HHS Public Access

Author manuscript

Plasmid. Author manuscript; available in PMC 2022 April 04.

Published in final edited form as:

Plasmid. 2022 ; 119-120: 102618. doi:10.1016/j.plasmid.2022.102618.

Complete sequence of classic F-type plasmid pRK100 shows unique conservation over time and geographic location

Marjanca Starčič Erjavec[a], **Karmen Jeseničnik**[a], **Lauren P. Elam**[b], **Andrej Kastrin**[c], **Luka Predojević**[a], **Tatyana A. Sysoeva**[b,*]

[a]Department of Biology, Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, SI-1000 Ljubljana, Slovenia

[b]Department of Biological Sciences, The University of Alabama in Huntsville, 301 Sparkman Dr, Huntsville, AL 35899, USA

[c]Institute for Biostatistics and Medical Informatics, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, SI-1000 Ljubljana, Slovenia

Abstract

Plasmids exhibit great diversity of gene content and host ranges and are famous for quick adaptation to the genetic background of the bacterial host cell. In addition to observing ever evolving plasmids, some plasmids have conserved backbones: a stable core composition and arrangement of genes in addition to variable regions. There are a few reports of extremely conserved plasmids. Here we report the complete sequence of pRK100 plasmid – a large, well-characterized conjugative F-like plasmid found in an *Escherichia coli* strain isolated from a urinary tract infection patient in 1990. The sequence shows that the 142 kb-long pRK100 plasmid is nearly identical to plasmids circulating in distant geographical locations and found in different host *E. coli* strains between 2007 and 2017. We also performed additional functional characterization of pRK100. Our results showed that pRK100 does not have a strong pathogenicity phenotype in porcine primary bladder epithelial cell culture. Moreover, the conjugation of pRK100 seems to strongly depend on recipient characteristics. These observations and identification of the pRK100 plasmid in different strain genotypes leave the extreme sequence conservation and broad distribution of this plasmid unexplained.

*Corresponding authors. tatyana.sysoeva@uah.edu (T.A. Sysoeva).

CRediT authorship contribution statement

Marjanca Starčič Erjavec: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. **Karmen Jeseničnik:** Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Lauren P. Elam:** Investigation, Writing – review & editing. **Andrej Kastrin:** Data curation, Formal analysis, Funding acquisition, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. **Luka Predojević:** Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Tatyana A. Sysoeva:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plasmid.2022.102618.

Keywords

pRK100 plasmid; F-like plasmid; Conjugative plasmid conservation

1. Introduction

Antibiotic resistance of bacterial pathogens has become an urgent threat in current medicine. Genes encoding such antibiotic resistance are often carried by extrachromosomal elements - plasmids. Plasmids and plasmid transmission were recognized as the most important phenomena involved in transfer of resistance genes among bacterial strains (Aslam et al., 2018; Johnson and Nolan, 2009; Smillie et al., 2010). Apart from carrying antibiotic resistance genes plasmids can also encode a remarkably diverse array of phenotypic traits of medical, agricultural, environmental, and commercial importance, *e.g.* virulence-associated genes (Kado, 1998; Thomas and Nielsen, 2005).

One type of plasmids that tend to carry a broad repertoire of the virulence genes are the colicin V (ColV) plasmids that encode the production of the toxin colicin V (Johnson, 2021; Johnson et al., 2006; Waters and Crosa, 1991). Virulence determinants of ColV plasmids include siderophores (aerobactin, salmochelin), iron metabolism traits (ABC transporters like Sit and Eit, a hemoglobin protease), hemolysins (Hly), increased serum survival (Iss, Bor) (Di Lorenzo and Stork, 2014; Nolan et al., 2003; Nowrouzian et al., 2001a, 2001b; Searle et al., 2015). Consistently, prior studies showed that ColV plasmids are important for bacterial strains carrying these plasmids for colonization of the host intestine and for increased virulence in diverse infection models, such as urinary tract infection, neonatal rat sepsis, chicken embryos (Aguero et al., 1989; Lemaître et al., 2013; Nolan et al., 2003; Skyberg et al., 2006).

Despite the discovery of conjugative plasmids back in the mid twentieth century, many aspects of plasmid biology remain unclear. For example, it was long thought that presence of large conjugative plasmids is a burden to the bacterial host due to the necessity of maintaining the large, non-essential, DNA molecule and spending energy on expression of the conjugation-associated transfer genes (San Millan and MacLean, 2017). Therefore, the systems that ensure stable plasmid maintenance through addiction modules, partitioning and conjugation regulation have been investigated in great details (Ebersbach and Gerdes, 2005; Frost and Koraimann, 2010; Kroll et al., 2010; Sengupta and Austin, 2011; Sysoeva et al., 2020). But recently more studies show that plasmid-carrying cells (1) often have higher fitness and (2) can undergo rapid genetic changes in the plasmid and in the host chromosome(s) to minimize the cost of carrying the plasmid. This way novel next generation sequencing (NGS) methods now allow determination of complete plasmid sequences and measurement of host chromosome and plasmid co-evolution in almost real time. Rapid co-evolution of plasmids and their hosts was observed already within 100–200 generations and concerns both plasmid and chromosomal genes (Hughes et al., 2012; Hülter et al., 2017; Stalder et al., 2017). On other hand, there are known examples when a plasmid co-exists with its host for extended times without undergoing significant changes. Moreover, with current ability to establish complete sequences of plasmids reports appear

on highly conserved plasmid sequences and backbone structures (Harmer and Hall, 2020; Schmitz-Esser et al., 2015; Weisberg et al., 2020).

The plasmid pRK100 was identified in 1990 in a clinical strain of *Escherichia coli* KS533 isolated from urine of a patient with urinary tract infection (UTI). Initial phenotypic and sequencing characterization of this plasmid found the presence of transfer genes responsible for conjugation of this plasmid, aerobactin iron uptake system, ColV colicin toxin and resistances to ampicillin and tetracycline (Žgur-Bertok et al., 1990). Later, the resistance genes were found to be part of the Tn*5431* transposon, which arose by transposition of Tn*3* carrying ampicillin resistance gene *bla* into Tn*1721* with its tetracycline resistance gene *tetA* (Žgur-Bertok et al., 1996; Žgur-Bertok et al., 1994). Further research revealed that pRK100 plasmid also encodes the ColIa toxin, two replicons, and can integrate into the chromosome. Following studies measured the size of pRK100 at ~145 kB, constructed the restriction map by hybridization experiments, and established nucleotide sequences of several regions of interest (Ambrožič et al., 1998). pRK100 plasmid contains two functional replication regions that were identified by constructing and testing several minireplicon plasmids - an F plasmid related RepFIB and a R1 plasmid related RepFIIA replication region (Starčič Erjavec et al., 2003; Starčič Erjavec and Žgur-Bertok, 2006). PCR and hybridization experiments further demonstrated that pRK100 harbors multiple IS2 and IS3 insertion sequences (Starčič Erjavec et al., 2003). Further it was discovered that the pRK100 *tra* region is the most similar to the *tra* region of the prototypic F plasmid (Starčič Erjavec et al., 2002). Functional studies identified cyclic AMP (cAMP), H-NS and Lrp as regulators of the transfer genes expression (Starčič Erjavec et al., 2003; Starčič et al., 2003).

The evolutionary mechanisms underlying sequence conservation *versus* rapid evolution in different plasmids are unclear. Hence more functional studies are needed to understand plasmid-gene functions and behavior in different hosts in conjunction with the NGS methods to connect the phenotypes with genotypes. The aim of this work was to establish the complete nucleotide sequence of pRK100 plasmid and to perform its further phenotypic characterization to expand our knowledge on this ColV plasmid.

2. Material and methods

2.1. Bacterial culture conditions

All strains were grown in LB (liquid medium or solid plates). When needed the growth medium was supplemented with antibiotics (ampicillin (Amp) 100 μg/ml; tetracycline (Tc) 10 μg/ml; streptomycin (Sm) 150 μg/ml). When grown in liquid medium, the culture was aerated (180 rpm). Bacterial strains used in the study are listed in Table 1.

2.2. DNA isolation and sequencing

Total genomic DNA from CL225 (HB101 strain carrying pRK100) was purified from 8 ml of dense LB-grown culture *via* MagAttract HMW DNA Kit (Qiagen, Germantown, MD). The obtained DNA was used by the Sequencing and Genomic Technologies Core Facility of the Duke University Center for Genomic and Computational Biology for library preparation and sequencing.

For short-read Illumina NextGen sequencing, 500 ng of genomic DNA was sheared to produce ~500 bp DNA fragments. DNA-Seq libraries were prepared using the Kapa BioSystem HyperPrep Library Kit (Cape Town, South Africa), amplified by PCR and purified using AMPure beads (Beckman Coulter, Indianapolis, IN). The library was sequenced on Illumina NextSeq 500 Mid-Output flow cell (Illumina, San Diego, CA, USA) configured for 150 bp reads.

For long-read PacBio sequencing, 500 ng total DNA was sheared to the average 10 kbp size and used for creating the multiplexed microbial SMRTbell Libraries for the PacBio Sequel system (Pacific Biosciences, Menlo Park, CA). The library was subjected to size selection using the BluePipin automated electrophoresis system with the lower cutoff of 4 kbp (SageScience, Beverly, MA) and subjected to sequencing on Sequel instrument. When two alternative assemblies for pRK100 revealed several single nucleotide mismatches, the consensus sequence was established *via* classic Sanger sequencing (Genewiz, NJ) using the specific primers (Table S1).

2.3. Assembly and annotation of the complete pRK100 sequence

Illumina and PacBio reads were used in hybrid data *de novo* assembly using default settings of the Unicycler program (Wick et al., 2017) at the HPC cluster of the Alabama Supercomputer. This assembly resulted in three contigs one of which represented the circularized sequence of pRK100 plasmid with 142,357 bp. The chromosome sequence was assembled as 2 contigs – 4,497,988 bp and 303 bp. The coverage of the pRK100 was at 1.73 copies per chromosomal contigs. The annotation of the resulting pRK100 sequence was done using automated NCBI Prokaryotic Genome Annotation Pipeline (PGAP) upon deposition of the assembled sequence onto GenBank database (Haft et al., 2018; Li et al., 2021; Tatusova et al., 2016).

2.4. Determination of the phylogenetic groups and plasmid similarity

The determination of the KS533 phylogenetic groups was performed according Clermont scheme (Clermont et al., 2013). The determination of the phylogenetic groups for other strains was done by the *in silico* Clermont typing using http://clermontyping.iame-research.center/ with default settings (Beghain et al., 2018; Clermont et al., 2019) and chromosomal assemblies of the relevant strains available online (2009_36 strain GCA_013780485.1, F2_14D strain GCA_013780445.1, 11.3_R3 strain GCA_002001945.1).

Comparison of plasmids was done by online BLASTn tool (Johnson et al., 2008), MAUVE program (Darling et al., 2004), and average nucleotide identity (ANI) calculator (Goris et al., 2007).

2.5. Solid medium conjugation assay

The strain CL225 (HB101) was used as the donor strain for the conjugative transfer of the pRK100 plasmid to recipient BJ69 strain and to MG1655 strain on an LB plate. The conjugative transfer was performed as described earlier (Starčič Erjavec et al., 2015), with

the exception of using minimal A medium plate supplemented with Amp and Tc as the selection plate.

2.6. Quantitative conjugation assays

Single colonies of donor (CL225) and recipient (MG1655, R33, R43, R53, R55) strains were inoculated in 5 ml of liquid LB with the appropriate antibiotics and incubated at 37 °C and 160–180 rpm overnight. Overnight cultures were diluted in a 1:100 ratio in 5 ml of liquid LB without antibiotics and incubated at 37 °C and 160–180 rpm for 2 h. 1 ml of recipient strain culture was centrifuged at 5000 rpm for 10 min, the supernatant was discarded and the harvested cells were resuspended in 400 μl of the donor strain culture. The whole volume was spread over an LB plate and incubated at 37 °C for 24 h. Subsequently, conjugation mixture was collected from the LB plate by resuspension in 1 ml of 1 × dPBS. Tenfold serial dilutions were prepared with 1 × dPBS and plated on selective plates for CFU counts of the transconjugants (minimal medium with glucose and tetracycline), the recipient strain (MacConkey medium) and donor strain (LB medium with streptomycin). All plates were incubated at 37 °C overnight and CFUs of the transconjugants, the recipient and the donor strain were assessed. Selectivity of plates was tested by streaking the donor and the recipient strains separately on selective plates and incubating overnight at 37 °C. The conjugation frequency was calculated as CFU of transconjugants over the CFU of recipient cells. Experiments were done in biological triplicates with consecutive calculation of arithmetic means and standard deviations of the conjugation frequencies.

2.7. Urothelium cell viability assay

For the assessment of the possible change in pathogenicity of the *E. coli* strain BJ69 due to the presence of the pRK100 plasmid an *in vitro* biomimetic pathogenicity model based on the porcine bladder urothelium primary cell culture was used. Firstly, the NPU cells were cultured on multi-well microtiter plastic plates using initially medium for proliferation and then medium for the differentiation of the cells in later stages of growth. Cell culture model was prepared for the inoculation when the NPU cells finally formed highly differentiated multi-layer urothelium. This *in vitro* model of urothelium was inoculated with bacterial cultures (with multiplicity of infection 10:1) and incubated for 3 h. Subsequently, the plate was washed, treated with a commercially available TrypLE™ Select Enzyme and individual urothelial cells were stained using trypan blue dye. Suspension of contrasted NPU cells was transferred to hemocytometer and the NPU cells were counted manually using inverted-light microscope. Viability of the urothelial cells was the main measure of the pathogenicity of *E. coli* strains in this model (Predojević et al., 2018).

2.8. Statistical analysis

For statistical analysis, either viability values obtained in urothelium cell viability assays or conjugational frequencies of strains, Analysis ToolPak in Excel was used. Single factor or one-way ANOVA were used to test the null hypothesis that the data means obtained in the repeated assays with different strains were all equal. When the null hypothesis was rejected by the ANOVA analysis, the F-test (one-tail) was performed in order to test the null hypothesis that the variances of two data sets were equal. Finally, depending on the F-test

results, either the *t*-Test (two-tail) assuming equal or unequal variances was performed to analyze the null hypothesis that the means of two data sets were equal.

3. Results and discussion

3.1. Overview of the complete pRK100 sequence

Using a hybrid sequencing and assembly approach, the complete sequence of the pRK100 plasmid was established. An annotated circular plasmid map of pRK100 is presented below (Fig. 1). The map shows the arrangement of key genes similar to the F plasmid with the transfer region spanning about 35 kb. Sequencing showed 1.73× coverage with the Illumina reads that is reflective of the low copy number of pRK100 predicted before.

The complete sequence (Fig. 1) identified the positions of previously established partial sequences with the exception of the IS*3* element (IS3 GenBank AY230885, Starčič Erjavec et al., 2003). This element was previously detected *via* PCR from pRK100 plasmid isolated from *E. coli* HB101 strain (Starčič et al., 2003). At this stage it is unclear whether this element was excised from the plasmid naturally or it was amplified in the prior PCR assay from contaminating chromosomal sequences, as the host strain HB101's chromosome appears to carry at least three IS*3* copies. Moran and Hall (2018) reported that in the pCERC4 plasmid (a related sublineage plasmid) the IS*3* is inserted in the *traS* gene and due to the polar effects of this IS*3* insertion the plasmid is not able to conjugate. Since pRK100 has already been reported to be capable of conjugative transfer (Žgur-Bertok et al., 1990), it is likely that no IS sequences were present disrupting the *tra* region of pRK100.

3.2. pRK100 plasmid is uniquely conserved and geographically widespread

BLAST search revealed that the complete pRK100 plasmid has four highly homologous plasmids in the GenBank: namely pCERC5, pRHBSTW-00004_2, p2009_36_F, and pF2_14D_F (Moran and Hall, 2018; Reid et al., 2019)(Table 2). In particular, pCERC5 plasmid carries only 16 single nucleotide polymorphisms and 2 gaps, being identical otherwise within the entire 142.4 kb sequences. pCERC5 is reported to have a slightly lower copy number ~ 1.3 in comparison with the estimated from sequence coverage 1.73 in this work. Strikingly, the homologous plasmids were found in strains isolated in Australia (2007–2010) and United Kingdom (2017) while pRK100 was isolated in Slovenia (1990). This shows a broad geographic distribution across continents with extreme sequence conservation for over two decades.

As the detailed comparisons of the pCERC5, p2009_36_F, and pF2_14D_F among themselves and with other closely related plasmids were recently published (Moran and Hall, 2018; Reid et al., 2019) we only briefly summarize the differences of the five most similar plasmids in the Table 2.

pCERC5, p2009_36_F, and pF2_14D_F were found in hosts with very close genotypes and thus Reid and co-workers hypothesized that the plasmid is unchanged because the host strains were closely related (Reid et al., 2019). pRK100 was found in B2 KS533 strain but then was propagated in HB101, which is a K12 clade strain.

Interestingly, p2009_36_F, and pF2_14D_F plasmids were found in strains with a larger, over 200 kb, IncHI2 plasmids, while the pCERC5 and pRK100 were the only plasmids in their respective carrier cells. The identified IncHI2 plasmids carry multiple mobile genetic elements (insertion elements and transposons) that could have potentially contributed to the observed changes in pRK100 homologues, but detailed studies will need to be conducted to test this hypothesis. In addition to this speculation, one can note (Table 2) that there are several changes in the transfer region of the pRK100 homologues that includes splitting of *traG* ORF, modifying *traG* ORF, inserting transposase or inverted repeats (between *traG* and *traS*; *trbA* and *trbN*). Additionally, two observed changes concern colicins'-proximal or interrupting locations. It is unclear at this point if these are just coincidences.

3.3. Phenotypic characterization of pRK100

3.3.1. pRK100 conjugation efficiency strongly depends on recipient cell identity—

As recently it became evident that different *E. coli* strains can differ in their recipient efficiency of a plasmid *via* conjugal transfer (Sysoeva et al., 2020; Kuznetsova et al., 2021), conjugation experiments with pRK100 into five different *E. coli* strains, four clinical uropathogenic *E. coli* strains and one laboratory *E. coli* strain (MG1655), were performed.

Conjugation assays revealed differences in conjugation frequency in mating assays with different recipient strains (Fig. 2, Table S2). As seen from Fig. 2 we were able to successfully perform conjugal transfer of plasmid pRK100 into three different recipients (R43, R53 and MG1655), albeit with very different conjugation frequencies. The arithmetic means and standard deviations of the conjugation frequencies obtained from three independent experiments were as follows: 6.09E-04 ± 3.76E-04 for the mating pair CL225 × MG1655; 6.34E-02 ± 7.74E-02 for the mating pair CL225 × R43 and 7.99E-06 ± 9.07E-06 for the mating pair CL225 × R53. Into the two other strains, R33 and R55 plasmid pRK100 could not be transferred.

The mechanism behind the observed differences in conjugation frequencies could be related to presence of prophages that could be triggered by the conjugation-induced SOS response, synthesis of bacteriocins, differences in biofilm formation abilities, differences in restriction-modification systems, surface exclusion due to recipient's possession of another conjugative plasmid or CRISPR-Cas system (Samson et al., 2015; Stalder and Top, 2016; Thomas and Nielsen, 2005). As the employed R strains were previously (Kuznetsova et al., 2021) analyzed for the presence of prophages, production of bacteriocins and presence of RepFIA sequences in the genome, it could be stated that both recipients into which pRK100 did not conjugate, harbored RepFIA sequence as revealed by PCR and hence probably the presence of another conjugative plasmid was either excluding or incompatible with the pRK100 conjugation. In addition, the phylogentic background of a strain could influence the ability to receive a plasmid. Recent study of transmission of plasmids conferring multi-drug resistance (MDR) suggested that acquisition and maintenance of MDR plasmids might be related to appearance of adaptive mutations in intergenic regions and selection on genes involved in anaerobic metabolism (Dunn et al., 2019). But further studies are needed to identify the underlying mechanisms.

3.3.2. pRK100 plasmid does not increase uropathogenicity of E. coli strains

—As the pRK100 plasmid has several features that could be connected with host strain's virulence, the KS533 strain with the plasmid and the KS533 strain that was cured of the plasmid (KS533-p) were tested for pathogenicity on an *in vitro* biomimetic pathogenicity model based on the porcine bladder urothelium primary cell culture (Predojević et al., 2018). In addition, a natural *E. coli* strain, isolated from feces of a healthy person (BJ69) and the BJ69 harboring pRK100 (BJ69 + p) were tested. For controls, we used the laboratory non-pathogenic *E. coli* strain MG1655 and highly uropathogenic strain J96 showing low and high cytotoxic activity, respectively (Blum et al., 1995).

As shown in Fig. 3 strains possessing the pRK100 plasmid (KS533 and BJ69 + p) caused a small, non-significant, increase in the viability of the urothelial cells of the *in vitro* model compared to the strains without the plasmid (KS533-p, BJ69). The pathogenicity level of all tested natural strains (KS533, KS533-p, BJ69, BJ69 + p) was moderate (Fig. 3) when compared to the two control MG1655 and J96 strains. This observation is in contrast with previously mentioned studies of other ColV plasmids that do increase virulence of carrying bacterial strain in diverse infection models including urinary tract colonization (Aguero et al., 1989; Lemaître et al., 2013; Nolan et al., 2003; Skyberg et al., 2006).

As presence of pRK100 plasmid in two different *E. coli* backgrounds does not increase their cytotoxic effect on the epithelial cells, it is possible that other properties of this plasmid are important for bacterial cells driving plasmid's maintenance and selection in populations. The strain possessing pRK100 plasmid does not grow appreciably slower or faster (not shown) and, therefore, other properties should be at play. If one were to speculate perhaps presence of pRK100 plasmid helps with changed growth or metabolic capacity in urine, bacterial adhesion, or it has increased epithelial cell invasion, as it was shown for a well characterized uropathogenic strain of *E. coli* UTI89 carrying F-type plasmid pUTI89 (Cusumano et al., 2010).

4. Contemporary issues in plasmid classification, taxonomy, and naming

Recent efforts were made to review, to organize and to propose unifying nomenclature of the conjugative transfer or plasmid genes (Fernandez-Lopez et al., 2017; Orlek et al., 2017; Thomas et al., 2017). Several types of classification of plasmids are currently in use or proposed: (1) based on replication mechanisms and incompatibility groups (Carattoli et al., 2005); (2) Mob families for conjugative plasmids using conservation of key transfer protein – relaxase (Francia et al., 2004; Orlek et al., 2017); (3) partitioning mechanism (Bousquet et al., 2015), (4) based on ANI *via* a taxonomic **c**lassifier **o**f **pla**smids tool (COPLA) for plasmid taxonomic units (PTUs) (Redondo-Salvo et al., 2021). Unfortunately, these classifications cannot capture all essential plasmid features, for example, Mob-based classification does not relate to nonconjugative plasmids; replicon typing gets complicated by abundant multireplicon plasmids, such as pRK100.

With current improved NGS and assembly methods, like the hybrid assembly used here, we are able to recover complete plasmid sequences (Douarre et al., 2020; Galata et al., 2018). It is highly likely that we will observe fast-changing, unique plasmids as well as those that

are more well-conserved, similar to the pRK100 plasmid described here. Therefore, as a community, we need to discuss how to better classify plasmids based on their conserved backbone homology and how to treat highly similar homologues in terms of naming, database deposition, and literature discussions. While clearly a generalized classification akin 16S rRNA gene analysis is not possible for plasmids, some classification resembling viral taxonomy might benefit the field. In this example, sequences of two identical plasmids (barring the few substitutions and gaps) were established (CP060383 in current study and KU664810 from Moran and Hall, 2018). It seems prudent to keep one name and database entry to allow for easier literature searches or to have specific notations included into the database entry indicating that the same or nearly the same plasmid was identified independently in another project.

5. Conclusions

Complete sequence of the classic F-type plasmid pRK100 was established using hybrid sequencing approach and *de novo* assembly. Bioinformatic analysis showed that pRK100 sequence is extremely conserved and this plasmid was detected in different sources, locations, and times spanning two decades. Our targeted phenotypic characterization of pRK100 plasmid did not provide an explanation of why this plasmid has the extreme sequence conservation in different genotypic *E. coli* backgrounds. Therefore, as of now it appears that the well characterized antibiotic resistances, conjugative transfer system, encoded colicins and several toxin/antitoxin pairs, hemolytic factors, and enhanced iron scavenging systems provide sufficient benefit to carrying pRK100 plasmid and pressure to keep unchanged sequence throughout this large, 142 kb plasmid. Further studies are needed to trace epidemiology of this plasmid and understand why under some conditions this plasmid remains intact while in other plasmids quickly evolve.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors are thankful to Eva Moreno, Christoph Beloin and Marina V. Kuznetsova for providing the strains used in the study. We thank the Sequencing and Genomic Technologies Core Facility of the Duke University Center for Genomic and Computational Biology for providing PacBio and Illumina sequencing services and Alabama Supercomputer Authority (David Young in particular) for providing computational power for genome assembly. This study was in part sponsored by K12 Duke KURe (DK100024 NIDDK) and UAH Startup funds (USA) and in part by the Slovenian Research Agency grants P10198 and P3-0154 (Slovenia)."

Data availability

Sequencing data is available under BioProject PRJNA657261 and BioSample SAMN15823135 with the complete annotated genome has accession number CP060383.

References

Aguero ME, de la Fuente G, Vivaldi E, Cabello F, 1989. ColV increases the virulence of *Escherichia coli* K1 strains in animal models of neonatal meningitis and urinary infection. Med. Microbiol. Immunol. 1784 (178), 211–216. 10.1007/BF00202554, 1989.

Ambrožič J, Ostroveršnik A, Starčič M, Kuhar I, Grabnar M, Žgur-Bertok D, 1998. *Escherichia coli* ColV plasmid pRK100: genetic organization, stability and conjugal transfer. Microbiology 144, 343–352. 10.1099/00221287-144-2-343. [PubMed: 9493372]

Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool MH, Nisar MA, Alvi RF, Aslam MA, Qamar MU, Salamat MKF, Baloch Z, 2018. Antibiotic resistance: a rundown of a global crisis. Infect. Drug Resist. 11, 1645. 10.2147/IDR.S173867. [PubMed: 30349322]

Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O, 2018. Clermon Typing: an easy-to-use and accurate *in silico* method for *Escherichia* genus strain phylotyping. Microb. Genomics 4. 10.1099/mgen.0.000192.

Blum G, Falbo V, Caprioli A, Hacker J, 1995. Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and a-hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. FEMS Microbiol. Lett. 126, 189–195. 10.1016/0378-1097(95)00009-T. [PubMed: 7705611]

Bousquet A, Henquet S, Compain F, Genel N, Arlet G, Decré D, 2015. Partition locus-based classification of selected plasmids in *Klebsiella pneumoniae*, *Escherichia coli* and *salmonella enterica* spp.: an additional tool. J. Microbiol. Methods 110, 85–91. 10.1016/J.MIMET.2015.01.019. [PubMed: 25623509]

Carattoli A, Bertini A, Villa L, Falbo V, Hopkins K, Threlfall E, 2005. Identification of plasmids by PCR-based replicon typing. J. Microbiol. Methods 63, 219–228. 10.1016/J.MIMET.2005.03.018. [PubMed: 15935499]

Clermont O, Christenson JK, Denamur E, Gordon DM, 2013. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ. Microbiol. Rep. 5, 58–65. 10.1111/1758-2229.12019. [PubMed: 23757131]

Clermont O, Dixit OVA, Vangchhia B, Condamine B, Dion S, Bridier-Nahmias A, Denamur E, Gordon D, 2019. Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. Environ. Microbiol. 21, 3107–3117. 10.1111/1462-2920.14713. [PubMed: 31188527]

Cusumano C, Hung C, Chen S, Hultgren S, 2010. Virulence plasmid harbored by uropathogenic *Escherichia coli* functions in acute stages of pathogenesis. Infect. Immun. 78, 1457–1467. 10.1128/IAI.01260-09. [PubMed: 20123719]

Darling ACE, Mau B, Blattner FR, Perna NT, 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394–1403. 10.1101/GR.2289704. [PubMed: 15231754]

Di Lorenzo M, Stork M, 2014. Plasmid-encoded iron uptake systems. Microbiol. Spectr. 2 10.1128/MICROBIOLSPEC.PLAS-0030-2014/ASSET/5017EBF0-7A54-4D1A-977D-3022A26982C8/ASSETS/GRAPHIC/PLAS-0030-2014-FIG1B.GIF.

Douarre PE, Mallet L, Radomski N, Felten A, Mistou MY, 2020. Analysis of COMPASS, a new comprehensive plasmiddatabase revealed prevalence of multireplicon and extensive diversity of IncF plasmids. Front. Microbiol. 11, 483. 10.3389/FMICB.2020.00483/BIBTEX. [PubMed: 32265894]

Dunn SJ, Connor C, McNally A, 2019. The evolution and transmission of multi-drug resistant *Escherichia coli* and *Klebsiella pneumoniae*: the complexity of clones and plasmids. Curr. Opin. Microbiol. 51, 51–56. 10.1016/J.MIB.2019.06.004. [PubMed: 31325664]

Ebersbach G, Gerdes K, 2005. Plasmid Segregation Mechanisms, vol. 39, pp. 453–479. 10.1146/annurev.genet.38.072902.091252.

Fernandez-Lopez R, Redondo S, Garcillan-Barcia MP, de la Cruz F, 2017. Towards a taxonomy of conjugative plasmids. Curr. Opin. Microbiol. 10.1016/j.mib.2017.05.005.

Francia MV, Varsaki A, Garcillan-Barcia MP, Latorre A, Drainas C, de la Cruz F, 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28, 79–100. 10.1016/J.FEMSRE.2003.09.001. [PubMed: 14975531]

Frost LS, Koraimann G, 2010. Regulation of bacterial conjugation: Balancing opportunity with adversity. In: Future Microbiology. Future Medicine Ltd. 10.2217/fmb.10.70.

Galata V, Fehlmann T, Backes C, Keller A, 2018. PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res. 47, 195–202. 10.1093/nar/gky1050.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM, 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81–91. 10.1099/IJS.0.64483-0. [PubMed: 17220447]

Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD, 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res. 46, D851–D860. 10.1093/nar/gkx1068. [PubMed: 29112715]

Harmer CJ, Hall RM. The Complete Nucleotidesequence of pZM3, a 1970 FIA:FIB:FII Plasmid Carrying Antibiotic Resistance and Virulence Determinants. https://home.liebertpub.com/mdr.

Hughes JM, Lohman BK, Deckert GE, Nichols EP, Settles M, Abdo Z, Topa EM, 2012. The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. MBio 3. 10.1128/mBio.00077-12.

Hülter N, Ilhan J, Wein T, Kadibalban AS, Hammerschmidt K, Dagan T, 2017. An evolutionary perspective on plasmid lifestyle modes. Curr. Opin. Microbiol. 10.1016/j.mib.2017.05.001.

Johnson TJ, 2021. Role of plasmids in the ecology and evolution of "high-risk" extraintestinal pathogenic *Escherichia coli* clones. EcoSal Plus. 10.1128/ECOSALPLUS.ESP-0013-2020/ASSET/96C9CE11-D0CE-4365-9D92-063DD163FB71/ASSETS/IMAGES/LARGE/ECOSALPLUS.ESP-0013-2020-F001.JPG.

Johnson TJ, Nolan LK, 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. Microbiol. Mol. Biol. Rev. 73, 750–774. 10.1128/MMBR.00015-09. [PubMed: 19946140]

Johnson TJ, Siek KE, Johnson SJ, Nolan LK, 2006. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. J. Bacteriol. 188, 745–758. 10.1128/JB.188.2.745-758.2006. [PubMed: 16385064]

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL, 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36, W5–W9. 10.1093/NAR/GKN201. [PubMed: 18440982]

Kado CI, 1998. Origin and evolution of plasmids. Antonie Van Leeuwenhoek 73, 117–126. 10.1023/A:1000652513822. [PubMed: 9602285]

Kroll J, Klinter S, Schneider C, Voβ I, Steinbüchel A, 2010. Plasmid Addiction Systems: Perspectives and Applications in Biotechnology. 10.1111/j.1751-7915.2010.00170.x.

Kuznetsova MV, Maslennikova IL, Pospelova JS, Žgur Bertok D, Starčič Erjavec M, 2021. Differences in recipient ability of uropathogenic *Escherichia coli* strains in relation with their pathogenic potential. Infect. Genet. Evol. 97, 105160. 10.1016/j.meegid.2021.105160. [PubMed: 34839025]

Lemaître C, Mahjoub-Messai F, Dupont D, Caro V, Diancourt L, Bingen E, Bidet P, Bonacorsi S, 2013. A conserved virulence plasmidic region contributes to the virulence of the multiresistant Escherichia coli meningitis strain S286 belonging to phylogenetic group C. PLoS One 8. 10.1371/JOURNAL.PONE.0074423.

Li W, O'Neill KR, Haft DH, Dicuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F, 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res. 49, D1020–D1028. 10.1093/nar/gkaa1105. [PubMed: 33270901]

Moran R, Hall R, 2018. Evolution of regions containing antibiotic resistance genes in FII-2-FIB-1 ColV-Colla virulence plasmids. Microb. Drug Resist. 24, 411–421. 10.1089/MDR.2017.0177. [PubMed: 28922058]

Nolan LK, Horne SM, Giddings CW, Foley SL, Johnson TJ, Lynne AM, Skyberg J, 2003. Resistance to serum complement, *iss*, and virulence of avian *Escherichia coli*. Vet. Res. Commun. 272(27), 101–110. 10.1023/A:1022854902700,2003.

Nowrouzian F, Adlerberth I, Wold AE, 2001a. P fimbriae, capsule and aerobactin characterize colonic resident *Escherichia coli*. Epidemiol. Infect. 126, 11–18. 10.1017/S0950268801005118. [PubMed: 11293669]

Nowrouzian F, Wold AE, Adlerberth I, 2001b. P fimbriae and aerobactin as intestinal colonization factors for *Escherichia coli* in Pakistani infants. Epidemiol. Infect. 126, 19–23. 10.1017/S095026880100512X. [PubMed: 11293677]

Orlek A, Phan H, Sheppard AE, Doumith M, Ellington M, Peto T, Crook D, Walker AS, Woodford N, Anjum MF, Stoesser N, 2017. Ordering the mob: insights into replicon and MOB typing schemes from analysis of a curated dataset of publicly available plasmids. Plasmid 91, 42–52. 10.1016/J.PLASMID.2017.03.002. [PubMed: 28286183]

Predojević L, Žgur-Bertok D, Keše D, Erdani-Kreft M, Starčič Erjavec M, 2018. New set up for pathogenicity evaluation of human uropathogenic *Escherichia coli* strains based on biomimmetic model of normal porcine urothelium in vitro. In: Ivsina IB (Ed.), Vysokie Tehnologii, Opredeljajuščie Kačestvo Žizni: Materialy II Meždunarodnoj Naučnoj Konferencii, Perm'. Perm University Press, Perm, Russia, pp.37–39.

Redondo-Salvo S, Bartomeus-Peñalver R, Vielva L, Tagg KA, Webb HE, Fernández-López R, de la Cruz F, 2021. COPLA, a taxonomic classifier of plasmids. BMC Bioinforma 221(22), 1–9. 10.1186/S12859-021-04299-X, 2021.

Reid CJ, Mckinnon J, Djordjevic SP, 2019. Clonal ST131-H22 *Escherichia coli* strains from a healthy pig and a human urinary tract infection carry highly similar resistance and virulence plasmids. Microb. Genomics 5. 10.1099/mgen.0.000295.

Samson JE, Magadan AH, Moineau S, 2015. The CRISPR-cas immune system and genetic transfers: reaching an equilibrium. Microbiol. Spectr. 3 10.1128/microbiolspec.plas-0034-2014.

San Millan A, MacLean R, 2017. Fitness costs of plasmids: a limit to plasmid transmission. Microbiol. Spectr. 5 10.1128/MICROBIOLSPEC.MTBP-0016-2017.

Schmitz-Esser S, Müller A, Stessl B, Wagner M, 2015. Genomes of sequence type 121 *Listeria monocytogenes* strains harbor highly conserved plasmids and prophages. Front. Microbiol. 6 10.3389/FMICB.2015.00380.

Searle LJ, Méric G, Porcelli I, Sheppard SK, Lucchini S, 2015. Variation in siderophore biosynthetic gene distribution and production across environmental and faecal populations of *Escherichia coli*. PLoS One 10. 10.1371/JOURNAL.PONE.0117906.

Sengupta M, Austin S, 2011. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. Infect. Immun. 79, 2502. 10.1128/IAI.00127-11. [PubMed: 21555398]

Skyberg JA, Johnson TJ, Johnson JR, Clabots C, Logue CM, Nolan LK, 2006. Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. Infect. Immun. 74, 6287–6292. 10.1128/IAI.00363-06. [PubMed: 16954398]

Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F, 2010. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 74, 434–452. 10.1128/MMBR.00020-10. [PubMed: 20805406]

Stalder T, Top E, 2016. Plasmid transfer in biofilms: a perspective on limitations and opportunities. npj Biofilms Microb. 10.1038/npjbiofilms.2016.22.

Stalder T, Rogers LM, Renfrow C, Yano H, Smith Z, Top EM, 2017. Emerging patterns of plasmid-host coevolution that stabilize antibiotic resistance. Sci. Rep. 7, 1–10. 10.1038/s41598-017-04662-0. [PubMed: 28127051]

Starčič Erjavec M, van Putten JPM, Gaastra W, Jordi BJAM, Grabnar M, Žgur-Bertok D, 2003. H-NS and Lrp serve as positive modulators of traJ expression from the Escherichia coli plasmid pRK100. Mol. Gen. Genomics. 270, 94–102. 10.1007/s00438-003-0908-1.

Starčič Erjavec M, Žgur-Bertok D, 2006. The RepFIIA replicon of the natural *Escherichia coli* plasmid pRK100. Acta Biol. Slov. 49, 3–12.

Starčič Erjavec M, Gaastra W, Žgur-Bertok D, 2002. *tra* region of the natural *Escherichia coli* plasmid pRK100 is F-like. Acta Biol. Slov. 45, 9–15.

Starčič Erjavec M, Gaastra W, van Putten J, Žgur-Bertok D, 2003. Identification of the origin of replications and partial characterization of plasmid pRK100. Plasmid 50, 102–112. 10.1016/S0147-619X(03)00016-7. [PubMed: 12932736]

Starčič Erjavec M, Jesenko B, Petkovšek Ž., Žgur-Bertok D, 2010. Prevalence and associations of *tcpC*, a gene encoding a toll/interleukin-1 receptor domain-containing protein, among *Escherichia coli* urinary tract infection, skin and soft tissue infection, and commensal isolates. J. Clin. Microbiol. 48, 966–968. 10.1128/JCM.01227-09. [PubMed: 20042631]

Starčič Erjavec M, Petkovšek Ž., Kuznetsova MV, Maslennikova IL, Žgur-Bertok D, 2015. Strain ŽP - the first bacterial conjugation-based "kill"-"anti-kill" antimicrobial system. Plasmid 82, 28–34. 10.1016/j.plasmid.2015.10.001. [PubMed: 26436830]

Starčič M, Žgur-Bertok D, Jordi BJAM, östen MMSM, Gaastra W, van Putten JPM, 2003. The cyclic AMP-cyclic AMP receptor protein complex regulates activity of the *traJ* promoter of the *Escherichia coli* conjugative plasmid pRK100. J. Bacteriol. 185, 1616–1623. 10.1128/JB.185.5.1616-1623.2003. [PubMed: 12591879]

Sysoeva TA, Kim Y, Rodriguez J, Lopatkin AJ, You L, 2020. Growth-stage-dependent regulation of conjugation. AICHE J. 66, e16848 10.1002/aic.16848.

Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J, 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614–6624. 10.1093/nar/gkw569. [PubMed: 27342282]

Thomas CM, Nielsen KM, 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 10.1038/nrmicro1234.

Thomas CM, Thomson NR, Cerdeñ-T˜áraga ÁM, Brown CJ, Top EM, Frost LS, 2017.A nnotationo fp lasmidg enes. Plasmid. 10.1016/j.plasmid.2017.03.006.

Waters VL, Crosa JH, 1991. Colicin V virulence plasmids. Microbiol. Rev. 55, 437–450. 10.1128/MR.55.3.437-450.1991. [PubMed: 1943995]

Weisberg AJ, DavisII EW, Tabima J, Belcher MS, Miller M, Kuo C-H, Loper JE, Grünwald NJ, Putnam ML, Chang JH, 2020. Unexpected conservation and global transmission of agrobacterial virulence plasmids. Science (80-.) 368. 10.1126/SCIENCE.ABA5256.

Wick RR, Judd LM, Gorrie CL, Holt KE, 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13, e1005595 10.1371/journal.pcbi.1005595. [PubMed: 28594827]

Žgur-Bertok D,Modrič E,Grabnar M, 1990. Aerobactin uptake system, ColV production, and drug resistance encoded by a plasmid from an urinary tract infection *Escherichia coli* strain of human origin. Can. J. Microbiol. 36, 297–299. 10.1139/m90-051. [PubMed: 2192786]

Žgur-Bertok D,Ambrožič J,Podlesek Z,Grabnar M, 1994. Tn*5431*,a new transposable element composed of Tn*1721* and Tn*3*-like genes.Plasmid32,95–99. 10.1006/plas.1994.1050. [PubMed: 7991678]

Žgur-Bertok D,AmbrožiČ J,Grabnar M,1996.Tn*5431* arose by transposition of Tn*3* into Tn*1721*. Can. J. Microbiol. 42, 1274–1276. 10.1139/m96-166. [PubMed: 8989866]

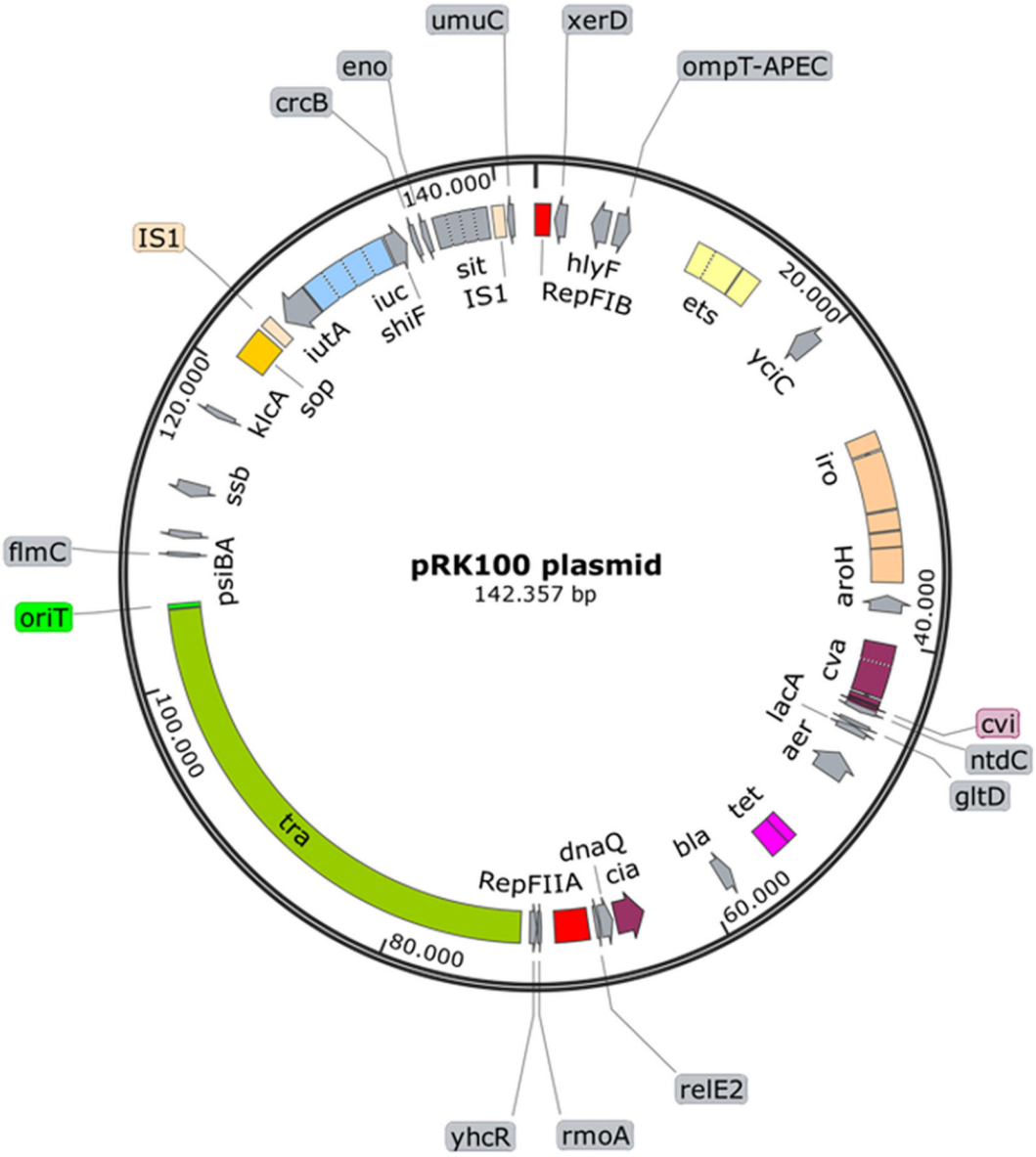

Fig. 1.

Map of conjugative plasmid pRK100. Selected annotated open reading frames encoding functional elements are shown. The map is prepared using SnapGene software (from Insightful Science), which is freely available at https://www.snapgene.com/snapgene-viewer/download. The *bla* and *tet* genes are part of the Tn*1207* (previously Tn*5431*) element spanning from 47,393 to 63,475 bp.

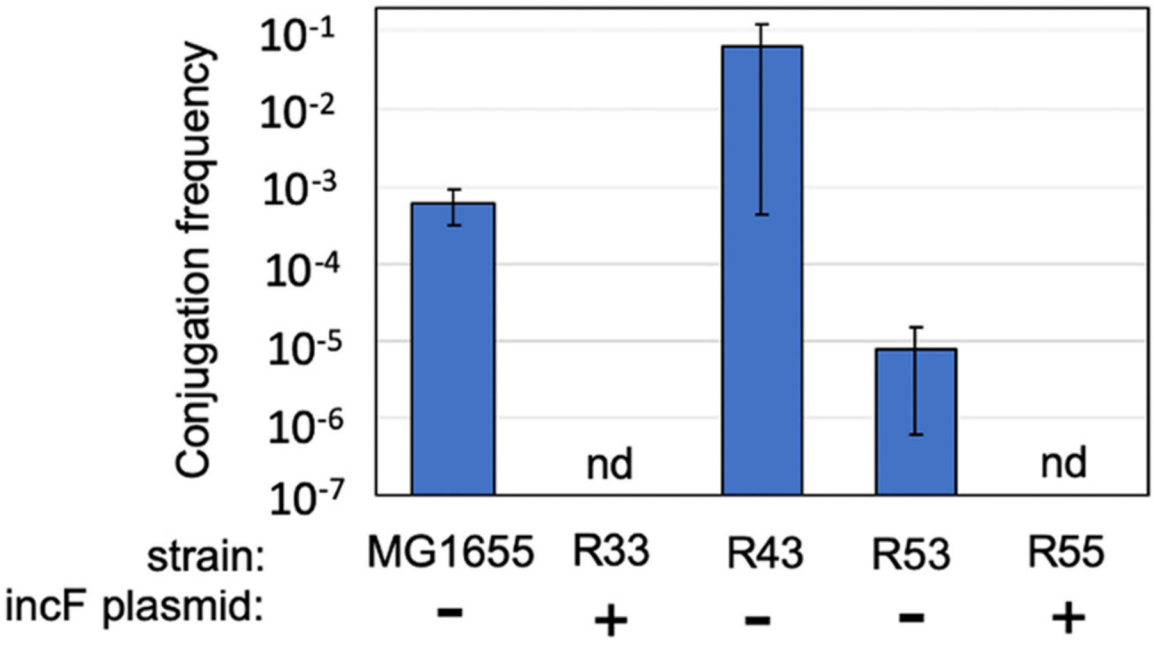

Fig. 2.

Recipient background strongly affects conjugative transfer of pRK100 plasmid. Four recent clinical isolates of uropathogenic *E. coli* and the laboratory MG1655 *E. coli* strain were used as recipients of pRK100 conjugation from CL225 strain. The average conjugation frequency, expressed as number of transconjugants (colony forming units/ml, CFU/ml) per number of recipients (CFU/ml), and respective standard deviation are calculated from three independent mating assays. In strains R33 and R53 the conjugation frequency was below the detection limit of the assay. Statistical analysis showed that the observed pairwise differences in conjugation frequencies were statistically significant (MG1655:R43 $p = 0.0161$; MG1655:R53 $p = 0.004$ and R43:R53 $p = 0.0008$). Presence of possibly excluding *incF* plasmid in recipient background is labeled based on PCR amplification data (Kuznetsova et al., 2021).

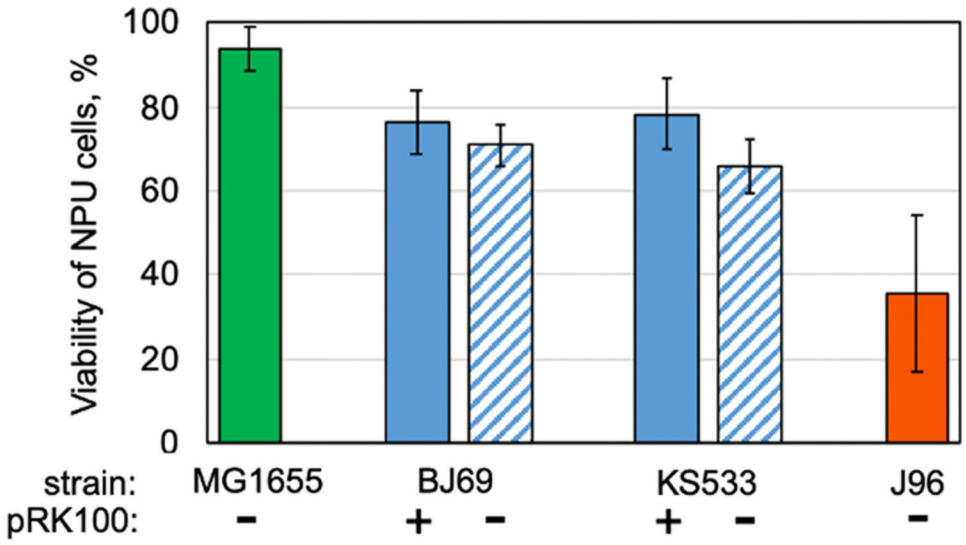

Fig. 3.

pRK100 plasmid does not increase virulence of *E. coli* bacteria towards urothelial cells. Viability of the porcine urothelial (NPU) cells was assessed after infecting them with various *E. coli* strains: non-pathogenic laboratory strain MG1655 (green); human uropathogenic strain J96 (red); human commensal strain BJ69 and urinary tract infection isolate KS533 with (solid fill) and without (slanted) pRK100 plasmid (blue). Unlike difference between non-pathogenic MG1655 and uropathogenic J96 strains, the differences between BJ69 with and without pRK100 as well as KS533 and KS533 without pRK100 were not statistically significant.

Table 1

Bacterial strains used in this study.

Bacterial strain	Phylogenetic group	Characteristics	Reference or source
KS533	B2	Natural *Escherichia coli* strain, isolated from urine of a patient with a urinary tract infection (UTI), original strain harboring pRK100	Žgur-Bertok et al., 1990
KS533-p	B2	KS533 strain without the plasmid pRK100 (KS533 was treated with SDS in order to lose the plasmid)	Ambrožič et al., 1998
CL225	A	*E. coli* HB101 strain with pRK100	Ambrožič et al., 1998
J96	B2	Natural *E. coli* strain, isolated from urine of a UTI patient, prototype of an uropathogenic *E. coli*	E. Moreno
MG1655	A	Model non-pathogenic laboratory K12 *E. coli* strain	C. Beloin
BJ69	B2	Commensal *E. coli* strain isolated from feces of a healthy person	Starčič Erjavec et al., 2010
BJ69 + p	B2	BJ69 strain with pRK100	This study
R33	B2	Clinical UTI *E. coli* isolate	M. Kuznetsova
R43	B1	Clinical UTI *E. coli* isolate	M. Kuznetsova
R53	C	Clinical UTI *E. coli* isolate	M. Kuznetsova
R55	B2	Clinical UTI *E. coli* isolate	M. Kuznetsova

Table 2

List of plasmids that show the closest homology to pRKlOO plasmid.

Plasmid name	Size, bp	GenBank accession number	ANI[a], %	Query coverage, %	Percent identity	Differences with pRKlOO	Isolation source	Isolation location	Isolation year	Host Phylogroup	Co-habitant plasmids	Reference or source
pRKlOO	142,357	CP060383	na[b]	na	na	na	Urine from patient with UTI	Slovenia	1990	B2	none	Žgur-Bertok et al., 1990
pCERC5	142,359	KU664810.1	99.99	100	99.99% (46 mismatches)	No large rearrangements/ insertions/ deletions	Feces of a healthy human	Australia	2010	B2	none	Moran and Hall, 2018
pRHBSTW-00004 2	143,134	CP056915.1	99.99	100	99.99% (49 mismatches)	ISl-like element IS1A family transposase in between traG and traS ORFs	Wastewater influent	United Kingdom	2017	B2	none	Shaw L.P., unpublished,
p2009_36_F	143,671	MK461929.1	99.99	99	99.99% (48 mismatches)	IS3 transposase is inserted in a colicin gene breaking it into two ORFs	Human catheter stream urine from person with UTI	Australia	2009	B2	IncHI2	Reid et al., 2019
pF2_14D_F	139,372	MK46192.1	99.95	97	99.90%; (45 mismatches)	Insertion sequences in *trbB* ORF (ISEc23), modified *traG*; Tnl721- derivative transposon	Fecal swab of a healthy piglet	Australia	2007	B2	IncHI2	Reid et al., 2019

[a]ANI - average nucleotide identity.

[b]na - not applicable.