Genetic and functional characterisation of a novel *Klebsiella pneumoniae* genomic island harbouring an accessory chaperone/usher fimbrial operon

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Jon Jurriaan van Aartsen, BSc (Hons)

Department of Infection, Immunity and Inflammation

University of Leicester

September 2011

Abstract

Genetic and functional characterisation of a novel *Klebsiella pneumoniae* genomic island harbouring an accessory chaperone/usher fimbrial operon

Jon J. van Aartsen

Strain-specific *Klebsiella pneumoniae* virulence determinants have been described but almost exclusively for hypervirulent liver abscess-associated strains. Island-tagging and fosmid-based marker rescue were used to capture and sequence KpGI-5, a novel genomic island integrated into the *met56* tRNA gene of *K. pneumoniae* KR116. This 14.0 kb island exhibited a genome-anomalous G+C content, possessed near-perfect 46 bp direct repeats, and encoded a putative γ_1-chaperone/usher fimbrial cluster (*fim2*). This island was shown to belong to a previously unknown KpGI-5-like island family and was hypothesized to have undergone substantial reductive evolution. Transcriptional analysis demonstrated expression of three *fim2* genes and suggested that the eight-gene *fim2A*-*fim2K* cluster comprised an operon. *In vivo* models of urinary tract and lung infection, and large intestinal colonization, in addition to *in vitro* assays were used to examine the role of *fim2* in pathogenesis by comparing KR2107, a streptomycin-resistant derivative of KR116, to three isogenic mutants (Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2*) constructed using optimized allelic exchange techniques. Although no statistically significant *in vivo* role for *fim2* was demonstrable, liver and kidney CFU counts for lung and urinary tract infection models, respectively, hinted at an ordered gradation of virulence as fimbrial clusters were lost: KR2107 (most virulent), KR2107Δ*fim2*, KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* (least virulent). Despite using several methods the putative Fim2 fimbriae could not be visualised. Furthermore, the *fim2*-encoded putative phosphodiesterase Fim2K was determined to alter several c-di-GMP-dependent phenotypes, including biofilm formation and exopolysaccharide production. Additionally, *fim2* was present in 14 % of *Klebsiella* strains studied and appeared intact in the majority of *fim2*-positive strains. Given the above findings, *fim2* may confer a niche-specific evolutionary advantage, potentially related to its ability to encourage dissemination. This thesis is the first study of *fim2* and KpGI-5-like islands and will aid further investigations into the full impact of these loci on the lifestyle of *Klebsiella* species.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Kumar Rajakumar for his commitment, encouragement, guidance and support throughout my PhD project. He has helped me develop, both as an individual and as a scientist.

Furthermore I would like to thank Dr. Ewan Harrison for his help and for keeping me sane, entertained and motivated during the last few years. I would especially like to thank all current and past members of Lab 212 including Luisa, Barbara, Subbu, Andrew, Mandira and Eva.

I would like to express my gratitude towards Prof. Karen Krogfelt, Dr. Carsten Struve, Steen Stahlhut and all members of their laboratory at the Staten Serum Institut in Denmark for receiving me with open arms during my two month visit. I am also indebted Prof. Zixin Deng, Dr. Hong Yu Ou and Dr. Xinyi He at Shanghai Jiao Tong University.

I am extremely grateful to the following people for their advice and assistance: Dr. Edouard Galyov, Dr. Primrose Freestone, Dr. Helen O'Hare, Dr. Richard Haigh, Stefan Hyman, Dr. Hasan Yesilkaya, Dr. Peter Fineran, Prof. George Salmond and Dr. Hüseyin Besir.

I also want to thank the University of Leicester, Medisearch Leicester, Wellcome Trust, The Association of Physicians, ESCMID and Society for General Microbiology for their financial support at various points during the course of my PhD.

Finally, I would like to thank Dad, Mum, Bob, Tijn and Ellie for their love and encouragement, and without whom this journey would not have been possible.

Declaration of joint efforts

The handling, infection and dissection of mice during the *in vivo* infection assays was performed by trained personnel at the Staten Serum Institut, Denmark, according to their local guidelines. I prepared the inocula, collected faecal samples, performed CFU counts, and analysed the data obtained.

Abbreviations

λ	Lambda
µg	Microgram
µl	Microliters
µM	Micromolar
aa	Amino acids
AB	Antibiotic resistance cassette
bp	Base pair
BSI	Blood stream infection
c-di-GMP	Bis-(3'-5')-cyclic dimeric guanosine monophosphate
cDNA	Complementary DNA
CDS	Coding Sequence
CI	Competitive index
CFU	Colony forming units
Ct	Cycle threshold
DA	Downstream arm of a genomic island
DF	Conserved downstream flank of a tRNA locus
DGC	Diguanylate cyclase
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DR	Direct repeat
EDTA	Ethylenediaminetetraacetic acid
ESBL	Extended-spectrum beta-lactamase
FCS	Fetal calf serum
FRT	Flp recombinase target
g	Grams
gDNA	Genomic DNA
GI	Genomic island
GP-RBC	Guinea pig red blood cells
GST	Glutathione S-transferase
h	Hours
HCN	High copy number
HPI	High pathogenicity island
IBC	Intracellular bacterial community
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	Kilobase pair
kDa	Kilodalton
l	Litres
LA	Lysogeny broth agar
LAG	Lysogeny broth agar supplemented with gentamycin
LAGS	Lysogeny broth agar supplemented with gentamycin and 6 % sucrose
LAS	Lysogeny broth agar supplemented with 6 % sucrose
LB	Lysogeny broth
LCN	Low copy number
LD_{50}	Lethal dose 50%
LF	Left homologous flank
LRR	Lambda Red recombination
M	Molar
Mb	Megabase pairs
MBP	Maltose binding protein
MCS	Multiple cloning site
mg	Milligram
MGE	Mobile genetic element
MIC	Minimum inhibitory concentration
min	Minutes
ml	Millilitres
mM	Millimolar
ng	Nanograms
nH_2O	Nanopure, PCR-grade water
nM	Nanomolar
°C	Degrees Celsius
OD_{xxx}	Optical density at XXXnm
ORF	Open reading frame
PAI	Pathogenicity island
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
PLA	Pyogenic liver abscess
pmol	Picomole
PTS	Phosphoenolpyruvate phosphotransferase system
qRT-PCR	Quantitative real time polymerase chain reaction
RF	Right homologous flank
RNA	Ribonucleic acid

RT-PCR	Reverse transcription polymerase chain reaction
s	Seconds
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SGSP-PCR	Single genome-specific primer-PCR
SLIC	Sequence and Ligation–Independent Cloning
SOE-PCR	Spliced overlap extension-PCR
tRIP-PCR	tRNA interrogation for pathogenicity islands, prophages and genomic islands-PCR
tRNAcc	tRNA gene content and context analysis
TAE	Tris-acetate-EDTA
T_m	Melting temperature
U	Units
UA	Upstream arm of a genomic island
UF	Conserved upstream flank of a tRNA locus
v/v	Volume per volume ratio
w/v	Weight per volume ratio

Table of Contents

List of Figures

List of Tables

Publications, presentations and prizes / grants

Publications

Chen N., H.-Y. Ou, J. J. van Aartsen, X. Jiang, M. Li, Z. Yang, Q. Wei, X. Chen, X. He, Z. Deng, K. Rajakumar, and Y. Lu.
The *pheV* phenylalanine tRNA gene in *Klebsiella pneumoniae* clinical isolates is an integration hotspot for possible niche-adaptation genomic islands.
Curr. Microbiol. 2010. 60:210-6.

Zhang J., J. J. van Aartsen, X. Jiang, Y. Shao, C. Tai, X. He, Z. Tan, Z. Deng, S. Jia, K. Rajakumar, and H.-Y. Ou.
Expansion of the known *Klebsiella pneumoniae* species gene pool by characterization of novel alien DNA islands integrated into *tmRNA* gene sites.
J. Microbiol. Methods. 2011. 84:283-9.

J. J. van Aartsen and K. Rajakumar (see Appendix 10)
An optimized method for suicide vector-based allelic exchange in *Klebsiella pneumoniae*.
J. Microbiol. Methods. 2011. 86:313-319.

Presentations

University of Leicester: Department of Infection, Immunity and Inflammation 1st Annual Postgraduate Student Conference, April 2009.
Oral presentation entitled: "*Klebsiella* genomics and pathogenesis: from genomes to function to disease-causing potential".

University of Leicester: Department of Infection, Immunity and Inflammation 2nd Annual Postgraduate Student Conference, April 2010.
Oral presentation entitled: "Characterisation of a *Klebsiella pneumoniae* genomic island bearing a novel type 1-like fimbrial operon".

Statens Serum Institut, Copenhagen, Denmark: November 2010.
Oral presentation entitled: "Characterisation of a *Klebsiella pneumoniae* genomic island bearing a novel type 1-like fimbrial operon".

Society for General Microbiology Spring Conference, Harrogate, United Kingdom: April 2011.

Poster presentation entitled: “A novel genomic island-borne type 1-like fimbrial operon is involved in biofilm formation in *Klebsiella pneumoniae*”.

Prizes and Grants

University of Leicester 50th Anniversary PhD Scholarship

PhD Studentship, September 2008

Wellcome Trust Advanced Courses

Attended the Genomics and Clinical Microbiology Course, 24 – 29 January 2010.

University of Leicester: Department of Infection, Immunity and Inflammation 2nd Annual Postgraduate Student Conference, 2010

Commended Oral Presentation.

The Association of Physicians of Great Britain and Ireland

Travel Grant, October 2010.

European Society of Clinical Microbiology and Infectious Diseases

Travel Grant, November – December 2010.

Society for General Microbiology

Postgraduate Student Conference Grant, April 2011.

Chapter 1. Introduction

The first account of a member of the *Klebsiella* genus was by the Austrian microbiologist Anton von Frisch who, in 1882, described an encapsulated bacillus that was isolated from patients with rhinoscleroma, a chronic granulomatous infection of the upper airways. This bacillus was named *Klebsiella rhinoscleromatis* in 1887 to honour Edwin Klebs, a German microbiologist who discovered *Corynebacterium diphtheria* (126, 417). In modern microbiology, *Klebsiella* strains are defined as capsulated Gram negative bacilli that form mucoid colonies and are non-motile (265).

In the following sections, the taxonomy, epidemiology and ecology of *Klebsiella* species will be described. Next, the high prevalence and mechanisms of antibiotic resistance will be discussed. An examination of the *K. pneumoniae* genome structure will follow and will emphasise the importance of strain-specific elements including plasmids and genomic islands. A synopsis of *K. pneumoniae* virulence factors will then be provided, with a particular emphasis on type 1 fimbriae. Finally, the preliminary data on the island sequenced and characterised in this study, KpGI-5, will be summarised and the aims and objectives will be discussed.

1.1 The *Klebsiella* species

1.1.1 Taxonomy

Since the first description of the *Klebsiella* genus many member species have been identified and the taxonomic structure has been determined using various biochemical (Table 1-1) and DNA methods by a number of authors, including Cowan (84), Bascomb (23) and Ørskov (457). However with the advent of sequenced-based taxonomic data confusion still exists regarding the correct classification (Figure 1-1).

Table 1-1 :- Biochemical tests used to identify *Klebsiella* species [a,b]

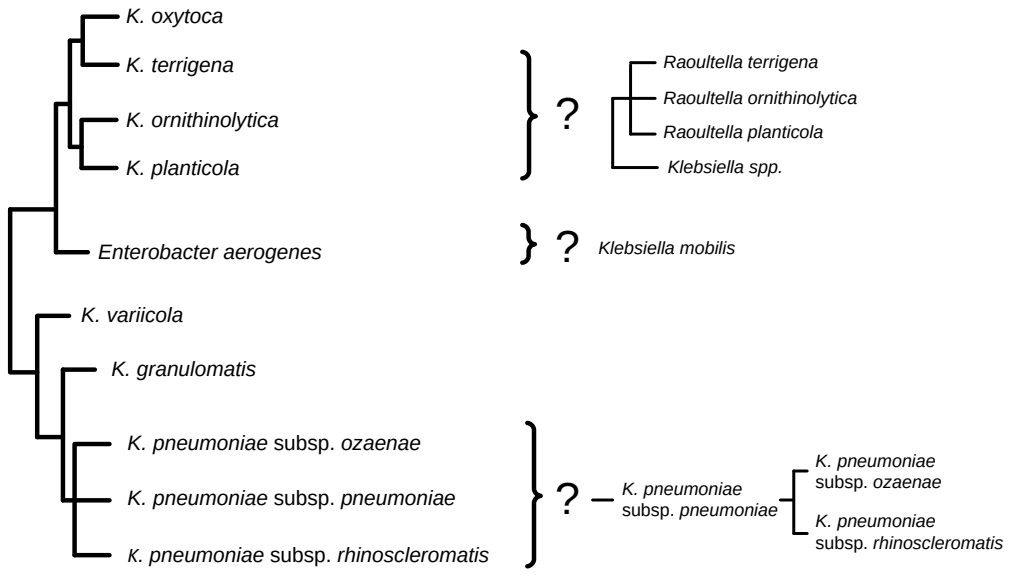
	Indole	Methyl red	Voges-Proskauer	Lysine decarboxylase	Ornithine decarboxylase	Malonate
K. pneumoniae subsp. pneumoniae	-	d	+	+	-	+
K. pneumoniae subsp. ozaenae	-	d	-	d	-	-
K. pneumoniae subsp. rhinoscleromatis	-	+	-	-	-	-
K. oxytoca	+	d	+	+	-	+
K. terrigena	-	+	+	+	-	+
K. planticola	d	+	+	+	-	+
K. ornithinolytica	+	+	d	+	+	d
E. aerogenes	-	-	+	+	+	+
K. variicola	Not reported in literature					

[a] +, 90 to 100% are positive; d, 11 to 89% are positive; - 0 to 10% are positive

[b] Table is a summary of data previously reported by Hansen *et al*. (153) and Brisse *et al*. (398).

The most recent taxonomic classification described eight distinct species (398). One of them, *Klebsiella pneumoniae*, is split into three subspecies: *pneumoniae*, *ozaenae* and *rhinoscleromatis*. Other species include *K. oxytoca*, *K. planticola*, *K. ornithinolytica*, *K. terrigena* and *K. mobilis*. The newest additions to the *Klebsiella* genus are *K. granulomatis* and *K. variicola* (55, 333). A recent study by Brisse *et al.* examined the existence of virulent *K. pneumoniae* clones using genomic and phenotypic techniques. They identified that *K. pneumoniae* subsp. *ozaenae* and *K. pneumoniae* subsp. *rhinoscleromatis* evolved from the *pneumoniae* subspecies and concluded that the current classification was incorrect and that they should be referred to as virulent clones of *K. pneumoniae* (40). It must also be noted that although *K. mobilis* has been included on the basis of its phenotype, 16S ribosomal DNA sequence and DNA-relatedness, it has also been grouped into another genus as *Enterobacter aerogenes*. Furthermore, based on 16S ribosomal DNA and *rpoB* sequence in addition to phenotypic data, Drancourt and colleagues have suggested transferring *K. planticola*, *K. ornithinolytica* and *K. terrigena* to a new genus called *Raoultella* (98).

Figure 1-1 :- Schematic of phylogeny within the *Klebsiella* genus



This schematic was generated from several phylogenetic trees previously constructed and published in the literature (36, 98, 333, 398). The right side of the schematic illustrates areas of *Klebsiella* phylogeny that are still under debate.

1.1.2 Ecology and epidemiology

Klebsiella spp. are ubiquitous within the environment and its species can be isolated from diverse locations including soil, water, trees and plants, roots, vegetables and industrial waste (124, 206, 303). Although data on disease and carriage in animals is limited, *Klebsiella* spp. are known to cause equine metritis (inflammation of the lining of the uterus) and can cause serious infections in dogs, guinea pigs, birds, and monkeys (125, 141, 199, 301, 441). A recent investigation has highlighted the genus as an important cause of mastitis (inflammation of breast tissue) and milk loss on dairy farms. A high prevalence of *Klebsiella* spp. was measured in rumen, water and soil samples and it was suggested that oral-faecal transmission of bacteria was occurring at the farm (450).

In humans *Klebsiella* spp. form part of the nasopharyngeal and gastrointestinal tract commensal flora. A survey of carriage rates in the community measured that 29 – 35 % of individuals carried *Klebsiella* spp. in their stool; a much lower rate was observed for

throat carriage (3 – 4 %). These numbers were only slightly increased in hospitalised patients (90). Both antibiotic therapy and length of hospital stay are directly correlated with an increased likelihood of colonization by *Klebsiella* spp. (307, 309). Carriage rates are important because it has been hypothesised that *Klebsiella*-related infections originate from bacterial reservoirs since capsule types found at both the infection site and gastrointestinal tract reservoir are identical (360).

The seriousness of *Klebsiella*-host interactions range from asymptomatic carriage, to opportunistic infections and more serious community-acquired infections, such as Friedländer's pneumonia and pyogenic liver abscesses (PLA). Friedländer's pneumonia is caused by *K. pneumoniae* and is a severe community-acquired pyogenic pneumonia that is often observed in chronic alcoholics and has distinctive radiographic features (54). Recent data suggests that although it has become rare in Europe, Australia and the United States, it has been observed more frequently in Asia and South Africa where it has become the most prevalent cause of community-acquired pneumonia (207).

Traditionally, PLAs were caused by polymicrobial infections (*Escherichia coli* being the most frequently isolated) and were associated with underlying hepatobiliary disorders. However, since the early 1990s *K. pneumoniae* has emerged as an important life-threatening community-acquired pathogen in eastern Asia and has taken over as the primary cause of PLAs (230, 315). Specifically, in Taiwan the incidence of *K. pneumoniae* PLA between 1996 and 2004 increased from 11.2 to 17.6 per 100 000 population (419). Interestingly, 70 – 75 % of *K. pneumoniae* PLA patients have underlying diabetes and often infection can lead to complications with poor prognoses including endophthalmitis and meningitis (207, 410). *K. pneumoniae* community-acquired meningitis can also occur without association with PLA, and rose from 8 % of bacterial meningitis cases (1981 to 1986) to 18 % of cases (1987 to 1995) in one Taiwanese hospital (207). Reports of *K. pneumoniae*-associated PLA and meningitis

have also been reported in the Europe, Canada and the United States (140, 194, 195, 272, 375).

K. pneumoniae subsp. *rhinoscleroma* and *ozaenae* are virulent clones which cause rhinoscleroma and atrophic rhinitis, respectively, two severe conditions of the upper respiratory tract. Rhinoscleroma is a chronic granulomatous infection which is endemic in tropical and subtropical regions and is usually not identified until advanced stages due to the low specificity of clinical signs. Successful treatment requires long-term antibiotic therapy (158). Atrophic rhinitis is mostly found in China, Africa and India and presents with three main symptoms: foul nasal odour, mucous membrane atrophy and crusting (91). When nasal discharge is cultured, *K. pneumoniae* subsp. *ozaenae* is the most frequently isolated organism (151). There have also been numerous reports of diarrheagenic *K. pneumoniae*, although this has not been investigated extensively (136, 145).

Klebsiella spp. cause 3 – 8 % of all nosocomial infections and are widely recognised as important pathogens in urinary tract infections, pneumonia, and wound, soft tissue and blood stream infections (BSI) (307, 321). The majority of opportunistic nosocomial *Klebsiella* infections are caused by *K. pneumoniae* and *K. oxytoca*, although occasionally *K. planticola* and *K. terrigena* have also been isolated (135, 154, 302, 305). Chronic alcoholism, diabetes mellitus, chronic cardiac, renal or pulmonary disease, malignancy and extremes of age are patient-specific factors predisposing to *Klebsiella* infection (245, 256).

The importance of *K. pneumoniae* as a nosocomial pathogen is highlighted by its inclusion as an “ESKAPE” pathogen, a group which includes *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species (328). These bacteria are involved in more than 40 % of intensive care unit infections and cause considerable

treatment difficulties due to high levels of antibiotic resistance (section 1.1.3). *Klebsiella* is ranked the second most common cause of Gram negative BSI, second only to *E. coli* (330, 447). However, recent data from the University Hospitals Leicester shows that the incidence of *K. pneumoniae* BSI has risen from 40 cases per year in 1999 to 114 in 2009 and that it is almost as frequently isolated as *Staphylococcus aureus*, locally the most common cause of ESKAPE BSI (D. Jenkins, personal communication). With regards to urinary tract infections, *Klebsiella* spp. account for 6 – 17 % of all infections and are commonly associated with catheterisation (330).

Epidemic outbreaks of nosocomial *Klebsiella* infections can occur and containment is often complicated by the widespread nature and dissemination of antibiotic resistance (section 1.1.3). Numerous reports of epidemics of highly resistant *K. pneumoniae* infections in neonatal intensive care units have been published and are cause for concern (17, 79, 86, 317). As mentioned earlier, the gastrointestinal tract forms a major reservoir of *K. pneumoniae* infection. In healthcare settings other infection reservoirs also exist and include medical equipment, blood products that have not undergone pathogen reduction treatments and the hands of hospital staff (142, 265, 318). This highlights the importance of good hand washing practice, proper maintenance and cleaning of medical equipment, and correct management of urinary catheters and other indwelling devices to reduce the spread of nosocomial *Klebsiella* infections.

Infection by *Klebsiella* has been associated with the development of ankylosing spondylitis, a chronic autoimmune inflammatory disease that affects spinal joints and large synovial joints. It is thought to be caused by molecular mimicry of the *K. pneumoniae* surface antigen K43 to human HLA-B27 antigens (280) and/or pullulanase enzyme to collagen fibre types I, III and IV (116). Recent reports have also suggested a possible link between *Klebsiella* infections and Crohn's disease, a chronic gastrointestinal auto-immune disease (106).

1.1.3 Antimicrobial resistance

Klebsiella strains possess native chromosomal β-lactamase genes, such as bla_{SHV}, bla_{LEN} and bla_{OXY}, which confer low-level resistance to aminopenicillin (e.g. ampicillin) and carboxypenicillin (e.g. carbenicillin) antimicrobials (59). In *K. oxytoca*, 10 – 20 % of strains possess mutations in promoter sequences that cause overexpression of bla_{OXY} resulting in increased resistance levels (123). Resistance to third and fourth generation cephalosporins are usually mediated by acquisition of plasmids encoding extended spectrum β-lactamases (ESBLs), such as the SHV1, TEM1 and TEM2 β-lactamases, of which over 50 variants have been described in *K. pneumoniae* (33). Within the Enterobacteriaceae family, *Klebsiella* spp. are the most frequently identified to produce ESBLs and a recent large centre study found ESBL production in 73 out of 455 *Klebsiella* BSI isolates (16 %) obtained from 12 hospitals in 7 countries. Additionally, most isolates produced two or more β-lactamases (289).

K. pneumoniae isolates have also been found to encode carbapenemases, which inactivate carboxy- and aminopenicillins, cephalosporins and carbapenems, which are last resort antimicrobials in many clinical infections. Different types of carbapenemases have been characterised, including metallo-β-lactamases (e.g. NDM-1), expanded-spectrum oxacillinases (e.g. OXA-48) and clavulanic acid inhibited β-lactamases (e.g. KPC) (276). The first report of a KPC carbapenemase-encoding *K. pneumoniae* was from the United States in 1996 and although prevalence in Europe is still limited, these strains are now endemic in Israel, Greece and China (49, 85, 276, 345). Moreover, KPC-encoding plasmids are spreading geographically and into other bacterial genera, including *E. coli*, *Enterobacter cloacae*, *Proteus mirabilis* and *P. aeruginosa* (276). New Delhi metallo-β-lactamase-1 (NDM-1) was first identified in 2008 in a *K. pneumoniae* isolate from a Swedish patient who had received treatment in a hospital in New Delhi, India (448). The enzyme inactivates all β-lactam antimicrobials apart from aztreonam and has rapidly disseminated worldwide (217, 448). Worryingly,

NDM-1 positive strains, including *K. pneumoniae*, *Shigella boydii* and *Vibrio cholerae*, could be isolated from drinking water and seepage samples (i.e. water pools in streets or small streams) in New Delhi, and these could act as reservoirs for the spread of NDM-1 (431).

Aminoglycoside resistance in *K. pneumoniae* is mainly determined by aminoglycoside modifying enzymes, which are regularly encoded by plasmid-based genes such as *aac(6')-Ib* and *aacA43* (287, 346). Quinolone resistance in *K. pneumoniae* is most often mediated by mutations in the chromosomal genes *gyrA* and *parC* (92), although there are reports of plasmid-based quinolone resistance mechanisms such as efflux pumps or DNA gyrase protection systems (129, 193).

Although the resistance mechanisms discussed above have been described in isolation, these genes often co-exist on plasmids where they collectively encode resistance to a number of antimicrobial classes and render host *K. pneumoniae* strains highly resistant (289). Fortunately, in clinical practice several antimicrobials can still be applied to treat *Klebsiella* infections, although agents used vary according to local resistance patterns. Table 1-2 summarises the antibiotics commonly used and avoided when treating *Klebsiella* infections at the University Hospitals of Leicester NHS Trust (425). The presence of drug resistance considerably limits infection treatment options and causes a significant impact on morbidity, mortality and the economy (83, 254, 422). Some strains have now been described which are pandrug resistant and only susceptible to polymyxins, antibiotic drugs with significant toxic side effects (111, 178). The information presented in this section highlights the increasing threat that *K. pneumoniae* poses to public health and the importance of elucidating its genome structure and function in terms of impact on both resistance and pathogenesis.

Table 1-2 :- Treatment of *Klebsiella* infections at University Hospitals of Leicester

Antimicrobials to which *Klebsiella* strains are commonly sensitive
- Cephalosporins
- Gentamicin
- Ciprofloxacin
- Piptaxobactam
- Carbapenems (meropenem or imipenem)
Antimicrobials to which *Klebsiella* strains are commonly resistant
- Amoxicillin
- Vancomycin

1.2 *K. pneumoniae* genome structure, plasticity and the involvement of mobile elements

Bacterial genomes consist of two parts: the core and flexible genome (257). The core genome encodes proteins involved in cellular maintenance, such as ATPases and DNA replication machinery. It is a relatively stable region and contains genes shared by all strains of the same species (148, 269). The remainder is the flexible genome and is highly variable between strains of a bacterial species. It plays an important evolutionary role by contributing to species diversity and encoding additional functions that may aid adaptation to new environments. Strain-to-strain variation in flexible genome content accounts for the genome size discrepancy between different strains of a bacterial species. Alterations in the flexible genome can occur by either gene loss or acquisition, which can happen on a small scale by homologous recombination or mutations. On a larger scale, deletion or gain of gene arrays can result in evolutionary ‘quantum leaps’ (95, 229).

Whole genome sequencing has allowed for detailed examination of the *Klebsiella* core and flexible genome. At the time of writing, four *Klebsiella* genomes have been fully sequenced, assembled and annotated, five have been uploaded as unassembled contigs and over thirty are in progress or have been targeted for sequencing. This

thesis makes many references to three sequenced *K. pneumoniae* strains: MGH78578, Kp342 and NTUH-K2044. Their characteristics are summarised in Table 1-3 in addition to six other *Klebsiella* strains with accessible genome sequence data.

Table 1-3 :- *Klebsiella* strains with freely accessible sequence data

Species, Strain and Description	Size / G+C content	Plasmids	GenBank Accession	Reference
K. pneumoniae				
Kp342 Nitrogen fixing endophyte isolate from maize plants	5 641 kb / 57.3 %	pKP187: 188 kb pKP91: 91 kb	CP000964.1	(124)
MGH78578 Pneumonia isolate	5 315 kb / 57.5 %	pKPN3: 176 kb pKPN4: 107 kb pKPN5: 89 kb pKPN6: 4.3kb pKPN7: 3.5 kb	CP000647.1	(407)
NTUH-K2044 Pyogenic liver abscess isolate	5 248 kb / 57.6 %	pK2044: 224 kb	AP006725.1	(444)
ATCC 13884 Airway isolate	5 281 kb / N/A	N/A	ACZD00000000.1	(175)
KpJH1 Antibiotic susceptible isolate	N/A / N/A	N/A	AFQK00000000	(216)
Kp1162281 Highly antibiotic resistant isolate	N/A / N/A	N/A	AFQL00000000	(216)
K. variicola				
At-22 Isolated from a leaf cutter ant fungus garden	5 458 kb / 57.6 %	None	CP001891.1	(299)
***Klebsiella* species**				
MS 92-3 Gut isolate	N/A / N/A	N/A	AFBO00000000.1	(408)
1_1_55 Gut isolate	5 417 kb / N/A	N/A	ACXA00000000.1	(176)

N/A, not available.

Comparative analyses using computational and wet science techniques show that, similar to many other Enterobacteriaceae, *K. pneumoniae* possesses an extremely plastic genome that consists of a conserved core genome that is interspersed by the strain-specific accessory genome (2, 60, 282). In a recent collaborative study, we identified that only 54.7 % of known *K. pneumoniae* genes are shared between three

sequenced isolates (Kp342, MGH78578 and NTUH-K2044) (454). The next section will highlight the mechanisms behind genome flexibility and the importance of the flexible genome in conferring strain-specific phenotypes to host strains (124, 444).

1.2.1 *K. pneumoniae* genome plasticity and mobile genetic elements

Large-scale acquisition of DNA occurs by horizontal gene transfer, via three different mechanisms: transformation, transduction and conjugation (94, 374). Transformation occurs when cells internalize DNA molecules directly from the surrounding environment. Conjugation is the cell-to-cell transfer of DNA by means of a specialized pilus structure. Finally, transduction is bacteriophage-mediated transfer of DNA that occurs when host bacterium DNA is accidentally incorporated into the phage head during packaging. Subsequent infection of a recipient cell results in the injection of the host cell DNA. Although these methods use different mechanisms of internalizing DNA, once inside either site-specific or homologous recombination needs to occur for chromosomal integration of the genetic element.

Mobile genetic elements (MGEs) are fragments of DNA that encode proteins involved in intracellular and/or extracellular DNA mobility, allowing DNA translocation within and between cells, respectively. They contribute significantly to horizontal gene transfer and constitute the majority of the bacterial flexible genome (127). MGEs exist in numerous forms, including plasmids, integrative conjugative elements (ICEs), transposons, bacteriophages and genomic islands (GIs). Often these elements encode pathogenicity or fitness factors, which may confer selective advantages to the host bacterium. Between these elements there is a potentially unlimited amount of variation in gene arrangement and composition (127).

Plasmids are self-replicating DNA molecules that are distinguishable from the chromosome because they are smaller and do not usually contain essential genes required to support bacterial life. Plasmid anatomy consists of genes required for

plasmid replication and a variable set of genes that encode accessory functions (374). Some plasmids facilitate their own horizontal transfer by encoding a conjugation system; these are called conjugative plasmids. In *K. pneumoniae* conjugative plasmids have been associated with the rapid spread of antibiotic resistance in both hospital settings and the community (section 1.1.3). Several *K. pneumoniae* plasmids have been sequenced and characterised; pLVPK and pMET1 are discussed here.

pLVPK is a 219 kb virulence plasmid isolated from *K. pneumoniae* CG43, a bacteraemia isolate (62). It possesses several gene clusters with putative roles in copper, silver, lead and tellurite resistance in addition to harbouring multiple iron-acquisition systems, traits which could be advantageous in certain niches. Furthermore, it harbours two capsule synthesis regulator genes (*rmpA* and *rmpA2*) that interact with the *cps* operon promoter and alter capsule expression levels (64, 221). Using an *in vivo* murine infection model, pLVPK was found to be involved in bacterial dissemination to extraintestinal sites and in subsequent abscess formation (401). Additionally, the hypervirulent strain NTUH-K2044 also possessed a plasmid similar to pLVPK (444).

The 41.7 kb self-transmissible plasmid pMET1 encodes a conjugation-associated type IV secretion system in addition to multiple antibiotic resistance genes (380). Interestingly, comparative analysis revealed that components of pMET1 were very similar to those of *Yersinia pestis* plasmid pCRY, highlighting the occurrence of horizontal gene transfer between *K. pneumoniae* and *Yersinia*. The pMET1 resistance genes were located on Tn*1331.2*, an 11.0 kb transposon containing a bla_{TEM-1} gene and a duplicated *aac(6')-Ib-aadA1-*bla_{OXA-9} gene cassette. This gene array conferred resistance to several aminoglycosides and β-lactams, including ampicillin, amikacin, kanamycin, streptomycin and tobramycin (380, 412) (section 1.1.3).

Transposons are MGEs which characteristically encode a transposase, an enzyme that catalyses transposition (movement from one genomic location to another) of DNA. Generally, transposons are flanked by short inverted repeats which are recognised by cognate transposases and are used to coordinate transposition, which can be site-specific, such as downstream of the *glmS* gene in the case of Tn*7* (293), or not, as with Tn*5* (327). In addition to the transposase, transposons encode accessory elements such as antibiotic resistance genes, as highlighted by Tn*1331.2* above*,* or genetic elements that confer unique phenotypic traits. Apart from adding new functions, transposons may also insert into and disrupt genes/operons and cause trait loss. Transposons which only encode transposition-related genes are referred to as insertion sequences (243).

Integrons are genetic elements that consist of an integrase gene, gene cassettes and an *att* site. Integrases are involved in the excision and integration of non-functional gene cassettes, which are subsequently assembled into large functional gene cassette arrays, commonly up to ten cassettes in size, although super integrons with over 100 gene cassettes have been isolated (121). Integrons themselves are non-mobile, but they are often found on mobile genetic elements such as plasmids and transposons and can therefore be horizontally transferred. Several *K. pneumoniae* resistance determinant-encoding integrons have been characterised (80, 99, 308).

Bacteriophages outnumber bacterial cells ten-fold and they represent an enormous reservoir of genetic diversity in the environment (44, 127, 170). Amongst their mosaic genetic assortment are genes that encode structural phage proteins as well as components that enable control of the host cell replicative machinery. Lytic bacteriophages replicate rapidly and cause host cell lysis, whilst temperate bacteriophages enter an alternative non-lytic phase called lysogeny. In this stage the phage genome becomes a chromosomally-integrated bacteriophage called a prophage. Under certain environmental conditions or stimuli, such as UV light or

antibiotic stress, prophages are triggered to enter into the lytic phase, involving rapid phage replication and packaging, and host cell lysis. Released phages are then able to re-infect and transduce their DNA into other host cells (127, 374). Recognition of host cell receptors often occurs via specific interactions with phage tail receptors, a property which has been exploited in phage typing of *K. pneumoniae* to distinguish between sporadic cases of infection (non-related strains) or outbreaks (related strains) (359). Analysis of the available sequenced genomes has identified several prophages, including at least two in Kp342 (36.3 kb and 48.6 kb), five in MGH78578 (11.0 kb to 58.3 kb) and one in NTUH-K2044 (23.9 kb) (124, 444). Several lytic phages have also been found in *K. pneumoniae* including Kpp95, which could also lyse *K. oxytoca*, *Enterobacter agglomerans* and *Serratia marcescens*, KPO1K2 and KP34 (100, 427, 445). The efficacy of these phages in bacteriophage therapy has yet to be examined.

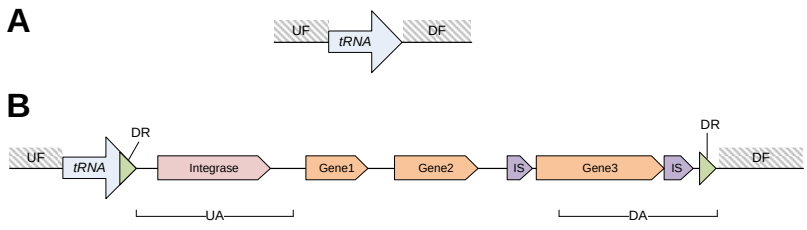
Some MGEs have been determined to possess both phage- and plasmid-like characteristics. These elements were named integrative conjugative elements (ICEs) because they were able to transfer by conjugation (similar to plasmids) and integrate into a host chromosome within which they are maintained (similar to phages) (48, 361). ICE*Kp1* is a 76 kb ICE from NTUH-K2044 which was determined to exist as both a circular element and a chromosomally-integrated form within *asn* tRNA genes (234, 444). It possessed a central region with putative virulence genes that were also present on plasmid pLVPK, and was flanked at the 5' end by genes similar to the yersiniabactin siderophore-encoding genes of the high pathogenicity island of *Y. pestis* and at the 3' end by genes determined to be involved in conjugation and transfer. Lin *et al.* also identified that the prevalence of ICE*Kp1* was significantly higher in PLA strains than non-invasive strains and concluded that the element may therefore contribute to virulence. Recently, an *asn* tRNA-associated ICE resembling ICE*Kp1* has been described for the colibactin *pks* GI, which is discussed in more detail in section 1.2.3 (277, 314).

As this project focuses on GIs, the next section will describe GI structure and function. Moreover, it will highlight how these MGEs contribute to novel traits and, in the subsequent section, specific examples of *K. pneumoniae* GIs will be discussed.

1.2.2 Genomic islands

The term GI is typically used to describe an element that can integrate and excise site-specifically but is unable to self-mobilize. They are often large 10 – 200 kb chromosomal regions that are tRNA gene-associated and possess a different base composition (also known as G+C content) compared to the genome average. Smaller genomic islets of 1-10 kb have also been found (95, 128, 149, 325). GIs frequently encode integrases, proteins involved in the site-specific integration and excision of GIs. Additionally, they are often flanked by short direct repeats (DR), which act as integrase recognition sites and are most likely artefacts generated from integration and/or excision, similar to the formation of *attL* and *attR* sites that flank prophages after genomic insertion (16, 58) (Figure 1-2). It is thought that GIs and ICEs are inter-related, but a recent phylogenetic analysis that compared integrases from numerous MGEs suggested that GIs are evolutionarily distinct from ICEs, plasmids, phages and integrons (35).

Figure 1-2 :- A typical genomic island



A) An 'empty' tRNA locus without an inserted GI. To the left and right of the tRNA gene are the conserved core genome upstream (UF) and downstream flank (DF), respectively. **B)** A typical GI inserted within a tRNA locus. It possesses genes that confer specific traits upon host bacteria in addition to an integrase gene. Transposons or insertions sequences (IS) may also be present on the island, which is flanked at its extremities by direct repeats (DR). Based on a diagram from Dobrindt *et al.* (95). UA, upstream arm ; DA, downstream arm.

Similar to all other MGEs, GIs can harbour an assortment of genes which encode various functions. Several gene combinations that confer advantageous traits and increase host cell fitness under specific environmental conditions have been described (95, 128, 149). The term pathogenicity island (PAI) was coined by Hacker and colleagues and has been used to describe GIs that harbour virulence-associated genes such as adhesins, invasins, exotoxins, iron acquisition systems, or secretion systems (149). Depending on their gene complement, GIs can also loosely be labelled as resistance, symbiosis or metabolic islands, but this classification can be troublesome as GIs often code for more than one function. Therefore, more broadly, islands without virulence-associated genes are often just referred to as GIs (94).

In Enterobacteria GI are often found immediately downstream of tRNA genes, and many integration 'hotspots' have been identified, including the *phe*, *arg*, *thr* and *sel* tRNA genes (139, 168). It has therefore been hypothesised that tRNA genes are universal targets for GI integration, a proposition supported by both the identification of tRNA segments in MGEs and the localisation of many GIs to tRNA loci (139, 168). Furthermore, it is thought that the tRNA 3' end is important during GI recombination into the chromosome. This is confirmed by the identification of DRs which are on average 16 – 20 bp (although larger repeats have been identified (2, 336)) in size and resemble the tRNA 3' end (149). Occasionally, tRNA genes can be truncated and rendered non-functional if inaccurate island excision occurs. It has therefore been suggested that tRNA genes with multiple copies (such as *asn* or *phe*) or those that are non-essential are more commonly used as tRNA integration sites (139, 168). Additionally, Germon *et al.* suggested that disruption of highly transcribed tRNA genes or those organised as polycistronic operons may cause severe effects on bacterial growth, and are therefore less often identified as GI insertion sites (139).

An elegant study by Touchon *et al.* has disputed the widely accepted view that GI integration most commonly occurs at tRNA loci (416). They compared 20 sequenced

E. coli strains and showed that large insertions and deletions occur at precisely the same locations across different genomes and that 83 % of these integration hotspots were not associated with tRNA genes. This raises the following question: what generates integration hotspots in bacterial genomes? Touchon and colleagues suggested that insertions within the majority of loci, including regulatory elements, genes, operons, or regions involved in DNA organization (such as nucleoid folding domains and macro-domains) are selected against as they would cause considerable loss of fitness (329, 416). They continued by suggesting that other local characteristics such as the presence of recombination machinery recognition sites or local DNA structure could also have an effect. However, the exact reasons behind hotspot locus generation remain speculative.

1.2.3 The repertoire of *K. pneumoniae* genomic islands and genomic island-like regions

K. pneumoniae genomes have been found to harbour many strain-specific GIs and GI-like regions (Table 1-4). An *in silico* genome comparison between the tRNA loci of strains MGH78578 and Kp342 using tRNAcc (tRNA gene content and context analysis) found that, respectively, at least 4.7 % (251 kb) and 7.3 % (404 kb) of each genome was strain-specific and associated with tRNA genes (281, 282). Eight tRNA genes were identified as island insertion hotspots, including *thr5*, *arg6*, *asn33* (*asnT*), *asn34*, *phe55* (*pheV*), *met56*, *leu82* and *tmRNA*, which have been examined in detail in follow-up studies (2, 60, 454).

Table 1-4 :- Summary of *K. pneumoniae* GI and GI-like regions

Name	Strain*	Size	G+C content	Locus	Function	Ref.
KpGI-1	KR116	3.6 kb	46.7 %	*pheV*	Unknown, but over-represented in sputum isolates	(2, 60)
KpGI-2	HS04160	6.4 kb	38.0 %	*pheV*	Role in growth regulation & possible role in pathogenesis	(60)
KpGI-3	MGH78578	12.6 kb	53.2 %	*pheV*	Unknown, but possible role in surface adhesion and colonization	(2, 60)
KpGI-4	KR173	Unknown		*pheV*	Unknown	(1)
tmGI_Kp20093	Kp20093	15.4 kb	43.4 %	*tmRNA*	Unknown	(454)
tmGI_Kp10011	Kp10011	17.5 kb	46.4 %	*tmRNA*	Unknown	(454)
tmGI_Kp49790	Kp49790	12.7 kb	48.6 %	*tmRNA*	Unknown	(454)
tmGI_Kp44	Kp44	9.0 kb	44.3 %	*tmRNA*	Unknown	(454)
tmGI_Kp63	Kp63	12.0 kb	50.3 %	*tmRNA*	Unknown	(454)
HPI[#]	RK75	35-43 kb	52.0%	*asnT*	Iron acquisition via yersiniabactin; virulence factor in murine respiratory tract infections	(16, 208)
pks	KpCF1	54 kb	53.0%	*asn*	Synthesis of colibactin, which can induce double-stranded DNA breaks and death in eukaryotic cells	(1, 314)
kfu/PTS	NTUH-K2044	20 kb	56.9 %	non-tRNA	Iron acquisition; virulence factor in murine infection model	(241)
allS island	NTUH-K2044	21.7 kb	54.2 %	non-tRNA	Allantoin metabolism; mutation in allS increased LD_{50} 10 to 100-fold.	(69)
Citrate fermentation locus	MGH78578	13.0 kb	59.2 %	non-tRNA	Improved ability to grow in the presence of citrate	(63)

* Strain in which GI was first identified. [#] High pathogenicity island

In previous work, we performed further analysis of five of these hotspots using techniques laid out in the MobilomeFINDER island identification algorithm (1, 282). Essentially, primers were designed that annealed to the conserved UF and DF (see Figure 1-2) and used in tRIP-PCR (tRNA site interrogation for pathogenicity islands, prophages and other genomic islands-PCR), in which an amplicon is obtained when the interrogated tRNA locus is empty (or has a small insertion). When the locus is occupied no amplicon is obtained, although other reasons for negative results exist, including PCR failure and UF/DF inversion (281, 282). Loci negative by tRIP-PCR were further investigated by a chromosome walking and sequencing technique called SGSP-PCR (single genome specific primer-PCR) (1).

Multiple studies have highlighted the role of the *phenylalanine* tRNA loci as GI integration hotspots (7, 139, 244, 336, 397). Using the above tRIP-PCR and SGSP-

PCR techniques, we previously identified that 95 out of 101 strains examined possessed putative insertions within their *pheV* tRNA locus (1, 2, 60). KpGI-1, a 3.6 kb genomic island, was first identified within the *pheV* locus of two *K. pneumoniae* strains (KR116 and KR164) isolated at the Leicester Royal Infirmary, United Kingdom. Two of four predicted coding sequences (CDS) were closely related to CDSs on a *pheR*-associated island in *Salmonella enterica* Typhi CT18 (1, 2). A follow up study showed this island was more common in sputum isolates than those from the urine, blood, wound or bile. The exact role of this island remains to be elucidated (60).

In the same tRIP-PCR screen, we identified a 6.4 kb GI in the urine isolate HS04160, called KpGI-2 (60). It possessed a G+C content of 38.0 %, which was remarkably lower than the *K. pneumoniae* average of ~57 %, and had five putative CDSs and a truncated integrase gene. When the *orf2-orf3* genes were expressed in *E. coli* in the presence of 5 mM cAMP, greater than 3-log cell killing occurred within 6 h of incubation. A subtler opposite phenotype was observed for *orf4-orf5*. However, a neutral phenotype was observed when the entire KpGI-2 locus was expressed in *E. coli*, suggesting a role in cell growth regulation. Interestingly, the predicted product from *orf5* was determined to contain a FIC (filamentation induced by cAMP) domain. FIC-domain containing proteins VopS and AnkX have recently been shown to disrupt specific host cell processes after infection into eukaryotic cells, suggesting that Orf5 may have a similar role in *K. pneumoniae* pathogenesis, although this remains to be determined (285, 446).

A 12.6 kb *pheV*-associated GI harbouring a putative fimbrial operon was first identified in the MGH78578 genome sequence using tRNAcc (1, 281). Located downstream of a putative integrase gene were seven genes predicted to encode subunits of a P pilus-like system. Very distant homologs of these subunits were found in *Yersinia frederiksenii* ATCC 33641 and *Y. intermedia* ATCC 29909. This island has subsequently been identified in isolates from the United Kingdom and China, including

a variant without an integrase gene (2, 60). Although it requires further characterisation, it is likely that this locus encodes a P pilus-like appendage involved in surface adhesion and colonization.

A fourth *pheV*-associated island was found in strain KR173, and is referred to as KpGI-4 for consistency with previous nomenclature (1, 2, 60). The KpGI-4 UA and DA possessed high homology to the UA and an internal region of the *pheR* island in *S.* Typhi CT18, respectively. This is the same island to which KpGI-1 shows similarity. However, sequence from further within KpGI-4 did not match any known Genbank entries and was predicted to encode a 303 aa hypothetical protein. Work is currently underway in Dr. Rajakumar's laboratory to sequence and characterise this island.

Extensive strain-to-strain variation has also been found at *tmRNA* gene loci (139, 244, 281). In a collaborative project with Dr. Hong-Yu Ou at Shanghai University, we used tRIP-PCR, SGSP-PCR and long-range PCR to investigate the *tmRNA* loci of 28 environmental and clinical strains (454). Seven different GIs were identified, including two islands that were present in MGH78578 and Kp342 and five novel islands. These islands were called: tmGI_Kp20093 (15.4 kb, G+C content 43.4 %), tmGI_Kp10011 (17.5 kb, G+C content 46.4 %), tmGI_Kp49790 (12.7 kb, G+C content 48.6 %), tmGI_Kp44 (9.0 kb, G+C content 44.3 %) and tmGI_Kp63 (12.0 kb, G+C content 50.3 %). Sequence analysis identified integrase genes in all five islands as well as multiple putative genes encoding phage components. However, the majority of predicted products matched to hypothetical proteins. This observation was also reported by Hsiao *et al.*, who analysed 63 bacterial genomes and found that the majority contained higher proportions of novel genes in GIs compared to the remaining genome (170). The exact role of these *tmRNA* GIs in the lifestyle of *K. pneumoniae* has not yet been investigated.

As mentioned ealier, *asn* tRNA genes are also island integration hotspots and numerous GIs have been found at this locus in *K. pneumoniae*. Koczura *et al.* reported that 6 out of 34 *K. pneumoniae* strains examined produced yersiniabactin, a siderophore first identified to be encoded by the *Yersinia* high pathogenicity island (HPI) (15, 208). Similar to *Yersinia*, the *K. pneumoniae* HPI is variable in size (35 – 43 kb) and located downstream of *asnT* (208) and two studies have identified yersiniabactin to be a virulence factor in *K. pneumoniae* respiratory tract infections (16, 228). The HPI is also present on ICE*Kp1* (section 0). In an elegant series of experiments, Lin *et al.* found that ICE*Kp1* can be transferred from NTUH-K2044 into any of the four paralogous *asn* tRNA genes of another *K. pneumoniae* strain (234). It has been hypothesised that ICE*Kp1* and related ICEs are involved in the dissemination of HPI.

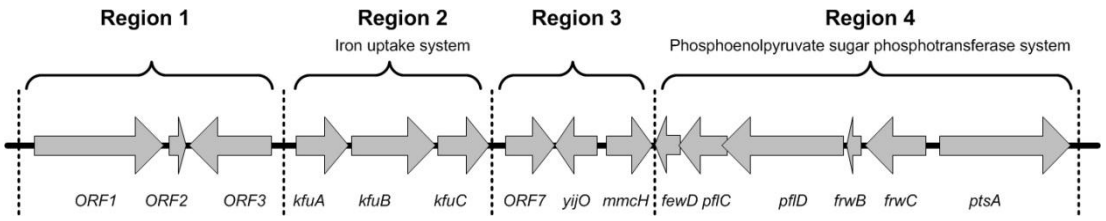
In 2006, Nougayrède and colleagues described a 54 kb GI that was associated with the *asnW* locus of *E. coli* phylogenetic group B2 strains, such as Nissle 1917 and CFT073 (277). This locus, *pks*, encoded proteins involved in the synthesis of colibactin, including eight nonribosomal peptide and polyketide synthases. *In vitro* the colibactin island was determined to have a cytopathic effect on eukaryotic cells by causing DNA double-strand breaks and arresting cells in the G2 phase of the cell cycle, after which cells underwent megalocytosis and cell death (277). It has therefore been hypothesised that *in vivo* this locus may predispose to the development of bowel cancer, although this has yet to be confirmed. Work in Dr. Rajakumar's laboratory using the tRIP-PCR and SGSP-PCR strategy also identified this island downstream of *asnT* in three *K. pneumoniae* strains (1). This finding and the aforementioned phenotype were confirmed by Putze *et al*. who found that 5 out of 141 *K. pneumoniae* strains possessed this locus and that the *pks* locus and *asn* tRNA gene were separated by a chromosomally-integrated ICE*Kp1* element (314). This colibactin-ICE*Kp1* tandem array was found at three different *asn* tRNA loci in the five strains examined. These findings

highlight a positive association between the colibactin and HPI islands. However, not all colibactin producing strains were yersiniabactin positive, and vice-versa (314). This is because both GIs can exist as unique single entities in addition to the ICE*Kp1* variants. It is therefore likely that the *pks* island, similar to HPI, also disseminates by excising and undergoing subsequent horizontal transfer with ICE*Kp1*.

The *K. pneumoniae met56* tRNA locus has also been examined extensively for strain-specific content. In MGH78578 this locus harbours a 33.1 kb prophage, whilst in Kp342 a 2.7 kb islet is present. Previous work using tRIP-PCR and SGSP-PCR in Dr. Rajakumar's laboratory identified a novel island inserted downstream of the *met56* tRNA in *K. pneumoniae* KR116 (1). The preliminary data for this putative island is discussed in detail later (section 1.4).

Although GIs frequently insert within tRNA sites, numerous non-tRNA loci in NTUH-K2044 have been identified to harbour strain-specific DNA. Using DNA microarray hybridization, Ma *et al.* identified a 20 kb NTUH-K2044-specific region (*kfu*/PTS) that possessed an iron uptake system (*kfuABC*), a phosphoenolpyruvate phosphotransferase system (PTS) and six CDSs to which no putative function could be attributed (241) (Figure 1-3). Whilst neither deletion of PTS nor the three CDSs comprising region 3 affected virulence in a murine infection model, the Δ*kfuABC* mutant displayed reduced virulence. An *in vitro* assay confirmed the *kfu* locus to be involved in iron acquisition, thus enhancing the ability of host *K. pneumoniae* strains to invade and survive within tissues (38, 241). However, the *kfu*/PTS locus did not match the classical description of GIs (section 1.2.2): G+C content was similar to the rest of the genome and no mobility genes or IS elements/transposons were present (38, 241).

Figure 1-3 :- *K. pneumoniae* NTUH-K2044 *kfu*/PTS locus



Genetic organization of the 20 kb *kfu*/PTS locus, a strain-specific region in *K. pneumoniae* NTUH-K2044. Image adapted from that provided in Ma *et al.* (241).

In a separate study, a DNA microarray was also used to compare the transcriptional profiles of 3 PLA strains and 3 non-PLA strains to identify genes associated with liver infection (69). A 21.7 kb region flanked by the *cap* and *purK* genes was not present in MGH78578 and was found to be expressed at high levels in PLA strains. BLAST analysis revealed that 14 of the 19 CDSs were similar to *E. coli* genes involved in the anaerobic assimilation of nitrogen and carbon from allantoin, including *allS* which encodes a putative regulator. This was confirmed by the ability of the NTUH-K2044 PLA isolate and inability of MGH78578 and NTUH-K2044Δ*allS* to grow in allantoin minimal medium (69). The wildtype strain was determined to be more virulent than the *allS* mutant in an BALB/c mouse infection model involving intragastric inoculation and the authors hypothesised that the allantoin-utilizing capability may help host-strains compete for sources of nitrogen (e.g. proteins and purines) when primary sources are unavailable. Similar to *kfu*/PTS, this strain-specific region was atypical of a GI. Although it harboured virulence genes and was large enough to be a GI, it was not integrated within a tRNA locus and possessed no integrase or mobility genes (69).

A comparison between the MGH78578, Kp342 and NTUH-K2044 genomes identified a 13 kb citrate fermentation gene locus that was not present in NTUH-K2044 and inserted upstream of the *dapB* (dihydrodipicolinate synthase) gene in MGH78578 and Kp342 (63). This GI-like region was also not located within or near any tRNA genes and did not possess repeat sequences or mobility genes. Intriguingly, nine out of

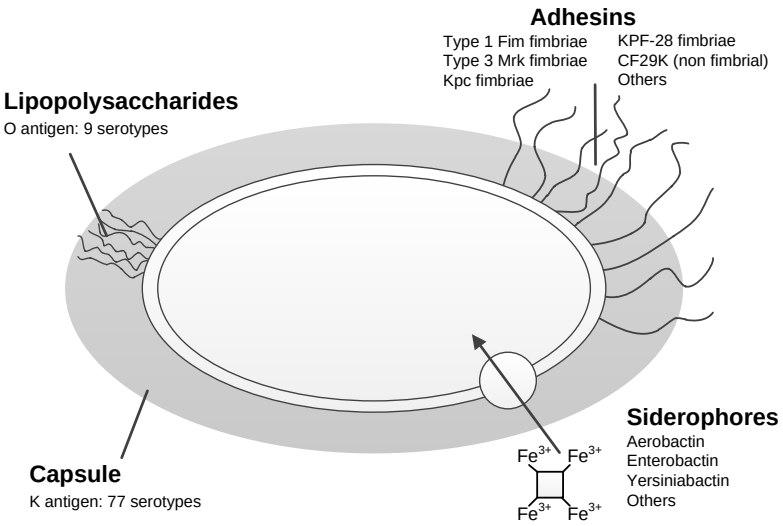
nineteen strains examined were found to have the cluster, and similarly organised but divergent clusters were found also found in *E. coli* and *Salmonella*. Using *in vitro* culturing techniques, the authors determined that this locus improved the ability of host strains to grow in the citrate-rich but oxygen deprived conditions of the urinary tract. However, the locus was not more common in urinary tract infection isolates (55 / 93 isolates) compared to those of other infections (53 / 94 isolates) (63). Further experiments, including *in vivo* virulence models, are necessary to describe the exact role of this locus in the host-pathogen interaction.

This section has described in detail the GIs and GI-like regions that have been identified and characterised in *K. pneumoniae*, highlighting the important role that the flexible genome plays in conferring strain-specific phenotypes to host-strains. The next section discusses the assortment virulence factors identified in *K. pneumoniae*.

1.3 *K. pneumoniae* virulence factors

K. pneumoniae strains possess a large arsenal of virulence factors (Figure 1-4). Most interestingly, environmental strains have the same arsenal as clinical isolates and are as virulent *in vivo* (124, 303, 390). This section will describe the immunoevasive properties of the surface-expressed capsular polysaccharide and lipopolysaccharide. It will then focus on iron-chelating siderophores that allow proliferation in iron-deficient conditions. Next, surface-adherent adhesins will be discussed and particular attention will be given to the type 1 and type 3 chaperone/usher fimbriae. Finally, the preliminary data of the novel KpGI-5 island and its putative virulence factor will be presented.

Figure 1-4 :- Schematic of *K. pneumoniae* virulence factors



K. pneumoniae strains encode a vast array of factors involved in pathogenicity, including a thick outer capsule, lipopolysaccharides, iron-chelating siderophores and many adhesins. These factors are discussed in detail in the text. Diagram based on a schematic by Podschun *et al.* (307).

1.3.1 Capsular polysaccharide: the K antigen

K. pneumoniae strains form thick outer capsules, which were the first virulence-associated factors to be described (411). The capsule is made from complex acidic polysaccharides that consist of repeating subunits of four to six sugars, including one uronic acid subunit (81), and it is responsible for the mucoid appearance of colonies on plates. 77 antigenically distinct capsular serotypes have been described and are included in the international K serotyping scheme (458).

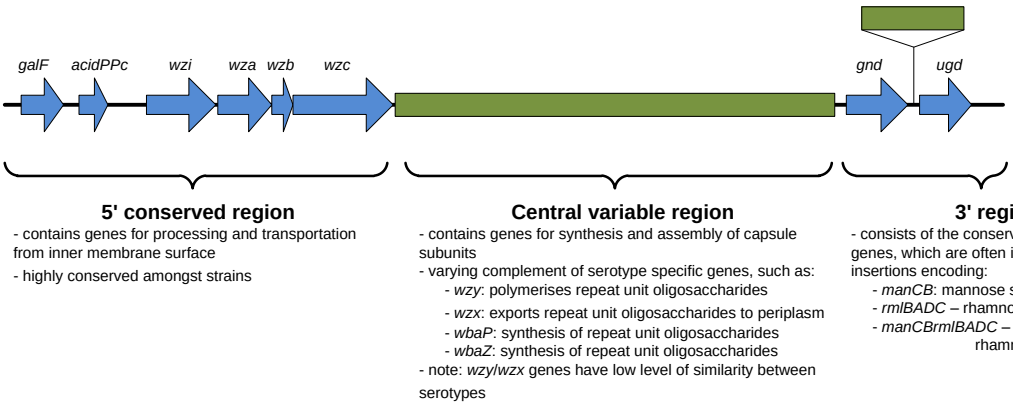
Several K serotyping schemes exist: 1) Quellung reaction, also known as the capsular swelling method; 2) the indirect immunofluorescence test; 3) double diffusion gel precipitation test; 4) latex agglutination test and 5) counter current immunoelectrophoresis technique (458). The latter is currently the method of choice at the WHO Collaborating Centre for Reference and Research on *Escherichia* and *Klebsiella* (Statens Serum Institut, Copenhagen, Denmark) (156). These techniques

are often limited to reference centres due to the lack of commercially available anti-capsule antiserum, observer bias and technique complexity.

Molecular methods have been developed in an effort to replace traditional K serotyping. In a technique developed by Brisse *et al.*, the capsule polysaccharide synthesis (*cps*) cluster was amplified by PCR, restriction digested with HincII and then separated on agarose gels (41). This generated C-patterns, and over 100 were described. C-patterns could discriminate between all K serotypes, apart from K22 and K37, and only 4.5 % of strains were non-typable, compared to 8 – 23 % of strains when classical K serotyping was used. An alternative PCR-based K typing method was also developed which used primers that annealed to *cps* locus serotype-specific *wzy* and *wzx* alleles (113). This technique could distinguish between the K1, K2, K5, K20, K54 and K57 serotypes. Moreover, no cross-reactions with the remaining reference serotype strains were identified. These primers have subsequently been used to develop a K serotyping multiplex PCR assay (423). Primers specific for other K serotypes have yet to be developed, and is hampered by the lack of *cps* region sequences. To date, only eleven *cps* loci with known K serotypes have been sequenced, in addition to another five with unknown serotypes.

The organization of the *K. pneumoniae cps* clusters resembles that of *E. coli* group 1 *cps* loci (316). To understand the diversity of the *cps* locus in *K. pneumoniae*, the loci of 12 strains with different capsular serotypes were compared by Shu *et al.* (363). The locus consisted of a conserved 5’ region with six transportation and processing genes and a 3’ region with two CDSs (Figure 1-5). The central region was highly divergent and encoded proteins involved in the synthesis and assembly of capsule subunits. It possessed a low G+C content compared to the genome average, a trait highly suggestive of horizontal gene transfer and homologous recombination events occurring at the *cps* locus, a notion supported in a study by Brisse and co-workers (40).

Figure 1-5 :- Genetic organisation of *K. pneumoniae cps* locus



The *K. pneumoniae* K antigen is encoded by the *cps* locus (21 – 30 kb), which consists of three distinct regions flanked by the *galF* and *ugd* genes (115, 363). The 5' region is highly conserved between K serotypes and contains six genes that encode proteins involved in processing and transportation of capsular subunits. The 3' region consists of the conserved *gnd* and *ugd* genes which, depending on the K serotype, can be interrupted by genes encoding proteins involved in either mannose and/or rhamnose biosynthesis (115, 363). The central region contains a highly variable complement of genes encoding proteins involved in the synthesis and assembly of capsular subunits and determine the composition and serotype of the capsule. The structure of this locus is similar to the *E. coli* group 1 *cps* locus (316, 436).

The role of the capsule in *K. pneumoniae* virulence has been investigated extensively. It has been identified as a virulence factor for murine urinary tract (389) and respiratory tract (227) infections in *in vivo* signature-tagged mutagenesis experiments; conflicting evidence has been reported regarding the effect on gastrointestinal colonization (114, 248). It has also been deemed essential to *K. pneumoniae* virulence in a murine pneumonia model (82) and metritis (398). Signature-tagged mutagenesis experiments have also shown that the capsule affects *in vitro* biofilm formation on abiotic surfaces (18) and extracellular matrix material (31). The capsule exerts a pro-virulence role by protecting bacteria from phagocytosis by polymorphonuclear leukocytes in addition to increasing resistance to serum by providing a barrier to serum bactericidal factors, such as opsonins (306, 307, 367). This was highlighted by a recent study which used isogenic capsule mutants and found that the rate of phagocytosis was inversely proportional to the amount of capsule produced, which hindered binding and

internalization of bacteria (109). Similarly, an isogenic NTUH-K2044 capsular mutant was serum sensitive and could be phagocytosed, unlike its wildtype parent strain (71).

Capsule types K1, K2, K4 and K5 are particularly virulent in murine models and are often identified in severe *K. pneumoniae* infections in humans (264, 368, 398). Serotype K1 strains are frequently isolated from patients with pneumonia and PLA, respectively (40, 113, 398). The remaining three 'hyper-virulent' serotypes are often found in community-acquired pneumonia (398). Interestingly, isolates that cause rhinoscleroma are always of serotype K3 (115). It is thought that the lower degree of virulence associated with some K antigens, such as K7 and K21a, is related to the presence of repetitive mannose sequences such as mannose-α-2/3-mannose and L-rhamnose-α-2/3-L-rhamnose (81, 307). These sequences can be recognized by carbohydrate binding domains (e.g. lectins) on macrophages and can lead to lectinophagocytosis.

More recently, it has been determined that virulence and infection type are associated with particular *K. pneumoniae* clones, rather than K serotype (40). Brisse *et al.* investigated the virulence gene content of two K1 clonal complexes, $CC23^{K1}$ and $CC82^{K1}$, and two K2 clonal complexes, $CC14^{K2}$ and $CC65^{K2}$, and found significant differences in virulence gene content even though K serotypes were shared. They then tested multiple strains of a particular clone in mice and identified that differences in virulence gene content between clonal complexes translated into differing *in vivo* virulence, and therefore concluded that "the long-held belief that K type is predictive of virulence should be discontinued" (40).

Regulation of the *K. pneumoniae cps* operon is not fully understood, but numerous regulatory genes have been characterised. Nassif *et al.* identified a gene they designated *rmpA* (regulator of mucoid phenotype) which was located on a 180 kb plasmid. When this gene was mutated, colony mucoidity was decreased and LD_{50} in

mice infections was increased 1000-fold (274). Mutation of *rmpA* has since been shown to reduce the expression of *cps* genes (64). A second gene, *rmpA2*, was found in plasmid pLVPK and was determined to enhance colony mucoidity by regulating capsule formation at the transcriptional level (221, 429). The two component regulatory system RcsAB and the ferric uptake regulator Fur also play a role in the intricate regulation of capsule synthesis (64, 233).

1.3.2 Lipopolysaccharide: the O antigen

Lipopolysaccharides are surface-expressed antigens composed of lipid A, a core oligosaccharide and an "O antigen" side chain. Nine O antigen types have been described: O1, O2, O2ac, O3, O4, O5, O7, O8 and O12. Of these, O1 antigens are the most frequently identified (155). Unlike K serotyping, O typing of *K. pneumoniae* is not routine practice because the capsule is heat stable and therefore unencapsulated mutants are necessary for typing (155), Additionally, K typing has higher discriminatory power than O typing (155).

K. pneumoniae lipopolysaccharides are encoded by the *waa* locus, a 12 gene cluster located between the *kbl* and *coaD* genes (324). Between O serotypes the lipopolysaccharide cluster appears well conserved; but this has not been investigated in detail (323). However, two types of core oligosaccharides can be produced at this locus, a property related to gene content differences: type 1 core clusters have genes *wabI* and *wabJ* while type 2 core clusters have genes *wabK* and *wabM* (324). Regué *et al.* investigated the effect of replacing a type 2 core with a type 1 core and found a two log increase in 50 % lethal dose value (324).

O antigens are important as they protect *Klebsiella* spp. from complement-mediated killing (413, 438), the effectiveness of which is related to O-antigen chain length (73). Using isogenic mutants and complementation experiments, Merino *et al.* identified a role for the O5 antigen in rat urinary tract colonization (258), and this is supported by *in*

vivo signature-tagged mutagenesis experiments where knockout mutants of multiple core lipopolysaccharide synthesis genes displayed attenuated virulence in murine urinary tract infection models (389). These mutants were also affected in their ability to colonize the murine gastrointestinal tract (389). In contrast, conflicting evidence exists on the involvement of the O antigen in pneumonia although this may be related to problems with experimental design and the close inter-relationship between the regulation of capsule synthesis and lipopolysaccharide production (82, 227). In a well-designed study by Lawlor *et al.*, a transposon mutant of the *wecA* synthesis gene was shown to have a 1000-fold greater LD_{50} compared to its wildtype in an intranasal murine infection model (227). Although *wecA* plays a key role in the synthesis of lipopolysaccharide and the Enterobacterial common antigen, it was determined that this attenuation was the result of a lipopolysaccharide-specific effect as Enterobacterial common antigen mutants were as lethal as wildtype strains in the same model (227).

1.3.3 Siderophores: high affinity iron binding compounds

Bacterial growth requires iron. In humans most iron is either sequestered intracellularly as haemoglobin or ferritin or bound to carrier molecules such as transferrin and lactoferrin (434), thus limiting the amount of iron available to bacteria. Certain human pathologies, such as the physiology post-chemotherapy and haemochromatosis, cause iron to become more freely available. Many pathogens can exploit this and cause significant morbidity and mortality on top of the underlying disorder (45, 46).

K. pneumoniae produces between four and six outer membrane proteins when grown under iron-limiting conditions, and these have been implicated in iron acquisition (437). Strains can encode the siderophores aerobactin, enterobactin (also known as enterochelin) and yersiniabactin, which have an extremely high affinity for iron and can remove iron from host carrier molecules (208, 292). Iron-siderophore complexes are recognised by cell-specific receptors and are transported into the cell via TonB-

dependent receptors, where the iron dissociates and is made available for use by the bacterium (260). Enterobactin was present in almost all strains, but evidence regarding the involvement in virulence is conflicting (303, 307). Aerobactin was synthesised by a plasmid-encoded operon found in approximately 16 % of *K. pneumoniae* isolates (208, 273). Heterologous expression of the aerobactin synthesis genes in an aerobactin negative strain was shown to confer virulence in a mouse peritonitis infection model (273). Yersiniabactin is coded for by the HPI island and related elements, such as ICE*Kp1*, and has been discussed elsewhere (section 1.2.3). In addition to yersiniabactin and aerobactin, NTUH-K2044 encodes eight putative iron uptake systems; the functionality of seven of these systems was shown to be dependent on TonB (171). Moreover, compared to the NTUH-K2044 wildtype an isogenic *tonB* mutant displayed greatly attenuated virulence, emphasising the importance of iron uptake mechanisms in *K. pneumoniae* pathogenesis (171).

1.3.4 Secretion systems

Secretion systems play a key role in bacterial interaction within biofilms and with eukaryotic hosts, in both pathogenic and mutualistic associations (24, 421). These systems enable secretion of a diverse assortment of substrates into the extracellular environment including toxins, proteases, haemophores, lipases, hydrolytic enzymes, DNA and effector proteins, which can alter cellular functions (24, 159, 421). Seven different secretion systems have been described in Gram negative bacteria. In type I, type III, type IV and type VI secretion systems the secreted proteins cross both the inner and outer membrane via a one-step export mechanism (24, 159, 421). Proteins transported via the chaperone/usher, type II and type V secretion systems are exported in a two-step mechanism (24, 159, 421). This is because the these secretion systems only traverse the outer membrane, and thus proteins secreted via these systems rely on an initial translocation across the inner membrane via the general Sec or twin-arginine (Tat) secretory pathways.

The most well-studied secretion system in the *Klebsiella* genus is the *K. oxytoca* type II secretion system (72, 312, 421). The system consists of several core components including pseudopilin subunits, an intracellular ATPase, integral inner membrane subunits and the outer membrane secretin PulD. Twelve to fifteen subunits of PulD form a pore through which the enzyme pullulanase (PulA) is secreted (72, 421). Pullulanase is an ~120 kDa lipoprotein secreted by *Klebsiella* oxytoca when grown in medium containing starch or maltose (105, 312). It hydrolyses α-1,6 glycosidic bonds in various polysaccharides to form maltotriose, which is subsequently transported into and metabolised by cells (312).

Several chaperone/usher secretion systems have also been described in Klebsiella spp. and these will be discussed further in Section 1.3.5. The remaining secretion systems have not been well described or investigated in *Klebsiella* spp.. Nevertheless, sequencing projects have identified core components of type I, type II, type IV and type VI secretion systems in the genomes of *K. pneumoniae* strains Kp342, MGH78578 and NTUH-K2044, and *K. variicola* At-22 (124, 299, 347, 444). However, there are no reports of type III or type V secretion systems in *Klebsiella* spp. (52). Further detailed investigations are needed to understand how these loci alter the interaction of *Klebsiella* spp. with their environment.

1.3.5 Adhesins: important surface expressed virulence factors

The first step in colonization and infection is surface adhesion, a process mediated by a diverse assortment of adhesins, each thought to possess a particular surface tropism. *K. pneumoniae* encodes large arsenal of fimbrial and afimbrial adhesins (307).

The terms pili and fimbriae are often used interchangeably to describe proteinaceous surface exposed filaments that are not part of flagella. They were first described on the surface of *E. coli* cells in 1950 (169) and the term ‘fimbriae’ was coined by Duguid and colleagues in 1955 (103). Duguid also performed the first studies of *Klebsiella* fimbriae

in the late 1950s (102). He demonstrated the presence of differing haemagglutination properties, which he used to classify fimbriae as type 1 or type 3. Type 1 fimbriae mediated mannose-sensitive haemagglutination of guinea pig red blood cells, whilst type 3 fimbriae mediated mannose-resistant agglutination of tannic acid treated ox erythrocytes (102, 138).

Traditionally, this phenotype-based historical classification has been useful but with the availability of whole genome sequences, large numbers of phenotypically uncharacterised putative fimbrial gene clusters have been found (278). Currently, fimbriae in Gram negative bacteria are distinguished according to their assembly mechanism. The following major classes exist: conjugative fertility (F) fimbriae, type IV fimbriae, fimbriae assembled by extracellular nucleation/precipitation pathway and fimbriae assembled by the chaperone/usher (CU) pathway (278). CU fimbriae can be further subdivided according to the sequence of the fimbrial usher protein, an outer membrane protein which acts as the fimbrial assembly platform (278). This classification defines six different fimbrial clades: α, β, γ, κ, π and σ. The γ fimbrial group is further split into four subgroups: γ_1, γ_2, γ_3 and γ_4. The type 1 and type 3 fimbriae have been allocated to the γ_1 and γ_4 groups, respectively.

In general, CU fimbrial genes are arranged into operons which code for at least three proteins: a major structural subunit, a chaperone and an usher. Some also contain additional genes encoding assembly proteins, minor fimbrial subunits and/or regulatory proteins (278). The *E. coli* P and type 1 fimbriae are the most well characterised CU assembly systems and it is assumed that similar mechanisms are used to assemble related CU fimbriae (430).

Using the P and type 1 fimbriae as archetypal examples, fimbrial biogenesis starts with the recognition of N-terminal signal sequences present on fimbrial assembly and structural proteins by the SecYEG general secretory pathway, and the subsequent

transport of these proteins across the cytoplasmic membrane (53, 278). Fimbrial subunits enter into the periplasm and interact with their cognate chaperones, which in turn ensure that subunits are folded correctly (53, 278). Structural studies have shown that fimbrial subunits possess a deep hydrophobic cleft because of an incomplete immunoglobulin-like fold that lacks a seventh β strand. When the chaperone binds it provides the missing β strand, called the 'donor' strand (418, 430). Next, the chaperone/subunit complex interacts with the usher, which is integrated within the outer membrane, resulting in the release of fimbrial subunits and their translocation through the usher channel. Assembly of subunits occurs by a donor strand exchange reaction in which the donor β strand of the chaperone is replaced with the N-terminal β strand of a second subunit (348, 418, 430). The tip adhesin subunit is separated into two folds which delineate the N-terminal and C-terminal domains. Whilst the C-terminal domain is similar to other subunits and is missing a β strand, the N-terminal domain forms a receptor binding domain that acts as an adhesin (185, 266, 295). Due to this assembly mechanism, chaperones and ushers interact in a system-specific manner and are either unable to interact with those of another system or do so very inefficiently (278).

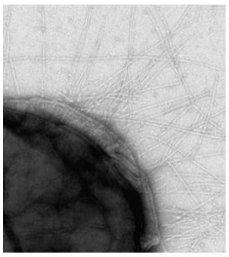
K. pneumoniae strains Kp342, MGH78578 and NTUH-K2044 contain nine, eleven and eight CU fimbrial operons, respectively, with the type 1 and type 3 fimbria-encoding operons being common to all three (443). Many of the remaining operons have not yet been characterised and it is likely that, similar to CU fimbrial operons in *E. coli* K-12, *E. coli* O157:H7 and *Salmonella* Typhi, the operons are tightly regulated and not expressed under the majority of *in vitro* conditions tested (177, 210, 238).

Type 3 fimbriae are 0.5 – 2.0 µm in length and 4 – 5 nm in width, and are coded for by the 5.6 kb chromosomal *mrk* operon (8, 212) (Figure 1-6). This gene cluster encodes a major fimbrial subunit (MrkA), a chaperone (MrkB), an usher (MrkC) and two minor fimbrial subunits (MrkF and MrkD). An alternative plasmid-encoded *mrk* gene cluster

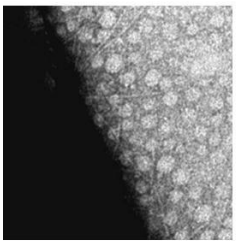
has also been described, and encodes an adhesin with a low degree of identity (60 %) to MrkD called $MrkD_P$ (388). The chromosomally-encoded *mrk* cluster is more prevalent than the plasmid-encoded variant, which was mainly limited to *K. oxytoca* strains (354, 358, 388). A recent PCR-screen identified that 95 % of *K. pneumoniae* strains possessed the chromosomal cluster (S. Stahlhut, personal communication).

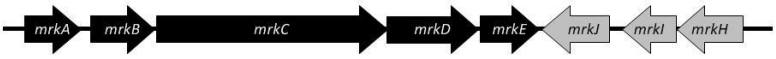
Figure 1-6 :- *K. pneumoniae* fimbrial loci and their associated fimbrial structures

Type 1 fimbriae – encoded and regulated by the 10.4 kb *fim* locus

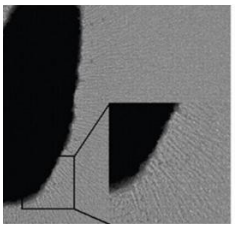


Type 3 fimbriae – encoded and regulated by the 7.8 kb *mrk* locus





Kpc fimbriae – encoded and regulated by the 5.8 kb *kpc* locus



This diagram illustrates the fimbrial structures and their associated *K. pneumoniae* fimbrial loci: type 1 fimbriae (387), type 3 fimbriae (388) and Kpc fimbriae (443). Grey arrows represent CDSs that encode regulatory proteins whilst the remaining structural protein- and chaperone/usher-encoding genes are indicated using black arrows. Transmission electron microscopy images were obtained from the aforementioned corresponding references.

Chromosomally-encoded type 3 fimbriae mediate bacterial attachment to human endothelial cell line EA-hy926 and bladder endothelium (402). Additionally, they affect

biofilm formation on both abiotic surfaces and surfaces coated by human extracellular matrix (31, 172, 352, 384). However, an isogenic *mrk* knockout strain was as virulent as its wildtype in a murine pneumonia and urinary tract infection model. Additionally, no alterations in the ability to colonize the murine gastrointestinal tract were identified (388). However, these models did not investigate the effects of urinary tract catheterisation and/or the presence of other abiotic devices on infection with the *K. pneumoniae* strain and its isogenic *mrk* mutant. This still remains to be investigated.

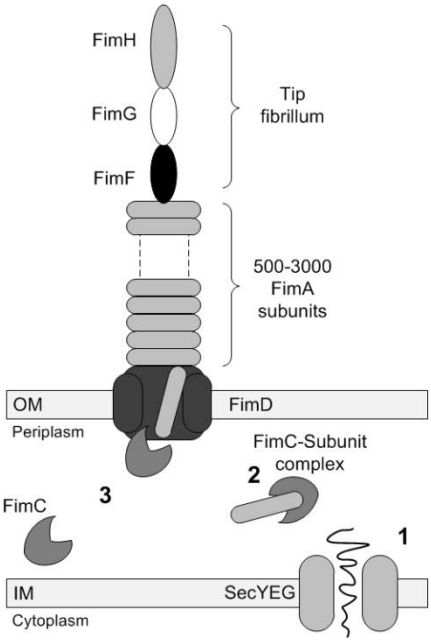
Several genes are implicated in the control of type 3 fimbrial expression. MrkJ is a functional phosphodiesterase enzyme with a conserved EAL domain (section 1.4). Deletion of *mrkJ* increased Mrk production and encouraged biofilm formation, an effect that occurred at the level of *mrkA* gene transcription (183). Mutants unable to produce the LuxR-type transcriptional regulator MrkI were deficient for Mrk production under aerobic conditions (184). An *mrkH* gene knockout did not produce type 3 fimbriae in any conditions. MrkH is a PilZ domain-containing protein that can bind c-di-GMP (bis-(3'-5')-cyclic dimeric guanosine monophosphate), the intracellular messenger hydrolysed by MrkJ (183, 184). Most interestingly, the *mrkHIJ* cluster is located adjacent to the *mrk* operon (Figure 1-6). However, the intricate mechanisms of how these elements coordinate and regulate type 3 fimbrial expression have not yet been investigated.

Type 1 fimbriae are 0.3 – 4.0 µm in length and 7 – 10 nm in width and were first described on the surface of *Klebsiella* strains in 1959 (102, 103). Phenotypically these fimbriae were determined to mediate mannose-sensitive agglutination of guinea pig red blood cells. *K. pneumoniae* type 1 fimbriae are encoded by eight genes which together constitute the *fim* operon (Figure 1-6) (75, 138, 313, 387). Although this arrangement is similar to the *E. coli fim* operon, the *fimK* gene is unique to *K. pneumoniae* and is involved in the regulation of fimbriation (331). Additionally, the products produced by the *K. pneumoniae fim* operon show 68 – 84 % homology to those of the *E. coli fim*

operon (387); however there are significant functional and serological differences between the two (332, 387).

The *fim* operon encodes a major fimbrial subunit (FimA), a chaperone (FimC) and usher (FimD), and three minor fimbrial subunits (FimF, FimG and FimH). The fimbria is assembled via the CU pathway and produces a fimbria structure that consists of a rod with 500-3000 helically arranged copies of FimA resting on FimD, and a tip fibrillum consisting of FimH, FimG, FimF subunits (Figure 1-7) (53).

Figure 1-7 :- Type 1 fimbria structure



The structure and arrangement of a type 1 fimbria and its subunits. 500 – 3000 copies of FimA form a helical arrangement which is capped by a tip fibrillum consisting minor subunits FimG, FimF and FimH. **(1)** Subunits are translolated across the inner membrane by means of the SecYEG system. **(2)** In the periplasm the translocated subunits bind to the chaperone FimC to prevent misfolding and facilitate delivery to the usher protein FimD. **(3)** The subunit dissociates from FimC and is secreted via FimD, which acts as a scaffold onto which the fimbial subunits are assembled. Image adapted from Capitani *et al*., 2006 (53), and Kine *et al*., 2010 (204). OM, outer membrane; IM, inner membrane.

The main adhesive subunit of type 1 fimbriae is FimH, although FimG and FimF have also been implicated (104). FimH has two domains that are connected by a linker of three amino acids (385). The C-terminal region (amino acids 1 – 156) is the pilin

incorporating domain, and is used to assemble FimH into the fimbrial structure. The N-terminal (amino acids 160 – 279) forms the mannose binding lectin / adhesin domain. In *E. coli* FimH this domain recognizes mannosylated glycoproteins that are present on the bladder urothelium, such as uroplakins and β-1 and α-3 integrin receptors (332).

Although the *K. pneumoniae* and *E. coli* FimH differ only by about 15 % of amino acid residues, considerable discrepancies between their binding specificities have been reported by multiple authors (332, 385). The FimH of *K. pneumoniae* TOP52 was unable to mediate the agglutination of guinea pig red blood cells but was involved in biofilm formation (332). However, the latter could have been a result of polar effects from the *fimH* mutation on *fimK* expression, which also regulates type 3 fimbrial production, and hence biofilm formation (C. Struve and S. Stahlhut, personal communication). Moreover, Stahlhut and colleagues identified that the *K. pneumoniae* FimH is unable to bind terminally-exposed Manα(1-2) residues, which is the basis of FimH-mediated adherence of *E. coli* to uroepithelial cells (385). Further investigation using recombinant *K. pneumoniae* and parallel plate flow chamber experiments identified that *K. pneumoniae* FimH was able to bind terminally-exposed Manα(1-3)Manβ(1-4)GlcNacβ1 trisaccharides. The strength of this interaction increased seven-fold under sheer force, a property mediated via a catch-bond mechanism, which was first described for the uropathogenic *E. coli* FimH protein (385, 409).

A population study of 65 *K. pneumoniae* FimH protein sequences revealed multiple pathoadaptive amino acid replacements that increase tropism towards different host membrane components (383). One of the replacements, S62A, converted the nature of FimH binding from sheer-dependent to sheer-independent, a conversion also described in *E. coli* and associated with highly virulent uropathogenic strains (376, 385). Fascinatingly, phylogenetic analyses have provided evidence that *fimH* is subject to frequent recombination via horizontal gene transfer (383).

More than 80 % of *K. pneumoniae* isolates produce type 1 fimbriae, as detected in screens looking for the ability of strains to agglutinate guinea pig red blood cells in a mannose-sensitive manner (303, 304). This number has recently been confirmed in a PCR-based assay for the *fim* locus (383). Most interestingly, the incidence in environmental strains is almost identical to that in clinical isolates (303, 383).

In vitro K. pneumoniae type 1 fimbriae can mediate attachment to ciliated hamster epithelial cells and bladder urothelium (110, 331). However, unlike *E. coli* type 1 fimbriae, a role in biofilm formation on abiotic surfaces has not been identified, as shown by heterologous expression experiments in the afimbriate *E. coli* strain HB101 and by comparisons between isogenic wildtype and *fim* knockout strains (352, 387). This is supported by the lack of *fim* genes amongst genome-wide screens for factors involved in biofilm formation (18, 31, 226, 384). The *in vivo* role of *fim* has been investigated using isogenic wildtype and *fim* knockout strains. Although the locus was not associated with pneumovirulence or intestinal colonization, it was determined to have a prominent role in urovirulence (387).

Similar to uropathogenic *E. coli*, *K. pneumoniae* urinary tract infections proceed via an intracellular pathway whereby bacterial internalization occurs though a type 1 fimbria-dependent mechanism after which biofilm-like intracellular bacterial communities (IBCs) are formed within superior facet cells of the bladder (331). Bacteria within IBCs can then reseed infections by fluxing out of epithelial cells and filamenting within the bladder lumen (331). Although not investigated in detail, based on similarities with uropathogenic *E. coli* it is likely that *K. pneumoniae* also cycles through multiple phases of IBC formation, fluxing and filamentation, after which they enter a quiescent phase during which they are hidden from antibiotics and the host immune system (187, 188, 442). It is thought that these quiescent phase bacteria are bacterial reservoirs for recurring infections of the urinary tract.

Expression of the *K. pneumoniae fim* operon is controlled by the reversible inversion of *fimS*, a 314 bp region upstream of *fimA* that is flanked by 9 bp inverted repeats (TTGGGGCCA) and harbours the *fim* promoter (387). Expression is turned “on” when the promoter is in the same orientation as the *fim* operon and turned “off” when the *fimS* region inverts. An identical mechanism has been described for the *E. coli fim* operon (5). Although not investigated in *K. pneumoniae*, from parallels with *E. coli* it is assumed that switch inversion is controlled by two site-specific recombinases encoded by *fimB* and *fimE*, which lie upstream of *fimS* (Figure 1-6). FimB directs the switch from “on” to “off” and from “off” to “on” whilst FimE is only involved in turning “off” expression (203, 253).

The *K. pneumoniae fimS* switch orientation has been examined *in vivo*: whilst the majority of the population was in phase “on” during urinary tract infections, the majority were in phase “off” in the respiratory and intestinal tract (387). The authors suggested that different environmental selection pressures act in distinct host niches and are involved in the selection of either fimbriated or non-fimbriated states of *K. pneumoniae*. Additionally, further work showed that *in vitro K. pneumoniae fim* expression at 37ºC was greater in serial shaking cultures than static, a result contradicted by the findings of Duguid who detected greater expression in serial static cultures (102, 387).

An impressive array of type 1 fimbrial expression regulators have been described in *E. coli* including RcsAB, HbiF, FimX, CRP, cAMP, integration host factor, Lrp and H-NS (76, 356). Apart from oxygen levels, several other environmental factors that affect type 1 fimbrial expression have been identified, including temperature, pH, osmolarity and glucose levels (130, 146, 355). It is likely that similar factors affect expression in *K. pneumoniae*, however these remain to be confirmed. Unique to *K. pneumoniae* is the FimK regulator. Rosen *et al.* showed that a *fimK* knockout of strain TOP52 possessed increased numbers of fimbriae, as observed by transmission electron microscopy, and increased biofilm formation (331). However, opposite results have

been obtained for other *K. pneumoniae* isolates (C. Struve, personal communication), possibly due to FimK-mediated regulation of *mrk* expression. The interplay between regulatory elements of surface exposed structures in *K. pneumoniae*, such as the type 1 and type 3 fimbriae, and the capsule is likely to involve a complex network of both shared and exclusive regulatory elements in addition to physicochemical interference, the latter which was identified in an elegant experiment by Schembri and colleagues (349).

Kpc fimbria-encoding operon was first identified in the genome sequence of NTUH-K2044 and was subsequently determined to be involved in biofilm formation on abiotic surfaces by heterologous expression of *kpc* in *E. coli* HB101 (443) (Figure 1-6). The *kpc* operon was found in 33 % of *K. pneumoniae* strains examined and was identified as a putative virulence factor of PLAs due to its skewed prevalence in K1 serotype isolates. Similar to the *fim* operon, *kpc* expression is controlled by inversion of a promoter region upstream of *kpcA*, a process mediated by the KpcI recombinase. The *in vivo* role of Kpc fimbriae has yet to be examined (443).

In 1992, Darfeuille-Michaud *et al.* identified the non-fimbrial adhesin CF29K, a 29 kDa protein involved in adhesion to Caco-2 human epithelial colorectal adenocarcinoma cells (87). It was coded for by a 185 kb conjugative plasmid that also possessed genes for antibiotic resistance and aerobactin synthesis and transport (87). *K. pneumoniae* strains can also produce KPF-28 fimbriae, which are 4 – 5 nm in diameter and 0.5 – 2.0 µm in length, and mediate adherence to Caco-2 cells (251). Similar to CF29K, KPF-28 fimbriae are coded for by an ESBL-positive conjugative plasmid.

Many of the remaining putative *K. pneumoniae* adhesins and CU fimbrial operons have not been characterised. It has been proposed that the collection of fimbriae and adhesins encoded by a bacterium are involved in specific-aspects of surface adhesion and tissue tropism (211). Therefore to fully understand the nature of the interaction

between *K. pneumoniae* and humans it is important to characterise and understand the contribution of these fimbriae and adhesins to bacterial colonization and virulence.

1.4 KpGI-5: a novel *K. pneumoniae* GI harbouring a putative fimbrial locus

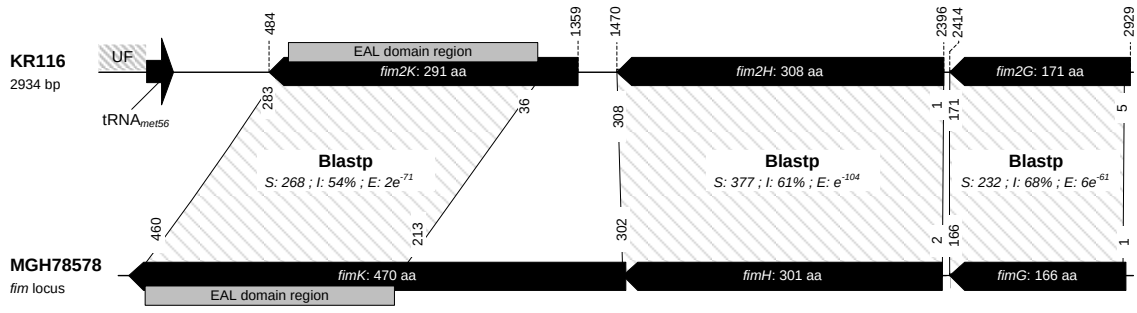
In earlier work, five tRNA gene insertion hotspots were screened for strain-specific DNA in sixteen clinical *K. pneumoniae* isolates (1). *K. pneumoniae* KR116, a strain isolated from the blood of a patient with pneumonia and neutropenic septicaemia, was found to possess an 'occupied' *met56* tRNA locus in a tRIP-PCR screen using primers PR601 and PR602, which were designed to amplify across an empty *met56* tRNA hotspot. SGSP-PCR primer walking from the conserved *met56* UF yielded ~3 kb of novel island UA sequence, whilst primer walking from the DF was repeatedly unsuccessful (1). This novel island was called KpGI-5.

Analysis of the 3 kb of UA sequence from KpGI-5 showed that although no GenBank nucleotide homologs could be identified using BLASTn (default search parameters), the UA possessed three genes with similar CDS organisation and predicted products as the *K. pneumoniae fim* operon (section 1.3.4) (Figure 1-8) (1). This putative operon was therefore called *fim2*. Two novel putative proteins, Fim2H and Fim2G, were 68 % and 61 % identical to their Fim homologs, respectively (1). This locus also encoded a homolog of the unique *K. pneumoniae* FimK protein, Fim2K, with 54 % BLASTp identity (1). Although both contained an EAL domain, the N-terminal of Fim2K lacked a putative helix-turn-helix domain.

Proteins containing EAL domains are often involved in the hydrolysis of c-di-GMP, an important bacterial intracellular messenger discovered by Ross *et al.* in 1987 as a result of its role in *Gluconacetobacter xylinus* cellulose biosynthesis regulation (334). EAL domain encoding phosphodiesterase (PDEs) enzymes degrade c-di-GMP into

linear pGpG (5'-phosphoguanyl-(3'-5')-guanosine), although inactive EAL domains have been described (319, 350, 404). Reductions in intracellular c-di-GMP levels are counteracted by the synthesis of c-di-GMP from two molecules of GTP, a reaction catalysed by GGDEF domain-containing diguanylate cyclases (DGCs) (350). Characterisation of PDEs and DGCs has revealed c-di-GMP is involved in important cellular functions including different forms of motility (275, 365), protein degradation and cell cycle progression (101), synthesis of cellulose, secondary metabolites, adhesins and exopolysaccharide matrix components (118, 134, 350), fimbrial expression (331, 381), and virulence regulation in important pathogens (220, 250, 379).

Figure 1-8 :- Schematic of the *K. pneumoniae* KR116 KpGI-5 upstream arm



Sequence analysis of ~3.0 kb sequence obtained from the UA of KpGI-5, the GI associated with the *met56* tRNA gene of *K. pneumoniae* KR116 (1). The sequence was not related to any known nucleotide sequences deposited in GenBank. Interrogation of predicted products with BLASTp revealed that the products of these putative CDSs were similar to fimbrial subunits encoded by the *K. pneumoniae* MGH78578 *fim* operon. Although the sizes of Fim2H and Fim2G were comparable to their counterparts, Fim2K was lacking the N-terminal putative helix-turn-helix region found in FimK. Cross-hatched boxes linking KR116 and MGH78578 indicate related protein regions, and the BLASTp results and coordinates are shown in the middle and corners, respectively. Black arrows indicate CDSs and the sizes of corresponding protein products are specified. The grey box represents and EAL domain. UF, upstream flank; S, score; I, identity; E, expect value.

The function(s) and mechanism of action of the putative PDEs FimK and Fim2K are still largely unknown. Although both have a C-terminal EAL domain, it is unknown whether they possesses PDE activity. Additionally, since Fim2K does not have a putative N-terminal helix-turn-helix domain, it seems unlikely that FimK and Fim2K function in a

similar manner. The role of Fim2K and potential interactions with FimK and/or c-di-GMP remains to be examined.

The gene content downstream of the KpGI-5 UA is unknown, but it is hypothesised that more genes encoding fimbrial subunits will be identified. A PCR screen of 42 *K. pneumoniae* strains for the region encoding *fim2G* to *fim2K* identified three other strains that possessed this region (KR173, KR396 and KR399) (J. J. van Aartsen and K. Rajakumar, unpublished data). When this is taken into account with the fact that the putative products encoded by the KpGI-5 UA show parallels to subunits of the well-characterised *K. pneumoniae* type 1 fimbria, it is likely that this locus plays an important role in either virulence, host colonization or environmental persistence.

1.5 Aims and objectives of this study

This project was focussed on the novel *K. pneumoniae* GI KpGI-5, which was found to encode part of a type 1-like fimbrial operon. The existence of this genetic locus and its role in infection and colonization were unknown at the commencement of this project, hence the study was aimed to characterise these aspects. We hypothesised that KpGI-5 possessed a full CU fimbrial operon that encoded functional fimbriae which conferred increased pathogenic and/or colonization potential to host *K. pneumoniae* strains and that Fim2K was a functional PDE enzyme involved in the regulation of biofilm formation and fimbrial expression.

This project aimed to:

1. ***Examine and refine techniques used for the genetic manipulation of K. pneumoniae***

   Both suicide vector- and lambda Red-based allelic exchange were examined as tools for allelic exchange in *K. pneumoniae*. Additionally, the application of site-

specifically integrated mini-Tn*7* transposons as a method for tagging strains and single-copy complementation in *K. pneumoniae* was investigated.

2. ***Obtain the entire sequence of KpGI-5***

   An antibiotic resistance marker was inserted into the KpGI-5 UA using allelic exchange. A fosmid library was generated and marker rescue was used to select for fosmid clones harbouring the tagged GI. Fosmid(s) were sequenced and the sequence was analysed using a selection of bioinformatics tools.

3. ***Investigate the prevalence and structure of KpGI-5 and KpGI-5-like islands in* Klebsiella *strains***

   Over 150 *K. pneumoniae* strains isolated from different sources and geographical locations were screened for the *fim2* locus using PCR. The KpGI-5 sequence was then used as a template for PCR-based mapping to understand strain-to-strain variation of KpGI-5-like island architecture in *fim2*-positive strains.

4. ***Characterise fim2, and identify its associated putative Fim2 fimbria***

   Isogenic *fim*, *fim2* and *fim fim2* mutants were generated using allelic exchange and examined using three *in vivo* murine infection models: large intestine colonization, urinary tract infection and respiratory tract infection. The mutants were also examined *in vitro* using biofilm formation and adhesion assays. Additionally, the *fim2* operon was heterologously expressed in *E. coli* HB101 and the *fim2*-negative *K. pneumoniae* strain C3091Δ*mrk*Δ*fim*, and the recombinant strains were examined by transmission electron microscopy for fimbrial production. The ability to mediate biofilm formation and guinea pig red blood cell agglutination was also investigated. Additionally, *fim2* expression profiles were examined in bacteria growing in different media and under distinct conditions.

5. Characterise Fim2K, and its partially characterised homolog FimK

Isogenic *fimK*, *fim2K* and *fimK fim2K* mutants were constructed using allelic exchange and examined using *in vitro* biofilm formation assays and agglutination assays. The effect of *fimK* and *fim2K* overexpression in *K. pneumoniae* KR116 was also examined using biofilm formation assays. To determine whether FimK and Fim2K possessed functional EAL domains, the ability of *fimK* and *fim2K* to modify prodigiosin production and motility in *Serratia* sp. was examined using a previously described model (118). Attempts were also made to express, purify and biochemically characterise FimK and Fim2K.

Chapter 2. Materials and methods

2.1 Bacterial strains and plasmids

K. pneumoniae isolates were obtained from clinical samples at Leicester Royal Infirmary and were stored at -20ºC / -80ºC in brain heart infusion broth with 30 % glycerol. Strains were routinely grown at 37ºC using lysogeny broth (LB) or agar (LA), unless otherwise specified. When required, *E. coli* growth medium was supplemented with 100 µg/ml ampicillin, 50 µg/ml kanamycin, 30 µg/ml chloramphenicol, 13 µg/ml gentamicin, 10 µg/ml tetracycline, 140 µg/ml hygromycin, 50 µg/ml streptomycin or 30 µg/ml apramycin. When required, *K. pneumoniae* strains were routinely grown in medium supplemented with 250 µg/ml ampicillin, 50 µg/ml kanamycin, 30 µg/ml chloramphenicol, 9 µg/ml gentamicin, 15 µg/ml tetracycline, 50 µg/ml streptomycin or 50 µg/ml apramycin, unless otherwise specified.

Bacterial strains used and constructed in this work are summarised in Appendix 3 and Appendix 4, respectively. Plasmids used and constructed in this work are summarised in Appendix 5 and Appendix 6, respectively.

2.2 DNA-related techniques and methods

2.2.1 Polymerase chain reaction (PCR)

The majority of PCR products in this work were generated using GoTaq DNA polymerase (Promega). Long-range PCR products were amplified using Phire Hot Start DNA polymerase (Finnzymes). KOD Hot Start DNA polymerase (Novagen) was used when high fidelity amplification of short- or long-range PCR products was required. All enzymes were used according to manufacturer's instructions. Representative cycling conditions for each enzyme type is shown in Table 2-1. Cycling conditions were altered according to primer melting temperature (T_m) and predicted amplicon size. Touchdown

cycling conditions are also included in Table 2-1 and were used to reduce non-specific bands when necessary (96).

Table 2-1 :- PCR cycling conditions of polymerases used in this work

A) GoTaq DNA polymerase

	Standard			Touchdown		
Denaturation	95ºC	30 s	30 cycles	95ºC	30 s	10 cycles
Annealing	(Lowest T_m)ºC	30 s		(Lowest T_m + 10) ºC ; decrease by 1ºC each cycle	30 s	
Extension	72ºC	1 min / kb		72ºC	1 min / kb	
Denaturation				95ºC	30 s	20-25 cycles
Annealing				(Lowest T_m)ºC	30 s	
Extension				72ºC	1 min / kb	

B) Phire Hot Start DNA polymerase

	Standard			Touchdown		
Denaturation	98ºC	5 s	30 cycles	98ºC	5 s	10 cycles
Annealing	(Lowest T_m)ºC	5 s		(Lowest T_m + 10) ºC ; decrease by 1ºC each cycle	5 s	
Extension	72ºC	20 s / kb		72ºC	20 s / kb	
Denaturation				98ºC	5 s	20-25 cycles
Annealing				(Lowest T_m)ºC	5 s	
Extension				72ºC	20 s / kb	

C) KOD Hot Start DNA polymerase

	Standard			Touchdown		
Denaturation	95ºC	20 s	30 cycles	95ºC	20 s	10 cycles
Annealing	(Lowest T_m)ºC	10 s		(Lowest T_m + 10) ºC ; decrease by 1ºC each cycle	10 s	
Extension	68ºC	20 s / kb		68ºC	20 s / kb	
Denaturation				95ºC	20 s	20-25 cycles
Annealing				(Lowest T_m)ºC	10 s	
Extension				68ºC	20 s / kb	

GoTaq DNA polymerase (A), Phire Hot Start DNA polymerase (B) and KOD Hot Start DNA polymerase (C) with their respective standard and touchdown PCR cycling programs. T_m, primer melting temperature.

2.2.2 Oligonucleotide design and synthesis

The sequences of oligonucleotides (primers) used in this project are listed in Appendix 7. Primers were designed using Perl Primer or Primer3 and synthesised by Sigma-Aldrich (249, 335). All primers were dissolved in nH_2O at 10 pmol/µl and stored at -20ºC.

2.2.3 Colony PCR

A single bacterial colony was resuspended in 35 µl of nH_2O and heated at 100ºC for 10 min. Cell debris was then pelleted at 12 000 x g for 1 min and 1 µl of the supernatant was used as template in PCR.

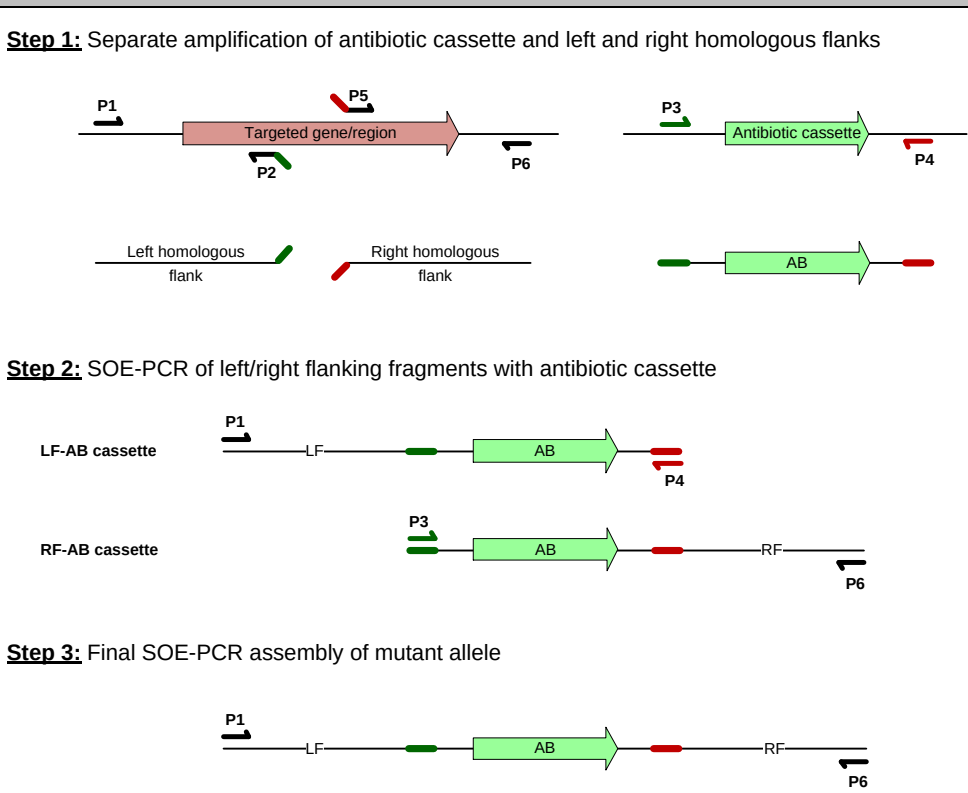
2.2.4 Splice overlap extension-PCR: mutant allele construction

Mutant alleles consisted of a left homologous flank (LF), an antibiotic resistance cassette (AB) and a right homologous flank (RF). Flanks ranged between 280 bp and 1 kb in size and were homologous to sequences upstream and downstream of the region targeted for deletion. Splice overlap extension-PCR (SOE-PCR) was used to simplify the construction of mutant alleles by providing a method whereby DNA molecules could be recombined at specific junctions, regardless of the nucleotide sequence or restriction sites present (167).

Initially, primer pairs P1/P2, P3/P4, and P5/P6 were used to separately amplify PCR fragments corresponding to the LF, AB and RF, respectively (Figure 2-1). All fragments were then gel purified, as described in section 2.2.6. Primers P2 and P5 were designed so that their ends contained ~20 bp of sequence complementary to the AB fragment primers P3 and P4, respectively. When products amplified using these primers are mixed, denatured and annealed in SOE-PCR, the complementary fragment ends anneal and act as primers for subsequent DNA polymerase extension reactions, producing a spliced product where two fragments have been merged. This phenomenon was exploited in step 2 (Figure 2-1). Briefly, 10 ng of LF fragment was

mixed with 10 ng of AB fragment and PCR amplified using primers P1 and P4, producing the LF-AB cassette. PCR cycling conditions were as mentioned in section 2.2.1, but primers were excluded from the initial reaction mixture and added during the extension phase of the 5th cycle. In a similar manner, the RF-AB cassette was produced using primer pair P3/P6 and fragments RF and AB. The LF-AB and RF-AB fragments were then gel purified and used as template for step 3. Finally, 10 ng of both the LF-AB and RF-AB fragments were mixed and subjected to amplification using primer pairs P1 and P6, producing an SOE-PCR assembled mutant allele.

Figure 2-1 :- Schematic of SOE-PCR for creating mutant alleles



SOE-PCR consists of three sequential steps. In **step 1** fragments corresponding to the left flank (LF) and right flank (RF) of the region targeted for allelic exchange and a selectable antibiotic resistance cassette (AB) are PCR amplified separately. The green line and red line on internal primers P2 and P5 correspond to 5' overlaps where sequence corresponds to the reverse complement of P3 and P4, respectively. In **step 2** the LF and RF fragments are spliced to the AB cassette using SOE-PCR to create the LF-AB and RF-AB cassettes. In **step 3** the final SOE-PCR assembly takes place by splicing LF-AB and RF-AB to produce the desired mutant allele.

Using the described protocol sixteen mutant alleles were produced, targeting nine different loci. Details regarding the construction of these alleles are presented in Table 2-2. Mutant alleles were either used directly for lambda Red-mediated allelic exchange (section 2.6.4) or cloned into lambda *pir*-based suicide vectors for plasmid-based allelic exchange (section 2.6.2).

2.2.5 Genomic and plasmid DNA extraction

Genomic DNA (gDNA) was routinely extracted from 250 µl of overnight culture using the 5Prime ArchivePure DNA Purification Kit (VWR). Plasmid DNA was routinely isolated from 5 ml overnight cultures using either the OmegaBiotek EZNA Plasmid Mini Kit I (VWR) or the GenElute Plasmid Miniprep Kit (Sigma-Aldrich). All kits were used according to manufacturer's protocol.

2.2.6 Gel electrophoresis, DNA purification and DNA sequencing

Gel electrophoresis for the visualization of DNA fragments was routinely performed in agarose gels made from 1x TAE buffer containing 0.5 µg/ml ethidium bromide. Agarose concentration varied from 0.7 % to 2.0 % depending on the DNA fragment size analysed. λ-*Hind*III, λ mix marker 19 or GeneRuler^{TM} (Fermentas) were used as DNA standards. When required, DNA fragments were excised from agarose gels with minimal ultraviolet light exposure using a clean scalpel and extracted using either the OmegaBiotek EZNA Gel Extraction Kit (VWR) or YORBIO Gel/PCR DNA purification kit (Geneflow). Sanger sequencing of PCR amplicons and plasmid inserts was routinely performed by either Eurofins MWG Operon or GATC Biotech. Roche 454 pyrosequencing was performed at the University of Leicester NUCLEUS Genomics Core Facility. The KpGI-5 sequence has been deposited in GenBank: JN181158.

Table 2-2 :- Mutant allele constructs made using SOE-PCR

Allele [a]	***fimK*::gen+**	***fimK*::gen-**	***fim2K*::kan+**	***fim2K*::kan-**	***fim*::gen+**	***fim*::gen-**	***fim*$_{KR161}$::gen**
Target	*K. pneumoniae* *fimK* gene	*K. pneumoniae* *fimK* gene	*K. pneumoniae* *fim2K* gene	*K. pneumoniae* *fim2K* gene	*K. pneumoniae* KR116 *fim* operon	*K. pneumoniae* KR116 *fim* operon	*K. pneumoniae* KR161 *fim* operon
Size (bp)	2965	2965	2368	2368	~2800	~2800	~2800
RE site [b]	*Not*I	*Not*I	*Xba*I	*Xba*I	*Not*I	*Not*I	*Not*I
P1	PR640	PR640	PR629	PR629	PR680	PR680	PR680
P2	PR642	PR643	PR627	PR625	PR681	PR701	PR681
Template	KR116	KR116	KR116	KR116	KR116	KR116	KR161
Size (bp)	984	984	538	538	~930 [c]	~930 [c]	~930 [c]
P3	GmF	GmR	#CF	#CR	GmF	GmR	GmF
P4	GmR	GmF	#CR	#CF	GmR	GmF	GmR
Template	pUC18R6k-mini-Tn7T-Gm	pUC18R6k-mini-Tn7T-Gm	pRT733	pRT733	pUC18R6k-mini-Tn7T-Gm	pUC18R6k-mini-Tn7T-Gm	pUC18R6k-mini-Tn7T-Gm
Cassette	*aacC1* [d]	*aacC1* [d]	*aph* [e]	*aph* [e]	*aacC1* [d]	*aacC1* [d]	*aacC1* [d]
FRT sites	yes	yes	no	no	yes	yes	yes
Size (bp)	1054	1054	1361	1361	1054	1054	1054
P5	PR645	PR644	PR628	PR626	PR703	PR702	PR703
P6	PR641	PR641	PR630	PR630	PR704	PR704	PR704
Template	KR116	KR116	KR116	KR116	KR116	KR116	KR116
Size (bp)	976	976	509	509	908	908	908

Table 2-2 :- Mutant allele constructs made using SOE-PCR

Allele [a]	***fim*$_{KR162}$::gen**	***fim*$_{KR173}$::gen**	***fim*::tet**	***fim2*::gen**	***fim2*::kan**	***fimH*::tet**	***fim2H*::kan**
Target	*K. pneumoniae* KR162 *fim* operon	*K. pneumoniae* KR173 *fim* operon	*K. pneumoniae* *fim* operon	*K. pneumoniae fim2* operon	*K. pneumoniae* *fim2* operon	*K. pneumoniae* *fimH* gene	*K. pneumoniae* *fim2H* gene
Size (bp)	~2800	~2800	Constructed previously by Dr. C. Struve and used in lambda Red recombination to make *K. pneumoniae* C3091Δ*mrk*Δ*fim* (388)	2735	3147	2237	1977
RE site [b]	*Not*I	*Not*I		*Not*I	None	None	None
P1	PR680	PR680		PR931	PR1257	PR1339	PR1344
P2	PR681	PR681		PR932	PR1258	PR1340	PR1345
Template	KR162	KR173		KR116	KR116	KR116	KR116
Size (bp)	~930 [c]	~930 [c]		856	841	305	281
P3	GmF	GmF		GmF	Kn1	GmF	Kn1
P4	GmR	GmR		GmR	Kn2	GmR	Kn3
Template	pUC18R6k-mini-Tn7T-Gm	pUC18R6k-mini-Tn7T-Gm	Amplified using Kp_opfim_fwd_Xba and Kp_opfim_rev_Xba	pUC18R6k-mini-Tn7T-Gm	pKD4	pJTOOL-4a	pKD4
Cassette	*aacC1* [d]	*aacC1* [d]		*aacC1* [d]	*aph* [e]	tet [g]	*aph* [e]
FRT sites	yes	yes		yes	yes	yes	yes
Size (bp)	1054	1054	tet cassette [f] ~ 3.1 kb	1054	1496	1707	1477
P5	PR703	PR703		PR933	PR1259	PR1341	PR1346
P6	PR704	PR704		PR934	PR1260	PR1342	PR1347
Template	KR116	KR116		KR116	KR116	KR116	KR116
Size (bp)	908	908		874	850	274	262

Table 2-2 :- Mutant allele constructs made using SOE-PCR

Allele [a]	mrk_{C3091}::kan	mrk_{KR116}::kan	KpGI-5::gen	$phe55_{KR173}$DA::gen
Target	*K. pneumoniae* C3091 *mrk* operon	*K. pneumoniae* KR116 *mrk* operon	*K. pneumoniae* KR116 KpGI-5 island	Downstream arm of KR173 GI at *phe55*
Size (bp)		2735	2778	2328
RE site [b]		*Not*I	*Not*I	*Not*I
P1	Constructed previously by Dr. C. Struve and used in lambda Red recombination to make *K. pneumoniae* C3091Δ*mrk*Δ*fim* (388)	PR1059	PR931	PR347
P2		PR1060	PR932	PR348
Template		KR116	KR116	KR173
Size (bp)		742	856	630
P3		GmF	GmF	GmF
P4		GmR	GmR	GmR
Template	Amplified using mrk_lf_NotI and mrk_rr_NotI	pUC18R6k-mini-Tn7T-Km	pUC18R6k-mini-Tn7T-Gm	pUC18R6k-mini-Tn7T-Gm
Cassette		*aph* [e]	*aacC1* [d]	*aacC1* [d]
FRT sites		yes	yes	yes
Size (bp)	*aph* [e] cassette	1208	1054	1054
P5		PR1061	PR935	PR350
P6	~ 3.1 kb	PR1062	PR936	PR349
Template		KR116	KR116	KR173
Size (bp)		834	917	669

[a] + and – correspond to identical or opposite orientation, respectively, of the central antibiotic resistance cassette relative to the orientation of the CDS/operon targeted for allelic exchange.

[b] Restriction enzyme site incorporated into the 5' end of primers P1 and P6.

[c] The exact sizes of the products from these reactions are unknown as a result of strain-to-strain variation at this end of the *fim* operon.

[d] Derived from the *aph* (aminoglycoside-3'-phosphotransferase) gene on transposon Tn*5*.

[e] Derived from the *aacC1* (aminoglycoside-3'-acetyltransferase) gene on transposon Tn*1696.*

[f] Derived from the omega interposon tetracycline resistance cassette in pUT-mini-Tn5-Tc.

[g] Tetracycline efflux protein gene.

2.2.7 Restriction enzyme digestion of DNA

Restriction enzymes were obtained from Promega, NEB and Roche, and were performed according to manufacturer's protocols. The reaction volume was routinely between 20 – 50 µl and was altered according to the enzyme chosen and DNA amount. When two enzymes were required for digestion, either the reaction was set up in a compatible buffer or was performed serially. When reactions were required for downstream applications, enzymes were EDTA- or heat-inactivated as per manufacturer's instructions.

2.2.8 Dephosphorylation of DNA

Plasmid DNA linearised by one restriction enzyme is prone to self-ligation in subsequent ligation reactions. This can be prevented by adding alkaline phosphatase which removes the 5' phosphate groups from DNA. Thermosensitive alkaline phosphatase (Promega) was added to plasmid restriction digests at 1 U/µg DNA. When the reaction was complete, TSAP was inactivated at 74ºC for 15 min.

2.2.9 Ligation of DNA fragments

Ligation reactions were performed in a final volume of 10 – 20 µl containing 3 U of T4 DNA ligase (Promega) and 1 – 2 µl of 10x ligation buffer. The reaction was incubated first at room temperature for 60 min and then overnight at 4ºC, and subsequently transformed into *E. coli*.

2.2.10 Fosmid cloning and marker rescue

A fosmid library of chromosomal DNA from KR116Δ*fim2K*::kan was constructed using the Copy Control Fosmid Library Production kit (Epicentre), with some minor modifications. Briefly, 2.5 µg of genomic DNA was sheared to ~40 kb fragments by pipetting through a 200 µl tip. After end repair the DNA was ligated into pCC2FOS and

packaged into phages using MaxPlax Lambda Packaging Extracts (Epicentre) which were used to infect *E. coli* EPI300-T1R cells. Marker rescue of kanamycin resistant fosmid clones bearing Δ*fim2K*::kan was performed by plating infected EPI300-T1R cells on LB plates supplemented with chloramphenicol and kanamycin.

2.3 RNA-related techniques and methods

2.3.1 RNA extraction and cDNA library preparation

Total RNA was prepared from KR2107 grown for 16 h in LB (37°C, 200 rpm) using the Norgen Total RNA Purification Kit. The Ambion TURBO DNA-free kit was used to remove residual DNA from RNA preparations and a cDNA library was synthesised using the QuantiTect Reverse Transcription kit (Qiagen) as recommended by the manufacturer. An identical reaction without reverse transcriptase was performed to assess DNA contamination. The above steps were repeated with KR2107 growing for 16 h (37ºC, 200 rpm) in M9, AUM, RPMI 1640 (GIBCO), RPMI 1640 with 10 % FCS, King's B and *Acanthamoeba* medium (Appendix 1). Additionally, the transcriptional profile of KR2107 grown in LB for 16 h at 25ºC (200 rpm) was also assessed.

2.3.2 RNA extraction from bladder and lung tissue samples

Approximately 1/10th of each bladder (section 2.7.3) was stored in RNAlater for subsequent RNA extraction and transcript analysis. To extract bacteria from tissues, the samples were pooled and homogenised in TE buffer. After centrifugation at 450 x g for 5 min, the supernatant was removed and centrifuged at 12 000 x g for 5 min in a fresh tube. The supernatant was removed and the remaining pellet resuspended in 50 µl TE buffer. Total RNA was extracted and cDNA libraries were constructed as described above (section 2.3.1).

2.3.3 Transcriptional analysis of the *fim*, *mrk* and *fim2* gene clusters

Each cDNA library and its reverse transcriptase negative counterpart were analysed for transcripts corresponding to the *fim*, *mrk* and *fim2* gene clusters. Regions corresponding to genes *fimA*, *fimH* and *fimK* were PCR amplified using primer pairs PR1601/PR1602 (233 bp), PR1603/PR1604 (233 bp), and PR1605/PR1606 (227 bp), respectively. The linking region spanning *fimH* to *fimK* was detected using primers PR1630 and PR1631 (372 bp). Regions corresponding to genes *mrkA* and *mrkF* were PCR amplified using primer pairs PR1622/PR1623 (211 bp) and PR1624/PR1625 (222 bp), respectively. Regions corresponding to genes *fim2A*, *fim2H* and *fim2K* were PCR amplified using primers pairs PR1607/PR1608 (221 bp), PR1609/PR1610 (241 bp), and PR1611/PR1612 (220 bp), respectively. The linking regions spanning *orf10* to *fim2A* and *fim2H* to *fim2K* were detected using primer pairs PR1626/PR1627 (predicted: 380 bp) and PR1628/PR1629 (316 bp), respectively.

2.3.4 qRT-PCR analysis of the *fim*, *mrk* and *fim2* gene clusters

For *in vitro* samples 2 µl of a 10^{-1} dilution of the cDNA library and its reverse transcriptase negative control were used for quantitative real-time PCR (qRT-PCR). 2 µl of neat sample was used for *in vivo* samples. 20 µl qRT-PCR reactions were set up using SensiMix Plus SYBR (Bioline) according to manufacturer's instructions. For *rpoD* (PR1701/PR1702; 164 bp) and *fimA* (PR1703/PR1704; 157 bp) 10 pmol of each primer was added to the reaction mix, whilst 7.5 pmol was used for *fim2A* (PR1705/PR1706; 165 bp) and *mrkA* (PR1707/PR1708; 155 bp). Cycling conditions consisted of 40 cycles of 95ºC for 15 s, 60ºC for 15 s and 72ºC for 20 s. Each qRT-PCR reaction was performed in triplicate. Additionally, reaction efficiency was measured in parallel using five serial ten-fold dilutions of KR116 gDNA as template and duplicate reactions. qRT-PCR reactions were performed using the Qiagen Rotor Gene 6000. The relative quantification of *mrk*, *fim* and *fim2* transcripts versus *rpoD* was determined according to the Pfaffl equation (294).

2.4 Protein-related techniques and methods

2.4.1 SDS-PAGE

SDS-PAGE gels were made according to standard protocols (78). To normalise the number of cells loaded into gels, 1 ml of culture was centrifuged at 12 000 x g for 1 min and resuspended in *x* µl of PBS, where *x* equals OD_{600} of culture divided by 0.015. 2 µl of 5x SDS sample loading buffer (Appendix 2) was added to 8 µl of normalized sample and incubated at 100ºC for 5 min. The entire sample was then loaded into an SDS-PAGE gel and separated at 60 mA in tris-glycine electrophoresis buffer (Appendix 2). Gels were stained as per manufacturer's instructions using PhastGel Blue (Sigma) and de-stained using de-staining solution (Appendix 2).

2.4.2 Separation of soluble and insoluble protein fractions

100 ml of bacterial culture was pelleted at 12 000 x g, resuspended in 5 ml lysis buffer (Appendix 2) with lysozyme (1 mg/ml) and incubated on ice for 30 min. The suspension was then sonicated using 7 pulses of 10 s duration each. Next, fractions were separated by centrifugation at 10 000 x g at 4ºC for 30 min. The supernatant contains the soluble protein fraction and was pipetted into a fresh tube. The remaining pellet represents the insoluble protein fraction and was resuspended in 5 ml of lysis buffer. Fractions were then analysed using SDS-PAGE.

2.4.3 Western Blotting

Western blotting was performed as previously described (14)**.** Briefly, proteins were transferred to a nitrocellulose membrane by electrophoresis at 250 mA for 1 h in 1x transfer buffer (25 mM Tris.HCl, pH 8.3, 192 mM glycine, 20 % (v/v) methanol). Protein transfer was confirmed by staining with 0.2 % Ponceau red (Sigma-Aldrich) in 1 % (v/v) glacial acetic acid. Subsequently, the nitrocellulose membrane was incubated for 1 h in PBS containing 5 % skimmed milk powder and washed three times for 5 min in PBS

containing 0.05 % Tween 20. Next, the membrane was incubated for 2 h with a 1:2000 concentration of anti-polyhistidine peroxidase conjugate (Sigma-Aldrich) in PBS containing 0.05 % Tween 20 and 1 % BSA, after which it was washed three times for 5 min in PBS containing 0.05 % Tween 20. The EZ-ECL Chemiluminescence Detection Kit (Geneflow) was used to identify antibody binding. The chemiluminescence reaction was exposed to an x-ray film in a film cassette to make a permanent film record. Films were developed manually using x-ray standard developed and fixer (14)**.**

2.5 Construction of plasmids

2.5.1 pJTOOL plasmids

This section describes the construction of useful plasmid-based tools (pJTOOL plasmid series) for the genetic manipulation of *K. pneumoniae* and other Gram negative bacteria.

2.5.1.1 pJTOOL-1 and pJTOOL-3: Lambda *pir*-based suicide vectors

A 5.2 kb amplicon produced by inverse PCR with KOD HS DNA Polymerase, the primers PR648 and PR649 and pDS132 as template was cloned into pJET1.2, according to manufacturer's instructions, to produce pJTOOL-1CJ. pJTOOL-1CJ was then digested with *Not*I and re-ligated to remove the pJET1.2 backbone and produce pJTOOL-1 (GenBank: JF756692).

pJTOOL-3 (GenBank: JF756693) was constructed from both pDS132 and pBluescript II KS+. The 5.2 kb pDS132-derived fragment was amplified from pDS132 using KOD HS DNA polymerase and primers PR621 and PR622. Similarly, a 0.2 kb pBluescript II KS+ fragment was produced using primers PR623 and PR624. Both fragments were then recombined using a modified SLIC (Sequence and Ligation–Independent Cloning) method (231). Briefly, both fragments were subjected to 3'-5' T4

DNA polymerase exonuclease digestion at room temperature for 30 min, after which the reaction was stopped by adding dCTP. The two fragments were subsequently annealed by mixing at a 1:1 molar ratio and incubating for 30 min at 37°C in 1 x STE buffer supplemented with 20 ng of RecA (Epicentre). The reaction was then transformed into *E. coli* DH5α harbouring pRDH137 and recombinants were selected on plates supplemented with kanamycin and chloramphenicol. Plasmid preparations obtained from these clones contained a mixture of pJTOOL-3 and pRDH137 that were subsequently separated by transforming the mix into *E. coli* CC118λ*pir* and plating on LB plates supplemented with chloramphenicol to select for pJTOOL-3 alone.

2.5.1.2 pJTOOL-4a and pJTOOL-6a/b: FRT-flanked tetracycline and minocycline resistance cassette plasmids

SOE-PCR was used to construct an FRT-flanked tetracycline resistance cassette. Briefly, the left and right FRT-flank fragments were obtained by PCR amplification from pUC18R6k-mini-Tn7T-Gm using primers pairs PR551/PR706 (176 bp) and PR550/PR705 (182 bp), respectively. The middle tetracycline resistance cassette was PCR amplified from pUCP26 using primers PR869 and PR870 (1394 bp). Next, both the left FRT flank and right FRT flank fragments were independently spliced to the tetracycline resistance cassette using SOE-PCR and primer pairs PR551/PR870 (1548 bp) and PR550/PR869 (1553 bp), respectively. The final FRT-flanked tetracycline (1707 bp) resistance cassette was constructed in an SOE-PCR reaction containing the aforementioned spliced templates and primers PR550 and PR551. The final product was cloned into pGEM-T easy to create pJTOOL-4a.

An FRT-flanked minocycline resistance cassette was constructed in a similar manner, with the primers PR551/PR712 and PR550/PR711 being used to amplify the left (178 bp) and right (177 bp) FRT flanks, respectively, and primers PR709/PR710 being used to amplify a minocycline resistance cassette from pCLL3468 (2275 bp). The final

product (2588 bp) was amplified using PR550 and PR551 and cloned into pGEM-T easy to create pJTOOL-6a and -6b.

2.5.1.3 pJTOOL-7: a pTRC99a derivative

To create an *Nco*I/*Not*I/*Eco*RI oligonucleotide linker (27 bp), 200 pmol of oligonucleotides PR1218 and PR1219 were mixed, heated to 95ºC for 5 min and allowed to cool slowly at room temperature. The linker was then digested using *Nco*I/*Eco*RI and cloned into *Nco*I/*Eco*RI digested pTRC99a. The multiple cloning site (MCS) of pJTOOL-7 was sequenced using primers PR1216 and PR1217 to confirm the predicted sequence. Although the added *Not*I site was present in the correct location, unexpectedly an extra *Eco*RI site was also generated, although this did not affect any downstream work.

2.5.1.4 pJTOOL-8 to pJTOOL-15: plasmids with mini-Tn*7* cassettes encoding kanamycin and/or streptomycin resistance determinants

Schematics of plasmids pJTOOL-8 to pJTOOL-15 are presented in Figure 2-2 and summarised in Appendix 6. pJTOOL-8 was constructed by PCR amplifying the FRT-flanked *aph* kanamycin resistance cassette from *K. pneumoniae* C3091Δ*fimK*::kan using primers PR1231 and PR1232 (1517 bp) and cloning the resulting amplicon into *Nsi*I/*Eco*RV digested pUC18R6kT-mini-Tn7T (66). pJTOOL-9 was constructed by PCR amplifying the *aadA1a* streptomycin resistance cassette from mini-Tn*5*::Sp-Sm using primers PR1235 and PR1236 (988 bp) and cloning the resulting amplicon into *Xho*I/*Kpn*I digested pUC18R6kT-mini-Tn7T. This fragment was also cloned into *Xho*I/*Kpn*I digested pJTOOL-8 to create pJTOOL-12.

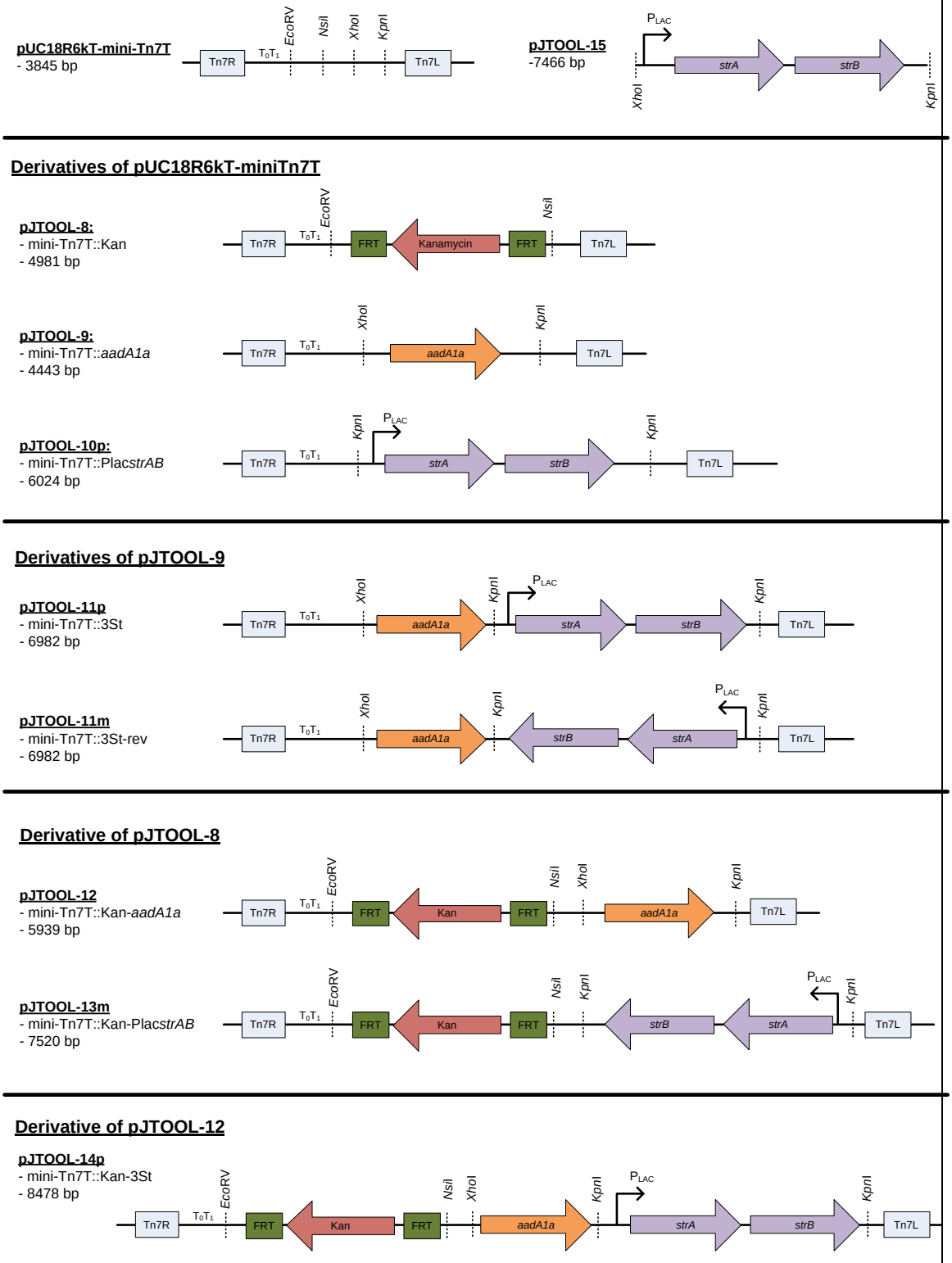
To create pJTOOL-15 a streptomycin resistance-encoding cassette from *K. pneumoniae* MGH78578 (*strAB*) was amplified using primers PR1237 and PR1238 (2060 bp) and cloned into *Kpn*I/*Xho*I digested pWSK29. pJTOOL-15 was then used as a template in a second PCR reaction using primers PR1269 and PR1238 (2551 bp) to

amplify the *strAB* cassette with the pWSK29 P_{LAC} promoter (P_{LAC}*strAB*). This was subsequently cloned into the *Kpn*I restriction site of pUC18R6kT-mini-Tn7T to create pJTOOL-10p. The P_{LAC}*strAB* fragment was also cloned into the *Kpn*I restriction sites of pJTOOL-8, pJTOOL-9 and pJTOOL-12 to create pJTOOL-13m, pJTOOL-11p/m and pJTOOL-14p, respectively.

2.5.1.5 pJTOOL-16: Flp recombinase-encoding plasmid

The 5162 bp *Sph*I/*Sac*I fragment from pFLP2 that encodes *sacB* and the *flp* recombinase gene was cloned into *Sph*I/*Sac*I digested pUCP24 to create pJTOOL-16, a counterselectable gentamicin-resistant plasmid that encodes Flp recombinase.

Figure 2-2 :- Constructed mini-Tn*7* cassettes encoding kanamycin and/or streptomycin resistance



This schematic depicts the kanamycin and/or streptomycin resistance-encoding mini-Tn*7* constructs made during this work. FRT, Flp recombinase target; Tn7R and Tn7L, right and left end of Tn7, respectively; T_o and T_1, transcriptional terminators from bacteriophage λ and *E. coli rrnB* operon, respectively.

2.5.2 Construction of pJKO suicide vectors

This section describes the construction of suicide vectors for various targets in *K. pneumoniae* by cloning SOE-PCR generated mutant alleles (section 2.2.4) into pDS132, pJTOOL-1 or pJTOOL-3. Suicide vectors were named according to the following structure: pJKO-[number][allele orientation]. Allele orientation was determined as orientation 'a' or 'b' via restriction digestion, a necessary distinction because allele orientation may affect homologous recombination efficiency (267, 268). pDS132, pJTOOL-1 and pJTOOL-3 and their derivatives were maintained in *E. coli* CC118λ*pir*.

Mutant alleles *fimK*::gen(+) and *fimK*::gen(-) were ligated into pJET1.2 and then subcloned into the *Not*I restriction site of pJTOOL-1 to construct pJKO-1a/b and pJKO-2a, respectively. Similarly, mutant alleles *fim2K*::kan(+) and *fim2K*::kan(-) were also ligated into pJET1.2 but were then subcloned into the *Xba*I restriction site of pDS132 to create pJKO-3a/b and pJKO-4a, respectively.

Vectors pJKO-5 to pJKO-9 were made by ligating *Not*I digested mutant alleles into *Not*I digested pWSK29 and then subcloning into *Not*I digested pJTOOL-1. Due to lack of sequence conservation upstream of *fimB* between *K. pneumoniae* strains[i], suicide vectors were constructed from strain-specific mutant alleles. pJKO-5 and -6 targeted the *K. pneumoniae* KR116 *fim* operon with mutant alleles fim_{KR116}::gen(+) and fim_{KR116}::gen(-), respectively. The *fim* operons of *K. pneumoniae* KR161, KR162 and KR173 were targeted by suicide vectors pJKO-7, -8 and -9 which possessed mutant alleles fim_{KR161}::gen(+), fim_{KR162}::gen(+) and fim_{KR173}::gen(+), respectively.

Plasmids belonging to the pJKO-11 to pJKO-14 series were constructed by directly ligating *Not*I digested mutant alleles into *Not*I digested pJTOOL-1 or pJTOOL-3. Plasmids pJKO-11, pJKO-12, pJKO-13 and pJKO-14b possessed the

[i] An *in silico*-PCR using primers PR680 and PR681 produces 978 bp and 932 bp fragments in *K. pneumoniae* Kp342 and MGH78578, respectively.

phe55$_{KR173}$DA::gen, *fim2*::gen, KpGI-5::gen and *mrk*$_{KR116}$::kan mutant alleles, respectively.

2.5.3 Construction of expression vectors: pJOE plasmids

A summary of the construction of each expression plasmid (pJOE vector series [**p**lasmid **J**on **o**ver**e**xpression]) can be found in Appendix 6. When necessary, the fragments cloned into expression vectors were sequenced and their integrity was confirmed by comparison to their reference sequence

2.6 Genetic manipulation

2.6.1 Preparation and transformation of electro-competent bacteria

2.6.1.1 *E. coli* and *K. pneumoniae*

A single colony was used to inoculate 5 ml of LB, which was then grown overnight at 37ºC, 200 rpm. This culture was used to inoculate fresh LB at a ratio of 1:100 (e.g. 500 µl to 50 ml) which was grown to an OD_{600} of 0.5 – 0.7 after which the cells were harvested by centrifugation (3 000 x g, 10 min, 4ºC) and gently washed three times with successively smaller volumes of ice cold 10 % (v/v) glycerol. After the final wash, the cells were resuspended in a volume of ice cold 10 % (v/v) glycerol equal to 1/100th of the initial culture volume and stored as 50 µl aliquots at -80ºC. 0.7 mM EDTA was added to the growth medium when higher competency was required for *K. pneumoniae* strains (122). Both *E. coli* and *K. pneumoniae* were transformed in 0.2 cm cuvettes using the Bio-Rad Gene Pulser system (Bio-Rad) at the following settings: 2.5 kV/cm, 25 µF and 200 Ω. Immediately post-transformation 450 – 950 µl of SOC (Appendix 1) was added and cells were incubated at 37ºC for 1 h. Transformed cells were then plated onto LA containing appropriate antibiotics and incubated at 37ºC, unless otherwise specified.

2.6.1.2 *Serratia* sp. ATCC 39006 [i]

Electrocompetent *Serratia* sp. ATCC 39006 cells were prepared as described in section 2.6.1.1, with minor modifications. Briefly, bacteria were grown at 30ºC to an OD_{600} of 0.8. After the final wash, cells were resuspended in a volume of ice cold 10 % (v/v) glycerol equal to 1/125th of the initial culture volume and 100 µl aliquots were immediately used for transformation. Prior to electroporation, the cells were incubated with DNA (~ 2 µg) on ice for 1 h. Cells were transformed in 0.2 cm cuvettes using the Bio-Rad Gene Pulser system (Bio-Rad) at the following settings: 2.5 kV/cm, 25 µF and 200 Ω. Immediately after electroporation 950 µl of SOC media was added and the mixture was incubated at 30ºC for 1.5 h, and then plated onto appropriate selective media.

2.6.2 Suicide vector-based allelic exchange

E. coli S17-1λ*pir* was implemented in conjugation as a donor strain for the mobilization of suicide vectors into *K. pneumoniae*. It encoded the broad-host-range *tra* (transfer) genes of the RP4 conjugation machinery which enabled mobilisation of plasmids that possessed the RP4 plasmid mobilisation (*mob RP4*) region, such as pUC18R6kT-mini-Tn7T and the pDS132 derivatives (296, 366) .

Briefly, recipient *K. pneumoniae* and donor *E. coli* S17-1λ*pir*/pJKO-X (where pJKO-X is a suicide vector (section 2.5.2)) strains were grown overnight in LB (plus necessary antibiotics) at 37ºC, 200 rpm. Next, overnight cultures were subcultured 1:50 (*E. coli* S17-1λ*pir*) or 1:100 (*K. pneumoniae* strains) into fresh LB without antibiotics and grown to an OD_{600} of 0.6 – 0.8. Donor and recipient strains were mixed at a 1:1 ratio (1 ml : 1 ml), pelleted at 3 000 x g for 5 min, resuspended in 20 µl of 10 mM $MgSO_4$ and spotted onto 3 cm diameter nitrocellulose filter papers (0.45 µm pore size) placed

[i] *Serratia* sp. ATCC 39006 possesses extremely low competence and this method only obtained low numbers of transformants (1 – 50 colonies)

aseptically on pre-warmed LA plates, and incubated at 37ºC overnight. When necessary, control conjugations were performed using a 1:1 ratio of donor:donor and recipient:recipient. Next, matings were resuspended from filter discs in 1.5 ml of 10 mM $MgSO_4$ and dilutions were plated onto M9 minimal medium (Appendix 1) supplemented with gentamicin (or kanamycin) and incubated overnight at 37ºC. Single colonies were then parallel patched onto LB only, LB plus chloramphenicol and LB supplemented with gentamicin (or kanamycin) to screen for integration of suicide plasmids. The merodiploid (single-crossover) genotype was confirmed by colony PCR to detect the suicide vector backbone (primers PR708 and PR637) and *aacC1* (primers PR552 and PR553) or *aph* (primers PR635 and PR636) cassette.

Next, a single merodiploid colony was grown overnight in LB and appropriate dilutions were plated on LAS media (Appendix 1), which were incubated overnight at 25ºC. Colonies were then screened by patching for the double-crossover resistance phenotype (gentamicin (or kanamycin)-resistant and chloramphenicol sensitive). Additionally, suicide vector loss was confirmed by colony PCR using primers PR708 and PR637. The preservation of either the *aacC1* or *aph* cassette was verified by PCR using primer pairs PR552/PR553 or PR635/PR636, respectively. Importantly, the introduced antibiotic resistance cassettes were definitively mapped to the targeted locus by PCR analysis using either an antibiotic resistance cassette-specific primer and a second primer that annealed to a flanking region just outside that present on the suicide vector, or two primers which annealed to left and right flanking regions not taking part in allelic exchange (Table 2-3).

Table 2-3 :- Primers used for PCR mapping: suicide vector-based allelic exchange

	Primer 1	Primer 2	Expected size
***fimK*::gen+**	PR654	PR641	3.5 kb
***fim2K*::kan+**	PR601	PR635	2.2 kb
***fim*$_{KR116}$::gen+**	PR639	PR550 or PR552	2.3 kb or 2.2 kb
***fim*$_{KR161}$::gen+**	PR639	PR550 or PR552	2.3 kb or 2.2 kb
***fim*$_{KR162}$::gen+**	PR639	PR550 or PR552	2.3 kb or 2.2 kb
***fim*$_{KR173}$::gen+**	PR639	PR550 or PR552	2.3 kb or 2.2 kb
***fim2*::gen**	PR1103	PR1104	3.4 kb
KpGI-5::gen	PR1103	PR1105	3.4 kb

2.6.3 Site-specific transposition of mini-Tn*7* elements

Plasmid-based mini-Tn*7* elements are flanked by Tn*7*L and Tn*7*R sites and can be site-specifically integrated downstream of *glmS* into the *att*Tn*7* site when the site-specific Tn7 transposition machinery (TnsABCD) is present (66, 293). In the method described by Choi *et al.*, the *tnsABCD* operon is provided using plasmid pTNS2.

Conjugal transfer of mini-Tn*7* delivery plasmids into *K. pneumoniae* was obtained by four-parental mating. Briefly, a mix consisting of a 1:1:1:1 ratio (100 µl of each) of *E. coli* DH5α*λpir*/pTNS2, *E. coli* HB101/pRK2013, *E. coli* CC118*λpir* harbouring a mini-Tn*7* delivery plasmid and *K. pneumoniae* recipient strain was prepared and conjugated overnight as described in section 2.6.2. When necessary, a control conjugation lacking *E. coli* DH5α*λpir*/pTNS2 was performed to identify background breakthrough growth and spontaneous transposition/recombination events. Next, matings were resuspended from filter discs in 1.5 ml of 10 mM $MgSO_4$ and dilutions were plated onto M9 minimal medium (Appendix 1) supplemented with appropriate antibiotics. Site-specific mini-Tn*7* transposition events were confirmed by PCR amplification across *att*Tn*7* using primers PR1307 and PR1308 (0.4 kb if insertion is not present).

2.6.4 Lambda Red recombination-based allelic exchange

Lambda Red recombination-based allelic exchange relies on the greatly enhanced rates of recombination seen when the phage λ Red recombinase system (*γ, β, exo*) is overexpressed in bacteria (88). When linear DNA fragments with homology to the bacterial chromosome, such as SOE-PCR generated mutant alleles (section 2.6.4), are transformed, homologous recombination can take place and chromosomal regions can be disrupted and/or deleted (58, 88, 268).

Initially, plasmid pKOBEG-Apra (58) was transformed (section 2.6.1.1) into *K. pneumoniae* KR2107 (or any other strain targeted for mutagenesis) and isolated on LA supplemented with apramycin and grown at 30ºC. The presence of pKOBEG-Apra was confirmed by colony PCR and primers PR1179 and PR1180 (1960 bp). A single colony was inoculated in 5 ml of LB plus apramycin and grown overnight at 30ºC, 200 rpm. This culture was inoculated at a ratio of 1:100 (e.g. 500 µl to 50 ml) into fresh LB supplemented with apramycin and 0.2 % arabinose. The culture was grown at 30ºC to an OD_{600} of 1.0, after which electrocompetent cells were prepared and transformed (section 2.6.1.1) with 1 – 3 µg of SOE-PCR generated mutant allele (section 2.2.4). After non-selective outgrowth for 1 h at 37ºC, the transformation mixture was plated onto LA containing appropriate antibiotics and incubated at 37ºC. Apart from PCR screening for loss of the temperature sensitive plasmid pKOBEG-Apra, introduced mutant alleles and antibiotic resistance cassettes were also mapped by PCR analysis to their target loci as described in section 2.6.2. Primers used for PCR mapping are summarised in Table 2-4.

Table 2-4 :- Primers used for PCR mapping: lambda Red recombination

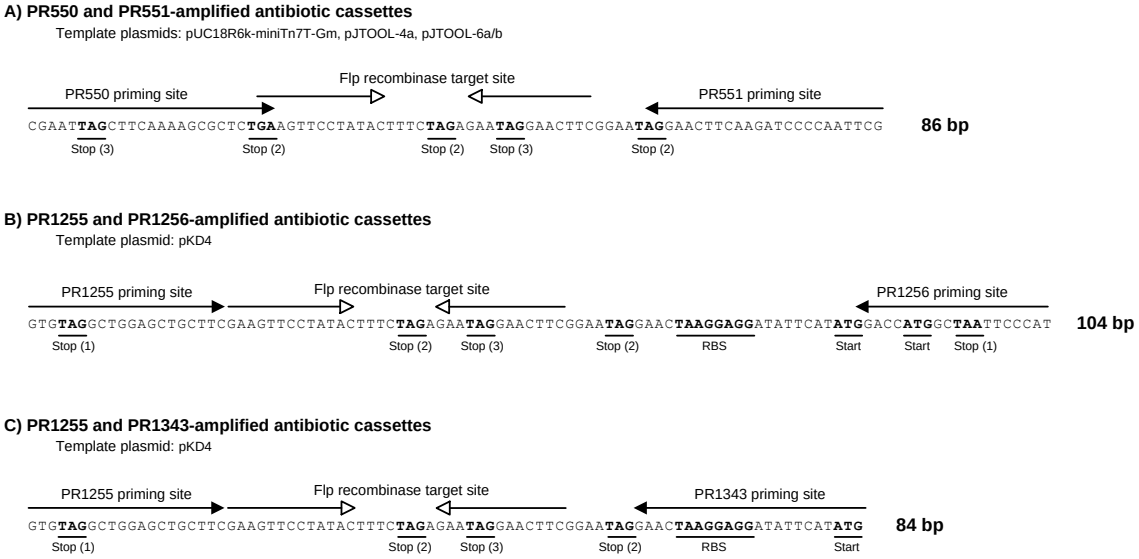
	Primer 1	Primer 2	Expected size
***fim*::tet**	PR1253	PR1148	0.8 kb
	PR1254	PR1149	1.2 kb
***fim2*::kan**	PR1103	PR1256	2.6 kb
***fimH*::tet**	PR1144	PR1145	2.5 kb
***fim2H*::kan**	PR669	PR1343	2.2 kb
	PR615	PR1255	2.7 kb
***mrk*$_{C3091}$::kan**	PR1363	PR1059	3.0 kb

2.6.5 Flp-mediated FRT recombination

A number of SOE-PCR generated mutant alleles have internal FRT (Flp recombinase target)-flanked antibiotic resistance cassettes which allow for the deletion of resistance cassettes by plasmid-based expression of Flp recombinase (163). Flp causes site-specific recombination between two FRT sites with subsequent deletion of the dividing sequence, resulting in an 'FRT scar'. The properties of FRT scars that were generated in this work are shown in Figure 2-3.

FRT-flanked antibiotic resistance cassettes (Table 2-2) were deleted as follows. Electrocompetent cells were prepared (section 2.6.1) and transformed with 10 ng of pFLP2 and grown on LB plates supplemented with ampicillin. A single ampicillin resistant colony was then streaked to single colonies on LAS plates and grown at 37ºC to cure pFLP2. Next, single colonies were screened by patch plating for loss of ampicillin resistance and the resistance associated with the FRT-flanked antibiotic resistance cassette. The genotype of antibiotic sensitive colonies was then confirmed using primers to screen for loss of pFLP2 (PR554/PR555) and the FRT-flanked antibiotic resistance cassette (see Table 2-2).

Figure 2-3 :- FRT scars left after Flp-mediated FRT recombination



This schematic details FRT scars left after Flp-mediated FRT recombination of FRT-flanked antibiotic resistance cassettes amplified using primers PR550/PR551 **(A)**, PR1255/PR1256 **(B)** and PR1255/PR1343 **(C)**. Although FRT scars are generated, they are benign and do not contain promoter sequences or transcriptional terminators. SOE-PCR generated mutant alleles can therefore be designed and constructed so that in-frame deletion mutants can be generated (68). Arrows with open arrowheads show the FRT site inverted repeats. Arrows with closed arrow heads show primer binding sites. Ribosomal binding site (RBS) and start codons are marked. Stop codons are also highlighted with their corresponding translational frame in brackets.

2.7 Other methods

2.7.1 Microtitre plate-based biofilm assay

Biofilm assays were performed using a modified microtitre plate-based method (283). Briefly, strains were grown for 16 h (37°C, 200 rpm) in 5 ml LB broth or M9 media (plus antibiotics if necessary) and were subcultured 1:100 into 96-well polystyrene (NUNC) or polyvinyl chloride (BD Falcon) microtitre plates with each well containing 100 µl of LB media supplemented with IPTG and ampicillin, if required. Plates were incubated statically for 24 or 48 h at 30°C or 37°C, after which the medium was removed and the plate was washed once with dH_2O. 125 µl of 0.1 % (v/v) crystal violet was added to each well and left to stain for ten minutes. The plate was then washed twice with dH_2O.

After drying thoroughly, 200 µl of 95 % ethanol was added to each well and the absorbance was measured at 595 nm (BioRad Model 680 Microplate reader). Each strain was tested in eight wells and three replicate experiments were performed.

2.7.2 Tissue culture & adhesion assays

Quantitative assessments of bacterial adhesion were performed using human HCT-8 ileocaecal and 5637 bladder epithelial cells. HCT-8 cells were subcultivated (1:10) twice a week in RPMI 1640 medium containing 25 mM HEPES, 2 mM glutamine, 1 mM pyruvate, 10 % FCS, 0.002 % neomycin and 0.01 % streptomycin. 5637 cells were cultivated similarly but no pyruvate was added to the medium.

Epithelial cells were seeded into two 24-well tissue culture plates (NUNC) and grown to confluent monolayers. After carefully washing each well three times with warm PBS, 1 ml of RPMI 1640 was added and inoculated with ~2×10^6 CFU from an overnight culture (37°C, 200 rpm). Plates were incubated for 3 h at 37°C. One plate was then used to determine the total number of bacteria at the end of 3 h incubation, as described previously (279). The wells in the second plate were carefully washed three times with PBS and then used to determine the total number of adherent bacteria. All assays were performed in duplicate and repeated independently four times.

2.7.3 *In vivo* bacterial competition assays

2.7.3.1 Mouse model of large intestinal colonization

Six- to eight-week-old female CFW1 mice (Harlan) were used for intestinal colonization experiments as described previously (232). Briefly, mice were provided with drinking water containing 5 g/l streptomycin sulphate for 24 h. Prior to inoculation faecal samples were tested for growth of streptomycin-resistant bacteria. All experiments were performed as co-infections by feeding the mice a 100 µl suspension containing ~10^9 CFU of wildtype and mutant strains mixed 1:1 in 20 % sucrose. On the indicated

days faecal pellets were collected, weighed and homogenised in 0.9 % NaCl. Dilutions were plated on MacConkey agar supplemented with appropriate antibiotics and the CFU/g faeces was determined.

2.7.3.2 Mouse lung infection model.

A previously described intranasal infection model was applied for co-infection studies (82, 343). Six- to eight-week-old female NMRi mice (Harlan) were anaesthetized and hooked on a string by their front teeth. 50 µl of bacterial suspension containing ~5×10^7 CFU of wildtype and mutant strains mixed 1:1 was dropped onto the nares and was readily aspirated. Mice were left hooked on the string for 10 minutes before being returned to their cages. Mice were sacrificed 30 h after inoculation and the lungs, spleen and liver were collected in 0.9 % NaCl and homogenised. Serial dilutions were plated on selective media and CFU per organ was determined.

2.7.3.3 Mouse model of ascending urinary tract infection

The model used has been described in detail previously (179, 387). Three days before inoculation, six- to eight week old female C3H mice (Harlan) were provided with drinking water containing 5 % glucose. The mice were anesthetised and a plastic catheter was carefully pushed horizontally through the urethral orifice until it reached the top of the bladder and 50 µl of bacterial suspension containing ~5×10^8 CFU was slowly injected. Mice were sacrificed 3 days after inoculation and the bladder and kidneys were collected in 0.9 % NaCl and homogenised. Serial dilutions were plated on selective media and CFU per organ was determined. Additionally, urine was collected and the number of CFU/ml was determined.

2.7.4 Agglutination assays

Agglutination assays were performed using either yeast [i] or guinea pig red blood cells (GP-RBCs). Dried yeast was prepared freshly as 0.5 % (w/v) suspensions in PBS, using multiple washing, centrifugation and resuspension steps. GP-RBCs were stored at 4ºC as 20 % solutions in Alsever's (Matrix Biologicals). When required, a 5 % suspension of GP-RBCs was freshly prepared by diluting the 20 % stock solution 1 in 4 in PBS. Briefly, agglutination assays were performed on microscope slides using 15 µl of overnight bacterial culture (37ºC, 200 rpm) and 15 µl of yeast or blood suspension. The slide was tipped back and forth and observed for signs of agglutination for up to 15 min. To test for mannose-sensitivity, 30 µl of 10 % (w/v) mannose was added to the agglutination reaction.

2.7.5 Prodigiosin quantification

Prodigiosin levels were quantified according to a protocol by Slater *et al.* (371). Briefly, 1 ml of sample was harvested by centrifugation at 13 000 rpm for 5 min, after determining the OD_{600}. Next, the pellet was resuspended in acidified ethanol (4 % (v/v) 1 M HCl in ethanol) and the cell debris was removed with a second centrifugation. Finally, the absorbance at 534 nm was determined for the remaining supernatant (prodigiosin shows strong absorption at 534 nm in acidified ethanol). Finally, the relative prodigiosin concentration was determined as amount expressed per cell density (absorbance at 534 nm per OD_{600}).

2.7.6 Swimming motility and exopolysaccharide production assay

Swimming motility was investigated on LA motility media supplemented with 0.1 mM IPTG and/or ampicillin, as required (Appendix 1). Exopolysaccharide biosynthesis was investigated similarly on congo red media (Appendix 1). Briefly, 2 µl of an overnight

[i] Type 3 fimbriae can agglutinate yeast cells but not GP-RBCs. In *K. pneumoniae* it is therefore essential to investigate for type 1 fimbriae using GP-RBCs (C. Struve, personal communication)

culture was spotted onto LA motility and/or congo red plates, which were incubated at 37ºC (*K. pneumoniae* and *E. coli*) or 30ºC (*Serratia* sp. ATCC 39006) for 48 h and then examined.

2.7.7 Transmission electron microscopy

5 µl of sample was applied to a hydrophilic Formvar-carbon coated copper grid (Agar Scientific) and allowed to adsorb for 5 min. After wicking excess liquid, the grid was washed once using distilled deionised water and then stained for 15 s with 5 µl of 1 % uranyl acetate (pH 4.5). Transmission electron microscopy was carried out on a JEOL JEM-1400 microscope at 80 kV by Stefan Hyman at the Centre for Core Biotechnology Services, University of Leicester.

2.7.8 Data analysis and statistics

The competitive index (CI) was calculated by dividing the ratio of *fim2*-positive to *fim2*-negative bacteria recovered from infected organs by the ratio of the corresponding bacteria in the initial inoculum. The non-parametric Mann-Whitney U test was used to analyse the data from infection studies and prodigiosin quantifications. Biofilm and cell-adhesion data were analysed using the non-parametric Kruskal-Wallis test and Dunn's posthoc analysis. *P* values less than 0.05 were considered statistically significant.

2.7.9 Bioinformatics

A variety of programs were used to perform various sequence analyses. These programs and their functions are listed in Table 2-5.

Table 2-5 :- Bioinformatics programs used in this work

Name	Functions	Reference
Database searches		
BLAST		(9)
BLASTn	Searches DNA database for similar DNA sequences	
BLASTp	Searches protein database for similar protein sequences	
BLASTx	Translates DNA in six frames and searches a protein database for similar sequences	
tBLASTx	Translates DNA in six frames and searches a database of nucleotide sequences also translated in six frames	
RPS-BLAST	Searches a query sequence against a database of position-specific scoring matrices	
CDD	Database of multiple sequence alignments for domains and full-length proteins	(246)
Pfam	Database of protein families	(120)
PROSITE	Database of protein domains, families and functional sites	(173)
Sequence analysis		
Aragorn	Searches a nucleotide sequence for tRNA/tmRNA coding sites	(224)
Artemis	Genome viewer and annotation tool	(337)
Glimmer 3.02	Putative CDS identification using interpolated Markov models	(93)
GeneMark.hmm	Putative CDS identification using Markov models	(240)
ISGA	User interface to the genome annotation process using Ergatis workflows	(282)
MobilomeFinder	Web-based tool for the identification and analysis of GIs	(282)
NEBCutter	Online tool to identify enzyme restriction sites in nucleotide sequences	(428)
ORF Finder	Tool that uses stop and start codons to find all open reading frames	(271)
PROSO	A sequence-based protein solubility evaluator	(372)
Others		
Autodimer	Tool to screen primers for primer-dimer and hairpin formation	(426)
ClustalX	Visual multiple sequence (nucleotide and protein) alignment tool	(223)
GeneDoc	Aids in shading of alignments generated by ClustalX	n/a
In silico PCR	Online program to perform *in silico* PCR	(27)
LaserGene DNAStar Suite	Suite with various programs for bioinformatics analysis, including restriction mapping, primer design and contig alignment	(47)
Oligo Calc	Oligonucleotide analysis	(196)
Perl Primer	Primer design	(249)
Primer3	Primer design	(335)
Rare Codon Calculator	Identification of rare codon clusters	(74)
Staden Tools	Sequence assembly and chromatogram analysis	(382)

Chapter 3. Genetic manipulation of the *K. pneumoniae genome*: development of tools and protocols

3.1 Introduction

Gene and operon function are often identified using reverse genetics. In this process, a region of interest is deleted or disrupted using allelic exchange and phenotypic analyses are carried out on derived mutants. It was essential to this project that a set of reliable genetic tools and protocols were available to perform reverse genetics in *K. pneumoniae*. Several tools developed for allelic exchange in *E. coli* and other Gram negative bacteria have been adapted for use in *K. pneumoniae*, including suicide vector- and lambda Red-based allelic exchange (77, 88, 331, 387).

Although the actual methodology of these two techniques varies considerably, they both require the generation of mutant alleles that permit homologous recombination and subsequent allelic exchange. Mutant alleles consist of an antibiotic resistance cassette flanked by regions with homology to sequences upstream and downstream of the region targeted for deletion. Occasionally, only a single homologous recombination event is required for insertional gene disruption. These mutant alleles consist of only an antibiotic resistance cassette and a left or right flanking region.

In suicide vector-based allelic exchange the mutant allele is cloned into a suicide vector, a plasmid which is unable to replicate in the strain targeted for mutagenesis. Multiple suicide vectors with unique useful features have been constructed including pDS132 (296), pKOV (331) and pUT (71). Suicide vectors with cloned mutant alleles are subsequently delivered into target bacteria using either transformation or conjugation. However, the low transformation efficiency of the majority of *K. pneumoniae* strains limits the effectiveness of plasmids such as pKOV, which

require transformation (331). Alternatively, vectors such as the pUT-based suicide vectors can be conjugated into *K. pneumoniae* (71). However, these vectors have only been used for insertional disruption of chosen genes by single homologous recombination. This is disadvantageous because following single-crossover two truncated copies of the target gene remain and may potentially mask effects that would be associated with complete loss or disruption of the gene. Moreover, larger deletions spanning entire operons, such as the KpGI-5 *fim2* operon, cannot be generated using this approach. These deletions are often obtained using a double homologous recombination technique. Initially, the suicide vector integrates into the chromosome and merodiploid strains are formed (also called single-crossover strain). Next, a second homologous recombination event occurs and either suicide vector and mutant allele are lost (abortive allelic exchange), or suicide vector and native allele are lost (successful allelic exchange). The latter event creates a knockout mutant strain.

The second homologous recombination event does not occur frequently and/or spontaneously and requires a suicide vector-encoded counterselectable marker to select for clones which have lost the suicide vector. Frequently researchers use the counterselectable *sacB* cassette, which confers sucrose-sensitivity to many Gram negative bacteria (30, 137). Several studies have reported using suicide vectors and mutant alleles in double homologous recombination events in *K. pneumoniae* but detailed protocols have not been published (69, 228, 331).

Lambda Red-based allelic exchange has been successfully applied to *K. pneumoniae*. In this method bacteria over-expressing the phage lambda Red recombination system are transformed with linear fragments of PCR-amplified mutant alleles, which then recombine into the host chromosome (88, 267). However, its applicability is limited to selected strains as there is considerable strain-to-strain variation in both capsule thickness and transformation efficiency (122). Moreover some strains are unable to

support the commonly used lambda Red recombination plasmid pKD46 (J. J. van Aartsen and K. Rajakumar, unpublished results) (88).

When reverse genetics is used to identify gene function it is often also necessary to perform complementation experiments, whereby the gene/operon targeted for disruption is reintroduced into the mutant strain and the obtained recombinant strain is assessed and compared to both the wildtype and mutant strains. Although this can be accomplished using multi-copy replicative plasmids, single-copy genome-based complementation is sometimes preferred as they are more representative of the native genetic condition of the region under investigation (66). Transposon Tn*7* inserts at high-frequency into an intergenic chromosomal region called *att*Tn*7*, located immediately downstream of the *glmS* gene which is highly conserved between species (66, 293). The broad host range and site-specificity of transposase Tn*7* has been exploited by Choi *et al.* (66) to develop a method for single copy complementation and strain tagging. In their method, suicide plasmids harbouring mini-Tn*7* cassettes and the Tn*7* recombination machinery genes are conjugated into recipient target strains. Tn*7* subsequently recognises the Tn*7*L and Tn*7*R sites that flank mini-Tn*7* elements and causes the site-specific transposition of these elements into *att*Tn*7*. Only one published paper has mentioned using the mini-Tn*7* transposition system in *K. pneumoniae* (236).

Both allelic exchange methods and the mini-Tn*7* system rely on positively selectable markers, such as antibiotic resistance cassettes, to isolate desired recombinant strains. *K. pneumoniae* strains are often highly antibiotic resistant thus limiting the availability of appropriate selectable markers (section 1.1.3) and hampering the generation of successive mutations in a previously generated recombinant strain background. Flp-mediated FRT-recombination can enable reiterative recycling of resistance markers that have been constructed with flanking FRT sites. Following expression of Flp, recombination occurs between FRT sites and the internal antibiotic cassette is deleted. Flp can be provided *in trans* by the plasmid pFLP2 (163). To date only two studies

have reported the use of FRT-flanked antibiotic resistance cassettes and FLP-mediated deletion in *K. pneumoniae* (13, 455).

This chapter will first discuss the development of an efficient suicide vector-based allelic exchange protocol for *K. pneumoniae* using suicide vector pDS132. Next, the unexpected technical problems that were observed when utilizing a commonly used gentamicin resistance cassette (*aacC1*) will be discussed. Subsequently, the applicability of lambda Red-based mutagenesis in the *fim2*-positive *K. pneumoniae* strain KR116 will be explored. Finally, the application of Flp recombinase-mediated deletion of FRT-flanked antibiotic resistance cassettes for the construction of markerless mutants will also be examined.

3.2 Results

3.2.1 *K. pneumoniae* suicide vector-based allelic exchange

This results section will initially describe the construction of two novel pDS132-based suicide vectors with optimized cloning sites. Next, the construction of mutant alleles and suicide vectors for mutagenesis of various *K. pneumoniae* loci will be summarised and a list of mutants constructed using suicide vector-based allelic exchange will be provided. Finally, evidence proving that 25ºC is the optimum temperature for *sacB* counterselection efficiency in *K. pneumoniae* will be discussed.

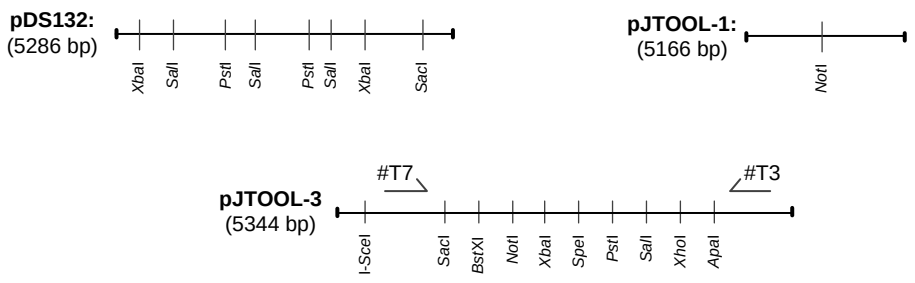
3.2.1.1 pJTOOL-1 and pJTOOL-3: pDS132-based vectors with optimized cloning sites

Suicide vector pDS132 possesses the *R6K* origin of replication which requires the phage lambda *pir*-encoded π protein to replicate (296). In host strains lacking *pir*, the plasmid form of pDS132 is spontaneously lost by ‘suicide’ allowing for the selection of clones that harbour genomically-integrated plasmids as a result of single homologous

recombination with a target locus. pDS132 also harbours the broad host range *mob RP4* origin of transfer, which allows for efficient conjugative transfer from *E. coli* strains that code for an intact Tra apparatus, such as SM10λ*pir* and S17-1λ*pir* (366). In addition, the vector-backbone carries *cat*, which encodes chloramphenicol resistance, and *sacB*, which confers sucrose-sensitivity to most Gram negative bacteria and thus functions as a counterselectable marker. However, the pDS132 MCS only offers four restriction sites: *Sal*I, *Pst*I, *Xba*I and *Sac*I. These restriction sites frequently occur within the *K. pneumoniae* genome. Therefore, to aid the cloning of *K. pneumoniae*-specific mutant alleles (section 2.2.4), two pDS132 derivatives carrying alternative cloning sites were constructed.

Firstly, pJTOOL-1 was constructed using inverse PCR to replace the pDS132 MCS with a single rare-cutting *Not*I restriction site (Figure 3-1). To offer further cloning utility, pJTOOL-3 was constructed using SLIC to replace the pDS132 cloning site with the 167 bp T7 and T3 primer annealing site-flanked MCS of pBluescript II KS+ (Figure 3-1). In addition to the original four sites, five further restriction sites were now available: *Apa*I, *Xho*I, *Spe*I, *Not*I and *Bst*XI. The homing endonuclease-specific I-*Sce*I site was also added, by incorporation within one of the SLIC primers. Because of its known absence in bacteria, it provides a unique site if all other restriction sites cannot be utilized.

Figure 3-1 :- Schematic of the MCS of pDS132, pJTOOL-1 and pJTOOL-3



The MCSs of pDS132, pJTOOL-1 and pJTOOL-3 are illustrated with the unique restriction sites that only occur within the MCS. Binding sites for the universal primers T7 and T3 flank the MCS in pJTOOL-3.

3.2.1.2 Construction of mutant alleles and suicide vectors for mutagenesis of targeted *K. pneumoniae* loci

As described in the introduction, the aim of this study was to characterise the putative *fim2* operon on the *K. pneumoniae* KR116 KpGI-5 island. It was therefore necessary to construct mutant alleles targeting *fim2*, *fim2K* and KpGI-5. Additional alleles were required to construct the necessary control strains for phenotype assays including those targeting *fim*, *fimK* and *mrk*. Details regarding the construction of mutant alleles, including primers, antibiotic resistance cassettes, fragment sizes and flanking restriction sites can be found in Table 2-2. Flanking restriction sites were chosen according to their absence in the assembled mutant allele and were incorporated into the 5' end of flanking primers. Mutant alleles were cloned into either the *Not*I or *Xba*I restriction site of pJTOOL-1 or pDS132, respectively. Details regarding each of the suicide vectors constructed in this work and their associated mutant alleles are presented in Appendix 6.

3.2.1.3 *K. pneumoniae* mutant construction: suicide vector-based allelic exchange

For purposes of explanation, only the suicide vector-based mutagenesis of the *K. pneumoniae* KR116 *fimK* gene using suicide vector pJKO-1b will be described in detail. This framework was also used for mutagenesis of other *K. pneumoniae* strains and different target loci. When necessary, the protocol was altered depending on the antibiotic resistant cassette used and suicide vectors were changed according to the region targeted for mutagenesis. A schematic of the double-crossover mediated deletion of the KR116 *fimK* gene using suicide vector pJKO-1b is presented in Figure 3-2.

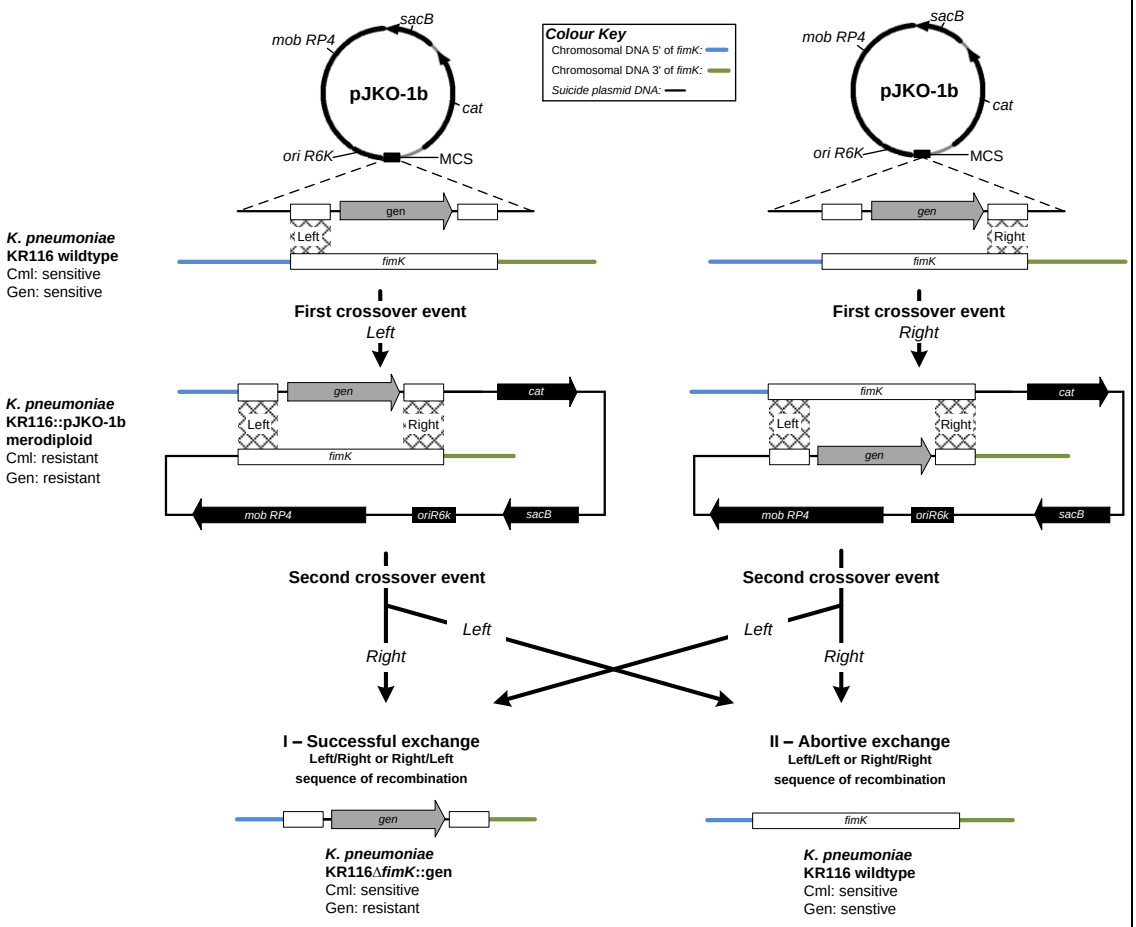
Initially, *E. coli* S17-1λ*pir* was used to conjugate pJKO-1b into *K. pneumoniae* KR116, which was shown to have low-level resistance to gentamicin, chloramphenicol and

kanamycin (data not shown). After overnight incubation, conjugation mixes were plated onto M9 medium agar supplemented with gentamicin (9 µg/ml) to select against *E. coli* S17-1λ*pir*, which required proline and thiamine for growth, and the *K. pneumoniae* wildtype recipient, which was gentamicin-sensitive. Numerous *K. pneumoniae* single-crossover merodiploids possessing integrated copies of pJKO-1b were readily obtained. Five putative merodiploids were examined and confirmed to have the expected resistance phenotype using patch-plating (gentamicin-resistant and chloramphenicol-resistant). Additionally, all five produced positive PCR bands for the suicide vector backbone (primers PR637 and PR638) and *fimK*::gen (primers PR640 and PR553), the mutant allele present on pJKO-1b (Figure 3-3A).

Next, serial dilutions containing 10^4 to 10^9 CFU of merodiploid KR116 *fimK*::pJKO-1b were plated on LAS medium and incubated overnight at 25ºC to select clones that had lost the *sacB* gene, and had therefore also lost pJKO-1b. Colonies obtained after *sacB*-counterselection were screened for gentamicin (9 µg/ml) and chloramphenicol (30 µg/ml) sensitivity by patching and analysed by PCR for loss of the suicide vector backbone (primers PR708 and PR637) and retention of the mutant allele (primers PR552 and PR553). The expected genotype was obtained in all eight randomly selected colonies that possessed the expected chloramphenicol-sensitive and gentamicin-resistant phenotype (Figure 3-3B). Finally, *fimK*::gen was definitively mapped to the *fimK* locus by PCR using primers PR654 and PR641 (Figure 3-3C).

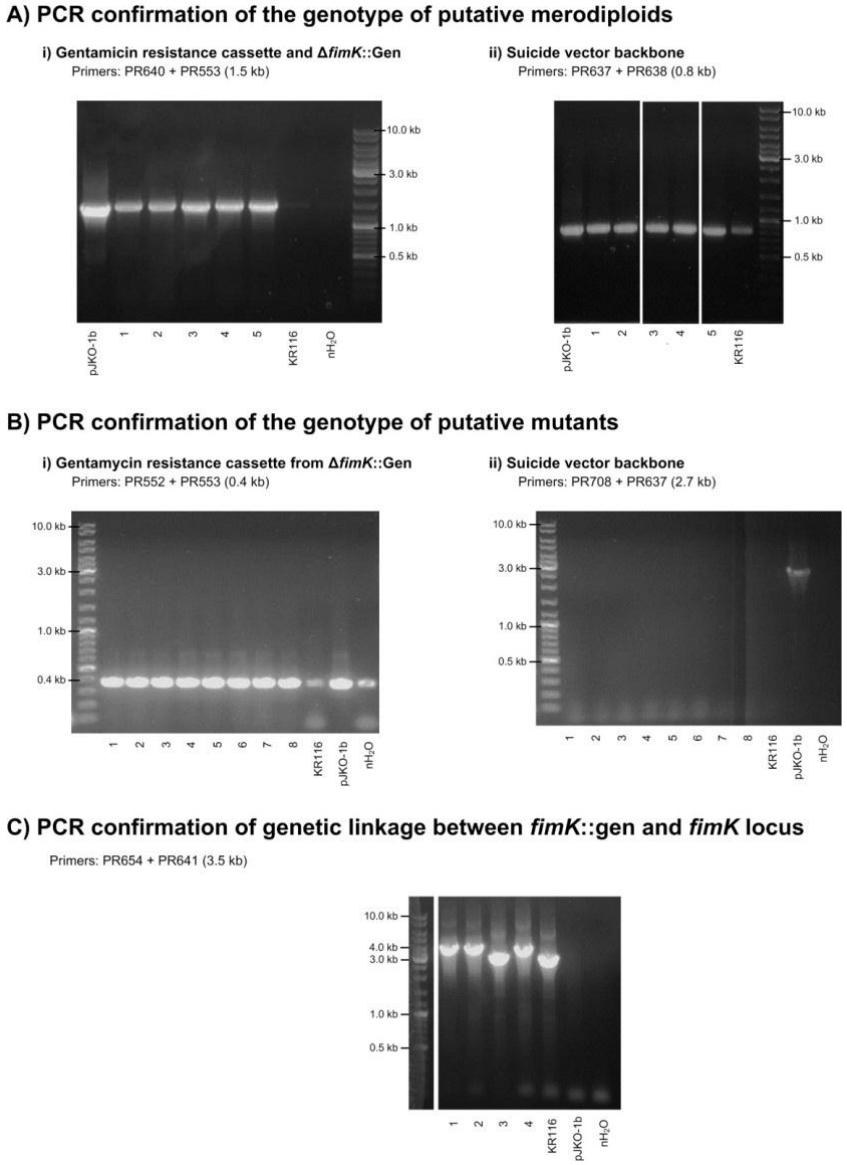
The above protocol has been used to genetically manipulate *K. pneumoniae* KR116 to create multiple knockout strains, which have been summarised in Table 3-1. Using the same protocol, three additional *K. pneumoniae* strains have also been successfully manipulated: KR161, KR162 and KR173 (Table 3-1). Notably, KR161 and KR162, which possess very thick capsules, yielded fewer single-crossover merodiploids than the other two strains. More details regarding strain construction, including the merodiploid intermediate strains, can be found in section 2.6.2 and Appendix 4.

Figure 3-2 :- Schematic representation of double-crossover mediated deletion of the *fimK* locus of *K. pneumoniae* KR116 using suicide vector pJKO-1b, a pJTOOL-1 derivative possessing *fimK*::gen



In step 1, merodiploids arising out of loss of the free plasmid and selection of clones harboring a chromosomally-integrated form of the suicide plasmid arising following left- or right-sided single-crossover homologous recombination are shown. Although both merodiploids have the same phenotype, they have distinct *fimK*-associated genetic organizations as shown. Importantly, both merodiploids possess both an intact *fimK* locus and a *fimK*::gen allele. After *sacB* counterselection on LAS plates, putative double-crossover mutants are identified by a chloramphenicol-sensitive, gentamicin-resistant phenotype and verified by colony PCR analysis. **(I)** Genetic configurations arising from Left/Right or Rright/Left [1st crossover/2nd crossover events] homologous recombination mediated successful allelic exchange. **(II)** Abortive allelic exchange resulting from Left/Left or Right/Right-sided recombination events. Abbreviations: Cml, chloramphenicol; Gen, gentamicin; and MCS, multiple cloning site.

Figure 3-3 :- PCRs used to confirm putative merodiploid and mutant genotypes obtained in mutagenesis experiments of *fimK* in *K. pneumoniae* KR116



A) Colony PCR analysis of five putative merodiploids (1 to 5) obtained by plating a *K. pneumoniae* KR116 and *E. coli* S17-1λ*pir*/pJKO-1b mating mixture on M9 media plates supplemented with gentamicin. Chloramphenicol-resistant and gentamicin-resistant merodiploids were analysed to identify the presence of the **(Ai)** mutant allele and the associated gentamicin resistance cassette, and **(Aii)** the suicide vector backbone. **B)** Colony PCR analysis of eight putative mutants (1 to 8) that were obtained by sucrose counterselection of KR116 *fimK*::pJKO-1b on LAS media at 25ºC. All eight mutants were phenotypically chloramphenicol-sensitive and gentamicin-resistant and PCR was used to identify the presence of the **(Bi)** mutant allele and the associated gentamicin resistance cassette, and the absence of **(Bii)** the suicide vector backbone. **C)** The genetic linkage of Δ*fimK*::gen to the *fimK* locus was also verified, and colony 3 was determined to be incorrect. It must be noted PCR contamination can be seen in Aii and Bi. These PCRs were repeated without contamination on a smaller scale using only one or two putative merodiploids or mutants; this data is not shown.

Table 3-1 :- Mutants created using suicide vector-based allelic exchange

KR catalogue number	Strain genotype
***K. pneumoniae* KR116 mutants**	
KR916	Δ*fimK*
KR1024	Δ*fim2K*
KR1023	Δ*fim2K*Δ*fimK*
KR1112	Δ*fim2K*Δ*fim*
KR1816	Δ*fim2*
KR1817	ΔKpGI-5
KR1864	Δ$fim2_{FRT}$Δ*fim*
***K. pneumoniae* KR161 mutants**	
KR978	Δ*fimK*
KR1143	Δ*fim*
***K. pneumoniae* KR162 mutants**	
KR1001	Δ*fimK*
KR1145	Δ*fim*
***K. pneumoniae* KR173 mutants**	
KR1005	Δ*fimK*
KR1147	Δ*fim*

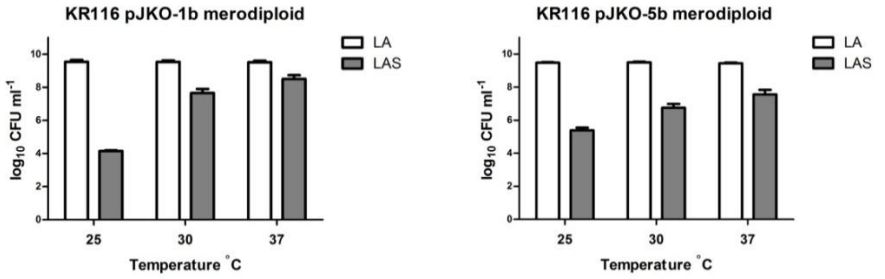
3.2.1.4 Efficient *sacB* counterselection in *K. pneumoniae* requires incubation at 25°C

Initial *sacB* counterselection experiments using protocols previously employed in *E. coli* and *P. aeruginosa* proved problematic (30, 67). Screening of over 2000 sucrose-resistant colonies growing on LAS derived from several independent sucrose-selection experiments yielded only three gentamicin-resistant chloramphenicol-sensitive clones that were subsequently shown by PCR to be genuine double-crossover mutants. Surprisingly, most colonies remained resistant to both antibiotics and retained the vector backbone. Additionally, no colonies were obtained when merodiploids were plated on LAS supplemented with 18 µg/ml gentamicin, a finding which will be discussed in section 3.2.2.1. Modifications previously employed in *E. coli* to reduce background, including decreasing NaCl concentration in medium and incubation at 30°C instead of 37°C (30), did not appear to significantly improve counterselection

efficiency in *K. pneumoniae*. Use of a range of media (LB agar, LB agar without added NaCl and M9 minimal medium with glycerol) and sucrose concentrations (5 %, 6 % and 10 %) also failed to resolve this problem

To investigate the effect of temperature on *sacB* counterselection in *K. pneumoniae* in more detail, an experiment was performed to test the ability of *K. pneumoniae* merodiploids to form colonies at 25°C, 30°C and 37°C on LB agar and LAS medium. As expected many orders fewer colonies were obtained on LAS plates as compared to LB plates at all three temperatures, confirming the sucrose-mediated negative-selection functionality of the *sacB* cassette in pJTOOL-1 (Figure 3-4). However, regardless of whether *fim* (pJKO-5b) or *fimK* (pJKO-1b) was targeted, growth on LAS as compared to LA at 25°C showed a 10^2- to 10^4-fold further reduction in CFU count than was apparent for plates incubated at 30°C or 37°C, suggesting that *sacB* counterselection at 25°C was much more effective. These results were further confirmed using independent merodiploids generated following repeat conjugation experiments using the same two suicide vectors (data not shown).

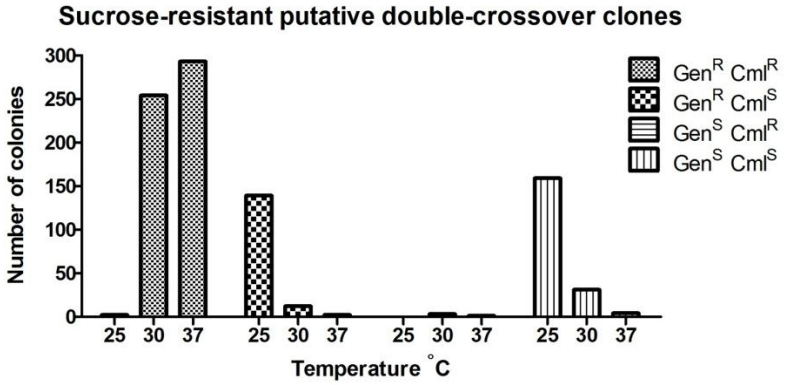
Figure 3-4 :- CFU counts of two *K. pneumoniae* KR116 merodiploids plated on LA and LAS media at incubated at different temperatures.



An overnight culture of each merodiploid was prepared by incubation at 37°C in 5 ml of LB plus gentamicin (9 µg/ml). Next, serial dilutions of each culture were plated onto LA and LAS in triplicate and incubated overnight at 25°C, 30°C or 37°C. The numbers of CFU/ml as detected by each medium were determined. Data shown was derived from duplicate experiments, each with two biological replicates, performed on separate days. Regardless of locus targeted, growth on LAS as compared to LA yields a 10^2- to 10^4-fold further reduction in CFU counts on plates incubated at 25°C versus those incubated at 30°C or 37°C. Error bars shown represent the standard error of the mean. LA, LB agar; LAS, LB agar plus 6 % (w/v) sucrose without added NaCl;

To investigate whether the effect of incubation temperature on CFU counts on LAS medium reflected improved efficiency of selection for *sacB*-minus colonies at 25°C, the gentamicin and chloramphenicol resistance phenotypes of 300 sucrose-resistant colonies obtained by growth of merodiploids on LAS at each temperature were determined by patching onto cognate media. The frequency histogram shown includes amalgamated data from replicate experiments performed on both pJKO-1b- and pJKO-5b-derived merodiploids in the *K. pneumoniae* KR116 background (Figure 3-5).

Figure 3-5 :- Phenotypic profiles of putative double-crossover clones derived from *K. pneumoniae* KR116 merodiploids.



The resistance phenotypes of 300 sucrose-resistant colonies derived from *K. pneumoniae* merodiploids plated onto LAS medium and incubated at 25°C, 30°C or 37°C was determined. The colonies were derived from duplicate experiments and were performed on both pJKO-1b- and pJKO-5b-derived merodiploids. Many false-positives colonies (sucrose-resistant, gentamicin-resistant and chloramphenicol-resistant) were obtained at 30°C and 37°C, while less than 1 % of colonies obtained from the LAS plate incubated at 25°C exhibited a false-positive phenotype. Also of note, at 25°C approximately 50 % of colonies possessed a phenotype consistent with successful allelic exchange (gentamicin-resistant, chloramphenicol-sensitive), while an equal proportion displayed a wildtype, abortive allelic exchange phenotype (gentamicin-sensitive, chloramphenicol-sensitive). Abbreviations: LA, LB agar; LAS, LB agar plus 6 % (w/v) sucrose without added NaCl; Gen, gentamicin; Cml, chloramphenicol; R, resistant; S, sensitive.

As noted, at 30°C and 37°C the vast majority of sucrose-resistant colonies growing on LAS plates still possessed the suicide vector backbone, as determined by the observed chloramphenicol resistance phenotype. By contrast, only two out of 300 colonies obtained on LAS medium following incubation at 25°C displayed this false-positive

phenotype. Indeed, based on phenotype (gentamicin-resistant, chloramphenicol-sensitive) 46 % were predicted to have undergone full allelic replacement and lost the pJTOOL-1 backbone, while 53 % had a phenotype (gentamicin-sensitive and chloramphenicol-sensitive) consistent with abortive allelic exchange and reversion to wildtype genotype. The genotype of up to 5 colonies obtained at each temperature that possessed the gentamicin-resistant and chloramphenicol-sensitive phenotype were analysed by colony PCR with primers to check for the suicide vector (primers PR708 and PR637) and mutant allele (primers PR552 and PR553). These confirmed loss of the suicide vector and retention of the mutant allele in all cases.

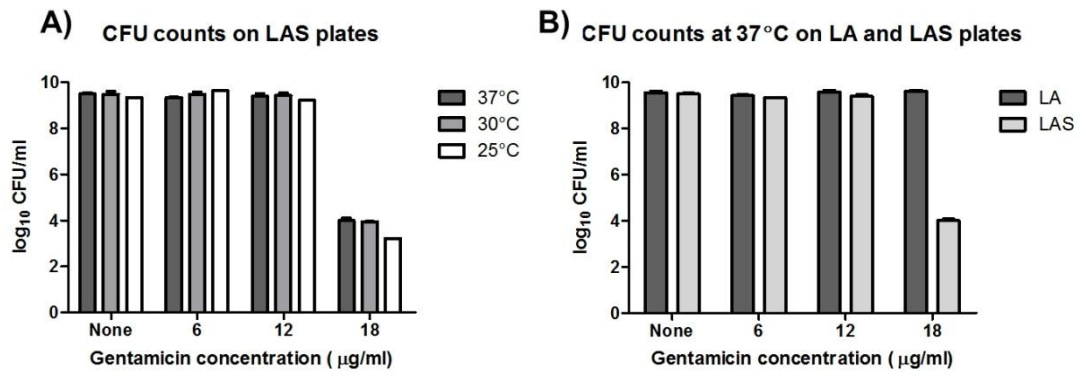
3.2.2 Unexpected technical problems

Once the mutants in Table 3-1 had been constructed, it was investigated whether the wildtype and mutant strains could be used for *in vivo* competition assays, which require plating on both selective and non-selective media to obtain CFU counts for each competing strains. These experiments will be discussed in this section as they raised important technical problems that needed resolving.

3.2.2.1 The *aacC1* resistance cassette confers only low-level gentamicin resistance in *K. pneumoniae*

Initial problems with the gentamicin resistance provided by the *aacC1* cassette were highlighted by the absence of colonies when merodiploids were plated on LAS plates supplemented with gentamicin for *sacB* counterselection. This was probably related to the markedly reduced CFU counts obtained when KR116Δ*fimK*::gen was plated on LAS medium supplemented with gentamicin as compared to sucrose-free LA supplemented with an identical concentration of gentamicin. This unexpected phenomenon was independent of temperature and resulted in $>10^4$-fold reduction in CFU counts when KR116Δ*fimK*::gen was plated on LAS supplemented with as little as 18 µg/ml gentamicin (Figure 3-6).

Figure 3-6 :- Impact of sucrose supplementation, incubation temperature and medium gentamicin concentration on CFU counts of *K. pneumoniae* KR116Δ*fimK*::*gen*



(A) CFU counts determined by plating dilutions of a 5 ml LB overnight culture of KR116Δ*fimK*::gen incubated at 37°C onto LAS plates supplemented with 0, 6, 12 and 18 µg/ml gentamicin and incubated at 25°C, 30°C and 37°C. **(B)** CFU counts determined by plating the same bacterial dilutions onto sucrose-free LA medium and LAS medium and incubating overnight at 37°C. Error bars shown represent the standard error of the mean. Abbreviations: LA, LB agar; LAS, LB agar plus 6 % (w/v) sucrose without added NaCl.

Moreover, when mutant strains were plated on LA supplemented with gentamicin concentrations greater than 30 µg/ml, large reductions in CFU counts and colony size were observed (data not shown). These observations were discussed with our collaborators at the Statens Serum Institut, Denmark, where the *in vivo* work would be carried out. They had previously observed differences in antibiotic resistance levels when comparing identical strains obtained from *in vivo* and *in vitro* samples. Paired with this observation, the associated low-level resistance problems and the fact that the *aacC1* resistance cassette was untested in *in vivo* models, they advised to reconstruct the mutants using the kanamycin and/or tetracycline resistance cassettes that they had previously used without problems (387, 388).

3.2.2.2 The *in vivo* murine intestinal colonization model requires isogenic streptomycin-resistant bacteria

As *K. pneumoniae* infections are preceded by intestinal colonization, it was important to determine the role of *fim2* in intestinal colonization. To investigate this *in vivo*,

streptomycin-treated mouse intestinal colonization models have been the method of choice and have been used extensively and successfully in *K. pneumoniae* (114, 387-390). Streptomycin treatment eradicates the resident intestinal flora, allowing for colonization of the large intestine by streptomycin-resistant bacteria (270). Initial experiments identified that *K. pneumoniae* KR116 was susceptible to streptomycin, and thus KR116 and its isogenic mutants could not be used in this model.

3.2.2.3 Solving the gentamicin and streptomycin resistance problems

Three methods were explored to solve the problems highlighted in the previous two sections, including:

A) Obtaining spontaneous streptomycin-resistant derivatives of existing mutant constructs.

B) Constructing mutant alleles with either a tetracycline or kanamycin resistance cassette and remaking the mutant strains in a spontaneous streptomycin-resistant mutant of KR116.

C) Tagging of the KR116 wildtype and isogenic mutant strains using mini-Tn*7* cassettes that code for streptomycin and kanamycin resistance.

Method A was tried initially, but it quickly became apparent that it would be difficult to maintain the isogenic nature of the strains due to the diverse nature of the genetic mechanisms that can confer streptomycin resistance (219). Moreover, large differences in growth curves were seen between mutants, a phenotype which could influence competition assay results. Finally, method A also did not solve the problems regarding gentamicin selection, as described earlier.

Method B would be time consuming as it would require repeating the entire process of generating the Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* knockout strains, including SOE-PCR mutant allele generation, suicide vector construction and allelic exchange, to be

repeated. In an attempt to shorten the process, lambda Red recombination-based allelic exchange in *K. pneumoniae* KR116 was investigated (section 3.2.3). Similarly, mini-Tn*7*-based tagging of *K. pneumoniae* KR116 and other strains (method C) had also not been extensively investigated previously and no pre-made mini-Tn*7* constructs harbouring both streptomycin and kanamycin resistance genes were available. Therefore, to maximise the likelihood of obtaining the required recombinant strains, both method B and C were pursued.

3.2.3 *K. pneumoniae* lambda Red recombination-based allelic exchange

3.2.3.1 Construction of lambda Red-compatible mutant alleles with alternative antibiotic resistance cassettes

Mutant alleles were constructed similarly to those used for suicide vector-based allelic exchange, although shorter left and right homologous flanking regions were used. Additionally, they contained alternative selectable cassettes that provided high-level resistance in *K. pneumoniae* when present as a single copy. These were the Tn*5*-derived FRT-flanked *aph(3')-II* kanamycin resistance cassette in pKD4 and the omega interposon tetracycline resistance cassette in pUT-mini-Tn5-Tc (88, 237). For purposes of this work, these cassettes were amplified from the recombinant strain *K. pneumoniae* C3091 Δ*mrk*::kan Δ*fim*::tet (KR2021) (388). Details regarding the construction of mutant alleles, including primers, antibiotic resistance cassettes and fragment sizes can be found in Table 2-2.

3.2.3.2 *K. pneumoniae* mutant construction: lambda Red-based allelic exchange

To ensure compatibility with the intestinal colonization model a spontaneous streptomycin-resistant derivative of KR116 was obtained by plating 10^{10} CFU onto LA plates supplemented with 500 µg/ml of streptomycin, which were grown overnight at

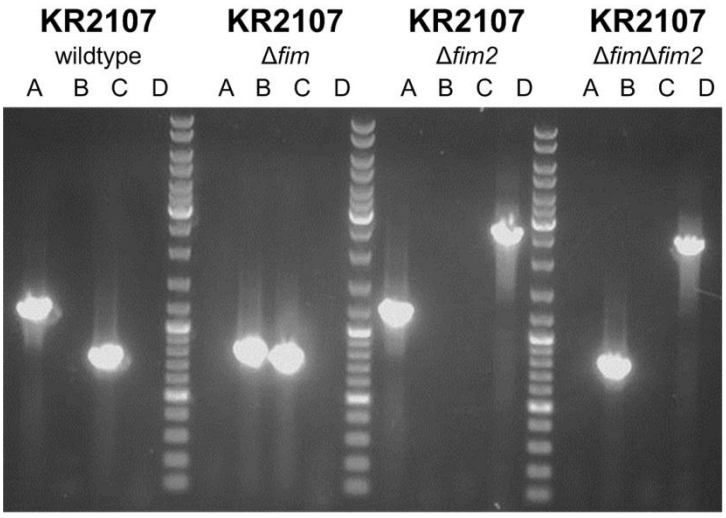
37ºC. One of the spontaneous streptomycin-resistant colonies obtained possessed a 24 h growth curve similar to KR116 and was designated KR2107 (data not shown). It also possessed a similar GP-RBC agglutination profile as the KR116 wildtype strain. Using Etest strips the MIC of streptomycin was determined to be 1.0 µg/ml for KR116 and >1024 µg/ml for KR2107. KR2107 was therefore used to construct isogenic Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* mutants.

Several unsuccessful attempts were made to transform KR2107 with pKD46, a plasmid encoding the lambda Red recombination machinery. Fortunately the transformation of pKOBEG-Apra, an alternative plasmid, was successful and more than 10^3 transformants were routinely obtained. pKOBEG-Apra is a temperature sensitive plasmid which is replicative at 30ºC but not 37ºC. Additionally, it possesses the phage lambda Red recombination system, the expression of which is controlled by the arabinose-inducible P_{BAD} promoter.

For purposes of explanation, only the lambda Red-based mutagenesis of the *K. pneumoniae* KR2107 *fim* operon will be described. This framework was also used for mutagenesis of other KR2107 loci and was only altered when a different antibiotic resistance cassette was present in the mutant allele. Briefly, an overnight culture of KR2107/pKOBEG-Apra was grown at 30ºC in LB and was used to inoculate a fresh culture of LB supplemented with apramycin and 0.2 % arabinose at a ratio of 1:100. The culture was grown to an OD_{600} of 1.0, and electrocompetent cells were prepared and transformed with 2 µg of the mutant allele *fim*::tet PCR product. After non-selective growth at 37ºC for 1 h, the transformation mixture was plated onto LA containing tetracycline and incubated overnight at 37ºC. Although this method was of low efficiency, 1 to 50 transformants were regularly obtained. Next, putative mutants were determined to be correct by PCR screening for introduced mutant alleles and antibiotic cassettes, as described in section 2.6.4. Details regarding the PCR-based genotype

confirmation of KR2107 and its isogenic Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* mutants are shown in Figure 3-7.

Figure 3-7 :- Confirmation of the genotype of KR2107 and its isogenic mutants



PCR A:
* PR1144 + PR1145 (1.1 kb)
* Screen for *fimH;* gene deleted during allelic exchange with Δ*fim*

PCR B:
* PR1253 + PR1148 (0.8 kb)
* Confirms correct genomic location of Δ*fim*

PCR C:
* PR1204 + PR1205 (0.7 kb)
* Screen for *fim2K*; gene deleted during allelic exchange with Δ*fim2*

PCR D:
* PR1103 + PR1256 (2.6 kb)
* Confirms correct genomic location of Δ*fim2*

KR2107 and its isogenic Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* mutants were screened using PCR to confirm their expected genotypes. As expected, KR2107 possessed both the native *fimH* and *fim2K* genes. KR2107Δ*fim*Δ*fim2* did not possess these native genes, as shown by negative results for PCR A and B, but did harbour the *fim* and *fim2* mutant alleles. Whilst KR2107Δ*fim* possessed a native *fim2K* gene, as expected the *fimH* gene was lost and replaced by Δ*fim.* Similarly, KR2107Δ*fim2* possessed a native *fimH* gene and *fim2K* was lost and replaced by Δ*fim2*. Although no negative controls (without template) have been shown, these were performed and were negative.

A total of ten mutants were constructed using the lambda Red-based allelic exchange method and these are summarised in Table 3-2. For more details, please refer to Appendix 4.

Table 3-2 :- Mutants created using lambda Red-based allelic exchange

KR catalogue number	Strain genotype
***K. pneumoniae* KR116 mutants**	
KR2106	Δ*fim*
***K. pneumoniae* KR2107 mutants**	
KR2128	Str^R Δ*fim*
KR2130	Str^R Δ*fim2*
KR2152	Str^R Δ*fim*Δ*fim2*
KR2305	Str^R Δ*fim*Δ*mrk*
KR2303	Str^R Δ*fimH*
KR2301	Str^R Δ*fim2H*
KR2324	Str^R Δ*fim2H*Δ*fimH*
KR2304	Str^R Δ*fim2*Δ*fimH*
KR2302	Str^R Δ*fim*Δ*fim2H*

3.2.4 Mini-Tn*7* tagging of *K. pneumoniae* strains

3.2.4.1 Construction of streptomycin- and/or kanamycin- resistance encoding mini-Tn*7* cassettes

The mini-Tn*7* system described by Choi *et al.* contains an extensive catalogue of mini-Tn*7* cassettes cloned into delivery vectors (66). Unfortunately, these were incompatible with this project as their delivery vectors either did not suicide in *K. pneumoniae* (*ColE1* origin of replication), did not possess kanamycin/streptomycin resistance genes and/or required electroporation. Therefore, mini-Tn*7* cassettes harbouring streptomycin and/or kanamycin resistance genes were constructed in the pUC18R6kT-mini-Tn7T delivery vector. This vector possessed the *R6K* origin of replication and hence was unable to replicate in *K. pneumoniae.* Also, this vector contained a *mob RP4* origin of transfer for high-efficiency conjugation. Additionally, the mini-Tn*7*T cassette is a mini-Tn*7* derivative which contains transcriptional terminators that prevent read-through from the *glmS* gene (66).

Four different resistance genes were used to assemble the mini-Tn*7*T constructs. The FRT-flanked Tn*5*-derived *aph(3')-II* kanamycin resistance gene was obtained from pKD4. This gene has previously been shown to confer high level (up to 200 µg/ml) resistance to kanamycin, even when present as a single, chromosomally-integrated copy (88). Three genes conferring streptomycin resistance were also used: *aadA1a*, *strA* and *strB*. The *aadA1a* gene encodes an aminoglycoside adenylyltransferase and has been used successfully at single copy in many mutagenesis studies (237). In this work, the *aadA1a* gene was amplified from the mini-Tn*5*Sm/Sp transposon in *Serratia* sp. strain ROP4S (118, 237). The *strAB* genes are genetically linked and encode two phosphotransferase enzymes (aminoglycoside-3'-phosphotransferase and aminoglycoside-6'-phosphotransferase) which inactivate streptomycin (65). In this work, *strAB* was obtained from the *K. pneumoniae* MGH78578 plasmid pKPN5, and was 100 % identical to the well characterised *strAB* genes from *Pseudomonas syringae* pv. *syringae* (394).

Schematics of the mini-Tn*7*T cassettes and matching construction details have been presented in an earlier section (section 2.5.1.4; Figure 2-2). In total, eight kanamycin and/or streptomycin mini-Tn*7*T cassettes were constructed using recombinant DNA techniques. Details of the mini-Tn7T delivery vectors pJTOOL-8 through to pJTOOL-14 and associated cassettes can be found in Appendix 6. Based on similarity to *strAB* from *P. syringae*, amplification of *strAB* from MGH78578 removed the native *strAB* promoter, predicted to be located ~100 bp proximal to the start codon of a resolvase gene lying upstream of *strA* (394). To ensure expression of *strAB* the locus was cloned between the *Xho*I and *Kpn*I restriction sites of pWSK129, downstream of the P_{LAC} promoter. This structure was called P_{LAC}*strAB* and was used instead of *strAB* in the construction of mini-Tn*7*T cassettes. It was previously determined that strains possessing both *strAB* and *aadA1a* have higher streptomycin resistance than those harbouring each cassette alone (393). Therefore a triplet streptomycin resistance

cassette (named 3St) with tandem *aadA1a*-P_{LAC}*strAB* genes was also constructed. In the alternative construct 3St-rev the *aadA1a* and P_{LAC}*strAB* cassettes were convergent.

To test the streptomycin resistance phenotype, the mini-Tn*7*T cassettes that encoded streptomycin resistance only were integrated downstream of *glmS* in *K. pneumoniae* KR116 using conjugation and Tn*7* transposition (as described in section 3.2.4.2). The level of streptomycin resistance was determined using Etest strips according to BSAC guidelines and the results are shown in Table 3-3. As expected, the triplet streptomycin resistance cassette conferred the strongest resistance phenotype upon *K. pneumoniae* KR116.

Table 3-3 :- *K. pneumoniae* KR116 strains tagged using the streptomycin resistance-encoding mini-Tn*7*T constructs and their associated MICs

Tagged KR116 strains	Streptomycin MICs	Phenotype*
Wildtype	1.0 µg/ml	Sensitive
*att*Tn*7*::mini-Tn*7*T::*aadA1a*	10 µg/ml	Resistant
*att*Tn*7*::mini-Tn*7*T::P_{LAC}*strAB*	10 µg/ml	Resistant
*att*Tn*7*::mini-Tn*7*T::3St	24 µg/ml	Resistant
*att*Tn*7*::mini-Tn*7*T::3St-rev	24 µg/ml	Resistant

* According to the BSAC Methods for Antimicrobial Susceptibility Testing (2009), Enterobacteria with a streptomycin MIC of >8 µg/ml should be considered streptomycin-resistant and those with an MIC of ≤8 µg/ml should be considered sensitive.

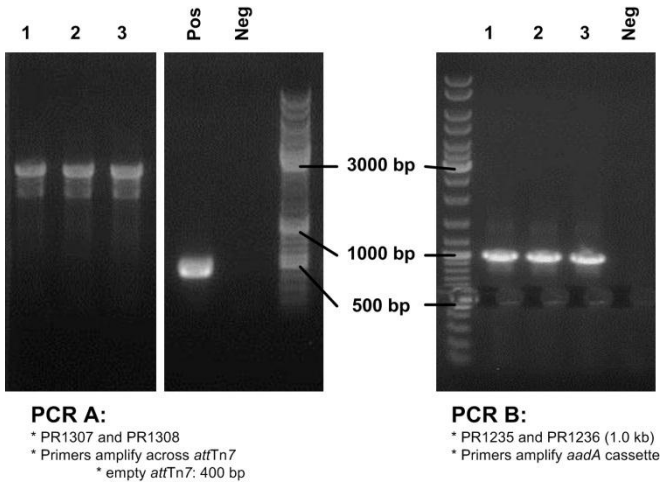
3.2.4.2 Site-specific integration of mini-Tn*7* cassettes into *att*Tn*7* in *K. pneumoniae*

In this section, only the mini-Tn*7* tagging of *K. pneumoniae* KR116 with the mini-Tn*7*T::*aadA1a* cassette from pJTOOL-9 will be described. Exactly the same framework was used for tagging other *K. pneumoniae* strains, but the antibiotic supplementation of growth media was altered when necessary.

Conjugal transfer of the mini-Tn*7*T::*aadA1a* delivery plasmid pJTOOL-9 into *K. pneumoniae* KR116 was obtained by four-parental mating. An equal mix of *E. coli*

DH5α*λpir*/pTNS2, *E. coli* HB101/pRK2013, *E. coli* CC118*λpir*/pJTOOL-9 and KR116 was prepared and conjugated overnight on filter discs placed aseptically on LA plates. Next, the matings were resuspended and dilutions were plated onto M9 minimal medium supplemented with 25 µg/ml of streptomycin and incubated overnight at 37ºC. M9 medium prevents the growth of all three conjugating *E. coli* strains and the streptomycin supplementation prevents growth of wildtype KR116. Three randomly selected colonies obtained after overnight incubation were analysed by colony PCR using primers PR1307 and PR1308 to confirm the site-specific integration of mini-Tn*7*T::*aadA1a* into the *att*Tn*7* site of KR116 (Figure 3-8). Additionally, the presence of the *aadA1a* cassette was confirmed with primers PR1235 and PR1236 (1.0 kb). All three colonies tested produced the expected bands in both PCR assays, thus confirming the construction of KR116/*aadA1a.*

Figure 3-8 :- PCR confirmation of site-specific transposition of mini-Tn*7*T::*aadA1a* into *att*Tn*7* of *K. pneumoniae* KR116



Three colonies of *K. pneumoniae* KR116 that had been putatively tagged with mini-Tn*7*T::*aadA1a* were colony PCR screened to confirm integration within *att*Tn*7* (PCR A) and the presence of the *aadA1a* cassette (PCR B). All three strains were confirmed to be site-specifically tagged using mini-Tn*7*T::*aadA1a*. Of note, the positive control in PCR A (wildtype KR116) highlights the size of the 'empty' *att*Tn*7* site.

This method routinely obtained between $1x10^4$ to $1x10^6$ putatively tagged *K. pneumoniae* colonies in strains KR116, KR162 and KR173. No insertions into sites other than *att*Tn*7* were observed in any of the nine mini-Tn*7* tagging experiments performed. However, no tagged strains could be constructed for KR161 and C3091. Since occupied *att*Tn*7* loci often prevent the site-specific transposition of Tn*7*-type elements via a local immunity effect, attempts were made to amplify across *att*Tn*7* of KR161 and C3091. Interestingly, no insertions were detected within the *att*Tn*7* sites of these strains and thus the basis of this Tn7-transposon immunity remains unknown. However, it is unlikely to be the result of a conjugation barrier as KR161 has been genetically altered using suicide vector-based allelic exchange (section 3.2.1) and C3091 has been mutated by plasmid-based conjugative delivery of mini-Tn*5* (384).

A total of nine tagged strains were constructed using the site-specific mini-Tn7 transposition system. These are summarised in Table 3-4. For more details, please refer to Appendix 4. Although originally the kanamycin and streptomycin resistance-conferring mini-Tn*7*T::Kan-3St cassette was constructed to tag KR116 and its isogenic mutants, when this system was ready the mutants had already been reconstructed in the spontaneous streptomycin-resistant mutant of KR116 (KR2107) using lambda Red recombination (section 3.2.3). Thus tagging of these strains was superfluous.

Nevertheless, an alternative use for the mini-Tn*7*T::Kan-3St cassette was found. Our Danish collaborators had previously observed that the *K. pneumoniae* C3091Δ*fimK* mutant strongly outcompeted the wildtype strain in an *in vitro* catheter biofilm competition assay (S. Stahlhut, personal communication) (386). The proportion of wildtype and Δ*fimK* strains present in biofilms were determined by differential plating on LA media and LA supplemented with kanamycin. However wildtype CFU was difficult to enumerate accurately as it lacked a selectable marker. They wanted to test the KR161, KR162 and KR173 strains and their respective *fimK* mutants in the same model, but unfortunately the gentamicin resistance cassette in *fimK*::gen was unsuitable for

selection, as described previously (section 3.2.2.1). To alleviate the aforementioned issues, both the wildtype and *fimK* mutant strains of KR162 and KR173 were tagged using the mini-Tn*7*T::Kan-3St cassette. As mentioned previously, KR161 appeared to show immunity to Tn*7* transposition and could therefore not be tagged. Next, the FRT-flanked kanamycin resistance cassettes were deleted from the tagged wildtype strains using Flp-mediated recombination (section 3.2.5). These alterations allowed easier and more accurate enumeration of the significantly outnumbered wildtype strains in the *in vivo* catheter biofilm competition assay.

Table 3-4 :- *K. pneumoniae* strains tagged using mini-Tn*7* system

KR catalogue number	Strain genotype
***K. pneumoniae* KR116**	
KR2257	KR116/*aadA1a*
KR2231	KR116/Kan-*aadA1a*
KR2258	KR116/P_{LAC}*strAB*
KR2259	KR116/3St
KR2281	KR116/3St-rev
***K. pneumoniae* KR162**	
KR2232	KR162/Kan-3St
KR2233	KR162Δ*fimK*/Kan-3St
***K. pneumoniae* KR173**	
KR2234	KR173/Kan-3St
KR2235	KR173Δ*fimK*/Kan-3St

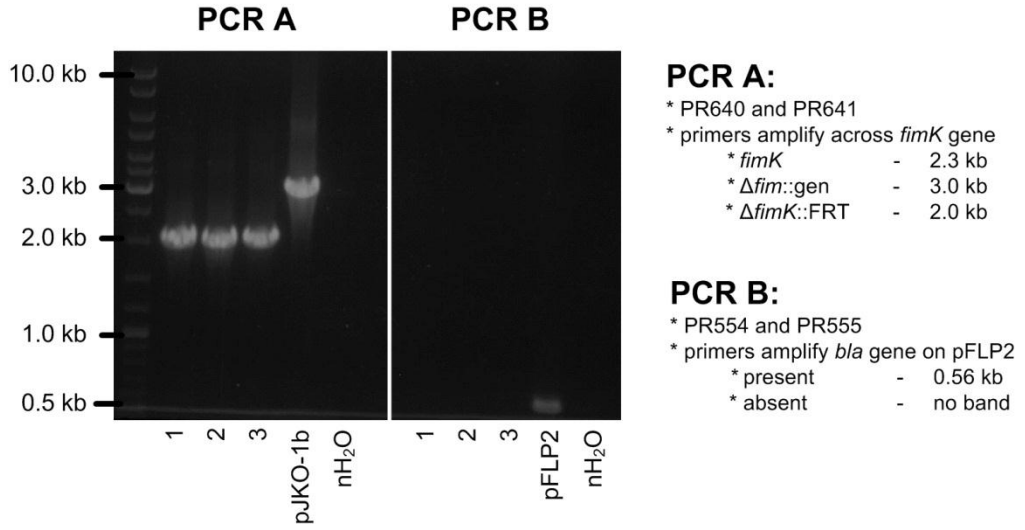
3.2.5 Flp-mediated FRT-recombination in *K. pneumoniae*

Many of the mutant alleles constructed in this study possessed antibiotic resistance cassettes flanked by FRT sites, which allow for deletion of the resistance cassette by the site-specific recombinase Flp. Details on which mutant alleles are compatible with Flp-mediated deletion can be found in Table 2-2. Flp recombinase can be provided *in trans* by the *sacB*-bearing plasmid pFLP2 (163).

As an example, the Flp-mediated deletion of the FRT-flanked gentamicin cassette from KR116Δ*fimK*::gen will be discussed. KR116Δ*fimK*::gen was transformed with pFLP2

and grown at 37ºC on LA supplemented with 200 µg/ml ampicillin, after which pFLP2 was cured by counterselection on LAS at 37ºC. Three colonies were then examined by colony PCR assays with primer pairs PR640/PR641 and PR554/PR555 to confirm the loss of the gentamicin resistance cassette and pFLP2, respectively (Figure 3-9). All three colonies were confirmed to have lost the gentamicin resistance cassette, and this was further confirmed by the inability of these colonies to grow on gentamicin-supplemented LA. Following Flp-mediated deletion of antibiotic cassettes remnant FRT scars are formed. A description of FRT-flanked resistance cassettes used in this work and their associated scars can be found in section 2.6.5.

Figure 3-9 :- Colony PCR confirmation of Flp-mediated loss of the FRT-flanked gentamicin resistance from KR116Δ*fimK*::gen



Three putative KR116Δ*fimK*::FRT colonies were screened for the Δ*fimK*::FRT allele (PCR A) and pFLP2 plasmid (PCR B). As expected, in PCR A all colonies had smaller bands compared to the positive control pJKO-1b, which harboured the *fimK*::gen allele. The product sizes matched those expected from amplification of *fimK*::FRT. The latter allele is 0.3 kb smaller than the wildtype *fimK* gene because an internal region is deleted during construction of the mutant allele, even though an 86 bp FRT scar remains (Table 2-2). PCR B confirms that curing of pFLP2 was successful.

The protocol described in this work has successfully removed FRT-flanked antibiotic cassettes from more than nine recombinant *K. pneumoniae* strains, the details of which can be found in Appendix 4. Deletion was successful in approximately 90 % of putative

deletants tested. Nevertheless, due to the inherent nature of *K. pneumoniae* strains to possess resistance to ampicillin, a non-ampicillin-dependent derivative of pFLP2 was constructed. To create pJTOOL-16, the *sacB* and *flp* genes from pFLP2 were cloned into the gentamicin resistant vector pUCP24, a high copy number vector that can be maintained and selected for in *K. pneumoniae* using growth media with 25 µg/ml of gentamicin (J. J. van Aartsen, unpublished observations). However, since all deletions using pFLP2 were successful the functionality of pJTOOL-16 has not been tested.

3.3 Discussion

3.3.1 *K. pneumoniae* suicide vector-based allelic exchange

In 2004 Philippe *et al.* described the construction of suicide vector pDS132, a pCVD442 derivative that lacked IS*1* and thus exhibited a reduced number of illegitimate plasmid recombination events (296). This work has built upon this improvement by constructing two pDS132-derivatives, pJTOOL-1 and pJTOOL-3, with enhanced and expanded MCSs. These new vectors were thoroughly tested in this work and were determined to be useful suicide vectors for allelic exchange in *K. pneumoniae* KR116, KR161, KR162 and KR173. These suicide vectors are currently being tested and used in *S. enterica* and *A. baumannii*.

An interesting observation during the allelic exchange experiments was that fewer merodiploids were obtained in *K. pneumoniae* strains KR161 and KR162. Upon visual inspection of colonies both strains had much thicker and more mucoid colonies than KR116 and KR173. It is therefore possible that the decreased efficiency was caused by reduced conjugative transfer of pDS132 derivatives from the *E. coli* donor due to physicochemical hindrance conferred by the thick capsule. Nevertheless, as only a single merodiploid was required this partial barrier did not impair the utility of the system.

At the start of this work it was identified that *sacB* counterselection in *K. pneumoniae* was highly inefficient, a previously unreported phenomenon. Although the protocols used in *E. coli* (30, 296), *P. aeruginosa* (67) and many other Gram negative bacteria were followed precisely, high numbers of false-positive sucrose-resistant colonies that had retained the integrated suicide vector were isolated rather than the desired double-crossover mutants. In the course of this project various aspects of the protocol were altered and their effect investigated. Although media type and sucrose concentration did not affect counterselection, it was identified that *sacB* counterselection in *K. pneumoniae* was much more effective at 25°C than at 30°C or 37°C. Most importantly, at 25ºC the number of false-positive sucrose-resistant colonies was greatly reduced and desired double-crossover mutants could easily be identified using colony PCR.

Although suicide vector-based allelic exchange is considerably more time consuming than the lambda Red-based method, this mutagenesis method is still highly applicable. Most importantly, after the second homologous recombination event there is a near 50:50 ratio of colonies with mutant and wildtype genotypes. This method should therefore readily permit markerless allelic exchange experiments, allowing for the switching of alternate functional alleles and other even more elegant genetic reengineering processes. This cannot be achieved using lambda Red recombination, which is highly inefficient in *K. pneumoniae* and relies on a selectable marker. Alternatively, this method serves as a backup for when lambda Red recombination fails either due to low transformation efficiency, the inability to support the lambda Red system-encoding plasmids or other factors.

Unfortunately, the Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* mutants of strain KR116 were constructed using the *aacC1* resistance cassette, which was shown in this work to confer only low-level gentamicin resistance in *K. pneumoniae* when present as a single copy. Indeed, Poteete *et al.* had previously described low-level gentamicin resistance

when the *aacC1* gene, originally derived from Tn*1696*, was present as only a single copy in *E. coli* (311). As the low level of resistance would cause difficulties in differentiating between competing strains in competition assays these mutant strains were reconstructed. In an attempt to speed up the mutagenesis process lambda Red-based mutagenesis of *K. pneumoniae* KR116 was explored.

3.3.2 *K. pneumoniae* lambda Red-based allelic exchange

Lambda Red-based allelic exchange is a much faster technique than suicide vector mutagenesis for two reasons: it does not require cloning of the mutant allele into a suicide vector and it does not require *sacB* counterselection. This work adds to the previous reports of lambda Red recombination in *K. pneumoniae* strains C3091 (387, 388), B5055 (77) and LM21 (13), by confirming the utility of the system in strain KR116 and its spontaneous streptomycin-resistant derivative KR2107. Although recombination reaction was highly inefficient and only 1 to 50 mutants were obtained per transformation, most importantly each transformation was successful at obtaining the desired recombinant strain. Minor protocol alterations have been described for lambda Red recombination in enterohaemorrhagic and enteropathogenic *E. coli* isolates; whether these improve efficiency in *K. pneumoniae* has not been investigated (268). Interestingly, KR116 was able to harbour the lambda Red plasmid pKOBEG-Apra but not pKD46. The reason for this is unknown as both plasmids have the same temperature sensitive origin of replication (*repA101ts*). The applicability of this system to the thick capsuled strains KR161 and KR162 remains to be explored.

Overexpression of recombination systems can lead to genome rearrangements or other undesired recombination events (164, 310). Therefore some protocols suggest transferring mutant alleles into a clean strain background using phage transduction. This was not done in this study as a *K. pneumoniae*-specific phage for generalised

transduction has not been fully characterised, although a workaround that generates *Klebsiella* strains sensitive to P1 transduction has been described (143).

This method allowed the relatively quick construction of Δ*fim*, Δ*fim2* and Δ*fim*Δ*fim2* mutants in KR2107. These mutants have subsequently been used for various *in vitro* and *in vivo* phenotypic analyses, which will be discussed in Chapter 5.

3.3.3 Mini-Tn*7* tagging of *K. pneumoniae* strains

The utility of mini-Tn*7* tagging of *K. pneumoniae* strains was also explored in this chapter. So far only one study has reported mini-Tn*7* tagging in *K. pneumoniae* (236). Using the methods described by Choi *et al.* (66), *K. pneumoniae* KR116, KR162 and KR173 were successfully tagged using mini-Tn*7* elements. In all cases transposition of the mini-Tn*7* element occurred site-specifically into the *att*Tn*7* site which was located immediately downstream of the *glmS* gene.

This useful system has dual functionality: it can be used for single copy gene complementation and for tagging strains with antibiotic resistance cassettes or fluorescent markers. Complementation experiments are most often performed by plasmid-based expression of the gene of interest. However, in many situations single copy chromosomal expression is more desirable. For example, both low and high copy number plasmids can cause non-physiological expression levels and can alter experimental results significantly (255). Additionally, chromosomally-integrated mini-Tn*7* elements are inherently stable as the Tn*7* transposase system-encoding plasmid is not maintained in the target strain and thus transposition of chromosomally-integrated elements cannot occur (255). The same cannot be said for plasmids, which must be maintained by a selection mechanism such as antibiotic selection and can be lost due to inaccurate segregation during bacterial replication (255). Moreover, in *S. enterica* even the mere presence of a non-virulence-related plasmid has been shown to indirectly affect virulence (205). Multiple studies have shown that the Tn*7*

insertion site, *att*Tn*7*, is a neutral chromosomal site since insertions do not affect host strain virulence or growth (66, 222, 255). Although the system has not been used for single copy complementation in this work, the confirmation of site-specific transposition of mini-Tn*7* elements into *att*Tn*7* in KR116, KR162 and KR173 confirms that, if necessary, the system can be applied for this purpose.

Only three out of the five *K. pneumoniae* strains tested were amenable to mini-Tn*7* transposition. No transposition occurred in KR161 and C3091. This effect is called target immunity and prevents insertions from occurring in DNA that already possesses a copy of the transposon. It has been well-described for Tn*7* and arises from the interaction of the Tn*7* proteins TnsB and TnsC: TnsB binds to integrated Tn*7* elements and accumulates to a high concentration, and displaces the TnsC transposase activator protein thus preventing transposition (293). Most intriguingly, the *att*Tn*7* site in KR161 and C3091 did not possess an insert. Although the majority of Tn*7* insertions are site-specific, it is possible that these strains both possess a Tn*7* or Tn*7*-like element at a different chromosomal region, similar to the alternative Tn*7* insertion sites described for *Pseudomonas putida* and *Burkholderia thailandensis* (66). The cause of Tn*7* target immunity in KR161 and C3091 remains to be investigated.

Additionally, this work described the construction of streptomycin and/or kanamycin mini-Tn*7*T derivatives. These cassettes were designed to confer high levels of streptomycin resistance to host strains, which could then be used in a streptomycin-treated mouse model of intestinal colonization. Strains harbouring the mini-Tn*7*T::3St cassette were shown to have higher streptomycin MICs than those harbouring cassettes possessing the *aadA1a* gene or *strAB* locus alone. This finding is in agreement with previously reported data, although the MICs measured in this work were considerably lower than those observed in the earlier study (393). It is unknown whether the 3St cassette confers enough streptomycin resistance to mediate survival of single-copy tagged bacteria in the intestinal tract of streptomycin-treated mice and

this awaits further investigation. Additionally, it remains to be determined whether IPTG induction of P_{LAC}*strAB* has an effect on streptomycin resistance levels, and with hindsight it would have been more appropriate to use a well-characterised constitutive promoter such as λP_R (210). Nonetheless, when these tags are used they can ensure an isogenic streptomycin resistance background between tagged strains that may be used in phenotypic assays and/or head-to-head competition assays.

In summary, the mini-Tn*7* tagging system described by Choi *et al.* is applicable to multiple *K. pneumoniae* strains, although Tn*7* target immunity may occur in some. The power of this system was not fully harnessed during this work, but it was shown that site-specific integration of mini-Tn*7* elements into the neutral *att*Tn*7* site in *K. pneumoniae* is possible and that the system can be used for single copy complementation and chromosomal tagging.

3.3.4 Flp-mediated FRT-recombination in *K. pneumoniae*

The versatile Flp-FRT recombination has an extremely broad host-range and has been used in bacteria, yeasts, plants and mammalian cells. It relies on the ability of the site-specific Flp recombinase from *Saccharomyces cerevisiae* to promote recombination between FRT sites (163). Similar systems have been described, such as the Cre-*lox* site-specific recombination system from phage P1 and the ParA-*res* system of plasmid RP4 (163).

This work investigated whether the system described by Hoang and colleagues for Flp-FRT recombination was applicable to *K. pneumoniae* (163). The protocol involved providing Flp recombinase *in trans* using pFLP2, and subsequently curing the plasmid using *sacB* counterselection (3, 163). The results show that this procedure was highly efficient at removing chromosomally-integrated FRT-flanked antibiotic resistance cassettes from *K. pneumoniae* strains. This approach enables resistance markers to be recycled in a reiterative fashion when generating subsequent mutations in a previously

generated mutant strains. Another advantage of this method is that FRT-reporter fusion constructs can be recombined into an FRT scar via Flp-recombination, thus allowing the expression and regulation of promoters upstream of the deleted region to be investigated (191). Unfortunately, as successively more FRT recognition sites accumulate within the genome the chance of large genome rearrangements increases (163). This was not observed in this work even though several strains were constructed with two FRT scars. Flp-mediated FRT-recombination is extremely valuable in *K. pneumoniae* as genetic manipulation of highly antibiotic resistant strains is limited by the reduced availability of appropriate selectable markers.

3.3.5 Future work

The work discussed in this chapter can be developed in many ways to further enhance the ability to genetically manipulate *K. pneumoniae* and related organisms. Potential aspects for future work are highlighted below:

Suicide vector-based allelic exchange

1/. Replacing the *cat* chloramphenicol resistance cassette of pJTOOL-1 and/or pJTOOL-3 with alternative antibiotic and non-antibiotic resistance markers will permit use of these suicide vectors in chloramphenicol-resistant and/or highly antibiotic resistant bacteria.

2/. Investigate whether *sacB* counterselection in closely-related bacteria, such as *Citrobacter* and *Enterobacter*, is also improved when performed at 25ºC.

Lambda Red-based allelic exchange

1/. Investigate whether *K. pneumoniae* KR161, KR162 and other strains that produce thick capsules are also amenable to lambda Red-based allelic exchange.

2/. Develop a generalized phage transduction system for transducing lambda-Red generated mutations into clean strain backgrounds.

3/. Test whether the protocol alterations described for enterohaemorrhagic and enteropathogenic *E. coli* improve efficiency of lambda Red recombination in *K. pneumoniae* (268).

Mini-Tn7 tagging

1/. Determine the cause of Tn*7* target immunity in KR161 and C3091. This could be done by creating a mini-Tn*5* knockout library of KR161/C3091, and using it to perform a mini-Tn*7* transposition experiment. Colonies that are obtained after mini-Tn*7* transposition should possess a mini-Tn*5* knockout of a gene involved in target immunity.

2/. Examine whether the triplet streptomycin resistance cassette 3St provides enough resistance to be used in the streptomycin-treated mouse intestinal colonization model.

Chapter 4. Marker rescue, sequence analysis, epidemiological profile and mobility probing of KpGI-5

4.1 Introduction

In earlier work, five tRNA gene insertion hotspots were screened for strain-specific DNA in sixteen clinical *K. pneumoniae* isolates (1). Strain KR116 was isolated from the blood of a patient with pneumonia and neutropenic septicaemia and was found to possess an 'occupied' *met56* tRNA locus by means of tRIP-PCR amplification across the *met56* tRNA locus using primers PR601 and PR602. Subsequent primer walking identified approximately 3 kb of novel island UA sequence which encoded putative fimbrial genes. This island was called KpGI-5. This chapter discusses the sequencing and sequence analysis of KpGI-5, in addition to the epidemiological profile and heterogeneity of KpGI-5 family islands within the *Klebsiella* genus. Finally, an investigation into the mobility of both a *met56*-associated phage in MGH78578 and KpGI-5 is presented.

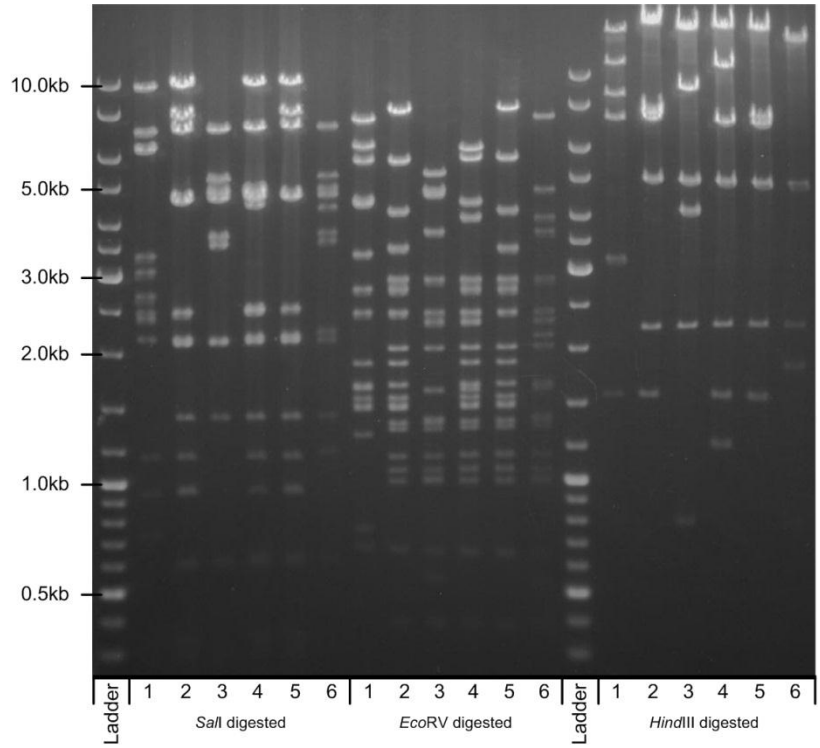
4.2 Results

4.2.1 Fosmid cloning, marker rescue and sequencing of KpGI-5

To capture and sequence the entirety of KpGI-5, the *fim2K* gene within the island UA was tagged with a kanamycin resistance cassette using allelic exchange. A fosmid library of this tagged strain (KR116Δ*fim2K*::kan) was created with the aim to isolate a fosmid clone harbouring an approximately 40 kb region encompassing both the novel KpGI-5 island and conserved UF and DF. To isolate kanamycin resistance cassette-bearing inserts using marker rescue, infected *E. coli* EPI300-T1R cells were plated on LB plates supplemented with chloramphenicol and kanamycin. Six kanamycin resistant

colonies that harboured Δ*fim2K*::kan were selected at random and their fosmids were extracted for further analysis (pJFos-1 to pJFos-6 (Appendix 6)). Initially each fosmid clone was restriction mapped using the enzymes *Sal*I, *Eco*RV and *Hind*III (Figure 4-1). pJFos-2 and -5 were identified to be highly similar, even more so than pJFos-3 and -6. pJFos-5 was therefore excluded from further analysis.

Figure 4-1 :- *Sal*I, *Eco*RV and *Hind*III restriction map of pJFos-1 to -6

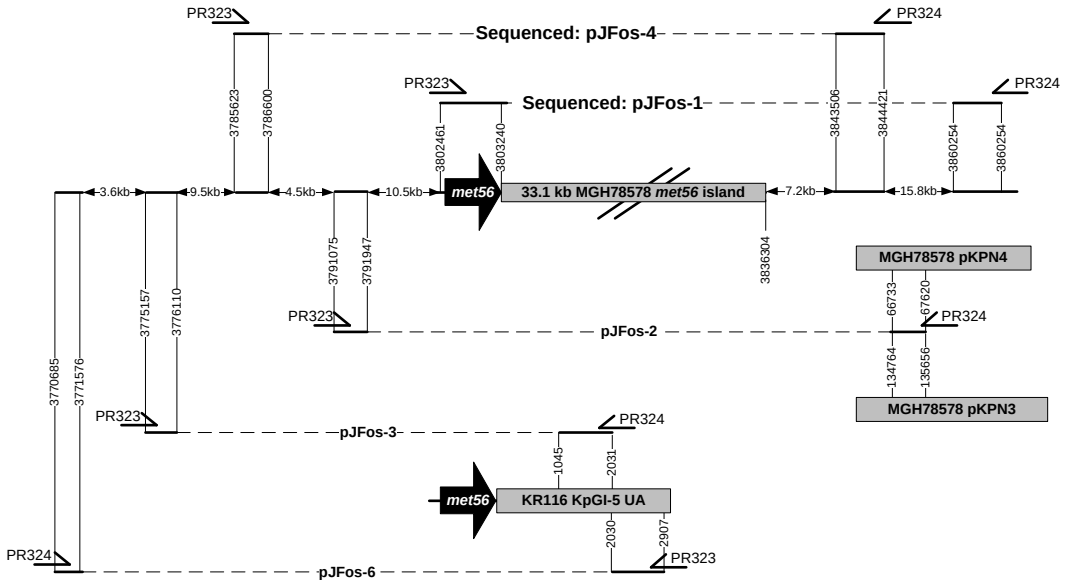


pJFos-X (where X corresponds to a number from 1 to 6 on the gel picture x axis (Appendix 6)) was restriction digested using enzymes *Sal*I, *Eco*RV and *Hind*III. Digested fosmids were subsequently resolved on a 0.7 % agarose gel to form a restriction map. The map identified that the fosmid pairs pJFos-2/pJFos-5 and pJFos-3/pJFos-6 share the majority of their restriction fragments.

To select clones for sequencing it was necessary to identify fosmids that possessed either the entire or the longest stretch of KpGI-5 island sequence. End sequencing of the inserts in pJFos-1 to -4 and -6 using primers PR323 and PR324 was performed to enable mapping of fosmid clones against the *met56* tRNA locus of *K. pneumoniae* MGH78578 (Figure 4-2). End sequencing of the insert in pJFos-1 and pJFos-4 showed 100 % identity to different regions of the conserved UF and DF of the MGH78578

met56 tRNA locus, suggesting these fosmids had both captured the entire KpGI-5 island. Based on end sequencing pJFos-3 and pJFos-6 were concluded to contain the conserved *met56* UF, whilst pJFos-2 appeared to only possess part of KpGI-5. Hence, pJFos-1 and pJFos-4 were selected as templates for 454 sequencing. Roche 454 pyrosequencing was performed at the University of Leicester NUCLEUS Genomics

Figure 4-2 :- Schematic showing fosmid end sequences mapped to the *K. pneumoniae* MGH78578 *met56* locus



DNA sequences obtained by end sequencing of inserts within fosmids pJFos-1 to -4 and -6 were scrutinised using BLASTn against the *K. pneumoniae* MGH78578 genome. pJFos-1 and pJFos-4 possessed inserts with end sequences that corresponded to both the UF and DF of *met56*, suggesting that they harboured the entire KpGI-5. pJFos-3 and pJFos-6 possessed the UF of *met56* and a short <3.0 kb segment of the KpGI-5 UA. The sequence obtained from pJFos-2 using PR324 matched to hypothetical genes in two previously sequenced plasmids from MGH78578, pKPN3 and pKPN4. UF, upstream flank; DF, downstream flank; UA, upstream arm.

Initial automated assembly from the 454 sequencing data generated five contigs: contig_1 (6643bp, 1867 reads), contig_2 (16810 bp, 1680 reads), contig_3 (16778 bp, 2744 reads), contig_4 (23819 bp, 6522 reads) and contig_5 (1334 bp, 374 reads). Contig_1 and contig_5 entirely matched the fosmid vector backbone and were excluded from further analysis. Contig_2 was 99 % identical (16740/16811 bp) to the MGH78578 *met56* locus conserved UF. Contig_3 was 99 % identical (16667/16780 bp)

to the conserved DF. On the other hand, contig_4 possessed the entire KpGI-5 island flanked by sequences which were 99 % (7797/7821 bp) and 98 % (845/860 bp) identical to the MGH78578 *met56* locus UF and DF, respectively. To finalise the assembly process, the junction between contig_2 and contig_4 was sequenced from a PCR product amplified using primers PR923 and PR925. Similarly, sequence corresponding to the junction between contig_3 and contig_4 was obtained using primer pair PR924/PR926. The final assembled sequence (57.4 kb) was constructed using the DNAStar contig assembly programme SeqMan.

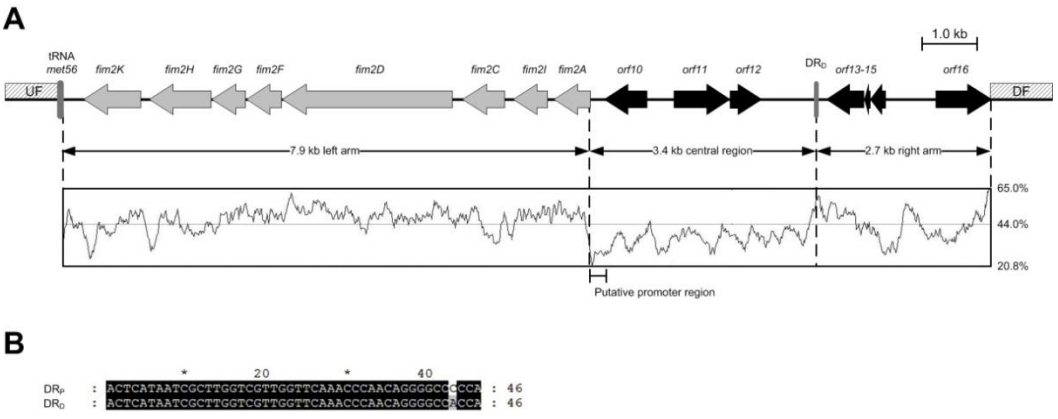
4.2.2 Sequence analysis of KpGI-5

KpGI-5 is a 14.0 kb insertion at the *met56* locus of KR116 with many features in common with typical GIs (Figure 4-3a). Firstly, the calculated G+C content (44.0 %) is much lower than the corresponding genome averaged values for *K. pneumoniae* MGH78578 (57.5 %) and Kp342 (57.3 %). Secondly, the island is present downstream of *met56*, which is a hotspot for GI integration in *K. pneumoniae* (454). And finally, the island possesses an almost perfect 46 bp direct repeat that corresponds to the 3' end of *met56* (Figure 4-3b). However, no putative integrase or mobility-associated genes were identified.

ORF Finder was used to analyse the KpGI-5 sequence to identify potential CDSs, which were defined as ORFs predicted to encode proteins larger than 100 aa (271). These results were compared to those obtained by Glimmer analysis of the KpGI-5 locus, and then manually curated (93). The closest protein homolog encoded by each CDS was identified using BLASTp (default parameters; nr database) and is presented in Table 4-1 and Appendix 8. Additionally, BLASTn (default parameters; nr/nt database) analyses were performed to identify closely related nucleotide sequences. The 2.7 kb segment mapping to the right arm of KpGI-5 was 90 % identical to a region immediately downstream of *met56* in *K. pneumoniae* Kp342 and was predicted to

encode two hypothetical proteins (Orf14 and Orf15), a metallo-β-lactamase family protein (Orf16) and a putative GCN5-related N-acetyltransferase (Orf13). The nucleotide sequence of a 3.4 kb central region did not match any GenBank entries and coded for three novel proteins; Orf10 and Orf11 exhibited weak matches to putative regulatory proteins from *Stigmatella aurantiaca* DW4/3-1 and *Serratia odorifera* DSM 4582, respectively. Orf10 also possessed a match to the Pfam domain Trans_reg_C (PF00486) which has been implicated in DNA binding, further suggesting a role for Orf10 in regulation.

Figure 4-3 :- Features of the novel KpGI-5 island in *K. pneumoniae* KR116



(A) Genetic organisation of KpGI-5 shown lying between the species-conserved upstream flank (UF) and downstream flank (DF) sequences. The eight putative fimbrial genes are labelled *fim2A*–*fim2K*. Closest BLASTp similarities for these and other predicted KpGI-5-encoded proteins are described in Table 4-1. KpGI-5 segments indicated by double arrows map to G+C % transitions as indicated by the profile below. **(B)** Alignment of the tRNA-proximal (DR_P) and tRNA-distal (DR_D) 46 bp direct repeat (DR) sequences associated with KpGI-5. DR_P comprises the 3' end of *met56*. The KpGI-5 sequence has been deposited in GenBank: JN181158

The 7.9 kb left arm of KpGI-5 harboured a novel eight-gene cluster that exhibited sequence similarity and organizational identity to the chromosomally-encoded *fim* operons of *Citrobacter koseri* ATCC BAA-895 (~60 %) and *K. pneumoniae* C3091 (~51 %). This cluster was named *fim2*. It coded for homologs of all structural and biosynthesis-associated components of the well characterized C3091 type 1 fimbrial system, including a major fimbrial subunit (named Fim2A), three minor fimbrial subunits

(Fim2F, Fim2G and Fim2H), and a chaperone (Fim2C) and usher (Fim2D) (387). Downstream of *fim2H* is *fim2K* which encoded a FimK homolog that possessed a matching EAL domain but lacked a FimK-equivalent N-terminal helix-turn-helix domain. Amino acid sequence identities between cognate *fim2* and *fim* products varied from 60 to 92 %. However, no homologs of the C3091 *fimB*, *fimE* or *fimS*-like invertible promoter switch could be identified upstream of *fim2*. *K. pneumoniae* KR116 was also confirmed to possess the species-conserved *fim* and *mrk* operons by PCR screening for the *fimH* and *mrkD* adhesin genes using primer pairs PR1144/PR1145 and PR1150/PR1151, respectively.

Table 4-1 :- BLASTp homologs of proteins predicted to be encoded by KpGI-5

Gene name	Coding region (bp)	Protein size (aa[a])	Closest BLASTp homolog[#]			
			Percentage identity (aa[a])	Organism	Possible function [GenBank Number]	E value
met56	180..255 (76)	100 % identical to methionine tRNA in *K. pneumoniae* MGH78578 [KPN_03476]				
fim2K	1385..528 (858)	285	60 % (165/276)	*C. koseri* ATCC BAA-895	Putative EAL domain protein [ABV14791.1]	$1e^{-94}$
fim2H	2440..1514 (927)	308	62 % (190/308)	*K. pneumoniae* sp15	Fimbrial adhesin (FimH) [ACL13802.1]	$1e^{-101}$
fim2G	2961..2458 (504)	167	72 % (120/167)	*C. koseri* ATCC BAA-895	Minor fimbrial subunit (FimG) [ABV14789.1]	$2e^{-65}$
fim2F	3501..2974 (528)	175	79 % (138/175)	*C. koseri* ATCC BAA-895	Minor fimbrial subunit (FimF) [ABV14788.1]	$1e^{-73}$
fim2D	6073..3515 (2559)	852	82 % (689/838)	*C. koseri* ATCC BAA-895	Outer membrane usher protein (FimD) [ABV14787.1]	0.0
fim2C	6858..6229 (630)	209	92 % (190/207)	*K. variicola* At-22	Fimbrial chaperone protein (FimC) [ADC56706.1]	$2e^{-107}$
fim2I	7519..6989 (531)	176	82 % (139/170)	*C. koseri* ATCC BAA-895	Fimbrial protein (FimI) [ABV14784.1]	$2e^{-80}$
fim2A	8148..7600 (549)	182	88 % (160/182)	*K. pneumoniae* MGH 78578	Major fimbrial protein (FimA) [ABR78685.1]	$1e^{-79}$
orf10	9002..8355 (648)	215	37 % (24/65)	*S. aurantiaca* DW4/3-1	Putative two component system regulatory protein [EAU69265.1]	0.019
orf11	9409..10254 (846)	281	28 % (77/277)	*S. odorifera* DSM 4582	Putative transcriptional regulatory protein [EFE96725.1]	$3e^{-20}$
orf12	10251..10727 (477)	158	29 % (38/130)	*S. odorifera* DSM 4582	Hypothetical protein [EFE96270.1]	$1e^{-13}$
orf13	12266..11694 (573)	190	97 % (184/190)	*Klebsiella* sp. 1_1_55	Putative GCN5-related N-acetyltransferase [EFD84432.1]	$1e^{-106}$
orf14	12387..12268 (120)	39	100 % (39/39)	*K. pneumoniae* 342	Hypothetical protein [ACI07992.1]	$1e^{-12}$
orf15	12616.. 12359 (234)	77	92 % (71/77)	*K. pneumoniae* 342	Hypothetical protein [ACI06987.1]	$1e^{-34}$
orf16	13342..14187 (846)	281	91 % (256/281)	*K. pneumoniae* 342	Metallo-beta-lactamase family protein [ACI07748.1]	$1e^{-151}$

[#] The BLASTp alignments for each of the proteins predicted to be encoded by KpGI-5 can be found in Appendix 8.

4.2.3 Epidemiological profile of the KpGI-5 genomic island family within *Klebsiella* spp.

For the purpose of epidemiological profiling, KpGI-5 and KpGI-5-like islands were defined as a family of *met56* tRNA-associated GIs that possessed a *fim2K* gene. Based on this definition the prevalence of the KpGI-5 island family was determined by PCR screening for *fim2K* using primer pair PR615/PR616. Although all investigated isolates had previously been biochemically characterised as *Klebsiella* species, these results were checked using PCR assays specific for *K. pneumoniae* subsp. *pneumoniae* and *K. pneumoniae* subsp. *ozaenae* (primer pair PR1261/PR1262), and *K. oxytoca* (primer pair PR1263/PR1264), as described previously (213, 235, 423). The former targets species-specific regions of the 16S-23S rDNA internal transcribed spacer whilst the latter amplifies the *K. oxytoca*-specific polygalacturonase *pehX* gene which is involved in pectate degradation (213, 235).

162 *Klebsiella* isolates were screened for the *fim2K* gene (Table 4-2), including 61 well-characterised isolates obtained from Dr. C. Struve's strain catalogue (383). 21 strains were *fim2K*-positive, including 16 strains identified as *K. pneumoniae* (13.0 % of all *K. pneumoniae* isolates examined) and two as *K. oxytoca* strains (11.1 % of all *K. oxytoca* isolates examined) (Table 4-2 and Table 4-3). Further examination using species-specific PCR revealed that two of the *fim2K*-positive strains (sp25 and sp28) had been probably misidentified as *K. pneumoniae* but were in fact *K. oxytoca*. Alternatively, it is possible that *pehX* was acquired horizontally by these strains although this is unlikely as there have been no reports of *Klebsiella* species other than *K. oxytoca* that are able to degrade pectin (153). The species identification of another *fim2K*-positive strain (KR2315) identified as *K. pneumoniae* could not be confirmed using the strain-specific PCR assays suggesting that it was either a *K. pneumoniae* subsp. *rhinoscleromatis* or an alternative *Klebsiella* species. The latter is more likely since *fimH* and *mrkD* could not be identified in KR2315, but were present in

K. pneumoniae subsp. *rhinoscleromatis* type strain ATCC 13884. Unexpectedly, the strains sp15 and sp30 which had been identified biochemically as *K. pneumoniae* produced amplicons with both species-specific PCR assays. Nevertheless, these results have shown that KpGI-5-like islands occur in approximately 10 – 15 % of *K. pneumoniae* and *K. oxytoca* strains.

Table 4-2 :- Prevalence of *fim2K* within the *Klebsiella* genus

	Total [#]	***fim2K* +**	
K. pneumoniae	123	16	13.0 %
K. species	19	3	6.3 %
K. oxytoca	18	2	9.0 %
K. ornithinolytica	1	0	0 %
K. planticola	1	0	0 %
All	162	21	13.0 %

[#] Numbers based on original biochemical *Klebsiella* species assignments

Table 4-3 :- *fim2K*-positive *Klebsiella* strains and associated characteristics

Strain identifier	Biochemical identification	PCR identification	Specimen type	Isolate origin*	*fimH* #	*mrkD* #	*fim2K* #
KR 116	*K. pneumoniae*	*K. pneumoniae*	Blood	UHL-LRI, UK	Pos	Pos	Pos
KR 392	*K. pneumoniae*	*K. pneumoniae*	Urine	UHL-LRI, UK	Pos	Pos	Pos
KR 396	*Klebsiella* sp.	*K. pneumoniae*	Urine	UHL-LRI, UK	Pos	Pos	Pos
KR 399	*Klebsiella* sp.	*K. pneumoniae*	Urine	UHL-LRI, UK	Pos	Pos	Pos
KR 518	*K. pneumoniae*	*K. pneumoniae*	Sputum	Shanghai, China	Pos	Pos	Pos
KR1120	*Klebsiella* sp.	*K. pneumoniae*	Unknown	UHL-LRI, UK	Pos	Pos	Pos
KR2159	*K. oxytoca*	*K. oxytoca*	Blood	UHL-LRI, UK	**Neg**	**Neg**	Pos
KR2163	*K. oxytoca*	PCRs negative	Wound swab	UHL-LRI, UK	**Neg**	**Neg**	Pos
KR2175	*K. pneumoniae*	*K. pneumoniae*	Peritoneal dialysis fluid	UHL-LRI, UK	Pos	Pos	Pos
KR2315	*K. pneumoniae*	PCRs negative	Unknown	USA	**Neg**	**Neg**	Pos
KR2317	*K. pneumoniae*	*K. pneumoniae*	Unknown	USA	Pos	Pos	Pos
KR2318	*K. pneumoniae*	*K. pneumoniae*	Unknown	USA	Pos	Pos	Pos
sp15	*K. pneumoniae*	*K. pneumoniae* or *K. oxytoca*	Blood	Cph, Denmark	Pos	Pos	Pos
sp25	*K. pneumoniae*	*K. oxytoca*	Blood	Cph, Denmark	**Faint**	Pos	Pos
sp28	*K. pneumoniae*	*K. oxytoca*	Blood	Cph, Denmark	**Faint**	Pos	Pos
sp29	*K. pneumoniae*	*K. pneumoniae*	Blood	Cph, Denmark	Pos	Pos	Pos
sp30	*K. pneumoniae*	*K. pneumoniae* or *K. oxytoca*	Blood	Cph, Denmark	Pos	Pos	Pos
sp34	*K. pneumoniae*	*K. pneumoniae*	Blood	Cph, Denmark	Pos	Pos	Pos
cas128	*K. pneumoniae*	*K. pneumoniae*	Water	S-H, Germany	Pos	Pos	Pos
cas664	*K. pneumoniae*	*K. pneumoniae*	Urine	S-H, Germany	Pos	Pos	Pos
cas669	*K. pneumoniae*	*K. pneumoniae*	Urine	S-H, Germany	Pos	Pos	Pos

* Cph, Copenhagen; S-H, Schleswig-Holstein; UHL-LRI, University Hospitals Leicester - Leicester Royal Infirmary; UK, United Kingdom. # Pos, positive; Neg, negative.

Further examination suggested that the specimen type from which a strain was obtained was not a predictor of the presence or absence of *fim2K* (Table 4-4). *fim2K* was identified in strains isolated from wounds, ascitic fluid, blood, sputum, urine and the environment. Although *fim2K* was not identified in strains derived from biliary fluid, cerebrospinal fluid, PLA or nasopharynx, this is probably a function of the low numbers of strains examined for these specimen types and the relatively low prevalence of *fim2K* in *Klebsiella* strains. In addition, inspection of the geographical origins of strains revealed that *fim2K*-positive strains were not limited to one geographical area (Table 4-3). Although *fim2K* was originally identified in KR116, a strain isolated in Leicester, *fim2K*-positive strains were subsequently found to have been isolated in Germany,

Denmark, USA and China. This represents a truly global spread of KpGI-5 island family.

Table 4-4 :- Prevalence of *fim2K* in strains from various isolation sources

	Total	***fim2K +***
Ascitic Fluid	9	1
Biliary Fluid	1	0
Blood	48	8
Cerebrospinal fluid	2	0
Environmental	11	1
PLA aspirates	11	0
Nasopharynx	3	0
Sputum	11	1
Unknown	20	4
Urine	45	5
Wound	1	1
All	162	21

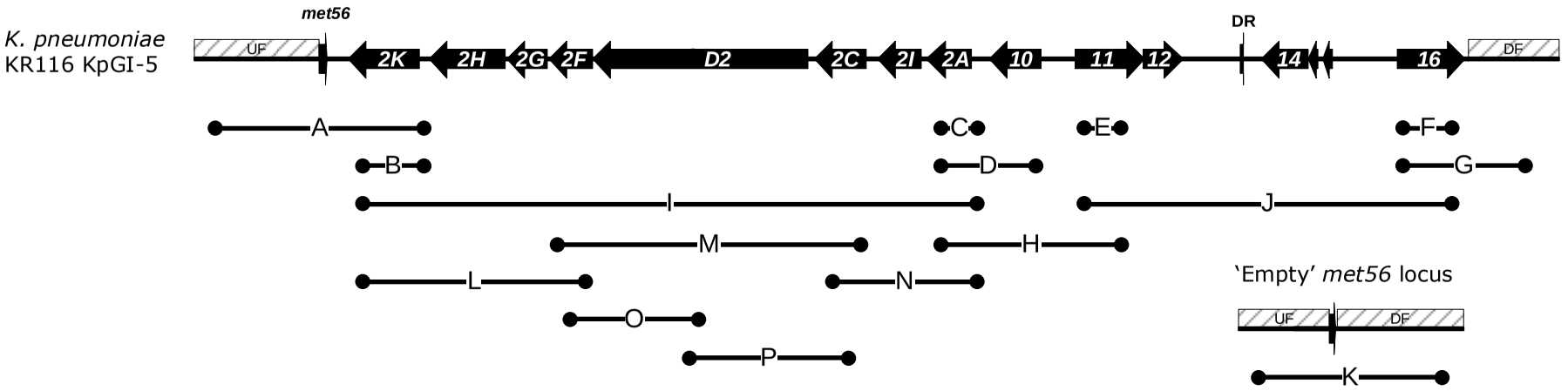
All 162 *Klebsiella* strains were also screened for the presence of the *K. pneumoniae* type 1 and type 3 fimbriae by PCR identification of the *fimH* (PR1144/PR1145) and *mrkD* (PR1150/PR1151) adhesin genes, respectively. 75.3 % (122 out of 162) of *Klebsiella* strains investigated were positive for the *fimH* gene, whilst 80 % (127 out of 162) were positive for the *mrkD* gene. Within the *fim2K*-positive strains, 76.0 % and 85.7 % of strains were found to harbour the *fimH* and *mrkD* genes, respectively. It should be noted that the majority of strains that did not have *fimH* or *mrkD* were identified as *Klebsiella* species other than *K. pneumoniae*. Interspecies variations in *mrkD* sequences have been identified (354, 388). It is therefore reasonable to assume that minor variations also occur within *fimH* genes and *fim2K* genes, which may lead to suboptimal primer annealing and failed PCR amplification. It is therefore possible that the actual prevalence of type 1, type 3 and *fim2* fimbrial gene clusters in the *Klebsiella* collection studied is higher than identified in this work.

Faint amplicons were obtained during *fimH* PCR screening of sp25 and sp28. These amplicons were sequenced and determined to be identical. The closet nucleotide match was the *fimH* gene of *K. variicola* At-22 (1076 / 1150 bp), whilst the predicted protein product was 99 % (301 / 302 aa) identical to FimH protein of *K. pneumoniae* subsp. *rhinoscleromatis* ATCC 13884. These strains had previously been biochemically identified as *K. pneumoniae* (154). However, PCR identification in this work identified these strains as *K. oxytoca*. To aid further identification, the indole test was repeated and was determined to be negative, inconsistent with classification as *K. oxytoca*. Intriguingly, PCRs using primers to amplify *fimB* (PR1142/PR1143) and *fimK* (PR1016/ PR1017) were negative whilst amplification of the *fimS* switch (PR561/PR562) was strongly positive.

4.2.4 Heterogeneity of KpGI-5 family members

Twenty one *fim2K*-positive *Klebsiella* strains were identified in the previous section, including *K. pneumoniae* KR116. These strains were further examined by PCR mapping to investigate whether *fim2K* was always associated with the remaining *fim2* and KpGI-5 genes, and to explore potential structural diversity of the KpGI-5 island family. Eighteen primer pairs were designed for KpGI-5 PCR-based mapping; primer pairs and their predicted amplicons, based on the known KpGI-5 sequence, are illustrated in Figure 4-4. A combination of normal PCR and long-range PCR was used.

Figure 4-4 :- Schematic of primers used to investigate the heterogeneity of KpGI-5 and KpGI-5-like islands

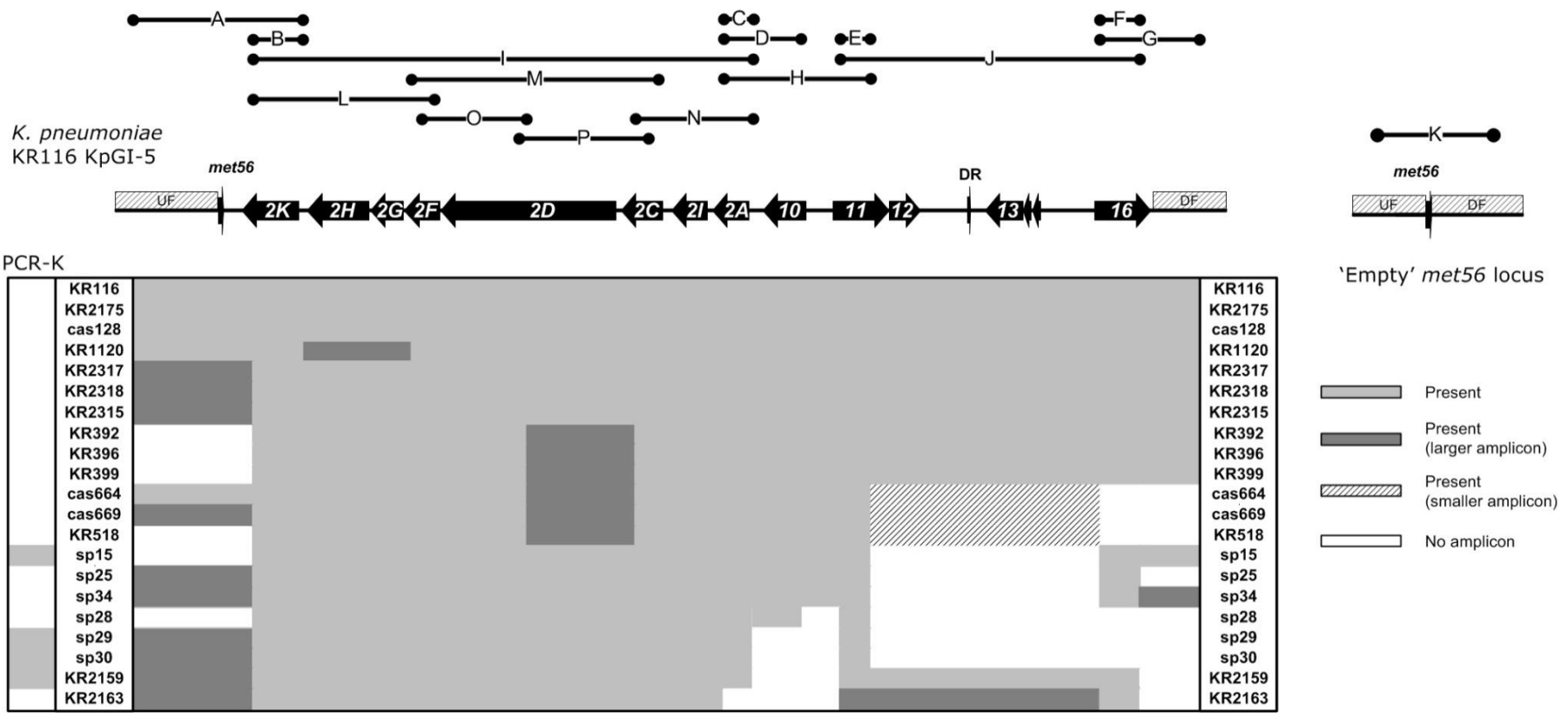


PCR	Primers	Expected Size
A	PR601 + PR616	1.0 kb
B	PR615 + PR616	0.5 kb
C	PR1367 + PR1368	0.6 kb
D	PR1368 + PR1223	1.5 kb
E	PR1114 + PR1115	0.4 kb
F	PR1106 + PR1107	0.4 kb
G	PR647 + PR1106	1.8 kb
H	PR1367 + PR1115	2.4 kb

PCR	Primers	Expected Size
I	PR615 + PR1367	8.8 kb
J	PR1114 + PR1107	4.4 kb
K	PR601 + PR647	2.1 kb
L	PR615 + PR1572	2.9 kb
M	PR1573 + PR1574	3.8 kb
N	PR1367 + PR1575	1.9 kb
O	PR1580 + PR1579	1.5 kb
P	PR1581 + PR1578	1.8 kb

This schematic highlights the eighteen PCR reactions and corresponding primer pairs used to investigate the structure of KpGI-5 island family members. Importantly PCR reactions B, C, D, H, I, L, M, N, O and P confirm the genetic structure of the *fim2* operon. Additionally, PCR reactions A and G detect the association of the KpGI-5 UA and DA with the conserved UF and DF, respectively. And finally, PCR K investigates the occupancy status of the *met56* locus. DR, direct repeat; UF, upstream flank; DF, downstream flank)

Figure 4-5 :- Heterogeneity of KpGI-5 and KpGI-5 like genomic islands



Comparison of the genetic structure of KpGI-5-like islands in *fim2K*-positive *Klebsiella* strains to the archetypal KpGI-5 island in KR116. Strains have been organised from most similar to KpGI-5 (top) to least similar to KpGI-5 (bottom). Homologous DNA regions are indicated in light grey. Heterogeneous regions that produced larger, smaller or no amplicons are illustrated using dark grey, diagonal shading or white, respectively. DR, direct repeat; UF, upstream flank; DF, downstream flank)

Results of the PCR mapping assay are shown in Figure 4-5. Out of 20 strains examined only two were found to have a PCR-profile identical to that of KpGI-5: KR2175 and cas128. The remaining strains showed significant differences compared to the KpGI-5 island, including variations within the *fim2* operon, upstream arm-flank junction, downstream arm and downstream arm-flank junction (terms defined in Figure 1-2).

Of the regions examined, the *fim2* operon was the most conserved. Strain KR1120 possessed a 2.1 kb insertion between the *fim2H* and *fim2G* genes, whilst six strains were found to have a 1.2 kb insertion within *fim2D*. The latter insertion in strain KR392 was sequenced and identified to be an IS*630* element. Previous studies have shown that IS*630* has a strong preference for inserting within 5'-CTAG-3' sequences (406). Consistent with this finding, IS*630* was found inserted within a 5'-CTAG-3' sequence 780 bp downstream of the *fim2D* start codon. Additionally, IS*630* encoded a stop codon in-frame with the *fim2D* gene.

The KpGI-5 central and right arm regions were both conserved in 45 % of strains. Whilst *orf10* could not be detected in four strains, the region spanning *orf11* to the 5' end of *orf16* (PCR J) was discrepant in ten strains. This latter region could not be amplified in six cases, whilst amplicons obtained with three strains were smaller than expected (2.2 kb). Direct sequencing of these amplicons was unsuccessful. In one strain, KR2163, PCR J yielded a product 1.6 kb larger than expected. This amplicon was not examined further.

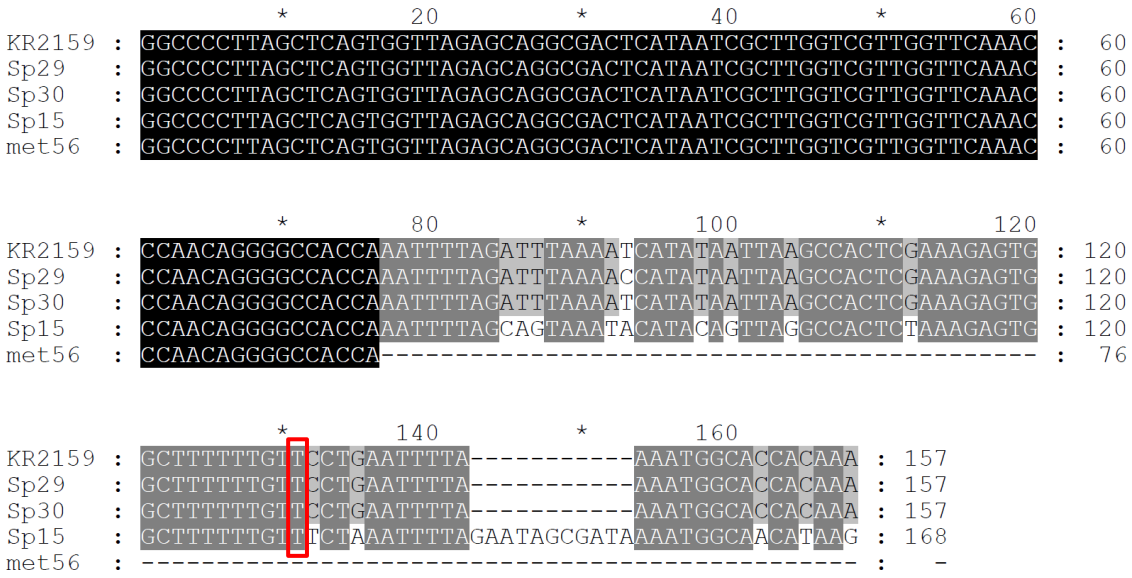
Examination of the association of *fim2K* with the UF revealed that only five strains produced PCR bands of the same size as would be expected for KpGI-5. No amplicons were obtained for six strains, whilst ten produced larger PCR amplicons that were between 1.8 kb and >10.0 kb in size. Several unsuccessful attempts were made to sequence selected amplicons. Eleven strains possessed KpGI-5-like elements that

were associated with the *met56* DA (PCR G). Although this association could not be confirmed in nine strains, a product 0.5 kb larger than expected was obtained for strain sp34. This product was sequenced from both ends and identified to be specific, confirming the association with the *met56* DA; the central portion of this amplicon still awaits sequencing.

PCR K was developed to investigate the occupancy status of the *met56* tRNA locus: no amplicon would be produced if KpGI-5, KpGI-5-like or other *met56*-associated GIs were located within the *met56* tRNA locus. After 35 PCR cycles no amplicon was obtained in 17 strains, consistent with the notion that KpGI-5-like elements or other GIs had inserted site-specifically into *met56*. Faint PCR bands corresponding to an unoccupied *met56* locus were obtained with four strains: sp15, sp29, sp30 and KR2159. Importantly, the island in sp15 was previously linked to the *met56* locus by PCR G, suggesting that a small proportion of sp15 cells did not possess an integrated KpGI-5-like element. These results indicated that the KpGI-5 island family includes mobile islands that are spontaneously excising.

Mobile islands often possess integrases that mediate site-specific integration. In order to seek out a putative KpGI-5 island family integrase, the faint amplicons from PCR K were sequenced and aligned using ClustalX (Figure 4-6). The alignment highlights that *met56* is identical in all four *Klebsiella* strains. Downstream of *met56* the sequence is almost identical in KR2159, sp29 and sp30. In turn, the predicted consensus sequence is 99 % identical to the empty *met56* locus in *K. pneumoniae* NTUH-K2044. However, sp15 shows 14 single nucleotide polymorphisms and an 11 bp insertion within this region.

Figure 4-6 :- ClustalX alignment of PCR K DNA sequences (*met56* loci) from strains with spontaneously excising KpGI-5-like elements



This alignment shows that the *met56* tRNA gene is 100 % conserved in all four strains. The region downstream of *met56* is also highly conserved in all strains but sp15, which possesses multiple single nucleotide polymorphisms and an 11 bp insertion. Based on analyses against the MGH78578 sequence (Figure 4-8) it is likely that islands integrate just downstream of a conserved 101 bp repeat which ends at base T-131, as highlighted by the red box.

4.2.5 Investigation into the mobility of *met56*-associated islands

In the previous section four *Klebsiella* strains were determined to possess spontaneously excising KpGI-5-like elements or *met56*-associated GIs. No integrase gene sequences were identified within the amplicons sequenced. However, due to limitations of the PCR mapping strategy used the possible existence of a KpGI-5 family integrase gene could not be ruled out. As an alternative, sequenced genome sequences were screened bioinformatically for integrases belonging to GIs or prophages that had inserted site-specifically into a *met56*-equivalent tRNA locus. Remarkably, the *met56* locus of *K. pneumoniae* MGH78578 was found to possess a 33.1 kb prophage with an integrase gene. However, this prophage did not exhibit DNA and/or protein similarity to KpGI-5, apart from the flanking repeat units. In the following

section the mobility of this prophage and the ability of its site-specific integrase to excise KpGI-5 from the chromosome were investigated.

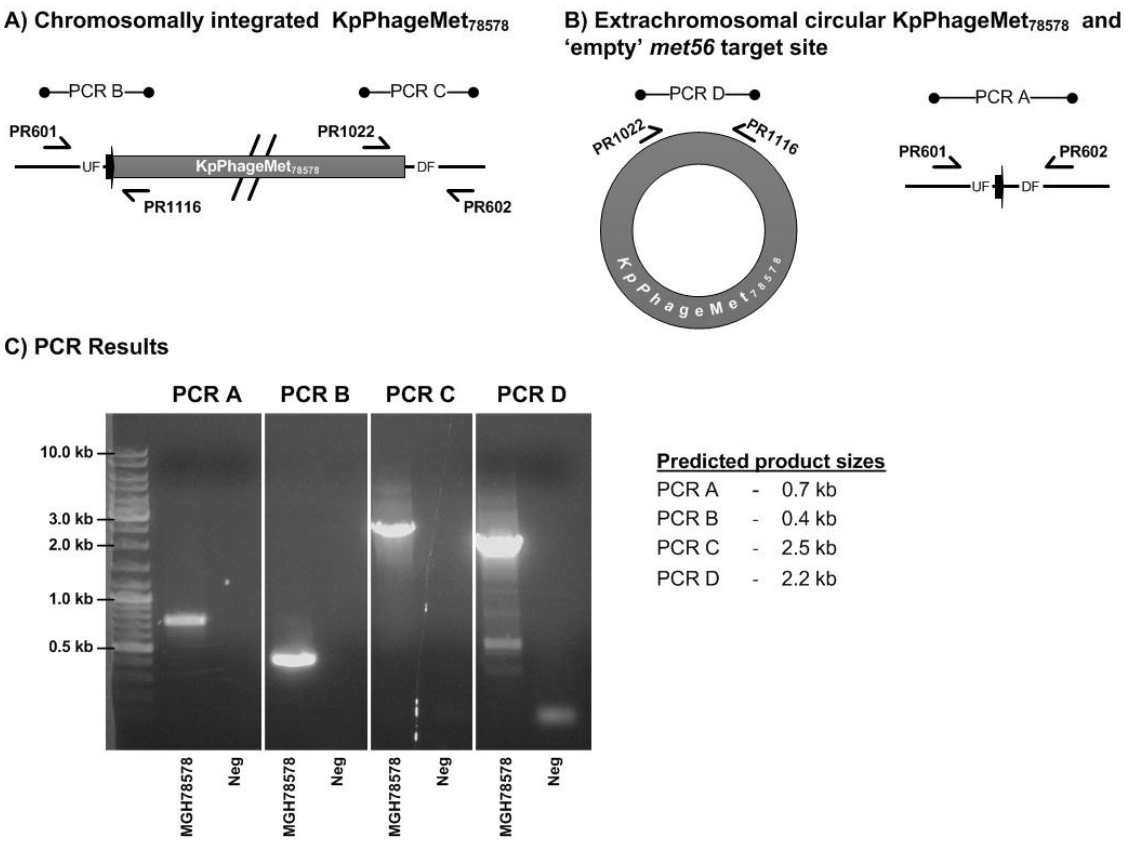
4.2.5.1 The *met56*-associated KpPhageMet$_{78578}$ element spontaneously excises from the MGH78578 genome

tRNAcc analysis revealed that the *met56* locus in MGH78578 possessed a 33.1 kb strain-specific insertion that encoded a putative integrase gene (KPN_04894) (281). Further examination of the sequence revealed the insertion was a chromosomally-integrated prophage that encoded hypothetical proteins, phage structural components and phage regulators. This prophage was named KpPhageMet$_{78578}$. A PCR-based probing assay was designed to investigate whether KpPhageMet$_{78578}$ was able to excise spontaneously and form an extrachromosomal circular element (Figure 4-7A and B). In addition to the chromosomally-integrated prophage, an 'unoccupied' *met56* tRNA locus and a KpPhageMet$_{78578}$ derived circular element were also detected in MGH78578 gDNA preparations (Figure 4-7C). The latter two amplicons were sequenced as these contained the chromosomal (PCR A; *attB*) and phage (PCR D; *attP*) attachment sites, respectively (374).

With the exception of one nucleotide polymorphism, the KpPhageMet$_{78578}$ *attB* site was identical to that described for strains Sp29, Sp30 and KR2159, which were likely to harbour spontaneously excising KpGI-5-like elements (Figure 4-7). Alternatively, these empty sites could be related to the spontaneous loss of another GI. The KpPhageMet$_{78578}$ *attB* and *attP* sites possessed an identical 101 bp unit which was flanked by conserved UF/DF or phage UA/DA sequence, respectively (Figure 4-8A and B). This unit comprised 46 bp of the 3' end of *met56* and 55 bp of adjacent downstream sequence. Integrase-mediated recombination occurs between *attB* and *attP* when phages and GIs integrate site-specifically, resulting in the formation of *attL* and *attR* sites at the left and right insertion junctions, respectively (374). Similar to *attB* and *attP*,

the KpPhageMet$_{78578}$ *attL* and *attR* sites also possessed the same 101 bp unit (Figure 4-8C). Together these 101 bp units form a direct repeat in the MGH78578 chromosome.

Figure 4-7 :- PCR assay to probe the mobility of KpPhageMet$_{78578}$

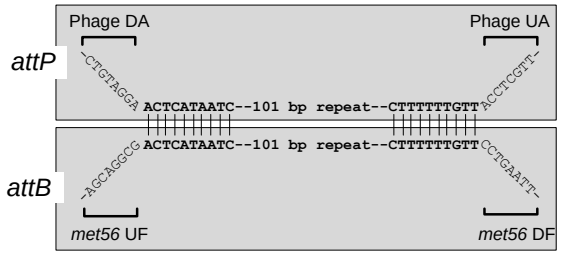


A) The chromosomally-integrated form of KpPhageMet$_{78578}$ can be detected using PCR B and/or PCR C. This form was present in the sequenced MGH78578 genome. **B)** The circular form of KpPhageMet$_{78578}$ and the resulting empty *met56* tRNA locus were identified by PCR D and A, respectively. **C)** Agarose gel showing results for PCRs A, B, C and D performed on MGH78578 gDNA. All predicted structures were detected by PCR, suggesting that KpPhageMet$_{78578}$ spontaneously excised and formed a circular element. UA, upstream arm; DA, downstream arm; UF, upstream flank; DF, downstream flank.

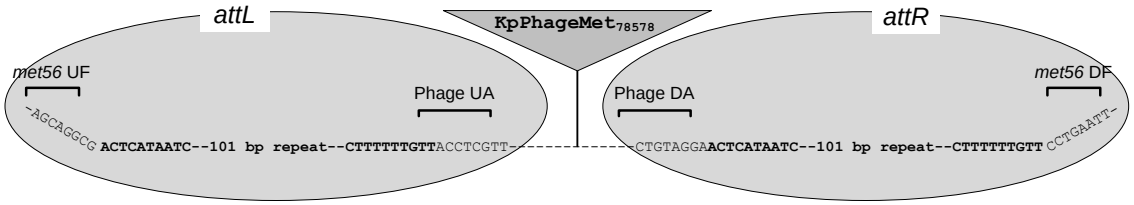
Figure 4-8 :- The *attB, attP, attL* and *attR* sites of KpPhageMet$_{78578}$

A) Conserved central 101 bp repeat unit of all KpPhageMet$_{78578}$ *att* sites

B) *attP* and *attB* sites of KpPhageMet$_{78578}$



C) *attL* and *attR* sites of KpPhageMet$_{78578}$



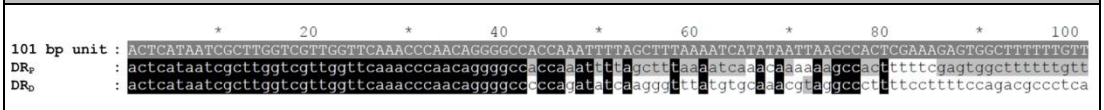
A) All KpPhageMet$_{78578}$ *att* sites possessed a central 101 bp unit that comprised 46 bp of the 3' end of *met56* (grey shading) and 55 bp of adjacent downstream sequence. The underlined bases have been used to represent this unit in diagrams B and C. **B)** The *attP* and *attB* sites consisted of this 101 bp unit flanked by an unknown length of phage DA/UA or *met56* UF/DF sequence, respectively. **C)** When site-specific recombination occurred between *attP* and *attB*, KpPhageMet$_{78578}$ integrated and flanking *attL* and *attR* sites were formed. The *attP* and *attB* sites also contained identical 101 bp units which formed the direct repeats that flanked the KpPhageMet$_{78578}$ prophage. UA, upstream arm; DA, downstream arm; UF, upstream flank; DF, downstream flank.

4.2.5.2 KpGI-5 cannot be excised by the *met56*-specific integrase from KpPhageMet$_{78578}$

Initial analysis found that KpGI-5 was flanked by 46 bp direct repeats, DR$_P$ and DR$_D$, which were almost identical to the 3' end of *met56* (Figure 4-3). As this showed similarities to the structure of the 101 bp conserved unit described for KpPhageMet$_{78578}$, the sequence downstream of DR$_P$ and DR$_D$ was also investigated for the 55 bp adjoining sequence. Within this region DR$_P$ possessed nine single nucleotide

polymorphisms, whilst DR_D was considerably more degenerate and few nucleotide similarities could be identified (Figure 4-9).

Figure 4-9 :- Comparison of DRs flanking KpGI-5 to the 101 bp conserved unit



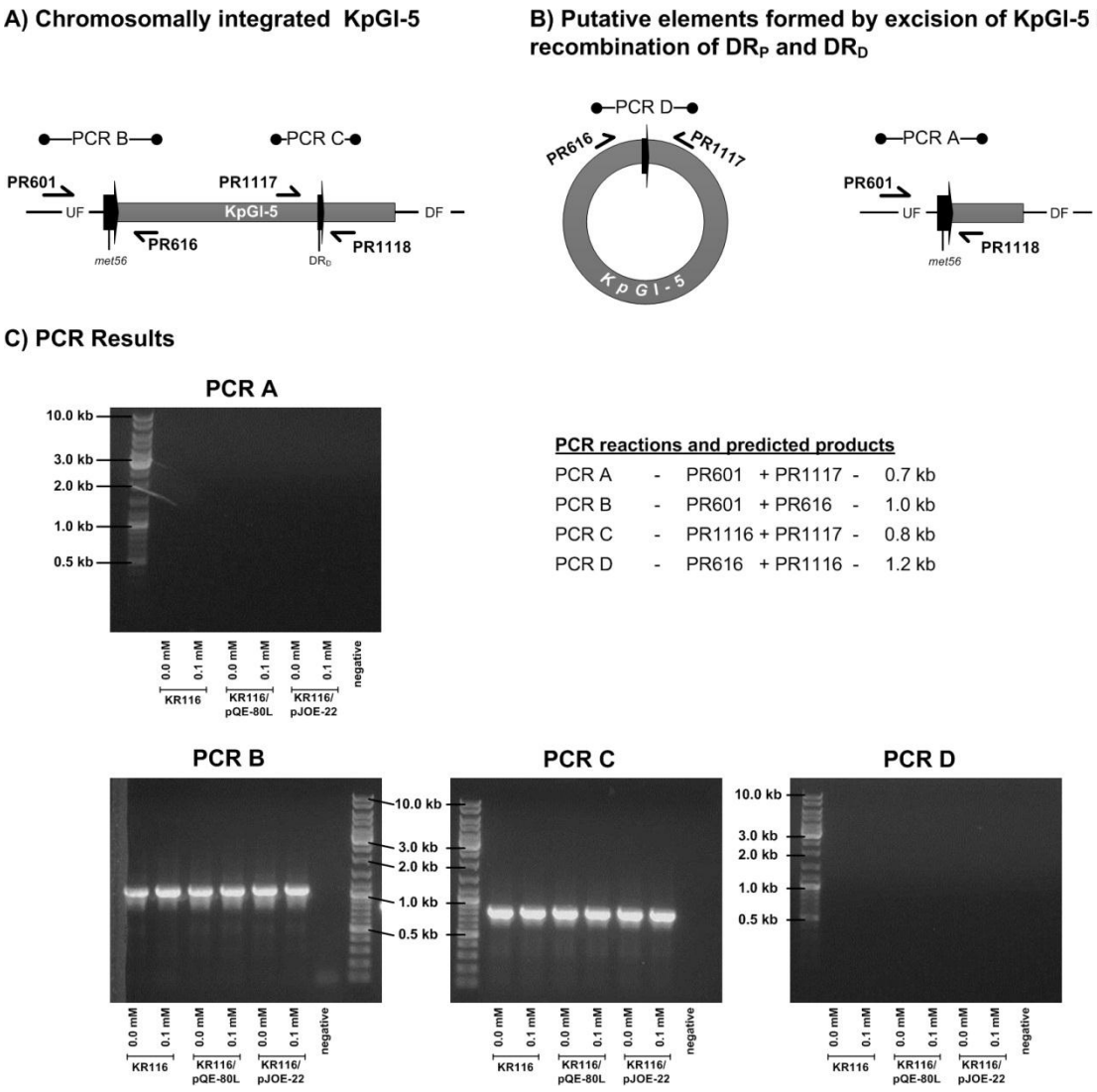
The DR_P and DR_D of KpGI-5 were compared to the 101 bp units that were conserved in the *att* sites of KpPhageMet$_{78578}$. The locations of the repeats in KpGI-5 are illustrated in Figure 4-3. All three possessed near identical 5' regions, which were similar to the 3' end of *met56*. Further downstream, DR_P possessed nine single nucleotide polymorphisms in comparison with the 101 bp unit, whilst DR_D shared only a few nucleotides. DR_P, direct repeat proximal; DR_D, direct repeat distal.

Based on the analysis of the repeat regions within the putative *attL* and *attR* sites of KpGI-5 and the lack of a cognate integrase gene, it was hypothesised that KpGI-5 was not spontaneously excisable. To confirm this, a PCR assay was designed to probe for structures that may be produced by site-specific recombination of the KpGI-5 repeats DR_P and DR_D (Figure 4-10). Neither the hypothetical circularised product formed from the site-specific recombination between DR_P and DR_D nor an unoccupied *met56* tRNA locus could be identified within KR116 gDNA extractions.

In an attempt to induce excision of KpGI-5, the integrase gene from KpPhageMet$_{78578}$ was expressed *in trans* in KR116. Initially, the integrase gene was cloned into expression vector pQE-80L to create pJOE-22. KR116, KR116/pQE-80L and KR116/pJOE-22 were then grown overnight at 37ºC in LB supplemented with 500 µg/ml ampicillin and IPTG, when required. gDNA was extracted from each culture and analysed by PCR. Again neither the hypothetical circularised product nor the unoccupied *met56* locus could be identified (Figure 4-10), even though large quantities of the integrase protein could be detected on SDS-PAGE gels. Alternative PCR reactions that probed for hypothetical structures that could be formed by recombination at locations other than the 46 bp *attL* and *attR* sites, including a 2.7kb KpGI-5 right arm circular element and an empty locus formed by deletion of the entire 14.0 kb KpGI-5

island (see Figure 4-3), were also negative (data not shown). These results suggest that KpGI-5 is not mobile and that it is permanently locked within the chromosome of KR116 under the conditions tested.

Figure 4-10 :- PCR assay to probe the mobility of KpGI-5



A) The chromosomally-integrated form of KpGI-5 can be detected using PCR A and/or PCR B. **B)** Diagram of the predicted unoccupied *met56* locus and hypothetical circular form of KpGI-5 that is formed by site-specific recombination between DR_P and DR_D, which constitute part of the KpGI-5 *attL* and *attR* sites, respectively. These can be identified using PCR A and D, respectively. Note, the unoccupied met56 locus still bears 2.7 kb of KpGI-5 sequence. **C)** Agarose gels showing results for PCRs A, B, C and D performed on gDNA extractions from KR116, KR116/pQE-80L and KR116/pJOE-22 grown for 24 h with 0.0 mM or 0.1 mM IPTG. UA, upstream arm; DA, downstream arm; UF, upstream flank; DF, downstream flank.

4.3 Discussion

4.3.1 Fosmid cloning, marker rescue and sequencing of KpGI-5

In this work a targeted approach using a fosmid library and marker rescue was implemented to obtain fosmids that harboured the entire *K. pneumoniae* KR116 KpGI-5 island for subsequent sequencing. Several studies have described a combinatorial approach for isolating regions of interest using fosmid libraries and Southern hybridization gene probing (182, 387). To the best of my knowledge, the antibiotic resistance cassette-based marker rescue of fosmids technique used in this work has not been described before. However, a similar method has been used to recover plasmids containing the entire *she* PAI sequence from *Shigella flexneri* 2a YSH6000T (7). This technique was particularly valuable as it allowed the rapid isolation of clones that possessed a region of interest without the need for Southern hybridization. Although locus tagging was necessary, the tagged strain KR116Δ*fim2K*::kan was subsequently also used to investigate the function of *fim2K* (Chapter 6).

It is important to note that next generation sequencing methods such as Roche 454 pyrosequencing, Illumina Solexa and ABI SOLiD can now be used to accurately sequence whole bacterial genomes rapidly at low cost, and have almost rendered targeted sequence approaches such as that described in this work redundant (259). However, these technologies tend to produce relatively short read lengths that, on average, range from 50 to 330 bp in size (247, 259). Consequently *de novo* assembly of genomes is often troublesome, although considerable improvements in both assembly algorithms and read lengths have recently been made (288). Correspondingly, novel GIs, prophages and mobile genetic elements may prove problematic to assemble, as was highlighted in this chapter by the presence of three contigs after sequence assembly and in a study by Kingsford *et al.* (198).

In summary, the targeted approach involving fosmid libraries, marker rescue and 454 sequencing was successful at obtaining the sequence of KpGI-5. Nevertheless, this approach has now almost completely been superseded by next generation sequencing of whole bacterial genomes, which is significantly faster and does not rely on the presence of chromosomal markers for marker rescue.

4.3.2 Sequence analysis of KpGI-5

KpGI-5 is a novel GI inserted within the *met56* gene of *K. pneumoniae* KR116, a strain isolated from the blood of a patient with pneumonia and neutropenic septicaemia. KpGI-5 was sequenced in this work and found to encode a putative γ_1-type CU fimbrial operon that has been named *fim2*. Although no integrase gene was identified, KpGI-5 was found to have features characteristic of GIs: lower G+C content than the genome average, located within an island integration hotspot (*met56*) and flanking direct repeats. The *fim2* operon resembles the genetic organization of the *K. pneumoniae fim* operon and contains homologs of all eight genes. Amino acid sequences of the eight *fim2* gene products showed 60 to 92 % identity to cognate *fim* proteins. Indeed, the two clusters would appear to be pseudoparalogs, homologs that appear to be paralogous but have ended up in the same genome by both vertical and horizontal gene transfer (209).

The unique evolutionary origins of the *fim* and *fim2* cluster are further highlighted by differences in transcriptional control. The *fim* cluster is largely controlled by the FimB and FimE recombinases which together switch transcription on and off by inverting a 314 bp promoter-containing sequence called *fimS* that lies upstream of *fimA* (387). No homologs of FimB, FimE or *fimS* were found within the KpGI-5 island, suggesting that *fim2* expression is not controlled via a *fimS* switch-like mechanism. Additionally, the *fim2K* gene of the *fim2* cluster encodes an EAL domain containing protein that is similar to the product of the *fimK* gene, which has previously been shown to regulate

fimbrial expression and is hypothesised to hydrolyse the intracellular messenger c-di-GMP (331). This messenger is known to regulate expression of virulence genes, motility, biofilm formation and cell cycle progression in certain other bacteria (160). Interestingly, Fim2K is lacking an N-terminal helix-turn-helix domain and its mechanism of action could be considerably different to that of FimK. An investigation into the function of *fim2K* will be presented in Chapter 6.

4.3.3 Epidemiological profile and heterogeneity of KpGI-5 family members within the *Klebsiella* genus.

A catalogue consisting of 162 *Klebsiella* isolates from both clinical and environmental sources were screened for the *fim2K* gene. The KpGI-5 family was defined as a group of related islands that harbour a *fim2* operon and *fim2K* gene. The majority of *Klebsiella* strains investigated were *K. pneumoniae* and *K. oxytoca*, however this is representative of their importance in *Klebsiella* infections (135, 154, 302, 305). In total, 21 isolates comprising *K. pneumoniae*, *K. oxytoca* and undefined *Klebsiella* spp. that had been obtained from five geographically distinct areas were identified as *fim2K*-positive. The global spread of KpGI-5-like islands strongly suggests that this locus is under positive selection and has been maintained within a subset of the *Klebsiella* population. Adhesion and colonization are often mediated by fimbriae and are essential processes in both infection and environmental persistence (409). It is therefore plausible that the putative *fim2*-encoded Fim2 fimbriae are under positive selection within habitats colonized by *Klebsiella*.

To investigate whether members of the KpGI-5 island family were associated with a particular niche, the isolation source of each *fim2K*-positive strain was examined. KpGI-5-like islands were found in *Klebsiella* strains from practically all sources, including sputum, blood and urine. Most interestingly the environmental isolate cas128 possessed a KpGI-5 island identical to that in KR116, suggesting a possible role in

environmental colonization as well. Indeed, as environmental *Klebsiella* isolates often possess similar virulence factors and are as pathogenic *in vivo* as their clinical counterparts, a role in both niche types cannot disregarded (124, 303, 390). Functional characterisation of the *fim2* operon will be discussed in more detail in Chapter 5.

The majority of *fim2*-positive strains were also found to possess the *fimH* (76.0%) and *mrkD* adhesin genes (85.7%). These strains were not examined for the presence of functional surface-exposed type 1 or type 3 fimbriae, and thus the structural integrity and functionality of these genes is unknown. Nevertheless, a previous investigation of 269 clinical and environmental *K. pneumoniae* isolates found that 87 % and 73 % of strains had agglutination phenotypes associated with type 1 and type 3 fimbriae, respectively (303). Since these numbers are approximately similar to the prevalence of *fimH* and *mrkD* within *fim2K*-positive *Klebsiella* isolates, it is reasonable to assume that the majority of detected *fimH* and *mrkD* genes are associated with functional surface expressed fimbriae.

Most intriguingly, identical faint bands were obtained during *fimH* PCR screening of sp25 and sp28. The DNA sequence of the *fimH* gene was novel, although the predicted product was almost identical to that of the related *K. pneumoniae* subsp. *rhinoscleromatis* ATCC 13884 strain. Most interestingly, the single amino acid change identified in this FimH protein was previously reported as a pathoadaptive mutation the *E. coli* FimH protein, although this was not experimentally verified (378). It was difficult to determine the species of these strains as they had been biochemically classified as *K. pneumoniae* (154), consistent with the negative indole test result on repeated testing, but paradoxically only the *K. oxytoca* species-specific PCR was positive. Further work is necessary to identify whether the sp25 and sp28 *fimH* is part of a third paralogous *fim*-like operon, or whether the associated operon is an ortholog of the *K. pneumoniae* type 1 fimbrial operon that is present in a novel and/or as-yet unsequenced *Klebsiella* species.

An investigation into the structural heterogeneity within the KpGI-5 family revealed that at least 14 different members could be identified. The *fim2* operon was the most conserved, whilst the island right arm was identical in only 45 % of strains. This is consistent with the hypothesis that the *fim2* operon is under positive selection. However, this should be interpreted with caution as alternative non-*fim2* harbouring *met56*-assocaited islands that possess other KpGI-5 elements were not sought during this epidemiological profiling.

Strain KR392 was found to possess an IS*630* element within its *fim2D* gene. A further five strains were also shown to harbour a similarly sized insertion within the same region. IS*630* is a 1.15 kb transposase-encoding element with terminal 32 bp inverted repeats and has previously been described in other Enterobacteria including *E. coli*, *S. enterica*, *Erwinia* spp. and *Shigella* spp. (364, 405, 406). The effect of this insertion on *fim2* expression is unknown, but based on bioinformatics analysis IS*630* encodes an in-frame stop codon and thus a truncated Fim2D usher protein is likely to be produced. Although some IS*630* relatives that possess strong promoters or transcriptional terminators have been described, these could not be identified within the IS*630* element identified (data not shown) (364). Due to the importance of usher proteins in fimbrial assembly, it is likely that this insertion prevents assembly of the putative *fim2* fimbriae (200, 266). Alternatively, since the subunit regions involved in donor-strand exchange in type 1 fimbrial assembly are almost identical to similar regions in Fim2 subunits, the Fim2 structure may assemble via the FimC/FimD chaperone/usher platform (data not shown) (418). Such a hybrid fimbria assembly mechanism has previously been reported with the chaperone and usher subunits of *E. coli* F1C and type 1 fimbriae, which exhibit only 34 – 60 % amino acid identity (202). This is much lower than the homology between the *K. pneumoniae* type 1 and putative Fim2 assembly components (82 – 92 %), thus increasing plausibility of a hybrid Fim/Fim2 assembly mechanism. Most interestingly, *fim2D* was disrupted in all five

urinary tract infection isolates; two were isolated in Germany whilst the remaining three were isolated in Leicester. It is possible that these *K. pneumoniae* strains are clonal and stem from a common ancestor. Alternatively, they could be independent insertions of IS*630* within *fim2D*, or even a combination of the two. On the assumption that *fim2D* disruption leads to a phenotypic loss, this suggests that the product of *fim2* is redundant or under negative selection within the urinary tract. The disrupting presence of an IS element is not unique to the *fim2* operon; indeed similar findings have been described for a multitude of other fimbrial operons including two out twelve CU fimbrial operons in *E. coli* K12 MG1655 (210, 278).

The results presented in this chapter allow speculation regarding the origin of KpGI-5. It can be hypothesized that the relatively small, immobile KpGI-5 island in *K. pneumoniae* stems from a larger KpGI-5-like island ancestor. By means of reductive evolution over time, the larger island may have gradually lost various functional modules, such as those mediating excision and conjugative transfer (95). Such an evolutionary process has previously been hypothesized for the ICE*St1* family elements from *Streptococcus thermophilus* (290) and ExoU island family in *P. aeruginosa* (215). Further examination and sequencing of the KpGI-5-like islands identified will help explain the evolution of this island family.

4.3.4 Investigation into the mobility of *met56*-associated islands

In this work the prophage KpPhageMet$_{78578}$ was demonstrated to be a spontaneously excising mobile element present within the *met56* tRNA locus of *K. pneumoniae* MGH78578. This system was used as a model to investigate the basis of site-specific recombination of *met56*-associated mobile elements, including KpPhageMet$_{78578}$ and KpGI-5. Site-specific recombination of KpPhageMet$_{78578}$ is analogous to the integration and excision mechanisms identified in many mobile elements such as the P2 prophage, the SXT ICEs of *Vibrio cholerae,* the ICE*Kp1* of *K. pneumoniae* NTUH-

K2044 and the *she* PAI of *S. flexneri* (26, 165, 234, 344). In all the above cases an integrase gene mediates the recombination of direct repeat sequences present within flanking *attL* and *attR* sites. This subsequently causes excision of the element from the chromosome and the formation of a circular extrachromosomal element. With respect to KpPhageMet$_{78578}$ the circular form may represent an intermediate in the phage packaging process, similar to the packaging of phages P2 and P4 (34). These phage particles can then transduce KpPhageMet$_{78578}$ DNA to new host strains. Alternatively, circular intermediates of GIs and ICEs are thought to be necessary for strain-to-strain transfer (95, 128). Consistent with this hypothesis, ICE*Kp1* requires a functional integrase gene for both excision and conjugative transfer (234). A similar system has been postulated for the SXT ICEs in *V. cholerae* (165) and PAPI-1 ICE of *P. aeruginosa* (56).

Examination of the *attL* and *attR* sites of KpGI-5 revealed that the island possessed almost identical 46 bp flanking DRs lying 11.3 kb apart. Based upon similarities with KpPhageMet$_{78578}$ and bacteriophage 186, it was hypothesized these 46 bp comprised the minimum repeat unit required for the integrase-mediated excision of KpGI-5. Attempts were made to excise KpGI-5 from the *met56* locus of KR116 by providing the KpPhageMet$_{78578}$ integrase *in trans*, but neither a circular element or reassembled *attC* site could be detected even though large quantities of integrase protein were produced. It is unknown whether the KpPhageMet$_{78578}$ integrase is sufficient for the integration and/or excision of KpPhageMet$_{78578}$. This integrase is a HP1-like tyrosine recombinase that is 96 % identical to the integrase of bacteriophage 186, which integrates site-specifically within the *ile* tRNA gene (144, 362). Since excision of the archetypal bacteriophage HP1 requires an excisionase, it is likely this is also required for KpPhageMet$_{78578}$ (108).

The inability of the KpPhageMet$_{78578}$ integrase to mediate site-specific recombination between the KpGI-5 direct repeats could be explained in two ways. Firstly, it is possible

that the KpPhageMet$_{78578}$ integrase does not mediate excision on its own. Possible candidate accessory recombination directionality factors in KpPhageMet$_{78578}$ include a gene downstream of the integrase gene encoding a hypothetical protein and the gene KPN_03477, which is identical to *apl* from bacteriophage 186. *apl* mutant prophages are defective in excision, a process mediated by direct excisionase action of Apl at phage *att* sites (322). Additional excision-mediating factors are not unique to bacteriophages and have previously been described for the *she* and HPI PAI which required the *rox* and *xis* genes for excision, respectively (344, 353). Alternatively, the KpPhageMet$_{78578}$ integrase may require the full 101 bp conserved unit identified within *attP* to be present within the *attL* and *attR* sites flanking KpGI-5. If this is the case, then the flanking DRs were too degenerate to mediate site-specific recombination by the KpPhageMet$_{78578}$ integrase. Both possibilities are under investigation in Dr. Kumar Rajakumar's laboratory.

Despite a massive surge in the number of whole genome sequences available, KpGI-5-like islands have not been identified previously even though many closely related Enterobacteriaceae have been sequenced, including over 100 *E. coli*, 40 *Salmonella* spp., 35 *Yersinia* spp. and nine *Klebsiella* spp. strains. This work found 21 occurrences of KpGI-5-like islands within 162 *Klebsiella* strains. Four strains possessed putatively mobile islands including one *K. pneumoniae*, one *K. oxytoca* and two undefined *Klebsiella* spp. isolates. Although no *in vitro* screening for *fim2K* was performed in other Enterobacteriaceae, the prevalence of *fim2K* and KpGI-5-like islands is likely to be limited to the *Klebsiella* genus and its close relatives due to the site-specific nature of the putative KpGI-5 integrase. Exact homologs of the 46 bp 3' end of *met56* can be found within strains belonging to the following genera: *Klebsiella*, *Cronobacter*, *Edwardsiella*, *Enterobacter*, *Escherichia* and *Dickeya*. However, only two out of 100 plus sequenced *E. coli* strains possess this 46 bp unit: strains UMNF18 and O103:H2 str. 12009. If the full 101 bp unit is taken as the minimum unit required for

island integration then the number of available insertion sites is further restricted to only the *Klebsiella* and *Enterobacter* genera. This unit also imposes intra-genus restrictions as neither *K. pneumoniae* Kp342 nor *K. variicola* At-22 possess this larger unit. Exact determination of the minimum conserved unit required for recombination of KpGI-5-like islands will aid further investigations into understanding the host range and dissemination of this novel island family.

4.3.5 Future work

Although *fim2* and *fim2K* were discussed in this chapter, their function was examined in greater detail and these results will be presented and discussed in Chapter 5 and Chapter 6, respectively. Potential aspects for future work include:

Understanding the* Klebsiella *sp. sp25 and sp28 faint fimH PCR results

1/. Amplify and sequence the 16S rRNA, *gyrA*, *parC* and *rpoB* genes from sp25 and sp28 and use data to molecularly identify the *Klebsiella* species and phylogenetic group according to the methods of Brisse *et al.* (39) and/or Drancourt *et al.* (98).

2/. Amplify and sequence the amplicon produced by long-range PCR from the sp25 and sp28 *fimS* switch to the *fimH* gene. This will allow classification of the putative operon as orthologous or paralogous to the *K. pneumoniae* type 1 fimbrial operon.

Sequencing and analysis of KpGI-5-like islands

1/. Sequence the remaining KpGI-5-like islands in their entirety, either using a shotgun whole genome sequencing approach or the targeted fosmid sequencing approach described in this chapter. Alternatively, attempts can be made to amplify these islands by long-range PCR (20 to 40 kb) using primers that bind to the UF/DF; these amplicons can subsequently be sequenced.

2/. The island sequences obtained can then be examined to:

a) Understand the evolutionary relationship between KpGI-5-like islands.

b) Investigate whether, similar to *fimH*, pathoadaptive mutations also occur in *fim2H*.

c) Determine whether mobile islands possess genes putatively involved in island mobility and/or conjugative transfer. Roles of such genes can subsequently be characterised in follow-up experiments.

Mobility analysis of KpGI-5 and KpGI-5 like islands

1/. Investigate whether the putative excisionase Apl from KpPhageMet$_{78578}$ may, either alone or in conjunction with the KpPhageMet$_{78578}$ integrase, cause excision of KpGI-5.

2/. The integrase gene and *apl* excisionase gene from KpPhageMet$_{78578}$ can be disrupted using allelic exchange and the effect on KpPhageMet$_{78578}$ can be examined.

3/. Replace the KpGI-5 *attL* and *attR* sites with corresponding sites from KpPhageMet$_{78578}$ using allelic exchange and examine the effect on mobility induced by the KpPhageMet$_{78578}$ integrase. This can also be investigated using a plasmid-based system where different combinations of *attL* and *attR* sites separated by an antibiotic resistance cassette can be constructed using SOE-PCR. Recombination of these sites can then be examined by PCR both in the presence and absence of the KpPhageMet$_{78578}$ integrase and/or putative Apl excisionase.

4/. Examine the *attL* and *attR* sites that flank the four mobile and 16 KpGI-5-like 'locked-in' islands to elucidate the minimum conserved unit required for island mobility.

5/. If one of the four mobile KpGI-5-like islands possesses putative conjugative machinery, experiments can be performed to transfer the GI from its host strain to a recipient *K. pneumoniae* strain with an 'empty' *met56* site. Transfer to another closely-related genus can also be examined.

Chapter 5. *In vitro* and *in vivo* characterisation of *fim2*, a novel operon present on KpGI-5 coding for a predicted γ_1-type fimbrial system

5.1 Introduction

In previous chapters the *K. pneumoniae* KR116 *fim2* operon was introduced as a novel γ_1-type fimbrial operon present on the KpGI-5 genomic island. Adhesion and colonization are essential steps in the infection process and are often mediated by fimbriae (300). To investigate whether this holds true for *fim2*, the *in vivo* and *in vitro* role of *fim2* was investigated. These experiments are the focus of this chapter.

Several genetic constructs were created to examine the function of *fim2*. Firstly, a *fim2* knockout was constructed of KR2107, a streptomycin-resistant derivative of KR116 (discussed in section 2.6.4). KR2107 *fim* and double *fim fim2* knockouts were also made. These strains have been used in this chapter to investigate the effect of *fim2* in the presence of and in isolation of *fim*, which can affect the function of other fimbriae (211). The latter two strains will also enable the identification of potential redundant roles between the two fimbriae. In parallel, *fim2* operons were cloned into plasmids and heterologously expressed in *E. coli* HB101, a strain that has been used extensively to investigate the *E. coli* type 1 and *K. pneumoniae* type 1 and type 3 fimbriae (201, 387, 388). The afimbriate nature of this strain enables the investigation of specific-fimbrial structures in isolation of other host-encoded fimbriae and has enabled visualisation of fimbrial structure in addition to the identification of associated phenotypic characteristics. Additionally, heterologous expression avoids potential physical hindrance on fimbrial function by the *K. pneumoniae* capsule, as observed previously for type 1 fimbriae (349, 391).

In this chapter, KR2107 and its three isogenic mutants (Δ*fim*, Δ*fim2*, and Δ*fim*Δ*fim2)* will be analysed using agglutination assays, biofilm assays, human bladder cell and intestinal cell adhesion assays, and *in vivo* competition-based murine models of intestinal colonization, lung infection and ascending urinary tract infection. In parallel, HB101 cells heterologously expressing the *fim2* operon will be analysed using agglutination assays, biofilm assays and transmission electron microscopy. Lastly, *K. pneumoniae* KR2107 cDNA libraries generated from *in vitro* and *in vivo* samples will be examined using normal PCR and qRT-PCR for expression of the *fim2* operon, which will be compared to that of the type 1 and type 3 fimbrial operons.

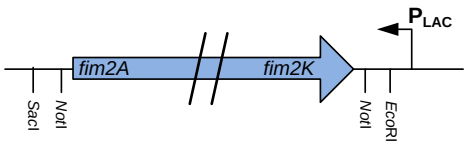
5.2 Results

5.2.1 Expression plasmids for the investigation of *fim2* function

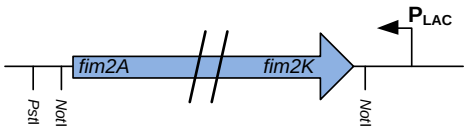
Several plasmids were constructed to enable the heterologous expression of *fim2* (Figure 5-1). Briefly, the *fim2* locus was PCR-amplified from KR116 and cloned into the high copy number vector pBluescript II KS+, the low copy number vector pWSK129 and the IPTG-inducible vector pJTOOL-7 to create pFim2-HCN, pFim2-LCN and pFim2-Ptrc, respectively. In the former two plasmids expression is controlled via the native *fim2* promoter, but greater than native expression levels are expected as the multi-copy nature of the plasmids will increase *fim2* copy number. In the latter plasmid *fim2* was cloned downstream of the IPTG-inducible P_{TRC} promoter (10). Similarly, the *fim2* loci from sp25 and KR518 were also PCR-amplified and cloned into pJTOOL-7 to create plasmids pFim2$_{sp25}$-Ptrc and pFim2$_{KR518}$-Ptrc, respectively. The latter *fim2* operon possessed an IS*630* element within *fim2D*, as discussed in section 4.2.4. When necessary, pFim-Ptrc and pCAS624 were used as control plasmids. These harboured the *K. pneumoniae* C3091 *fim* operon in vectors pJTOOL-7 and pUC18, respectively.

Figure 5-1 :- Expression plasmids constructed to investigate Fim2 function

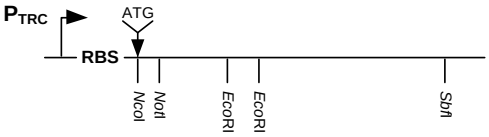
pFim2-HCN
- KR116 *fim2* operon cloned into pBluescript II KS+
  - high copy number vector
- expression controlled via native promoter



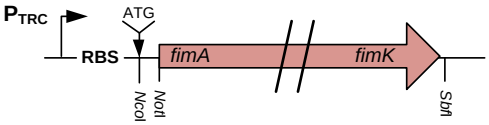
pFim2-LCN
- KR116 *fim2* operon cloned into pWSK129
  - low copy number vector
- expression controlled via native promoter



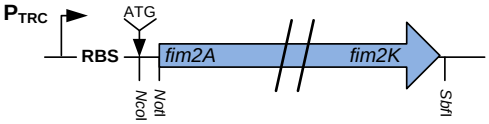
pJTOOL-7
- pTRC99a derivative with NotI cloning site
- P_{TRC} promoter



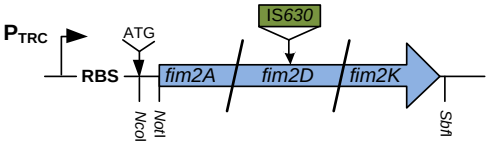
pFim-Ptrc
- C3091 *fim* operon cloned into pJTOOL-7
  - native *fim* promoter removed
  - *fimA* is present as an in-frame translational fusion
- expression controlled via P_{TRC}



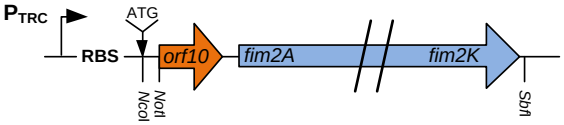
pFim2-Ptrc and pFim2$_{sp25}$-Ptrc
- KR116 or Sp25 *fim2* operon cloned into pJTOOL-7
  - native *fim2* promoter removed
  - *fim2A* is present as an in-frame translational fusion
- expression controlled via P_{TRC}



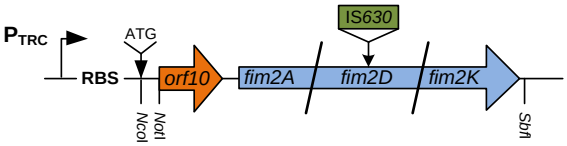
pFim2$_{KR518}$-Ptrc
- KR518 *fim2* operon cloned into pJTOOL-7
  - native *fim2* promoter removed
  - *fim2A* is present as an in-frame translational fusion
- expression controlled via P_{TRC}



pFim2(orf10-2K)-Ptrc and pFim2(orf10-2K)$_{sp25}$-Ptrc
- KR116 or Sp25 *orf10* gene and *fim2* operon cloned into pJTOOL-7
  - native *orf10* promoter removed; *fim2* promoter likely to be intact
  - *orf10* is present as an in-frame translational fusion
- expression controlled via P_{TRC} and native *fim2* promoter



pFim2(orf10-2K)$_{KR518}$-Ptrc
- KR518 *orf10* gene and *fim2* operon cloned into pJTOOL-7
  - native *orf10* promoter removed; *fim2* promoter likely to be intact
  - *orf10* is present as an in-frame translational fusion
- expression controlled via P_{TRC} and native *fim2* promoter



The schematic depicts the various expression plasmids constructed to investigate the function of the *fim2* operon. Detailed construction information can be found in Appendix 6. RBS, ribosomal binding site ; ATG, start codon ; P_{TRC}, *trc* promoter ; P_{LAC}, *lac* promoter.

5.2.2 Expression of *fim2* does not cause guinea pig red blood cell or yeast agglutination

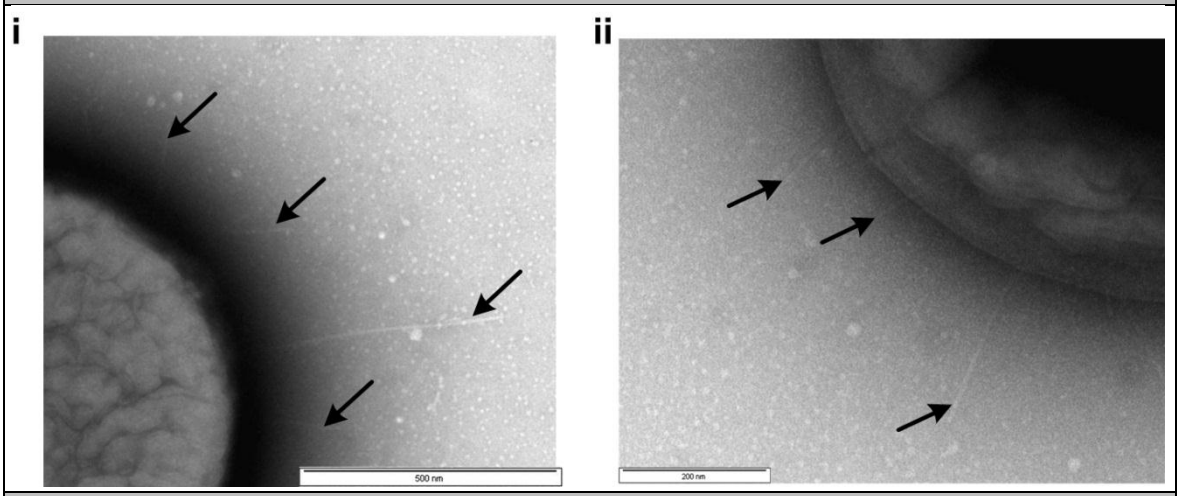
Plasmids pFim2-HCN, pFim2-LCN and pFim2-Ptrc were transformed into the afimbriate *E. coli* strain HB101 to examine whether they conferred the ability to agglutinate GP-RBCs or yeast cells, as described for the *fim* operon (387). Unlike the *fim*-bearing control plasmid, none of the three *fim2* plasmids conferred a positive agglutination phenotype. Plasmids pFim2$_{sp25}$-Ptrc and pFim2$_{KR518}$-Ptrc were also unable to mediate agglutination. Consistent with these findings, the *fim*-positive *fim2* knockout strain KR2107Δ*fim2* agglutinated GP-RBCs whilst the *fim2*-positive *fim* knockout strain KR2107Δ*fim* did not.

5.2.3 Heterologous expression of *fim2* does not result in visualisable host fimbriation

HB101/pFim2-HCN, HB101/pFim2-LCN and HB101/pFim2-Ptrc were further examined by transmission electron microscopy in an attempt to visualise the putative Fim2 fimbriae. Despite use of multiple induction methods and over 100 cells being viewed per strain, no definite fimbrial structures could be identified on the bacterial surfaces examined. By contrast, HB101/pFim-Ptrc expressed abundant and highly characteristic type 1 fimbriae on its outer surface. When the sp25 *fim2* operon was expressed in HB101/pFim2$_{sp25}$-Ptrc, two cells from independent cultures were detected with up to ten fimbria-like structures projecting from their surface (Figure 5-2). However, since HB101 has been shown to sporadically display type 1 fimbriae, this finding may reflect rare native HB101 fimbriae rather than the putative Fim2 fimbriae (29, 107). Nevertheless, no equivalent cells were seen with the HB101/pJTOOL-7 control only, despite scanning more than 100 control cells. In an attempt to verify whether Fim2 subunits were produced when the *fim2* operon was overexpressed, cell lysates of HB101/pFim2-HCN, HB101/pFim2-LCN, HB101/pFim2-Ptrc and HB101/pFim2$_{sp25}$-Ptrc were analysed by SDS-PAGE. No additional bands could be identified.

Negative staining of samples was performed using uranyl acetate. This stain has a low pH (4.2 to 4.5) and could denature the putative Fim2 fimbrial structure (personal communication, Stefan Hyman). To test this hypothesis, the relatively neutral pH negative stains ammonium molybdate (pH 7.0) and sodium phosphotungstic acid (pH 7.5) were used. Again, no fimbriae could be identified on the surface of HB101/pFim2-Ptrc whilst abundant type 1 fimbriae were present on the surface of the control strain HB101/pFim-Ptrc when these alternative stains were used (data not shown).

Figure 5-2 :- Fimbria-like structures present on *E. coli* HB101/pFim2$_{sp25}$-Ptrc



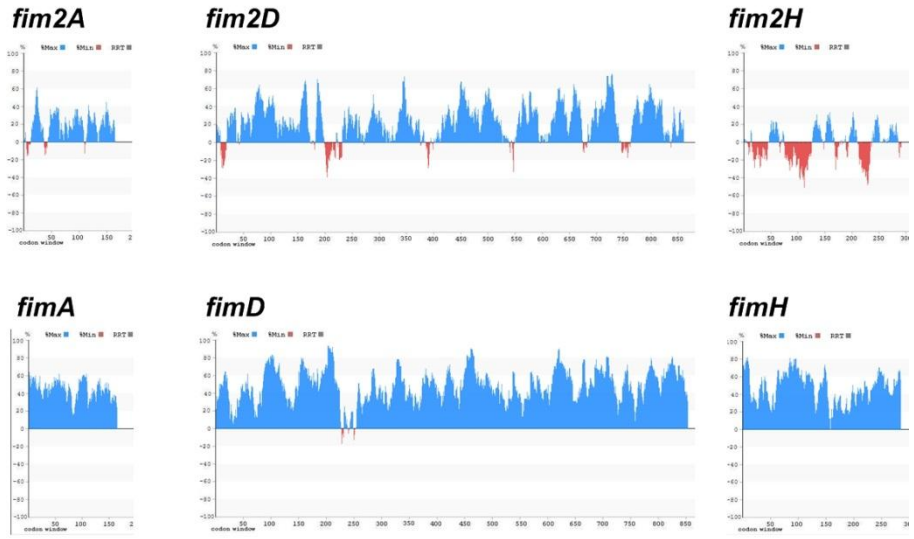
Two cells, **i** and **ii**, from independent cultures of *E. coli* HB101/pFim2$_{sp25}$-Ptrc possessed up to ten fimbria-like structures projecting from their surface.

5.2.4 IPTG induction of pFim2-Ptrc greatly reduces HB101 growth

During the course of the experiments described in the previous two sections it was observed that 16 h cultures of HB101/pFim2-Ptrc supplemented with 0.05 mM IPTG grew to considerably lower culture densities than when grown without IPTG supplementation. In fact, the OD_{600} value obtained after 18 h of incubation was reduced by 50 to 75% compared to the uninduced cultures. This effect was more pronounced with higher concentrations of IPTG and was not observed for control strains HB101, HB101/pJTOOL-7 and HB101/pFim-Ptrc. Of note, cultures containing HB101/pFim2-

HCN and HB101/pFim2-LCN grew normally and plasmid extraction yields were as expected, suggesting that plasmid copy number was maintained. A growth defect was also identified with IPTG induction of plasmids pFim2$_{sp25}$-Ptrc, pFim2$_{KR518}$-Ptrc, pFim2(orf10-2K)-Ptrc and pFim2(orf10-2K)$_{KR518}$-Ptrc, of which the latter two possessed the *fim2* operon and *orf10*. Most intriguingly, pFim2(orf10-2K)$_{sp25}$-Ptrc did not affect HB101 growth; the reason for this remains unexplained but is probably related to PCR amplification errors of the cloned product.

Figure 5-3 :- Rare codon clusters within *K. pneumoniae fim* and *fim2* genes

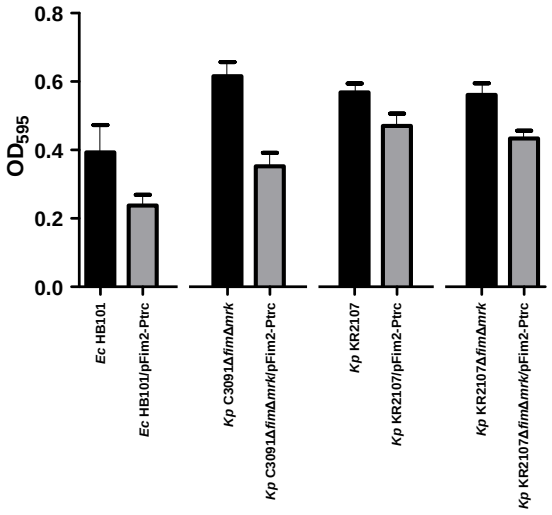


All genes comprising the *K. pneumoniae fim* and *fim2* operons were scanned for rare codon clusters using the Rare Codon Calculator and an *E. coli* K-12 codon usage table (74). Results are plotted as %MinMax values that have been averaged over an 18 codon sliding window. Clusters consisting mainly of common codons appear as blue positive peaks whilst those containing mainly rare codons appear as red negative peaks. The results for *fim2A*, *fim2D* and *fim2H* and their *fim* relatives are illustrated above. Whilst almost no rare codon clusters were identified in the *fim* genes, multiple rare codon clusters were located within *fim2* genes. These clusters can impede ribosomal translation of the sequence (74).

IPTG induction of pFim2-Ptrc is assumed to produce high levels of *fim2* mRNA. Since *fim2* has a G+C content lower than both the *E. coli* and *K. pneumoniae* genome averages, and as tRNAs have been optimised for translation of transcripts with host-similar G+C content levels, it is plausible that problems may occur when translating the vast quantities of *fim2* mRNA in HB101 (74, 190). To investigate whether translation of

fim2 mRNA could have been hampered, *fim2* genes were examined for rare codon clusters using the Rare Codon Calculator with an *E. coli* K-12 codon usage table (74). All *fim2* genes possessed multiple major rare codon clusters whilst considerably fewer rare codon clusters occurred within the *K. pneumoniae fim* genes (Figure 5-3 and data not shown). The translational efficiency hypothesis was further tested *in vitro* by transforming HB101/pFim2-Ptrc with pRARE, a plasmid that encodes tRNAs that interact with codons rarely used in *E. coli*. Despite increasing the availability of rare tRNAs, IPTG induction of HB101/pFim2-Ptrc still resulted in an equally dramatic reduction in growth. A similar IPTG-dependent negative effect on growth was seen in *K. pneumoniae* when pFim2-Ptrc was induced in the type 1 and type 3 fimbrial knockout strain C3091Δ*fim*Δ*mrk*, the double knockout strain KR2107Δ*fim*Δ*mrk* and the wildtype KR2107 strain (Figure 5-4). Almost identical results were obtained when strains were incubated at 30ºC for 48 h (data not shown).

Figure 5-4 :- Effect of IPTG induction on four host strains containing pFim2-Ptrc

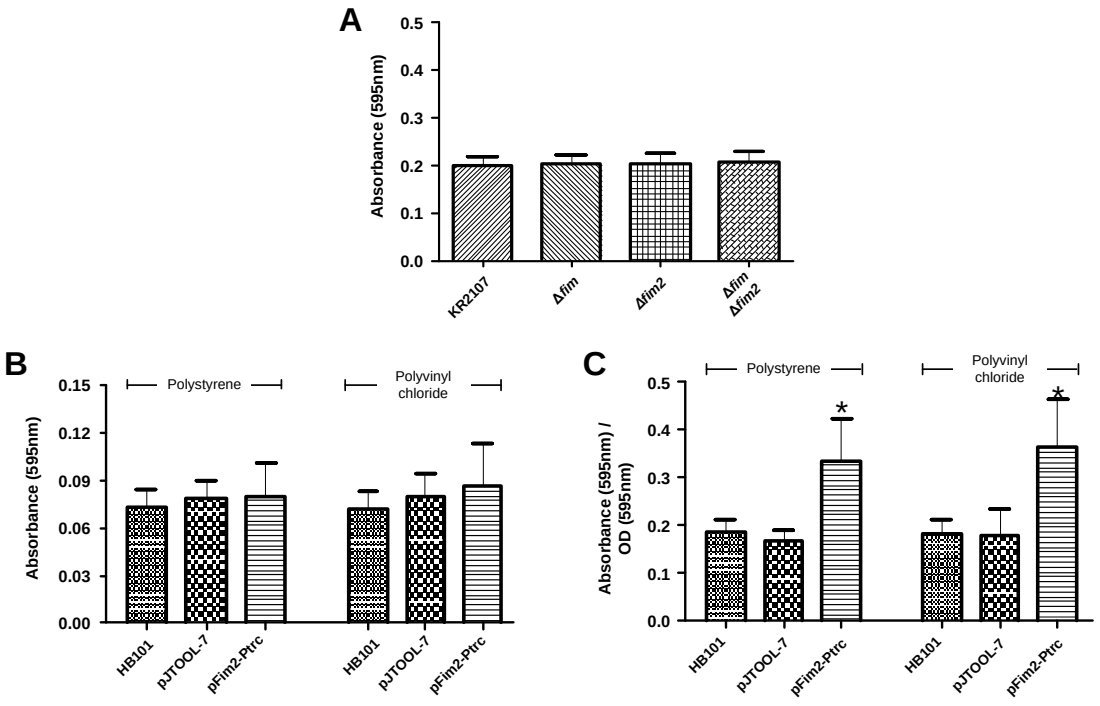


The effect of IPTG induction on four host strains containing pFim2-Ptrc was quantified as follows. Briefly, strains were grown for 16 h (37°C, 200 rpm) in 5 ml LB broth supplemented with 250 µg/ml ampicillin, and then subcultured 1:100 into 96-well polystyrene microtitre plates with each well containing 100 µl of LB media supplemented with 0.05 mM IPTG and 250 µg/ml ampicillin. Each strain was tested in eight wells. 96-well plates were incubated at 37ºC for 48 h and OD_{595} was determined. In all cases the data shown represents the mean and standard deviation obtained from eight replicate wells. *Ec*, *E. coli*; *Kp*, *K. pneumoniae*.

5.2.5 Expression of *fim2* influences biofilm formation

K. pneumoniae readily colonizes and forms biofilms on abiotic surfaces such as urinary catheters and tracheal tubes (252, 390). As surface expressed structures play a key role in biofilm formation, the ability of KR2107 and its isogenic mutants to form biofilms was examined. The absence of *fim2* and/or *fim* had no effect on biofilm formation on polystyrene surfaces as assayed at 24 h under static growth conditions in LB or M9 media at either 37°C or 30°C (Figure 5-5A; data not shown).

Figure 5-5 :- The *fim2* locus from KR116 appears to contribute to biofilm formation when expressed in *E. coli* HB101.



(A) Assay results for biofilm formation on polystyrene for KR2107 and its three *fim* and/or *fim2* isogenic mutants as determined by crystal violet absorbance data. Equivalent results, suggestive of no strain-to-strain differences, were obtained for assays on polyvinyl chloride plates (data not shown). **(B)** Biofilm formation assay based on heterologous expression of *fim2* in *E. coli* HB101/pFim2-Ptrc. HB101 and HB101 carrying an empty pJTOOL-7 served as controls. Biofilm formation was quantified using crystal violet staining and absorbance was measured at 595 nm. Non-normalized crystal violet absorbance data are shown. **(C)** Biofilm formation assay results shown in (B) were normalized to take account of pre-wash total cell numbers based on OD_{595} readings performed at 48 h, just prior to washing off non-surface adherent cells and crystal violet staining. Data shown in all cases represent means and standard deviations of three biological replicates, each assayed in eight wells (n = 24). An asterisk indicates a highly significant difference ($P < 0.0001$) from HB101 and HB101/pJTOOL-7. Statistical analyses were performed using the non-parametric Kruskal-Wallis test and Dunn's posthoc analysis.

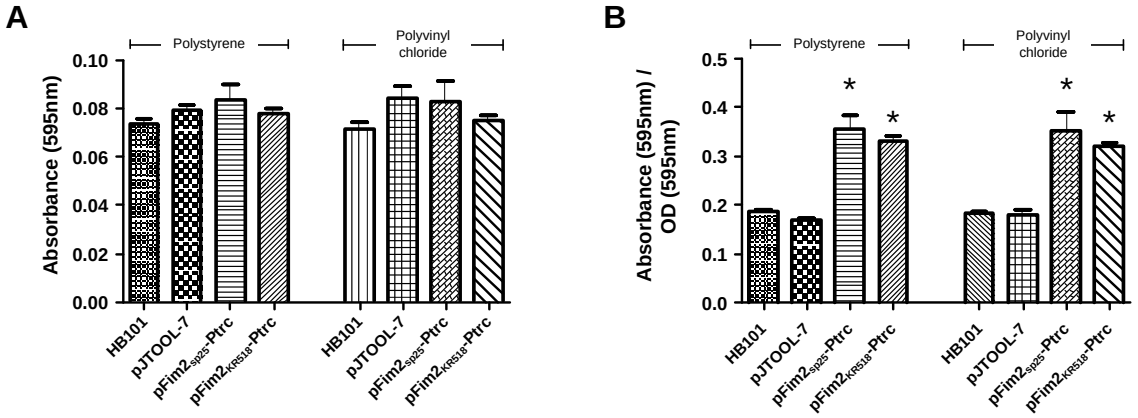
To detect a potential contribution to biofilm formation that may have been masked by low-level *fim2* expression or capsule-related physical hindrance of fimbrial function (349), *fim2* from KR116 was expressed from pFim2-Ptrc using 0.05 mM IPTG in *E. coli* HB101. Compared to HB101 carrying the empty pJTOOL-7 vector, HB101/pFim2-Ptrc exhibited similar biofilm formation at 48 h on polystyrene wells as assessed by post-washing crystal violet staining (Figure 5-5B). On the other hand, expression of *fim2* in HB101 resulted in marginally denser biofilm in polyvinyl chloride wells as compared to the vector-only control, but this was not statistically significant (P = 0.464; Figure 5-5B).

As HB101/pFim2-Ptrc grew to a much lower OD_{595} at 48 h than the other two strains, the biofilm data was also analysed as a ratio of crystal violet staining intensity to the pre-wash OD_{595} measurement that reflected total growth, as described previously (197, 449). This analysis suggested that the proportion of HB101/pFim2-Ptrc cells comprising biofilm growth as opposed to total growth (biofilm and planktonic cells) was almost twice that of HB101 and the vector-only control strain (Figure 5-5C). Indeed, based on this ratio, *fim2* expression in HB101 exerted a highly significant increase in biofilm formation on both surfaces (P < 0.001 in each case). By contrast when *fim2* was expressed in the Mrk- and Fim-deficient *K. pneumoniae* C3091Δ*fim*Δ*mrk* using this same system, no statistically significant accentuation in biofilm formation on either surface was observed (data not shown).

To examine whether the above results could be extrapolated to *fim2* operons from other strains, similar assays were performed by heterologous expression of the *fim2* operons from sp25 and KR518. Again, similar to HB101/pFim2-Ptrc, HB101/pFim2$_{sp25}$-Ptrc and HB101/pFim2$_{KR2415}$-Ptrc grew to a much lower OD_{595} at 48 h than the HB101 and HB101/pJTOOL-7 control strains (data not shown). When these OD_{595} variations were taken into account the expression of *fim2*$_{sp25}$ and *fim2*$_{KR518}$ in HB101 resulted in a highly significant positive impact on biofilm formation (P < 0.0001 in cases). This improved biofilm forming phenotype was identified on polystyrene and polyvinyl

chloride surfaces (Figure 5-6). Most intriguingly, the presence of IS*630* within the *fim2D* gene of *fim2*$_{KR518}$ did not alter the influence of *fim2* on biofilm formation or alleviate the poor growth observed when *fim2* was expressed in HB101 (section 4.2.4).

Figure 5-6 :- The *fim2* loci from sp25 and KR518 contribute to biofilm formation when expressed in *E. coli* HB101.



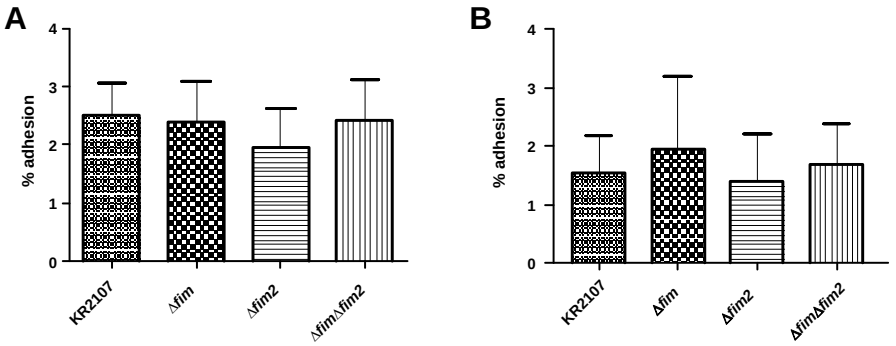
(A) Biofilm formation assay based on heterologous expression of *fim2*$_{sp25}$ *and fim2*$_{KR518}$ in *E. coli* HB101 using plasmids pFim2$_{sp25}$-Ptrc and pFim2$_{KR518}$-Ptrc, respectively. HB101 and HB101 carrying an empty pJTOOL-7 served as controls. Biofilm formation was quantified using crystal violet staining and absorbance was measured at 595 nm. Non-normalized crystal violet absorbance data are shown. **(B)** The biofilm assay results shown in (A) were OD_{595} normalised and are displayed. Data shown in all cases represent means and standard deviations of three biological replicates, each assayed in eight wells (n = 24). Data shown in all cases represent means and standard error of the mean of three biological replicates, each assayed in eight wells (n = 24). An asterisk indicates a highly significant difference ($P < 0.0001$) from HB101 and HB101/pJTOOL-7. Statistical analyses were performed using the non-parametric Kruskal-Wallis test and Dunn's posthoc analysis.

5.2.6 Deletion of *fim2* does not affect adhesion to human HCT-8 ileocaecal or 5637 bladder epithelial cells

In vitro adhesion assays were performed to further investigate whether KR2107 and its three isogenic mutants (KR2107Δ*fim*, KR2107Δ*fim2* and KR2107Δ*fim*Δ*fim2*) exhibited equivalent cell adhesion properties. Human HCT-8 ileocecal and human 5637 bladder epithelial cell lines were chosen to investigate adherence to intestine- and bladder-derived cells, respectively. No significant differences were detectable by these *in vitro* tissue culture assays (Figure 5-7). Furthermore, in contrast to the previously reported

impaired urovirulence of a *fim*-negative *K. pneumoniae* strain (387), the KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* mutants examined in this study did not display any defect in adherence to bladder epithelial cells relative to KR2107 or KR2107Δ*fim2*. It is possible that *fim* expression was insignificant under the *in vitro* conditions used or that the *K. pneumoniae* capsule interfered with fimbrial function (349, 391).

Figure 5-7 :- Comparison of the cell adherence properties of *K. pneumoniae* KR2107 and its isogenic mutants lacking *fim* and/or *fim2*.



(A) *In vitro* adhesion assays to human HCT-8 ileocaecal cells. **(B)** *In vitro* adhesion assays to human 5637 bladder epithelial cells. In both cases percentages of bacteria that remained adherent to cell monolayers after 3 h of incubation at 37°C followed by careful washing are shown. All assays were performed in duplicate wells and repeated independently four times. Bars represent means and standard deviations.

5.2.7 Investigation into the *in vivo* role of *fim2* using murine infection models

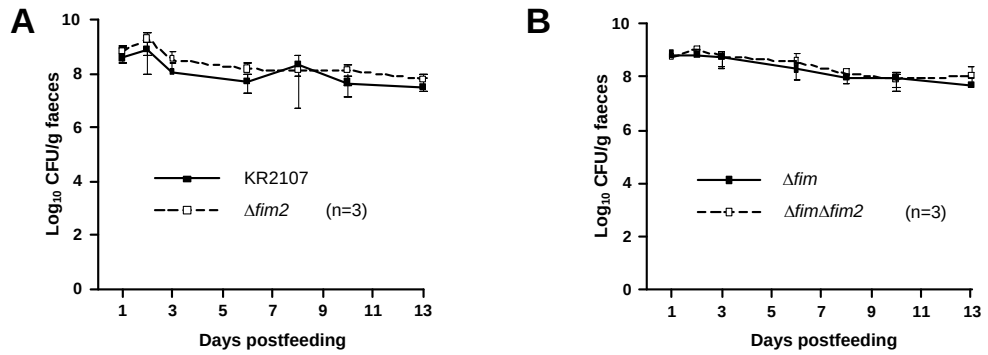
The *in vivo* role of *fim2* was investigated using a series of previously established murine models. KR2107 and its isogenic *fim* and/or *fim2* mutants were used in head-to-head competition assays to investigate whether the presence of *fim2* affected the ability of *K. pneumoniae* KR2107 to colonize the intestine and cause urinary tract and lung infections.

5.2.7.1 Murine intestinal colonization

Epidemiological studies have elucidated that the majority of *K. pneumoniae* infections are preceded by gastrointestinal tract colonization (265). To investigate whether *fim2*

influences this initial step, a 1:1 mixture of KR2107 and KR2107Δ*fim2* was fed to three mice and faecal CFU counts were monitored for 13 days. To exclude potential type 1 fimbriae-related masking, a competition experiment between KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* was also performed. As assessed by faecal CFU counts, no strain exhibited an obvious competitive advantage and all four strains were found to readily colonize the large intestine in similar numbers (~10^8 – 10^9 CFU/g) throughout the experiment (Figure 5-8). Apart from confirming that *fim* does not affect intestinal colonization (387), these results also suggested that *fim2* does not play a significant role in murine intestinal colonization by *K. pneumoniae*.

Figure 5-8 :- Murine intestinal colonization of *K. pneumoniae* KR2107 and its isogenic *fim* and/or *fim2* mutants



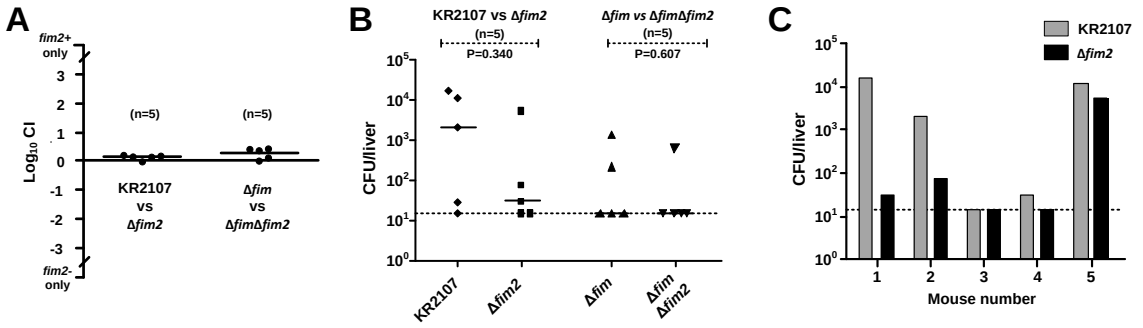
(A) Intestinal co-colonization following oral feeding with a 1:1 mixture of KR2107 and KR2107Δ*fim2* or **(B)** a 1:1 mixture of KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* (approximately 10^9 CFU in total). Three mice were fed for each co-colonization experiment. Mean CFU/g faeces and corresponding standard deviation values are shown.

5.2.7.2 Murine lung infection model

K. pneumoniae is a clinically important cause of lung infections and various potential virulence factors have been determined (225, 227). The influence of *fim2* on pneumovirulence was investigated by intranasal inoculation of five mice with a mixture comprising equal numbers of KR2107 and KR2107Δ*fim2.* A competition experiment between KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* was also performed. 30 h post-infection all mice displayed significant signs of disease and were sacrificed. High numbers of

K. pneumoniae were found in the lungs of all mice (5×10^5 – 1×10^7 CFU/lung). Similar lung CFU counts were obtained for both competition assays, suggesting that the KR2107 *fim* locus does not play a significant role in lung infection. Furthermore, no significant deviation in *fim2*-positive to *fim2*-negative strain ratios was evident for either competition assay (Figure 5-9A). Total liver and spleen CFU counts were used as a measure of the ability of bacteria to disseminate from the lungs into the bloodstream. Much lower numbers and greater mouse-to-mouse variation occurred in CFU counts for the livers (<15 – 1.6×10^4) and spleens (<20 – 200) of these mice. The median CFU count per liver for KR2107 (2.1×10^3) was elevated compared to that of KR2107Δ*fim2* (3.0×10^1), although this difference was not significant (P = 0.340) (Figure 5-9B). When the liver CFU counts were examined individually for each mouse there was a tendency, especially at higher total CFU counts, for livers to possess more KR2107 than KR2107Δ*fim2* (Figure 5-9C). No difference was found between the median CFU counts per liver for KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* (1.5×10^1). Thus, although *fim2* does not appear to influence the pneumovirulence of *K. pneumoniae* in a murine lung infection model, *fim2* may play a role in the dissemination of infection.

Figure 5-9 :- Murine lung infection model studies on *K. pneumoniae* KR2107 and its isogenic *fim* and/or *fim2* mutants.

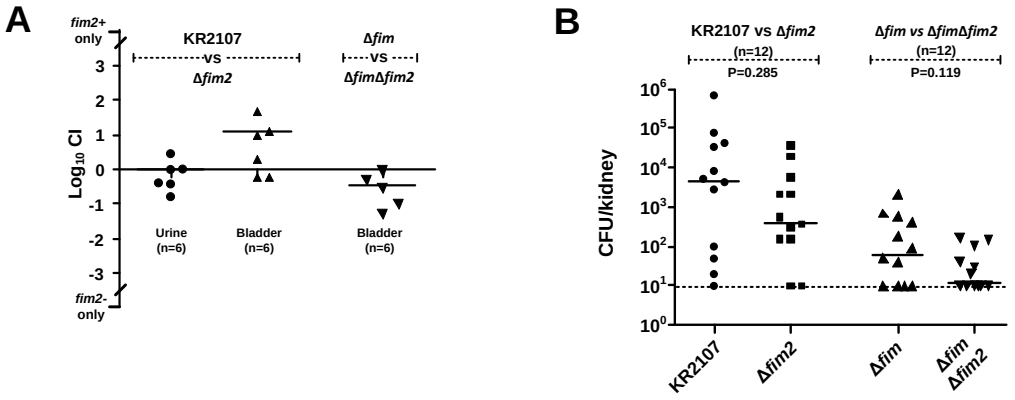


(A) Comparison of the ability of KR2107 and its isogenic mutants to infect the lungs as assessed by two head-to-head competition assays. A mixture containing an equal ratio of each competing strain was inoculated intranasally into five mice (5×10^7 CFU). The competitive index (CI) is the ratio of the number of *fim2*-positive to *fim2*-negative bacteria recovered from infected organs divided by the equivalent ratio as present in the intranasal inoculum. **(B)** Differential CFU counts for each of the competing strains in the liver at 30 h post-inoculation; the lower limit of detection is represented by the dotted line. **(C)** Liver CFU counts obtained in the five mice used for the competition assay between KR2107 and its isogenic *fim2* mutant. In A and B horizontal bars represent the median, with data points for each mouse as indicated. *P* values were calculated using the Mann-Whitney U test.

5.2.7.3 Murine urinary tract infection model

Type 1 fimbriae are a well-established virulence factor of *K. pneumoniae* urinary tract infections (387, 388). To assess the role of *fim2* in urinary tract infections, a group of six mice was inoculated transurethrally with a 1:1 mixture of KR2107 and its *fim2* mutant and sacrificed 3 days post-inoculation. All urine and bladder samples were found to be colonized and a median CFU count of 8.7×10^5 per bladder and 5.0×10^4 per ml of urine was obtained. In all mice the infection had ascended into the kidneys producing a median bacterial count of 5.3×10^3 per kidney (n = 12). The average CI value obtained for bladder samples indicates a 10-fold difference between KR2107 and *fim2* mutant CFU counts (Figure 5-10A). These values are supported by the median kidney CFU count which was 10-fold higher for the wildtype (4.8×10^3) than the *fim2* mutant (4.8×10^2), although this difference is not statistically significant (P = 0.285) (Figure 5-10B). Therefore, the data indicate that *fim2* may exert a subtle influence on the urovirulence of *K. pneumoniae*.

Figure 5-10 :- Murine ascending urinary tract infection model studies on *K. pneumoniae* KR2107 and its isogenic *fim* and/or *fim2* mutants.



(A) Comparison of the urovirulence of KR2107 and its isogenic mutants as assessed by two head-to-head competition assays. A mixture containing approximately equal numbers of each competing strain was inoculated into the bladders of six mice (5×10^8 CFU). The competitive index (CI) is the ratio of the number of *fim2*-positive to *fim2*-negative bacteria recovered from urine or bladder divided by the equivalent ratio as present in the infecting inoculum. **(B)** Differential CFU counts for each of the competing strains in the left and right kidneys at 3 days post-inoculation; the lower limit of detection is represented by the dotted line. In both of the above analyses horizontal bars represent the median, with data points for each mouse as indicated. *P* values were calculated using the Mann-Whitney U test.

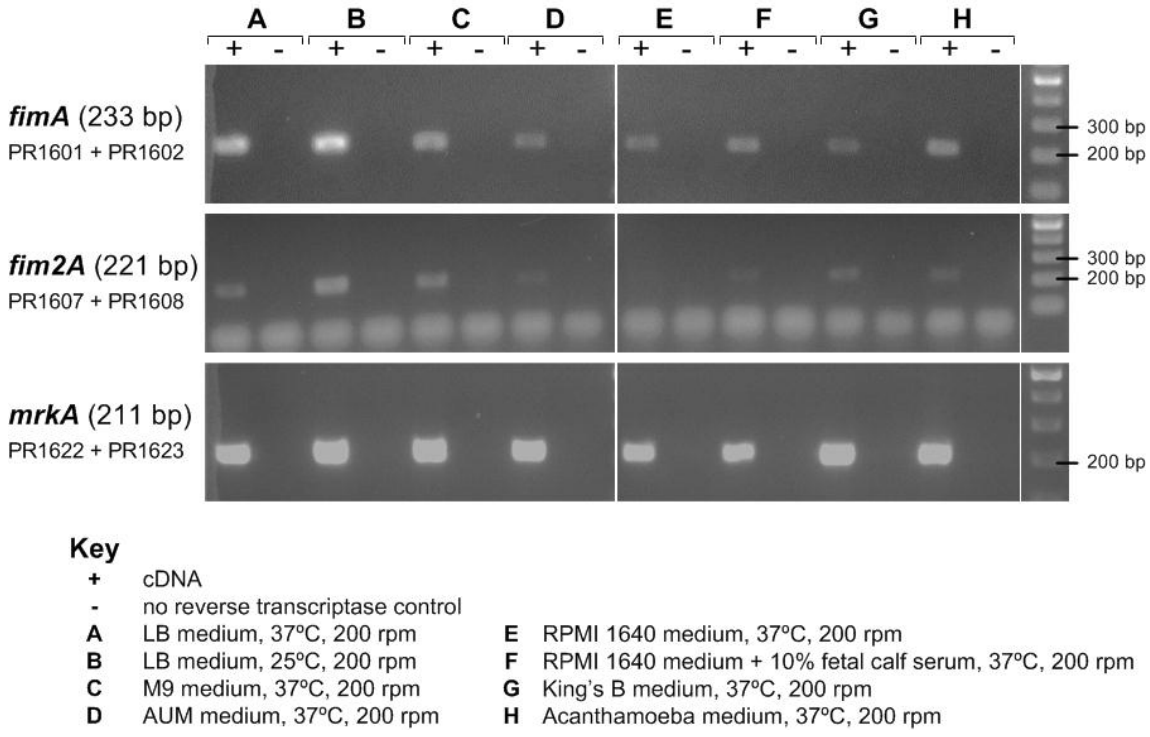
To investigate potential genetic redundancy or functional masking between *fim* and *fim2*, the competition assay was repeated in a *fim*-negative background. Consistent with previous data (388) and compared to the initial competition assay, bacterial counts were considerably lower in this *fim*-negative background experiment. Infection was established in the bladders of five mice (n = 6) with a median bacterial count of 1.0×10^2. At time of sacrifice, infection had ascended into nine kidneys (n = 12) with a median CFU count of 1.7×10^2. However, in all cases no bacteria were isolated from the urine suggesting counts of less than 50 per ml. The average CI value obtained for bladder samples show that CFU counts for KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2* did not differ significantly (Figure 5-10A). However, the median kidney CFU counts were on average 5-fold higher for the KR2107Δ*fim* (7.0×10^1) than KR2107Δ*fim*Δ*fim2* mutant (1.5×10^1), and although similar to the results obtained in the *fim*-positive background these results were also not statistically significant (*P* = 0.119) (Figure 5-10B). These

results have confirmed the importance of *fim* in *K. pneumoniae*-mediated urovirulence and suggest a potential subtle accessory role for *fim2*.

5.2.8 *In vitro* and *in vivo* expression of *fim2*, *fim* and *mrk* operons

Many chaperone/usher operons are poorly expressed under laboratory conditions (177, 210). To investigate *fim2* expression, RNA was isolated from KR2107 which had been cultured in LB medium for 16 h (37°C, 200 rpm) and a cDNA library was constructed using random primer-based RT-PCR. This process was repeated with KR2107 grown for 16 h at 25°C in LB medium as well as at 37°C in media that mimiced certain conditions: M9 (minimal), AUM (urine), RPMI 1640 ('physiological' medium), RPMI 1640 plus 10 % FCS ('physiological' medium plus serum), King's B and *Acanthamoeba* medium. PCR analysis of these cDNA libraries detected transcripts corresponding to *fimA*, *fim2A* and *mrkA* in all media types, while reverse transcriptase-free cDNA libraries did not yield any products, confirming the absence of DNA carryover (Figure 5-11). Transcripts corresponding to *fimH*, *fimK*, *fim2H*, *fim2K* and *mrkD* were also detected in all media types (data not shown). These results indicate that all three operons are expressed in all media types examined.

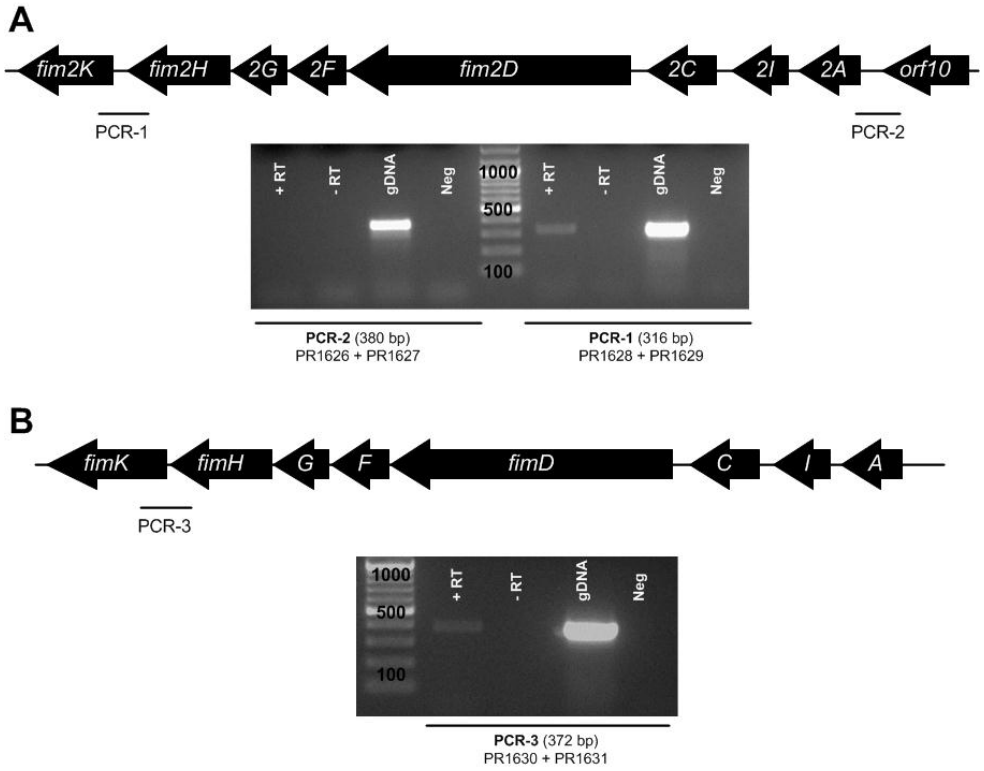
Figure 5-11 :- PCR analysis of KR2107 cDNA libraries for *fimA*, *fim2A* and *mrkA* transcripts



RNA was extracted from *K. pneumoniae* KR2107 grown in eight different media types and was used to construct cDNA libraries in addition to complementary no reverse transcriptase controls to measure DNA contamination. Transcripts corresponding to *fimA*, *fim2A* and *mrkA* could be detected in all media types. As all three target genes were amplified from identical cDNA libraries and assuming that PCR efficiencies of all reactions were similar, both *mrkA* and *fimA* appeared to be expressed at higher levels than *fim2A*.

To examine the organisation of *fim2* mRNA, possible mRNA linkage between *fim2K* and *fim2H*, and *orf10* and *fim2A* was investigated by PCR analysis of the cDNA library of KR2107 grown in LB medium (37ºC, 200 rpm). PCR analysis of the latter region did not yield a product, while amplification of the *fim2H-fim2K* linking region produced a specific band (Figure 5-12A). By analogy with the C3091 *fim* operon, this strongly suggests that the *fim2* cluster comprises eight genes which are expressed as a single mRNA transcript. Furthermore, the results displayed in Figure 5-12B confirm that *fimK* and *fimH* are genetically linked within the KR2107 *fim* operon. Therefore, *fimK* is expressed concurrently with the remaining structural subunit-encoding type 1 fimbrial genes; this linkage had not previously been demonstrated.

Figure 5-12 :- Investigation into the structure of the *fim2* and *fim* operons.



The cDNA and mock cDNA library generated from RNA extracted from *K. pneumoniae* KR2107 (LB, 16 h 37ºC, 200 rpm) was used to investigate the structure of the *fim2* and *fim* operons. **A)** PCR-1 proves *fim2K* is part of the *fim2* transcript, whilst PCR-2 confirms that *orf10* is not. The *fim2* promoter is therefore predicted to be positioned within the region amplified by PCR-2. **B)** PCR-3 proves *fimK* is part of the type 1 fimbrial operon transcriptional unit. -RT, no reverse transcriptase mock cDNA library ; +RT, plus reverse transcriptase cDNA library ; gDNA, genomic DNA ; Neg, negative.

Media type and alterations in environmental conditions affect the expression of *E. coli* type 1 and other related CU fimbriae (130, 355, 387). To quantify the effect of media conditions on the expression levels of *fim2A, fimA and mrkA*, qRT-PCR experiments were performed on six of the aforementioned cDNA libraries. Each qRT-PCR reaction was performed in triplicate and average Ct values were used to calculate relative expression levels with the Pfaffl method (294) (Figure 5-13A). Raw data Ct values are presented in Appendix 8. qRT-PCRs on RT negative control cDNA libraries were negative, confirming the absence of DNA contamination. To normalize Ct values between cDNA libraries, housekeeping gene *rpoD* was used as an endogenous PCR control, as described previously (183). PCR efficiencies were determined to be 91 %,

90 %, 89 % and 91 % for *rpoD*, *fimA*, *fim2A* and *mrkA*, respectively, and were taken into account when calculating relative expression levels.

qRT-PCR analysis identified that, regardless of growth media used, there was an ordered gradation in the relative expression levels of the fimbrial genes examined; the highest quantity of transcript was always detected for *mrkA*, followed by *fimA* and then *fim2A* (Figure 5-13B). Furthermore, the actual relative expression ratio differed considerably between media types. *mrkA* expression levels were determined to be 15- to 170-fold higher than *fim2A*, whilst those of *fimA* were 1.5- to 30-fold higher.

To investigate the effect of media type and environmental conditions on the expression levels of individual genes, the relative expression levels under each condition were compared to the expression level in LB at 37ºC (Figure 5-13C). Expression of *fim2A* was temperature sensitive and increased 3-fold when KR2107 was cultured in LB at 25ºC instead of at 37ºC. Furthermore, it was 5-fold greater in M9 minimal medium than LB. The expression of *fimA* did not appear temperature sensitive, but was reduced 4- to -8 fold when grown in non-LB type media. Finally, *mrkA* expression was only slightly increased in LB 25ºC and RPMI 1640 cultures. Remarkably, when the latter was supplemented with 10% FCS a 25-fold reduction in *mrkA* transcript level was detected.

Figure 5-13 :- qRT-PCR analysis of *fimA*, *fim2A* and *mrkA* expression

A) The Pfaffl equation

$$relative\ ratio = \frac{(E_{target})^{\Delta CP_{target}\ (control\ -sample\ )}}{(E_{ref})^{\Delta CP_{ref}\ (control\ -sample\ )}}$$

E_{target} = Target PCR efficiency = $10^{(-1/slope)}$

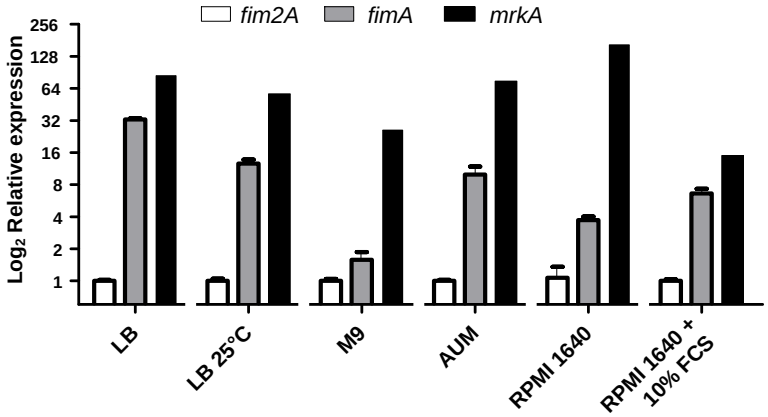
E_{ref} = Reference PCR efficiency = $10^{(-1/slope)}$

ΔCP_{target} = Target PCR (average $Ct_{control}$ – average Ct_{sample})

ΔCP_{ref} = Reference PCR (average $Ct_{control}$ – average Ct_{sample})

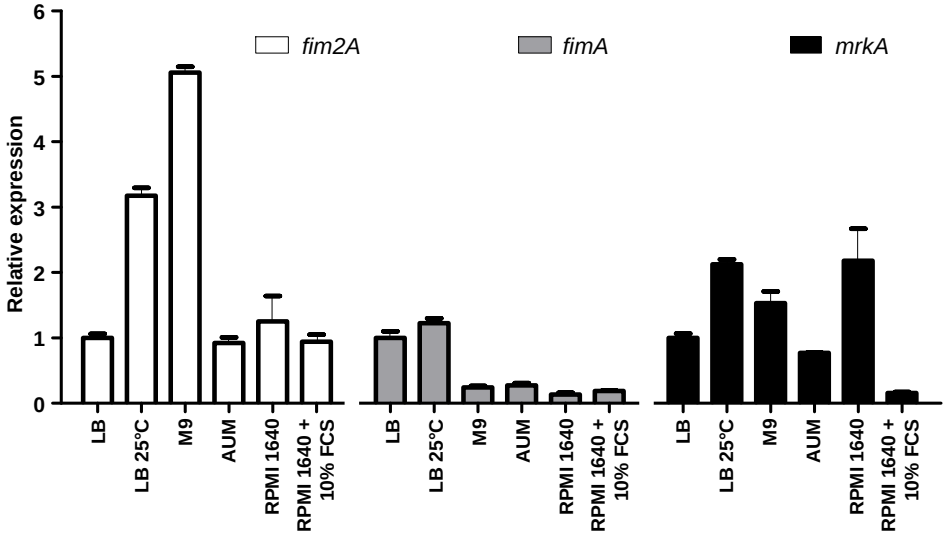
B) Relative expression of *fim2A*, *fimA* and *mrkA* in different media types

Relative expression levels have been calculated using *fim2A* as the reference



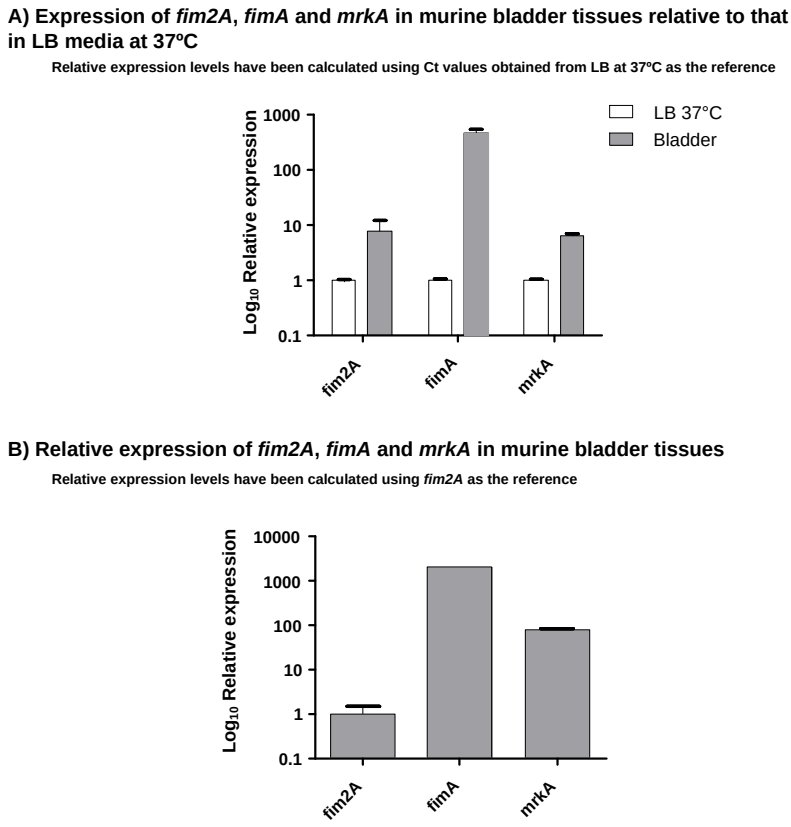
C) Effect of media type on expression of *fim2A*, *fimA* and *mrkA*

Relative expression levels have been calculated using Ct values obtained from LB at 37ºC as the reference



cDNA libraries were made from RNA extracted from *K. pneumoniae* KR2107 grown in various media types (16 h, 37ºC, 200 rpm, unless stated otherwise). Triplicate qRT-PCR reactions were used to examine the relative expression levels of *fim2A*, *fimA* and *mrkA*. **(A)** The Pfaffl equation was used to calculate relative expression ratios (294). **(B)** Relative expression of *fimA* and *mrkA* compared to *fim2A* in different media types. **(C)** Relative expression levels of *fim2A*, *fimA* and *mrkA* in six growth media compared to LB medium. Error bars represent means and standard deviations. Ct, cycle threshold.

Figure 5-14 :- qRT-PCR analysis of the expression levels of *fim2A*, *fimA* and *mrkA* in murine bladder tissues.



Murine bladders were isolated 3 days after the start of an *in vivo* competition assay between *K. pneumoniae* KR2107 and KR2107Δ*fim*. Bacteria were isolated from the bladder tissue, RNA was extracted and cDNA libraries constructed. Triplicate qRT-PCR reactions was used to examine the relative expression levels of *fim2A*, *fimA* and *mrkA*. **A)** Relative expression levels of *fim2A*, *fimA* and *mrkA* in bladder tissues compared to growth in LB medium at 37ºC. **B)** Relative expression of *fimA* and *mrkA* compared to *fim2A* in murine bladder tissues. Error bars represent means and standard deviations.

Further qRT-PCRs were performed to examine potential *in vivo* upregulation of *fim2A*, *fimA* and/or *mrkA* expression. Six bladder samples were harvested during the urinary tract infection competition assay between KR2107 and KR2107Δ*fim*. Approximately one tenth of each sample (containing approximately 1×10^6 CFU[i]) was retained to construct a cDNA library representative of *K. pneumoniae* KR2107 gene expression in the murine bladder. Bladder samples were pooled and homogenised, and the tissue-

[i] Although this was a mixed KR2107 and KR2107Δ*fim* sample, the wildtype had outcompeted the mutant 1000-fold. This sample was therefore almost identical to a pure KR2107 sample.

associated bacteria were extracted, their RNA isolated, cDNA libraries constructed and qRT-PCR analysis performed. Within the bladder, all three fimbrial genes were upregulated compared to their expression levels in LB media at 37ºC (Figure 5-14A). Most strikingly, *fimA* expression was 490-fold higher in the bladder. When the relative quantities of *fimA*, *fim2A* and *mrk* transcript in bladder samples were compared, 2000-fold more *fimA* than *fim2A* transcript was detected, whilst *fimA* expression was 26-fold higher than that of *mrkA* (Figure 5-14B). Together, these results indicate that type 1 fimbriae are a major virulence factor in urinary tract infections, and that the products of the *fim2* and *mrk* operons, if relevant, may play only a subtle accessory role.

5.3 Discussion

Bacterial adhesion to and colonization of host cells is frequently mediated by a diverse assortment of afimbrial and fimbrial adhesins, each thought to possess a particular tissue tropism (300). The vast majority of *K. pneumoniae* strains are able to produce type 1 fimbriae (303, 390). These structures are associated with mannose-sensitive agglutination of GP-RBCs, a phenotype caused by interaction of the adhesin subunit FimH with terminally-exposed mannose residues in N-linked oligosaccharides on cell surfaces (385). Previously it has been shown that the mono- and tri-mannose residues in FimH are highly conserved (385). Although *E. coli* HB101 heterologously expressing the *fim* operon could mediate GP-RBC agglutination, HB101 expressing the *fim2* operon could not. The specific binding properties of Fim2H, the putative Fim2 adhesin, remain to be identified but it is unlikely to bind mannose residues as it possesses only four out of 14 mono- and tri-mannose binding residues present in FimH.

Despite multiple attempts fimbrial structures could not be visualised using transmission electron microscopy when the *fim2* operon was overexpressed in *E. coli* HB101. Paradoxically, biofilm forming ability appeared to be enhanced in the analysed *E. coli* strain. These results are similar to those of a study in which constitutive expression of

four out of seven *E. coli* CU fimbrial operons was shown to cause phenotypic alternations despite the fact that fimbrial appendages could not be visualized by transmission electron microscopy (211). It is possible that the *fim2* operon may code for an afimbrial adhesin or a short and/or thin fimbrial structure that is not readily visualized by transmission electron microscopy, or one that is extremely fragile. Alternatively, the *fim2* locus may produce extracellular filaments similar to the curli fibres expressed by many *E. coli* and *Salmonella* spp. strains (22). However, this is highly speculative as there are no known CU fimbrial operons that encode afimbrial structures. Additionally, although the presence of the Fim2 structure could not be confirmed using electron microscopy, ideally these proteins should be detected using either anti-Fim2 antibodies in Western Blotting or immunogold labelling experiments.

Remarkably, overexpression of the *fim2* locus using the IPTG-inducible P_{TRC} promoter-based vector pJTOOL-7 caused considerable growth defects in HB101 and three *K. pneumoniae* strains. This effect did not occur with the *fim* operon, was independent of the strain origin of the *fim2* locus and was not altered by the presence of IS*630* in *fim2D*. However, in the latter case induction may have inadvertently overexpressed the putative IS*630* transposase gene, leading to subsequent excision of IS*630* and restoration of the *fim2* operon, hence the observation of a similar phenotype. Whether this had occurred requires further investigation. Inhibition of cell growth and protein synthesis was previously described in an *E. coli* strain overexpressing the phage lambda *int* gene, which contains multiple rare codons (451, 452). By supplementing the strain with rare tRNAs this growth defect was overcome. Similar methods have been used to express other problematic genes (74, 190, 399). Likewise, the *fim2* operon genes also possessed large clusters of rare codons but, unfortunately, when rare tRNAs were provided by the co-transformation of plasmid pRARE the growth of HB101 strains expressing *fim2* did not improve. It is possible that the observed growth retardation was related to the depletion of other molecules used in transcription and

translation, such as amino acids or ATP. This increased metabolic burden can result in a stress response in *E. coli*, which may downregulate genes involved in transcription and translation (399). However, the latter seems unlikely as normal growth occurred during IPTG-induced expression of the *fim* operon which, based on predicted protein products, would be expected to form a similar metabolic burden on host cells. Similarly, reducing temperature (and therefore growth rate and protein synthesis) to 30ºC did not alleviate the issue, and neither did decreasing the IPTG concentration from 0.1 mM to 0.01mM. Lastly, the growth defect persisted when *fim2* was expressed within *K. pneumoniae* strains KR2107 and C3091, suggesting that the phenotype observed in HB101 was not due to *fim2* expression requiring a factor unique to *K. pneumoniae*. Further experimentation will be required to identify the exact cause and nature of the observed growth inhibition.

Analysis of the three sequenced *K. pneumoniae* strains has revealed that, in addition to the *fim* and *mrk* operons, these genomes collectively encode at least six other CU fimbrial systems (387, 388), one or more of which may perform an as-yet characterised role in adhesion to target tissues. To investigate the role of the *fim2* CU system in virulence, isogenic *fim2* mutants were constructed and examined in three murine models, each focussed on primary infection of a distinct clinically-relevant anatomical site. Surprisingly, despite many fimbrial systems having been clearly implicated in virulence, no clear evidence of attenuation (murine lung and urinary tract infection models) or reduction in colonizing ability (murine intestinal colonization model) was detected in the *fim2*-negative strains studied.

Intriguingly, examination of bladder CFU count-based CIs for the urinary tract infection experiments hinted at a subtle role for *fim2* in the colonization of bladder and kidney tissues. In both tissues, median wildtype CFU counts were approximately ten-fold higher than those of the *fim2* mutant, although when performed in a *fim* negative background this difference was reduced in both the bladder and kidney samples. The

latter are likely to be coincidental due to the low CFU counts obtained in the *fim* negative background. Additionally, *fim2* transcripts levels were almost ten-fold higher in murine bladder samples than in a standard LB culture, again suggesting a minor role for *fim2* in urovirulence. As shown by neutral CI values in the lung tissue but 100-fold higher median liver CFU counts for KR2107 as compared to its isogenic *fim2* mutant, the *fim2* locus may also be involved in the dissemination of *K. pneumoniae* from the respiratory tract. As an aside, the previously demonstrated dramatic positive contribution of *fim* to urovirulence in this murine model was also shown to be the case in the KR2107 background (387, 388). This was validated by identifying that *fim* expression was greatly upregulated in bladder tissues compared to *in vitro* cultures. Finally, at an overview level, based on total CFU counts per liver and per kidney for the lung infection and ascending urinary tract infection models, respectively, there was a suggestion, though not supported statistically, of an ordered gradation amongst the four isogenic strains with the most-to-least virulent as follows: KR2107, KR2107Δ*fim2*, KR2107Δ*fim* and KR2107Δ*fim*Δ*fim2*.

The *fim2* cluster was also assessed for its ability to contribute to biofilm formation. Gene knock-out experiments in KR2107 failed to reveal any role for *fim2* in biofilm formation. However, similar to type 1 fimbriae, the adhesive functionality conferred by the product of *fim2* may have been masked due to physical interference of the *K. pneumoniae* capsule or low level native expression in *in vitro* cultures (349, 391). Therefore, heterologous expression of *fim2* in the afimbriate *E. coli* strain HB101 and the bald *fim2*-negative *K. pneumoniae* strain C3091Δ*fim*Δ*mrk* was pursued. Evidence of a *fim2*-associated phenotype was elusive and only apparent in HB101 when crystal violet-staining data was standardised for total pre-wash cell numbers. Hence, although it would appear that expression of *fim2* in HB101 contributes to greater adherence to the well-surface and/or direct participation in biofilm formation, the biological significance of this finding remains to be examined.

The genomes of *E. coli* K-12, *E. coli* O157:H7 and *Salmonella* Typhi possess numerous cryptic CU fimbrial operons that are tightly regulated and not expressed under the majority of *in vitro* conditions tested (177, 211, 238). In this work, *fim2*-specific transcript was identified in standard LB cultures but the amount detected was 30- to 90-fold lower than that identified for *fim* and *mrk*, respectively. Transcripts corresponding to *fim2A*, *fimA* and *mrkA* were also be detected in other culture media. Most importantly, regardless of the growth medium used, there was an ordered gradation in the relative amount of transcript present: *mrkA* (most), *fimA*, *fim2A* (least). It is unknown whether this is proportional to the number of fimbriae of each type present on the surface of *K. pneumoniae* cells.

Compared to the *K. pneumoniae* genome-averaged AT content (~43 %), *fim2* is AT-rich (53 %) and the putative promoter region upstream of *fim2A* possesses an even higher AT-content (73 %). As moderate-to-marked upregulation of seven CU fimbrial operons has been reported in an *E. coli* K-12 H-NS mutant (211), the finding of an AT-rich *fim2* promoter region suggests that the H-NS protein may bind within this region and play a role in controlling this operon as well. Moreover, H-NS has been shown to preferentially bind to regions of horizontally-acquired DNA in *Salmonella enterica* serovar Typhimurium, and it is therefore possible this also occurs with KpGI-5 (239). In addition to the putative PDE Fim2K, KpGI-5 also encoded two other potential regulators one or more of which could alter *fim2* expression. Alternatively, other unknown regulators within the KR2107 genome may be involved. Similar to the *E. coli* type 1 fimbriae, temperature appears to regulate transcription of *fim2A*, as shown by a minor increase in transcript at 25ºC compared to 37ºC (97, 130). By analogy with other CU systems, it is likely that upregulation of *fim2* expression and biosynthesis of the putative Fim2 fimbriae is triggered by specific environmental conditions and involves a complex interplay of multiple transcriptional regulators such as H-NS, Fim2K and/or FimK and other surface components, such as the capsule (76, 211, 331, 349).

5.3.1 Future work

Despite the extensive work carried out in this chapter, several questions still remain. Firstly, although the *fim2* operon was expressed in KR2107, it is not known whether the associated protein products have been synthesised. Next, some progress has been made regarding understanding the transcriptional control of *fim2*, but the effect of specific environmental stimuli, and known and unknown fimbrial regulators, remains to be investigated. Lastly, it is still unknown what causes the growth retardation associated with IPTG-induced *fim2* expression from pFim2-Ptrc; it would be useful to solve this problem and re-examine some of the findings presented in this chapter. Strategies for examining these questions are presented below:

Confirmation of the presence of* fim2 *products on the surface of native KR2107

Raise an antibody against the putative major fimbrial subunit Fim2A. The antibody can then be employed in Western blotting and/or immunogold electron microscopy to identify *Fim2A* on both native *fim2*-positive *K. pneumoniae* cells and *E. coli* cells heterologously expressing the *fim2* operon.

Attempt to express* fim2 *without causing a growth defect

1/. A constitutive promoter, such as the λP_R promoter that was used in the investigation of *E. coli* K-12 fimbrial operons, could be inserted upstream of *fim2A* (210).

2/. The P_{TRC} promoter in pJTOOL-7 is a very strong promoter. A weaker promoter such as the arabinose-inducible P_{BAD} could be trialled (147).

Using these constructs the electron microscopy and biofilm assays could then be repeated. They could also be used in tissue adhesion assays.

Identify the transcriptional start site of the* fim2 *operon

In this work the *fim2* promoter was localised between *orf10* and *fim2A*, but the exact location remains unknown. Towards the end of the project two unsuccessful attempts

were made to identify the transcriptional start site using 5'RACE (357). With some protocol optimisations, specific 5'RACE products could be obtained in future work to map the *fim2* transcriptional start site and promoter.

Further examine the conditions that encourage* fim2 *expression

Although the qRT-PCR method was useful in this work, it is costly and time consuming when screening a large number of conditions. To pinpoint conditions that increase *fim2* expression it would be essential to examine the effect of growth phase, pH, NaCl concentration, temperature variation, glucose concentration, oxygen availability and others. Two simple systems can be constructed to investigate regulation of the *fim2* promoter:

1/. A plasmid-based operon fusion with *lacZYA* can be constructed by cloning the putative *fim2* promoter region upstream of a promoterless *lac* operon, as described previously (355, 414). Promoter activity can then be measured using a standard β-galactosidase activity assay using *o*-nitrophenyl-β-*d*-galactopyranoside as the substrate.

2/. Allelic exchange can be used to create a chromosomal *fim2A*::*lacZYA* transcriptional fusion, as described previously (28). Promoter activity can be measured as above.

Test various genes for their effect on* fim2 *transcription:

Fimbrial operons are controlled by a complex network of regulators. The role of various known and putative regulators in *fim2* expression can be investigated by deleting/disrupting their corresponding genes in the reporter strain KR116Δ*lacZYA* Δ*fim2A*::*lacZYA* using allelic exchange*,* and measuring the resulting effect on *fim2* expression using a β-galactosidase activity assay. Candidate proteins include FimK, Fim2K, MrkJ, Lrp, IHF, H-NS and Ler (76, 415). Alternatively, the genes of candidate

proteins can be cloned into expression plasmids, artificially overexpressed, and the resultant effect on *fim2A* expression measured.

Chapter 6. Preliminary characterisation of FimK and Fim2K, two novel, fimbrial operon-encoded putative phosphodiesterases

6.1 Introduction

The secondary intracellular messenger bis-(3'-5')-cyclic dimeric GMP (c-di-GMP) was first discovered in 1987 during work by Ross *et al.* to elucidate the mechanism of cellulose biosynthesis in *Gluconacetobacter xylinus* (334). In recent years c-di-GMP has been recognized as an important bacterial signal that coordinates the shift between planktonic and sessile lifestyles (365). Intracellular c-di-GMP levels are modulated by diguanylate cyclases (DGCs), which possess GGDEF domains, and phosphodiesterases (PDEs), which possess either EAL or HD-GYP domains. Whilst DGCs synthesise c-di-GMP from two GTP molecules, PDEs catalyse degradation of c-di-GMP into the linear dinucleotide pGpG by ester bond hydrolysis (180, 339, 342, 350).

Bioinformatics analyses of sequenced genomes have shown that the number of genes involved in the regulation of c-di-GMP increases with genome size in a non-linear manner. Thus, although most bacteria possess genes encoding DGCs and PDEs, there is substantial variation in the number of such genes between bacterial species and strains (131, 339). This highlights the likely complexity of c-di-GMP signalling in bacteria and the potential for considerable pathway variations.

So far only a few of the vast numbers of proteins predicted to be involved in c-di-GMP regulation have been characterized. These studies have identified that DGCs and PDEs both positively and negatively regulate important cellular functions including different forms of motility (275, 365), protein degradation and cell cycle progression (101), and synthesis of cellulose, secondary metabolites, adhesins and

exopolysaccharide matrix components (118, 134, 350). In addition, examples of c-di-GMP regulated fimbrial and virulence factor expression have been reported (220, 250, 331, 379, 381).

EAL domains can be classified as active or inactive based on their ability to hydrolyse c-di-GMP. Active domains are monomeric and require either Mg^{2+} or Mn^{2+} for hydrolytic activity, and are inhibited by Ca^{2+} and Zn^{2+} (350, 351). As revealed by crystal structures, EAL domains consist of a $(\beta/\alpha)_8$ barrel fold, a structure found in many hydrolases and isomerase enzymes (262). Similar to these enzymes, EAL domains also possess a flexible loop 6 between the β_6 strand and α_6 helix that acts as an active site "lid" for segregating substrates, keeping out solvents and releasing reaction products (186). Several essential catalytic amino acid residues have been determined for EAL domains. By structural analysis and site-directed mutagenesis, Rao *et al.* identified that residues E175, E265, D295 and N233 of the *P. aeruginosa* protein RocR were important for binding Mg^{2+} (319, 320). In a follow-up study by Tchigvintsev *et al*., these and an additional six charged/polar residues were found to be conserved in all hydrolytically active EAL domains on the basis of a sequence alignment of 13 EAL domains with confirmed hydrolytic activity and 3700 EAL domains with unknown functionality (404).

Although several hydrolytically active EAL domains have been identified, many examined domains are non-enzymatically active and their physiological role is still unclear. It must be emphasised that this does not mean that these proteins are non-functional, as highlighted by the *E. coli* BLUF-EAL protein YcgF (420). Although YcgF is unable to hydrolyse c-di-GMP, when irradiated with blue light it interacts with the DNA-associated repressor YcgE and releases YcgE from cognate operator sites, subsequently inducing a regulon of eight genes involved in biofilm formation and curli synthesis (420). Furthermore, the *E. coli* protein YhdA (also called CsrD) which possesses degenerate EAL and GGDEF domains can stimulate degradation of the

small regulator RNAs CsrB and CsrC, which in turn regulate expression of the flagellar master regulator FlhDC (396).

DGC activity can be controlled by input signals from a multitude of environmental and internal sources, including light, oxygen and redox conditions, starvation, and extracellular substances such as antibiotics or signalling molecules (160). Activity modulation can occur via a direct protein linkage of an EAL domain to a signal input domain, such as blue light sensing FAD-associated (BLUF) domains, helix-turn-helix domains or haem/flavin-associated PAS domains (132, 420). Alternatively, EAL domains may be linked to two-component signal transduction cascades which can sense environmental signals via sensory histidine kinases that autophosphorylate and subsequently transfer their phosphate residue to REC (two component receiver) domains, which often co-exist with EAL domains on proteins (160, 350). This alters the activity of receiver protein EAL domains and has been well described in the *V. cholerae* VieSAB system (250).

The *K. pneumoniae fim* and *fim2* operons encode FimK and Fim2K, respectively, which are two putative EAL domain-containing PDEs. Unlike FimK, Fim2K does not possess an N-terminal helix-turn-helix domain, which may be involved in DNA binding or the regulation of the associated EAL domain (1, 331, 387). It is therefore unlikely that FimK and Fim2K function in a similar manner. Rosen and colleagues have shed some light on the biological role of FimK, expression of which appears to decrease type 1 fimbrial expression and biofilm formation by an as-yet unknown mechanism (331).

This chapter will first discuss the identification of domains and conserved catalytic residues in the sequences of FimK and Fim2K. Next, attempts to express soluble recombinant FimK and Fim2K for biochemical analysis will be discussed. Subsequently, *fimK* and *fim2K* will be expressed in *K. pneumoniae* strains and the effects on GP-RBC agglutination, biofilm formation and exopolysaccharide production

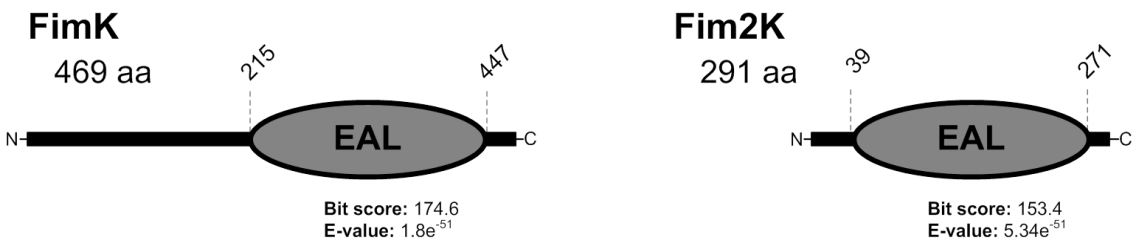
will be explored. Finally, the ability of *fimK* and *fim2K* to modulate prodigiosin production and motility in *Serratia* sp. ATCC 39006 when heterologously expressed in this species will also be examined.

6.2 Results

6.2.1 Bioinformatics analysis of FimK, Fim2K and cognate EAL domains

The predicted protein products of the *K. pneumoniae fimK* and *fim2K* genes were analysed for conserved domains using Pfam (120). FimK and Fim2K were 469 and 291 aa in length, respectively, and possessed a conserved C-terminal EAL domain (Figure 6-1). The EAL domain regions of FimK and Fim2K were 68 % similar and 55 % identical. Of note, the N-terminal of FimK possessed a low significance match to the GerE family, which included helix-turn-helix domain regions present in transcriptional regulators belonging to the LuxR/FixJ response regulator family (133).

Figure 6-1 :- Domain analysis of FimK and Fim2K

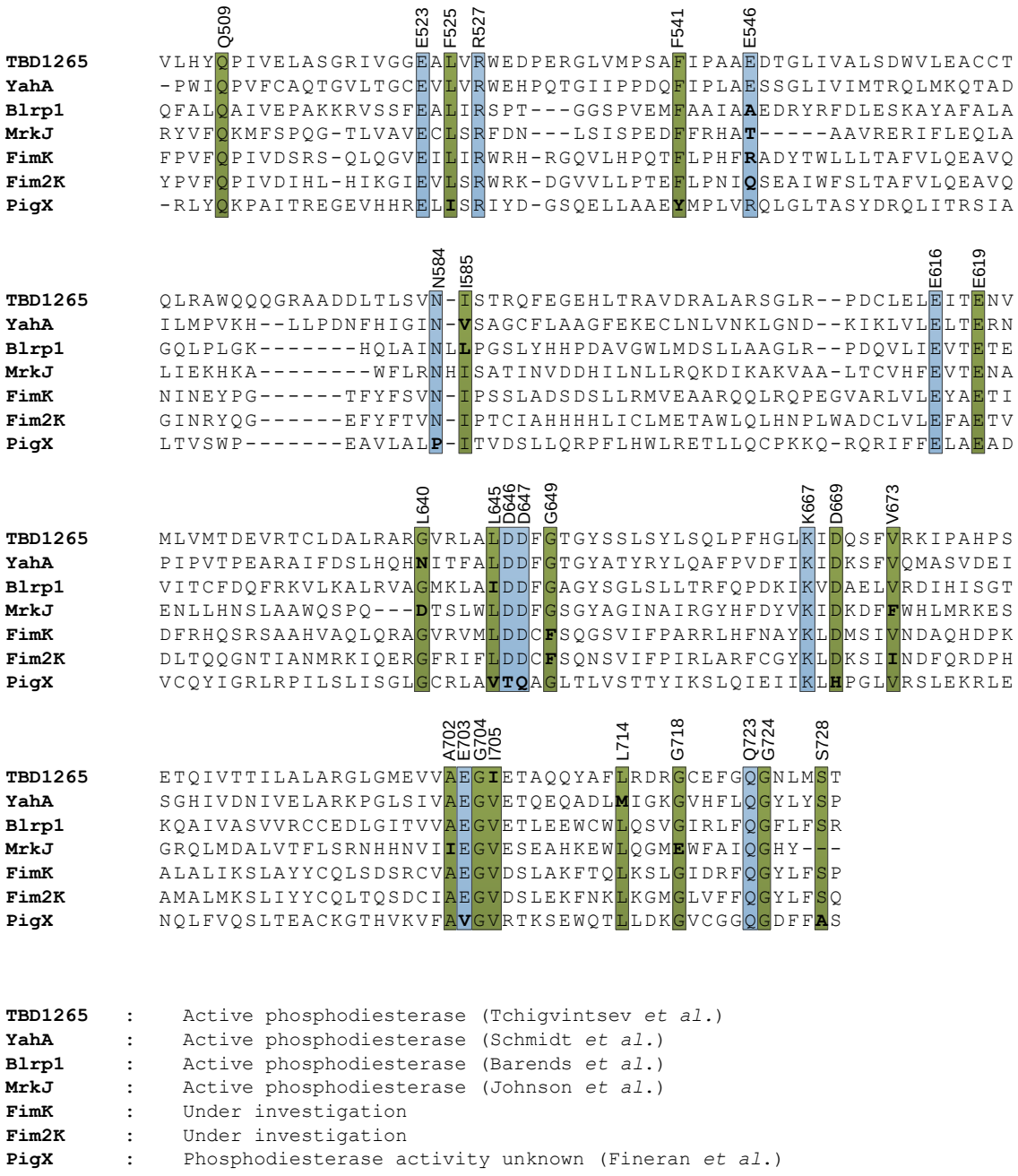


FimK and Fim2K were analysed using Pfam. Both proteins possessed significant matches to the EAL domain family. Of note, the N-terminal of FimK possessed an insignificant match (Bit score: 13.1, E-value: 0.036) to the GerE family, which includes helix-turn-helix domains present in transcriptional regulators of the LuxR/FixJ response regulator family. aa, amino acids.

To identify amino acid residues previously determined to be essential for c-di-GMP hydrolysis and recognise potential PDE activity, FimK and Fim2K were aligned with the protein sequences of TBD1265, YahA, Blrp1 and MrkJ using ClustalW. Nine of the ten

conserved residues were also found in FimK and Fim2K, whilst alternative amino acids were present at residue E546, which is involved metal ion binding (usually Mn^{2+} or Mg^{2+}) required for catalysis (Figure 6-2) (20, 404). Based on amino acid properties, the substitutions in FimK and Fim2K were likely to be neutral to protein function (21). The substitutions were also unlikely to affect PDE activity since the hydrolytically active *K. pneumoniae* protein Blrp1 also possessed an amino acid substitution at E546 (20). Intriguingly, E546 was also altered in MrkJ, a *K. pneumoniae* protein that can hydrolyse c-di-GMP and upregulate type 3 fimbrial production (183). These findings therefore suggest that the ten 'conserved' residues may not all be required to form a hydrolytically active EAL domain and that FimK and Fim2K may be active PDEs.

Figure 6-2 :- ClustalW alignment of the TBD1265, YahA, Blrp1, MrkJ, FimK, Fim2K and PigX EAL domains.



ClustalW alignment of the EAL domains from proteins TBD1265 (404), YahA (351), Blrp1 (20), MrkJ (183), FimK, Fim2K and PigX (118). In this alignment the ten conserved residues identified by Tchigvintsev *et al*. (404) have been highlighted in blue. Amino acid residues that were not identified by Tchigvintsev *et al.* but were present in at least five out of the seven proteins in this alignment were designated as conserved residues and were highlighted in green. Discrepant amino acid residues have been emphasized in bold. Residues E523, N584, E616, D646 and D647 have a role in binding metal ions, whilst E703 is thought to coordinate catalytic water molecules in conjunction with K677. R527 and E546 are thought to interact with c-di-GMP. The function of Q723 is unknown. The ClustalW alignment was produced using default parameters (223).

6.2.2 Genetic constructs to investigate *fimK* and *fim2K* function

Several plasmids were constructed for both recombinant protein expression and/or overexpression of *fimK* and *fim2K*. Briefly, *fimK*, *fim2K*, *fimK_EAL and fim2K_EAL* were cloned into a series of expression vectors with unique properties (Table 6-1). The latter two inserts only possessed the EAL domain-encoding portions of *fimK* and *fim2K*, respectively (section 6.2.3). More details regarding insert size and PCR primer pairs are presented in Appendix 6. Expression from vectors pET-28a (6xHis fusions), pETM-22 (6xHis and thioredoxin protein fusions), pETM-44 (6xHis and maltose binding protein (MBP) fusions) and pETM-66 (6xHis and NusA fusions) is driven from the T7 promoter and requires T7 RNA polymerase, which is not encoded by native *E. coli*, *Serratia* and *K. pneumoniae* strains (392). Typically, the recombinant host strain *E. coli* BL21(DE3) is used for expression as it has been genetically engineered to manufacture T7 RNA polymerase in the presence of IPTG (392). Expression from pGEX-2T is induced using IPTG and produces a glutathione S-transferase (GST) fusion protein. For IPTG-inducible plasmid-based expression in *K. pneumoniae*, *E. coli* and *Serratia* the plasmid pQE-80L (6xHis fusions) was employed.

When necessary, control plasmids that contained *yahA_EAL* (the EAL domain-encoding portion of the *E. coli* K-12 MG1655 *yahA*) or the entire *pigX* gene were used. The former plasmid encoded the biochemically characterised active PDE YahA (351), whilst the latter plasmid (pTA40) was gifted by Prof. G. Salmond and has been shown to influence prodigiosin formation in *Serratia* sp. ATCC 39006 (118).

The construction of *fimK* and/or *fim2K* knockouts in *K. pneumoniae* strains was discussed in a previous chapter (section 3.2.1.3). To investigate the function of *fimK* knockouts were made in three *fim*-positive, *fim2*-negative strains to produce KR161Δ*fimK*, KR162Δ*fimK* and KR173Δ*fimK*. Similarly, for functional analysis of *fimK* and *fim2K*, knockouts were also made in the *fim*- and *fim2*-positive strain KR116: KR116Δ*fimK,* KR116Δ*fim2K,* KR116Δ*fim2K*Δ*fimK.*

Table 6-1 :- Expression plasmids used in the characterisation of *fimK* and *fim2K*

pJOE number	Genotypic description	T7 promoter *	Product
fimK (FimK – 53.1 kDa)			
pJOE-2	pET28a::*fimK*	Yes	FimK::6xHis (55.3 kDa)
pJOE-10	pGEX-2T::*fimK*	No	FimK::GST (79.3 kDa)
pJOE-23	pETM-22::*fimK*	Yes	FimK::6xHis::Trx (67.5 kDa)
pJOE-27	pETM-44::*fimK*	Yes	FimK::6xHis::MBP (95.8 kDa)
pJOE-31	pETM-66::*fimK*	Yes	FimK::6xHis::NusA (110.2 kDa)
pJOE-20	pQE-80L::*fimK*	No	FimK::6xHis (54.4 kDa)
***fimK*_EAL** (FimK_EAL – 28.0 kDa)			
pJOE-24	pETM-22::*fimK_EAL*	Yes	FimK_EAL::6xHis::Trx (42.6 kDa)
pJOE-28	pETM-44::*fimK_EAL*	Yes	FimK_EAL::6xHis::MBP (70.8 kDa)
pJOE-32	pETM-66::*fimK_EAL*	Yes	FimK_EAL::6xHis::NusA (85.3 kDa)
fim2K (Fim2K – 33.5 kDa)			
pJOE-1	pET28a::*fim2K*	Yes	Fim2K::6xHis (35.0 kDa)
pJOE-9	pGEX-2T::*fim2K*	No	Fim2K::GST (59.1 kDa)
pJOE-25	pETM-22::*fim2K*	Yes	Fim2K::6xHis::Trx (47.1 kDa)
pJOE-29	pETM-44::*fim2K*	Yes	Fim2K::6xHis::MBP (75.4 kDa)
pJOE-33	pETM-66::*fim2K*	Yes	Fim2K::6xHis::NusA (89.9 kDa)
pJOE-21	pQE-80L::*fim2K*	No	Fim2K::6xHis (34.1 kDa)
***fim2K*_EAL** (Fim2K_EAL – 29.1 kDa)			
pJOE-26	pETM-22::*fim2K_EAL*	Yes	Fim2K_EAL::6xHis::Trx (43.5 kDa)
pJOE-30	pETM-44::*fim2K_EAL*	Yes	Fim2K_EAL::6xHis::MBP (71.8 kDa)
pJOE-34	pETM-66::*fim2K_EAL*	Yes	Fim2K_EAL::6xHis::NusA (86.2 kDa)
Control plasmids			
pJOE-3	pET28a::*yahA_EAL*	Yes	YahA_EAL::6xHis (34.2 kDa)
pJOE-19	pQE-80L::*yahA_EAL*	No	YahA_EAL::6xHis (34.2 kDa)

* Genes cloned downstream of the T7 promoter can only be induced in strains that encode the T7 RNA polymerase, such as *E. coli* BL21(DE3). GST, glutathione S-transferase ; MBP, maltose binding protein ; Trx, thioredoxin

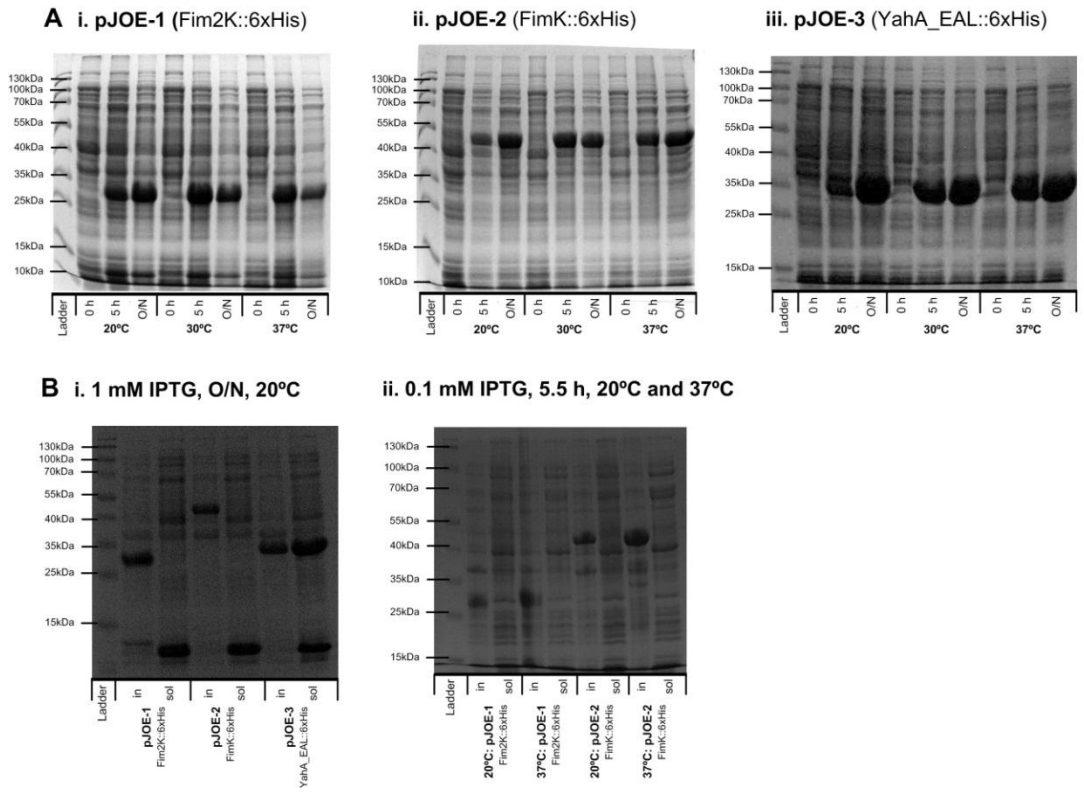
6.2.3 FimK::6xHis and Fim2K::6xHis are highly insoluble proteins

Several PDEs that cleave c-di-GMP have been purified and biochemically characterised, including YahA, RocR, and BlrP1 (20, 320, 351). With the aim to investigate the ability of FimK and Fim2K to hydrolyse c-di-GMP, *fimK* and *fim2K* were recombinantly expressed to produce high levels of the desired proteins for purification.

To aid subsequent purification steps FimK and Fim2K were produced as 6xHis tag fusions. The ability of histidine residues to strongly bind nickels ions is exploited in affinity purification of 6xHis-tagged recombinant proteins (166). Initial expression experiments in BL21(DE3) (1 mM IPTG, 37ºC) using pJOE-1, pJOE-2 and pJOE-3 produced vast quantities of Fim2K::6xHis, FimK::6xHis and YahA_EAL::6xHis, respectively (Figure 6-3A). Similar quantities were produced at 20ºC and 30ºC. It was observed that Fim2K::6xHis did not run at its exact molecular weight on a polyacrylamide gel; it resolved as an ~30 kDa protein. Anomalous SDS-PAGE migration has previously been observed for other proteins (Dr. Helen O'Hare, personal communication). A Western blot using an anti-polyhistidine peroxidase conjugate antibody confirmed the identity and integrity of the 6xHis-containing N-terminal of Fim2K::6xHis. Similarly, the N-termini of FimK::6xHis and YahA_EAL::6xHis were also intact (Figure 6-4).

Overexpression of recombinant proteins produces soluble protein and/or insoluble incorrectly folded protein aggregates, also known as inclusion bodies. The latter can be solubilized and then refolded to prepare active proteins, although this is often inefficient. It is therefore preferable to obtain correctly folded (soluble) protein for biochemical characterisation. To investigate the solubility of the examined proteins, overnight cultures of BL21(DE3) pJOE-1, pJOE-2 and pJOE-3 (1 mM IPTG, 20ºC) were pelleted and resuspended in lysis buffer with lysozyme. The suspension was sonicated and soluble and insoluble fractions were separated and analysed by SDS-PAGE. Neither FimK::6xHis nor Fim2K::6xHis could be detected in the soluble fraction, whilst control protein YahA_EAL::6xHis was present in both fractions (data not shown). Reducing IPTG concentration from 1 mM to 0.1 mM, induction time from overnight to 5.5 h and induction temperature from 37ºC to 20ºC did not affect protein solubility (Figure 6-3B).

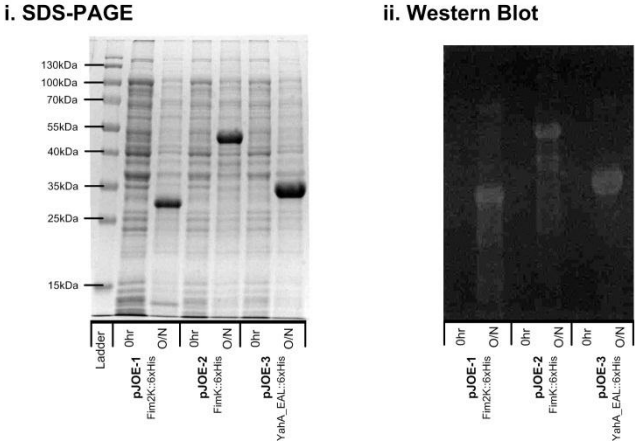
Figure 6-3 :- SDS-PAGE gel analysis of BL21(DE3) expressing recombinant proteins from plasmids pJOE-1, pJOE-2 and pJOE-3



A) *E. coli* BL21(DE3) strains harbouring pJOE-1, pJOE-2 and pJOE-3 were induced at 20ºC, 30ºC and 37ºC using 1 mM IPTG. Samples were taken at 0 h (before induction), at 5 h and after overnight expression. OD_{600} normalised volumes of whole cell lysates were loaded into 12 % polyacrylamide gels for analysis. PhastGel Blue R was used for gel staining. In all cases no protein was produced at 0 h. At 5 h and after overnight culture, large amounts of protein were produced that roughly corresponded to the predicted molecular weight of Fim2K::6xHis (35.0 kDa), FimK::6xHis (55.3 kDa) and YahA_EAL::6xHis (34.2 kDa). **B)** Separation of the insoluble and soluble protein fractions of *E. coli* BL21(DE3) expressing recombinant FimK::6xHis, Fim2K::6xHis and YahA_EAL::6xHis. Reducing IPTG concentration from 1 mM to 0.1 mM did not improve solubility of FimK::6xHis or Fim2K::6xHis, nor did decreasing time of induction or temperature. O/N, overnight; h, hours.

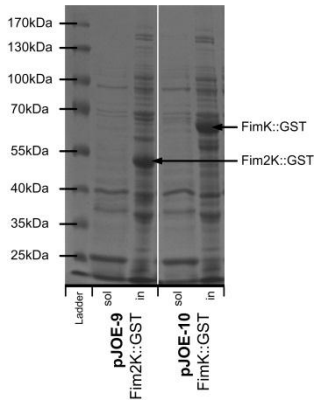
Commonly insoluble recombinant proteins are fused to GST to improve protein solubility and due to its ability to bind glutathione GST can also be used for affinity purification (373). GST fusion proteins FimK::GST and Fim2K::GST were produced in BL21(DE3) cells harbouring pJOE-10 and pJOE-9, respectively. When expression was induced for 4 h at 20ºC using 0.1 mM IPTG highly insoluble recombinant GST-fused proteins were produced (Figure 6-5).

Figure 6-4 :- Western blot of Fim2K::6xHis, FimK::6xHis and YahA_EAL::6xHis



A) *E. coli* BL21(DE3) strains harbouring pJOE-1, pJOE-2 and pJOE-3 were induced at 30ºC using 1 mM IPTG. Samples were taken at 0 h (before induction) and after overnight expression. OD_{600} normalised whole cell lysates were loaded into 12 % polyacrylamide gels for analysis. For all samples, induction of recombinant Fim2K::6xHis (35.0 kDa), FimK::6xHis (55.3 kDa) and YahA_EAL (34.2 kDa) could be identified. **B)** Negative image of nitrocellulose film. Chemiluminescence was very intense, and a correctly exposed x-ray film could not be obtained. However, areas of very intense chemiluminescence damaged the nitrocellulose film resulting in an 'image' on the film. These areas can be seen in the image and correspond to the anti-polyhistidine peroxidase conjugate antibody bound to Fim2K::6xHis, FimK::6xHis and YahA_EAL::6xHis, from left to right.

Figure 6-5 :- SDS-PAGE gel analysis of soluble/insoluble fractions of BL21(DE3) expressing Fim2K::GST and FimK::GST from pJOE-9 and pJOE-10, respectively.



E. coli BL21(DE3) strains harbouring pJOE-9 and pJOE-10 were induced for 4 h at 20ºC using 0.1 mM IPTG. Bacteria were harvested from cultures, sonicated, separated into soluble and insoluble fractions and analysed on an 8 % polyacrylamide gel. No visual traces of Fim2K::GST or FimK::GST could be detected in the extracted soluble fractions, whilst large amounts of protein that roughly corresponded to the predicted molecular weight of Fim2K::GST (59.1 kDa) and FimK::GST (79.3 kDa) were detected in the insoluble fraction.

These results strongly suggest that FimK and Fim2K are produced in an insoluble form under the conditions tested and that synthesis of soluble protein could not be encouraged by systematic alterations in induction temperature, duration or strength (IPTG concentration). To further examine the solubility of FimK and Fim2K, the native proteins and their 6xHis and GST-fused counterparts were analysed using an upgraded version of the PROtein SOlubility evaluator (372). This programme compares subtle differences between soluble and insoluble proteins deposited in TargetDB and enables solubility predictions of proteins with unknown solubility. A solubility score ranging from 0 to 1 is calculated (cut-off value at 0.6); the higher the score the more likely the analysed protein is soluble. In agreement with earlier solubility/insolubility analyses, the native and 6xHis or GST tagged versions FimK and Fim2K were predicted insoluble. Several other fusion tags have also been used to improve protein solubility including thioredoxin (218), MBP (192) and NusA (89). When FimK and Fim2K were fused to these tags in addition to 6xHis the solubility score increased considerably, although they were still predicted as insoluble (Table 6-2). Intriguingly, the control protein YahA_EAL::6xHis had a solubility score of 0.496, suggesting insolubility.

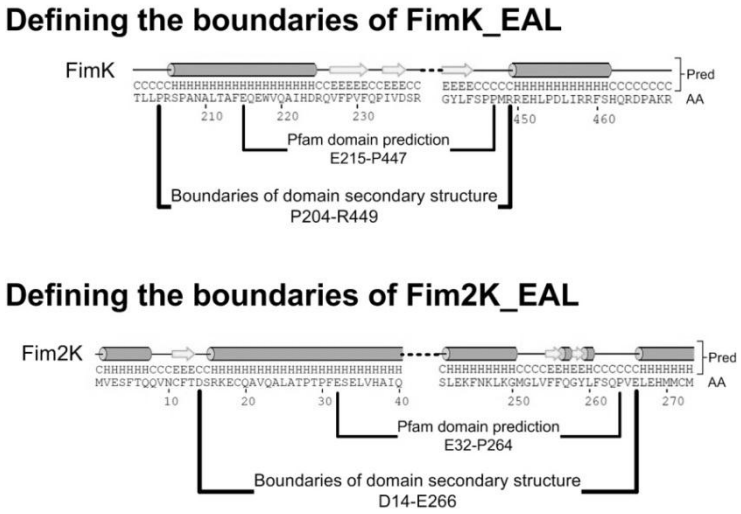
Table 6-2 :- PROtein SOlubility evaluation of FimK, Fim2K and their corresponding EAL domain-only versions fused with various tags

	FimK	**FimK_EAL**	**Fim2K**	**Fim2K_EAL**
None	insoluble = 0.377 *N/A*	insoluble = 0.521 *N/A*	insoluble = 0.422 *N/A*	insoluble = 0.489 *N/A*
6xHis	insoluble = 0.399 *pJOE-2*		insoluble = 0.357 *pJOE-1*	
GST	insoluble = 0.419 *pJOE-10*		insoluble = 0.398 *pJOE-9*	
6xHis + Trx [#]	insoluble = 0.425 *pJOE-23*	insoluble = 0.485 *pJOE-24*	insoluble = 0.395 *pJOE-25*	insoluble = 0.444 *pJOE-26*
6xHis + MBP	insoluble = 0.564 *pJOE-27*	**soluble = 0.635** ***pJOE-28***	insoluble = 0.583 *pJOE-29*	**soluble = 0.615** ***pJOE-30***
6xHis + NusA	insoluble = 0.565 *pJOE-31*	**soluble = 0.625** ***pJOE-32***	insoluble = 0.581 *pJOE-33*	**soluble = 0.607** ***pJOE-34***

[#] Trx, thioredoxin.

Since EAL domains comprise only part of the FimK and Fim2K proteins, it was investigated whether truncated proteins consisting of only the EAL domain would possess improved solubility. The process of defining domain boundaries for FimK and Fim2K_EAL is discussed in Figure 6-6. Compared to their intact parent proteins, these truncated proteins displayed increased predicted solubility, which was further increased when fused with MBP and NusA (Table 6-2). Intriguingly, FimK_EAL and Fim2K_EAL fusions with thioredoxin were calculated as less soluble than their unfused counterparts.

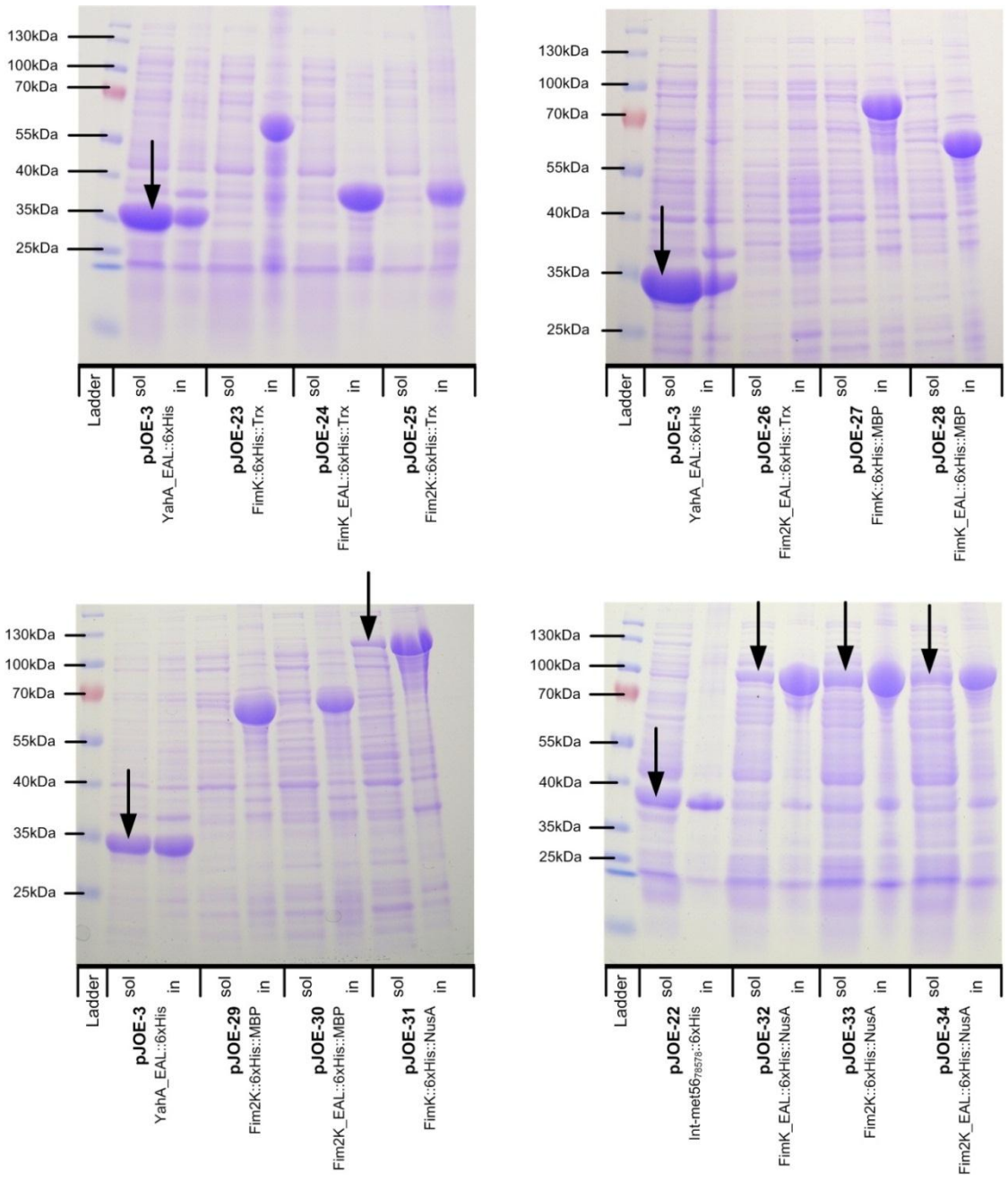
Figure 6-6 :- Defining the domain boundaries for FimK_EAL and Fim2K_EAL



Domain boundaries were defined using Pfam (120) and were extended to encompass the entire secondary structure of the domain region, as predicted by PSIPRED (43). FimK_EAL comprises FimK residues 204 (E – proline) to 449 (R – arginine). Similarly, Fim2K_EAL comprises Fim2K residues 14 (D – aspartic acid) to 266 (E – glutamic acid). Predicted secondary structure (Pred) is either helical (H), coiled (C) or stranded (E). AA, amino acid.

As the thioredoxin, MBP, and NusA tags were predicted to, in general, slightly improve solubility of the proteins under examination, pETM-22, pETM-44 and pETM-66 derivatives harbouring *fimK*, *fimK_EAL*, *fim2K* and *fim2K_EAL* fragments were constructed (Table 6-2 and Appendix 6 for details). These derivatives were expressed in BL21(DE3) (0.1 mM IPTG, 25ºC, overnight) and the solubility of recombinant protein products was examined. No soluble products could be detected for proteins fused to thioredoxin or MBP. Nevertheless, the soluble protein fractions of BL21(DE3) expressing pJOE-31, -32, -33 and -34 possessed minor bands that corresponded in size to FimK::6xHis::NusA, FimK_EAL::6xHis::NusA, Fim2K::6xHis::NusA and Fim2K_EAL::6xHis::NusA, respectively (see black arrows in Figure 6-7). Due to time constraints, the identities of these soluble products were not examined any further.

Figure 6-7 :- Examining the solubility of recombinant FimK, FimK_EAL, Fim2K and Fim2K_EAL fused to 6xHis and thioredoxin, MBP or NusA



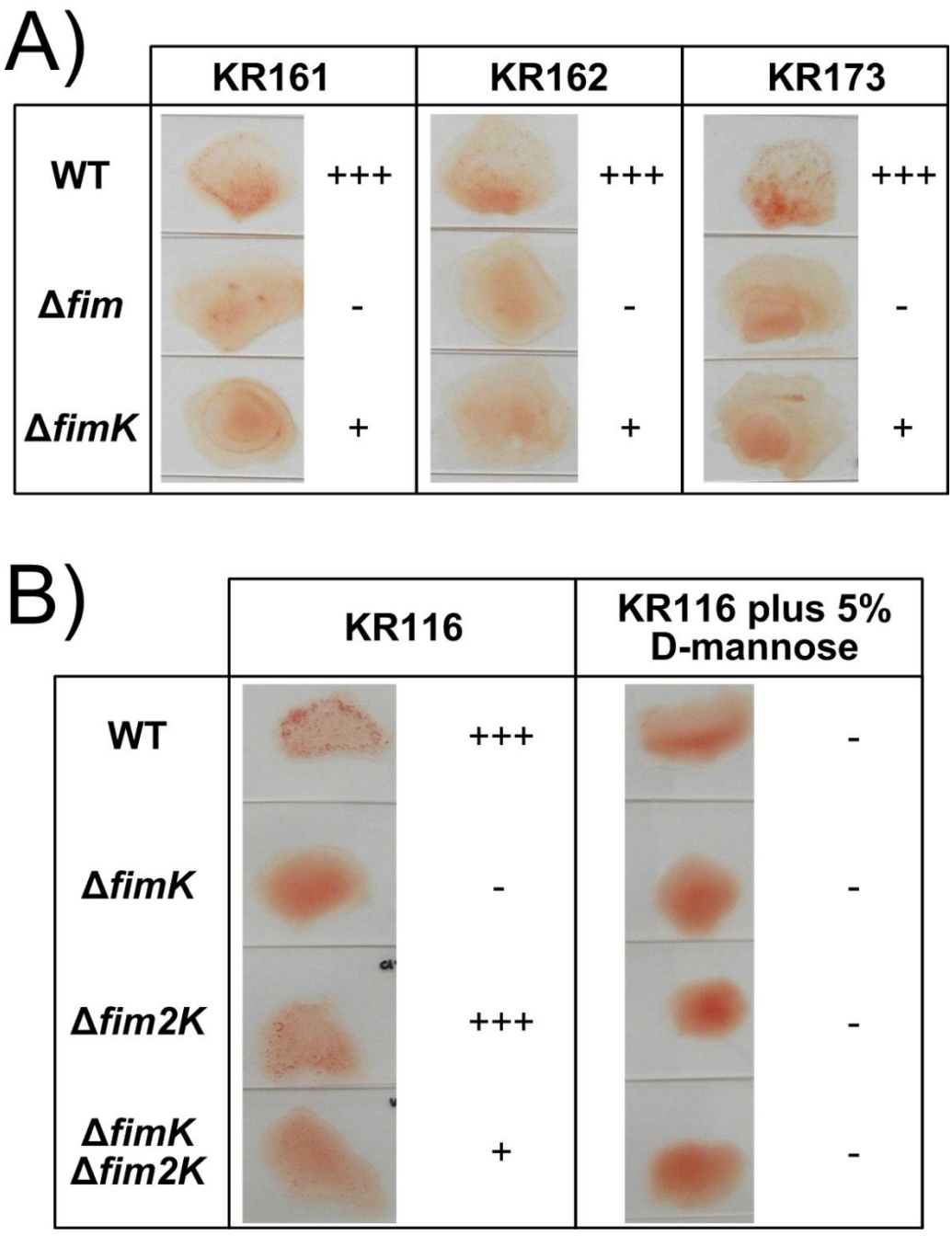
Plasmids pJOE-23 to pJOE-34 were expressed in BL21(DE3) (0.1 mM IPTG, 25ºC, overnight) and the solubility of produced recombinant proteins was analysed. Plasmids pJOE-3 (producing soluble YahA_EAL::6xHis) and pJOE-22 (producing soluble Int-met56$_{78578}$::6xHis) were expressed in parallel as methodological controls for expression, sonication and solubility fractionation. Black arrows highlight recombinant proteins within soluble fractions. Of note, no product was produced in the *E. coli* strain harbouring pJOE-26; this was not investigated further.

6.2.4 Deletion of *fimK*, but not *fim2K*, affects the ability to agglutinate GP-RBC

Expression of the *K. pneumoniae* TOP52 *fimK* gene was previously shown to reduce the surface expression of type 1 fimbriae (331). The effect of *fimK* and/or *fim2K* on type 1 fimbrial expression was therefore also investigated in *K. pneumoniae* strains KR161, KR162, KR173 and KR116. Compared to their wildtype parents, the ability of KR161Δ*fimK*, KR162Δ*fimK* and KR173Δ*fimK* to agglutinate GP-RBCs was decreased, as observed in duplicate agglutination reactions from triplicate overnight cultures at 37ºC and 200 rpm (Figure 6-8). Agglutination in all strains was completely inhibited when 5 % D-mannose was added to the reaction (data not shown), confirming the involvement of the mannose-sensitive type 1 fimbriae. These results suggest that FimK is required for high-level type 1 fimbrial expression and are contradictory to earlier reports using strain TOP52Δ*fimK* (331). However, the results presented here are in agreement with findings in *K. pneumoniae* C3091 and its associated *fimK* knockout (Dr. C. Struve, unpublished results). Similarly, the *fimK* knockout of KR116 (*fim*- and *fim2*-positive) also possessed reduced ability to agglutinate GP-RBCs, as did the double knockout strain KR116Δ*fim2K*Δ*fimK*. Again, this confirms the role of *fimK* in type 1 fimbrial expression. In contrast, KR116Δ*fim2K* was not attenuated in its ability to agglutinate GP-RBCs, potentially a function of the low level expression of *fim2* and therefore *fim2K* in 37ºC LB cultures.

Attempts were made using dilutions of overnight *K. pneumoniae* cultures to quantify the ability of strains to agglutinate GP-RBCs and minimize observer bias, as has been described for *E. coli* (174). These were unsuccessful as the agglutination strength of wildtype *K. pneumoniae* strains was much weaker than that observed with *E. coli* (data not shown).

Figure 6-8 :- Investigation into the ability of KR161, KR162, KR173, KR116 and selected knockout strains to agglutinate GP-RBCs

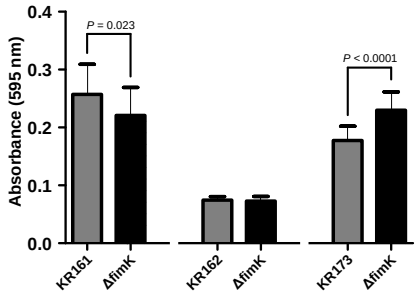


A) Overnight cultures (LB, 37ºC, 200 rpm) of KR161, KR162, KR173 and their Δ*fim* and Δ*fimK* derivatives were mixed with an equal volume of GP-RBCs on glass slides, gently rocked from side-to-side and observed for 15 min to identify the presence and strength of an agglutination phenotype. Whilst the wildtype strains agglutinated GP-RBCs strongly, *fimK* knockouts weakly agglutinated GP-RBCs. As expected, the Δ*fim* strains did not agglutinate GP-RBCs. The phenotype was D-mannose sensitive (data not shown). **B)** The agglutination phenotypes of KR116 and its Δ*fimK*, Δ*fim2K* and Δ*fimK*Δ*fim2K* derivatives were examined similarly. Whilst the wildtype and *fimK* knockout possessed identical GP-RBC agglutination phenotypes, that observed in the Δ*fimK* and Δ*fimK*Δ*fim2K* strains was markedly reduced. WT , wildtype ; +++ , strong agglutination ; + , weak agglutination ; - , no agglutination observed.

6.2.5 The effect of *fimK* and *fim2K* on biofilm formation

Biofilm formation in *K. pneumoniae* hypothesized to be mediated by interactions between the polysaccharide capsule, type 1 fimbriae and type 3 fimbriae, together with other known and unknown factors (31, 349, 352, 388, 391). It has previously been reported that deletion of *fimK* in *K. pneumoniae* TOP52 encourages biofilm formation (331). However, since discrepant results were observed in the previous section, the biofilm forming ability of wildtype KR161, KR162 and KR173 strains and their isogenic knockout mutants was assessed after 24 h of growth at 37ºC in 96-well polystyrene plates in LB medium. In the highly encapsulated strain KR161 loss of *fimK* significantly decreased biofilm formation ($P = 0.023$), whilst the opposite was observed with KR173 and KR173Δ*fimK* ($P < 0.001$) (Figure 6-9). Intriguingly, loss of *fimK* did not affect biofilm formation in KR162, a second thickly encapsulated strain. Of note, the capsule serotype of KR161 and KR173 could not be defined by a PCR assay that covered six K serotypes, whilst that of KR162 was determined to be serotype K2 (data not shown).

Figure 6-9 :- Alterations in biofilm formation phenotype are different in three distinct *K. pneumoniae* strains



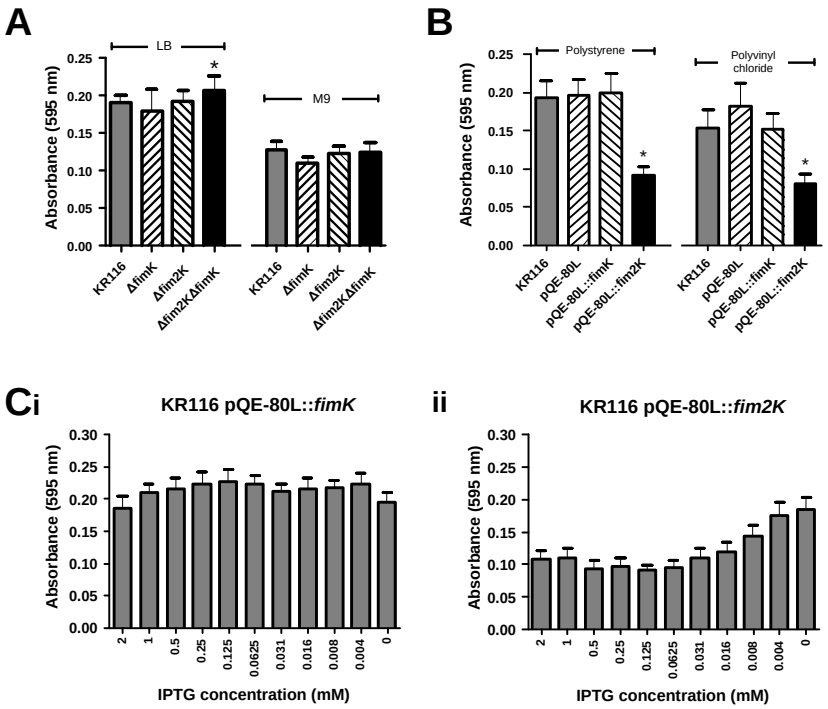
Assay results for biofilm formation on polystyrene (37ºC, LB, 24 h) surfaces for KR161, KR162 and KR173 and their isogenic *fimK* mutants as determined by crystal violet absorbance data. Data shown in all cases represent means and standard deviations of three biological replicates, each assayed in eight wells (n = 24). Statistical analyses were performed using the non-parametric Kruskal-Wallis test and Dunn's posthoc analysis.

A similar biofilm assay was performed using KR116, KR116Δ*fimK*, KR116Δ*fim2K* and, KR116Δ*fim2K*Δ*fimK*. Biofilm formation in both LB and M9 media was similar in the *fimK*

and *fim2K* knockouts and wildtype KR116 strain (Figure 6-10A). However, when both genes were disrupted biofilm formation in LB, but not in M9, was significantly increased ($P < 0.001$). Similar observations were made when the assay was performed at 30ºC. It is important to note that whilst KR116Δ*fimK* biofilm formation appeared slightly decreased, the significance of this finding is unclear. Firstly, KR116Δ*fimK* colonies were smaller in size than those of KR116. Secondly, the strain grew to a lower overnight culture OD_{600} than the wildtype KR116 strain, although similar bacterial counts per ml of culture were obtained. Intriguingly, this was observed in three independent colonies from the same allelic exchange experiment. Lastly, these features could not be reversed by complementing with pQE-80L::*fimK* (data not shown). Further investigations are required to identify whether these observations are related to the disruption of *fimK* or a non-related mutation/change.

Since the basal level of *fim2* expression is low (see section 6.2.5) and as *fim2K* is part of the *fim2* operon, it is likely that *fim2K* is expressed at low levels as well. As a result, obvious phenotypic changes may have been difficult to observe. Therefore, the KR116 *fim2K* gene was overexpressed in KR116 using pQE-80L::*fim2K* and the effect on biofilm formation investigated (37ºC, 24 h, 0.05 mM IPTG). Similarly, to clarify the minor reduction in biofilm formation observed with KR116Δ*fimK, fimK* was also overexpressed using pQE-80L::*fimK.* Whilst KR116/pQE-80L::*fimK* possessed similar biofilm densities as wildtype KR116 and KR116/pQE-80L empty vector control, expression of *fim2K* resulted in a significant reduction in biofilm formation ($P < 0.0001$) (Figure 6-10B). Similar effects were observed on both polyvinyl chloride and polystyrene surfaces and did not alter when the incubation temperature was 30ºC (data not shown).

Figure 6-10 :- The effect of *fimK* and *fim2K* expression on biofilm formation in *K. pneumoniae* KR116



A) Assay results for biofilm formation on polystyrene (37ºC, 24 h) surfaces for KR116 and its three *fimK* and/or *fim2K* isogenic mutants grown in LB or M9 medium as determined by crystal violet absorbance data. **B)** Results from biofilm formation assay on two abiotic surfaces that is based on overexpression of *fimK* and *fim2K* in KR116. KR116 and KR116 carrying an empty pQE-80L vector served as controls (37ºC, 24 h, 0.05 mM IPTG). **C)** Investigation into the effect of IPTG concentration, and therefore FimK or Fim2K concentration, on biofilm formation in KR116 overexpressing *fimK* (i) and *fim2K* (ii) (37ºC, 24 h). Of note, wildtype KR116 and KR116 bearing pQE-80L possessed similar IPTG dilution series biofilm profiles to that shown for KR116 pQE-80L::*fimK*. An asterisk indicates a highly significant difference ($P < 0.001$). Data shown in all cases represent means and standard deviations of three biological replicates, each assayed in eight wells ($n = 24$). Statistical analyses were performed using the non-parametric Kruskal-Wallis test and Dunn's posthoc analysis.

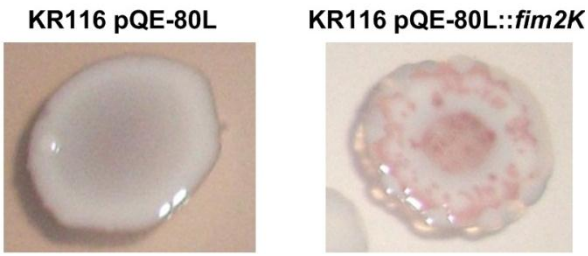
Finally, biofilm formation on polystyrene surfaces was assayed using a two-fold IPTG dilution series in LB media (37ºC, 24 h). This was to investigate whether the effects of *fimK* and *fim2K* varied with IPTG concentration, which itself would be expected to correlate with FimK and Fim2K concentration. As IPTG concentration increased, the biofilm density of KR116/pQE-80L::*fim2K* decreased and was reduced to a minimum at concentrations higher than 0.0625 mM IPTG (Figure 6-10Ci). Such an effect was not observed with KR116/pQE-80L::*fimK* (Figure 6-10Cii), whose IPTG dilution series

biofilm formation profile was similar to that observed for the wildtype and pQE-80L vector-only control (data not shown). Moreover, expression of *fimK* and *fim2K* did not significantly alter culture optical density, regardless of the IPTG concentration (data not shown). The results presented in this section suggest that *fimK* is involved in regulating biofilm formation, but that the effect is strain-dependent. These results also demonstrate that overexpression of *fim2K* in KR116 reduces in a dose-dependent manner the biofilm forming capacity of the recombinant strain.

6.2.6 Expression of *fim2K*, but not *fimK*, increases exopolysaccharide production in *K. pneumoniae* KR116

Several EAL and GGDEF domain-encoding proteins influence biofilm formation by altering the production of exopolysaccharide components (341, 424). To determine whether *fimK* and *fim2K* affect the biosynthesis of these components, strains were plated on congo red medium which is able to stain glucose-rich exopolysaccharides, such as cellulose (242). Strains KR116, KR116Δ*fimK*, KR116Δ*fim2K*, KR116Δ*fim2K*Δ*fimK*, KR116/pQE-80L and KR116/pQE-80L::*fimK* displayed similar smooth colony morphotypes in the presence and absence of IPTG. On the other hand, KR116/pQE-80L::*fim2K* displayed a pigmented and more rugged colony morphotype on congo red medium consistent with greater exopolysaccharide production (340, 341). This phenotype was IPTG-dependent.

Figure 6-11 :- Exopolysaccharide production is increased in *K. pneumoniae* KR116 expressing *fim2K*, but not *fimK*



When grown on congo red medium in the presence and absence of IPTG, strain KR116, KR116Δ*fimK*, KR116Δ*fim2K*, KR116Δ*fim2K*Δ*fimK*, KR116/pQE-80L and KR116/pQE-80L::*fimK* displayed similar smooth colony morphotypes. Whilst this morphotype was also observed when KR116/pQE-80L::*fim2K* was plated on plain congo red medium, when plates were supplemented with 0.1 mM IPTG an alternative morphotype was observed: the colony was more rugged and pigmented.

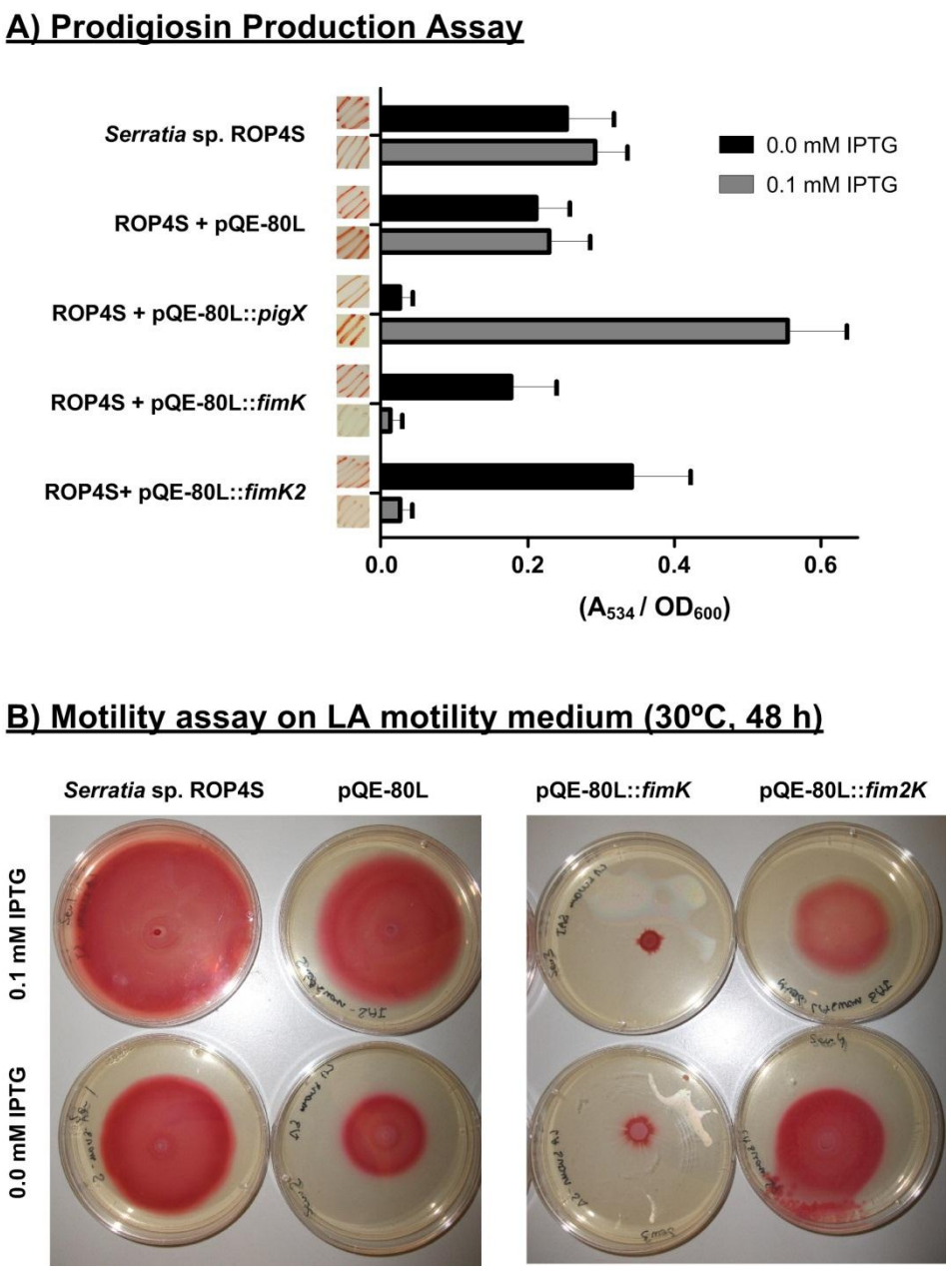
6.2.7 Heterologous expression of *fimK* and *fim2K* in *Serratia* sp. ATCC 39006 strain ROP4S reduces prodigiosin production and decreases motility

The synthesis of prodigiosin, a red coloured secondary metabolite, in *Serratia* sp. ATCC 39006 is controlled by the EAL domain component of the GGDEF-EAL domain-containing protein PigX (118). To determine whether *fimK* and/or *fim2K* possessed EAL domains that could influence prodigiosin synthesis, the *pigX* knockout strain *Serratia* sp. ROP4S was transformed with pQE-80L, pQE-80L::*pigX*, pQE-80L::*fimK* and pQE-80L::*fim2K* and prodigiosin levels were measured. In the absence of IPTG, prodigiosin production in ROP4S/pQE-80L::*fimK* appeared to be reduced although this was not significant (P = 0.310) (Figure 6-12A). Unexpectedly, levels were slightly increased in ROP4S/pQE-80L::*fim2K* (P = 0.015). When expression was induced using 0.1 mM IPTG, corresponding predicted increased levels of Fim2K and FimK caused a seven (P = 0.0022) and fifteen-fold (P = 0.0048) reduction in prodigiosin levels, respectively, compared to the empty vector control strain ROP4S/pQE-80L (Figure 6-12A). As reported previously, when IPTG was not present in cultures of ROP4S/pQE-80L::*pigX* prodigiosin synthesis was decreased compared to

ROP4S/pQE-80L (*P* = 0.0022) (118). However, when higher levels of PigX were produced by inducing expression with 0.1 mM IPTG, prodigiosin levels increased 3.1-fold compared to the vector-only control strain (*P* = 0.0022), strongly suggestive of a dose-dependent PigX-mediated effect on prodigiosin synthesis (Figure 6-12A). Of note, it was also observed that IPTG-induced expression of the full length PigX protein in *E. coli* using pQE-80L::*pigX* was inhibitory to growth, most likely due to the presence of two transmembrane helices (118). This has previously been observed in the PigX homolog CsrD (396).

It has previously been shown that PigX is unable to alter the swimming motility of *Serratia* sp. ATCC 39006 (440). As *K. pneumoniae* strains are non-motile, the effect of *fimK* and *fim2K* on swimming motility was investigated in *Serratia* sp. ROP4S. When *fimK* was heterologously expressed in ROP4S a reduction in swimming motility was observed after 48 h of incubation at 30ºC, as compared to the wildtype and vector-only control (Figure 6-12). *fim2K* affected swimming motility in a more subtle manner; whilst the presence of IPTG greatly enhanced the motility of ROP4S and ROP4S/pQE-80L, this increase was not observed in ROP4S/pQE-80L::*fim2K*, which also possessed a slightly reduced swimming diameter. It must be noted that, although interesting, these are preliminary experiments, and that further replicates in addition to accurate diameter measurements are required to confirm these observations.

Figure 6-12 :- Phenotypic effects of *fimK* and *fim2K* expression in *Serratia* sp. ATCC 39006



A) Prodigiosin levels were measured in cultures of *Serratia* sp. ROP4S, ROP4S/pQE-80L, ROP4S/pQE-80L::*pigX*, ROP4S/pQE-80L::*fimK* and ROP4S/pQE-80L::*fim2K* by growing cultures for 24 h at 30ºC in LB media (200 rpm) supplemented with either 0.0 or 0.1 mM IPTG and ampicillin/spectinomycin, when necessary. Three prodigiosin extractions were performed on two independent cultures. The amount of prodigiosin produced was also examined by visual comparison on solid media and is presented adjacent to the graph Y-axis. Genes *pigX*, *fimK* and *fim2K* influence prodigiosin production. **B)** Motility assays were performed in duplicate at 30ºC on LA motility medium. When grown in the presence of IPTG, ROP4S/pQE-80L::*fimK* and ROP4S/pQE-80L::*fim2K* possessed decreased motility as compared to the control strains; the effect was more marked in the *fimK*-bearing as compared to the *fim2K*-bearing strain (see text).

6.3 Discussion

EAL domain-containing proteins can change intracellular c-di-GMP concentrations directly by hydrolysing c-di-GMP or indirectly by influencing other DGCs and PDEs (160). It has therefore been recommended to only assign PDE function to EAL domain-containing proteins that have been purified and biochemically identified to be active *in vitro* (160). Initial attempts to heterologously express and purify FimK and Fim2K identified that both proteins were highly insoluble regardless of induction strength, duration and temperature. Recombinant protein insolubility is a frequent bottleneck in functional experiments. In fact, approximately half of all recombinant bacterial proteins expressed in *E. coli* are insoluble (70). Protein solubility can be improved by creating fusions with highly soluble 'carrier' proteins, although the mechanism of action of these carriers is still unclear (192, 218, 373). In an attempt to improve the solubility of FimK and Fim2K, both the full proteins and EAL domain-encoding portions of FimK and Fim2K were expressed as fusions to the solubility-enhancing proteins GST, thioredoxin, MBP and NusA. Despite testing several fusion proteins only those fused to NusA were produced as soluble protein, and even then only at low concentrations. Unfortunately, due to time constrains the purification and biochemical characterisation of the FimK and Fim2K NusA fusions was not pursued.

Biofilms are three dimensional aggregates of bacteria enmeshed in an extracellular matrix composed of exopolysaccharides, water, proteins, nucleic acids, lipids, nutrients and metabolites (11, 25). These multi-cellular communities display increased resistance to stress, antibiotics and the host immune response, properties which cause considerable problems in medical and industrial venues (11). In few bacterial species, switching between planktonic and sessile lifestyles has been shown to be triggered by c-di-GMP levels, and several PDEs and DGCs involved in controlling biofilm formation have been characterised (25, 160). Therefore, the role of FimK and Fim2K in biofilm formation was investigated.

FimK influenced biofilm formation on polystyrene surfaces in a strain-specific manner: whilst no changes were observed in KR116 and KR162, biofilm formation was increased in KR173 but decreased in KR161. This is unlikely to be related to the effect of *fimK* on type 1 fimbrial expression, which was decreased in all four knockout strains. Additionally, *K. pneumoniae* type 1 fimbriae do not mediate biofilm formation on polystyrene surfaces (352, 387). It is speculated that FimK may regulate several genes involved in making biofilms, and that these regulon gene complements differ between strains of *K. pneumoniae* due to strain-specific flexible genome repertoires.

Potential FimK regulon candidate genes include those involved in synthesis of the polysaccharide capsule, type 3 fimbriae, Kpc fimbriae, the presently hypothesised Fim2 cell appendages and the novel putative cell surface protein KPN_00994 (352, 384, 388, 443). Expression of transcriptional regulators such as OxyR and MrkI could also potentially be subject to FimK control (161, 184). When *fimK* is interrupted the surface expression of type 3 fimbriae is also altered in a strain-specific manner: in KR161 expression is increased, whilst in KR173 it is decreased and in KR162 it is unaltered (Steen Stahlhut, personal communication). This is inconsistent with the biofilm assay results, and suggests that FimK also regulates a number of other genes involved in *K. pneumoniae* biofilm formation. The observed discrepancies may also reflect differences in the regulation and structure of the different K serotypes represented amongst the strains examined. As the capsule has been shown to mask the function of type 1 fimbriae and is hypothesised to mask those of other surface organelles, capsule-specific differences may alter the shielding effect observed when the FimK regulon is activated hence altering biofilm production (349, 391). Additionally, the quantity of capsule produced may influence biofilm structure and density, as described previously (18, 31). Together, the factors highlighted in this paragraph may explain why opposite phenotypes were observed in the Rosen *et al.* study as compared to the results presented in this chapter (331).

Biofilm formation decreased significantly when *fim2K* was overexpressed in KR116, the magnitude of which was IPTG concentration-dependent. It is likely that biofilm formation was reduced due to Fim2K-mediated regulation of type 3 fimbriae, which exert a strong positive effect on biofilm formation, or other biofilm-associated genes (18, 19, 31, 384). It is plausible that as adhesion between bacteria within biofilms decreases, biofilm structure weakens and bacterial shedding increases enabling the spread of bacteria to foreign sites, such as the kidneys and liver. This hypothesis is consistent with the *fim2 in vivo* data presented in Chapter 5. Intriguingly, Fim2K also increased exopolysaccharide formation, which has been associated with increased biofilm forming capability (395, 424). This is paradoxical to the biofilm results presented in this chapter. Potentially, *fim2K* expression may encourage the production of an exopolysaccharide with anti-biofilm activity, such as the A101 exopolysaccharide described in *Vibrio* sp. QY101 (181, 326, 456). However, it is more likely that the increased congo red staining in KR116/pQE-80L::*fim2K* resulted from a decrease in the production of capsule and/or lipopolysaccharide, which act as protective barriers against congo red staining (51).

Serratia sp. ATCC 39006 synthesises prodigiosin, a red coloured secondary metabolite with antimicrobial, immunosuppressive and anticancer properties (118, 440). Regulation of the prodigiosin biosynthetic pathway is complex and includes the GGDEF and EAL domain-containing protein PigX (118, 119, 440). When *pigX* was interrupted prodigiosin synthesis increased 350 %, a phenotype specifically mediated by the EAL domain component (118). It was concluded that this was a c-di-GMP dependent effect as the increase in prodigiosin synthesis could be reversed by complementing with the active PDE YahA. In this chapter it was determined whether FimK and Fim2K could influence prodigiosin production, and hence potentially intracellular c-di-GMP concentration, when their cognate genes were overexpressed in the *pigX* knockout strain *Serratia* sp. ROP4S. As expected, IPTG-induced expression of *fimK* and *fim2K*

decreased prodigiosin production, confirming that FimK and Fim2K were functional proteins with significant biological function(s). No effect on prodigiosin synthesis was observed when IPTG was omitted from the ROP4S/pQE-80L::*fimK* cultures whilst, paradoxically, prodigiosin production in ROP4S/pQE-80L::*fim2K* was slightly increased. It is possible that at low concentrations Fim2K encourages prodigiosin synthesis. Alternatively minor differences between batches of rich culture media can differ, potentially changing the levels of substances known to alter prodigiosin synthesis including glucose, salt, phosphate and iron (439). However, the latter cannot explain the stark differences that were observed in prodigiosin synthesis when ROP4S/pQE-80L::*pigX* was cultured with 0.1 mM compared to without IPTG.

It is hypothesized that PigX may exert its influence in a similar manner to CsrD, although a similar dose-dependent effect has not been observed previously (396). At low concentrations of PigX, the EAL domain activity may predominate leading to CsrB and CsrC degradation, two RNA molecules which sequester CsrA. This results in an increase in CsrA activity, and repression of prodigiosin synthesis (12). As the concentration of PigX increases, the likelihood of GGDEF domain-mediated PigX dimerization increases which may physically hinder or cause conformational changes that reduces ability of the PigX EAL domain to mediate degradation of CsrB and CsrC (12, 396). These RNAs may then accumulate, and bind to and sequester CsrA, which subsequently increases prodigiosin synthesis levels. This hypothesis is highly speculative and requires further experimentation.

It is unlikely that FimK and Fim2K use a similar mechanism to PigX as their EAL domains are less degenerate and possess nine out of the ten residues required for c-di-GMP hydrolysis. Additionally, both proteins influence swimming motility whilst PigX does not, suggesting that the phenotypic influence of PigX is more limited than that of FimK and Fim2K. However, it must be noted that although the anti-swimming effects of FimK are clear, the phenotype observed with Fim2K is more subtle and requires further

experimentation. The ability of EAL domains to regulate motility and/or biofilm formation has been observed in many PDEs, including the *P. aeruginosa* EAL protein BifA, and is strongly supportive of a role for FimK and Fim2K within the c-di-GMP regulatory network (160, 214, 365).

Of note, FimK appeared to exert a stronger influence on prodigiosin synthesis and motility than Fim2K, assuming that the concentrations of Fim2K and FimK were equivalent at the same concentration of IPTG. It is possible that the EAL domain of FimK is more efficient at hydrolysing c-di-GMP than Fim2K (320, 350). The inability of FimK to regulate biofilm formation in strain KR116 could be explained by post-translational regulation of FimK, which may be inactive due to the lack of an 'activator' molecule or the presence of a 'repressor' molecule. It is likely that this molecule interacts with the N-terminal region of FimK, which is most closely related to the GerE domain family, which consist of helix-turn-helix domains. Based on similarities with GerE domain family members, FimK could be activated (or inactivated) by a two-component sensory transduction system, quorum sensing molecules or ligands (133). An unknown 'activator' could therefore bind to the N-terminal of FimK, cause a conformational change and activate the FimK EAL domain. Further work is required to investigate the nature of this N-terminal domain and its interaction with the FimK EAL domain.

6.3.1 Future work

Despite the extensive work carried out in this chapter, several questions still remain. Firstly, it is still not known whether FimK and Fim2K are able to alter c-di-GMP concentration directly or do so indirectly by regulating PDEs and potentially DGCs. Next, although some phenotypic effects of FimK and Fim2K on KR116 have been identified, the basis by which these proteins exert their influence is still unknown. Methods for examining some of these queries are presented below:

Expression of fimK *and* fimK2: *concluding experiments*

To confirm the phenotypes observed when *fimK* and *fim2K* were expressed in KR116, biofilm formation and exopolysaccharide production should also be examined in KR161, KR162 and KR173 expressing *fimK* and *fim2K*.

Purification and biochemical characterisation of FimK and Fim2K

The 6xHis::NusA fusions of FimK, FimK_EAL, Fim2K and Fim2K_EAL can be purified using affinity purification and subsequently investigated for PDE activity, as described previously (20, 351, 400, 422).

Site directed mutagenesis of FimK and Fim2K

The conserved residues required for c-di-GMP hydrolytic activity that were identified in this chapter would be ideal targets for site directed mutagenesis. The effects of amino acid substitutions could then be investigated using either phenotypic, biochemical assays and/or by directly measuring cell c-di-GMP concentration.

Investigate the FimK/Fim2K regulon

This can be examined using genetic and/or proteomic approaches to compare the wildtype KR116 strain to those overexpressing FimK and/or Fim2K, or the *fimK* and/or *fim2K* knockouts. Genetic techniques such as microarrays or RNA-Seq analysis could be used to characterise strain transcriptomes (291, 298, 433). Alternatively, protein-based techniques such as 2D-DiGE or mass spectroscopy could be employed (150, 291).

Examine the effect of FimK and Fim2K on virulence

This can initially be investigated using the *K. pneumoniae fimK* and/or *fim2K* knockouts in individual and/or competition-based infection assays of *Galleria mellonella* larvae, *Caenorhabditis elegans* or *Dictyostelium* and extended to murine models (157, 263, 286). Note: KR116Δ*fimK* will need to be reconstructed to confirm the observed small colony phenotype is a direct result of the *fimK* deletion and not some other mutation.

Examine the nature of other K. pneumoniae *EAL domain-encoding proteins*

Based on the available sequenced genomes, it is likely that a typical *K. pneumoniae* strain encodes more than fifteen different EAL proteins (1). Using the techniques described in this thesis, the corresponding genes can be disrupted and/or overexpressed and the phenotypic result investigated. Large scale analysis of all EAL and GGDEF domain-encoding proteins has previously been described for *Salmonella* and *Clostridium difficile* (32, 379). Plasmids have already been constructed for expression of the MGH78578 EAL protein KPN_00268 and the CsrD / PigX relative KPN_03660 (Appendix 6).

Chapter 7. Conclusion

This thesis has examined in detail the genetic manipulation of clinical *K. pneumoniae* isolates using mini-Tn*7* insertions, suicide plasmid- and lambda Red-based allelic exchange, and FRT recombination. These techniques have been essential to understanding the function of genes, operons and larger regions of DNA in many Gram negative bacteria, and it is highly likely that the recombineering methods described here will also enable the detailed exploration and characterisation of such regions in *K. pneumoniae* and related species. The techniques examined will also facilitate functional investigations into regions of interest uncovered by the plethora of sequence data that has been obtained from whole genome sequencing projects (124, 284). The applications of these techniques and tools are vast and only limited by imagination. As highlighted by the study of *fim2* and *fim2K* in this work, these techniques could be used to study putative virulence factors, thus revealing insights into functions and associated phenotype(s) linked with these determinants, and allowing for robust molecular Koch's postulates type examination of hypotheses (112). Alternatively, genetic recombineering may be used to produce novel strains or improve those currently used in industrial applications and/or bioremediation (398, 453).

This thesis contains the first description of the *met56* tRNA-associated KpGI-5 genomic island, which was sequenced and identified to harbour a putative fimbrial gene cluster and code for several hypothetical proteins. This island belongs to a much larger family of KpGI-5-like islands which occur in 10 – 15 % of *K. pneumoniae* and *K. oxytoca* strains, two of the most clinically relevant *Klebsiella* spp. (307). In time, the KpGI-5 island family may become a genomic marker for epidemiological studies and/or diagnostics. Indeed, bacterial pathogen profiling, which is the process of consolidating proteomics, metabolomics, transcriptomics and genomics data, is currently being explored for use in disease management and/or surveillance (6, 369). As a speculative

example, the *fim2* locus may function as a marker indicating increased ability of host bacteria to disseminate and form biofilms, thus potentially warranting added clinical attention when a patient is infected by a *Klebsiella* strain with *fim2*-positive profile.

This thesis has also described the first *in vitro* and *in vivo* analysis of *fim2*, the KpGI-5-associated novel γ_1-type CU fimbrial operon. Whilst *fim2*-specific mRNA is detectable at low levels under diverse growth conditions, *fim2* expression appeared to be about nine-fold upregulated in bacteria isolated from murine bladders. This finding, coupled with other observations stemming from the *in vivo* infection model studies, suggests a possible role for *fim2* in promoting dissemination of bacteria from the bladder and lungs into the kidneys and liver, respectively. If so, in future, the protein products of *fim2* may serve as targets for antivirulence therapies, which aim to disrupt virulence properties of bacteria without a direct impact on bacterial viability (57). In comparison to traditional antibiotics, these therapies are thought to reduce the evolutionary pressure on bacteria to develop resistance and may also minimize perturbations of the host microbiota (57, 370). Antivirulence molecules have been developed that inhibit the assembly of *E. coli* P pili and type 1 fimbriae, resulting in decreased bladder cell adherence and disrupted biofilm formation (4, 297). These molecules are called pilicides and bind to conserved regions of the fimbrial assembly chaperone proteins, and thus may also be applicable to the closely related *K. pneumoniae* type 1 fimbriae and putative Fim2 fimbriae (4).

Finally, the putative phosphodiesterases FimK and Fim2K were identified to regulate c -di-GMP dependent phenotypes, including motility and prodigiosin synthesis. Whilst FimK also regulated type 1 fimbrial production, Fim2K altered the production of exopolysaccharides and formation of biofilms. As these comprise virulence traits, it is thought that proteins involved in the regulation of intracellular c-di-GMP concentration, such as FimK and Fim2K, are prime candidates for antivirulence therapies, although research into this is still in its infancy (370). More immediately, these findings represent a starting point to enable further dissection of how EAL and GGDEF domains are

involved in the regulation of *K. pneumoniae* fimbrial expression, biofilm formation, exopolysaccharide production and other virulence factors, to enhance our understanding of the c-di-GMP signal transduction pathway and its involvement in the life cycle of *K. pneumoniae*.

In conclusion, based on the findings presented in this work and on those drawn from extensive investigations into fimbrial systems, it is hypothesized that specific environmental and/or host conditions trigger the expression of *fim2*. As levels of *fim2K* increase, *K. pneumoniae* biofilms weaken and bacterial dissemination increases. These planktonic cells express the putative Fim2 structures on their surfaces, which may subsequently mediate adherence to and colonization of new surfaces, such as those present in the liver and kidney. Based on this hypothesis, *fim2*-positive strains may have a niche-specific evolutionary advantage over those that lack *fim2*. This thesis is the first report of *fim2* and it will guide further investigations into the full impact of this enigmatic operon.

References

1. **Aartsen J. J. van**. 2007. Identifying and analysing novel genomic sequences present in *Klebsiella* species that cause human blood stream and urinary tract infections. BSc Thesis.

2. **Aartsen J. J. van**. 2008. The *Klebsiella pheV* tRNA locus: a hotspot for integration of alien genomic islands. Bioscience Horizons **1**:51-60.

3. **Aartsen J. J. van, and K. Rajakumar**. 2011. An optimized method for suicide vector-based allelic exchange in *Klebsiella pneumoniae*. J. Microbiol. Methods **86**:313-319.

4. **Aberg V., and F. Almqvist**. 2007. Pilicides-small molecules targeting bacterial virulence. Org. Biomol. Chem. **5**:1827-34.

5. **Abraham J. M., C. S. Freitag, J. R. Clements, and B. I. Eisenstein**. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **82**:5724-7.

6. **Ahmed N., U. Dobrindt, J. Hacker, and S. E. Hasnain**. 2008. Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. Nat. Rev. Microbiol. **6**:387-394.

7. **Al-Hasani K., K. Rajakumar, D. Bulach, R. Robins-Browne, B. Adler, and H. Sakellaris**. 2001. Genetic organization of the *she* pathogenicity island in *Shigella flexneri* 2a. Microb. Pathog. **30**:1-8.

8. **Allen B. L., G. F. Gerlach, and S. Clegg**. 1991. Nucleotide sequence and functions of *mrk* determinants necessary for expression of type 3 fimbriae in *Klebsiella pneumoniae*. J. Bacteriol. **173**:916-20.

9. **Altschul S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman**. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389-3402.

10. **Amann E., B. Ochs, and K. J. Abel**. 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. Gene **69**:301-15.

11. **Anderson G. G., and G. A. O'Toole**. 2008. Innate and induced resistance mechanisms of bacterial biofilms. Curr. Top. Microbiol. Immunol. **322**:85-105.

12. **Ang S., Y. T. Horng, J. C. Shu, P. C. Soo, J. H. Liu, W. C. Yi, H. C. Lai, K. T. Luh, S. W. Ho, and S. Swift**. 2001. The role of RsmA in the regulation of swarming motility in *Serratia marcescens*. J. Biomed. Sci. **8**:160-9.

13. **Araujo C. De, D. Balestrino, L. Roth, N. Charbonnel, and C. Forestier**. 2010. Quorum sensing affects biofilm formation through lipopolysaccharide synthesis in *Klebsiella pneumoniae*. Res. Microbiol. **161**:595-603.

14. **Ausubel F. M., R. Brent, R. E. Kingston, D. Moore, J. G. Seidman, and K. Struhl**. 2004. Current Protocols in Molecular Biology. John Wiley & Sons.

15. **Bach S., A. de Almeida, and E. Carniel**. 2000. The *Yersinia* high-pathogenicity island is present in different members of the family *Enterobacteriaceae*. FEMS Microbiol. Lett. **183**:289-94.

16. **Bachman M. A., J. E. Oyler, S. H. Burns, M. Caza, F. Lepine, C. M. Dozois, and J. N. Weiser**. 2011. *Klebsiella pneumoniae* Yersiniabactin predisposes to respiratory tract infection through evasion of Lipocalin 2. Infect. Immun. **79**:3309-16.

17. **Bagattini M., V. Crivaro, A. Di Popolo, F. Gentile, A. Scarcella, M. Triassi, P. Villari, and R. Zarrilli**. 2006. Molecular epidemiology of extended-spectrum beta-lactamase-producing *Klebsiella pneumoniae* in a neonatal intensive care unit. J. Antimicrob. Chemother. **57**:979-82.

18. **Balestrino D., J.-M. Ghigo, N. Charbonnel, J. A. J. Haagensen, and C. Forestier**. 2008. The characterization of functions involved in the establishment and maturation of *Klebsiella pneumoniae in vitro* biofilm reveals dual roles for surface exopolysaccharides. Environ. Microbiol. **10**:685-701.

19. **Balestrino D., J. Haagensen, C. Rich, and C. Forestier**. 2005. Characterization of Type 2 Quorum Sensing in *Klebsiella pneumoniae* and Relationship with Biofilm Formation. J. Bacteriol. **187**:2870-2880.

20. **Barends T. R. M., E. Hartmann, J. J. Griese, T. Beitlich, N. V. Kirienko, D. a Ryjenkov, J. Reinstein, R. L. Shoeman, M. Gomelsky, and I. Schlichting**. 2009. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. Nature **459**:1015-8.

21. **Barnes M. R.** 2007. Amino acid properties and consequences of substitutions., p. 554. *In* Bioinformatics for Geneticists. John Wiley and Sons.

22. **Barnhart M. M., and M. R. Chapman**. 2006. Curli biogenesis and function. Annu. Rev. Microbiol. **60**:131-47.

23. **Bascomb S., S. P. Lapage, W. R. Willcox, and M. A. Curtis**. 1971. Numerical classification of the tribe Klebsielleae. J. Gen. Microbiol. **66**:279-95.

24. **Beeckman D. S. A., and D. C. G. Vanrompay**. 2010. Bacterial secretion systems with an emphasis on the chlamydial Type III secretion system. Curr. Issues Mol. Biol. **12**:17-41.

25. **Beloin C., A. Roux, and J. M. Ghigo**. 2008. *Escherichia coli* biofilms. Curr. Top. Microbiol. Immunol. **322**:249-89.

26. **Bertani L. E., and E. W. Six.** 1988. The P2-like phages and their parasite, P4, p. 73-143. *In* R. Calendar (ed.), The Bacteriophages, 2nd ed. Plenum Publishing Corp., New York.

27. **Bikandi J., R. San Millán, A. Rementeria, and J. Garaizar**. 2004. *In silico* analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. Bioinformatics **20**:798-9.

28. **Blomfield I. C., P. J. Calie, K. J. Eberhardt, M. S. McClain, and B. I. Eisenstein**. 1993. Lrp stimulates phase variation of type 1 fimbriation in *Escherichia coli* K-12. J. Bacteriol. **175**:27-36.

29. **Blomfield I. C., M. S. McClain, and B. I. Eisenstein**. 1991. Type 1 fimbriae mutants of *Escherichia coli* K12: characterization of recognized afimbriate strains and construction of new *fim* deletion mutants. Mol. Microbiol. **5**:1439-45.

30. **Blomfield I. C., V. Vaughn, R. F. Rest, and B. I. Eisenstein**. 1991. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. Mol. Microbiol. **5**:1447-57.

31. **Boddicker J. D., R. A. Anderson, J. Jagnow, and S. Clegg**. 2006. Signature-tagged mutagenesis of *Klebsiella pneumoniae* to identify genes that influence biofilm formation on extracellular matrix material. Infect. Immun. **74**:4590-4597.

32. **Bordeleau E., L.-C. Fortier, F. Malouin, and V. Burrus**. 2011. c-di-GMP turn-over in *Clostridium difficile* is controlled by a plethora of diguanylate cyclases and phosphodiesterases. PLoS Genet. **7**:e1002039.

33. **Bouza E., and E. Cercenado**. 2002. *Klebsiella* and *Enterobacter*: antibiotic resistance and treatment implications. Semin. Respir. Infect. **17**:215-30.

34. **Bowden D. W., and P. Modrich**. 1985. *In vitro* maturation of circular bacteriophage P2 DNA. Purification of *ter* components and characterization of the reaction. J. Biol. Chem. **260**:6999-7007.

35. **Boyd E. F., S. Almagro-Moreno, and M. A. Parent**. 2009. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. Trends Microbiol. **17**:47-53.

36. **Boye K., and D. S. Hansen**. 2003. Sequencing of 16S rDNA of Klebsiella: taxonomic relations within the genus and to other Enterobacteriaceae. Int J Med Microbiol, 2003/03/15 ed. **292**:495-503.

37. **Boyer H. W., and D. Roulland-Dussoix**. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459-72.

38. **Braun V.** 2001. Iron uptake mechanisms and their regulation in pathogenic bacteria. Int. J. Med. Microbiol. **291**:67-79.

39. **Brisse S., T. van Himbergen, K. Kusters, and J. Verhoef**. 2004. Development of a rapid identification method for *Klebsiella pneumoniae* phylogenetic groups and analysis of 420 clinical isolates. Clin. Microbiol. Infect. **10**:942-945.

40. **Brisse S., C. Fevre, V. Passet, S. Issenhuth-Jeanjean, R. Tournebize, L. Diancourt, and P. Grimont**. 2009. Virulent clones of *Klebsiella pneumoniae*: identification and evolutionary scenario based on genomic and phenotypic characterization. PLoS One **4**:e4982.

41. **Brisse S., S. Issenhuth-Jeanjean, and P. A. D. Grimont**. 2004. Molecular serotyping of *Klebsiella* species isolates by restriction of the amplified capsular antigen gene cluster. J. Clin. Microbiol. **42**:3388-98.

42. **Brooks T., and C. W. Keevil**. 1997. A simple artificial urine for the growth of urinary pathogens. Lett. Appl. Microbiol. **24**:203-6.

43. **Bryson K., L. J. McGuffin, R. L. Marsden, J. J. Ward, J. S. Sodhi, and D. T. Jones**. 2005. Protein structure prediction servers at University College London. Nucleic Acids Res. **33**:W36-8.

44. **Brüssow H., C. Canchaya, and W.-D. Hardt**. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol. Rev. **68**:560-602.

45. **Bullen J., E. Griffiths, H. Rogers, and G. Ward**. 2000. Sepsis: the critical role of iron. Microbes Infect. **2**:409-15.

46. **Bullen J., H. J. Rogers, P. B. Spalding, and C. G. Ward**. 2006. Natural resistance, iron and infection: a challenge for clinical medicine. J. Med. Microbiol. **55**:251-8.

47. **Burland T. G.** 2000. DNASTAR's Lasergene sequence analysis software. Methods Mol. Biol. **132**:71-91.

48. **Burrus V., and M. K. Waldor**. 2004. Shaping bacterial genomes with integrative and conjugative elements. Res. Microbiol. **155**:376-386.

49. **Cai J. C., H. W. Zhou, R. Zhang, and G.-X. Chen**. 2008. Emergence of *Serratia marcescens*, *Klebsiella pneumoniae*, and *Escherichia coli* Isolates possessing the plasmid-mediated carbapenem-hydrolyzing beta-lactamase KPC-2 in intensive care units of a Chinese hospital. Antimicrob. Agents Chemother. **52**:2014-8.

50. **Campbell A.** 2003. Prophage insertion sites. Res. Microbiol. **154**:277-282.

51. **Camprubí S., S. Merino, J. Benedí, P. Williams, and J. M. Tomás**. 1992. Physicochemical surface properties of *Klebsiella pneumoniae*. Curr. Microbiol. **24**:31-33.

52. **Cano V., D. Moranta, E. Llobet-Brossa, J. A. Bengoechea, and J. Garmendia**. 2009. *Klebsiella pneumoniae* triggers a cytotoxic effect on airway epithelial cells. BMC Microbiol. **9**.

53. **Capitani G., O. Eidam, R. Glockshuber, and M. G. Grutter**. 2006. Structural and functional insights into the assembly of type 1 pili from *Escherichia coli*. Microbes Infect. **8**:2284-2290.

54. **Carpenter J. L.** 1990. *Klebsiella* pulmonary infections: occurrence at one medical center and review. Rev. Infect. Dis. **12**:672-82.

55. **Carter J. S., F. J. Bowden, I. Bastian, G. M. Myers, K. S. Sriprakash, and D. J. Kemp**. 1999. Phylogenetic evidence for reclassification of *Calymmatobacterium granulomatis* as *Klebsiella granulomatis* comb. nov. Int. J. Syst. Bacteriol. **49**:1695-700.

56. **Carter M. Q., J. Chen, and S. Lory**. 2010. The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a novel type IV pilus. J. Bacteriol. **192**:3249-58.

57. **Cegelski L., G. R. Marshall, G. R. Eldridge, and S. J. Hultgren**. 2008. The biology and future prospects of antivirulence therapies. Nat. Rev. Microbiol. **6**:17-27.

58. **Chaveroche M., J. Ghigo, and C. D'Enfert**. 2000. A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. Nucleic Acids Res. **28**:E97.

59. **Chaves J., M. G. Ladona, C. Segura, A. Coira, R. Reig, and C. Ampurdanés**. 2001. SHV-1 beta-lactamase is mainly a chromosomally encoded species-specific enzyme in *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **45**:2856-61.

60. **Chen N., H.-Y. Ou, J. J. van Aartsen, X. Jiang, M. Li, Z. Yang, Q. Wei, X. Chen, X. He, Z. Deng, K. Rajakumar, and Y. Lu**. 2010. The *pheV* phenylalanine tRNA gene in *Klebsiella pneumoniae* clinical isolates is an integration hotspot for possible niche-adaptation genomic islands. Curr. Microbiol. **60**:210-6.

61. **Chen S. L., C. S. Hung, J. Xu, C. S. Reigstad, A. Sabo, V. Magrini, D. Blasiar, T. Bieri, R. R. Meyer, P. Ozersky, J. R. Armstrong, R. S. Fulton, J. P. Latreille, J. Spieth, T. M. Hooton, E. R. Mardis, S. J. Hultgren, and J. I. Gordon**. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. Proc. Natl. Acad. Sci. U.S.A. **103**:5977-5982.

62. **Chen Y. T., H. Y. Chang, Y. C. Lai, C. C. Pan, S. F. Tsai, and H. L. Peng**. 2004. Sequencing and analysis of the large virulence plasmid pLVPK of *Klebsiella pneumoniae* CG43. Gene **337**:189-198.

63. **Chen Y.-T., T.-L. Liao, K.-M. Wu, T.-L. Lauderdale, J.-J. Yan, I.-W. Huang, M.-C. Lu, Y.-C. Lai, Y.-M. Liu, H.-Y. Shu, J.-T. Wang, I.-J. Su, and S.-F. Tsai**. 2009. Genomic diversity of citrate fermentation in *Klebsiella pneumoniae*. BMC Microbiol. **9**:168.

64. **Cheng H. Y., Y. S. Chen, C. Y. Wu, H. Y. Chang, Y. C. Lai, and H. L. Peng**. 2010. RmpA regulation of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* CG43. J. Bacteriol. **192**:3144-58.

65. **Chiou C. S., and a L. Jones**. 1995. Expression and identification of the *strA-strB* gene pair from streptomycin-resistant *Erwinia amylovora*. Gene **152**:47-51.

66. **Choi K. H., J. B. Gaynor, K. G. White, C. Lopez, C. M. Bosio, R. R. Karkhoff-Schweizer, and H. P. Schweizer**. 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat. Methods **2**:443–448.

67. **Choi K. H., and H. P. Schweizer**. 2005. An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. BMC Microbiol. **5**:30.

68. **Choi K.-H., T. Mima, Y. Casart, D. Rholl, A. Kumar, I. R. Beacham, and H. P. Schweizer**. 2008. Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. Appl. Environ. Microbiol. **74**:1064-75.

69. **Chou H. C., C. Z. Lee, L. C. Ma, C. T. Fang, S. C. Chang, and J. T. Wang**. 2004. Isolation of a chromosomal region of *Klebsiella pneumoniae* associated with allantoin metabolism and liver infection. Infect. Immun. **72**:3783-3792.

70. **Christendat D., A. Yee, A. Dharamsi, Y. Kluger, M. Gerstein, C. H. Arrowsmith, and A. M. Edwards**. 2000. Structural proteomics: prospects for high throughput sample preparation. Prog. Biophys. Mol. Biol. **73**:339-45.

71. **Chuang Y. P., C. T. Fang, S. Y. Lai, S. C. Chang, and J. T. Wang**. 2006. Genetic determinants of capsular serotype K1 of *Klebsiella pneumoniae* causing primary pyogenic liver abscess. J. Infect. Dis. **193**:645-654.

72. **Cianciotto N. P.** 2005. Type II secretion: a protein secretion system for all seasons. Trends Microbiol. **13**:581-8.

73. **Ciurana B., and J. M. Tomás**. 1987. Role of lipopolysaccharide and complement in susceptibility of *Klebsiella pneumoniae* to nonimmune serum. Infect. Immun. **55**:2741-6.

74. **Clarke T. F., and P. L. Clark**. 2008. Rare codons cluster. PLoS One **3**:e3412.

75. **Clegg S., B. K. Purcell, and J. Pruckler**. 1987. Characterization of genes encoding type 1 fimbriae of *Klebsiella pneumoniae*, *Salmonella typhimurium*, and *Serratia marcescens*. Infect. Immun. **55**:281-7.

76. **Clegg S., J. Wilson, and J. Johnson**. 2011. More than one way to control hair growth: regulatory mechanisms in enterobacteria that affect fimbriae assembled by the chaperone/usher pathway. J. Bacteriol. **193**:2081-8.

77. **Clements A., D. Tull, A. W. Jenney, J. L. Farn, S.-H. Kim, R. E. Bishop, J. B. McPhee, R. E. W. Hancock, E. L. Hartland, M. J. Pearse, O. L. C. Wijburg, D. C. Jackson, M. J. McConville, and R. A. Strugnell**. 2007. Secondary acylation of *Klebsiella pneumoniae* lipopolysaccharide contributes to sensitivity to antibacterial peptides. J. Biol. Chem. **282**:15569-77.

78. **Coligan J. E., B. M. Dunn, D. W. Speicher, and P. T. Wingfield**. 2001. Current Protocols in Protein Science. John Wiley & Sons, Inc., Hoboken, NJ, USA.

79. **Cordero L., R. Rau, D. Taylor, and L. W. Ayers**. 2004. Enteric gram-negative bacilli bloodstream infections: 17 years' experience in a neonatal intensive care unit. Am. J. Infect. Control **32**:189-95.

80. **Correia M., F. Boavida, F. Grosso, M. J. Salgado, L. M. Lito, J. M. Cristino, S. Mendo, and A. Duarte**. 2003. Molecular characterization of a new class 3 integron in *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **47**:2838-43.

81. **Corsaro M. M., C. De Castro, T. Naldi, M. Parrilli, J. M. Tomás, and M. Regué**. 2005. 1H and 13C NMR characterization and secondary structure of the K2 polysaccharide of *Klebsiella pneumoniae* strain 52145. Carbohydr. Res. **340**:2212-7.

82. **Cortés G., N. Borrell, B. de Astorza, C. Gómez, J. Sauleda, and S. Albertí**. 2002. Molecular analysis of the contribution of the capsular polysaccharide and the lipopolysaccharide O side chain to the virulence of *Klebsiella pneumoniae* in a murine model of pneumonia. Infect. Immun. **70**:2583-90.

83. **Cosgrove S. E., and Y. Carmeli**. 2003. The impact of antimicrobial resistance on health and economic outcomes. Clin. Infect. Dis. **36**:1433-7.

84. **Cowan S. T., K. J. Steel, C. Shaw, and J. P. Duguid**. 1960. A classification of the *Klebsiella* group. J. Gen. Microbiol. **23**:601-12.

85. **Cuzon G., T. Naas, M. C. Demachy, and P. Nordmann**. 2008. Plasmid-mediated carbapenem-hydrolyzing beta-lactamase KPC-2 in *Klebsiella pneumoniae* isolate from Greece. Antimicrob. Agents Chemother. **52**:796-7.

86. **Damjanova I., A. Tóth, J. Pászti, M. Jakab, H. Milch, A. Bauernfeind, and M. Füzi**. 2007. Epidemiology of SHV-type beta-lactamase-producing *Klebsiella* spp. from outbreaks in five geographically distant Hungarian neonatal intensive care units: widespread dissemination of epidemic R-plasmids. Int. J. Antimicrob. Agents **29**:665-71.

87. **Darfeuille-Michaud A., C. Jallat, D. Aubel, D. Sirot, C. Rich, J. Sirot, and B. Joly**. 1992. R-plasmid-encoded adhesive factor in *Klebsiella pneumoniae* strains responsible for human nosocomial infections. Infect. Immun. **60**:44-55.

88. **Datsenko K., and B. L. Wanner**. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. **97**:6640-5.

89. **Davis G. D., C. Elisee, D. M. Newham, and R. G. Harrison**. 1999. New fusion protein systems designed to give soluble expression in *Escherichia coli*. Biotechnol Bioeng. **65**:382-8.

90. **Davis T. J., and J. M. Matsen**. 1974. Prevalence and characteristics of *Klebsiella* species: relation to association with a hospital environment. J. Infect. Dis. **130**:402-405.

91. **DeShazo R. D., and S. P. Stringer**. 2011. Atrophic rhinosinusitis: progress toward explanation of an unsolved medical mystery. Curr. Opin. Allergy Clin. Immunol. **11**:1-7.

92. **Deguchi T., A. Fukuoka, M. Yasuda, M. Nakano, S. Ozeki, E. Kanematsu, Y. Nishino, S. Ishihara, Y. Ban, and Y. Kawada**. 1997. Alterations in the GyrA subunit of DNA gyrase and the ParC subunit of topoisomerase IV in quinolone-resistant clinical isolates of *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **41**:699-701.

93. **Delcher A., K. Bratke, E. Powers, and S. Salzberg**. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23**:673-679.

94. **Dobrindt U., and J. Hacker**. 2001. Whole genome plasticity in pathogenic bacteria. Curr. Opin. Microbiol. **4**:550-557.

95. **Dobrindt U., B. Hochhut, U. Hentschel, and J. Hacker**. 2004. Genomic islands in pathogenic and environmental microorganisms. Nat. Rev. Microbiol. **2**:414-424.

96. **Don R. H., P. T. Cox, B. J. Wainwright, K. Baker, and J. S. Mattick**. 1991. “Touchdown” PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. **19**:4008.

97. **Dorman C. J., and N. Ní Bhriain**. 1992. Thermal regulation of *fimA*, the *Escherichia coli* gene coding for the type 1 fimbrial subunit protein. FEMS Microbiol. Lett. **78**:125-30.

98. **Drancourt M., C. Bollet, A. Carta, and P. Rousselier**. 2001. Phylogenetic analyses of *Klebsiella* species delineate *Klebsiella* and *Raoultella* gen. nov., with description of *Raoultella ornithinolytica* comb. nov., *Raoultella terrigena* comb. nov. and *Raoultella planticola* comb. Int. J. Syst. Evol. Microbiol. **51**:925-32.

99. **Dropa M., L. C. Balsalobre, N. Lincopan, E. M. Mamizuka, V. C. Cassettari, G. R. Matté, and M. H. Matté**. 2010. Emergence of *Klebsiella pneumoniae* carrying the novel extended-spectrum beta-lactamase gene variants *bla(SHV-40)*, *bla(TEM-116)* and the class 1 integron-associated *bla(GES-7)* in Brazil. Clin. Microbiol. Infect. **16**:630-2.

100. **Drulis-Kawa Z., P. Mackiewicz, A. Kęsik-Szeloch, E. Maciaszczyk-Dziubinska, B. Weber-Dąbrowska, A. Dorotkiewicz-Jach, D. Augustyniak, G. Majkowska-Skrobek, T. Bocer, J. Empel, and A. M. Kropinski**. 2011. Isolation and characterisation of KP34--a novel φKMV-like bacteriophage for *Klebsiella pneumoniae*. Appl. Microbiol. Biotechnol. **90**:1333-45.

101. **Duerig A., S. Abel, M. Folcher, M. Nicollier, T. Schwede, N. Amiot, B. Giese, and U. Jenal**. 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes Dev. **23**:93-104.

102. **Duguid J. P.** 1959. Fimbriae and adhesive properties in *Klebsiella* strains. J. Gen. Microbiol. **21**:271-86.

103. **Duguid J. P., I. W. Smith, G. Dempster, and P. N. Edmunds**. 1955. Non-flagellar filamentous appendages (fimbriae) and haemagglutinating activity in *Bacterium coli*. J. Pathol. Bacteriol. **70**:335-48.

104. **Duncan M. J., E. L. Mann, M. S. Cohen, I. Ofek, N. Sharon, and S. N. Abraham**. 2005. The distinct binding specificities exhibited by enterobacterial type 1 fimbriae are determined by their fimbrial shafts. J. Biol. Chem. **280**:37707-37716.

105. **D'Enfert C., and A. P. Pugsley**. 1989. *Klebsiella pneumoniae pulS* gene encodes an outer membrane lipoprotein required for pullulanase secretion. J. Bacteriol. **171**:3673-9.

106. **Ebringer A., T. Rashid, H. Tiwana, and C. Wilson**. 2007. A possible link between Crohn's disease and ankylosing spondylitis via *Klebsiella* infections. Clin. Rheumatol. **26**:289-97.

107. **Elliott S. J., N. Nandapalan, and B. J. Chang**. 1991. Production of type 1 fimbriae by *Escherichia coli* HB101. Microb. Pathog. **10**:481-6.

108. **Esposito D., and J. J. Scocca**. 1994. Identification of an HP1 phage protein required for site-specific excision. Mol. Microbiol. **13**:685-95.

109. **Evrard B., D. Balestrino, A. Dosgilbert, J.-L. J. Bouya-Gachancard, N. Charbonnel, C. Forestier, and A. Tridon**. 2010. Roles of capsule and lipopolysaccharide O antigen in interactions of human monocyte-derived dendritic cells and *Klebsiella pneumoniae*. Infect. Immun. **78**:210-9.

110. **Fader R. C., K. Gondesen, B. Tolley, D. G. Ritchie, and P. Moller**. 1988. Evidence that *in vitro* adherence of *Klebsiella pneumoniae* to ciliated hamster tracheal cells is mediated by type 1 fimbriae. Infect. Immun. **56**:3011-3.

111. **Falagas M. E., and I. A. Bliziotis**. 2007. Pandrug-resistant Gram-negative bacteria: the dawn of the post-antibiotic era? Int. J. Antimicrob. Agents.

112. **Falkow S.** 2004. Molecular Koch's postulates applied to bacterial pathogenicity--a personal recollection 15 years later. Nat. Rev. Microbiol. **2**:67-72.

113. **Fang C.-T., S.-Y. Lai, W.-C. Yi, P.-R. Hsueh, K.-L. Liu, and S.-C. Chang**. 2007. *Klebsiella pneumoniae* genotype K1: an emerging pathogen that causes septic ocular or central nervous system complications from pyogenic liver abscess. Clin. Infect. Dis. **45**:284-93.

114. **Favre-Bonté S., T. R. Licht, C. Forestier, and K. A. Krogfelt**. 1999. *Klebsiella pneumoniae* capsule expression is necessary for colonization of large intestines of streptomycin-treated mice. Infect. Immun. **67**:6152-6.

115. **Fevre C., V. Passet, A. Deletoile, V. Barbe, L. Frangeul, A. S. Almeida, P. Sansonetti, R. Tournebize, and S. Brisse**. 2011. PCR-Based Identification of *Klebsiella pneumoniae* subsp. *rhinoscleromatis*, the Agent of Rhinoscleroma. PLoS Negl. Trop. Dis. **5**:e1052.

116. **Fielder M., S. J. Pirt, I. Tarpey, C. Wilson, P. Cunningham, C. Ettelaie, A. Binder, S. Bansal, and A. Ebringer**. 1995. Molecular mimicry and ankylosing spondylitis: possible role of a novel sequence in pullulanase of *Klebsiella pneumoniae*. FEBS Lett. **369**:243-8.

117. **Figurski D. H.** 1979. Replication of an Origin-Containing Derivative of Plasmid RK2 Dependent on a Plasmid Function Provided *in trans*. Proc. Natl. Acad. Sci. U.S.A. **76**:1648-1652.

118. **Fineran P. C., N. R. Williamson, K. S. Lilley, and G. P. Salmond**. 2007. Virulence and prodigiosin antibiotic biosynthesis in *Serratia* are regulated pleiotropically by the GGDEF/EAL domain protein, PigX. J. Bacteriol. **189**:7653-7662.

119. **Fineran P. C., H. Slater, L. Everson, K. Hughes, and G. P. C. Salmond**. 2005. Biosynthesis of tripyrrole and beta-lactam secondary metabolites in *Serratia*: integration of quorum sensing with multiple new regulatory components in the control of prodigiosin and carbapenem antibiotic production. Mol. Microbiol. **56**:1495-517.

120. **Finn R. D., J. Tate, J. Mistry, P. C. Coggill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. Sonnhammer, and A. Bateman**. 2008. The Pfam protein families database. Nucleic Acids Res. **36**:0-8.

121. **Fluit A. C., and F.-J. Schmitz**. 2004. Resistance integrons and super-integrons. Clin. Microbiol. Infect. **10**:272-88.

122. **Fournet-Fayard S., B. Joly, and C. Forestier**. 1995. Transformation of wild type *Klebsiella pneumoniae* with plasmid DNA by electroporation. J. Microbiol. Methods **24**:49–54.

123. **Fournier B., C. Y. Lu, P. H. Lagrange, R. Krishnamoorthy, and A. Philippon**. 1995. Point mutation in the pribnow box, the molecular basis of beta-lactamase overproduction in *Klebsiella oxytoca*. Antimicrob. Agents Chemother. **39**:1365-8.

124. **Fouts D. E., H. L. Tyler, R. T. DeBoy, S. Daugherty, Q. Ren, J. H. Badger, A. S. Durkin, H. Huot, S. Shrivastava, S. Kothari, R. J. Dodson, Y. Mohamoud, H. Khouri, L. F. Roesch, K. A. Krogfelt, C. Struve, E. W. Triplett, and B. A. Methe**. 2008. Complete genome sequence of the N2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. PLoS Genet. **4**:1-18.

125. **Fox J. G., and M. W. Rohovsky**. 1975. Meningitis caused by *Klebsiella* spp in two rhesus monkeys. J. Am. Vet. Med. Assoc. **167**:634-6.

126. **Frisch A. Von**. 1882. Zur aetiologie des rhinoskleroms. Med. Wschr. Jahrg. **32**:969–972.

127. **Frost L. S., R. Leplae, A. O. Summers, and A. Toussaint**. 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3**:722-732.

128. **Gal-Mor O., and B. B. Finlay**. 2006. Pathogenicity islands: a molecular toolbox for bacterial virulence. Cell. Microbiol. **8**:1707-19.

129. **Galani I., M. Souli, N. Mitchell, Z. Chryssouli, and H. Giamarellou**. 2010. Presence of plasmid-mediated quinolone resistance in *Klebsiella pneumoniae* and *Escherichia coli* isolates possessing *blaVIM-1* in Greece. Int. J. Antimicrob. Agents **36**:252-4.

130. **Gally D. L., J. a Bogan, B. I. Eisenstein, and I. C. Blomfield**. 1993. Environmental regulation of the *fim* switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. J. Bacteriol. **175**:6186-93.

131. **Galperin M. Y.** 2005. A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. BMC Microbiol. **5**:35.

132. **Galperin M. Y.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. J. Bacteriol. **188**:4169-82.

133. **Gao R., and A. M. Stock**. 2009. Biological insights from structures of two-component proteins. Annu. Rev. Microbiol. **63**:133-54.

134. **García B., C. Latasa, C. Solano, F. Garcia-del Portillo, C. Gamazo, and I. Lasa**. 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. Mol. Microbiol. **54**:264-277.

135. **García de la Torre M., J. Romero-Vivas, J. Martínez-Beltrán, A. Guerrero, M. Meseguer, and E. Bouza**. 1985. *Klebsiella* bacteremia: an analysis of 100 episodes. Rev. Infect. Dis. **7**:143-50.

136. **Gassama-Sow A., M. H. Diallo, A. A. Wane, A. Seck, B. Samb-Ba, P. S. Sow, and A. Aïdara-Kane**. 2010. Genetic determinants of antibiotic resistance in diarrheagenic *Klebsiella pneumoniae* subspecies *ozaenae*: an emerging enteropathogen in Senegal. Clin. Infect. Dis. **50**:453-5.

137. **Gay P., D. Le Coq, M. Steinmetz, T. Berkelman, and C. I. Kado**. 1985. Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. J. Bacteriol. **164**:918-21.

138. **Gerlach G. F., S. Clegg, and B. L. Allen**. 1989. Identification and characterization of the genes encoding the type 3 and type 1 fimbrial adhesins of *Klebsiella pneumoniae*. J. Bacteriol. **171**:1262-1270.

139. **Germon P., D. Roche, S. Melo, S. Mignon-Grasteau, U. Dobrindt, J. Hacker, C. Schouler, and M. Moulin-Schouleur**. 2007. tDNA locus polymorphism and ecto-chromosomal DNA insertion hot-spots are related to the phylogenetic group of *Escherichia coli* strains. Microbiology **153**:826-837.

140. **Giobbia M., P. G. Scotton, A. Carniato, M. Cruciani, A. Farnia, E. Daniotti, G. Scarpa, and A. Vaglia**. 2003. Community-acquired *Klebsiella pneumoniae* bacteremia with meningitis and endophthalmitis in Italy. Int. J. Infect. Dis. **7**:234-5.

141. **Glickman L. T.** 1981. Veterinary nosocomial (hospital-acquired) *Klebsiella* infections. J. Am. Vet. Med. Assoc. **179**:1389-92.

142. **Goetz A. M., J. D. Rihs, J. W. Chow, N. Singh, and R. R. Muder**. 1995. An outbreak of infusion-related *Klebsiella pneumoniae* bacteremia in a liver transplantation unit. Clin. Infect. Dis. **21**:1501-3.

143. **Goldberg R. B., R. A. Bender, and S. L. Streicher**. 1974. Direct selection for P1-sensitive mutants of enteric bacteria. J. Bacteriol. **118**:810-4.

144. **Goodman S. D., and J. J. Scocca**. 1989. Nucleotide sequence and expression of the gene for the site-specific integration protein from bacteriophage HP1 of *Haemophilus influenzae*. J. Bacteriol. **171**:4232-40.

145. **Guerin F., C. Le Bouguenec, J. Gilquin, F. Haddad, and F. W. Goldstein**. 1998. Bloody diarrhea caused by *Klebsiella pneumoniae*: a new mechanism of bacterial virulence? Clin. Infect. Dis. **27**:648-9.

146. **Gutierrez-Ríos R. M., J. A. Freyre-Gonzalez, O. Resendis, J. Collado-Vides, M. Saier, and G. Gosset**. 2007. Identification of regulatory network topological units coordinating the genome-wide transcriptional response to glucose in *Escherichia coli*. BMC Microbiol. **7**:53.

147. **Guzman L. M., D. Belin, M. J. Carson, and J. Beckwith**. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. **177**:4121-4130.

148. **Hacker J., and E. Carniel**. 2001. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. EMBO Rep. **2**:376-381.

149. **Hacker J., and J. B. Kaper**. 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. **54**:641-79.

150. **Hamon E., P. Horvatovich, E. Izquierdo, F. Bringel, E. Marchioni, D. Aoudé-Werner, and S. Ennahar**. 2011. Comparative proteomic analysis of *Lactobacillus plantarum* for the identification of key proteins in bile tolerance. BMC Microbiol. **11**:63.

151. **Han-Sen C.** 1982. The ozena problem. Clinical analysis of atrophic rhinitis in 100 cases. Acta Otolaryngol. **93**:461-4.

152. **Hanahan D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557-80.

153. **Hansen D. S., H. M. Aucken, T. Abiola, and R. Podschun**. 2004. Recommended test panel for differentiation of *Klebsiella* species on the basis of a trilateral interlaboratory evaluation of 18 biochemical tests. J. Clin. Microbiol. **42**:3665-3669.

154. **Hansen D. S., A. Gottschau, and H. J. Kolmos**. 1998. Epidemiology of *Klebsiella* bacteraemia: a case control study using *Escherichia coli* bacteraemia as control. J. Hosp. Infect. **38**:119-32.

155. **Hansen D. S., F. Mestre, S. Alberti, S. Hernández-Allés, D. Alvarez, A. Doménech-Sánchez, J. Gil, S. Merino, J. M. Tomás, and V. J. Benedí**. 1999. *Klebsiella pneumoniae* lipopolysaccharide O typing: revision of prototype strains and O-group distribution among clinical isolates from different sources and countries. J. Clin. Microbiol. **37**:56-62.

156. **Hansen D. S., R. Skov, J. V. Benedi, V. Sperling, and H. J. Kolmos**. 2002. *Klebsiella* typing: pulsed-field gel electrophoresis (PFGE) in comparison with O:K-serotyping. Clin. Microbiol. Infect. **8**:397-404.

157. **Harrison E. M., M. E. K. Carter, S. Luck, H.-Y. Ou, X. He, Z. Deng, C. O'Callaghan, A. Kadioglu, and K. Rajakumar**. 2010. Pathogenicity islands PAPI-1 and PAPI-2 contribute individually and synergistically to the virulence of *Pseudomonas aeruginosa* strain PA14. Infect. Immun. **78**:1437-46.

158. **Hart C. A., and S. K. Rao**. 2000. Rhinoscleroma. J. Med. Microbiol. **49**:395-6.

159. **Henderson I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, and D. Ala'Aldeen**. 2004. Type V protein secretion pathway: the autotransporter story. Microbiol. Mol. Biol. Rev. **68**:692-744.

160. **Hengge R.** 2009. Principles of c-di-GMP signalling in bacteria. Nat. Rev. Microbiol. **7**:263-73.

161. **Hennequin C., and C. Forestier**. 2009. *oxyR*, a LysR-type regulator involved in *Klebsiella pneumoniae* mucosal and abiotic colonization. Infect. Immun. **77**:5449-57.

162. **Herrero M., V. de Lorenzo, and K. N. Timmis**. 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. J. Bacteriol. **172**:6557-6567.

163. **Hoang T. T., R. R. Karkhoff-Schweizer, A. J. Kutchma, and H. P. Schweizer**. 1998. A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene **212**:77-86.

164. **Hobman J. L., M. D. Patel, G. A. Hidalgo-Arroyo, S. J. L. Cariss, M. B. Avison, C. W. Penn, and C. Constantinidou**. 2007. Comparative genomic hybridization detects secondary chromosomal deletions in *Escherichia coli* K-12 MG1655 mutants and highlights instability in the *flhDC* region. J. Bacteriol. **189**:8786-92.

165. **Hochhut B., and M. K. Waldor**. 1999. Site-specific integration of the conjugal *Vibrio cholerae* SXT element into *prfC*. Mol. Microbiol. **32**:99-110.

166. **Hochuli E., W. Bannwarth, H. Döbeli, R. Gentz, and D. Stüber**. 1988. Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. Bio/Technology **6**:1321-1325.

167. **Horton R. M., H. D. Hunt, S. N. Ho, J. K. Pullen, and L. R. Pease**. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77**:61-8.

168. **Hou Y. M.** 1999. Transfer RNAs and pathogenicity islands. Trends Biochem. Sci. **24**:295-298.

169. **Houwink A. L., and W. van Iterson**. 1950. Electron microscopical observations on bacterial cytology; a study on flagellation. Biochim. Biophys. Acta **5**:10-44.

170. **Hsiao W. W., K. Ung, D. Aeschliman, J. Bryan, B. B. Finlay, and F. S. Brinkman**. 2005. Evidence of a large novel gene pool associated with prokaryotic genomic islands. PLoS Genet. **1**:e62.

171. **Hsieh P.-F., T.-L. Lin, C.-Z. Lee, S.-F. Tsai, and J.-T. Wang**. 2008. Serum-induced iron-acquisition systems and TonB contribute to virulence in *Klebsiella pneumoniae* causing primary pyogenic liver abscess. J. Infect. Dis. **197**:1717-27.

172. **Huang Y.-J., H.-W. Liao, C.-C. Wu, and H.-L. Peng**. 2008. MrkF is a component of type 3 fimbriae in *Klebsiella pneumoniae*. Res. Microbiol. **160**:71-9.

173. **Hulo N., A. Bairoch, V. Bulliard, L. Cerutti, E. De Castro, P. S. Langendijk-Genevaux, M. Pagni, and C. J. Sigrist**. 2006. The PROSITE database. Nucleic Acids Res. **34**.

174. **Hultgren S. J., W. R. Schwan, A. J. Schaeffer, and J. L. Duncan**. 1986. Regulation of production of type 1 pili among urinary tract isolates of *Escherichia coli*. Infect. Immun. **54**:613-620.

175. **Human Microbiome Project**. 2010. *Klebsiella pneumoniae* subsp. *rhinoscleromatis* ATCC 13884. http://www.hmpdacc-resources.org/cgi-bin/img_hmp/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=647000262.

176. **Human Microbiome Project**. 2011. *Klebsiella* sp. 1_1_55. http://www.hmpdacc-resources.org/cgi-bin/img_hmp/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=647533174.

177. **Humphries A. D., M. Raffatellu, S. Winter, E. H. Weening, R. a Kingsley, R. Droleskey, S. Zhang, J. Figueiredo, S. Khare, J. Nunes, L. G. Adams, R. M. Tsolis, and A. J. Bäumler**. 2003. The use of flow cytometry to detect expression of subunits encoded by 11 *Salmonella enterica* serotype Typhimurium fimbrial operons. Mol. Microbiol. **48**:1357-1376.

178. **Humphries R. M., T. Kelesidis, J. Dien Bard, K. W. Ward, D. Bhattacharya, and M. A. Lewinski**. 2010. Successful treatment of pan-resistant *Klebsiella pneumoniae* pneumonia and bacteraemia with a combination of high-dose tigecycline and colistin. J. Med. Microbiol. **59**:1383-6.

179. **Hvidberg H., C. Struve, K. A. Krogfelt, N. Christensen, S. N. Rasmussen, and N. Frimodt-Møller**. 2000. Development of a long-term ascending urinary tract infection mouse model for antibiotic treatment studies. Antimicrob. Agents Chemother. **44**:156-63.

180. **Jenal U., and J. Malone**. 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. Annu. Rev. Genet. **40**:385-407.

181. **Jiang P., J. Li, F. Han, G. Duan, X. Lu, Y. Gu, and W. Yu**. 2011. Antibiofilm activity of an exopolysaccharide from marine bacterium *Vibrio* sp. QY101. PLoS One **6**:e18514.

182. **Jogler C., W. Lin, A. Meyerdierks, M. Kube, E. Katzmann, C. Flies, Y. Pan, R. Amann, R. Reinhardt, and D. Schüler**. 2009. Toward cloning of the magnetotactic metagenome: identification of magnetosome island gene clusters in uncultivated magnetotactic bacteria from different aquatic sediments. Appl. Environ. Microbiol. **75**:3972-9.

183. **Johnson J. G., and S. Clegg**. 2010. Role of MrkJ, a phosphodiesterase, in type 3 fimbrial expression and biofilm formation in *Klebsiella pneumoniae*. J. Bacteriol. **192**:3944-50.

184. **Johnson J. G., C. N. Murphy, J. Sippy, T. J. Johnson, and S. Clegg**. 2011. Type 3 Fimbriae and Biofilm Formation Are Regulated by the Transcriptional Regulators MrkHI in *Klebsiella pneumoniae*. J. Bacteriol. **193**:3453-60.

185. **Jones C. H., J. S. Pinkner, R. Roth, J. Heuser, a V. Nicholes, S. N. Abraham, and S. J. Hultgren**. 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the *Enterobacteriaceae*. Proc. Natl. Acad. Sci. U.S.A. **92**:2081-5.

186. **Joseph D., G. A. Petsko, and M. Karplus**. 1990. Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop. Science **249**:1425-8.

187. **Justice S. S., C. Hung, J. A. Theriot, D. A. Fletcher, G. G. Anderson, M. J. Footer, and S. J. Hultgren**. 2004. Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. Proc. Natl. Acad. Sci. U.S.A. **101**:1333-8.

188. **Justice S. S., D. A. Hunstad, P. C. Seed, and S. J. Hultgren**. 2006. Filamentation by *Escherichia coli* subverts innate defenses during urinary tract infection. Proc. Natl. Acad. Sci. U.S.A. **103**:19884-9.

189. **Kalivoda E. J., J. Horzempa, N. a Stella, A. Sadaf, R. P. Kowalski, G. J. Nau, and R. M. Q. Shanks**. 2011. New Vector Tools with a Hygromycin Resistance Marker for Use with Opportunistic Pathogens. Mol. Biotechnol. **48**:7-14.

190. **Kane J.** 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr. Opin. Biotechnol. **6**:494-500.

191. **Kang Y., M. H. Norris, A. R. Barrett, B. A. Wilcox, and T. T. Hoang**. 2009. Engineering of tellurite-resistant genetic tools for single-copy chromosomal analysis of *Burkholderia* spp. and characterization of the *Burkholderia thailandensis betBA* operon. Appl. Environ. Microbiol. **75**:4015-27.

192. **Kapust R. B., and D. S. Waugh**. 1999. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci. **8**:1668-74.

193. **Karah N., L. Poirel, S. Bengtsson, M. Sundqvist, G. Kahlmeter, P. Nordmann, A. Sundsfjord, and Ø. Samuelsen**. 2010. Plasmid-mediated quinolone resistance determinants *qnr* and *aac(6')-Ib-cr* in *Escherichia coli* and *Klebsiella* spp. from Norway and Sweden. Diagn. Microbiol. Infect. Dis. **66**:425-31.

194. **Karama E. M., F. Willermain, X. Janssens, M. Claus, S. Van den Wijngaert, J.-T. Wang, C. Verougstraete, and L. Caspers**. 2008. Endogenous endophthalmitis complicating *Klebsiella pneumoniae* liver abscess in Europe: case report. Int. Ophthalmol. **28**:111-3.

195. **Keynan Y., J. A. Karlowsky, T. Walus, and E. Rubinstein**. 2007. Pyogenic liver abscess caused by hypermucoviscous *Klebsiella pneumoniae*. Scand. J. Infect. Dis. **39**:828-30.

196. **Kibbe W. A.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. **35**:W43-6.

197. **Kim Y., X. Wang, Q. Ma, X.-S. Zhang, and T. K. Wood**. 2009. Toxin-antitoxin systems in *Escherichia coli* influence biofilm formation through YjgK (TabA) and fimbriae. J. Bacteriol. **191**:1258-67.

198. **Kingsford C., M. C. Schatz, and M. Pop**. 2010. Assembly complexity of prokaryotic genomes using short reads. BMC Bioinformatics **11**:21.

199. **Kinkler R. J., J. E. Wagner, R. E. Doyle, and D. R. Owens**. 1976. Bacterial mastitis in guinea pigs. Lab. Anim. Sci. **26**:214-7.

200. **Klemm P., and G. Christiansen**. 1990. The *fimD* gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. Mol. Gen. Genet. **220**:334-8.

201. **Klemm P., B. J. Jørgensen, I. van Die, H. de Ree, and H. Bergmans**. 1985. The *fim* genes responsible for synthesis of type 1 fimbriae in *Escherichia coli*, cloning and genetic organization. Mol. Gen. Genet. **199**:410-4.

202. **Klemm P., B. J. Jørgensen, B. Kreft, and G. Christiansen**. 1995. The export systems of type 1 and F1C fimbriae are interchangeable but work in parental pairs. J. Bacteriol. **177**:621-7.

203. **Klemm P.** 1986. Two regulatory *fim* genes,*fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. EMBO J. **5**:1389-1393.

204. **Kline K. A., K. W. Dodson, M. G. Caparon, and S. J. Hultgren**. 2010. A tale of two pili: assembly and function of pili in bacteria. Trends Microbiol. **18**:224-32.

205. **Knodler L. A., A. Bestor, C. Ma, I. Hansen-Wester, M. Hensel, B. A. Vallance, and O. Steele-Mortimer**. 2005. Cloning vectors and fluorescent proteins can significantly inhibit *Salmonella enterica* virulence in both epithelial cells and macrophages: implications for bacterial pathogenesis studies. Infect. Immun. **73**:7027-31.

206. **Knowles R., R. Neufeld, and S. Simpson**. 1974. Acetylene reduction (nitrogen fixation) by pulp and paper mill effluents and by *Klebsiella* isolated from effluents and environmental situations. Appl. Microbiol. **28**:608-13.

207. **Ko W.-C., D. L. Paterson, A. J. Sagnimeni, D. S. Hansen, A. Von Gottberg, S. Mohapatra, J. M. Casellas, H. Goossens, L. Mulazimoglu, G. Trenholme, K. P. Klugman, J. G. McCormack, and V. L. Yu**. 2002. Community-acquired *Klebsiella pneumoniae* bacteremia: global differences in clinical patterns. Emerging Infect. Dis. **8**:160-6.

208. **Koczura R., and A. Kaznowski**. 2003. Occurrence of the *Yersinia* high-pathogenicity island and iron uptake systems in clinical isolates of *Klebsiella pneumoniae*. Microb. Pathog. **35**:197-202.

209. **Koonin E. V.** 2005. Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. **39**:309-38.

210. **Korea C.-G., R. Badouraly, M.-C. Prevost, J.-M. Ghigo, and C. Beloin**. 2010. *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. Environ. Microbiol. **12**:1957-1977.

211. **Korea C.-G., J.-M. Ghigo, and C. Beloin**. 2011. The sweet connection: Solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli* : Multiple *E. coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tropism. BioEssays **33**:300-11.

212. **Korhonen T. K., E. Tarkka, H. Ranta, and K. Haahtela**. 1983. Type 3 fimbriae of *Klebsiella* sp.: molecular characterization and role in bacterial adhesion to plant roots. J. Bacteriol. **155**:860-5.

213. **Kovtunovych G., T. Lytvynenko, B. Sylvain, V. Negrutska, and N. Kozyrovska**. 2003. Identification of *Klebsiella oxytoca* using a specific PCR assay targeting the polygalacturonase *pehX* gene. Res. Microbiol. **154**:587-592.

214. **Kuchma S. L., K. M. Brothers, J. H. Merritt, N. T. Liberati, F. M. Ausubel, and G. A. O'Toole**. 2007. BifA, a cyclic-di-GMP phosphodiesterase, inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. J. Bacteriol. **189**:8165-78.

215. **Kulasekara B. R., H. D. Kulasekara, M. C. Wolfgang, L. Stevens, D. W. Frank, and S. Lory**. 2006. Acquisition and evolution of the *exoU* locus in *Pseudomonas aeruginosa*. J. Bacteriol. **188**:4037-50.

216. **Kumar V., P. Sun, J. Vamathevan, Y. Li, K. Ingraham, L. Palmer, J. Huang, and J. R. Brown**. 2011. Comparative Genomics of *Klebsiella pneumoniae* strains with Different Antibiotic Resistance Profiles. Antimicrob. Agents Chemother. **55**:4267-76.

217. **Kumarasamy K. K., M. A. Toleman, T. R. Walsh, J. Bagaria, F. Butt, R. Balakrishnan, U. Chaudhary, M. Doumith, C. G. Giske, and S. Irfan**. 2010. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. Lancet. Infect. Dis. **10**:597-602.

218. **LaVallie E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, and J. M. McCoy**. 1993. A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the *E. coli* Cytoplasm. Bio/Technology **11**:187-193.

219. **Lacoste L., M. Lacaille, and L. Brakier-Gingras**. 1976. New types of streptomycin-resistant mutants of *Escherichia coli*. Biochim. Biophys. Acta **442**:88-97.

220. **Lai T.-H., Y. Kumagai, M. Hyodo, Y. Hayakawa, and Y. Rikihisa**. 2009. The *Anaplasma phagocytophilum* PleC histidine kinase and PleD diguanylate cyclase two-component system and role of cyclic di-GMP in host cell infection. J. Bacteriol. **191**:693-700.

221. **Lai Y.-C., H.-L. Peng, and H.-Y. Chang**. 2003. RmpA2, an activator of capsule biosynthesis in *Klebsiella pneumoniae* CG43, regulates K2 *cps* gene expression at the transcriptional level. J. Bacteriol. **185**:788-800.

222. **Lambertsen L., C. Sternberg, and S. Molin**. 2004. Mini-Tn*7* transposons for site-specific tagging of bacteria with fluorescent proteins. Environ. Microbiol. **6**:726-32.

223. **Larkin M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins**. 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947-8.

224. **Laslett D., and B. Canback**. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. **32**:11-16.

225. **Lau H. Y., S. Clegg, and T. A. Moore**. 2007. Identification of *Klebsiella pneumoniae* genes uniquely expressed in a strain virulent using a murine model of bacterial pneumonia. Microb. Pathog. **42**:148-55.

226. **Lavender H. F., J. R. Jagnow, and S. Clegg**. 2004. Biofilm formation *in vitro* and virulence *in vivo* of mutants of *Klebsiella pneumoniae*. Infect. Immun. **72**:4888-4890.

227. **Lawlor M. S., J. Hsu, P. D. Rick, and V. L. Miller**. 2005. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. Mol. Microbiol. **58**:1054-1073.

228. **Lawlor M. S., C. O'Connor, and V. L. Miller**. 2007. Yersiniabactin is a virulence factor for *Klebsiella pneumoniae* during pulmonary infection. Infect. Immun. **75**:1463-72.

229. **Lawrence J. G., and H. Hendrickson**. 2005. Genome evolution in bacteria: order beneath chaos. Curr. Opin. Microbiol. **8**:572-578.

230. **Lederman E. R., and N. F. Crum**. 2005. Pyogenic liver abscess with a focus on *Klebsiella pneumoniae* as a primary pathogen: an emerging disease with unique clinical characteristics. Am. J. Gastroenterol. **100**:322-331.

231. **Li M., and S. J. Elledge**. 2007. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. Nat. Methods **4**:251-256.

232. **Licht T. R., K. A. Krogfelt, P. S. Cohen, L. K. Poulsen, J. Urbance, and S. Molin**. 1996. Role of lipopolysaccharide in colonization of the mouse intestine by *Salmonella typhimurium* studied by *in situ* hybridization. Infect. Immun. **64**:3811-7.

233. **Lin C.-T., C.-C. Wu, Y.-S. Chen, Y.-C. Lai, C. Chi, J.-C. Lin, Y. Chen, and H.-L. Peng**. 2011. Fur regulation of the capsular polysaccharide biosynthesis and iron-acquisition systems in *Klebsiella pneumoniae* CG43. Microbiology **157**:419-29.

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

235. **Liu Y., C. Liu, W. Zheng, X. Zhang, J. Yu, Q. Gao, Y. Hou, and X. Huang**. 2008. PCR detection of *Klebsiella pneumoniae* in infant formula based on 16S-23S internal transcribed spacer. Int. J. Food Microbiol. **125**:230-5.

236. **Llobet E., C. March, P. Giménez, and J. A. Bengoechea**. 2009. *Klebsiella pneumoniae* OmpA confers resistance to antimicrobial peptides. Antimicrob. Agents Chemother. **53**:298-302.

237. **Lorenzo V. de, M. Herrero, U. Jakubzik, and K. N. Timmis**. 1990. Mini-Tn*5* Transposon Derivatives for Insertion Mutagenesis, Promoter Probing, and Chromosomal Insertion of Cloned DNA in Gram-Negative Eubacteria. J. Bacteriol. **172**:6568-6572.

238. **Low A. S., N. Holden, T. Rosser, A. J. Roe, C. Constantinidou, J. L. Hobman, D. G. E. Smith, J. C. Low, and D. L. Gally**. 2006. Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. Environ. Microbiol. **8**:1033-47.

239. **Lucchini S., G. Rowley, M. D. Goldberg, D. Hurd, M. Harrison, and J. C. D. Hinton**. 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog. **2**:e81.

240. **Lukashin A. V., and M. Borodovsky**. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26**:1107-1115.

241. **Ma L. C., C. T. Fang, C. Z. Lee, C. T. Shun, and J. T. Wang**. 2005. Genomic heterogeneity in *Klebsiella pneumoniae* strains is associated with primary pyogenic liver abscess and metastatic infection. J. Infect. Dis. **192**:117-128.

242. **Ma L., M. Conover, H. Lu, M. R. Parsek, K. Bayles, and D. J. Wozniak**. 2009. Assembly and development of the *Pseudomonas aeruginosa* biofilm matrix. PLoS Pathog. **5**:e1000354.

243. **Mahillon J., and M. Chandler**. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62**:725-74.

244. **Mantri Y., and K. P. Williams**. 2004. Islander: a database of integrative islands in prokaryotic genomes, the associated integrases and their DNA site specificities. Nucleic Acids Res. **32**:D55-8.

245. **Marchaim D., R. Zaidenstein, T. Lazarovitch, Y. Karpuch, T. Ziv, and M. Weinberger**. 2008. Epidemiology of bacteremia episodes in a single center: increase in Gram-negative isolates, antibiotics resistance, and patient age. Eur. J. Clin. Microbiol. **27**:1045-51.

246. **Marchler-Bauer A., J. B. Anderson, M. K. Derbyshire, C. DeWeese-Scott, N. R. Gonzales, M. Gwadz, L. Hao, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, D. Krylov, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, N. Thanki, R. A. Yamashita, J. J. Yin, D. Zhang, and S. H. Bryant**. 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res. **35**:D237-40.

247. **Mardis E. R.** 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. **24**:133-41.

248. **Maroncle N., D. Balestrino, C. Rich, and C. Forestier**. 2002. Identification of *Klebsiella pneumoniae* genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis. Infect. Immun. **70**:4729-4734.

249. **Marshall O. J.** 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. Bioinformatics **20**:2471-2.

250. **Martinez-Wilson H. F., R. Tamayo, A. D. Tischler, D. W. Lazinski, and A. Camilli**. 2008. The *Vibrio cholerae* hybrid sensor kinase VieS contributes to motility and biofilm regulation by altering the cyclic diguanylate level. J. Bacteriol. **190**:6439-47.

251. **Martino P. Di, V. Livrelli, D. Sirot, B. Joly, and A. Darfeuille-Michaud**. 1996. A new fimbrial antigen harbored by CAZ-5/SHV-4-producing *Klebsiella pneumoniae* strains involved in nosocomial infections. Infect. Immun. **64**:2266-73.

252. **Martino P. Di, N. Cafferini, B. Joly, and A. Darfeuille-Michaud**. 2003. *Klebsiella pneumoniae* type 3 pili facilitate adherence and biofilm formation on abiotic surfaces. Res. Microbiol. **154**:9-16.

253. **McClain M. S., I. C. Blomfield, and B. I. Eisenstein**. 1991. Roles of *fimB* and *fimE* in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. J. Bacteriol. **173**:5308-14.

254. **McGowan J.** 2001. Economic impact of antimicrobial resistance. Emerging Infect. Dis. **7**:286-292.

255. **McKenzie G. J., and N. L. Craig**. 2006. Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn*7* without need for selection of the insertion event. BMC Microbiol. **6**:39.

256. **Meatherall B. L., D. Gregson, T. Ross, J. D. D. Pitout, and K. B. Laupland**. 2009. Incidence, risk factors, and outcomes of *Klebsiella pneumoniae* bacteremia. Am. J. Med. **122**:866-73.

257. **Medini D., C. Donati, H. Tettelin, V. Masignani, and R. Rappuoli**. 2005. The microbial pan-genome. Curr. Opin. Genet. Dev. **15**:589-594.

258. **Merino S., M. Altarriba, L. Izquierdo, M. M. Nogueras, M. Regué, and J. M. Tomás**. 2000. Cloning and sequencing of the *Klebsiella pneumoniae* O5 *wb* gene cluster and its role in pathogenesis. Infect. Immun. **68**:2435-40.

259. **Metzker M. L.** 2010. Sequencing technologies - the next generation. Nat. Rev. Genet. **11**:31-46.

260. **Miethke M., and M. A. Marahiel**. 2007. Siderophore-based iron acquisition and pathogen control. Microbiol. Mol. Biol. Rev. **71**:413-51.

261. **Miller V. L., and J. J. Mekalanos**. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170**:2575-83.

262. **Minasov G., S. Padavattan, L. Shuvalova, and JS**. 2009. Crystal structures of YkuI and its complex with second messenger cyclic di-GMP suggest catalytic mechanism of phosphodiester bond cleavage by EAL domains. J. Biol. Chem. **284**:13174-84.

263. **Miyata S., M. Casey, D. Frank, and F. Ausubel**. 2003. Use of the *Galleria mellonella* caterpillar as a model host to study the role of the type III secretion system in *Pseudomonas aeruginosa* pathogenesis. Infect. Immun. **71**:2404-2413.

264. **Mizuta K., M. Ohta, M. Mori, T. Hasegawa, I. Nakashima, and N. Kato**. 1983. Virulence for mice of *Klebsiella* strains belonging to the O1 group: relationship to their capsular (K) types. Infect. Immun. **40**:56-61.

265. **Montgomerie J. Z.** 1979. Epidemiology of *Klebsiella* and hospital-associated infections. Rev. Infect. Dis. **1**:736-53.

266. **Munera D., S. Hultgren, and L. A. Fernandez**. 2007. Recognition of the N-terminal lectin domain of FimH adhesin by the usher FimD is required for type 1 pilus biogenesis. Mol. Microbiol. **64**:333-346.

267. **Murphy K. C., K. G. Campellone, and A. R. Poteete**. 2000. PCR-mediated gene replacement in *Escherichia coli*. Gene **246**:321-330.

268. **Murphy K., and K. Campellone**. 2003. Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. BMC Mol. Biol. **4**:1-12.

269. **Mushegian A. R., and E. V. Koonin**. 1996. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc. Natl. Acad. Sci. U.S.A. **93**:10268-10273.

270. **Myhal M. L., D. C. Laux, and P. S. Cohen**. 1982. Relative colonizing abilities of human fecal and K 12 strains of *Escherichia coli* in the large intestines of streptomycin-treated mice. Eur. J. Clin. Microbiol. **1**:186-92.

271. **NCBI**. ORF Finder. http://www.ncbi.nlm.nih.gov/projects/gorf/orfig.cgi.

272. **Nadasy K. A., R. Domiati-Saad, and M. A. Tribble**. 2007. Invasive *Klebsiella pneumoniae* Syndrome in North America. Clin. Infect. Dis. **45**:e25-e28.

273. **Nassif X., and P. J. Sansonetti**. 1986. Correlation of the virulence of *Klebsiella pneumoniae* K1 and K2 with the presence of a plasmid encoding aerobactin. Infect. Immun. **54**:603-8.

274. **Nassif X., J. Fournier, J. Arondel, and PJ**. 1989. Mucoid phenotype of *Klebsiella pneumoniae* is a plasmid-encoded virulence factor. Infect. Immun. **57**:546-552.

275. **Navarro M. V. A. S., N. De, N. Bae, Q. Wang, and H. Sondermann**. 2009. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. Structure **17**:1104-16.

276. **Nordmann P., G. Cuzon, and T. Naas**. 2009. The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. Lancet. Infect. Dis. **9**:228-236.

277. **Nougayrède J.-P., S. Homburg, F. Taieb, M. Boury, E. Brzuszkiewicz, G. Gottschalk, C. Buchrieser, J. Hacker, U. Dobrindt, and E. Oswald**. 2006. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. Science **313**:848-51.

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

279. **Oelschlaeger T. A., and B. D. Tall**. 1997. Invasion of cultured human epithelial cells by *Klebsiella pneumoniae* isolated from the urinary tract. Infect. Immun. **65**:2950-8.

280. **Ogasawara M., D. H. Kono, and D. T. Yu**. 1986. Mimicry of human histocompatibility HLA-B27 antigens by *Klebsiella pneumoniae*. Infect. Immun. **51**:901-8.

281. **Ou H. Y., L. L. Chen, J. Lonnen, R. R. Chaudhuri, A. B. Thani, R. Smith, N. J. Garton, J. Hinton, M. Pallen, M. R. Barer, and K. Rajakumar**. 2006. A novel strategy for the identification of genomic islands by comparative analysis of the contents and contexts of tRNA sites in closely related bacteria. Nucleic Acids Res. **34**:e3.

282. **Ou H. Y., X. He, E. M. Harrison, B. R. Kulasekara, A. B. Thani, A. Kadioglu, S. Lory, J. C. Hinton, M. R. Barer, Z. Deng, and K. Rajakumar**. 2007. MobilomeFINDER: web-based tools for *in silico* and experimental discovery of bacterial genomic islands. Nucleic Acids Res. **35**:W97-W104.

283. **O'Toole G., and R. Kolter**. 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol. Microbiol. **28**:449-61.

284. **Pallen M. J., and B. W. Wren**. 2007. Bacterial pathogenomics. Nature **449**:835-42.

285. **Pan X., A. Lührmann, A. Satoh, M. A. Laskowski-Arce, and C. R. Roy**. 2008. Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. Science **320**:1651-4.

286. **Pan Y.-J., T.-L. Lin, C.-R. Hsu, and J.-T. Wang**. 2011. Use of a *Dictyostelium* model for isolation of genetic loci associated with phagocytosis and virulence in *Klebsiella pneumoniae*. Infect. Immun. **79**:997-1006.

287. **Partridge S. R., L. C. Thomas, A. N. Ginn, A. M. Wiklendt, P. Kyme, and J. R. Iredell**. 2011. A Novel Gene Cassette, *aacA43*, in a Plasmid-Borne Class 1 Integron. Antimicrob. Agents Chemother. **55**:2979-82.

288. **Paszkiewicz K., and D. J. Studholme**. 2010. *De novo* assembly of short sequence reads. Brief. Bioinform. **11**:457-72.

289. **Paterson D. L., K. M. Hujer, A. M. Hujer, B. Yeiser, M. D. Bonomo, L. B. Rice, and R. A. Bonomo**. 2003. Extended-spectrum beta-lactamases in *Klebsiella pneumoniae* bloodstream isolates from seven countries: dominance and widespread prevalence of SHV- and CTX-M-type beta-lactamases. Antimicrob. Agents Chemother. **47**:3554-60.

290. **Pavlovic G., V. Burrus, B. Gintz, B. Decaris, and G. Guédon**. 2004. Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICESt1-related elements from *Streptococcus thermophilus*. Microbiology **150**:759-74.

291. **Perkins T. T., R. A. Kingsley, M. C. Fookes, P. P. Gardner, K. D. James, L. Yu, S. A. Assefa, M. He, N. J. Croucher, D. J. Pickard, D. J. Maskell, J. Parkhill, J. Choudhary, N. R. Thomson, and G. Dougan**. 2009. A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. PLoS Genet. **5**:e1000569.

292. **Perry R. D., and C. L. San Clemente**. 1979. Siderophore synthesis in *Klebsiella pneumoniae* and *Shigella sonnei* during iron deficiency. J. Bacteriol. **140**:1129-32.

293. **Peters J. E., and N. L. Craig**. 2001. Tn*7*: smarter than we thought. Nat. Rev. Mol. Cell Biol. **2**:806-14.

294. **Pfaffl M. W.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29**:e45.

295. **Phan G., H. Remaut, T. Wang, W. J. Allen, K. F. Pirker, A. Lebedev, N. S. Henderson, S. Geibel, E. Volkan, J. Yan, M. B. A. Kunze, J. S. Pinkner, B. Ford, C. W. M. Kay, H. Li, S. J. Hultgren, D. G. Thanassi, and G. Waksman**. 2011. Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. Nature **474**:49-53.

296. **Philippe N., J. P. Alcaraz, E. Coursange, J. Geiselmann, and D. Schneider**. 2004. Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. Plasmid **51**:246-255.

297. **Pinkner J. S., H. Remaut, F. Buelens, E. Miller, V. Aberg, N. Pemberton, M. Hedenstrom, A. Larsson, P. Seed, G. Waksman, S. J. Hultgren, and F. Almqvist**. 2006. Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. Proc. Natl. Acad. Sci. U.S.A. **103**:17897-17902.

298. **Pinto A. C., H. P. Melo-Barbosa, A. Miyoshi, A. Silva, and V. Azevedo**. 2011. Application of RNA-seq to reveal the transcript profile in bacteria. Genet. Mol. Res. **10**:1707-18.

299. **Pinto-Tomás A. A., M. A. Anderson, G. Suen, D. M. Stevenson, F. S. T. Chu, W. W. Cleland, P. J. Weimer, and C. R. Currie**. 2009. Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. Science **326**:1120-3.

300. **Pizarro-Cerdá J., and P. Cossart**. 2006. Bacterial adhesion and entry into host cells. Cell **124**:715-27.

301. **Platt H., J. G. Atherton, and I. Orskov**. 1976. *Klebsiella* and *Enterobacter* organisms isolated from horses. J .Hyg. **77**:401-8.

302. **Podschun R., A. Fischer, and U. Ullman**. 2000. Expression of putative virulence factors by clinical isolates of *Klebsiella planticola*. J. Med. Microbiol. **49**:115-9.

303. **Podschun R., S. Pietsch, C. Holler, and U. Ullmann**. 2001. Incidence of *Klebsiella* species in surface waters and their expression of virulence factors. Appl. Environ. Microbiol. **67**:3325-3327.

304. **Podschun R., and H. Sahly**. 1991. Hemagglutinins of *Klebsiella pneumoniae* and *K. oxytoca* isolated from different sources. Zentralbl. Hyg. Umweltmed. **191**:46-52.

305. **Podschun R., and U. Ullmann**. 1992. Isolation of *Klebsiella terrigena* from clinical specimens. Eur. J. Clin. Microbiol. **11**:349-52.

306. **Podschun R., and U. Ullmann**. 1992. *Klebsiella* capsular type K7 in relation to toxicity, susceptibility to phagocytosis and resistance to serum. J. Med. Microbiol. **36**:250-4.

307. **Podschun R., and U. Ullmann**. 1998. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin. Microbiol. Rev. **11**:589-603.

308. **Poirel L., I. Le Thomas, T. Naas, A. Karim, and P. Nordmann**. 2000. Biochemical sequence analyses of GES-1, a novel class A extended-spectrum beta-lactamase, and the class 1 integron In52 from *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **44**:622-32.

309. **Pollack M., P. Charache, R. E. Nieman, M. P. Jett, J. A. Reimhardt, and P. H. Hardy**. 1972. Factors influencing colonisation and antibiotic-resistance patterns of Gram-negative bacteria in hospital patients. Lancet **2**:668-71.

310. **Poteete A. R., A. C. Fenton, and A. Nadkarni**. 2004. Chromosomal duplications and cointegrates generated by the bacteriophage lamdba Red system in *Escherichia coli* K-12. BMC Mol. Biol. **5**:22.

311. **Poteete A., C. Rosadini, and C. St. Pierre**. 2006. Gentamicin and other cassettes for chromosomal gene replacement in *Escherichia coli*. BioTechniques **41**:261-264.

312. **Pugsley A. P., O. Francetic, K. Hardie, O. M. Possot, N. Sauvonnet, and A. Seydel**. 1997. Pullulanase: Model protein substrate for the general secretory pathway of gram-negative bacteria. Folia Microbiol. Springer Netherlands **42**:184-192.

313. **Purcell B. K., J. Pruckler, and S. Clegg**. 1987. Nucleotide sequences of the genes encoding type 1 fimbrial subunits of *Klebsiella pneumoniae* and *Salmonella typhimurium*. J. Bacteriol. **169**:5831-4.

314. **Putze J., C. Hennequin, J.-P. Nougayrède, W. Zhang, S. Homburg, H. Karch, M.-A. Bringer, C. Fayolle, E. Carniel, W. Rabsch, T. a Oelschlaeger, E. Oswald, C. Forestier, J. Hacker, and U. Dobrindt**. 2009. Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. Infect. Immun. **77**:4696-703.

315. **Rahimian J., T. Wilson, V. Oram, and R. S. Holzman**. 2004. Pyogenic liver abscess: recent trends in etiology and mortality. Clin. Infect. Dis. **39**:1654-9.

316. **Rahn, A. Drummelsmith J., and C. Whitfield**. 1999. Conserved organization in the *cps* gene clusters for expression of *Escherichia coli* group 1 K antigens: relationship to the colanic acid biosynthesis locus and the *cps* genes from *Klebsiella pneumoniae*. J. Bacteriol. **181**:2307-13.

317. **Randrianirina F., S. Vedy, D. Rakotovao, C.-E. Ramarokoto, H. Ratsitohaina, J. F. Carod, E. Ratsima, M. Morillon, and A. Talarmin**. 2009. Role of contaminated aspiration tubes in nosocomial outbreak of *Klebsiella pneumoniae* producing SHV-2 and CTX-M-15 extended-spectrum beta-lactamases. J. Hosp. Infect. **72**:23-9.

318. **Ransjö U., Z. Good, K. Jalakas, I. Kühn, I. Siggelkow, B. Aberg, and E. Anjou**. 1992. An outbreak of *Klebsiella oxytoca* septicemias associated with the use of invasive blood pressure monitoring equipment. Acta. Anaesthesiol. Scand. **36**:289-91.

319. **Rao F., Y. Qi, H. S. Chong, M. Kotaka, B. Li, J. Li, J. Lescar, K. Tang, and Z.-X. Liang**. 2009. The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase. J. Bacteriol. **191**:4722-31.

320. **Rao F., Y. Yang, Y. Qi, and Z.-X. Liang**. 2008. Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. J. Bacteriol. **190**:3622-31.

321. **Reacher M. H., A. Shah, D. M. Livermore, M. C. Wale, C. Graham, A. P. Johnson, H. Heine, M. A. Monnickendam, K. F. Barker, D. James, and R. C. George**. 2000. Bacteraemia and antibiotic resistance of its pathogens reported in England and Wales between 1990 and 1998: trend analysis. BMJ **320**:213-6.

322. **Reed M. R., K. E. Shearwin, L. M. Pell, and J. B. Egan**. 1997. The dual role of Apl in prophage induction of coliphage 186. Mol. Microbiol. **23**:669-81.

323. **Regué M., N. Climent, N. Abitiu, N. Coderch, S. Merino, L. Izquierdo, M. Altarriba, and J. M. M. Tomás**. 2001. Genetic characterization of the *Klebsiella pneumoniae waa* gene cluster, involved in core lipopolysaccharide biosynthesis. J. Bacteriol. **183**:3564.

324. **Regué M., L. Izquierdo, S. Fresno, N. Piqué, M. M. Corsaro, T. Naldi, C. De Castro, D. Waidelich, S. Merino, and J. M. Tomás**. 2005. A second outer-core region in *Klebsiella pneumoniae* lipopolysaccharide. J. Bacteriol. **187**:4198-206.

325. **Reiter W. D., P. Palm, and S. Yeats**. 1989. Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res. **17**:1907-1914.

326. **Rendueles O., L. Travier, P. Latour-Lambert, T. Fontaine, J. Magnus, E. Denamur, and J.-M. Ghigo**. 2011. Screening of *Escherichia coli* species biodiversity reveals new biofilm-associated antiadhesion polysaccharides. mBio **2**:e00043-11.

327. **Reznikoff W. S.** 2008. Transposon Tn*5*. Annu. Rev. Genet. **42**:269-86.

328. **Rice L. B.** 2010. Progress and challenges in implementing the research on ESKAPE pathogens. Infect. Control Hosp. Epidemiol. **31**:S7-10.

329. **Rocha E. P. C.** 2008. The organization of the bacterial genome. Annu. Rev. Genet. **42**:211-33.

330. **Ronald A.** 2003. The etiology of urinary tract infection: traditional and emerging pathogens. Dis. Mon. **49**:71-82.

331. **Rosen D. A., J. S. Pinkner, J. M. Jones, J. N. Walker, S. Clegg, and S. J. Hultgren**. 2008. Utilization of an intracellular bacterial community pathway in *Klebsiella pneumoniae* urinary tract infection and the effects of FimK on type 1 pilus expression. Infect. Immun. **76**:3337-3345.

332. **Rosen D. A., J. S. Pinkner, J. N. Walker, J. S. Elam, J. M. Jones, and S. J. Hultgren**. 2008. Molecular variations in *Klebsiella pneumoniae* and *Escherichia coli* FimH affect function and pathogenesis in the urinary tract. Infect. Immun. **76**:3346-56.

333. **Rosenblueth M., L. Martínez, J. Silva, and E. Martínez-Romero**. 2004. *Klebsiella variicola*, a novel species with clinical and plant-associated isolates. Syst. Appl. Microbiol. **27**:27-35.

334. **Ross P., H. Weinhouse, Y. Aloni, D. Michaeli, P. Weinberger-Ohana, R. Mayer, S. Braun, E. de Vroom, G. A. van der Marel, J. H. van Boom, and M. Benziman**. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. Nature **325**:279-81.

335. **Rozen S., and H. Skaletsky**. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132**:365-86.

336. **Rumer L., J. Jores, P. Kirsch, Y. Cavignac, K. Zehmke, and L. H. Wieler**. 2003. Dissemination of *pheU*- and *pheV*-located genomic islands among enteropathogenic (EPEC) and enterohemorrhagic (EHEC) *E. coli* and their possible role in the horizontal transfer of the locus of enterocyte effacement (LEE). Int. J. Med. Microbiol. **292**:463-75.

337. **Rutherford K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell**. 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944-945.

338. **Ruzin A., D. Keeney, and P. a Bradford**. 2007. AdeABC multidrug efflux pump is associated with decreased susceptibility to tigecycline in *Acinetobacter calcoaceticus-Acinetobacter baumannii* complex. J. Antimicrob. Chemother. **59**:1001-4.

339. **Ryan R. P., Y. Fouhy, J. F. Lucey, and J. M. Dow**. 2006. Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. J. Bacteriol. **188**:8327-34.

340. **Römling U.** 2001. Genetic and phenotypic analysis of multicellular behavior in *Salmonella typhimurium*. Methods Enzymol. **336**:48-59.

341. **Römling U., W. D. Sierralta, K. Eriksson, and S. Normark**. 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD promoter*. Mol. Microbiol. **28**:249-64.

342. **Römling U., M. Gomelsky, and M. Y. Galperin**. 2005. C-di-GMP: the dawning of a novel bacterial signalling system. Mol. Microbiol. **57**:629-39.

343. **Saeland E., G. Vidarsson, and I. Jonsdottir**. 2000. Pneumococcal pneumonia and bacteremia model in mice for the analysis of protective antibodies. Microb. Pathog. **29**:81-91.

344. **Sakellaris H., S. N. Luck, K. Al-Hasani, K. Rajakumar, S. A. Turner, and B. Adler**. 2004. Regulated site-specific recombination of the *she* pathogenicity island of *Shigella flexneri*. Mol. Microbiol. **52**:1329-36.

345. **Samra Z., O. Ofir, Y. Lishtzinsky, L. Madar-Shapiro, and J. Bishara**. 2007. Outbreak of carbapenem-resistant *Klebsiella pneumoniae* producing KPC-3 in a tertiary medical centre in Israel. Int. J. Antimicrob. Agents **30**:525-9.

346. **Sarno R., G. McGillivary, D. J. Sherratt, L. A. Actis, and M. E. Tolmasky**. 2002. Complete nucleotide sequence of *Klebsiella pneumoniae* multiresistance plasmid pJHCMW1. Antimicrob. Agents Chemother. **46**:3422-7.

347. **Sarris P. F., C. Zoumadakis, N. J. Panopoulos, and E. V. Scoulica**. 2010. Distribution of the putative type VI secretion system core genes in *Klebsiella* spp. Infect. Genet. Evol. 1-10.

348. **Sauer F. G., H. Remaut, S. J. Hultgren, and G. Waksman**. 2004. Fiber assembly by the chaperone-usher pathway. Biochim. Biophys. Acta **1694**:259-67.

349. **Schembri M. A., J. Blom, K. A. Krogfelt, and P. Klemm**. 2005. Capsule and fimbria interaction in *Klebsiella pneumoniae*. Infect. Immun. **73**:4626-4633.

350. **Schirmer T., and U. Jenal**. 2009. Structural and mechanistic determinants of c-di-GMP signalling. Nat. Rev. Microbiol. **7**:724-35.

351. **Schmidt A. J., D. A. Ryjenkov, and M. Gomelsky**. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. J. Bacteriol. **187**:4774-4781.

352. **Schroll C., K. B. Barken, K. a Krogfelt, and C. Struve**. 2010. Role of type 1 and type 3 fimbriae in *Klebsiella pneumoniae* biofilm formation. BMC Microbiol. **10**:179.

353. **Schubert S., A. Rakin, and J. Heesemann**. 2004. The *Yersinia* high-pathogenicity island (HPI): evolutionary and functional aspects. Int. J. Med. Microbiol. **294**:83-94.

354. **Schurtz T. A., D. B. Hornick, T. K. Korhonen, and S. Clegg**. 1994. The type 3 fimbrial adhesin gene (*mrkD*) of *Klebsiella* species is not conserved among all fimbriate strains. Infect. Immun. **62**:4186-4191.

355. **Schwan W. R., J. L. Lee, F. A. Lenard, B. T. Matthews, and M. T. Beck**. 2002. Osmolarity and pH growth conditions regulate *fim* gene transcription and type 1 pilus expression in uropathogenic *Escherichia coli*. Infect. Immun. **70**:1391-402.

356. **Schwan W. R., S. Shibata, S.-I. Aizawa, and A. J. Wolfe**. 2007. The two-component response regulator RcsB regulates type 1 piliation in *Escherichia coli*. J. Bacteriol. **189**:7159-63.

357. **Scotto-Lavino E., G. Du, and M. a Frohman**. 2006. 5' end cDNA amplification using classic RACE. Nat. Protoc. **1**:2555-62.

358. **Sebghati T. A., T. K. Korhonen, D. B. Hornick, and S. Clegg**. 1998. Characterization of the type 3 fimbrial adhesins of *Klebsiella* strains. Infect. Immun. **66**:2887-94.

359. **Sechter I., F. Mestre, and D. S. Hansen**. 2000. Twenty-three years of *Klebsiella* phage typing: a review of phage typing of 12 clusters of nosocomial infections, and a comparison of phage typing with K serotyping. Clin. Microbiol. Infect. **6**:233-8.

360. **Selden R., S. Lee, W. L. Wang, J. V. Bennett, and T. C. Eickhoff**. 1971. Nosocomial *Klebsiella* infections: intestinal colonization as a reservoir. Ann. Intern. Med. **74**:657-64.

361. **Seth-Smith H., and N. J. Croucher**. 2009. Genome watch: breaking the ICE. Nat. Rev. Microbiol. **7**:328-9.

362. **Shearwin K. E., and J. B. Egan**. 2000. Establishment of lysogeny in bacteriophage 186. DNA binding and transcriptional activation by the CII protein. J. Biol. Chem. **275**:29113-22.

363. **Shu H.-Y., C.-P. Fung, Y.-M. Liu, K.-M. Wu, Y.-T. Chen, L.-H. Li, T.-T. Liu, R. Kirby, and S.-F. Tsai**. 2009. Genetic diversity of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* clinical isolates. Microbiology **155**:4170-83.

364. **Siguier P., J. Filée, and M. Chandler**. 2006. Insertion sequences in prokaryotic genomes. Curr. Opin. Microbiol. **9**:526-31.

365. **Simm R., M. Morr, A. Kader, M. Nimtz, and U. Romling**. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Mol. Microbiol. **53**:1123-1134.

366. **Simon R., U. Priefer, and A. Pühler**. 1983. A Broad Host Range Mobilization System for *In Vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Bio/Technology **1**:784-791.

367. **Simoons-Smit A. M., A. M. Verweij-van Vught, and D. M. MacLaren**. 1986. The role of K antigens as virulence factors in *Klebsiella*. J. Med. Microbiol. **21**:133-7.

368. **Simoons-Smit A. M., A. M. Verwey-van Vught, I. Y. Kanis, and D. M. MacLaren**. 1984. Virulence of *Klebsiella* strains in experimentally induced skin lesions in the mouse. J. Med. Microbiol. **17**:67-77.

369. **Sintchenko V., J. Iredell, and G. Gilbert**. 2007. Pathogen profiling for disease management and surveillance. Nat. Rev. Microbiol. **5**:464-470.

370. **Sintim H. O., J. A. I. Smith, J. Wang, S. Nakayama, and L. Yan**. 2010. Paradigm shift in discovering next-generation anti-infective agents: targeting quorum sensing, c-di-GMP signaling and biofilm formation in bacteria with small molecules. Future Med. Chem. **2**:1005-35.

371. **Slater H., M. Crow, L. Everson, and G. P. Salmond**. 2003. Phosphate availability regulates biosynthesis of two antibiotics, prodigiosin and carbapenem, in *Serratia* via both quorum-sensing-dependent and -independent pathways. Mol. Microbiol. **47**:303-320.

372. **Smialowski P., A. J. Martin-Galiano, A. Mikolajka, T. Girschick, T. A. Holak, and D. Frishman**. 2007. Protein solubility: sequence based prediction and experimental verification. Bioinformatics **23**:2536-42.

373. **Smith D. B., and K. S. Johnson**. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67**:31-40.

374. **Snyder L., and W. Champness**. 2007. Molecular genetics of bacteria. ASM Press.

375. **Sobirk S. K., C. Struve, and S. G. Jacobsson**. 2010. Primary *Klebsiella pneumoniae* Liver Abscess with Metastatic Spread to Lung and Eye, a North-European Case Report of an Emerging Syndrome. Open Microbiol. J. **4**:5-7.

376. **Sokurenko E. V., V. Chesnokova, D. E. Dykhuizen, I. Ofek, X. R. Wu, K. A. Krogfelt, C. Struve, M. A. Schembri, and D. L. Hasty**. 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. Proc. Natl. Acad. Sci. U.S.A. **95**:8922-6.

377. **Sokurenko E. V., H. S. Courtney, D. E. Ohman, P. Klemm, and D. L. Hasty**. 1994. FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor sequence variations among *fimH* genes. J. Bacteriol. **176**:748-755.

378. **Sokurenko E. V., M. Feldgarden, E. Trintchina, S. J. Weissman, S. Avagyan, S. Chattopadhyay, J. R. Johnson, and D. E. Dykhuizen**. 2004. Selection footprint in the FimH adhesin shows pathoadaptive niche differentiation in *Escherichia coli*. Mol. Biol. Evol. **21**:1373-83.

379. **Solano C., B. García, C. Latasa, A. Toledo-Arana, V. Zorraquino, J. Valle, J. Casals, E. Pedroso, and I. Lasa**. 2009. Genetic reductionist approach for dissecting individual roles of GGDEF proteins within the c-di-GMP signaling network in *Salmonella*. Proc. Natl. Acad. Sci. U.S.A. **106**:7997-8002.

380. **Soler Bistue A., D. Birshan, A. Tomaras, M. Dandekar, T. Tran, J. Newmark, D. Bui, N. Gupta, K. Hernandez, R. Sarno, and others**. 2008. *Klebsiella pneumoniae* multiresistance plasmid pMET1: similarity with the *Yersinia pestis* plasmid pCRY and integrative conjugative elements. PLoS One **3**:e1800.

381. **Sommerfeldt N., A. Possling, G. Becker, C. Pesavento, N. Tschowri, and R. Hengge**. 2009. Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in *Escherichia coli*. Microbiology **155**:1318-31.

382. **Staden R., K. F. Beal, and J. K. Bonfield**. 2000. The Staden package, 1998. Methods Mol. Biol. **132**:115-30.

383. **Stahlhut S. G., S. Chattopadhyay, C. Struve, S. J. Weissman, P. Aprikian, S. J. Libby, F. C. Fang, K. A. Krogfelt, and E. V. Sokurenko**. 2009. Population variability of the FimH type 1 fimbrial adhesin in *Klebsiella pneumoniae*. J. Bacteriol. **191**:1941-50.

384. **Stahlhut S. G., C. Schroll, M. Harmsen, C. Struve, and K. A. Krogfelt**. 2010. Screening for genes involved in *Klebsiella pneumoniae* biofilm formation using a fosmid library. FEMS Immunol. Med. Microbiol. **59**:521-4.

385. **Stahlhut S. G., V. Tchesnokova, C. Struve, S. J. Weissman, S. Chattopadhyay, O. Yakovenko, P. Aprikian, E. V. Sokurenko, and K. A. Krogfelt**. 2009. Comparative structure-function analysis of mannose-specific FimH adhesins from *Klebsiella pneumoniae* and *Escherichia coli*. J. Bacteriol. **191**:6592-601.

386. **Stickler D., N. Morris, and C. Winters**. 1999. Simple physical model to study formation and physiology of biofilms on urethral catheters. Methods Enzymol. **310**:494-501.

387. **Struve C., M. Bojer, and K. A. Krogfelt**. 2008. Characterization of *Klebsiella pneumoniae* type 1 fimbriae by detection of phase variation during colonization and infection and impact on virulence. Infect. Immun. **76**:4055-4065.

388. **Struve C., M. Bojer, and K. A. Krogfelt**. 2009. Identification of a conserved chromosomal region encoding *Klebsiella pneumoniae* type 1 and type 3 fimbriae and assessment of the role of fimbriae in pathogenicity. Infect. Immun. **77**:5016-5024.

389. **Struve C., C. Forestier, and K. A. Krogfelt**. 2003. Application of a novel multi-screening signature-tagged mutagenesis assay for identification of *Klebsiella pneumoniae* genes essential in colonization and infection. Microbiology **149**:167-176.

390. **Struve C., and K. A. Krogfelt**. 2004. Pathogenic potential of environmental *Klebsiella pneumoniae* isolates. Environ. Microbiol. **6**:584–590.

391. **Struve C., and K. A. Krogfelt**. 2003. Role of capsule in *Klebsiella pneumoniae* virulence: lack of correlation between *in vitro* and *in vivo* studies. FEMS Microbiol. Lett. **218**:149-54.

392. **Studier F. W., and B. A. Moffatt**. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113-30.

393. **Sunde M., and M. Norström**. 2005. The genetic background for streptomycin resistance in *Escherichia coli* influences the distribution of MICs. J. Antimicrob. Chemother. **56**:87-90.

394. **Sundin G. W., and C. L. Bender**. 1995. Expression of the *strA-strB* streptomycin resistance genes in *Pseudomonas syringae* and *Xanthomonas campestris* and characterization of IS*6100* in *X. campestris*. Appl. Environ. Microbiol. **61**:2891-7.

395. **Sutherland I.** 2001. Biofilm exopolysaccharides: a strong and sticky framework. Microbiology **147**:3-9.

396. **Suzuki K., P. Babitzke, S. R. Kushner, and T. Romeo**. 2006. Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. Genes & development **20**:2605-17.

397. **Swenson D. L., N. O. Bukanov, D. E. Berg, and R. A. Welch**. 1996. Two pathogenicity islands in uropathogenic *Escherichia coli* J96: cosmid cloning and sample sequencing. Infect. Immun. **64**:3736-3743.

398. **Sylvain B., F. Grimont, and P. Grimont**. 2006. The Genus *Klebsiella*., p. 159-196. *In* Prokaryotes.

399. **Sørensen H. P., and K. K. Mortensen**. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotechnol. **115**:113-28.

400. **Tamayo R., A. D. Tischler, and A. Camilli**. 2005. The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. J. Biol. Chem. **280**:33324-30.

401. **Tang H.-L., M.-K. Chiang, W.-J. Liou, Y.-T. Chen, H.-L. Peng, C.-S. Chiou, K.-S. Liu, M.-C. Lu, K.-C. Tung, and Y.-C. Lai**. 2010. Correlation between *Klebsiella pneumoniae* carrying pLVPK-derived loci and abscess formation. Eur. J. Clin. Microbiol. **29**:689-98.

402. **Tarkkanen A. M., R. Virkola, S. Clegg, and T. K. Korhonen**. 1997. Binding of the type 3 fimbriae of *Klebsiella pneumoniae* to human endothelial and urinary bladder cells. Infect. Immun. **65**:1546-9.

403. **Taylor R. K., C. Manoil, and J. J. Mekalanos**. 1989. Broad-host-range vectors for delivery of Tn*phoA*: use in genetic analysis of secreted virulence determinants of *Vibrio cholerae*. J. Bacteriol. **171**:1870-1878.

404. **Tchigvintsev A., X. Xu, A. Singer, C. Chang, G. Brown, M. Proudfoot, H. Cui, R. Flick, W. F. Anderson, A. Joachimiak, M. Y. Galperin, A. Savchenko, and A. F. Yakunin**. 2010. Structural insight into the mechanism of cyclic di-GMP hydrolysis by EAL domain phosphodiesterases. J. Mol. Biol. **402**:524-38.

405. **Tenzen T., S. Matsutani, and E. Ohtsubo**. 1990. Site-specific transposition of insertion sequence IS*630*. J. Bacteriol. **172**:3830-6.

406. **Tenzen T., and E. Ohtsubo**. 1991. Preferential transposition of an IS*630*-associated composite transposon to TA in the 5'-CTAG-3' sequence. J. Bacteriol. **173**:6207-12.

407. **The Genome Institute. Washington University School of Medicine.** 2005. *Klebsiella pneumoniae* MGH78578. http://genome.wustl.edu/genomes/view/klebsiella_pneumoniae.

408. **The Genome Institute. Washington University School of Medicine.** 2010. *Klebsiella* sp. MS 92-3. http://genome.wustl.edu/genomes/view/klebsiella_sp._ms_92-3.

409. **Thomas W. E., E. Trintchina, M. Forero, V. Vogel, and E. V. Sokurenko**. 2002. Bacterial adhesion to target cells enhanced by shear force. Cell **109**:913-23.

410. **Thomsen R. W., P. Jepsen, and H. T. Sørensen**. 2007. Diabetes mellitus and pyogenic liver abscess: risk and prognosis. Clin. Infect. Dis. **44**:1194-201.

411. **Toenniessen E.** 1914. Ueber Vererbung und Variabilitgt bei Bakterien mit besonderer Beriicksichtigung der Virulenz. Zentralblatt fur Bakteriologl. I. Abt. Orig. **73**:241-277.

412. **Tolmasky M. E., R. M. Chamorro, J. H. Crosa, and P. M. Marini**. 1988. Transposon-mediated amikacin resistance in *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **32**:1416-20.

413. **Tomás J. M., V. J. Benedí, B. Ciurana, and J. Jofre**. 1986. Role of capsule and O antigen in resistance of *Klebsiella pneumoniae* to serum bactericidal activity. Infect. Immun. **54**:85-9.

414. **Torres A. G., J. A. Giron, N. T. Perna, V. Burland, F. R. Blattner, F. Avelino-Flores, and J. B. Kaper**. 2002. Identification and characterization of *lpfABCC'DE*, a fimbrial operon of enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. **70**:5416-27.

415. **Torres A. G., G. N. López-Sánchez, L. Milflores-Flores, S. D. Patel, M. Rojas-López, C. F. Martínez de la Peña, M. M. P. Arenas-Hernández, and Y. Martínez-Laguna**. 2007. Ler and H-NS, regulators controlling expression of the long polar fimbriae of *Escherichia coli* O157:H7. J. Bacteriol. **189**:5916-28.

416. **Touchon M., C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguénec, M. Lescat, S. Mangenot, V. Martinez-Jéhanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Médigue, E. P. C. Rocha, and E. Denamur**. 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. **5**:e1000344.

417. **Trevisan V.** 1887. Sul Micrococco della rabbia e sulla possiblità di riconoscere durante il periode d'incubazione, dall'esame del sangue della persona moricata, se ha contratta l'infezione rabbica. R. C. Ist. Lombardo. **20**:88-105.

418. **Trong I. Le, P. Aprikian, B. A. Kidd, W. E. Thomas, E. V. Sokurenko, and R. E. Stenkamp**. 2010. Donor strand exchange and conformational changes during *E. coli* fimbrial formation. J. Struct. Biol. **172**:380-8.

419. **Tsai F.-C.** 2008. Pyogenic Liver Abscess as Endemic Disease, Taiwan. Emerging Infect. Dis. **14**:1592-1600.

420. **Tschowri N., S. Busse, and R. Hengge**. 2009. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. Genes Dev. **23**:522-34.

421. **Tseng T.-T., B. M. Tyler, and J. C. Setubal**. 2009. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. BMC Microbiol. **9 Suppl 1**:S2.

422. **Tumbarello M., T. Spanu, M. Sanguinetti, R. Citton, E. Montuori, F. Leone, G. Fadda, R. Cauda, and R**. 2006. Bloodstream Infections Caused by Extended-Spectrum-{beta}-Lactamase-

Producing *Klebsiella pneumoniae*: Risk Factors, Molecular Epidemiology, and Clinical. Antimicrob. Agents Chemother. **50**:498-504.

423. **Turton J. F., C. Perry, S. Elgohari, and C. V. Hampton**. 2010. PCR characterization and typing of *Klebsiella pneumoniae* using capsular type-specific, variable number tandem repeat and virulence gene targets. J. Med. Microbiol. **59**:541-7.

424. **Ueda A., and T. K. Wood**. 2009. Connecting quorum sensing, c-di-GMP, pel polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885). PLoS Pathog. **5**:e1000483.

425. **University Hospitals of Leicester NHS Trust.** 2011. Guide to Antimicrobial Use in the Leicestershire NHS Health Community.

426. **Vallone P. M., and J. M. Butler**. 2004. AutoDimer: a screening tool for primer-dimer and hairpin structures. BioTechniques **37**:226-31.

427. **Verma V., K. Harjai, and S. Chhibber**. 2009. Characterization of a T7-like lytic bacteriophage of *Klebsiella pneumoniae* B5055: a potential therapeutic agent. Curr. Microbiol. **59**:274-81.

428. **Vincze T., J. Posfai, and R. J. Roberts**. 2003. NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Res. **31**:3688-91.

429. **Wacharotayankun R., Y. Arakawa, M. Ohta, K. Tanaka, T. Akashi, M. Mori, and N. Kato**. 1993. Enhancement of extracapsular polysaccharide synthesis in *Klebsiella pneumoniae* by RmpA2, which shows homology to NtrC and FixJ. Infect. Immun. **61**:3164-74.

430. **Waksman G., and S. J. Hultgren**. 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. Nat. Rev. Microbiol. **7**:765-74.

431. **Walsh T. R., J. Weeks, D. M. Livermore, and M. A. Toleman**. 2011. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. Lancet. Infect. Dis. **11**:355-62.

432. **Wang R. F., and S. R. Kushner**. 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. Gene **100**:195-9.

433. **Wang Z., M. Gerstein, and M. Snyder**. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. **10**:57-63.

434. **Weinberg E. D.** 1978. Iron and infection. Microbiol. Rev. **42**:45-66.

435. **West S. E., H. P. Schweizer, C. Dall, A. K. Sample, and L. J. Runyen-Janecky**. 1994. Construction of improved *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19 and sequence of the region required for their replication in *Pseudomonas aeruginosa*. Gene **148**:81-6.

436. **Whitfield C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem. **75**:39-68.

437. **Williams P., H. Chart, E. Griffiths, and P. Stevenson**. 1987. Expression of high affinity iron uptake systems by clinical isolates of *Klebsiella*. FEMS Microbiol. Lett. **44**:407-12.

438. **Williams P., P. A. Lambert, M. R. Brown, and R. J. Jones**. 1983. The role of the O and K antigens in determining the resistance of *Klebsiella aerogenes* to serum killing and phagocytosis. J. Gen. Microbiol. **129**:2181-91.

439. **Williamson N. R., P. C. Fineran, F. J. Leeper, and G. P. C. Salmond**. 2006. The biosynthesis and regulation of bacterial prodiginines. Nat. Rev. Microbiol. **4**:887-99.

440. **Williamson N. R., P. C. Fineran, W. Ogawa, L. R. Woodley, and G. P. C. Salmond**. 2008. Integrated regulation involving quorum sensing, a two-component system, a GGDEF/EAL domain protein and a post-transcriptional regulator controls swarming and RhlA-dependent surfactant biosynthesis in *Serratia*. Environ. Microbiol. **10**:1202-17.

441. **Wilson R. B.** 1994. Hepatic hemosiderosis and *Klebsiella* bacteremia in a green aracari (*Pteroglossus viridis*). Avian Dis. **38**:679-81.

442. **Wright K. J., P. C. Seed, and S. J. Hultgren**. 2007. Development of intracellular bacterial communities of uropathogenic *Escherichia coli* depends on type 1 pili. Cell. Microbiol. **9**:2230-41.

443. **Wu C.-C., Y.-J. Huang, C.-P. Fung, and H.-L. Peng**. 2010. Regulation of the *Klebsiella pneumoniae* Kpc fimbriae by the site-specific recombinase KpcI. Microbiology **156**:1983-92.

444. **Wu K.-M., L.-H. Li, J.-J. Yan, N. Tsao, T.-L. Liao, H.-C. Tsai, C.-P. Fung, H.-J. Chen, Y.-M. Liu, J.-T. Wang, C.-T. Fang, S.-C. Chang, H.-Y. Shu, T.-T. Liu, Y.-T. Chen, Y.-R. Shiau, T.-L. Lauderdale, I.-J. Su, R. Kirby, and S.-F. Tsai**. 2009. Genome sequencing and comparative analysis of *Klebsiella pneumoniae* NTUH-K2044, a strain causing liver abscess and meningitis. J. Bacteriol. **191**:4492-501.

445. **Wu L.-T., S.-Y. Chang, M.-R. Yen, T.-C. Yang, and Y.-H. Tseng**. 2007. Characterization of extended-host-range pseudo-T-even bacteriophage Kpp95 isolated on *Klebsiella pneumoniae*. Appl. Environ. Microbiol. **73**:2532-40.

446. **Yarbrough M. L., Y. Li, L. N. Kinch, N. V. Grishin, H. L. Ball, and K. Orth**. 2009. AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. Science **323**:269-72.

447. **Yinnon A. M., A. Butnaru, D. Raveh, Z. Jerassy, and B. Rudensky**. 1996. *Klebsiella* bacteraemia: community versus nosocomial infection. QJM **89**:933-41.

448. **Yong D., M. A. Toleman, C. G. Giske, H. S. Cho, K. Sundman, K. Lee, and T. R. Walsh**. 2009. Characterization of a new metallo-beta-lactamase gene, *bla(NDM-1)*, and a novel erythromycin esterase gene carried on a unique genetic structure in *Klebsiella pneumoniae* sequence type 14 from India. Antimicrob. Agents Chemother. **53**:5046-54.

449. **Yoon M. Y., K.-M. Lee, Y. Park, and S. S. Yoon**. 2011. Contribution of cell elongation to the biofilm formation of *Pseudomonas aeruginosa* during anaerobic respiration. PLoS One **6**:e16105.

450. **Zadoks R. N., H. M. Griffiths, M. A. Munoz, C. Ahlstrom, G. J. Bennett, E. Thomas, and Y. H. Schukken**. 2011. Sources of *Klebsiella* and *Raoultella* species on dairy farms: be careful where you walk. J. Dairy Sci. **94**:1045-51.

451. **Zahn K.** 1996. Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth. J. Bacteriol. **178**:2926-33.

452. **Zahn K., and A. Landy**. 1996. Modulation of lambda integrase synthesis by rare arginine tRNA. Mol. Microbiol. **21**:69-76.

453. **Zhang G., B. Ma, X. Xu, C. Li, and L. Wang**. 2007. Fast conversion of glycerol to 1,3-propanediol by a new strain of *Klebsiella pneumoniae*. Biochem. Eng. J. **37**:256-260.

454. **Zhang J., J. J. van Aartsen, X. Jiang, Y. Shao, C. Tai, X. He, Z. Tan, Z. Deng, S. Jia, K. Rajakumar, and others**. 2010. Expansion of the known *Klebsiella pneumoniae* species gene pool by characterization of novel alien DNA islands integrated into tmRNA gene sites. J. Microbiol. Methods **84**:283-9.

455. **Zheng P., J. Sun, J. van den Heuvel, and A.-P. Zeng**. 2006. Discovery and investigation of a new, second triose phosphate isomerase in *Klebsiella pneumoniae*. J. Biotechnol. **125**:462-73.

456. **Zogaj X., W. Bokranz, M. Nimtz, and U. Römling**. 2003. Production of cellulose and curli fimbriae by members of the family Enterobacteriaceae isolated from the human gastrointestinal tract. Infect. Immun. **71**:4151-8.

457. **Ørskov I**. 1984. Bergey's Manual of Systematic Bacteriology, 1st ed. Williams & Wilkins, Baltimore.

458. **Ørskov I., and F. Ørskov**. 1984. Serotyping of *Klebsiella*. Methods in Microbiology **14**:143-164.

Appendices

Appendix 1. Culture media

Acanthamoeba medium (AUM)

USE: Growth medium for bacterial cultures

20 g of biosate (Becton Dickinson), 5 g of glucose, 0.3 g KH_2PO_4, 10 µg vitamin B12, and 15 mg L-methionine were dissolved in a final volume of 1 l of dH_2O and pH adjusted to 6.5-6.6. Media was autoclaved before use.

Artificial urine media (AUM)

USE: Growth medium for bacterial cultures

This complex medium was prepared as described by Brooks and Keevil (42).

Brain heart infusion (BHI) broth plus 30 % glycerol

USE: Storage of bacterial stocks at -20 and -80ºC

47 g of brain heart infusion broth powder (Oxoid) was dissolved in a final volume of 1 l of dH_2O containing 30 % (v/v) glycerol and was autoclaved before use.

Congo Red medium

USE: Assessment of cellulose production

Congo red medium was prepared by dissolving 4 g of tryptone, 2 g of yeast extract and 2 g of NaCl, 40 µg/ml congo red and 20 µg/ml brilliant blue and 1.5 % w/v of agar in dH_2O to a final volume of 400 ml. Media was autoclaved before use.

King's B medium

USE: Growth medium for bacterial cultures

8 g of peptone, 0.6 g K_2HPO_4, 0.6 g $MgSO_4.7H_2O$ and 4 ml of glycerol were dissolved in a final volume of 400 ml dH_2O and autoclaved before use.

Lysogeny broth (LB) and agar (LA)

USE: Standard liquid (LB) and solid growth (LA) medium for bacterial cultures

Lysogeny broth (LB) was prepared by dissolving 4 g of tryptone, 2 g of yeast extract and 2 g of NaCl in dH_2O to a final volume of 400 ml. LB agar (LA) was prepared as for LB with 1.5 % w/v of agar. Both media types were autoclaved before use.

LAS and LAGS

USE: LAS and LAGS are used as counterselection media for allelic exchange

LB agar with 6 % sucrose (LAS) was prepared by dissolving 4 g of tryptone, 2 g of yeast extract, 6 g of agar powder and 24 g of sucrose in dH_2O to a final volume of 400 ml. LAGS is LAS supplemented with 9 µg/ml gentamicin.

LA motility

USE: To test swimming motility

LA motility was prepared by dissolving 4 g of tryptone, 2 g of yeast extract, 2 g of NaCl, 2 g of glucose and 1.6 g of agar powder in dH_2O to a final volume of 400 ml.

M9 broth and agar

USE: Minimal media for selection of desired merodiploid strains

A 5x M9 media stock was prepared by dissolving 15 g NA_2HPO_4, 7.5 g KH_2PO_4, 2.5 g NH_4Cl, 1.25 g NaCl and 7.5 mg $CaCl_2$ in dH_2O to a final volume of 500 ml and autoclaved. 100 ml of 5x M9 media stock, 5 ml of 20 % glycerol (autoclaved) and 500 µl of 1 M $MgSO_4.7H_2O$ (filter sterilized) were added to 395 ml of autoclaved dH_2O to make M9 broth. When M9 agar was required 1.5 % w/v agar was added.

MacConkey agar

USE: Selection for Gram negative bacteria from murine stool samples

52 g of MacConkey agar powder (Oxoid) was suspended in 1 l of dH_2O and sterilised by autoclaving. It contains bile salts and inhibits most Gram positive bacterial growth.

SOC (Super Optimal broth with Catabolite repression)

USE: Broth for non-selective outgrowth of bacteria post-transformation

SOC was prepared by dissolving 5 g of tryptone, 2.5 g of yeast extract and 5 g of NaCl into 200 ml of dH_2O. After autoclaving, 50 µl of 2 M $MgCl_2$ (filter sterilized) and 200 µl of 1M glucose (filter sterilized) were added to 1 ml of medium.

Appendix 2. Reagents

Phosphate buffered saline (PBS)

8 g NaCl
0.2 g KCl
1.15 g $Na_2HPO_4.7H_2O$
0.20 g KH_2PO_4
pH 7.4
Make up to 500 ml in dH_2O
Sterilized by autoclaving

50x Tris-acetate-EDTA (TAE) buffer

2 M Tris-HCl
2 M Acetic acid
50 mM EDTA

Tris-EDTA (TE) buffer

10 mM Tris-HCl, pH 8.0
1 mM EDTA, pH 8.0

1x Sodium chloride-Tris-EDTA (STE) buffer

0.1 M NaCl
10 mM Tris-HCl, pH 8.0
1 mM EDTA

5x SDS sample loading buffer

10 % (w/v)SDS
5 % (v/v) β-mercaptoethanol
20 % (v/v) glycerol
0.2 M Tris-HCl pH 6.8
0.05 % (w/v) bromophenolblue

De-staining solution

40 % (v/v) methanol
10 % (v/v) glacial acetic acid
60 % (v/v) dH_2O

Transfer buffer

3.0 g Tris base
14.4 g glycine
20 % (v/v) methanol
Add dH_2O to 1 l

10x Tris-Glycine electrophoresis buffer

30.2 g Tris-HCl pH 8.3
144 g glycine
10.0 g SDS
Add dH_2O to 1000ml
Dilute to 1x for working stock

Lysis buffer (native)

6.9 g $NaH_2PO_4.H_2O$
17.54 g NaCl
0.68 g imidazole
Adjust pH to 8.0 using NaOH
Add water to 1 l

Lysis buffer (denaturing)

13.8 g $NaH_2PO_4.H_2O$
1.2 g Tris base
480.5 g urea
Adjust pH to 8.0 using NaOH
Add water to 1 l

Appendix 3. *E. coli* and *Klebsiella* wildtype strains

Strains	KR number[a]	Relevant characteristics[b]	Reference
***E. coli* wildtype strains**			
DH5α	KR1231	F- φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17*(rK-mK+) *phoA supE*44 λ- *thi*-1 *gyrA96 relA1*	(152)
EPI300-T1R	KR1151	F- φ80d*lacZ*ΔM15 *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Δ*lacX74 recA1 endA1 araD139* Δ(*ara-leu*)7697 *galU galK* λ- *rpsL nupG trfA tonA dhfr*	Epicentre
JM109	KR1230	*endA1 recA1 gyrA96 thi*-1 *hsdR17*(rK-mK+) *relA1 supE44* Δ(*lac-proAB*), [F´ *traD36*, *proAB*, $lacI_QZ\Delta M15$].	Promega
BL21	KR1859	F- *dcm ompT hsdS20* (rB-mB-) *gal*	(392)
BL21(DE3)	KR1418	F- *dcm ompT hsdS20* (rB-mB-) *gal* λ(DE3)	(392)
HB101	KR667	F- *mcrB mrr hsdS20*(rB-mB-) *recA13 leuB6 ara*-14 *proA2 lacY1 galK2 xyl*-5 *mtl*-1 *rpsL20*(Sm^R) *glnV44* λ-	(37)
CC118λpir	KR144	Δ(*are-leu*) *araD* Δ*lacX74 galE galK phoA20 thi*-1 *rpsE rpoB argE*(Am) *recA1* λpir	(162)
SM10λpir	KR1934	*thi thr leu tonA lacY supE recA*::RP4-2-Tc::Mu Km λ*pir*	(261)
S17-1λpir	KR145	F' *thi pro hsdR*$^-$ *recA*::RP4-2-Tc::Mu Km::Tn7 λ*pir*	(366)
UTI89	KR1066	Uropathogenic *E. coli* strain.	(61)
***Klebsiella* wildtype strains**			
KR116	-	*K. pneumoniae* blood stream infection isolate, UHL-LRI	This work
KR162	-	*K. pneumoniae* blood stream infection isolate, UHL-LRI	This work
KR163	-	*K. pneumoniae* blood stream infection isolate, UHL-LRI	This work
KR173	-	*K. pneumoniae* blood stream infection isolate, UHL-LRI	This work
KR392	-	*K. pneumoniae* urinary tract infection isolate, UHL-LRI	This work
KR396	-	*Klebsiella* sp. urinary tract infection isolate, UHL-LRI	This work
KR399	-	*Klebsiella* sp. urinary tract infection isolate, UHL-LRI	This work
KR518	-	*K. pneumoniae* sputum isolate, Shanghai, China	This work

Strains	KR number[a]	Relevant characteristics[b]	Reference
KR1120	-	*Klebsiella* sp., unknown isolation site, UHL-LRI	This work
KR2159	-	*K. oxytoca* blood stream infection isolate, UHL-LRI	This work
KR2163	-	*K. oxytoca* tissue sample isolate, UHL-LRI	This work
KR2175	-	*K. pneumoniae* ascetic fluid isolate, UHL-LRI	This work
KR2315	-	*K. pneumoniae*, unknown isolation site, USA	This work
KR2317	-	*K. pneumoniae*, unknown isolation site, USA	This work
KR2318	-	*K. pneumoniae*, unknown isolation site, USA	This work
sp15	KR2414	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
sp25	KR2415	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
sp28	KR2416	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
sp29	KR2417	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
sp30	KR2418	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
sp34	KR2419	*K. pneumoniae* blood stream infection isolate, Copenhagen, Denmark	(154)
cas128	KR2420	*K. pneumoniae* environmental isolate (water), Schleswig-Holstein, Germany	(383)
cas664	KR2421	*K. pneumoniae* urinary tract infection isolate, Copenhagen, Denmark	(383)
cas669	KR2422	*K. pneumoniae* urinary tract infection isolate, Copenhagen, Denmark	(383)
C3091	KR2019	*K. pneumoniae* blood stream infection isolate	(279)
Kp342	KR1039	*K. pneumoniae* environmental isolate that is able to fix nitrogen; sequenced	(124)
NTUH-K2044	KR1730	*K. pneumoniae* pyogenic liver abscess isolate; sequenced	(444)
MGH78578	KR640	*K. pneumoniae* sputum isolate; sequenced	Washington University

[a] KR number corresponds to a physical catalogue in Dr. Kumar Rajakumar's laboratory. [b] UHL-LRI, University Hospitals Leicester-Leicester Royal Infirmary

Appendix 4. *Klebsiella*, *E. coli* and *Serratia* mutant strains

KR number[a]	Strain designation	Relevant characteristics[b]	Resistance[c]	Method[d]	Reference
***K. pneumoniae* KR116**					
Merodiploids					
KR921	KR116 *fim2K*::pJKO-4a[a]	pJKO-4a inserted within the *fim2K* locus of KR116; orientation a	Kan, Cml	AE	This work
KR922	KR116 *fim2K*::pJKO-4a[b]	pJKO-4a inserted within the *fim2K* locus of KR116; orientation b	Kan, Cml	AE	This work
KR1026	KR116 *fimK*::pJKO-1b[a]	pJKO-1b inserted within the *fimK* locus of KR116; orientation a	Gen, Cml	AE	This work
KR1027	KR116 *fimK*::pJKO-1b[b]	pJKO-1b inserted within the *fimK* locus of KR116; orientation b	Gen, Cml	AE	This work
KR1076	KR116 *fim*::pJKO-5b[a]	pJKO-5b inserted within the *fim* locus of KR116; orientation a	Gen, Cml	AE	This work
KR1104	KR116Δ*fim2K fim*::pJKO-5b [a]	pJKO-5b inserted within the *fim* locus of KR116Δ*fim2K::*Kan; orientation a	Gen, Cml, Kan	AE	This work
KR1724	KR116 *fim2*::pJKO-12a	pJKO-12a inserted within the *fim2* locus of KR116; unknown orientation	Gen, Cml	AE	This work
KR1725	KR116 KpGI-5::pJKO-13b	pJKO-13b inserted within the KpGI-5 locus of KR116; unknown orientation	Gen, Cml	AE	This work
KR1863	KR116Δ*fim2*$_{FRT}$ *fim*::pJKO-5b	pJKO-5b inserted within the *fim2* locus of KR116Δ*fim2::*FRT; unknown orientation	Gen, Cml	AE	This work
Mutants					
KR916	KR116Δ*fimK*	KR116 with an insertion/deletion mutation of *fimK*; Δ*fimK*::gen	Gen	AE	This work
KR919	KR116Δ*fimK*$_{FRT}$	KR116 with an insertion/deletion mutation of *fimK*; Δ*fimK*::FRT		AE	This work
KR1023	KR116Δ*fim2K*Δ*fimK*	KR116 with insertion/deletion mutations of *fim2K* and *fimK*; Δ*fim2K*::kan, Δ*fimK*::gen	Kan, Gen	AE	This work
KR1024	KR116Δ*fim2K*	KR116 with an insertion/deletion mutation of *fim2K*; Δ*fim2K*::kan	Kan	AE	This work
KR1112	KR116Δ*fim2K*Δ*fim*	KR116 with insertion/deletion mutations of *fim2K* and *fim*; Δ*fim2K*::kan, Δ*fim*::gen	Kan, Gen	AE	This work
KR1816	KR116Δ*fim2*	KR116 with an insertion/deletion mutation of *fim2*; Δ*fim2*::gen	Gen	AE	This work
KR1817	KR116ΔKpGI-5	KR116 with an insertion/deletion mutation of the KpGI-5 island; ΔKpGI-5::gen	Gen	AE	This work
KR1861	KR116Δ*fim2*$_{FRT}$	KR116 with an insertion/deletion mutation of *fim2*; Δ*fim2*::FRT		AE	This work

KR number[a]	Strain designation	Relevant characteristics[b]	Resistance[c]	Method[d]	Reference
KR1864	KR116Δ*fim2*$_{FRT}$Δ*fim*	KR116 with insertion/deletion mutations of *fim2* and *fim;* Δ*fim2*::FRT Δ*fim*::gen	Gen	AE	This work
KR1958	KR116Δ*fim2*$_{FRT}$Δ*fim*$_{FRT}$	KR116 with insertion/deletion mutations of *fim2* and *fim;* Δ*fim2*::FRT Δ*fim*::FRT		AE	This work
KR2106	KR116Δ*fim*	KR116 with an insertion/deletion mutation of *fim*; Δ*fim*::tet	Tet	LRR	This work
KR2107	-	Spontaneous streptomycin-resistant derivative of KR116	St		
KR2128	KR2107Δ*fim*	KR2107 with an insertion/deletion mutation of *fim*; Δ*fim*::tet	St, Tet	LRR	This work
KR2130	KR2107Δ*fim2*	KR2107 with an insertion/deletion mutation of *fim2*; Δ*fim2*::kan	St, Kan	LRR	This work
KR2152	KR2107Δ*fim*Δ*fim2*	KR2107 with insertion/deletion mutations of *fim* and *fim2;* Δ*fim2*::kan Δ*fim*::tet	St, Kan, Tet	LRR	This work
KR2301	KR2107Δ*fim2H*	KR2107 with an insertion/deletion mutation of *fim2H*; Δ*fim2H*::kan	St, Kan	LRR	This work
KR2302	KR2107Δ*fim*Δ*fim2H*	KR2107 with insertion/deletion mutations of *fim* and *fim2H;* Δ*fim*::tet Δ*fim2H*::kan	St, Kan, Tet	LRR	This work
KR2303	KR2107Δ*fimH*	KR2107 with an insertion/deletion mutation of *fimH*; Δ*fimH*::tet	St, Tet	LRR	This work
KR2304	KR2107Δ*fim2*Δ*fimH*	KR2107 with insertion/deletion mutations of *fim2* and *fimH;* Δ*fim2*::kan Δ*fimH*::tet	St, Kan, Tet	LRR	This work
KR2305	KR2107Δ*fim*Δ*mrk*	KR2107 with insertion/deletion mutations of *fim* and *mrk;* Δ*fim*::tet Δ*mrk*$_{C3091}$::kan	St, Kan, Tet	LRR	This work
KR2324	KR2107Δ*fim2H*Δ*fimH*	KR2107 with insertion/deletion mutations of *fim2H* and *fimH;* Δ*fim2H*::kan Δ*fimH*::tet	St, Kan, Tet	LRR	This work
KR2325	KR2107Δ*fim2H*$_{FRT}$	KR2107 with an insertion/deletion mutation of *fim2H*; Δ*fim2H*::FRT	St	LRR	This work
KR2326	KR2107Δ*fim*Δ*fim2H*$_{FRT}$	KR2107 with insertion/deletion mutations of *fim* and *fim2H;* Δ*fim*::tet Δ*fim2H*::FRT	St, Tet	LRR	This work
KR2327	KR2107Δ*fimH*$_{FRT}$	KR2107 with an insertion/deletion mutation of *fimH;* Δ*fimH*::FRT	St	LRR	This work
KR2328	KR2107Δ*fim2*$_{FRT}$Δ*fimH*$_{FRT}$	KR2107 with insertion/deletion mutations of *fim2* and *fimH;* Δ*fim2*::FRT Δ*fimH*::FRT	St	LRR	This work
KR2329	KR2107Δ*fim2H*$_{FRT}$Δ*fimH*$_{FRT}$	KR2107 with insertion/deletion mutations of *fim2H* and *fimH;* KR2107 Δ*fim2H*::FRT Δ*fimH*::FRT	St	LRR	This work
Mini-Tn7 insertions					
KR2231	KR116/Kan-*aadA1a*	KR116 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::Kan-*aadA1a*	Kan, St	mini-Tn*7*	This work

KR number[a]	Strain designation	Relevant characteristics[b]	Resistance[c]	Method[d]	Reference
KR2257	KR116/*aadA1a*	KR116 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::*aadA1a*	St	mini-Tn*7*	This work
KR2258	KR116/P_{LAC}*strAB*	KR116 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::P_{LAC}*strAB*	St	mini-Tn*7*	This work
KR2259	KR116/3St	KR116 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::3St	St	mini-Tn*7*	This work
KR2281	KR116/3St-rev	KR116 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::3St-rev	St	mini-Tn*7*	This work
***K. pneumoniae* KR161**					
Merodiploids					
KR980	KR161 *fimK*::pJKO-1b	pJKO-1b inserted within the *fimK* locus of KR161; orientation unknown	Cml, Gen	AE	This work
KR1077	KR161 *fim*::pJKO-7b	pJKO-7b inserted within the *fim* locus of KR161; orientation unknown	Cml, Gen	AE	This work
Mutants					
KR978	KR161Δ*fimK*	KR161 with an insertion/deletion mutation of *fimK;* Δ*fimK*::gen	Gen	AE	This work
KR1143	KR161Δ*fim*	KR161 with an insertion/deletion mutation of *fim;* Δ*fim*::gen	Gen	AE	This work
***K. pneumoniae* KR162**					
Merodiploids					
KR1003	KR162 *fimK*::pJKO-1b	pJKO-1b inserted within the *fimK* locus of KR162; orientation unknown	Cml, Gen	AE	This work
KR1079	KR162 *fim*::pJKO-8b	pJKO-8b inserted within the *fim* locus of KR162; orientation unknown	Cml, Gen	AE	This work
Mutants					
KR1001	KR162Δ*fimK*	KR162 with an insertion/deletion mutation of *fimK;* Δ*fimK*::gen	Gen	AE	This work
KR1145	KR162Δ*fim*	KR162 with an insertion/deletion mutation of *fim;* Δ*fim*::gen	Gen	AE	This work
Mini-Tn7 insertions					
KR2232	KR162/Kan-3St	KR162 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::Kan-3St	Kan, St	mini-Tn*7*	This work
KR2233	KR162Δ*fimK*/Kan-3St	KR162Δ*fimK* with a mini-Tn7 insertion within the *att*Tn*7* site; Δ*fimK*::gen, *att*Tn*7*::Kan-3St	Kan, St, Gen	mini-Tn*7*	This work
KR2236	KR162Δ*fimK*/FRT-3St	KR162Δ$fimK_{FRT}$ with a mini-Tn7 insertion within the *att*Tn*7* site; Δ*fimK*::FRT, *att*Tn*7*::FRT-3St	St, Gen	mini-Tn*7*	This work

KR number[a]	Strain designation	Relevant characteristics[b]	Resistance[c]	Method[d]	Reference
***K. pneumoniae* KR173**					
Merodiploids					
KR1007	KR173 *fimK*::pJKO-1b	pJKO-1b inserted within the *fimK* locus of KR173; orientation unknown	Cml, Gen	AE	This work
KR1081	KR173 *fim*::pJKO-9b	pJKO-9b inserted within the *fim* locus of KR173; orientation unknown	Cml, Gen	AE	This work
Mutants					
KR1005	KR173Δ*fimK*	KR173 with an insertion/deletion mutation of *fimK;* Δ*fimK*::gen	Gen	AE	This work
KR1147	KR173Δ*fim*	KR173 with an insertion/deletion mutation of *fim;* Δ*fim*::gen	Gen	AE	This work
Mini-Tn7 insertions					
KR2234	KR173/Kan-3St	KR173 with a mini-Tn7 insertion within the *att*Tn*7* site; *att*Tn*7*::Kan-3St	Kan, St	mini-Tn*7*	This work
KR2235	KR173Δ*fimK*/Kan-3St	KR173Δ*fimK* with a mini-Tn7 insertion within the *att*Tn*7* site; Δ*fimK*::gen, *att*Tn*7*::Kan-3St	Kan, St, Gen	mini-Tn*7*	This work
KR2237	KR173Δ*fimK*/FRT-3St	KR173Δ$fimK_{FRT}$ with a mini-Tn7 insertion within the *att*Tn*7* site; Δ*fimK*::FRT, *att*Tn*7*::FRT-3St	St, Gen	mini-Tn*7*	This work
Other mutant strains					
KR1549	ROP4S	*Serratia* sp. ATCC 39006 *lac*$^-$ *pigX*::mini-Tn5Sm/Sp; insertion at + 259bp in *pigX* transcript	Sp	mini-Tn*5*	(118)
KR2020	C3091Δ*fimK*	*K. pneumoniae* C3091 with an insertion/deletion mutation of *fimK*; Δ*fimK*::kan	Kan	LRR	C. Struve
KR2021	C3091Δ*mrk*Δ*fim*	*K. pneumoniae* C3091 with insertion/deletion mutations of *mrk* and *fim;* Δmrk_{C3091}::kan Δ*fim*::tet	Kan, Tet	LRR	(388)
KR2576	KBC216Δ*fimH*	*E. coli* strain KBC216 with an interrupted *fimH* gene		Unknown	C. Struve
KR2577	MG1655Δ*fimH*	*E. coli* strain MG1655 with an interrupted *fimH* gene		Unknown	C. Struve

[a] KR number corresponds to a physical catalogue in Dr. Kumar Rajakumar's laboratory. [b] Relevant details on the construction of suicide vectors and mutant alleles can be found in section 2.5.2. [c] Cml, chloramphenicol; Gen, gentamicin; Kan, kanamycin; Tet, tetracycline; Sp, spectinomycin; St, streptomycin. [d] AE, suicide vector-based allelic exchange; LRR, lambda Red recombination.

Appendix 5. Commercial and published plasmid tools used in this work

Plasmid name	pKR number[a]	Relevant characteristics[b]	Resistance[c]	Reference
Plasmids for cloning				
pBluescript II KS+	pKR1	HCN cloning vector; P_{LAC}*lacZα*	Amp	Strategene
pGEM-T easy	-	HCN cloning vector; P_{LAC}*lacZα*	Amp	Promega
pJET1.2	-	HCN cloning vector	Amp	Fermentas
pUCP20	pKR31	HCN cloning vector; P_{LAC}*lacZα*	Amp	(435)
pUCP24	pKR32	HCN cloning vector; P_{LAC}*lacZα*	Gen	(435)
pWSK29	pKR11	LCN cloning vector; P_{LAC}*lacZα*	Amp	(432)
pWSK129	pKR12	LCN cloning vector; P_{LAC}*lacZα*	Kan	(432)
pCC2FOS	-	Fosmid cloning vector	Cml	Epicentre
Plasmids for gene expression				
pET-28a	- (KR 908)	Vector with phage T7 promoter for expression of N-terminal 6xHis tag fusion proteins	Kan	Novagen
pETM-22	pKR361	Vector with phage T7 promoter for expression of N-terminal 6xHis and thioredoxin fusion-proteins	Kan	EMBL Pepcore
pETM-44	pKR362	Vector with phage T7 promoter for expression of N-terminal 6xHis and maltose binding protein fusion proteins	Kan	EMBL Pepcore
pETM-66	pKR363	Vector with phage T7 promoter for expression of N-terminal 6xHis and NusA fusion proteins	Kan	EMBL Pepcore
pGB17	pKR460	Vector with P_{BLA} promoter for constitutive gene expression; possesses a cloned *fimH* gene	Cml	(377)
pGEX-2T	pKR324	Vector with P_{TAC} promoter for expression of N-terminal GST-tagged proteins	Amp	GE Healthcare
pMQ300	pKR431	Expression vector with P_{LAC}*lacZα*; Broad host range pBBR1 origin of replication	Hyg	(189)
pTRC99a	- (KR446)	Expression vector with P_{TRC} promoter and $lacI^Q$	Amp	(10)
pQE-80L	pKR313	Expression vector with phage T5 promoter and $lacI^Q$	Amp	Qiagen

Plasmid name	pKR number[a]	Relevant characteristics[b]	Resistance[c]	Reference
Plasmids for genetic manipulation				
pDS132	pKR14	Lambda *pir*-based suicide vector	Cml	(296)
pFLP2	pKR19	FLP recombinase-expressing plasmid	Amp	(163)
pKOBEG-Apra	pKR316	Lambda Red expression plasmid, P_{BAD} promoter	Apr	(58)
pRK2013	- (KR777)	Helper plasmid for mobilisation of non-self-transmissible plasmids	Kan	(117)
pTNS2	- (KR740)	Transposition helper plasmid, encoding the site-specific TnsABC+ D Tn*7* transposition pathway	Amp	(66)
pUC18R6kT-mini-Tn7T	- (KR743)	Lambda *pir*-based mini-Tn*7* delivery plasmid	Amp	(66)
Plasmids with other functions				
pCLL3468	pKR240	pCR2.1-TOPO harbouring a minocycline resistance cassette from *Staphylococcus aureus* Mu3	Min, Kan	(338)
pCAS624	pKR360	pUC18 clone harbouring the *fim* operon of *K. pneumoniae* C3091	Amp	(387)
pRARE	pKR408	Plasmid expressing rare tRNA genes	Cml	Novagen
pRDH137	pKR34	Plasmid encoding *pir* gene	Kan	R. Haigh, unpublished
pRT733	pKR36	Source of kanamycin resistance cassette for SOE-PCR	Amp, Kan	(403)
pTA40	pKR314	Plasmid containing the *Serratia* sp. ATCC 39006 *pigx* gene	Amp	(118)
pUC18R6k-mini-Tn7T-Gm	- (KR744)	Source of gentamicin resistance cassette for SOE-PCR	Amp, Gen	(66)
pUC18R6k-mini-Tn7T-Km	- (KR747)	Source of kanamycin resistance cassette for SOE-PCR	Amp, Kan	(66)

[a] pKR number corresponds to a physical plasmid catalogue in Dr. Kumar Rajakumar's laboratory. [b] HCN, high copy number; LCN, low copy number. [c] Amp, ampicillin; Apr, apramycin; Cml, chloramphenicol; Gen, gentamicin; Hyg, hygromycin; Kan, kanamycin; Min, minocycline.

Appendix 6. Plasmids constructed in this work

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
Plasmid tools (pJTOOL-X)				
pJTOOL-1	-	pKR10	Lambda *pir*-based suicide vector with *Not*I cloning site. Derived from pDS132	Cml
pJTOOL-3	-	pKR35	Lambda *pir*-based suicide vector with expanded multiple cloning site. Derived from pDS132	Cml
pJTOOL-4a	-	pKR216	pGEM-T easy with the pUCP26 tetracycline resistance cassette flanked by FRT sites	Amp, Tet
pJTOOL-6a	-	pKR260	pGEM-T easy with the pCLL3468 minocycline resistance cassette flanked by FRT sites	Amp, Min
pJTOOL-6b	-	pKR261	pGEM-T easy with the pCLL3468 minocycline resistance cassette flanked by FRT sites	Amp, Min
pJTOOL-7	-	pKR377	pTRC99a-based expression vector with a *Not*I site added between the *Nco*I and *Eco*RI restriction sites	Amp
pJTOOL-8	pUC18R6kT-miniTn7T::Kan	pKR378	pUC18R6kT-miniTn7T harbouring an FRT-flanked kanamycin resistance cassette within the transposable mini-Tn7 site	Amp, Kan
pJTOOL-9	pUC18R6kT-miniTn7T::*aadA1a*	pKR379	pUC18R6kT-miniTn7T harbouring an *aadA1a* streptomycin resistance cassette within the transposable mini-Tn7 site	Amp, St
pJTOOL-10p	pUC18R6kT-mini-Tn7T::P_{LAC}*strAB*	pKR380	pUC18R6kT-miniTn7T harbouring a streptomycin resistance cassette (*strAB*) from *K. pneumoniae* MGH78578 under control of P_{LAC} within the transposable mini-Tn7 site	Amp, St
pJTOOL-11m	pUC18R6kT-mini-Tn7T::3St-rev	pKR404	pUC18R6kT-miniTn7T harbouring a streptomycin resistance cassette (*strAB*) from *K. pneumoniae* MGH78578 under control of P_{LAC} and an *aadA1a* streptomycin resistance cassette within the transposable mini-Tn7 site. *aadA1a* and *strAB* are divergent	Amp, St
pJTOOL-11p	pUC18R6kT-mini-Tn7T::3St	pKR405	pUC18R6kT-miniTn7T harbouring a streptomycin resistance cassette (*strAB*) from *K. pneumoniae* MGH78578 under control of P_{LAC} and an *aadA1a* streptomycin resistance cassette within the transposable mini-Tn7 site. The three cassettes are in tandem	Amp, St
pJTOOL-12	pUC18R6kT-miniTn7T::Kan-*aadA1a*	pKR381	pUC18R6kT-miniTn7T harbouring an FRT-flanked kanamycin resistance cassette and an *aadA1a* streptomycin resistance cassette within the transposable mini-Tn7 site	Amp, Kan, St
pJTOOL-13m	pUC18R6kT-mini-Tn7T::Kan-P_{LAC}*strAB*	pKR406	pUC18R6kT-miniTn7T harbouring an FRT-flanked kanamycin resistance cassette and a streptomycin resistance cassette (*strAB*) from *K. pneumoniae* MGH78578 under control of P_{LAC} within the transposable mini-Tn7 site	Amp, Kan, St

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJTOOL-14p	pUC18R6kT-mini-Tn7T::Kan-3St	pKR407	pUC18R6kT-miniTn7T harbouring an FRT-flanked kanamycin resistance cassette, a streptomycin resistance cassette (*strAB*) from *K. pneumoniae* MGH78578 under control of P_{LAC} and an *aadA1a* streptomycin resistance cassette within the transposable mini-Tn7 site	Amp, Kan, St
pJTOOL-15	pWSK29::*strAB*	pKR402	pWSK129 harbouring a streptomycin resistance cassette from *K. pneumoniae* MGH78578	Amp, St
pJTOOL-16	pUCP24::*sacB*-FLP2	pKR427	pUCP24 harbouring the *Sph*I/*Sac*I fragment from pFLP2 which encodes the *sacB* and *flp* recombinase genes	Gen
Suicide vectors (pJKO-X)				
pJKO-1a	pJTOOL-1::*fimK*::gen(+)[a]	pKR22	pJTOOL-1 harbouring the 2.95 kb *fimK*::gen(+) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-1b	pJTOOL-1::*fimK*::gen(+)[b]	pKR23	pJTOOL-1 harbouring the 2.95 kb *fimK*::gen(+) SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-2a	pJTOOL-1::*fimK*::gen(-)[a]	pKR25	pJTOOL-1 harbouring the 2.95 kb *fimK*::gen(-) allele within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-3a	pDS132::*fim2K*::gen(+)[a]	pKR27	pDS132 harbouring the 2.36 kb *fim2K*::kan(+) SOE-PCR product within the *Xba*I restriction site; orientation a	Cml, Kan
pJKO-3b	pDS132::*fim2K*::gen(+)[b]	pKR28	pDS132 harbouring the 2.36 kb *fim2K*::kan(+) SOE-PCR product within the *Xba*I restriction site; orientation b	Cml, Kan
pJKO-4a	pDS132::*fim2K*::gen(-)[a]	pKR30	pDS132 harbouring the 2.36 kb *fim2K*::kan(-) SOE-PCR product within the *Xba*I restriction site; orientation a	Cml, Kan
pJKO-5a	pJTOOL-1::fim_{KR116}::gen(+)[a]	pKR233	pJTOOL-1 harbouring the 2.9 kb fim_{KR116}::gen(+) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-5b	pJTOOL-1::fim_{KR116}::gen(+)[b]	pKR234	pJTOOL-1 harbouring the 2.9 kb fim_{KR116}::gen(+) SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-6a	pJTOOL-1::fim_{KR116}::gen(-)[a]	pKR235	pJTOOL-1 harbouring the 2.9 kb fim_{KR116}::gen(-) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-7a	pJTOOL-1::fim_{KR161}::gen(+)[a]	pKR243	pJTOOL-1 harbouring the 2.9 kb fim_{KR161}::gen(+) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJKO-7b	pJTOOL-1::*fim*$_{KR161}$::gen(+)[b]	pKR244	pJTOOL-1 harbouring the 2.9 kb *fim*$_{KR161}$::gen(+) SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-8a	pJTOOL-1::*fim*$_{KR162}$::gen(+)[a]	pKR245	pJTOOL-1 harbouring the 2.9 kb *fim*$_{KR162}$::gen(+) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-8b	pJTOOL-1::*fim*$_{KR162}$::gen(+)[b]	pKR246	pJTOOL-1 harbouring the 2.9 kb *fim*$_{KR162}$::gen(+) SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-9a	pJTOOL-1::*fim*$_{KR173}$::gen(+)[a]	pKR247	pJTOOL-1 harbouring the 2.9 kb *fim*$_{KR173}$::gen(+) SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-9b	pJTOOL-1::*fim*$_{KR173}$::gen(+)[b]	pKR248	pJTOOL-1 harbouring the 2.9 kb *fim*$_{KR173}$::gen(+) SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-12a	pJTOOL-3::*fim2*::gen[a]	pKR327	pJTOOL-3 harbouring the 2.72 kb *fim2*::gen SOE-PCR product within the *Not*I restriction site; orientation a	Cml, Gen
pJKO-12b	pJTOOL-3::*fim2*::gen[b]	pKR328	pJTOOL-3 harbouring the 2.72 kb *fim2*::gen SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-13b	pJTOOL-3::KpGI-5::gen[b]	pKR329	pJTOOL-3 harbouring the 2.77 kb KpGI-5::gen SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Gen
pJKO-14b	pJTOOL-1::*mrk*$_{KR116}$::kan[b]	pKR341	pJTOOL-1 harbouring the 2.72 kb *mrk*$_{KR116}$::kan SOE-PCR product within the *Not*I restriction site; orientation b	Cml, Kan
Expression plasmids (pJOE-X)				
pJOE-1	pET28a::*fim2K*	pKR273	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR501 and PR502 (980 bp) cloned into *Nde*I/*Not*I site of pET-28a; produces Fim2K::6xHis	Kan
pJOE-2	pET28a::*fimK*	pKR272	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR503 and PR504 (1415 bp) cloned into *Nde*I/*Not*I site of pET-28a; produces FimK::6xHis	Kan
pJOE-3	pET28a::*yahA_EAL*	pKR274	EAL domain-encoding portion of the *yahA* gene from *E. coli* K12 MG1655 amplified using primers PR505 and PR506 (869 bp) cloned into *Nde*I/*Not*I site of pET-28a; produces YahA_EAL::6xHis	Kan
pJOE-4	pET23a::*fim2K*	pKR277	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR501 and PR502 (980 bp) cloned into *Nde*I/*Not*I site of pET-23a; produces Fim2K::6xHis	Amp

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-5a	pFim$_{UTI89}$-HCN	pKR315	*fim* operon from *E. coli* UTI89 amplified using primers PR565 and PR566 (10.5 kb) and cloned into the *Xba*I site of pUCP20	Amp
pJOE-5LCNa	pFim$_{UTI89}$-LCN	pKR321	10.5 kb XbaI fragment from pJOE-5a containing the *E. coli fim* operon cloned into the *Xba*I site of pWSK129	Kan
pJOE-6	pET28a::*pigx*	pKR317	2.0 kb *Hind*III/*Eco*RI fragment from pTA40 containing the *Serratia* sp. ATCC 39006 *pigx* gene cloned into the *Hind*III/*Eco*RI site of pET28a; produces PigX::6xHis	Kan
pJOE-7a	pFim-HCN	pKR320	*fim* operon from *K. pneumoniae* KR116 amplified using primers PR563 and PR564 (10.5 kb) and cloned into the *Xba*I site of pUCP20	Amp
pJOE-7LCNb	pFim-LCN	pKR322	10.5 kb *Xba*I fragment from pJOE-7a containing the *K. pneumoniae fim* operon cloned into the *Xba*I site of pWSK129	Kan
pJOE-9	pGEX-2T::*fim2K*	pKR325	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR929 and PR930 (981 bp) cloned into *Bam*HI/*Eco*RI site of pET-28a; produces Fim2K::GST	Amp
pJOE-10	pGEX-2T::*fimK*	pKR326	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR927 and PR928 (1413 bp) cloned into *Bam*HI/*Eco*RI site of pET-28a; produces FimK::GST	Amp
pJOE-11b	pFim2-HCN	pKR330	*fim2* operon (*fim2A* to *fim2K*) from *K. pneumoniae* KR116 amplified using primers PR937 and PR938 (8.2 kb) and cloned into the *Not*I site of pBluescript II KS+	Amp
pJOE-11LCNa	pFim2-LCN[a]	pKR332	8.2 kb *Not*I fragment from pJOE-11b containing the *K. pneumoniae fim2* operon (*fimA* to *fimK*) cloned into the *Not*I site of pWSK129; orientation a	Kan
pJOE-11LCNb	pFim2-LCN[b]	pKR333	8.2 kb *Not*I fragment from pJOE-11b containing the *K. pneumoniae fim2* operon (*fimA* to *fimK*) cloned into the *Not*I site of pWSK129; orientation b	Kan
pJOE-12LCNb	pFim2(orf10-2K)-LCN	pKR331	*fim2* operon (*orf10* to *fim2K*) from *K. pneumoniae* KR116 amplified using primers PR937 and PR939 (9.1 kb) and cloned into the *Not*I site of pBluescript II KS+	Kan
pJOE-13	pET-28a::*int-met56*$_{78578}$	pKR334	*met56*-associated putative integrase gene from *K. pneumoniae* MGH78578 amplified using primers PR1022 and PR1021 (1114 bp) and cloned into the *Nde*I/*Not*I site of pET-28a; produces Int-met56$_{78578}$::6xHis	Kan
pJOE-14	pGEX-2T::*pigX_GGDEF_EAL*	pKR335	*Serratia* sp. ATCC 39006 *pigX* GGDEF and EAL domain-encoding region amplified using primers PR1019 and PR1020 (1332 bp) and cloned into the *Eco*RI/*Bam*HI site of pGEX-2T; produces PigX_GGDEF_EAL::GST	Amp

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-15	pGEX-2T::*pigX*_EAL	pKR336	*Serratia* sp. ATCC 39006 *pigX* EAL domain-encoding region amplified using primers PR1019 and PR1020 (834 bp) and cloned into the *Eco*RI/*Bam*HI site of pGEX-2T; produces PigX_EAL::GST	Amp
pJOE-19	pQE-80L::*yahA*_EAL	pKR340	EAL domain-encoding portion of the *yahA* gene from *E. coli* K-12 MG1655 amplified using primers PR1069 and PR1070 (866 bp) cloned into *Sac*I/*Pst*I site of pQE-80L; produces YahA_EAL::6xHis	Amp
pJOE-20	pQE-80L::*fimK*	pKR357	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1136 and PR1137 (1417 bp) cloned into *Bam*HI/*Kpn*I site of pQE-80L; produces FimK::6xHis	Amp
pJOE-21	pQE-80L::*fim2K*	pKR358	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1134 and PR1135 (982 bp) cloned into *Bam*HI/*Kpn*I site of pQE-80L; produces Fim2K::6xHis	Amp
pJOE-22	pQE-80L::*int-met56*$_{78578}$	pKR359	*met56*-associated putative integrase gene from *K. pneumoniae* MGH78578 amplified using primers PR1138 and PR1139 (1121 bp) and cloned into the *Bam*HI/*Kpn*I site of pQE-80L; produces Int-met56$_{78578}$::6xHis	Amp
pJOE-23	pETM-22::*fimK*	pKR365	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1206 and PR1207 (1415 bp) cloned into *Nco*I/*Bam*HI site of pETM-22; produces FimK::6xHis::Trx	Kan
pJOE-24	pETM-22::*fimK_EAL*	pKR366	EAL domain-encoding portion of *fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1208 and PR1209 (752 bp) cloned into *Nco*I/*Bam*HI site of pETM-22; produces FimK_EAL::6xHis::Trx	Kan
pJOE-25	pETM-22::*fim2K*	pKR367	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1202 and PR1203 (977 bp) cloned into *Nco*I/*Bam*HI site of pETM-22; produces Fim2K::6xHis::Trx	Kan
pJOE-26	pETM-22::*fim2K_EAL*	pKR368	EAL domain-encoding portion of *fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1204 and PR1205 (770 bp) cloned into *Nco*I/*Bam*HI site of pETM-22; produces Fim2K_EAL::6xHis::Trx	Kan
pJOE-27	pETM-44::*fimK*	pKR369	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1206 and PR1207 (1415 bp) cloned into *Nco*I/*Bam*HI site of pETM-44; produces FimK::6xHis::MBP	Kan
pJOE-28	pETM-44::*fimK_EAL*	pKR370	EAL domain-encoding portion of *fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1208 and PR1209 (752 bp) cloned into *Nco*I/*Bam*HI site of pETM-44; produces FimK_EAL::6xHis::MBP	Kan
pJOE-29	pETM-44::*fim2K*	pKR371	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1202 and PR1203 (977 bp) cloned into *Nco*I/*Bam*HI site of pETM-44; produces Fim2K::6xHis::MBP	Kan

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-30	pETM-44::*fim2K_EAL*	pKR372	EAL domain-encoding portion of *fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1204 and PR1205 (770 bp) cloned into *Nco*I/*Bam*HI site of pETM-44; produces Fim2K_EAL::6xHis::MBP	Kan
pJOE-31	pETM-66::*fimK*	pKR373	*fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1206 and PR1207 (1415 bp) cloned into *Nco*I/*Bam*HI site of pETM-66; produces FimK::6xHis::NusA	Kan
pJOE-32	pETM-66::*fimK_EAL*	pKR374	EAL domain-encoding portion of *fimK* gene from *K. pneumoniae* KR116 amplified using primers PR1208 and PR1209 (752 bp) cloned into *Nco*I/*Bam*HI site of pETM-66; produces FimK_EAL::6xHis::NusA	Kan
pJOE-33	pETM-66::*fim2K*	pKR375	*fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1202 and PR1203 (977bp) cloned into *Nco*I/*Bam*HI site of pETM-66; produces Fim2K::6xHis::NusA	Kan
pJOE-34	pETM-66::*fim2K_EAL*	pKR376	EAL domain-encoding portion of *fim2K* gene from *K. pneumoniae* KR116 amplified using primers PR1204 and PR1205 (770 bp) cloned into *Nco*I/*Bam*HI site of pETM-66; produces Fim2K_EAL::6xHis::NusA	Kan
pJOE-35	pFim-Ptrc	pKR409	*fim* operon (*fimA* to *fimK*) from *K. pneumoniae* C3091 amplified using primers PR1268 and PR1221 (8.1 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-36	pFim2-Ptrc	pKR401	*fim2* operon (*fim2A* to *fim2K*) from *K. pneumoniae* KR116 amplified using primers PR1222 and PR1224 (7.8 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-37	pFim2(orf10-2K)-Ptrc	pKR403	*fim2* locus (*orf10* to *fim2K*) from *K. pneumoniae* KR116 amplified using primers PR1223 and PR1224 (8.7 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-38	pfimH	pKR410	*fimH* gene from *K. pneumoniae* KR116 amplified using primers PR1313 and PR1318 (919 bp) and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces FimH	Amp
pJOE-39	pfimH_FLAG225	pKR411	*fimH*::$FLAG_{225}$ SOE-PCR product (943 bp) derived from *K. pneumoniae* KR116 and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces FimH::$FLAG_{225}$	Amp
pJOE-40	pfimH_FLAG258	pKR412	*fimH*::$FLAG_{258}$ SOE-PCR product (943 bp) derived from *K. pneumoniae* KR116 and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces FimH::$FLAG_{258}$	Amp
pJOE-41	pfim2H	pKR413	*fim2H* gene from *K. pneumoniae* KR116 amplified using primers PR1319 and PR1324 (1038 kb) and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces Fim2H	Amp
pJOE-42	pfim2H_FLAG225	pKR414	*fim2H*::$FLAG_{225}$ SOE-PCR product (1062 bp) derived from *K. pneumoniae* KR116 and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces Fim2H::$FLAG_{225}$	Amp

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-43	pfim2H_FLAG258	pKR415	*fim2H*::$FLAG_{258}$ SOE-PCR product (1062 bp) derived from *K. pneumoniae* KR116 and cloned into the *Nco*I/*Kpn*I site of pTRC99a; produces Fim2H::$FLAG_{258}$	Amp
pJOE-44	pQE-80L::*pigX_GGDEF*	pKR416	GGDEF domain-encoding region of the *pigX* gene from *Serratia* sp. ATCC 39006 amplified using primers PR1325 and PR1326 (507 bp) cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces PigX_GGDEF::6xHis	Amp
pJOE-45	pQE-80L::*pigX_EAL*	pKR417	EAL domain-encoding region of the *pigX* gene from *Serratia* sp. ATCC 39006 amplified using primers PR1327and PR1328 (832 bp) cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces PigX_EAL::6xHis	Amp
pJOE-47	pQE-80L::*kpn03660_EAL*	pKR419	EAL domain-encoding region of the *kpn03660* gene from *K. pneumoniae* MGH78578 amplified using primers PR1335and PR1336 (867 bp) cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces Kpn03660_EAL::6xHis	Amp
pJOE-48	pQE-80L::*rocR_EAL*	pKR420	EAL domain-encoding region of the *rocR* gene from *P. aeruginosa* PA01 amplified using primers PR1337 and PR1338 (801 bp) cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces RocR_EAL::6xHis	Amp
pJOE-49	pQE-80L::*fimX_EAL*	pKR422	EAL domain-encoding region of the *fimX* gene from *P. aeruginosa* PA01 amplified using primers PR1329 and PR1330 (806 bp) cloned into the *Bam*HI/*Sal*I site of pQE-80L; produces FimX_EAL::6xHis	Amp
pJOE-50	pQE-80L::*kpn00268_EAL*	pKR423	EAL domain-encoding region of the *kpn00268* gene from *K. pneumoniae* MGH78578 amplified using primers PR1333 and PR1334 (912 bp) cloned into the *Bam*HI/*Sal*I site of pQE-80L; produces Kpn00268_EAL::6xHis	Amp
pJOE-51	pQE-80L::*kpn03660_GGDEF*	pKR424	GGDEF domain-encoding region of the *kpn03660* gene from *K. pneumoniae* MGH78578 amplified using primers PR1364 and PR1365 (535 bp) cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces Kpn03660_GGDEF::6xHis	Amp
pJOE-52	pQE-80L::*kpn03660_EAL_ GGDEF*	pKR425	EAL and GGDEF domain-encoding region of the *kpn03660* gene from *K. pneumoniae MGH78578* amplified using primers PR1364 and PR1336 (1329 bp) and cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces Kpn03660_EAL_GGDEF::6xHis	Amp
pJOE-53	pQE-80L::*kpn3660*	pKR426	*kpn3660* gene from *K. pneumoniae* MGH78578 amplified using primers PR1366 and PR1336 (1965 bp) and cloned into the *Sac*I/*Kpn*I site of pQE-80L; produces Kpn03660::6xHis	Amp
pJOE-54-1	p$Fim2_{sp25}$-Ptrc	pKR436	*fim2* operon (*fim2A* to *fim2K*) from *K. pneumoniae* KR2415 amplified using primers PR1222 and PR1224 (7.8 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-55-1 to -6	pFim2$_{KR518}$-Ptrc	pKR436-42	*fim2* operon (*fim2A* to *fim2K*) from *K. pneumoniae* KR518 amplified using primers PR1222 and PR1224 (9.0 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-56-1	pFim2(orf10-2K)$_{sp25}$-Ptrc	pKR443	*fim2* locus (*orf10* to *fim2K*) from *K. pneumoniae* KR2415 amplified using primers PR1223 and PR1224 (8.7 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-57-1 to -5	pFim2(orf10-2K)$_{KR518}$-Ptrc	pKR444-48	*fim2* locus (*orf10* to *fim2K*) from *K. pneumoniae* KR518 amplified using primers PR1223 and PR1224 (9.9 kb) and cloned into the *Not*I/*Sbf*I site of pJTOOL-7	Amp
pJOE-58a	pMQ300::*fimH*[a]	pKR449	*fimH* amplified from pJOE-38 using primers PR1632 and PR1633 (963 bp) and cloned into the *Sph*I site of pMQ300; same orientation as P_{LAC} (orientation a); produces FimH	Hyg
pJOE-58b	pMQ300::*fimH*[b]	pKR450	*fimH* amplified from pJOE-38 using primers PR1632 and PR1633 (963 bp) and cloned into the *Sph*I site of pMQ300; opposite orientation to P_{LAC} (orientation b)	Hyg
pJOE-59a	pMQ300::*fimH_FLAG225*[a]	pKR451	*fimH*::FLAG$_{225}$ SOE-PCR product amplified from pJOE-39 using primers PR1632 and PR1633 (987 bp) and cloned into the *Sph*I site of pMQ300; same orientation as P_{LAC} (orientation a); produces FimH::FLAG$_{225}$	Hyg
pJOE-59b	pMQ300::*fimH_FLAG225*[b]	pKR452	*fimH*::FLAG$_{225}$ SOE-PCR product amplified from pJOE-39 using primers PR1632 and PR1633 (987 bp) and cloned into the *Sph*I site of pMQ300; opposite orientation to P_{LAC} (orientation b)	Hyg
pJOE-60a	pMQ300::*fim2H*[a]	pKR453	*fim2H* amplified from pJOE-41 using primers PR1632 and PR1633 (1082 bp) and cloned into the *Sph*I site of pMQ300; same orientation as P_{LAC} (orientation a); produces Fim2H	Hyg
pJOE-60b	pMQ300::*fim2H*[b]	pKR454	*fim2H* amplified from pJOE-41 using primers PR1632 and PR1633 (1082 bp) and cloned into the *Sph*I site of pMQ300; opposite orientation to P_{LAC} (orientation b)	Hyg
pJOE-61a	pMQ300::*fim2H_FLAG225*[a]	pKR455	*fimH*::FLAG$_{225}$ SOE-PCR product amplified from pJOE-42 using primers PR1632 and PR1633 (1106 bp) and cloned into the *Sph*I site of pMQ300; same orientation as P_{LAC} (orientation a); produces Fim2H::FLAG$_{225}$	Hyg
pJOE-62	pGB17::*fimH*	pKR456	*fimH* amplified from pJOE-38 using primers PR1667 and PR1668 (963 bp) and cloned into the *Apa*LI/*Hind*III site of pGB17; produces FimH	Cml
pJOE-63	pGB17::*fimH_FLAG225*	pKR457	*fimH*::FLAG$_{225}$ SOE-PCR product amplified from pJOE-39 using primers PR1667 and PR1668 (987 bp) and cloned into the *Apa*LI/*Hind*III site of pGB17; produces FimH::FLAG$_{225}$	Cml
pJOE-64	pGB17::*fim2H*	pKR458	*fim2H* amplified from pJOE-41 using primers PR1667 and PR1668 (1082 bp) and cloned into the *Apa*LI/*Hind*III site of pGB17; produces Fim2H	Cml

Plasmid name	Alternative name[a]	pKR number[b]	Description	Resistance[c]
pJOE-65	pGB17::*fim2H_FLAG225*	pKR459	*fim2H*::$FLAG_{225}$ SOE-PCR product amplified from pJOE-42 using primers PR1667 and PR1668 (1106 bp) and cloned into the *Apa*LI/*Hind*III site of pGB17; produces Fim2H::$FLAG_{225}$	Cml
Fosmid clones (pJFos-X)				
pJFos-1	-	pKR252	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. Fully sequenced	Cml, Kan
pJFos-2	-	pKR253	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. Only end sequencing of insert	Cml, Kan
pJFos-3	-	pKR254	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. Only end sequencing of insert	Cml, Kan
pJFos-4	-	pKR255	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. Fully sequenced	Cml, Kan
pJFos-5	-	pKR256	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. No sequence available.	Cml, Kan
pJFos-6	-	pKR257	Kanamycin resistant fosmid clone obtained by marker rescue from a fosmid library generated from strain *KR116*Δ*fim2K* and pCC2FOS. Only end sequencing of insert	Cml, Kan

[a] Alternative name is a descriptive plasmid name and may be used in the text to improve readability. [b] pKR number corresponds to a physical plasmid catalogue in Dr. Kumar Rajakumar's laboratory. [c] Amp, ampicillin; Apr, apramycin; Cml, chloramphenicol; Gen, gentamicin; Hyg, hygromycin; Kan, kanamycin; Min, minocycline; St, streptomycin; Tet, tetracycline.

Appendix 7. Primer table

Primer name	Catalogue number[a]	Sequence[b]	Special features
T3	PR78	AATTAACCCTCACTAAAGGG	-
T7	PR79	GTAATACGACTCACTATAGGGC	-
pCC2Fwd	PR323	GTACAACGACACCTAGAC	-
pCC2Rev	PR324	CAGGAAACAGCCTAGGAA	-
55pheDA_lf_NotI	PR347	GC**GCGGCCGC**ATCAGGCATGAACCAAATCC	*Not*I
55pheDA_lr_GmR	PR348	**AGGAACTTCAAGATCCCCAATTCG**TTGAGCAACGCATCATTATGG	SOE-PCR overlap
55pheDF_rr_NotI	PR349	GC**GCGGCCGC**GATCAATAAGATCGCGTTCGG	*Not*I
55pheDF_rf_GmF	PR350	**TCAGAGCGCTTTTGAAGCTAATTCG**TAAGATGCTATTTGGCCATCTG	SOE-PCR overlap
fimK2_fwd_NdeI	PR501	GGAATTC**CATATG**GTTGAGTCTTTTACGCAGC	*Nde*I
fimK2_rev_NotI	PR502	ATAAGAAT**GCGGCCGC**GAATAACTCGTCCTTTATTCCC	*Not*I
fimK_fwd_NdeI	PR503	GGAATTC**CATATG**ACCGATTATATCCTCTCGCCC	*Nde*I
fimK_rev_NotI	PR504	ATAAGAAT**GCGGCCGC**TCAACGTTTCGCCGGATCGC	*Not*I
yahAEAL_fwd_NdeI	PR505	GGAATTC**CATATG**CGCGATATTTTCTTTCAGTAC	*Nde*I
yahAEAL_rev_NotI	PR506	ATAAGAAT**GCGGCCGC**TCAACCACCTGCTTTCATTAC	*Not*I
GmF	PR550	CGAATTAGCTTCAAAAGCGCTCTGA	-
GmR	PR551	CGAATTGGGGATCTTGAAGTTCCT	-
aacC1F	PR552	GACATAAGCCTGTTCGGTT	-
aacC1R	PR553	CTCCGAACTCACGACCGA	-
BlaF	PR554	CAGTGCTGCAATGATACCGCG	-
BlaR	PR555	CCCGAAGAACGTTTTCCAATG	-
T7 term	PR558	CTAGTTATTGCTCAGCGGT	-
UTI89_fimS_fwd	PR559	CCGTAACGCAGACTCATCCTC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
UTI89_fimS_rev	PR560	GACAGAACAACAATTGCCAG	-
KR116_fimS_fwd	PR561	GGGACAGATACGCGTTTGAT	-
KR116_fimS_rev	PR562	GGCCTAACTGAACGGTTTGA	-
Kp_opfim_fwd_Xba	PR563	CTAG**TCTAGA**TGCGGGTATCATCAAGAGG	*Xba*I
Kp_opfim_rev_Xba	PR564	CTAG**TCTAGA**CGATAACACCCGCGAATACGAC	*Xba*I
Ec_opfim_fwd_XbaI	PR565	CTAG**TCTAGA**ACGAGATTTGCAATTCGTGTCACAAAAT	*Xba*I
Ec_opfim_rev_XbaI	PR566	CTAG**TCTAGA**GTTTGGGCCGTACCAGCGCC	*Xba*I
56metU	PR601	GGTTGATACCGCAAAAGACAA	-
56metD	PR602	TGCCAGACAGGTACTGATTGATATA	-
116_met56_fimK2_walk	PR614	CACTTCAATGCCTTTAATATGCAG	-
fimK2F	PR615	TCAACCGGCTGCGAAAACAAAT	-
fimK2R	PR616	CAAGAAGCGGTGCAGGGTATCAA	-
pDS132_acw_SceI	PR621	TAGGGATAACAGGGTAATCACATATACCTGCCGTTCAC	-
pDS132_cw	PR622	GTTATCCGCTCACAATTCCA	-
T7_SceI	PR623	ATTACCCTGTTATCCCTAGTAATACGACTCACTATAGGGC	-
T3_pDS132_cw	PR624	TGGAATTGTGAGCGGATAACAATTAACCCTCACTAAAGGG	-
fimK2_lr_Kn5R	PR625	TCCGGTTCGCATGCATTCCCCAGCGTGATTTTTCCGATAAG	SOE-PCR overlap
fimK2_rf_Kn5F	PR626	TCTCATGAGTATGCATTCCCGGCCCACAGCGGATTATG	SOE-PCR overlap
fimK2_lr_Kn5F	PR627	TCTCATGAGTATGCATTCCCCAGCGTGATTTTTCCGATAAG	SOE-PCR overlap
fimK2_rf_Kn5R	PR628	TCCGGTTCGCATGCATTCCCGGCCCACAGCGGATTATG	SOE-PCR overlap
fimK2_lf_XbaI	PR629	GC**TCTAGA**GTAATGGGATGGCGACAGG	*Xba*I
fimK2_rr_XbaI	PR630	GC**TCTAGA**GGCGTCGTCCAAACTCTAC	*Xba*I
pirF	PR633	ACTCAAGGTCATGATGGACG	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
pirR	PR634	GGTACGGTTTCATCAAATACCT	-
#CR	PR635	GGGAATGCATGCGAACCGGAATTGC	-
#CF	PR636	GGGAATGCATACTCATGAGATGCC	-
catF	PR637	GGTCACAGCTTGTCTGTAAG	-
catR	PR638	TCAACATAAAGGTGAATCCCAT	-
opfimcontrol	PR639	GGTAATCAGGATATGGCAGG	-
fimK_lf_NotI	PR640	GC**GCGGCCGC**GCCAACGTCTACGTTAACCT	*Not*I
fimK_rr_NotI	PR641	GC**GCGGCCGC**TGCCAAACATGAATTCGATAACAC	*Not*I
fimK_lr_GmF	PR642	**TCAGAGCGCTTTTGAAGCTAATTCG**GGATCGTCAGGGAGATACAC	SOE-PCR overlap
fimK_lr_GmR	PR643	**AGGAACTTCAAGATCCCCAATTCG**GGATCGTCAGGGAGATACAC	SOE-PCR overlap
fimK_rf_GmF	PR644	**TCAGAGCGCTTTTGAAGCTAATTCG**TGGAATGTCATCCGCATCTG	SOE-PCR overlap
fimK_rf_GmR	PR645	**AGGAACTTCAAGATCCCCAATTCG**TGGAATGTCATCCGCATCTG	SOE-PCR overlap
56metD_chkF	PR646	TATATCAATCAGTACCTGTCTGGC	-
56metD_chkR	PR647	GGCGAGGATATTATTCATTCCC	-
pDS132_cw_NotI	PR648	CAGGTATATGTG**GCGGCCGC**GTTATCCGCTCACAATTCCA	*Not*I
pDS132_acw_NotI	PR649	**GCGGCCGC**CACATATACCTGCCGTTCAC	*Not*I
fimKrecombchk	PR654	CGATCTCTATACCTTTAGCC	-
116_56metU_wlk2	PR655	TCACCGTAATGTTGATATCG	-
116_56metU_wlk3	PR656	TACGGGTCAGGGTTATCAGC	-
169asn33Dwalk2	PR657	GGTTATGGATAATACCTCTGG	-
9metU	PR658	CGTTGTACGGATGGGGTATC	-
9metD	PR659	GCGATACCCCAATTGTCTGT	-
13metU	PR660	CATGAACGATCCATCAGGC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
13metD	PR661	GATACCCCATCCGTACAACG	-
51-3metU	PR662	TGAAGATAACAAAGCGCCG	-
51-3metD	PR663	CATTCTTGTTGGCCAGGG	-
57metU	PR664	ATGTCCCGTCTTCGGTACC	-
57metD	PR665	GGCGTATAGCCCAGTTATTCTG	-
33asnUU	PR666	CAGCAGGTTAATTTTAGTGTTGACA	-
51-3metUU	PR667	TACAAAAATGTCGATTTGGTCG	-
56metDD	PR668	TGATCGTCAATGATTATCGC	-
fimG2R	PR669	CCGATATCAACATTACGGTG	-
116_56metU_wlk4	PR670	TACTGCCGCTCTGACTGCT	-
fimK2_UF_wlk	PR679	ATCTTTTATCGATGGCTGCG	-
opfim_lf_NotI	PR680	GC**GCGGCCGC**TTTGCAGTGAATGCTATCCA	*Not*I
opfim_lr_GmF	PR681	**TCAGAGCGCTTTTGAAGCTAATTCG**ACAGGAGTGACGCAGCAT	SOE-PCR overlap
opfim_lr_GmR	PR701	**AGGAACTTCAAGATCCCCAATTCG**ACAGGAGTGACGCAGCAT	SOE-PCR overlap
opfim_rf_GmF	PR702	**TCAGAGCGCTTTTGAAGCTAATTCG**GATGCTGGACGACTGCTTT	SOE-PCR overlap
opfim_rf_GmR	PR703	**AGGAACTTCAAGATCCCCAATTCG**GATGCTGGACGACTGCTTT	SOE-PCR overlap
opfim_rr_NotI	PR704	GC**GCGGCCGC**ATAAGGAATTACTGGAAGCGGG	*Not*I
FRT_fwd_tetR	PR705	**CGCCTTTGAGTGAGCTGATACCG**CATACGCTACTTGCATTACAG	SOE-PCR overlap
FRT_rev_tetF	PR706	**CGCTAACGGATTCACCACTCCA**GTTGTTCGGTAAATTGTCAC	SOE-PCR overlap
fimK2_UF_seq	PR707	GAGAATTGCCTGGCGGT	-
pDS132_acw	PR708	CACATATACCTGCCGTTCAC	-
tetMF	PR709	TACAAATATGCTCTTACG	-
tetMR	PR710	TTTCATCTTATTTAATAAGAAACC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
FRT_fwd_tetMF	PR711	**CGTAAGAGCATATTTGTA**CATACGCTACTTGCATTACAG	SOE-PCR overlap
FRT_rev_tetMR	PR712	**GGTTTCTTATTAAATAAGATGAAA**GTTGTTCGGTAAATTGTCAC	SOE-PCR overlap
tetMF_NotI	PR713	GC**GCGGCCGC**TACAAATATGCTCTTACG	*Not*I
KpGI3_UA_PiI_XbaI	PR801	CTAG**TCTAGA**CTTCGTAATCCATTTCTCGG	*Xba*I
KpGI3_DA_PiI_XbaI	PR802	CTAG**TCTAGA**GTAGATAGTTTCTCACCTCAC	*Xba*I
tetF	PR869	TGGAGTGGTGAATCCGTTAGCG	-
tetR	PR870	CGGTATCAGCTCACTCAAAGGCG	-
116met56contiggap2	PR923	ATCGCCAAAGAGCCTATCTC	-
116met56contiggap3	PR924	CGGTGGATACGTTTATTCCG	-
116met56contiggap4b	PR925	ATTAGCGATATCCTTGCGCC	-
116met56contiggap4a	PR926	GGTTCATTATCGGCATCCTG	-
fimK_GST_fwd_BamHI	PR927	CG**GGATCC**ACCGATTATATCCTCTCGCCC	*Bam*HI
fimK_GST_rev_EcoRI	PR928	CG**GAATTC**TCAACGTTTCGCCGGATCGC	*Eco*RI
fimK2_GST_fwd_BamHI	PR929	CG**GGATCC**ATGGTTGAGTCTTTTACGCAGC	*Bam*HI
fimK2_GST_rev_EcoRI	PR930	CG**GAATTC**GAATAACTCGTCCTTTATTCCC	*Eco*RI
opfim2_lf_NotI	PR931	GC**GCGGCCGC**CATCGTCCAGTTCACGATAGG	*Not*I
opfim2_lr_GmF	PR932	**TCAGAGCGCTTTTGAAGCTAATTCG**AACGGGATGTTTACAAAGCA	SOE-PCR overlap
opfim2_rf_GmR	PR933	**AGGTACTTCAAGATCCCCAATTCG**ATCTTCATTTAACCGCCTTCAG	SOE-PCR overlap
opfim2_rr_NotI	PR934	GC**GCGGCCGC**GCTGAAAGGTATGTTCTTGGTG	*Not*I
kpGI5_rf_GmR	PR935	**AGGTACTTCAAGATCCCCAATTCG**CGGGAAATAAATGTGAAAGC	SOE-PCR overlap
kpGI5_rr_NotI	PR936	G**CGCGGCCGC**TGATCGTCAATGATTATCGC	*Not*I
fim2op_fwd_NotI	PR937	GC**GCGGCCGC**GCAATTCTCTGACATTGATGCT	*Not*I
fim2op_rev_fimA_NotI	PR938	GC**GCGGCCGC**ACAACTTAGCTATCAAACTCGG	*Not*I

Primer name	Catalogue number[a]	Sequence[b]	Special features
fim2op_rev_hypo_NotI	PR939	GC**GCGGCCGC**CTCCTCTTCAAGCATGTTGC	*Not*I
Kp_fimS_internal	PR1015	CCACCACACCAGATTATATTACC	-
fimK_fwd_screen	PR1016	TACCGATGTTGATTCTGAGCC	-
fimK_rev_screen	PR1017	TACCCATTCCTGTTCAAACG	-
pigx_fwd_GE_BamHI	PR1018	CG**GGATCC**GATACATTGATTCGCGCTTTC	*Bam*HI
pigx_fwd_E_BamHI	PR1019	CG**GGATCC**TGGTGTGTGTATGACCGCCAG	*Bam*HI
pigx_rev_EcoRI	PR1020	CG**GAATTC**TTCTACGCCAATTCGCAC	*Eco*RI
78578int_rev_NotI	PR1021	ATAAGAAT**GCGGCCGC**ATAATTTCGAAGGCTTGC	*Not*I
78578int_fwd_NdeI	PR1022	GGAATTC**CATATG**ACCGTGCGTAAAAATCC	*Nde*I
mrk_fwd_NotI	PR1057	GC**GCGGCCGC**CGCAAAGGCAGATTCAGACT	*Not*I
mrk_rev_NotI	PR1058	GC**GCGGCCGC**AGTATTGAACAACTGGTCGC	*Not*I
mrk_lf_NotI	PR1059	GC**GCGGCCGC**AGATTCAGACTTATCGCGATGG	*Not*I
mrk_lr_GmF	PR1060	**TCAGAGCGCTTTTGAAGCTAATTCG**CCAGTTTGCTTACGTCATCC	SOE-PCR overlap
mrk_rf_GmR	PR1061	**AGGTACTTCAAGATCCCCAATTCG**TGTTCACCTATTACGTTGGCT	SOE-PCR overlap
mrk_rr_NotI	PR1062	GC**GCGGCCGC**CTGTACGCTACGTCTTTCAG	*Not*I
fimK2_rev_PstI	PR1064	TGCA**CTGCAG**GAATAACTCGTCCTTTATTCCC	*Pst*I
fimK_rev_HindIII	PR1066	CCC**AAGCTT**TCAACGTTTCGCCGGATCGC	*Hind*III
78578int_rev_SalI	PR1068	ACGC**GTCGAC**GCATAATTTCGAAGGCTTGC	*Sal*I
yahAEAL_fwd_SacI	PR1069	ACGC**GAGCTC**CGCGATATTTTCTTTCAGTAC	*Sac*I
yahAEAL_rev_PstI	PR1070	TGCA**CTGCAG**CAACCACCTGCTTTCATTAC	*Pst*I
Bact-8F	PR1080	AGAGTTTGATCCTGGCTCAG	-
Bact-1391R	PR1081	GACGGGCGGTGTGTRCA	-
opfim2KpGI5mutantcheck	PR1103	TATCGACATGAACACCGACC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
opfim2mutantchk_rev	PR1104	GAACCGACTATATTCATTCGAG	-
kpGI5mutantchk_rev	PR1105	GGCGAGGATATTATTCATTCCC	-
MBL_fwd	PR1106	TGGGCATAACGATCATACCT	-
MBL_rev	PR1107	CTGTGACTACCACCAGTCC	-
hypo1_fwd	PR1114	GTAATGTAATTGATGGGCGA	-
hypo1_rev	PR1115	GTATCATACCGTGCTTGTACC	-
mob78578_ua_rev	PR1116	GTGGCTCTGATAAATTTCCGA	-
met56_KR116mobfwd	PR1117	CAGGGAGGAGTTCATATTGG	-
met56_KR116mobrev	PR1118	CATGGCAAGTATTCTCAACGA	-
fimK2_f_BamHI	PR1134	CG**GGATCC**GTTGAGTCTTTTACGCAG	*Bam*HI
fimK2_r_KpnI	PR1135	GC**GGTACC**GAATAACTCGTCCTT	*Kpn*I
fimK_f_BamHI	PR1136	CG**GGATCC**ACCGATTATATCCTCTCG	*Bam*HI
fimK_r_KpnI	PR1137	GC**GGTACC**TCAACGTTTCGCCGG	*Kpn*I
78578int_f_BamHI	PR1138	CG**GGATCC**ATGACCGTGCGTAAAAATCC	*Bam*HI
78578int_rev_KpnI	PR1139	GC**GGTACC**GCATAATTTCGAAGGCTTGC	*Kpn*I
fimB_F	PR1142	AGGCTCGGAAAACCAAAAAT	-
fimB_R	PR1143	CGGAAATTTTTCCCACCAAT	-
FimH_fwd	PR1144	CACGCAAGGCACCATTC	-
FimH_rev	PR1145	GCTCAGAATCAACATCGGTAAC	-
UpfimB_F2	PR1146	TCACCACGCTGTTTCAAAAT	-
DwfimH_R2	PR1147	AGCAGGGCGAGAGGATATAA	-
Upfim	PR1148	AGCCCAAGCCTTCATCATCA	-
Dwfim	PR1149	CACCGGCGAACACCTCAGCATTAT	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
mrkDF	PR1150	AATAGCAGCCACGCGATAGT	-
mrkDR	PR1151	TGTTTATCAGCGATGCGAAC	-
EBGNHe	PR1179	CCCGCTAGCGAAAAGATGTTTCGTGAAGC	-
EBGh3	PR1180	GGGAAGCTTATTATCGTGAGGATGCGTCA	-
fimK2_f_NcoI	PR1202	CG**CCATGG**TTGAGTCTTTTACGCAG	*Nco*I
fimK2_r_BamHI	PR1203	GC**GGATCC**GAATAACTCGTCCTT	*Bam*HI
fimK2_f_D14_NcoI	PR1204	CG**CCATGG**ATAGCCGAAAAGAATGCC	*Nco*I
fimK2_r_E266_BamHI	PR1205	GC**GGATCC**TTATTATTCAACCGGCTGCGAAAAC	*Bam*HI
fimK_f_NcoI	PR1206	CG**CCATGG**TAACCGATTATATCCTCTCG	*Nco*I
fimK_r_BamHI	PR1207	GC**GGATCC**TCAACGTTTCGCCGG	*Bam*HI
fimK_f_P204_NcoI	PR1208	CG**CCATGG**TACCGCGCTCACCCGCAAACG	*Nco*I
fimK_r_R449_BamHI	PR1209	GC**GGATCC**TTATTACCGCATCGGTGGTGAAAAC	*Bam*HI
pTrc99A.seq.F	PR1216	TGCAGGTCGTAAATCACTGC	-
pTrc99A.seq.R	PR1217	CTGGCAGTTCCCTACTCTCG	-
pTRC-NotILink	PR1218	GCATGCCATGGCGGCCGCGAATTCCGG	*Nco*I/*Not*I/*Eco*RI
pTRC-NotIComp	PR1219	CCGGAATTCGCGGCCGCCATGGCATGC	*Nco*I/*Not*I/*Eco*RI
Kp_fim_fwd_NotI	PR1220	GGC**GCGGCCGC**GAAAATCAAAACACTGGCA	*Not*I
Kp_opfim_rev_SbfI	PR1221	CCA**CCTGCAGG**CGATAACACCCGCGAATACGAC	*Sbf*I
fim2op_rev_fimA2_NotI	PR1222	GGC**GCGGCCGC**TAATGCAAAAATTTTTGTAATTGC	*Not*I
fim2op_rev_fimQ2_NotI	PR1223	GGC**GCGGCCGC**TAAGATAGTCGTTATTAGTGAC	*Not*I
fim2op_fwd_SbfI	PR1224	CCA**CCTGCAGG**CAATTCTCTGACATTGATGCT	*Sbf*I
Kn1_NsiI	PR1231	CCA**ATGCAT**GTGTAGGCTGGAGCTGCTTC	*Nsi*I
Kn2_EcoRV	PR1232	CCG**GATATC**ATGGGAATTAGCCATGGTCC	*Eco*RV

Primer name	Catalogue number[a]	Sequence[b]	Special features
aadA_miniTn5R_StrABF	PR1233	**CGAAGATTCCGAACACCATG**AGACATTATTTGCCGACTAC	SOE-PCR overlap
strAB_KpF	PR1234	CATGGTGTTCGGAATCTTCG	-
aadA_miniTn5R_KpnI	PR1235	CGG**GGTACC**AGACATTATTTGCCGACTAC	*Kpn*I
aadA_miniTn5F_XhoI	PR1236	CGG**CTCGAG**GTAACGGCGCAGTGGCGGTT	*Xho*I
strAB_KpF_XhoI	PR1237	GCGG**CTCGAG**CATGGTGTTCGGAATCTTCG	*Xho*I
strAB_KpR_KpnI	PR1238	CGG**GGTACC**CGGTAAGATTGATGTGTTCC	*Kpn*I
TcU	PR1253	GCGCTGTGGGGCATTTTACTTTA	-
TcD	PR1254	TTAGCGGGTCTTGGTCTTTTACAC	-
Kn1	PR1255	GTGTAGGCTGGAGCTGCTTC	-
Kn2	PR1256	ATGGGAATTAGCCATGGTCC	-
opfim2_lf_red	PR1257	CATCGTCCAGTTCACGATAGG	-
opfim2_lr_red_Kn1	PR1258	**GAAGCAGCTCCAGCCTACAC**AACGGGATGTTTACAAAGCA	SOE-PCR overlap
opfim2_rf_red_Kn2	PR1259	**GGACCATGGCTAATTCCCAT**GAAAGCAACAGACGCTTTGG	SOE-PCR overlap
opfim2_rr_red	PR1260	GAAAGTTGTATGTTGGCATTGG	-
KpneuSp-F	PR1261	ATTTGAAGAGGTTGCAAACGAT	-
KpneuSp-R	PR1262	TTCACTCTGAAGTTTTCTTGTGTTC	-
KoyxSp-F	PR1263	GATACGGAGTATGCCTTTACGGTG	-
KoyxSp-R	PR1264	TAGCCTTTATCAAGCGGATACTGG	-
Kp_fim_fwd2_NotI	PR1268	GGC**GCGGCCGC**GAAAATCAAAACACTGGCAATGATTGTTG	*Not*I
strAB_plac_fwd_KpnI	PR1269	CGG**GGTACC**CACATGTTCTTTCCTGCG	*Kpn*I
oriR6kT_F	PR1270	GCTGCATAACCCTGCTTCGGGG	-
oriR6kT_R	PR1271	GCCATGAGAGCTTAGTACGTTAGCC	-
glmS_Kp_fwd	PR1307	TGCAGCTGTTGGCTTATCAC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
glmS_Kp_rev	PR1308	GGTACGCATAACGTTCATAATGTC	-
Kp_FimH_fwd_NcoI	PR1313	CATG**CCATGG**TGAAAAAAATAATCCCCCTGTTCACC	*Nco*I
Kp_FimH_Flag225_lr	PR1314	**CTTGTCGTCGTCGTCCTTGTAGTC**CTGAACGCCTATCCCCTGC	SOE-PCR overlap
Kp_FimH_Flag225_rf	PR1315	**GACTACAAGGACGACGACGACAAG**CTGACGCGCAACGGCAGC	SOE-PCR overlap
Kp_FimH_Flag258_lr	PR1316	**CTTGTCGTCGTCGTCCTTGTAGTC**GGCATAGGTGGCCGTCAG	SOE-PCR overlap
Kp_FimH_Flag258_rf	PR1317	**GACTACAAGGACGACGACGACAAG**CGGACCACAGGCCAGGTT	SOE-PCR overlap
Kp_FimH_rev_KpnI	PR1318	CGG**GGTACC**TCGGTCATCATTGATAGACAAAGG	*Kpn*I
Kp_FimH2_fwd_NcoI	PR1319	CATG**CCATGG**ATAAACTTATCCCTCAGTTTACCG	*Nco*I
Kp_FimH2_Flag225_lr	PR1320	**CTTGTCGTCGTCGTCCTTGTAGTC**TTGCACGCCAACGCCCTG	SOE-PCR overlap
Kp_FimH2_Flag225_rf	PR1321	**GACTACAAGGACGACGACGACAAG**ATGACACGTAATGGAGCGATTG	SOE-PCR overlap
Kp_FimH2_Flag258_lr	PR1322	**CTTGTCGTCGTCGTCCTTGTAGTC**GGCATAATTCGCAGTTAAACCG	SOE-PCR overlap
Kp_FimH2_Flag258_rf	PR1323	**GACTACAAGGACGACGACGACAAG**CGTACCATAGGTCAGGTTACT	SOE-PCR overlap
Kp_FimH2_rev_KpnI	PR1324	CGG**GGTACC**TAAAGTAGAGTTTGGACGACG	*Kpn*I
PigX_GGDEF_fwd_SacI	PR1325	CGC**GAGCTC**GATACATTGATTCGCGCTTTC	*Sac*I
PigX_GGDEF_rev_KpnI	PR1326	CGG**GGTACC**TCAGCCATTCCCTCCTTGTAATAC	*Kpn*I
PigX_EAL_fwd_SacI	PR1327	CGC**GAGCTC**TGGTGTGTGTATGACCGCCAG	*Sac*I
PigX_EAL_rev_KpnI	PR1328	CGG**GGTACC**CTACGCCAATTCGCAC	*Kpn*I
FimX_EAL_fwd_BamHI	PR1329	CGC**GGATCC**AACCCGGCGGAAGAGCTTG	*Bam*HI
FimX_EAL_rev_SalI	PR1330	AACGC**GTCGAC**CCCGTTCTTCATTCGTCTCC	*Sal*I
yahA_EAL_fwd_SacI	PR1331	CGC**GAGCTC**CGCGATATTTTCTTTCAGTAC	*Sac*I
yahA_EAL_rev_KpnI	PR1332	CGG**GGTACC**TCAACCACCTGCTTTCATTAC	*Kpn*I
KPN00268_EAL_fwd_BamHI	PR1333	CGC**GGATCC**TTTTGCTTCTCATTTTTGACCCTG	*Bam*HI
KPN00268_EAL_rev_SalI	PR1334	AACGC**GTCGAC**CTTCTGTTAAGAATGCGTC	*Sal*I

Primer name	Catalogue number[a]	Sequence[b]	Special features
KPN03660_EAL_fwd_SacI	PR1335	CGC**GAGCTC**CCGGAGAAAGGTCGCGGC	*Sac*I
KPN03660_EAL_rev_KpnI	PR1336	CGG**GGTACC**GAAATCACGCTAAACGGCAG	*Kpn*I
RocR_EAL_fwd_SacI	PR1337	CGC**GAGCTC**CGCCAGGACCTTCCGCG	*Sac*I
RocR_EAL_rev_KpnI	PR1338	CGG**GGTACC**CGGCAGTTGTATGGAAAGGA	*Kpn*I
Kp_fimH_lf	PR1339	AATAATCCCCCTGTTCACCA	-
Kp_fimH_lr_GmF	PR1340	**TCAGAGCGCTTTTGAAGCTAATTCG**CACGGTGCCTGAAAAACTC	SOE-PCR overlap
Kp_fimH_rf_GmR	PR1341	**AGGAACTTCAAGATCCCCAATTCG**CGCGATCTTCACCAATACC	SOE-PCR overlap
Kp_fimH_rr	PR1342	AGCAGGGCGAGAGGATATAAT	-
Kn3	PR1343	CATATGAATATCCTCCTTAGTTC	-
Kp_fimH2_lf	PR1344	CTTATCCCTCAGTTTACCGTC	-
Kp_fimH2_lr_Kn1	PR1345	**GAAGCAGCTCCAGCCTACAC**CCCACCATACATCGAACCA	SOE-PCR overlap
Kp_fimH2_rf_Kn3	PR1346	**GAACTAAGGAGGATATTCATATG**CTCTCCGGTACGACTACAG	SOE-PCR overlap
Kp_fimH2_rr	PR1347	AATGTTACTCCGATGAGGGAC	-
fimI2_int_fwd	PR1348	GCTGATTCCCTTGAATGGTC	-
fimC2_int_rev	PR1349	GGTGGTGTAATGACGAAACGA	-
mrk_LRR_screen	PR1363	TATCCTCTGCCTATTGTCTG	-
KPN03660_GGDEF_fwd_SacI	PR1364	CGC**GAGCTC**AGCCGGATGGATACGCTTATTC	*Sac*I
KPN03660_GGDEF_rev_KpnI	PR1365	CGG**GGTACC**TCAGCTGTTGCTGCCCTGC	*Kpn*I
KPN03660_fwd_SacI	PR1366	CGC**GAGCTC**CGATTAACGACGAAGTTCTCAGC	*Sac*I
Kp_FimA2_fwd	PR1367	AATGCAAAAATTTTTGTAATTGCAGC	-
Kp_FimA2_rev	PR1368	CATTCCTTGCATGTTATTACCC	-
FimF2_fwd	PR1572	ATACACCACTTAACCTTCTGAG	-
FimF2_rev	PR1573	TACTGGAATTCAAGGGTAAAGG	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
FimC2_fwd	PR1574	GGGTTATTTATCCAGCTGGAC	-
FimC2_rev	PR1575	AATACACCATTCATCGTTGGAG	-
FimA2_fwd	PR1576	AATGCAAAAATTTTTGTAATTGCAGC	-
FimA2_rev	PR1577	CATTCCTTGCATGTTATTACCC	-
fimD2_int_rev	PR1578	TAACATTCACCTGTACTTTGCC	-
fimD2_int_fwd	PR1579	GAAGAGCCTAAGTTCTTTCAG	-
fimF2_rev_int	PR1580	GCTGATTGATATGGTACTGTC	-
fimC2_fwd_int	PR1581	CTTCAGTTAAATTGCCTCCAG	-
fimA_RTPCR_fwd	PR1601	ACCGTTCAGTTAGGCCAGGT	-
fimA_RTPCR_rev	PR1602	CCGGTATTGTCGAGGATCTG	-
fimH_RTPCR_fwd	PR1603	CGGAAACGATCACCGACTAC	-
fimH_RTPCR_rev	PR1604	GATTAACGATCCTGCGGTGA	-
fimK_RTPCR_fwd	PR1605	GTGGAATGTCATCCGCATCT	-
fimK_RTPCR_rev	PR1606	AGAAAGGTCTGGGGGTGAAG	-
fimA2_RTPCR_fwd	PR1607	CGTTAATGGCGGTACAGTGC	-
fimA2_RTPCR_rev	PR1608	CTGAGGTACCGGAGAAAGCA	-
fimH2_RTPCR_fwd	PR1609	CAGCGAGTAGTGTCGCGATT	-
fimH2_RTPCR_rev	PR1610	TGTGCGCAATGAACGGTAAC	-
fimK2_RTPCR_fwd	PR1611	GCCATTCAGAACGGTCAGGT	-
fimK2_RTPCR_rev	PR1612	ATTCGCCCTGATAGCGATTG	-
gyrB_RTPCR_fwd	PR1613	CCCTTCCACCAGGTACAGT	-
gyrB_RTPCR_rev	PR1614	GGAGCAGCAGATGAACGAAC	-
IS_56met_fwd	PR1620	GATCAATCGCCAAACCAACC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
IS_56met_rev	PR1621	CTACACTTTTGCCTGACCATG	-
mrkA_RTPCR_fwd	PR1622	CGATGCGAACGTTTACCTGT	-
mrkA_RTPCR_rev	PR1623	CAGGTAGCCCTGTTGTTTGC	-
mrkF_RTPCR_fwd	PR1624	CAATACCTTGCCGGATGGA	-
mrkF_RTPCR_rev	PR1625	GTTGCCCAGCTGGTAATCG	-
fimQ_RTPCR_IG_fwd	PR1626	TTATACCTTAAATAAGCCAGTTGA	-
fimA2_RTPCR_IG_rev	PR1627	GTGGTATTAGCCAAAGCCGC	-
fimH2_RTPCR_IG_fwd	PR1628	TCTTTGGGGACGGTAGGC	-
fimK2_RTPCR_IG_rev	PR1629	TGGCATTCTTTTCGGCTATC	-
fimH_RTPCR_IG_fwd	PR1630	CCAACGCGATCTTCACCAAT	-
fimK_RTPCR_IG_rev	PR1631	TACACCACGATCCGCTTCAC	-
fimH-H2-transfer_fwd_SphI	PR1632	GACAT**GCATGC**ACACAGGAAACAGACCATGG	*Sph*I
fimH-H2-transfer_rev_SphI	PR1633	GACAT**GCATGC**GGTCGACTCTAGAGGATCC	*Sph*I
fimQ_RTPCR_fwd	PR1634	GAGCTCAAAGAGTTATGATGGAG	-
fimQ_RTPCR_rev	PR1635	TTGCCAATGCCAACATACAAC	-
QE80L_fwd	PR1643	CCCGAAAAGTGCCACCTG	-
QE80L_rev	PR1644	CAAGACTCCAGTAATGACC	-
rpoD_RTPCR_fwd	PR1665	AAGACGGCATCAATCAGGTG	-
rpoD_RTPCR_rev	PR1666	TTCGTCTTCGTCTTCGTCGT	-
fimH-H2-transfer_fwd_ApaLI	PR1667	GGGG**GTGCAC**ACACAGGAAACAGACCATGG	*Apa*LI
fimH-H2-transfer_rev_HindIII	PR1668	GGGG**AAGCTT**GGTCGACTCTAGAGGATCC	*Hind*III
rpoD_qPCR_fwd	PR1701	CACAGCTGAAGCTTCTTGTCC	-
rpoD_qPCR_rev	PR1702	ATCCGGTGCTTCTTCCATC	-

Primer name	Catalogue number[a]	Sequence[b]	Special features
fimA_qPCR_fwd	PR1703	AGTTAGGCCAGGTGCGTTC	-
fimA_qPCR_rev	PR1704	ACGGTGGTATTGCTGCTGTC	-
fimA2_qPCR_fwd	PR1705	CAAAGCGTCTGTTGCTTTCTC	-
fimA2_qPCR_rev	PR1706	GTCGCTGCACTGAAGGATG	-
mrkA_qPCR_fwd	PR1707	TGATGGCACTAAACAGGATGAC	-
mrkA_qPCR_rev	PR1708	CGTTGTCAGTAGACAGCACCAG	-
pMQ300_MCS_fwd	PR1709	TGAGCGGATAACAATTTCACAC	-
pMQ300_MCS_rev	PR1710	CGTTTCACTTCTGAGTTCGG	-

[a] Catalogue number corresponds to a physical primer catalogue in Dr. Kumar Rajakumar's laboratory. [b] Primer sequence (5' to 3' orientation) with special features highlighted in bold and underlined text.

Appendix 8. BLASTp alignment results for Table 4-1

Query = Fim2K (285 aa)

```
>gb|ABV14791.1|  hypothetical protein CKO_03715 (FimK) [Citrobacter koseri ATCC BAA-895]
Length=468

 Score =  350 bits (899),  Expect = 5e-117, Method: Compositional matrix adjust.
 Identities = 165/276 (60%), Positives = 209/276 (76%), Gaps = 0/276 (0%)

Query  2    MVESFTQQVNCFTDSRKECQAVQALATPTPFESELVHAIQNGQVYPVFQPIVDIHLHIKG  61
            MV S    +  F  S+ +    Q +A  +PFE E +HAI + QV+PVFQPI D HL ++G
Sbjct  184  MVASHPHLITRFPGSQSKTHNNQLIAALSPFEREFIHAIHSQQVFPVFQPITDGHLRLQG  243

Query  62   IEVLSRWRKDGVVLLPTEFLPNIQSEAIWFSLTAFVLQEAVQGINRYQGEFYFTVNIPTC  121
            +E+LSRWR+   VLLP EFLP I +E  W  LTAFVLQ A+Q IN++QG+F+F++NIP C
Sbjct  244  VEILSRWRRGDNVLLPGEFLPQIHAEYAWLLLTAFVLQIAIQNINQHQGKFWFSINIPPC  303

Query  122  IAHHHHLICLMETAWLQLHNPLWADCLVLEFAETVDLTQQGNTIANMRKIQERGFRIFLD  181
            IA+H +L+ +META  QL  P W+  LVLEFAETV+L QQG T  NM KIQ +GFRIFLD
Sbjct  304  IANHENLLRMMETARQQLQQPQWSGRLVLEFAETVNLHQQGRTAENMDKIQRQGFRIFLD  363

Query  182  DCFSQNSVIFPIRLARFCGYKLDKSIINDFQRDPHAMALMKSLIYYCQLTQSDCIAEGVD  241
            DCFS +SV+FP+R  RF GYKLD SI+NDFQRDPHA+AL+KSL+YYCQLTQS CIAEGVD
Sbjct  364  DCFSHSSVMFPVRTIRFSGYKLDMSIVNDFQRDPHALALIKSLLYYCQLTQSRCIAEGVD  423

Query  242  SLEKFNKLKGMGLVFFQGYLFSQPVELEHMMCMIKK  277
            SLEKFN+LK +G+  FQGYLFS P+  + +  +I++
Sbjct  424  SLEKFNQLKALGVDRFQGYLFSPPITHDRLPEIIQQ  459
```

Query = Fim2H (308 aa)

```
>gb|ACL13802.1|  FimH [Klebsiella pneumoniae]
Length=302

 Score =  374 bits (959),  Expect = 3e-128, Method: Compositional matrix adjust.
 Identities = 190/308 (62%), Positives = 232/308 (75%), Gaps = 7/308 (2%)

Query  1    MNKLIPQFTVLLLAGWSINVWAFSCMDASGQTLHSMAGPGSVNVYVNLQPTIAVGQNLVV  60
            M K+IP FT LLL GWS+N W+F+C  A+G T+    G GS NVYVNL PT+ VGQNLVV
Sbjct  2    MKKIIPLFTTLLLLGWSMNAWSFACKTATGATIP--IGGGSANVYVNLTPTVNVGQNLVV  59

Query  61   DISRSVVCRNDAPTHRNDNVMMLFGSMYGGALSNFTGTLRYYGSSYTFPLRSATSPHNFT  120
            D+S  + C ND P    D V +  GS YGG LS+F+GT++Y G+SY FP  + T+   +
Sbjct  60   DLSTQIFCHNDYPETITDYVTLQRGSAYGGVLSSFSGTVKYNGTSYPFPTTTETARVIYD  119

Query  121  SGSYVPWNTQLYLTPISAASSVAIKGGTRFAQLVMYQVGSDITSGGNIQTVKFTWNLYSM  180
            S +  PW   LYLTP+S A  VAI  G+  A L+++Q     T+  N  + +F WN+Y+
Sbjct  120  SRTDKPWPAVLYLTPVSTAGGVAITAGSLIAVLILHQ-----TNNYNSDSFQFIWNIYAN  174

Query  181  RDVVVPTGGCDVSARDVTMTLPNYPGSMAVPVTVHCAQNQNLSYYLSGTTTDSANSIFAN  240
             DVVVPTGGCDVSARDVT+TLP+YPGSMAVP+TVHCAQ+Q L YYLSGTT DSAN+IF N
Sbjct  175  NDVVVPTGGCDVSARDVTVTLPDYPGSMAVPLTVHCAQSQQLGYYLSGTTADSANAIFTN  234

Query  241  TASASPAQGVGVQMTRNGAIVSANNTISLGTVGPSPVNLGLTANYARTIGQVTAGNVQSL  300
            TASASPAQG+GVQ+TRNG+ V AN+T+SLGTVG SPVNLGLTA YART GQVTAGNVQS+
Sbjct  235  TASASPAQGIGVQLTRNGSAVPANSTVSLGTVGTSPVNLGLTATYARTTGQVTAGNVQSI  294

Query  301  IGVTFIYQ  308
            IG+TF+YQ
Sbjct  295  IGITFVYQ  302
```

Query = Fim2G (167 aa)

```
>gb|ABV14789.1| hypothetical protein CKO_03713 (FimG) [Citrobacter koseri ATCC BAA-895]
Length=167

 Score =  251 bits (642),  Expect = 3e-84, Method: Compositional matrix adjust.
 Identities = 120/167 (72%), Positives = 140/167 (84%), Gaps = 0/167 (0%)

Query  1    MKYRFAGALLVGLLSMSVSTSQGADINITVNGRVVARPCNIATTNASIDLGDLFTFSLKS  60
            MK  + G LL  +L+ + S  Q ADI ITVNG+VVA+PC ++TTNA++DLGDL+TFSL +
Sbjct  1    MKRNYVGWLLASVLTAASSGLQAADITITVNGKVVAKPCTVSTTNAAVDLGDLYTFSLVT  60

Query  61   AGSTSAWHDTALSLTNCPIGTSRVVATFSGMVDNTGYYKNQGTAGNIQLELQDDSGRRLN  120
            AGS SAWH TAL LTNCPIGTSRV+ATFSG  D+TGYYKNQGTAGNIQLELQDD+G  LN
Sbjct  61   AGSASAWHSTALELTNCPIGTSRVIATFSGTADSTGYYKNQGTAGNIQLELQDDAGTTLN  120
```

```
Query  121  NGAAKLVQVDDSSQSASFPLKVRVLSVKGGATQGTIQAVINVTYSYA  167
            NGA K VQVDDS+QS  FPL+VR L+V GGATQGTIQAVI+VTY+YA
Sbjct  121  NGATKSVQVDDSTQSTRFPLQVRALTVNGGATQGTIQAVISVTYTYA  167
```

Query = Fim2F (175 aa)

```
>gb|ABV14788.1|  hypothetical protein CKO_03712 (Fim2F) [Citrobacter koseri ATCC BAA-895]
Length=177

 Score =  282 bits (721),  Expect = 6e-96, Method: Compositional matrix adjust.
 Identities = 139/177 (79%), Positives = 153/177 (86%), Gaps = 0/177 (0%)

Query  1    MIMDTPLNLLRAALILICANAFAADSTISISGYVRDNACAVAIESKEFTVDLMNNAVKQF  60
            M+M  PL  L A L L C N FAADSTI+ISGYVRDNACAVA ESK+FTVDLMNNA KQF
Sbjct  1    MMMKKPLYFLGAVLTLACVNVFAADSTITISGYVRDNACAVAGESKDFTVDLMNNAAKQF  60

Query  61   SAVGSVTPALPFRIVLSPCGSSATAVKVGFTGTPDSNNDKLLKIDSDAFAASGMAVQILD  120
            +AVG+ TP +PFRIVLSPCG+S TAVKVGFTGT DS+N  LLKIDS A AA+GM VQIL+
Sbjct  61   NAVGATTPLVPFRIVLSPCGNSVTAVKVGFTGTEDSSNTCLLKIDSGASAAAGMGVQILN  120

Query  121  NQQTMLPINAAASSIAWTNLTPGQTNILGFYARLMATRVPVTAGHVYATATFTLEFQ  177
            NQQTMLP+NAA+SSIAWT LTPGQTNIL FYARLMAT++PVTAGHV ATATFTLEFQ
Sbjct  121  NQQTMLPLNAASSSIAWTTLTPGQTNILNFYARLMATQIPVTAGHVNATATFTLEFQ  177
```

Query = Fim2D (852 aa)

```
>gb|ABV14787.1|  hypothetical protein CKO_03711 (FimD) [Citrobacter koseri ATCC BAA-895]
Length=874

Score = 1431 bits (3704),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 689/838 (82%), Positives = 762/838 (91%), Gaps = 0/838 (0%)

Query  15   GHAQTEIYFNPRFLADDPASVADLSDFENGQEVPPGTYRVDIYLNDGFITTRDVTFNLGE  74
            G A+ ++YFNPRFLADDP++VADLS FENGQEVPPG YRVDIYLNDGFITTRDVTFN G
Sbjct  37   GFARADLYFNPRFLADDPSAVADLSGFENGQEVPPGAYRVDIYLNDGFITTRDVTFNAGA  96

Query  75   KGHGLDPCLARSQLAGMGVNIEGIKGIDRPTADTCVPLTKLINGATTHFDVGLQRLYLTV  134
             GH L+PCL RSQLAGMGV+   I G+D   +D CVPLT++  ATT FDV  QRLYL+V
Sbjct  97   NGHRLEPCLTRSQLAGMGVSTSAITGMDTLASDACVPLTEMVKDATTRFDVSQQRLYLSV  156

Query  135  PQAFMGNRARGYIPPEQWDAGINAGLLNYNFTGNNVHNDTGGGSNYAWLNLQSGLNLGAW  194
            PQAFMGNRARGYIPPE WD GINAGLLNYNFTGNNVHND    SNYA+LNLQSGLNLGAW
Sbjct  157  PQAFMGNRARGYIPPELWDDGINAGLLNYNFTGNNVHNDASDSSNYAYLNLQSGLNLGAW  216

Query  195  RLRDNTTWNYTSDDGTSRNKNKWQHVNTWLERDITPLRSRLTLGDSYTNGDIFDGFNFRG  254
            RLRDNT W+Y+S    S N+NKWQHVN+WLERDITPLR+RLTLGDSYTNGDIFDG NFRG
Sbjct  217  RLRDNTIWSYSSGGSASSNENKWQHVNSWLERDITPLRARLTLGDSYTNGDIFDGINFRG  276

Query  255  TQLASDDNMLPDSQRGFAPVIHGIARSTAQVSIKQNGYEIYQSTVPPGPFAINDLYAASS  314
             QLASDDNMLPDSQ+GFAPVIHGIAR TAQVSIKQNGYEIYQSTVPPGPF INDLYAA +
Sbjct  277  AQLASDDNMLPDSQKGFAPVIHGIARGTAQVSIKQNGYEIYQSTVPPGPFTINDLYAAGN  336

Query  315  GGDLQVTIKENDGTRQVFSVPWSTVPVLQREGYTRYALTAGEYRSGNVQQEEPKFFQGTV  374
            GGDLQVTIKE DGT QVFSVPWSTVP+LQREG+TRYA+TAGEYRSGN QQE+PKFFQ T+
Sbjct  337  GGDLQVTIKEADGTSQVFSVPWSTVPMLQREGHTRYAVTAGEYRSGNDQQEKPKFFQSTL  396

Query  375  LHGLPTGWTIYGGAQLAERYRAVNLGMGKNMGDFGALSLDITQANATLSDYSDHQGQSIR  434
            LHGLP GWT+YGG QLA+ YR+ NLG+GKNMG+FGA+SLD TQANATL D S HQGQSIR
Sbjct  397  LHGLPAGWTLYGGTQLADNYRSFNLGVGKNMGEFGAVSLDATQANATLPDDSSHQGQSIR  456

Query  435  FLYNKALSDVGTNIQLVGYRYSTKGYYNFADTTYRRMSGYDVQTQDGVIQVKPKFTDYYN  494
            FLYNK+L + GTN+QLVGYRYST+GYY+FADTTYRRMSGYDV+TQDGVIQVKPKFTDYYN
Sbjct  457  FLYNKSLYETGTNVQLVGYRYSTQGYYSFADTTYRRMSGYDVETQDGVIQVKPKFTDYYN  516

Query  495  LTWNKRGKVQVNVTQQLGRTATLYLSGSNQTYWTTGKADQQLQAGLSAAVEDINWTLSYS  554
            L +NKRGKVQ++VTQQLGRTATLY++GS QTYW+T KAD+QLQ GL+AAV+DINWTLSYS
Sbjct  517  LAYNKRGKVQMSVTQQLGRTATLYVNGSRQTYWSTNKADEQLQLGLNAAVDDINWTLSYS  576

Query  555  LTKNAWQQGRDQMLAVNVNIPFSHWLRSDSQSLWRHTSASYSMSHDLNGRMTNLAGIYGT  614
            LTKNAWQQGRDQMLA NVNIPFSHWLRSDS+S+WRH SASYSMSHDL+GRMTNLAG+YGT
Sbjct  577  LTKNAWQQGRDQMLAFNVNIPFSHWLRSDSKSVWRHASASYSMSHDLDGRMTNLAGLYGT  636

Query  615  LLEDNNLSYSVQTGYANGGETDTSSTGYAALNYRGGYGSANLGYSRNDGFRQLYYGVSGG  674
            LLEDNNLSYS+QTGYA+GGE +   TGYAA+NYRGGYG+AN+GYS +DGF+QLYYG+SGG
Sbjct  637  LLEDNNLSYSMQTGYASGGEGNNGGTGYAAMNYRGGYGNANVGYSHSDGFKQLYYGMSGG  696

Query  675  VLAHGDGVTLSQPLNDTVVLVKAPGADGVKVENQTGVTTDWRGYAVVPYATEYRENRVAL  734
            VLAH +G+TLSQPLNDTVVL+KAPGA  VKVENQTGV TDWRGYAV+PYATEYRENRVAL
Sbjct  697  VLAHANGITLSQPLNDTVVLIKAPGAGDVKVENQTGVRTDWRGYAVMPYATEYRENRVAL  756

Query  735  DTNTLADNVDLDDAVVSVVPTHGAIVRADFKAHVGVKLLITLTRKGKPVPFGAMVTSGSS  794
            DTNTLADNVDLDDAVVSVVPTHGAIVRA+FKAHVGVK+L+TLT  GKPVPFGAMVTSG +
Sbjct  757  DTNTLADNVDLDDAVVSVVPTHGAIVRAEFKAHVGVKILMTLTHNGKPVPFGAMVTSGDN  816
```

```
Query  795  QSGSIVADNGQAYMSGMPLAGSLQVKWGEGVGASCVARYRLPEESQRQSLSQLSAECR  852
            QSGSIVADNGQ Y+SGMPLAGS+QVKWGEG GA+C+ARYRLPEESQ+  LSQL AECR
Sbjct  817  QSGSIVADNGQVYLSGMPLAGSVQVKWGEGPGANCMARYRLPEESQKHVLSQLFAECR  874
```

Query = Fim2C (209 aa)

```
>gb|ADC56706.1| Pili assembly chaperone, N-terminal protein (FimC) [K. variicola At-22]
Length=241

 Score =  392 bits (1007),  Expect = 5e-138, Method: Compositional matrix adjust.
 Identities = 190/207 (92%), Positives = 199/207 (96%), Gaps = 0/207 (0%)

Query  3    QAGVALGATRVIYPAGQKQVQLGVTNNDDSNTYLIQSWIENSENEKDGRFVITPPLFAMQ  62
            QAGVALGATRVIYPAGQKQVQL VTNNDD++T+LIQSW+EN++ ++DGRFVITPPLFAMQ
Sbjct  35   QAGVALGATRVIYPAGQKQVQLAVTNNDDNSTWLIQSWVENADGQRDGRFVITPPLFAMQ  94

Query  63   GKKENILRIIDATNNQLPQDRESLFWVNVKAIPSMDKSKLSDNTLQLAIISRIKLYYRPA  122
            GKKEN LRIIDATNNQLPQDRESLFW+NVKAIPSMDKSKLSDNTLQLAIISRIKLYYRP
Sbjct  95   GKKENTLRIIDATNNQLPQDRESLFWMNVKAIPSMDKSKLSDNTLQLAIISRIKLYYRPG  154

Query  123  KLALAPDQAAEKLTFSRSGSSLTLTNPTPYYLTVTELNAGTRILENALVPPMGKTSVKLP  182
            KLAL PDQAAEKLTFSRSGSSLTLTNPTPYYLTVTELNAGTRILENALVPPMGKTSVKLP
Sbjct  155  KLALPPDQAAEKLTFSRSGSSLTLTNPTPYYLTVTELNAGTRILENALVPPMGKTSVKLP  214

Query  183  PDAGSRITYRTINDYGALTPTMNGVLR  209
             DAGS ITYRTINDYGALTP MNGVLR
Sbjct  215  ADAGSTITYRTINDYGALTPKMNGVLR  241
```

Query = Fim2I (176 aa)

```
>gb|ABV14784.1|  hypothetical protein CKO_03708 (FimI)[Citrobacter koseri ATCC BAA-895]
Length=179

Score =  301 bits (771),  Expect = 1e-103, Method: Compositional matrix adjust.
 Identities = 139/170 (82%), Positives = 157/170 (92%), Gaps = 0/170 (0%)

Query  1    MFLLPPLALAGDHWNVMVAGGSLQFHGVIIAESCRVEVGDQQMTVNMGQVSSNRFHSTGE  60
            +FL P LALAG+HWNV + GG+++F GVIIAESCRVE GDQQMTVNMGQ+SSNRFHS+GE
Sbjct  10   IFLFPSLALAGNHWNVTLPGGNMRFQGVIIAESCRVEAGDQQMTVNMGQISSNRFHSSGE  69

Query  61   DASPVPFDIHLQDCSTKVSQQVGVSFRGVADGKNPDVLSVGEGPGIATGIGIAFFNKEDQ  120
            DA+PVPFDIHLQDCST VSQ+VGVSFRGVADGKNP+VLSVGEGPGIATG+GIA F+K++Q
Sbjct  70   DANPVPFDIHLQDCSTAVSQRVGVSFRGVADGKNPEVLSVGEGPGIATGVGIALFDKDNQ  129

Query  121  LIPLNGPPSLWQRLYTGPTTLHFVAKYRATGKQVTGGMANAQVWFSLTYQ  170
            LIPLN PP  W +LY+GPTTLHFVAKYRATGKQVTGG ANAQ+WFSLTYQ
Sbjct  130  LIPLNSPPGAWTQLYSGPTTLHFVAKYRATGKQVTGGAANAQIWFSLTYQ  179
```

Query = Fim2A (182 aa)

```
>gb|ABR78685.1|  major type 1 subunit fimbrin (FimA) [Klebsiella pneumoniae subsp.
pneumoniae MGH 78578]
Length=182

 Score =  298 bits (764),  Expect = 2e-102, Method: Compositional matrix adjust.
 Identities = 160/182 (88%), Positives = 171/182 (94%), Gaps = 0/182 (0%)

Query  1    MNAKIFVIAAMSALSLSSAAALANTTTVNGGTVHFKGEVVNAACAVDAGSIDQTVQLGQV  60
            M  K   +  +SALSLSS AALA+TTTVNGGTVHFKGEVVNAACAVDAGSIDQTVQLGQV
Sbjct  1    MKIKTLAMIVVSALSLSSTAALADTTTVNGGTVHFKGEVVNAACAVDAGSIDQTVQLGQV  60

Query  61   RSAKLATAGSTSSAVGFNIQLDDCDTSVATKASVAFSGTSVDTTNTSVLALQNSAAGSAT  120
            RSAKLATAGSTSSAVGFNIQLDDCDT+VATKASVAF+GT++D++NT+VLALQNSAAGSAT
Sbjct  61   RSAKLATAGSTSSAVGFNIQLDDCDTTVATKASVAFAGTAIDSSNTTVLALQNSAAGSAT  120

Query  121  NVGVQILDKTGTPLALDGASFSAATTLNDGTNIIPFQARYYATGAATAGTANADATFKVQ  180
            NVGVQILD TGTPLALDGA+FSAATTLNDG NIIPFQARYYATGAATAG ANADATFKVQ
Sbjct  121  NVGVQILDNTGTPLALDGATFSAATTLNDGPNIIPFQARYYATGAATAGIANADATFKVQ  180

Query  181  YE  182
            YE
Sbjct  181  YE  182
```

Query = Orf10 (215 aa)

```
>gb|ADO74380.1|  Two-component system, regulatory protein [Stigmatella aurantiaca DW4/3-1]
Length=213

 Score = 43.5 bits (101),  Expect = 0.011, Method: Compositional matrix adjust.
 Identities = 24/65 (37%), Positives = 38/65 (58%), Gaps = 1/65 (2%)

Query  119  PEKIRGSEWSTMLYLMSGISNRQLAYMLKTSEKTLSGRVNNLAIKLGLVGFNKALQLRAM  178
            PE +   E   M  +  G+SNR++A +L T+E T+  + +N+  KLG+    +A+ LRAM
Sbjct  150  PEGLTRRELEVMRLIARGLSNREIATVLGTAEGTVKNQTSNILAKLGVRDRTRAV-LRAM  208

Query  179  NLFYL  183
             L  L
Sbjct  209  ELGCL  213
```

Query = Orf11 (281 aa)

```
>gb|EFE96725.1|  hypothetical protein HMPREF0758_1584 [Serratia odorifera DSM 4582]
Length=264

 Score =  103 bits (257),  Expect = 1e-23, Method: Compositional matrix adjust.
 Identities = 77/277 (28%), Positives = 123/277 (44%), Gaps = 30/277 (11%)

Query  15   YTIEGNIHFCPNENTL----RNVIDGRTVTLLSTASECFKILIENQGRVISRDEMKELVW  70
            Y I   + F P E +L     ++     + L +  S+C  +LIE  G V+S+DE+ EL+W
Sbjct  5    YIINDRVIFDPEERSLILTGEHIGSTSKLILHTPTSQCLALLIERHGDVLSQDELLELIW  64

Query  71   GKRGVIVSSNTFYQNMLNLRRGLEKVGAGKRLISTHYGKGVTIDNGLDITPIVTSVSNC-  129
             K+GV+VSS+T YQN+  LRR L ++G    +I T   +GVT+   + IT +    ++
Sbjct  65   RKKGVVVSSSTVYQNISLLRRSLNQLGLYDDVIITVPRQGVTLSRHIKITLLAEQHADVR  124

Query  130  ----LEPMNGTKNIPFSTSEAINTADEHFLNNFKQLEEKGIIKKNRYKHGMILITNLLLL  185
                L  + G    P  T+    +                 K  +Y  G+I   ++ L+
Sbjct  125  QQELLPAVAGDVAAPTGTTPISPS------------------KGKKYLPGLIAGLSIALI  166

Query  186  FFCILMFYTAHNLKRENYFSHYTKTKLKVKNCDIYTDPALLNEHELNELLKQNPPDCEEG  245
              CI  FY  HN    NY   Y      +  C +Y +   +    L  ++K +  DC
Sbjct  167  MLCIGAFYLGHN-NSSNYMDSYDLAPTSLSECHMYVNDN-MPTGRLEYIVKNSAFDCATR  224

Query  246  ETLYLSSVYPIQRMSVIRCIGTFLPGE-ECESDYYLE  281
               Y+++   + R SVI C+         C S +YLE
Sbjct  225  PFNYVTAFAGVNRYSVISCVRDIAGARASCRSSFYLE  261
```

Query = Orf12 (148 aa)

```
>gb|EFE96270.1| conserved hypothetical protein [Serratia odorifera DSM 4582]
Length=155

 Score = 79.3 bits (194),  Expect = 9e-17, Method: Compositional matrix adjust.
 Identities = 38/130 (29%), Positives = 67/130 (52%), Gaps = 0/130 (0%)

Query  28   KAQHSDTMTCDAIINVHKEHQLLSIRLKYSFLGHEGVANLSGVLNNGNKGIFSISRTVLF  87
            + ++ +  +C A          ++  +++ F   +G+A L G + +G   + ++SR VLF
Sbjct  26   EVKNDNQFSCTARYTTQMNDVQINSIIRFIFNDGDGIATLDGSMIDGPGPLKTVSRKVLF  85

Query  88   KYKRINEGVYLENTTTTISLQDTAGPDGLKKLLPEFYTEHGARTSFQVYAQKPGGYIFVK  147
            KY R N   +L  T  ++ L D +  D +KK +P+FY    A  +F +Y  K   Y+F
Sbjct  86   KYDRKNTSYFLMTTNASVQLDDESNTDEMKKNVPDFYLIKDATINFDIYPLKQNSYLFTS  145

Query  148  EFTPAFYCTE  157
               P+FYCT+
Sbjct  146  RKIPSFYCTK  155
```

Query = Orf13 (190 aa)

```
>gb|EFD84432.1|  hypothetical protein HMPREF0485_03210 [Klebsiella sp. 1_1_55]
Length=190

 Score =  389 bits (999),  Expect = 6e-138, Method: Compositional matrix adjust.
 Identities = 184/190 (97%), Positives = 189/190 (99%), Gaps = 0/190 (0%)

Query  1    MAMVNENIVIRLAQDVDIPEVKQLLERYHAKNLAGEQRANGFVTTDMTEQQLSELSSAES  60
            MAMVNENIVIRLAQDVDIPEVKQLLERYHAKNLAGEQRANGFVTTDMT+QQLSELSSAES
Sbjct  1    MAMVNENIVIRLAQDVDIPEVKQLLERYHAKNLAGEQRANGFVTTDMTQQQLSELSSAES  60

Query  61   GVVIAVDRSCDKVIGLLLGGSWEFLSPWPMFKYMASILNEYRYQGEKLDAVSSYQYGPIC  120
            GVVIAVDRSC+KVIGLLLGGSWEFLSPWPMFKYMASILNEYRYQG+KLDA SSYQYGPIC
Sbjct  61   GVVIAVDRSCNKVIGLLLGGSWEFLSPWPMFKYMASILNEYRYQGKKLDAASSYQYGPIC  120
```

```
Query  121  VAEEYRGRGVGEMLLEYQRKVFAPRYPVIVTFVNVLNPRSYAFHTRNQFEDVGFFNFNGN  180
            VAEEYRG+GVGE+LLEYQRKVFAPRYPVIVTFVNVLNPRSYAFHTRNQFEDVGFFNFNGN
Sbjct  121  VAEEYRGQGVGELLLEYQRKVFAPRYPVIVTFVNVLNPRSYAFHTRNQFEDVGFFNFNGN  180

Query  181  KYHMMALPTS  190
            KYHMMALPTS
Sbjct  181  KYHMMALPTS  190
```

Query = Orf14 (39 aa)

```
>gb|ACI07992.1|  hypothetical protein KPK_0640 [Klebsiella pneumoniae 342]
Length=39

 Score = 76.6 bits (187),  Expect = 8e-19, Method: Compositional matrix adjust.
 Identities = 39/39 (100%), Positives = 39/39 (100%), Gaps = 0/39 (0%)

Query  1   MALMVLPAIKTNGEIIRKVLFVTVGAMMVTLSVKYFTGV  39
           MALMVLPAIKTNGEIIRKVLFVTVGAMMVTLSVKYFTGV
Sbjct  1   MALMVLPAIKTNGEIIRKVLFVTVGAMMVTLSVKYFTGV  39
```

Query = Orf15 (77 aa)

```
>gb|ACI06987.1|  hypothetical protein KPK_0639 [Klebsiella pneumoniae 342]
Length=77

 Score =  149 bits (377),  Expect = 1e-46, Method: Compositional matrix adjust.
 Identities = 71/77 (92%), Positives = 71/77 (92%), Gaps = 0/77 (0%)

Query  1   MFMKNGIILSIADRIKDAFPIPTDEFRKRLEEDCYVIWKLPQFYTWQITLSSIIPACTDY  60
           MFMKNGIILSIADRIKDA PIPTDEFRKRLEEDCYVIWKLPQFY  QIT SSIIP CTDY
Sbjct  1   MFMKNGIILSIADRIKDALPIPTDEFRKRLEEDCYVIWKLPQFYIRQITQSSIIPTCTDY  60

Query  61  IEAINAEVHGAYGSARH  77
           IEA NAEVHGAYGSARH
Sbjct  61  IEAKNAEVHGAYGSARH  77
```

Query = Orf16 (281 aa)

```
>gb|ACI07748.1|  metallo-beta-lactamase family protein [Klebsiella pneumoniae 342]
Length=281

 Score =  539 bits (1389),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 256/281 (91%), Positives = 270/281 (96%), Gaps = 0/281 (0%)

Query  1    MRLTVLVDNNTLIDRYLIGEPGVSYLIEYDGQKILFDTGYSDVFLKNAQTLKIDLTSIDS  60
            M+LTVLVDNNTLIDRYL+GEPGVSYLIE+DGQKILFDTGYSDVFL+NAQTL+IDLT+IDS
Sbjct  1    MQLTVLVDNNTLIDRYLVGEPGVSYLIEHDGQKILFDTGYSDVFLQNAQTLRIDLTTIDS  60

Query  61   IVFSHGHNDHTWGLNHLVQYYDRINFIPERKINLICHPDALTPKYFDAKAIGINYRFDQN  120
            IV SHGHNDHTWGLNHL Q+YDR+NF PERKINL+CHPDAL PKYFDAK+IGINYRFDQN
Sbjct  61   IVLSHGHNDHTWGLNHLTQHYDRLNFTPERKINLVCHPDALNPKYFDAKSIGINYRFDQN  120

Query  121  DPFFDKICSKKPYPLTENIIFLGQIPRDNKFEMLTPVGQTVDDKSEITDDYVMDDSALAI  180
            D FFDKICSKKPYPLTENIIFLGQIPRDNKFEMLTPVGQTVDD  EITDDYVMDDSALAI
Sbjct  121  DLFFDKICSKKPYPLTENIIFLGQIPRDNKFEMLTPVGQTVDDNGEITDDYVMDDSALAI  180

Query  181  KTEDGLVVVTGCSHAGIANIIEYAKQVTGEKHIVSVIGGFHLQNADEDRLTQTGDYLKVL  240
            KTE+GLVVVTGCSHAGIANIIEYAKQVTGE HIVSVIGGFHLQNADEDRLTQTGDYLK L
Sbjct  181  KTEEGLVVVTGCSHAGIANIIEYAKQVTGENHIVSVIGGFHLQNADEDRLTQTGDYLKAL  240

Query  241  SPDSVYPCHCTDLAAKVSLARFIKINEVGVGLTLNFSATGK  281
            SPDS+YPCHCTDLAAK+SLARFIKINEVGVGL L+FSA GK
Sbjct  241  SPDSLYPCHCTDLAAKISLARFIKINEVGVGLKLDFSAAGK  281
```

Appendix 9. qRT-PCR results

		rpoD		*fimA*		*fimA2*		*mrkA*	
PCR efficiency		91 %		90 %		89 %		91 %	
		Ct values		**Ct values**		**Ct values**		**Ct values**	
		Replicates	**Average**	**Replicates**	**Average**	**Replicates**	**Average**	**Replicates**	**Average**
LB	**1**	24.00		24.96		30.45		23.58	
37ºC, 200 rpm	**2**	23.74	23.89	24.98	24.95	30.34	30.38	23.38	23.45
	3	23.94		24.90		30.35		23.38	
LB	**1**	24.55		25.30		29.24		22.99	
25ºC 200 rpm	**2**	24.73	24.65	25.56	25.39	29.47	29.33	23.06	23.04
	3	24.67		25.31		29.29		23.06	
M9	**1**	25.84		29.28		29.79		24.70	
37ºC, 200 rpm	**2**	25.66	25.80	28.77	29.08	29.66	29.77	24.76	24.70
	3	25.91		29.20		29.87		24.65	
AUM	**1**	26.02		29.28		32.69		25.96	
37ºC, 200 rpm	**2**	25.88	26.00	28.77	29.08	32.67	32.65	25.83	25.95
	3	26.09		29.20		32.59		26.05	
RPMI 1640	**1**	30.17		34.60		35.99		28.71	
37ºC, 200 rpm	**2**	30.33	30.47	34.73	34.72	37.17	36.77	28.91	28.84
	3	30.92		34.83		37.14		28.91	
RPMI + FCS	**1**	25.45		29.11		32.27		27.90	
37ºC, 200 rpm	**2**	25.65	25.56	29.42	29.23	32.10	32.17	28.17	27.97
	3	25.57		29.17		32.15		27.83	
Bladder sample	**1**	32.00		23.50		34.46		28.62	
	2	31.80	32.01	23.51	23.50	36.31	35.42	28.56	28.64
	3	32.23		23.49		35.48		28.74	

Appendix 10. Journal of Microbiological Methods Paper

J. J. van Aartsen and K. Rajakumar

An optimized method for suicide vector-based allelic exchange in *Klebsiella pneumoniae*.

J. Microbiol. Methods. 2011. 86:313-319.

Note 18/12/2011:

Paper has been removed to comply with Leicester Research Archive Copyright rules.