons in Whole-Genome Sequencing (WGS). Methods Mol Biol. 2020;2075:285-94.

[25] Zhou Z, Alikhan NF, Mohamed K, Fan Y, Agama Study G, Achtman M. The EnteroBase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. Genome Res. 2020;30:138-52.

[26] Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44:W16-21.

[27] Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing G, Lau BY, Hoad G, et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 2017;45:W30-W5.

[28] Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 2006;34:D32-6.

[29] Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014;15:524.

[30] Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. Nat Rev Microbiol. 2020;18:344-59.

[31] Peirano G, Chen L, Kreiswirth BN, Pitout JDD. Emerging Antimicrobial-Resistant High-Risk *Klebsiella pneumoniae* Clones ST307 and ST147. Antimicrob Agents Chemother. 2020;64.

[32] Chakraborty T, Sadek M, Yao Y, Imirzalioglu C, Stephan R, Poirel L, et al. Cross-Border Emergence of *Escherichia coli* Producing the Carbapenemase NDM-5 in Switzerland and Germany. J Clin Microbiol. 2021;59.

[33] Alba P, Taddei R, Cordaro G, Fontana MC, Toschi E, Gaibani P, et al. Carbapenemase IncF-borne $bla_{\text{NDM-5}}$ gene in the *E. coli* ST167 high-risk clone from canine clinical infection, Italy. Vet Microbiol. 2021;256:109045.

[34] Bibbolino G, Di Lella FM, Oliva A, Lichtner M, Del Borgo C, Raponi G, et al. Molecular epidemiology of NDM-5-producing *Escherichia coli* high-risk clones identified in two Italian hospitals in 2017-2019. Diagn Microbiol Infect Dis. 2021;100:115399.

[35] Patino-Navarrete R, Rosinski-Chupin I, Cabanel N, Zongo PD, Hery M, Oueslati S, et al. Specificities and Commonalities of Carbapenemase-Producing *Escherichia coli* Isolated in France from 2012 to 2015. mSystems. 2022;7:e0116921.

[36] Liu Z, Wang Z, Lu X, Peng K, Chen S, He S, et al. Structural Diversity, Fitness Cost, and Stability of a $bla_{\text{NDM-1}}$-Bearing Cointegrate Plasmid in *Klebsiella pneumoniae* and *Escherichia coli*. Microorganisms. 2021;9.

[37] Li X, He J, Jiang Y, Peng M, Yu Y, Fu Y. Genetic Characterization and Passage Instability of a Hybrid Plasmid Co-Harboring $bla_{\text{IMP-4}}$ and $bla_{\text{NDM-1}}$ Reveal the Contribution of Insertion Sequences During Plasmid Formation and Evolution. Microbiol Spectr. 2021;9:e0157721.

[38] Ambrose SJ, Harmer CJ, Hall RM. Evolution and typing of IncC plasmids contributing to antibiotic resistance in Gram-negative bacteria. Plasmid. 2018;99:40-55.

[39] Porse A, Schonning K, Munck C, Sommer MO. Survival and Evolution of a Large Multidrug Resistance Plasmid in New Clinical Bacterial Hosts. Mol Biol Evol. 2016;33:2860-73.

[40] Beyrouthy R, Robin F, Delmas J, Gibold L, Dalmasso G, Dabboussi F, et al. IS1R-mediated plasticity of IncL/M plasmids leads to the insertion of bla OXA-48 into the *Escherichia coli* Chromosome. Antimicrob Agents Chemother. 2014;58:3785-90.

[41] Brouwer MSM, Jurburg SD, Harders F, Kant A, Mevius DJ, Roberts AP, et al. The shufflon of IncI1 plasmids is rearranged constantly during different growth conditions. Plasmid. 2019;102:51-5.

[42] Nicoloff H, Perreten V, Levy SB. Increased genome instability in *Escherichia coli* lon mutants: relation to emergence of multiple-antibiotic-resistant (Mar) mutants caused by insertion sequence elements and large tandem genomic amplifications. Antimicrob Agents Chemother. 2007;51:1293-303.

[43] Ferrari C, Corbella M, Gaiarsa S, Comandatore F, Scaltriti E, Bandi C, et al. Multiple *Klebsiella pneumoniae* KPC Clones Contribute to an Extended Hospital Outbreak. Front Microbiol. 2019;10:2767.

[44] Brilhante M, Gobeli Brawand S, Endimiani A, Rohrbach H, Kittl S, Willi B, et al. Two high-risk clones of carbapenemase-producing *Klebsiella pneumoniae* that cause infections in pets and are present in the environment of a veterinary referral hospital. J Antimicrob Chemother. 2021;76:1140-9.

[45] Roer L, Hansen F, Thomsen MCF, Knudsen JD, Hansen DS, Wang M, et al. WGS-based surveillance of third-generation cephalosporin-resistant *Escherichia coli* from bloodstream infections in Denmark. J Antimicrob Chemother. 2017;72:1922-9.

[46] Hawkey J, Paranagama K, Baker KS, Bengtsson RJ, Weill FX, Thomson NR, et al. Global population structure and genotyping framework for genomic surveillance of the major dysentery pathogen, *Shigella sonnei*. Nat Commun. 2021;12:2684.

[47] Petkau A, Mabon P, Sieffert C, Knox NC, Cabral J, Iskander M, et al. SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. Microb Genom. 2017;3:e000116.

[48] Jamin C, De Koster S, van Koeveringe S, De Coninck D, Mensaert K, De Bruyne K, et al. Harmonization of whole-genome sequencing for outbreak surveillance of Enterobacteriaceae and Enterococci. Microb Genom. 2021;7.

[49] Long H, Miller SF, Strauss C, Zhao C, Cheng L, Ye Z, et al. Antibiotic treatment enhances the genome-wide mutation rate of target cells. Proc Natl Acad Sci U S A. 2016;113:E2498-505.

[50] Gutierrez A, Laureti L, Crussard S, Abida H, Rodriguez-Rojas A, Blazquez J, et al. beta-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. Nat Commun. 2013;4:1610.

[51] Jee J, Rasouly A, Shamovsky I, Akivis Y, Steinman SR, Mishra B, et al. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. Nature. 2016;534:693-6.

[52] Song LY, Goff M, Davidian C, Mao Z, London M, Lam K, et al. Mutational Consequences of Ciprofloxacin in *Escherichia coli*. Antimicrob Agents Chemother. 2016;60:6165-72.

[53] Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet. 2016;17:704-14.

[54] Maljkovic Berry I, Melendrez MC, Bishop-Lilly KA, Rutvisuttinunt W, Pollett S, Talundzic E, et al. Next Generation Sequencing and Bioinformatics Methodologies for Infectious Disease Research and Public Health: Approaches, Applications, and Considerations for Development of Laboratory Capacity. J Infect Dis. 2020;221:S292-S307.