Mating pair stabilization mediates species specificity during bacterial conjugation

Wen Wen Low
Department of Life Sciences
Imperial College London

A thesis submitted for the degree of
Doctor of Philosophy

July 2022

Candidate's statement of originality

Unless referenced below or in the text, I hereby confirm that the work presented in this thesis is my own.

Wen Wen Low
July 2022

Figure 3.1. and 3.4. The pKpGFP reporter plasmid and Δ35/Δ36::35 were constructed with the assistance of Dr Joshua Wong.
Figure 3.10. Purification and analysis of the conjugative pilus from pKpQIL was performed by Dr Alejandro Peña.
Figure 3.22. *In silico* modelling of the single amino acid L3 insertions in OmpK36 was performed by Dr Konstantinos Beis.
Figure 3.27. The Δ*ompW* strain was constructed with the assistance of Ms Tatiana Bizeau.
Figures 3.35. and 3.41. AlphaFold predictions were generated with the assistance of Dr Konstantinos Beis.
Figure 3.37. Purification of TraN and OmpK36 and size exclusion chromatography experiments were performed by Ms Chloe Seddon.
Figure 3.38. Cryo-EM analysis of the purified TraN-OmpK36 complex was performed by Ms Letitia Beltran and Dr Fengbin Wang.
Appendices Curation of the conjugative IncF plasmid dataset and classification of TraN sequence variants was performed by Dr Sophia David.

List of publications arising from this work

Low, W. W. *et al.* Mating pair stabilization mediates bacterial conjugation species specificity. *Nature Microbiology* **7**, 1016–1027 (2022).

Zheng, W. *et al.* Cryoelectron-microscopic structure of the pKpQIL conjugative pili from carbapenem-resistant *Klebsiella pneumoniae*. *Structure* **28**, 1321-1328.e2 (2020).

Abstract

Bacterial conjugation is a contact-dependent form of horizontal gene transfer where DNA is transferred in a unidirectional manner from a donor to recipient bacterium. It is also a key driver of the spread of antimicrobial resistance plasmids within clinically important pathogens. The intimate attachment of cells within a conjugating pair is crucial for efficient DNA transfer. However, the mechanism underlying the formation of tight mating junctions, which occurs through a process termed mating pair stabilization (MPS), is unclear. This work describes how variants of the plasmid-encoded outer membrane protein TraN interact with different receptors on recipient cells to mediate MPS. Using a reporter plasmid generated from the *Klebsiella pneumoniae* carbapenem resistance plasmid pKpQIL, mutations in the major outer membrane porin OmpK36 were found to reduce plasmid uptake in recipients. Meanwhile, substitution of *traN* on this plasmid with *traN* from two related plasmids, the *Shigella flexneri* resistance plasmid, R100-1 and the prototypical F plasmid, revealed that these TraN variants mediate dependency on recipient OmpW and OmpA respectively instead. Structural analysis showed that TraN from pKpQIL forms a complex with OmpK36 via the insertion of a β-hairpin structure into one of the subunits of the trimeric porin. Combining bioinformatic analysis and structural predictions using AlphaFold, four additional TraN variants were identified. All seven TraN variants could be classified into four groups based on their structural similarity and associated receptors: TraNα (OmpW), TraNβ (OmpK36), TraNγ (OmpA) and TraNδ (OmpF). Species specificity was also observed during MPS as not all homologues of a receptor are recognized by each TraN. This specificity was reflected in the real-world host distribution of conjugative IncF plasmids suggesting that MPS plays an influential role in shaping the host range of these plasmids. These findings provide a precedent for developing strategies that target MPS to mitigate resistance gene dissemination.

Acknowledgements

I have been extremely lucky to have had an incredible support system throughout this PhD. First and foremost, I want to acknowledge members of the Frankel lab, both past and present, for their part in making these past few years so enjoyable. Thank you, of course, to Professor Gad Frankel for being a great mentor and supervisor. I am very grateful that you always encouraged me to pursue whatever piqued my curiosity and reminded me to be confident in myself. I will remember to always 'trust the data'. To Josh, without whom this project would not have even materialized, thank you for being not just a supervisor but a great friend. To Qiyun and Elena, I'm so grateful for our seemingly endless conversations and I hope we continue to have many of them. Thank you also to (in no particular order): Louise, for being my night shift buddy; Julia, for celebrating my wins like they're her own; Alejandro, for being the biggest GFP-DD hype man; Sabrina, for nerding out with me; Carrie, for diligently documenting my Asian flush; Sarah, for picking me up when I was going through a rough patch; Anya, for finding the swamp jokes so funny; Priyanka, for all the times I would turn around for a chat in the lab; Jonny, for being so chill (even when I sent you the wrong sequence); Audrey, for the helpful travel tips; and Rita, for always being so encouraging and supplying us with incredible baked goods.

It would not have been possible to turn the data in this study into a proper story without the help of all the collaborators involved. Thank you, Dr Konstantinos Beis, for essentially being my unofficial co-supervisor throughout this project. I'm grateful to both you and Chloe for all the discussions and time you both invested in this work. Thank you also to Dr Sophia David for all the help with the bioinformatics and for being open to testing out all my different ideas. My thanks also go out to Professor Edward Egelman, Leti Beltran, and Dr Fengbin Wang at the University of Virginia for contributing their cryo-EM expertise to this project. To Dr Min Chen and Dr Bach Pham, thank you for playing a role in trying to help us solve this puzzle. To my progress review panel members, Dr Angelika Grundling and Dr Vladimir Pelicic, I am grateful for the constructive comments and useful discussions. Focusing on TraN was indeed,

as you both advised, the way to go. Lastly, to Chloe (again) and Tatiana, thank you for making my supervision journey a fun experience and for contributing so much to this work.

I also want to acknowledge all the people who have played a part in guiding me towards where I am today. To my science teachers growing up, Puan Sal, Mr Mahendran, Mr Lim and Madam Hor, thank you for cultivating the sense of curiosity that continues to drive me and my work today. To Dr Federica Bernadini, Dr Matthew Child and Dr Abigail Clements, thank you for being the most encouraging supervisors at a time when I still wasn't sure about doing a PhD. You all played a part in shaping me as a researcher and I will always remember each of those projects fondly.

This work was supported by my parents in more ways than one and I will never be able to thank them enough or repay them for everything they have done for me. I am so lucky to be their daughter and this thesis is dedicated to them. Thank you to my sister who listened to me talk about this project for 3 hours as she got a crash course in microbiology on a plane ride to Spain. To my brother, sister-in-law and nieces, thank you for the constant supply of wholesome family videos that bring me so much joy. I love you, my familia.

Lastly, to Jer, who listened to every rant about a failed PCR, who forces my imposter syndrome out the door whenever it comes knocking and who never fails to make me laugh, thank you for being my best friend across continents and time zones.

Table of Contents

Table of Figures

List of Tables

1. Introduction

1.1 *Enterobacteriaceae*

Enterobacteriaceae is a large family of rod-shaped Gram-negative bacteria comprising more than 30 genera. Of these, species from about 10 genera including *Escherichia*, *Klebsiella*, *Salmonella*, *Enterobacter* and *Citrobacter* are considered 'core' members of this family – defined as being 40-50% related to the type species, *Escherichia coli*[1]. Several species can asymptomatically colonize the gastrointestinal tracts of both human and animals[2–4]. However, genera which are less commonly associated with vertebrates can be associated with insects and found in a wide range of environmental niches including on vegetation and in aquatic habitats[1]. The core genera include many clinically important human pathogens which cause a wide range of infections including septicaemia, pneumonia, and diarrheal disease both in the community and hospital setting[5].

In addition to inhabiting a wide range of environmental niches, enterobacterial species also frequently participate in horizontal exchange of DNA which contributes to their ability to adapt quickly to different surroundings[6]. This has also resulted in the ability to rapidly adapt to the use of antibiotics. Recently, *Klebsiella pneumoniae* and *Enterobacter* spp., have been included within the ESKAPE list of pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *K. pneumoniae*, *Acinetobacter baumanii*, *Pseudomonas aeruginosa,* and *Enterobacter* spp.)[7]. These six bacterial species are highlighted as emerging leading causes of multidrug resistant hospital acquired infections (HAIs)[8]. In a recent systematic analysis, three enterobacterial species were among the top ten bacterial causes of death attributed or associated with antimicrobial resistance (AMR): *E. coli* (1), *K. pneumoniae* (3) and *Enterobacter* spp. (9)[9]. Among the top 22 species in this analysis, 9 were from the *Enterobacteriaceae* family.

1.1.1 The bacterial outer membrane

A defining feature of Gram-negative bacteria is a cell envelope comprising two membranes surrounding a thin cell wall made of peptidoglycan (PG) (Figure 1.1)[10]. The outer membrane (OM) is a semi-permeable barrier which regulates the movement of various compounds into and out of the membrane-enclosed periplasmic space. This semi-permeable quality can be attributed to the unique asymmetrical architecture of the membrane – the inner leaflet of the OM is composed of phospholipids while the outer leaflet consists largely of lipopolysaccharide (LPS)[11]. This asymmetry is maintained by the Mla system which transports mislocalized phospholipids from the outer leaflet to the inner leaflet[12]. Many bacteria are also surrounded by a layer called the glycocalyx comprised of exopolysaccharides (EPS) and capsular polysaccharides (CPS) which is associated with the OM[13]. This layer can aid in the formation of biofilms and in protecting against the host immune system[14].

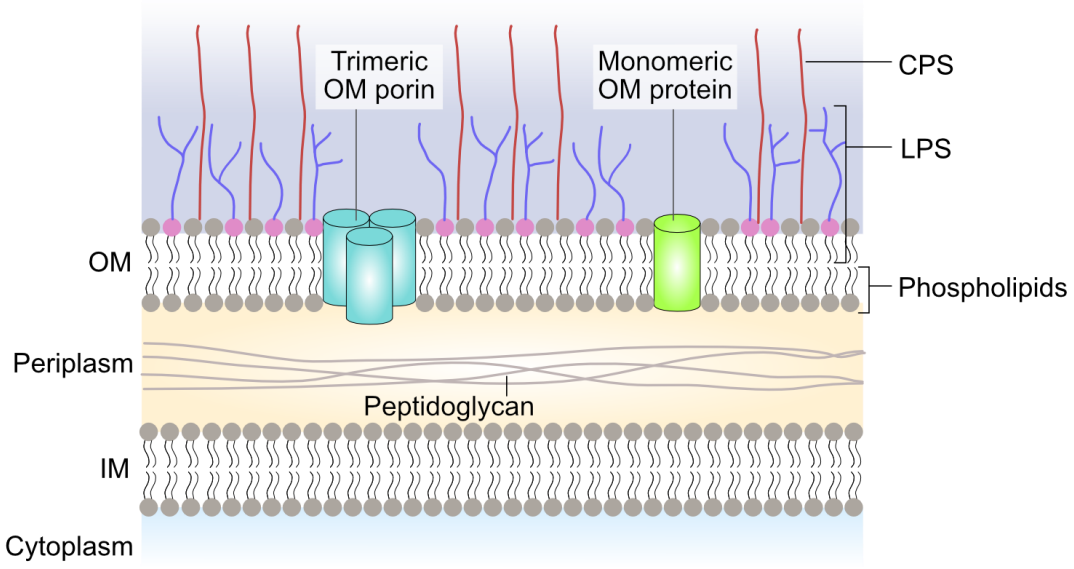


Figure 1.1. The cell envelope of Gram-negative bacteria.

Gram-negative bacteria are surrounded by two membranes enclosing the periplasmic space and peptidoglycan cell wall. The outer membrane (OM) is asymmetrical, with an inner leaflet comprised of phospholipids and an outer leaflet containing mostly lipopolysaccharides (LPS). The OM contains also contains many integral proteins and may be closely associated with capsular polysaccharides (CPS). IM = inner membrane.

The OM contains an abundance of proteins which are involved in various functions including signalling, cell adhesion and the transport of molecules across the membrane[15]. Most of these proteins are OM-embedded β-barrel proteins which range in size from containing 8 to 24 β-strands per monomer[15]. Several examples are listed in Table 1.1. One of the most well-studied Gram-negative OM proteins is OmpA, a monomeric β-barrel protein that is important for stabilizing the OM[16]. In addition to its role as a structural protein, OmpA has also been shown to mediate the formation of biofilms, is a receptor for several bacteriophages and plays a role in immune evasion among other functions[17–19]. Other major components of the OM are the OM porins. Porins are defined as non-selective trimeric channel-forming proteins which allow for the diffusion of small, hydrophilic compounds across the OM[20]. Thus, porins play an important role in controlling the permeability of the OM and the expression of these proteins is tightly regulated based on conditions like environmental osmolality[21]. Most early porin research was performed on the three 'classical' porins of *E. coli*: OmpF, OmpC and PhoE[22]. Alongside OmpA, OmpF and OmpC are among the most highly expressed components of the bacterial OM[23].

Table 1.1. Examples of outer membrane proteins.

Name	Structural features	Function	Other names
OmpA	Monomeric, 8-stranded monomer	Multiple functions including adhesion, invasion, and evasion of the host immune system	OprF in *Pseudomonas aeruginosa*
OmpW	Monomeric, 8-stranded monomer	Protection against environmental stress	Encoded by *yciD*
PldA	Dimeric, 12-stranded monomer	Phospholipase	OMPLA
OmpC	Trimeric, 16-stranded monomer	Non-specific porin	OmpK36 in *Klebsiella pneumoniae*
OmpF	Trimeric, 16-stranded monomer	Non-specific porin	OmpK35 in *Klebsiella pneumoniae*
PhoE	Trimeric, 16-stranded monomer	Non-specific porin	OmpE
LamB	Trimeric, 18-stranded monomer	Maltosaccharide-specific channel protein	Maltoporin

1.2 *Klebsiella pneumoniae*

Klebsiella pneumoniae (KP) is an encapsulated, facultative anaerobic bacterium that was first isolated in 1882 by Carl Friedlander from a patient who had succumbed to pneumonia[24]. Species belonging to the genus *Klebsiella* are ubiquitous in nature and can be found in the environment in soil, surface water and associated with animals[25]. KP is generally considered an opportunistic pathogen which asymptomatically colonises the mucosal surfaces of the nasopharynx, oropharyngeal cavity, and gastrointestinal tract while causing infection in hospitalised or immunocompromised individuals[26]. The most common infections caused by KP include those within the urinary tract, respiratory tract, and bloodstream. In the United States, it is recognised as the third leading cause of hospital acquired infections (HAIs), particularly ventilator-associated pneumonia in intensive care patients[27]. The mortality rate associated with pneumonia caused by KP has been reported to be as high as 50%[26].

Unlike most pathogenic bacteria, KP can successfully initiate infections despite the absence of virulence factors such as toxins and host-targeting enzymes. In addition, it is generally not associated with effector delivering secretion systems such as the type III secretion system (T3SS) for host immune modulation. Instead, it relies largely on host immune evasion as a strategy for survival[28]. Capsule is perhaps the most well-studied immune evasion factor associated with KP and consists of a polysaccharide matrix which coats the cell surface and protects against both serum killing and opsonophagocytosis[29,30]. Accordingly, multiple studies have shown that capsule null mutants are highly attenuated during infection[31–33]. At least 78 different capsular polysaccharides or ‘K-antigens’ have been identified in KP[34]. However, of these, only 25 serotypes account for over 70% of clinical isolates suggesting that some capsule serotypes may be associated with increased virulence[35].

1.2.1 Antimicrobial Resistance in *Klebsiella pneumoniae*

Treatment of infections caused by KP can be made more complicated due to the emergence of multidrug resistance against clinically available antibiotics. It has been estimated, that in

some countries, up to 80% of KP isolates are resistant to commonly used first-line antibiotics such as quinolones and aminoglycosides[36]. Another important group of antibiotics that are commonly used to treat KP infections are the β-lactams. Their widespread use in clinical settings is driven by their relatively low toxicity and high efficacy[37]. These antibiotics function by inhibiting synthesis of the peptidoglycan cell wall and are classified according to their chemical structure into several subgroups including aminopenicillins, cephalosporins and carbapenems. A chromosomal copy of the gene encoding the SHV-1 β-lactamase (*bla*$_{SHV-1}$) is considered ubiquitous in all KP strains, conferring intrinsic resistance to ampicillin[38]. As a result, third and fourth generation cephalosporins are commonly used to treat infections caused by this pathogen. However, due to the acquisition of genes encoding extended spectrum β-lactamases (ESBLs), many KP strains now exhibit resistance to cephalosporins[39].

1.2.2 Carbapenem resistant *K. pneumoniae*

Carbapenems are a last line treatment against multidrug resistant strains, but their utility is also rapidly being threatened by the emergence of carbapenem resistant KP (CRKP). Consequently, the World Health Organization has highlighted carbapenem resistant *Enterobacteriaceae* including CRKP as a pathogen of critical priority for the development of new antimicrobials[40]. Several high-risk sequence types tend to be responsible for the majority of CRKP outbreaks. Isolates from sequence type 258 (ST258) were first reported in the United States in the early 2000s and are now endemic in many regions around the world[41,42]. Resistance to carbapenems in ST258 strains of KP is generally attributed to two synergistic mechanisms – plasmid-mediated resistance and reduced OM permeability (Figure 1.2). These are discussed in more detail in the following sections.

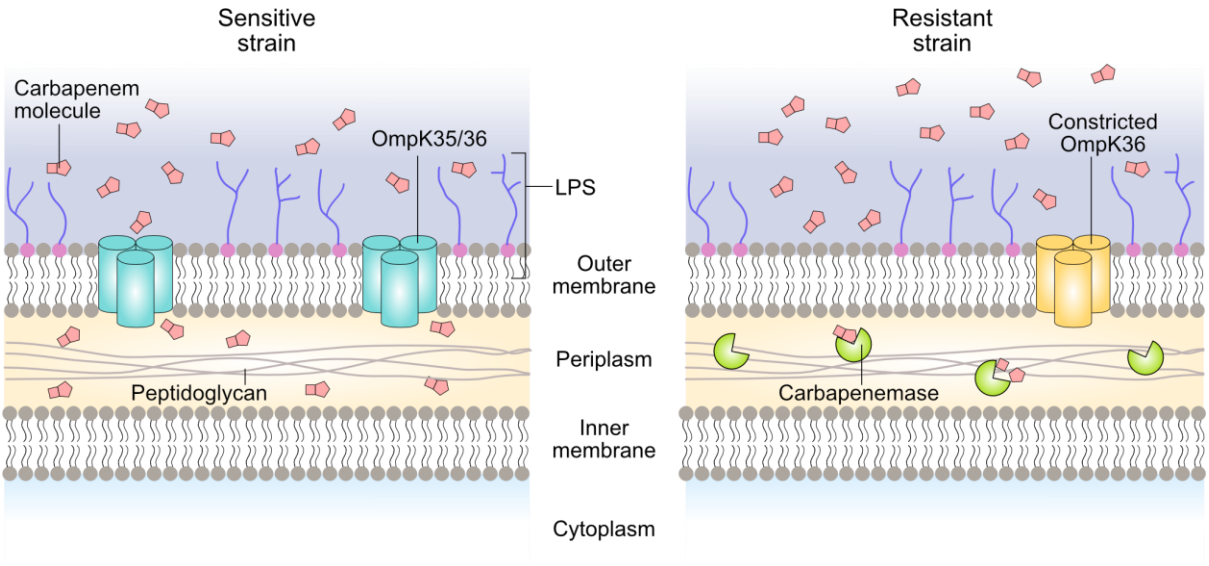


Figure 1.2. Mechanisms of carbapenem resistance in ST258 strains of *K. pneumoniae*.

Carbapenem sensitive strains of KP express the major OM porins, OmpK35 and OmpK36. These permit the entry of small hydrophilic molecules such as carbapenems into the bacterial periplasm where they interfere with peptidoglycan synthesis. In ST258 CRKP strains, mutations in *ompK35* and *ompK36* can result in the loss of expression of the porins or structural modifications which reduce OM permeability to carbapenems. This can be augmented by the expression of plasmid-encoded carbapenemases which rapidly hydrolyse carbapenems in the periplasm.

1.2.3 OmpK35 and OmpK36

Carbapenem resistance can result from mutations which alter the expression level or structural properties of the major KP OM porins, OmpK35 and OmpK36 (Figure 1.2). OmpK35 and OmpK36 are homologous to OmpF and OmpC respectively and have been studied extensively with regards to their role in AMR in KP[43,44].

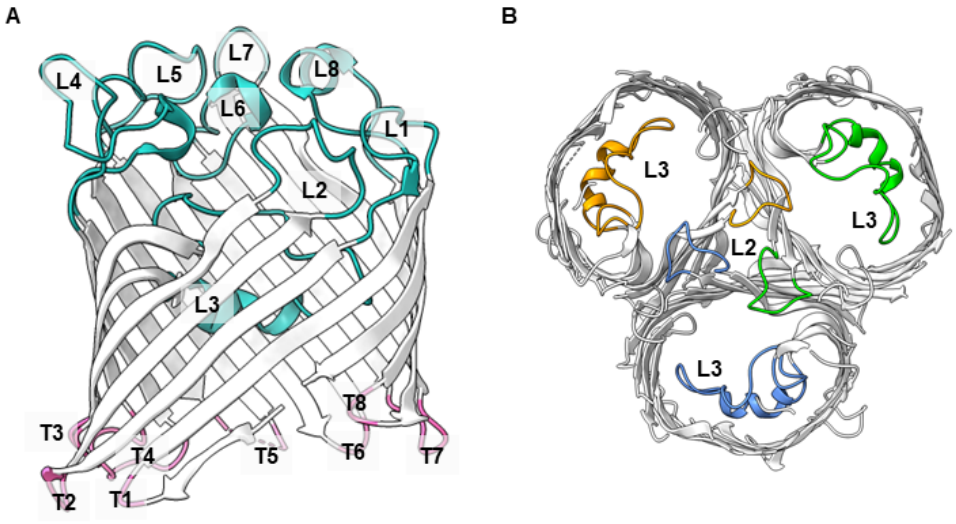


Figure 1.3. Structural features of classical outer membrane porins.

A. The extracellular loops (L) and periplasmic turns (T) on a monomer of the *K. pneumoniae* OM porin, OmpK36 (PDB: 5079)[45] are highlighted in pink and teal respectively. Loops and turns have been numbered according to the convention in the literature. **B.** Top view of OmpK36 showing the positions of L2 and L3 within the porin trimer. Loops from the same monomer share the same colour.

Although OM porins are generally considered non-specific, minor sequence differences can lead to structural changes which alter selectivity for solutes in terms of their charge and size[22]. The membrane-spanning portion of each porin monomer is made up of a 16 stranded β-barrel and each β-strand is connected to another by long extracellular loops and shorter periplasmic turns (Figure 1.3A). Several of these loops play important structural roles such as loop 2 which stabilizes the contacts between the porin monomers[46] (Figure 1.3B). Loop 3 (L3) is one of the longer loops and is unique in that it folds back into the barrel of the monomer to form a constriction point, also referred to as an 'eyelet', midway through the porin channel[47] (Figure 1.3B). Thus, L3 is a crucial determinant of size and ion selectivity, directly influencing the influx of compounds including antibiotics. Classical porins in *Enterobacteriaceae* have been found to contain a conserved PEFGGD motif in L3[48]. This motif is found in OmpK36 and the quiescent porin, OmpK37, which is structurally similar but not typically associated with carbapenem resistance[49]. OmpK35 contains a VEWGGD motif in L3 instead.

In ST258 strains of CRKP, mutations are frequently reported in the genes encoding both porins. OmpK35 has a larger channel diameter and is more permissive to the entry of carbapenems compared to OmpK36[44]. However, several studies have shown that OmpK35 is less abundant than OmpK36 even when strains are grown in low osmolality media which increases OmpK35 expression[43,44,50]. Loss of expression of OmpK35 is ubiquitous in ST258 strains and on its own results in an 8-fold change in the minimum inhibitory concentration (MIC) of the carbapenem, meropenem[51,52]. This loss of expression is commonly brought about by a frame-shift mutation that results in the introduction of a premature stop codon. The resulting mature form of the protein is 19 amino acids long, non-functional and no longer localizes to the OM.

Loss of expression of OmpK36 is much less common. Instead, amino acid insertions are seen in L3 which alter the structural properties of the porin channel without resulting in significant changes to bacterial fitness[52,53]. These insertions were determined based on comparison to the OmpK36 isoform expressed in a carbapenem sensitive laboratory strain of KP, ATCC 43816 which also serves as the parental wild-type (WT) strain in this work. Three L3 insertions have been described in ST258 strains: a glycine-aspartic acid (GD) insertion, threonine-aspartic acid (TD) insertion and a single aspartic acid (D) insertion[54]. The GD insertion results in a conformational change to L3 that reduces the diameter of the porin channel by 26%[52]. This shift is stabilised by a salt bridge between D114 and R127 on the barrel face of the porin subunit. Consequently, this is associated with a 16-fold increase in the MIC of meropenem. Structural analysis of the TD and D insertions reveal that they result in a 41% and 8% pore constriction respectively but had a similar effect on meropenem MIC as the GD insertion[54].

Although mutations resulting in structural modifications in OmpK36 are more prevalent, there have since been other reports which highlight alternative mechanisms that reduce the expression of this porin to enhance carbapenem resistance[55,56]. One mutation that has been

shown to emerge *de novo* in response to carbapenem treatment is a synonymous c>t mutation at position 24[57]. This base substitution results in the formation of a secondary mRNA hairpin structure that has been proposed to obstruct efficient translation of the *ompK36* transcript resulting in greatly reduced expression of OmpK36 and a corresponding increase in carbapenem MIC.

1.2.4 Plasmid-mediated resistance

Plasmid-mediated resistance is an additional mechanism of resistance which is particularly prevalent in KP as this species appears to support plasmid uptake and carriage much better than many other hospital-associated pathogens[58]. *In vitro* studies show that plasmid carriage has a lower fitness cost in KP compared to *E. coli* and KP strains are capable of long-term plasmid maintenance *in vivo*[59–61]. Consequently, analysis of KP isolates deposited in GenBank revealed that strains from this species are associated with a greater number of plasmids than several other species of interest[58]. This elevated plasmid load combined with the ability to inhabit a wide range of environmental niches makes KP an important species facilitating the flow of AMR genes from environmental microbes to clinically important pathogens including other ESKAPE pathogens.

Unsurprisingly, carbapenem resistance in KP is also largely attributed to the acquisition of plasmids encoding carbapenemases, β-lactamase enzymes which hydrolyse and inactivate carbapenems (Figure 1.2). Low level resistance to carbapenems can also result from the expression of other β-lactamases when present in combination with other mechanisms of resistance such as reduced OM permeability[62]. Many carbapenemases are intimately associated with specific transposable elements or insertion sequences which integrate themselves into conjugative plasmids thus enabling their rapid dissemination via horizontal gene transfer[63,64]. The acquisition of these plasmids by specific strains can give rise to high-risk resistant sequence types such as ST258[65].

β-lactamases are encoded by *bla* genes and can be classified into one of four Ambler classes based on their sequence similarity and mode of action. Enzymes belonging to classes A, C and D are serine β-lactamases while class B enzymes are metallo-β-lactamases whose catalytic activity is reliant on a bivalent metal ion (Table 1.2)[66]. Carbapenemases typically fall into classes A, B and D. Enzymes which are commonly associated with CRKP include the Class A KP carbapenemases or 'KPC'-family enzymes, the Class B New Delhi metallo-β-lactamases (NDM) and the Class D OXA-family β-lactamases[67].

Table 1.2. Ambler molecular classification of common β-lactamases.

Ambler Class	Enzymes
A	Class A carbapenemases (KPC, GES) ESBLs (TEM, CTX-M)
B	Metallo-β-lactamases (NDM, VIM, IMP)
C	AmpC
D	Oxacillinases (OXA-1, OXA-15) Carbapenem-hydrolysing class D β-lactamases (OXA-23, OXA-48)

1.2.5 pKpQIL

The first *bla*$_{KPC}$-carrying plasmid to be sequenced, pKpQIL, was identified in an ST258 clinical isolate of KP in Israel in 2006[68]. Strains carrying this plasmid displayed high levels of carbapenem resistance associated with the expression of the KPC-3 carbapenemase[69]. Sequencing of pKpQIL revealed it to be approximately 113 kb in size, with sequence similarity to two other KP plasmids, pKPN4 and pNYC. A large region (76%) of the plasmid displays 99% identity with pKPN4, one of five plasmids isolated and sequenced from a multidrug-resistant isolate, MGH78578, in 1994. This region includes genes involved in bacterial conjugation, resistance, and plasmid maintenance. Although pKPN4 encodes other resistance genes (*aadA*, *bla*$_{OXA-9}$ and *bla*$_{TEM-1}$), only *bla*$_{TEM-1}$ remains intact and encodes a functional product in pKpQIL[68]. The remaining 14% which corresponds to part of pNYC contains the Tn*4401*a transposable element housing *bla*$_{KPC-3}$[70].

Since the initial discovery and characterization of pKpQIL, many pKpQIL-like plasmids have been recovered from KP isolates. Most of these isolates belong to ST258 suggesting a strong co-adaptation between the plasmid and these strains. Indeed, several reports attribute the expansion of pKpQIL-bearing ST258 isolates to clonal expansion[71,72]. Despite this, the presence of pKpQIL-like plasmids among multiple sequence types of KP and other species of *Enterobacteriaceae* in England's North-West, suggests that horizontal gene transfer also plays a major role in the dissemination of these plasmids[73].

1.3 Bacterial conjugation

Bacterial conjugation is a form of horizontal gene transfer (HGT) that was first described by Lederberg and Tatum in the 1940s following the discovery of the F plasmid[74]. During this process, DNA is transferred from one cell (the 'donor') to another (the 'recipient') in a contact-dependent, unidirectional manner. Like other forms of HGT such as transformation and transduction, conjugation plays an important role in the adaptation and evolution of bacterial strains to their environment[75]. It is also clinically important as it facilitates the acquisition and dissemination of many virulence and AMR genes[76,77].

The steps leading up to recipient acquisition of DNA can be mechanistically divergent depending on the bacteria involved and the type of system facilitating conjugation. This work focuses primarily on conjugative systems expressed in Gram-negative bacteria, specifically enterobacterial species. In these bacteria, DNA transfer is facilitated by three main complexes – a type IV secretion system (T4SS), the relaxosome complex, and the type IV coupling protein (T4CP)[78]. Functionally, the T4SS is involved in mating pair formation (MPF), the relaxosome facilitates DNA transfer and replication, and the T4CP recruits the relaxosome to the T4SS, linking the two complexes[79]. The genes encoding the subunits which comprise these complexes are often carried on mobile genetic elements (MGEs) including conjugative or 'self-transmissible' plasmids. Mobilizable plasmids which encode part of or a complete relaxosome

as well as an origin of transfer (*oriT*) may also utilize a compatible T4SS encoded *in trans* within the same cell to be transferred via conjugation[80].

1.3.1 Plasmid classification

As previously mentioned, conjugative plasmids can be mobilized by distinct mechanisms depending on their encoded transfer machinery. Plasmid classification has greatly facilitated the ability to predict the transfer properties and host range of these MGEs. The most widely employed form of plasmid classification is incompatibility (Inc) grouping complemented by its modern counterpart, replicon sequence typing.

Following the identification of the F plasmid in (also referred to as the 'F fertility factor' or, simply, 'F'), other plasmids were isolated including colicinogenic (Col) plasmids and resistance gene-bearing R plasmids. Early classification of these plasmids was based on the concept of incompatibility which refers to the inability for two plasmids to be stably propagated within the same host cell and results from the relatedness of replication elements in those plasmids[81]. To test for incompatibility, an uncharacterised plasmid carrying a selection marker would be introduced into a host cell carrying a resident plasmid with a different selection marker by conjugation. Cells would then be selected for uptake of the introduced plasmid and assessed for the loss of the resident plasmid. Where this occurred, the plasmids would be assigned to the same incompatibility group[82].

As more plasmids were identified, incompatibility testing no longer served as a feasible means of classification due to the time-consuming nature of this protocol. An additional limitation of this method is that not all plasmids can be conjugated into the host cell due to factors such as surface exclusion, which inhibits conjugative entry of a plasmid into cells already carrying the same plasmid and is distinct from incompatibility[82]. Moreover, today it is known that single base changes in the elements which regulate incompatibility, or the presence of multiple

replicons can render two plasmids compatible[83]. Thus, the results of incompatibility testing may not accurately reflect plasmid genetic relatedness.

Modern methods of incompatibility testing now rely on PCR-based replicon typing (PBRT) which uses sequence similarity of the plasmid-encoded replication (*rep*) region DNAs as a classification rule[84]. Doing so has revealed that many plasmids may harbour multiple or recombinant replicons which can have implications on plasmid host replication range. At present, commercially available PBRT kits can identify up to 28 different replicons[85]. This includes 6 different subtypes from the incompatibility group F (IncF) plasmid family to discern between replicons that are predominantly associated with a specific bacterial species, namely FII_K of *Klebsiella* spp., FII_Y of *Yersinia* spp. and FII_S of *Salmonella* spp.

1.3.2 Classification of conjugative systems

Conjugative plasmids belonging to the same replicon types generally encode genetically similar conjugative machinery. However, it may be difficult to infer the type of conjugative machinery present on plasmids with multiple replicons. PBRT is also limited to the currently available array of primers that have been designed based on known-plasmid types[86]. Thus, many emerging plasmids may not be easily typed. Moreover, as incompatibility groupings rely on the replication properties of plasmids for classification, they do not necessarily reflect the mobilization properties of these MGEs.

Other classification schemes have been proposed for conjugative plasmids including those based on the mobilization or 'MOB' region or the relatedness of regulatory proteins[87,88]. For this work, the most informative classification scheme groups conjugative systems into one of four MPF groups – MPFF, MPFT, MPFI and MPFG[80]. This classification considers the homology of the components making up the T4SS and additional accessory factors which are involved in MPF. Although this method of classification requires that the plasmid of interest,

or at the very least its transfer genes, be fully sequenced, it is perhaps the most informative with regards to understanding the conjugative properties of a plasmid.

The representative T4SSs from each group are derived from F (IncFIB) for MPFF, the *vir* system for MPFT, R64 (IncI) for MPFI and ICEHIN1056 for MPFG[80]. Of the 4, MPFG systems are poorly described as they are rarely found in plasmids from clinical isolates. Transfer gene products that are conserved across the three systems are listed in Table 1.3 along with their known functions. Apart from structural genes which make up the conjugative T4SS, most conjugative proteins are not conserved across the three systems.

Table 1.3. Transfer gene homologues from various conjugative systems.

Function/Description	MPFF (F)	MPFT (pTi)	MPFI (R64)	Reference
T4SS components				
IMC protein	TraL	VirB3		[89]
Periplasmic protein	TraE	VirB5	TraM	[88,90,91]
OMC protein	TraK	VirB9	TraN	[90]
OMC protein	TraB	VirB10	TraO	[89]
OMC lipoprotein	TraV	VirB7	TraI	[90]
ATPase	TraC	VirB4	TraU	[89]
IMC protein	TraG	VirB6/VirB8		[88,92]
T4CP, ATPase	TraD	VirD4	TrbC	[89]
ATPase	Absent	VirB11	TraJ	[89]
Pilus synthesis and dynamics				
Pilin precursor	TraA	VirB2	TraX	[90]
Pilin chaperon	TraQ			
Pilin N-acetylation	TraX			
Pilus extension	TraW			
Pilus extension	TraF			
Pilus extension	TraH			
DNA transfer	TraU			
Pilus retraction	TrbI			
Unknown	TrbC		TrbB	[90]
Unknown	TrbB			
Others				
Mating pair stabilization	TraN			
Mating pair stabilization and entry exclusion	TraG			
Entry exclusion	TraS			
Surface exclusion	TraT			
Relaxase	TraI	VirD2		[92]
Lytic glycosyltransferase	Orf169	VirB1		[88]

MPFF conjugative systems express long, flexible pili which permit conjugation in both liquid media and on solid substrates[90,93]. These systems account for more than a third of conjugative plasmids isolated from Gammaproteobacterial species and are particularly highly represented amongst *Enterobacteriaceae* species[80]. IncF, IncH and IncA/C plasmids are typically transferred via these systems. While the transfer genes in conjugative IncF plasmids make up a contiguous operon of approximately 40 kb, in IncH and IncA/C plasmids, they are split into

two operons (Figure 1.4A)[94–96]. The host replication range of plasmids in these groups vary with IncF plasmids occupying a narrow host range, IncH plasmids an intermediate range and IncA/C plasmids predicted to have a broad host range[97].

MPFT systems express short, rigid pili and generally only support conjugation on solid media[93]. These are the simplest conjugative systems and are often associated with intermediate or broad host range plasmids from incompatibility groups N, W and P[98]. Unlike other MPF groups, MPFT systems can also facilitate DNA transfer into non-prokaryotic cells as exemplified by the prototypical Vir system which mediates transfer of the tumour-inducing (Ti) plasmid of *Agrobacterium tumefaciens* into plant cells[99]. Plasmids such as R388 (IncW) have served as model plasmids for both functional and structural studies and many cloning vectors have been engineered to carry transfer elements from RP4 (IncP) to take advantage of the broad transfer range of this plasmid[91,100–102]. The transfer genes of several MPFT systems are shown in Figure 1.4B. Notably, genes are arranged into two operons in RP4, and as a single contiguous transfer operon of approximately 15 kb in R388[103,104].

MPFI systems can express both a thick rigid pilus which facilitates surface mating and a thin flexible type IV pilus (T4P) which facilitates transfer in liquid media[105]. These systems are often encoded on plasmids from the IncI, IncL and IncM groups and include several proteins which are homologous to the effector-delivering Dot/Icm T4SS from *Legionella pneumophila*[106]. This group is represented by the *Salmonella* Typhimurium plasmid R64 which encodes the T4P from the *pil* locus located upstream of the *tra* operon (Figure 1.4C)[107]. The T4P does not appear to be a universal feature of MPFI systems as the *Citrobacter freundii* IncM resistance plasmid pCTX-M-3 does not contain a *pil* locus[108]. Interestingly, despite sharing a majority of the *tra* genes found on R64 (Figure 1.1C), pCTX-M3 is capable of conjugation in liquid media without expressing the T4P demonstrating heterogeneity in MPFI conjugation mechanisms[108]. The OXA-48 carbapenemase-encoding plasmid pOXA-48a carries a transfer operon that is highly similar to pCTX-M3 and similarly lacks genes encoding the T4P[63].

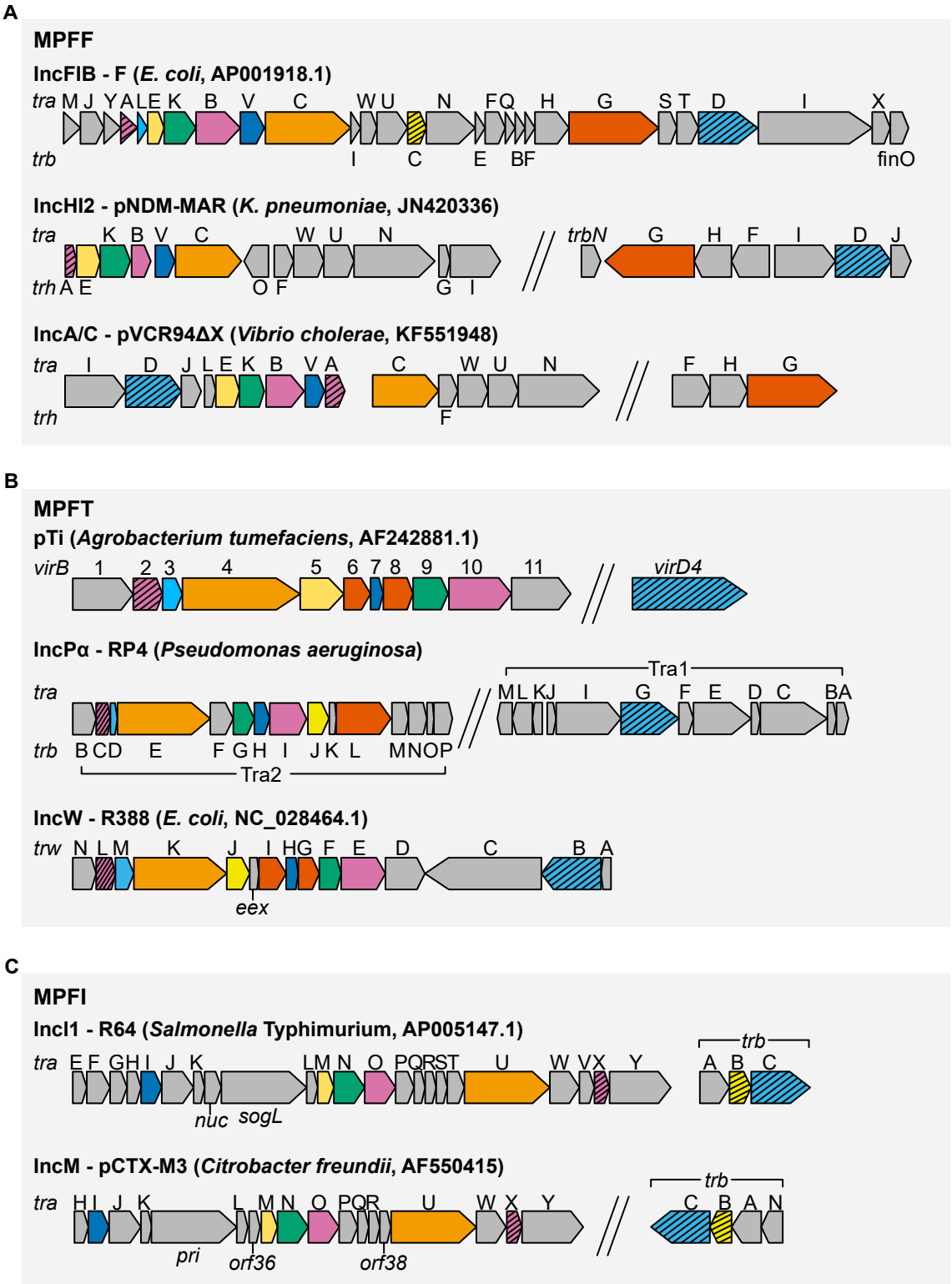


Figure 1.4. Genetic arrangement of transfer genes from different conjugative systems.

Schematic of transfer gene operons from (**A**) MPFF, (**B**) MPFT and (**C**) MPFI conjugative systems. Genes have been coloured to show conserved homologues as listed in Table 1.1. The length of the genes shown are not proportional to actual gene length. Accession IDs for each plasmid have been included where they are available. The schematic for the RP4 transfer operons was adapted from Lawley et al., (2003)[103].

1.3.3 An overview of IncF plasmid conjugation

IncF plasmids are found almost exclusively in *Enterobacteriaceae* species and include well-known plasmids such as the pMAR7 [enteropathogenic *Escherichia coli* (EPEC)] and pSLT (*Salmonella* Typhimurium) virulence plasmids, and the R100 resistance plasmid (*Shigella flexneri*)[109–111]. Many IncF plasmids also serve as model plasmids for conjugation studies, such as the prototypical F plasmid (IncFIB), R100-1 (IncFII) a derepressed variant of R100, and pED208 (IncFV)[112,113]. These studies have revealed the steps involved during conjugation, from the initiation of DNA transfer to the formation of a transconjugant cell (Figure 1.5).

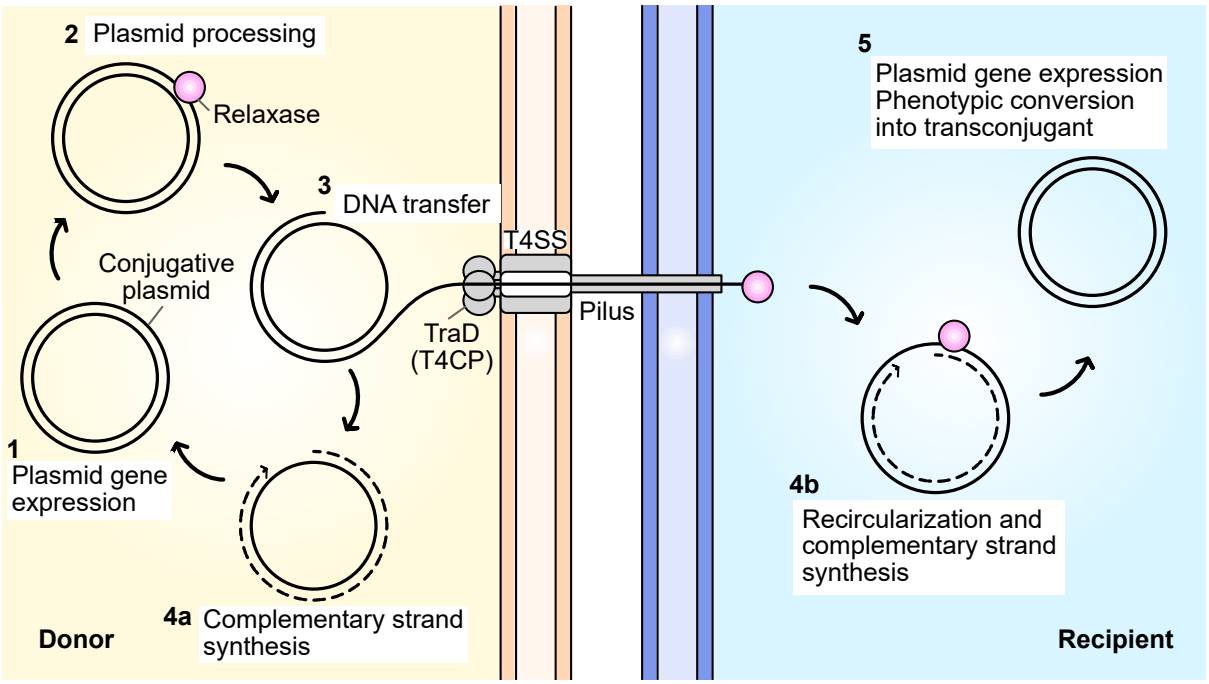


Figure 1.5. Schematic of IncF plasmid conjugative transfer.

(**1**) Conjugation begins with the expression of transfer genes involved in the assembly of the conjugative T4SS. The T4SS and its associated pilus is required for establishing contact with a recipient cell resulting in mating pair formation. (**2**) Other transfer gene products which constitute the relaxosome bind to the *oriT* and the relaxase (TraI) catalyses a nicking reaction to prepare the plasmid for conjugative transfer. (**3**) The TraI-bound transfer strand is recruited to the T4SS by the T4CP (TraD) which initiates DNA transfer into the recipient cell. (**4a**) In the donor cell, a complementary strand is synthesised by rolling circle replication to convert the ssDNA plasmid into dsDNA. (**4b**) In the recipient, TraI facilitates recircularization of the single-stranded transfer strand and recipient DNA polymerase is recruited to initiate complementary strand synthesis. (**5**) Gene expression off the newly formed dsDNA plasmid leads to the phenotypic conversion of the recipient into a transconjugant.

1.3.4 Transcriptional regulation of the *tra* operon

The genes required for IncF plasmid conjugation are contained within a single contiguous operon referred to as the *tra* operon with the majority of these genes under the control of the P_Y promoter[114]. Transcriptional regulation of this operon is highly regulated to ensure that genes are only expressed during optimal conditions as constitutive expression is believed to impart a substantial fitness cost on donor cells[115].

The most well-described regulation mechanism involves the post-transcriptional regulation of TraJ, a positive regulator of *tra* operon transcription, by the FinOP fertility inhibition system (Figure 1.6)[116]. FinO is a 21.2 kDa RNA-binding protein which binds to the antisense RNA, FinP, and protects it from degradation by RNaseE[117]. FinP is constitutively transcribed from the complementary strand encoding the 5' non-translated region of *traJ* which includes the ribosome binding site. It folds into a structure consisting of two stem loops with the first stem loop being complementary to the ribosome binding site present on the *traJ* mRNA[118]. The formation of this duplex structure prevents *traJ* mRNA from being translated leaving the expression of the *tra* operon in a repressed state.

The F plasmid harbours a naturally derepressed *tra* operon resulting from the disruption of *finO* by an IS3 insertion element[119]. The absence of FinO results in the rapid degradation of FinP, lowering intracellular concentrations of the antisense RNA. As a result, *traJ* mRNA is free to be translated, activating the P_Y promoter and *tra* operon transcription. In derepressed systems, all plasmid-containing cells are conjugation-proficient which greatly increases plasmid transfer efficiency. This contrasts with plasmids with repressed transfer like R1 and R100, of which only 0.1% of cells participate in conjugative transfer of the plasmid into recipients[114].

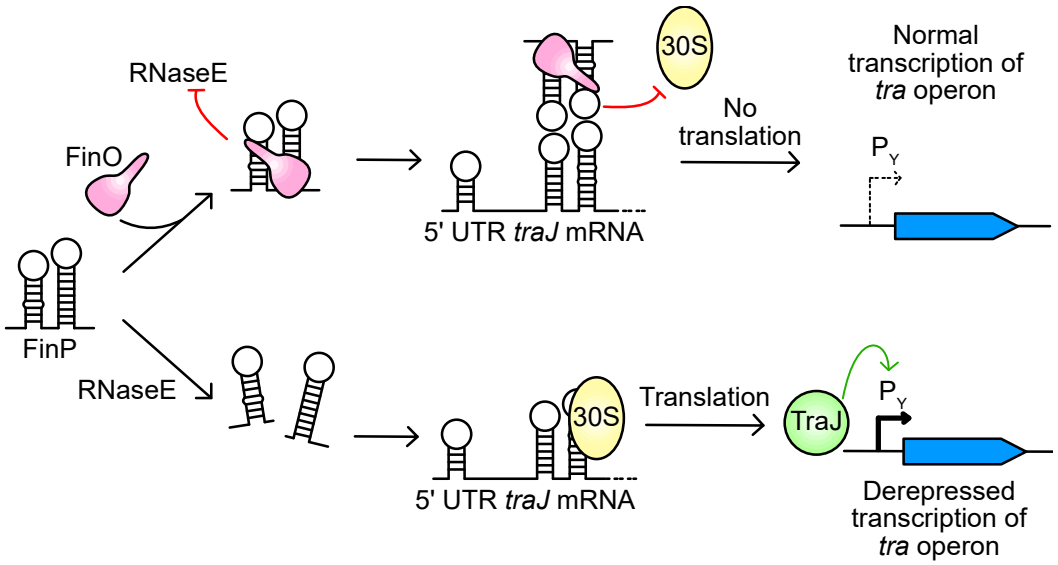


Figure 1.6. Schematic of FinOP system.

The FinP antisense RNA is protected from RNAseE degradation by the RNA-binding protein FinO. This allows it to form a duplex with the 5' untranslated region (UTR) of *traJ* occluding 30S ribosomal subunit from the ribosome binding site. This prevents translation of *traJ* mRNA and there is no positive regulation of the *tra* operon. In the absence of FinO, positive regulation of the *tra* operon by TraJ results in derepressed transcription of the *tra* operon.

1.3.5 The conjugative pilus

The majority of *tra* operon encoded gene products are involved in the biogenesis and assembly of the T4SS-associated conjugative pilus (Figure 1.7A). Recently, the structure of the pED208-encoded T4SS was determined *in situ* by cryo-electron tomography (cryo-ET) while the structure of the OM core complex was determined by cryo-electron microscopy (cryo-EM)[113,120]. Based on these structures, the overall architecture of the T4SS and the role of various conjugation subunits during pilus biogenesis is becoming increasingly well understood (Figure 1.7B).

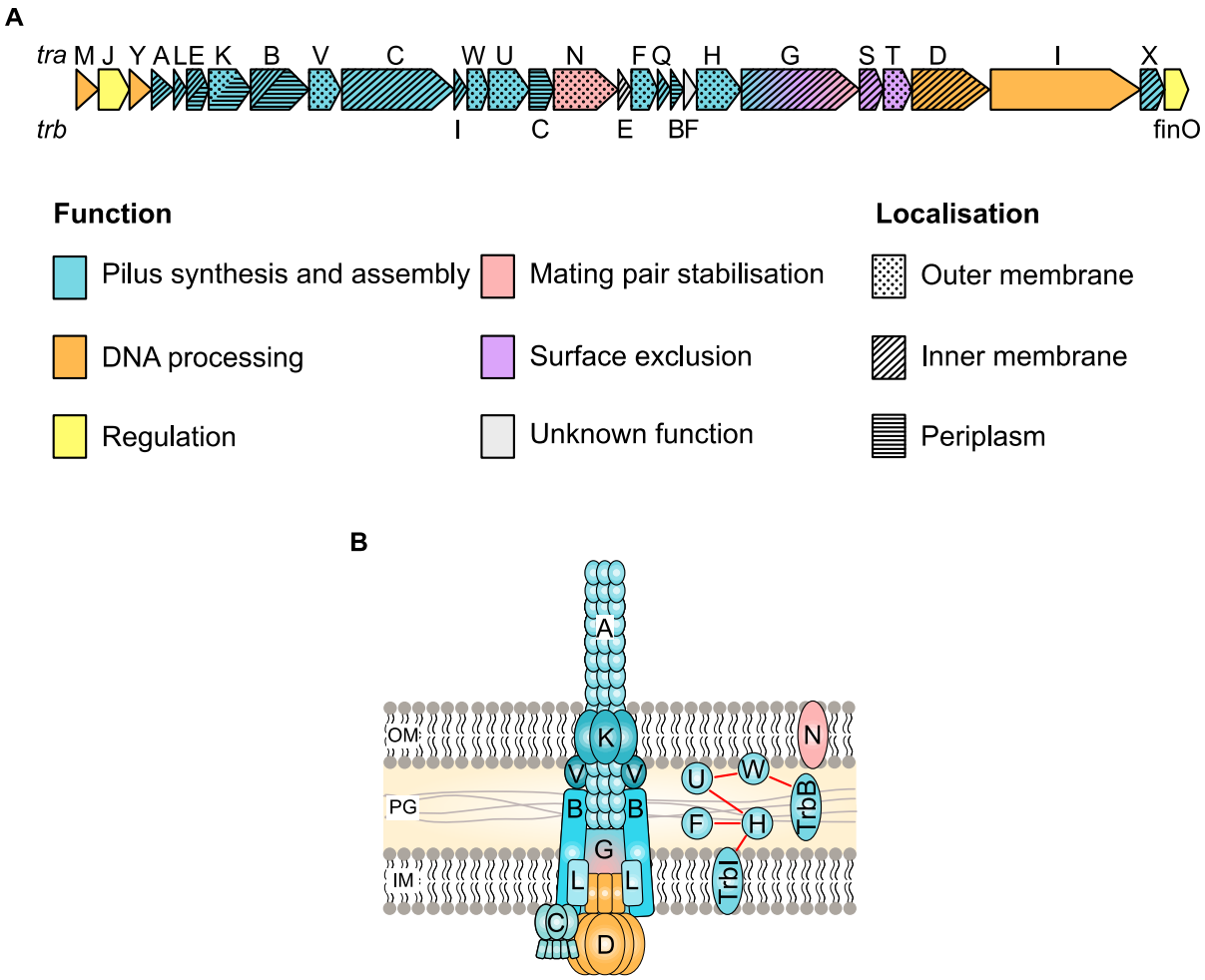


Figure 1.7. Function and localisation of F plasmid transfer gene products.

A. Genetic arrangement of the *tra* operon in IncF plasmids. The function and subcellular localization of each gene product is shown. **B**. The architecture of the F plasmid conjugative T4SS. Subunits have been coloured according to function to match panel **A**. Subunits which were not found to associate with the T4SS directly are shown on the right with the red lines indicating previously determined interactions. PG = peptidoglycan. Figure adapted from Hu, Khara and Christie (2019)[113].

The pilus is an essential component required for IncF plasmid conjugation and is assembled as a thin, flexible filament comprising polymerised pilin subunits encoded by *traA* (Figure 1.7B)[121]. F plasmid *traA* is expressed as a 121 amino acid pro-pilin peptide which is processed into a mature 70 amino acid pilin subunit by TraQ and TraX[122]. These mature pilin subunits accumulate in the inner membrane (IM) prior to being assembled into the pilus by 11 *tra* gene products. TraL, E, K, C, and G play a role in the assembly of the pilus tip and TraB, V, W, F

and H play a role in pilus extension[123]. While TrbC is essential for pilus biogenesis, its exact function remains unknown[124]. Three other *tra* proteins support pilus structure and function – TraP acts to stabilise the extended filament, TraU has a role in DNA transfer and TrbI is required for pilus retraction[125–127].

The structures of F-like conjugative pili encoded by the F plasmid, pED208 and the KP carbapenem resistance plasmid pKpQIL (IncFII$_K$) have been solved by cryo-EM, revealing that pilin subunits form helical assemblies with phospholipid molecules at a stochiometric ratio of 1:1[128,129]. The assembled filament has a lumen of ~28 Å in diameter with an external diameter of ~87 Å and extends to a length of approximately 10 nm[113,128,130].

While the structure of the pilus has been described, its function remains a contentious subject in the field. It is widely accepted that conjugative pili are important for establishing initial contacts with recipient cells during the initial stages of MPF[131]. How these contacts are formed remains unknown as a receptor for the pilus on recipient cells has not been identified. Some have speculated that, like the type VI secretion system, initial contacts established by the conjugative pilus of the T4SS may not rely on a specific recipient cell receptor[132]. The identification of a pilus receptor is complicated by the fact that the nature of the pilus tip is unknown. As it is homologous to VirB5 which localizes to the tip of the Ti-pilus, TraE was proposed to serve as an adhesin which caps the pilus[133]. However, there is no direct evidence to confirm this theory. Nevertheless, using microscopy to observe F-pili in live cells has shown that it is a highly dynamic structure which can retract upon contacting a recipient cell, drawing it closer to the donor cell to facilitate the formation of a mating pair[134].

It is unclear if the role of the pilus is limited to these initial stages of MPF and several theories have posited that the pilus has a secondary role as the conduit through which DNA is transferred into the recipient. This is supported by observations of low frequency plasmid acquisition by recipients which are not intimately attached to donor cells[130]. Moreover, based

on the cryo-EM structures of the pilus, the phosphate groups of the lipid molecules were found to face the lumen resulting in the overall surface being moderately negatively charged[128]. This could facilitate the repulsion of DNA through the pilus as a similar feature was observed in the contractile tail of the T4 phage which functions to deliver dsDNA into cells[135]. In addition, the luminal diameter is sufficiently wide to allow for the passage of single-stranded DNA (ssDNA). As the transferred ssDNA is covalently linked to TraI, the luminal diameter would also need to accommodate TraI in its unfolded state[136]. The luminal diameter of the T4SS is very similar to the diameter of the *espA*-encoded filament from the type III secretion system (T3SS) which delivers unfolded effector proteins into host cells[137]. This would suggest that the conjugative pilus could accommodate the relaxase-bound transfer strand. Despite these, the role of the pilus as a DNA conduit is challenged by proponents of an alternative mechanism for DNA traversal across the recipient OM through a 'mating bridge' or 'mating pore'. This alternative theory may stem from observations of conjugative systems in Gram-positive bacteria which do not tend to express extracellular pili, thereby requiring an alternative conduit for DNA entry into the recipient cell[138].

1.3.6 Mating pair stabilization

When conjugating cells were observed by electron microscopy, they were seen to form mating aggregates containing two or more cells[139]. These aggregates were resistant to disruption by shear forces and the addition of sodium dodecyl sulphate (SDS). Closer inspection of the cells revealed that they form tight 'mating junctions' characterised by intimate wall-to-wall contact[140]. Initially, these junctions were assumed to be the result of interactions between the tip of the conjugative pilus and a recipient cell receptor. However, it was later found that mutations in two genes, *traN* and *traG*, specifically affect the formation of mating aggregates without abolishing pilus biogenesis, suggesting that this process is distinct from pilus-mediated initial contacts with recipient cells[141]. The process which leads to the formation of these tight mating junctions is now commonly referred to in the literature as mating pair stabilization (MPS).

In the context of the F plasmid, MPS is not essential for conjugation as several mutations in *traN* and *traG* reduced but did not abolish plasmid transfer[142,143]. Moreover, the reduction in conjugation efficiency was markedly greater when cells were allowed to mate in liquid media as opposed to on a solid substrate, suggesting that MPS plays a more crucial role in liquid mating[142]. This is similar to what has been observed with MPFI plasmids such as R64 which rely on the T4P for efficient mating specifically in liquid media and, as recent studies suggest, in the gut microbiome[144,145].

Since being identified as MPS proteins, several studies have been performed to characterise both TraN and TraG. The F plasmid TraN is an OM protein which, when expressed, contains 602 amino acids[146]. Following cleavage of the N-terminal signal sequence, the mature protein contains 584 amino acids. Analysis of TraN homologues from different IncF conjugative plasmids revealed that 22 cysteine residues are conserved amongst homologues of this protein, of which 6 were shown experimentally, to be important for optimum plasmid transfer[147]. OM proteins with such a high number of cysteine residues are extremely uncommon which has made predicting the structure of TraN difficult. Moreover, it shares little sequence identity with other known OM proteins that may facilitate cell-to-cell contacts such as adhesins. Using epitope fusion experiments, Klimke et al., provided insight on the topology of TraN revealing three extracellular loops which were predicted to be involved in receptor recognition[147]. These loops corresponded to a region of approximately 200 amino acids in TraN which shared very low sequence similarity between TraN from F and TraN from R100-1 (Figure 1.8). The C-terminal domain of the protein, however, is highly conserved.

TraG is a multifunctional *tra* gene product which localizes to the IM and is approximately 102 kDa. Through genetic analysis, it was determined that the N-terminus is required for pilus assembly while the C-terminus plays a role in MPS[141]. Initially, *traG* was thought to encode for two products, full length TraG and a smaller cleavage product, TraG*, containing the C-terminal which localizes to the bacterial periplasm[143]. However, it was later determined that

only full length TraG is functional[148]. Not much else is understood about the role of TraG during MPS and its relationship with TraN is not known as attempts to show that both proteins interact have thus far been unsuccessful[147]. As the N- and C-terminal domains are homologous to VirB6 and VirB8 from MPFT systems respectively, TraG may similarly form part of the inner membrane complex and the stalk structure which acts as a base for pilus biogenesis[90,91].

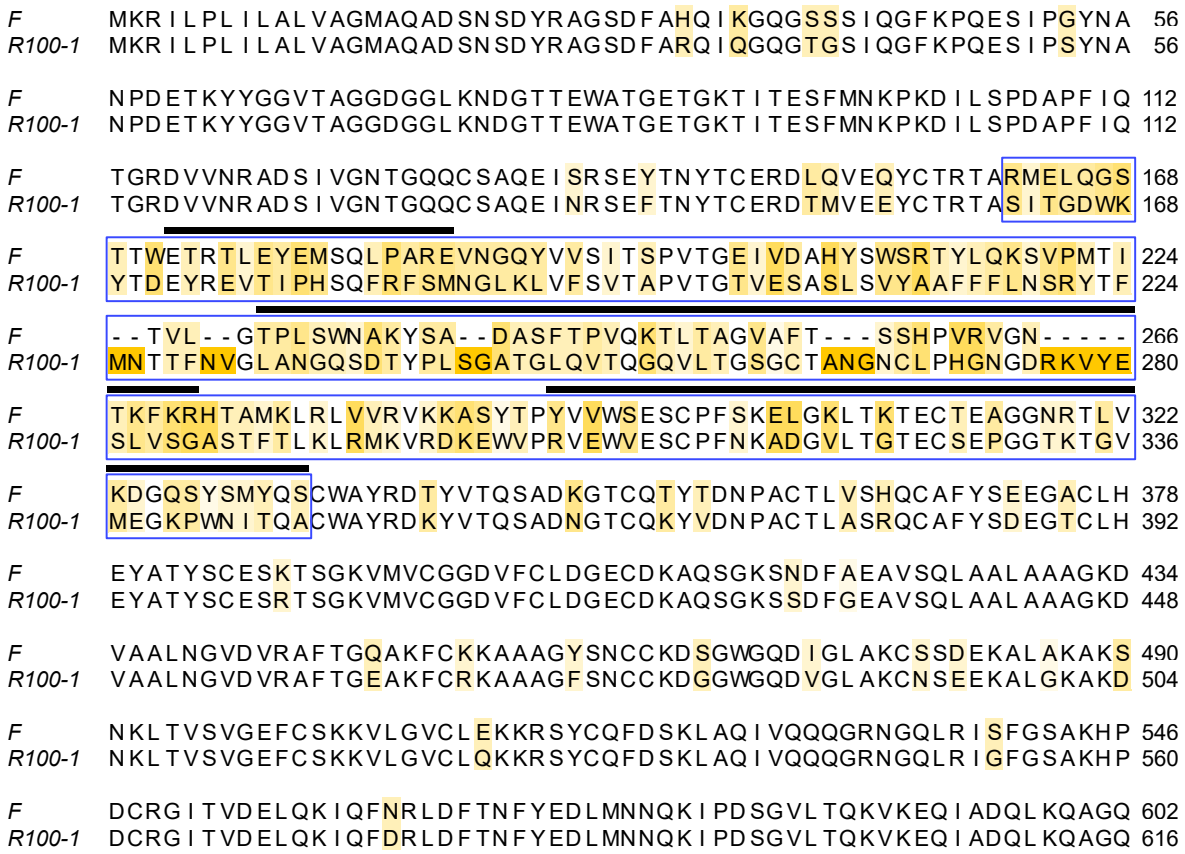


Figure 1.8. Sequence alignment of TraN from F and TraN from R100-1.

Global amino acid sequence alignment of full length TraN from the F plasmid and R100-1. Residues were coloured according to amino acid conservation with the darker shading representing less conserved residues. The variable region identified by Klimke and Frost is highlighted in the blue box and includes residues 162-333 in TraN from F. The black bars above residues in the variable region indicate the likely position of surface exposed loops as determined from topology experiments.

1.3.7 The role of recipients in bacterial conjugation

Despite decades of research characterizing the various *tra* gene products, there is still much to learn about their roles during IncF plasmid conjugation. Nonetheless, recipient involvement during plasmid transfer is even less well understood. The earliest studies aiming to elucidate the role of recipient cells investigated the effect of nutrient starvation on plasmid uptake[149]. Restricting nutrient availability to recipients during conjugation resulted in a reduction in the overall yield of transconjugants. This effect was modest but suggested that recipients also play an active role during conjugation that can affect the overall efficiency of plasmid transfer.

Another area of significant interest has been in elucidating the role of recipients during MPF, specifically identifying a receptor for the conjugative pilus which has remained elusive. In the 1970s, several *E. coli* mutants were isolated which displayed resistance to F plasmid-induced lethal zygosis[150]. Characterization of these strains termed 'Con$^-$ mutants' showed that they carried mutations in the genes encoding the OM protein OmpA and the inner core of LPS[151]. Three classes of OmpA mutants which were affected in F plasmid uptake were isolated. These were: 1) mutants which did not express OmpA, 2) mutants which expressed reduced amounts of OmpA and 3) mutants which showed normal OmpA expression but contained missense mutations that interfered with conjugation[152]. These recipient-associated mutations had a more pronounced effect than what was observed during nutrient starvation with conjugation frequency being reduced by approximately 2-3 log folds[150].

Interestingly, mutations in OmpA and LPS specifically affect transfer of the F plasmid but not R100-1[150]. Although Con$^-$ mutants were similarly isolated for R100-1, the identity of the recipient factor affecting transfer could not be determined from OM protein analysis[153]. Taking advantage of this specificity, OmpA and LPS were determined not to be the receptors for the F-encoded conjugative pilus as substitution of *traA* on F with *traA* from R100-1 did not abrogate the effect of OmpA mutations[154]. Instead, they were found to affect MPS as dependency on these recipient factors was associated with TraN, specifically the variable

central region of the protein (Figure 1.8)[112]. Despite strong genetic evidence to suggest that TraN from F interacts with recipient OmpA to mediate MPS, attempts to show that these proteins interact directly using co-immunoprecipitation and yeast two hybrid approaches were unsuccessful[147]. Thus, it is not known if the role of OmpA during MPS is direct or indirect. Similarly, it is unclear how TraN and LPS cooperate during MPS but the effect of LPS mutations had a more modest effect on plasmid uptake compared to OmpA mutations[112].

Studies have also been performed to identify recipient factors for plasmids expressing non-MPFF conjugative systems. Transfer of the IncI plasmid R64 which encodes an MPFI conjugative system was also shown to be affected by recipient LPS during conjugation in liquid implying that these differences can be attributed to effects on MPS[144]. MPS in these systems is mediated by PilV adhesins located at the tip of the T4P. The C-terminal segments of the PilV adhesins can be exchanged by DNA recombination within the R64 shufflon to form 7 different adhesins[105]. Genetic analysis determined that specific moieties of LPS were found to act as receptors for the various adhesins[155]. This was proven using filter overlay experiments which demonstrated the direct binding of specific PilV C-terminal segments to LPS[156]. These findings were recently replicated using the IncI plasmid TP114 which showed that the T4P also plays a role in increasing conjugation efficiency in the gut[145].

In contrast to MPFF and MPFI systems, analysis of MPFT system-expressing plasmids failed to identify specific recipient factors required for conjugation. High throughput assays were developed to assess conjugative transfer of the IncW plasmid R388 and the IncP plasmid RP4[100,157]. These plasmids were tested against recipients from the Keio collection which contains 3908 mutants, each with a single in-frame deletion covering most of the non-essential *E. coli* genes[158]. None of the recipients showed significantly reduced R388 uptake when compared to a negative control recipient expressing R388 surface exclusion proteins. Following validation by selection-based plating of transconjugants, the greatest reduction in conjugation frequency was found to be only approximately 5-fold compared to the WT strain

and was associated with the *rfaC* mutant[100]. RfaC is a heptosyltransferase involved in biosynthesis of the LPS inner core which could suggest that LPS has a conserved role during conjugation as it also serves as a recipient factor during F and R64 transfer. However, no recipient mutants in the Keio collection were identified as showing reduced transfer efficiency of RP4[157].

As demonstrated above, conjugation has typically been characterized using prototypical plasmids and laboratory strains of bacteria. Comparatively few studies focus on investigating the transfer properties of clinically relevant plasmids. Recently, studies aimed at identifying factors which influence conjugative transfer of the KP resistance plasmid pKpQIL determined that environmental factors such as temperature and mating substrate can affect conjugation efficiency[159]. The relatedness of donor and recipient species also appeared to influence overall transfer frequency. However, these *in vitro* findings did not show a strong correlation with clinical epidemiological data suggesting that other factors play a more important role in influencing conjugation *in vivo*[159]. Although this study did not identify factors of interest, it highlights the advantage of studying clinically relevant plasmids where epidemiological data is available as a benchmark for determining the significance of *in vitro* findings.

1.4 Aims and Objectives

There is considerable interest in identifying factors which affect conjugation of the KPC-encoding plasmid pKpQIL to better understand how this plasmid disseminates within clinical isolates. Based on the close association of pKpQIL-like plasmids with clinical ST258 KP isolates, the aim was to investigate if the ST258 OM porins, OmpK35 and OmpK36 play a role in pKpQIL conjugation. In line with that, the objectives were to:

1) Determine the effect of ST258 porin mutations during conjugative plasmid uptake
2) Identify the donor component which mediates the observed effect
3) Describe the mechanism underlying donor-recipient interactions
4) Apply the findings to publicly available plasmid datasets

2. Materials and Methods

Several of the methods outlined in this section have been similarly described in the publications which arose from this work[129,160]

2.1 Bacterial strains and plasmids

The bacterial strains and conjugative IncF plasmids used in this work are listed in Table 2.1 and Table 2.2 respectively. Unless otherwise stated, bacteria were cultured in lysogeny broth (LB) at 37°C, 200 RPM. When needed, antibiotics were used at the following concentrations: ertapenem (0.5 µg/ml), streptomycin (50 µg/ml), kanamycin (50 µg/ml), gentamicin (10 µg/ml).

Table 2.1. List of bacterial strains used in this work.

Strain	Description	Source
Cloning strains		
CC118λpir	Expresses the Pi protein for the replication of plasmids with the R6K origin.	Lab stock
E. coli 1047 pRK2013	Triparental conjugation helper strain. Kanamycin resistant	Lab stock
***K. pneumoniae* strains**		
ICC8001	Parental strain of *K. pneumoniae* ATCC43816 serially passaged *in vitro* on Rifampicin (100 µg/ml) followed by two passages in BALB/c mice	Lab stock
Donor strains		
ICC8001*lacI*	Donor strain for conjugation assays. Tagged with constitutive Biofab promoter-driven *lacI* construct at 3' end of *glmS* gene	This study
GFP-D	ICC8001*lacI* carrying pKpGFP	This study
GFP-DD	ICC8001*lacI* carrying pKpGFP-D	This study
GFP-DΔ*traN*	ICC8001*lacI* carrying pKpGFPΔ*traN*	This study
GFP-DDΔ*traN*	ICC8001*lacI* carrying pKpGFP-DΔ*traN*	This study
GFP-D*traN*$_{R100-1}$	ICC8001*lacI* carrying pKpGFP*traN*$_{R100-1}$	This study
GFP-DD*traN*$_{R100-1}$	ICC8001*lacI* carrying pKpGFP-D*traN*$_{R100-1}$	This study
GFP-D*traN*$_{Ch1}$	ICC8001*lacI* carrying pKpGFP*traN*$_{Ch1}$	This study
GFP-DD*traN*$_{Ch1}$	ICC8001*lacI* carrying pKpGFP-D*traN*$_{Ch1}$	This study
GFP-DtraN$_F$	ICC8001*lacI* carrying pKpGFP*traN*$_F$	This study
GFP-DD*traN*$_F$	ICC8001*lacI* carrying pKpGFP-D*traN*$_F$	This study
GFP-D*traN*$_{C242S}$	ICC8001*lacI* carrying pKpGFP*traN*$_{C242S}$	This study
GFP-D*traN*$_{Ch2}$	ICC8001*lacI* carrying pKpGFP*traN*$_{Ch2}$	This study
GFP-DD*traN*$_{Ch2}$	ICC8001*lacI* carrying pKpGFP-D*traN*$_{Ch2}$	This study
GFP-D*traN*$_{Ch3}$	ICC8001*lacI* carrying pKpGFP*traN*$_{Ch3}$	This study
GFP-DD*traN*$_{Ch3}$	ICC8001*lacI* carrying pKpGFP-D*traN*$_{Ch3}$	This study

Recipient strains		
35_{ST258}	ICC8001 expressing ST258 variant of *ompK35*	Lab stock
35_{ST258}dTomato	Recipient strain used for live microscopy experiment. Tagged with Biofab-promoter driven *dTomato* construct at 3' end of *glmS* site	This study
$35_{ST258}/36_{ST258}$	ICC8001 expressing ST258 variants of both *ompK35* and *ompK36*	Lab stock
35_{ST258}/Δ36	ICC8001 expressing ST258 variant of *ompK35,* Δ*ompK36*	Lab stock
36_{ST258}	ICC8001 substitution of *ompK36* ORF with *ompK36* ORF from ST258 *K. pneumoniae* clinical isolate	Lab stock
Δ36	ICC8001 Δ*ompK36*	Lab stock
Δ35	ICC8001 Δ*ompK35*	Lab stock
Δ35/Δ36	ICC8001 Δ*ompK35, ΔompK36*	Lab stock
Δ35/Δ36::35	ICC8001 Δ*ompK35,* substitution of *ompK36* ORF with *ompK35* ORF	This study
$Δ35/36_{WTc>t}$	ICC8001 Δ*ompK35, ompK36*$_{WTc>t}$	Lab stock
Δ37	ICC8001 Δ*ompK37*	This study
Δ35/Δ36/Δ37	ICC8001 Δ*ompK35, ΔompK36*, Δ*ompK37*	This study
Δ*wcaJ*	ICC8001 Δ*wcaJ*	This study
Δ*mlaA*	ICC8001 Δ*mlaA*	This study
Δ*waaL*	ICC8001 Δ*waaL*	This study
$Δ35/36_{ST258ΔGD}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with L3 GD deletion	Lab stock
$Δ35/36_{WT+GD}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 GD insertion	Lab stock
$Δ35/36_{ST258ΔLSP}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with L4 LSP deletion	This study
$Δ35/36_{ST258R127A}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with R127A substitution	Lab stock
$Δ35/36_{ST258R127K}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with R127K substitution	Lab stock
$Δ35/36_{ST258D114N}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with D114N substitution	This study
$Δ35/36_{ST258D116N}$	ICC8001 Δ*ompK35*, *ompK36*$_{ST258}$ with D166N substitution	This study
$Δ35/36_{WT+D}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 D insertion	Lab stock
$Δ35/36_{WT+TD}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 TD insertion	Lab stock
$Δ35/36_{WT+E}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 E insertion	This study
$Δ35/36_{WT+K}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 K insertion	This study
$Δ35/36_{WT+H}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 H insertion	This study
$Δ35/36_{WT+R}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 R insertion	This study
$Δ35/36_{WT+C}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 C insertion	This study
$Δ35/36_{WT+M}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 M insertion	This study
$Δ35/36_{WT+N}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 N insertion	This study
$Δ35/36_{WT+Q}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 Q insertion	This study
$Δ35/36_{WT+S}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 S insertion	This study
$Δ35/36_{WT+T}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 T insertion	This study
$Δ35/36_{WT+G}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 G insertion using GGT codon	This study

Δ35/36$_{WT+G2}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 G insertion using GGC codon	This study
Δ35/36$_{WT+A}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 A insertion	This study
Δ35/36$_{WT+P}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 P insertion	This study
Δ35/36$_{WT+V}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 V insertion	This study
Δ35/36$_{WT+I}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 I insertion	This study
Δ35/36$_{WT+L}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 L insertion	This study
Δ35/36$_{WT+F}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 F insertion	This study
Δ35/36$_{WT+Y}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 Y insertion	This study
Δ35/36$_{WT+W}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L3 W insertion	This study
Δ*ompA*	ICC8001 Δ*ompA*	This study
Δ*ompW*	ICC8001 Δ*ompW*	This study
Δ*phoE*	ICC8001 Δ*phoE*	This study
36$_{L4ins}$	ICC8001 Δ*ompK35*, *ompK36*$_{WT}$ with L4 FTSG insertion	This study
OmpA$_{G>H}$	ICC8001 OmpA with L4 G>H substitution	This study
***E. coli* strains**		
DH5α R100-1	*E. coli* donor for R100-1	This study
MG1655	*E. coli* K-12 strain	Lab stock
Δ*ompA*	MG1655 Δ*ompA*	This study
Δ*ompF*	MG1655 Δ*ompF*	This study
Δ*ompC*	MG1655 Δ*ompC*	This study
Δ*ompF*/Δ*ompC*	MG1655 Δ*ompF*/Δ*ompC*	This study
Other strains		
S. Typhimurium LT2 SV3081	*S. enterica* serovar Typhimurium strain. Cured of pSLT	Gift from Josep Casadesús
E. cloacae ATCC 13047	*E. cloacae* recipient strain	Gift from Avinash Shenoy

Table 2.2. List of conjugative IncF plasmids used in this work.

Plasmid	Description	Source
pKpQIL	*bla*$_{KPC-2}$-encoding IncFII$_K$ plasmid. Shares sequence with pKpQIL-UK (Genbank accession no. KY798507)	Lab stock
R100-1	Derepressed variant of the IncFII plasmid R100. Also referred to as NR1 in the literature	Gift from Fernando de la Cruz
pOX38	F plasmid derivative	Gift from Tiago RD Costa
pSLT	*S.* Typhimurium virulence plasmid	Lab stock
Synthetic reporter conjugative plasmids		
pKpGFP	pKpQIL with P*lac-sfGFP* construct inserted at disrupted *aadA* gene. Parental reporter plasmid.	This study
pKpGFP-D	pKpGFPΔ*finO.* Derepressed reporter plasmid	This study
pKpGFPΔ*traN*	pKpGFPΔ*traN*	This study
pKpGFP-DΔ*traN*	pKpGFPΔ*finO*Δ*traN*	This study
pKpGFP*traN*$_{R100-1}$	pKpGFPΔ*traN*::*traN*$_{R100-1}$	This study
pKpGFP-D*traN*$_{R100-1}$	pKpGFPΔ*finO*Δ*traN*::*traN*$_{R100-1}$	This study
pKpGFP*traN*$_{Ch1}$	pKpGFPΔ*traN*::*traN*$_{Ch1}$. Contains variable region from *traN*$_{R100-1}$	This study
pKpGFP-D*traN*$_{Ch1}$	pKpGFPΔ*finO*Δ*traN*::*traN*$_{Ch1}$. Contains variable region from *traN*$_{R100-1}$	This study
pKpGFP*traN*$_F$	pKpGFPΔ*traN*::*traN*$_F$	This study
pKpGFP-D*traN*$_F$	pKpGFPΔ*finO*Δ*traN*::*traN*$_F$	This study
pKpGFP*traN*$_{C242S}$	pKpGFPΔ*traN*::*traN*$_{C242S}$	This study
pKpGFP*traN*$_{Ch2}$	pKpGFPΔ*traN*::*traN*$_{Ch2}$. Contains variable region from *traN*$_{pSLT}$	This study
pKpGFP-D*traN*$_{Ch2}$	pKpGFPΔ*finO*Δ*traN*::*traN*$_{Ch2}$. Contains variable region from *traN*$_{pSLT}$	This study
pKpGFP*traN*$_{Ch3}$	pKpGFPΔ*traN*::*traN*$_{Ch3}$. Contains variable region from *traN*$_{MV1}$	This study
pKpGFP-D*traN*$_{Ch3}$	pKpGFPΔ*finO*Δ*traN*::*traN*$_{Ch3}$. Contains variable region from *traN*$_{MV1}$	This study
Other conjugative plasmids		
pKpQIL-D	pKpQILΔ*finO*. Derepressed pKpQIL without sfGFP reporter construct.	This study

2.2 Generation of mutants

Mutagenesis vectors (Table 2.3) were generated by Gibson Assembly (New England Biolabs) on the pSEVA612S backbone and maintained in *E. coli* CC118λpir cells. Site-directed mutagenesis on previously generated vectors was performed according to the Q5 Site-Directed Mutagenesis Kit protocol (New England Biolabs). Primers used to generate the mutagenesis vectors and for screening are listed in Table 2.4. The variable region of $TraN_{pMV1}$ was synthesised by GeneArt Gene synthesis (ThermoFisher Scientific).

Genomic mutations were made in ICC8001, a rifampicin resistant derivative of *K. pneumoniae* ATCC43816[52] and *E. coli* MG1655, a K-12 derivative using a two-step recombination methodology. Briefly, 20 µl of CC118λpir cells carrying a mutagenesis vector was mixed with 20 µl of helper (*E. coli* 1047 pRK2013) and incubated on LB agar at 37°C for 2 h. Subsequently, 40 µl of recipient carrying the pACBSR recombineering plasmid (expressing the λ-red system and I-SceI endonuclease) was added and incubated for a further 4 h. Merodiploids were selected on LB agar containing gentamicin and streptomycin. For the second recombination step, merodiploids were grown in LB containing streptomycin and 0.4% L-arabinose at 37°C for 6 h to induce expression of the I-SceI endonuclease. PCR and sequencing were used to confirm the intended mutations and pACBSR was cured through serial passaging in LB.

Mutations in pKpQIL were introduced using the same methodology with several modifications. In the second recombination step, ertapenem was added in the culture media to maintain the plasmid. All subsequent mutations were introduced onto the parental pKpGFP. Deletion of *finO* was always performed after the introduction of other mutations in pKpGFP to generate derepressed plasmid variants. Furthermore, serial passaging to cure pACBSR was only performed where pKpGFP was not derepressed as this was found to result in the loss of the derepressed phenotype. In the case of derepressed plasmids, conjugation was used to transfer the generated plasmid into a donor strain which did not carry pACBSR.

Table 2.3. Plasmids and mutagenesis vectors used in this work.

Plasmid	Description	Source
pACBSR	SmR; expresses I-SceI and lambda-red induced by L-Ara.	Lab stock
pSEVA612S	GmR; integrative plasmid (ori R6K) that harbours the oriT for conjugation.	Lab stock
pSEVA471	SmR; for selecting recipient cells in conjugation assays	Lab stock
pSEVA612S-Kp-N2-sfGFP	pSEVA612S derivative; inserts Biofab promoter driven *sfGFP* construct at 3' end of *glmS*	Lab stock
pET-28a(+)	KanR; used as template to amplify *lacI* gene	Lab stock
pCSCMV-dTomato	Used as template to amplify *dTomato*	Lab stock
pSEVA612S_OmpK36$_{WT}$	Template for site directed mutagenesis of OmpK36$_{WT}$	Lab stock
pSEVA612S_OmpK36$_{ST258}$	Template for site directed mutagenesis of OmpK36$_{ST258}$	Lab stock
pSEVA612s derivatives		
Cloning vectors for chromosomal mutagenesis		
pSEVA612Ssub35	Substitutes the ORF of *ompK36* with the ORF for *ompK35*	This study
pSEVA612SΔ37	Deletes *ompK37* from ICC8001	This study
pSEVA612SΔwcaJ	Deletes *wcaJ* from ICC8001	This study
pSEVA612SΔmlaA	Deletes *mlaA* from ICC8001	This study
pSEVA612SΔwaaL	Deletes *waaL* from ICC8001	This study
pSEVA612Sbiofabdtomato	Inserts Biofab promoter driven *dTomato* construct at 3' end of *glmS*	This study
pSEVA612S_D114N	Substitutes codon encoding D114 residue in OmpK36$_{ST258}$ with codon for asparagine	This study
pSEVA612S_D116N	Substitutes codon encoding D116 residue in OmpK36$_{ST258}$ with codon for asparagine	This study
pSEVA612S_36+E	Inserts E115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+K	Inserts K115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+H	Inserts H115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+R	Inserts R115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+C	Inserts C115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+M	Inserts M115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+N	Inserts N115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+Q	Inserts Q115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+S	Inserts S115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+T	Inserts T115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+G	Inserts G115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+A	Inserts A115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+P	Inserts P115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+V	Inserts V115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+I	Inserts I115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+L	Inserts L115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+F	Inserts F115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612S_36+Y	Inserts Y115 residue in L3 of OmpK36$_{WT}$	This study

pSEVA612S_36+W	Inserts W115 residue in L3 of OmpK36$_{WT}$	This study
pSEVA612SΔompAKP	Deletes *ompA* from ICC8001	This study
pSEVA612SΔompW	Deletes *ompW* from ICC8001	This study
pSEVA612SΔphoE	Deletes *phoE* from ICC8001	This study
pSEVA612SΔompAEC	Deletes *ompA* from MG1655	This study
pSEVA612S36L4ins	Inserts bases coding for FTSG residues in L4 of OmpC into L4 of OmpK36	This study
pSEVA612SompAG>H	Substitutes codon for glycine with codon for histidine in L4 of OmpA in ICC8001	This study
pSEVA612SΔompF	Deletes *ompF* from MG1655	This study
pSEVA612SΔompC	Deletes *ompC* from MG1655	This study
Cloning vectors for IncF plasmid mutagenesis		
pSEVA612SbiofablacI	Inserts Biofab promoter driven *lacI* construct at 3' end of *glmS* in *K. pneumoniae* ATCC 43816	This study
pSEVA612SPlacsfGFP	Inserts *lac* promoter driven *sfGFP* at disrupted *aadA* gene on pKpQIL	This study
pSEVA612SOXA-sfGFP	Inserts *lac* promoter driven sfGFP at disrupted *tir* gene on pOXA-48a	This study
pSEVA612SΔfinO	Deletes *finO* from pKpQIL	This study
pSEVA612SΔtraN	Deletes *traN* from pKpQIL	This study
pSEVA612StraNR100	Substitutes the ORF of *traN* in pKpQIL with the ORF for *traN* from R100-1	This study
pSEVA612StraNCh1	Substitutes the ORF of *traN* in pKpQIL with *traN*$_{Ch1}$	This study
pSEVA612StraNF	Substitutes the ORF of *traN* in pKpQIL with the ORF for *traN* from F	This study
pSEVA612StraNCh2	Substitutes the ORF of *traN* in pKpQIL with a chimeric *traN* containing the variable region from *traN*$_{pSLT}$	This study
pSEVA612StraNCh3	Substitutes the ORF of *traN* in pKpQIL with a chimeric *traN* containing the variable region from *traN*$_{MV1}$	This study
pSEVA612StraNC242S	Substitutes the ORF of *traN* in pKpQIL with *traN* containing the C242S mutation	This study

Table 2.4. Primers used in this work.

Primer	Sequence (5' → 3')	Description
pSEVA612S_F	ATTACCCTGTTATCCCTATACTG	Amplifies linear pSEVA612S
pSEVA612S_R	TAGGGATAACAGGGTAATCCG	
lacIvector_F	TAAGGATCCAACAGGGTTC	Amplifies linear pSEVA612S with homology regions flanking the 3' end of *glmS* and Biofab promoter from pSEVA612S-Kp-N2-sfGFP
lacIvector_R	TTTTTTTTTACCTCCTTAAACTCC	
lacI_F	TTTAAGGAGGTAAAAAAAAAGTGG TGAATGTGAAACCAGTAAC	Amplifies *lacI*
lacI_R	AGAACCCTGTTGGATCCTTATCAC TGCCCGCTTTCCAG	
glmS_F	GGTCAGGATGCGTCTATCG	Checks for insertion at 5' end of *glmS*
glmS_R	CCTGAGTCAGTTTGTCATC	
aadAUPHR_F	GGATTACCCTGTTATCCCTACAAA CGCGAAGGCCGGTG	Amplifies upstream homology region flanking disrupted *aadA* on pKpQIL
aadAUPHR_R	TTTTCTCGACGCGCGAGGCCAAGC GATC	
aadADNHR_F	GTACAAATAAGCAGATCAGTTGGA AGAATTTG	Amplifies downstream homology region flanking disrupted *aadA* on pKpQIL
aadADNHR_R	TATAGGGATAACAGGGTAATGCAA GATTCCACTATCAAAC	
plac_F	GGCCTCGCGCGTCGAGAAAATTTA TCAAAAAGAGTG	Amplifies lac promoter
plac_R	CTTTACGCATACGTATCCTCCAAG CCTG	
sfGFP_F	GAGGATACGTATGCGTAAAGGCGA AGAG	Amplifies *sfGFP*
sfGFP_R	ACTGATCTGCTTATTTGTACAGTTC ATCCATACC	
placsfGFP_F	TGGCTGGGCGGTCGAGAAAATTTA TCAAAAAGAG	Amplifies P*lac-sfGFP* construct
placsfGFP_R	GTTATCAGCGTTATTTGTACAGTTC ATCCATAC	
tirUPHR_F	GGATTACCCTGTTATCCCTACGTA CTCAACATCGGCGAAAGAG	Amplifies upstream homology region flanking disrupted *tir* on pOXA-48a
tirUPHR_R	TTTTCTCGACCGCCCAGCCACATC GTCC	
tirDNHR_F	GTACAAATAACGCTGATAACGTCT TTGCTG	Amplifies downstream homology region flanking disrupted *tir* on pOXA-48a
tirDNHR_R	TATAGGGATAACAGGGTAATCGCA TCCAGCCATTCACG	
36vector_F	TTTGTTATGCAGCTTGCAACTTAGA ACTGGTAAACGATAC	Amplifies linear pSEVA612S with homology regions flanking *ompK36*
36vector_R	GTTATTAACCCTCTGTTTGTTATAT G	
ompK35_F	ACAAACAGAGGGTTAATAACATGA AGCGCAATATTCTG	Amplifies *ompK35*
ompK35_R	TTTGTTATGCAGCTTGCAACTTAGA ACTGGTAAACGATAC	
ompK36ext_F	GCCGACTGATTAGAAGGGTAATC	

ompK36ext_R	GAGTATACCAGCGAGGTTAAACC	Checks for deletions/insertions/substitutions in *ompK36* ORF
ompK37UPHR_F	GGATTACCCTGTTATCCCTAGAGC AACGTTTCCGCCGC	Amplifies upstream homology region flanking *ompK37*
ompK37UPHR_R	GGCAGGCGCAGAAGAAATTCCTTT AACTGTTTTTATGTTTTAACTTCCG	
ompK37DNHR_F	GAATTTCTTCTGCGCCTGCCCGCC GGTT	Amplifies downstream homology region flanking *ompK37*
ompK37DNHR_R	TATAGGGATAACAGGGTAATGCAC CTCGGCGTCGTCAATCCG	
ompK37ext_F ompK37ext_R	CTTATATCGTAACAATTG GACTATATCATTAAACGG	Checks for deletion of *ompK37*
wcaJUPHR_F	GGATTACCCTGTTATCCCTAAATAA CCGCGGTGAAATG	Amplifies upstream homology region flanking *wcaJ*
wcaJUPHR_R	GAGCATCTAAAATTCAATCACTCAT TTATAAACAAG	
wcaJDNHR_F	TGATTGAATTTTAGATGCTCCTTAA GACAAGG	Amplifies downstream homology region flanking *wcaJ*
wcaJDNHR_R	TATAGGGATAACAGGGTAATGGTT ACGACGGATGGTGTC	
wcaJext_F wcaJext_R	GCAAGGGTTATCAAAAG GTGATGATAATGTTAGCCC	Checks for deletion of *wcaJ*
mlaAUPHR_F	GGATTACCCTGTTATCCCTAGGCC GGCCGCAATCGCTGAC	Amplifies upstream homology region flanking *mlaA*
mlaAUPHR_R	CTTTTTTTATCATTCGCCGAAGATG TCTCCCTGGTTTTTATGGCTTTCGC	
mlaADNHR_F	TAAAAACCAGGGAGACATCTTCGG CGAATGATAAAAAAAGGGT	Amplifies downstream homology region flanking *mlaA*
mlaADNHR_R	TATAGGGATAACAGGGTAATGCTG TCCGGCTATAACCGCG	
mlaAext_F mlaAext_R	CAACGTGCTGTTACAATCGC GGCGCTAACTTCAACTACCG	Checks for deletion of *mlaA*
waaLUPHR_F	GGATTACCCTGTTATCCCTACGCC ATTAACCCTTTTACCGCC	Amplifies upstream homology region flanking *waaL*
waaLUPHR_R	TTTTCAGCGCGATTTTTTGCCAAAA AGGGCCGTCAGCGGC	
waaLDNHR_F	GCCGCTGACGGCCCTTTTTGGCAA AAAATCGCGCTGAAAA	Amplifies downstream homology region flanking *waaL*
waaLDNHR_R	TATAGGGATAACAGGGTAATGTTG CTATAAGATTTACCAGCC	
waaLext_F waaLext_R	CCGTGATTAAAGAGCAGGCC CGAAAAGGTTGCGTTTAATGAG	Checks for deletion of *waaL*
finOUPHR_F	GGATTACCCTGTTATCCCTACCCG TGGTATTCCGGGAATATTC	Amplifies upstream homology region of *finO*
finOUPHR_R	GTAAATATAAAACAATTGCCTATCG TTCAGTTAATAAG	
finODNHR_F	GGCAATTGTTTTATATTTACCCATT CCTGATAATTATACCTGGG	Amplifies downtream homology region of *finO*
finODNHR_R	TATAGGGATAACAGGGTAATCGGC AACATCGTCTCCCC	
finOext_F finOext_R	GTTCTATGCTGTGCACCTGG GTTATGATGCCGCAGCCTG	Checks for deletion of *finO*
traNUPHR_F	GGATTACCCTGTTATCCCTACCGC CAGTTTATCGATAATCTG	

traNUPHR_R	GCAGCATGGTTTCTGCCCTCCCTCATCC	Amplifies upstream homology region of *traN* from pKpQIL
traNDNHR_F	GAGGGCAGAAACCATGCTGCCTAATAAAGAG	Amplifies downtream homology region of *traN* from pKpQIL
traNDNHR_R	TATAGGGATAACAGGGTAATGGAATAGCGGCATGCTCAG	
traNext_F	GGAGAAAGTGGCACAAACCG	Checks for deletion or substitution of *traN* on pKpQIL
traNext_R	CTTCCCGACGTCCCTTTGAC	
dTomato_F	TTTAAGGAGGTAAAAAAAAAATGGTGAGCAAGGGCGAG	Amplifies *dTomato*
dTomato_R	AGAACCCTGTTGGATCCTTATTACTTGTACAGCTCGTCCATG	
D114N_F	AACGGCGACACCTACGGTTC	Used for SDM to generate pSEVA612S_D114N
D114N_R	GCCGCCGAATTCCGGC	
D116N_F	AACACCTACGGTTCTGACAACTTCCT	Used for SDM to generate pSEVA612S_D116N
D116N_R	GCCGTCGCCGCCGAAT	
OmpK36SDM_R	GTCGCCGCCGAATTCCGG	Reverse primer to generate E,H,R,C,N,Q,S,G,G2,A,P,V, L,F,Y and W insertion vectors
OmpK36+E_F	GAGACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+E
OmpK36+H_F	CACACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+H
OmpK36+R_F	CGTACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+R
OmpK36+C_F	TGCACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+C
OmpK36+N_F	AACACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+N
OmpK36+Q_F	CAGACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+Q
OmpK36+S_F	TCCACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+S
OmpK36+G_F	GGTACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+G
OmpK36+G2_F	GGCACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+G2
OmpK36+A_F	GCAACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+A
OmpK36+P_F	CCGACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+P
OmpK36+V_F	GTTACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+V
OmpK36+L_F	CTGACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+L
OmpK36+F_F	TTCACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+F
OmpK36+Y_F	TACACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+Y
OmpK36+W_F	TGGACCTACGGTTCTGACAACTTCCTG	Used for SDM to generate pSEVA612S_36+W

OmpK36SDM_R2	AATTCCGGCAGAACGTCG	Reverse primer to generate M,T and I insertion vectors
OmpK36+M_F	CGGCGGCGACATGACCTACGGTT	Used for SDM to generate pSEVA612S_36+M
OmpK36+T_F	CGGCGGCGACACCACCTACGGTT	Used for SDM to generate pSEVA612S_36+T
OmpK36+I_F	CGGCGGCGACATCACCTACGGTT	Used for SDM to generate pSEVA612S_36+I
OmpK36+K_F OmpK36+K_R	CGGCGGCGACAAAACCTACGGTT AATTCCGGCAGAACGTCGG	Used for SDM to generate pSEVA612S_36+K
traNvector_F traNvector_R	AGGACAGTAAACCATGCTGCCTAA TAAAGAG TACGTTTCATTTCTGCCCTCCCTCA TCC	Amplifies linear pSEVA612S with homology regions flanking *traN* from pKpQIL
traNR100_F traNR100_R	GAGGGCAGAAATGAAACGTATTTT ACCTCTG GCAGCATGGTTTACTGTCCTGCCT GTTTC	Amplifies *traN* from R100-1
KPCtraN_F KPCtraN_R	AGGGATGAGGGAGGGCAGAAATG AAGACGGTTATTTCCG TCTTTATTAGGCAGCATGGTTTATT GCGCGGATTGCTG	Amplifies *traN* from pKpQIL
TraNchimera_F TraNchimera_R	ACCCTGGTGATGGAAGAAAC CCGTGTGCAAAAATTTTCC	Amplifies linear pSEVA612S with portion of *traN* from pKpQIL flanking the variable region
traNR100var_F traNR100var_R	TGGAAAATTTTTGCACACGGACTG CCAGTATCACCGGG GTTTCTTCCATCACCAGGGTCAGC GTAAAGGTGGAAGC	Amplifies variable region from *traN* of R100-1
ompAUPHR_F ompAUPHR_R	GGATTACCCTGTTATCCCTAGGAG TTAACCGCTGACGAAC CGGTTATAACTTTTTGCGCCTCATT ATCATCC	Amplifies upstream homology region flanking *ompA* from *K. pneumoniae*
ompADNHR_F ompADNHR_R	GGCGCAAAAAGTTATAACCGATAA AAAAACCCGCTTC TATAGGGATAACAGGGTAATCCCG CTACATTGAGGCCAG	Amplifies downstream homology region flanking *ompA* from *K. pneumoniae*
ompAext_F ompAext_R	CTTACGCTGCATGTATCAG CAGGTAGGATCGTCGAC	Checks for deletion of *ompA* in *K. pneumoniae*
ompWUPHR_F ompWUPHR_R	GGATTACCCTGTTATCCCTACCGG TTTTCATAAATAGTGC ACAGAAGAATATCCACTTCCTCATT ATGG	Amplifies upstream homology region flanking *ompW*
ompWDNHR_F ompWDNHR_R	GGAAGTGGATATTCTTCTGTAAAC TGGCCAACG TATAGGGATAACAGGGTAATGGCC AGGGGAGACCTATG	Amplifies downstream homology region flanking *ompW*
ompWext_F ompWext_R	CTGAGGACTTAGTGTGATC CCCTGCTCAACATGTATCAC	Checks for deletion of *ompW*
phoEUPHR_F phoEUPHR_R	GGATTACCCTGTTATCCCTAAGGC GATGGTGGCGGGCA CTGCGGTTAATATTCAGTCCTGGT GATTTATTTATACGCGCTATTCAAT TGCG	Amplifies upstream homology region flanking *phoE*

phoEDNHR_F	GGACTGAATATTAACCGCAGAACCACCC	Amplifies downstream homology region flanking *phoE*
phoEDNHR_R	TATAGGGATAACAGGGTAATATTGATAGCGGATCGGAC	
phoEext_F	CGGCGTTAAAAAACCTCC	Checks for deletion of *phoE*
phoEext_R	CTGCCGAAGGAGTATAAC	
traNF_F	GAGGGCAGAAATGAAACGTATTTTACCTCTG	Amplifies *traN* from F
traNF_R	GCAGCATGGTTTACTGTCCTGCCTGTTTC	
ompAECUPHR_F	GGATTACCCTGTTATCCCTAGACTGAAGAAGAGCATGC	Amplifies upstream homology region flanking *ompA* from *E. coli*
ompAECUPHR_R	AGACGAGAACTTTTTGCGCCTCGTTATC	
ompAECDNHR_F	GGCGCAAAAAGTTCTCGTCTGGTAGAAAAAC	Amplifies downstream homology region flanking *ompA* from *E. coli*
ompAECDNHR_R	TATAGGGATAACAGGGTAATGAAAGCGGTTGGAAATGG	
OmpK36L4SDM_F	AGTGGCGCGACCAACAACGGTCGTGG	Used for SDM to insert coding sequence for FTSG residues into L4 of $OmpK36_{WT}$
OmpK36L4SDM_R	AGTAAAGCCTTCGCCGCTGACGCT	
OmpAKP_F	GGATTACCCTGTTATCCCTAATGAAAAAGACAGCTATCGCGATTG	Amplifies *ompA* ORF from *K. pneumoniae*
OmpAKP_R	TATAGGGATAACAGGGTAATTTAAGCCGCCGGCTGAGT	
OmpASDM_F	CGGCGACGCGCACACCGTGGGT	Used for SDM to generate $pSEVAompA_{G>H}$
OmpASDM_R	ATGTTGTTAACCCACTGGTATTCCAG	
C242S_F	GAATCGATGTCATCAGGCGGTC	Used for SDM to mutate C242 of TraN to a serine residue
C242S_R	TGAACGTGGCCAGTTGAT	
traNpSLT_F	TGGAAAATTTTTGCACACGGACGGCCACCATCACCGGC	Amplifies variable region from $traN_{pSLT}$
traNpSLT_R	GTTTCTTCCATCACCAGGGTCAGCGTGAACGTCGTTCTCCC	
traNMV1_F	AGGGATGAGGGAGGGCAGAAATGAAGACGGTTATTTCCG	Amplifies variable region from $traN_{MV1}$
traNMV1_R	GTTTCTTCCATCACCAGGGTAACGGTGAAGCTGTAGCGC	
ompFUPHR_F	GGATTACCCTGTTATCCCTACGATCATCCTGTTACGGAATATTAC	Amplifies upstream homology region flanking *ompF* from *E. coli*
ompFUPHR_R	AGGTGTGCTATATTTATTACCCTCATGGTTTTTTTTATG	
ompFDNHR_F	GTAATAAATATAGCACACCTCTTTGTTAAATGCCGAAAAAACAGGACTTTG	Amplifies downstream homology region flanking *ompF* from *E. coli*
ompFDNHR_R	TATAGGGATAACAGGGTAATCGCCAGTGCCCCCCGGAG	
ompFext_F	GCAGACACATAAAGACAC	Checks for deletion of *ompF*
ompFext_R	GAGATTGCTCTGGAAG	
ompCUPHR_F	GGATTACCCTGTTATCCCTAGTGAAATAGTTAACAAGCG	Amplifies upstream homology region flanking *ompC* from *E. coli*
ompCUPHR_R	TCAATCGAGAGTTATTAACCCTCTGTTATATGC	

ompCDNHR_F	GGTTAATAACTCTCGATTGATATCG AACAAAGGGC	Amplifies downstream homology region flanking *ompC* from *E. coli*
ompCDNHR_R	TATAGGGATAACAGGGTAATGATT CACCAGCGGCCCGA	
ompCext_F	GTATCATATTCGTGTTGG	Checks for deletion of *ompC*
ompCext_R	GTACGCTGAAAACAATG	

2.3 Selection-based conjugation assays

To confirm plasmid acquisition by conjugation, plasmids studied in this work were tagged with a superfolder green fluorescence protein (sfGFP) construct which is selectively expressed in transconjugant cells. For experiments using donors carrying the tagged pOXA-48a and pKpGFP and its derivatives, recipients were transformed with pSEVA471, a low-copy number plasmid encoding streptomycin resistance. To quantify conjugation of R100-1, DH5α carrying R100-1 was used as the donor and recipients were transformed with pUltra-sfGFP which confers gentamicin resistance. For all experiments, overnight cultures of donor and recipient bacteria were washed in phosphate-buffered saline (PBS). Donor and recipient cells were mixed at a ratio of 8:1, which had previously been determined to result in the maximum transfer frequency for pKpQIL and diluted in PBS (1 in 25 v/v)[161]. 40 µl of the final conjugation mixture was spotted onto LB agar plates and incubated for 6 h at 37°C. The spots were collected and resuspended in 1 ml of sterile PBS for serial dilution. Recipient colonies were selected on streptomycin or gentamicin-containing LB agar plates. Transconjugants were selected on plates supplemented with streptomycin and ertapenem for pKpGFP and pOXA-48a experiments and streptomycin and gentamicin for R100-1 conjugation experiments. Plates were visualised on a Safe Imager 2.0 Blue Light Transilluminator (ThermoFisher Scientific) to confirm uptake of tagged plasmids by GFP expression and colony forming units (CFU) for recipients and transconjugants were determined. Conjugation frequency was calculated as a ratio of CFU/ml of transconjugants to CFU/ml recipients and the data was $\log_{10}$ transformed prior to statistical analysis. All conjugation experiments were performed in biological triplicate.

2.4 Preparation and visualisation of outer membrane proteins

To purify bacterial OM proteins, overnight cultures grown in LB were washed and resuspended in 10mM HEPES (pH 7.4). Cells were sonicated at 25% amplitude for 10 bursts of 15 seconds each on a Fisherbrand™ Model 705 Sonic Dismembrator (ThermoFisher Scientific). The sonicated cells were centrifuged at 3000 x *g* for 10 min at 4°C to pellet cellular debris. Next, the supernatant was centrifuged at 14000 x *g* for 30 min at 4°C and the pellet was resuspended in 0.4 ml of 1% sarcosine/10 mM HEPES (pH 7.4) (w/v) followed by incubation on a tube roller for 30 min at RT. The mixture was centrifuged at 14000 x *g* for 30 min at 4°C and the pellet was resuspended in water. The concentration of purified OM proteins was determined on a Qubit4 (Invitrogen). 10 µg of protein was separated by SDS-PAGE using 12% acrylamide Mini-protean TGX precast gels (Bio-Rad) and stained with Coomassie.

2.5 Fluorescence imaging

Two-dimensional fluorescence imaging of conjugation mixtures on agar plates was performed using an IVIS Spectrum CT system (Perkin Elmer). Images were processed on LivingImage v4.3.1 and radiance was quantified using the region of interest tool.

2.6 Real-time conjugation system (RTCS) assays

Conjugation mixtures were prepared by mixing PBS-washed overnight cultures of donors carrying derepressed reporter plasmids and recipient bacteria. Donor and recipient bacteria were mixed at the ratios determined to provide maximal fluorescence emission after 6 hours (1:1 with no dilution for pKpGFP-D plasmids; 1:1 with 1 in 25 (v/v) dilution in PBS for the tagged pOXA-48a). 8 µl of the conjugation mixture was spotted onto 270 µl of LB agar in a 96-well black microtiter plate in technical triplicate. The plates were incubated for 6 h at 37°C with fluorescence readings taken at 10-min intervals on a FLUOstar Omega (BMG Labtech). Fluorescence data at each timepoint was calculated by normalizing the raw GFP emission at

that timepoint to the minimum GFP emission recorded for each sample over the 6 h time course. Arbitrary fluorescence units (AFU) were determined by calculating the log fold change of fluorescence at *t* = 300 min for each mutant recipient strain (X) against the WT recipient i.e., AFU = $\log_{10}(\text{fluorescence}_X/\text{fluorescence}_{WT})$. To compare the efficiency of plasmid transfer mediated by different TraN variants into various recipient species, endpoint measurements were taken after 24 hours of incubation. For these experiments, a negative control mixture (-GFP) was included for statistical analysis. The negative control mixture contained a donor strain carrying the untagged but derepressed pKpQIL, pKpQIL-D, and the respective recipient strain being tested. All RTCS assays were performed in biological triplicate. Representative graphs are shown for fluorescence emission over time data.

2.7 Growth curves

Strains were assessed for growth at 37°C in both LB and M9 minimal media supplemented with 0.4% glucose (w/v). Saturated overnight cultures were diluted 1:100 and OD_{600} measurements were obtained every 30 min for 12 hours on a FLUOstar Omega (BMG Labtech) with shaking at 200 RPM in between readings. Growth assays were performed in biological and technical triplicate and the data at each timepoint is presented as the mean ± standard deviation (SD).

2.8 Purification of pKpQIL conjugative pilus and generation of polyclonal antibodies

A 2 L overnight culture of ICC8001pKpQIL-D (pKpQILΔ*finO*) was harvested by centrifugation at 7000 x *g* for 20 min and resuspended in 40 ml of cold PBS. The resuspended cells were passed through a 25G needle 30 times before centrifugation at 50000 x *g* for 1 h. The supernatant was mixed with 5% PEG 6000 with constant stirring for 1 h at 4°C. Pili were precipitated by centrifugation at 50000 x *g* for 30 min. The pellet was resuspended in buffer containing 50 mM Tris pH 8, 1 M NaCl and dialysed against the same buffer. Purified pili were visualised by negative stain electron microscopy to assess for pilus integrity and purity. Rat

polyclonal antibodies were raised againt the purified pili (ThermoFisher Scientific). The polyclonal antibodies were serially adsorbed against ICC8001 fixed in paraformaldehyde (PFA) to isolate antibodies specific to the pKpQIL conjugative pilus.

2.9 Western Immunoblotting

Bacterial cell lysates were prepared by harvesting cells from overnight cultures adjusted to an OD_{600} of 0.5 and resuspending in 2X Laemlli buffer (from 5X stock, 312.5 mM Tris-HCL, pH 6.8, 10% w/v SDS, 50% v/v glycerol, 0.5 M DTT, 0.05% w/v bromophenol blue). Samples were boiled at 100°C for 5 min and centrifuged briefly prior to SDS-PAGE separation on 4-20% acrylamide Mini-protean TGX precast gels (Bio-Rad). Western blot analysis was performed using custom anti-pili polyclonal antibodies (ThermoFisher Scientific) and HRP-conjugated donkey anti-rat IgG antibodies (Jackson Immunoresearch) and blots were imaged on a ChemiDoc™ XRS+ imaging system (Bio-Rad).

2.10 Immunofluorescence microscopy

Overnight cultures were diluted 1 in 20 (v/v) in fresh LB and 300 µl was added to glass coverslips placed in a 24-well plate before incubation at 37°C for 1.5 h to allow the bacteria to adhere to the surface of the coverslips. Excess media was removed, and the coverslips were washed with PBS before fixation in 4% PFA for 20 min at room temperature (RT). Fixed samples were washed in PBS and blocked in 2% bovine serum albumin (BSA) in PBS (w/v). Samples were washed 3 times before incubation with anti-pili antibodies (1:100 in 2% BSA/PBS) for 1 h at RT. Samples were washed 3 times in PBS and incubated with Alexa Fluor 488 conjugated donkey anti-rat IgG antibodies (Jackson Immunoresearch, 1:1000 in 2% BSA/PBS) for 1 h at RT. Coverslips were washed three times in PBS and incubated with FM4-64™ (Invitrogen, 1:100 in water) for 5 minutes at RT. Following this, coverslips were dried and mounted onto glass slides using VECTASHIELD® Hardset™ Antifade Mounting Medium with DAPI (Vector Laboratories) according to the manufacturer's instruction. Slides were analysed

using a x100 objective on a Zeiss Axio Observer 7 microscope and images were processed on Zen 2.3 (Blue Version; Zeiss).

2.11 Agarose pad live microscopy

Bacterial conjugation was visualised over time on a Celldiscoverer 7 live cell imaging microscope (Zeiss). For these experiments, the GFP-DD donor strain was mixed with 36_{WT}-expressing recipients which constitutively express dTomato. Overnight cultures of donor and recipient bacteria were washed in PBS, mixed in a 1:1 ratio and 8 µl was spotted onto a 1 cm^2 2% agarose (w/v) pad supplemented with M9 salts and 0.4% glucose (w/v). The pad was inverted into a 2-well µ-Slide fitted with a polymer coverslip (Ibidi). The sample was maintained at 37°C throughout live imaging. Images were acquired every 10 min for 3.5 h and processed using Zen 2.3 (Blue Version; Zeiss).

2.12 Generation of TraN AlphaFold models

In the absence of homologous TraN structures, *ab initio* models were generated by AlphaFold v2.0[162]. TraN sequences were submitted to the AlphaFold Colab server with the default settings; the signal peptide was removed from all sequences prior to modelling. Each structural model was validated by analysing the confidence score as generated by the predicted local-distance difference test (pLDDT). PDB files generated by AlphaFold were used for further analysis and for generating molecular graphics on UCSF ChimeraX[163].

2.13 Overexpression and purification of TraN and OmpK36

OmpK36 was produced as previously described[52] and exchanged into 50 mM NaCl, 10 mM HEPES pH 7.0 and 0.03% *n*-dodecyl-β-D-maltoside (DDM; Anatrace). The mature TraN gene (D28 to Q651) from pKpQIL was subcloned into the pTAMAHISTEV vector with an N-terminal His_7-tag and a tobacco etch virus (TEV) cleavage site using the NcoI and XhoI restriction

enzyme sites. The construct was transformed into *E. coli* C43 (DE3) competent cells and expressed in Terrific Broth (TB) medium (Formedium) supplemented with 100 µg/mL ampicillin. Cultures were grown to an OD_{600} of 0.6–0.8 at 37°C, then induced with 0.5 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG) and maintained for 16 h at 25°C. Outer membranes were produced as previously described[52], then solubilised overnight in 1% DDM in 1X PBS. Unsolubilised material were pelleted at 131,000 × *g* for 1 h, and the supernatant was combined with 30 mM imidazole and loaded onto an Econo-Column (Bio-Rad) containing 5 mL Ni-NTA resin (Qiagen). The column was washed with 5 column volumes (CV) of wash buffer (1X PBS, 30 mM imidazole and 0.1% DDM). TraN-His_7 eluted from the resin in wash buffer containing 100–250 mM imidazole. TraN-His_7 was dialysed against 50 mM NaCl, 10 mM HEPES pH 7.0 and 0.03% DDM (buffer A) whilst undergoing incubation with His_6-tagged TEV protease for 16–18 h at 4°C. The dialysed sample was passed over a 5 mL His-Trap column (Cytiva) and the cleaved protein was collected from the flowthrough. Fractions containing TraN were combined and further purified using anion-exchange chromatography (Mono Q 5/10 GL column; Cytiva) using an ÄKTA pure system (Cytiva). The column was equilibrated with buffer A and eluted using a linear gradient with 500 mM NaCl, 10 mM HEPES pH 7.0 and 0.03% DDM (buffer B) over 20 CV. TraN eluted in 18% buffer B and was concentrated to 1 mg/mL for SEC analysis.

2.14 SEC analysis of TraN-OmpK36

TraN and OmpK36 were dialysed against buffer A (16–18 h at 4°C) and then were combined at a 1:2 molar ratio respectively at 1 mg/mL and incubated for 16 h at room temperature. The sample was injected onto a Superose 6 10/300 GL column (Cytiva) and eluted at a flow rate of 0.3 mL/min in buffer A, whilst monitoring the absorbance at 280 nm. This was followed by separate injections of TraN and OmpK36 onto the column at the same molar concentrations as previously described for the comparison of retention volumes. Fractions were collected and analysed by SDS-PAGE.

2.15 Cryo-EM Sample Preparation and Data Collection

Sample containing OmpK36-TraN at a concentration of 0.33 mg/mL was diluted 1:6 in buffer A. In brief, a 4 µL aliquot of sample was applied to a plasma-cleaned graphene oxide-coated Cu 300 mesh 1.2/1.3 holey carbon grid (Quantifoil), blotted with force 6 for 4.5 s at 90% humidity and flash frozen in liquid ethane using a Vitrobot Mark IV (FEI). The dataset used for structure determination was collected at the Molecular Electron Microscopy Core at the University of Virginia on a Titan Krios EM operated at 300 keV, equipped with an energy filter and K3 direct electron detector (Gatan). An energy filter slit width of 10 eV was used during data collection and was aligned automatically every hour. All 13,668 movies were collected in counting mode at a magnification of 81K, pixel size of 1.08 Å, and a defocus range from -2.2 to -1.2 µm. Data collection was performed using a total dose of 50 e- $Å^{-2}$ across 40 frames at a rate of 4.78 s/movie.

2.16 Data Processing

Unless otherwise stated, all data processing was completed using cryoSPARC v3.2[164]. Movies were corrected for full-frame motion using Patch Motion Correction followed by Gctf constant transfer function (CTF) Estimation. After CTF estimation, micrographs were sorted and selected based on estimated resolution (better than 4 Å), defocus (-1 to -2.5 µm), ice thickness, and total full-frame motion. Initial particles were automatically picked using ‘Blob picker’ with minimum and maximum particle diameters of 200 and 256 Å, respectively. Particles were extracted at a box size of 256 pixels, followed by 2D classification. Class averages of trimeric OmpK36 alone, and OmpK36 with TraN were selected for template-based particle picking. A total of 13,780,567 particles were extracted using a box diameter of 256 Å. These particles were sorted using three iterative rounds of 2D classification with 50 classes each, number of online-EM iterations set to 100 and a batch size of 1000 per class. The final iteration of 2D classification yielded a subset of 3,412,946 particles.

To differentiate particles containing only OmpK36 or OmpK36 + TraN, multiple 3D maps were generated using '*Ab-initio* reconstruction', class size set to 4. Output 3D maps were inspected for the presence of TraN. Particles were further refined using two iterations of heterogeneous refinement with input volumes created by multi-class *ab initio*. The highest resolution class from the second iteration of heterogeneous refinement contained 359,314 particles, which allowed for a ~2.6 Å map to be reconstructed using 'Non-uniform refinement'.

2.17 Model Building and Refinement

The density for the trimeric OmpK36 allowed us to trace the entire backbone and build most side chains throughout the structure. The OmpK36 crystal structure (PDB ID: 6RD3) was used for building the cryo-EM model which only had small differences relative to the starting model. The predicted TraN AlphaFold model was used for initial interpretation of the loop-shaped density found within the lumen of one porin channel. The loop and the two β-strands on either side of the hairpin of the AlphaFold model could be fit into the density. Two cysteines at either side of the hairpin fit into the TraN density and were used as a starting point for matching larger sidechains within the density. Model building including adjusting side chains were performed in Coot[165]. The model was refined in Phenix, using Real-space refinement with `ignoring symmetry conflicts` turned on[166]. Refinement included global minimization, B-factor optimization, and applied secondary structure and Ramachandran restraints. The final model had a MolProbity score of 1.39, with 96% and 0.1% in the Ramachandran favored and outlier regions, respectively. The coordinates and structure factors of the TraN-OmpK36 complex are available at the Protein Data Bank and Electron Microscopy Data Bank with ID codes 7SZI and 25567 respectively.

2.18 Bioinformatic analysis of TraN variants

Bioinformatic analysis was performed on a previously described dataset of 14,029 sequenced plasmids deposited in GenBank up to August 2018. This dataset was analysed using the

Plascad tool for plasmid characterization[167]. Briefly, Plascad identifies and distinguishes putative conjugative plasmids from mobilizable and non-mobilizable plasmids based on the presence of relaxase, T4CP and T4SS genes. Further characterization of the T4SS genes into the four archetypal MPF groups is performed for the putative conjugative plasmids. For this work, conjugative MPFF plasmids were extracted from the dataset. Next, the bacterial host family associated with each plasmid was determined by querying the NCBI Taxonomy database via the "ncbi_taxonomy" module in the Python toolkit "ETE" (v3.0) and manually curating the results. Lastly, PlasmidFinder was used to determine the replicon types of each plasmid and non-IncF plasmids were manually excluded[168]. 824 putative conjugative IncF plasmids found in an *Enterobacteriaceae* host met the inclusion criteria and were screened for *traN* variants. A tBLASTn[169] was performed and required sequences to share ≥90% amino acid similarity with TraN from pKpQIL, R100-1 or F and possess ≥75% of the reference length to categorize them as part of the same *traN* type. To identify other sequence variants, plasmids were analysed for annotated *traN* genes and a tBLASTn was performed to group sequence variants together.

2.19 Multiple sequence alignments of TraN and OM proteins

Multiple sequence alignments were generated in Clustal Omega 1.2.4. TraN sequences were obtained from the following reference plasmids: pKpQIL (Accession ID: KY798507.1), R100-1 (Accession ID: DQ364638.1), F (Accession ID: NC_002483.1), pSLT (Accession ID: AE006471.2), MV1 (Accession ID: NZ_CP016763.1), MV2 (Accession ID: AP014954.1), MV3 (Accession ID: NZ_CP023348.1). OM protein sequences were obtained from the following published genomes: *K. pneumoniae* ATCC 43816 (Accession ID: CP009208), *E. coli* MG1655 (Accession ID: U00096.3), *S.* Typhimurium LT2 (Accession ID: AE006468), *E. cloacae* ATCC 13047 (Accession ID: CP001918). Phylogenetic trees were calculated using the neighbour joining method and visualized in Jalview 2.11.2.2.

2.20 Statistical analysis

All statistical analysis was performed on Prism 9 (GraphPad Software Inc). Except for representative RTCS graphs for fluorescence emission over time, all quantitative data are presented as the mean ± standard deviation (SD) of three biological replicates. The statistical tests applied to the data are indicated in the corresponding figure legends and *P* values <.05 were considered significant.

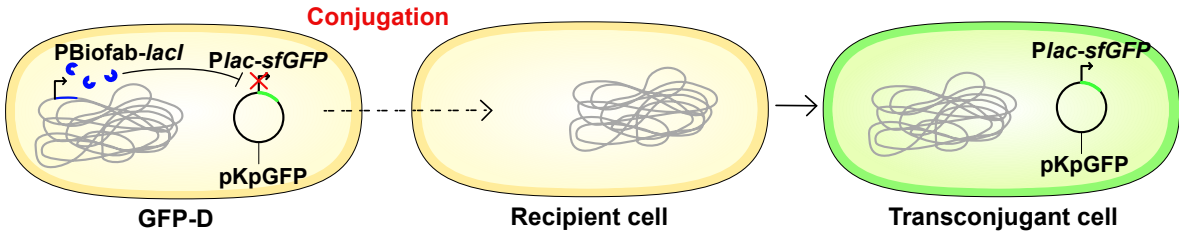
3. Results

3.1 Generating a reporter pKpQIL for measuring conjugation frequency

To study conjugative transfer of pKpQIL, a reporter plasmid was generated by adapting the tagging strategy used by Pérez-Mendoza and de la Cruz to evaluate R388 and F plasmid conjugation[100]. A schematic of the adapted tagging strategy is shown in Figure 3.1A. Briefly, a construct consisting of the superfolder green fluorescence protein gene (*sfGFP*) driven by a *lac* promoter (P_{lac}) was inserted into the *aadA* gene which is truncated by the insertion of the IS*26* insertion sequence. This engineered reporter plasmid was named pKpGFP. The parental WT strain of KP used in these studies, a derivative of ATCC 43816 called ICC8001, naturally expresses the lac inhibitor, LacI. However, the endogenous levels of expression were found to be insufficient to inhibit expression of *sfGFP*. Therefore, a strain was engineered from ICC8001 to constitutively overexpress *lacI* by inserting a construct containing *lacI* driven by a synthetic Biofab promoter downstream of the highly conserved *glmS* gene, a site that is commonly targeted for Tn*7* integration in KP. When pKpGFP was transferred into this strain, fluorescence emission was significantly reduced compared to the parental strain (Figure 3.1B). This *lacI*-overexpressing strain was used as the donor strain for all reporter plasmids generated in this work. The donor carrying pKpGFP was named GFP-D (GFP-donor).

For selection-based assays, GFP-D was mixed with streptomycin resistant recipient strains which do not constitutively express *lacI*. Transconjugants selected on agar containing both streptomycin and the carbapenem, ertapenem, were visibly fluorescent when viewed on a transilluminator confirming uptake of pKpGFP (Figure 3.1C). Occasionally, non-fluorescent colonies were detected which were excluded from transconjugant counts used to assess conjugation frequency. In these assays, conjugation frequency is calculated as the ratio of CFU/ml of transconjugants to CFU/ml of recipients and statistical analysis was performed on the $\log_{10}$ transformed data.

A



B

PBiofab-*lacI*	Total Radiant Efficiency [p/s]/[μW/cm^2]
Absent	1.56X10^{10}
Present	4.65X10^9

C

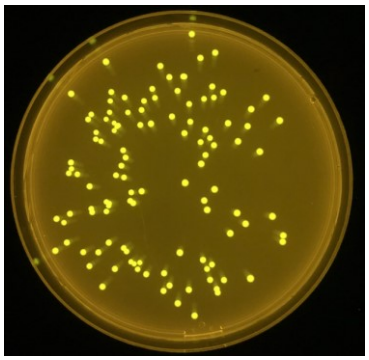


Figure 3.1. Engineering a reporter system to assess conjugation frequency of pKpQIL

A. Schematic illustrating the fluorescence reporter system for confirming the acquisition of pKpGFP by recipient cells. GFP-D donor cells are constructed to constitutively express *lacI* which represses *lac* promoter (P*lac*)-driven *sfGFP* expression from the tagged pKpQIL, pKpGFP. Following transfer of pKpGFP into recipients which do not constitutively express *lacI*, *sfGFP* is expressed. **B.** The effect of the chromosomal insertion of the synthetic PBiofab-*lacI* construct on repression of *sfGFP* expression from pKpGFP was compared by measuring the total radiant efficiency from pKpGFP-carrying cells. **C.** Image of selective agar plate showing fluorescent transconjugant colonies carrying pKpGFP.

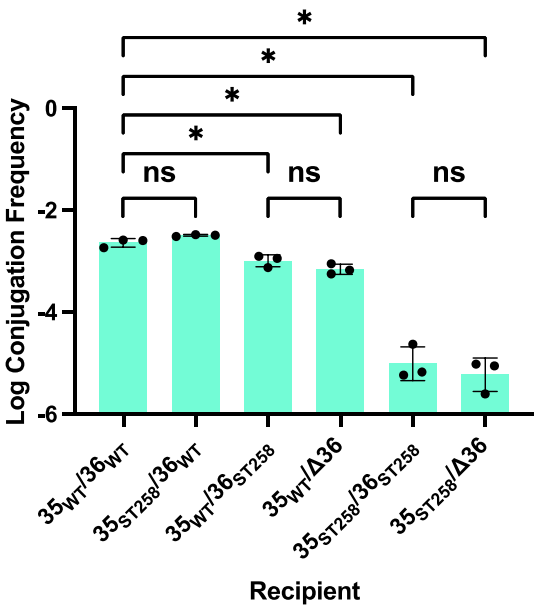
3.2 pKpGFP conjugation is dependent on OmpK35 and OmpK36 in the recipient

GFP-D was used to assess the effect of the ST258 mutations in OmpK35 and OmpK36 on conjugation. Previously, a panel of isogenic mutant strains was generated from ICC8001 to express a combination of WT and ST258 isoforms of the two porins[52]. These strains were used as recipients in the selection-based assay and the log conjugation frequency of pKpGFP into WT recipients ($35_{WT}/36_{WT}$) was observed to be approximately -2.6 (Figure 3.2A&B). This equates to plasmid uptake by less than 1% of recipients. When only the truncated, non-functional ST258 isoform of OmpK35 (35_{ST258}) was expressed, there was no significant

difference in pKpGFP conjugation frequency compared to the WT recipient.

In contrast, when the ST258 isoform of OmpK36 (36_{ST258}) was expressed in recipients, a significant reduction in pKpGFP conjugation frequency was observed (Figure 3.2A). There was no significant difference between strains expressing 36_{ST258} and those in which *ompK36* had been deleted (Δ36) which suggests that the ST258-associated mutations confer a similar effect on conjugation as a complete loss of expression of OmpK36. Lastly, although expression of 35_{ST258} alone does not have a significant effect on pKpGFP transfer, when co-expressed with 36_{ST258} or Δ36, there was an additive reduction in conjugation frequency of approximately 2 log-fold or 100-fold (Figure 3.2B).

A



B

OmpK35	OmpK36	Log conjugation frequency	Mean difference compared to WT
WT	WT	-2.64 ± 0.08	-
ST258	WT	-2.49 ± 0.02	0.15
WT	ST258	-2.99 ± 0.12	-0.35
WT	Δ	-3.16 ± 0.10	-0.52
ST258	ST258	-5.01 ± 0.33	-2.37
ST258	Δ	-5.22 ± 0.33	-2.59

Figure 3.2. ST258 porin mutations reduce conjugative uptake of pKpGFP

A. Isogenic mutants expressing different isoform combinations of OmpK35 (35_{WT}, 35_{ST258}) and OmpK36 (36_{WT}, 36_{ST258}, Δ36) were used as recipients in pKpGFP conjugation. Log transformed conjugation frequency of pKpGFP into each recipient is plotted. Data analysed by RM one-way ANOVA with Tukey's multiple comparison test applied. * = $p < 0.05$, ns = non-significant, error bars represent SD ($n = 3$). **B.** Table showing log conjugation frequency values associated with each recipient strain and the mean difference in values between each mutant and the WT strain.

To determine if the effect of the porin mutations on conjugation is specific to the recipient, a donor strain expressing both 35_{ST258} and 36_{ST258} was generated and plasmid transfer was compared with the GFP-D donor which expresses the WT isoforms of both porins. No significant difference in conjugation frequency was observed between these donors into the recipients tested suggesting that this effect is recipient specific (Figure 3.3).

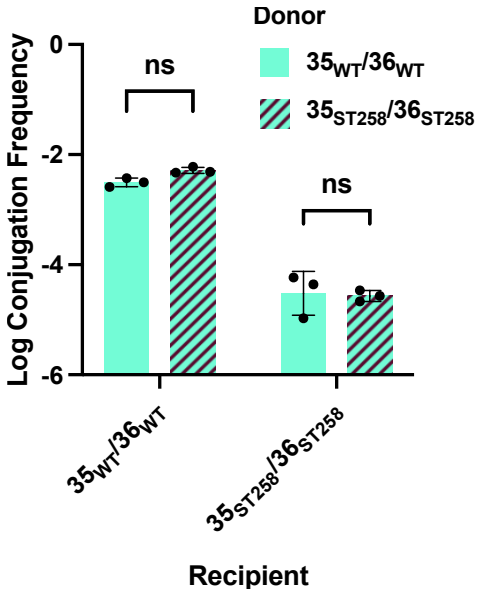


Figure 3.3. The effect of the ST258 mutations on conjugation is recipient specific

Plasmid transfer was compared between donors expressing the WT and ST258 isoforms of OmpK35 and OmpK36. Data analysed by 2way ANOVA with Bonferroni's multiple comparison test. ns = non-significant, error bars represent SD (*n* = 3).

3.3 Changes in abundance of OmpK35 but not OmpK36 affects conjugation

From the above experiment, OmpK36 appears to play a more important role in conjugative uptake of pKpGFP when compared to OmpK35. It was hypothesised that this might be associated with the lower expression level of OmpK35 compared to OmpK36. To test this, the open reading frame (ORF) of *ompK36* was substituted with the *ompK35* ORF in a strain in which both porin genes had been deleted (Δ35/Δ36) to generate Δ35/Δ36::35_{WT}. OM proteins were isolated and analysed by SDS-PAGE to confirm that the abundance of OmpK35 increased when expressed off the promoter for *ompK36* than when it is expressed off its

endogenous promoter (Figure 3.4A). These strains were then used as recipients with the GFP-D donor. Although not statistically significant, the reduction in pKpGFP conjugation frequency between $35_{WT}/36_{WT}$ and $35_{WT}/\Delta36$ suggests that endogenous expression of *ompK35* cannot fully compensate for the deletion of *ompK36* during pKpGFP conjugation (Figure 3.4B). However, overexpression of *ompK35* was able to increase conjugation frequency to levels that were comparable to the WT strain suggesting that changes in the abundance of OmpK35 impact conjugation frequency.

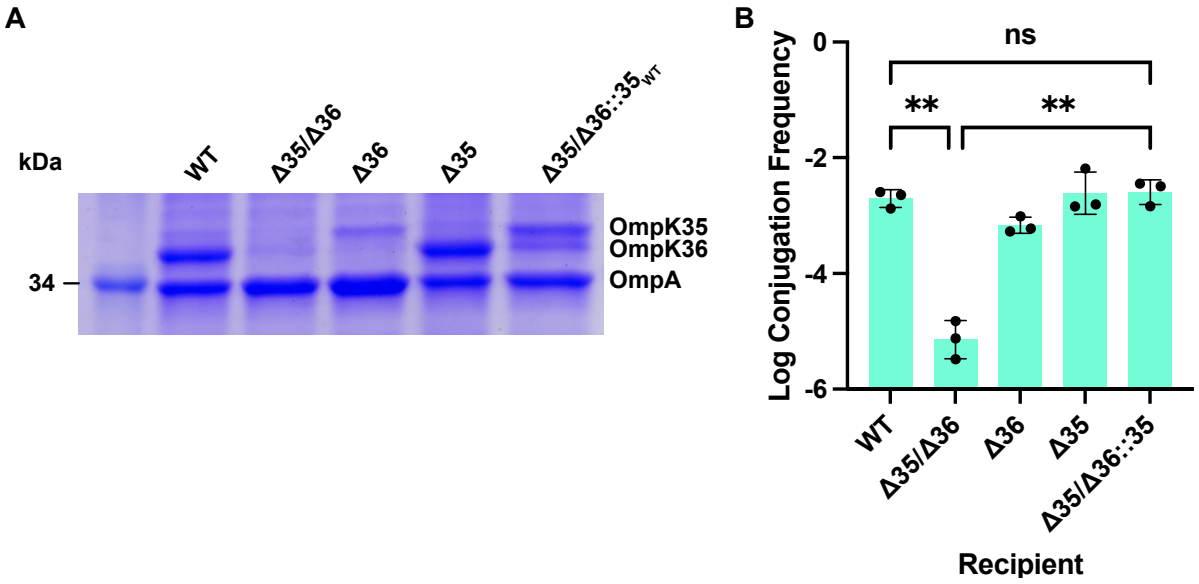


Figure 3.4. The effect of changes in OmpK35 abundance on pKpGFP conjugation

A. OM proteins were isolated from the isogenic mutant strains indicated and separated by SDS-PAGE. Bands corresponding to OmpK35, OmpK36 and OmpA are shown. **B.** Log conjugation frequency of pKpGFP into the isogenic mutants is shown. Data analysed by RM one-way ANOVA with Tukey's multiple comparison test applied. ** = $p < 0.01$, ns = non-significant. Error bars represent SD ($n = 3$).

Next, the impact of changes in OmpK36 abundance on pKpGFP conjugative uptake were assessed. In these experiments, the Δ35 strain was used as the parental strain. A synonymous c>t mutation in *ompK36* was found to greatly reduce the expression of this porin through the formation of a secondary mRNA structure that interferes with translation[57]. This

reduction in expression was confirmed through SDS-PAGE analysis of isolated OM proteins from a strain carrying this mutation (Figure 3.5A). However, there was no significant difference in pKpGFP conjugation frequency between recipients expressing 36_{WT} and $36_{WTc>t}$ (Figure 3.5B). Taken together, this shows that pKpGFP uptake is affected by changes in abundance in OmpK35 but not OmpK36 in the recipient.

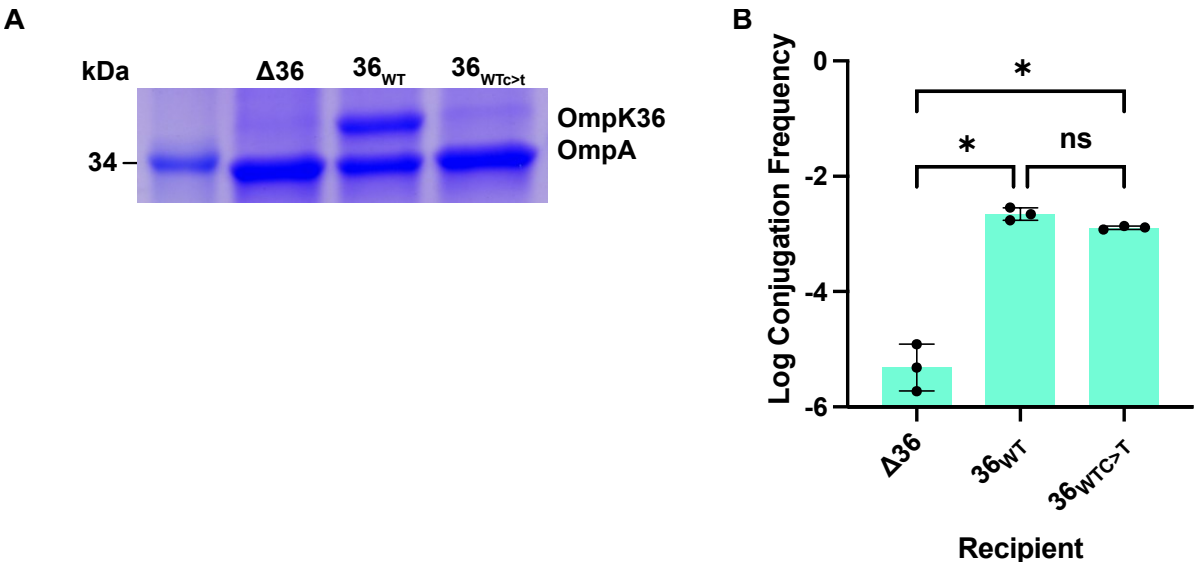


Figure 3.5. The effect of changes in OmpK36 abundance on pKpGFP conjugation

A. OM proteins were isolated from strains expressing Δ36, 36_{WT} and $36_{WTc>t}$ and separated by SDS-PAGE. Bands corresponding to OmpK36 and OmpA are shown. **B.** Log conjugation frequency of pKpGFP into Δ36, 36_{WT} and $36_{WTc>t}$ recipients are shown. Data analysed by RM one-way ANOVA with Tukey's multiple comparison test applied. * = $p < 0.05$, ns = non-significant. Error bars represent SD ($n = 3$).

3.4 OmpK37 does not play a role in pKpGFP conjugation

OmpK37 is a structural homologue of OmpK35 and OmpK36 which shares the classical porin L3 motif with OmpK36[170]. Unlike OmpK35 and OmpK36, OmpK37 expression is quiescent and its role in β-lactam resistance is less well understood[49]. In the absence of OmpK37-specific antibodies, OmpK37 expression in ICC8001 could not be confirmed. To determine if OmpK37 could support conjugative uptake of pKpGFP, *ompK37* was deleted from both the WT and Δ35/Δ36 strains. No significant change in conjugation frequency was observed

between strains in which *ompK37* had been deleted and the parental strain (Figure 3.6). This suggests that even if it is expressed in the parental ICC8001 strain, at the conditions used, OmpK37 is not involved in pKpGFP conjugation.

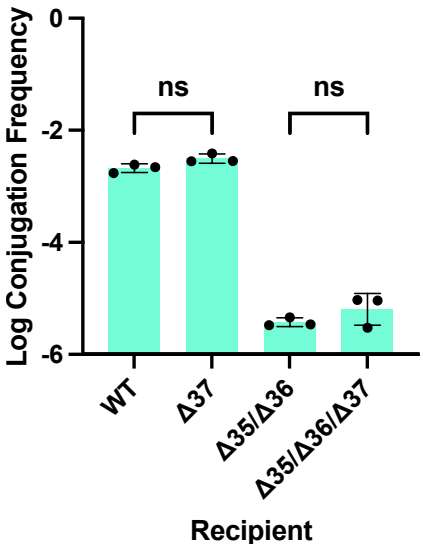


Figure 3.6. Recipient OmpK37 is not involved in pKpGFP conjugation

The effect of deleting *ompK37* on its own or in combination with *ompK35* and *ompK36* on log conjugation frequency is shown. Data analysed by RM one-way ANOVA with Tukey's multiple comparison test. ns = non-significant. Error bars represent SD (*n* = 3).

3.5 Investigating the role of other recipient surface components

To determine if the effects on pKpGFP conjugation were specific to mutations in OmpK35 and OmpK36, the role of other features of the bacterial cell surface during conjugation were investigated. Specifically, the role of capsule, OM asymmetry and the O-antigen were assessed.

To study the role of capsule, a capsule-null Δ*wcaJ* recipient was generated from the WT ICC8001 strain. WcaJ acts as the terminal glycosyltransferase involved in capsule synthesis[33]. No significant difference in conjugation frequency between WT and Δ*wcaJ* recipients was observed, showing that the capsule does not affect pKpGFP conjugation (Figure 3.7A). OM

asymmetry was investigated by deleting *mlaA* from ICC8001. MlaA functions to maintain the asymmetry of the OM by removing excess phospholipids from the outer leaflet of the membrane[171]. No significant difference in conjugation frequency was associated with deletion of *mlaA* suggesting that OM asymmetry does not play a role in pKpGFP conjugation (Figure 3.7B). Lastly, *waaL* was deleted from ICC8001 to investigate the role of the O-antigen. WaaL is a ligase required in the formation of a complete LPS molecule and functions to attach an O-antigen molecule to a lipid A-core oligosaccharide[172]. No significant difference in conjugation frequency was observed (Figure 3.7C), suggesting that the O-antigen portion of LPS does not play a role during pKpGFP conjugation.

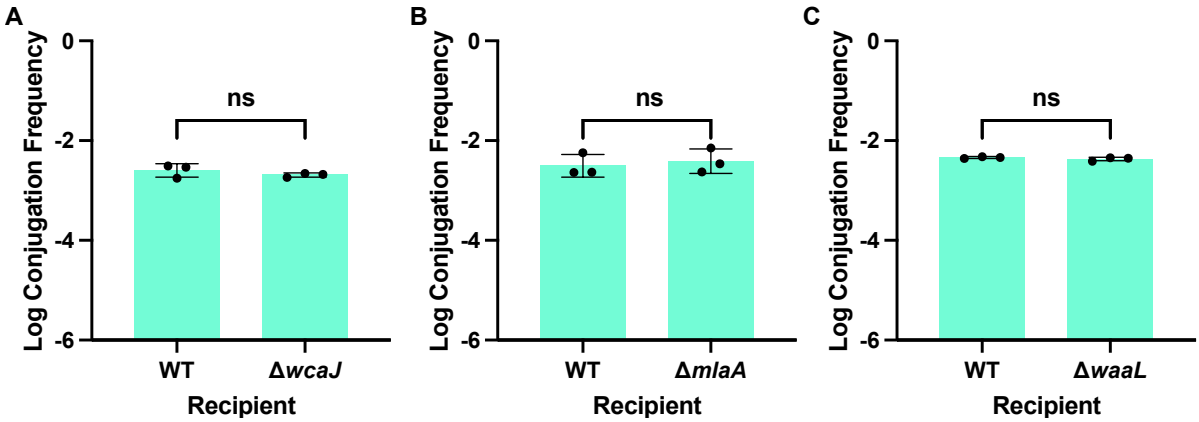


Figure 3.7. Assessing the role of WcaJ, MlaA and WaaL on pKpGFP conjugation

The effect of deleting (**A**) *wcaJ*, (**B**) *mlaA* and (**C**) *waaL* from ICC8001 on log conjugation frequency of pKpGFP is shown. Data analysed by a paired *t*-test. ns = non-significant. Error bars represent SD (*n* = 3).

3.6 High-throughput real time conjugation system (RTCS)

The previous results suggest that the major OM porins are important recipient factors for pKpGFP conjugative uptake with OmpK36 serving as the primary porin facilitating efficient plasmid. Thus, the work hereafter, focuses mainly on investigating the role of OmpK36 during conjugation and uses recipients which do not express *ompK35* (Δ35).

Mutations in OmpK36 are known to increase the MIC of the strain to β-lactam antibiotics including carbapenems, which are used for the purpose of selecting transconjugants[43].

Therefore, an alternative method to quantify pKpGFP conjugation in the absence of selection pressure was developed. This would address any potential bias that could arise from differences in carbapenem resistance profiles due to the introduction of mutations in *ompK36*.

The tagging strategy described by Pérez-Mendoza and de la Cruz had been originally designed for the purpose of performing high-throughput screens of mutant libraries. An additional advantage of the system is that it can be used to assess conjugation without the need for selection[100]. For this study, a similar assay was developed in which fluorescence emission is measured over time and compared between each conjugation mixture. This assay was named real-time conjugation system (RTCS). A pilot experiment was performed to measure fluorescence emission over time from pKpGFP into WT ICC8001 recipient cells. However, no increase in fluorescence was detected.

The RTCS was hypothesised to lack the sensitivity required to detect low efficiency transfer like in the case of pKpGFP which naturally conjugates at low frequencies (<1% of recipients) (Figure 3.2B). To confirm this, the system was recapitulated in the naturally derepressed IncL/M plasmid, pOXA-48a[63]. pOXA-48a conjugates at higher frequencies into WT recipients but does not display OmpK36-dependency (Figure 3.8A). pOXA-48a was tagged with the sfGFP construct at the disrupted *tir* gene and was transferred into the *lacI*-expressing donor strain. Total fluorescence emission was measured from conjugation mixtures prepared at various ratios of donor and WT recipients to determine the conditions which would result in the maximum fluorescence emission following 6 hours of incubation at 37°C (Figure 3.8B). From this, it was determined that a mixture consisting of a 1:1 ratio of donor and recipient which had been diluted 1 in 25 v/v in PBS be used for further experiments. Using donors carrying the tagged pOXA-48a, an increase in fluorescence over time was observed into recipients expressing both the WT and ST258 isoforms of OmpK36 (Figure 3.8C). In addition, calculating the log ratio of fluorescence emission from the 36_{ST258}-expressing strain against

the WT recipient at *t* = 300 min, herein referred to as arbitrary fluorescence units (AFU), validated that plasmid uptake into this strain is comparable to the WT recipient (Figure 3.8D).

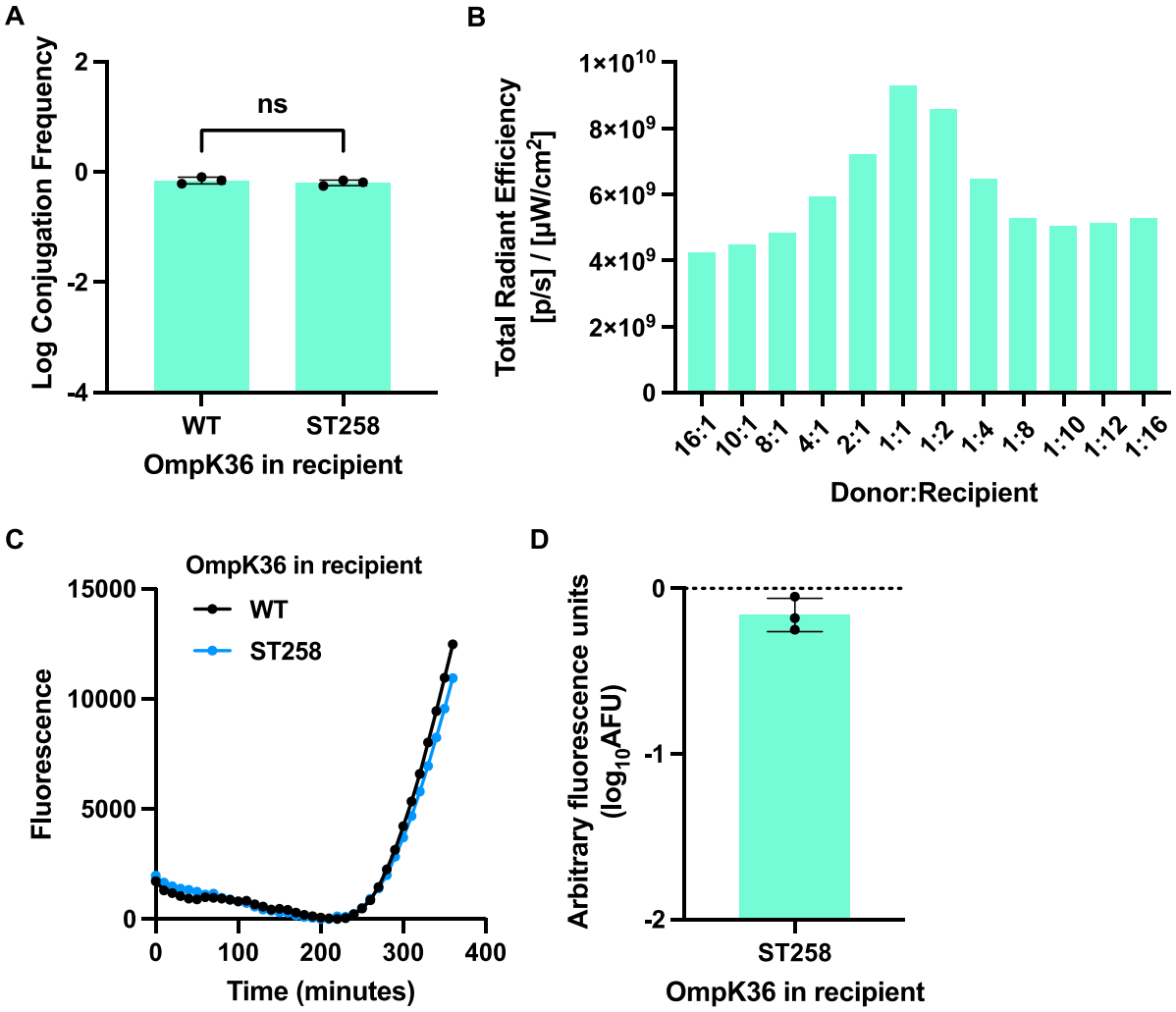


Figure 3.8. The RTCS can be used to assess transfer of pOXA-48a

A. Log conjugation frequency of pOXA-48a into recipients expressing $OmpK36_{WT}$ or $OmpK36_{ST258}$. Data analysed by paired *t*-test. ns = non-significant. Error bars represent SD (*n* = 3). **B.** Total radiant efficiency was measured from conjugation mixtures containing different ratios (v/v) of donors carrying the tagged pOXA-48a and WT recipients. **C.** Representative graph of fluorescence emission over time from conjugation mixtures of donors carrying tagged pOXA-48a and recipients expressing $OmpK36_{WT}$ or $OmpK36_{ST258}$. **D.** Arbitrary fluorescence units (AFU) for the $OmpK36_{ST258}$-expressing recipient were calculated as the $\log_{10}$ fold change from the recipient expressing the WT isoform. Error bars represent SD (*n* = 3).

Based on these findings, it was concluded that the RTCS can be used to monitor and compare conjugation frequency of a tagged plasmid in real-time. However, as the sensitivity of this assay is much lower when compared to quantification of transconjugants on selective media, this setup is unsuitable for studying the transfer of plasmids which conjugate at frequencies below the sensitivity threshold.

3.7 Characterisation of GFP-DD and preparation of anti-pili antibodies

To allow for the investigation of pKpGFP transfer using the RTCS, *finO* was deleted to derepress *tra* gene expression and to increase the overall conjugation frequency of the plasmid. This plasmid was named pKpGFP-D (derepressed) and transferred into the *lacI*-expressing donor strain to generate the GFP-derepressed donor, GFP-DD. GFP-DD was characterised to determine the effects of derepression of the *tra* operon in this plasmid. Growth curves comparing GFP-D and GFP-DD were performed in both rich (LB) and minimal (M9 + 0.4% glucose) media (Figure 3.9). The doubling time of GFP-DD in both rich and minimal media was significantly lower than the GFP-D strain, suggesting that derepression of the *tra* operon incurs a significant fitness cost on donor cells.

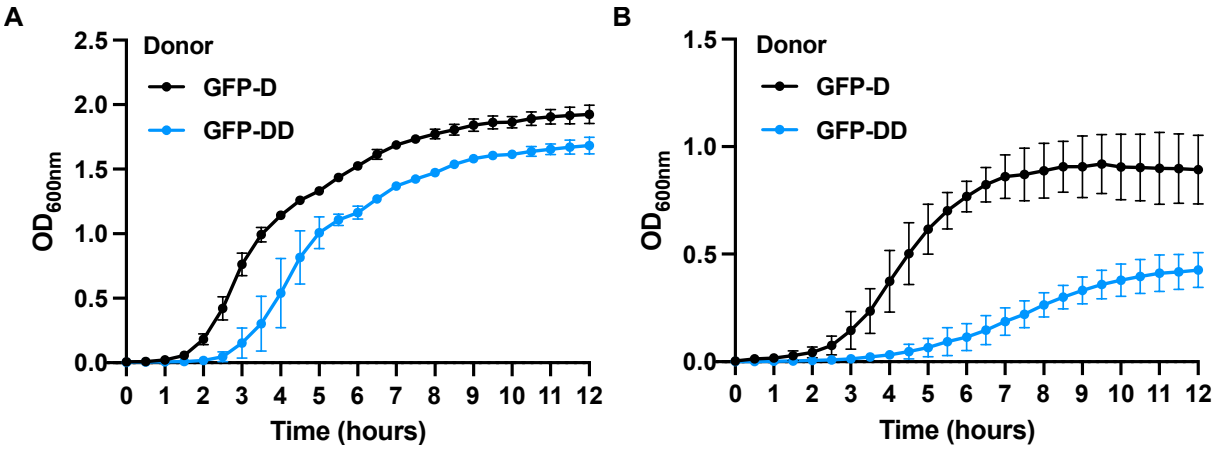


Figure 3.9. Derepression of pKpGFP has a fitness cost on donor cells

Growth of GFP-D and GFP-DD cells in (**A**) rich LB and (**B**) minimal M9 media supplemented with 0.4% glucose is shown. Absorbance was measured at 600 nm every 30 min for 12 h. Error bars represent SD (*n* = 3).

Next, conjugative pili were purified from GFP-DD cells and imaged with the assistance of Dr Alejandro Peña. SDS-PAGE analysis and mass spectrometry of pilin subunits confirmed that the purified filaments were conjugative pili originating from pKpGFP-D (Figure 3.10A). The quality of the pili was further assessed by negative stain transmission electron microscopy (Figure 3.10B).

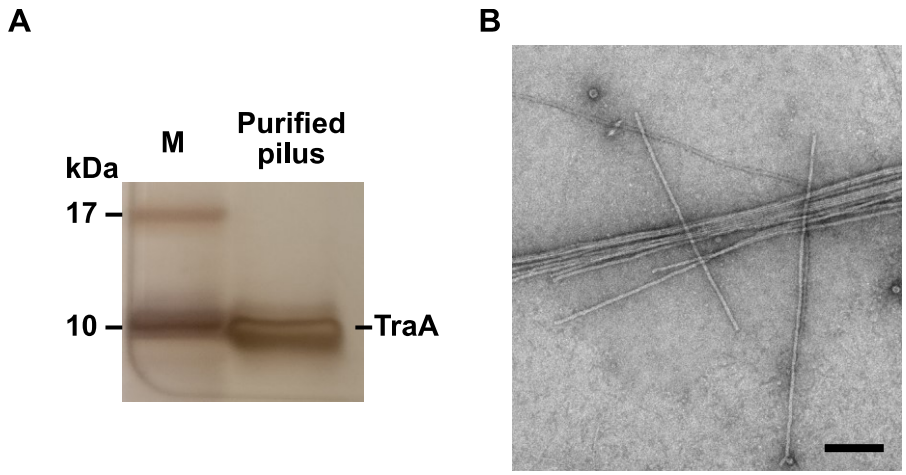


Figure 3.10. Purification of conjugative pili from the GFP-DD cells

A. Pili purified from GFP-DD were separated by SDS-PAGE and visualized by silver stain. A band corresponding to the approximate size of a TraA (pilin) subunit is observed. M = protein marker. **B.** Negative stain EM image of purified pili. Scale bar = 200 nm.

This pilus preparation was used to raise rat polyclonal antibodies. Western blot analysis was performed to assess the specificity of antisera derived from two immunised rats, hereafter referred to as rat #1 and rat #2 (Figure 3.11A). Antisera from rat #1 was found to be highly specific for the TraA pilin subunit as seen by comparing between the lanes containing GFP-DD cell lysate and purified pili. Expression of pilin subunits from the GFP-D cell lysate was not detected. Antisera from rat #2 also bound pilin from GFP-DD but appeared to bind two other low molecular weight products that do not correspond to the pilin subunit, as these bands were also present in cell lysate prepared from ICC8001 which was not carrying any plasmid (Figure 3.11A). Although significantly more unspecific binding was observed with antisera from rat #2, it was used for further experiments as it displayed higher affinity for the pilin subunit.

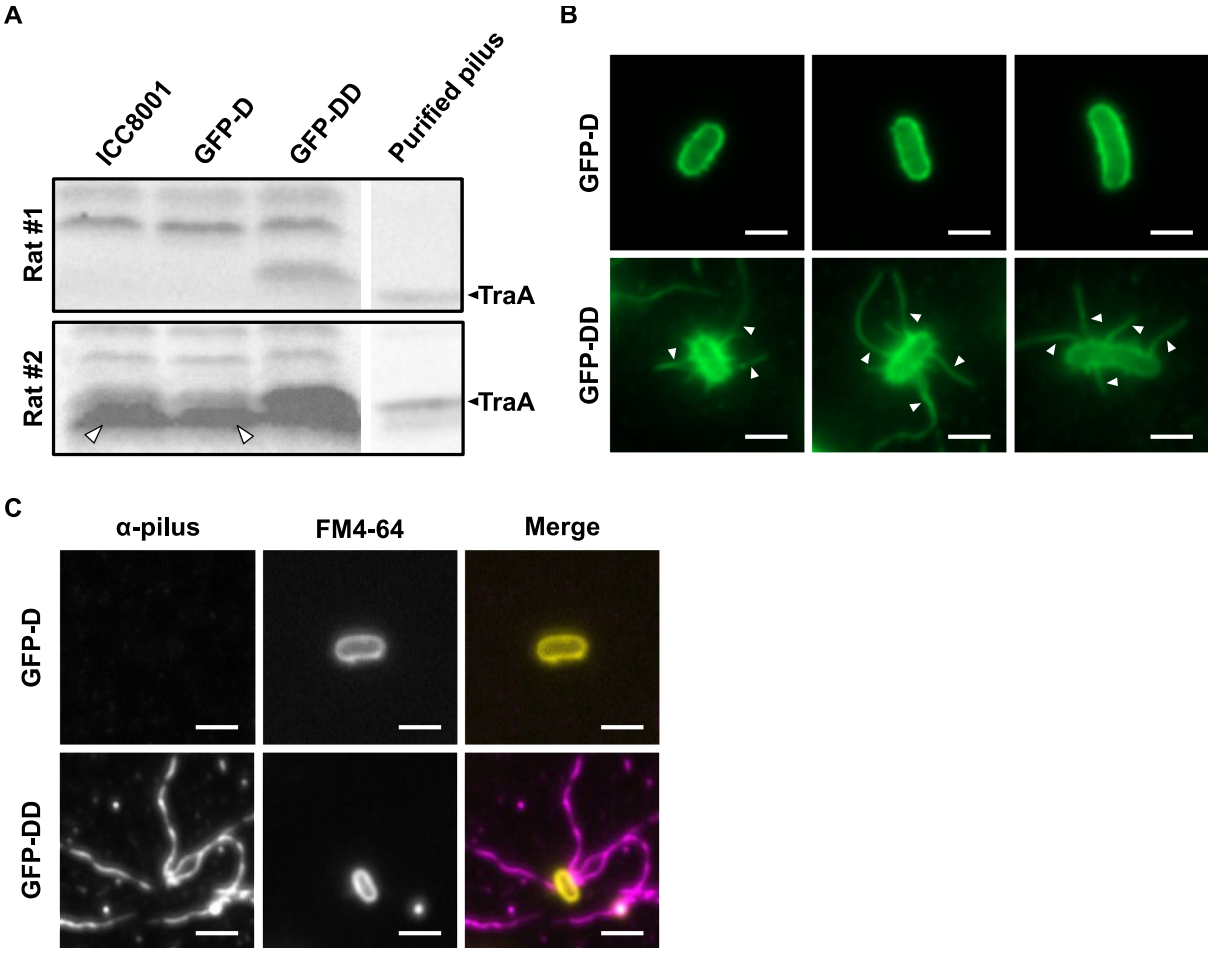


Figure 3.11. Antibodies raised against purified pilus

A. Antibodies raised against the purified pilus were assessed for specificity and affinity for TraA by Western blot. Cell lysates were prepared from ICC8001 (no plasmid control), GFP-D and GFP-DD. A sample of purified pilus was also included as a TraA control. Arrows indicate non-specific bands. **B.** Immunofluorescence staining of several GFP-D and GFP-DD cells with antisera from Rat #2. Conjugative pili are indicated with the white arrows. **C.** Immunofluorescence microscopy of GFP-D and GFP-DD cells with adsorbed anti-pilus antibodies (magenta) and FM4-64 membrane dye (yellow). Scale bars = 2 μm.

We next used pure cultures of donor cells to visualise the pili by immunofluorescence (IF) microscopy with the raised antibodies. GFP-DD cells appeared piliated while GFP-D cells did not (Figure 3.11B). This observation aligns with the Western blots performed on cell lysates and may suggest that FinO restricts pilus biogenesis in the absence of recipient cells in this system. To improve the specificity of the raised antibodies, the antisera were adsorbed against PFA-fixed ICC8001 which did not carry pKpGFP. A total of 6 overnight adsorptions were

performed and the IF microscopy was repeated (Figure 3.11C). The bacterial membrane was counterstained with the FM4-64^{TM} membrane dye. Adsorption was found to significantly reduce the presence of antibodies which were not specific to the pilus, allowing for better visualization of these structures. Again, no pili were detected in slides containing GFP-D cells using the adsorbed antibodies (Figure 3.11C).

The GFP-DD strain was also used in live microscopy to observe real-time transfer of the tagged, derepressed plasmid at the single cell level. GFP-DD cells were mixed with *dTomato*-expressing recipients and the mixture was spotted onto agarose pads supplemented with M9 salts and 0.4% glucose. Transconjugant cells were observed from onwards of 60 minutes of incubation at 37°C (Figure 3.12).

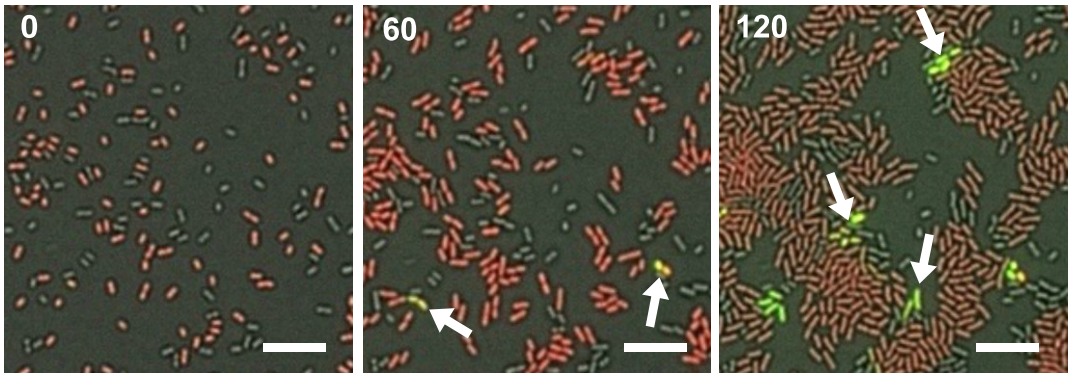


Figure 3.12. Visualising transfer of pKpGFP-D by live microscopy

Live microscopy stills captured at 0-, 60- and 120-min showing acquisition of pKpGFP-D by recipient cells. GFP-DD cells are not fluorescent, *d-Tomato*-expressing recipient cells appear red and transconjugant cells (indicated with white arrows) appear green due to *sfGFP* expression off pKpGFP-D. Scale bar = 10 μm.

3.8 Conjugation experiments performed with GFP-DD

To determine the impact of *tra* operon derepression on conjugation frequency, selection-based assays were used to compare the GFP-D and GFP-DD donors. Deletion of *finO* resulted in an overall increase in conjugation frequency regardless of the OmpK36 isoform

that is expressed in the recipient (Figure 3.13A). However, the relative difference in conjugation frequency of approximately 2 log-fold between these recipients was similar to what was previously observed with pKpGFP. This shows that derepression does not abrogate the dependency on recipient OM porins, allowing for the use of GFP-DD to investigate the impact of the ST258 mutations on conjugation.

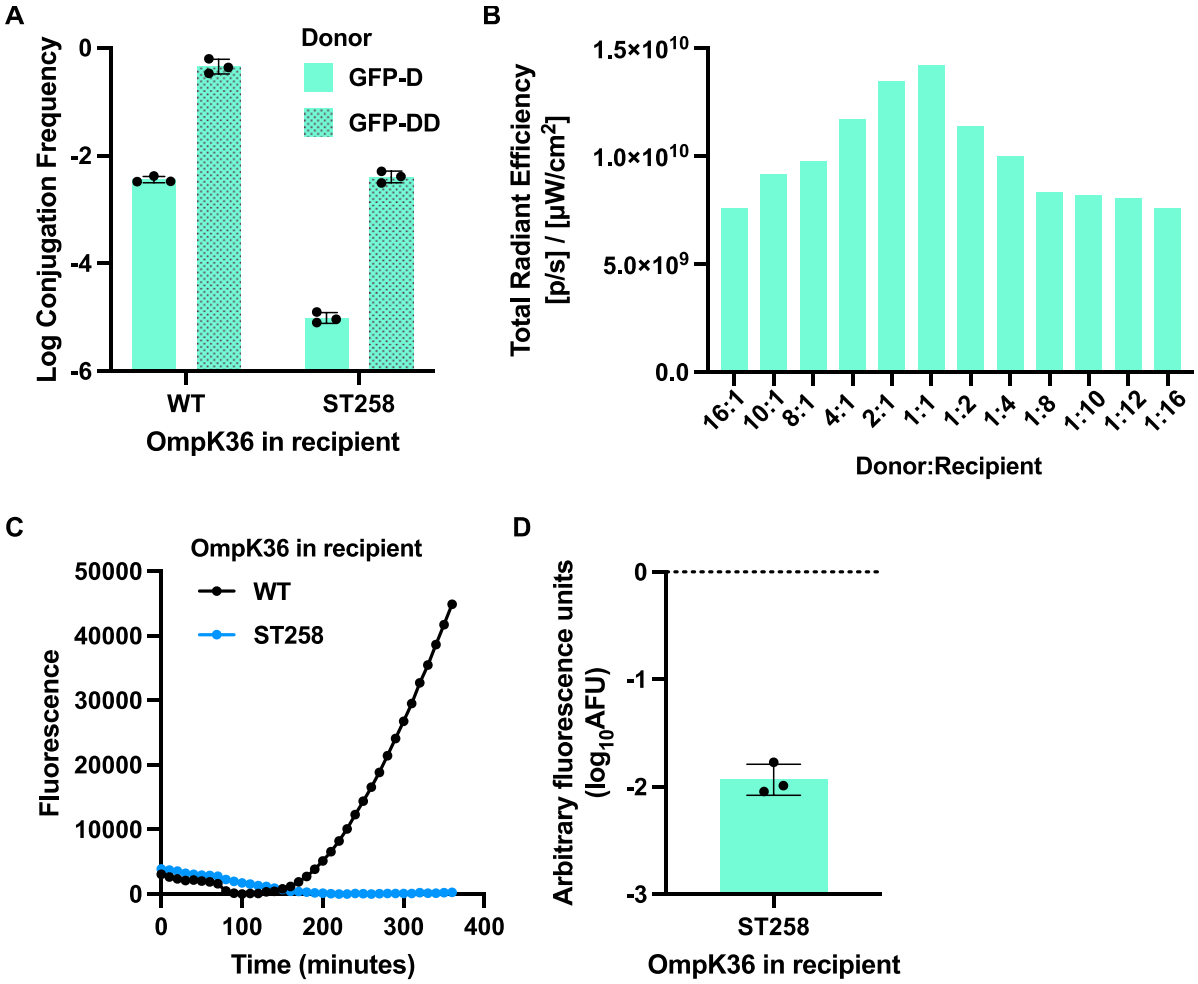


Figure 3.13. Derepression of pKpGFP for use in the RTCS

A. Log conjugation frequency from GFP-D and GFP-DD was compared into recipients expressing the WT and ST258 isoforms of OmpK36. Error bars represent SD (*n* = 3). **B.** Total radiant efficiency was measured from conjugation mixtures containing different ratios (v/v) of GFP-DD and WT recipients. **C.** Representative graph of fluorescence emission over time from conjugation mixtures containing GFP-DD and recipients expressing $OmpK36_{WT}$ or $OmpK36_{ST258}$. **D.** Arbitrary fluorescence units (AFU) for the $OmpK36_{ST258}$-expressing recipient were calculated as the $\log_{10}$ fold change from the recipient expressing the WT isoform. Error bars represent SD (*n* = 3).

Using GFP-DD, a 1:1 ratio of neat donor and WT recipient cultures gave the maximal fluorescence emission (Figure 3.13B). Under these conditions, transfer of pKpGFP-D into recipients expressing either 36_{WT} or 36_{ST258} was compared. Fluorescence emission over time increased into the recipient expressing 36_{WT} but not the 36_{ST258}-expressing recipient (Figure 3.13C). These findings aligned with the results obtained from the selection-based assays and was further validated by calculating the AFU for the mutant at t = 300 min (Figure 3.13D).

3.9 pKpGFP conjugation is disrupted by the L3 GD insertion in OmpK36$_{ST258}$

When compared to OmpK36$_{WT}$, OmpK36$_{ST258}$ contains several differences in its amino acid sequence. In addition to several substitutions found throughout the porin, there are two insertions that are present in OmpK36$_{ST258}$. The first is the L3 GD insertion which is associated with increased carbapenem resistance in the ST258 isolates and the second is a leucine-serine-proline (LSP) insertion in the extracellular loop, loop 4 (L4) (Figure 3.14). The crystal structures of OmpK36$_{WT}$ and OmpK36$_{ST258}$ are shown in Figure 3.15 highlighting the effect of these mutations on the overall 3-dimensional structure of the porin.

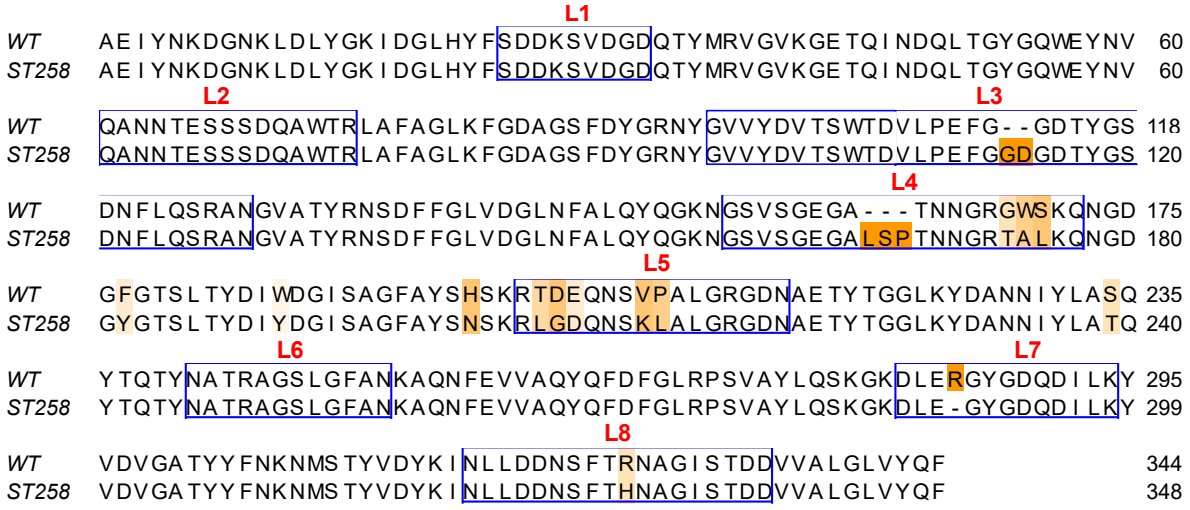


Figure 3.14. Multiple sequence alignment of WT and ST258 isoforms of OmpK36

Globally aligned amino acid sequences of OmpK36 isoforms. The N-terminal signal sequence has been removed and amino acid residues that make up the extracellular loops have been labelled in blue boxes. Orange shading represents non-conserved residues with the darker shading indicating lower degree of conservation and gaps.

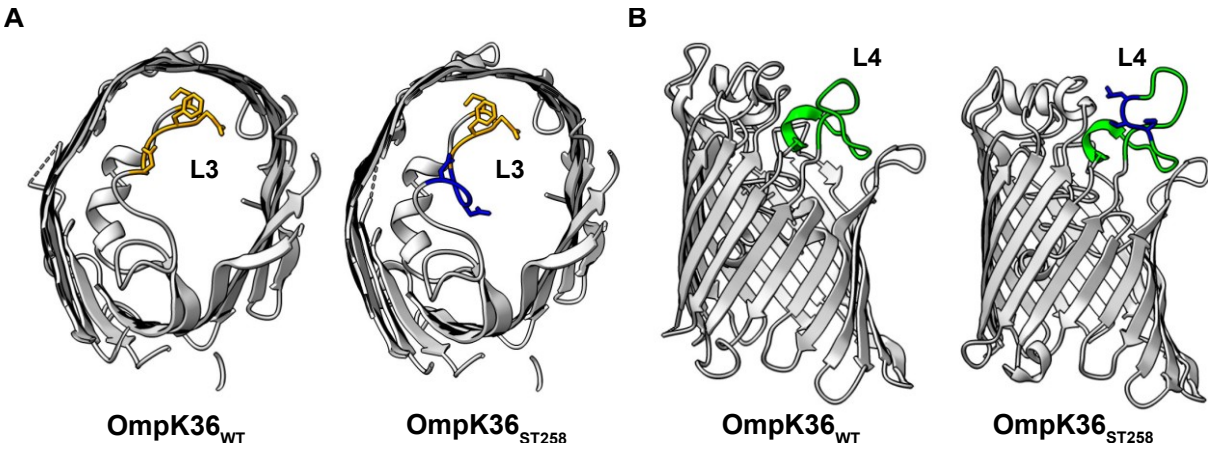


Figure 3.15. Structural comparisons of the WT and ST258 isoforms of OmpK36

Top (**A**) and side (**B**) views of OmpK36$_{WT}$ (PDB ID: 6RD3) and OmpK36$_{ST258}$ (PDB ID: 6RCP) monomers as determined by X-Ray crystallography. Side chains from the conserved PEFGGD motif in L3 are highlighted in orange in (**A**) with side chains from the GD insertion coloured in blue. L4 is shown in green in (**B**) with the LSP insertion side chains shown in blue.

To determine which insertion results in the reduction in pKpGFP conjugation, recipients that express an OmpK36$_{ST258}$ isoform where the insertions have been deleted (36$_{ST258\Delta GD}$ and 36$_{ST258\Delta LSP}$) were used. The L4 insertion, being surface exposed, was hypothesised to be involved in conjugation. This is based on previous studies which found that mutations in the surface exposed L4 in *E. coli* OmpA affect F plasmid transfer[173]. However, using the RTCS revealed that deletion of the L4 LSP insertion had no effect on conjugation frequency compared to 36$_{ST258}$-expressing recipient. In contrast, conjugation was observed to be affected by the L3 GD insertion as deletion of these amino acids restored the calculated AFU to levels seen with the 36$_{WT}$-expressing strain (Figure 3.16A). The effect of the GD insertion could be recapitulated in 36$_{WT}$ by inserting these amino acids into L3 (36$_{WT+GD}$) (Figure 3.15A). These findings were corroborated using selection-based assays (Figure 3.16B).

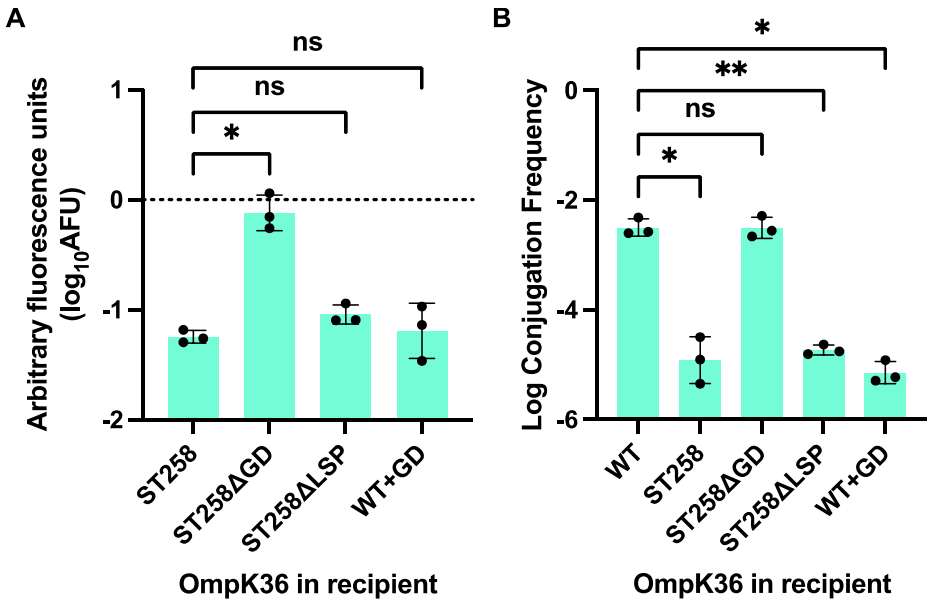


Figure 3.16. The L3 GD insertion is responsible for reduced plasmid uptake

A. The RTCS was used to assess plasmid transfer from mixtures containing GFP-DD and recipients expressing the WT, ST258, ST258ΔGD, ST258ΔLSP and WT+GD isoforms of OmpK36. AFU data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test comparing to the $OmpK36_{ST258}$-expressing recipient. **B.** Log conjugation frequency of pKpGFP was calculated for the same panel of recipients. Data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test comparing to the $OmpK36_{WT}$-expressing recipient. * = $p < 0.05$, ** = $p < 0.01$, ns = non-significant. Error bars represent SD ($n = 3$).

3.10 Differences in nutrient availability do not account for reduced plasmid uptake

Previous work has shown that the L3 GD insertion can impair uptake of some nutrient molecules via OmpK36[52]. Specifically, liposome swelling assays showed that while the diffusion of glucose, a monosaccharide, was unaffected by this insertion, uptake of larger nutrient molecules like lactose was significantly reduced when compared to uptake through $OmpK36_{WT}$. Moreover, III et al., determined that starvation of recipient cells during conjugation experiments yielded fewer transconjugants compared to when recipients were not starved[149]. Therefore, it was hypothesised that the L3 GD insertion results in reduced conjugation frequency by limiting the uptake of complex nutrient molecules in recipient cells. To test this, conjugation mixtures were prepared with GFP-D donors and incubated on M9 agar

supplemented with 0.4% glucose, sucrose, or lactose as the sole carbon source to control for nutrient availability (Figure 3.17).

Based on the hypothesis, the L3 GD insertion should not result in impaired plasmid uptake when mixtures were incubated on glucose as uptake of this carbon source is not affected by the pore constriction. In contrast, conjugation frequency is expected to be reduced when mixtures were incubated on lactose which does not diffuse freely through the isoforms containing the GD insertion. From the conjugation data, it was observed that the L3 GD insertion results in a reduction in pKpGFP uptake on all the carbon sources tested, including glucose (Figure 3.17A). This suggests that differences in nutrient uptake do not account for the reduction in conjugation frequency.

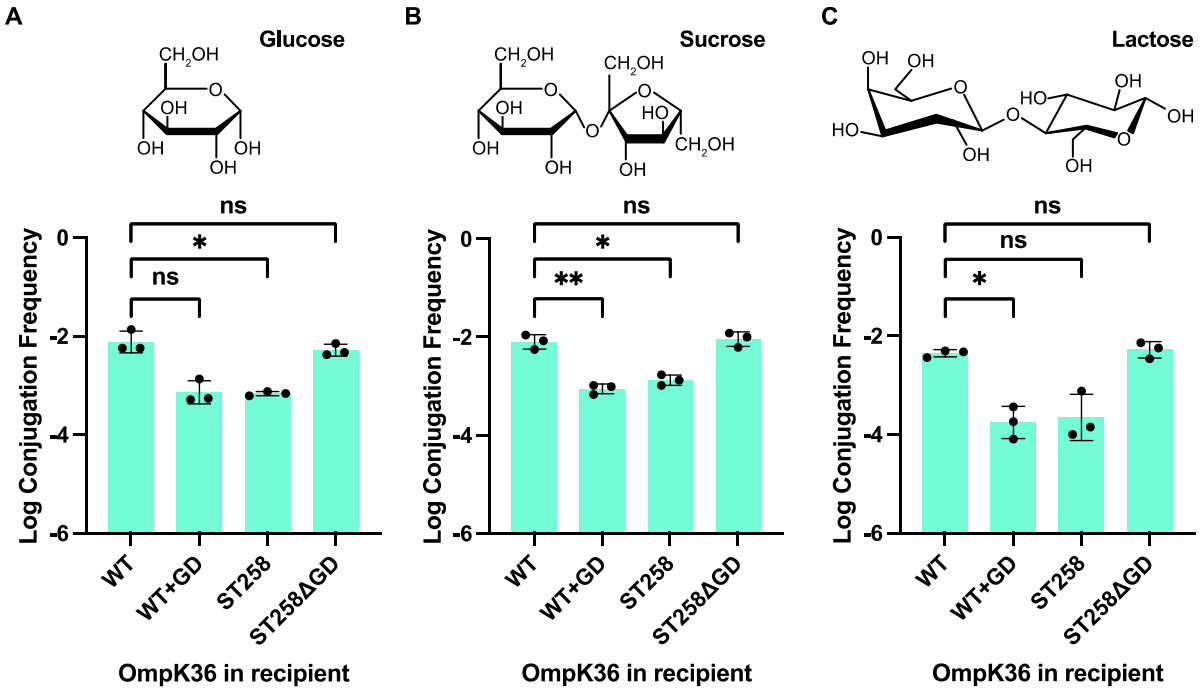


Figure 3.17. Investigating the effect of nutrient availability on pKpGFP transfer

Log conjugation frequency was calculated from mixtures of GFP-D donors and recipients expressing the WT, WT+GD, ST258 and ST258ΔGD isoforms of OmpK36 incubated on M9 minimal media supplemented with 0.4% (**A**) glucose, (**B**) sucrose and (**C**) lactose. The chemical structure of each carbon source is shown. Conjugation data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test. * = $p < 0.05$, ** = $p < 0.01$, ns = non-significant. Error bars represent SD ($n = 3$).

3.11 The salt bridge does not affect plasmid uptake

From analysis of the crystal structure of OmpK36$_{WT+GD}$, it was determined that the GD insertion in L3 results in a conformational change in the loop that leads to the formation of a salt bridge between D114 in L3 and residue R127 in the opposing face of the barrel (Figure 3.18A)[52]. It was hypothesised that the formation of this salt bridge interferes with conjugation. This was assessed using recipient strains in which the R127 residue has been mutated into an alanine (R127A) or lysine (R127K). The R127A mutation removes the charged residue and is expected to disrupt the salt bridge, while the R127K mutation changes the size of the side chain while preserving the charge properties of the residue. Neither mutation had a significant impact on plasmid transfer compared to the OmpK36$_{ST258}$-expressing recipient (Figure 3.18B) suggesting that salt bridge formation does not impair conjugation.

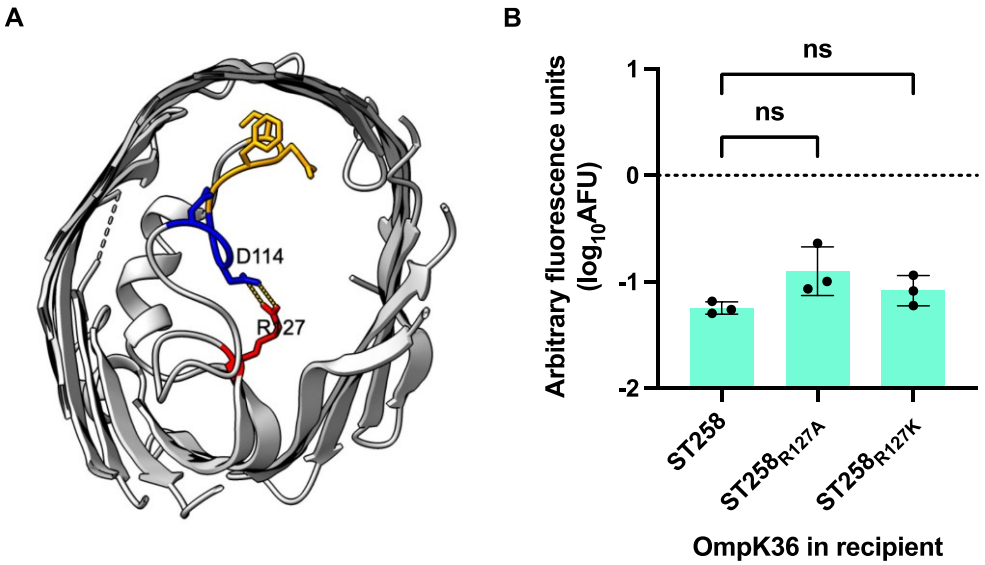


Figure 3.18. Assessing the effect of the salt bridge on conjugation

A. Top view of an OmpK36$_{ST258}$ monomer (PDB ID: 6RCP) showing the L3 PEFGGD motif in orange and the residues involved in the formation of the salt bridge (D114 in blue; R127 in red). The salt bridge is illustrated as lines linking both side chains. **B.** The RTCS was used to assess plasmid transfer from mixtures containing GFP-DD and recipients expressing the WT, ST258, ST258$_{R127A}$ and ST258$_{R127K}$ isoforms of OmpK36. AFU data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test. ns = non-significant. Error bars represent SD (*n* = 3).

3.12 Charge effects associated with the L3 GD insertion do not affect conjugation

It was further hypothesised that the GD insertion results in alterations to the charge distribution at the eyelet of OmpK36, and the resultant changes in the electrophysiological properties of the lumen affects the efficiency of pKpGFP uptake. To test this, two site-directed mutants were generated on the background of the 36_{ST258} isoform – D114N and D116N. Using both the RTCS (Figure 3.19A) and selection-based assays (Figure 3.19B), no significant difference in conjugation frequency between the parental 36_{ST258}-expressing strain and the generated mutants was observed suggesting that conjugation was not impaired due to a charge effect.

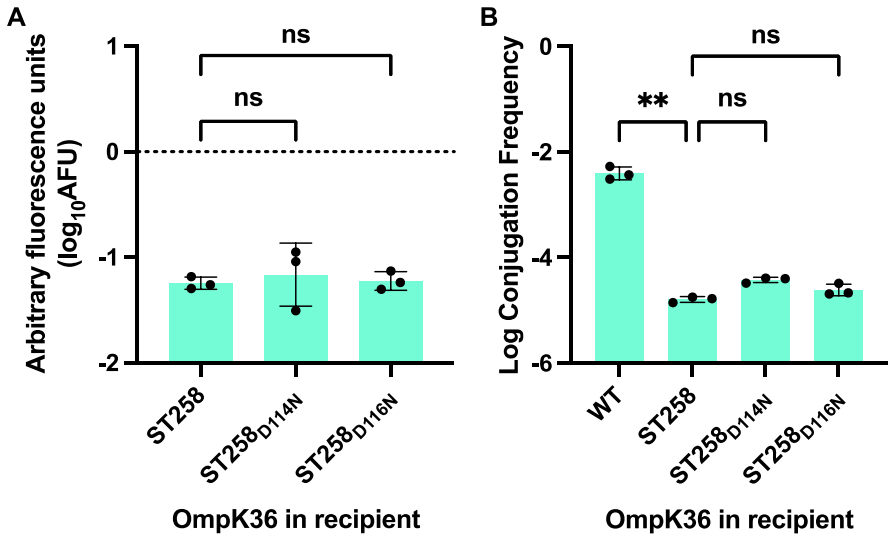


Figure 3.19. Investigating the effect of charge on plasmid uptake

(**A**) The RTCS and (**B**) selection-based assays were used to assess the effect of the D114N and D116N substitutions in $OmpK36_{ST258}$ on plasmid transfer. **A.** AFU data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test. **B.** Log conjugation frequency of pKpGFP into recipients expressing the WT, ST258, $ST258_{D114N}$ and $ST258_{D116N}$ isoforms of OmpK36 is shown. Data was analysed by RM one-way ANOVA with Tukey's multiple comparison test. ** = $p < 0.01$, ns = non-significant. Error bars represent SD ($n = 3$).

3.13 A single amino acid insertion in L3 has an intermediate effect on conjugation

In addition to the GD insertion, analysis of *ompK36* genes in a collection of ST258 strains revealed two other L3 insertions – a threonine-aspartic acid (TD) insertion and a single aspartic acid (D) insertion[54]. Structural analysis of OmpK36 containing these insertions

revealed that the TD and D insertions result in a 41% and 8% constriction of the WT pore respectively. It was hypothesised that pKpGFP uptake may be inversely correlated with the extent of the pore constriction. To test this, the RTCS was used to assess pKpGFP-D transfer into recipients expressing $OmpK36_{WT+TD}$ or $OmpK36_{WT+D}$ alongside a recipient expressing $OmpK36_{WT+GD}$ and a Δ*ompK36* recipient (Figure 3.20A). While there was no significant difference in AFU between the Δ*ompK36* recipient and those expressing OmpK36 with the TD and GD insertions, the recipient expressing $OmpK36_{WT+D}$ had significantly higher AFU. Selection-based assays confirmed that the TD and GD insertions result in a similar reduction in conjugation frequency compared to the WT isoform of OmpK36 (Figure 3.20B). In contrast, the reduction associated with the D insertion was approximately a log-fold less than the reduction attributed to the diamino acid insertions, resulting in an intermediate phenotype.

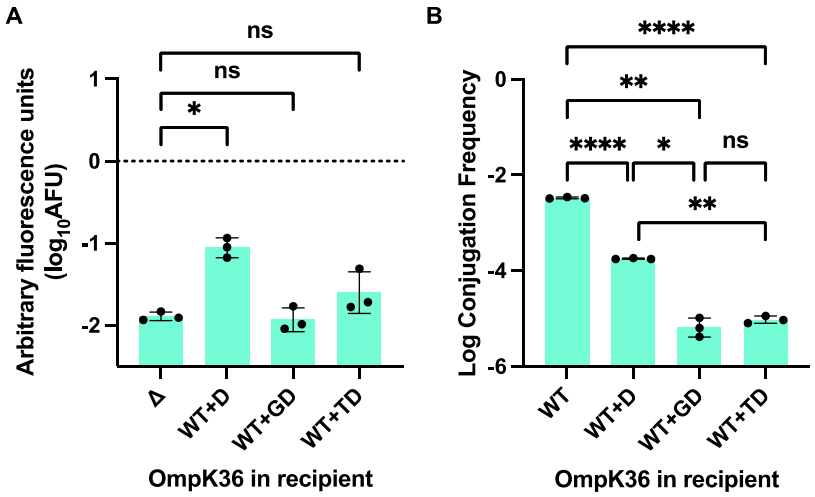


Figure 3.20. A single amino acid L3 insertion has an intermediate effect on conjugation

A. The RTCS was used to assess transfer of pKpGFP-D into recipients with various L3 insertions and a Δ*ompK36* recipient. AFU data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test comparing to the Δ*ompK36* recipient. **B.** The selection-based assay was used to quantify the effect of the different OmpK36 L3 insertions. Data was analysed by RM one-way ANOVA with Tukey's multiple comparison test. * = $p < 0.05$, ** = $p < 0.01$, **** = $p < 0.0001$, ns = non-significant. Error bars represent SD ($n = 3$).

3.14 Single amino acid insertions have varying effects on plasmid uptake

As the single D insertion results in an intermediate phenotype with regards to pKpGFP uptake, it was hypothesised that single amino acid insertions may provide a better understanding of how changes in L3 impact conjugation. This contrasts with diamino acid insertions, where mutating the L3 GD insertion to a GN or TD insertion did not result in any significant changes to plasmid uptake and results in conjugation frequency that is comparable to Δ*ompK36* recipients.

A panel of OmpK36$_{WT+X}$-expressing strains were generated by site-directed mutagenesis with X representing each of the remaining 19 naturally occurring amino acids (Figure 3.21A). OM proteins were isolated from each of these strains and analysed by SDS-PAGE to confirm that OmpK36 was being expressed normally. Normal expression levels of OmpK36 were seen in all strains generated except for the strain expressing OmpK36$_{WT+G}$ in which OmpK36 expression was greatly reduced (Figure 3.21B). A second strain with this insertion was generated (OmpK36$_{WT+G2}$) using an alternative glycine codon (GGC) which restored OmpK36 expression to WT levels. As it was not within the scope of this project, reduced porin expression resulting from the use of the GGT codon in OmpK36$_{WT+G}$ was not investigated.

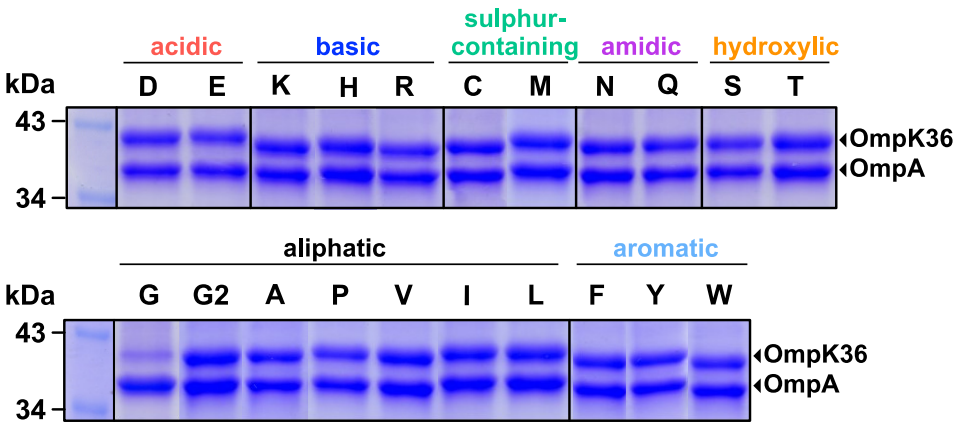
The RTCS was used to assess the impact on conjugation of each of the amino acid insertions in L3 (Figure 3.21C). The mean AFU calculated for the OmpK36$_{WT+D}$ recipient was used as a reference against which to compare the other amino acid insertions. Both the side chain properties and molecular weight of the residues were considered when analysing the effect of different insertions on pKpGFP-D uptake (Figure 3.21A).

A

Properties	Acidic		Basic			Sulphur-containing		Amidic		Hydroxylic	
Amino acid	D	E	K	H	R	C	M	N	Q	S	T
Mw	133.1	147.1	146.2	155.2	174.2	121.2	149.2	132.1	146.2	105.1	119.1

Properties	Aliphatic						Aromatic		
Amino acid	G	A	P	V	I	L	F	Y	W
Mw	75.1	89.1	115.1	117.2	131.2	131.2	165.2	181.2	204.2

B



C

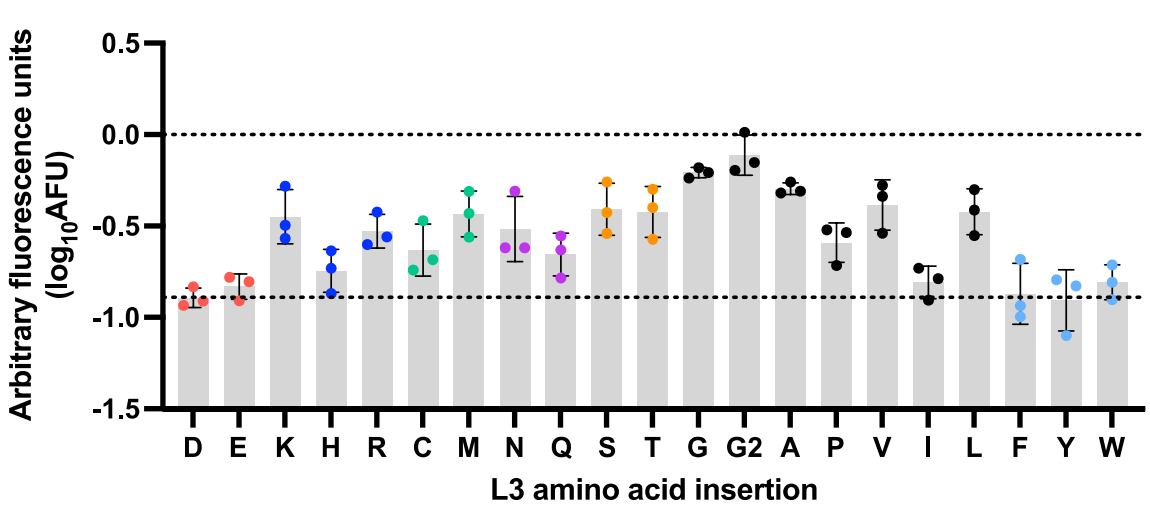


Figure 3.21. Single L3 amino acid insertions have varied effects on conjugation

A. Table of amino acids arranged according to side chain properties and in order of increasing molecular weight (Mw). **B.** Coomassie stained SDS-PAGE gel of OM proteins from OmpK36 mutants. Bands corresponding to OmpK36 and OmpA are indicated. **C.** The RTCS was used to assess the effect of each insertion on pKpGFP-D uptake. Arbitrary fluorescence units (AFU) were calculated at t = 300 min. The data points are coloured according to amino acid properties and arranged in order of increasing Mw. The lower dotted line represents the average AFU calculated for the $OmpK36_{WT+D}$-expressing recipient. Error bars represent SD (n = 3).

Firstly, insertion of a negatively charged glutamic acid (E) resulted in relatively similar AFU values compared to the D insertion, despite the higher molecular weight (Figure 3.21C). Next, higher AFU values were observed when positively charged basic residues were inserted into L3, apart from histidine (H) which showed no increase. However, these values do not appear to correspond to the molecular weight of the residue. Insertion of the sulphur-containing residues, cysteine (C) and methionine (M), led to an increase in AFU and this appeared to be correlated with residue molecular weight. However, this correlation between molecular weight of the residue and AFU was not seen with the amidic or hydroxylic residues which were all associated with increase plasmid uptake. Interestingly, insertion of a single asparagine (N) residue was associated with higher AFU values than observed with insertion of aspartic acid (D). This contrasts with previous experiments where the D114N and D116N mutations did not lead to a significant difference in conjugation frequency in the context of the two amino acid GD insertion (Figure 3.19).

In recipients expressing OmpK36 isoforms with aliphatic amino acid insertions, a general trend in which AFU was inversely related to the molecular weight of the residue was observed (Figure 3.21C). Isoleucine (I), however, appears to be an outlier in this regard as it shows a similar conjugation frequency to the D insertion despite having the same molecular weight as leucine (L). In addition, both strains containing glycine insertions (G and G2) had similar AFU values. This finding supports previous results showing that differences in OmpK36 abundance do not affect pKpGFP uptake. Finally, insertion of all three aromatic residues were associated with similar AFU values to the D insertion.

3.15 In silico modelling of single amino acid insertions

To gain a better understanding of how these insertions may be affecting pKpGFP-D conjugation, in silico modelling was performed with the help of Dr Konstantinos Beis using the crystal structure of $OmpK36_{WT+D}$ for 7 amino acid insertions of interest (Figure 3.22). Steric clashes with L295 were predicted in the models of OmpK36 containing the R, H and I

insertions that could not be alleviated by their different rotamers (Figure 3.22). This would likely require that L3 adopt a different conformation which further constricts the pore and provides a possible explanation for the significantly lower AFU calculated for the isoform containing the I insertion compared to the L insertion. However, where the insertion contains a basic or acidic residue, these steric clashes (or the lack thereof) do not correspond to the trend in AFU observed. Taken together, these results suggest that conjugation frequency may be directly correlated with the extent of pore constriction resulting from insertions in L3. However, this correlation does not apply to the insertion of charged residues.

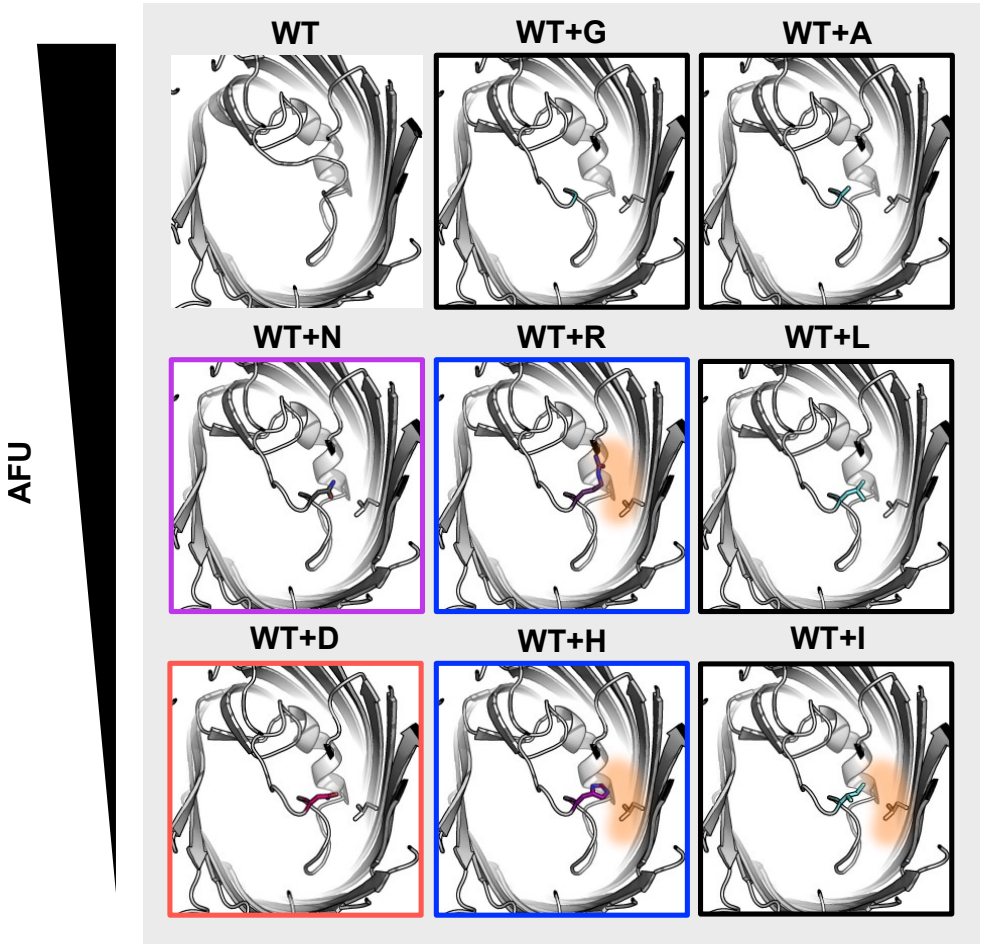


Figure 3.22. Modelling the L3 amino acid insertions

Seven L3 amino acid insertions were computationally modelled using the crystal structure of $OmpK36_{WT+D}$ (PDB ID: 7Q3T) as a scaffold. The colour of the borders indicates the side chain properties of each inserted amino acid (black = aliphatic, purple = amidic, blue = basic, red = acidic). Orange indicates a predicted clash of side chains between the inserted residue and residues present in the barrel of OmpK36. Predicted structures are arranged from top to bottom in order of decreasing AFU based on the data shown in Figure 3.21.

3.16 Investigating plasmid transfer on different mating substrates

The preceding data strongly supports the direct involvement of OmpK36 during pKpGFP conjugation. This dependency is similar to OmpA dependency observed during F plasmid transfer which was mediated by the plasmid encoded TraN[112]. Therefore, it was hypothesised that recipient OmpK36 may play a role in MPS via TraN from pKpQIL ($TraN_{pKpQIL}$).

Previously, it was proposed that MPS is important for improving conjugation efficiency in liquid media as experiments performed with the *traN548* amber mutant revealed a more pronounced reduction in transfer efficiency in liquid media compared to on solid media[141]. Moreover, OmpA dependency could be alleviated when F plasmid conjugation was allowed to occur on a solid surface[100].

In this work, all previously described experiments had been incubated on LB agar. Based on the hypothesis that OmpK36 plays a role in MPS, it was predicted that incubating conjugation mixtures in liquid media would result in a greater difference in conjugation frequency between recipients expressing $OmpK36_{WT}$ and $OmpK36_{WT+GD}$. However, the difference in conjugation frequency between both recipient strains was less than 2 log-fold (Figure 3.23A) which is relatively smaller than the difference observed on LB agar (Figure 3.16). The overall conjugation frequency of pKpGFP in liquid was also observed to be much lower than on solid media using the same ratio of donor to recipient cells.

Next, an alternative method was tested where conjugation mixtures were incubated on filter paper. This method is proposed to provide additional support for conjugating bacteria and reduces the role of mating pair stabilization in limiting conjugation efficiency[174]. pKpGFP transfer was compared into recipients expressing the WT and WT+GD isoforms of OmpK36 but a significantly lower conjugation frequency of approximately 2-log fold was still associated with the GD insertion (Figure 3.23B).

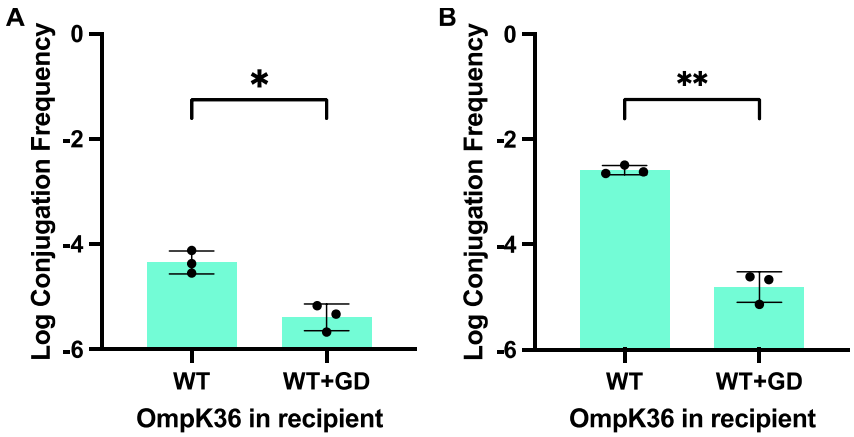


Figure 3.23. Investigating the effect of the GD insertion during liquid and filter mating

GFP-D was incubated with recipients expressing the WT and WT+GD isoforms of OmpK36 (**A**) in liquid media and (**B**) on sterile filters. The log conjugation frequency from each mixture was determined and is shown. Data was analysed by paired *t*-test. * = $p < 0.05$, ** = $p < 0.01$. Error bars represent SD ($n = 3$).

3.17 TraN may be essential for pilus biogenesis

Although the previous observations suggested that OmpK36 does not play a role in MPS, due to the similarities observed with OmpA dependency during conjugation, the role of TraN in mediating this dependency was investigated. To do so, *traN* was first deleted from pKpGFP and pKpGFP-D to generate pKpGFPΔ*traN* and pKpGFP-DΔ*traN* respectively. Despite previous observations that disruption of *traN* does not completely interfere with pilus biogenesis and thus still allows for low efficiency plasmid transfer[112,175], no transconjugants could be detected above the limit of detection from conjugation mixtures containing the pKpGFPΔ*traN* donor (Figure 3.24A). IF microscopy of donor cells carrying the derepressed pKpGFP from which *traN* had been deleted (pKpGFP-DΔ*traN*) revealed that the cells were not piliated (Figure 3.24B). This suggests that TraN may play a more essential role in during conjugative transfer of pKpQIL compared to other IncF plasmids.

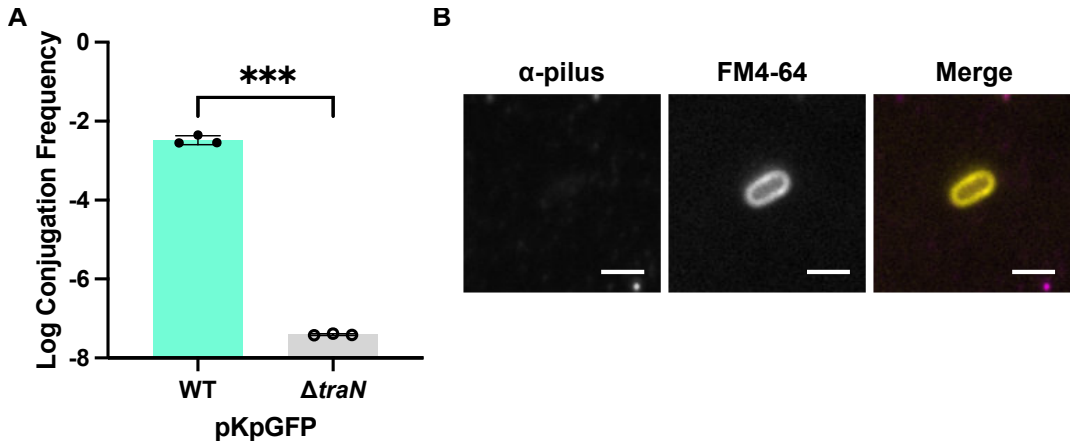


Figure 3.24. Deletion of traN abolishes donor cell piliation and transfer of pKpGFP

A. Immunofluorescence staining of GFP-DDΔ*traN* cells. Scale bar = 2 µm. **B.** Log conjugation frequency data from mixtures containing WT recipients and GFP-D or GFP-DΔ*traN* donors. Data analysed by paired *t*-test. *** = $p < 0.001$. Error bars represent SD ($n = 3$).

3.18 TraN mediates OmpK36 dependency during pKpGFP conjugation

Next, to determine if $TraN_{pKpQIL}$ specifically mediates OmpK36 dependency, the effect of mutations in recipient OmpK36 were assessed during conjugation of the related IncFII plasmid, R100-1. There was no significant difference in conjugation frequency between both recipients, confirming that R100-1 conjugation is not OmpK36-dependent (Figure 3.25).

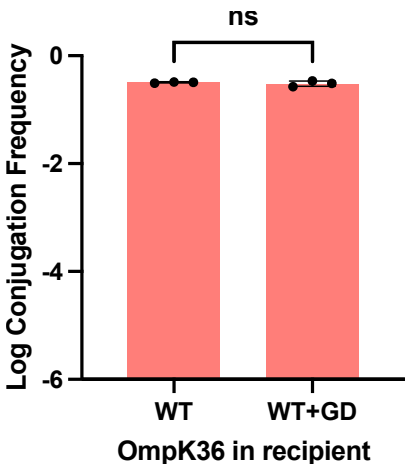


Figure 3.25. R100-1 conjugation is not affected by the L3 GD insertion in OmpK36

DH5α carrying R100-1 was incubated with recipients expressing the WT and WT+GD isoforms of OmpK36 and the log conjugation frequency was determined. Data was analysed by paired *t*-test. ns = non-significant. Error bars represent SD ($n = 3$).

The gene encoding TraN from R100-1 (*traN*$_{R100-1}$) was inserted into pKpGFPΔ*traN* and pKpGFP-DΔ*traN* to generate pKpGFP*traN*$_{R100-1}$ and pKpGFP-D*traN*$_{R100-1}$ respectively. Donor cells carrying pKpGFP-D*traN*$_{R100-1}$ were piliated (Figure 3.26A). Using the RTCS, fluorescence emission over time was seen to increase in all recipients tested regardless of the OmpK36 isoform being expressed and there was no significant difference in AFU (Figure 3.26B&C). This was validated using selection-based assays (Figure 3.26D). Taken together, these results suggest that TraN$_{R100-1}$ can functionally complement the *traN* deletion in pKpGFP and it also abrogates OmpK36 dependency like what was observed with OmpA dependency during F plasmid conjugation[112]. By extension, this would suggest that TraN$_{pKpQIL}$ cooperates with OmpK36 on recipient cells to facilitate MPS.

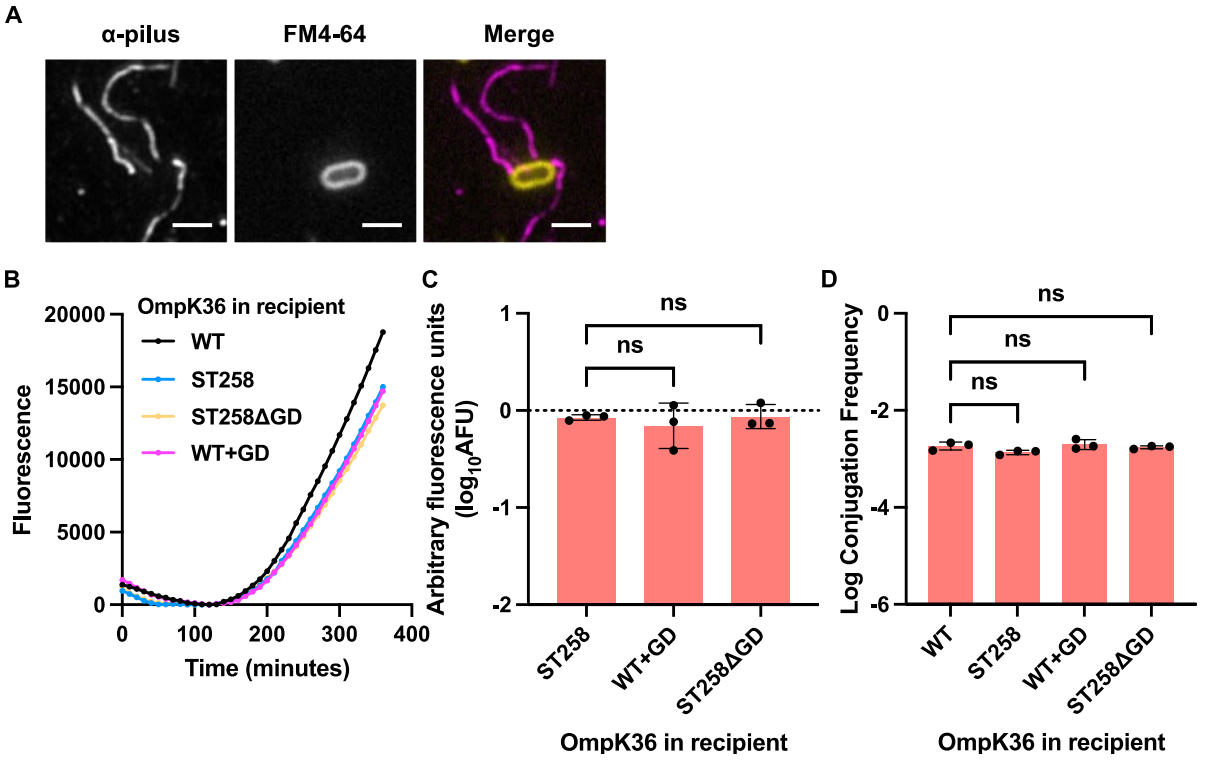


Figure 3.26. Substitution of traN abrogates effect of L3 GD insertion on conjugation

A. IF staining of GFP-DD*traN*$_{R100-1}$ cells with anti-pilus antibodies (magenta) and FM4-64 (yellow). Scale bar = 2 μm. **B.** Representative RTCS graph showing fluorescence emission over time from mixtures containing GFP-DD*traN*$_{R100-1}$ and the panel of isogenic OmpK36 mutant recipients. **C.** AFU data analysed by RM one-way ANOVA with Dunnett's multiple comparison test. **D.** Log conjugation frequency data from mixtures containing GFP-D*traN*$_{R100-1}$ and the panel of isogenic OmpK36 mutant recipients. Data analysed by RM one-way ANOVA with Dunnett's multiple comparison test. ns = non-significant. Error bars represent SD (*n* = 3).

3.19 TraN from R100-1 mediates dependency on recipient OmpW

As the conjugation frequency of pKpGFP*traN*$_{R100-1}$ into KP expressing various combinations of OmpK35 and OmpK36 was comparable to that of pKpGFP into WT KP recipients, it was hypothesised that TraN$_{R100-1}$ mediates MPS via another recipient OM protein. Two OM proteins that are conserved in *Enterobacteriaceae* were identified – PhoE and OmpW and Δ*phoE* and Δ*ompW* strains were generated[176,177]. A Δ*ompA* mutant was also generated as a control as R100-1 conjugation was previously shown to be unaffected by mutations in OmpA[112]. Using GFP-DD*traN*$_{R100-1}$ in the RTCS, conjugation frequency was measured into these isogenic recipient strains alongside the single porin deletion strains (Δ*ompK36,* Δ*ompK35* and Δ*ompK37*). The AFU calculated for the Δ*ompW* strain was significantly lower than all other strains tested (Figure 3.27A). Moreover, there was no increase in fluorescence over time specifically into the Δ*ompW* recipient (Figure 3.27B). This suggests that TraN$_{R100-1}$ mediates MPS with recipient OmpW during conjugation.

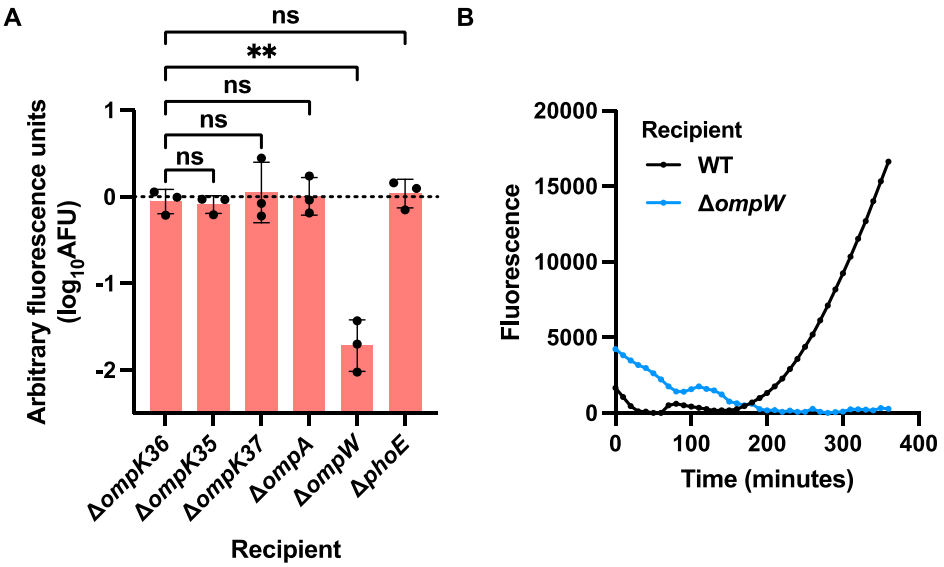


Figure 3.27. TraN$_{R100-1}$ mediates dependency on recipient OmpW during conjugation

A. The RTCS was used to determine the OM protein which cooperates with TraN$_{R100-1}$ during MPS. AFU data analysed by RM one-way ANOVA with Dunnett's multiple comparison test. ns = non-significant. Error bars represent SD (*n* = 3). **B.** Representative RTCS graph showing fluorescence emission over time from mixtures containing GFP-DD*traN*$_{R100-1}$ and WT and Δ*ompW* recipients.

3.20 OM protein specificity is associated with the variable domain of TraN

To determine the region of TraN which confers specificity for different OM proteins, the amino acid sequences from $TraN_{pKpQIL}$ and $TraN_{R100\text{-}1}$ were aligned and a region of low sequence identity between the two homologues was identified (Figure 3.28). This domain was found to be located within the variable region identified by Klimke et al., on $TraN_F$ which mediates specificity for recipient OmpA[112].

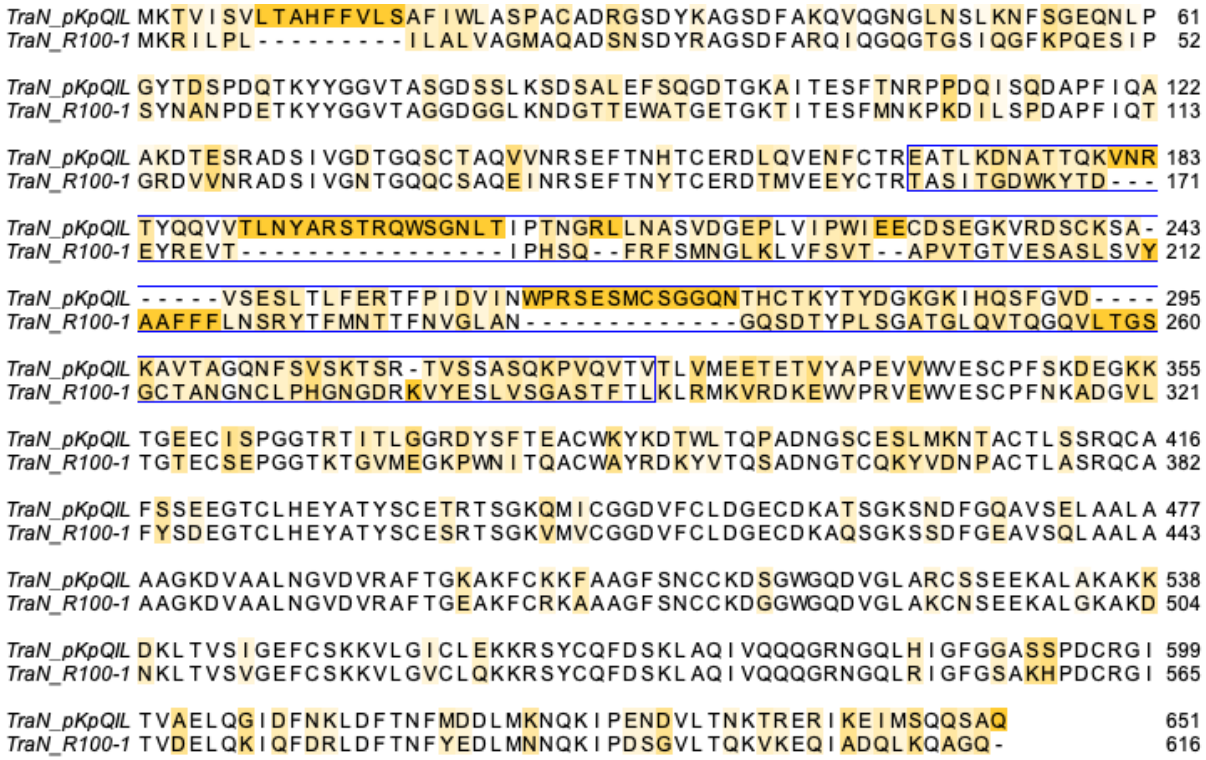


Figure 3.28. Sequence alignment of TraN homologues

Globally aligned amino acid sequences of TraN from pKpQIL and R100-1. Darker shading indicates lower residue conservation in the homologues. The variable region which was identified for substitution is highlighted in the blue box.

Based on the sequence alignment, a chimeric TraN ($TraN_{Ch1}$) was generated by substituting the nucleotide sequence coding for amino acids 169 to 325 from $traN_{pKpQIL}$ with the nucleotides which encode amino acids 160 to 291 in $traN_{R100\text{-}1}$. GFP-DD cells expressing $TraN_{Ch1}$ were piliated suggesting that expression of this chimeric protein does not disrupt pilus biogenesis (Figure 3.29A). Using this donor strain in the RTCS, fluorescence over time was observed to

increase into WT and Δ*ompK36* recipients but not Δ*ompW* recipients (Figure 3.29B). Moreover, the AFU calculated into Δ*ompW* recipients was significantly lower compared to the Δ*ompK36* recipients (Figure 3.29C). This was validated using selection-based assays, where conjugation frequency was significantly lower into the Δ*ompW* recipient compared to the WT recipient (Figure 3.29D). The relative reduction in conjugation frequency, of approximately 2 log-folds was similar to the reduction in conjugation frequency of pKpGFP into Δ*ompK36* or recipients which expressed OmpK36 isoforms containing the L3 GD insertion (Figure 3.16).

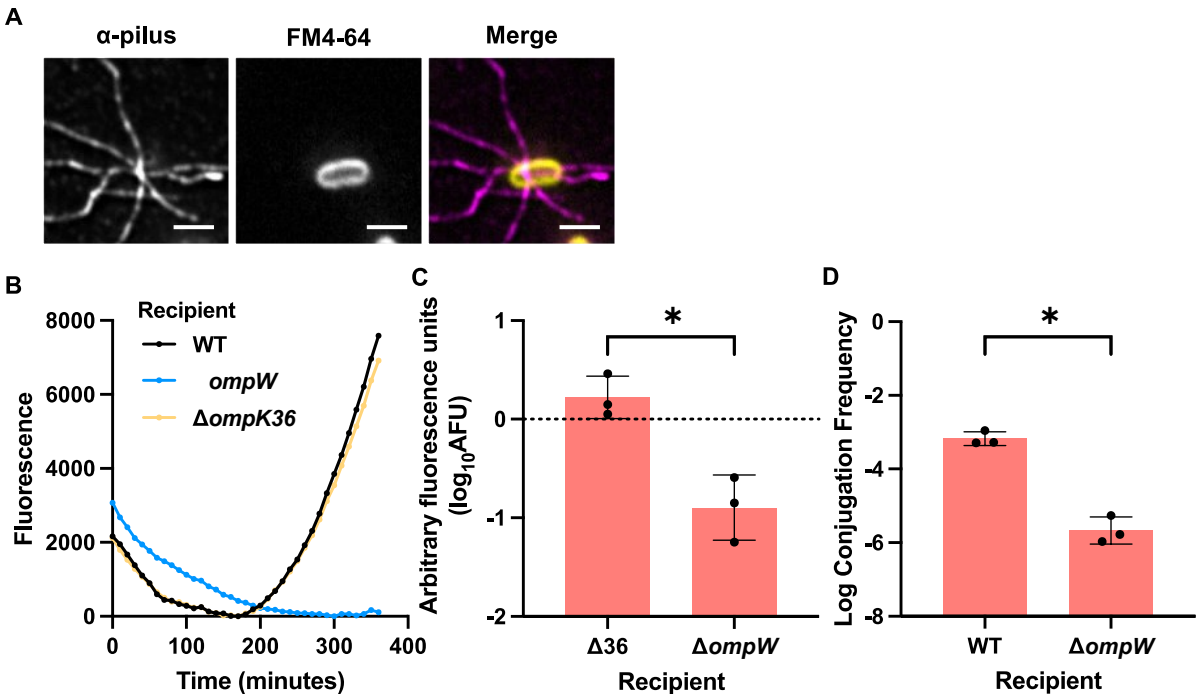


Figure 3.29. The variable region of TraN mediates recipient OM protein specificity

A. Immunofluorescence staining of GFP-DD*traN*$_{Ch1}$ cells with anti-pilus antibodies (magenta) and FM4-64 (yellow). Scale bar = 2 μm. **B.** Representative RTCS graph of fluorescence emission over time from conjugation mixtures containing GFP-DD*traN*$_{Ch1}$ and WT, Δ*ompK36* and Δ*ompW* recipients. **C.** AFU data was calculated for the Δ*ompK36* and Δ*ompW* recipients and analysed by paired *t*-test. **D.** The selection-based assay was used to confirm dependency on OmpW. Log conjugation frequency of pKpGFP*traN*$_{Ch1}$ was compared between mixtures containing WT and Δ*ompW* recipients. Data analysed by paired *t*-test. * = $p < 0.05$. Error bars represent SD ($n = 3$).

3.21 TraN mediates recipient species specificity during conjugation

$TraN_F$ was previously shown to mediate dependency on recipient OmpA. To validate this observation, *traN* on pKpGFP and pKpGFP-D was substituted with *traN*$_F$ to generate pKpGFP*traN*$_F$ and pKpGFP-D*traN*$_F$. Although donors carrying pKpGFP-D*traN*$_F$ (GFP-DD*traN*$_F$) were piliated (Figure 3.30A), no increase in fluorescence emission over time into WT KP recipients was observed with GFP-DD*traN*$_F$ in the RTCS (Figure 3.30B). Moreover, a negative fold change in AFU was calculated when compared to GFP-DD (Figure 3.30C).

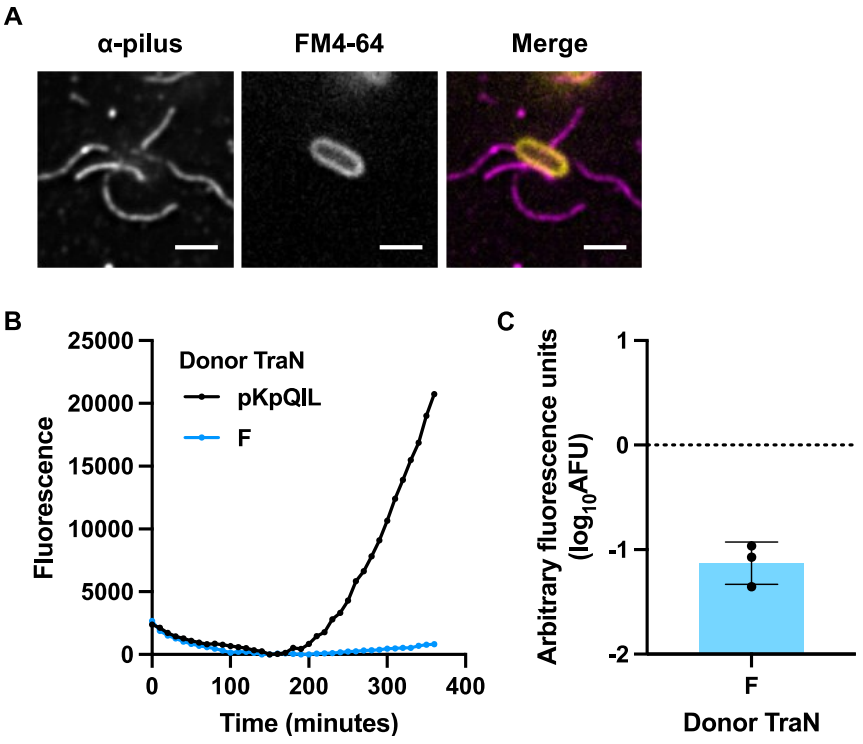


Figure 3.30. $TraN_F$ does not mediate efficient plasmid transfer into KP

A. Immunofluorescence staining of GFP-DD*traN*$_F$ cells with anti-pilus antibodies (magenta) and FM4-64 (yellow). Scale bars = 2 μm. **B.** Representative RTCS graph of fluorescence emission over time from conjugation mixtures containing GFP-DD or GFP-DD*traN*$_F$ with WT KP recipients. **C.** AFU data was calculated for GFP-DD*traN*$_F$. Error bars represent SD (*n* = 3).

Due to the limited sensitivity of the RTCS, low efficiency conjugation and the complete absence of plasmid transfer could not be discriminated. Thus, selection-based assays were used to determine if $TraN_F$ can support conjugation into KP. Donor strains expressing

$TraN_{pKpQIL}$ and $TraN_{R100-1}$ were included for comparison. Although transconjugants were detected from mixtures containing donors expressing $TraN_F$, the calculated log conjugation frequency was approximately 2 log-folds lower when compared to the mixtures containing donors expressing $TraN_{pKpQIL}$ or $TraN_{R100-1}$ (Figure 3.31A).

As F plasmid conjugation experiments were largely performed in *E. coli* K-12, the selection-based assays were repeated into *E. coli* MG1655 (EC) recipients. Although $TraN_F$ was now observed to mediate increased plasmid transfer into EC, the conjugation frequency associated with $TraN_{pKpQIL}$ was now significantly lower (Figure 3.31B). In contrast, $TraN_{R100-1}$ mediated comparable levels of conjugative transfer into both species of recipient. These results suggest that different TraN homologues are associated with a species-specific effect on conjugative plasmid transfer.

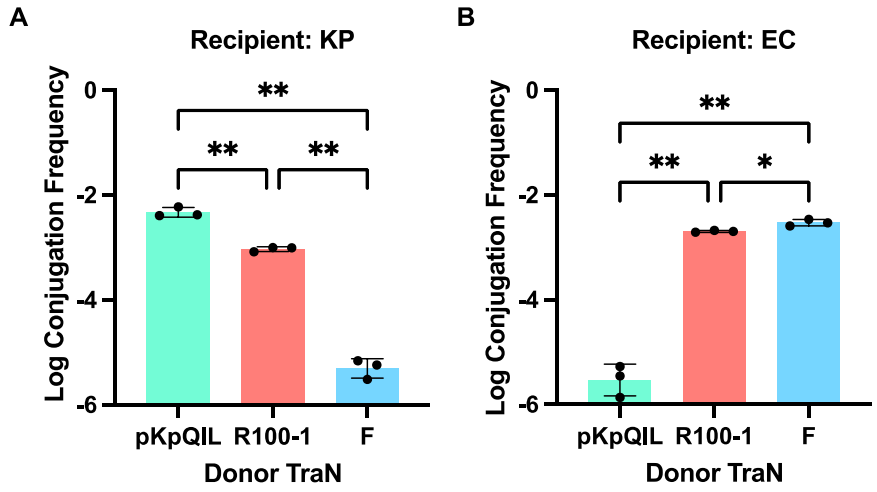


Figure 3.31. TraN mediates species specific transfer of pKpGFP

Log conjugation frequency of pKpGFP carrying *traN* from pKpQIL, R100-1 and F into WT (**A**) KP and (**B**) EC recipients was determined. Data analysed by RM one-way ANOVA with Tukey's multiple comparison test. * = $p < 0.05$, ** = $p < 0.01$. Error bars represent SD ($n = 3$).

To validate that $TraN_F$ displays specificity for OmpA expressed on recipient EC, the conjugation frequency of pKpGFP*traN*$_F$ was compared into WT and Δ*ompA* MG1655

recipients. No increase in fluorescence emission over time into Δ*ompA* recipients was observed using the RTCS and the AFU calculated at *t* = 300 min showed a negative fold change compared to the WT recipient (Figure 3.32A&B). Quantification of conjugation frequency by selection-based assays showed a significant reduction by approximately 2-log fold into the Δ*ompA* strain compared to the WT recipients confirming that $TraN_F$ mediates dependency on OmpA from EC during conjugation (Figure 3.32C).

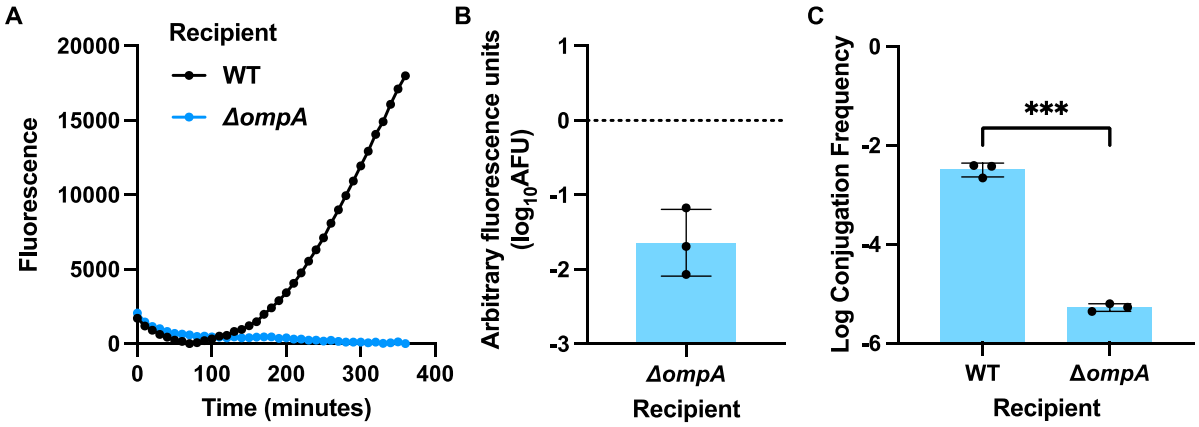


Figure 3.32. $TraN_F$ mediates dependency on OmpA from EC

A. Representative RTCS graph of fluorescence emission over time from mixtures containing GFP-DD*traN*$_F$ with WT or Δ*ompA* EC recipients. **B.** Arbitrary fluorescence units (AFU) at *t* = 300 min was determined for the Δ*ompA* recipient. **C.** Log conjugation frequency of pKpGFP*traN*$_F$ is into WT and Δ*ompA* recipients was determined. Data analysed by paired *t*-test. *** = $p < 0.001$. Error bars represent SD ($n = 3$).

3.22 Reconstituting loop sequences in OM protein homologues

Based on our findings, we hypothesised that differences in amino acid sequence between OM protein homologues in KP and EC account for the specificity observed during MPS. OmpK36 and its EC homologue, OmpC, share 79% amino acid sequence identity and a sequence alignment reveals domains in the proteins which are less well conserved (Figure 3.33A). Although L3 appeared to be important for $TraN_{pKpQIL}$ recognition based on the effects observed due to L3 insertions, it is relatively well conserved between the two porin homologues. In contrast, L4 of OmpC contains an additional four amino acids which was hypothesised to

interfere with $TraN_{pKpQIL}$ recognition. To test this, an OmpK36 isoform was generated to contain these additional amino acids in L4. However, the conjugation frequency of pKpGFP into recipients expressing this isoform of OmpK36 was not significantly different compared to recipients expressing $OmpK36_{WT}$ (Figure 3.33B).

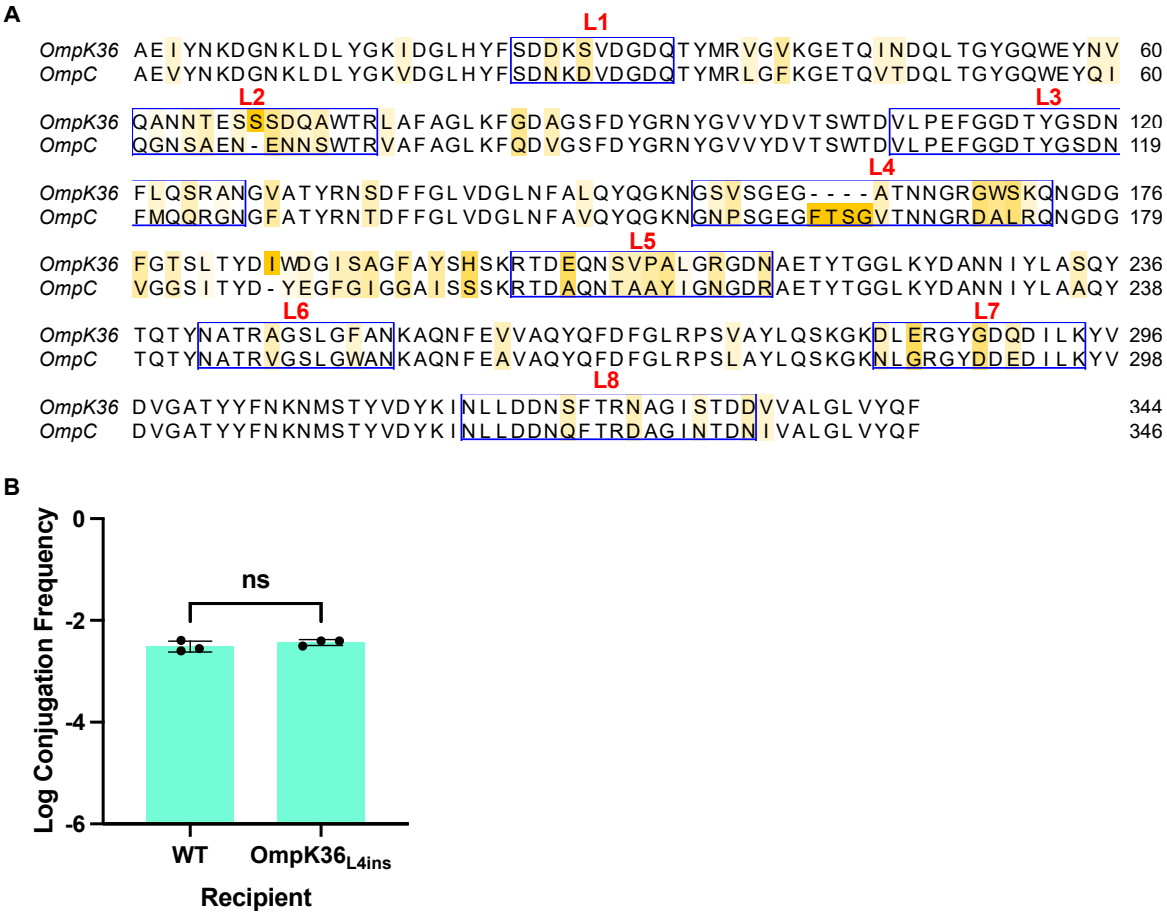


Figure 3.33. The L4 insertion in OmpC does not reduce $TraN_{pKpQIL}$ mediated conjugation

A. Amino acid sequence alignment of $OmpK36_{WT}$ and OmpC. Extracellular loops have been highlighted in boxes and darker shading indicates lower residue conservation. **B.** Log conjugation frequency of pKpGFP into WT and $OmpK36_{L4ins}$ KP recipients. Data analysed by paired *t*-test. ns = non-significant. Error bars indicate SD (n = 3).

Next, OmpA homologues found in KP ($OmpA_{KP}$) and EC ($OmpA_{EC}$) were compared. Previous work found that mutations in L4 of $OmpA_{EC}$ reduced F plasmid uptake in a TraN-specific manner, suggesting that this loop is important for $TraN_F$ recognition[173]. Thus, although $OmpA_{KP}$

contains additional amino acids in L1 and L3, differences that were specific to L4 were targeted. Of note was a glycine residue in $OmpA_{KP}$ that aligned to a histidine residue in $OmpA_{EC}$ (Figure 3.34A). Site-directed mutagenesis was used to generate an $OmpA_{KP}$ isoform expressing histidine in place of glycine ($OmpA_{G>H}$). However, no significant difference in pKpGFP*traN*$_F$ conjugation frequency was observed compared to a recipient expressing the WT $OmpA_{KP}$ isoform (Figure 3.34B).

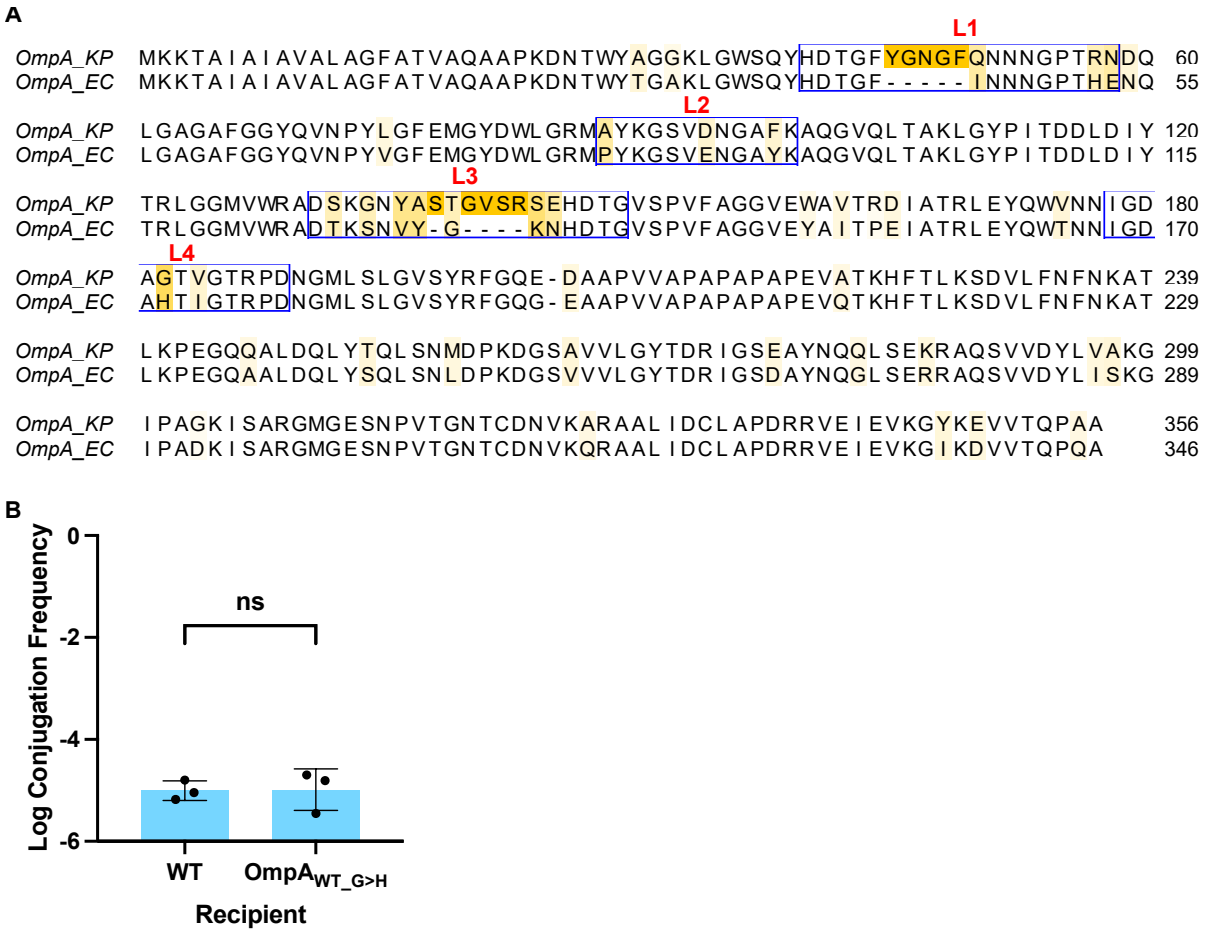


Figure 3.34. L4 G>H mutation does not increase TraN$_F$-mediated conjugation in KP

A. Amino acid sequence alignment of OmpA from KP and EC. Extracellular loops have been highlighted in boxes and darker shading indicates lower residue conservation. **B.** Log conjugation frequency of pKpGFP*traN*$_F$ into WT and $OmpA_{G>H}$ KP recipients. Data analysed by paired *t*-test. ns = non-significant. Error bars indicate SD (*n* = 3).

3.23 AlphaFold predictions

Due to the challenges in identifying regions of the OM proteins which are involved in TraN recognition, the focus was shifted towards understanding the structural basis of TraN specificity. In the absence of an experimentally determined 3D structure for TraN, AlphaFold was used to predict the structure of the three variants studied thus far[162]. A significant proportion of all three structures were predicted with a per-residue confidence score (pLDDT score) of above 90% which indicate that they are expected to be modelled to high accuracy (Figure 3.35A). As would be predicted from the amino acid sequence alignment of $TraN_{pKpQIL}$ and $TraN_{R100\text{-}1}$ shown in Figure 3.27 and the alignment published by Klimke and Frost for $TraN_F$ and $TraN_{R100\text{-}1}$[112], the N- and C-terminals of all three proteins appear structurally conserved. In contrast, the domains aligning to the variable region of TraN which mediates receptor specificity showed substantial structural diversity (Figure 3.34A; highlighted).

Closer analysis of these regions which are hereafter referred to as the tip domains of TraN show that all three homologues share a structurally conserved β-sandwich domain (Figure 3.34B). However, structural diversity is observed in the loops connecting the β-sheets including several defining features such as an extended β-hairpin and α-helix seen in the tips of $TraN_{pKpQIL}$ and $TraN_{R100\text{-}1}$ respectively which we propose accounts for differences in receptor specificity (Figure 3.35B).

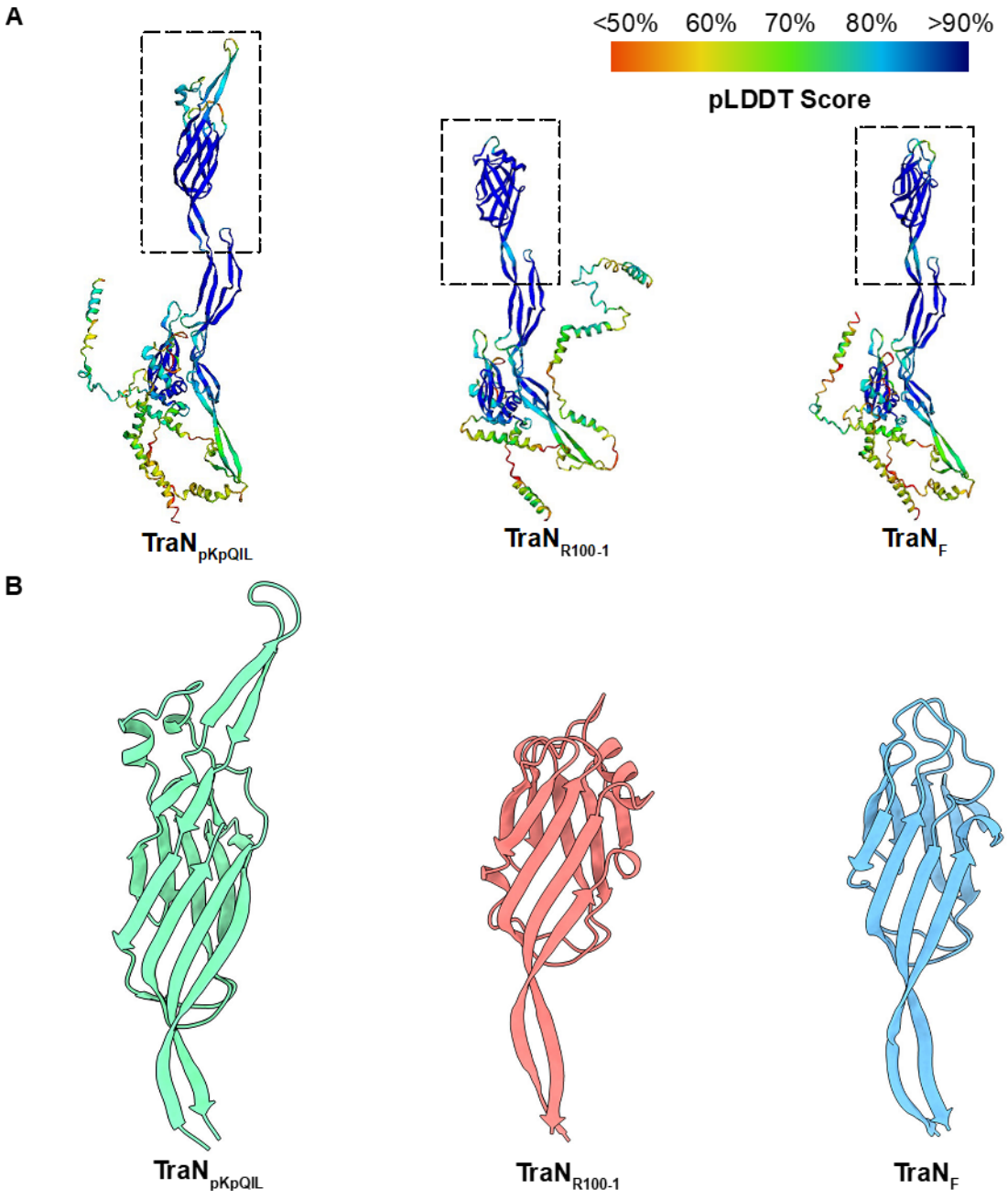


Figure 3.35. AlphaFold predictions reveal structural diversity in TraN 'tip' domains

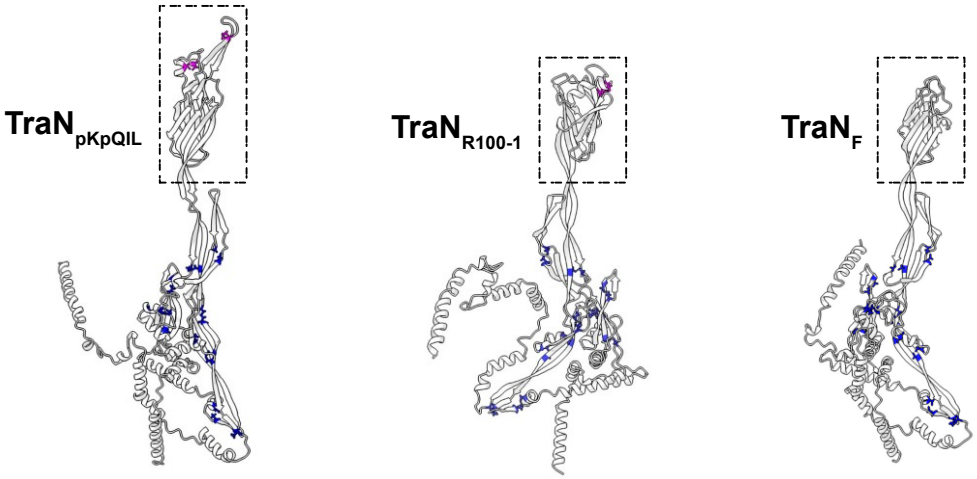
A. AlphaFold structural predictions for $TraN_{pKpQIL}$, $TraN_{R100-1}$ and $TraN_F$. Predictions are coloured according to the pLDDT score for each residue which indicates the confidence in the modelling accuracy. Tip domains are highlighted within the dashed lines. **B.** Zoomed in structure of the tip domains from the three TraN homologues.

3.24 Analysis of predicted TraN structure

From the predicted TraN structures, the distribution of the conserved cysteine residues could be analysed. The distance between spatially adjacent cysteine residues was measured in ChimeraX to confirm that all 22 conserved cysteines could be engaged in intramolecular disulphide bonds (Figure 3.36A). Additional cysteine residues predicted within the tip domains of $TraN_{pKpQIL}$ and $TraN_{R100\text{-}1}$ could also form disulphide bonds. The residues which form each of these bonds are shown in Figure 3.36B.

The predicted structure of $TraN_{pKpQIL}$ has two pairs of cysteine residues located within the tip domain one of which (C242 and C250) was hypothesized to stabilize the extended β-hairpin through a disulphide bond (Figure 3.36C). To test this, C242 in $TraN_{pKpQIL}$ was mutagenized into a serine residue (C242S). This mutation did not have a significant effect on the conjugation frequency of pKpGFP suggesting that this residue is not required to preserve the structure and function of the $TraN_{pKpQIL}$ tip domain (Figure 3.36D).

A



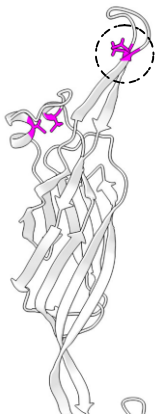
B

CYS #1	CYS #2
114	417
129	372
139	319
203	213
242	250
333	355
381	406
388	397
423	428
474	485
484	499
522	530
538	569

CYS #1	CYS #2
114	392
129	347
139	294
244	250
308	330
356	381
363	372
398	403
449	460
459	474
497	505
513	544

CYS #1	CYS #2
114	378
129	333
139	280
294	316
342	367
349	358
384	389
435	446
445	460
483	491
499	530

C



D

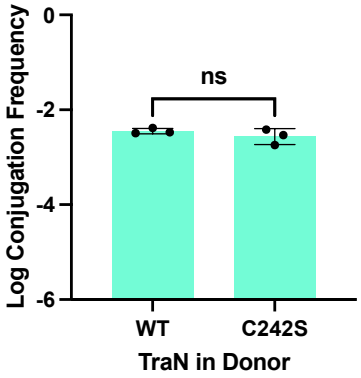


Figure 3.36. Analysis of cysteine residue in predicted TraN structures.

A. Cysteines in each predicted TraN structure are highlighted in blue with those present in the tip domains (dashed box) of $TraN_{pKpQIL}$ and $TraN_{R100-1}$ highlighted in magenta. **B.** Table for each TraN variant showing the residue number of cysteines engaged in disulphide bonds. **C.** Zoomed in structure of the $TraN_{pKpQIL}$ tip domain. C242 and C250 are highlighted in the dashed circle. **D.** Log conjugation frequency was calculated to compare the effect of the C242S mutation in TraN on pKpGFP transfer. Data analysed by paired *t*-test. ns = non-significant. Error bars indicate SD (n = 3).

3.25 $TraN_{pKpQIL}$ forms a complex with OmpK36

The structural differences in the TraN tip domains appear to provide the basis for specificity for structurally different recipient OM proteins and suggest that TraN interacts with each protein to facilitate MPS. Historically, attempts to show that $TraN_F$ interacts with $OmpA_{EC}$ via co-immunoprecipitation and yeast two hybrid approaches were unsuccessful[147]. Thus, $TraN_{pKpQIL}$ and OmpK36 were used to demonstrate complex formation during MPS. This was done in collaboration with members of Dr Konstantinos Beis' group. Both proteins were purified separately and mixed overnight prior to separation by size exclusion chromatography (SEC) by Chloe Seddon. A shift in the retention volume was observed with the mixture compared to the retention volumes associated with the individual proteins strongly suggesting the formation of a complex between $TraN_{pKpQIL}$ and $OmpK36_{WT}$ (Figure 3.37A and B). When the SEC was repeated with a mixture containing $TraN_{pKpQIL}$ mixed with $OmpK36_{WT+GD}$, no shift in the retention volume was observed (Figure 3.37C). This implies that the GD insertion interferes with the formation of a complex, aligning with the functional conjugation data where conjugation frequency of pKpGFP is significantly reduced due to this insertion.

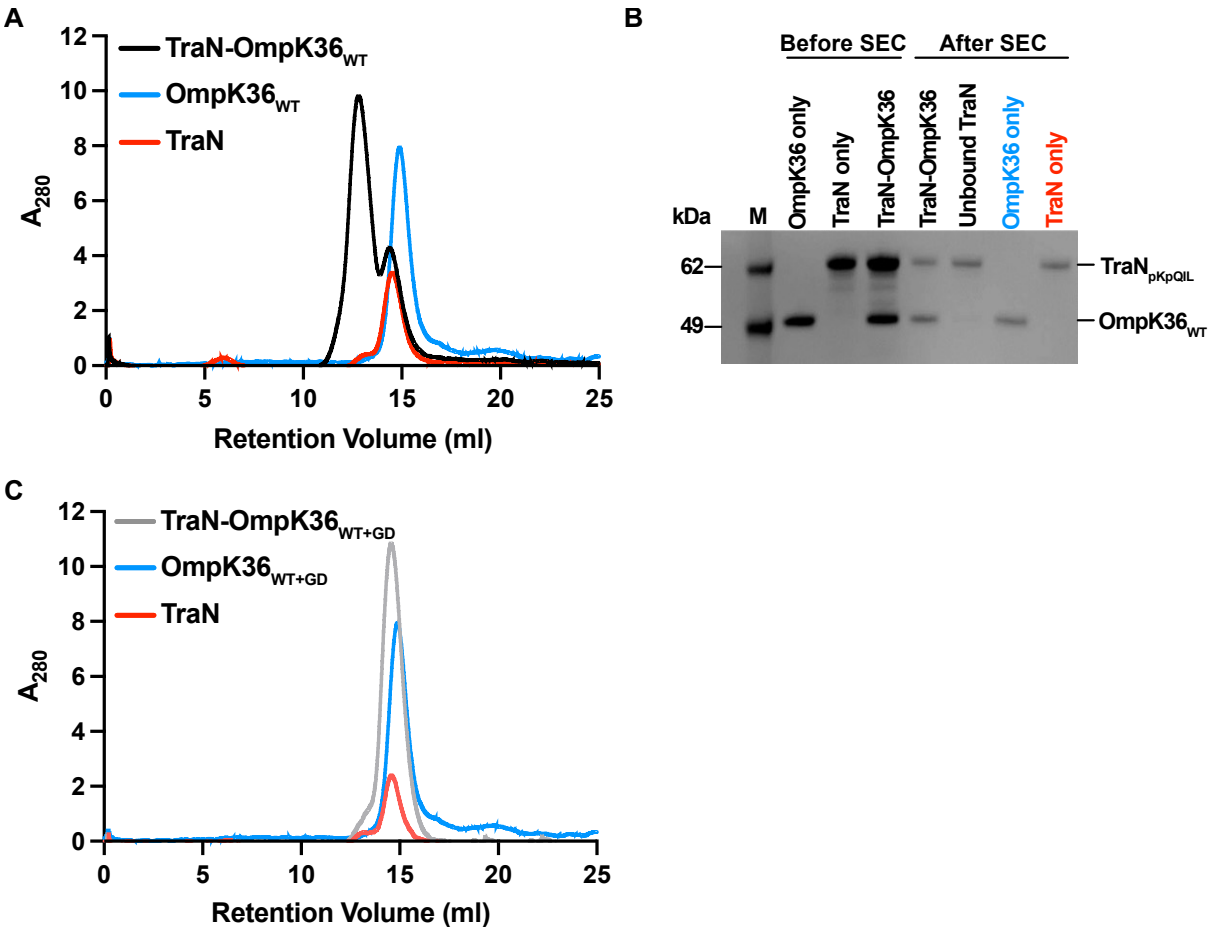


Figure 3.37. $TraN_{pKpQIL}$ forms a complex with $OmpK36_{WT}$

A. A mixture containing $TraN_{pKpQIL}$ and $OmpK36_{WT}$ was separated by size exclusion chromatography alongside $TraN_{pKpQIL}$ and $OmpK36_{WT}$ only controls. **B.** Complex formation was verified by SDS-PAGE. Lanes contain samples before and after SEC with label colours corresponding to the chromatogram peaks; Lane M, marker in kDa. **C.** Size exclusion chromatography was repeated for a mixture containing $TraN_{pKpQIL}$ and $OmpK36_{WT+GD}$ alongside the appropriate controls.

3.26 Cryo-electron microscopy of the TraN-OmpK36 complex

To understand how TraN and OmpK36 interact to form a complex, a collaboration was established with Professor Edward Egelman's group to analyse the purified TraN-OmpK36 complex via cryo-electron microscopy (cryo-EM). The data presented in this section were obtained and analysed by Leticia Beltran and Fengbin Wang.

Density was generated for both proteins resulting in a 3D reconstruction with an overall resolution of 2.6 Å. The crystal structure of OmpK36 (PDB ID: 6RD3) could be fitted inside the density with a root-mean-square deviation (rmsd) of 0.6 Å over 180 C_α atoms (Figure 3.38A). Density that was not assigned to OmpK36 was assigned to TraN and was present below the OmpK36 trimer and inside the channel of one of the porin monomers. The density below OmpK36 was present at a low threshold and featureless and, thus, this portion of TraN was not modelled. The density within the porin channel was more well defined and placing the AlphaFold model within this density revealed that it corresponded to the predicted β-hairpin of the $TraN_{pKpQIL}$ tip (Figure 3.38B).

Due to the poor density associated with TraN, additional validation of the complex was performed by generating an *ab initio* complex using the AlphaFold complex server. The prediction was superimposable with the experimentally obtained data with an rmsd of 0.45 Å over 480 C_α OmpK36 atoms and an rmsd of 1.2 Å for the 9 C_α β-hairpin atoms and supports the role of the extended β-hairpin in the tip of $TraN_{pKpQIL}$ in complex formation (Figure 3.38C).

To understand how the GD insertion interferes with complex formation, the TraN-$OmpK36_{WT}$ model was used as a reference for structural comparison. By superimposing the crystal structure for $OmpK36_{WT+GD}$ (PDB ID: 6RCK) onto the reference model, it was found that the GD insertion clashes with the S243 and G244 residues in the TraN β-hairpin (Figure 3.38D). This suggests the molecular basis for the lack of complex formation between TraN and this porin isoform.

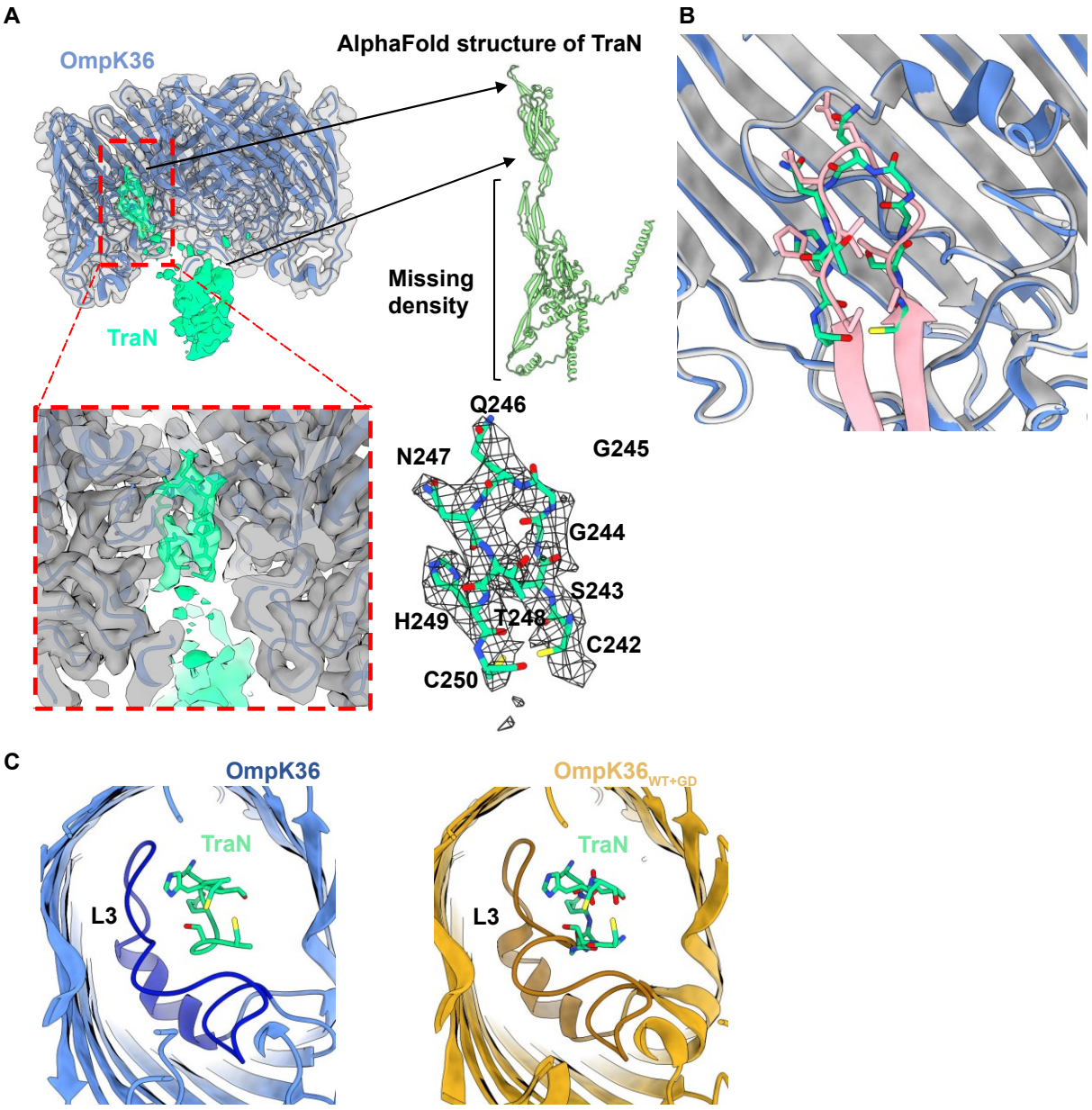


Figure 3.38. Cryo-EM structure of the TraN$_{pKpQIL}$-OmpK36 complex.

A. Coulomb potential density for the OmpK36 trimer (grey) and part of TraN (green) are shown. The front view of the reconstruction perpendicular to the OM has been omitted to reveal density for TraN within the porin channel. Atomic models for OmpK36 (blue; PDB ID: 6RD3) and the TraN tip were fitted into the reconstruction. Density is missing for the remainder of the predicted AlphaFold structure. Close-up view of the TraN β-hairpin model fitted inside the density (red box and bottom right panel). **B.** Superimposition of the *ab initio* complex predicted by AlphaFold (OmpK36 in grey/TraN in pink) and the experimentally determined complex (OmpK36 in blue/TraN in green). **C.** Top view showing the TraN β-hairpin within the channel of OmpK36$_{WT}$ (left panel). The TraN β-hairpin was modelled within the channel of OmpK36$_{WT+GD}$ (PDB ID: 6RCK; right panel).

3.27 Specific TraN-OMP pairings influence host species distribution of plasmids

Having established that TraN interacts with a recipient OM protein to facilitate MPS, the implications of TraN specificity observed *in vitro* on real world conjugative plasmid distribution were investigated. The bioinformatic analysis in this section was performed in collaboration with Dr Sophia David. Plasmids within a dataset that had been generated using the Plascad tool for plasmid classification were analysed[167]. In this dataset, fully sequenced plasmids that had been deposited in GenBank up to August 2018 were categorized according to their predicted mobilization properties based on the presence of several conjugation-associated genes. Predicted conjugative plasmids were further categorized based on protein homology into mating pair formation (MPF) groups.

A subset of plasmids which transfer via an MPFF conjugative system were shortlisted from this dataset. This was further curated to only include plasmids isolated from species within the *Enterobacteriaceae* family to account for the known permissive host range of IncF plasmids[178]. Lastly, the dataset was manually curated to exclude plasmids which had been identified to have a non-IncF replicon using PlasmidFinder[168]. The resulting dataset contained 824 plasmids isolated from host species representing at least 6 genera. Table 3.1. Analysis of plasmid hosts in curated IncF plasmid dataset.Table 3.1 lists the percentage of host isolates representing different genera and species. Where less than 10 isolates of a genus were present in the dataset, these were categorized as ‘Other’. The species was not listed if less than 10 isolates were present in the dataset.

Table 3.1. Analysis of plasmid hosts in curated IncF plasmid dataset.

Host	Percentage (%)
Escherichia	**40.8**
Escherichia coli	40.0
Klebsiella	**38.7**
Klebsiella pneumoniae	34.3
Klebsiella oxytoca	1.7
Klebsiella variicola	1.6
Salmonella	**11.0**
Salmonella enterica	10.8
Enterobacter	**4.2**
Enterobacter cloacae	3.0
Shigella	**1.7**
Citrobacter	**1.6**
Other	**1.9**

The plasmids in the dataset were analysed for *traN* sequences using tBLASTn and those expressing *traN* gene products which share at least 90% amino acid sequence similarity were grouped together[169]. It was found that 265 (32.2.%), 166 (20.1%) and 178 (21.6%) plasmids encoded TraN from pKpQIL, R100-1 and F respectively (Figure 3.39; Appendix 1-3). These accounted for 74% (*n* = 609) of the total plasmids. Importantly, when the host species from which these plasmids had been isolated were analysed, each variant was found to be associated with a small number of species including a single dominant species. Of note, 89.1% of plasmids encoding $TraN_{pKpQIL}$ were isolated from KP while 92.1% of plasmids encoding $TraN_{F}$ were found in EC (Figure 3.39). Although most plasmids encoding $TraN_{R100\text{-}1}$ were isolated in EC (72.9%), a significant proportion (16.9%) were also found in KP isolates. These results reflect the findings presented in Figure 3.30 which suggested that while $TraN_{R100\text{-}1}$ mediates efficient conjugation into both EC and KP, $TraN_{pKpQIL}$ and $TraN_{F}$ are associated with species-specific transfer into KP and EC respectively. This supports the hypothesis that species specificity during MPS affects conjugative plasmid host distribution.

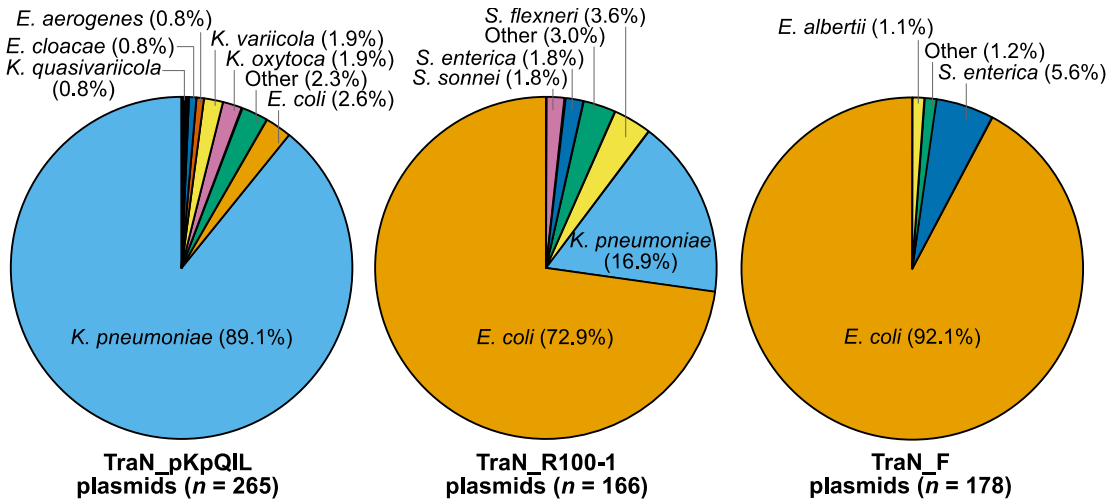


Figure 3.39. Host distribution for plasmids carrying similar TraN variants

Host species distribution associated with IncF plasmids encoding a *traN* gene product with ≥ 90% amino acid sequence similarity to TraN from pKpQIL (Accession ID: KY798507.1), R100-1 (Accession ID: DQ364638.1) and F (Accession ID: NC_002483.1) was determined. Where less than two isolates from a species was found, it was categorised as 'Other'.

3.28 Identification of additional TraN sequence variants

The remaining 215 plasmids in the dataset were analysed for annotated *traN* sequences and the same 90% sequence identity threshold was used to group plasmids together. From this analysis 16 other sequence variants were identified of which 4 were associated with at least 10 plasmids (Figure 3.40; Appendix 4-7). Further analysis was restricted to these four variants, of which one was found to align to TraN from the *S.* Typhimurium (ST) virulence plasmid, pSLT (Figure 3.40A). As the three remaining sequence variants were not associated with well-known plasmids, these were termed Minor Variants 1 to 3 (MV1-3). Of the plasmids encoding these 4 variants, those encoding $TraN_{pSLT}$ were isolated exclusively from *S. enterica* with the majority found in isolates from the Typhimurium serovar while plasmids encoding $TraN_{MV3}$ were isolated almost exclusively from EC (Figure 3.40B). In contrast, plasmids encoding the remaining two TraN variants displayed a much more varied host distribution with no clear dominant host species.

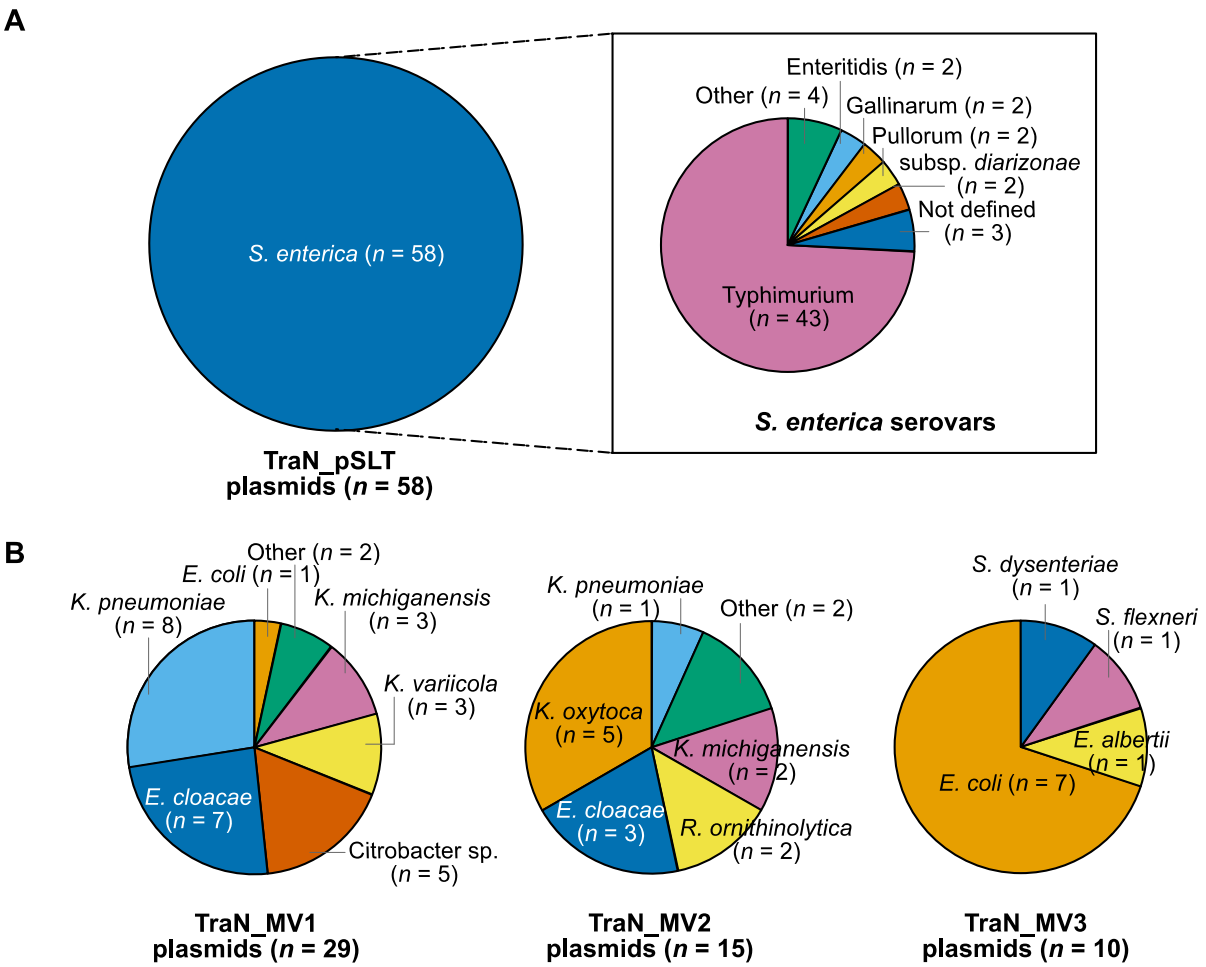


Figure 3.40. Additional TraN sequence variants identified from plasmid dataset

A. Host species distribution was determined for plasmids encoding TraN from pSLT (Accession ID: AE006471.2). The distribution of these plasmids within different serovars of *S. enterica* is shown in the inset. **B.** Host species distribution was determined for plasmids encoding minor variant TraNs. Plasmids were grouped based on reference TraN sequences from the following plasmids: MV1 (Accession ID: NZ_CP016763.1), MV2 (Accession ID: AP014954.1) and MV3 (Accession ID: NZ_CP023348.1). Where less than two isolates from a species was found, it was categorised as 'Other'.

3.29 Structural homology determines receptor specificity

AlphaFold was used to predict the structures of the four additional *traN* sequence variants (Figure 3.41A). As these variants share less than 90% amino acid sequence identity with the homologues from pKpQIL, R100-1 and F, it was hypothesised that the tips would be

structurally different. Instead, $TraN_{pSLT}$ was found to share a structurally similar tip with $TraN_{R100-1}$ and the tip of $TraN_{MV2}$ is structurally homologous to $TraN_{pKpQIL}$ (Figure 3.41B). The tip domains of $TraN_{MV1}$ and $TraN_{MV3}$ were also found to share a similar structure that was dissimilar from all previously analysed variants.

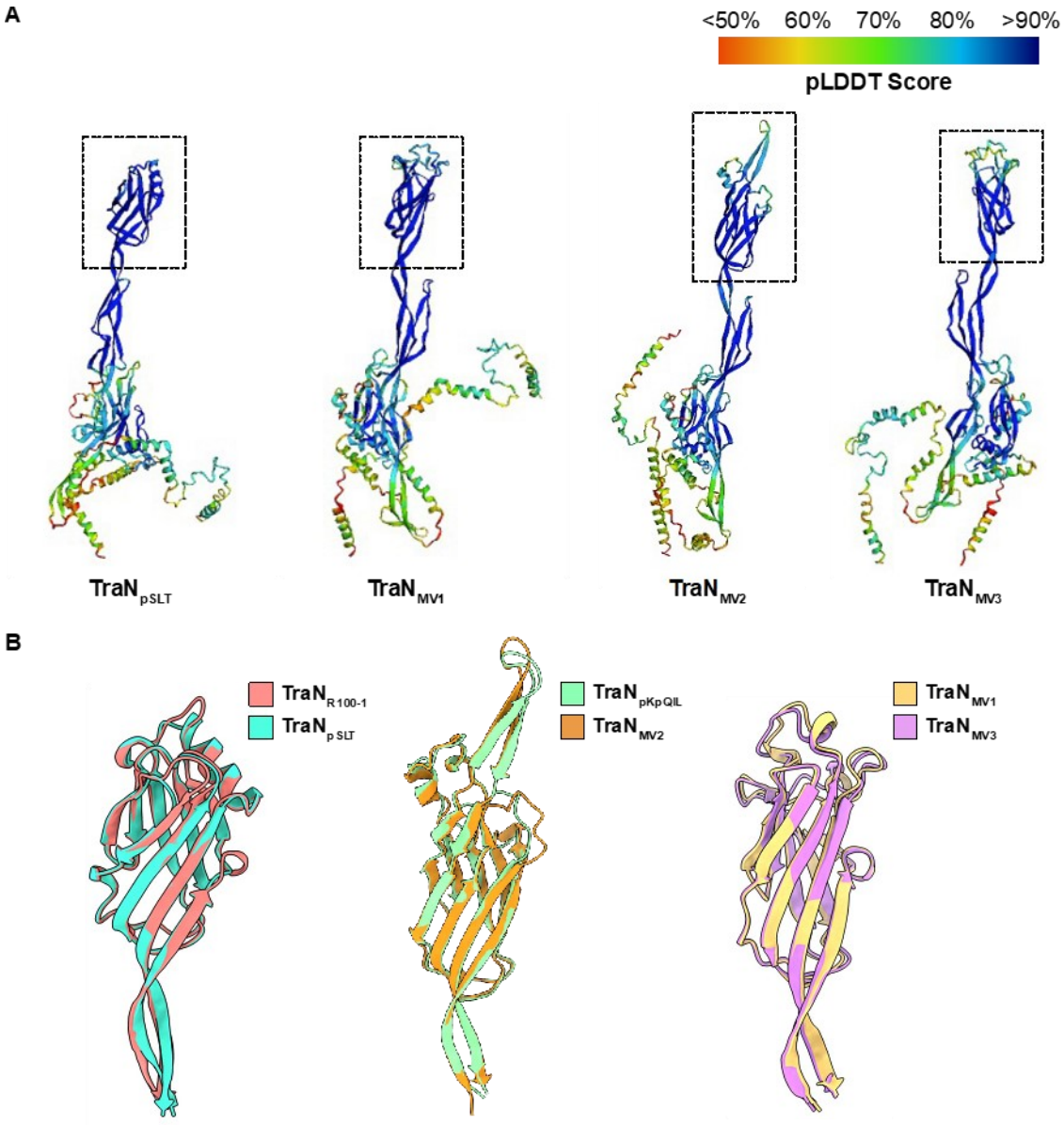


Figure 3.41. Structural homology is observed amongst TraN variants

Superimpositions of TraN tip domains predicted using AlphaFold from (**A**) R100-1 and pSLT, (**B**) pKpQIL and MV2, and (**C**) MV1 and MV3.

TraN homologues that share structurally homologous tip domains were hypothesised to mediate MPS via the same recipient OM protein. To test this, a chimeric TraN ($TraN_{Ch2}$) which contains the variable region of $TraN_{pSLT}$ was generated. Based on the hypothesis, $TraN_{Ch2}$

should mediate dependency on recipient OmpW for MPS as it shares a similar tip structure to $TraN_{R100\text{-}1}$. Using the RTCS, no increase in fluorescence over time was observed from donors expressing $TraN_{Ch2}$ into Δ*ompW* recipients compared to WT recipients (Figure 3.42A). In addition, the AFU calculated for Δ*ompW* recipients was also negative supporting the hypothesis (Figure 3.42B).

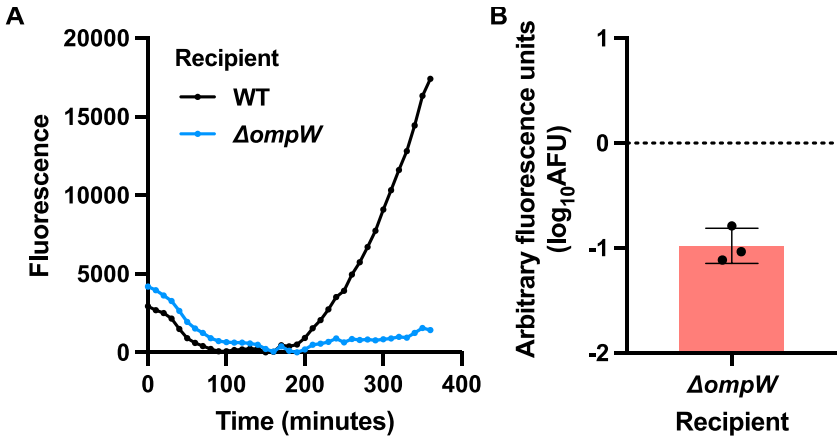


Figure 3.42. $TraN_{pSLT}$ mediates dependency on recipient OmpW

A. Representative RTCS graph showing fluorescence emission over time from conjugation mixtures containing GFP-DD*traN*$_{Ch2}$ donors and WT and Δ*ompW* KP recipients. **B.** AFU data calculated at *t* = 300 min for the Δ*ompW* recipient. Error bars represent SD (*n* = 3).

3.30 Identifying the receptor for $TraN_{MV1}$

As $TraN_{MV1}$ and $TraN_{MV3}$ are structurally dissimilar from the previously studied TraN variants, these variants were hypothesised to interact with a different recipient OM protein. To investigate this, an additional chimeric TraN ($TraN_{Ch3}$) containing the variable region from $TraN_{MV1}$ was generated. From initial experiments, pKpGFP encoding $TraN_{Ch3}$ facilitated efficient transfer into EC but not KP recipients (data not shown). Thus, GFP-DD*traN*$_{Ch3}$ was used in the RTCS to assess the role of different EC OM proteins during conjugation. This panel of recipients included mutants which do not express the 3 major EC OM proteins – OmpF, OmpC and OmpA. A mutant in which both OmpF and OmpC was also included to investigate if the loss of expression of both porins would have an additive effect on conjugation as in the case of $TraN_{pKpQIL}$.

The AFU data calculated for the Δ*ompF* recipient was lower than the AFU calculated for the other single mutants tested and was similar to the data obtained for the double deletion mutant (Figure 3.43A). This suggests that $TraN_{Ch3}$ mediates dependency on OmpF and that, unlike what was observed with $TraN_{pKpQIL}$, there is no additive effect resulting from the deletion of both major OM porins. Moreover, compared to the WT recipient, the increase in fluorescence emission over time was not as great from the Δ*ompF* recipients (Figure 3.43B).

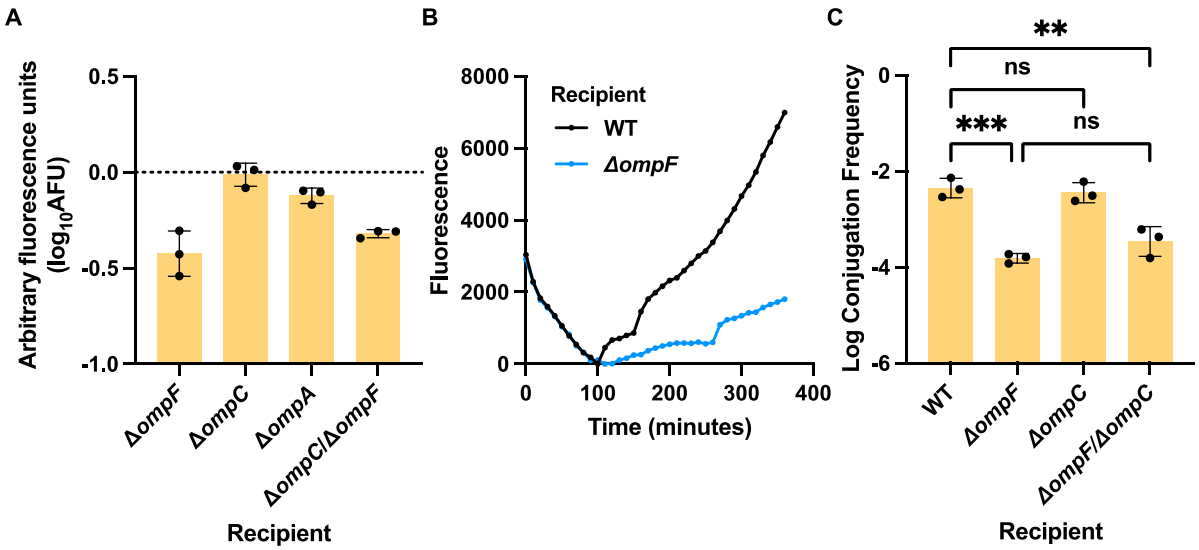


Figure 3.43. OmpF is a receptor for $TraN_{MV1}$

A. AFU data calculated at *t* = 300 min for a panel of isogenic EC recipients. **B.** Representative graph of fluorescence emission over time from conjugation mixtures containing GFP-DD*traN*$_{Ch3}$ with WT and Δ*ompF* EC recipients. **C.** Log conjugation frequency was calculated into WT, Δ*ompF*, Δ*ompC* and Δ*ompF*/Δ*ompC* recipients. Data was analysed by RM one-way ANOVA with Tukey's multiple comparison test. ** = $p < 0.01$, *** = $p < 0.001$, ns = non-significant. Error bars represent SD ($n = 3$).

The role of OmpF was validated using selection-based assays. A significant reduction in log conjugation frequency was observed into the Δ*ompF* recipient but not Δ*ompC* recipient and there was no significant difference between the Δ*ompF* recipient and the double deletion mutant (Figure 3.43C). However, the reduction in conjugation frequency was less than the 2-log fold difference that was observed with other TraN variants which was indicative of the lack

of MPS. Taken together, these results suggest that OmpF is the major receptor for $TraN_{MV1}$. However, as the difference in conjugation frequency was less than 2-log fold, this may indicate the presence of an additional receptor on the surface of EC recipients for this TraN variant.

3.31 Reclassification of TraN homologues according to structural similarity

From the above experiments, structural homology in the TraN tip domain appears to be a key determinant of receptor specificity during conjugation. Thus, the 7 TraN sequence variants were reclassified based on structural homology into the following groups: TraNα ($TraN_{R100\text{-}1}$ and $TraN_{pSLT}$), TraNβ ($TraN_{pKpQIL}$ and $TraN_{MV2}$), TraNγ ($TraN_F$) and TraNδ ($TraN_{MV1}$ and $TraN_{MV3}$). Plasmid host species distribution was revised to reflect this reclassification (Figure 3.44A). Phylogenetic analysis of the 7 TraN sequences revealed clustering that aligns with the classification system (Figure 3.44B).

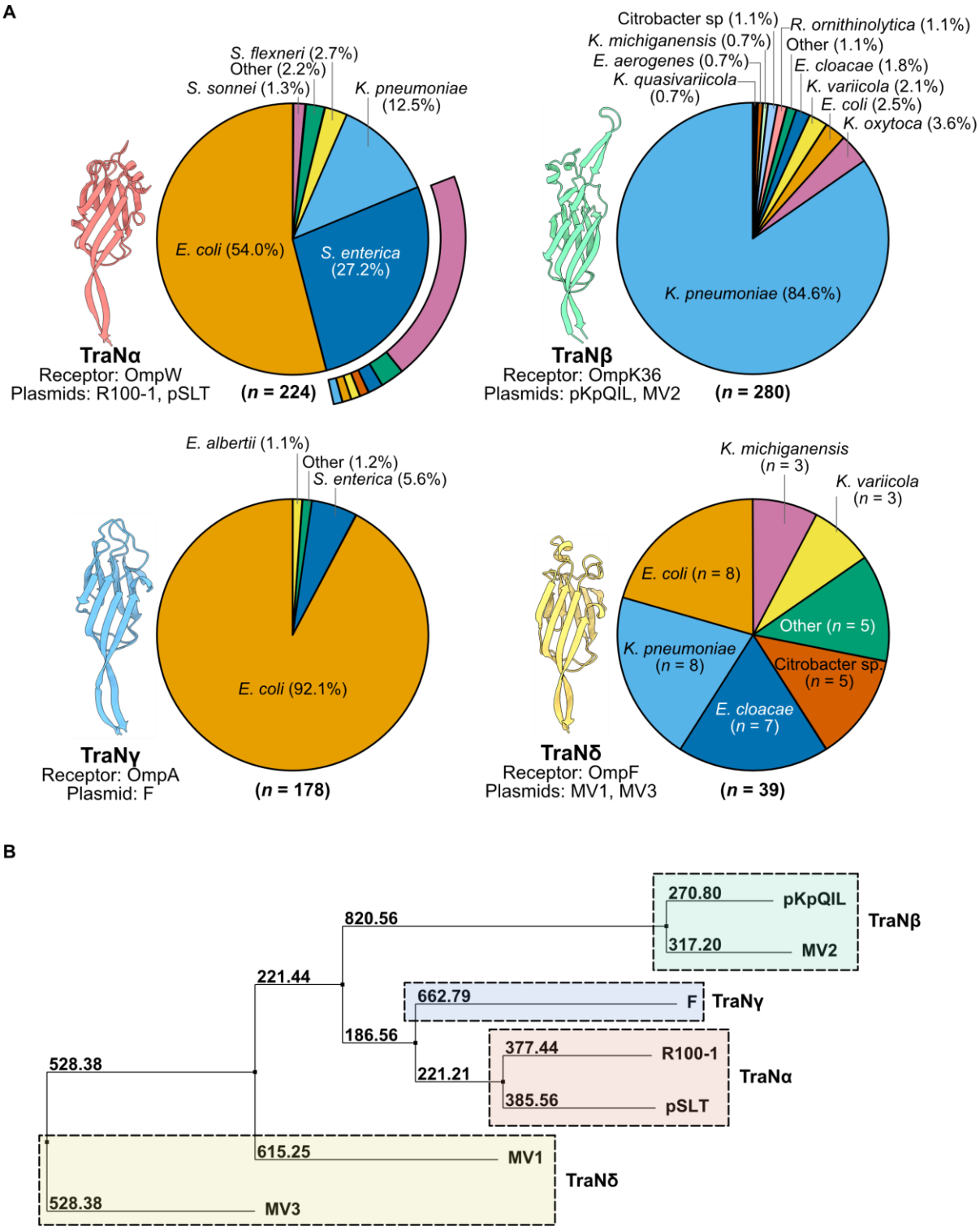


Figure 3.44. Reclassification of TraN variants according to structural homology

A. Plasmid host species distribution for the 4 structural variants of TraN. Where there were less than two isolates from a species, it was categorised as 'Other'. The serovar breakdown for *S. enterica* isolates expressing TraNα is shown in the outer ring with the colours corresponding to the inset in Figure 3.40. **B.** Phylogenetic analysis of TraN sequence variants from the 7 reference plasmids. Dashed boxes indicate variants in the same structural groups.

3.32 Assessing species specificity with other bacterial recipients

As a substantial proportion of plasmids encoding TraNα were isolated from ST, it was hypothesised that variants in this group would facilitate efficient conjugation into recipients of this species. In contrast, plasmids encoding TraNδ were isolated from a diverse range of host species including EC, KP and *Enterobacter cloacae* (ECl) which suggest that homologues in this group mediate MPS with a broad range of recipient species. The RTCS was used to analyse transfer of pKpGFP-D as mediated by different TraN homologues into WT recipients from 4 different species – KP, EC, ST and ECl.

Fluorescence measurements were performed as before over 6 h. While increases in fluorescence were observed with several TraN variants into KP, EC and ECl recipients, fluorescence emission did not appear to increase for all mixtures containing ST recipients (Figure 3.45). Excluding ST, TraN-mediated species specificity could be observed in the other recipient species. Only $TraN_{pKpQIL}$, $TraN_{R100\text{-}1}$ and $TraN_{pSLT}$ facilitated efficient transfer into KP while all TraN variants except $TraN_{pKpQIL}$ facilitated efficient conjugation into EC. In the case of ECl, all TraN variants except $TraN_F$ mediated efficient plasmid transfer.

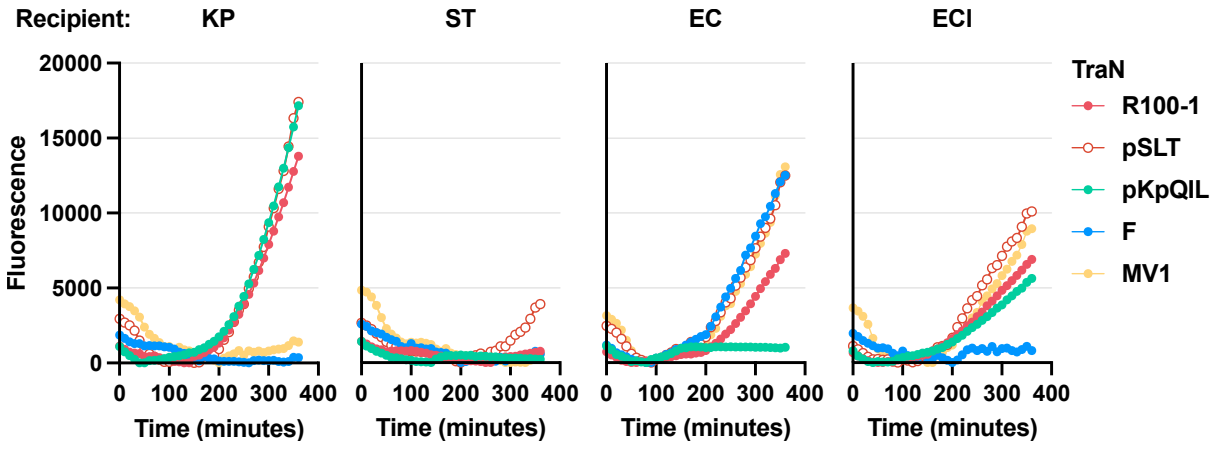


Figure 3.45. Using the RTCS to assess TraN-mediated conjugation species specificity

Representative RTCS graphs showing fluorescence emission over time from conjugation mixtures containing GFP-DD donors expressing different TraN homologues into WT KP, ST, EC and ECl recipients.

To account for the lag in fluorescence measurements observed with ST recipients, endpoint fluorescence readings were taken for all mixtures following 24 h of incubation. Measurements were also obtained from negative control mixtures containing ICC8001 donors carrying an untagged, derepressed pKpQIL, and the respective recipient species (-GFP). Mixtures from which there is significantly higher fluorescence emission compared to the negative control are taken as indicative of MPS-mediated efficient conjugation. The results obtained for KP, EC and ECl recipients were consistent with the results obtained after 6 h (Figure 3.45). However, after 24 h, mixtures containing ST recipients with donors expressing $TraN_{R100\text{-}1}$, $TraN_{F}$ and $TraN_{pSLT}$ were now found to have statistically higher fluorescence emission than the negative control and in terms of raw emission measurements showed higher readings than mixtures containing EC and ECl recipients.

$TraN_{R100\text{-}1}$ and $TraN_{pSLT}$ showed similar results in terms of species specificity, supporting the hypothesis that structural conservation in the tip domains is associated with functional conservation during conjugation (Figure 3.46A). The remaining TraN variants were seen to mediates species specific transfer: $TraN_{pKpQIL}$ mediated MPS with KP and ECl, $TraN_{F}$ with ST and EC and $TraN_{MV1}$ with EC and ECl. Overall, the findings appear to align well with real-world plasmid host distribution, emphasizing the role of TraN-mediated MPS in influencing conjugative plasmid range. An exception to this finding would be the significant proportion of plasmids isolated from KP which encode $TraN_{MV1}$ despite the *in vitro* data suggesting that this variant does not facilitate MPS with this species. This discrepancy may be a result of the smaller sample size of plasmids which encode TraNδ. Phylogenetic analysis of OM proteins from each of the four species revealed clustering of homologues which interact with TraN based on the functional assays (Figure 3.46B). This, however, does not apply to homologues of OmpK36/OmpC where the clustering does not align with the species specificity observed during pKpGFP conjugation.

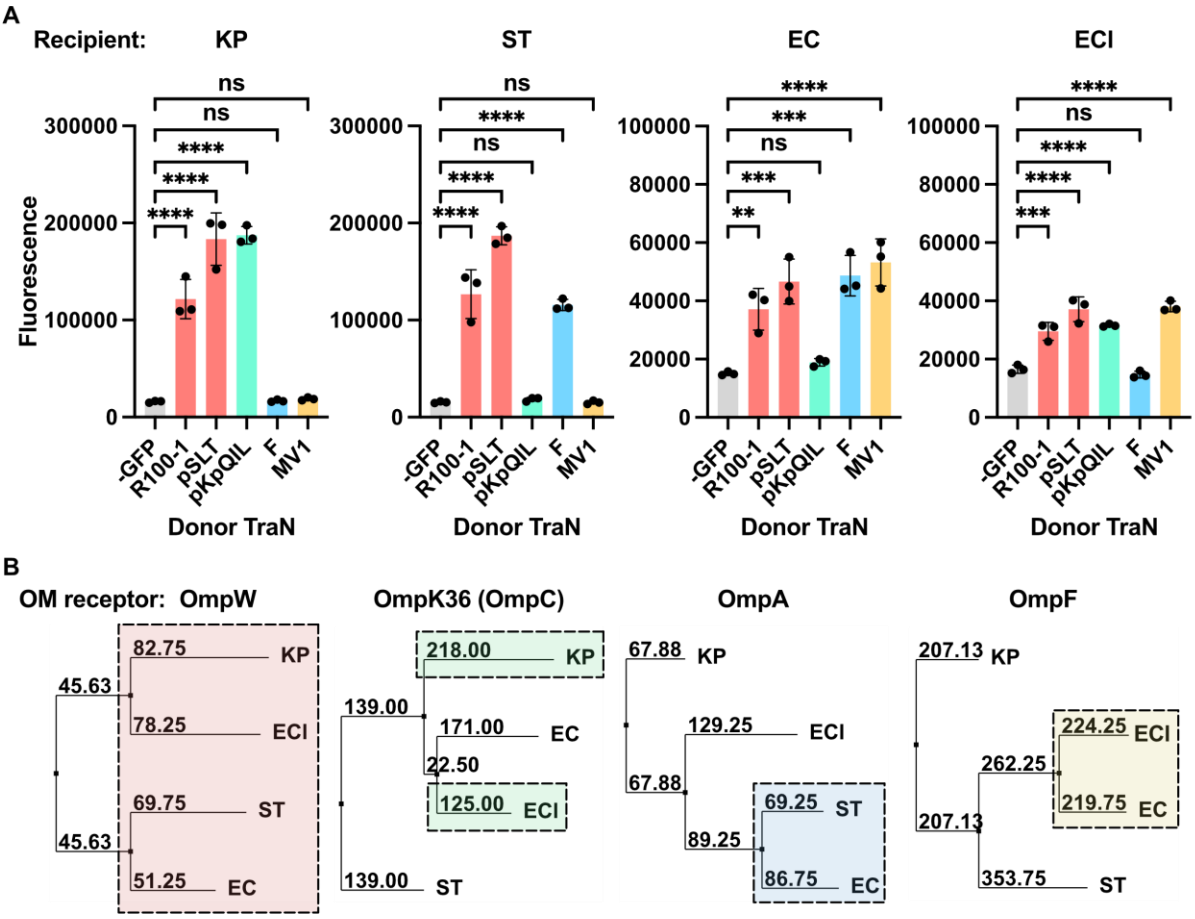


Figure 3.46. Different TraN variants mediate species specificity during conjugation

End-point fluorescence measurements were obtained from conjugation mixtures containing GFP-DD donors expressing different TraN homologues into WT KP, ST, EC and ECI recipients. A negative control (-GFP) was included for statistical analysis. Data was analysed by RM one-way ANOVA with Dunnett's multiple comparison test comparing to the negative control. ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, ns = non-significant. Error bars represent SD ($n = 3$). **B.** Phylogenetic trees of OM protein homologues from the four recipient species. Dashed boxes are used to highlight the homologues which facilitate MPS as determined from the *in vitro* data.

4. Discussion

This work describes the role of MPS during conjugation and the influence it has on the host distribution of conjugative IncF plasmids. MPS is revealed to be mediated by interactions between the plasmid-encoded donor subunit TraN and an OM protein on recipient cells. Different TraN variants recognize and form complexes with distinct receptors depending on the structure of the TraN tip domain. The variants identified in this study were classified based on structural similarity into four groups denoted TraN α-δ which interact with OmpW, OmpK36, OmpA and OmpF on recipient cells respectively (Figure 4.1). OmpK35 of KP can also serve as a receptor for TraNβ. Importantly, TraN interactions are specific to certain homologues of each receptor resulting in a species-specific effect on conjugation that is reflected in the real-world host distribution of TraN-encoding plasmids.

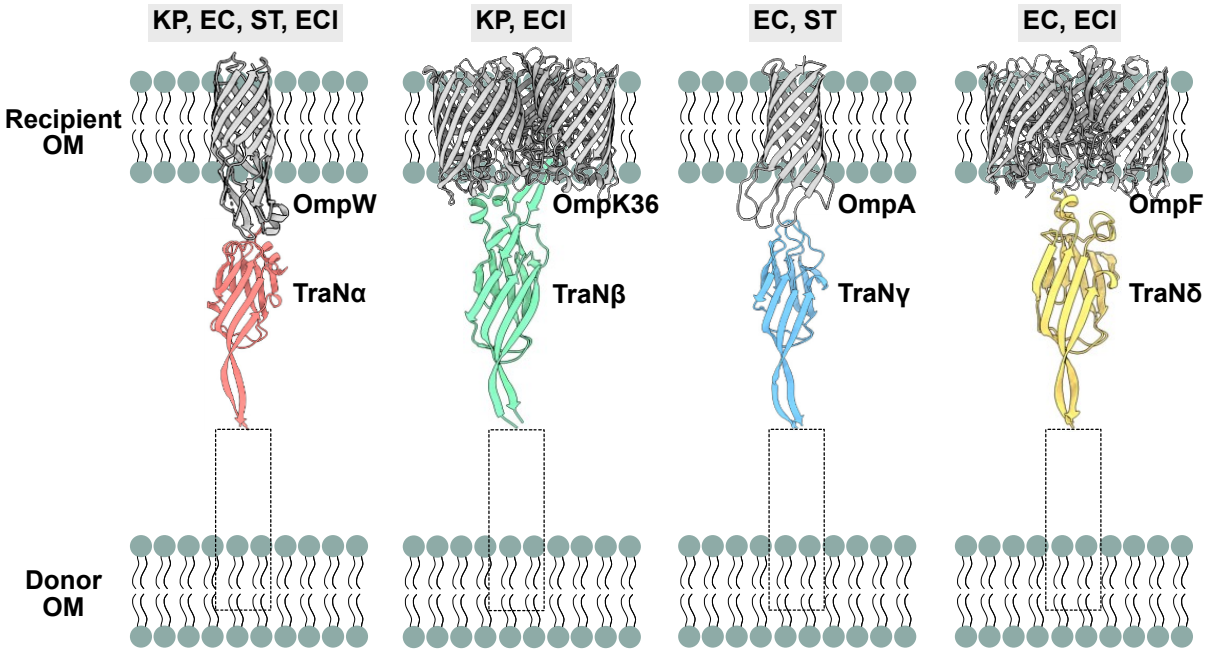


Figure 4.1. Structural variants of TraN and their recipient OM protein receptors.

The tip domains of the 4 TraN structural variants described in this study is shown with their corresponding receptors in the recipient OM. The rest of TraN is represented by the dashed box as information on how the protein is associated with the donor OM is not available. Species which express OM protein homologues that can interact with each TraN variant are listed above the structure of each receptor. KP = *K. pneumoniae*; EC = *E. coli*; ST = *S.* Typhimurium; ECl = *E. cloacae*.

4.1 The role of recipient cells during bacterial conjugation

Recipient cell involvement during bacterial conjugation has been relatively poorly understood in comparison to plasmid-bearing donor cells. Recently, studies investigating recipient involvement during conjugation of the MPFT plasmids R388 and RP4 conjugation found that no non-essential proteins affected transfer of these plasmids[100,157]. The combination of these findings and the previously undetermined nature of OmpA involvement during F plasmid conjugation led to the notion that recipient cells are passive participants in an otherwise highly complex and active process.

In this work, the role of recipient cells during conjugation was investigated using the KP carbapenemase-encoding plasmid pKpQIL. It was hypothesised that ST258-associated mutations in the major OM porins OmpK35 and OmpK36 would increase conjugative uptake of pKpQIL due to the close association of this plasmid with strains from this sequence type. However, these mutations instead reduced uptake of the pKpQIL reporter plasmid, pKpGFP by 2-3 log folds. OmpK36 appears to play a more important role in conjugation as a loss of expression of only OmpK35 does not significantly affect transfer, but mutations affecting only OmpK36 do. Furthermore, conjugation frequency was affected by changes in OmpK35 expression but not a reduction in OmpK36 expression. This implies that substantially less OmpK36 is needed to facilitate efficient conjugation compared to OmpK35. Reduced expression of OmpA had also been linked to lower conjugation frequency of the F plasmid[152]. However, without a system which allows for precise modulation and quantification of porin expression, it was not possible to determine the minimum expression level of each porin required to facilitate efficient conjugation of pKpGFP.

The role of other cell surface components during pKpGFP conjugation was also investigated. Capsule, specifically, has been hypothesised to act as a barrier to conjugative plasmid uptake and this was previously demonstrated as deletion of *wcaJ* in KP was accompanied by increased conjugation frequency of a plasmid mobilized *in trans* by the IncP plasmid RK2

which encodes an MPFT conjugative system[179,180]. This finding may explain why hypercapsulated, hypervirulent KP strains were initially less commonly associated with conjugative resistance plasmids[181]. However, in this work, there was no significant difference in pKpGFP conjugation when comparing the WT recipient with the same acapsular mutant. This suggests that the role of capsule during conjugation cannot be generalized to all conjugative systems.

After showing that several other cell surface components did not affect pKpGFP conjugation, the mechanism underlying OmpK36 dependency was investigated. The L3 GD insertion in $OmpK36_{ST258}$ was determined to be the mutation which impairs plasmid uptake. The location of this loop within the porin channel contrasts with previous work which described mutations in the surface exposed L4 of OmpA that reduced F plasmid uptake[173]. Thus, it was initially hypothesised that the role of OmpK36 during pKpGFP transfer may be indirect through pleiotropic effects resulting from the constriction of the porin channel. Indeed, a slight trend in conjugation frequency that was inversely related to the extent of pore constriction was observed as single amino acid insertions in L3 had an intermediate effect on pKpGFP uptake. However, impaired nutrient uptake did not appear to account for the differences in conjugation frequency. As changes in the OM proteome resulting from this insertion were not assessed, pleiotropic effects on the expression of other OM proteins could not be confirmed.

4.2 MPS mediates species specificity during conjugation

Although the role of OmpK36 was still unclear, the identification of the pKpGFP-encoded TraN as the donor component mediating this dependency suggests that the porins play a similar role in pKpQIL conjugation as OmpA during F plasmid conjugation[112]. TraN mediates MPS during conjugation and, by extension, this would also implicate OmpK36 as being involved during this stage of MPF. However, while both OmpA and OmpK36 are β-barrel proteins which are major constituents of the OM, they are structurally and functionally dissimilar. OmpA is expressed as a monomer while OmpK36 is trimeric. Furthermore, although the role of

OmpK36 as a non-specific diffusion channel is well established, the pore of OmpA appears to fluctuate between open and closed states and its role as an OM channel in enterobacterial species appears to be minor[182,183]. The receptor for $TraN_{R100-1}$ was later identified as the monomeric OM protein, OmpW. Analysis of the crystal structure for OmpW from EC revealed that it forms a long, narrow hydrophobic channel which can serve as an ion channel across the OM[177]. Unlike OmpA and OmpK36, OmpW is expressed at very low levels under standard laboratory growth conditions, potentially explaining why it could not be identified as the recipient factor for R100-1 from Con^- mutants that were previously isolated[153,184].

The identification of three structurally and functionally different OM proteins that cooperate with different TraN variants strongly suggested that these play a direct role in conjugation by serving as interaction partners for TraN. A seminal finding of this work supported this by showing that $TraN_{pKpQIL}$ and OmpK36 interact to form a complex. The structure of this complex rationalises the importance of L3 for efficient conjugation as the extended β-hairpin in the $TraN_{pKpQIL}$ tip inserts midway into the channel of an OmpK36 monomer and is stabilized by this loop. Mutations such as the GD insertion interfere with complex formation as they result in conformational changes to L3 which appear to result in clashes with residues in the $TraN_{pKpQIL}$ β-hairpin. Modelling the effect of single amino acid insertions on complex formation could further determine how these mutations result in the intermediate phenotype observed.

The structural basis of receptor specificity was determined by predicting the structures of TraN from pKpQIL, R100-1 and F using AlphaFold[162]. The tip domains of all 3 variants were highly dissimilar and experiments with chimeric TraNs established that these domains mediate specificity for different OM proteins. The ability to isolate the $TraN_{pKpQIL}$-OmpK36 complex where previous attempts to isolate $TraN_F$-OmpA failed could be a virtue of the unique way in which the proteins in the former complex interact. Based on their predicted structures, both $TraN_{R100-1}$ and $TraN_F$ likely interact with their respective receptors via surface exposed loops which may make the isolation of individual complexes more challenging. Alternative strategies

besides SEC may be required to validate the ability of these two TraN variants to interact with their identified receptors. A preliminary experiment has demonstrated that OmpW can be captured via pull-down assays using immobilized $TraN_{R100\text{-}1}$ suggesting that this technique may be an appropriate strategy to purify other MPS complexes for structural analysis.

4.3 The driving force of conjugation species specificity

In the process of investigating OmpA dependency during F plasmid conjugation, TraN was determined to not only mediate receptor specificity but also species specificity during conjugation. Comparing homologues of each TraN receptor suggests that specificity results from minor differences in the amino acid sequences of these proteins, presumably within regions of the protein that form the interaction interface. Previous experiments have shown that a single amino acid substitution in L4 of OmpA is sufficient to interrupt MPS[173]. Similarly, in this work, single or diamino acid insertions in L3 of OmpK36 were found to interfere with $TraN_{pKpQIL}$ binding. While these residues and loops appear to be important in the context of those homologues (EC OmpA and KP OmpK36 respectively), based on sequence alignments with other receptor homologues they are unlikely to be the only determinants of successful TraN binding. In fact, L3 of OmpK35 shares very little sequence similarity with L3 found in various OmpK36 homologues yet $TraN_{pKpQIL}$ can mediate MPS via OmpK35 but not OmpC from EC or ST. Moreover, targeted attempts to promote or inhibit MPS by reconstituting specific residues in other receptor homologues were unsuccessful in this study. This may suggest that different combinations of surface exposed residues can form appropriate binding interfaces for TraN.

As major components of the OM, the surface exposed loops of OM proteins are likely to undergo substantial selective pressure from various sources, resulting in differences that drive species specificity during MPS. All the receptors identified in this work have also been demonstrated to act as receptors in other contexts. OmpW serves as a receptor for colicin S4 while OmpA, OmpC and OmpF have been shown to serve as receptors for various

bacteriophages[185]. OmpC and OmpF are also utilized for the internalization of several colicins and play a role in the influx of various antimicrobial compounds[186,187]. In addition, all 4 receptors have been demonstrated to be highly immunogenic and have been proposed as vaccine candidates in various species[188–190]. Nevertheless, OmpW is relatively well conserved and consequently homologues from all 4 species tested were able to mediate MPS with TraNα. This may indicate that there is less selection pressure exerted on the surface exposed loops of OmpW potentially because, unlike the other TraN receptors which are generally highly expressed, expression of OmpW fluctuates greatly depending on environmental cues[191].

4.4 The role of TraN during conjugation

The absence of a receptor for TraN on recipient cells greatly reduces conjugative plasmid uptake but does not abolish it completely. Conversely, it appears that deleting *traN* from pKpQIL abolishes pilus biogenesis and by extension plasmid transfer. This suggests that it plays an essential role in this system unlike previously studied plasmids. An early study using the F plasmid derivative pOX38 found that disruption of *traN* did not abolish pilus biogenesis but greatly reduced plasmid transfer[112]. More recently, fluorescence imaging and quantification of conjugative pili determined that deletion of *traN* in both pOX38 and pED208 greatly reduced the number of pili expressed per cell[175]. These findings suggest that the function of TraN is not limited to MPS. This is supported by observations made using cryo-ET, which found that Δ*traN* mutants appear impaired in their ability to support the transition of the T4SS from its 'closed' state into its pilus biogenesis state[113,175]. The mechanism underlying this transition remains unclear.

Expanding upon this, TraN has yet to be shown to interact directly with any other component of the *tra* network. This puts into question the role of TraN and MPS specifically within the MPFF conjugative system. The ability to isolate the TraN-OmpK36 complex by mixing its purified constituents shows that TraN can bind its receptor in the absence of other *tra* gene products including TraG which was previously shown to be required for MPS[141]. The

successful formation of a stable mating pair has been proposed to precede the transmission of a 'signal' which initiates DNA transfer from the donor cell[147]. Presumably, this leads to the increased efficiency of plasmid transfer observed when TraN interacts with its receptor. However, the nature of this signal and how it is transduced from TraN to other conjugation proteins is unknown. A recent study describes the use of fluorescence live microscopy to track the localization of the R388 conjugative ATPase when donor cells were mixed with recipients[192]. Adapting this technique for IncF plasmids may aid in determining if MPS promotes the initiation of DNA transfer in these systems. Alternatively, this could be investigated using a temperature-sensitive *traD* mutant[193]. This would be followed up with mutagenesis of TraN to decouple receptor binding from downstream signal transduction to identify domains that mediate interactions with other Tra proteins.

Additional insights into pilus biogenesis and the role of MPS during conjugation were obtained through the derepression of pKpGFP for use in the RTCS. Immunofluorescence microscopy revealed that donors carrying the derepressed pKpGFP were piliated while donors carrying pKpGFP with an intact *finO* were not. A similar observation was seen through Western blot analysis of donor cell lysates using the same anti-pilus antibodies. This implies that, when FinO is expressed, pilus biogenesis is tightly regulated in pure cultures of donor cells. By extension this suggests that an additional signal is required to initiate expression of the pilin subunits and potentially other components of the transfer machinery. Such a signal would likely be associated with the presence of appropriate recipient cells. While this is beyond the scope of MPS, it would present an interesting insight into an additional level of regulation that limits the fitness cost of conjugation on donor cells.

Derepression was also found to be associated with an increase in conjugation frequency even when recipients do not express a suitable OM protein receptor for TraN. The possibility that there are other recipient factors which can participate in MPS certainly cannot be ruled out. In particular, the role of the inner core of LPS was not explored in this study. Mutations in several

genes in the *rfa* gene cluster have previously been shown to affect the transfer of conjugative IncF plasmids like F and colB2[154]. However, the effect of these mutations varies depending on which *rfa* gene has been mutated. While Klimke and Frost reported a modest reduction in F plasmid uptake into *rfaP* null mutants, Perez-Mendoza and de la Cruz did not identify this mutant in their screen of the Keio collection but instead identified a mutant in *rfaC*[100,112]. Interestingly, mutations in either *rfaP* or *rfaC* were previously shown to result in the loss of expression of OmpA[194]. In another study, defects in the LPS inner core greatly reduced the expression of OmpF and OmpC but not OmpA[195]. Collectively, these findings show that mutations in the *rfa* gene cluster have considerable pleiotropic effects on the OM[196]. Thus, to validate the direct involvement of LPS during MPS, the OM proteome will also have to be thoroughly analysed to ensure that the expression of known TraN receptors has not been affected.

If the OM proteins identified in this study are indeed the only receptors for TraN, this would imply that MPS is not essential during conjugation. While this did not apply to experiments conducted with pKpGFP, several groups have demonstrated that OmpA dependency during F plasmid conjugation is alleviated when cells were incubated on solid media[100]. Thus, MPS appears to function mainly to increase conjugation efficiency above the baseline level of transfer which is determined by the expression level of conjugation proteins and donor piliation. This proposed model is illustrated in Figure 4.2.

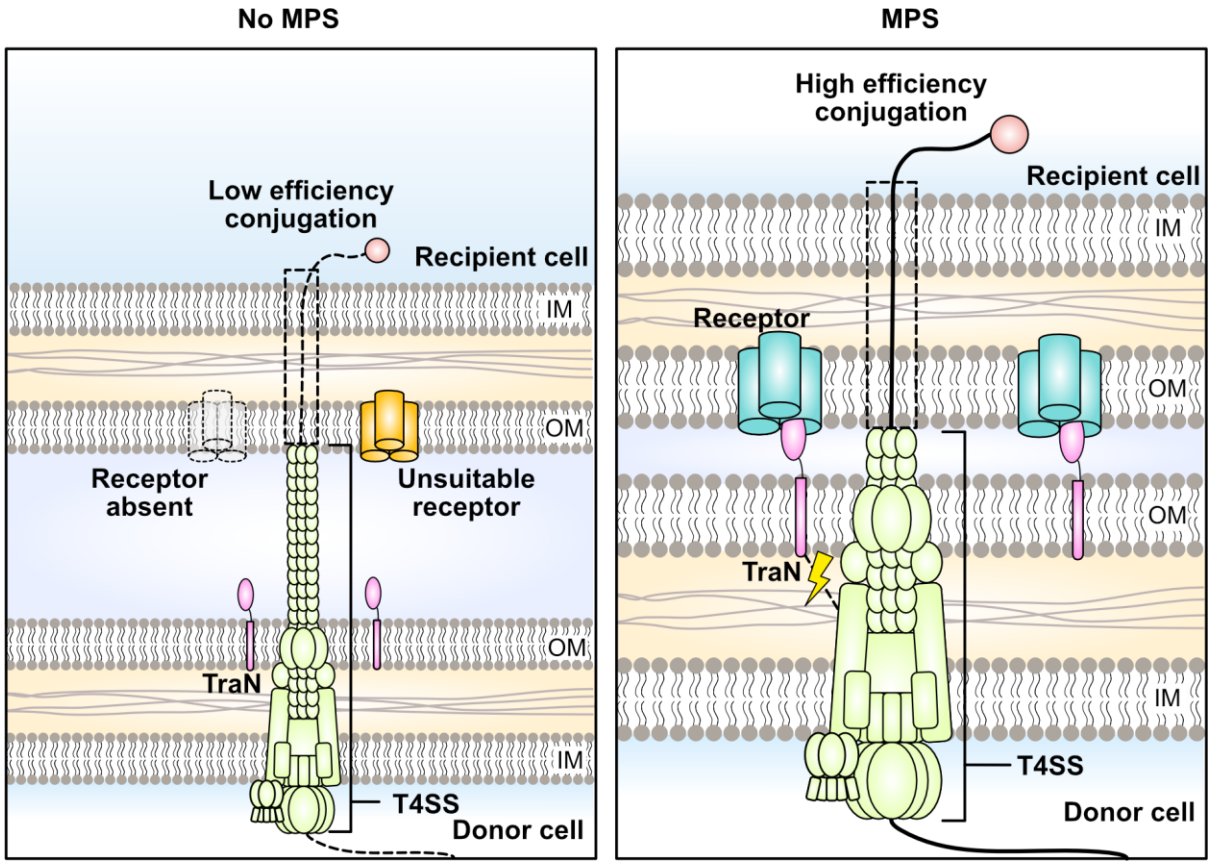


Figure 4.2. Proposed model illustrating the role of MPS during mating pair formation.

Recipient cells which do not express a suitable receptor for TraN cannot partake in MPS and are not intimately attached to donor cells. DNA transfer can still proceed into these recipients albeit at a greatly reduced, baseline frequency. Where recipients are expressing an appropriate receptor, there is MPS and intimate attachment. The binding of TraN to its receptor may transmit a signal to the T4SS to initiate DNA transfer. Conjugation proceeds at much higher frequencies into these recipient cells. As the conduit for DNA was not determined in this study, it is represented using the dashed box.

The model does support, in part, the notion that recipient cells are passive participants during conjugation. The inability to identify a specific receptor for the conjugative pilus from multiple attempts suggests that recipient cells cannot avoid this stage of mating pair formation. This aligns with the 'shoot and pump' model for conjugation proposed by Llosa and colleagues[197]. The establishment of contacts between the pilus and the recipient cell surface may be sufficient to drive low levels of conjugative transfer without the intimate attachment of cells. This model, by extension, supports the role of the pilus as the conduit for relaxase-bound

ssDNA into the recipient. Indeed, more evidence is accumulating in support of this in the context of IncF conjugative systems in Gram-negative bacteria. Conjugation has been observed to occur between spatially distanced cells and structural analysis of the conjugative pilus suggests that its physical and biochemical properties would not only permit but promote DNA transport[128,130]. However, as the nature of the conduit was not confirmed in this work, the proposed model acknowledges that this remains a hypothetical mechanism of DNA entry into recipients. Where MPS does occur, mating pairs become intimately attached and this facilitates increased efficiency of plasmid transfer into the recipient. As discussed earlier, this increased efficiency may result purely from the mechanical stabilization of cells within the mating pair or from the transduction of a signal from TraN to the core transfer machinery following receptor binding.

While this work provides insights on how bacteria become intimately attached during conjugation, it is unclear how these interactions are resolved for transconjugants to become disengaged from within a mating aggregate. The surface exclusion protein, TraT, has been hypothesised to play a role in this step[198]. Functional characterization of TraT determined that it prevents plasmid-bearing cells from engaging in additional conjugation events involving the same plasmid by interfering with MPS[199]. It was initially proposed to interact with and mask the surface exposed loops of OmpA as TraT expression protects cells against phage which utilize OmpA as a receptor[200]. Based on these findings, TraT was hypothesised to mediate surface exclusion by interfering with MPS through the occlusion of the TraN binding site. By extension, the expression of TraT from newly formed transconjugants was proposed to similarly disrupt TraN-OM receptor interactions within the mating pair[198]. Although this is a compelling hypothesis, the effect of surface exclusion could not be overcome by the substitution of *traN* from F with *traN* from R100-1 which does not utilize OmpA as a receptor for MPS[112]. Thus, the mechanism of both surface exclusion and the resolution of MPS remains unclear and warrants further investigation.

4.5 Species specificity during MPS influences plasmid host range

Despite the non-essential nature of MPS and the inability of recipient cells to avoid conjugative transfer completely, the specificity of interactions between TraN and recipient OM proteins appeared to influence the real-world host distribution of conjugative plasmids. Conjugative IncF plasmids were analysed to determine if host range was reflective of the species which express receptors recognized by their encoded TraNs. This was found to be the case for plasmids encoding TraN from pKpQIL, R100-1 and F, demonstrating the MPS is non-essential but highly consequential. The functional assays performed throughout this work have been optimized to maximize plasmid transfer efficiency as only 2 pure cultures of bacteria are mixed and incubated on nutrient rich media. This potentially leads to an underestimation of the importance of MPS in polymicrobial environments where cells are not immobilized and there is substantial diversity in the species present, such as in the gut or in wastewater. Investigating the role of MPS using models for these hotspots of genetic exchange may fully elucidate the importance of this process in a more physiologically relevant setting.

Although plasmid distribution largely mirrored the *in vitro* conjugation data, there were several exceptions. $TraN_{pSLT}$ mediates efficient transfer into all 4 species of recipients tested but plasmids encoding this variant were only isolated from *S. enterica* strains. The pSLT plasmid is a virulence plasmid that is closely associated with several strains of ST[201]. In some strains, this plasmid is no longer conjugative but remains mobilizable suggesting that several essential *tra* gene products are non-functional[110,202]. Phylogenetic analysis of several *Salmonella* virulence plasmids also strongly supports the case that these plasmids are mostly inherited vertically within serovars[203]. As it is not possible to predict conjugative ability based solely on sequence data, the hypothesis that many of the plasmids encoding $TraN_{pSLT}$ are non-conjugative could not be confirmed. Nevertheless, the presence of several plasmids within other serovars of *S. enterica* do suggest they encoded a functional conjugative system at some point. Thus, the absence of these plasmids in other species is likely due to other factors

which restrict plasmid host range. This could be tested through experiments investigating the long-term carriage of pSLT in other enterobacterial species.

In addition, plasmids encoding TraNδ were isolated from a broad range of host species including KP. However, the *in vitro* data suggests that this variant does not mediate efficient conjugation into KP recipients. This discrepancy highlights the potential overrepresentation of several species, particularly KP and EC in this dataset, with this effect being more apparent in groups containing few plasmids such as TraNδ (*n* = 39). The interpretation of plasmid host distribution relies entirely on previously sequenced data which would be expected to largely represent clinical isolates associated with AMR or increased disease burden in developed countries. Thus, in addition to providing a limited idea of the full range of IncF plasmid-carrying species, many species and plasmids may be disproportionately represented in this dataset.

This study was initiated on a hypothesis based upon the close association of pKpQIL-like plasmids with ST258 strains. As the findings did not support the hypothesis that the ST258-associated porins promote plasmid uptake, they suggest that expansion of pKpQIL-harbouring ST258 strains is indeed largely driven by clonal expansion rather than HGT[65]. As host genomic data is not available for each $TraN_{pKpQIL}$-encoding plasmid, KP strain diversity within this subsection of the dataset could not be determined. However, phylogenetic analysis of the plasmid sequences suggests that there is substantial diversity within the group. Ultimately, while the expansion of pKpQIL-like plasmids within ST258 strains may be attributed to clonal expansion, it does not the discount the role of HGT in the dissemination of these plasmids into other KP sequence types which may express isoforms of OmpK36 that can mediate MPS.

4.6 MPS in other conjugative systems

While the analysis in this work was restricted to IncF plasmids, there are also non-IncF plasmids which encode an MPFF conjugative system and TraN. A preliminary analysis of IncH and IncA/C plasmids in the plasmid dataset revealed that the TraN homologues (occasionally

annotated as TrhN) encoded on these plasmids are approximately 400 amino acids longer and share very low sequence identity with the IncF variants. IncA/C plasmids in particular are found in a much wider range of host species across many different families, but it is unclear if this correlates to specificity observed during MPS as seen in this work. Functional characterization of the TraN homologues *in vitro* would be required to understand if the role of MPS during conjugation is conserved in these groups of plasmids and to characterize plasmid transfer range and specificity.

When compared to MPFI systems, MPFF systems appear to be more restricted and less adaptable in terms of host transfer range. In MPFI systems encoding the T4P, rapid recombination of the DNA segments encoding the C-terminal region of the PilV adhesin facilitates the recognition of different moieties in recipient LPS[105,145]. This switching can occur rapidly during growth in different conditions making this system highly adaptable[204]. This posits an interesting alternative to TraN-mediated MPS where receptor specificity is less plastic and is likely to have greater ramifications on plasmid host distribution. In line with this, host distribution of IncI plasmids which encode the T4P-encoding *pil* locus and which are also generally isolated from *Enterobacteriaceae* species can be compared with the IncF plasmid data presented in this work. The host range of IncL/M plasmids which encode MPFI systems but not T4P for MPS can also be analysed to compare the effect of T4P-mediated MPS within this MPF group on plasmid host distribution.

Despite observations of intimately attached cells during RP4 conjugation, proteins which specifically mediate MPS in MPFT conjugative systems have yet to be identified. Donor cells which do not express pili, and thus could not transfer DNA into recipients, were still able to form aggregates with recipients but the junction between mating cells appeared less robust when compared to aggregates containing donors expressing pili[205]. This suggests that the pilus may play a minor, albeit non-essential, role in MPS. Although a donor component for MPS could not be identified, RP4 carriage has been shown to alter donor cell membrane

physiology[206]. This was proposed to play a role in the eventual formation of tight mating junctions during MPS via an unknown mechanism. Interestingly, MPFT systems support conjugation into non-prokaryotic cells, but are generally restricted to supporting transfer on solid surfaces[207,208]. This may allude to a potential compromise that the lack of a dedicated MPS mechanism presents, where host transfer range is considerably broader at the expense of a restricted transfer niche.

5. Conclusion and Future Perspectives

Bacterial conjugation was a foundational cornerstone in the field of bacterial genetics and while it has provided countless insights in this area, there is still much to learn about the process itself. Recently, there has been a welcome resurgence in interest in conjugation driven in part by advances in structural biology and plasmid genomics. This renaissance is aptly timed as the threat of widespread AMR grows, and it becomes increasingly important to understand how MGEs contributing to resistance disseminate by HGT. The findings presented in this work provide insights on a key stage of DNA transfer with potential implications on developing strategies which may interfere with conjugation.

Conjugation has often been viewed as a promiscuous mechanism of DNA transfer with plasmid host range being largely determined downstream by replication associated factors[197]. However, the work presented here offers an additional perspective by demonstrating that the host range of plasmids belonging to the same replicon type can also be influenced during the transfer process itself. TraN-mediated recipient specificity may be a necessary adaptation to the acquisition of a system which supports conjugation in a wide variety of niches such as within the intestinal microbiota or in wastewater[209,210]. The ability to discriminate between recipient populations within microbe-rich environments may be key to ensuring the successful propagation of a plasmid within a suitable host species.

Future work could be aimed at understanding how specificity has evolved during MPS using *in vitro* evolution experiments. This can be done to understand how the binding interface at the TraN tip domain evolves to mediate interactions with receptor homologues that are not currently recognized by selecting for plasmids with improved conjugation efficiency. On the flipside, *in vitro* evolution experiments using a negative selection system could be used to understand how mutations are accumulated in OM proteins to inhibit TraN binding and thus plasmid uptake. These experiments could aid in identifying important residues required for facilitating the interaction between MPS proteins via a genetic rather than structural approach.

In addition, additional insights could be gained on the importance of the genetic background of the host and the plasmid backbone in influencing the evolution of specificity.

The identification of a direct interaction between a donor and recipient bacterium which drives efficient conjugation presents a potential target for the inhibition of this process. As alluded to earlier, conjugation is increasingly recognized as an important mechanism driving the dissemination of AMR genes[76]. In line with this, several groups have proposed the development of 'conjugation inhibitors', compounds which can block plasmid transfer by interfering with key steps in the conjugation process[211]. Conceptually, blocking either the assembly of the conjugation machinery or the subsequent process of MPF is seen as a preferred strategy since the downstream processes of plasmid processing and maintenance would involve essential recipient cell factors such as recipient DNA polymerase[212]. Although several compounds appear to inhibit conjugation, many of these do not directly target conjugative processes but rather affect bacterial growth or DNA synthesis[213,214]. Recently, synthetic fatty acids have been shown to specifically inhibit plasmid transfer by 1-2 log fold in an *in vivo* conjugation study, supporting the feasibility of conjugation blocking strategies[215].

Future work could look at the feasibility of targeting TraN as an IncF plasmid-specific conjugation blocking strategy. As a proof-of-concept experiment, TraN-specific antibodies or small molecule compounds that mask the receptor binding interface could be used to assess the impact of blocking MPS on conjugation efficiency. This could be further assessed within a polymicrobial community, first *in vitro* and later *in vivo* using mouse models, to simulate plasmid transmission dynamics more accurately within physiologically relevant conditions. Establishing polymicrobial models to assess conjugation may also reveal a more diverse transfer range of host species and help to uncover previously unknown reservoirs for conjugative IncF plasmids within important HGT hotspots.

Lastly, this work demonstrates the value of combining rapid structure prediction with genomic analysis of large plasmid datasets. Many conjugation studies still report findings that are based on the characterization of prototypical conjugative systems within model laboratory organisms. However, this work demonstrates that there is significant diversity in the mechanisms employed to drive transfer even amongst closely related plasmids. Understandably, it is not feasible to characterize each conjugative system in detail, especially via structural studies. This is where technology such as AlphaFold can be used to improve the understanding of these systems beyond what can be learned from model systems. In fact, recent advancements in this technology now allow for the prediction of multi-chain protein complexes which was used to validate the cryo-EM structure of the TraN-OmpK36 complex[216]. In time, it is reasonable to expect that larger complexes may also be modelled *in silico* potentially aiding in elucidating the role of TraN within the transfer gene product network.

As plasmid sequencing becomes more accessible and cost-effective, it will present more opportunities for the improved surveillance of MGEs mediating the dissemination of AMR genes in clinical settings. Expanding this to include surveillance of plasmids circulating in environmental isolates for MPFF-like transfer genes and *traN* in particular may be useful for the identification of plasmids which may pose a risk in future as a suitable scaffold for housing virulence/resistance genes. This could enable the pre-emptive design of intervention strategies targeted against specific conjugative plasmids to complement other efforts aimed at prolonging our ability to utilize vital, life-saving antimicrobial compounds.

References

1. Janda, J. M. & Abbott, S. L. The changing face of the family *Enterobacteriaceae* (Order: '*Enterobacterales*'): New members, taxonomic issues, geographic expansion, and new diseases and disease syndromes. *Clin Microbiol Rev* **34**, 1–45 (2021).
2. Tenaillon, O., Skurnik, D., Picard, B. & Denamur, E. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* **8**, 207–217 (2010).
3. Osbelt, L. *et al. Klebsiella oxytoca* causes colonization resistance against multidrug-resistant *K. pneumoniae* in the gut via cooperative carbohydrate competition. *Cell Host Microbe* **29**, 1663-1679.e7 (2021).
4. Mullineaux-Sanders, C. *et al. Citrobacter amalonaticus* inhibits the growth of *Citrobacter rodentium* in the gut lumen. *mBio* **12**, (2021).
5. Paterson, D. L. Resistance in Gram-negative bacteria: *Enterobacteriaceae*. *American Journal of Medicine* **119**, S20–S28 (2006).
6. Redondo-Salvo, S. *et al.* Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat Commun* **11**, 1–13 (2020).
7. Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *J Infect Dis* **197**, 1079–1081 (2008).
8. Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. & Pardesi, K. R. Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review. *Front Microbiol* **10**, (2019).
9. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, (2022).
10. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414 (2010).
11. Whitfield, C. & Stephen Trent, M. Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* **83**, 99–128 (2014).

12. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. *Proceedings of the National Academy of Sciences* **106**, 8009–8014 (2009).

13. Cuthbertson, L., Mainprize, I. L., Naismith, J. H. & Whitfield, C. Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in Gram-negative bacteria. *Microbiology and Molecular Biology Reviews* **73**, 155–177 (2009).

14. Limoli, D. H., Jones, C. J. & Wozniak, D. J. Bacterial extracellular polysaccharides in biofilm formation and function. *Microbiol Spectr* **3**, MB-0011-2014 (2015).

15. Fairman, J. W., Noinaj, N. & Buchanan, S. K. The structural biology of β-barrel membrane proteins: A summary of recent reports. *Curr Opin Struct Biol* **21**, 523–531 (2011).

16. Wang, Y. The function of OmpA in *Escherichia coli*. *Biochem Biophys Res Commun* **292**, 396–401 (2002).

17. Orme, R., Douglas, C. W. I., Rimmer, S. & Webb, M. Proteomic analysis of *Escherichia coli* biofilms reveals the overexpression of the outer membrane protein OmpA. *Proteomics* **6**, 4269–4277 (2006).

18. Morona, R., Klose, M. & Henning, U. *Escherichia coli* K-12 outer membrane protein (OmpA) as a bacteriophage receptor: Analysis of mutant genes expressing altered protein. *J Bacteriol* **159**, 570–578 (1984).

19. Prasadarao, N. v., Blom, A. M., Villoutreix, B. O. & Linsangan, L. C. A novel interaction of outer membrane protein A with C4b binding protein mediates serum resistance of *Escherichia coli* K1. *The Journal of Immunology* **169**, 6352–6360 (2002).

20. Nikaido, H. Transport across the bacterial outer membrane. *J Bioenerg Biomembr* **25**, 581–589 (1993).

21. Slauch, J. M., Garrett, S., Jackson, D. E. & Silhavy, T. J. EnvZ functions through OmpR to control porin gene expression in *Escherichia coli* K-12. *J Bacteriol* **170**, 439–441 (1988).

22. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews* **67**, 593–656 (2003).

23. Weiner, J. H. & Li, L. Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim Biophys Acta Biomembr* **1778**, 1698–1713 (2008).

24. Friedlaender, C. Ueber die Schizomyceten bei der acuten fibrösen Pneumonie. *Archiv für Pathologische Anatomie und Physiologie und für Klinische Medicin* **87**, 319–324 (1882).

25. Bagley, S. T. Habitat association of Klebsiella species. *Infection Control* **6**, 52–58 (1985).

26. Podschun, R. & Ullmann, U. *Klebsiella* spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev* **11**, 589–603 (1998).

27. Magill, S. S. *et al.* Multistate point-prevalence survey of health care–associated infections. *New England Journal of Medicine* **370**, 1198–1208 (2014).

28. Paczosa, M. K. & Mecsas, J. *Klebsiella pneumoniae*: Going on the offense with a strong defense. *Microbiology and Molecular Biology Reviews* **80**, 629–661 (2016).

29. Álvarez, D., Merino, S., Tomás, J. M., Benedí, V. J. & Albertí, S. Capsular polysaccharide is a major complement resistance factor in lipopolysaccharide O side chain-deficient *Klebsiella pneumoniae* clinical isolates. *Infect Immun* **68**, 953–955 (2000).

30. Domenico, P., Salo, R. J., Cross, A. S. & Cunha, B. A. Polysaccharide capsule-mediated resistance to opsonophagocytosis in *Klebsiella pneumoniae*. *Infect Immun* **62**, 4495–4499 (1994).

31. Lawlor, M. S., Hsu, J., Rick, P. D. & Miller, V. L. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. *Mol Microbiol* **58**, 1054–1073 (2005).

32. Yoshida, K. *et al.* Role of bacterial capsule in local and systemic inflammatory responses of mice during pulmonary infection with *Klebsiella pneumoniae*. *J Med Microbiol* **49**, 1003–1010 (2000).

33. Tan, Y. H., Chen, Y., Chu, W. H. W., Sham, L. T. & Gan, Y. H. Cell envelope defects of different capsule-null mutants in K1 hypervirulent Klebsiella pneumoniae can affect bacterial pathogenesis. *Mol Microbiol* **113**, 889–905 (2020).

34. Pan, Y. J. *et al.* Capsular types of *Klebsiella pneumoniae* revisited by *wzc* sequencing. *PLoS One* **8**, e80670 (2013).

35. Cryz, S. J., Mortimer, P. M., Mansfield, V. & Germanier, R. Seroepidemiology of *Klebsiella* bacteremic isolates and implications for vaccine development. *J Clin Microbiol* **23**, 687–690 (1986).

36. Pitout, J. D. D., Nordmann, P. & Poirel, L. Carbapenemase-producing *Klebsiella pneumoniae*, a key pathogen set for global nosocomial dominance. *Antimicrob Agents Chemother* **59**, 5873–5884 (2015).

37. Bush, K. & Bradford, P. A. β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harb Perspect Med* **6**, (2016).

38. Babini, G. S. & Livermore, D. M. Are SHV β-lactamases universal in *Klebsiella pneumoniae*? *Antimicrob Agents Chemother* **44**, 2230 (2000).

39. Paterson, D. L. & Bonomo, R. A. Extended-spectrum β-lactamases: A clinical update. *Clin Microbiol Rev* **18**, 657–686 (2005).

40. Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**, 318–327 (2018).

41. Kitchel, B. *et al.* Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* isolates in the United States: Clonal expansion of multilocus sequence type 258. *Antimicrob Agents Chemother* **53**, 3365–3370 (2009).

42. DeLeo, F. R. *et al.* Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proceedings of the National Academy of Sciences* **111**, 4988–4993 (2014).

43. Tsai, Y. K. *et al. Klebsiella pneumoniae* outer membrane porins OmpK35 and OmpK36 play roles in both antimicrobial resistance and virulence. *Antimicrob Agents Chemother* **55**, 1485–1493 (2011).

44. Sugawara, E., Kojima, S. & Nikaido, H. *Klebsiella pneumoniae* major porins OmpK35 and OmpK36 allow more efficient diffusion of β-lactams than their *Escherichia coli* homologs OmpF and OmpC. *J Bacteriol* **198**, 3200–3208 (2016).

45. Acosta-Gutiérrez, S. *et al.* Getting drugs into Gram-negative bacteria: Rational rules for permeation through general porins. *ACS Infect Dis* **4**, 1487–1498 (2018).

46. Phale, P. S. *et al.* Stability of trimeric OmpF porin: The contributions of the latching loop L2. *Biochemistry* **37**, 15663–15670 (1998).

47. Cowan, S. W. *et al.* Crystal structures explain functional properties of two *E. coli* porins. *Nature* **358**, 727–733 (1992).

48. Jeanteur, D., Lakey, J. H. & Pattus, F. The bacterial porin superfamily: sequence alignment and structure prediction. *Mol Microbiol* **5**, 2153–2164 (1991).

49. Doménech-Sánchez, A., Hernández-Allés, S., Martínez-Martínez, L., Benedí, V. J. & Albertí, S. Identification and characterization of a new porin gene of Klebsiella pneumoniae: Its role in β-lactam antibiotic resistance. *J Bacteriol* **181**, 2726–2732 (1999).

50. Hernández-Allés, S. *et al.* Porin expression in clinical isolates of *Klebsiella pneumoniae*. *Microbiology (N Y)* **145**, 673–679 (1999).

51. Bowers, J. R. *et al.* Genomic analysis of the emergence and rapid global dissemination of the clonal group 258 *Klebsiella pneumoniae* pandemic. *PLoS One* **10**, e0133727 (2015).

52. Wong, J. L. C. *et al.* OmpK36-mediated Carbapenem resistance attenuates ST258 *Klebsiella pneumoniae* in vivo. *Nat Commun* **10**, 3957 (2019).

53. Fajardo-Lubián, A., ben Zakour, N. L., Agyekum, A., Qi, Q. & Iredell, J. R. Host adaptation and convergent evolution increases antibiotic resistance without loss of virulence in a major human pathogen. *PLoS Pathog* **15**, e1007218 (2019).

54. David, S. *et al.* Widespread emergence of OmpK36 loop 3 insertions among multidrug-resistant clones of *Klebsiella pneumoniae*. *PLoS Pathog* **18**, e1010334 (2022).

55. Wang, M. *et al.* High osmotic stress increases OmpK36 expression through the regulation of KbvR to decrease the antimicrobial resistance of *Klebsiella pneumoniae*. *Microbiol Spectr* **10**, e00507-22 (2022).

56. Tian, X. *et al.* First description of antimicrobial resistance in carbapenem-susceptible *Klebsiella pneumoniae* after imipenem treatment, driven by outer membrane remodeling. *BMC Microbiol* **20**, 1–11 (2020).

57. Wong, J. L. C. *et al.* Recurrent emergence of carbapenem resistance in Klebsiella pneumoniae mediated by an inhibitory ompK36 mRNA secondary structure. *bioRxiv* (2022) doi:10.1101/2022.01.05.475072.

58. Wyres, K. L. & Holt, K. E. *Klebsiella pneumoniae* as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Curr Opin Microbiol* **45**, 131–139 (2018).

59. Göttig, S., Riedel-Christ, S., Saleh, A., Kempf, V. A. J. & Hamprecht, A. Impact of blaNDM-1 on fitness and pathogenicity of *Escherichia coli* and *Klebsiella pneumoniae*. *Int J Antimicrob Agents* **47**, 430–435 (2016).

60. Löhr, I. H. *et al.* Persistence of a pKPN3-like CTX-M-15-encoding IncFIIK plasmid in a *Klebsiella pneumoniae* ST17 host during two years of intestinal colonization. *PLoS One* **10**, e0116516 (2015).

61. Conlan, S. *et al.* Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. *mBio* **7**, (2016).

62. Findlay, J., Hamouda, A., Dancer, S. J. & Amyes, S. G. B. Rapid acquisition of decreased carbapenem susceptibility in a strain of *Klebsiella pneumoniae* arising during meropenem therapy. *Clinical Microbiology and Infection* **18**, 140–146 (2012).

63. Poirel, L., Bonnin, R. A. & Nordmann, P. Genetic features of the widespread plasmid coding for the carbapenemase OXA-48. *Antimicrob Agents Chemother* **56**, 559–562 (2012).

64. Cuzon, G., Naas, T. & Nordmann, P. Functional characterization of Tn*4401*, a Tn*3*-based transposon involved in bla_{KPC} gene mobilization. *Antimicrob Agents Chemother* **55**, 5370–5373 (2011).

65. Chen, L. *et al.* Comparative genomic analysis of KPC-encoding pKpQIL-like plasmids and their distribution in New Jersey and New York hospitals. *Antimicrob Agents Chemother* **58**, 2871–2877 (2014).

66. Queenan, A. M. & Bush, K. Carbapenemases: The versatile β-lactamases. *Clin Microbiol Rev* **20**, 440–458 (2007).

67. Navon-Venezia, S., Kondratyeva, K. & Carattoli, A. *Klebsiella pneumoniae*: A major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol Rev* **41**, 252–275 (2017).

68. Leavitt, A., Chmelnitsky, I., Carmeli, Y. & Navon-Venezia, S. Complete nucleotide sequence of KPC-3-encoding plasmid pKpQIL in the epidemic *Klebsiella pneumoniae* sequence type 258. *Antimicrob Agents Chemother* **54**, 4493–4496 (2010).

69. Leavitt, A., Chmelnitsky, I., Ofek, I., Carmeli, Y. & Navon-Venezia, S. Plasmid pKpQIL encoding KPC-3 and TEM-1 confers carbapenem resistance in an extremely drug-resistant epidemic *Klebsiella pneumoniae* strain. *Journal of Antimicrobial Chemotherapy* **65**, 243–248 (2009).

70. Naas, T. *et al.* Genetic structures at the origin of acquisition of the β-lactamase *blaKPC* gene. *Antimicrob Agents Chemother* **52**, 1257–1263 (2008).

71. Peirano, G. *et al.* Importance of clonal complex 258 and $IncF_{K2\text{-}like}$ plasmids among a global collection of *Klebsiella pneumoniae* with *blaKPC*. *Antimicrob Agents Chemother* **61**, (2017).

72. David, S. *et al.* Integrated chromosomal and plasmid sequence analyses reveal diverse modes of carbapenemase gene spread among *Klebsiella pneumoniae*. *Proceedings of the National Academy of Sciences* **117**, 25043–25054 (2020).

73. Doumith, M. *et al.* Major role of pKpQIL-like plasmids in the early dissemination of KPC-type carbapenemases in the UK. *Journal of Antimicrobial Chemotherapy* **72**, 2241–2248 (2017).

74. Lederberg, J. & Tatum, E. L. Gene Recombination in Escherichia coli. *Nature* **158**, 558 (1946).

75. Ochman, H., Lawrence, J. G. & Grolsman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).

76. Mazel, D. & Davies, J. Antibiotic resistance in microbes. in *Cellular and Molecular Life Sciences* vol. 56 742–754 (1999).

77. Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F. & Arechaga, I. Towards an integrated model of bacterial conjugation. *FEMS Microbiol Rev* **39**, 81–95 (2015).

78. de La Cruz, F., Frost, L. S., Meyer, R. J. & Zechner, E. L. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**, 18–40 (2010).

79. Waksman, G. From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep* **20**, (2019).

80. Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C. & de la Cruz, F. Mobility of Plasmids. *Microbiology and Molecular Biology Reviews* **74**, 434–452 (2010).

81. Novick, R. P. Plasmid incompatibility. *Microbiol Rev* **51**, 381–395 (1987).

82. Couturier, M., Bex, F., Bergquist, P. L. & Maas, W. K. Identification and classification of bacterial plasmids. *Microbiol Rev* **52**, 375 (1988).

83. Bergquist, P. L., Saadi, S. & Maas, W. K. Distribution of basic replicons having homology with RepFIA, RepFIB, and RepFIC among IncF group plasmids. *Plasmid* **15**, 19–34 (1986).

84. Carattoli, A. *et al.* Identification of plasmids by PCR-based replicon typing. *J Microbiol Methods* **63**, 219–228 (2005).

85. Carloni, E. *et al.* Comparative analysis of the standard PCR-Based Replicon Typing (PBRT) with the commercial PBRT-KIT. *Plasmid* **90**, 10–14 (2017).

86. Orlek, A. *et al.* Plasmid classification in an era of whole-genome sequencing: Application in studies of antibiotic resistance epidemiology. *Front Microbiol* **8**, 182 (2017).

87. Francia, M. V. *et al.* A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev* **28**, 79–100 (2004).

88. Koraimann, G. Spread and persistence of virulence and antibiotic resistance genes: A ride on the F plasmid conjugation module. *EcoSal Plus* **8**, (2018).

89. Guglielmini, J. *et al.* Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res* **42**, 5715–5727 (2014).

90. Lawley, T. D., Klimke, W. A., Gubbins, M. J. & Frost, L. S. F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* **224**, 1–15 (2003).

91. Macé, K. *et al.* Cryo-EM structure of a type IV secretion system. *Nature* **607**, 191–196 (2022).

92. Virolle, C., Goldlust, K., Djermoun, S., Bigot, S. & Lesterlin, C. Plasmid transfer by conjugation in Gram-negative bacteria: From the cellular to the community level. *Genes (Basel)* **11**, 1239 (2020).

93. Bradley, D. E., Taylor, D. E. & Cohen, D. R. Specification of surface mating systems among conjugative drug resistance plasmids in *Escherichia coli* K-12. *J Bacteriol* **143**, 1466–1470 (1980).

94. Fernandez-Lopez, R., de Toro, M., Moncalian, G., Garcillan-Barcia, M. P. & de la Cruz, F. Comparative Genomics of the Conjugation Region of F-like Plasmids: Five Shades of F. *Front Mol Biosci* **3**, 71 (2016).

95. Villa, L., Poirel, L., Nordmann, P., Carta, C. & Carattoli, A. Complete sequencing of an IncH plasmid carrying the blaNDM-1, blaCTX-M-15 and qnrB1 genes. *Journal of Antimicrobial Chemotherapy* **67**, 1645–1650 (2012).

96. Fricke, W. F. *et al.* Comparative genomics of the IncA/C multidrug resistance plasmid family. *J Bacteriol* **191**, 4750–4757 (2009).

97. Suzuki, H., Yano, H., Brown, C. J. & Top, E. M. Predicting plasmid promiscuity based on genomic signature. *J Bacteriol* **192**, 6045–6055 (2010).

98. Encinas, D. *et al.* Plasmid conjugation from proteobacteria as evidence for the origin of xenologous genes in cyanobacteria. *J Bacteriol* **196**, 1551–1559 (2014).

99. Gordon, J. E. & Christie, P. J. The Agrobacterium Ti Plasmids. *Microbiol Spectr* **2**, (2014).

100. Pérez-Mendoza, D. & de La Cruz, F. Escherichia coli genes affecting recipient ability in plasmid conjugation: Are there any? *BMC Genomics* **10**, (2009).

101. Blatny, J. M., Brautaset, T., Winther-Larsen, H. C., Haugan, K. & Valla, S. Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon. *Appl Environ Microbiol* **63**, 370–379 (1997).

102. Babic, A., Guérout, A. M. & Mazel, D. Construction of an improved RP4 (RK2)-based conjugative system. *Res Microbiol* **159**, 545–549 (2008).

103. Lawley, T. D., Gilmour, M. W., Gunton, J. E., Tracz, D. M. & Taylor, D. E. Functional and mutational analysis of conjugative transfer region 2 (Tra2) from the IncHI1 plasmid R27. *J Bacteriol* **185**, 581–591 (2003).

104. de Toro, M., Garcilláon-Barcia, M. P. & de La Cruz, F. Plasmid diversity and adaptation analyzed by massive sequencing of *Escherichia coli* plasmids. *Microbiol Spectr* **2**, (2014).

105. Komano, T., Kim, S. R., Yoshida, T. & Nisioka, T. DNA rearrangement of the shufflon determines recipient specificity in liquid mating of IncI1 plasmid R64. *J Mol Biol* **243**, 6–9 (1994).

106. Komano, T., Yoshida, T., Narahara, K. & Furuya, N. The transfer region of IncI1 plasmid R64: Similarities between R64 *tra* and *Legionella icm/dot* genes. *Mol Microbiol* **35**, 1348–1359 (2000).

107. Sampei, G. *et al.* Complete genome sequence of the incompatibility group I1 plasmid R64. *Plasmid* **64**, 92–103 (2010).

108. Dmowski, M., Golębiewski, M. & Kern-Zdanowicza, I. Characteristics of the conjugative transfer system of the IncM plasmid pCTX-M3 and identification of its putative regulators. *J Bacteriol* **200**, e00234-18 (2018).

109. Brinkley, C. *et al.* Nucleotide sequence analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid pMAR7. *Infect Immun* **74**, 5408–5413 (2006).

110. Ahmer, B. M. M., Tran, M. & Heffron, F. The virulence plasmid of *Salmonella typhimurium* is self-transmissible. *J Bacteriol* **181**, 1364–1368 (1999).

111. Womble, D. D. & Rownd, R. H. Genetic and Physical Map of Plasmid NR1: Comparison with Other IncFII Antibiotic Resistance Plasmids. *Microbiol Rev* **52**, 433–451 (1988).

112. Klimke, W. A. & Frost, L. S. Genetic analysis of the role of the transfer gene, traN, of the F and R100-1 plasmids in mating pair stabilization during conjugation. *J Bacteriol* **180**, 4036–4043 (1998).

113. Hu, B., Khara, P. & Christie, P. J. Structural bases for F plasmid conjugation and F pilus biogenesis in Escherichia coli. *Proceedings of the National Academy of Sciences* **116**, 14222–14227 (2019).

114. Zatyka, M. & Thomas, C. M. Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiol Rev* **21**, 291–319 (1998).

115. Frost, L. S. & Koraimann, G. Regulation of bacterial conjugation: Balancing opportunity with adversity. *Future Microbiol* **5**, 1057–1071 (2010).

116. Mark Glover, J. N. *et al.* The FinO family of bacterial RNA chaperones. *Plasmid* **78**, 79–87 (2015).

117. Lee, S. H., Frost, L. S. & Paranchych, W. FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO. *MGG Molecular & General Genetics* **235**, 131–139 (1992).

118. Koraimann, G., Koraimann, C., Koronakis, V., Schlager, S. & Högenauer, G. Repression and derepression of conjugation of plasmid R1 by wild-type and mutated *finP* antisense RNA. *Mol Microbiol* **5**, 77–87 (1991).

119. Cheah, K. C. & Skurray, R. The F plasmid carries an IS3 insertion within *finO*. *J Gen Microbiol* **132**, 3269–3275 (1986).

120. Amin, H., Ilangovan, A. & Costa, T. R. D. Architecture of the outer-membrane core complex from a conjugative type IV secretion system. *Nat Commun* **12**, 1–12 (2021).

121. Frost, L. S., Paranchych, W. & Willetts, N. S. DNA sequence of the F traALE region that includes the gene for F pilin. *J Bacteriol* **160**, 395–401 (1984).

122. Maneewannakul, K., Maneewannakul, S. & Ippen-Ihler, K. Synthesis of F pilin. *J Bacteriol* **175**, 1384 (1993).

123. Anthony, K. G., Klimke, W. A., Manchak, J. & Frost, L. S. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: Insights into the mechanism of conjugation. *J Bacteriol* **181**, 5149–5159 (1999).

124. Maneewannakul, S., Maneewannakul, K. & Ippen-Ihler, K. Characterization of *trbC*, a new F plasmid tra operon gene that is essential to conjugative transfer. *J Bacteriol* **173**, 3872–3878 (1991).

125. Anthony, K. G., Kathir, P., Moore, D., Ippen-Ihler, K. & Frost, L. S. Analysis of the traLEKBP sequence and the TraP protein from three F-like plasmids: F, R100-1, and ColB2. *J Bacteriol* **178**, 3194–3200 (1996).

126. Moore, D. *et al.* Characterization of the F-plasmid conjugative transfer gene *traU*. *J Bacteriol* **172**, 4263–4270 (1990).

127. Maneewannakul, S., Maneewannakul, K. & Ippen-Ihler, K. Characterization, localization, and sequence of F transfer region products: The pilus assembly gene product TraW and a new product, TrbI. *J Bacteriol* **174**, 5567–5574 (1992).

128. Costa, T. R. D. *et al.* Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* **166**, 1436-1444.e10 (2016).

129. Zheng, W. *et al.* Cryoelectron-microscopic structure of the pKpQIL conjugative pili from carbapenem-resistant *Klebsiella pneumoniae*. *Structure* **28**, 1321-1328.e2 (2020).

130. Babić, A., Lindner, A. B., Vulić, M., Stewart, E. J. & Radman, M. Direct visualization of horizontal gene transfer. *Science (1979)* **319**, 1533–1536 (2008).

131. Ou, J. T. & Anderson, T. F. Role of pili in bacterial conjugation. *J Bacteriol* **102**, 648–654 (1970).

132. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343–347 (2011).

133. Aly, K. A. & Baron, C. The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology (N Y)* **153**, 3766–3775 (2007).

134. Clarke, M., Maddera, L., Harris, R. L. & Silverman, P. M. F-pili dynamics by live-cell imaging. *Proceedings of the National Academy of Sciences* **105**, 17978–197981 (2008).

135. Zheng, W. *et al.* Refined cryo-EM structure of the T4 tail tube: Exploring the lowest dose limit. *Structure* **25**, 1436-1441.e2 (2017).

136. Trokter, M. & Waksman, G. Translocation through the conjugative type IV secretion system requires unfolding of its protein substrate. *J Bacteriol* **200**, e00615-17 (2018).

137. Daniell, S. J. *et al.* 3D structure of EspA filaments from enteropathogenic *Escherichia coli*. *Mol Microbiol* **49**, 301–308 (2003).

138. Grohmann, E., Muth, G. & Espinosa, M. Conjugative plasmid transfer in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews* **67**, 277–301 (2003).

139. Achtman, M., Morelli, G. & Schwuchow, S. Cell-cell interactions in conjugating *Escherichia coli*: Role of F pili and fate of mating aggregates. *J Bacteriol* **135**, 1053–1061 (1978).

140. Dürrenberger, M. B., Villiger, W. & Bächi, T. Conjugational junctions: Morphology of specific contacts in conjugating *Escherichia coli* bacteria. *J Struct Biol* **107**, 146–156 (1991).

141. Manning, P. A., Morelli, G. & Achtman, M. TraG protein of the F sex factor of *Escherichia coli* K-12 and its role in conjugation. *Proceedings of the National Academy of Sciences* **78**, 7487 (1981).

142. Kishida, K. *et al.* Contributions of F-specific subunits to the F plasmid-encoded type IV secretion system and F pilus. *Mol Microbiol* **117**, 1275–1290 (2022).

143. Firth, N. & Skurray, R. Characterization of the F plasmid bifunctional conjugation gene, *traG*. *MGG Molecular & General Genetics* **232**, 145–153 (1992).

144. Komano, T., Kim, S. R. & Yoshida, T. Mating variation by DNA inversions of shufflon in plasmid R64. *Adv Biophys* **31**, 181–193 (1995).

145. Allard, N., Neil, K., Grenier, F. & Rodrigue, S. The type IV pilus of plasmid TP114 displays adhesins conferring conjugation specificity and is important for DNA transfer in the mouse gut microbiota. *Microbiol Spectr* **10**, e02303-21 (2022).

146. Maneewannakul, S., Kathir, P. & Ippen-Ihler, K. Characterization of the F plasmid mating aggregation gene *traN* and of a new F transfer region locus *trbE*. *J Mol Biol* **225**, 299–311 (1992).

147. Klimke, W. A. *et al.* The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. *Microbiology (N Y)* **151**, 3527–3540 (2005).

148. Audette, G. F., Manchak, J., Beatty, P., Klimke, W. A. & Frost, L. S. Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology (N Y)* **153**, 442–451 (2007).

149. 3rd, R. C., Charamella, L. J., Stallions, D. R. & Mays, J. A. Parental Functions During Conjugation in Escherichia coli K-12. *Bacteriol Rev* **32**, 320–348 (1968).

150. Skurray, R. A., Hancock, R. E. W. & Reeves, P. Con- mutants: class of mutants in *Escherichia coli* K-12 lacking a major cell wall protein and defective in conjugation and adsorption of a bacteriophage. *J Bacteriol* **119**, 726–735 (1974).

151. Achtman, M., Schwuchow, S., Helmuth, R., Morelli, G. & Manning, P. A. Cell-cell interactions in conjugating *Escherichia coli*: Con- mutants and stabilization of mating aggregates. *MGG Molecular & General Genetics* **164**, 171–183 (1978).

152. Manoil, C. & Rosenbusch, J. P. Conjugation-deficient mutants of Escherichia coli distinguish classes of functions of the outer membrane OmpA protein. *MGG Molecular & General Genetics* **187**, 148–156 (1982).

153. Havekes, L., Hoekstra, W. & Kempen, H. Relation between F, R1, R100 and R144 *Escherichia coli* K-12 donor strains in mating. *MGG Molecular & General Genetics* **155**, 185–189 (1977).

154. Anthony, K. G., Sherburne, C., Sherburne, R. & Frost, L. S. The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. *Mol Microbiol* **13**, 939–953 (1994).

155. Ishiwa, A. & Komano, T. Thin pilus PilV adhesins of plasmid R64 recognize specific structures of the lipopolysaccharide molecules of recipient cells. *J Bacteriol* **185**, 5192–5199 (2003).

156. Ishiwa, A. & Komano, T. PilV adhesins of plasmid R64 thin pili specifically bind to the lipopolysaccharides of recipient cells. *J Mol Biol* **343**, 615–625 (2004).

157. Moriguchi, K. *et al.* Targeting antibiotic resistance genes is a better approach to block acquisition of antibiotic resistance than blocking conjugal transfer by recipient cells: A genome-wide screening in *Escherichia coli*. *Front Microbiol* **10**, 2939 (2020).

158. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* **2**, 2006.0008 (2006).

159. Hardiman, C. A. *et al.* Horizontal transfer of carbapenemase-encoding plasmids and comparison with hospital epidemiology data. *Antimicrob Agents Chemother* **60**, 4910–4919 (2016).

160. Low, W. W. *et al.* Mating pair stabilization mediates bacterial conjugation species specificity. *Nat Microbiol* **7**, 1016–1027 (2022).

161. Buckner, M. M. C. *et al.* Clinically relevant plasmid-host interactions indicate that transcriptional and not genomic modifications ameliorate fitness costs of *Klebsiella pneumoniae* carbapenemase-carrying plasmids. *mBio* **9**, e02303-17 (2018).

162. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

163. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, (2004).

164. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).

165. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).

166. Afonine, P. v. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531–544 (2018).

167. Che, Y. *et al.* Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proceedings of the National Academy of Sciences* **118**, e2008731118 (2021).

168. Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* **58**, 3895–3903 (2014).

169. Gerts, E. M., Yu, Y. K., Agarwala, R., Schäffer, A. A. & Altschul, S. F. Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biol* **4**, 1–14 (2006).

170. Rocker, A. *et al.* Global trends in proteome remodeling of the outer membrane modulate antimicrobial permeability in Klebsiella pneumoniae. *mBio* **11**, (2020).

171. Chong, Z.-S., Woo, W.-F. & Chng, S.-S. Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in *E scherichia coli*. *Mol Microbiol* **98**, 1133–1146 (2015).

172. Ruan, X., Loyola, D. E., Marolda, C. L., Perez-Donoso, J. M. & Valvano, M. A. The WaaL O-antigen lipopolysaccharide ligase has features in common with metal ion-independent inverting glycosyltransferases. *Glycobiology* **22**, 288–299 (2012).

173. Ried, G. & Henning, U. A unique amino acid substitution in the outer membrane protein OmpA causes conjugation deficiency in Escherichia coli K-12. *FEBS Lett* **223**, 387–390 (1987).

174. Alderliesten, J. B. *et al.* Effect of donor-recipient relatedness on the plasmid conjugation frequency: A meta-analysis. *BMC Microbiol* **20**, (2020).

175. Kishida, K. *et al.* Contributions of F-specific subunits to the F plasmid-encoded type IV secretion system and F pilus. *Mol Microbiol* **117**, 1275–1290 (2022).

176. van der Ley, P., Bekkers, A., van Meersbergen, J. & Tommassen, J. A comparative study on the *phoE* genes of three enterobacterial species: Implications for structure-function relationships in a pore-forming protein of the outer membrane. *Eur J Biochem* **164**, 469–475 (1987).

177. Hong, H., Patel, D. R., Tamm, L. K. & Berg, B. van den. The outer membrane protein OmpW forms an eight-stranded β-barrel with a hydrophobic channel. *Journal of Biological Chemistry* **281**, 7568–7577 (2006).

178. Villa, L., García-Fernández, A., Fortini, D. & Carattoli, A. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy* **65**, 2518–2529 (2010).

179. Haudiquet, M., Buffet, A., Rendueles, O. & Rocha, E. P. C. Interplay between the cell envelope and mobile genetic elements shapes gene flow in populations of the nosocomial pathogen Klebsiella pneumoniae. *PLoS Biol* **19**, e3001276 (2021).

180. Figurski, D. H., Pohlman, R. F., Bechhofer, D. H., Prince, A. S. & Kelton, C. A. Broad host range plasmid RK2 encodes multiple kil genes potentially lethal to *Escherichia coli* host cells. *Proceedings of the National Academy of Sciences* **79**, 1935–1939 (1982).

181. Bialek-Davenet, S. *et al.* Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg Infect Dis* **20**, 1812–1820 (2014).

182. Bond, P. J., Faraldo-Gómez, J. D. & Sansom, M. S. P. OmpA: A pore or not a pore? Simulation and modeling studies. *Biophys J* **83**, 763–775 (2002).

183. Reusch, R. N. Insights into the structure and assembly of *Escherichia coli* outer membrane protein A. *FEBS Journal* **279**, 894–909 (2012).

184. Xiao, M., Lai, Y., Sun, J., Chen, G. & Yan, A. Transcriptional regulation of the outer membrane porin gene *ompW* reveals its physiological role during the transition from the aerobic to the anaerobic lifestyle of *Escherichia coli*. *Front Microbiol* **7**, (2016).

185. Kortright, K. E., Chan, B. K. & Turner, P. E. High-throughput discovery of phage receptors using transposon insertion sequencing of bacteria. *Proceedings of the National Academy of Sciences* **117**, 18670–18679 (2020).

186. Housden, N. G. *et al.* Directed epitope delivery across the *Escherichia coli* outer membrane through the porin OmpF. *Proc Natl Acad Sci U S A* **107**, 21412–21417 (2010).

187. Choi, U. & Lee, C. R. Distinct roles of outer membrane porins in antibiotic resistance and membrane integrity in *Escherichia coli*. *Front Microbiol* **10**, 953 (2019).

188. Isibasi, A. *et al.* Active protection of mice against *Salmonella typhi* by immunization with strain-specific porins. *Vaccine* **10**, 811–813 (1992).

189. Wu, X. bin *et al.* Outer membrane protein OmpW of *Escherichia coli* is required for resistance to phagocytosis. *Res Microbiol* **164**, 848–855 (2013).

190. Pore, D. & Chakrabarti, M. K. Outer membrane protein A (OmpA) from *Shigella flexneri* 2a: A promising subunit vaccine candidate. *Vaccine* **31**, 3644–3650 (2013).

191. Brambilla, L., Morán-Barrio, J. & Viale, A. M. Expression of the *Escherichia coli ompW* colicin S4 receptor gene is regulated by temperature and modulated by the H-NS and StpA nucleoid-associated proteins. *FEMS Microbiol Lett* **352**, 238–244 (2014).

192. Carranza, G. *et al.* Monitoring bacterial conjugation by optical microscopy. *Front Microbiol* **12**, 2982 (2021).

193. Panicker, M. M. & Minkley, E. G. DNA transfer occurs during a cell surface contact stage of F sex factor-mediated bacterial conjugation. *J Bacteriol* **162**, 584 (1985).

194. Beher, M. G. & Schnaitman, C. A. Regulation of the ompA outer membrane protein of *Escherichia coli*. *J Bacteriol* **147**, 972–985 (1981).

195. Ried, G., Hindennach, I. & Henning, U. Role of lipopolysaccharide in assembly of *Escherichia coli* outer membrane proteins OmpA, OmpC, and OmpF. *J Bacteriol* **172**, 6048–6053 (1990).

196. Pagnout, C. *et al.* Pleiotropic effects of rfa-gene mutations on *Escherichia coli* envelope properties. *Sci Rep* **9**, 1–16 (2019).

197. Llosa, M., Gomis-Ruth, F. X., Coll, M. & Cruz, F. de la. Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45**, 1–8 (2002).

198. Arutyunov, D. & Frost, L. S. F conjugation: Back to the beginning. *Plasmid* **70**, 18–32 (2013).

199. Achtman, M., Kennedy, N. & Skurray, R. Cell--cell interactions in conjugating *Escherichia coli*: role of traT protein in surface exclusion. *Proceedings of the National Academy of Sciences* **74**, 5104 (1977).

200. Riede, I. & Eschbach, M. L. Evidence that TraT interacts with OmpA of *Escherichia coli*. *FEBS Lett* **205**, 241–245 (1986).

201. Helmuth, R. *et al.* Epidemiology of virulence-associated plasmids and outer membrane protein patterns within seven common *Salmonella* serotypes. *Infect Immun* **48**, 175–182 (1985).

202. García-Quintanilla, M. & Casadesús, J. Virulence plasmid interchange between strains ATCC 14028, LT2, and SL1344 of *Salmonella enterica* serovar Typhimurium. *Plasmid* **65**, 169–175 (2011).

203. Feng, Y. *et al.* Inheritance of the *Salmonella* virulence plasmids: Mostly vertical and rarely horizontal. *Infection, Genetics and Evolution* **12**, 1058–1063 (2012).

204. Brouwer, M. S. M. *et al.* The shufflon of IncI1 plasmids is rearranged constantly during different growth conditions. *Plasmid* **102**, 51–55 (2019).

205. Samuels, A. L., Lanka, E. & Davies, J. E. Conjugative junctions in RP4-mediated mating of *Escherichia coli*. *J Bacteriol* **182**, 2709 (2000).

206. Daugelavičius, R., Bamford, J. K. H., Grahn, A. M., Lanka, E. & Bamford, D. H. The IncP plasmid-encoded cell envelope-associated DNA transfer complex increases cell permeability. *J Bacteriol* **179**, 5195 (1997).

207. Bates, S., Cashmore, A. M. & Wilkins, B. M. IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: Involvement of the Tra2 mating system. *J Bacteriol* **180**, 6538–6543 (1998).

208. Bradley, D. E., Taylor, D. E. & Cohen, D. R. Specification of surface mating systems among conjugative drug resistance plasmids in *Escherichia coli* K-12. *J Bacteriol* **143**, 1466–1470 (1980).

209. Licht, T. R., Christensen, B. B., Krogfelt, K. A. & Molin, S. Plasmid transfer in the animal intestine and other dynamic bacterial populations: The role of community structure and environment. *Microbiology (N Y)* **145**, 2615–2622 (1999).

210. Rahube, T. O., Viana, L. S., Koraimann, G. & Yost, C. K. Characterization and comparative analysis of antibiotic resistance plasmids isolated from a wastewater treatment plant. *Front Microbiol* **5**, 558 (2014).

211. Cabezón, E., de la Cruz, F. & Arechaga, I. Conjugation Inhibitors and Their Potential Use to Prevent Dissemination of Antibiotic Resistance Genes in Bacteria. *Front Microbiol* **8**, 2329 (2017).

212. Graf, F. E., Palm, M., Warringer, J. & Farewell, A. Inhibiting conjugation as a tool in the fight against antibiotic resistance. *Drug Dev Res* **80**, 19–23 (2019).

213. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W. & Redinbo, M. R. Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proceedings of the National Academy of Sciences* **104**, 12282–12287 (2007).

214. Nash, R. P., McNamara, D. E., Keith Ballentine, W., Matson, S. W. & Redinbo, M. R. Investigating the impact of bisphosphonates and structurally related compounds on bacteria containing conjugative plasmids. *Biochem Biophys Res Commun* **424**, 697–703 (2012).

215. Palencia-Gándara, C. *et al.* Conjugation inhibitors effectively prevent plasmid transmission in natural environments. *mBio* **12**, (2021).

216. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* (2022) doi:10.1101/2021.10.04.463034.

Appendices

Appendix 1. Plasmids (*n* = 265) encoding $TraN_{pKpQIL}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
NZ_CP022698.1	107249	IncFII(K); IncR;	*Citrobacter farmeri*
NZ_CP026212.1	126940	IncFII(K);	*Citrobacter sp.*
KY093013.1	134521	IncFII(K);	*Enterobacter aerogenes*
KY093014.1	108772	IncFII(K);	*Enterobacter aerogenes*
NZ_LT991960.1	65087	FII(pBK30683);	*Enterobacter cloacae*
NZ_CP018815.1	139942	FIA(pBK30683); FII(pBK30683);	*Enterobacter cloacae*
NZ_CP010363.1	139941	FIA(pBK30683); FII(pBK30683);	*Enterobacter hormaechei*
KU295132.1	101915		*Escherichia coli*
NZ_CP018955.1	101186	IncFIB(pQil); IncFII;	*Escherichia coli*
KU295133.1	115706	IncFIB(pQil); IncFII(K);	*Escherichia coli*
KY798506.1	111742	IncFII(K); IncR;	*Escherichia coli*
NC_025167.1	99142	IncFIB(pQil); IncFII(K);	*Escherichia coli*
NZ_CP014669.1	113852	IncFIB(pQil); IncFII(K);	*Escherichia coli*
NZ_CP018992.1	110226	IncFIB(pQil); IncFII(K);	*Escherichia coli*
NZ_CP026271.1	73495	IncFII(Yp);	*Klebsiella oxytoca*
AP014952.1	209081	IncFIB(K);	*Klebsiella oxytoca*
NZ_CP027425.1	195881	IncFIB(K); IncFII(K);	*Klebsiella oxytoca*
NZ_CP026280.1	152041	IncFII(K);	*Klebsiella oxytoca*
NZ_CP008791.1	205586	IncFII(K); IncFII(Yp);	*Klebsiella oxytoca*
NZ_CP026174.1	93680	FII(pBK30683);	*Klebsiella pneumoniae*
NZ_CP026181.1	94405	FII(pBK30683);	*Klebsiella pneumoniae*
LT882698.1	321455	IncA/C2; IncFIB(pQil);	*Klebsiella pneumoniae*
KY940546.1	88127	IncFII(Yp);	*Klebsiella pneumoniae*
CP025040.1	83541	IncFII(Yp);	*Klebsiella pneumoniae*
JMSX01000005.1	117489	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP006922.1	86232	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP011621.1	77808	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP018670.1	183432	IncFIA(HI1); IncFII;	*Klebsiella pneumoniae*
NZ_CP020064.1	83376	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP021545.1	125536	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP021753.1	116187	IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP021857.1	106331	IncFII(Yp);	*Klebsiella pneumoniae*
CP024547.1	71587	IncFII(Yp);	*Klebsiella pneumoniae*
CP024546.1	110374	FIA(pBK30683); IncFII;	*Klebsiella pneumoniae*
NZ_CP020851.1	161100	IncFIA(HI1); IncFII;	*Klebsiella pneumoniae*
CP008701.1	186323	IncFIB(K); IncFII(K); IncR;	*Klebsiella pneumoniae*
LT576116.1	114208	IncFIA(HI1); IncFII(K);	*Klebsiella pneumoniae*
CP023503.1	135191	IncFIB(pQil); IncFII;	*Klebsiella pneumoniae*
NZ_CP023948.1	219350	IncA/C2; IncFIB(pQil); IncFII;	*Klebsiella pneumoniae*
HF545434.1	133191	IncFII(K);	*Klebsiella pneumoniae*
HQ202266.1	95626	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
FJ876826.1	94219	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP014005.1	101030	IncFII(K);	*Klebsiella pneumoniae*

NZ_CP024876.1	129299	IncFII(K); IncQ1delta;	*Klebsiella pneumoniae*
NZ_CP024917.1	126018	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP025966.2	130691	IncFII(K); IncQ1delta;	*Klebsiella pneumoniae*
NZ_CP026588.1	126149	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028389.2	166034	IncFII(K); IncQ1delta;	*Klebsiella pneumoniae*
NZ_CP028717.1	134869	IncFII(K); IncQ1delta;	*Klebsiella pneumoniae*
CP023488.1	223434	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KP987218.1	105008		*Klebsiella pneumoniae*
NC_023332.1	213019	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012988.3	190072	FIA(pBK30683); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012993.2	190072	FIA(pBK30683); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP017386.1	155781		*Klebsiella pneumoniae*
KJ721789.1	169774	FIA(pBK30683); FII(pBK30683);	*Klebsiella pneumoniae*
CP025039.1	146689	FIA(pBK30683); FII(pBK30683);	*Klebsiella pneumoniae*
NC_025131.1	139941	FIA(pBK30683); FII(pBK30683);	*Klebsiella pneumoniae*
NZ_CP015387.1	139933	FIA(pBK30683); FII(pBK30683);	*Klebsiella pneumoniae*
NZ_CP026396.1	144072	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP027616.1	129985	FIA(pBK30683); FII(pBK30683);	*Klebsiella pneumoniae*
NZ_CP021946.1	120345		*Klebsiella pneumoniae*
NC_019390.1	207819	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP007734.1	338850	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011623.1	250396	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012884.1	194742	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP013323.1	257944	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015132.1	224457	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015135.1	158741	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP026397.1	220406	IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP027613.1	191856	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KX839208.1	187349	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP025516.1	159714	IncFIB(K);	*Klebsiella pneumoniae*
FJ628167.2	151188	IncFII(K); IncR;	*Klebsiella pneumoniae*
LK391770.1	94893	IncFII;	*Klebsiella pneumoniae*
KX636095.1	335317	IncA/C2; IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KT896504.1	186474	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KY495890.1	147442	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP009879.1	178563	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP010390.1	198371	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015383.1	182846	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
JN233704.1	208191	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KJ721790.1	113637	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
LT009689.1	106320	IncFII(K); IncR;	*Klebsiella pneumoniae*
CP012000.1	113626	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
CP025010.1	118202	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
HG969995.1	113642	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
HG969996.1	98184	IncFII(K); IncR;	*Klebsiella pneumoniae*
HG969997.1	114173	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
HG969998.1	103164	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
JNBM01000005.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*

KU665642.1	111926	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
KY271403.1	116499	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
KY798505.1	117903	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
KY798507.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_009649.1	175879	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_009650.1	107576	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_014016.1	113637	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_023903.1	113640	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_023904.1	113622	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_023905.1	116047	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_023906.1	113640	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_025166.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NC_025187.1	114464	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP008830.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP009875.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011986.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011991.1	113638	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP014765.1	106559	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015393.1	202696	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015824.1	103147	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018340.1	81939	IncFIA(HI1); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP019774.1	114073	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022693.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023443.1	117859	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023922.1	129106	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023928.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP024040.1	122082	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP027158.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028177.1	116768	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029102.1	135655	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029740.1	103250	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP030342.1	113639	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_LT216439.1	214718	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP009115.1	118061	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
KT203286.1	211813	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP006800.1	103694	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP016922.1	103694	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP020069.1	213013	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020855.1	221428	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021710.1	214114	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022613.1	102915	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP026752.1	211813	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028930.1	103694	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP020843.1	97202	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021541.1	111531	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021959.1	89382	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012566.1	127690	IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP012571.1	126863	IncFII(K); IncR;	*Klebsiella pneumoniae*

CP025468.1	129684	IncFII(K); IncR;	*Klebsiella pneumoniae*
KX236178.1	121348	IncFII(K);	*Klebsiella pneumoniae*
KY174332.1	137060	IncFII(K); IncR;	*Klebsiella pneumoniae*
NC_016846.1	111195	IncFII(K); IncR;	*Klebsiella pneumoniae*
NC_016966.1	220824	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_023333.1	246176	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_023334.1	203577	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012567.1	196706	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP012572.1	163420	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020903.1	200365	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023916.1	130348	IncFIA(HI1); IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023937.1	279104	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP024459.1	211313	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028805.1	131028	IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP029387.1	226524	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029588.1	216772	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029591.1	219996	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_022609.1	146695	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021166.1	177145	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP025142.1	152230	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP025145.1	149158	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP025148.1	161986	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP026175.1	172259	Col440I; IncFIB(K);	*Klebsiella pneumoniae*
NZ_FO834904.1	95087	IncFIA(HI1); IncFII;	*Klebsiella pneumoniae*
NZ_CP021686.1	183376	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021951.1	181589	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022125.1	187721	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_021199.1	141545	FIA(pBK30683);	*Klebsiella pneumoniae*
NZ_CP029583.1	112141	FIA(pBK30683);	*Klebsiella pneumoniae*
KY271404.1	227989	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
JX283456.1	107748	IncFII(K); IncR;	*Klebsiella pneumoniae*
KP125893.1	133031	IncFII(K);	*Klebsiella pneumoniae*
NZ_LT216438.1	117916	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028804.1	323934	IncA/C2; IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP018351.1	184940	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP018675.1	181436	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP023489.1	124323	IncFII(K);	*Klebsiella pneumoniae*
CP023555.1	169145	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024537.1	147932	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024558.1	147945	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024572.1	147932	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP025006.1	167373	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP025009.1	176049	ColRNAI; IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
JMMY01000002.1	130548	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
JMSW01000002.1	137733	IncFII(K); IncR;	*Klebsiella pneumoniae*
JMSX01000002.1	161408	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
JNBM01000002.1	162535	IncFII(K); IncR;	*Klebsiella pneumoniae*
LT009688.1	166484	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*

NC_020132.1	165295	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_021654.1	245869	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_024992.1	182204	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP008800.1	194877	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP008829.1	243824	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP008930.1	187571	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP009777.1	212192	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP010393.1	207543	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP010574.1	167664	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011577.1	130719	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011977.1	218836	IncFIB(K); IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP011981.1	130552	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP011990.1	162533	IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP014121.1	200203	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015386.1	199497	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015395.1	116419	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015823.1	205221	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018355.1	207543	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018365.1	260772	IncFIB(K); IncFII(K); IncR;	*Klebsiella pneumoniae*
NZ_CP018430.1	208225	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018434.1	208225	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018441.1	180027	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018460.1	232181	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018693.1	180027	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP018886.1	217456	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP019773.1	208528	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020072.1	209424	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020109.1	169952	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020838.1	232540	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020842.1	180210	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020904.1	120533	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021540.1	201874	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP021544.1	208223	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021713.1	208224	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021752.1	209550	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP021834.1	212837	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022145.1	197670	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022574.1	170415	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP022692.1	207349	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023912.1	167814	IncFIB(K);	*Klebsiella pneumoniae*
NZ_CP023943.1	196733	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP023953.1	197671	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP027152.1	208849	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP027161.1	207543	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028181.1	227967	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028543.1	159355	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028784.1	182097	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028995.1	152176	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*

NZ_CP029000.1	125913	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029101.1	208035	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029136.1	190416	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP030343.1	207546	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP006927.1	142788	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP027614.1	100759		*Klebsiella pneumoniae*
NZ_CP015754.1	165275	FIA(pBK30683); IncFIB(K); IncFII;	*Klebsiella pneumoniae*
NZ_CP016160.1	165276	FIA(pBK30683); IncFIB(K); IncFII;	*Klebsiella pneumoniae*
CP024508.1	228353	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024510.1	98344	IncFII(K);	*Klebsiella pneumoniae*
NC_021231.1	113685	IncFII(Yp);	*Klebsiella pneumoniae*
CP024483.1	243634	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024500.1	243620	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP024516.1	227807	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KY271407.1	116325	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NC_019389.1	127508	IncFII(K);	*Klebsiella pneumoniae*
NZ_CP016810.1	307743	IncFIB(K); IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP028954.1	214704	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP029724.1	178561	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
CP025457.1	223274	IncFIB(K); IncFII;	*Klebsiella pneumoniae*
NZ_CP022825.1	97896	IncFIA(HI1); IncFII(K);	*Klebsiella quasivariicola*
NZ_CP022824.1	240771	IncFIB(K);	*Klebsiella quasivariicola*
NZ_CP017281.1	92219	FIA(pBK30683); IncFII(K);	*Klebsiella variicola*
NZ_CP017286.1	92231	FIA(pBK30683); IncFII(K);	*Klebsiella variicola*
NZ_CP017851.1	92232	FIA(pBK30683); IncFII(K);	*Klebsiella variicola*
NZ_CP009275.1	162706	IncFIB(K); IncFII;	*Klebsiella variicola*
NZ_CP028553.1	117244	IncFII(K);	*Klebsiella variicola*
CP014777.1	147881		*Pluralibacter gergoviae*
NZ_CP023893.1	231294	IncFIB(K); IncFII;	*Raoultella ornithinolytica*
NZ_CP026048.1	127313	IncFII(K);	*Raoultella planticola*

Appendix 2. Plasmids (*n* = 166) encoding $TraN_{R100-1}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
KU987452.1	74385	IncFII;	*Citrobacter freundii*
NZ_CP029734.1	76953	IncFII;	*Citrobacter sp.*
CP024284.1	59626	IncFII(29);	*Escherichia albertii*
NZ_CP023358.1	100008	IncFIA(HI1); IncFII(pSE11);	*Escherichia coli*
NZ_CP023355.1	72470	IncFII;	*Escherichia coli*
NC_025177.1	76197	IncFII;	*Escherichia coli*
CP027674.1	133420	IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
NZ_CP010209.1	61005	IncFII(pCoo);	*Escherichia coli*
NZ_CP010215.1	61058	IncFII(pCoo);	*Escherichia coli*
NZ_CP010223.1	61063	IncFII(pCoo);	*Escherichia coli*
AB255435.1	120730	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
JQ418522.1	79478	IncFII;	*Escherichia coli*
NZ_CP015070.1	111851	IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
NZ_CP015077.1	133735	IncFIA; IncFIB(AP001918); IncFII(pRSB107); IncQ1delta;	*Escherichia coli*
NZ_CP016035.1	129085	IncFIA; IncFIB(AP001918); IncFII(pAMA1167-NDM-5);	*Escherichia coli*
NZ_CP018952.1	132345	Col156; IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
NZ_CP019009.1	130905	IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
NZ_CP023816.1	145883	IncFIA; IncFIB(AP001918); IncFII(pAMA1167-NDM-5); IncQ1delta;	*Escherichia coli*
NZ_CP027537.1	100279	IncFIA; IncFII(pRSB107);	*Escherichia coli*
NZ_CP029580.1	178078	Col156; IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
CP027377.1	70152	IncFII(pCoo);	*Escherichia coli*
NZ_CP027367.1	125561	IncFIA; IncFIB(AP001918); IncFII(pHN7A8);	*Escherichia coli*
NZ_CP015239.1	125977	IncFIA; IncFIB(AP001918); IncFII(pHN7A8);	*Escherichia coli*
NZ_CP027343.1	131410	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
JN232517.1	76878	IncFII(pHN7A8);	*Escherichia coli*
KT990220.1	106886	IncFII(pHN7A8); IncX1;	*Escherichia coli*
KX503323.1	74046	IncFII(pHN7A8);	*Escherichia coli*
KX608544.1	66924	IncFII(pHN7A8);	*Escherichia coli*
CP024823.1	64564	IncFII(pHN7A8);	*Escherichia coli*
KJ020575.1	99868	IncFII(pHN7A8); IncN;	*Escherichia coli*
KT879914.1	81498	IncFII(pHN7A8);	*Escherichia coli*
KU321583.1	138718	IncFII(pHN7A8); IncN; IncX1;	*Escherichia coli*
KY865321.1	92545	IncFII(pHN7A8);	*Escherichia coli*
KY865322.1	77822	IncFII(pHN7A8);	*Escherichia coli*
NZ_CP017633.1	120528	IncFII(pHN7A8); IncN; IncX1;	*Escherichia coli*
NZ_CP017981.1	91451	IncFII(pHN7A8);	*Escherichia coli*
NZ_CP025949.1	73313	IncFII(pHN7A8);	*Escherichia coli*

NZ_CP026724.1	129627	IncFIB(pB171); IncFII(pRSB107);	*Escherichia coli*
CM004378.1	131906	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP027375.1	118863	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP027458.1	107796	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP027581.1	118822	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
CP027641.1	126957	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NC_013362.1	63365	IncFII(29);	*Escherichia coli*
NZ_CP027314.1	59928	IncFII(29);	*Escherichia coli*
CP027765.1	66388	IncFII(pSE11);	*Escherichia coli*
NZ_CP027455.1	73262	IncFII(pCoo);	*Escherichia coli*
NZ_CP012501.1	242187	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP029577.1	139191	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
JN087528.1	70382	IncFII;	*Escherichia coli*
NC_017630.1	114550	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
CP024888.1	76754	IncFII;	*Escherichia coli*
NZ_CP014112.1	69183	IncFII;	*Escherichia coli*
NZ_CP018107.1	61085	IncFII(29);	*Escherichia coli*
NZ_CP018113.1	62157	IncFII(29);	*Escherichia coli*
NZ_CP014668.1	101201	IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP027554.1	70129	IncFII(pHN7A8);	*Escherichia coli*
CP024242.1	74644	IncFII(pCoo);	*Escherichia coli*
NC_017627.1	113346	IncFIC(FII); IncFII(pHN7A8);	*Escherichia coli*
AP014876.1	71214	IncFII;	*Escherichia coli*
DQ364638.1	94289	IncFII;	*Escherichia coli*
FJ449539.1	72946	IncFII;	*Escherichia coli*
HM355591.2	70262	IncFII;	*Escherichia coli*
AP018136.1	84133	IncFII;	*Escherichia coli*
AP018144.1	94643	IncFII;	*Escherichia coli*
CP000913.1	91019	IncFII;	*Escherichia coli*
CP007652.1	66289	IncFII;	*Escherichia coli*
CP009233.1	154789	Col156; IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
JN087529.1	69812	IncFII;	*Escherichia coli*
KR078259.1	80206	IncFII;	*Escherichia coli*
KU288634.1	78962	IncFII;	*Escherichia coli*
KU932024.1	71656	IncFII;	*Escherichia coli*
NC_005327.1	92353	IncFII;	*Escherichia coli*
NC_019090.1	73607	IncFII;	*Escherichia coli*
NC_019095.1	76635	IncFII;	*Escherichia coli*
NC_019424.1	69768	IncFII;	*Escherichia coli*
NC_020278.2	76626	IncFII;	*Escherichia coli*
NZ_CP010174.1	106274	IncFII; IncR;	*Escherichia coli*
NZ_CP010239.1	184614	Col156; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP011065.1	82288	IncFII;	*Escherichia coli*
NZ_CP012632.1	93074	IncFII;	*Escherichia coli*

NZ_CP013657.1	143748	IncFIA; IncFII;	*Escherichia coli*
NZ_CP019910.1	68786	IncFII; IncN;	*Escherichia coli*
NZ_CP021203.1	183508	IncFII; p0111;	*Escherichia coli*
NZ_CP023895.1	91648	IncFII;	*Escherichia coli*
NZ_CP027385.1	118259	IncFII(pRSB107);	*Escherichia coli*
NZ_CP029974.1	71718	IncFII;	*Escherichia coli*
NZ_HG941719.1	135602	IncFIA; IncFII;	*Escherichia coli*
NC_016039.1	70060	IncFII;	*Escherichia coli*
NZ_CP018973.1	70152	IncFII;	*Escherichia coli*
CM007889.1	114233	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NC_007941.1	114230	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NC_011749.1	122301	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NC_013175.1	114222	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NC_013655.1	122345	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP007150.1	114231	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP012634.1	114221	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP014110.1	115764	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP030770.1	149680	Col156; IncFIB(AP001918); IncFII(29); IncQ1delta;	*Escherichia coli*
CP024295.1	103995	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP016498.1	151033	Col156; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP021289.1	147172	Col156; IncFIB(AP001918); IncFII;	*Escherichia coli*
KU932028.1	80057	IncFII;	*Escherichia coli*
CM007909.1	73943	IncFII(29);	*Escherichia coli*
CP024133.1	190293	IncFIB(AP001918); IncFII; IncR;	*Escherichia coli*
NZ_CP027596.1	87855	IncFII(pRSB107);	*Escherichia coli*
NZ_CP026579.1	100229	IncFIB(AP001918); IncFII;	*Escherichia coli*
NC_010409.1	151002	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP010192.1	162720	IncFIB(AP001918); IncFII;	*Escherichia coli*
KU695535.1	70093	IncFII;	*Escherichia coli*
NZ_CP012498.1	213847	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NC_019037.1	82676	IncFII;	*Escherichia coli*
NZ_CP009579.1	141804	IncFIB(AP001918); IncQ1delta;	*Escherichia coli*
NZ_CP018958.1	53198	IncFII(pSE11);	*Escherichia coli*
CP024652.2	107274	IncFII;	*Escherichia coli*
NZ_CP025253.1	105801	IncFII;	*Escherichia coli*
NZ_CP021684.1	84929	IncFII;	*Escherichia coli*
CP024157.1	88736	IncFII;	*Escherichia coli*
NZ_CP027469.1	62881	IncFII(pCoo);	*Escherichia coli*
NZ_CP012499.1	223952	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP012500.1	225292	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
CP014496.1	114223	Col156; IncFIB(AP001918); IncFII(29);	*Escherichia coli*

NZ_CP015160.1	212180	Col156; IncFIB(AP001918); IncFII; IncFII(29);	*Escherichia coli*
KX023260.1	85126	IncFII(pHN7A8);	*Escherichia coli*
FJ876827.1	147416	Col156; IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Klebsiella pneumoniae*
KX928752.1	216976	IncFIA; IncFIB(AP001918); IncFII(pRSB107); IncY;	*Klebsiella pneumoniae*
LN897474.2	76863	IncFII(pHN7A8);	*Klebsiella pneumoniae*
LN897475.2	74768	IncFII(pHN7A8);	*Klebsiella pneumoniae*
KT185451.1	151466	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP018455.1	162552	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP023942.1	187926	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP025952.1	170821	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP026584.1	156099	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP028541.2	177516	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
NZ_CP029381.1	146790	IncFII(pHN7A8); IncR;	*Klebsiella pneumoniae*
KU318420.1	71141	IncFII(pCoo);	*Klebsiella pneumoniae*
KX839209.1	95701	IncFII;	*Klebsiella pneumoniae*
KY751926.1	110970	IncFII; IncR;	*Klebsiella pneumoniae*
CP024836.1	96185	IncFII;	*Klebsiella pneumoniae*
CP024840.1	96185	IncFII;	*Klebsiella pneumoniae*
KT725788.1	95005	IncFII;	*Klebsiella pneumoniae*
KT725789.1	89745	IncFII;	*Klebsiella pneumoniae*
KT818627.1	172679	IncFII; IncR;	*Klebsiella pneumoniae*
KY130431.1	96549	IncFII;	*Klebsiella pneumoniae*
KY288024.1	82841	IncFII;	*Klebsiella pneumoniae*
NZ_CP015388.1	89067	IncFII;	*Klebsiella pneumoniae*
NZ_CP015503.1	84941	IncFII;	*Klebsiella pneumoniae*
NZ_CP021711.1	78638	IncFII;	*Klebsiella pneumoniae*
NZ_CP024430.1	131243	IncFII; IncR;	*Klebsiella pneumoniae*
NZ_LN824138.1	84940	IncFII;	*Klebsiella pneumoniae*
NC_013542.1	70057	IncFII;	*Klebsiella pneumoniae*
NZ_CP015131.1	142858	Col156; IncFIB(AP001918); IncFII(29);	*Klebsiella pneumoniae*
NZ_CP017721.1	280421	IncFII(S);	*Salmonella enterica*
NZ_CP016389.1	68117	IncFII;	*Salmonella enterica*
NZ_CP019444.1	84565	IncFII;	*Salmonella enterica*
KX646543.1	111559	IncFII;	*Shigella boydii*
KT754164.1	111227	IncFIA; IncFII(pRSB107);	*Shigella dysenteriae*
CP012139.1	86407	IncFII;	*Shigella flexneri*
KJ201887.1	75335	IncFII;	*Shigella flexneri*
NZ_CP020340.1	73096	IncFII;	*Shigella flexneri*
NC_002134.1	94281	IncFII;	*Shigella flexneri*
CP024475.1	82833	IncFII(pSE11);	*Shigella flexneri*
LN624486.1	73047	IncFII(pSE11);	*Shigella flexneri*
KX008967.1	88976	IncFII;	*Shigella sonnei*
NC_013727.1	70275	IncFII;	*Shigella sonnei*
LT174531.1	68999	IncFII(pSE11);	*Shigella sonnei*

Appendix 3. Plasmids (*n* = 178) encoding $TraN_F$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
CP024287.1	124142	IncFII;	*Escherichia albertii*
CP024286.1	113727	IncFIB(AP001918); IncFII;	*Escherichia albertii*
NZ_CP025329.1	286854	IncFIB(AP001918); IncFIC(FII); IncFII(pHN7A8);	*Escherichia coli*
CP024296.1	97297	IncFII(pCoo);	*Escherichia coli*
CP006001.1	66341	IncFII(29);	*Escherichia coli*
NZ_CP023351.1	77345	IncFII(29); IncFII(pCoo);	*Escherichia coli*
NZ_CP024662.1	160351	IncFIB(AP001918); IncFII(29);	*Escherichia coli*
AY545598.5	184501	IncFIB(AP001918); IncFII;	*Escherichia coli*
KT845955.1	108837	IncFII(pSE11); IncR;	*Escherichia coli*
KY865320.1	67077	IncFII(pCoo);	*Escherichia coli*
AY214164.3	101375	IncFII(pRSB107);	*Escherichia coli*
NZ_CP011916.1	110961	IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
NZ_CP024975.1	86835	IncFII(29);	*Escherichia coli*
CP025626.1	144225	IncFIA; IncFII;	*Escherichia coli*
KU043115.1	124378	IncFIA; IncFII;	*Escherichia coli*
NZ_CP011064.1	94712	IncFIA; IncFII;	*Escherichia coli*
NZ_CP019076.1	127772	IncFIA; IncFII;	*Escherichia coli*
NZ_CP021737.1	117703	IncFIA; IncFII;	*Escherichia coli*
NZ_CP023871.1	149485	IncFIA; IncFII;	*Escherichia coli*
NZ_CP029113.1	114412	IncFIA; IncFII;	*Escherichia coli*
CP024258.1	96607	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
CM007715.1	67502	IncFIA; IncFIC(FII);	*Escherichia coli*
CM007847.1	67493	IncFIA; IncFIC(FII);	*Escherichia coli*
NC_002483.1	99159	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP011496.1	67545	IncFIA; IncFIC(FII);	*Escherichia coli*
NZ_CP014271.1	232954	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP014273.1	213924	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP027256.1	103336	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP019001.1	272202	IncA/C2; IncFIA; IncFII;	*Escherichia coli*
KU355873.1	141534	IncFIA; IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
CP009231.1	155456	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
CP009232.1	172280	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
CP012197.1	185134	Col156; IncFIA; IncFIB(AP001918); IncFII; IncFII(pSE11);	*Escherichia coli*
CP017221.1	147225	IncFIA; IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
CP024852.1	167974	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
CP024856.1	181436	Col156; IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
CP024860.1	172588	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*

FQ482074.1	165657	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
KP789020.1	138692	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
LN850163.1	167198	IncFIA; IncFIB(AP001918); IncFII; IncFII(pAMA1167-NDM-5);	*Escherichia coli*
LO017736.1	124392	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
LO017737.1	175264	Col156; IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
LO017738.1	143225	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP009167.1	108501	IncFIA; IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP010372.1	151583	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP013027.1	174564	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP018990.1	150994	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP021733.1	114267	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP021880.1	169449	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP022227.1	150823	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP023354.1	62312	IncFII;	*Escherichia coli*
NZ_CP023372.1	109207	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP026754.1	90045	IncFIA(HI1); IncFII(29);	*Escherichia coli*
NZ_CP027130.1	170292	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP027703.1	173882	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP028587.1	150853	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
AP017611.1	168972	IncFIB(AP001918); IncFII; IncX1;	*Escherichia coli*
DQ381420.1	174241	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
KU932025.1	143590	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NC_004998.1	125491	IncFIB(AP001918); IncFII;	*Escherichia coli*
NC_024956.1	94587	IncFII;	*Escherichia coli*
NC_025139.1	144925	IncFIB(AP001918); IncFII;	*Escherichia coli*
NC_025175.1	166594	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NZ_CP010158.1	128612	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP019007.1	128248	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP027357.1	131463	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP029744.1	117192	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP027448.1	170848	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
CP024816.1	176274	IncFIA; IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
CP024831.1	216181	IncFIA; IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
NC_010558.1	168113	IncFIA; IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
NZ_CP021204.1	92438	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP023378.1	123289	IncFIA; IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NZ_CP025708.1	184098	IncFIA; IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
NZ_CP028193.1	52918	IncFII(pSE11);	*Escherichia coli*
NC_014384.1	118525	IncFIA; IncFII;	*Escherichia coli*
NZ_CP010138.1	83922	IncFII;	*Escherichia coli*

NZ_CP011493.1	140232	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP013832.1	145221	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
CP024249.1	167230	IncFIB(AP001918); IncFII(29);	*Escherichia coli*
NZ_CP029104.1	76919	IncFII;	*Escherichia coli*
NZ_CP019561.1	148529	IncFII;	*Escherichia coli*
KU254579.1	68812	IncFII(29);	*Escherichia coli*
NC_014382.1	73801	IncFII;	*Escherichia coli*
AP018140.1	81783	IncFII(pHN7A8);	*Escherichia coli*
NC_023315.1	85507	IncFII;	*Escherichia coli*
NZ_CP021843.1	78672	IncFII;	*Escherichia coli*
NZ_CP021847.1	84091	IncFII;	*Escherichia coli*
NC_017640.1	81475	IncFIB(AP001918);	*Escherichia coli*
CP024145.1	106478	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NC_010719.1	225683	IncFII;	*Escherichia coli*
KR653209.1	351717	IncFII; IncHI2; IncHI2A;	*Escherichia coli*
AP018139.1	90294	IncFII;	*Escherichia coli*
AP018147.1	102071	IncFII;	*Escherichia coli*
JQ364967.1	87021	IncFII;	*Escherichia coli*
KT002541.1	86015	IncFII;	*Escherichia coli*
KX246267.1	83435	IncFII;	*Escherichia coli*
KY463220.1	104353	IncFII;	*Escherichia coli*
NC_010488.1	130440	IncFIB(AP001918); IncFII;	*Escherichia coli*
NC_025106.1	87619	IncFII;	*Escherichia coli*
NZ_CP010232.1	155425	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP026201.1	100989	IncFII;	*Escherichia coli*
NZ_CP029243.1	146268	IncFIB(AP001918); IncFII;	*Escherichia coli*
LT838203.1	60622	IncFII(pCoo);	*Escherichia coli*
KY416992.1	114916	IncFII(pCoo);	*Escherichia coli*
NZ_CP013026.1	63800	IncFII(pCoo);	*Escherichia coli*
NZ_CP015914.1	129035	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP023367.1	111523	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP005931.1	194170	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP023384.1	147369	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP026494.1	201203	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
DQ388534.1	101558	IncFIB(pB171); IncFII(pCoo);	*Escherichia coli*
NC_011603.1	97978	IncFIB(pB171); IncFII(pCoo);	*Escherichia coli*
NZ_CP010345.1	100061	IncFIA; IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
KX276657.1	225069	IncFIB(AP001918); IncFIC(FII); IncN;	*Escherichia coli*
NZ_CP029748.1	276880	IncFIB(AP001918); IncFIC(FII); IncN;	*Escherichia coli*
CP024140.1	166233	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
CP024827.1	122937	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
KR827684.1	144344	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP006635.1	161511	IncFIB(AP001918); IncFIC(FII); IncQ1delta;	*Escherichia coli*
NZ_CP010141.1	174041	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP010149.1	199494	IncFIB(AP001918); IncFIC(FII); IncR;	*Escherichia coli*

NZ_CP020934.1	151137	IncFIB(AP001918); IncFIC(FII); IncQ1delta;	*Escherichia coli*
NC_012944.1	138362	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP012626.1	110808	IncFIB(AP001918); IncFII;	*Escherichia coli*
CP021198.1	157585	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
CP031107.1	203856	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NC_025179.1	153231	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP019018.1	158911	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP014320.1	130603	IncFIB(pB171); IncFII;	*Escherichia coli*
NZ_CP023959.1	128761	IncFIB(pB171); IncFII;	*Escherichia coli*
NZ_CP030329.1	128762	IncFIB(pB171); IncFII;	*Escherichia coli*
NZ_CP013833.1	126302	IncFIA; IncFII;	*Escherichia coli*
NZ_CP018978.1	117844	IncFIA; IncFII;	*Escherichia coli*
CP027485.1	147394	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
CP012113.1	132464	IncFIB(AP001918); IncFII;	*Escherichia coli*
CP014489.1	149732	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
CP025403.1	133281	IncFIB(AP001918); IncFII;	*Escherichia coli*
CU928146.1	133853	IncFIB(AP001918); IncFII;	*Escherichia coli*
KP398867.1	147837	IncFIB(AP001918); IncFII;	*Escherichia coli*
KU578032.1	143812	IncFIB(AP001918); IncFII;	*Escherichia coli*
KU664810.1	142359	IncFIB(AP001918); IncFII;	*Escherichia coli*
KY007017.1	135661	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP012636.1	149683	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NZ_CP018994.1	154853	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NZ_CP023363.1	134388	IncFIB(AP001918); IncFII; IncQ1delta;	*Escherichia coli*
NZ_CP026854.1	133843	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010316.1	144859	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP027439.1	107188	IncFII(pHN7A8);	*Escherichia coli*
CP024151.1	97858	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP019909.1	85009	IncFIC(FII);	*Escherichia coli*
NC_009602.1	121239	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP023389.1	74094	IncFII;	*Escherichia coli*
NZ_CP019456.1	159821	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010123.1	214445	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP026403.1	103161	IncFII;	*Escherichia coli*
NZ_CP027551.1	162810	IncFIB(AP001918); IncFII(pHN7A8); IncFII(pRSB107);	*Escherichia coli*
NZ_CP019028.1	147346	IncFIA; IncFIB(AP001918); IncFII; IncFII(pCoo);	*Escherichia coli*
NC_010720.1	75089	IncFII;	*Escherichia coli*
NZ_CP010126.1	143617	IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
AP014877.1	65140		*Escherichia coli*
NZ_CP029575.1	156518	IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*

NZ_CP027311.1	133438	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP021536.1	244955	IncA/C2; IncFIA; IncFIB(AP001918); IncFII;	*Escherichia coli*
CP024238.1	274465	IncFII(pCoo); IncFII(pSE11);	*Escherichia coli*
NZ_CP015836.1	136327	IncFIB(pB171); IncFII;	*Escherichia coli*
KP453775.1	267645	IncFIA; IncFII; IncFII(pCoo); IncY;	*Klebsiella pneumoniae*
NZ_CP017726.1	58302	IncFII(pCoo);	*Salmonella enterica*
NC_019117.1	117278	IncFIA; IncFIB(AP001918); IncFII(pCoo);	*Salmonella enterica*
NC_019122.1	120524	IncFIA; IncFIB(AP001918); IncFII(pCoo);	*Salmonella enterica*
KY749247.1	97566	IncFII;	*Salmonella enterica*
CP019190.1	186451	IncFII(S);	*Salmonella enterica*
NZ_CP012928.1	62920	IncFII;	*Salmonella enterica*
NZ_CP012683.1	112639	IncFIB(AP001918); IncFIC(FII);	*Salmonella enterica*
NC_011076.1	146811	IncFIB(AP001918); IncFII;	*Salmonella enterica*
NC_018954.1	146811	IncFIB(AP001918); IncFII;	*Salmonella enterica*
NC_018966.1	146002	IncFIB(AP001918); IncFII;	*Salmonella enterica*
KT754160.1	129841	IncFIA; IncFIB(pB171); IncFIC(FII);	*Shigella dysenteriae*

Appendix 4. Plasmids (*n* = 58) encoding $TraN_{pSLT}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species	Serovar
HF969016.1	93046	IncFIB(S); IncFII(S);	*Salmonella enterica*	Bovismorbificans
NZ_CP012345.1	119113	IncFIB(S); IncFII(S);	*Salmonella enterica*	Choleraesuis
AY517905.1	80156	IncFII(S);	*Salmonella enterica*	Dublin
NZ_CP018658.1	94039	IncFIB(S); IncFII(S);	*Salmonella enterica*	Enteritidis
NZ_CP018634.1	92831	IncFIB(S); IncFII(S);	*Salmonella enterica*	Enteritidis
CM001154.1	87371	IncFII(S);	*Salmonella enterica*	Gallinarum
HG970001.1	88350	IncFII(S);	*Salmonella enterica*	Gallinarum
NZ_CP030176.1	148530	IncFII(S);	*Salmonella enterica*	Milwaukee
CP015599.1	94046	IncFIB(S); IncFII(S);	*Salmonella enterica*	Not defined
NZ_CP020923.1	116677	IncFIB(S); IncFII(S);	*Salmonella enterica*	Not defined
NZ_CP030198.1	134274		*Salmonella enterica*	Not defined
NC_019112.1	86686	IncFII(S);	*Salmonella enterica*	Pullorum
NZ_CP012348.1	86650	IncFII(S);	*Salmonella enterica*	Pullorum
NZ_CP022137.1	66457	IncFII(S);	*Salmonella enterica*	subsp *diarizonae*
CP030027.1	169096		*Salmonella enterica*	subsp *diarizonae*
NZ_CP017729.1	93826	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
AB605179.1	132842	IncFIB(AP001918); IncFII(S);	*Salmonella enterica*	Typhimurium
AE006471.2	93933	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
KX777254.1	93807	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
LN794247.1	128925	IncFIB(S); IncFII(S); IncQ1delta;	*Salmonella enterica*	Typhimurium
LN999012.1	93844	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
CP014962.1	93850	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
CP020113.1	140518	IncFIB(AP001918); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_013437.1	117047	IncFIB(S); IncFII(S); IncQ1delta;	*Salmonella enterica*	Typhimurium
NC_014476.2	112670	IncFII(S);	*Salmonella enterica*	Typhimurium
NC_016855.1	93832	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_016861.1	106510	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_016864.1	93277	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_019108.1	112673	IncFII(S);	*Salmonella enterica*	Typhimurium
NC_019109.1	133216	IncFIB(AP001918); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_021155.1	148711	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_022570.1	94034	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_AP014566.1	132611	IncFIB(AP001918); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP007489.1	93965	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP007582.1	91739	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP008745.1	93833	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP013721.1	93925	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014050.2	93933	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014357.1	93837	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014537.1	93844	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium

NZ_CP014577.1	93930	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014968.1	94016	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014970.1	94014	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014973.1	93960	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014976.1	94019	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP014980.1	94034	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP015158.1	93829	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP016390.1	93855	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP021464.1	115864	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP025556.1	93832	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP026701.1	93832	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP027413.1	93832	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP028200.1	93933	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP029596.1	94001	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_LT855377.1	93862	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_016858.1	93842	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NC_017054.1	93877	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium
NZ_CP017618.1	93055	IncFIB(S); IncFII(S);	*Salmonella enterica*	Typhimurium

Appendix 5. Plasmids (*n* = 29) encoding $TraN_{MV1}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
NZ_CP022697.1	149207	IncFIB(pB171); IncFII(Yp);	*Citrobacter farmeri*
NZ_CP016763.1	127005	IncFIB(K); IncFII(Yp);	*Citrobacter freundii*
NZ_CP026241.1	200651	IncFIB(pB171); IncFII(Yp);	*Citrobacter freundii*
NZ_CP026698.1	82058	IncFIB(pB171); IncFII(Yp);	*Citrobacter koseri*
NZ_CP029731.1	108148	IncFIB(pB171); IncFII(Yp);	*Citrobacter sp.*
KU318419.1	90351	IncFII(Yp);	*Enterbacter aerogenes*
CM008904.1	106482	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
CM008911.1	110766	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KC887917.2	99435	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KJ812998.1	110786	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KP868647.1	109353	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KY399974.1	109361	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KY399975.1	109325	IncFIB(pB171); IncFII(Yp);	*Enterobacter cloacae*
KC887916.2	110786	IncFIB(pB171); IncFII(Yp);	*Escherichia coli*
NC_021501.1	110781	IncFIB(pB171); IncFII(Yp);	*Klebsiella michiganensis*
NZ_CP022350.1	106140	IncFIB(pB171); IncFII(Yp);	*Klebsiella michiganensis*
NZ_CP023187.1	106579	IncFIB(pB171); IncFII(Yp);	*Klebsiella michiganensis*
CM008882.1	110432	IncFIB(pB171); IncFII(Yp);	*Klebsiella pneumoniae*
CP023554.1	212326	IncFIB(pB171); IncFII(Yp); IncR;	*Klebsiella pneumoniae*
KR351290.1	106844	IncFIB(S); IncFII(Yp);	*Klebsiella pneumoniae*
NC_025184.1	110786	IncFIB(pB171); IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP018366.1	94434	IncFIB(pB171); IncFII(Yp);	*Klebsiella pneumoniae*
NZ_CP023914.1	106853	IncFIB(pB171); IncFII(Yp);	*Klebsiella pneumoniae*
KU987453.1	85862	IncFII(Yp);	*Klebsiella pneumoniae*
KY020154.1	142876	IncFIB(K); IncFII(Yp);	*Klebsiella pneumoniae*
CM008891.1	120078	IncFIB(pB171); IncFII(Yp);	*Klebsiella variicola*
CM008917.1	120406	IncFIB(pB171); IncFII(Yp);	*Klebsiella variicola*
KX868552.1	77843		*Klebsiella variicola*
KP900016.1	99276	IncFIB(pB171); IncFII(Yp);	*Leclercia adecarboxylata*

Appendix 6. Plasmids (*n* = 15) encoding $TraN_{MV2}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
NZ_CP022275.1	144825	FIA(pBK30683); IncFII(Yp);	*Citrobacter freundii*
KP868646.1	88214	IncFII(Yp);	*Enterobacter cloacae*
KY399972.1	88213	IncFII(Yp);	*Enterobacter cloacae*
KY399973.1	88213	IncFII(Yp);	*Enterobacter cloacae*
NZ_CP022349.1	126877	FIA(pBK30683); IncFII(Yp);	*Klebsiella michiganensis*
NZ_CP023186.1	126846	FIA(pBK30683); IncFII(Yp);	*Klebsiella michiganensis*
AP014954.1	81481		*Klebsiella oxytoca*
NZ_CP020359.1	110705	IncFII(Yp);	*Klebsiella oxytoca*
NZ_CP018361.1	277616	FIA(pBK30683); IncFIB(K);	*Klebsiella oxytoca*
NZ_CP008789.1	133397	IncFIA(HI1);	*Klebsiella oxytoca*
NZ_CP026272.1	144055		*Klebsiella oxytoca*
CP024509.1	140704		*Klebsiella pneumoniae*
NZ_CP028554.1	89738	IncFII(Yp);	*Klebsiella variicola*
NZ_CP021328.1	446611	IncFIB(pQil); IncFII(Yp); IncU;	*Raoultella ornithinolytica*
NZ_CP023890.1	89860	IncFII(Yp);	*Raoultella ornithinolytica*

Appendix 7. Plasmids (*n* = 10) encoding $TraN_{MV3}$.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
NZ_CP030780.1	80627	IncFII(pHN7A8);	*Escherichia albertii*
CP024250.1	99099	IncFII(pCoo);	*Escherichia coli*
NZ_CP024977.1	62669		*Escherichia coli*
NC_022333.1	165311	IncFII(pCoo);	*Escherichia coli*
NZ_CP023348.1	88757	IncFII(pCoo);	*Escherichia coli*
NZ_CP024669.1	110612	IncFII(pCoo);	*Escherichia coli*
NZ_CP021212.1	90229	IncFII(pSE11);	*Escherichia coli*
CP024244.1	152012	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
KT754161.1	83692	IncFII(pSE11);	*Shigella dysenteriae*
NZ_CP020338.1	95633	IncFII(pHN7A8);	*Shigella flexneri*

Appendix 8. Putative conjugative IncF plasmids (*n* = 103) encoding other TraNs.

Plasmid Accession ID	Size (bp)	Replicons (PlasmidFinder)	Host Species
NZ_CP017614.1	98627		*Candidatus Hamiltonella*
NZ_CP017608.1	72453		*Candidatus Hamiltonella*
NZ_CP017612.1	67631		*Candidatus Hamiltonella*
NZ_CP011610.1	135117	IncFIB(pB171); IncFII(S);	*Citrobacter freundii*
NZ_CP026699.1	42930	IncFII(Yp);	*Citrobacter koseri*
NZ_CP029732.1	68384	IncFII(Yp);	*Citrobacter sp.*
NC_015963.1	166725	IncFIB(pENTAS01); IncFII(pENTA);	*Enterobacter asburiae*
KX912253.1	108672	IncFII(pECLA); IncR;	*Enterobacter asburiae*
NZ_CP029719.1	138933	IncFIB(pECLA); IncFII(pECLA);	*Enterobacter cloacae*
NZ_CP021777.1	70882		*Enterobacter cloacae*
NZ_CP017187.1	131604	IncFII(pECLA);	*Enterobacter cloacae*
NC_014107.1	199562	IncFIB(pECLA); IncFII(pECLA);	*Enterobacter cloacae*
CP017992.1	79045	IncFII(SARC14); IncFII(p14);	*Enterobacter cloacae*
NZ_CP017473.1	169226	IncFII(Yp);	*Enterobacter cloacae*
KX786187.1	165469	IncFII(Yp);	*Enterobacter cloacae*
CP017991.1	144025	IncFIB(pENTAS01); IncFII(Yp);	*Enterobacter cloacae*
KX858825.1	89970	IncFII(pECLA);	*Enterobacter cloacae*
CP017413.1	85398		*Enterobacter cloacae*
NC_014108.1	84653		*Enterobacter cloacae*
NZ_LT991959.1	142732		*Enterobacter cloacae*
NZ_CP017185.1	151583	IncFII(Yp);	*Enterobacter cloacae*
NZ_CP012169.1	210894	IncFIB(pECLA); IncFII(pECLA);	*Enterobacter hormaechei*
NC_021492.1	116007		*Enterobacter sp.*
NC_009425.1	157749	IncFIB(pENTE01);	*Enterobacter sp.*
NZ_CP030785.1	81110	IncFII(pHN7A8);	*Escherichia albertii*
CP025677.1	129356	IncFII(pHN7A8); IncN; IncX1;	*Escherichia albertii*
NZ_CP010181.1	201930	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP010184.1	200925	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP011019.1	207265	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
NZ_CP010187.1	146575	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010197.1	146524	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010214.1	146496	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010222.1	146552	IncFIB(AP001918); IncFII;	*Escherichia coli*
NC_011413.1	91158	IncFII(pSE11);	*Escherichia coli*
NZ_CP010124.1	86032	IncFII;	*Escherichia coli*
CP024864.1	88768	IncFII(pSE11);	*Escherichia coli*
NZ_CP027106.1	106432	IncFIB(AP001918); IncFII(pHN7A8);	*Escherichia coli*
NZ_CP030922.1	77405		*Escherichia coli*
AP014804.1	146197	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*
CP013836.1	140502	IncFIA; IncFIB(pB171); IncFII;	*Escherichia coli*
CP024281.1	84894	IncFII;	*Escherichia coli*
LT906556.1	157174	IncFIB(AP001918); IncFII(pCoo);	*Escherichia coli*

NZ_CP010882.1	122966	Col156; IncFIA; IncFIB(AP001918); IncFII(pRSB107);	*Escherichia coli*
NZ_CP021210.1	75553	IncFII;	*Escherichia coli*
NZ_CP027536.1	72482	IncFII;	*Escherichia coli*
NZ_CP029578.1	70197	IncFII(pSE11);	*Escherichia coli*
CP012377.1	125121	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP022156.1	113793	IncFII(S); IncFII(SARC14);	*Escherichia coli*
NZ_CP023369.1	66332	IncFII(pRSB107);	*Escherichia coli*
NZ_CP027545.1	101089	IncFIB(AP001918);	*Escherichia coli*
NZ_CP022166.1	173624	IncFIA; IncFIB(AP001918); IncFIC(FII);	*Escherichia coli*
NZ_CP019904.1	108340	IncFII(Yp);	*Escherichia coli*
NZ_CP027580.1	54544	IncFII(pHN7A8);	*Escherichia coli*
NZ_CP028383.1	67750	IncFII(pHN7A8);	*Escherichia coli*
CP024291.1	111697	IncFIB(AP001918); IncFII;	*Escherichia coli*
NZ_CP010201.1	116581	IncFIB(AP001918); IncFII(pSE11);	*Escherichia coli*
NZ_CP028952.1	451422	IncFIB(K);	*Klebsiella aerogenes*
NZ_CP008842.1	204468		*Klebsiella michiganensis*
NZ_CP008843.1	183015	IncFII(Yp);	*Klebsiella michiganensis*
NZ_CP011615.1	113105	IncFII(S); IncFII(SARC14);	*Klebsiella oxytoca*
NZ_CP026276.1	282476	IncFIA(HI1); IncFIB(K); IncFII(K);	*Klebsiella oxytoca*
NZ_CP026273.1	179680		*Klebsiella oxytoca*
NZ_CP011633.1	150318	IncFIB(K); IncFII(Yp);	*Klebsiella oxytoca*
NZ_CP007736.1	113440	IncFIA(HI1); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP015396.1	110092	ColRNAI; IncFIA(HI1); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP020062.1	152697	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP026179.1	237571	IncFIA(HI1); IncFII(K); IncN;	*Klebsiella pneumoniae*
JN233705.2	115300	IncFIB(pQil); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP017387.1	225962	IncA/C2; IncFIB(pKPHS1);	*Klebsiella pneumoniae*
NZ_CP026172.1	243967	IncFIB(Mar); IncHI1B;	*Klebsiella pneumoniae*
CP025038.1	199686	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
KY271406.1	212319	IncFIB(K); IncFII(K);	*Klebsiella pneumoniae*
NZ_CP017283.1	108867		*Klebsiella variicola*
NZ_CP017285.1	108857		*Klebsiella variicola*
NZ_CP017850.1	108965		*Klebsiella variicola*
CM008916.1	216150	IncFIB(Mar);	*Klebsiella variicola*
NZ_CP011600.1	215092	IncFIB(K); IncFII(pCRY);	*Kluyvera intermedia*
NZ_CP019446.1	113443		*Kosakonia cowanii*
NZ_CP028521.1	126628		*Lelliottia sp.*
KT225520.1	166620	IncFII(Yp);	*Raoultella ornithinolytica*
CP013339.1	152010	IncFIB(K);	*Raoultella ornithinolytica*
NZ_CP026049.1	206855		*Raoultella planticola*
NC_021871.1	89986	IncFIB(pB171);	*Salmonella bongori*
NZ_CP022035.1	136771	IncFII(p14);	*Salmonella enterica*
NZ_CP022660.1	290957		*Salmonella enterica*
CP030220.1	159279		*Salmonella enterica*
NC_021817.1	78193	IncFII(S);	*Salmonella enterica*

NZ_CP029998.1	61198	IncFII(Yp);	*Salmonella enterica*
CP019187.1	100629	IncFII(S); IncFII(SARC14);	*Salmonella enterica*
NZ_CP030210.1	92624	IncFII(S); IncFII(SARC14);	*Salmonella enterica*
NZ_CP030177.1	131435	IncFII(pSE11);	*Salmonella enterica*
NZ_CP030197.1	141502	IncFII(pSE11);	*Salmonella enterica*
CP024167.1	92392	IncFII(Yp);	*Salmonella enterica*
NC_022372.1	121723	IncFIB(AP001918);	*Salmonella enterica*
CP022071.1	93843	IncFIB(S); IncFII(S);	*Salmonella enterica*
NZ_CP022071.1	93862	IncFIB(S); IncFII(S);	*Salmonella enterica*
NZ_CP029594.1	94034	IncFIB(S); IncFII(S);	*Salmonella enterica*
JQ418541.1	108459	IncFII(Yp);	*Salmonella enterica*
NZ_CP029997.1	86128		*Salmonella enterica*
NZ_CP019175.1	94727	IncFIB(pB171);	*Salmonella enterica*
NZ_CP022118.1	416444		*Salmonella enterica*
NC_019342.1	130175	IncFII(p14);	*Salmonella sp.*
NZ_CP026189.1	171306		