Norwegian University
of Life Sciences

Master's Thesis 2020 60 ECTS
Faculty of Chemistry, Biotechnology and Food Science

IncI1 Resistance Plasmids: Studies on Fitness Cost, Long-term Stability, and Shufflon Rearrangements

Milan Stosic
MSc Biotechnology, Microbiology

Acknowledgments

This master's thesis was performed either at the Norwegian University of Life Sciences, Faculty of Chemistry, Biotechnology and Food Science with Professor Knut Rudi as a supervisor, or at the Norwegian Veterinary Institute (NVI) in Oslo, with Senior Researcher Marianne Sunde as a supervisor. The work done at NVI was a part of the ARDIG project (Antibiotic resistance Dynamics: the influence of geographic origin and management system on resistance gene flows within humans, animals and the environment).

I would like to thank my supervisor Knut Rudi for giving me the opportunity to explore the world of plasmids and long-range sequencing. Thank you for your feedback, suggestions and help. Special thanks to Laboratory Engineers at NMBU Ida Ormaasen and Inga Leena Angell for all the help, tips and positive energy in your lab.

I would also like to extend heartfelt thanks to my secondary supervisor Marianne Sunde. Thank you for your support and guidance, your positive attitude, abundant feedback on my work, and always keeping your office door open for me. A big thanks to Solveig Sølverød Mo, Researcher at NVI, for always answering my cries for help, for all the tips on conjugation experiments, and for always being the objective. I would also like to thank Amar Anandrao Telke, Senior Researcher at NVI, for providing me with valuable lessons on genome assembly, annotation, and read mapping, and for showing me how to transfer plasmid sequencing data into beautiful figures. My time at NVI would not have been the same without Fiona Franklin-Alming and Thongpan Leangapichart - Win my office-buddies. Thank you for wonderful moments filled with every possible human and bacterial emotion.

I would like to thank my fiancée, Aleksandar Vugdelija for keeping up with my emotional rollercoasters and offering all the support I needed over the years. A special thanks to my friends, Andrijana Pavlovic and Milena Stefanovic, who were forced to learn about plasmids, plasmid inflicted fitness cost, plasmid stability and shufflons.

And finally, a big thanks to my family for encouraging and supporting me throughout the last five years.

Ås, June 2019

Milan Stosic

Abstract

The plasmid-mediated spread of antibiotic resistance genes is considered a major dissemination pathway between related and non-related bacteria. IncI1 plasmids have been found to be carriers of a wide variety of bacterial accessory genes that code for different abilities such as antibiotic resistance, virulence, utilization of different nutrients, and heavy metal tolerance. Without selection for these abilities, the plasmid could impose a burden to its host, thereby reducing its fitness. One of the aims of the thesis was to examine to what extent IncI1 plasmids carrying ESBL/AmpC genes inflict a fitness cost to their new hosts under conditions not selecting for the plasmid encoded traits. The new hosts for the selected plasmids were three Quinolone Resistant E. coli (QREC) strains with three different STs, and one Avian Pathogenic E. coli (APEC) strain. During single strain growth assay, the carrying capacity of most of QREC transconjugants was significantly lower than carrying capacity of their respective plasmid-free recipients, while APEC transconjugants exhibited no such difference compared to their plasmid-free recipient. The competitive growth experiment of selected transconjugant/plasmid-free recipient pairs revealed reduced competitive fitness of transconjugants. However, the magnitude of the competitive fitness reduction appeared to be both plasmid and host dependent.

A previous study revealed that IncI1 plasmids were stable during competitive growth of plasmid-containing cells with their plasmid-free counterpart even when the initial number of plasmid-containing cells was 100x lower. This study obtained comparable results. However, instead of a gradual increase of the number of plasmids throughout the experiment, an instantaneous increase of the number of plasmids in all competing pairs after only 24 hours was detected, further indicating a high rate of plasmid transfer between plasmid-free and plasmid-containing cells. However, the strain chosen to examine the stability of selected plasmids was a laboratory DH5α rif^R E. coli strain with an exceptional ability to accept and maintain plasmids. Future long-term stability studies of IncI1 plasmids in wildtype E. coli strains rather than laboratory model strain are needed to examine the actual stability of the IncI1 plasmids.

Finally, this study also attempted to quantitatively analyse rearrangements of insertion-sequence interrupted shufflons and uninterrupted shufflon during single strain growth. An insertion sequence (ISEcp1) harbouring $bla_{CTX\text{-}M\text{-}1}$ was a part of the B shufflon segment. To further examine whether the host of plasmids with the interrupted shufflon affects its rearrangement, plasmids were grown in their original host or in the QREC 2773(ST162) strain. Results confirmed that the interrupted shufflons generated fewer variants compared to the uninterrupted shuflon, although both shufflon types exhibited a predominance of certain plasmid-specific variants, regardless of the host or sampling time point. Additionally, shufflon variants with a deletion of one or two segments were detected in both shufflon types. The predominant truncated variants of the interrupted shufflon suffered from the deletion of the B segment alone or B and C segments together. This finding implicates that Rci, site-specific recombinase, was challenged when inverting a 3kbp longer segment that often resulted in segment deletion. Finally, the most abundant segment found to complete the pilV ORF was the A segment as previously reported. As long-read sequencing was found suitable for the structural shufflon rearrangement analysis, future studies should focus on uncovering whether the same pattern of shufflon rearrangements observed during single strain growth would also be observed during mating.

Sammendrag

Den plasmid-medierte spredningen av antibiotikaresistens gener ansees som en viktig spredningsvei mellom beslektede og ikke-beslektede bakterier. IncI1-plasmider er funnet til å være bærere av et bredt utvalg av bakterielle tilbehørgener som koder for forskjellige evner som antibiotikaresistens, virulens, utnyttelse av forskjellige næringsstoffer, og tungmetalltoleranse. Uten påført seleksjon for disse evnene, kan plasmidet påføre verten en belastning, og dermed redusere dens fitness. Denne avhandlingen hadde som mål å undersøke i hvilken grad IncI1-plasmider, bærere av ESBL / AmpC-gener, påvirker fitness av sine nye verter under forhold som ikke påfører seleksjon for plasmidkodede egenskaper. De nye vertene for de valgte plasmidene var tre Qunilone Resistant E. coli (QREC) stammer med tre forskjellige ST-er, og en Avian Pathogenic E. coli (APEC) stamme. Under singel-stamme vekstanalyse var bæreevnen til de fleste av QREC-transkonjuganter statistisk betydelig lavere enn bæreevnen til deres respektive plasmidfrie resipientstammer, mens APEC-transkonjuganter ikke utviste en statistisk betydelig forskjell sammenlignet med deres plasmidfrie resipientstamme. Konkurrerende vekstekseriment av utvalgte transkonjugant/resipient par utviste redusert konkurranseevne av transkonjuganter. Imidlertid ser omfanget på fitness-reduksjon ut til å være både plasmid- og vertsavhengig.

En tidligere studie avslørte at IncI1-plasmider var stabile under konkurrerende vekst av plasmidholdige celler med deres plasmidfrie motpart selv når det opprinnelige antallet av plasmidholdige celler var 100 ganger lavere. Sammenlignbare resultater ble oppnådd i dette studiet. Interessant nok, i stedet for en gradvis økning av antall plasmider gjennom hele eksperimentet, ble en øyeblikkelig økning av antall plasmider i alle konkurrerende par påvist etter bare 24 timer som videre indikerer en høy hastighet av plasmidoverføring mellom plasmidfrie og plasmidholdige celler. Imidlertid var en DH5a rif^R E. coli den utvalgte stammen for å undersøke stabiliteten til utvalgte plasmider. Denne stammen er en kjent labbstammen med en eksepsjonell evne til å akseptere og opprettholde plasmider. Fremtidig langsiktig stabilitetsstudier av IncI1-plasmider i villtype E. coli stammer i stedet for laboratoriemodellstamme er nødvendig for å undersøke den faktiske stabiliteten til IncI1 plasmidene.

Til slutt forsøkte denne studien også å kvantitativt analysere rearrangering av shufflon-er som var avbrutt av en insertion sekvens og et uavbrutt shufflon under singel-stamme vekst. Den insertion sekvensen (ISEcp1) som var også bæreren av $bla_{CTX-M-1}$, var en del av B shufflon-segmentet. For å ytterligere undersøke om verten av plasmider med den avbrutte shuffloner påvirker dets rearrangering ble plasmidene dyrket i sine opprinnelige verter eller i QREC 2773 (ST162) stammen. Resultatene bekreftet at de avbrutte shufflon-ene genererte færre varianter sammenlignet med det uavbrutte shuflon-et selv om begge shufflon-typene utpekte en overvekt av visse plasmidspesifikke varianter, uavhengig av verten eller prøvetakingstidspunktet. I tillegg ble shufflon-varianter med delesjonen av ett eller to segmenter påvist i begge shufflon-typer. De dominerende avkortede variantene av det avbrutte shufflon-et hadde delesjon av B-segmentet alene eller B- og C-segmentene sammen. Dette impliserer at Rci, stedsspesifikk rekombinase, ble utfordret når den inverterte et 3kbp lengre segment som ofte resulterte i delesjon av segmentet. Til slutt ble det funnet ut at segmentet som i de fleste tilfellene fullførte pilV ORF, var A-segmentet. Dette ble rapportert tidligere. Siden long-read sekvensering ble funnet egnet for analysen av shufflonets strukturell rearrangering, bør fremtidige studier fokuseres på å avdekke om det samme mønsteret av shufflon rearrangering som ble observert under singel-stamme vekst også ville bli observert under bakteriell parring.

Abbreviations

AmpC – Broad spectrum beta lactamase
AMR – Antimicrobial resistance
APEC – Avian Pathogenic *E. coli*
ARG – Antimicrobial resistance gene
BLAST – Basic Local Alignment Search Tool
CFU – Colony forming units
CI – Confidence interval
CRISPR-Cas – Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR associated protein
DNA – Deoxyribonucleic acid
cDNA – copy DNA
ESBL – Extended Spectrum β-lactamase
HGT – Horizontal Gene Transfer
K – Carrying capacity
LB – Luria-Bertani
LPS – Lipopolysaccharide
MH – Mueller Hinton
MLST – Multilocus Sequence Typing
MOA – Mode of Action
MGE – Mobile Genetic Element
OD – Optical Density
ON – Overnight
PCR – Polymerase Chain Reaction
qPCR – quantitative PCR
r_{max} – Maximum growth rate
RNA – Ribonucleic acid
rRNA – ribosomal RNA
tRNA – transport RNA
RPA – Relative Plasmid Abundance
s – selection rate constant
SCAI – Simmons citrate 1% inositol
SD – Standard Deviation
TF – Transfer Frequency
QREC – Quinolone Resistant *E. coli*

Table of content

1. Introduction

Once a miracle drug, antibiotics have begun to lose their therapeutic effects due to the emergence of antimicrobial-resistance (AMR) in different bacterial species (Lerminiaux & Cameron, 2019), forcing humanity back into the pre-antibiotic era (Argudin et al., 2017). As synthesis of new antibiotics become rarer (Deng, 2018), and the number of antibiotic-resistant bacteria rises worldwide, AMR is now considered a serious public health hazard (CDC, 2018).

Antimicrobial resistance genes (ARGs), which are not an unusual feature of chromosomes of naturally occurring microorganisms, are also found on mobile genetic elements (MGEs). Different horizontal gene transfer (HGT) pathways mediate the transfer of these elements between related and non-related bacteria. Conjugation, most commonly mediated by conjugative plasmids, is considered the major pathway ARGs are being disseminated between different bacterial species (Buckner et al., 2018). The IncI1 group of plasmids have been isolated from both human and animal originating bacteria and are often found to be carriers of different ARGs (Argudin et al., 2017). The scientific interest in this plasmid group became greater with the discovery that these plasmids are important vectors of ESBL/AmpC gene exchange within the *Enterobacteriaceae* family (Partridge et al., 2018). Additionally, these plasmids persist in the bacterial population even with conditions not selecting for plasmid encoded ARGs. Under these conditions, the maintenance of the plasmid and its conjugative transfer could impose a burden to the host cell by inducing the non-beneficial cell resource utilization. The outcome would be a reduced fitness of the plasmid host (San Millan & Craig maclean, 2019).

To halt the spread of IncI1 conjugative plasmids and, thereby, dissemination of ARGs carried by them, greater understanding of these unique bacterial extrachromosomal DNA formations and their mechanisms is needed. Especially mechanisms that govern the persistence of the plasmid in a bacterial population, their transfer to and maintenance within different hosts, as well as effects that the plasmid could induce in the host cell.

1.1. Groups of antibiotics and their mode of action

Antibiotics used nowadays are either modified naturally produced antibiotics (semi-synthetic) or synthetically designed antibiotics with no known parallel in nature, the former being in greater use (Bhattacharjee, 2016; C Reygaert, 2018).

Based on the cellular component targeted by the antibiotic, and mechanism of antibiotic activity, antibiotics are classified into six groups: I) bacterial cell wall synthesis inhibitors, II) disruptors of cell membrane, III) metabolite synthesis inhibitors, IV) DNA synthesis inhibitors, V) RNA synthesis

inhibitors and VI) protein synthesis inhibitors (Bhattacharjee, 2016). Table 1.1 shows modes of action (MOA) of each antibiotic class/substance, together with their respective examples, while Figure 1.1 illustrates the cell components targeted by the different antibiotic groups.

Another antibiotic classification is based on the final effect that antibiotics exert on a bacterial cell: a bactericidal or bacteriostatic effect. While bactericidal antibiotics have lethal effects on a bacteria, bacteriostatic antibiotics stall growth of the targeted bacterial population, helping the immune system as it battles with the ongoing infection (Bhattacharjee, 2016).

Antibiotics can also be classified as broad and narrow spectrum antibiotics (Bhattacharjee, 2016).

The choice of antibiotic treatment for the ongoing bacterial infection is based both on the effects of the selected antibiotic, the bacterial species causing the infection, and its antibiotic susceptibility pattern (Bhattacharjee, 2016).

Table 1.1. Main antibiotic classes/substances, their targeted cell components, MOA and effect (Bhattacharjee, 2016).

Antibiotic class		**Antibiotic**	**MOA and target**	**Effect**
I) Cell-wall synthesis inhibitors, three subgroups	I.a) Cytosolic phase of synthesis	Fosfomycin	Inhibits the enzyme catalysing the first step of cell wall synthesis (conversion of PEP to UDP-nag)	Bactericidal
	I.b) Cell wall phase of synthesis	B-lactams - Penicillin	Inhibits transpeptidase enzyme catalysing the cross-linking of peptidoglycan strands.	Bactericidal
		B-lactams - Cephalosporins	Similar to penicillin	Bactericidal
		B-lactams – Monobactams (Aztreonam)	Similar to penicillin	Bactericidal
		B-lactams - Carbapenems	Similar to penicillin	Bactericidal
	I.c) Membrane phase of synthesis	Glycopeptides – Vancomycin	Inhibits transglycosylation by blocking the substrate, not enzyme.	Bactericidal
II) Cell membrane disruptors - Antimicrobial peptides		Polymyxin	Due to its amphiphilic structure they are inserted into the membrane (inner or outer in Gram-negatives) increases its permeability thus causing leakage of cytoplasmic content.	Bactericidal
III) Metabolite synthesis inhibitors		Sulfonamides	Folic acid synthesis inhibition that causes the inhibition of DNA synthesis.	Bacteriostatic
		Trimethoprim	Bacterial DHFR inhibition that causes the inhibition of DNA synthesis.	
IV) DNA synthesis inhibitors		Quinolones and fluoroquinolones	DNA gyrase inhibition	Bactericidal
V) RNA synthesis inhibitors		Rifampicin	Inhibits the RNA elongation by binding to RNA polymerase	Bactericidal
VI) Protein synthesis inhibitors		Aminoglycosides. Streptomycin included in the group although its chemical structure differs from the members of the group.	Main MOA: Binds to 30S ribosome subunit and prevents entry of tRNA to the A site. Effects: Blocking of ribosome, misreading of genetic code, membrane damage, irreversible uptake of antibiotics	Bactericidal. (Streptomycin bacteriostatic)
		Macrolides	Binds to 23S rRNA of 50S ribosome subunit and blocks the exit of growing peptide.	Bacteriostatic (Bactericidal at higher concentrations)
		Tetracyclines	Similar to aminoglycosides	Bacteriostatic
		Chloramphenicol	Reversable binding to 50S ribosome subunit and inhibits peptidyl transferase activity.	Bacteriostatic

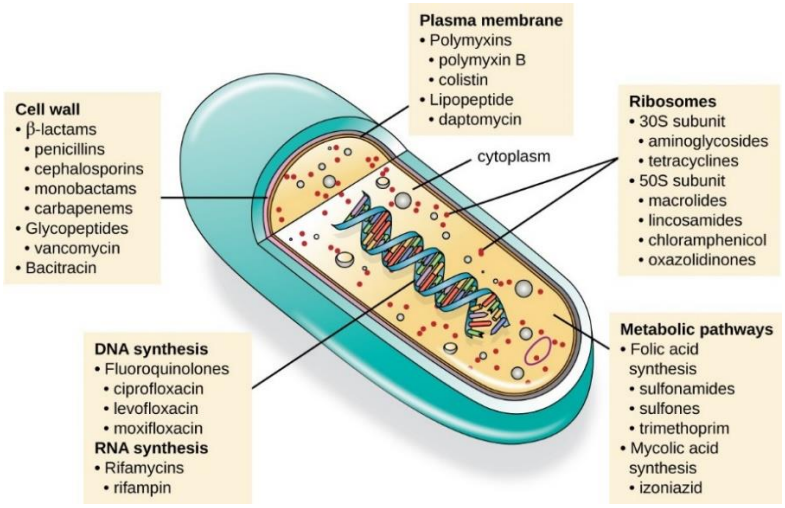


Figure 1.1. Simplified overview of different cell structures targeted by different classes of antibiotics (Lumen-lerning).

1.2. Antimicrobial resistance

AMR is not an unusual phenomenon in nature. Some fungi and bacterial species are among the best-known natural antibiotic producers. To protect themselves, these organisms also possess an appropriate ARG granting them invulnerability to the produced antibiotic. Selective pressure generated by the natural antibiotic producers promotes the survival of the mutated strains, and strains that acquired ARGs (Allen et al., 2010).

In addition to natural antibiotic producers, certain bacterial species are intrinsically resistant to some antibiotic classes. The intrinsic resistance within a bacterial species is not induced by previous exposure to antibiotic or the HGT. An example is all Gram-positive bacteria that are intrinsically resistant to aztreonam due to the poor binding of these antibiotics to its target in these bacteria (C Reygaert, 2018).

Different factors have contributed to the emergence and spread of genes coding for different AMR mechanisms. In humane medicine, antibiotics have been used both in therapy and prophylaxis, while in animals, antibiotics have been used as therapeutics, prophylactics, metaphylactics, and subtherapeutics for growth promotion (Argudin et al., 2017). AMR in human bacteria is considered to be directly caused by the inappropriate use of broad-spectrum vs. narrow-spectrum antibiotics, overuse of one type of antibiotic, improper dosage and treatment of non-bacterial infections with antibiotics. Metaphylactic approach in bacterial infection treatments in animals includes the treatment of the whole herd or flock as soon as clinical symptoms appear in a few individual members. This approach has led to frequent exposure of entire group of animals to antimicrobial agents. The frequent exposure can, in turn, cause greater occurrence and survival of mutant, antibiotic resistant strains in addition to strains with acquired ARGs. Although banned in Europe in 2006, antimicrobials in sub-therapeutic doses have been used as growth promoters in animals raised for human consumption (Argudin et al., 2017).

1.2.1. Mechanisms of AMR

Four main classes of antimicrobial resistance mechanisms are listed in Table 1.2 (Argudin et al., 2017; Bhattacharjee, 2016; C Reygaert, 2018). Due to the different structures of Gram-negative and Gram-positive bacteria, Gram-negatives can utilize all four mechanisms, although the limiting of a drug uptake and active drug efflux are less common in Gram-negatives (Argudin et al., 2017; C Reygaert, 2018).

Table 1.2. Mechanisms of AMR and their respective examples

Mechanisms of AMR	***Examples of AMR mechanisms***
I) Drug uptake limitation	Regulation of the porin channel number in the outer membrane of Gram-negatives
II) Efflux of active drugs	The action of efflux pumps decreases the concentration of cell penetrating antibiotic and prevents its accumulation in the cell.
III) Target modification	Structure alteration of PBPs in Gram-positives reduces the beta-lactam's binding affinity for these cell structures.
IV) Drug inactivation	Chemical group transfer (acetylation, phosphorylation, adenylation) to the antibiotic rendering it inactive.
	Degradation of the antibiotic by beta-lactamases (ESBL and AmpC production)

1.2.1.1. Drug degradation by β -lactamases

The β-lactam ring represents the core structure of all β-lactam antibiotics. This four-member ring contains three carbons, one with a carboxy group, and nitrogen atom at a β position. By hydrolysing the ring, causing it to open, β-lactamases inactivate the drug, preventing it to bind to its targeted cell structure (Bhattacharjee, 2016). As more than 50% of all antibiotics used in all sorts of bacterial infection treatments, belong to the β-lactam antibiotics, resistance has become a concern (C Reygaert, 2018). A wide variety of naturally occurring β-lactamases has been reported. The number of variants exceeds 2800 (Bush, 2018).

Since the discovery of β-lactamases, different classification systems have been created. Based on the amino acid sequence, Ambler-classification separates the β-lactamases into four distinct groups (A-D). Based on the active-site mechanism, β-lactamases can be further classified into two broader groups: serine β-lactamases and zinc-dependent, so called metallo-β-lactamases. According to their function, β-lactamases can be classified into three groups: 1) cephalosporinases, 2) broad-spectrum (AmpC) and extended-spectrum β-lactamases (ESBL), and 3) metallo-β-lactamases (Bhattacharjee, 2016). As this study primarily includes plasmid-borne AmpC and ESBL, these groups of β-lactamases will be discussed in more detail.

1.2.1.2. AmpC β-lactamases and ESBL

Both AmpC and ESBL are serine β-lactamases. According to Ambler-classification AmpC β-lactamases belong to group C, while ESBL belong to group A (C Reygaert, 2018). Both the structure

and the mechanism of action are similar in these two groups. The mechanism involves active serine at the active site, that, by nucleophilic attack, opens the β-lactam ring (Majiduddin et al., 2002) (Figure 1.2).

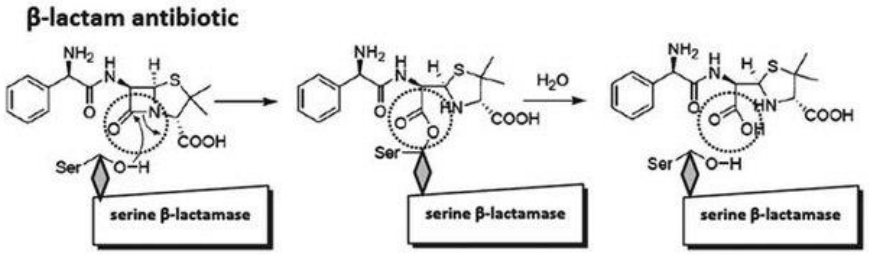


Figure 1.2. Mechanism of action of β-lactamases with Ser in their active site (Sacha et al., 2008).

ESBL producing bacteria are resistant to 3rd and 4th cephalosporin generations, and aztreonam, while cephamycins and carbapenems are not degraded by these β-lactamases. Activity of ESBLs are inhibited by β-lactamase inhibitors, such as clavulanate, sulbactam and tazobactam. The three main families of ESBLs are TEM, SHV and CTX-M types, all of these can be found both chromosomally and plasmid encoded (Seiffert et al., 2013).

AmpC β-lactamases grants the bacterium resistance to 3rd generation cephalosporins and the combination of β-lactam/β-lactamase inhibitor, although still inactive against carbapenems. Genes for AmpC can be found both on bacterial chromosomes and plasmid integrated. In recent years, plasmid encoded AmpC (pAmpC) have been frequently reported among members of *Enterobacteriaceae*, especially CMY- AmpC family (Seiffert et al., 2013).

1.3. Vertical and horizontal gene transfer

HGT is one of the main pathways of ARG dissemination between bacteria. While vertical gene transfer implies genetic material transfer from parent to offspring (in bacteria, by binary fission), HGT represents the transfer of genetic material between unrelated cells (Lawrence, 2005). Although considered threatening to humans, due to the growing spread rate of AMR, HGT is, in fact, extremely beneficial to bacteria. Without HGT, the bacterial genome diversification would be achieved at very slow rates by random point-mutations. HGT grants the bacteria instant access to new genes, thus creating new strains within the same species with special abilities (Soucy et al., 2015).

In all complex microbial communities, the evolution and individual function of each member is modified by the driving force of HGT. In these communities, the total amount of genes available to the members of one prokaryotic community could be described as a super-genome. Furthermore, the "fixed" genes found only on the prokaryotic chromosomes could be designated the private gene pool, while the communal gene pool, also called mobilome, would then encompass all the mobilizable genetic material accessible to permissive prokaryotes (Norman et al., 2009).

Genetic material can be transferred by three main HGT pathways: transformation, transduction and conjugation (Soucy et al., 2015).

1.3.1. Main pathways of HGT

Transformation is restricted to competent cells and involves the uptake of DNA segments from the environment, and its eventual incorporation into the chromosome after homologous recombination. Unlike the other two HGT pathways, transformation does not depend on the extrachromosomal mobile genetic element coding for its own transfer (Soucy et al., 2015).

Transduction relies on the transfer and incorporation of the bacteriophage during temperate phage infection (Soucy et al., 2015). Upon the phage entry into the bacterial cell, it could start with the immediate replication and production of new phage particles, or it could become inserted into the bacterial genome becoming the prophage. This dormant form is then replicated and transferred vertically, from a mother to daughter cells. When reactivated, the phage enters the lytic part of its life cycle. Once released from the bacterial genome, phage genes are replicated, transcribed and translated, leading to the production of phage particles which are ultimately packed with newly replicated phage genomes. At the end of the lytic cycle, the cell is lysed, and phage particles are released (Doss et al., 2017). During the phage genome-packing step, chromosomal fragments or complete plasmids could be mispacked into the phage particles and thus transferred to a new host (Valero-Rello et al., 2017).

Conjugation is the only mechanism of HGT that requires establishing of physical contact between two bacterial cells, a donor and a recipient of the genetic material (Soucy et al., 2015). The transfer is mediated by the hair-like appendage, pilus, that forms a bridge between the donor and the recipient. MGEs that promote their-own transfer by conjugation are integrative conjugative elements and conjugative plasmids (Partridge et al., 2018).

With the exception of the transformation, transduction and conjugation pathways are mediated by the mobile genetic elements (MGE), enabling them to cross between non-related bacterial cells (Norman et al., 2009). Furthermore, MGEs are considered to be the true drivers of HGT (Soucy et al., 2015).

1.3.2. Mobile genetic elements

MGEs could be roughly divided into two large groups, intracellular and intercellular. Examples of intracellular MGEs are insertion sequences, transposons and gene cassettes. In addition to genes that promote their allocation within a cell, intracellular MGEs are also important carriers of ARGs. On the other hand, intercellular MGEs encode the machinery for their own transfer between cells, such as conjugative plasmids and phages. The discovery of chimeric MGEs composed of two or more different

MGEs, this straight-forward classification becomes somewhat inadequate. Further, transfer machinery encoded by the intercellular MGEs is not reserved only for these elements, thus leading to the co-transfer of intracellular MGEs to a new host. Different combinations of MGE co-transfer have been reported (Norman et al., 2009; Partridge et al., 2018).

As conjugative plasmids have a major role in this study, a detailed explanation of plasmid structure and mechanisms involved in their transfer will be laid out in subsequent chapters.

1.4. Plasmids

Plasmids are extra-chromosomal, circular DNA units that replicate autonomously and are considered to be non-essential to their hosts (Lerminiaux & Cameron, 2019; Norman et al., 2009). Length varies from several kbp to more than 1Mbp (Partridge et al., 2018). Plasmids are found in both Gram-negative and Gram-positive bacteria (Partridge et al., 2018). Based on their mobility, plasmids can be roughly classified into three categories: conjugative, mobilizable and non-mobilizable plasmids (Getino & De la cruz, 2019). While conjugative plasmids possess genes that promote their own transfer, mobilizable conjugative plasmids can utilize the machinery produced by the conjugative plasmids for their own transfer. Non-mobilizable plasmids, on the other hand, neither have the necessary genes for intracellular transfer nor they use the transfer machinery of the conjugative plasmids (Getino & De la cruz, 2019).

1.4.1. Conjugative plasmids

Conjugative plasmids are longer compared to mobilizable and non-mobilizable plasmids, owing to additional genes encoding for the conjugation machinery, which increases their length to up to a couple of hundred kbp (Norman et al., 2009). Due to their length, the number of copies of conjugative plasmids per cell is usually under 10 (Thomas, 2000). Based on whether a conjugative-plasmid can be transferred into, and maintained in distantly related bacterial hosts, they can be classified as narrow and broad host range plasmids (Klümper et al., 2015).

Another interesting feature of these plasmids is their inability to share the host with another closely related plasmid. The coexistence of two related plasmids is jeopardized by their similar replication initiation system which “overloads” the plasmid copy number control, leading to a reduced number of copies of one or both plasmids, and to the possible plasmid loss during segregation (Partridge et al., 2018). Due to this fact, the conjugative plasmids are classified into several distinctive incompatibility, Inc, groups. Although previous incompatibility typing involved different laborious laboratory methods, nowadays most of the plasmids found during whole genome sequencing (WGS) are

grouped based on the sequence homology with previously sequenced and Inc-group determined plasmids (Partridge et al., 2018).

Frequently reported Inc groups found in *Enterobacteriaceae* family are A/C, F, G, HI1, HI2, I, I2, J, L/M, N, P, Q-1, Q-3, R, T, U, W, X, Y, and ColE1. Certain Inc groups are further divided into several sub-groups (Partridge et al., 2018; Rozwandowicz et al., 2018).

This study was focused on the IncI1 plasmid group, a subgroup of the IncI group. Additionally, IncI group also encompasses IncIγ, IncB, IncO, IncK and IncZ (Partridge et al., 2018). In general, IncI plasmids are conjugative, low-copy number plasmids with a narrow-host range found exclusively in the family *Enterobacteriaceae*, with length varying between 50 and 250kbp. In addition, they are highly stable and are found to maintain ARGs in their host without any external selective pressure (Carattoli et al., 2018; Partridge et al., 2018; Zhang et al., 2019).

1.4.2. IncI1 plasmid backbone

Several conserved regions can be found as a backbone in any IncI1 plasmid (Carattoli et al., 2018). The pR64, an IncI1 plasmid originally isolated from *Salmonella enterica* serovar *Typhimirium* (accession number AP005147) in 1966 (Carattoli et al., 2018), is considered the prototype for the IncI1 plasmid group containing all conserved regions of these plasmids (Figure 1.3).

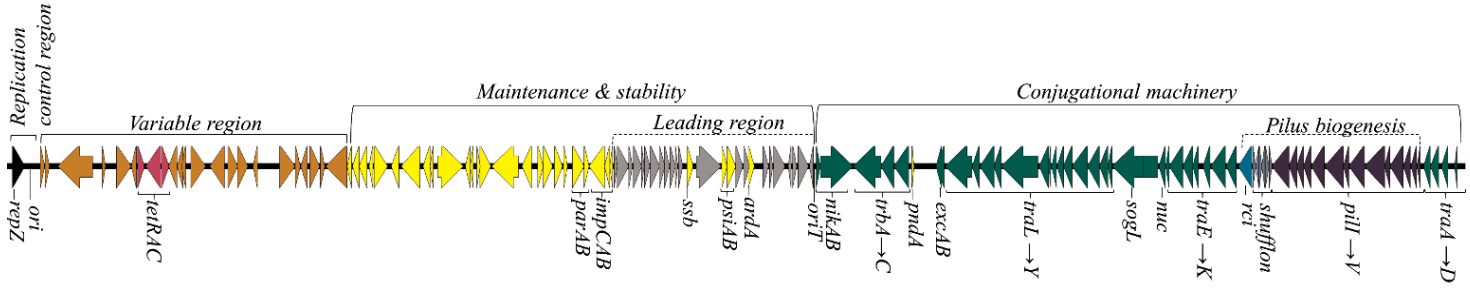


Figure 1.3. Schematic representation of the R64 plasmid. Coloured arrows indicate genes and their respective transcription direction. Names of the most important genes are given below the gene arrows. Colours denote the following: black- plasmid replicon type, orange- insertion sequences, purple- AMR genes, yellow- genes involved in maintenance and stability, grey-genes part of the first segment to enter the cell, green- genes encoding for the products involved in plasmid transfer, light blue- shufflon region, dark-purple- genes involved in pilus biogenesis.

IncI1 plasmids have the *repZ* replicon (Figure 1.3). *RepZ* gene is in close proximity of the origin of plasmid replication (*ori* site) and its control elements. The *repZ* encodes the plasmid replicase. Its expression is strictly regulated. The organization and nucleotide sequence of replication region are conserved in all IncI1 plasmids. (Carattoli et al., 2018; Thomas, 2000).

The replication region is followed by the variable region, also called the accessory module (Carattoli et al., 2018; Norman et al., 2009). Depending on the plasmid, the content of this region can differ, including different insertion sequences, transposons, integrons and gene cassettes. Due to these inserted elements, plasmids are carriers of variable accessory genes, thus providing the host with a new set of "special" abilities (Carattoli et al., 2018). These "unusual" traits include AMR, virulence, heavy

metal tolerance, and the catabolism of unusual nutrients (Carroll & Wong, 2018). pR64 contains tetracycline resistance genes in this region (*tetR*, *tetA* and *tetC*).

The maintenance and stability region is common for all low copy number plasmids. This region is responsible for both the plasmid distribution to daughter cells, and the protection of the plasmid after its successful transfer to a recipient cell (Carattoli et al., 2018). One of the most important genes found in the maintenance and stability region are the *parA* and *parB* genes, which encodes the components of the active partitioning mechanism (Kaur et al., 2011). These are directly responsible for equal plasmid distribution to daughter cells (Kaur et al., 2011). Although not found in this region, *pndAC* genes by their functionality belong in this group (Partridge et al., 2018). PndAC are involved in post segregational killing, a back-up mechanism to active partitioning ensures the survival of the offspring that have successfully received the plasmid (Kroll et al., 2010).

A subpart of the maintenance and stability region, called leading region, represents the first DNA segment of the plasmid that enters the recipient cell during the conjugation process (Carattoli et al., 2018). The region is found between *impCAB* operon and origin of transfer (*oriT*). This region is also conserved and contains genes encoding for factors that counteracts the defence response of the recipient cell upon entry of the single stranded plasmid (Carattoli et al., 2018). Single stranded proteins (SSBs) are considered to be non-essential for conjugative transfer, although are assumed to be involved in plasmid stability after the successful transfer to a new recipient cell (Jain et al., 2012). Another important factor encoded by this region is ardA, anti-restriction protein. ArdA mitigate the activity of the recipient encoded endonuclease EcoK1 (Carattoli et al., 2018). Protection is also provided by psiAB proteins which inhibit the recipient SOS response (Carattoli et al., 2018).

A complete conjugational machinery is encoded by a 54kbp region, composed of several functionally clustered genes (Carattoli et al., 2018). Products of *nikAB* genes, found next to the *oriT* site, are directly responsible for plasmid transfer initiation (Carattoli et al., 2018). Products of the *traA-D* and *trbA-C* are considered essential for plasmid transfer (Carattoli et al., 2018). IncI1 plasmids also have a controlled selection of possible recipient cells. This is accomplished by the *excAB* and *traY* genes whose products, when combined, make up the surface-exclusion system, preventing the entry of another IncI1 plasmid (Carattoli et al., 2018; Partridge et al., 2018). One of the largest clusters found in this region is a type IV thin pilus encoding cluster. The cluster contains 14 gens (*pilJ-V*). At the end of the *pilV* gene is a shufflon region that, together with the site-specific DNA recombinase (Rci), introduce variability in the 3' end of the *pilV* gene (Carattoli et al., 2018).

1.4.3. Plasmid mechanisms that promote plasmid stability in bacterial populations

Conjugative plasmids can be seen as selfish genetic elements due to their ability to efficiently maintain constant number of copies in a cell, to efficiently ensure their segregation during cell division, and a conjugation mechanism that enables them to be stably preserved in a bacterial community (Liu et

al., 2015; San Millan et al., 2014). While chromosomal genes act together in order to improve the fitness of the organism, selfish plasmids have their own control over the expression of their genes. Due to their selfishness, their presence can affect the fitness of their hosts, in both positive and negative way (San Millan et al., 2014). Furthermore, accessory genes found on the plasmid that provide the host with special abilities can be transferred from the plasmid to the chromosome, making the plasmid useless to the host. Therefore, some authors agree on defining the plasmid as parasitic DNA which uses different mechanisms to ensure its survival (San Millan et al., 2014).

1.4.3.1. Replication control

Replication control serves the purpose of keeping the number of plasmids constant in a cell (Carattoli et al., 2018). This mechanism is the first in line that ensures plasmid stability in one host. Maintaining their number under 10 per cell eases the maintenance burden an increase of total amount of DNA would pose on the host (San Millan et al., 2014).

IncI1 plasmids control the replication by controlling the expression of the replication protein RepZ (Carattoli et al., 2018). Expression is controlled by the positive and negative regulator factors, *Inc* and *repY*, at transcriptional level. *RepY* precedes the *repZ* and needs to be translated in order for repZ to be translated. *RepY* acts as a positive regulator of the *repZ* expression. *Inc* RNA blocs the translation of *repY* thereby hindering the translation of *repZ*. When replication is activated, it still depends on the host's replication machinery. Further interactions and compatibility of plasmid encoded replication initiation factors and host's replication machinery are crucial for the plasmid's survival in the host. These interactions are among the factors that dictate the compatibility of the plasmid and the host and further maintenance of the plasmid within the host (Asano et al., 1999).

1.4.3.2. Active partitioning and post-segregational killing

While high copy number plasmids rely on chance to be distributed to both daughter cells during cell division, low-copy number plasmids possess a mechanism ensuring equal plasmid segregation to daughter cells (Münch et al., 2019). The mechanism is called active partitioning (Carattoli et al., 2018) (Figure 1.4).

Partitioning involves relocation of plasmids towards the poles of the dividing cell. In IncI1 plasmids, the system is comprised of *parA-parB* operon and downstream cis-acting site *pars*. *Pars* is the site of the assembly of the partitioning mechanism and is considered a prokaryotic centromere. ParB recognizes the *pars*, while ParA, ATPase protein, binds to ParB and hydrolyses the ATP as it moves the plasmid towards the pole of the cell (Kaur et al., 2011; Tolmasky & Alonso, 2015).

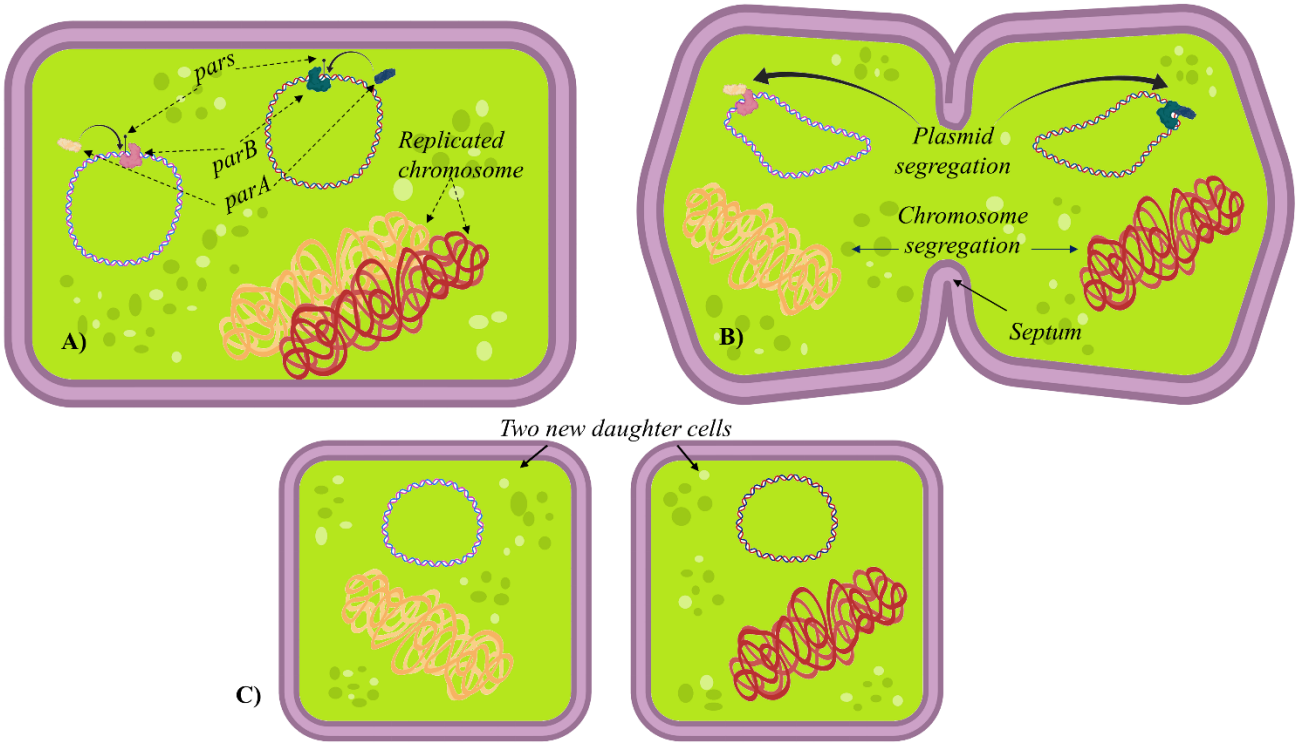


Figure 1.4. Active partitioning mechanism. A) Initial phase of cell division and assembly of parAB on pars. B) Formation of the septum, chromosome copies and plasmid copies relocate towards the poles of the dividing cell. C) Two newly formed daughter cells each with a copy of the plasmid.(Illustration: M. Stosic)

If active partitioning fails to ensure plasmid segregation during cell division, a back-up mechanism is activated (Carattoli et al., 2018). The post-segregational killing system is composed of a very stable mRNA translated into a cell toxin that damages the cell membrane from within, and a small labile antisense RNA that prevents the translation of the toxin (Nielsen & Gerdes, 1995). Both toxin mRNA and antitoxin (antisense RNA) are inherited from the mother cell. With no plasmid present, rapid decay of the inherited anti-sense RNA and no production of the new ones, leads to the translation of the toxin mRNA. The produced toxin kills the plasmid-cured offspring.

While both active partitioning and post-segregational killing ensure the plasmid preservation during the clonal expansion, conjugation provides new hosts for the plasmids.

1.4.4. Conjugation mechanism

Conjugation is much more complex than the simple preservation of the plasmid during cell division and consists of several phases (Figure 1.5). Conjugation begins with gene expression of the type IVB secretion system (T4SS B), *tra/trb* genes (Voth et al., 2012). The T4SS complex consists of several functionally distinctive parts: the cytoplasmic ATPase, the inner membrane platform, the core channel and the pilus (Getino & De la cruz, 2019).

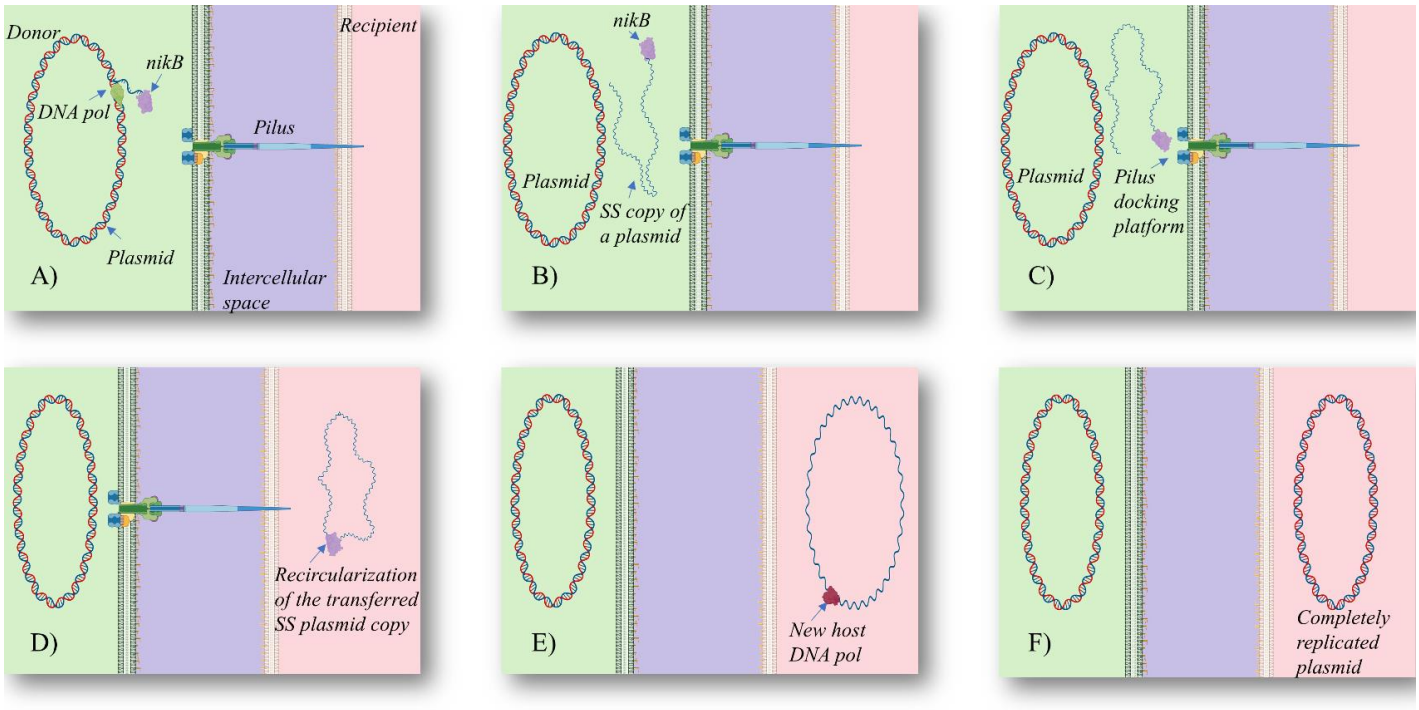


Figure 1.5. Simplified representation of the conjugation phases. A) nikB-single stranded complex was created while replication is replacing the unwinding strand. B) Single stranded copy of the plasmid released from its original. C) nikB-ss plasmid complex interaction with a pilus docking protein preparing it for the transfer. D) nikB reverse nicking process recircularizes the ss plasmid copy. E) New host assembly of the replication machinery that will regenerate the missing strand. F) Completion of the conjugation, plasmid is now present in both donor and recipient cell. (Illustration: M. Stosic)

The pilus is a hair-like appendage that extends from the donor to the recipient cell and is responsible for securing the physical connection between the two cells (Getino & De la cruz, 2019). Depending on the medium conjugation takes place in, the pilus encoded by the IncI1 plasmid can be both thin and thick (Partridge et al., 2018). While a thick pilus is necessary to stabilize the conjugation on the solid surfaces, the thin pilus enables the conjugation that takes place in the liquid medium. Thin pilus is encoded by 14 genes, *pilI* through *pilV* (Figure 1.3), and 12 of them are crucial for pilus biogenesis (Dudley et al., 2006).

After the contact between the donor and the recipient cell has been established, the next step prepares the plasmid DNA for intracellular transport (Getino & De la cruz, 2019). The function of the relaxase encoded by the *nikB* is nicking the plasmid DNA at *nic* site, found within the *oriT* site. The 5' strand of the nicked plasmid becomes covalently attached to the relaxase, thereby creating a nucleoprotein complex called relaxosome (Getino & De la cruz, 2019; Partridge et al., 2018). The relaxase stays attached to the single stranded plasmid until the transfer is complete. Next, plasmid replication begins at the 3' end of the nicked strand by using the complementary still circularized strand as a template by simultaneously unwinding the strand attached to the relaxase. Upon completion of the replication, the relaxase introduces a second cleavage at the *nic* site releasing the complete single stranded plasmid. The relaxosome is further transported to the inner membrane platform of T4SS where transport is initiated (Getino & De la cruz, 2019).

By using the ATPase provided energy, the relaxosome is then "fired" through the core channel of T4SS and pilus. Upon the relaxosome entry, relaxase recognizes again the *nic* site at the 3' end, and by the reverse nicking reaction recircularizes the plasmid, while rolling-circle replication regenerates the double stranded plasmid formation (Getino & De la cruz, 2019).

1.4.4.1. Rearrangement of the pilV shufflon region

PilV is found on the tip of the thin conjugative pilus and is an adhesin-type protein that interacts with the recipient's cell surface by recognizing specific lipopolysaccharide structures (LPS) (Carattoli et al., 2018). PilV is considered an important factor in selecting a recipient cell (Gyohda et al., 2006). The deletion of the *pilV* leads to a significant reduction of conjugation rates of IncI1 in a liquid medium. A special feature found as a part of *pilV* is a shufflon, a multiple inversion system that introduces the variability into the PilV C-terminal domain (Brouwer et al., 2019).

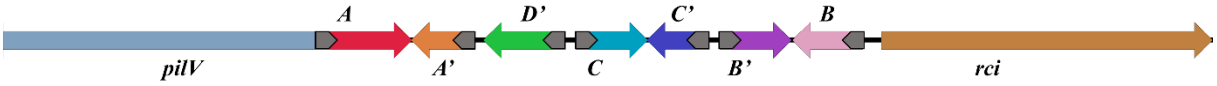


Figure 1.6. Structure of the pR64 shufflon, coloured arrows (red, orange, green, light blue, dark blue, purple and pink) represent partial ORFs of the pilV. Grey arrows within each segment represent sfx repeats (illustration: M. Stosic)

Due to the shufflon rearrangements, PilV of the pR64 can have seven different C-terminal domains (Figure 1.6), and it is assumed that seven different LPS structures can be recognized by the pilV adhesin during conjugation (Brouwer et al., 2019). The pR64 shufflon is composed of four segments which, with the exception of the D' segment, contain two partial open reading frames encoding for two different C-terminal ends of the PilV (Brouwer et al., 2015). Segment A therefore contains *A* and *A'* ORFs, segment B contains *B* and *B'* ORFs, and C contains *C* and *C'* ORFs. Segments A, B and C are flanked by so-called s*fx* repeats composed of a highly conserved 19bp repeat and 12 bp of non-conserved sequence. The core site (7bp) found at the end of the 19 bp conserved region is a conserved spacer, a site where DNA crossover occurs. Segment D has only one s*fx* repeat and is flanked by the direct repeats (Figure 1.6) (Brouwer et al., 2015; Gyohda et al., 2006).

Rearrangement of segments is mediated by the site-specific recombinase of the tyrosine family Rci. Rci recognizes any pair of inverted s*fx* repeats and inverts the fragment found between them. Binding of the fragment flanked by the direct s*fx* repeats would lead to an excision of the fragment. While the shufflon of pR64 have four segments and seven partial ORFs, the number of segments varies between shufflons of different IncI1 plasmids (Brouwer et al., 2015).

Due to the long-range sequencing being able to generate reads that cover the whole shufflon region, the rearrangements of the shufflon can be studied more extensively. So far, only one study analysed the rearrangement of IncI1 shufflons by employing the long-range sequencing. The results imply that rearrangement is constant, but not random. The *pilV-A* and *pilV-A' ORFs* were

overrepresented in the investigated plasmids. Furthermore, according to their results, the rearrangement of the shufflon was not dependent on the growth stage of the plasmid host, or on the stress factors imposed on the growing bacterial culture (Brouwer et al., 2019).

1.4.4.2. Mechanisms activated in the recipient cell upon plasmid entry

The recipient cell is not passive during the invasion of the plasmid. Different mechanisms activate in order to prevent the survival of the new DNA. The restriction modification system (RM) and the CRISPR-Cas system detect foreign DNA and prevents its stable acquisition (Getino & De la cruz, 2019).

The RM system, a primitive two-component immunity system in bacteria, recognizes and cleaves the foreign DNA. A method this system uses to distinguish itself from non-self DNA is methylation. Methyltransferase methylates a specific adenine or cytosine at specific sites, while restriction endonuclease recognizes and cleaves unmethylated foreign DNA. Parasitic DNA molecules have developed different strategies to escape the cleavage by RM. IncI1 plasmids encodes ArdA (antirestriction) proteins that alleviate the activity of REase (Getino & De la cruz, 2019).

In addition to the RM system, certain bacterial species also possess an adaptive immunity system, CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR associated protein). In the so-called immunization phase, the sequences of the invading genomes are collected and integrated into the CRISPR array. In the immunity phase, the actual battle with the incoming foreign DNA, the CRISPR array is transcribed and processed generating crRNAs. Upon entry of foreign DNA whose sequence has been previously integrated into the CRISPR array, the produced crRNA interacts with the foreign DNA by complementary attachment, guiding the Cas nuclease that would further cleave the invading DNA (Getino & De la cruz, 2019).

Prior to activation of RM and CRISPR-Cas, the first response of the recipient cell on the plasmid entry is the SOS system response. This system is activated due to the entry of large amounts of the single stranded DNA. SOS system halts cell division and promotes the production of mutagenic DNA polymerase V. If the SOS system is not alleviated, the response could be potentially harmful for the cell. To prevent this, IncI1 plasmids encode PsiAB proteins that inhibit the full response of the SOS system preventing the infliction of possible damage to the host (Petrova et al., 2009).

1.4.5. Plasmid fitness cost

While beneficial under the conditions requiring certain plasmid genes, plasmids impose a fitness cost to their hosts during intervals when these genes are unnecessary. Still, plasmids are very stable. The persistence of plasmids in their hosts even in the absence of the selective pressure is called plasmid

paradox. Under these conditions, plasmids behave completely as parasites (Carroll & Wong, 2018). Using the energy of the host for their own mechanisms, causes the host to struggle to grow at a normal rate, as its competitiveness becomes weakened (San Millan, 2018).

Different cell mechanisms can be affected by the plasmid. Upon the plasmid's introduction into the new host, cell division stalls due to the transient activation of the SOS system (Carattoli et al., 2018). Stopping the cell division directly affects the cell growth rate (San Millan, 2018). Furthermore, energy costly plasmid processes are not under any regulation immediately upon plasmid entry, leading to so-called overshooting of the plasmid genes, draining the energy resources of the host. The expression of the conjugation machinery is considered the most expensive and it must be strictly controlled (San Millan, 2018).

After reestablishing plasmid gene regulation, the plasmid' replication also imposes an energy cost to the cell. Due to the low number of copies of the IncI1 plasmids in the cell, maintenance of an additional amount of DNA should not represent a considerable burden to the cell. However, mobilization of the DNA replication machinery provided by the host could impact the chromosomal replication and further reduce the host's growth rate (San Millan, 2018).

Plasmid gene expression also affects the normal functioning of the cell, although transcription is considered not to impose too much burden to the cell. On the other hand, translation and difference in codon usage between host genes and plasmid genes, can lead to a depletion of certain tRNAs from the host tRNA pool. This depletion can in turn, also affect translation of the host genes, eventually leading to reduced growth rate. Finally, any plasmid produced protein could potentially alter the host protein network. An example of the plasmid encoded protein that makes multiple interactions with host-encoded proteins is the Rep replication initiation protein (San Millan, 2018).

Fitness cost induced by a plasmid was found to be dependent on the genetic background of the host. The same plasmid could induce variable effects in hosts with different genetic backgrounds, suggesting that the plasmid fitness cost is host genotype dependent (Carroll & Wong, 2018).

Furthermore, a bacterial cell does not host one MGE. Different MGEs can reside in one cell. Their interaction with each other, and interaction with the host also affects normal functioning of the host (Carroll & Wong, 2018). A study reveals that simultaneous presence of different plasmids in a host imposed lower fitness cost, then when only one of the plasmids was present (Carroll & Wong, 2018).

In order to reduce the impact of the plasmid on its own fitness, the host cell needs to adapt. So-called compensatory mutations have been reported to turn the fitness cost inflicted by the plasmid into a benefit. It was reported that chromosomal mutations found in accessory helicase and the RNA polymerase subunit β in *Pseudomonas* sp.H2 were key mutations that alleviated the burden imposed by the plasmid (Loftie-Eaton et al., 2017). Another study found that compensatory mutation was often found in GacA/GacS regulatory system in the conjugative plasmid carrying host (Harrison et al., 2016). Furthermore, compensatory mutations could also be found on the plasmids, however they were found more frequently on the chromosome. Although not thoroughly investigated, it is hypothesized that the

compensatory mutations could be potentially deleterious for the host if the plasmid is lost (Carroll & Wong, 2018).

While a number of studies report that the plasmid fitness cost was minimal or not present, other studies report opposite results. However, it is assumed that these differences are due to a different experimental setup, bacterial strains, and plasmids used in the individual experiments. A group of experiments was conducted with model bacterial laboratory strains and non-clinically relevant plasmids. The other encompassed wild type strains with clinically relevant plasmids (Benz et al., 2019; Carroll & Wong, 2018).

This inconsistency in the results implies that greater understanding of the complex interactions between the conjugative plasmid and the host is needed (Benz et al., 2019).

1.5. Methods to study plasmid fitness cost and plasmid stability

1.5.1. Single strain bacterial growth

The growth of a bacterial culture can be divided into several stages. During the initial phase, called lag phase, growth of the culture is not observable. The exponential phase is characterised by constant cell division rate. In the deceleration phase, the cell division rate ceases, while in the stationary phase the culture have consumed all available space and/or nutrients and ceased growing (Rockwood, 2015). In the death phase, cell death rate exceeds the cell division rate. Finally culture reaches the long stationary phase which could be extend for years (Hall et al., 2014; Rolfe et al., 2012).

Bacterial growth can be monitored by using different methods. While certain methods are based on monitoring only the increase of the number of living cells, other methods monitor the increase in the biomass (McBirney et al., 2016). The biomass could be described as the total amount of both living cells, dead cells, cell debris and cell products per volume of the cell culture (Schinner et al., 2012). While determining the count of living cells requires laborious laboratory work, this method yields the greatest accuracy when describing bacterial growth (McBirney et al., 2016).

On the other hand, the increase of the bacterial biomass can be represented by the change of the optical density (OD) of a culture over time. An OD measurement at a particular wavelength measures the amount of light lost due to scattering and/or absorption. In bacterial growth analysis, OD measurements with the wavelength of 600nm are frequently used. However, the bacterial culture encompasses not only the cells and cell debris, but also the medium cells are grown in, as well as bacterial products and by-products excreted into the medium. These additional factors increase the inaccuracy of the OD value. Although measurements could be corrected by performing an OD measurement on the pure medium, it is not possible to correct the values with OD measurements of the bacterial products. In addition, bacterial cells can form small clusters by attaching to each other, which

also affects the OD measurement. Finally, due to the actual size of the bacteria, smaller cells are considered “poor-scattereres” of light at 600nm wavelength (McBirney et al., 2016).

Both counting of the living cells and OD measurements are widely used to monitor the growth of bacterial cultures, although it is necessary to be familiar with their advantages and disadvantages in order to correctly interpret the results.

Regardless of the method, bacterial growth can be represented with both an exponential and a logistic model. Both models start with a differential equation (Rockwood, 2015):

$$\frac{dN}{dt} = rN \qquad \textit{(Equation 1.1)}$$

wherein the change in bacterial population N over time interval t, is equal to the increase of the initial population size by the growth rate r, r representing the difference between the birth and death rates (Rockwood, 2015).

The exponential model is somewhat simpler and assumes that the bacterial growth is limitless and unconstrained by space and available nutrients, while the logistic model assumes that the population size of any species is constrained by available resources and/or space (Figure 1.7) (Rockwood, 2015). The maximum of the population size is represented as the carrying capacity (K).

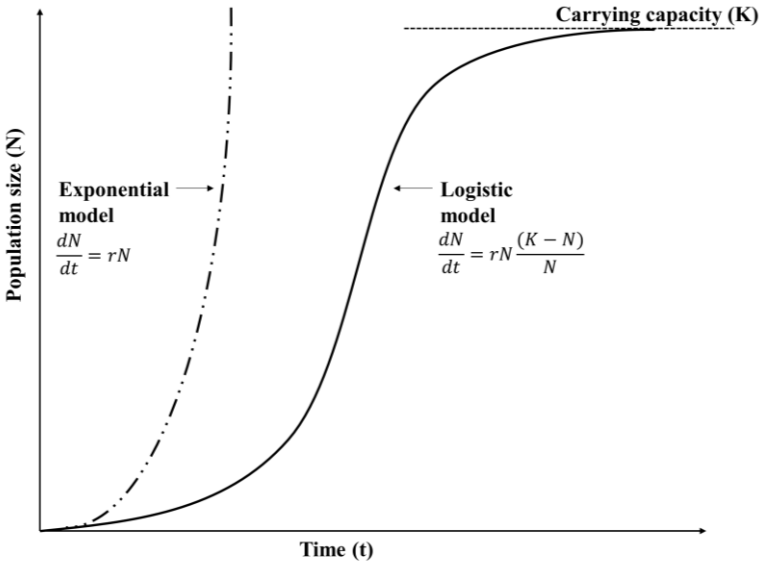


Figure 1.7. Growth curves represented by exponential and logistic model and their respective differential equations (Illustration: M. Stosic).

1.5.2. Competitive growth

Competition could be defined as a biological interaction between two or more individuals for a resource in short supply (Rockwood, 2015). As in terms of bacterial competition, the limiting resources would be the space and nutrients accessible to the competing strains. The assumption is that the observed growth differences in a single strain growth would be more pronounced during the pairwise growth

(Lenski et al., 1994; Ram et al., 2019). The bacterial strain with the lower growth rate would be outcompeted by the strain with the higher growth rate, owing to the fact that it would reach higher numbers and thereby occupy larger space and have better access to the nutrients (Lenski et al., 1994; Rockwood, 2015). The bacterial strain would have an increased competitive fitness. The population size of the strain with the lower growth rate would be reduced due to the reduced access to nutrients forcing it to terminate its growth and enter the premature stationary phase. This bacterial strain would then have a reduced competitive fitness (Lenski et al., 1994).

A competitive growth assay could be modified to monitor the plasmid stability. The competitors are the plasmid free strain and its transconjugant mixed in different initial ratios without any selective pressure that would favour the growth of the plasmid carrying cells. The emphasis is not on the individual competitors and whether their competitive fitness changes, but rather on the change in the number of plasmids in the whole culture during the competitive growth. The number of conjugative plasmids in a culture is affected by three independent mechanisms, clonal expansion of their hosts, plasmid replication within each host cell and conjugation. When the transconjugants are competing with its plasmid free counterpart in much greater numbers, the assumption is that the increase of the number of plasmids would be most affected by the conjugative transfer rate (Hagbø et al., 2019).

1.6. Methods to study shufflon rearrangements – Nanopore sequencing

Until the commercialised utilization of long-read technologies such as Oxford Nanopore Technologies MinION, studying structural variation of the shufflon region was hampered (Brouwer et al., 2019). It is assumed that the length of the MinIon produced reads is limited only by the strand length (Jain et al., 2016). Long reads can cover the whole region of the shufflon and, thus, enable the analysis of structural variations in this region (Brouwer et al., 2019). The remaining problem is the high error rate in these reads that could reach up to 30% in practice for Oxford Nanopore (Morisse et al., 2018).

The preparation for sequencing and the sequencing process involves several steps. The adapter ligation to both ends of DNA or cDNA fragments ensures concentrating the DNA substrate at the surface of the flow cell in a close proximity to the nanopores, while hairpin adapters enables the continuous sequencing of both strands of the dsDNA. A motor enzyme captures the 5'- end of one strand and performs single—nucleotide displacement along the strand through the nanopore. A flow cell is composed of up to 2048 individual protein nanopores. Shifting of the passing nucleotides through the nanopore creates disruption of the ionic current detected by the sensor and computationally interpreted as a sequence of 3-6 nucleotide long kmers. Information from the template and complement strand reads is combined giving raise to high-quality 2D reads (Jain et al., 2016).

However, the ligation step does not need to include the ligation of the hairpin adapter. In this case, the nanopore reads only one strand allowing for the higher throughput from a flow cell, although the accuracy for 1D reads produced in this manner are lower (Jain et al., 2016).

1.7. Aims and hypothesis of the thesis

This study was conducted:

- to examine whether selected IncI1 plasmids with ESBL/AmpC encoding genes inflict a fitness cost in different bacterial hosts;
- to assess the stability of the selected IncI1 plasmids during a competitive growth between a plasmid-containing strain and its plasmid-free counterpart without antibiotic selective pressure;
- to examine the structural variation of shufflons from selected plasmids when grown in different hosts without antibiotic selective pressure.

As plasmid could interfere with the normal growth of it host, it was expected that the growth rate and/or carrying capacity of the transconjugants would be significantly lower compared to the same parameters of their respective plasmid-free recipient strains during the single strain growth. In addition, it was further hypothesized that differences between the selected pairs of the transconjugant and its respective plasmid-free recipient strain would be more pronounced in the competitive pairwise growth assay than in the single strain growth assay.

The plasmid stability experiment was the repetition of a previous study (Hagbø et al., 2019). It was expected that the plasmid number would increase gradually during the competitive growth by the means of both conjugation and clonal expansion of plasmid-containing cells even when their initial CFU/ml of was 1% of the mating mixture with plasmid-free counterparts.

And finally, the structural variation of shufflons was studied during different bacterial growth phases, and in different plasmid hosts without any selective pressure, similarly as it has been conducted in the previous study (Brouwer et al., 2019). However, as this study encompassed shufflons interrupted with the insertion sequence, it was expected that the rearrangement of this shufflon type would be hindered compared to the reference control plasmid with the uninterrupted sequence. As the number of cells increases during bacterial growth, an assumption was made that the number of variants would increase with every consecutive sampling point. In addition, due to mobility of the given insertion sequence, it was also investigated whether the shufflon variants without the insertion sequence would be generated.

2. Materials and methods

All the experiments presented here were performed either at the Norwegian Veterinary Institute (NVI) in Oslo or at the Norwegian Life and Science University (NMBU) in Ås. All plasmids used in this study are conjugative IncI1 plasmids with length ranging from 89kb to 168kb. Schematic overview of experiments conducted in this study are shown in Figure 2.1.

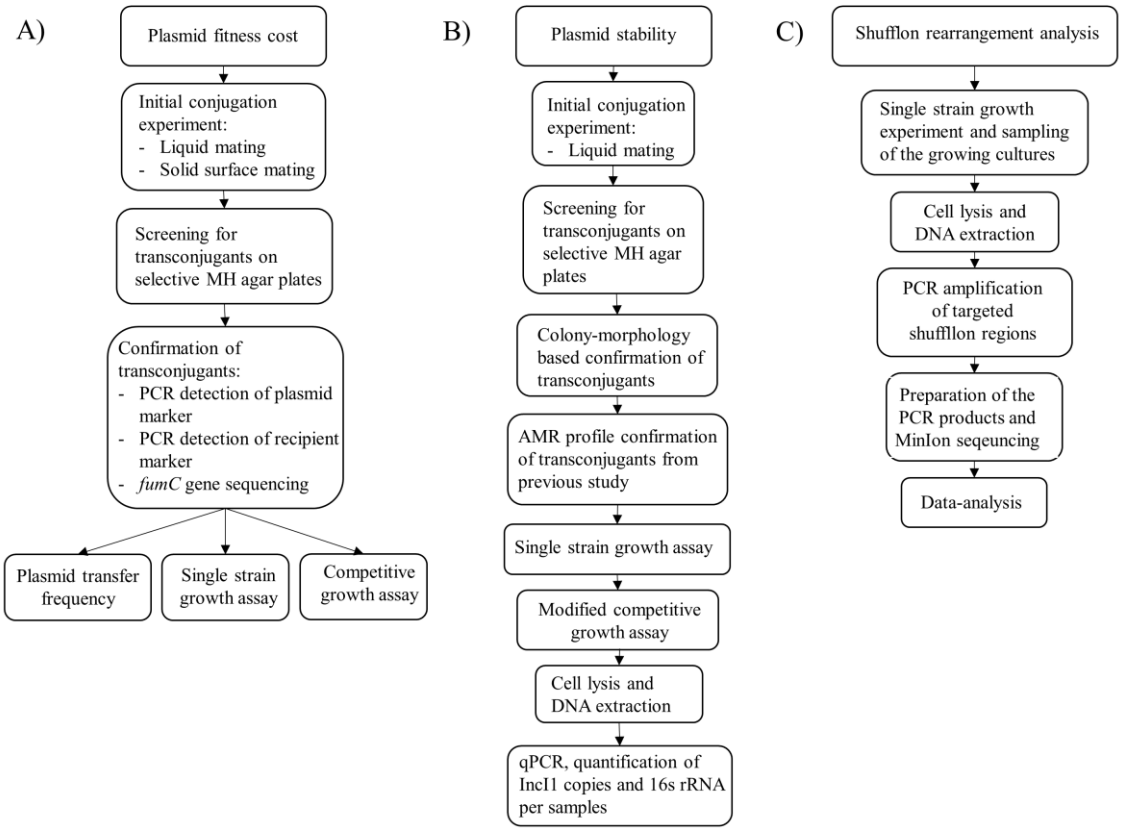


Figure 2.1. Flowcharts of the experimental set-up. Flowchart A – Plasmid fitness cost, flowchart B – plasmid stability, flowchart C – Shufflon rearrangements analysis.

2.1. Bacterial strains

All bacterial strains included in this study, both plasmid hosts and plasmid-recipient strains, are listed in Table 2.1 and 2.2, while their antimicrobial resistance profiles are shown in Table A1-A5 (Supplementary materials, part A). All poultry-originating strains (both *E. coli* and *Klebsiella pneumoniae*) were collected in previous studies and /or surveillance programs performed by the NVI in the period 2006-2018. *K. pneumoniae* isolated from human and pig faecal samples was collected as a part of the REDUCEAMU project (courtesy of Thongpan Leangapichart, post.doc. at NVI).

E. coli of human origin, original host for the three DH5α rif^{R} transconjugants (LII-22, LII-30, LII-55), was extracted from stool samples of premature twins, in Spain, in 2015. These transconjugants further differ in terms of the additional MGEs listed in Table A6 (Supplementary materials, part A).

Table 2.1. Plasmid-carrying strains, their origin, resistance phenotype, Sequence Type and phylotype, sequence type of their respective plasmid and ESBL/AmpC gene found on the plasmid.

Plasmid-host strains	***Host origin***	***Sequence type***	***Host phylotype***	***Plasmid replicon type***	***Plasmid sequence type***	***Plasmid length***	***Plasmid marker***	***ESBL/AmpC gene***	***Reference***
2016-40-17437	Poultry	ST57	D	*IncI1/IncFIB*	pMLST 3	*168kb*	Ctx^{R}	*bla*$_{CTX-M-1}$	*(Mo et al., 2020)*
2016-40-20481	Poultry	ST57	D	*IncI1*	pMLST 42	*102kb*	Ctx^{R}	*bla*$_{CTX-M-1}$	*(Mo et al., 2020)*
2016-40-22638	Poultry	ST1638	A	*IncI1*	pMLST 3	*118kb*	Ctx^{R}	*bla*$_{CTX-M-1}$	*(Mo et al., 2020)*
2012-01-2798	Poultry	ST3249	A	*IncI1*	pMLST 12	*94kb*	Ctx^{R}	*bla*$_{CMY-2}$	*Mo et al. (2016)*
2006-01-1248	Poultry	ND	A	*IncI1*	pMLST 5	*89kb*	Ctx^{R}	*bla*$_{TEM-20}$	*Sunde et al. (2009)*
2016-40-20426	Poultry	ST641	D	*IncI1*	pMLST 3	*102kb*	Ctx^{R}	*bla*$_{CTX-M-1}$	*(Mo et al., 2020)*
DH5α rif^{R} pLII-22*	Original host: *E. coli*, human origin	NR	NR	*IncI1*	pMLST 3	*98kb*	NR	*bla*$_{TEM-1B}$, *bla*$_{SHV-12}$**	*(Hagbø et al., 2019)*
DH5α rif^{R} pLII-30*		NR	NR				*NR*	*bla*$_{TEM-1B}$, *bla*$_{SHV-12}$**	*(Hagbø et al., 2019)*
DH5α rif^{R} pLII-55*		NR	NR				*NR*	*bla*$_{TEM-1B}$, *bla*$_{SHV-12}$**	*(Hagbø et al., 2019)*

Abbreviations: NR-not relevant; ND-not determined

*DHA5α rif^{R} with *pLII-22, pLII-30 and pLII-55 are transconjugants of one E. coli isolate. Transconjugants contain different MGE in addition to the same plasmid transported during the initial mating with the original wild type E. coli host.*

***Due to the plasmid sequence assembly performed only by Illumina short pair-end reads, it is unresolved whether the bla-genes are part of the plasmid or other MGE.*

Table 2.2 Plasmid-free strains, their origin, resistance phenotype, sequence type and phylotype.

Plasmid-recipient strains	***Origin***	***Recipient marker***	***Recipient Sequence type***	***Recipient Phylotype***
2011-01-1173 (APEC)	Poultry	Nal^{R}	Unknown	D
2014-01-2070 (QREC)	Poultry	Nal^{R}	ST355	B2
2014-01-4539 (QREC)	Poultry	Nal^{R}	ST355	B2
2014 -01-6043-1s (QREC)	Poultry	Nal^{R}	ST355	B2
2014-01-2145 (QREC)	Poultry	Nal^{R}	ST162	B1
2014-01-2773 (QREC)	Poultry	Nal^{R}	ST162	B1
2014-01-7133 (QREC)	Poultry	Nal^{R}	ST162	B1
2009-01-3815 (QREC)	Poultry	Nal^{R}	ST602	B1
2009-01-4618-2 (QREC)	Poultry	Nal^{R}	ST602	B1
2011-01-3460-5 (QREC)	Poultry	Nal^{R}	ST602	B1
2014-01-2069 (QREC)	Poultry	Nal^{R}	ST453	B1
2014-01-6924 (QREC)	Poultry	Nal^{R}	ST453	B1
2014-01-7234-1 (QREC)	Poultry	Nal^{R}	ST453	B1
Klebisella pneumoniae 2018-01-715	Poultry	Nal^{R}	NR	NA
Klebisella pneumoniae 152 CK	Human	NR	NR	NA
Klebisella pneumoniae 27PK	Pig	NR	NR	NA
E. coli DH5α rif^{R}	laboratory strain	Rif^{R}	NR	ND

Abbreviations: NR-not relevant; ND-not determined; NA-not applicable; APEC – avian pathogenic E. coli; QUREC- quinolone resistant E. coli; Nal^{R} – nalidixic acid resistance; Rif^{R}- rifampicin resistance.

2.2. Conjugation experiment

The conjugation experiment involved mating a plasmid-host strain (donor) with a recipient strain in liquid and/or on solid-surface non-selective growth medium, and screening for the transconjugants on the selective agar medium.

2.2.1. Liquid mating

Liquid mating was performed as previously described (Mo et al., 2017) when the donor/recipient were wild-type *E. coli* strains. In short, pure cultures of a donor and a recipient were grown separately in Luria-Bertani (LB) broth overnight (ON) at 37ºC. The next day, 1ml of donor and 1 ml of recipient preculture (ON culture) were mixed in 4ml fresh LB broth. Screening for the transconjugants was conducted by sampling the mating mixture after 24hour incubation at 37ºC. A sample (100 µl) of mating mixture was plated on the Mueller Hinton (MH) agar, supplemented with two different antibiotics. One of the selected antibiotics promotes the growth of only recipient strains due the ARG found on the recipient chromosome. The second antibiotic selects for the growth of the plasmid-carrying strain, due to the plasmid-carried ARG. The selective environment produced in such a way allows the growth of transconjugants only. The selection of antibiotics was based on the antibiotic resistance profiles of both recipient and donor (Table A1-A5, Supplementary materials, part A). In mating experiments where *K. pneumoniae* was the recipient, 100µl of the mating sample was plated on both MH agar supplemented with transconjugant selective antibiotics (when possible), and on Simmons citrate 1% inositol agar (SCAI agar), selective for *Klebsiella* spp. (Van Kregten et al., 1984), supplemented with cefotaxime selecting specifically for the growth of *K. pneumoniae* transconjugants.

In liquid mating experiments where the DH5α rif^R was the plasmid recipient, a modified liquid conjugation experiment was used (Mo et al., 2016). In short, 500µl of the recipient preculture and 10µl of the donor preculture were mixed in 4 ml LB broth. Sampling and screening for the transconjugants was performed as described above. Antibiotics used to promote the growth of transconjugants were rifampicin and cefotaxime.

2.2.2. Solid surface mating

Solid surface mating was carried out by mixing one colony of the recipient and one colony of the donor directly on a blood agar plate (Mo et al., 2017). Samples were taken by swiping an inoculation loop through the mating mixture on the agar surface and plated directly on the MH agar supplemented with two different antibiotics (for *E. coli* strains) or SCAI agar supplemented with cefotaxime (for *K. pneumoniae*), thus selecting for the growth of transconjugants.

2.2.3. Plasmid transfer frequency

Plasmid transfer frequency was calculated by dividing the number of transconjugants with the number of recipients after 24 hours incubation (Mo et al., 2017). In short, the experiment was conducted as described in the section 2.2.1. After 24 hours of incubation, a sample was taken and serially diluted

(10^{-1} – 10^{-6}). From each diluted bacterial culture from the series, 100µl was plated on MH agar supplemented with antibiotics that promote only the growth of the transconjugants, and MH agar supplemented with antibiotics that promote the growth of recipient cells (both plasmid-carriers and plasmid-free). Antibiotics were selected based on the AMR profiles of both donor and recipient (Table A1-A5, Supplementary materials, part A). Individual colonies formed on MH agar plates were counted and CFU/ml of both transconjugants and recipients was calculated. Based on these values, the plasmid-transfer frequency between the donor strain and the recipient strain was calculated for each mating pair.

2.3. Antimicrobial susceptibility testing by disc diffusion method

The disc diffusion test was conducted as described in the EUCAST[1] disc test manual[2]. A selected bacterial colony was suspended into 5ml 0,9% saline while the turbidity was adjusted to 0,5 McFarland. The bacterial suspension was inoculated on the surface of MH agar plates (Sigma Aldrich, Norway). Selected antimicrobial discs were placed on the agar surface and the plate was incubated at 35±1°C for 18±2h. Antimicrobial discs were MEM (meropenem) 10µg, TE (tetracycline) 30 µg, C (chloramphenicol) 30µg, S (streptomycin) 10µg, W (trimethoprim) 5µg, SXT (sulfamethozole-trimethoprim) 25µg, Amp (ampicillin) 10µg, CTX (cefotaxime) 5 µg, CAZ (ceftazidime) 10µg, CPD (cefpodoxime) 10µg (Oxoid, ThermoFisher Scientific, USA). The inhibition zone diameters created around the discs were measured, and the results were interpreted according to the current clinical breakpoints provided by EUCAST. For streptomycin an in-house breakpoint was used (NVI), due to the lack of breakpoint from EUCAST.

The reference strain *E. coli* ATCC 25922 was used as a quality control. The control was treated in the same manner as described above. The inhibition zones were measured, and values were compared to the range of the inhibition zone diameters previously determined for that strain[3].

2.4. Bacterial growth

In this study both single strain bacterial growth, competitive growth and modified competitive growth assays were conducted.

[1] http://www.eucast.org
[2] http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/2020_manuals/Manual_v_8.0_EUCAST_Disk_Test_2020.pdf
[3] http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/QC/v_10.0_EUCAST_QC_tables_routine_and_extended_QC.pdf

2.4.1. Single strain growth assay

2.4.1.1. High through-put automated OD_{600} measurements of growing bacterial culture

Initial steps of single strain growth assay were performed as described (Alanazi, 2016). Strains were grown separately in liquid nutrient medium in triplicates, ON at 37ºC with shaking 180rpm. Each ON culture was then diluted by the fresh nutrient medium by previously calculated dilution factor. Further, 200ul of each culture was transferred to a 96 well-plate (NUNCLON D-Thermo Fisher Scientific, Denmark). Wells containing only fresh nutrient medium broth (the first and the last row of the 96 well-plate) were included as control blanks. The plate was covered with sealing foil to prevent contamination, and incubated in the SPECTROStar Nano (BMG-labtech, Germany), for 24h at 37ºC. Absorbance (OD_{600}) was measured every 10 minutes with 5 seconds of orbital shaking prior to each measurement. Orbital shaking ensured the aeration of the growing culture and prevented the aggregation of the cells, thereby increasing the accuracy of the measurements.

As this bacterial growth assay was used in different parts of the study, the liquid nutrient medium was either LB or MH broth, while the dilution factor was either the same for all the bacterial cultures included in the experiment, or it was determined for each bacterial culture separately. The determined dilution factor was based on previously determined CFU/ml of the ON culture, thereby ensuring more accurate dilution of the bacterial culture to a desired CFU/ml.

2.4.1.2. Manual OD_{600} measurements of the growing culture

The experiment was conducted similarly as previously described (Brouwer et al., 2019). Single strain bacterial cultures were grown separately in LB broth, ON at 37ºC with shaking, 180 rpm. Subsequently, the cultures were diluted 1:100 with a fresh LB broth and incubated at 37ºC with shaking 180rpm. Samples were taken from the growing cultures at previously determined cell densities measured on Ultrospec™ 10 Cell Density Meter (OD_{600}).

2.4.1.3. Interpretation of the OD_{600} measurements by the logistic model

A data set containing OD_{600} measurements from each well during the 24 hours incubation was collected, and then adapted to a format accepted by the R-package, growthcurver (Sprouffske & Wagner, 2016). Background correction to remove the media absorbance signal was performed by the "min" method implemented in the package. This method corresponds to the correction performed by identifying the lowest OD_{600}-measured value form a set of OD_{600} measurements for a given well. The

identified minimum value was then subtracted from all measurements for the same well. Growthcurver fits the empirical data into the logistic equation (Sprouffske & Wagner, 2016):

$$N_t = \frac{K}{1+(\frac{K-N_0}{N_0})e^{-rt}} \quad \textit{(Equation 2.1)}$$

where N_t is the number of individuals at time t, the N_0 is the initial number of individuals in a growing population, while K (carrying capacity) is the maximum number of individuals that the growing population reaches. The r-parameter represents the maximum growth rate a population reaches during the exponential phase. Although the original model is based on the actual number of bacteria in the culture, the growthcurver takes the OD_{600} values as the directly proportional parameter to the number of bacteria.

Growthcurver generated parameters, N_0 and K, were used to calculate the number of generations produced by a single strain culture (Rockwood, 2015) by Equation 2.2:

$$n = \frac{ln(^K/_{N_0})}{ln\ 2} \quad \textit{(Equation 2.2)}$$

2.4.1.4. Interpretation of the OD_{600} measurements by the exponential model

The data set was manually corrected by the same method described in previous section (growthcurver, correction method "min"), ln-transformed and plotted against time. An exponential growth phase (linear relationship between $lnOD_{600}$ and time) was identified, a trendline for the linear interval was generated together with the trendline formula where slope represents the maximum growth rate of the given culture calculated by the exponential growth model (Equation 2.3) (Rockwood, 2015).

$$lnN_t = lnN_0 + rt \quad \textit{(Equation 2.3)}$$

Parameters N_t, N_0 and r are the same as in the exponential model.

2.4.2. Competitive growth assay

The competitive growth experiment was conducted as described (Lenski et al., 1994; Tietgen et al., 2018). In short, selected pairs of competing strains were grown separately ON, 37°C with shaking 180rpm in 1/3 LB broth. Subsequently, the OD_{600} of ON cultures were measured and cultures of the two competitors were mixed in a 1:1 ratio, based on the measured optical density. The competitive mixture was then diluted in a fresh 1/3 LB broth (Kimura et al., 2012), in a $1:10^4$ ratio. A sample of the newly diluted competitive mixture was taken immediately after mixing, serially diluted and spread on MacConkey agar, supplemented with 20mg/l nalidixic acid, and on MacConkey agar, supplemented with 20mg/l nalidixic acid and 0,5 mg/l cefotaxime. Further, the plate counting method was used to determine CFU/ml of plasmid-carrying cells and the total number of cells. To increase the accuracy of the plate counting method, sample plating was performed in triplicates.

Competing mixtures were then incubated for 5 days, with daily dilutions of the culture (1:10^4) with a fresh 1/3 LB medium. A sample was taken from each culture prior to diluting, then serially diluted and spread on the selective agar as previously described. This procedure was followed throughout the three biological replicates of the experiment.

The ratio of the abundances (R) was calculated every 24 hours of the competition experiment (Lenski et al., 1994), according to Equation 2.4:

$$R = \ln\left(\frac{\frac{CFU}{ml}\,experimental\ strain}{\frac{CFU}{ml}\,reference\ strain}\right) \qquad \textit{(Equation 2.4.)}$$

The values of R were plotted against time (days), and a trend line with its equation was generated. The trend line has the form *R(t)= R(0)+st*, where *s*, the selection rate constant, represents the measured difference between two competing strains during the competition experiment. *s* was calculated for each biological replicate. Values of *s* higher than 0 indicate the increase of the fitness of a strain compared to its reference competing strain, 0 indicates that the fitness of both competing strains is the same, while *s* values lower then 0, indicate fitness loss, compared to its reference competing strain (Lenski et al., 1994; Tietgen et al., 2018).

2.4.3. Modified competitive growth assay for an assessment of plasmid stability

The modified competition experiment was conducted as described in the previous section, similarly as described by Hagbø et al (2019). Briefly, single strain cultures were grown in MH broth, at 37°C, with shaking (180rpm) ON, prior to mixing each competing pair in three different ratios as shown in Table 2.3. The dilution factor for each strain of a competing pair was calculated based on the previously determined CFU/ml of the ON single strain culture. Competing mixtures were grown for 13 days, and the growing culture was diluted daily with a fresh MH broth. Dilution ensured that ON-culture was diluted to ~$1{,}5\times10^6$ CFU/ml which was the initial CFU/ml of each competing culture.

Table 2.3. Plasmid-carrying strain / plasmid -free counterpart mixing ratio in each parallel and their initial CFU/ml.

	Plasmid-carrying strain / plasmid -free counterpart mixing ratio	*Initial CFU/ml of the plasmid-free strain*	*Initial CFU/ml of the competing plasmid-carrying strain*	*Final CFU/ml in the MH broth*
Experimental Parallel 1	1:100	$1{,}5\times10^4$ CFU/ml	$1{,}485\times10^6$ CFU/ml	$1{,}5\times10^6$ CFU/ml
Experimental Parallel 2	100:1	$1{,}485\times10^6$ CFU/ml	$1{,}5\times10^4$ CFU/ml	$1{,}5\times10^6$ CFU/ml
Control Parallel	100% plasmid-carrying culture	/	$1{,}5\times10^6$ CFU/ml	$1{,}5\times10^6$ CFU/ml

2.5. Cell lysis and DNA-extraction

2.5.1. DNA extraction by the boil-lysis method

A bacterial colony grown on a non-selective agar medium was suspended in 300µl milli-Q water. Then, suspensions were lysed by boiling at 100ºC for 15 minutes, and centrifuged for 10 minutes at 5000 rpm, thus separating released DNA from the cell pellet. The supernatant was transferred to a new Eppendorf tube.

2.5.2. Mechanical lysis by bead beating and DNA extraction from the samples

Mechanical lysis and DNA extraction from bacterial samples was conducted as described by Hagbø et al. (2019). In short, bead beating was performed by FastPrep 96 (MP Biomedicals, USA), while degradation of the protein content and DNA extraction from lysed cells was performed in the KingFisher Flex instrument (Thermo Scientific, USA). The procedure ProteinaseLGCmini was used for the protein degradation, which entailed incubating the mixture of the cell lysate with the proteinase at 50ºC for 10min. The automated DNA extraction procedure performed on KingFisher Flex is based on the immobilization of the DNA fragments by paramagnetic silica beads (MagMiniLGC procedure). Details about the reagents and individual steps in both protein degradation and DNA extraction are listed in the Mag Mini LGC protocol[4].

In addition to experimental samples, S.T.A.R. buffer (Roche, Oslo, Norway) was used as a negative control.

Following the DNA extraction, the concentration of three randomly selected samples and negative control was measured on Qubit™ fluorometer (Life technologies, USA), thus ensuring that the DNA extraction was successful.

2.6. Plasmid extraction and electroporation

Plasmid extraction and electroporation were conducted for only one plasmid of special interest, p17437. Extraction was performed as recommended by GenJet Plasmid MidiPrep Kit-protocol[5], while preparation of electrocompetent cells and electroporation was performed as described in the instruction

[4] https://biosearch-cdn.azureedge.net/assetsv6/mag-mini.pdf
[5] https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0012653_GeneJET_Plasmid_Midiprep_UG.pdf&title=VXNlciBHdWlkZTogR2VuZUpFVCBQbGFzbWlkIE1pZGlwcmVwIEtpdCwgSzA0ODE=

manual[6]. The plasmid was extracted from the DH5α transconjugant as recommended by the producer. Recipient cells was made electrocompetent and electroporation was performed on Gene Pulser Xcell™, Electroporation System (BioRad). As a positive control for electroporation, pUC57[7], a high copy number plasmid (ampicillin resistance marker) was used. Electroporated cells were grown ON on selective agar medium supplemented with antibiotics that promote the growth of cells that accepted extracted or control plasmid.

2.7. Qualitative PCR protocols

2.7.1. Phylogenetic grouping

The phylogenetic grouping of a bacterial strain was carried out by multiplex PCR reaction targeting four conserved regions (Clermont et al., 2000; Doumith et al., 2012), i.e. phylogenetic markers, *gadA*, *chuA*, *yjaA* and TSPE4.C2. Primer sequences for each of the markers, their respective amplicon length, and the PCR protocol used are listed in Table A7 (Supplementary materials, part A). PCR products were separated and visualized by gel-electrophorese.

The negative PCR control contained no DNA template, while positive control was phylotype B2 *E. coli* strain containing all phylogenetic markers (Sunde et al., 2015). Interpretation of the *E. coli* phylotype was conducted based on the absence/presence of the genetic markers as demonstrated in Table A8 (Supplementary materials, part A).

2.7.2. PCR based detection of plasmid-carried ESBL/AmpC genes

PCR based detection of ESBL/pAmpC genes includes the amplification of the respective $bla_{\text{CTX-M}}$, bla_{TEM} (ESBL genes) (Briñas et al., 2002; Hasman et al., 2005), or $bla_{\text{CMY-2}}$ gene segments (pAmpC gene) (Pérez-Pérez & Hanson, 2002). Table A9 (Supplementary materials, part A) shows primer sequences used to amplify the segment of $bla_{\text{CTX-M}}$, bla_{TEM} and $bla_{\text{CMY-2}}$ genes, their respective annealing temperatures, length of the segments, and a detailed PCR protocol. Negative control was MilliQ water added instead of the template, while positive control was the original plasmid-host confirmed to contain the respective ESBL/AmpC gene.

PCR products were separated by length using gel electrophorese.

[6] http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4006217A.pdf
[7] https://www.snapgene.com/resources/plasmid-files/?set=basic_cloning_vectors&plasmid=pUC57

2.7.3. PCR amplification of *fumC*-gene

Table A10 (Supplementary materials, part A) lists primers' sequences, their annealing temperature, the PCR produced amplicon length, as well as a detailed description of the PCR protocol (Wirth et al., 2006). Negative control was MilliQ water added instead of the template.

Sanger sequencing of the PCR product was performed at the Molecular Biology section at NVI. The *fumC* gene sequence was searched against the MLST[8] *E. coli* database[9], where the allele variant was determined.

2.7.4. PCR amplification of IncI1 plasmid targeted sequence and 16s rRNA gene targeted sequence

Targeted sequences were amplified by separate PCR reactions by using the primers listed in Table A11 (Supplementary materials, part A). A detailed protocol for these PCR reactions can be found in Table A12 (Supplementary materials, part A). Negative control was MilliQ water added instead of the template.

The DNA concentration of amplified products, as well as the negative control, was measured on Qubit™ fluorometer.

2.7.5. PCR amplification of the shufflon region

Shufflon regions have been amplified by tailed primers (Brouwer et al., 2019) with universal sequences[10]. The amplified shufflon region wass flanked on one side with the 539bp of $pilV_{\text{N-terminus}}$ and 319bp of *rci* on the other. Tailed sequences would allow attachment of the barcode sequence in the next barcoding PCR reaction. Tailed-primer sequences, amplicon size, as well as the detailed protocol are listed in Table A13 (Supplementary materials, part A). MilliQ water was used as negative control.

The DNA concentration of amplified products, as well as the negative control, was measured on the Qubit™ fluorometer. PCR products have been purified with 1x AMPure beads, as described in section 2.7.

[8] https://pubmlst.org/general.shtml
[9] https://pubmlst.org/bigsdb?db=pubmlst_escherichia_seqdef&page=sequenceQuery
[10] https://community.nanoporetech.com/protocols/pcr-96-barcoding-amplicons/v/PBAC96_9069_v109_revN_14Aug2019

2.8. Purification of PCR products

The amplified products were purified with AMPure XP beads (Beckman Coulter, USA). Beads enable size selective binding of the PCR product. In the presence of the magnet, the beads are immobilized while several washing steps take place. In the final step, the now purified PCR product is released from the beads. Due to different sizes of the amplified product, the final concentration of beads was 1,5x and 0,8x for 16s rRNA amplicon and IncI1 amplicon, respectively. All the steps were performed as recommended in the protocol[11].

After the purification step, the DNA concentration of purified PCR products was measured on Qubit™ fluorometer.

2.9. Quantitative PCR

Quantitative PCR (qPCR) was used in order to quantify the amount of the IncI1 plasmid and, the targeted 16s rRNA gene sequence per sample, (Hagbø et al., 2019). Table A11 (Supplementary materials, part A) shows primer sequences targeting selected gene segments, their amplicon lengths, respective annealing temperatures and details about qPCR protocol used (Carattoli et al., 2005; Yu et al., 2005).

2.9.1. qPCR of standard DNA dilutions

Standard DNA dilutions with a known concentration, were created by serially diluting purified PCR products of targeted DNA segments, from 10^{-1} to 10^{-8} in triplicates. The DNA concentration was calculated for each dilution. qPCR of standards was performed as described in section 2.9.

2.9.2. Quantification of targeted genes

For each standard DNA dilution of purified PCR products, the number of targeted gene copies per sample was calculated based on the DNA concentrations measured with Qubit™ fluorometer. The following equation was used:

$$Number\ of\ coipes = \frac{(\boldsymbol{X}\frac{ng}{\mu l}\times 1\mu l\times 6.0221\times 10^{23}\frac{molecules}{mole})}{\left(N\times 660\frac{g}{mole}\right)\times 10^{9}\frac{ng}{g}} \qquad \textit{(Equation 2.5)}$$

X represents the concentration of purified amplicon (ng/µl), 1µl is the volume of the PCR product used in a qPCR reaction, 6,0221*10^{23} molecules/mole is the Avogadro's number, N represents

[11] https://genome.med.harvard.edu/documents/sequencing/Agencourt_AMPure_Protocol.pdf

the length of the amplicon (bp), 660 g/mol is the average mass of 1bp dsDNA, while 10^9 ng/g converts ng to g. The result is the number of molecules present in the standard solution. A copy number of each targeted gene sequence was calculated with the aid of the dilution factor used for preparing standard dilutions (1:10).

Constructed standard curves represent the linear relationship between Cq-values obtained by the qPCR of standards of each targeted sequence, and log copy numbers of amplified DNA fragments per standard solution, as calculated by the Equation 2.5. The standard curve equation was applied to convert Cq values of the experimental samples to log copy numbers of targeted sequences per sample. Prior to conversion, the threshold for the Cq values of experimental samples was manually adjusted to the thresholds of the standard dilution Cq-values for each of the respective amplicons.

2.10. MinIon amplicon sequencing

The preparation of PCR amplified products for MinIon sequencing is described in detail in the Nanopore Protocol SQK-LSK109[12]. In short, the PCR products from different samples were barcoded by barcoding PCR reaction (Table A14, Supplementary materials, part A). The barcoded PCR products were run on the gel to verify the success of the barcoding, while DNA concentrations of the products were measured on the Qubit™ fluorometer. For the next step, the products were purified with 1x AMPure beads, as described in section 2.8.

Barcoded libraries produced in the previous step were pooled in equal DNA amounts. The concentration of the pooled library was measured on the Qubit™ fluorometer. The final concentration of the pooled library was adjusted to 1µg DNA in 47µl milliQ water.

The next step consisted of DNA repair and end-prep. This step was followed by the 1x AMPure beads purification (as described in section 2.7.) where the purified product was eluted with 61µl of milliQ water in the final step.

This was followed by adapter ligation and clean-up. After the adapter ligation, the pooled library was washed with AMPure beads (x0,4). The immobilized library bound to the AMPure beads was washed twice with a short fragment buffer, and finally eluted with 15µl elution buffer. The concentration of the purified pooled library was measured on the Qubit™.

In the final step, the flow cell was primed, and the library was loaded on the flow cell (FLO-MIN106D type R9.4.1.). The number of active nanopores was verified in a quality check step prior to loading the library. The sequencing was run for 48h.

[12] https://community.nanoporetech.com/protocols/pcr-96-barcoding-amplicons/v/PBAC96_9069_v109_revN_14Aug2019

2.11. Bioinformatics

2.11.1. Plasmid sequence assembly

Two out of seven plasmid-host genomes used in this study have been previously assembled. For plasmid-host 2012-01-2798, both short reads (Illumina) and long reads (PacBio) were available, while for plasmid-host 2016-01-20481 only short reads were available. Depending on the type of available reads, two different assembling methods were used.

2.11.1.1. Unicycler, hybrid assembly

Unicycler, an assembly pipeline used for bacterial genome assembly, is applicable both on short reads only, long reads only and for hybrid assembly (combination of short and long reads) (Wick et al., 2017). Hybrid assembly constructs assembly graphs generated by the SPAdes using only the short reads. Contigs created in this way can have multiple unresolved bridges to other contigs due to often-present repeated regions at their ends. These multiple bridges are resolved by the employment of the long reads resulting in a possibly complete separate contig of a bacterial chromosome sequence, and if present, separate contigs of present MGEs.

The whole genome of the strain 2012-01-2798 was assembled by the Unicycler. Contigs were searched against the BLAST[13] (Basic Local Alignment Search Tool) data base, where a separate contig of the assembled plasmid sequence was identified.

2.11.1.2. Short reads assembly

De novo assembly using only short reads can introduce an error to the assembly. In order to avoid this, a reference assembly was conducted to increase the accuracy of the assembled plasmid sequence. Short reads of 2016-01-20481 were first assembled into contigs which were then blasted, and reference plasmid were identified with the highest identity percentage. Next, Bowtie2 (Langmead & Salzberg, 2012) was used to map short reads on to the reference plasmid, while Ugene was used to visualize the mapped reads and to extract the best consensus plasmid sequence (Okonechnikov et al., 2012).

[13] https://blast.ncbi.nlm.nih.gov/Blast.cgi

2.11.2. Annotation of assembled plasmids

Annotation of the assembled plasmids was performed by Prokka: rapid prokaryotic genome annotation (Seemann, 2014). The assembled plasmid sequences were searched against the BLAST nucleotide database, and the identified similar plasmid sequences (similarity >95%) were further used as annotation reference. The Table 2.4. lists reference annotated sequences for each newly assembled plasmid. Prokka uses the "--protein" command to search for homologous coding sequences (CDSs) from the reference in the non-annotated sequence. Annotated plasmids were visualized and manually edited by Artemis (Carver et al., 2005). The manual editing of annotated CDSs ensured uniform naming and color-coding of the same CDS across all plasmids. The CDSs annotated as "hypothetical proteins" were investigated by pair-wise alignment with the in-more-detail annotated plasmid in Artemis. When the alignment indicated that the orientation and position of a "hypothetical protein" CDS matched the better annotated CDSs from another annotated plasmid, the nucleotide sequence of the hypothetical protein was pair-wise BLASTed against its correctly annotated CDSs. If the query cover and percent identity were 100%, the hypothetical protein was re-annotated manually in Artemis by changing the gene-name after its corresponding CDSs found on the fully annotated plasmid.

Although pL-II was previously annotated, the plasmid was annotated once more due to the large non-annotated regions found to be important in this study with Prokka, as described above. A reference plasmid sequence used for the re-annotation of this plasmid is listed in the Table 2.4.

Table 2.4. Accession number of the plasmids used as annotation references.

Plasmid	Reference plasmid sequences used in annotation/assembly (accession number)
p2798	CP012929
p1248	KJ484639
p20481	p20426 (no accession number)
pL-II	MK070495

2.11.3. Alignment, comparation and visualization of plasmids

Easyfig with integrated BLAST-based alignment was used to visualize and compare the plasmid annotated sequences (Sullivan et al., 2011).

2.11.4. Shufflon rearrangement analysis workflow

The obtained sequencing database was demultiplexed by EPI2ME (Metrichor, Oxford, UK). The adapter sequences were removed by the Porechop (Wick, 2018). The FASTQ files were converted to the FASTA files. The number of possible shufflon variants without segment duplication was

calculated with the following equation where n represents the number of shufflon segments (Sekizuka et al., 2017):

Number of possible structure variants = 2^n x n! *(Equation 2.6)*

Sequences of all possible shufflon variants (48 variants in total) were generated. Due to a previous study (Brouwer et al., 2019) where the shufflon variants with the truncated structure (deleted segments) were detected, the shufflon structure variants with a deletion of one (24 variants in total), two (6 variants in total) or all the segments (1 variant) were also generated as a FASTA file format. Total number of shufflon variants without the duplication of the segments were 79. By merging the FASTA sequence files of possible shufflon variants the FASTA sequence database was created.

As this study included plasmids with the same original shufflon structure (three segments A, B and C), but where two out of three plasmid shufflons harboured the insertion-sequence interrupted B segment, two different databases with shufflon variant sequences were generated. The one containing the sequences without the interrupted B sequence, and the other with the B sequence harbouring insertion sequence. Finally, the database of shufflon variants sequences with the interrupted B segment included all variants with the uninterrupted B segment, thereby increasing the number of variants for this shufflon to 145.

In the following step BLAST+ (sub-option megaBLAST) was used to align all the reads from a sample to all shufflon sequence variants with reads as query sequences and shufflon variants as a reference database. The output was loaded into R studio where the correct alignments were filtered by the following criteria: the length of the alignment had to match the length of a shufflon variant the read was aligned against, the start and the end of the correct alignment had to match a shufflon variant a read was aligned to, the percentage of the identical matches in the alignment had to exceed 90%. Correctly aligned reads to each of the shufflon variants were counted.

3. Results

3.1. Characteristics of IncI1 plasmids included in this study

As evident in pR64, the distinct conserved IncI1 backbone is identical for all the plasmids encompassed by this study (Figure 3.1). The greatest variation between plasmids is in the variable region, except for p20481 and p20426, which appear to be completely identical, although found in two different hosts. In plasmids that carry *bla*$_{\text{CTX-M-1}}$ (p20481, p20426, p22638 and p17437), this ARG is not part of the variable region, but is instead inserted into the B-segment of the shufflon. These plasmids carry the *sul2* (sulfonamide resistance gene) in the variable region, while p22638 carries *tetA* that encodes for tetracycline resistance. The p1248, another ESBL encoding plasmid, carries *bla*$_{\text{TEM-20}}$ in the variable region, while p2798 carries an AmpC gene (*bla*$_{\text{CMY-2}}$) in the same region.

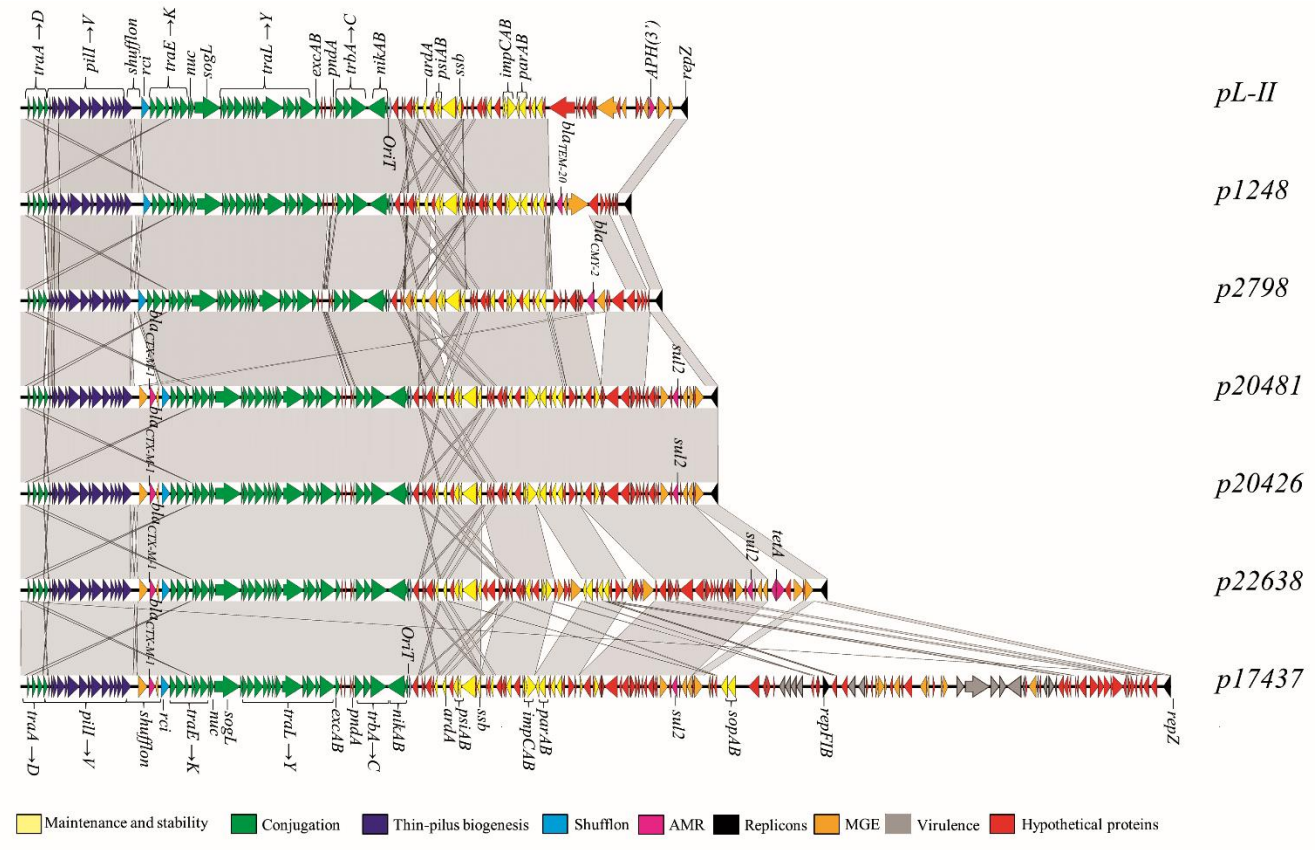


Figure 3.1. Genetic organization of plasmids pL-II, p1248, p2798, p20481, p20426, p22638, p17437. Arrows indicate genes their respective direction and length. Colour and their meanings are as follows: yellow – genes involved in maintenance and stability, green- genes involved in conjugative transport, purple- genes involved in thin pilus biogenesis, blue-shufflon region, pink-AMR, black- replicon, orange- MGE, grey-virulence genes, red – hypothetical proteins. Areas shaded grey indicate homologous regions, nucleotide identity threshold >95%.

Although it was found in an ESBL-producing host, the pL-II sequence indicates that it is not a carrier of an ESBL encoding gene. The *aph(3')-Ia* that encodes for the resistance to aminoglycosides, is the only ARG encoded by this plasmid.

The p17437 is unique among the plasmids, in that it contains two replicons, IncI1 and IncFIB. The variable region of this plasmid appears to be cointegrated with the variable region of an IncFIB plasmid, as shown in the previous study (Mo et al., 2020). Presence of the *sopA* and *sopB* genes, commonly found on IncF plasmids, and functionally similar to *parA* and *parB* genes, is another indication that approximately 60kpb originates from an IncFIB plasmid.

3.2. Initial conjugation experiments

In order to study the plasmid fitness cost in wild-type *E.coli* strains, p17437, p22638, p20481, p1248 and p2798 were initially transferred to new hosts by the liquid/solid surface conjugation (Table 3.1). The aim of the initial mating was to detect at least one strain capable of accommodating the selected plasmids out of each of the three selected ST of quinolone resistant *E. coli* (QREC). The same attempt was made with the 2011-01-1173 avian pathogenic *E. coli* (APEC) as a recipient strain (henceforth referred to as APEC).

Details about all conjugation experiments are shown in Table B1 (Supplementary materials, part B), while Table 3.1 lists only mating pairs that resulted in successfully confirmed transconjugants. Confirmation of transconjugants was based on both the transconjugant's colony morphology on non-selective medium (blood agar), the phylotype, and the presence of the ESBL/AmpC gene. When 2011-01-1173(APEC) was conjugated with the 2016-01-20481 plasmid host, phylotyping was an inappropriate method of transconjugant confirmation, because the recipient and the donor both belong to the same phylogroup. This was circumvented by amplifying and sequencing a segment of the *fumC* gene of the transconjugant. The allele variant of this gene was determined and compared to the *fumC* allele variant of the donor strain (whole genome sequencing data were available). Different alleles of the *fumC* gene segments between the transconjugant and the donor confirmed the transconjugant. Potential transconjugants identified on the double selective agar whose phyolotype and presence of the ESBL/AmpC gene were not tested by the PCR based assays, were not considered confirmed.

None of the *K. pneumoniae* strains were able to successfully accept and maintain any of the selected plasmids. An attempt to isolate plasmid DNA of p17437 and electroporate the APEC strain with the isolated plasmid was unsuccessful, although the electroporation of electrocompetent APEC cells with the pUC57 plasmid vector was successful.

Table 3.1. Successful conjugation pairs with confirmed transconjugants.

Donor	Recipient	Conjugation surface	Successfully confirmed transconjugant
2016-01-17437	→2014-01-2773 (QREC/ST162)	L	p17437/2773 (ST162)
2016-01-20481	→2011-01-1173 (APEC)	L	p20481/1173
	→2014-01-2773 (QREC/ST162)	L	p20481/2773 (ST162)
	→2011-01-3460-5 (QREC/ST602)	L	p20481/3460-5(ST602)
	→2014-01-6924 (QREC/ST453)	L	p20481/6924 (ST453)
2016-01-22638	→2011-01-1173 (APEC)	L	p22638/1173
	→2014-01-2773 (QREC/ST162)	L	p22638/2773(ST162)
	→2009-01-3815 (QREC/ST602)	L	p22638/3815 (ST602)
	→2009-01-4618-2 (QREC/ST602)	L	p22638/4618-2(ST602)
	→2011-01-3460-5 (QREC/ST602)	L	p22638/3460-5(ST602)
	→2014-01-6924 (QREC/ST453)	L	p22638/6924 (ST453)
	→2014-01-7234-1 (QREC/ST453)	L	p22638/7234-1 (ST453)
2012-01-2798	→2011-01-1173 (APEC)	L/S	p2798/1173 (ST453)
	→2014-01-2773 (QREC/ST162)	L	p2798/2773 (ST162)
	→2009-01-3815 (QREC/ST602)	L	p2798/3815 (ST602)
	→2009-01-4618-2 (QREC/ST602)	L	p2798/4618-2 (ST602)
	→2011-01-3460-5 (QREC/ST602)	L	p2798/3460-5 (ST602)
	→2014-01-6924 (QREC/ST453)	L	p2798/6924 (ST453)
	→2014-01-7234-1 (QREC/ST453)	L	p2798/7234-1 (ST453)
2006-01-1248	→2011-01-1173 (APEC)	L	p1248/1173
	→2014-01-2773 (QREC/ST162)	S	p1248/2773 (ST162)
	→2009-01-3815 (QREC/ST602)	L	p1248/3815 (ST602)
	→2009-01-4618-2 (QREC/ST602)	L	p1248/4618-2 (ST602)
	→2011-01-3460-5 (QREC/ST602)	L	p1248/3460-5 (ST602)
	→2014-01-6924 (QREC/ST453)	L	p1248/6924 (ST453)

Abbreviation: L-liquid, S-Solid surface

The *E.coli* ST162 strain 2014-01-2773 (hereafter termed 2773(ST162) strain) was the only strain capable of accepting p17437, in addition to being capable of accepting and maintaining plasmids from the rest of the donor strains. Three QREC strains in total, belonging to three different sequence types (STs), ST162 (2773(ST162)), ST453 (2014-01-6924, henceforth referred to as 6924(ST453)) and ST602 (2011-01-3460-5, henceforth referred to as 3460-5(ST453)) successfully accepted and maintained p20481, p22638, p2798 and p1248. Transconjugants of the APEC strain, p22638/APEC, p20481/APEC, p2798/APEC and p1248/APEC were also confirmed. Plasmid transfer frequency was determined for:

- 2016-01-22638(p22638) →2773(ST162),
- 2016-01-17437(p17437) → 2773(ST162),
- 2016-01-20481(p20481) → 2773(ST162),
- 2016-01-22638(p22638) → APEC, and
- 2016-01-20481(p20481) → APEC.

3.3. Transfer frequency

Transfer frequency (TF) was calculated for the selected pairs after the 24 hours mating in LB broth. The experiment was conducted in three biological replicates, while each replicate comprised three technical replicates. Technical replicates improved the accuracy of the plate counting method used to determine the number of transconjugants and the number of recipient cells. TFs of each biological

replicate were used to calculate the mean TF value and standard deviation (SD) for each conjugation pair.

Although, mean TF values varied between conjugation pairs (Figure 3.2, Table B2, Supplementary materials, part B), none of the observed differences were statistically significant (p> 0,05, Table B3, Supplementary materials, part B) owing to high TF's SD values.

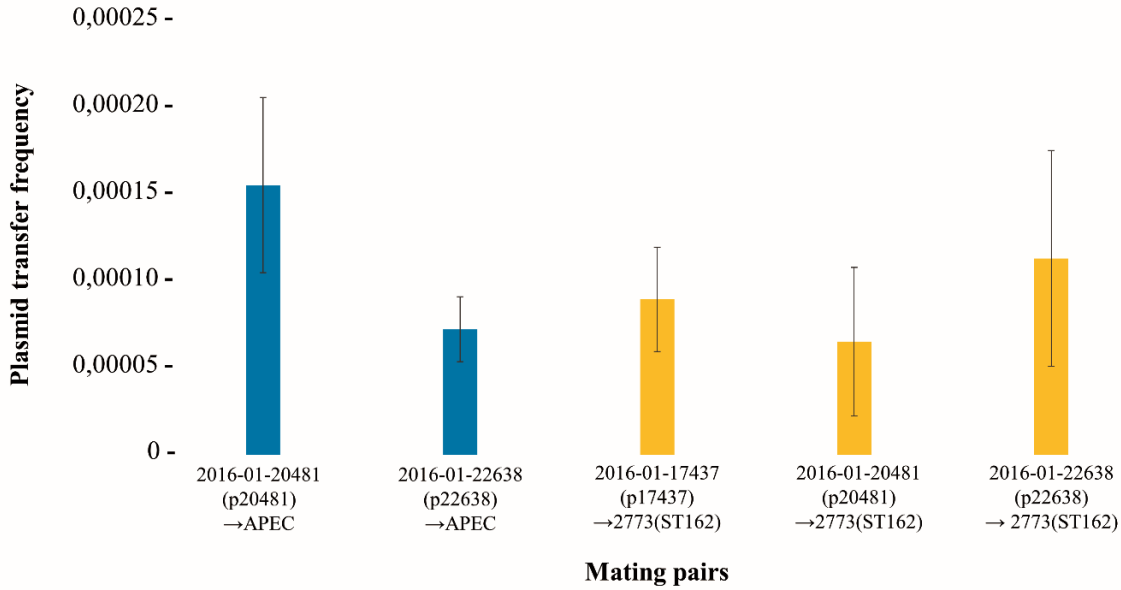


Figure 3.2. Mean transfer frequencies of selected mating pairs calculated after 24h incubation of mating pairs, bars indicate SD.

3.4. Plasmid fitness cost

Plasmid fitness cost was examined using the single strain growth assay, and the competitive growth assay. The r_{max} and K (the maximum OD_{600} value the culture reached during 24h growth) were the parameters, generated by the growthcurver R-package (Sprouffske & Wagner, 2016), chosen to define single strain growth of the selected strains. Observed differences between selected recipients and their respective transconjugants were further investigated in the competitive growth assay.

3.4.1. Single strain growth assay

Pure cultures of selected strains were incubated for 24h in LB broth, while their OD_{600} was measured every 10 minutes. The collected data were interpreted with both the exponential and the logistic population growth model. The logistic model is implemented into the growthcurver and generates both r_{max} and K for each growing culture. The exponential model generates only the r_{max}. Although r_{max} parameters from these models should be the same, differences in values were substantial. The assumption that these differences were caused by the growthcurver package when it attempted to interpret the whole data set and fits it into the logistic model in the best possible way, was not

investigated further. Therefore, r_{max} values from the manually applied exponential model for each culture dataset were used, rendering the r_{max} values from different cultures comparable.

Growth curves generated based on both empirical OD_{600} measurements and generated based on the logistic model by the growthcurver, for each strain are presented in Figures B1-B4 (Supplementary materials, part B). Based on the comparison of the empirical growth curve and the logistic model growth curve, the K values were assumed not to be affected in the same manner as r_{max} values. K-values were therefore found suitable to describe the culture growth.

The single strain growth assay comprised biological triplicates of each culture. Calculated r_{max} and K values from the culture replicates were used to calculate mean r_{max} and K, and their respective SD values, presented in Figure 3.3. The values of these parameters for each growing culture triplicate are shown in Table B4 (Supplementary materials, part B).

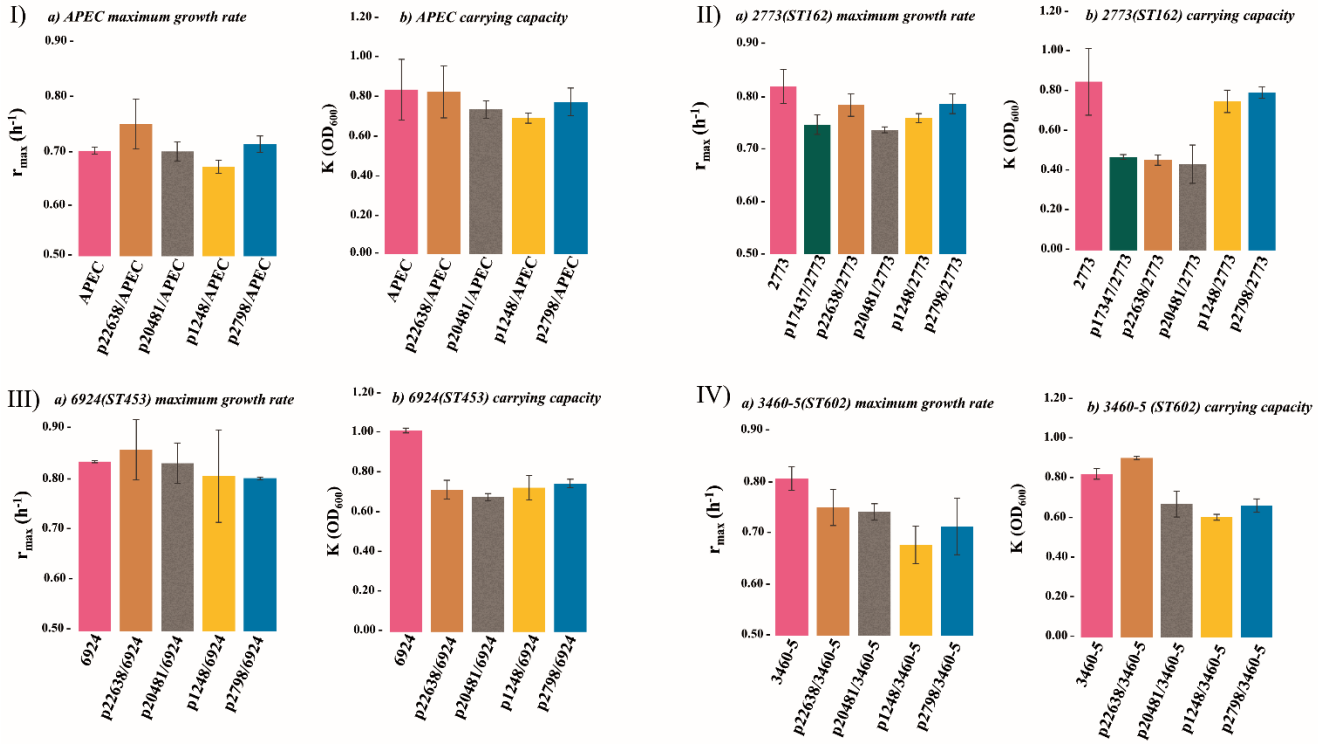


Figure 3.3. Mean maximum growth rates (r_{max}) and carrying capacities (K) of the recipient strains and its respective transconjugants, bars indicate SD values. I) mean r_{max} (a) and K (b) of APEC and its transconjugants, II) mean r_{max} (a) and K (b) of 2773(ST162) and its transconjugants, III) mean r_{max} (a) and K (b) of 6924(ST453) and its transconjugants, IV) mean r_{max} (a) and K (b) of 3460-5(ST602) and its transconjugants.

In general, mean r_{max} of all strains varied between 0,67 and 0,83 indicating that no greater difference in r_{max} has been observed between any of the strains (Table B4, Supplementary materials, part B).

In the APEC strain and its transconjugants (Figure 3.3 Ia), the mean r_{max} varied between 0,69 and 0,81 (Table B4, Supplementary materials, part B). Although not statistically significant ($p>0,05$, Table B5, Supplementary materials, part B), the p22638/APEC exhibited higher r_{max} than its respective recipient strain. Furthermore, the p1248/APEC was the only APEC transconjugant that had significantly lower r_{max} ($p<0,05$, Table B5, Supplementary materials, part B) than its recipient, although the difference was only 0,03 (Table B4, Supplementary materials, part B).

In the 2773(ST162) strain, the mean r_{max} of the transconjugants had a tendency to be lower than the mean r_{max} of the recipient strain (Figure 3.3 IIa). However, a significantly lower r_{max} was confirmed in the p17437/2773(ST162) and p20481/2773(ST162) ($p<0,05$, Table B5, Supplementary materials, part B). The mean r_{max} of both the recipient and the transconjugants varied between 0,73 and 0,83 (Table B4, Supplementary materials, part B).

The mean r_{max} of the 6924(ST453) strain and its transconjugants exhibited minor variations between 0,8-0,86 (Figure 3.3 IIIa, Table B4, Supplementary materials, part B). The only transconjugant that had a significantly lower mean r_{max} was p2798/6924(ST453) ($p<0,05$, Table B5, Supplementary materials, part B), although difference was only 0,03 (Table B5, Supplementary materials, part B).

Regarding the 3460-5(ST602) and its transconjugants, the tendency of transconjugants to have a lower growth rate than the recipient strain was observed once again (Figure 3.3 IVa). The difference was statistically significant between the recipient strain and the p20481/3460-5(ST602), and between the recipient and p1248/3460-5(ST602) ($p<0,05$. Table B5, Supplementary materials, part B). The variation of the mean r_{max} in this group was again minor, where the mean r_{max} values varied between 0,67 and 0,80 (Figure 3.3, Table B4, Supplementary materials, part B).

In addition, r_{max} value of the APEC strain was found to be significantly lower than the r_{max} values of the QREC strains ($p<0,05$, Table B6, Supplementary materials, part B).

On the other hand, differences in the carrying capacity (K) between the recipients and their respective transconjugants were more emphasized (Figure 3.3 Ib-IVb, Table B4., Figure B1-B4, Supplementary materials, part B).

Mean K values of the APEC and its transconjugants were the most uniform compared to other groups (Figure 3.3. Ib). In addition to being minor, the observed differences were not statistically significant ($p>0,05$, Table B5, Supplementary materials, part B).

The transconjugants of the 2773(ST162) strain, p17437/2773(ST162), p22638/2773(ST162) and p20481/2773(ST162), all had a nearly identical mean K value (Figure 3.3 IIb). Their K values were approximately two times lower than the K value of the recipient (Table B4, Supplementary materials, part B). Based on the Student's T-test, the mean K values of these transconjugants were significantly lower than the mean K-value of the recipient ($p<0,05$ Table B5, Supplementary materials, part B). Transconjugants p1248/2773(ST162) and p2798/2773(ST162) reached a mean K-values closer to the recipient strain and were not significantly different ($p>0,05$, Table B5, Supplementary materials, part B).

A similar trend was observed in 6924(ST453), where all transconjugants had mean K values ranging from 0,67 and 0,76 (Figure 3.3 IIIb, Table B5, Supplementary materials, part B). The observed difference between the transconjugants and the recipient was significantly different ($p<0,05$ Table B6, Supplementary materials, part B).

With the exception of p22638/3460-5(ST602), the mean K values of the 3460-5(ST602) transconjugants were significantly lower ($p<0,05$, Table B5, Supplementary materials, part B) than the

recipient's mean K value (Figure 3.3). The identified difference ranged between 0,16-0,23 depending on the transconjugant (Table B4, Supplementary materials, part B). On the other hand, p22638/3460-5(ST602) had a significantly higher mean K value than its recipient ($p<0,05$, Table B5, Supplementary materials, part B), although the difference was only 0,08 (Table B4, Supplementary materials, part B).

3.4.2. Competitive growth

In this experiment, a recipient and its transconjugant were grown pairwise with the same initial CFU/ml. The assumption was that the growth differences between the two strains would be more emphasized when growing together and competing for the limited nutrients and space.

Selected pairs included:

- p17437/2773(ST162) + 2773(ST162)
- p22638 /2773(ST162) + 2773(ST162)
- p22638/APEC + APEC.

These pairs were selected based on the results from the single strain growth assay. In addition to the fact that the p17437 was the largest plasmid included in this study, the single strain growth experiment indicated that the mean r_{max} and K of p17437/ 2773(ST162) were significantly lower than the parameters obtained for the recipient. To further investigate this difference, it was considered beneficial to examine how this transconjugant would behave when grown in pair with its recipient. Since p22638/2773(ST162) transconjugant had mean r_{max} value somewhat higher than p17437/2773(ST162) and mean K value in the same range as p17437/2773(ST162), this transconjugant was also chosen to compete with the recipient. On the other hand, mean r_{max} and K values of the p22638APEC were in the same range as mean r_{max} and K of the recipient. The difference was not statistically significant, although the transconjugant exhibited somewhat higher mean r_{max} than the recipient. Therefore, the p22638/APEC was also included in the competitive growth experiment in order to examine potential differences caused by the same plasmid in different hosts.

In the pairwise competition growth assay, the transconjugant would also be a plasmid donor and, thereby, the conjugation initiator. This would lead to higher number of transconjugants in the competition mixture not only by the means of the clonal expansion. However, based on mean TF values (Figure 3.2), the assumption was that the emergence of the new transconjugants would not significantly affect the results.

Each competing pair was grown in biological triplicates under the same conditions. Within each of the triplicates, three technical replicates of the plate counting method were used to increase the accuracy of the calculated CFU/ml of each competitor in the mixture. Biological replicates were used to calculate the mean R values and its respective SD of each consecutive time point the sample was taken. R values represent the ln-transformed CFU/ml ratio of the transconjugant and the recipient found after every 24 hours of pairwise competitive growth. The R value has a negative value when the CFU/ml

ratio of the transconjugant and the recipient was <1 , R is zero when this ratio is 1, and R has positive values when this ratio is >1.

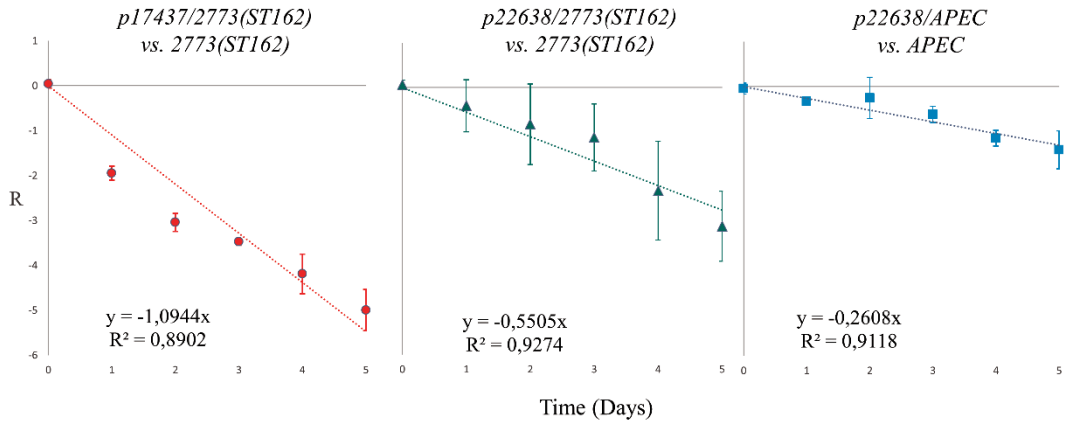


Figure 3.4. Competitive growth experiment. Mean R values represents the ln-transformed CFU/ml ratio of a transconjugant and its respective recipient calculated for each day during the 5 days. SD values are represented as bars. The trendline equation and its respective R^2 are displayed for each competing pair.

Mean R values for each competing pair (Table B7, Supplementary materials, part B) was used to generate the trendline, where the slope of the trendline represents the selection rate constant, *s*. Negative *s* values in all competing pairs indicate the reduction of the number of transconjugant compared to the recipient strain. The lowest *s* value was found in the p17437/2773(ST162) vs 2773(ST162) competing pair, while the highest was identified in the p22638/APEC vs. APEC competing pair. Compared to other two competing pairs, the p22638/2773(ST162) vs. 2773(ST162) pair exhibited the greatest variation of the individual R values (indicated as high SD).

3.5. Plasmid stability

Plasmid stability was examined in a modified competition assay where each competing pair encompassed a plasmid free (p^-) DH5α rif^R and one of its plasmid-containing (p^+) counterparts (DH5α rif^R transconjugants). Three of the p^+ DH5α rif^R strains were generated in the previous study (Hagbø et al., 2019). Each of these strains are carriers of the same pL-II plasmid (pL-II-22/DH5α rif^R, pL-II-30/DH5α rif^R and pL-II-55/DH5α rif^R). These transconjugants differ from one another in terms of presence/absence of additional MGEs, and AMR profiles. AMR profiles determined in the previous study were confirmed with the disc-diffusion method prior to the growth experiments.

Two DH5α rif^R transconjugants carrying p2798 and p24026 were successfully generated in a liquid mating experiment. The transconjugants were not confirmed with the PCR method.

Henceforth, the pure culture of DH5α rif^R will be referred to as the p^- DH5α rif^R strain, while transconjugants of DH5α rif^R carriers of different plasmids will be referred to as p^+ DH5α rif^R strains.

3.5.1. Single strain growth curves

A single strain growth experiment was conducted in order to estimate the number of generations a pure culture is capable of producing within 24 hours. The cultures were grown in MH broth. The number of days required to produce 50 generations (with daily dilution of the growing culture) can be calculated by estimating the number of generations produced daily. Each pure culture was grown in triplicates with both $1{,}5x10^4$ and $1{,}5x10^6$ initial CFU/ml. The pure cultures were grown for 24h, while OD_{600} were measured every 10 minutes.

Mean K and N_0, as generated by Growthcurver, were used to calculate the number of generations for each single strain culture. Table B8 (Supplementary materials, part B) lists the mean K and N_0 values based on the triplicates of the growth experiment, and the number of generations for each single strain culture with the initial $1{,}5x10^6$ CFU/ml. Table B9 (Supplementary materials, part B) shows the same parameters calculated for each single strain culture with the initial $1{,}5x10^4$ CFU/ml. For cultures with a higher initial CFU/ml, 13 days were sufficient to produce ~50 bacterial generations.

Figure B6 and B7 (Supplementary materials, part B) summarizes growth curve triplicates of each culture, with both initial $1{,}5x10^6$ and $1{,}5x10^4$ CFU/ml. Different initial CFU/ml of the cultures resulted in a different lag-phase duration.

The competitive growth of the p^+ and the p^- DH5α rif^R strains mixed in a 1:100 ratio was the main experimental variable of the experiment. Single strain growth curve triplicates of the p^- DH5α rif^R strain with the initial $1{,}5x10^6$, and single strain growth curve triplicates of all p^+ DH5α rif^R strains with the initial $1{,}5x10^4$ CFU/ml are compared in Figure B8 (Supplementary materials, part B). Owing to the shorter lag-phase, less time is needed for the p^- DH5α rif^R strain culture to reach the stationary phase, compared to p^+ DH5α rif^R cultures with a 100 times lower initial CFU/ml. For this reason, the assumption was that any p^+ DH5α rif^R strain would be highly unlikely to outcompete the p^- DH5α rif^R strain during the pairwise competitive growth.

This assumption was further examined by comparing the calculated mean r_{max} of the single strain cultures (Figure 3.5). The mean r_{max} values were caluclated based on r_{max} values (exponential growth model) of each biological replicate (Table B10, Supplementary materials, part B).

The p^- DH5α rif^R r_{max} values were significantly higher compared to the mean r_{max} values of all p^+ DH5α rif^R ($p<0{,}05$, Table B11, Supplementary materials, part B). By being the majority in the main experimental variable, the assumption was that the p^+ DH5α rif^R would grow and reach the stationary phase faster than all the donors confirming once again that the p^+ DH5α rif^R strain would not be able to outcompete the p^+ DH5α rif^R exclusively by clonal expansion.

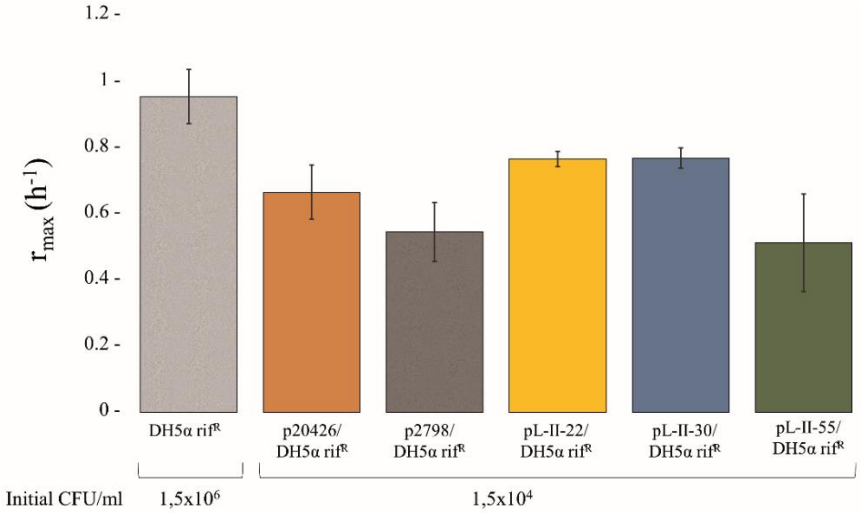


Figure 3.5. Mean maximum growth rates of the pure cultures, participants of the main experimental variable in the modified competitive growth experiment. Vertical bars represent standard deviation for each calculated mean r_{max}.

3.5.2. Competitive growth assay and determining the number of plasmid copies and the number of 16s rRNA gene copies per sample

The competitive growth experiment was conducted by mixing the p$^+$ DH5α rif^R culture and the p$^-$ DH5α rif^R culture in three different ratios in MH broth. In experimental parallel 1, the p$^+$ DH5α rif^R / p$^-$ DH5α rif^R was 1:100, in experimental parallel 2 the p$^+$ DH5α rif^R / p$^-$ DH5α rif^R ratio were 100:1, while control parallel contained only the pure p$^+$ DH5α rif^R culture. Regardless of the parallel, the initial CFU/ml was 1,5x10^6. The competing mixtures were grown for 13 days. At 24 hours intervals the competing mixtures were diluted to approximately 1,5x10^6 CFU/ml, and further incubated. Sampling of the competing mixtures took place every 24 hours prior to the dilution.

The samples were mechanically lysed, DNA extracted and submitted to qPCR with the appropriate primers to determine the number of IncI1 plasmids and the number of 16s rRNA gene copies per sample. The ratio of the log-transformed number of IncI1 plasmid copies and log-transformed number of 16s rRNA copies per sample was used to determine the relative plasmid abundance (RPA) per samples.

The number of 16s rRNA copies per *E.coli* DH5α genome is seven (accession number: NZ_JRYM00000000), while the number of IncI1 plasmid copies per cell is under 10. Based on this, the assumption was that the RPA value per sample close to 1 indicates that the majority of cells in the sample are p$^+$ DH5α rif^R cells, while RPA values close to 0 indicate a low number of p$^+$ DH5α rif^R cells in the sample.

The RPA values calculated for each sampling time point from each competing pair are presented in Figure 3.6 (Table B12, Supplementary materials, part B). The RPA values were above 0,80 after the

first 24 hours in all samples regardless of the initial mixing ratio (parallel 1 with 1% p^+ DH5α rif^R, and parallel 2 with 99% p^+ DH5α rif^R) and regardless of the competing pair.

RPA values of the control parallels were seemingly within a similar range as RPA values of their respective experimental parallels. This observation was further investigated by conducting t-test comparison. Each set of RPA values of the experimental parallels was compared to the RPA set of values of their respective control parallels.

According to the t-test, none of the experimental parallels 1 (1% p^+ DH5α rif^R) were significantly different from their respective control parallels, regardless of the competing pair ($p>0,05$, Figure 3.6). The same result was obtained when the experimental parallels 2 (99% p^+ DH5α rif^R) were compared to their respective control parallels, with the exception of the competing pair pL-II-55/DH5α rif^R vs. DH5α rif^R, where the p-value of the T-test ($p<0,05$) indicated a significant difference between these two parallels. (Figure 3.6).

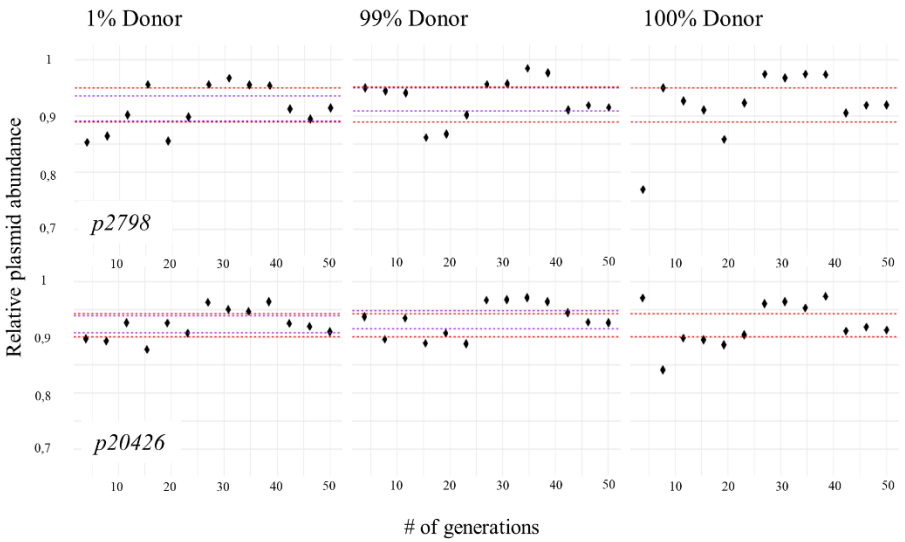
In order to further verify the similarities between the experimental parallels and their control parallels, the confidence intervals (CI) of 95% were calculated for each parallel of competing strains, according to the formula:

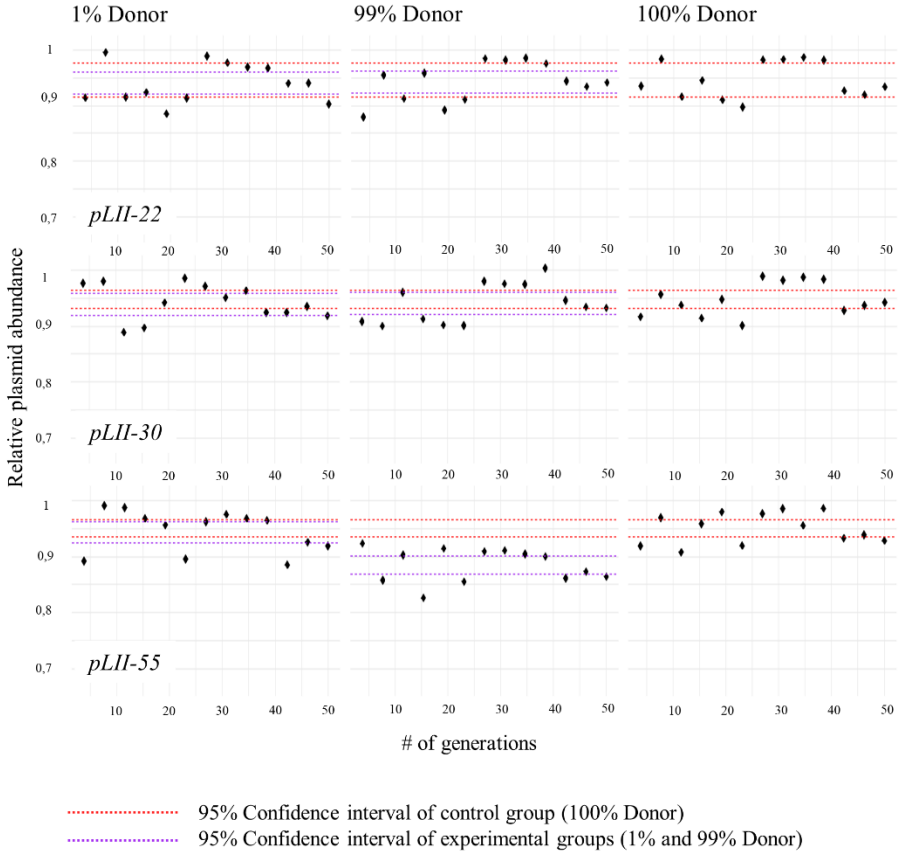
$$\bar{x} \pm Z\frac{s}{\sqrt{n}} \qquad \textit{(Equation 3.1)}$$

The $\bar{x}$ represents the mean value of the whole dataset, Z-score for the 95% CI (1,96), s is standard deviation, while n is the number of observations.

These intervals comprise a range of RPA values that would, with 95% certainty, contain the true value of the RPA for the given parallel. The overlapping CIs of an experimental parallel and its respective control parallel further indicate that the true RPA values of these parallels would be found within the same range.

CIs of all control parallels overlaps with the CIs of their respective experimental parallel 1 and experimental parallel 2, regardless of the competing pair (Figure 3.6., Table B12, Supplementary materials, part B), with the notable exception of the experimental parallel 2 (99% donor) of the DH5α rif^R/ DH5α rif^RpL-II-55 competing pair. As indicated by the T-test comparison of the RPA values between the experimental parallel 2 and the control parallel in this competing pair, CI of these parallels did not overlap.





	Student's T-test (p-value)		
Plasmid	1%Donor vs.100%Donor	99%Donor vs. 100%Donor	1%Donor vs. 99%Donor
p2798	0,73	0,35	0,62
p20426	0,32	0,24	0,78
pLII-22	0,70	0,78	0,92
pLII-30	0,40	0,58	0,72
pLII-50	0,69	0,00001	0,00013

Figure 3.6. Modified competitive experiment. In all mating pairs included plasmid-containing and plasmid-free DH5α rif^R. Points on the graphs represent individual RPA values for each sampling time point. The CIs for the control parallel are marked with red dashed line, while CIs of the experimental parallels are marked by purple dashed lines. Control parallel CI of each respective competing pair is presented in each experimental parallel showing the overlap between CI of the control and the experimental parallel.

3.6. Shufflon rearrangement analysis

The plasmids selected for this study were p17437, p20481 and pL-II whose shufflons are composed of three segments, A, B and C, each of them encoding for two different C' ends of the pilV protein (Figure 3.7). The shufflons of the p17437 and p20481 are completely identical, both with an insertion sequence, IS*Ecp1*, carrying the $bla_{\text{CTX-M-1}}$ gene (IS*Ecp1*+ $bla_{\text{CTX-M-1}}$), that disrupts the *B* ORF of the B segment (Mo et al., 2020). Compared to pL-II shufflon, the shufflon of p17437 and p20481 comprise an extra region of approximately 3kbp. The *B'* ORF of the B segment is not affected by the insertion and was considered to be fully functional. In this part of the study, the activity of the shufflon was investigated during different bacterial growth phases without any selective pressure. Furthermore, to determine whether the shufflon activity of p17437 and p20481 was affected by the host of the plasmids, plasmids were grown in their original hosts and in *E.coli* 2773(ST162) strain. The shufflon of pL-II was used as a reference control in this study, and its host for this experiment was DH5α rif^R *E.coli* strain.

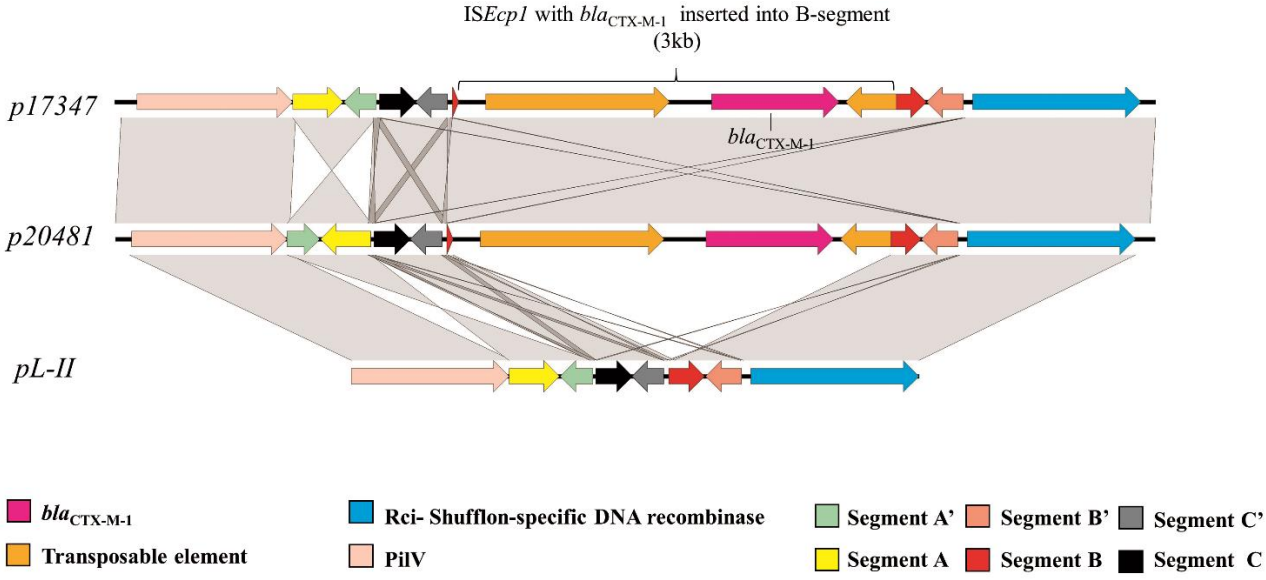


Figure 3.7. Aligned shufflon structures of p17437, p20481 and pL-II. Areas shaded grey indicate homologous regions, nucleotide identity threshold >95%.

The ON cultures of 2016-40-17347 (original host for p17437), p17437/2773(ST162), 2016-40-20481 (original host for p20481), p20481/2773(ST162) and pL-II/DH5α rif^R were diluted 1:100 with fresh LB broth. Samples were taken immediately after the dilution, at the early exponential phase (0,3 OD_{600}) at the middle of the exponential phase (0,5 OD_{600}) and after 24h of incubation, when the culture had reached the stationary phase. DNA was extracted from the samples, purified and the shufflon region amplified by PCR, followed by the nanopore sequencing preparation and sequencing for 48h.

The number of shufflon structure combinations was calculated, and two sequence databases were generated containing sequences of shufflon variants with and without inserted IS*Ecp1*+$bla_{\text{CTX-M-1}}$. Each database contained 79 sequences of shufflon combinations without duplication of segments. The

sequence database for the shufflon variants with the interrupted B segment also included the variants with the uninterrupted *B* ORF of the B segment. The number of shufflon variants of these plasmids increased to 145. To detect and quantify the structural variation of the selected shufflons, amplicon reads of the shufflon regions were aligned to possible shufflon structure sequences using BLAST+. The alignment between a read and a structure variant was considered correct if the start and the end of these sequences matched, if the alignment length matched the length of the corresponding sequence of the shufflon structure variant, and if the percentage of the identical matches exceeded 90%. These conditions would ensure that only the reads flanked by the part of the *pilV* and part of the *rci* would be considered correct. The percentage of the identical matches was expected to be below 95% due to the fact that the raw reads created by the nanopore sequencing have approximately 85% accuracy (Laver et al., 2015). The number of correct alignments was then counted for each of the possible shufflon structure variants, enabling the quantification of the shufflon analysis.

The samples that had less than 1000 correctly aligned reads were discarded. In general, the number of reads satisfying the given conditions spanned from approximately 1000 to 15 000 per sample.

3.6.1. Shufflon rearrangement during different phases of bacterial growth

Analysis of the reads of shufflon amplicons demonstrated that the number of variants of the pL-II shufflon, regardless of the sampling time point, was substantially higher than the number of p17437 and p20481 shufflon variants (Figure 3.8, Table B13, Supplementary materials, part B). The number of different variants ranged between 47 and 65 (total number of included variants were 79) for the pL-II shufflon depending on the sampling time point.

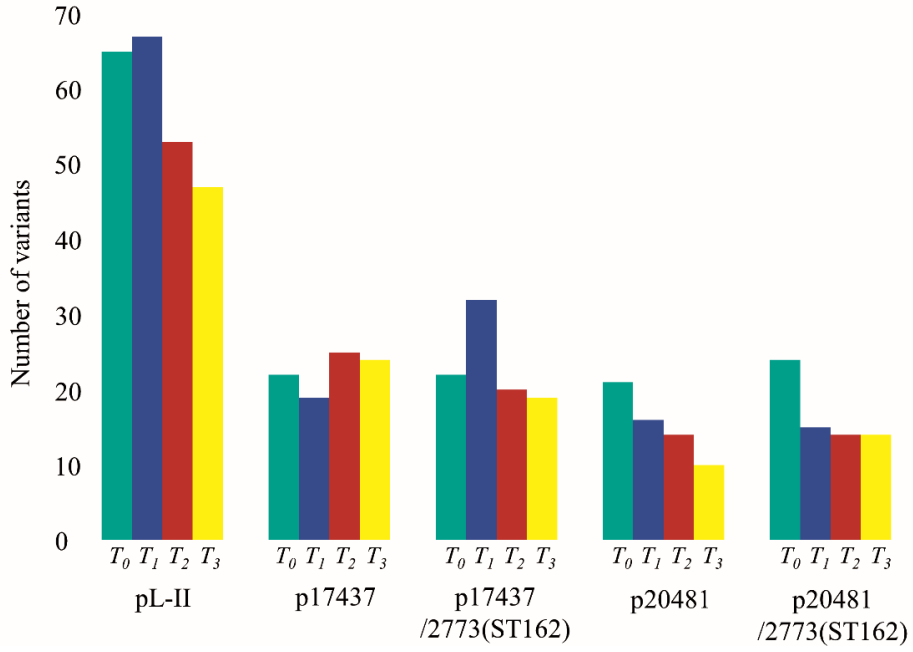


Figure 3.8. Number of shufflon variations of pL-II, p17437 (both in original host and in E.coli 2773(ST162)) and p20481 (both in original host and E.coli 2773(ST162)) detected at each sampling time point (T_0, T_1, T_2,T_3).

The total number of variants included for IS*Ecp1*+*bla*$_{CTX-M-1}$ interrupted plasmids was 145. The detected number of different shufflon variants of p17437 was between 19 and 32 over all time points, regardless of the plasmid host. The same was observed for p20481, whose shufflon generated between 10 and 24 variants at all time points, regardless of the plasmid host. Although it was expected that the number of variants should show greater variability with the increase of the number of bacterial cells during the growth of the culture, this was not observed. On the contrary, what was observed was either the reduction in the number of variants or a stable number of shufflon variants during bacterial growth.

3.6.2. Relative distribution of shufflon variants based on the number of reads during different bacterial growth phases

Regardless of the time point, the most prevalent reads from the pL-II shufflon contained either a complete set of segments or only one segment (deletion of two segments) (Figure 3.9, Table B14-B17, Supplementary materials, part B). While reads with the complete set of segments dominated in T_0 and T_1 (over 55% of all correct reads), the last two sampling time points were dominated by reads with a deletion of two segments (over 60% of all the reads). Relative abundance of the reads with a deletion of one segment varied between 3,4 % and 11,2 % over all time points.

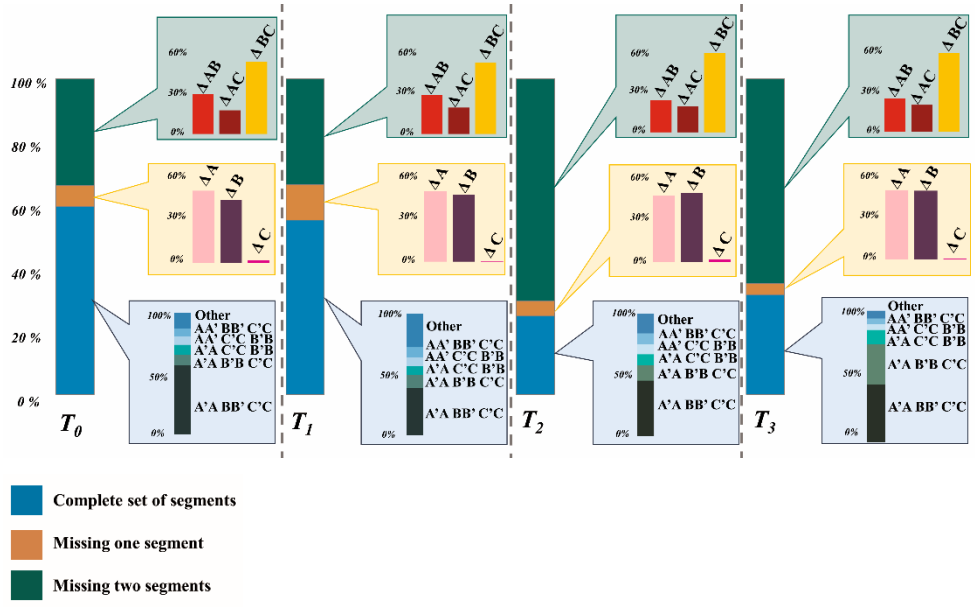


Figure 3.9. Structural variation analysis of pL-II shufflon. Main stack bar chart represents the distribution of reads containing a complete set of shufflon segments (blue), reads missing one segment (marked orange) and reads missing two segments (marked green). The distribution of each individual read group are presented in separate panels marked with the matching colour.

Of all the reads with a complete set of shufflon segments, five shufflon combinations (A'A BB' C'C , A'A B'B C'C, A'A C'C B'B, AA' C'C B'B and AA' BB' C'C) were the most prevalent over all time points (over 70% of all the reads with a complete set of segments). When it comes to reads with a

deletion of two segments, the most prevalent remaining segment was segment A over all sampling time points (over 53%). Among the reads with a deletion of one segment, the most often deleted segment was either segment A or segment B in all time points.

When p17437 was grown in its original host, the most prevalent number of reads contained, either a complete set of segments or only one segment (deletion of two segments) over all time points (Figure 3.10-A, Table B18-B21, Supplementary materials, part B). The relative abundance of reads with a complete set of segments dropped with each consecutive time point, from 96,7% at T_0 to 38,9% at T_3. Simultaneously, the relative abundance of reads with a deletion of two segments increased from 1,5% reaching 45% at the last time point. The relative abundance of reads with a deletion of one segment varied from 1,7% to 16,1% over all time points. Two shufflon combinations accounted for more than 97,5% of all the reads with the complete set of shufflon segments over all time points, namely A'A CC' BB' and AA' CC' BB'. All reads found to have one segment, had only segment A over all time points. The deletion of A and B, and A and C was not observed. Among the reads that had two shufflon segments, the B segment was found to be deleted most frequently (76,7-98% over all time points), while the C segment was never lost.

When p17437 was grown in *E.coli* 2773(ST162) strain (Figure 3.10-B, Table B22-B25 Supplementary materials, part B), a similar pattern was observed, although the relative abundance of reads with a complete set of segments was over 71% at all time points. The same two shufflon variants as when the p17437 was grown in the original host (A'A CC' BB' and AA' CC' BB') accounted for more than 96% of reads with the complete shufflon set of segments. The second most abundant group of reads was again the one with a deletion of two segments. More than 94% of all reads from this group had only segment A over all time points. Reads with a deleted B segment dominated the group of reads with a deletion of one segment.

The small number of reads with the B segment not interrupted by the IS*Ecp1*+$bla_{\text{CTX-M-1}}$ was detected at all time points regardless on the p17437 host. The relative abundance of these reads was between 0,038% and 0,21%, regardless of the plasmid host (Table B34, Supplementary materials, part B).

The relative abundance of reads from the p20481 shufflon with the complete set of shufflon segments was again the highest (Figure 3.10-C, Table B26-B29, Supplementary materials, part B, and Figure 3.10-D, Table B30-B33, Supplementary materials, part B). However, the relative abundance of these reads was much higher when the p20481 was grown in its original host (94.9% - 99,9% over all time points) compared to the relative abundance of these reads when p20481 was grown in 2773(ST162) (94,3% - 39,2% over all time point). The relative abundance of reads with a deletion of one segment when p20481 was grown in *E.coli* 2773(ST162) strain, was much higher than for any other plasmid (28,2% of reads at T_3).

The two most prevalent shufflon variants (>90%) found in p20481 shufflon reads with the complete set of segments were A'A BB' C'C and AA' BB' C'C regardless of the plasmid host and

sampling time point. More than 88% of the reads with a deletion of one segment lacked segment B regardless of the time point and the plasmid host. In the group of reads with a deletion of two segments, 100% of the reads had only the A segments regardless of the sampling time point and the plasmid host.

Between 0,04% and 0,3% of p20481 shufflon reads had the B segment without the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$, regardless of the time point and the host (Table B34, Supplementary materials, part B).

In addition, the shufflon variants with all deleted segments were never detected regardless of the plasmid, its host and sampling time point.

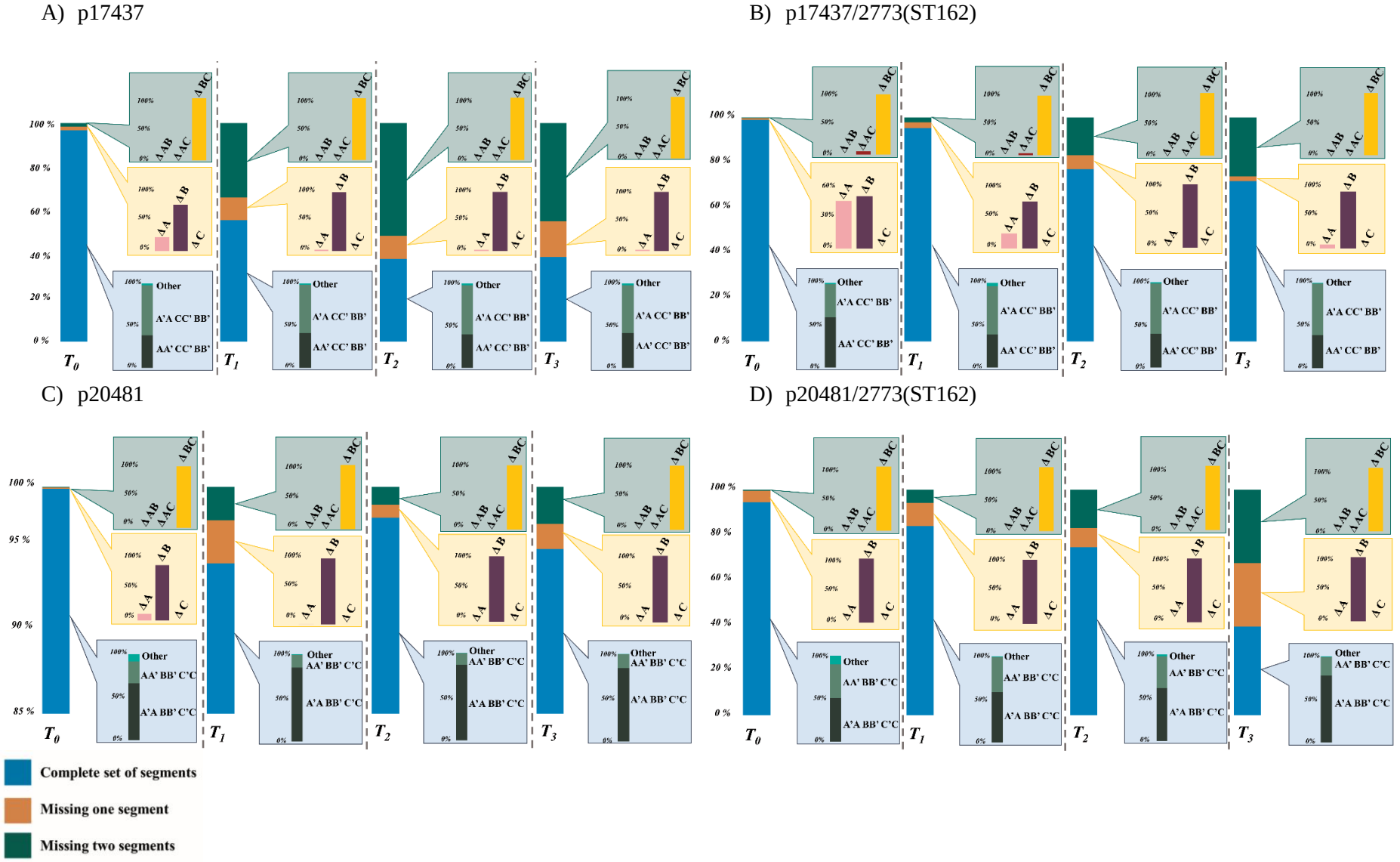


Figure 3.10. Structural variation analysis of shufflons of p17437 (A-p17437 in the original host, B-p17437 in the E. coli 2773(ST162) strain) and p20481 (C-p20481 in the original host, D-p20481 in the E. coli 2773 strain). Main stack bar charts represent the distribution of reads containing a complete set of shufflon segments (blue), reads missing one segment (marked orange) and reads missing two segments (marked green). The distribution of each individual read group are presented in separate panels marked with the matching colour.

Finally, the analysis of the 3' end of the *pilV* gene was conducted by calculating the relative abundance of the six different variants of the *pilV* gene (Figure 3.11, Table B35-B39, Supplementary materials, part B). Regardless of the plasmid, host strain, and sampling time points, the most prevalent *pilV* ORFs were *pilV-A* and *pilV-A'*. While the relative abundance of these two ORFs found among the reads from pL-II sample was between 67,4% and 76,2% in all time point, their relative abundance was much higher in p17437 and p20481 (over 90% regardless of the time point). The relative abundance of reads showing the *pilV-B* and *pilV-B'* configuration in the samples with p17437 and p20481 varied between 0% and 3,5% depending on the time point. The relative abundance of reads with these two variants varied between time points from 10,7% to 15,4% in pL-II.

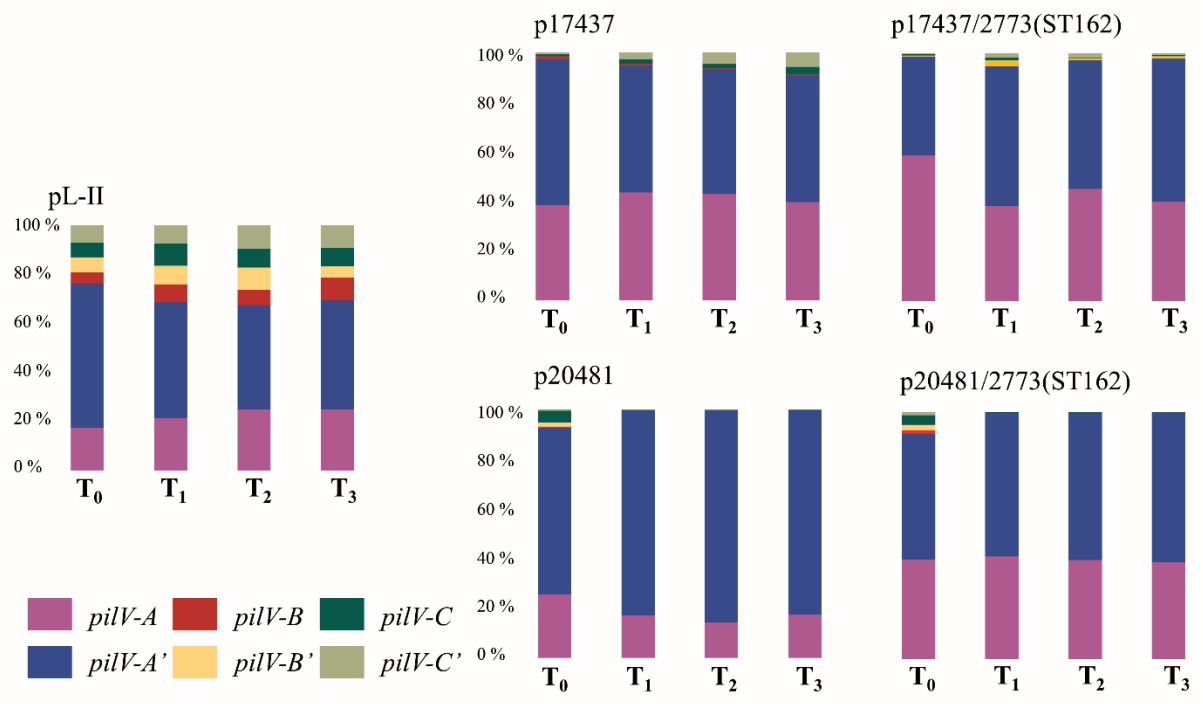


Figure 3.11. Relative abundance of the PilV ORFs found over different sampling time points for pL-II, p17437 (in the original host and 2773 E. coli strain) and p20481 (in the original host and 2773 E. coli strain).

4. Discussion

4.1. Plasmid fitness cost

One of the aims of this study was to examine whether the selected conjugative IncI1 plasmids induce the fitness cost in their new hosts. All selected plasmids were clinically relevant plasmids that encode either ESBL or AmpC. Their genetic organization was found to be similar, while the greatest differences were observed in their variable regions harbouring different accessory genes. Of special interest was p17437 which, compared to other plasmids, was much longer with a cointegrated part of an IncFIB plasmid. The IncFIB part contained several virulence genes (Mo et al., 2020).

The initial conjugation experiment generated successful transconjugants with the APEC strain and three QREC strains with different sequence types (ST162, ST602, ST453). Both APEC and three QREC strains were able to accept and maintain four out of five selected plasmids (p22638, p20481, p1248, p2798) during a pairwise mating with plasmid donors. Only one strain, 2773(ST162), was able to successfully accept and maintain p17437.

The ability of the selected plasmids to be transferred and replicated within different *E. coli* hosts varied. In addition, variable outcome of pairwise mating of the same plasmid donor with different QREC strains within the same ST group was also observed. None of the selected recipients harboured IncI1 plasmids and the conjugation should, therefore, not be inhibited by surface exclusion (Sakuma et al., 2013). This observation could suggest that individual interplay between the plasmid and the new recipient governs the outcome of donor/recipient mating (San Millan & Craig maclean, 2019). The inability of almost all included recipients to accept p17437 could further be evidence of this. It can be assumed that a single stranded copy of p17437 was transferred to a new host during conjugation, but that recircularization of the plasmid and synthesis of the complementary strand was not accomplished. On the other hand, the defence mechanisms in the host triggered by the presence of foreign DNA (such as RM or SOS) could have had various rates of success in preventing the stable maintenance of the newly acquired plasmid. However, due to a lack of concrete evidence, this remains speculation.

None of the plasmids were found to be successfully transferred to and replicated within any of the *K. pneumoniae* strains. An explanation for this observation could be found in the fact that conjugative IncI1 plasmids are not the predominant plasmid class in *Klebsiella* species (Wyres et al., 2020). However, this study encompassed only a limited number of *K. pneumoniae* strains, and further experiments are needed to investigate this observation.

The transfer frequency in selected mating pairs showed that only a small fraction of transconjugants were generated over a 24 hours period. At the endpoint, the number of available recipient cells was between four and five orders of magnitude higher than the number of transconjugants.

A previous study reported comparable IncI1 plasmid transfer-frequency results between *E. coli* - *E. coli* mating pairs during the same time span of the experiment, although in biofilms (Mo et al., 2017). A similar range of IncI1 plasmid transfer frequency during a liquid *E. coli – E. coli* mating was also previously reported (Benz et al., 2019). However, the correct interpretation of this value could be biased by the fact that the transconjugants also grows by clonal expansion, and the fact that transconjugants can also act as plasmid donors (Mo et al., 2017). Taking this into account, it is possible that the plasmid transfer rate directly from the donor to the recipient was lower than the calculated one.

The investigation of the possible burden that the plasmids could impose on their hosts was examined in single strain and competitive growth assay. During the single strain growth assay, the maximum growth rate of the transconjugants was rarely found to be significantly different from their respective recipient strains. However, the mean r_{max} of the transconjugants showed a tendency to be lower than mean r_{max} values of their respective recipient strains. In natural environments a variety of bacterial strains compete for nutrients and space and even a minimal difference in r_{max} is pivotal (Komori et al., 2018).

On the other hand, with an exception of the APEC transconjugants, most of the QREC transconjugants differed significantly from their respective recipients in K value. Greater K value of recipient pure cultures indicates that a much higher number of individuals was produced during the exponential phase than their respective transconjugants, and that the same number can be sustained in the stationary phase. On the other hand, all bacteria are considered r-strategists because they invest all the available energy to populate an area as quickly as possible (high maximum growth rate) (Song et al., 2017). Under the conditions used in this assay the only limiting factors were space and nutrients. However, plasmid maintenance requires energy (San Millan & Craig maclean, 2019). An assumption was made that, while the recipients were investing all the energy into their growth, transconjugants had to use an additional amount of energy for plasmid maintenance. Consequently, an exhaustion of energy resources would force the growing culture into the premature stationary phase (Jaishankar & Srivastava, 2017). However, further research is needed to confirm that the transconjugants are metabolically more active and that they need extra energy for the plasmid maintenance. In addition, the plasmid-caused impact on the K-value of its host was not the same in different transconjugants. Individual compatibility of the host and the plasmid could explain variations in K value in different transconjugants stemming from the same recipient.

Interestingly, while most QREC transconjugants exhibited lower K values than their respective recipients, the K values of APEC transconjugants were not significantly different from K values of their recipient. Furthermore, mean K values of transconjugants and the recipient showed minor variation. In addition, the APEC strain exhibited a significantly lower r_{max} than QREC recipient strains. The APEC possesses different virulence factors that enable their attachment to the host tissue, their survival in the host fluids, and evasion of the host immune system. The expression of virulence genes is under strict control (Barbieri et al., 2017; Collingwood et al., 2014). On the other hand, QREC strains included in

this study, had chromosomally encoded quinolone resistance and no virulence genes. They could be considered originally commensal *E. coli* strains, and their developed mutation the result of frequent exposure to antibiotic selective pressure, as it was previously outlined (de Lastours et al., 2013). However, as these isolates originated from Norwegian broilers in which production quinolones are not used, the exact cause of the chromosomal mutation is unknown. Nonetheless, the differences between APEC and QREC could explain the difference in the r_{max} between these strains. The additional virulence genes could have already imposed a fitness cost in the APEC strain leading to its reduced r_{max}. However, it is expected that most of the virulence genes are not constitutively expressed, thus reducing the impact on the host's fitness (Barbieri et al., 2017). It was also noted in an earlier study that the virulent *E. coli* strains have a reduced competitive fitness compared to its avirulent counterpart (Kitamoto et al., 2016). Consequently, it can be assumed that the APEC that is already adapted to the "extra" virulence genes and their controlled expression, could tackle the plasmid burden better than QREC strains. However, with no available genome sequence for the APEC strain from this study, and no evidence on how the controlled expression of the virulence genes could minimize the plasmid fitness cost, this explanation is merely a speculation. On the other hand, no conclusion could be drawn based on only one APEC strain examined in this study.

The competitive growth assay was expected to be more sensitive, and the differences observed in single strain growth assay to be more pronounced. For this experiment, three competing pairs of transconjugant/recipient were selected. As expected, the greatest reduction in competitive fitness was observed in p17437/2773(ST162) compared to its recipient. The p22638/2773(ST162) also exhibited competitive fitness reduction compared to the same recipient, however this reduction was lower than in p17437/2773(ST162). On the other hand, the competitive fitness reduction was minor when APEC was the carrier of p22638. These results are in compliance with the results from the single strain growth assay. Although both p17437/2773(ST162) and p22638/2773(ST162) had comparable mean r_{max} and K values, the mean r_{max} value of p22638/2773(ST162) was not significantly lower than the recipient r_{max}. As was previously mentioned, the strain with an even slightly lower r_{max} is under risk of being outcompeted during competitive growth. This could have led to a much greater competitive fitness reduction of p17437/2773(ST162) than of p22638/2773(ST162). A plausible explanation is that p17437 imposes a much greater burden than p22638 due to its size, additional virulence factors and two replicons.

Although a limited number of competing pairs was investigated here, the existence of a plasmid-imposed fitness cost was evident. The toll bacteria must pay for the increased chance of survival under conditions that select for plasmid encoded genes, is a reduction in their competitive fitness in conditions without selective pressure. However, taking also into account results from single strain growth experiments, the fitness reduction caused by different IncI1 plasmids in the same host, and vice versa, varied. This observation could be explained by the individual plasmid-host compatibility. A previous study reported that the impact on the host's fitness by IncN and IncP1 plasmids differed greatly

even when plasmids shared closely related backbone (Humphrey et al., 2012). Furthermore, a significant variation of fitness impact of the same plasmid in different hosts was also noted. These findings further suggest that individual "genetic compatibility" between a plasmid and its host plays a major role in determining to what extent the host's fitness would be affected. Thereby, the magnitude of plasmid-imposed fitness cost cannot be assessed from the observations of other plasmid-host combinations (Humphrey et al., 2012).

Further studies encompassing a wider variety of both commensal and pathogenic strains are needed to confirm these observations. As reported in previous studies that the compensatory mutations could be responsible for the fitness cost reduction (Carroll & Wong, 2018), future studies should also focus on the identifying of these mutations with an emphasis on identifying a possible plasmid- and/or host-dependant pattern of compensatory mutations.

4.2. Plasmid stability

The stability of selected plasmids was investigated during 50 bacterial generations long competitive growth of DH5α rif^R transconjugants with its plasmid free counterpart. The results of the experiment indicate that the number of plasmids in competing mixtures where the initial mixing ratio of p^- DH5α rif^R / p^+ DH5α rif^R was 100:1, have reached the same numbers as if it was a pure culture of the p^+ DH5α rif^R strain within the first 24 hours. A gradual increase in the plasmid numbers per sample with every successive sampling time point was expected due to the results from the previous study (Hagbø et al., 2019). Both plasmid encoded active partitioning, addiction system, and conjugation would lead to the increase in numbers of plasmid-carrying cells as well as clonal expansion. However, the instantaneous increase in the number of plasmids in all competing pairs could indicate that after the first 24 hours all the cells were plasmid-carriers which was not expected. A simple explanation would be that the plasmid-carrying cells had a much higher growth rate than the plasmid-free cells, leading to the complete overgrowth of plasmid-carrying cells. However, the results of the single strain growth assay counters this by providing the evidence that none of the plasmid-carrying cells had a r_{max} value higher than their plasmid-free counterpart. With a higher initial number of cells and a significantly higher r_{max}, plasmid-free cells would hardly be outcompeted by the plasmid-carrying cells.

Previously detected massive overshooting of plasmid genes right upon entry of the plasmid into a new host could elucidate the high number of plasmids (San Millan & Craig maclean, 2019). As plasmid encoded repressors are still not produced, the conjugation machinery could be expressed at higher rates. This would stall the clonal expansion of the new plasmid-carrying cell but would also promote the conjugational transfer of the plasmid to a new suitable cell (San Millan & Craig maclean, 2019). However, the plasmid gene repressors are not exclusively encoded by the plasmid. A chromosomally-encoded histone-like proteins are found to repress the expensive plasmid encoded processes such as

conjugation (Aznar et al., 2013). Consequently, overshooting of expensive plasmid encoded genes is highly unlikely to be the only factor causing this observation.

The most plausible explanation is that the high rates of plasmid transfer between the plasmid-carrying and plasmid-free cell were caused by the bacterial strain used in the experiment. The DH5α rif^R strain is a laboratory strain with a several crucial mutations which makes it an outstanding plasmid acceptor (Taylor et al., 1993). A mutation-inactivated endonuclease prevents the degradation of the plasmid upon entry and could further lead to an assumption that every conjugation attempt between a donor and a recipient cell resulted in a successful plasmid transfer.

Further investigation of plasmid stability in wildtype bacterial strains, rather than in laboratory *E.coli* strain, known to be an excellent plasmid acceptor, would be more beneficial. The rate of the plasmid transfer between wildtype bacteria would be much lower. When a selected QREC transconjugant was grown together with its plasmid free counterpart with initial mixing ratio 1:1 in the previous experiment, the number of transconjugants had a tendency to be lower with every consecutive sampling point. Mixed in a 1:100 ratio, the transconjugant will probably be facing the extinction.

The maintenance and stability of the conjugative IncI1 plasmids in bacterial cultures depend on both the benefits they provide to the host and the fitness cost they impose. It appears that the balance of plasmid benefits and plasmid-imposed costs are both plasmid- and strain-dependent. However, this balance is also affected by environmental conditions (Platt et al., 2012). The strains that are able to alleviate the fitness cost inflicted by the plasmid will thrive under both conditions with no selective pressure, and conditions demanding plasmid-encoded genes. These strains could be considered a strong plasmid source (Platt et al., 2012). On the other hand, environments with fluctuating conditions with frequently imposed selective pressure that promote the growth of plasmid carrying strains would be suitable for strains in which the plasmid had reduced the host's fitness. The same strain would not have any benefits from the plasmid in conditions with no selective pressure and could be overgrown by other strains. With no suitable recipient in the close proximity, both plasmid and its host may become extinct from the environment. Under these conditions, these strains could be considered strong plasmid sinks (Platt et al., 2012).

4.2.1. Methodological considerations

A method-introduced bias could also have affected the results to a certain extent. First, qPCR that targets plasmid sequences and 16s rRNA sequences, does not differentiate between the dead and the living cells. Second, the estimate of the number of plasmid copies and number of 16s rRNA gene copies per sample is based on standard curves made by using the pure amplified regions of the respective DNA fragments. As the sample is not composed of only pure primer-targeted DNA fragments, the amplification efficiency during the qPCR of the samples would be lower than calculated with the standard curves (Brankatschk et al., 2012). This would further lead to the overestimation of the

calculated numbers of both plasmid copies and 16s rRNA copies per sample. However, as estimates of both plasmid copies and 16s rRNA copies would presumably be overestimated in the same manner, their ratio would not be greatly affected.

It could be speculated that the results would have been improved if monocopy gene per bacterial genome was used to estimate the number of cells instead of the 16s rRNA (San Millan et al., 2014).

Although the attempt to quantify the results can be seen as biased, qualitatively speaking the results provide evidence that the plasmid was present in all mating cultures throughout the experiment, even when the plasmid-carrying cells were initially present in 100x lower numbers than their plasmid-free counterparts.

4.3. Shufflon rearrangement

The main objective of the shufflon rearrangement analysis was to investigate whether the shufflon interrupted by the insertion sequence would be able to generate as many variants as the non-interrupted one. The previous study found that shufflon rearranges constantly during different bacterial growth phases, although some variants are detected more often than the others. The assumption that the shufflon rearrangement is hampered by the insertion sequence was confirmed. The non-interrupted shufflon generated twice as many variants. As in the previous study (Brouwer et al., 2019), it was expected that the number of variants would increase with every new sampling time point due to the increasing number of cells in the sample. This would also imply that the shufflon rearranges constantly. However, the number of variants was either stable after the first time point or it showed a tendency to be reduced towards the last sampling time point. It is possible that in the absence of conjugation and with an increasing number of cells in a sample the rearrangement of the shufflon could be reduced, as it would represent a waste of energy. The difference in the results between this study and study of Brouwer et al. (2019) could be caused by the different methodology used to determine the number of variants in a sample. Brouwer et al. (2019) rarified the number of reads from each sample and included rare shufflon variants with a segment duplication which were omitted here.

Interestingly, each shufflon had a preferred configuration that dominated the variants with a complete set of segments regardless of the plasmid host and timepoint. While the uninterrupted shufflon had five different preferred variants, the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ interrupted shufflons had two. This finding also confirms the hypothesis of the non-random shufflon rearrangement. Another interesting find was the increasing relative abundance of the reads with the deletion of two segments throughout bacterial growth in both shufflon types. In addition, more than 94% of reads from the interrupted shufflons with the deletion of two segments had a deletion of segment B and C regardless of its host, and time point. Likewise, the B segment was also found to be most often deleted in reads with a deletion of one segment. As the B segment is the interrupted one, it is possible that the Rci found it more difficult to invert the B

segment than the other segments, and that the segment was lost during the rearrangement. However, the Rci is a tyrosine recombinase and its mode of action never involves double-stranded breaks while a segment is being inverted (Watson, 2004). On the other hand, a possible incomplete inversion would still introduce single strand breaks and could potentially lead to deletion. Further investigation of Rci mediated inversion of long segments such as the IS*Ecp1*+$bla_{\text{CTX-M-1}}$ interrupted B segment could give a better explanation to what caused its deletion.

Nevertheless, without the B segment, the plasmids lose the $bla_{\text{CTX-M-1}}$ and becomes unnecessary to its host under the conditions demanding for its presence (Carroll & Wong, 2018; San Millan & Craig maclean, 2019). However, the deletion of this segment could decrease plasmid induced fitness cost on its host, since ARG could also inflict a fitness cost in the ARG-carrying strain (Hernando-Amado et al., 2017). In addition, the loss of the IS*Ecp1*+$bla_{\text{CTX-M-1}}$ insertion sequence alone was also detected although in a minor number of reads at all time points, regardless of the plasmid and its host. It could be speculated that the insertion sequence has been moved to the chromosome rendering the plasmid a useless parasite (Carattoli et al., 2018).

This deletion of the segments would probably not be registered that often during the mating experiment, since the plasmid would lose the ability to express additional variants for PilV adhesin-like protein, further reducing the plasmid's ability to be transferred to a new host (Brouwer et al., 2019). Further research on the subject could reveal if the deletion of shufflon segments is the product of the random error introduced by the Rci or whether it is the product of the experimental conditions. Additionally, it should also be investigated whether the truncated shufflons inhibits the plasmid transfer to different host types with different LPS structures.

The analysis of different 3' ends of the *pilV* gene showed that A is the most preferable segment to complete the *pilV* ORF. This was observed in both interrupted and uninterrupted shufflons, regardless of the host and the sampling time point. The observation that segment A is the most favourable indicates its possible importance in conjugation. The same results were obtained in previous studies of the shufflon rearrangements of IncI1 and IncI2 shufflons (Brouwer et al., 2019; Sekizuka et al., 2017). However, evidence of whether this variant is important in conjugation or whether it is generated more often by pure chance, is missing. On the other hand, although these ORF variants were found to be dominant, there is no confirmation that they are also more often transcribed or produced.

4.3.1. Methodological considerations

The presented results have to be taken with caution, due to the possible bias each step of the process could have introduced. The number of reads aligned correctly to the shufflon sequence variants used in an attempt to quantify the shufflon structural variation, represents only a minor fraction of reads obtained from each sample. The true number of plasmids, and thus shufflons, in each sample does not correspond to the number of correct reads due to variable DNA extraction rates from each sample, and

the two PCR amplification steps prior to sequencing. Furthermore, the shufflon regions in the sample with the deletion of the B segment with the cointegrated IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ are shorter by approximately 3kbp than the complete shufflon regions. In the first amplification reaction that was adjusted to amplify the 5,5kbp region, the shorter shufflon variants would probably be amplified more frequently than the longer shufflon variants. This would lead to overrepresentation of the shorter amplicons and, further, possibly a higher number of their reads. Bias could have been introduced during the sequencing step. The accuracy of MinIon sequencing is not 100% (Morisse et al., 2018), which also led to a variable number of correct reads per sample. This fact calls into question whether comparation of number of reads between samples is plausible. Thereby, it was found suitable to compare the distribution of the shufflon variants and discuss their relative abundances. However, a normalization of the number of reads across the samples could have improved the results.

Nonetheless, even if the attempt at quantification of shufflon structural variation by the method used here could be questionable, if taken only qualitatively, these results provided new insights. First, a limited number of variants was observed in IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ interrupted shufflons compared to the number of variants from the non-interrupted shufflon. Second, the deletion of shufflon segments was confirmed, and in most of the samples with the plasmids harbouring the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ interrupted shufflon, the only deleted segment was either B alone or B and C together. And third, variants with a "clean" B segment, free from the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$, was also detected.

5. Conclusion and future perspectives

The IncI1 plasmids examined in this study were found to impose a fitness cost in their new hosts. The results indicate that the plasmid inflicted fitness cost is greater in QREC strains than in the APEC strain, pointing to the conclusion that the magnitude of impact is not exclusively linked to plasmids. The relevance of the host's genetic background and its ability to cope with the newly acquired plasmid is yet to be determined in future research.

The long-lasting persistence of selected IncI1 plasmids has been confirmed, although only in a laboratory *E. coli* strain. An assessment of the plasmid stability in wildtype *E. coli* strains in future studies is required.

The results presented here call for further examination of the individual interplay between a plasmid, its host and the environment, with special emphasis on uncovering factors that could alleviate the fitness cost inflicted by the plasmid.

Shufflon rearrangement analysis confirmed that generation of shufflon variants is not a random process in both interrupted and non-interrupted shufflons. Both types showed the predominance of certain plasmid-specific variants throughout the growth of the monoclonal plasmid carrying culture. In addition, in both shufflon types, the *pilV-A'* and *pilV-A* were the most abundant variants of the complete *pilV* ORF. However, there is no evidence that these variants of *pilV* ORFs are actually transcribed more often than the others.

The study also confirmed that the rearrangement of the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ interrupted shufflon was reduced. However, there is no evidence that this reduced rearrangement activity inhibits the conjugation with certain recipients.

The most interesting finding was that the truncated forms of interrupted shufflons suffered most often from a deletion of B segment alone or in combination with the C segment. Although found in a small number of reads, the excision of the IS*Ecp*+*bla*$_{\text{CTX-M-1}}$ from the shufflon region was also confirmed. The generated truncated shufflon would probably induce less fitness cost to its host. However, without knowledge of rate of loss of either IS*Ecp*+*bla*$_{\text{CTX-M-1}}$ interrupted B segment or the IS*Ecp1*+*bla*$_{\text{CTX-M-1}}$ alone, it is impossible to predict to what extent the plasmid fitness cost is reduced.

Further studies should investigate: 1) the importance of the A segment and whether its deletion leads to the reduced plasmid-transfer success to different hosts; 2) whether the pattern of observed shufflon variants differs during conjugation, and 3) the expression rate of the Rci and, if possible, its activity during monoclonal growth and during mating. These findings would be of great value for understanding the shufflon and its possible importance during the bacterial mating. Furthermore, the obtained knowledge could be applied in the development of strategies to target the shufflon and its rearrangement activity in order to prevent conjugation, as both pilV and shufflon regions are considered to be crucial during the selection of a new suitable host for the plasmid.

6. References

Alanazi, M. (2016). *Understanding the fitness burden of ESBL plasmids in Avian Pathogenic Escherichia Coli (APEC)*: University of Surrey.

Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J. & Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol*, 8 (4): 251-9. doi: 10.1038/nrmicro2312.

Argudin, M. A., Deplano, A., Meghraoui, A., Dodemont, M., Heinrichs, A., Denis, O., Nonhoff, C. & Roisin, S. (2017). Bacteria from Animals as a Pool of Antimicrobial Resistance Genes. *Antibiotics (Basel)*, 6 (2). doi: 10.3390/antibiotics6020012.

Asano, K., Hama, C., Inoue, S.-i., Moriwaki, H. & Mizobuchi, K. (1999). The Plasmid ColIb-P9 Antisense Inc RNA Controls Expression of the RepZ Replication Protein and Its Positive Regulator repYwith Different Mechanisms. *Journal of Biological Chemistry*, 274 (25): 17924-17933.

Aznar, S., Paytubi, S. & Juarez, A. (2013). The Hha protein facilitates incorporation of horizontally acquired DNA in enteric bacteria. *Microbiology*, 159 (Pt_3): 545-554.

Barbieri, N. L., Vande Vorde, J. A., Baker, A. R., Horn, F., Li, G., Logue, C. M. & Nolan, L. K. (2017). FNR Regulates the Expression of Important Virulence Factors Contributing to the Pathogenicity of Avian Pathogenic Escherichia coli. *Front Cell Infect Microbiol*, 7: 265. doi: 10.3389/fcimb.2017.00265.

Benz, F., Huisman, J. S., Bakkeren, E., Herter, J. A., Stadler, T., Ackermann, M., Diard, M., Egli, A., Hall, A. R. & Hardt, W.-D. (2019). Extended-spectrum beta-lactamase antibiotic resistance plasmids have diverse transfer rates and can be spread in the absence of selection. *bioRxiv*: 796243.

Bhattacharjee, M. K. (2016). *Chemistry of antibiotics and related drugs*, vol. 8: Springer.

Briñas, L., Zarazaga, M., Sáenz, Y., Ruiz-Larrea, F. & Torres, C. (2002). β-Lactamases in ampicillin-resistant Escherichia coli isolates from foods, humans, and healthy animals. *Antimicrobial agents and chemotherapy*, 46 (10): 3156-3163.

Brouwer, M. S., Tagg, K. A., Mevius, D. J., Iredell, J. R., Bossers, A., Smith, H. E. & Partridge, S. R. (2015). IncI shufflons: assembly issues in the next-generation sequencing era. *Plasmid*, 80: 111-117.

Brouwer, M. S., Jurburg, S. D., Harders, F., Kant, A., Mevius, D. J., Roberts, A. P. & Bossers, A. (2019). The shufflon of IncI1 plasmids is rearranged constantly during different growth conditions. *Plasmid*, 102: 51-55.

Buckner, M. M. C., Ciusa, M. L. & Piddock, L. J. V. (2018). Strategies to combat antimicrobial resistance: anti-plasmid and plasmid curing. *FEMS Microbiol Rev*, 42 (6): 781-804. doi: 10.1093/femsre/fuy031.

Bush, K. (2018). Past and Present Perspectives on β-Lactamases. *Antimicrobial Agents and Chemotherapy*, 62 (10). doi: 10.1128/aac.01076-18.

C Reygaert, W. (2018). An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiology*, 4 (3): 482-501. doi: 10.3934/microbiol.2018.3.482.

Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L. & Threlfall, E. J. (2005). Identification of plasmids by PCR-based replicon typing. *Journal of microbiological methods*, 63 (3): 219-228.

Carattoli, A., Villa, L., Fortini, D. & García-Fernández, A. (2018). Contemporary IncI1 plasmids involved in the transmission and spread of antimicrobial resistance in Enterobacteriaceae. *Plasmid*.

Carroll, A. C. & Wong, A. (2018). Plasmid persistence: costs, benefits, and the plasmid paradox. *Canadian journal of microbiology*, 64 (5): 293-304.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G. & Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics*, 21 (16): 3422-3423.

CDC. (2018). *Antibiotic Resistance: A Global Threat* Available at: https://www.cdc.gov/drugresistance/solutions-initiative/stories/ar-global-threat.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Ffeatures%2Fantibiotic-resistance-global%2Findex.html.

Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of theEscherichia coli phylogenetic group. *Appl. Environ. Microbiol.*, 66 (10): 4555-4558.

Collingwood, C., Kemmett, K., Williams, N. & Wigley, P. (2014). Is the Concept of Avian Pathogenic Escherichia coli as a Single Pathotype Fundamentally Flawed? *Front Vet Sci*, 1: 5. doi: 10.3389/fvets.2014.00005.

de Lastours, V., Bleibtreu, A., Chau, F., Burdet, C., Duval, X., Denamur, E. & Fantin, B. (2013). Quinolone-resistant Escherichia coli from the faecal microbiota of healthy volunteers after ciprofloxacin exposure are highly adapted to a commensal lifestyle. *Journal of Antimicrobial Chemotherapy*, 69 (3): 761-768. doi: 10.1093/jac/dkt422.

Deng, H. (2018). New Antibiotics: Where Are They? *Biomedical Journal of Scientific & Technical Research*, 10 (1). doi: 10.26717/bjstr.2018.10.001909.

Doss, J., Culbertson, K., Hahn, D., Camacho, J. & Barekzi, N. (2017). A review of phage therapy against bacterial pathogens of aquatic and terrestrial organisms. *Viruses*, 9 (3): 50.

Doumith, M., Day, M., Hope, R., Wain, J. & Woodford, N. (2012). Improved multiplex PCR strategy for rapid assignment of the four major Escherichia coli phylogenetic groups. *Journal of clinical microbiology*, 50 (9): 3108-3110.

Dudley, E. G., Abe, C., Ghigo, J. M., Latour-Lambert, P., Hormazabal, J. C. & Nataro, J. P. (2006). An IncI1 plasmid contributes to the adherence of the atypical enteroaggregative Escherichia coli strain C1096 to cultured cells and abiotic surfaces. *Infect Immun*, 74 (4): 2102-14. doi: 10.1128/IAI.74.4.2102-2114.2006.

Getino, M. & De la cruz, F. (2019). Natural and artificial strategies to control the conjugative transmission of plasmids. *Microbial Transmission*: 33-64.

Gyohda, A., Zhu, S., Furuya, N. & Komano, T. (2006). Asymmetry of shufflon-specific recombination sites in plasmid R64 inhibits recombination between direct sfx sequences. *Journal of Biological Chemistry*, 281 (30): 20772-20779.

Hagbø, M., Ravi, A., Angell, I. L., Sunde, M., Ludvigsen, J., Diep, D. B., Foley, S. L., Vento, M., Collado, M. C. & Perez-Martinez, G. (2019). Experimental support for multidrug resistance transfer potential in the preterm infant gut microbiota. *Pediatric research*: 1-9.

Hall, B. G., Acar, H., Nandipati, A. & Barlow, M. (2014). Growth rates made easy. *Molecular biology and evolution*, 31 (1): 232-238.

Harrison, E., Dytham, C., Hall, J. P., Guymer, D., Spiers, A. J., Paterson, S. & Brockhurst, M. A. (2016). Rapid compensatory evolution promotes the survival of conjugative plasmids. *Mobile genetic elements*, 6 (3): 2034-2039.

Hasman, H., Mevius, D., Veldman, K., Olesen, I. & Aarestrup, F. M. (2005). β-Lactamases among extended-spectrum β-lactamase (ESBL)-resistant Salmonella from poultry, poultry products and human patients in The Netherlands. *Journal of Antimicrobial Chemotherapy*, 56 (1): 115-121.

Hernando-Amado, S., Sanz-García, F., Blanco, P. & Martínez, J. L. (2017). Fitness costs associated with the acquisition of antibiotic resistance. *Essays in biochemistry*, 61 (1): 37-48.

Humphrey, B., Thomson, N. R., Thomas, C. M., Brooks, K., Sanders, M., Delsol, A. A., Roe, J. M., Bennett, P. M. & Enne, V. I. (2012). Fitness of Escherichia coli strains carrying expressed and partially silent IncN and IncP1 plasmids. *BMC microbiology*, 12 (1): 53.

Jain, M., Olsen, H. E., Paten, B. & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology*, 17 (1): 239.

Jain, S., Zweig, M., Peeters, E., Siewering, K., Hackett, K. T., Dillard, J. P. & van der Does, C. (2012). Characterization of the single stranded DNA binding protein SsbB encoded in the Gonoccocal Genetic Island. *PLoS One*, 7 (4): e35285. doi: 10.1371/journal.pone.0035285.
Jaishankar, J. & Srivastava, P. (2017). Molecular Basis of Stationary Phase Survival and Applications. *Frontiers in Microbiology*, 8 (2000). doi: 10.3389/fmicb.2017.02000.
Kaur, T., Al Abdallah, Q., Nafissi, N., Wettig, S., Funnell, B. E. & Slavcev, R. A. (2011). ParAB-mediated intermolecular association of plasmid P1 parS sites. *Virology*, 421 (2): 192-201. doi: 10.1016/j.virol.2011.09.027.
Kimura, A., Yuhara, S., Ohtsubo, Y., Nagata, Y. & Tsuda, M. (2012). Suppression of pleiotropic phenotypes of a Burkholderia multivorans fur mutant by oxyR mutation. *Microbiology*, 158 (5): 1284-1293.
Kitamoto, S., Nagao-Kitamoto, H., Kuffa, P. & Kamada, N. (2016). Regulation of virulence: the rise and fall of gastrointestinal pathogens. *Journal of gastroenterology*, 51 (3): 195-205.
Klümper, U., Riber, L., Dechesne, A., Sannazzarro, A., Hansen, L. H., Sørensen, S. J. & Smets, B. F. (2015). Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. *The ISME journal*, 9 (4): 934-945.
Komori, T., Shibai, A., Saito, H., Akeno, Y., Germond, A., Horinouchi, T., Furusawa, C. & Tsuru, S. (2018). Enhancement of K-strategy evolution in histidine utilization using a container with compartments. *Genes Cells*, 23 (10): 893-903. doi: 10.1111/gtc.12640.
Kroll, J., Klinter, S., Schneider, C., Voss, I. & Steinbuchel, A. (2010). Plasmid addiction systems: perspectives and applications in biotechnology. *Microb Biotechnol*, 3 (6): 634-57. doi: 10.1111/j.1751-7915.2010.00170.x.
Langmead, B. & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9 (4): 357.
Laver, T., Harrison, J., O'neill, P., Moore, K., Farbos, A., Paszkiewicz, K. & Studholme, D. J. (2015). Assessing the performance of the oxford nanopore technologies minion. *Biomolecular detection and quantification*, 3: 1-8.
Lawrence, J. G. (2005). Horizontal and vertical gene transfer: the life history of pathogens. *Contributions to microbiology*, 12: 255-271.
Lenski, R. E., Simpson, S. C. & Nguyen, T. T. (1994). Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. *Journal of bacteriology*, 176 (11): 3140-3147.
Lerminiaux, N. A. & Cameron, A. D. S. (2019). Horizontal transfer of antibiotic resistance genes in clinical environments. *Canadian Journal of Microbiology*, 65 (1): 34-44. doi: 10.1139/cjm-2018-0275.
Liu, Y.-T., Sau, S., Ma, C.-H., Kachroo, A. H., Rowley, P. A., Chang, K.-M., Fan, H.-F. & Jayaram, M. (2015). The partitioning and copy number control systems of the selfish yeast plasmid: an optimized molecular design for stable persistence in host cells. *Plasmids: Biology and Impact in Biotechnology and Discovery*: 325-347.
Loftie-Eaton, W., Bashford, K., Quinn, H., Dong, K., Millstein, J., Hunter, S., Thomason, M. K., Merrikh, H., Ponciano, J. M. & Top, E. M. (2017). Compensatory mutations improve general permissiveness to antibiotic resistance plasmids. *Nature ecology & evolution*, 1 (9): 1354-1363.
Lumen-lerning. Available at: https://courses.lumenlearning.com/microbiology/chapter/mechanisms-of-antibacterial-drugs/.
Majiduddin, F. K., Materon, I. C. & Palzkill, T. G. (2002). Molecular analysis of beta-lactamase structure and function. *Int J Med Microbiol*, 292 (2): 127-37. doi: 10.1078/1438-4221-00198.
McBirney, S. E., Trinh, K., Wong-Beringer, A. & Armani, A. M. (2016). Wavelength-normalized spectroscopic analysis of Staphylococcus aureus and Pseudomonas aeruginosa growth rates. *Biomedical optics express*, 7 (10): 4034-4042.
Mo, S. S., Slettemeas, J. S., Berg, E. S., Norstrom, M. & Sunde, M. (2016). Plasmid and Host Strain Characteristics of Escherichia coli Resistant to Extended-Spectrum Cephalosporins in the

Norwegian Broiler Production. *PLoS One*, 11 (4): e0154019. doi: 10.1371/journal.pone.0154019.
Mo, S. S., Sunde, M., Ilag, H. K., Langsrud, S. & Heir, E. (2017). Transfer Potential of Plasmids Conferring Extended-Spectrum-Cephalosporin Resistance in Escherichia coli from Poultry. *Appl Environ Microbiol*, 83 (12). doi: 10.1128/AEM.00654-17.
Mo, S. S., Telke, A. A., Osei, K. O., Sekse, C., Slettemeås, J. S., Urdahl, A. M., Ilag, H. K., Leangapichart, T. & Sunde, M. (2020). blaCTX–M–1/IncI1-Iγ Plasmids Circulating in Escherichia coli From Norwegian Broiler Production Are Related, but Distinguishable. *Frontiers in Microbiology*, 11: 333.
Morisse, P., Lecroq, T. & Lefebvre, A. (2018). Hybrid correction of highly noisy long reads using a variable-order de Bruijn graph. *Bioinformatics*, 34 (24): 4213-4222.
Münch, K., Münch, R., Biedendieck, R., Jahn, D. & Müller, J. (2019). Evolutionary model for the unequal segregation of high copy plasmids. *PLoS computational biology*, 15 (3): e1006724.
Nielsen, A. K. & Gerdes, K. (1995). Mechanism of post-segregational killing by hok-homologue pnd of plasmid R483: two translational control elements in the pnd mRNA. *Journal of molecular biology*, 249 (2): 270-282.
Norman, A., Hansen, L. H. & Sørensen, S. J. (2009). Conjugative plasmids: vessels of the communal gene pool. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364 (1527): 2275-2289.
Okonechnikov, K., Golosova, O., Fursov, M. & Team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28 (8): 1166-1167.
Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. (2018). Mobile genetic elements associated with antimicrobial resistance. *Clinical microbiology reviews*, 31 (4): e00088-17.
Pérez-Pérez, F. J. & Hanson, N. D. (2002). Detection of plasmid-mediated AmpC β-lactamase genes in clinical isolates by using multiplex PCR. *Journal of clinical microbiology*, 40 (6): 2153-2162.
Petrova, V., Chitteni-Pattu, S., Drees, J. C., Inman, R. B. & Cox, M. M. (2009). An SOS inhibitor that binds to free RecA protein: the PsiB protein. *Molecular cell*, 36 (1): 121-130.
Platt, T. G., Bever, J. D. & Fuqua, C. (2012). A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. *Proceedings of the Royal Society B: Biological Sciences*, 279 (1734): 1691-1699.
Ram, Y., Dellus-Gur, E., Bibi, M., Karkare, K., Obolski, U., Feldman, M. W., Cooper, T. F., Berman, J. & Hadany, L. (2019). Predicting microbial growth in a mixed culture from growth curve data. *Proceedings of the National Academy of Sciences*, 116 (29): 14698-14707.
Rockwood, L. L. (2015). *Introduction to population ecology*: John Wiley & Sons.
Rolfe, M. D., Rice, C. J., Lucchini, S., Pin, C., Thompson, A., Cameron, A. D., Alston, M., Stringer, M. F., Betts, R. P. & Baranyi, J. (2012). Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *Journal of bacteriology*, 194 (3): 686-701.
Rozwandowicz, M., Brouwer, M. S. M., Fischer, J., Wagenaar, J. A., Gonzalez-Zorn, B., Guerra, B., Mevius, D. J. & Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *Journal of Antimicrobial Chemotherapy*, 73 (5): 1121-1137. doi: 10.1093/jac/dkx488.
Sacha, P., Wieczorek, P., Hauschild, T., Zórawski, M., Olszańska, D. & Tryniszewska, E. (2008). Metallo-beta-lactamases of Pseudomonas aeruginosa--a novel mechanism resistance to beta-lactam antibiotics. *Folia Histochemica et cytobiologica*, 46 (2): 137-142.
Sakuma, T., Tazumi, S., Furuya, N. & Komano, T. (2013). ExcA proteins of IncI1 plasmid R64 and IncIγ plasmid R621a recognize different segments of their cognate TraY proteins in entry exclusion. *Plasmid*, 69 (2): 138-145.
San Millan, A., Peña-Miller, R., Toll-Riera, M., Halbert, Z., McLean, A., Cooper, B. & MacLean, R. (2014). Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. *Nature communications*, 5 (1): 1-11.

San Millan, A. (2018). Evolution of plasmid-mediated antibiotic resistance in the clinical context. *Trends in microbiology*, 26 (12): 978-985.
San Millan, A. & Craig maclean, R. (2019). Fitness costs of plasmids: a limit to plasmid transmission. *Microbial Transmission*: 65-79.
Schinner, F., Öhlinger, R., Kandeler, E. & Margesin, R. (2012). *Methods in soil biology*: Springer Science & Business Media.
Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30 (14): 2068-2069. doi: 10.1093/bioinformatics/btu153.
Seiffert, S. N., Hilty, M., Perreten, V. & Endimiani, A. (2013). Extended-spectrum cephalosporin-resistant Gram-negative organisms in livestock: an emerging problem for human health? *Drug Resist Updat*, 16 (1-2): 22-45. doi: 10.1016/j.drup.2012.12.001.
Sekizuka, T., Kawanishi, M., Ohnishi, M., Shima, A., Kato, K., Yamashita, A., Matsui, M., Suzuki, S. & Kuroda, M. (2017). Elucidation of quantitative structural diversity of remarkable rearrangement regions, shufflons, in IncI2 plasmids. *Scientific reports*, 7 (1): 1-10.
Song, H.-K., Song, W., Kim, M., Tripathi, B. M., Kim, H., Jablonski, P. & Adams, J. M. (2017). Bacterial strategies along nutrient and time gradients, revealed by metagenomic analysis of laboratory microcosms. *FEMS Microbiology Ecology*, 93 (10). doi: 10.1093/femsec/fix114.
Soucy, S. M., Huang, J. & Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nature Reviews Genetics*, 16 (8): 472-482.
Sprouffske, K. & Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC bioinformatics*, 17 (1): 172.
Sullivan, M. J., Petty, N. K. & Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics*, 27 (7): 1009-1010. doi: 10.1093/bioinformatics/btr039.
Sunde, M., Tharaldsen, H., Slettemeås, J. S., Norström, M., Carattoli, A. & Bjorland, J. (2009). Escherichia coli of animal origin in Norway contains a bla TEM-20-carrying plasmid closely related to bla TEM-20 and bla TEM-52 plasmids from other European countries. *Journal of antimicrobial chemotherapy*, 63 (1): 215-216.
Sunde, M., Simonsen, G. S., Slettemeås, J. S., Böckerman, I. & Norström, M. (2015). Integron, plasmid and host strain characteristics of Escherichia coli from humans and food included in the Norwegian antimicrobial resistance monitoring programs. *PLoS One*, 10 (6).
Taylor, R. G., Walker, D. C. & McInnes, R. (1993). E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic acids research*, 21 (7): 1677.
Thomas, C. M. (2000). *Horizontal gene pool: bacterial plasmids and gene spread*: CRC Press.
Tietgen, M., Semmler, T., Riedel-Christ, S., Kempf, V. A. J., Molinaro, A., Ewers, C. & Gottig, S. (2018). Impact of the colistin resistance gene mcr-1 on bacterial fitness. *Int J Antimicrob Agents*, 51 (4): 554-561. doi: 10.1016/j.ijantimicag.2017.11.011.
Tolmasky, M. E. & Alonso, J. C. (2015). *Plasmids: biology and impact in biotechnology and discovery*, vol. 46: John Wiley & Sons.
Valero-Rello, A., López-Sanz, M., Quevedo-Olmos, A., Sorokin, A. & Ayora, S. (2017). Molecular mechanisms that contribute to horizontal transfer of plasmids by the bacteriophage SPP1. *Frontiers in Microbiology*, 8: 1816.
Van Kregten, E., Westerdaal, N. & Willers, J. (1984). New, simple medium for selective recovery of Klebsiella pneumoniae and Klebsiella oxytoca from human feces. *Journal of clinical microbiology*, 20 (5): 936-941.
Voth, D. E., Broederdorf, L. J. & Graham, J. G. (2012). Bacterial Type IV secretion systems: versatile virulence machines. *Future microbiology*, 7 (2): 241-257.
Watson, J. D. (2004). *Molecular biology of the gene*, vol. 1: Pearson Education India.
Wick, R. (2018). *Porechop*. Available at: https://github.com/rrwick/Porechop.
Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS computational biology*, 13 (6): e1005595.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L. H., Karch, H., Reeves, P. R., Maiden, M. C. & Ochman, H. (2006). Sex and virulence in Escherichia coli: an evolutionary perspective. *Molecular microbiology*, 60 (5): 1136-1151.

Wyres, K. L., Lam, M. M. C. & Holt, K. E. (2020). Population genomics of Klebsiella pneumoniae. *Nat Rev Microbiol*, 18 (6): 344-359. doi: 10.1038/s41579-019-0315-1.

Yu, Y., Lee, C., Kim, J. & Hwang, S. (2005). Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnology and bioengineering*, 89 (6): 670-679.

Zhang, D., Zhao, Y., Feng, J., Hu, L., Jiang, X., Zhan, Z., Yang, H., Yang, W., Gao, B. & Wang, J. (2019). Replicon-based typing of IncI-complex plasmids, and comparative genomics analysis of IncIγ/K1 plasmids. *Frontiers in microbiology*, 10: 48.

7. Supplementary materials

7.1. Part A

Table A1. Antibiotic resistance profiles of the plasmid-recipient strains. MIC values are represented as mg/l

	MIC of selected antibiotics (mg/l)																
Plasmid-recipient strains	*SMX*	*TMP*	*CIP*	*TET*	*MEM*	*AZI*	*NAL*	*CTX*	*CHL*	*TGC*	*CAZ*	*CST*	*AMP*	*GEN*	*KAN*	*FLO*	*STR*
2011-01-1173 (APEC)	1024	16	1	32	ND	ND	**128**	**0,12**	64	ND	1	1	128	1	8	4	256
2014-01-2070 (QREC)	16	0,25	0,12	2	0,03	4	**128**	**0,2**	8	0,25	0,5	1	4	0,5	ND	ND	ND
2014-01-4539 (QREC)	8	0,25	0,25	2	0,03	4	**128**	**0,25**	8	0,25	0,5	1	4	0,25	ND	ND	ND
2014 -01-6043(QREC)	32	0,5	0,25	2	0,03	4	**128**	**0,25**	8	0,25	0,5	1	2	1	ND	ND	ND
2014-01-2145 (QREC)	2048	64	0,25	2	0,03	2	**128**	**0,25**	8	0,25	0,5	1	128	1	ND	ND	ND
2014-01-2773 (QREC)	16	0,25	0,25	64	0,03	4	**128**	**0,25**	8	0,25	0,5	1	4	1	ND	ND	ND
2014-01-7133 (QREC)	2048	64	8	128	0,03	4	**256**	**0,25**	256	0,25	0,5	1	128	0,5	ND	ND	ND
2009-01-3815 (QREC)	16	0,5	0,5	1	0,015	16	**256**	**0,12**	4	0,12	0,25	0,5	4	1	ND	ND	ND
2009-01-4618-2 (QREC)	4	0,12	0,25	1	0,015	8	**128**	**0,12**	4	0,12	0,25	0,5	4	0,25	ND	ND	ND
2011-01-3460-5 (QREC)	16	0,5	0,25	1	0,015	8	**128**	**0,12**	4	0,12	0,25	0,5	4	1	ND	ND	ND
2014-01-2069 (QREC)	2048	0,5	0,12	128	0,03	4	**256**	**0,25**	8	0,25	0,5	1	128	0,5	ND	ND	ND
2014-01-6924 (QREC)	2048	64	1	2	0,03	2	**64**	**0,25**	8	0,25	0,5	1	128	0,5	ND	ND	ND
2014-01-7234-1 (QREC)	8	0,25	1	2	0,03	2	**64**	**0,25**	8	0,25	0,5	1	128	0,5	ND	ND	ND
K. pneumoniae 2018-01-715	1024	32	0,5	64	0,03	16	8	**0,25**	8	0,5	0,5	1	64	0,5	ND	ND	ND

Abbreviations: SMX- sulfamethoxazole, TMP- trimethoprim, CIP- ciprofloxacin, TET- tetracycline, MEM- meropenem , AZI- azithromycin , NAL- nalidixic acid , CTX- cefotaxime , CHL- chloramphenicol , TGC- tigecycline, CAZ- Ceftazidime, CST- colistin , AMP- ampicillin, GEN- gentamycin, KAN- kanamycin, FLO- florfenicol , STR- streptomycin.

Table A2. Antibiotic resistance profiles of K. pneumoniae recipient strains.

	Antibiotic discs									
	AMP 10	FEP 30	CAZ 10	**CTX 5**	MEM 10	CN 10	AK	CIP 5	SXT 25	C 30
K. pneumoniae 152 CK	R	S	S	**S**	S	S	S	R	S	S
K. pneumoniae 27 PK	R	S	S	**S**	S	S	S	R	R	S

Abbreviations: AMP- ampicillin, FEP- Cefepime, *CAZ- Ceftazidime, CTX- cefotaxime, MEM – meropenem, CN-gentamycin, AK- amikacin, CIP- ciprofloxacin. SMX- sulfamethoxazole, C- chloramphenicol.*

Table A3. Antibiotic resistance profiles of the plasmid-donor strains from 2016. MIC values are represented as mg/l.*

	MIC of selected antibiotics (mg/l)																				
Plasmid-donor strains	*SMX*	*TMP*	*CIP*	*TET*	*MER*	*AZI*	*NAL*	*CTX*	*CHL*	*TGC*	*CAZ*	*CST*	*AMP*	*GEN*	*FOX*	*FEP*	*TOM*	*ETP*	*IPM*	*CTX+ Clavulanic acid*	*CAZ + Clavulanic acid*
2016-40-17437	1024	0,25	0,015	2	0,03	4	4	**64**	8	0,25	2	1	64	0,5	8	4	8	0,015	0,12	0,06	0,12
2016-40-20481	1024	0,25	0,015	2	0,03	4	4	**32**	8	0,25	2	1	64	0,5	8	8	8	0,015	0,12	0,06	0,25
2016-40-22638	1024	32	0,015	64	0,03	8	4	**64**	8	0,25	2	1	64	0,5	4	32	8	0,015	0,12	0,06	0,25
2016-40-20426																					

Abbreviations: *SMX- sulfamethoxazole, TMP- trimethoprim, CIP- ciprofloxacin, TET- tetracycline, MEM- meropenem, AZI- azithromycin, NAL- nalidixic acid, CTX- cefotaxime, CHL- chloramphenicol, TGC- tigecycline, CAZ- Ceftazidime, CST- colistin , AMP- ampicillin, GEN- gentamycin, FOX- cefoxitin, FEP- cefepime, ETP- ertapenem, IPN- imipenem.*

*Difference in antibiotics used to determine antibiotic resistance profiles is due to the year strains have been isolated and routine of antibiotic resistance profile determination.

Table A4. Antibiotic resistance profiles of the plasmid-donor strains from 2006 and 2012. MIC values are represented as mg/l.*

	MIC of selected antibiotics (mg/l)														
Plasmid-donor strains	*SMX*	*TMP*	*CIP*	*TET*	*NAL*	*CTX*	*CHL*	*CAZ*	*TIO*	*AMP*	*GEN*	*STR*	*FLO*	*KAN*	*CST*
2006-01-1248	16	0,5	0,03	2	4	**1**	8	ND	4	32	1	8	8	4	ND
2012-01-2798	32	0,5	0,06	64	4	**0,12**	4	0,5	ND	128	1	8	4	8	0,5

Abbreviations: *SMX- sulfamethoxazole, TMP- trimethoprim, CIP- ciprofloxacin, TET- tetracycline, NAL- nalidixic acid, CTX- cefotaxime, CHL- chloramphenicol, CAZ- Ceftazidime, TIO-ceftiofur, AMP- ampicillin, GEN- gentamycin, STR- streptomycin, FLO- florfenicol, KAN- kanamycin, CST- colistin.*

*Difference in antibiotics used to determine antibiotic resistance profiles is due to the year strains have been isolated and routine of antibiotic resistance profile determination.

Table A5. Antibiotic profiles of plasmid-donor strains from NMBU (Hagbø et al., 2019)

	Antibiotic discs (µg)									
	MEM 5	*TET 30*	*CHL 30*	*STR 10*	*TMP 5*	*SMX 25*	*AMP 10*	*CTX 5*	*CAZ 10*	*CPD 10*
DH5α rif^R pLII-22	S	S	S	R	R	R	R	S	S	S
DH5α rif^R pLII-30	S	S	R	R	R	R	R	R	R	R
DH5α rif^R pLII-55	S	S	S	R	S	S	R	R	R	R

Abbreviations: MEM- meropenem, TET-tetracycline, CHL- chloramphenicol, STR- streptomycin, TMP- trimethoprim, SMX- sulfamethoxazole, AMP- ampicillin, CTX- cefotaxime, CAZ-ceftazidime, CPD- cefpodoxime, RIF-rifampicin.

Table A6. Identified MGEs present in DH5α rif^R pLII-22, DH5α rif^R pLII-30, DH5α rif^R pLII-55(Hagbø et al., 2019) .

	MGEs			
Transconjugant	IncI1	Int1	P1	ColEI
DH5α rif^R pLII-22	Pos	Pos	Neg	Neg
DH5α rif^R pLII-30	Pos	Pos	Pos	Neg
DH5α rif^R pLII-55	Pos	Neg	Neg	Pos

Table A7. Genetic markers used for E. coli phylotyping, their respective primer sequences used to amplify targeted regions, and length of the amplified product.

Genetic marker	Primer sequence	Amplicon length (bp)
gadA	Forward 5' GATGAAATGGCGTTGGCGCAAG 3'	373
	Revers 5' GGCGGAAGTCCCAGACGATATCC 3'	
chuA	Forward 5' ATGATCATCGCGGCGTGCTG3'	281
	Revers 5' AAACGCGCTCGCGCCTAAT 3'	
yjaA	Forward 5' TGTTCGCGATCTTGAAAGCAAACGT 3'	216
	Revers 5' ACCTGTGACAAACCGCCCTCA 3'	
TSPE4.C2	Forward 5' GCGGGTGAGACAGAAACGCG 3'	152
	Revers 5' TTGTCGTGAGTTGCGAACCCG 3'	
Primer-mix was made by mixing 10µl of each primer (100 uM stock) with 20µl Milli-Q water. Final volume of each PCR reaction was 25µl, while final concentration of the reagents was as follows, 1x Qiagen Multiplex PCR mix and 0,2µM primer-mix. Reagents were suspended in 10µl milliQ water per reaction. Two µl of extracted DNA were added to each reaction mixture.		
The PCR protocol included 15min of initial denaturation and polymerase activation at 95ºC, followed by 30 cycles of denaturation for 30sec at 95ºC, annealing for 30sec at 60ºC, elongation for 30sec at 72ºC. The final elongation at the end of the last cycle was at 72ºC for 5min, followed by the indefinite hold of the PCR product at 8ºC. PCR was run on Sure cycler 8800.		

Table A8. Phylotype grouping based on the presence/absence of four phylogenetic markers.

Phylotype	*gadA*	*chuA*	*yjaA*	*TSPE4.C2*
A	+	-	+/-	-
B1	+	-	-	+
B2	+	+	+	+/-
D	+	+	-	+/-

Table A9. blaCTX-M, blaTEM, blaCMY-2 primers, their respective annealing temperature, and length of their amplicons.

Primes			***Length of the amplicon***	***Annealing temperature***
***bla*$_{\text{CTX-M}}$**	blaCTX.F	ATGTGCAGYACCAGTAARGTKATGGC	593bp	60ºC
	blaCTX.R	TGGGTRAARTARGTSACCAGAAYCAGCGG		
***bla*$_{\text{TEM}}$**	blaTEM.F	TTCTTGAAGACGAAAGGGC	1150bp	60ºC
	blaTEM.R	ACGCTCAGTGGAACGAAAAC		
***bla*$_{\text{CMY-2}}$**	CITM-F	TGGCCAGAACTGACAGGCAAA	462bp	66ºC
	CITM-R	TTTCTCCTGAACGTGGCTGGC		
Bla$_{\text{CTX-M}}$ / *bla*$_{\text{TEM}}$ PCR protocol	Final volume of each PCR reaction was 25µl, while final concentration of the reagents was as follows, 1x 10xPCR buffer, 0,2mM of each 10mM dNTP stock, 0,2µM of forward primer and reverse primer, 0,5U of 5U/ul Taq DNA polymerase. Reagents were suspended in 18,2 µl milliQ water. Two µl of extracted DNA were added to each reaction mixture.			
	The PCR protocol included 5 minutes of initial denaturation and polymerase activation at 95ºC, followed by 30 cycles of denaturation for 30sec at 95ºC, annealing for 30sec at 60ºC and elongation for 1minute at 72ºC. The final			

	elongation at the end of the last cycle was at 72ºC for 7min, followed by indefinite hold of PCR product at 8ºC. PCR was run on Sure cycler 8800.
bla$_{CMY-2}$	Final volume of each PCR reaction was 25 µl, while final concentration of the reagents was as follows, 1x Qiagen Multiplex PCR mix, 0,2 µM forward primer, 0,2µM reverse primer. Reagents were suspended in 10µl milliQ water per reaction. Two µl of extracted DNA were added to each reaction mixture.
	The PCR protocol included 15 minutes of initial denaturation and polymerase activation at 95ºC, followed by 25 cycles of denaturation for 30sec at 95ºC, annealing for 30sec at 66ºC, elongation for 1minute at 72ºC. The final elongation at the end of the last cycle at 72ºC for 10min, followed by indefinite hold of PCR product at 8ºC. PCR was run on Sure cycler 8800.

Table A10. Primers targeting fumC, their annealing temperature and length of the amplicon

Primes			*Length of the amplicon*	*Annealing temperature*
fumC	fumC-F	5'-TCCCGGCAGATAAGCTGTGG-3'	806bp	54ºC
	fumC-R	5'-TCACAGGTCGCCAGCGCTTC-3'		
Final volume of each PCR reaction was 50µl, while final concentration of the reagents was 1x PCR buffer (10x), 0,2 mM of each dNTP (dNTP mix stock 10mM), 0,4µM of forward and reverse primer, 2,5U of 5U/ul Taq DNA polymerase. Reagents were suspended in 40,5µl miliQ water per reaction. Three µl extracted DNA was added to the reaction mixture.				
PCR protocol included 5min of initial denaturation and polymerase activation at 95º C, followed by 30 cycles of denaturation for 1min at 95º C, annealing for 1min at 54ºC and elongation for 2min at 72º C. The final elongation at the end of the last cycle was at 72º C for 5min, followed by indefinite hold of PCR product at 8º C. PCR was run on Sure cycler 8800.				

Table A11. Primer pairs targeting 16s rRNA gene segment and IncI1, their sequences, annealing temperatures and amplicon length produced targeted by the primer pairs.

Primer target	*Primer sequence*	*Amplicon size (bp)*	*Annealing temperature(ºC)*
Bacterial 16s rRNA gene	Fw 5'- CCATACGGGRBGCASCAG - 3'	450	55
	Rev 5'- GGACTACYVGGGTATCTAAT - 3'		
IncI1 plasmid	Fw 5'- CGAAAGCCGGACGGCAGAA - 3'	141	60
	Rev 5' - TCGTCGTTCCGCCAAGTTCGT- 3'		
Final volume of each PCR reaction was 20 µl, while the Final concentrations of the reagents was 1x HOT FIREPol® EvaGreen qPCR supermix and 0.2 µM of both forward and reverse primer. The reagents were suspended in 14,2 µl of milliQ water. One µl of extracted DNA was added to each reaction mixture. qPCR was performed in a C1000 Touch™ Thermal Cycler (BioRad).			
PCR protocol included 15 min of initial denaturation and polymerase activation at 95ºC, followed by 40 cycles of denaturation for 30sec at 95ºC, annealing for 30sec, and elongation for 30sec at 72ºC. In addition, the melting temperature of the amplified product was also measured, thus providing quality control of the amplified segments.			

Table A12. PCR reaction protocol used for amplification of targeted sequences of 16s rRNA and IncI1.*

Final volume of each PCR reaction was 25µl, while final concentration of the reagents was 1x 5xHOT FIREPol bled MasterMix Ready to Load , 0,2 mM of forward and reverse primer, and 5µl of 0,1-10ng template DNA. Reagents were suspended in 14µl milliQ water per reaction.
PCR protocol included 15min of initial denaturation and polymerase activation at 95º C, followed by 35 cycles of denaturation for 30sec at 95º C, annealing for 30sec, and elongation for 30sec at 72º C. The final elongation at the end of the last cycle was at 72º C for 7min, followed by indefinite hold of PCR product at 8º C. PCR was run on 2720 Thermal Cycler.

**The primer sequences and their respective annealing temperatures are shown in Table A.11*

Table A13. PCR amplification of the sufflon region

Primes		*Length of the amplicon*	*Annealing temperature*
Nanopore_shufflon_Fw	5' *TTTCTGTTGGTGCTGATATTGC*-**ATGACAGAAGGGCGAGTTCA** 3'	Shufflon region of the pL-II : 2,5kbp	60ºC
Nanopore_shufflon_Rev	5' *ACTTGCCTGTCGCTCTATCTTC*-**GGTGCATTACGTTCCTGGTC** 3'	Shufflon region of the p17437, p20481 and p22638 : 5,5kbp	
Final volume of each PCR reaction was 50µl, while final concentration of the reagents was 0,4µM of forward and reverse primer, 1x LongAmp Taq 2x MasterMix, and <1ng template DNA. Reagents were suspended in miliQ water.			
-PCR protocol adapted for 2,5 kbp amplicon: 30sec of initial denaturation and polymerase activation at 94º C, followed by 30 cycles of denaturation for 30sec at 94º C, annealing for 30sec at 60ºC and elongation for 2,5min at 65º C. The final elongation at the end of the last cycle was at 65º C for 10min, followed by indefinite hold of PCR product at 8º C. -PCR protocol adapted for 5,5 kbp amplicon: 30sec of initial denaturation and polymerase activation at 94º C, followed by 30 cycles of denaturation for 30sec at 94º C, annealing for 30sec at 60ºC and elongation for 5min at 65º C. The final elongation at the end of the last cycle was at 65º C for 10min, followed by indefinite hold of PCR product at 8º C. Both types of PCR reactions were run on Sure cycler 8800.		

*Bold marked part of the primer sequence – shufflon targeting primer sequence; Italic marked part of the primer sequence – tailed primer sequence.

Table A14. barcoding PCR protocol.

Final volume of each PCR reaction was 50µl, containing 25µl of LongAmp Taq 2x MasterMix, 1µl of PCR barcode (one BC01-096 per sample) and 24µl of the previously amplified and purified PCR products.
-PCR protocol adapted for 2,5 kbp amplicon: 30sec of initial denaturation and polymerase activation at 94º C, followed by 15 cycles of denaturation for 30sec at 94º C, annealing for 15sec at 62ºC and elongation for 2,5min at 65º C. The final elongation at the end of the last cycle was at 65º C for 10min, followed by indefinite hold of PCR product at 8º C. -PCR protocol adapted for 5,5 kbp amplicon: 30sec of initial denaturation and polymerase activation at 94º C, followed by 30 cycles of denaturation for 30sec at 94º C, annealing for 15sec at 62ºC and elongation for 5min at 65º C. The final elongation at the end of the last cycle was at 65º C for 10min, followed by indefinite hold of PCR product at 8º C. Both types of PCR reactions were run on Sure cycler 8800.

7.2. Part B

Table B1. All attempted conjugation pairs, the medium the conjugation experiment was performed in, the selective medium used for transconjugant screening, and the transconjugant confirmation method used.

					Confirmation method			
Donor	Recipient	Liquid conjugation	Solid surface conjugation	Selective medium	Phylotyping	ESBL/AmpC gene confirmation	Alternative Conformation method	Result
2016-40-17437 (p17437)	→2011-01-1173 (APEC)	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l ; 2) MH agar with CTX 0,5mg/l and STR 100mg/l	NR	NR		-
	→2014-01-2070 (QREC/ST355)	-	-	MH agar with CTX 0,5mg/l and NAL 20mg/l	NR	NR		-
	→2014-01-4539 (QREC/ST355)	-	-		NR	NR		-
	→2014 -01-6043 (QREC/ST355)	-	-		NR	NR		-
	→2014-01-2145 (QREC/ST162)	-	-		NR	NR		-
	→2014-01-2773 (QREC/ST162)	+	NR		B2	bla_{CTX-M1}		OK
	→2014-01-7133 (QREC/ST162)	-	-		NR	NR		-
	→2009-01-3815 (QREC/ST602)	?	NR		B1	-		-
	→2009-01-4618-2 (QREC/ST602)	?	NR		B1	-		-
	→2011-01-3460-5 (QREC/ST602)	?	-		B1	-		-
	→2014-01-2069 (QREC/ST453)	?	NR		B1	-		-
	→2014-01-6924 (QREC/ST453)	?	-		B1	-		-
	→2014-01-7234-1 (QREC/ST453)	-	NR		B1	-		-
	→K. pneumoniae 2018-01-715	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l; 2) SCAI agar with cefotaxime 0,5mg/l	NR	NR		-
	→K. pneumoniae 152 CK	-	-		NR	NR		-
	→K. pneumoniae 27PK	-	-		NR	NR		-
2016-40-20481 (p20481)	→2011-01-1173 (APEC)	+	NR	MH agar with CTX 0,5mg/l and NAL 20mg/l	NR	bla_{CTX-M1}	fumC sequencing	OK
	→2014-01-2070 (QREC/ST355)	?	NR		B2	-		-
	→2014-01-4539 (QREC/ST355)	+	NR		ND	ND		
	→2014 -01-6043 (QREC/ST355)	+	NR		ND	ND		
	→2014-01-2145 (QREC/ST162)	+	NR		ND	ND		
	→2014-01-2773 (QREC/ST162)	+	NR		B1	bla_{CTX-M1}		OK
	→2014-01-7133 (QREC/ST162)	+	NR		ND	ND		
	→2009-01-3815 (QREC/ST602)	?	ND		B1	-		-
	→2009-01-4618-2 (QREC/ST602)	?	ND		B1	-		-
	→2011-01-3460-5 (QREC/ST602)	+	NR		B1	bla_{CTX-M1}		OK
	→2014-01-2069 (QREC/ST453)	-	ND		NR	NR		-
	→2014-01-6924 (QREC/ST453)	?	NR		B1	bla_{CTX-M1}		OK
	→2014-01-7234-1 (QREC/ST453)	?	ND		B1	-		-
	→K. pneumoniae 2018-01-715	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l 2) SCAI agar with cefotaxime 0,5mg/l	NR	NR		-
	→K. pneumoniae 152 CK	-	-		NR	NR		-
	→K. pneumoniae 27PK	-	-		NR	NR		-
2016-40-22638 (p22638)	→2011-01-1173 (APEC)	+	NR	MH agar with CTX 0,5mg/l and NAL 20mg/l	D	bla_{CTX-M1}		OK
	→2014-01-2070 (QREC/ST355)	?	NR		B2	-		-
	→2014-01-4539 (QREC/ST355)	+	NR		ND	ND		
	→2014 -01-6043 (QREC/ST355)	+	NR		ND	ND		
	→2014-01-2145 (QREC/ST162)	+	NR		ND	ND		
	→2014-01-2773 (QREC/ST162)	+	NR		B1	bla_{CTX-M1}		OK
	→2014-01-7133 (QREC/ST162)	+	NR		ND	ND		

	→2009-01-3815 (QREC/ST602)	+	NR		B1	bla_{CTX-M1}	OK
	→2009-01-4618-2 (QREC/ST602)	+	NR		B1	bla_{CTX-M1}	OK
	→2011-01-3460-5 (QREC/ST602)	+	NR		B1	bla_{CTX-M1}	OK
	→2014-01-2069 (QREC/ST453)	-	ND		ND	ND	-
	→2014-01-6924 (QREC/ST453)	+	NR		B1	bla_{CTX-M1}	OK
	→2014-01-7234-1 (QREC/ST453)	+	NR		B1	bla_{CTX-M1}	OK
	→K. pneumoniae 2018-01-715	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l 2) SCAI agar with cefotaxime 0,5mg/l	NR	NR	-
	→K. pneumoniae 152 CK	-	-		NR	NR	-
	→K. pneumoniae 27PK	-	-		NR	NR	-
2012-01-2798 (p2798)	→2011-01-1173 (APEC)	+	+	MH agar with CTX 0,5mg/l and NAL 20mg/l	D	bla_{CMY-2}	OK
	→2014-01-2070 (QREC/ST355)	-	-		NR	NR	-
	→2014-01-4539 (QREC/ST355)	-	-		NR	NR	-
	→2014 -01-6043 (QREC/ST355)	-	-		NR	NR	-
	→2014-01-2145 (QREC/ST162)	+	NR		ND	ND	
	→2014-01-2773 (QREC/ST162)	+	NR		B1	bla_{CMY-2}	OK
	→2014-01-7133 (QREC/ST162)	-	+		ND	ND	
	→2009-01-3815 (QREC/ST602)	+	NR		B1	bla_{CMY-2}	OK
	→2009-01-4618-2 (QREC/ST602)	+	NR		B1	bla_{CMY-2}	OK
	→2011-01-3460-5 (QREC/ST602)	+	NR		B1	bla_{CMY-2}	OK
	→2014-01-2069 (QREC/ST453)	-	ND		ND	ND	
	→2014-01-6924 (QREC/ST453)	+	NR		B1	bla_{CMY-2}	OK
	→2014-01-7234-1 (QREC/ST453)	+	NR		B1	bla_{CMY-2}	OK
	→K. pneumoniae 2018-01-715	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l 2) SCAI agar with cefotaxime 0,5mg/l	NR	NR	-
	→K. pneumoniae 152 CK	-	-		NR	NR	-
	→K. pneumoniae 27PK	-	-		NR	NR	-
2006-01-1248 (p1248)	→2011-01-1173 (APEC)	+	NR	MH agar with CTX 0,5mg/l and NAL 20mg/l	D	bla_{TEM-20}	OK
	→2014-01-2070 (QREC/ST355)	-	-		ND	ND	-
	→2014-01-4539 (QREC/ST355)	-	ND		ND	ND	
	→2014 -01-6043 (QREC/ST355)	-	ND		ND	ND	
	→2014-01-2145 (QREC/ST162)	-	+		ND	ND	
	→2014-01-2773 (QREC/ST162)	-	+		B1	bla_{TEM-20}	OK
	→2014-01-7133 (QREC/ST162)	+	NR		ND	ND	
	→2009-01-3815 (QREC/ST602)	+	NR		B1	bla_{TEM-20}	OK
	→2009-01-4618-2 (QREC/ST602)	+	NR		B1	bla_{TEM-20}	OK
	→2011-01-3460-5 (QREC/ST602)	+	NR		B1	bla_{TEM-20}	OK
	→2014-01-2069 (QREC/ST453)	-	ND		ND	ND	
	→2014-01-6924 (QREC/ST453)	+	NR		B1	bla_{TEM-20}	OK
	→2014-01-7234-1 (QREC/ST453)	?	NR		B1	-	-
	→K. pneumoniae 2018-01-715	-	-	1) MH agar with CTX 0,5mg/l and NAL 20mg/l 2) SCAI agar with cefotaxime 0,5mg/l	NR	NR	-
	→K. pneumoniae 152 CK	-	-		NR	NR	-
	→K. pneumoniae 27PK	-	-		NR	NR	-

Abbreviations: ?- colony morphology appearing more like mutant strains than transconjugants; "+" – detected transconjugants on a selective medium; "-" – non-detectable transconjugants on the selective medium or not confirmed transconjugants by the PCR; NR- not relevant, ND-not determined.

Table B2. Transfer frequencies (TF) when p20481 and p22638 plasmid original hosts were conjugated with APEC, and when p17437, p20481 and p22638 plasmid original hosts were conjugated with 2773(ST162) E.coli strain, in biological triplicates., the calculated mean TF and SD for each conjugating pair.

	TF of biological replicates				
Mating pairs	I	II	III	Mean TF	SD
2016-40-20481(p20481) → APEC	0,00015	0,00019	$9{,}79 \times 10^{-5}$	0,00015	$4{,}83 \times 10^{-5}$
2016-04-22638(p22638) →APEC	$5{,}25 \times 10^{-5}$	$8{,}8 \times 10^{-5}$	$6{,}72 \times 10^{-5}$	$6{,}92 \times 10^{-5}$	$1{,}78 \times 10^{-5}$
2016-40-17347(p17437)→2773(ST162)	$6{,}06 \times 10^{-5}$	0,00012	$7{,}93 \times 10^{-5}$	$8{,}57 \times 10^{-5}$	$2{,}88 \times 10^{-5}$
2016-40-20481(p20481)→2773(ST162)	$8{,}3 \times 10^{-5}$	$1{,}54 \times 10^{-5}$	$8{,}89 \times 10^{-5}$	$6{,}24 \times 10^{-5}$	$4{,}09 \times 10^{-5}$
2016-40-22638(p22638)→2773(ST162)	0,00016	$4{,}41 \times 10^{-5}$	0,00012	0,00011	$5{,}94 \times 10^{-5}$

Table B3. p-value of the T-test comparison of TFs of selected mating pairs.

Compared pairs	p-value
2016-40-20481(p20481) → APEC vs 2016-04-22638(p22638) →APEC	0,090
2016-40-17347(p17437)→2773(ST162) vs 2016-40-20481(p20481)→2773(ST162)	0,47
2016-40-17347(p17437)→2773(ST162) vs 2016-40-22638(p22638)→2773(ST162)	0,60
2016-40-20481(p20481)→2773(ST162) vs 2016-40-22638(p22638)→2773(ST162)	0,34
2016-40-20481(p20481) → APEC vs 2016-40-20481(p20481)→2773(ST162)	0,079
2016-04-22638(p22638) →APEC vs 2016-40-22638(p22638)→2773(ST162)	0,37

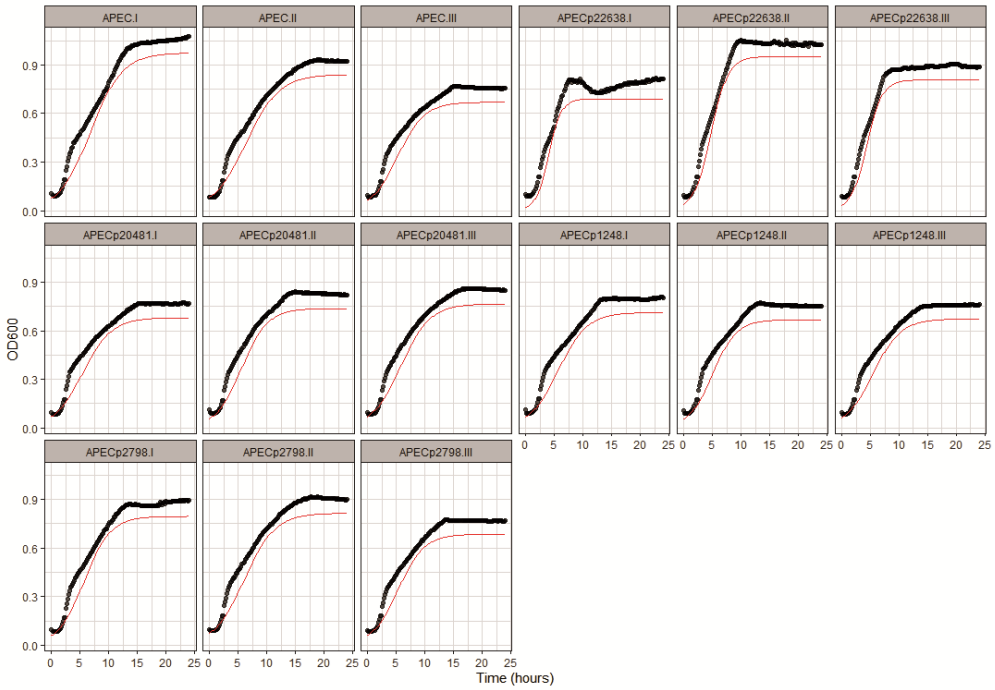


Figure B1. Triplicates of single strain growth curves of APEC and its respective transconjugants

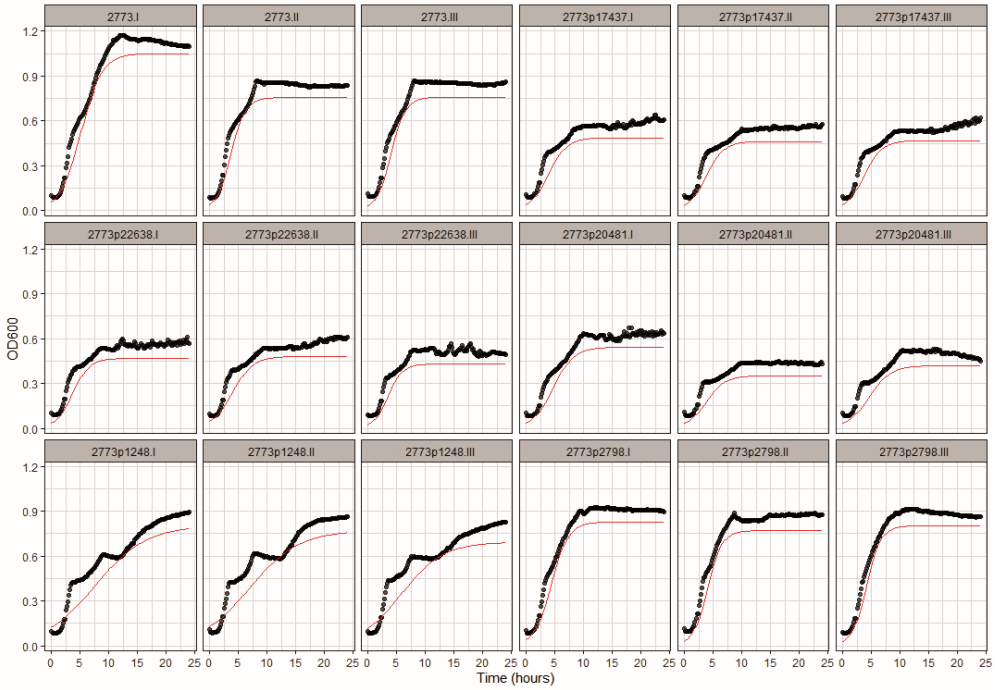


Figure B2. triplicates of single strain growth curves of 2773(ST162) and its respective transconjugants

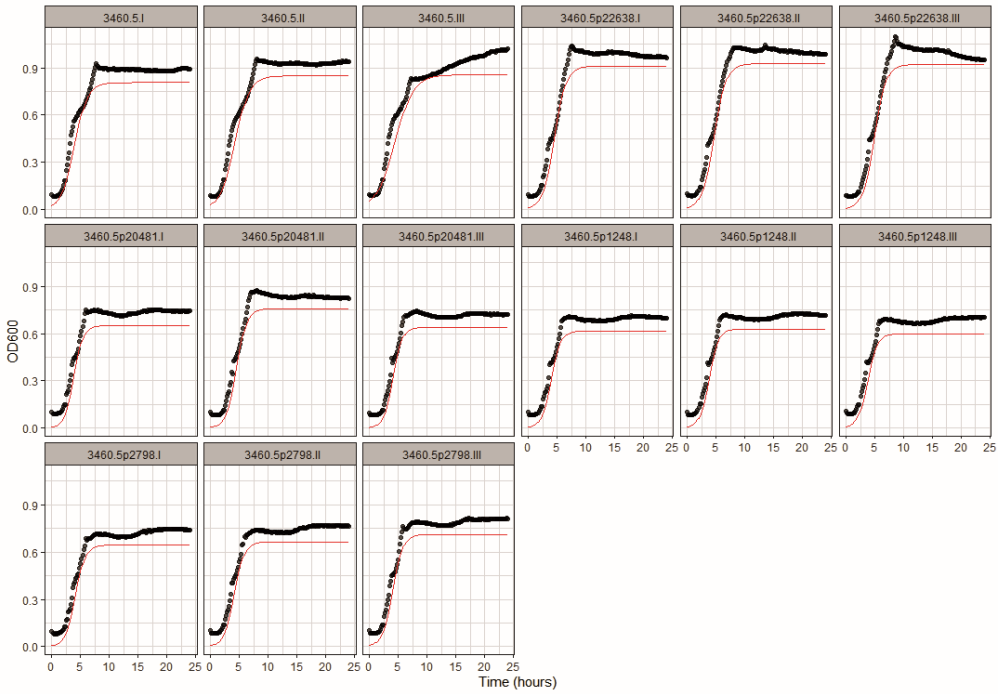


Figure B3. Triplicates of single strain growth curves of 3460-5(ST602) and its respective transconjugants.

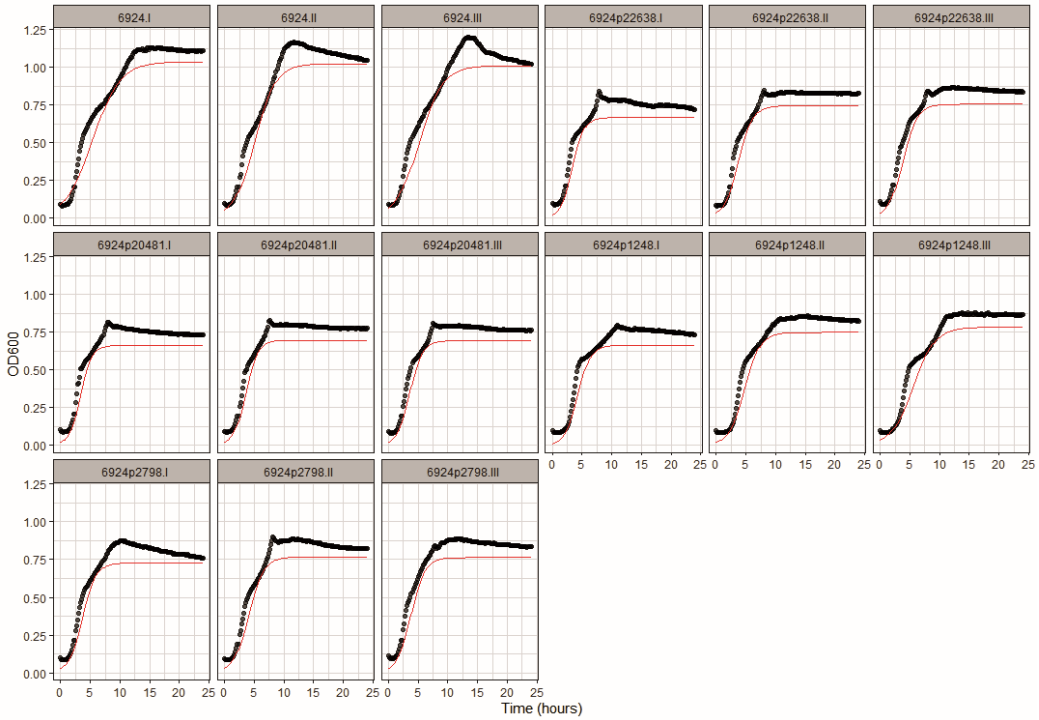


Figure B4. Triplicates of single strain growth curves of 6924 (ST453) and its respective transconjugants.

Table B4. Maximum growth rates (r_{max}) and carrying capacities (K) of the triplicates of each single strain growing culture, as well as mean and SD values.

	r_{max}					K				
Strain	I	II	III	Mean	SD	I	II	III	Mean	SD
APEC	0,69	0,70	0,69	0,70	0,0063	0,97	0,84	0,67	0,83	0,15
p22638/APEC	0,70	0,79	0,74	0,74	0,046	0,69	0,95	0,81	0,82	0,13
p20481/APEC	0,71	0,67	0,70	0,69	0,018	0,68	0,74	0,77	0,73	0,043
p1248/APEC	0,66	0,67	0,66	0,67	0,0047	0,71	0,67	0,67	0,68	0,025
p2798/APEC	0,72	0,70	0,70	0,71	0,015	0,80	0,82	0,69	0,77	0,070
2773(ST162)	0,80	0,85	0,79	0,81	0,032	1,05	0,76	0,76	0,86	0,17
p17437/2773(ST162)	0,72	0,75	0,76	0,74	0,018	0,49	0,46	0,47	0,47	0,013
p22638/2773(ST162)	0,76	0,78	0,80	0,78	0,021	0,47	0,48	0,43	0,46	0,026
p20481/2773(ST162)	0,73	0,73	0,74	0,73	0,0053	0,54	0,35	0,42	0,44	0,10
p1248/2773(ST162)	0,77	0,77	0,79	0,78	0,0093	0,80	0,77	0,69	0,76	0,056
p2798/2773(ST162)	0,80	0,77	0,77	0,78	0,018	0,83	0,77	0,80	0,80	0,029
6924(ST453)	0,83	0,83	0,83	0,83	0,0020	1,04	1,02	1,01	1,02	0,013
p22638/6924(ST453)	0,83	0,92	0,81	0,86	0,059	0,67	0,74	0,75	0,72	0,048
p20481/6924(ST453)	0,87	0,80	0,81	0,83	0,040	0,66	0,70	0,69	0,68	0,019
p1248/6924(ST453)	0,87	0,84	0,70	0,80	0,090	0,66	0,75	0,78	0,73	0,062
p2798/6924(ST453)	0,80	0,80	0,80	0,80	0,0025	0,73	0,77	0,76	0,75	0,022
3460-5 (ST602)	0,83	0,79	0,79	0,80	0,022	0,81	0,85	0,85		0,027
p22638/3460-5(ST602)	0,77	0,76	0,71	0,75	0,035	0,91	0,93	0,92	0,92	0,0080
p20481/3460-5(ST602)	0,72	0,75	0,75	0,74	0,016	0,65	0,76	0,64	0,68	0,066
p1248/3460-5(ST602)	0,64	0,67	0,71	0,67	0,036	0,62	0,63	0,60	0,61	0,015
p2798/3460-5(ST602)	0,68	0,68	0,77	0,71	0,054	0,64	0,67	0,71	0,67	0,035

Table B5. Student's T-test comparison of r_{max} values and K values between the recipient strain and its respective transconjugants.

Pairs compared	Growth rate T-test pairwise comparison (p-value)	Carrying capacity T-test pairwise comparison (p-value)
APEC vs. p22638/APEC	0,20	0,93
APEC vs. p20481/APEC	0,92	0,37
APEC vs. p1248/APEC	0,0035	0,24
APEC vs. p2798/APEC	0,30	0,57
2773(ST162) vs. p17437/2773(ST162)	0,04	0,059
2773(ST162) vs. p22638/2773(ST162)	0,20	0,052
2773(ST162) vs. p20481/2773(ST162)	0,045	0,030
2773(ST162) vs. p1248/2773(ST162)	0,20	0,42
2773(ST162) vs. p2798/2773(ST162)	0,22	0,63
6924(ST453) vs. p22638/6924(ST453)	0,57	0,0056
6924(ST453) vs. p20481/6924(ST453)	0,91	$3,62x10^{-5}$
6924(ST453) vs. p1248/6924(ST453)	0,64	0,012
6924(ST453) vs. p2798/6924(ST453)	$8,75x10^{-5}$	0,00023
3460-5(ST602) vs. p22638/3460(ST602)	0,094	0,025
3460-5(ST602) vs. p20481/3460(ST602)	0,019	0,040
3460-5(ST602) vs. p1248/3460(ST602)	0,010	0,00081
3460-5(ST602) vs. p2798/3460(ST602)	0,082	0,0039

Table B6. Student's T-test comparison of r_{max} values between the APEC strain and QREC recipient strains.

Paris compared	T-test
APEC vs. 2773(ST162)	0,021
APEC vs. 6924(ST453)	0,00024
APEC vs. 3460-5(ST602)	0,0099

Table B7. Mean R values and their respective SD of selected competing pairs calculated for each day during the competitive growth assay.

Competing pairs	Time (days)	Mean R	SD
p17437/2773(ST162) vs. 2773 (ST162)	0	0,04	0,10
	1	-1,95	0,15
	2	-3,04	0,20
	3	-3,47	0,08
	4	-4,19	0,44
	5	-5,00	0,46
p22638/2773(ST162) vs. 2773(ST162)	0	0,06	0,11
	1	-0,41	0,59
	2	-0,82	0,90
	3	-1,12	0,75
	4	-2,32	1,11
	5	-3,12	0,79
p22638/APEC vs APEC	0	-0,04	0,12
	1	-0,33	0,06
	2	-0,25	0,46
	3	-0,62	0,18
	4	-1,15	0,17
	5	-1,41	0,42

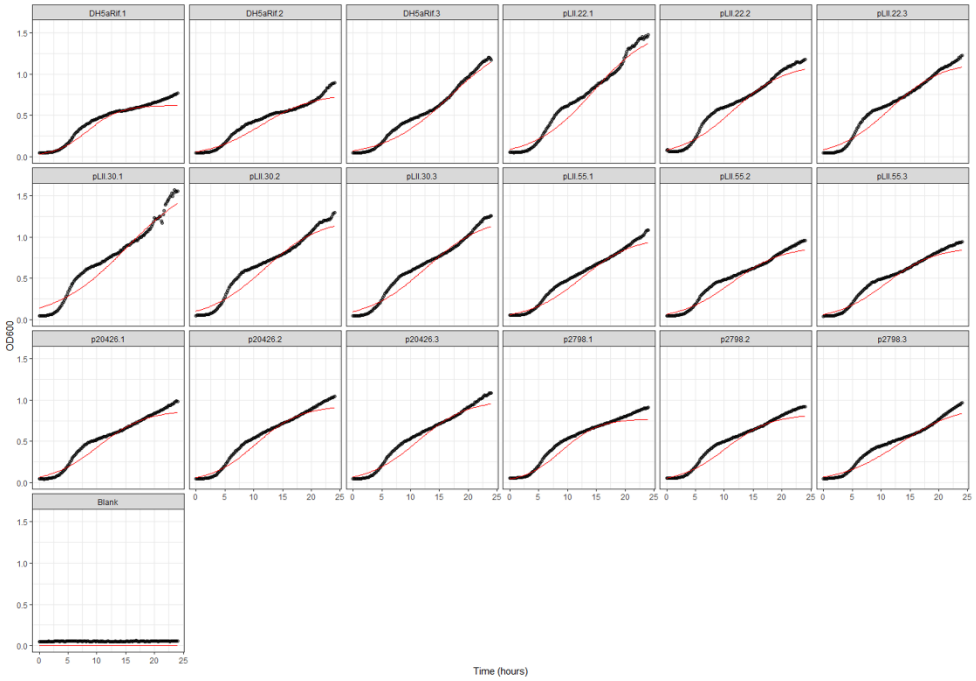


*Figure B6. Growth curve triplicates of DH5α rif^R, pL-II-22/DH5α rif^R, pL-II-30/DH5α rif^R, pL-II-55/DH5α rif^R, p20426/DH5α rif^R and p2798/DH5α rif^R with initial 1,5*10^6 CFU/ml. Black dotted line represents the empirical OD_{600} measurements while red line represents the estimated growth curve generated by the growthcurver.*

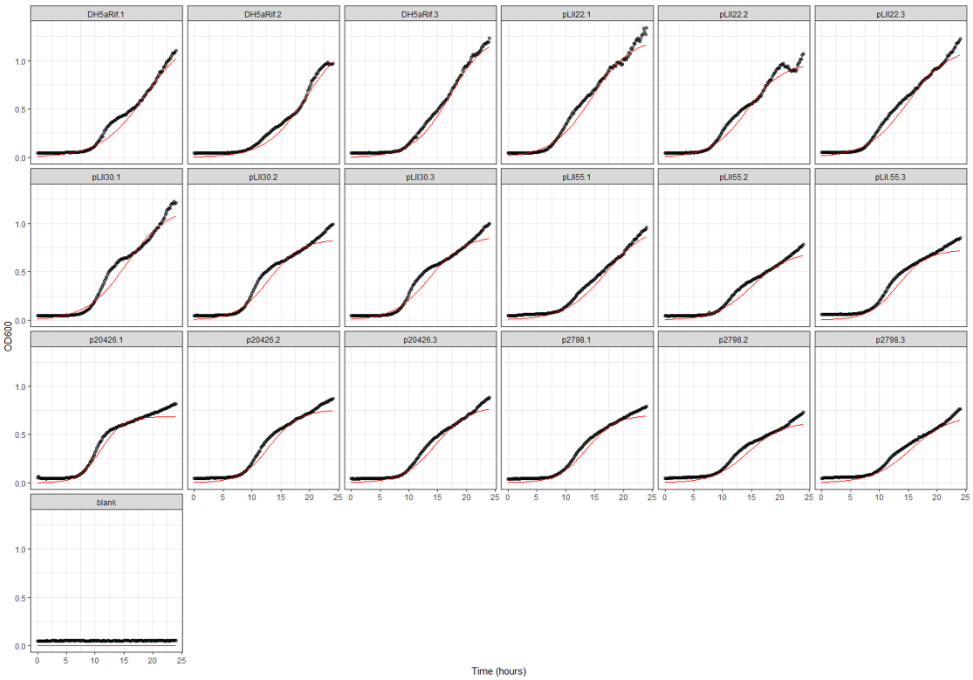


*Figure B7. Growth curve triplicates of DH5α rif^R, pL-II-22/DH5α rif^R, pL-II-30/DH5α rif^R, pL-II-55/DH5α rif^R, p20426/DH5α rif^R and p2798/DH5α rif^R with initial 1,5*10^4 CFU/ml. Black dotted line represents the empirical OD_{600} measurements while red line represents the estimated growth curve generated by the growthcurver.*

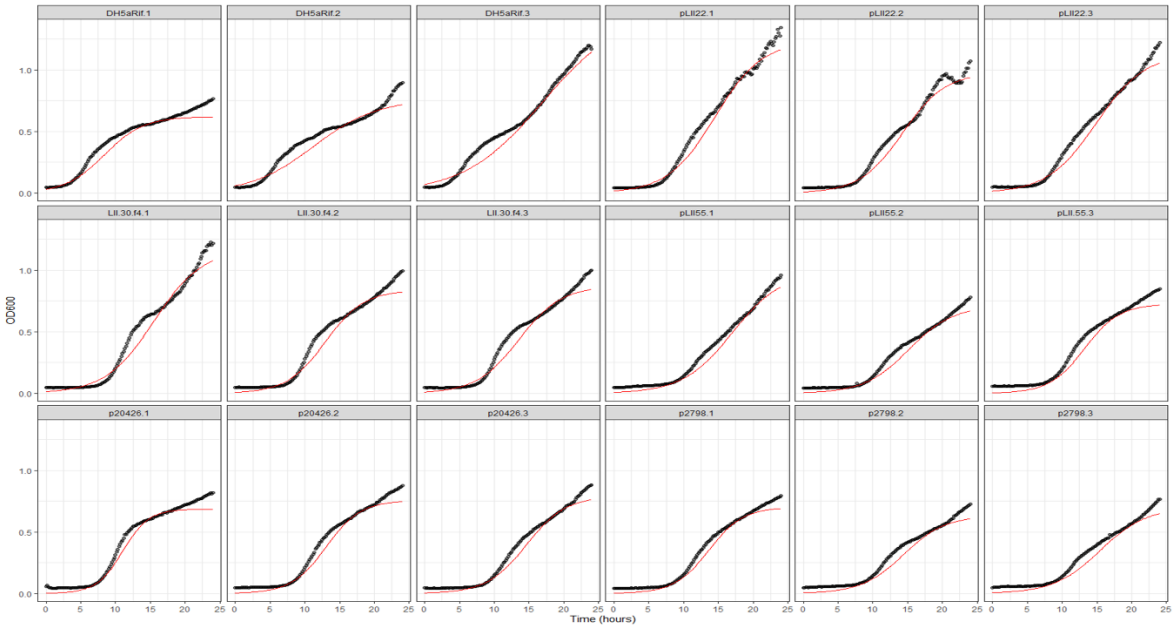


Figure B8. Growth curve triplicates of DH5α rifR with initial 1,5x10^6 CFU/ml and DH5α rifR plasmid-carrier cultures with initial concentration of 1,5x10^4 CFU/ml. Black dotted line represents the empirical OD_{600} measurements while red line represents the estimated growth curve generated by the growthcurver.

Table B8. Number of generations produced during 24hours incubation period when initial CFU/ml for each culture was 1,5x10^6. Number of generations were calculate based on the average value of K and N_0 generated by the growthcurver.

Growing culture	Average K	Average N_0	Number of generations pr. 24h
DH5α rif^R	0,95	0,05	4,16
pL-II-22/DH5α rif^R	1,29	0,08	3,93
pL-II-30/DH5α rif^R	1,41	0,11	3,64
pL-II-55/DH5α rif^R	0,92	0,06	3,86
p20426/DH5α rif^R	0,94	0,07	3,84
p2798/DH5α rif^R	0,85	0,06	3,90

Table B9. Number of generations produced during 24hours incubation period when initial CFU/ml for each culture was 1,5x10^4. Number of generations were calculate based on the average value of K and N_0 generated by the growthcurver.

Growing culture	Average K	Average N_0	Number of generations pr. 24h
DH5α rif^R	1,27	0,007	7,51
pL-II-22/DH5α rif^R	0,94	0,009	6,79
pL-II-30/DH5α rif^R	0,96	0,010	6,52
pL-II-55/DH5α rif^R	0,81	0,006	7,09
p20426/DH5α rif^R	0,74	0,004	7,72
p2798/DH5α rif^R	0,68	0,005	7,12

Table B10. r_{max} of triplicates of the DH5α rif^R and its respective transconjugants, where initial CFU/ml of the recipient was 100x times higher than the initial CFU of the transconjugants

Strain	Replicates	r_{max}	Average r_{max}	SD
DH5α rif^R	I	0,96	0,95	0,081
	II	0,87		
	III	1,03		
pL-II-22/DH5αrif^R	I	0,77	0,77	0,023
	II	0,74		
	III	0,79		
pL-II-30/DH5α rif^R	I	0,73	0,77	0,031
	II	0,79		
	III	0,77		
pL-II-55/DH5α rif^R	I	0,34	0,51	0,15
	II	0,57		
	III	0,62		
p20426/DH5α rif^R	I	0,74	0,66	0,074
	II	0,65		
	III	0,59		
p2798/DH5α rif^R	I	0,65	0,55	0,089
	II	0,49		
	III	0,49		

Table B11. Student's T-test p-values of the growth rate comparation between the p- DH5αrif^R and p$^+$DH5αrif^R, calculated based on the values from the Table B9.

p- DH5αrif^R vs. p$^+$DH5αrif^R	Student's T-test p-value
DH5α rif^R vs. pL-II-22/DH5α rif^R	0,047
DH5α rif^R vs. pL-II-30/DH5 rif^R	0,043
DH5α rif^R vs. pL-II-55/DH5α rif^R	0,018
DH5α rif^R vs. p20426/DH5α rif^R	0,011
DH5α rif^R vs. p2798/DH5α rif^R	0,004

Table B12. Relative plasmid abundance calculated for each parallel of each competing pair for each sampling point, as well as their respective mean RPA values, SD values, and confidence intervals (upper and lower values).

	Relative plasmid abundance (RPA)														
	p2798			p20426			pL-II-22			pL-II-30			p-L-II-55		
Sample №	1% Donor	99% Donor	100% Donor	1% Donor	99% Donor	100% Donor	1% Donor	99% Donor	100% Donor	1% Donor	99% Donor	100% Donor	1% Donor	99% Donor	100% Donor
1	0,85	0,95	0,77	0,90	0,94	0,97	0,91	0,88	0,93	0,98	0,91	0,92	0,89	0,92	0,92
2	0,86	0,95	0,95	0,98	0,90	0,84	1,00	0,95	0,98	0,98	0,90	0,96	0,99	0,86	0,97
3	0,90	0,94	0,93	0,93	0,93	0,90	0,92	0,91	0,92	0,89	0,96	0,94	0,99	0,90	0,91
4	0,96	0,86	0,91	0,88	0,89	0,90	0,92	0,96	0,95	0,90	0,91	0,91	0,97	0,83	0,96
5	0,86	0,87	0,86	0,93	0,91	0,89	0,89	0,89	0,91	0,94	0,90	0,95	0,96	0,91	0,98
6	0,90	0,90	0,92	0,91	0,89	0,90	0,91	0,91	0,90	0,99	0,90	0,90	0,90	0,86	0,92
7	0,96	0,96	0,97	0,96	0,97	0,96	0,99	0,99	0,98	0,97	0,98	0,99	0,96	0,91	0,98
8	0,97	0,96	0,97	0,95	0,97	0,96	0,98	0,98	0,98	0,95	0,98	0,98	0,98	0,91	0,99
9	0,96	0,99	0,97	0,95	0,97	0,95	0,97	0,99	0,99	0,96	0,97	0,99	0,97	0,90	0,96
10	0,95	0,98	0,97	0,96	0,96	0,97	0,97	0,98	0,98	0,82	1,00	0,98	0,96	0,90	0,99
11	0,91	0,91	0,91	0,93	0,94	0,91	0,94	0,94	0,93	0,92	0,95	0,93	0,89	0,86	0,93
12	0,90	0,92	0,92	0,92	0,93	0,92	0,94	0,93	0,92	0,94	0,93	0,94	0,93	0,87	0,94
13	0,91	0,92	0,92	0,91	0,93	0,91	0,90	0,94	0,93	0,92	0,93	0,94	0,92	0,86	0,93
Mean SR	0,95	0,88	0,95	0,93	0,93	0,92	0,94	0,94	0,95	0,94	0,94	0,95	0,95	0,88	0,95
SD	0,037	0,030	0,028	0,030	0,030	0,039	0,035	0,036	0,033	0,046	0,035	0,030	0,037	0,030	0,028
CI (upper	0,94	0,95	0,95	0,94	0,95	0,94	0,96	0,96	0,94	0,96	0,96	0,96	0,97	0,90	0,97
and lower value)	0,89	0,91	0,89	0,91	0,92	0,90	0,92	0,92	0,90	0,92	0,92	0,93	0,93	0,87	0,94

Table B13. Total number of shufflon variants of pL-II, p17437 (original host), p17437/2773(ST162), p20481 (original host), and p20481/2773(ST162) in different time points.

	Number of shufflon variants				
Time point	pL-II	17347	2773p17437	20481	2773p20481
T0	65	22	22	21	24
T1	67	19	32	16	15
T2	53	25	20	14	14
T3	47	24	19	10	14

Table B14. Main distribution of pL-II shufflon reads divided into three main categories, reads with the full set of segments, reads with a deletion of one segment, and reads with a deletion of 2 segments over different time points.

Subgroup of reads		Complete set of segments	Δ 1 segment	Δ 2 segments	Total
T0	№ reads	3528	398	1998	5924
	%	59,55	6,72	33,73	100
T1	№ reads	3662	743	2231	6636
	%	55,18	11,20	33,62	100
T2	№ reads	2789	548	7926	11263
	%	24,76	4,87	70,37	100
T3	№ reads	4943	538	10169	15650
	№ reads	31,58	3,44	64,98	100

Table B15. Distribution of the pL-II shufflon reads with the complete set of segments over different time points.

Complete set of segments		A'A BB' C'C	A'A B'B C'C	A'A C'C B'B	AA' C'C B'B	AA' BB' C'C	Other	Total
T0	№ reads	2016	304	285	241	229	453	3528
	%	57,14	8,62	8,08	6,83	6,49	12,84	100
T1	№ reads	1412	401	262	248	316	1023	3662
	%	38,56	10,95	7,15	6,77	8,63	27,94	100
T2	№ reads	1270	349	248	235	247	449	2798
	%	45,39	12,47	8,86	8,40	8,83	16,05	100
T3	№ reads	2159	1532	523	217	220	292	4943
	№ reads	43,68	30,99	10,58	4,39	4,45	5,91	100

Table B16. Distribution of the pL-II shufflon reads with a deletion of one segment over different time points.

Δ 1 segment		Δ A	Δ B	Δ C	Total
T0	№ reads	209	183	6	398
	%	52,51	45,98	1,51	100

T1	№ reads	379	361	3	743
	%	51,01	48,59	0,40	100
T2	№ reads	264	275	9	548
	%	48,18	50,18	1,64	100
T3	№ reads	269	266	3	538
	№ reads	50,00	49,44	0,56	100

Table B17. Distribution of the pL-II shufflon reads with a deletion of two segments over different time points.

Δ 2 segments		Δ AB	Δ AC	Δ CB	Total
T0	№ reads	588	349	1061	1998
	%	29,43	17,47	53,10	100
T1	№ reads	639	429	1163	2231
	%	28,64	19,23	52,13	100
T2	№ reads	1840	1484	4602	7926
	%	23,21	18,72	58,06	100
T3	№ reads	2455	1959	5755	10169
	№ reads	24,14	19,26	56,59	100

Table B18. Main distribution of p17437 (original host) shufflon reads divided into three main categories, reads with the full set of segments, reads with a deletion of one segment, and reads with a deletion of 2 segments over different time points.

Subgroup of reads		Complete set of segments	Δ 1 segment	Δ 2 segments	Total
T0	№ reads	3988	73	63	4124
	%	96,70	1,77	1,53	100
T1	№ reads	945	175	580	1700
	%	55,59	10,29	34,12	100
T2	№ reads	996	277	1353	2626
	%	37,93	10,55	51,52	100
T3	№ reads	874	362	1012	2248
	№ reads	38,88	16,10	45,02	100

Table B19. Distribution of the p17437 (original host) shufflon reads with the complete set of segments over different time points.

Complete set of segments		AA' CC' BB'	A'A CC' BB'	Other	Total
T0	№ reads	1535	2373	80	3988
	%	38,49	59,50	2,01	100
T1	№ reads	387	540	18	945
	%	40,95	57,14	1,90	100
T2	№ reads	397	578	21	996
	%	39,86	58,03	2,11	100
T3	№ reads	358	497	19	874
	№ reads	40,96	56,86	2,17	100

Table B20. Distribution of the p17437 (original host) shufflon reads with a deletion of one segment over different time points.

Δ 1 segment		Δ A	Δ B	Δ C	Total
T0	№ reads	17	56	0	73
	%	23,29	76,71	0	100
T1	№ reads	4	171	0	175
	%	2,29	97,71	0	100
T2	№ reads	5	272	0	277
	%	1,81	98,19	0	100
T3	№ reads	7	355	0	362
	№ reads	1,93	98,07	0	100

Table B21. Distribution of the p17437 (original host) shufflon reads with a deletion of two segments over different time points.

Δ 2 segments		Δ AB	Δ AC	Δ CB	Total
T0	№ reads	0	0	63	63
	%	0	0	100	100
T1	№ reads	0	0	580	580
	%	0	0	100	100
T2	№ reads	0	0	1353	1353
	%	0	0	100	100

T3	№ reads	0	1	1011	1012
	№ reads	0	0,099	99,90	100

Table B22. Main distribution of p17437 (original host) shufflon reads divided into three main categories, reads with the full set of segments, reads with a deletion of one segment, and reads with a deletion of 2 segments over different time points.

Subgroup of reads		Complete set of segments	Δ 1 segment	Δ 2 segments	Total
T0	№ reads	5186	40	19	5245
	%	98,88	0,76	0,36	100
T1	№ reads	5701	150	139	5990
	%	95,18	2,50	2,32	100
T2	№ reads	875	60	165	1100
	%	79,55	5,45	15,00	100
T3	№ reads	1019	33	377	1429
	№ reads	71,31	2,31	26,38	100

Table B23. Distribution of the p17437/2773(ST162) shufflon reads with the complete set of segments over different time points.

Complete set of segments		AA' CC' BB'	A'A CC' BB'	Other	Total
T0	№ reads	3073	2032	81	5186
	%	59,26	39,18	1,56	100
T1	№ reads	2222	3265	217	5701
	%	38,98	57,27	3,81	100
T2	№ reads	392	471	10	873
	%	44,90	53,95	1,15	100
T3	№ reads	393	608	8	1019
	№ reads	38,57	59,67	0,79	100

Table B24. Distribution of the p17437/2773(ST162) shufflon reads with a deletion of one segments over different time points.

Δ 1 segment		Δ A	Δ B	Δ C	Total
T0	№ reads	19	21	0	41
	%	46,34	51,22	0	100
T1	№ reads	36	114	0	150
	%	24,00	76,00	0	100
T2	№ reads	0	60	0	60
	%	0,00	100,00	0	100
T3	№ reads	2	29	0	31
	№ reads	6,45	93,55	0	100

Table B25. Distribution of the p17437/2773(ST162) shufflon reads with deletion of two segments over different time points.

Δ 2 segments		Δ AB	Δ AC	Δ CB	Total
T0	№ reads	0	1	18	19
	%	0	5,26	94,74	100
T1	№ reads	0	5	134	139
	%	0	3,60	96,40	100
T2	№ reads	0	0	164	164
	%	0	0	100	100
T3	№ reads	0	0	376	376
	№ reads	0	0	100	100

Table B26. Main distribution of p20481(original host) shufflon reads divided into three main categories, reads with the full set of segments, reads with a deletion of one segment, and reads with a deletion of 2 segments over different time points.

Subgroup of reads		Complete set of segments	Δ 1 segment	Δ 2 segments	Total
T0	№ reads	10383	9	3	10395
	%	99,89	0,087	0,029	100
T1	№ reads	4956	149	115	5220
	%	94,94	2,85	2,20	100
T2	№ reads	4002	35	48	4085
	%	97,97	0,86	1,18	100
T3	№ reads	2542	44	65	2651
	№ reads	95,89	1,66	2,45	100

Table B27. Distribution of the p20481 (original host) shufflon reads with the complete set of segments over different time points.

Complete set of segments		A'A BB' C'C	AA' BB' C'C	Other	Total
T0	№ reads	6904	2634	845	10383
	%	66,49	25,37	8,14	100
T1	№ reads	4203	737	16	4956
	%	84,81	14,87	0,32	100
T2	№ reads	3457	532	13	4002
	%	86,38	13,29	0,32	100
T3	№ reads	2128	406	8	2542
	№ reads	83,71	15,97	0,31	100

Table B28. Distribution of the p20481 (original host) shufflon reads with a deletion of one segment.

Δ 1 segment		Δ A	Δ B	Δ C	Total
T0	№ reads	1	8	0	9
	%	11,11	88,89	0	100
T1	№ reads	1	148	0	149
	%	0,67	99,33	0	100
T2	№ reads	0	35	0	35
	%	0	100	0	100
T3	№ reads	0	44	0	44
	№ reads	0	100	0	100

Table B29. Distribution of the p20481 (original host) shufflon reads with a deletion of two segments over different time points.

Δ 2 segments		Δ AB	Δ AC	Δ CB	Total
T0	№ reads	0	0	3	3
	%	0	0	100	100
T1	№ reads	0	0	115	115
	%	0	0	100	100
T2	№ reads	0	0	48	48
	%	0	0	100	100
T3	№ reads	0	0	65	65
	№ reads	0	0	100	100

Table B30. Main distribution of p20481/2773(ST162) shufflon reads divided into three main categories, reads with the full set of segments, reads with a deletion of one segment, and reads with a deletion of 2 segments over different time points.

Subgroup of reads		Complete set of segments	Δ 1 segment	Δ 2 segments	Total
T0	№ reads	3906	215	22	4143
	%	94,28	5,18	0,54	100
T1	№ reads	868	105	62	1035
	%	83,87	10,14	5,99	100
T2	№ reads	850	97	195	1142
	%	74,43	8,49	17,08	100
T3	№ reads	552	396	458	1406
	№ reads	39,26	28,16	32,58	100

Table B31. Distribution of the p20481/2773(ST162) shufflon reads with the complete set of segments over different time points.

Complete set of segments		A'A BB' C'C	AA' BB' C'C	Other	Total
T0	№ reads	1987	1542	377	3906
	%	50,87	39,48	9,65	100
T1	№ reads	511	349	7	868
	%	58,87	40,21	0,81	100
T2	№ reads	521	310	19	850
	%	61,29	36,47	2,23	100
T3	№ reads	429	119	4	552
	№ reads	77,71	21,56	0,73	100

Table B32. Distribution of the p20481/2773(ST162) shufflon reads with a deletion of one segment over different time points.

Δ 1 segment		Δ A	Δ B	Δ C	Total
T0	№ reads	0	215	0	215
	%	0	100	0	100
T1	№ reads	0	105	0	105
	%	0	100	0	100
T2	№ reads	0	97	0	97
	%	0	100	0	100
T3	№ reads	0	396	0	396
	№ reads	0	100	0	100

Table B33. Distribution of the p20481/2773(ST162) shufflon reads with a deletion of two segments over different time points.

Δ 2 segments		Δ AB	Δ AC	Δ CB	Total
T0	№ reads	0	0	22	22
	%	0	0	100	100
T1	№ reads	0	0	62	62
	%	0	0	100	100
T2	№ reads	0	0	195	195
	%	0	0	100	100
T3	№ reads	0	0	458	458
	№ reads	0	0	100	100

Table B34. Number of shufflon reads with the B segment not interrupted by the ISEcp1 found in p17437 and p20481 grown in original and 2773 E. coli strain over all time points. % of reads was calculated based on the total number of correct reads found per sample.

	p17437		2773p17437		p20481		2773p20481	
Sampling time point	Total	%	Total	%	Total	%	Total	%
T0	6	0,15	2	0,038	1	0,0096	10	0,24
T1	1	0,059	6	0,10	3	0,057	3	0,29
T2	4	0,15	3	0,27	2	0,049	2	0,17
T3	2	0,089	3	0,21	1	0,037	1	0,071

Table B35. Distribution of the pL-II pilV ORFs based on the all the correct reads over different time points

pilV ORFs		*PilV-A*	*PilV-A'*	*PilV-B*	*PilV-B'*	*PilV-C*	*PilV-C'*	Total
T0	№ reads	1031	3484	274	358	355	422	5924
	%	17,40	58,81	4,63	6,04	5,99	7,12	100
T1	№ reads	1409	3152	481	497	592	505	6636
	%	21,23	47,50	7,25	7,49	8,92	7,61	100
T2	№ reads	2794	4792	713	1016	869	1079	11263
	%	24,81	42,55	6,33	9,02	7,72	9,58	100
T3	№ reads	3904	6952	1440	734	1169	1451	15650
	№ reads	24,95	44,42	9,20	4,69	7,47	9,27	100

Table B36. Distribution of the p17437 (original host) pilV ORFs based on the all the correct reads over different time points.

pilV ORFs		*PilV-A*	*PilV-A'*	*PilV-B*	*PilV-B'*	*PilV-C*	*PilV-C'*	Total
T0	№ reads	1580	2423	58	0	26	37	4124
	%	38,31	58,75	1,41	0,00	0,63	0,90	100
T1	№ reads	739	869	14	0	30	48	1700
	%	43,47	51,12	0,82	0,00	1,76	2,82	100
T2	№ reads	1131	1315	10	0	46	124	2626
	%	43,07	50,08	0,38	0,00	1,75	4,72	100
T3	№ reads	889	1150	12	0	64	133	2248
	№ reads	39,55	51,16	0,53	0,00	2,85	5,92	100

Table B37. Distribution of the p17437/2773(ST162) pilV ORFs based on the all the correct reads over different time points.

PilV ORFs		*PilV-A*	*PilV-A'*	*PilV-B*	*PilV-B'*	*PilV-C*	*PilV-C'*	Total
T0	№ reads	3092	2075	1	32	33	12	5245
	%	58,95	39,56	0,02	0,61	0,63	0,23	100
T1	№ reads	2308	3370	5	139	68	100	5990
	%	38,53	56,26	0,08	2,32	1,14	1,67	100
T2	№ reads	497	569	0	7	4	20	1097
	%	45,31	51,87	0,00	0,64	0,36	1,82	100

T3	№ reads	576	822	0	12	6	10	1426
	№ reads	40,39	57,64	0,00	0,84	0,42	0,70	100

Table B38. Distribution of the p20481 (original host) pilV ORFs based on the all the correct reads over different time points.

pilV ORFs		*PilV-A*	*PilV-A'*	*PilV-B*	*PilV-B'*	*PilV-C*	*PilV-C'*	Total
T0	№ reads	2643	6997	40	172	485	58	10395
	%	25,43	67,31	0,38	1,65	4,67	0,56	100
T1	№ reads	897	4309	0	5	8	1	5220
	%	17,18	82,55	0,00	0,10	0,15	0,02	100
T2	№ reads	580	3495	0	3	7	0	4085
	%	14,20	85,56	0,00	0,07	0,17	0,00	100
T3	№ reads	464	2182	0	0	1	4	2651
	№ reads	17,50	82,31	0,00	0,00	0,04	0,15	100

Table B39. Distribution of the p20481/2773(ST162) pilV ORFs based on the all the correct reads over different time points.

pilV ORFs		*PilV-A*	*PilV-A'*	*PilV-B*	*PilV-B'*	*PilV-C*	*PilV-C'*	Total
T0	№ reads	1664	2117	60	86	166	50	4143
	%	40,16	51,10	1,45	2,08	4,01	1,21	100
T1	№ reads	430	602	0	0	3		1035
	%	41,55	58,16	0,00	0,00	0,29	0,00	100
T2	№ reads	458	677	0		5	2	1142
	%	40,11	59,28	0,00	0,00	0,44	0,18	100
T3	№ reads	552	850	0	0	3	1	1406
	№ reads	39,26	60,46	0,00	0,00	0,21	0,07	100

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway