NORTHWESTERN UNIVERSITY

Genome-wide identification of shared and strain-specific factors necessary for

Klebsiella pneumoniae enteric colonization

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Driskill Graduate Training Program in Life Sciences

By

Bettina H. Cheung

EVANSTON, ILLINOIS

September 2023

ABSTRACT

Gastrointestinal (GI) colonization by *Klebsiella pneumoniae* is a risk factor for subsequent infection as well as transmission to other patients. Additionally, colonization is achieved by many strain types that exhibit high diversity in genetic content. However, how *K. pneumoniae* achieves colonization and whether the genetic factors it uses differ by strain is not well understood. In this dissertation, we developed a mouse model of GI colonization which supported studies into how strains differ in their colonization capacities and how colonization factors are shared or strain-specific. We found that strain-to-strain differences in colonization fitness could be quantified in competition *in vivo*. Additionally, we created saturating transposon mutant libraries in three clinical isolates of *K. pneumoniae* to screen for colonization factors *in vivo*. With these genomic screens, we identified a core colonization program as well as strain-specific colonization factors. Furthermore, we validated three shared colonization factors (*acrA*, *carAB*, and *tatABCD*), one factor which provided a colonization advantage when disrupted (*malT*), and two factors which were strain specific (*hha* and *scrY*) Thus, our data provide insight into how a global pathogen of concern establishes gastrointestinal colonization. Our approach reveals that understanding pathogenesis in species with high genomic diversity requires examination of multiple genetically distinct strains.

ACKNOWLEDGEMENTS

This work would not have been possible without all of the wonderful members of the Hauser Lab. I would like to thank my labmates for all of the good times and advice and support during the not-so-good times. In particular, I would like to thank Sophie Nozick for being a fantastic friend and for her incredible work supporting the entire lab logistically, Travis Kochan for his advice on *Klebsiella* and being a scientist, and Sumitra Mitra, Marine Lebrun Corbin, Tim Turner, Rachel Medernach, Aliki Valdes, and Christopher Axline.

I would also like to thank Dr. Alan Hauser for his exceptional mentorship. I have enjoyed learning how to investigate scientific questions rigorously and continue to appreciate his commitment to considering all the possibilities and implications raised by our data. He has empowered me on my path to becoming a physician scientist, especially as we experienced numerous challenges and project hiccups along the way. I am also grateful that Al has always been supportive of me as a whole person and has been genuinely interested in all of my endeavors—even those outside of science.

On that note, I would like to thank Dr. Gregory Brisson and all his support for me as a student as well as for the Northwestern Medical Orchestra. Co-founding this organization with fellow medical student Michael Wang and playing in and leading it through both wonderful and difficult times has been an integral part of my time at Feinberg. I would like to thank Michael for years of friendship, constant and seamless teamwork, and indefatigable support. I can't imagine how I would have been able to complete my graduate work without the orchestra, and I thank Dean Eric Neilson for his support of this organization.

I would also like to acknowledge the teachers and professors who guided me towards this path and encouraged me to pursue my passions. In particular at Ardsley School District, Alan

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 *Klebsiella pneumoniae* is a major cause of nosocomial infections and multidrug-resistant infections

     Klebsiella pneumoniae is a Gram-negative bacterium belonging to the family Enterobacteriaceae. While it is most studied as a human pathogen, *K. pneumoniae* is also known to inhabit a wide range of environments. It is found in soil, water sources, companion animals, and livestock[1-4]. This ability to survive in such diverse environments also hints at its metabolic versatility. Another line of research outside of *K. pneumoniae* pathogenesis is metabolic engineering to produce products such as 2,3-butanediol and biofuels such as 2-butanol for commercial uses[5,6]. Furthermore, *K. pneumoniae* is known to colonize abiotic surfaces, forming biofilms in places relevant to clinical care such as sink drains and prosthetic devices[7,8].

     As a human pathogen, *K. pneumoniae* primarily causes opportunistic infections in the United States[9]. As its name suggests, *K. pneumoniae* frequently infects the respiratory tract to cause pneumonia[9]. However, the bacteria can also infect other sites, causing urinary tract infections, liver abscesses, skin and soft tissue infections, and bacteremia[10]. Infections by *K. pneumoniae* are usually limited to immunocompromised patients and linked to healthcare settings in the US[9,11]. However, the Centers for Disease Control and Prevention recently named *K. pneumoniae* an urgent concern, as many strains are multidrug-resistant[12]. The prevalence of *K. pneumoniae* resistant to the most commonly used drugs for Gram-negative infections, such as third-generation cephalosporins and carbapenems, has risen from 5.3% to 11.5% between 1999 and 2010 in the United States[13]. These figures match those from more recent studies in which 11.2% of *K. pneumoniae* isolates from the United States in the SMART global surveillance program between 2015 and 2019 were found to be extended-spectrum beta-lactamase

producers[14]. These strains present a significant challenge for treatment, forcing clinicians to use

last-resort antibiotics with significant toxicity[15]. Additionally, case reports have begun to surface

of *K. pneumoniae* non-susceptible to every antibiotic currently available for medical use[16], with

some strains exhibiting resistance to seldom-used antibiotics they have not been exposed to them

previously[17]. To highlight the clinical impact of multidrug-resistant *K. pneumoniae*, carbapenem-

resistant *K. pneumoniae* bacteremia and pneumonia cause 27% of excess mortality in

hospitalized patients, or mortality that can be attributable to the infection rather than non-

infectious causes[18].

      Of these multidrug-resistant *K. pneumoniae*, some strains are referred to as "high-risk

clones." This term is used to describe strains with global spread and an increased propensity for

pathogenicity and antimicrobial resistance[19,20]. As many antimicrobial resistance genes are

encoded on mobile genetic elements such as plasmids[19,21,22], strains that pick up these elements

along with adaptations allowing them to spread more efficiently are thus "higher risk" for being

vehicles for worldwide dissemination of multidrug-resistant infections. Two high-risk clones of

K. pneumoniae of note for this project are the ST45 and ST258 multi-locus sequence types which

belong to a clonal group (closely related ST), CG258[23]. Multi-locus sequence typing (MLST)

uses the sequences of 7 housekeeping genes to group strains together which are likely to be more

closely related to each other[24]. This nucleotide-based method is more convenient for analyses of

large collections than methods such as pulsed-field gel electrophoresis (PFGE) which may be

more specific but more time-consuming and technically challenging[24].

      The first high-risk clone mentioned above, ST45, frequently produces extended-spectrum

beta-lactamases (ESBLs), making it resistant to most antimicrobials except carbapenems. One

example of these ESBLs is the CTX-M family. This family can be carried on a variety of mobile

genetic elements and plasmids[22], leading to CTX-M ESBLs becoming the most common ESBLs worldwide[25,26]. The second high-risk clone, ST258, is notorious for producing carbapenemases and causing a significant portion of carbapenem-resistant infections worldwide[27]. Many carbapenem-resistant *K. pneumoniae* carry one of the <u>*Klebsiella pneumoniae* carbapenemases</u> (bla$_{KPC}$) genes, but some strains carry the genes for other carbapenemases such as New Delhi metallo-beta-lactamases (NDM) or Verona Integron-encoded Metallo-beta-lactamase (VIM)[28]. In addition to resistance to carbapenems and beta-lactams, ST258 isolates frequently encode for resistance to other classes of antibiotics. For instance, one study reports that ST258 isolates carry an average of 15.4 antimicrobial resistance genes. These range from resistance to aminoglycosides to quinolones to sulfonamides, and more[29]. Finally, genomic analysis of over 300 *K. pneumoniae* strains demonstrates that there is a bimodal distribution of antimicrobial resistance genes encoded by strains. In other words, strains either do not show any antimicrobial resistance or carry a large number of antimicrobial resistance genes (mode of 10)[30].

Phylogenetic and genomic analysis of the ST258 lineage reveals that persistent subclades (or those which can be isolated repeatedly over time) harbor a specific integrative conjugative element: ICE*Kp10*, which contains both the siderophore yersiniabactin and the genotoxin colibactin[29]. This overabundance in the persistent isolates and presence in two distinct clades of ST258 indicates that there is likely a fitness advantage associated with ICE*Kp10*. As such, the acquisition of this mobile genetic element may be demonstrative of a high-risk clone continuing to evolve and acquire extra factors that increase fitness.

Concurrent with this rise in antibiotic resistance, cases of "hypervirulent" *K. pneumoniae* are increasing worldwide. In contrast to the strains causing hospital-acquired infections in immunocompromised patients (termed "classical" strains), hypervirulent *K. pneumoniae* can

infect young, healthy hosts in the community[31]. These strains also cause a larger proportion of invasive infections such as bacteremia, pyogenic liver abscesses, meningitis, and endophthalmitis[32]. While most cases of hypervirulent *K. pneumoniae* infection are concentrated along the Asian Pacific Rim, and the large majority are responsive to conventional antibiotics, there is concern these strains will spread and converge with increased antibiotic resistance. Unfortunately, case reports of hypervirulent, antimicrobial-resistant strains have already appeared in the literature. Hypervirulent *K. pneumoniae* infections have surfaced in the United States[33,34], and carbapenem-resistant hypervirulent *K. pneumoniae* has been detected in China[35].

As we are facing a pathogen that is becoming increasingly virulent and showing signs of becoming impossible to treat with current therapies, research efforts must be focused towards uncovering the mechanisms by which *K. pneumoniae* causes infection. This, in turn, will guide new prevention and treatment strategies.

1.2 Gastrointestinal colonization by *K. pneumoniae*

K. pneumoniae is known to colonize both the nasopharynx and gastrointestinal tract. Sequence-based approaches such as the Human Microbiome Project detected *K. pneumoniae* in the mouth, nares, and stool[36]. Additionally, sequencing of nasal swabs prompted by the COVID-19 pandemic demonstrated the presence of *K. pneumoniae* in the nasopharynx of both healthy controls and SARS-CoV-2 positive individuals[37,38]. These data are corroborated by culture-based methods from two studies from Indonesia and Vietnam demonstrating the prevalence of oropharyngeal carriage to be between 7 to 15%[39,40].

While the significance of nasopharyngeal carriage of *K. pneumoniae* is not well described in the literature, the prevalence and consequences of gastrointestinal colonization has been more extensively investigated. In this context, GI colonization refers to asymptomatic carriage of *K.*

pneumoniae in the GI tract. Detection of community prevalence of colonization ranges is

estimated to be around 6%[36,41]. However, in populations with recent healthcare contact, carriage

jumps up to around 20%[41,42]. Furthermore, a systematic review estimated the prevalence of GI

colonization with carbapenem-resistant *K. pneumoniae* at 5.3%, but included studies ranged from

0.13% to 22%[43]. Additionally, previously hospitalized patients may also remain colonized by

carbapenem-resistant strains for several years after exposure[44].

Recent studies have shown that patients who are colonized by *K. pneumoniae* are at

increased risk for subsequent *K. pneumoniae* infection[45]. In one study, ICU patients who were

not colonized with *K. pneumoniae* at baseline had a 3% incidence of infections. This percentage

jumped to 16% developing infections in those who did carry *K. pneumoniae*[41]. Additionally,

through whole-genome sequencing, up to 80% of these infections can be traced back to the

colonizing strain, confirming that indeed, patients are infected by strains they carry in their GI

tract[41]. The use of whole-genome sequencing also led to the identification of transmission events,

where patients with overlapping stays who had negative initial rectal swabs for *K. pneumoniae*

converted to positive for strains carried by other patients on the same clinical unit[41].

What, then, are the risk factors for being colonized with *K. pneumoniae*? In the many

epidemiological studies on colonization by *K. pneumoniae*, the focus is primarily on

carbapenem-resistant *K. pneumoniae*, as these strains constitute a major public health concern.

The main risk factor found in many of these studies is prior antibiotic administration[46-48]. As

antibiotics can perturb the gut microbiota, which is responsible in part, for colonization

resistance, this is not a surprising finding. Previous studies in animals have shown the gut can be

made permissive to GI colonization when antibiotic administration eradicates a protective phyla

of the microbiota[49]. However, unless patients are admitted to an ICU which regularly collects

surveillance swabs for carbapenem-resistant *K. pneumoniae*, they are not regularly screened for *K. pneumoniae* carriage. Thus, the consequences of different antibiotic regimens on increasing the risk of colonization in non-critically ill patients are not well defined. Other than antibiotic therapy, higher proportions of patients who needed repositioning by staff converted to positive carbapenem-resistant *K. pneumoniae* on rectal swab despite increased infection control measures, highlighting the risk for transmitting carriage in the hospital[50].

Previous models of *K. pneumoniae* GI colonization

As colonization of the GI tract is a multifactorial process influenced by host factors and the microbiome, animal models thus far have been in mice rather than other non-mammalian model systems. As in humans, GI colonization in this context refers to asymptomatic carriage of *K. pneumoniae* in the GI tract. However, considerable variation still exists in experimental approaches developed by various groups to achieve colonization in animals. These differ in any and all of the following characteristics: (1) administration of antibiotics/other types of microbiome alterations (2) host immune modulation (3) administration of *K. pneumoniae* and (4) type of mice. These different experimental conditions mimic varying host scenarios for carriage of *K. pneumoniae*, whether that be colonization of non-immunosuppressed individuals or populations most at risk for infection, such as hospitalized patients receiving antimicrobial therapy (**Figure 1**). Additionally, the two pathotypes of *K. pneumoniae* (classical and hypervirulent) exhibit different patterns of acquisition (hospital vs community, respectively) and levels of infection invasiveness (low vs high). The latter can influence selection of experimental conditions to limit the dissemination of hypervirulent bacteria from the GI tract into other tissues, so as to more accurately mimic asymptomatic colonization versus productive infection.
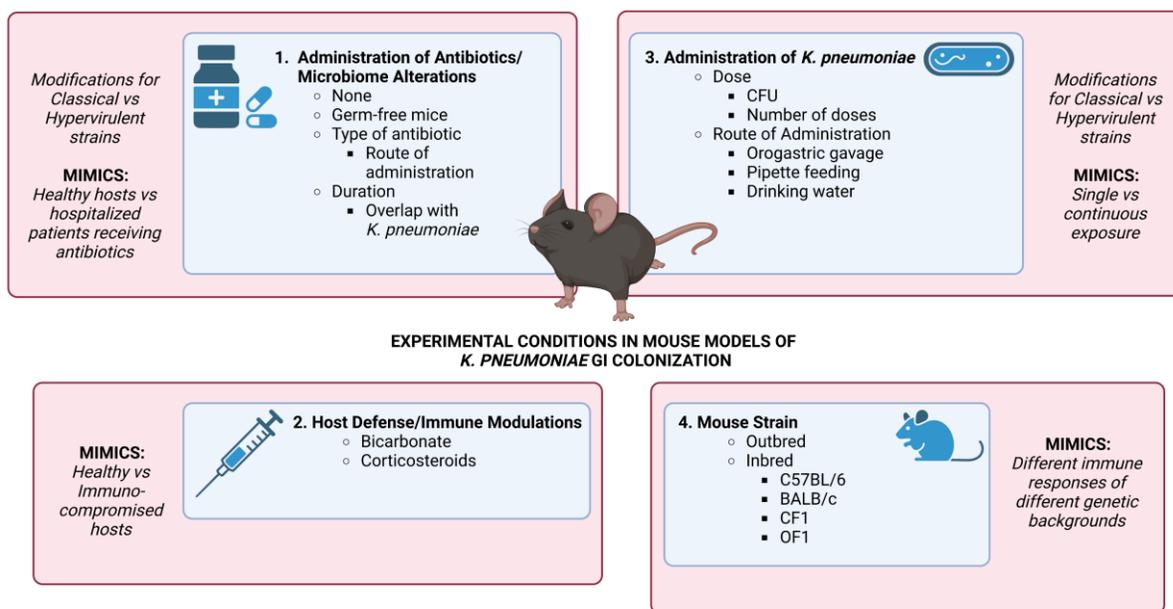
Figure 1 Differences in experimental approaches to modeling *K. pneumoniae* GI colonization in mice

Experimental approaches to modeling *K. pneumoniae* GI colonization in mice can differ greatly to represent the wide spectrum of *K. pneumoniae* colonization states observed in humans. The modifiable characteristics are found in the blue boxes while the relevant populations approximated are in the surrounding red boxes.

The first characteristic that is commonly included is the use of antibiotics. To establish susceptibility to colonization, mice are first treated with antimicrobials to perturb the microbiome and reduce colonization resistance to achieve dense colonization by *K. pneumoniae*. The microbiome contributes to colonization resistance in a variety of ways: stimulation of host immunity, direct antagonism of pathogenic colonizing bacteria, and competition for nutrients. First, *Bacteroides thetaiotamicron*, a member of the microbiota, is known to stimulate the host to increase production of antimicrobial peptides[51]. It also releases a soluble factor to suppress Shiga

toxin production by enterohemorrhagic *Escherichia coli*[52]. Additionally, microbiota-produced butyrate, a short-chain fatty acid (SCFA), is known have host-protective effects as well as suppressing virulence factor expression by *Salmonella enterica*[53]. Next, members of the microbiome can produce bacteriocins, or proteins that can kill or inhibit growth by other bacteria to prevent colonization[54]. Finally, *B. thetaiotamicron* has been shown to outcompete the pathogen *Citrobacter rodentium* for utilization of monosaccharides, effectively preventing colonization by depletion of preferred nutrients[55]. As such, most *K. pneumoniae* colonization experiments aim to suppress the microbiome and its protective role to establish high levels of *K. pneumoniae* in the gut.

However, some hypervirulent strains are able to colonize the gut without the use of antibiotics. For instance, the lab strain KPPR1 and the hypervirulent clinical isolate hvKP1 are shed at 10^7 CFU/g feces for two weeks post-inoculation[56]. Fecal burden is a commonly used metric for colonization, as it is a convenient and nonlethal way to estimate the presence of *K. pneumoniae* within the GI tract. This successful antibiotic-free colonization by hypervirulent strains is in contrast to the far lower (undetectable to $10^{4.5}$ CFU/g feces) carriage of a classical strain[56]. While this fecal burden may be too low to serve as a useful readout for experiments requiring quantification of changes in colonization, another study demonstrated that classical strains can still be present in the gut despite being undetectable in feces[57]. In this study of occult colonization, outgrowth of a classical strain which was initially undetectable in feces could be achieved by administration of antibiotics anytime within 2 weeks after initial inoculation[57]. Intriguingly, even this low level of occult colonization was associated with changes in the host microbiome.

For experimental approaches which utilize antimicrobials, regimens range from single oral doses of streptomycin prior to inoculation with *K. pneumoniae* to continuous administration of a cocktail (metronidazole and vancomycin) in the drinking water prior to inoculation and through the duration of the experiment[58-60]. As timing and duration of antibiotic administration varies considerably between studies, it is difficult to compare head-to-head the effects of different antibiotics on colonization. However, a few studies have made these direct comparisons. For instance, Perez et al (2011), measured fecal burden of a classical carbapenem-resistant strain following administration of six different antibiotics from two days prior to inoculation to 5 days after. In this study, clindamycin supported the highest levels of shedding (10^{10} CFU/g) for the first week which then decreased to $10^{5.5}$ CFU/g by Day 11 post-inoculation. Tigecycline and piperacillin/tazobactam produced the next highest levels of shedding, 10^7 CFU/g, and ertapenem, cefepime, and ciprofloxacin were associated with the lowest levels of shedding (10^3 CFU/g)[61]. The authors noted that the antibiotics (clindamycin and piperacillin/tazobactam) which decreased anaerobes, including Bacteroides species, were associated with the highest levels of *K. pneumoniae* colonization. This observation is in line with the findings that Bacteroidetes are protective against colonization[49]. This study also demonstrated that in germ-free mice without any colonization resistance from the microbiome, classical *K. pneumoniae* were able to achieve dense colonization to 10^9 to 10^{10} CFU/g feces—a phenotype that was abolished by the introduction of Bacteroidetes but not Firmicutes or Actinobacteria[49]. In addition to the type of antibiotic used, the timing of administration may also affect colonization success. Le Guern et al (2019) demonstrated that clindamycin (one of the antibiotics shown by several studies to support high levels of colonization[61-64]) could only induce colonization by *K. pneumoniae* if administered within a week before or after inoculation[65]. As all

antibiotics have different pharmacokinetics as well as differing effects on the microbiome, further experiments would have to be conducted to determine the timeline for when individuals may be at risk for colonization when treated with other antibiotics. Nevertheless, this study demonstrates that susceptibility to *K. pneumoniae* colonization induced by antibiotics is a time-limited effect.

While not precisely host immune modulation, neutralization of stomach acid as a defense against *K. pneumoniae* reaching the more distal parts of the GI tract has also been used by some groups to facilitate colonization[66]. With this experimental condition, the hypervirulent strain CIP52.145 was shed in the feces at 10^2 CFU/g. To investigate more directly the role of the immune system in colonization resistance against *K. pneumoniae*, Sequeira et al (2020) administered dexamethasone (a corticosteroid) to mice which had not received antibiotics. This immunosuppressive treatment supported subsequent colonization by *K. pneumoniae* (10^8 CFU/g feces) at similar levels achieved by administration of an antibiotic cocktail consisting of metronidazole, neomycin, vancomycin, and ampicillin in the drinking water[49]. *K. pneumoniae* infections are more common among individuals in more immunosuppressed states, such as the elderly and alcoholics, and using immunosuppressive treatments such as dexamethasone could be used to mimic these patients even in the absence of antibiotic treatment.

The third factor varied in experimental approaches is the administration of *K. pneumoniae*. The inoculation dose can vary widely between studies: anywhere from 10^3 CFU to 10^8 CFU. Next is the route of administration. While most groups use orogastric gavage, some use oral pipette feeding[56] or exposure in drinking water[65]. The last two methods are less challenging technically than orogastric gavage, but both gavage and pipette feeding have an advantage over drinking water exposure in delivering a specified inoculum to each mouse. Next, number of *K*

pneumoniae doses can also be varied. Although most studies use a single orogastric gavage of *K.*

pneumoniae, some use an additional gavage dose a few days after the initial inoculation[60,62].

Unfortunately, most studies do not specify why certain dosing regimens were chosen or provide

data from pilot experiments to demonstrate the difference between size or timing of

inoculation(s). A last general strain factor that differs between studies is the type of *K.*

pneumoniae strain used. Of note, many studies use a lab train of *K. pneumoniae* known as

KPPR1. This strain is of the hypervirulent pathotype and exhibits a hypermucoid capsule[67].

Using KPPR1 has the benefits of using any particular lab strain. For instance, it is a genetically

tractable strain and tools such as ordered transposon mutant libraries are available[67]. However, as

hypervirulent strains are not the most common pathotype encountered in the United States,

results of studies using KPPR1 may not be generalizable to classical strains which lack

hypervirulent features. In contrast, there is no standard lab strain used to study classical strains.

However, the clinical isolate KPNIH1, an ST258 strain which caused an outbreak at the NIH in

2011[68], has been used in several recent studies—especially because an ordered transposon library

created with this strain is available[56,69].

Finally, the last characteristic that can be altered is the type of mice used. Different

groups have used both inbred and outbred mice. Within inbred mice, most studies use the

C57BL/6 strain whereas a few others use BALB/c, CF1, or OF1 mice. While different strains of

inbred mice are known to have some differences in immune response[70,71], no studies have

compared the same colonization regimens in different strains of mice. However, in two studies—

one with outbred mice[62] and the other with C57BL/6 mice[49]—levels of colonization at day 3

post-inoculation are similar (10^9 CFU/g feces). While the antibiotic regimens used prior to

inoculation differed, this may suggest that the effect of antibiotic treatments may be greater than whatever effect is exerted by the differences in strain of mouse used.

In conclusion, though daunting to keep track of, the wide diversity of experimental approaches to establish *K. pneumoniae* GI colonization in animals allows researchers to predict the impact of this bacterium on different patient groups and to assess risk factors.

Known *K. pneumoniae* virulence and GI colonization factors

Known virulence factors have mainly been characterized in infection but not colonization However, we will discuss the main virulence factors below alongside what is known about their role in colonization to provide context in interpretation of the factors identified in this study.

Capsule

The virulence factors of *K. pneumoniae* are noteworthy in that they mainly involve immune evasion and adhesion rather than toxin production or direct damage to host tissues. One such example is the polysaccharide capsule that many strains produce which coats the exterior of the cell. Capsule production is encoded by genes in the *cps* locus which varies widely between strains[72]. For the purposes of *K. pneumoniae* pathogenesis, capsule has been implicated in evading the immune system through multiple mechanisms: inhibition of phagocytosis[73], evasion of complement[74], and protection against antimicrobial peptides[75]. The capsule is required for full virulence (mortality) in a mouse model of pneumonia[76], and a capsule deletion mutant elicits an earlier innate immune response and cytokine production. Hypervirulent strains of *K. pneumoniae* produce a hypermucoid capsule that can be identified by a positive string test. The hypercapsule has also been shown to increase virulence in liver abscess in mice[77]. However, the role of capsule in colonization is less clear. Older studies suggest capsule aids in adhesion to intestinal mucosa

and persistence in the GI tract[78], but recent work demonstrates that certain methods for capsule

knockout used in these older studies may disrupt other factors as well (such as the cell

membrane) and affect viability[79]. This work by Tan et al (2020) also showed that various capsule

mutants (*rmpA*, *wcaJ*, *wza*, and *wzy*) do not have a competitive defect in the GI tract in

antibiotic-treated mice[79]. However, some studies which have made use of *manC* knockouts

(biosynthesis of the capsular polysaccharide colonic acid) in hypervirulent strains and have

demonstrated that inoculation of this mutant into antibiotic-free mice lead to lower fecal

shedding[56] and organ burdens[66]. Thus, further studies are needed to determine the effects of

different types of capsule knockout on colonization.

The polysaccharide content of capsules can differ between strains. While capsular

polysaccharide typing (K typing) is a widely used method for categorizing capsule types, the

technical difficulty of doing so for large collections has led to the adoption of a sequence-based

methods. A first method relied on comparing sequences of the gene *wzi*[80]. However, *wzi*, which

is responsible for surface attachment of capsule polysaccharides[81], is not essential for capsule

production and is not present in all capsule loci[82]. As such, the software "Kaptive" was

developed to take into account the full *cps* locus to categorize *K. pneumoniae* capsule types, and

the newest version "Kaptive 2.0" can identify and assign 150 codified K types from whole

genome sequencing inputs[82,83].

LPS

Lipopolysaccharide (LPS), a component of the outer membrane, is another surface

virulence factor for *K. pneumoniae*, as it is for other Gram-negative bacteria. Full-length O

antigen (smooth LPS) contributes to serum resistance and virulence in a mouse model of

pyogenic liver abscess[84]. In colonization, enzymes involved in both core antigen (*waa*) and O-antigen biosynthesis (*wbb*) were identified as factors important for persistence in the mouse GI tract 3 days after inoculation, but the mechanism by which they act in this process has not been elucidated.[58]

Fimbriae and biofilm formation

      K. pneumoniae produce both Type I and Type III fimbriae, adhesive structures on the surface of the cell[85,86]. Type I fimbriae have been implicated in the pathogenesis of UTI but are not expressed in the gut[87], and Type III fimbriae can facilitate binding to plastic surfaces, suggesting a role in colonizing devices such as catheters or endotracheal tubes[88].

      In gut colonization, fimbriae seem to be dispensable. A targeted deletion of *fimH*, responsible for adhesion of the fimbrial tip to host cells, did not have a colonization defect, and a deletion mutant of *fimD*, necessary for assembly of fimbriae, did not have a significant defect as well[56,59].

Siderophores

      Siderophores are iron-binding small molecules important for bacterial pathogenesis[89]. While all strains produce the siderophore enterobactin to obtain iron in limiting conditions, host defenses include the production of lipocalin-2 to bind and render it unusable. In addition, enterobactin triggers the immune system and production of inflammatory molecules. To evade these defenses, some strains produce alternative siderophores such as yersiniabactin or salmochelin, which cannot be bound by lipocalin-2[90]. The presence of yersiniabactin is overrepresented in respiratory clinical isolates, indicating this siderophore may increase

virulence in the lung[89]. In contrast, salmochelin seems to be enriched in nasopharyngeal colonization isolates[91]. However, the role of siderophores in gut colonization have yet to be investigated.

Outer Membrane Protein A (ompA)

Outer membrane protein A (ompA) is a membrane-spanning protein which has been implicated in pathogenesis of other Gram-negative bacteria[92]. Deletion of *ompA* attenuates virulence in the lung and increases the production of the proinflammatory cytokine, IL-8[93]. The deletion of the gene encoding for the associated peptidoglycan-associated lipoprotein (Pal) decreases fitness in murine intraperitoneal infection[84]. This effect may be due to decreased protection against neutrophils. In a screen for GI colonization factors in mice, an *ompA* mutant was found to have a colonization defect at Day 3[58], but the mechanism behind this phenotype has not been explored.

Efflux pumps

Efflux pumps are important during pathogenesis to export toxic compounds and antibiotics. One such efflux pump is AcrAB-TolC which is known to efflux lipophilic molecules, antibiotics, and bile acids[94]. The efflux pump component AcrB contributes to virulence in pneumonia, and deletion leads to a decreased bacterial burden in the lungs[94]. This may be mediated by its role in resistance to antimicrobial peptides, and an AcrB mutant is more susceptible to BAL fluid.

EefABC is a separate tripartite efflux pump characterized in *Enterobacter aerogenes* for its ability to increase resistance to several antibiotics[95]. This efflux pump was identified as a

colonization factor for *K. pneumoniae*[96]. In addition to displaying a colonization defect *in vivo*, an *eefA* deletion mutant had decreased resistance to hydrochloric acid. This suggests that survival in the acidic environment of the stomach may be a key step in successful colonization.

Type VI Secretion System

Type VI Secretion Systems consist of a needle-like system for injection of toxins into other cells[97]. *K. pneumoniae* possesses a Type VI Secretion System (T6SS), which has been shown to aid in inter-species competition as well as modulate host responses to infection[66,98]. Intriguingly, the T6SS has also been implicated in survival of oxidant stress from host inflammation, but the effectors responsible have not yet been identified[99]. The T6SS also appears to be conserved in many pulmonary isolates as well as being overrepresented in pyogenic liver abscess isolates versus colonization isolates[99,100]. Furthermore, a T6SS mutant was less virulent in liver abscess induced in mice[101].

Studies in hypervirulent strains found that deletion of integral genes in the T6SS apparatus resulted in colonization defects[66,100]. Another using a classical strain demonstrated that T6SS was necessary for long-term colonization[102].

Metabolic Genes

A few studies have identified colonization factors through transposon mutant screens. Two studies made use of signature-tagged mutagenesis to identify these factors which included metabolic genes for biosynthesis of phospholipid and fatty acids as well as a urease[58,103]. One study has confirmed the importance of the fucose operon in colonization by the lab strain KPPR1[104], and another has found that xylose utilization may also be important for colonization[57].

Transcriptional Regulators

      The role of the transcriptional regulator OxyR in colonization was investigated due to its importance in oxidative stress resistance in a similar species, *E. coli*[105]. An *oxyR* deletion mutant could not establish colonization in the GI tract and also demonstrated decreased resistance to hydrogen peroxide, bile, and inorganic acid as well as decreased biofilm formation. Other transcriptional regulators such as *ntrC* (nitrogen metabolism) and *gcvR* (glycine metabolism regulator) have also been identified in screens for colonization factors[103].

1.3 Approaches for identifying colonization factors

Investigating specific targets vs. agnostic methods

Agnostic methods such as screens with large pools of transposon insertion mutants have been popular for many recent studies to identify colonization factors[58,59,103]. These methodologies allow for a high-throughput method to detect virulence factors across the entire genome, making them a powerful hypothesis-generating tool. However, these types of screens are not without limitations. Secreted factors can be missed, as a mutant that cannot produce them may still benefit from the activity of intact mutants around them. This is termed trans-complementation. Another limitation is technical: the analysis of such large datasets can be difficult, and arbitrary cut-off points for significance may cause both false positives and negatives.

As such, targeting known secreted virulence factors even when they do not appear in the results of screens can be a valuable method complementary to large screens for identifying colonization factors. For instance, neither components nor effectors of the Type VI secretion system appeared in three transposon insertion mutant screens for GI colonization factors[58,59,103]. However, several groups have demonstrated that Type VI secretion is important in colonization by both classical and hypervirulent strains[66].

Signature-tagged mutagenesis

A key methodology used in earlier screens for *K. pneumoniae* colonization factors was signature-tagged mutagenesis (STM)[106]. This method employs tags of unique 40 base pair sequences that are attached to transposons and inserted into the genome. Mutants are pooled and screened in the desired experimental conditions (e.g., GI colonization of mice) before being

recovered for genomic DNA extraction. The input and output pools are then hybridized with complementary probes, and the sequences which hybridize in the input pool but not the output pool are considered insertions that cause attenuation in the screening condition. Finally, sequencing of the flanking regions in the mutants with those tags reveals the genes in which insertion disrupts fitness. One drawback of this method is that limitations of prior sequencing modalities meant that some insertion sites could not be identified and the methodology itself is quite laborious[58]. However, this problem has largely been solved by newer methods described below.

Transposon-insertion sequencing

Several methods have been developed to take advantage of the massively parallel screening that could be achieved by next generation sequencing. In most methods, a transposon with a drug-resistance marker is used to insert into genes across the genome, saturating for insertions in all non-essential genes. These subsequent transposon insertion libraries are subjected to the desired experimental conditions (e.g., infection, colonization, antibiotic pressure, etc.), and the resulting surviving mutants are sequenced. Where these methods differ is in how the insertion sites in these mutants are recognized and sequenced. The most well-known methods are as follows: transposon sequencing (TnSeq), Insertion Sequencing (INSeq), high-throughput insertion tracking by deep sequencing (HITS), and transposon-directed insertion site sequencing (TraDIS).

TnSeq and INSeq make use of restriction enzyme recognition sites in the transposon for enzymes which cut several nucleotides away from the transposon. For example, MmeI, the enzyme used in INSeq cuts 20 bp away from its recognition site[107]. Adapters and barcodes for

sequencing are ligated, and the desired fully library prepped product is isolated by gel purification.

In contrast, HITS (high-throughput insertion tracking by deep sequencing) and TraDIS (transposon-directed insertion sequencing) do not use restriction enzymes in preparation of insertion sites for sequencing. Instead, they utilize random DNA shearing and end repair with generation of a polyA tail and ligation of adapters[108].

More modifications and innovations in transposon sequencing have continued to be made. One drawback of INSeq is the technical difficulty involved in the protocol. Yields can be low due to difficulty handling the product associated with streptavidin beads throughout the process and loss of product during a gel extraction at the end. A recent protocol by Kazi et al (2020) which is similar to that found in Stacy et al (2016) is significantly less time-intensive as well as less technically challenging. This method, like HITS and TraDIS, begins with random shearing of DNA followed by end repair—this time with the addition of a poly-C tail. Then, a biotinylated primer annealing to the transposon and one to the poly-C tail are used to amplify the insertion-adjacent DNA. Pulldown of product containing both the transposon and a poly-C tail is then achieved with streptavidin beads, and a second PCR is performed to further amplify these fragments and add barcodes to each library. A final purification using magnetic beads (instead of a gel purification as in INSeq) is used to clean up the product for sequencing.

Each of the protocols and approaches to insertion sequencing described above have been used successfully in screens to identify and validate factors responsible for different aspects of pathogenesis. However, the number of validated genes in each study is usually a very small portion of those detected in the screen. As such, what remains to be seen is the degree of

concordance between results from these different methods of insertion sequencing and more in depth investigations into the technical limitations of each method.

1.4 Genomic diversity of *K. pneumoniae*

One challenge in studying the pathogenesis of *K. pneumoniae* is the vast genomic diversity of the species. In population genomics, the genes common to most strains in a species is termed the "core genome." Different analyses place the threshold for a gene being "core" at presence in anywhere from 90% to 100% of all genomes available for a species[109,110]. When performing this type of analysis, we must also consider how similar genes must be to be considered the same gene. One can consider either differences in amino acids (cut off at >30% or >10% divergence) or sequence similarity (>85%)[109,110]. Any genes that are not core are then designated "accessory genes." Together, the core and accessory genes make up the "pangenome" of a species[111].

To quantify these for *K. pneumoniae* as a species, one study analyzed the sequences of 283 *K. pneumoniae* strains from sources distributed across the globe[109]. As the addition of each additional strain to the analysis resulted in the addition of new genes not represented in all others previously analyzed, the authors concluded that *K. pneumoniae* has an "open pangenome." This term refers to a pangenome where an extremely large number of strains would need to be sequenced to capture the full pangenome[111]. Bacterial species with open genomes tend to be those which live in diverse environments and which have methods for picking up new genetic material through horizontal transfer[112].

Additionally, while the typical *K. pneumoniae* genome contained about 5,000 to 6,000 genes, only about 1700 of these genes were shared by all other strains[109]. As such, a large

proportion of each individual strain's genome is composed of accessory genes. These accessory genes belong to a variety of different groups: from carbohydrate metabolism and other metabolic factors to resistance to antibiotics and heavy metals[109]. However, a third of the genes in the species do not have known functions, possibly indicating *K. pneumoniae* may have an even larger repertoire of adaptations to different environments than currently described. This high diversity across different strains of *K. pneumoniae* is also highlighted by their analysis demonstrating that sequencing of hundreds of additional strains from diverse sources is needed to encompass the full diversity of the species[109].

As of June 2023, there are 49,120 deposited *K. pneumoniae* genomes hosted by the National Center for Biotechnology Information (NCBI). More groups are turning to whole genome sequencing to analyze their strain collections, detect outbreaks, and identify virulence factors[113]. To aid in this, the software "Kleborate" uses genome assemblies to predict several characteristics: sequence type, capsule type, genes encoding for siderophores and other markers of hypervirulence, and genes encoding for antimicrobial resistance[23]. Recent genomics studies have made use of Kleborate for surveillance of high risk clones (by MLST) and virulence determinants[113], and these large-scale analyses have continued to emphasize the diversity of *K. pneumoniae* as a species.

1.5 Introduction to this work

To summarize the previous sections, *K. pneumoniae* is a major cause of highly drug-resistant nosocomial infections, and enteric colonization predisposes patients to subsequent infection and possibly transmission to others. As treatment options are becoming limited due to antimicrobial resistance, therapeutics designed to eradicate colonization may the key to preventing difficult-to-treat infections. However, the mechanisms behind GI colonization are not well-understood, and the high genomic diversity of the species begs the question of whether enteric colonization mechanisms differs from strain to strain. Thus, to maximize the potential utility of any anti-colonization therapeutics, we aimed to identify a core colonization program between *K. pneumoniae* strains.

In this study, we developed a mouse model of GI colonization with *K. pneumoniae* with high levels of fecal shedding in both male and female mice for at least 60 days post-inoculation. With this model, we demonstrated that different strains of *K. pneumoniae* have different competitive colonization abilities. We also created transposon mutant libraries in three strains of *K. pneumoniae* with different epidemicity and antibiotic resistance and identified and validated both shared and strain-specific genes used to achieve asymptomatic carriage in the GI tract.

2. Results

2.1 *K. pneumoniae* GI colonization in mice in the presence of an intact microbiome

Initially, we studied GI colonization in the absence of any perturbations to the microbiome with the goal of mimicking the asymptomatic colonization of healthy populations such as healthcare workers who are exposed to *K. pneumoniae* and who develop carriage but not infection. We investigated both GI colonization by hypervirulent and classical strains of *K. pneumoniae* in mice.

Gavage of hypervirulent strains leads to unpredictable fecal shedding and mortality

To investigate the carriage of hypervirulent strains, we administered the strain hvKP2 to mice by orogastric gavage. We quantified asymptomatic carriage by fecal burden, or colony forming units (CFU) recoverable from fecal pellets. *K. pneumoniae* was selectively cultured from feces by plating fecal homogenates on lysogeny broth (LB) agar supplemented with carbenicillin, an antibiotic to which all *K. pneumoniae* are resistant due to a chromosomally-encoded beta-lactamase (SHV-1)[109]. No dose-dependent fecal shedding was observed (**Figure 2**). While one to two mice inoculated with $10^{3.5}$ CFU had recoverable fecal shedding, there was only one instance at which any colonies were detected in the 10^4 CFU inoculation group.

Furthermore, there was at least one mortality in most of the groups (**Figure 2**). Like the fecal shedding observed, there was no clear dose-dependent effect on mortality. While one mouse died at Day 2 in the 10^4 CFU group, no mice died in the 10^5 CFU group. As we were attempting to mimic asymptomatic carriage rather than disease (or mortality), we moved from hypervirulent strains to classical strains.
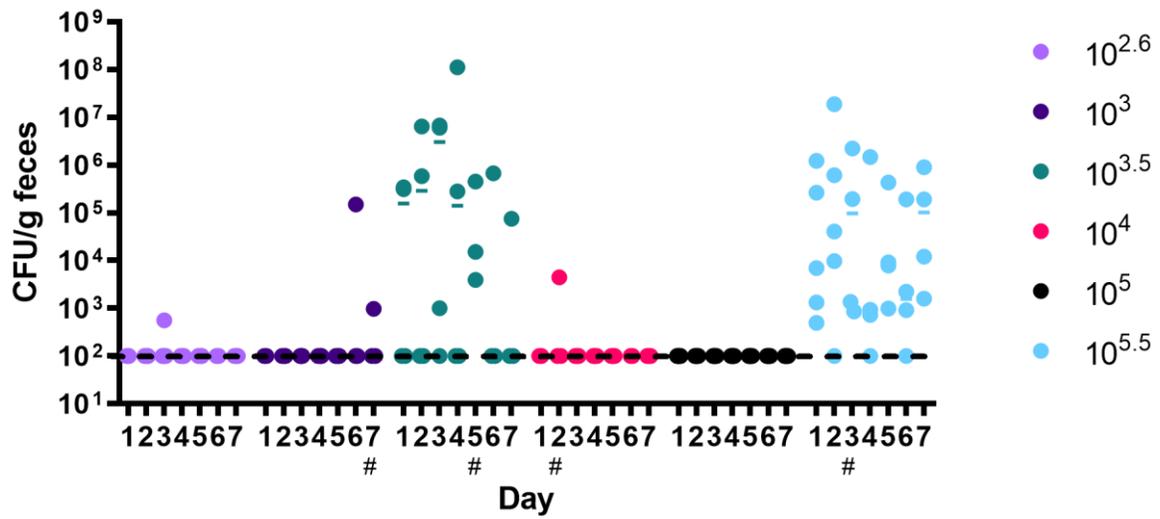
Figure 2 Fecal shedding of hypervirulent *K. pneumoniae*

Mice were inoculated by oral gavage with a range of doses of the hypervirulent strain hvKP2. Fecal pellets were collected and CFU enumerated. $n \geq 4$ for each dose. # indicates a single mortality at the indicated day. Dashed line indicates limit of detection.

Classical strains of *K. pneumoniae* exhibit variable colonization in the presence of an intact microbiome

We selected a range of classical strains to determine whether colonization with strains of high-risk clone lineages or those with virulence factors would result in greater fecal shedding than with non-epidemic strains. Each strain contained yersiniabactin—a siderophore associated with pneumonia isolates[89].

Strain	ST	Yersiniabactin	Colibactin	Aerobactin
CRE-001	ST258[#]	X	X	
CRE-098	ST13	X	X	
CRE-233	ST231	X		X
S007	ST111	X		

Table 1 Features of classical strains used for dose-response pilots of GI colonization
MLST and presence of genes encoding for yersiniabactin, colibactin, and aerobactin were detected from whole genome sequences of each strain by the software Kleborate[23]. High-risk sequence types are denoted with a #.

Each strain was inoculated into mice at two doses: 10^6 and 10^8 CFU. There were no clear dose-dependent relationships with fecal burden of *K. pneumoniae*. For CRE-001, there appeared to be greater fecal burden with the higher dose, but for CRE-233, the opposite was true (**Figure 3**). However, we completed subsequent experiments with the higher dose of 10^8 CFU, as the mice did not exhibit signs of sickness (ruffling, hunching, tachypnea).
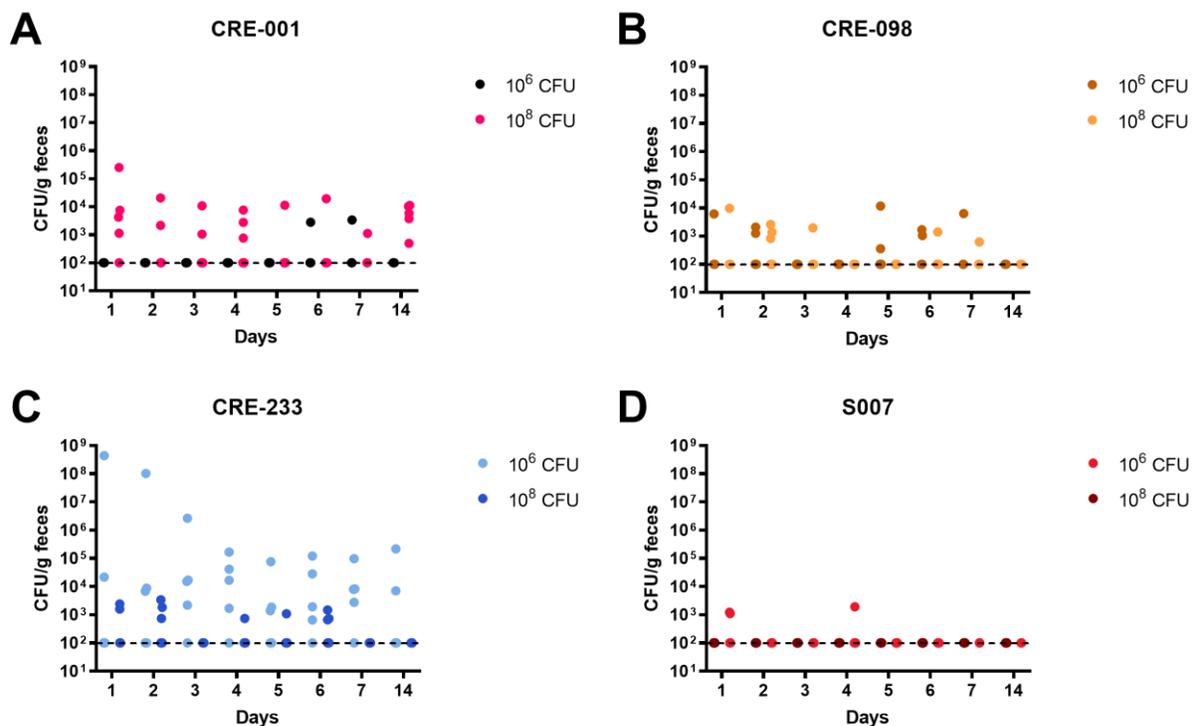
Figure 3 Effect of dose on fecal shedding of classical strains of *K. pneumoniae*

For four different classical strains of *K. pneumoniae*, two different doses (10^6 CFU or 10^8 CFU) were administered by orogastric gavage to separate groups of mice. Fecal samples were collected and CFU were enumerated. $n = 5$ per dose per strain. There was no statistical significance between fecal shedding for the two doses at any timepoint by multiple unpaired *t*-tests with correction for multiple comparisons by the Holm-Sidak method.

To determine whether strains of the same sequence type might behave similarly and thus make our results generalizable across a larger amount of isolates, we inoculated mice with four strains of the ST16 sequence type: CRE-015, CRE-058, CRE-177, and CRE-255 (**Table 2**).

Strain	ST	Capsule type	Yersiniabactin	Colibactin	Aerobactin
CRE-015	ST16	KL15			X
CRE-058	ST16	KL15			X
CRE-177	ST16	KL149			
CRE-255	ST16	-	X	X	X

Table 2 Characteristics of ST16 strains used for GI colonization

MLST, capsule type, and presence of genes encoding for yersiniabactin, colibactin, and aerobactin were detected from whole genome sequences of each strain by the software Kleborate[23].

We selected two strains (CRE-015 and CRE-058) which had similar virulence determinants (same capsule type and siderophore-encoding genes). However, only one of three mice inoculated with CRE-015 exhibited detectable fecal shedding across all days whereas two of three inoculated with CRE-058 had detectable *K. pneumoniae* (**Figure 4A**). Additionally, fecal shedding was not stable across different days. For CRE-015, shedding varied across three logs of magnitude between different days. There was also evidence of colonization that varied from detectable to undetectable across different days. For instance, one mouse inoculated with CRE-015, shed 10^5 CFU/g feces at Day 5 but did not have detectable carriage for any of the other timepoints. Thus, our limit of detection and the variability of shedding across different days made it difficult to draw conclusions about gastrointestinal carriage between different strains.
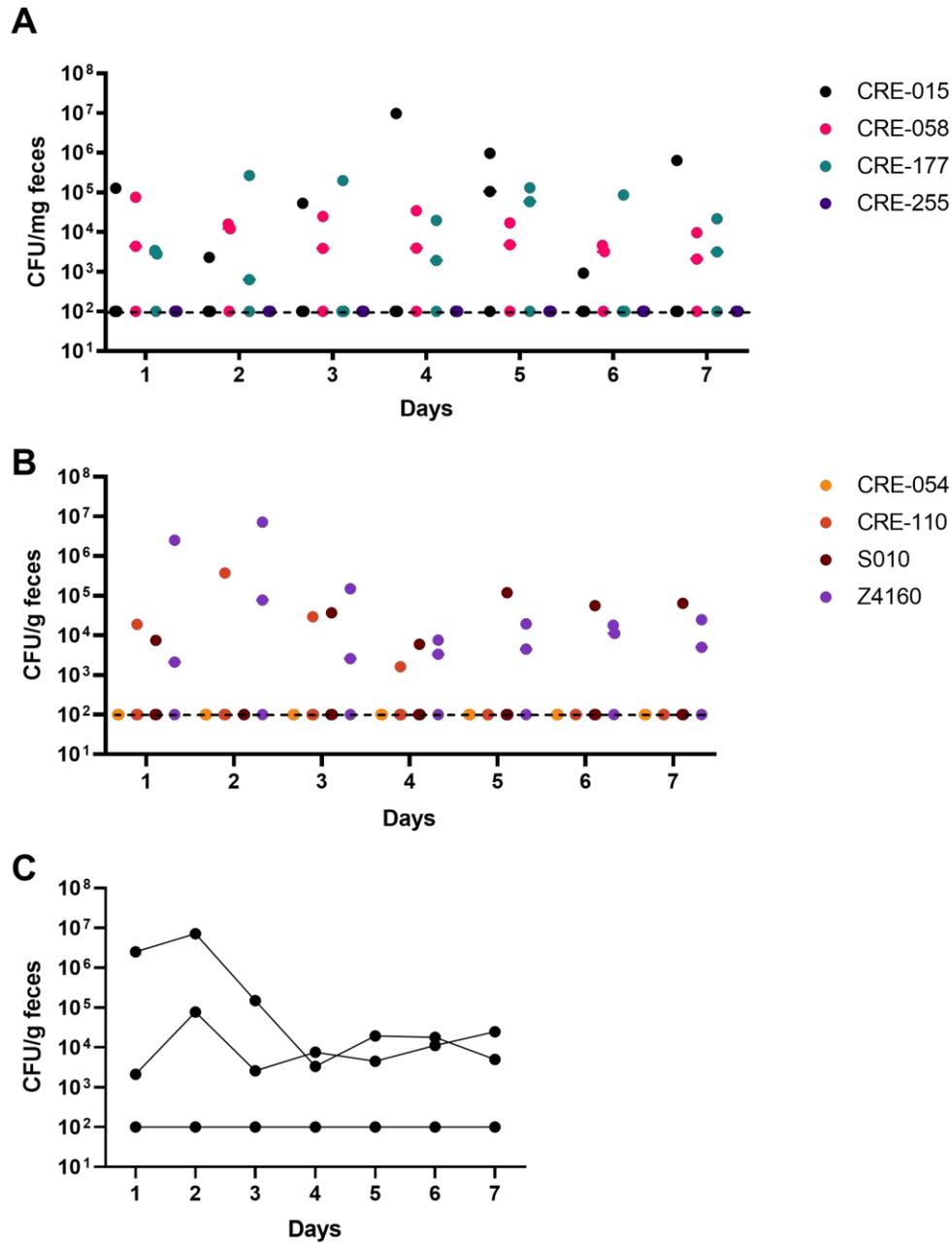
Figure 4 GI colonization of classical strain of *K. pneumoniae*

Mice were inoculated with 10^8 CFU by orogastric gavage of (A) four ST16 strains or (B) four strains with pairwise differences in certain features (see **Table 2**). Feces were collected and CFU enumerated. The fecal burden for each mouse inoculated with the strain Z4160 are shown in (C) $n = 3$ for each strain.

Similarly, while the capsule-null mutant, CRE-255 did not exhibit any fecal shedding (**Figure 4A**), we were unable to determine whether this was a result of the lack of capsule or was a product of the low number of biological replicates and variability in proportion of mice in which colonization could be detected.

Additional experiments with four more classical strains (CRE-054, CRE-110, S004, and Z4160, **Table 3**) demonstrated similarly variable success in establishing detectable levels of fecal shedding (**Figure 4B**). This variability in successful engraftment is highlighted by **Figure 4C** where one of three mice never exhibited detectable shedding of the strain Z4160 even though the other two mice had around 10^4 CFU/g of fecal burden at Day 7.

Strain	ST	Capsule type	Yersiniabactin	Colibactin	Aerobactin
CRE-054	ST258	KL106	X	X	
CRE-110	ST258	KL107		X	
S004	ST16	Unassigned	X		X
Z4160	ST45	KL24	X		

Table 3 Features of additional classical strains used for GI colonization pilots
MLST, capsule type, and presence of genes encoding for yersiniabactin, colibactin, and aerobactin were detected from whole genome sequences of each strain by the software Kleborate[23].

While variability may have been ameliorated by inclusion of additional biological controls, we determined that the low magnitude of fecal burden as well as high variability provided too little power to distinguish differences between strains. As such, we piloted the following changes to our colonization protocol to determine whether they would produce more stable levels of detectable colonization.

Administration of bicarbonate does not increase GI colonization of *K. pneumoniae*

The administration of sodium bicarbonate is known to support colonization of the gut by other pathogens such as *Vibrio cholerae*[114]. As such, we tested whether bicarbonate treatment before gavage with *K. pneumoniae* would increase fecal shedding. However, fecal burden one day post-gavage was not significantly different between groups gavaged with *K. pneumoniae* with or without sodium bicarbonate pre-treatment (**Figure 5**).
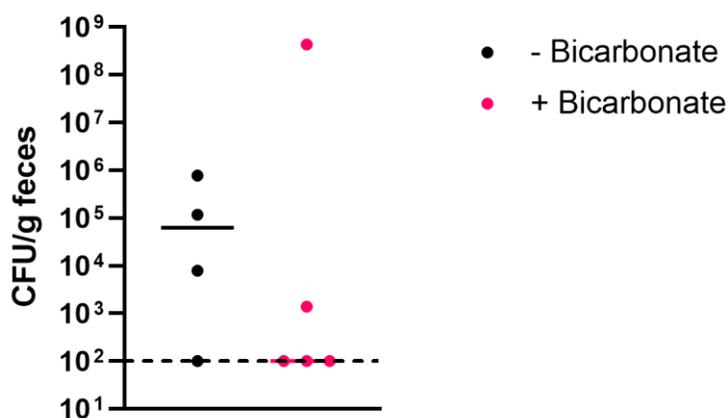
Figure 5 Effect of bicarbonate administration on fecal shedding

A solution of 8.5% sodium bicarbonate was administered by orogastric gavage prior to gavage with 10^8 CFU of the classical *K. pneumoniae* strain CRE-233. Feces were collected 1-day post-gavage and CFU enumerated. $n \geq 4$ for each group. No significant difference between groups by Student's *t*-test.

Sucrose feeding is not superior to orograstric gavage for inducing GI colonization of classical *K. pneumoniae* strains

Another method for inoculating mice with *K. pneumoniae* is sucrose feeding, and one group achieved more stable GI colonization with this method versus gavage[56]. As such, we compared both methods with one classical strain of *K. pneumoniae*: KPN46. However, there was no significant difference between levels of shedding between either method, and some mice from the sucrose feeding group still exhibited undetectable levels of shedding at two timepoints (Day 2 and 4, **Figure 6**). Additionally, shedding in the sucrose feeding group still spanned about two logs of magnitude at each timepoint, raising concerns that our power would still be too low with this method to detect differences between strains.

Other technical considerations also prompted us to continue with gavage as the method of administration. While gavage could be completed in less than a minute per mouse, sucrose feeding required restraint for several minutes, increasing procedure time greatly. Thus, as sucrose feeding did not increase reliability of gavage and also increased technical difficulty, all mice in subsequent experiments were inoculated by gavage.
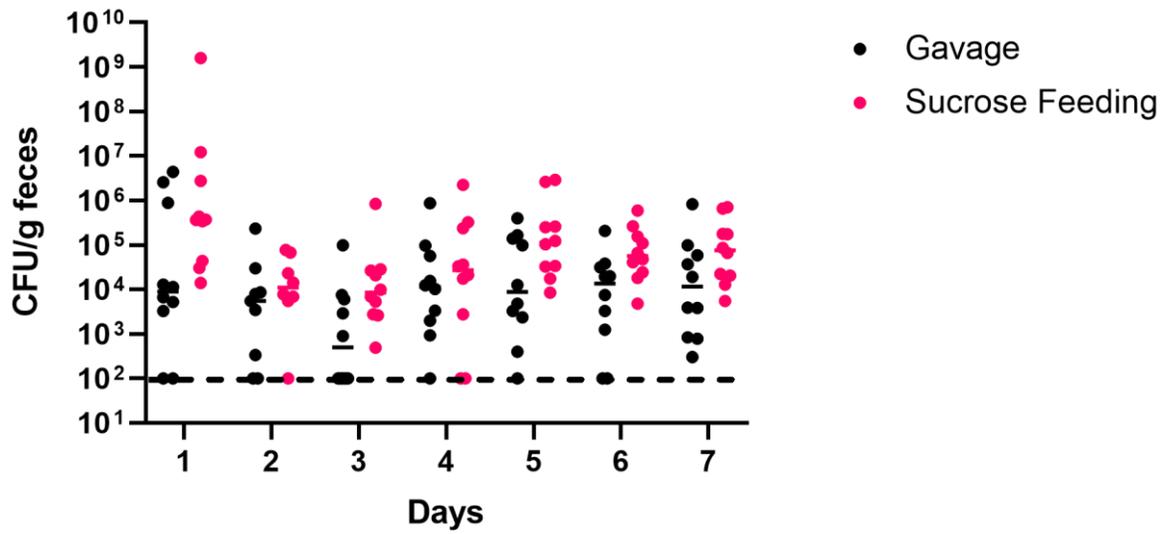
Figure 6 GI colonization following administration of *K. pneumoniae* via orogastric gavage or sucrose feeding

Mice were inoculated with 10^8 CFU of the strain KPN46 by gavage or sucrose feeding. $n = 10$ for each group. There was no significant difference between groups by two-way ANOVA.

2.2 Investigation of genetic determinants of GI colonization of mice with an intact microbiome

A siderophore-microcin does not play a major role in GI colonization in the presence of an intact microbiome

While the previous GI colonization pilot experiments indicated differences between strains may be difficult to quantify, we were also interested in whether the role of specific genetic factors could be investigated in GI colonization of mice. As the microbiome of the mice were intact in these colonization experiments, we hypothesized that antibacterial effectors that *K. pneumoniae* produces may aid it in combating other bacteria in the gut to establish colonization. One of these factors produced by some of our strains is microcin E492. Microcins are small peptides produced and secreted to kill other bacteria. Microcin E492 is a pore-forming peptide linked to the siderophore salmochelin. This siderophore is produced by glucosylation of enterobactin, and the genes in the microcin E492 locus encodes for the enzyme to perform this step as well as the linkage to and production of the rest of the toxin[115,116].

In prior population genomic analyses, microcin E492 was noted to be associated with hypervirulent strains[31]. As such, we screened strains in our collection for inhibitory activity against the *E. coli* strain TOP10 using overlay assays. The *K. pneumoniae* strains were spotted onto LB agar plates, and a soft 0.75% agar was spiked with TOP10 and overlayed on top. After incubation overnight at 37°C, zones of inhibition around the spots of *K. pneumoniae* could be observed. We found that hvKP2 and hvKP3 exhibited zones of inhibition whereas hvKP1, hvKP4, hvKP5, KPPR1, and NTUH-K2044 did not (**Figure 7**). Furthermore, a Type 6 Secretion System deletion mutant of hvKP2 (Δ*vgrG*) could still produce a zone of inhibition, indicating there was an additional antibacterial factor produced by hvKP2. Finally, the classical strains

CRE-001, CRE-098, CRE-233, S004, and S007 did not produce a zone of inhibition in similar assays.
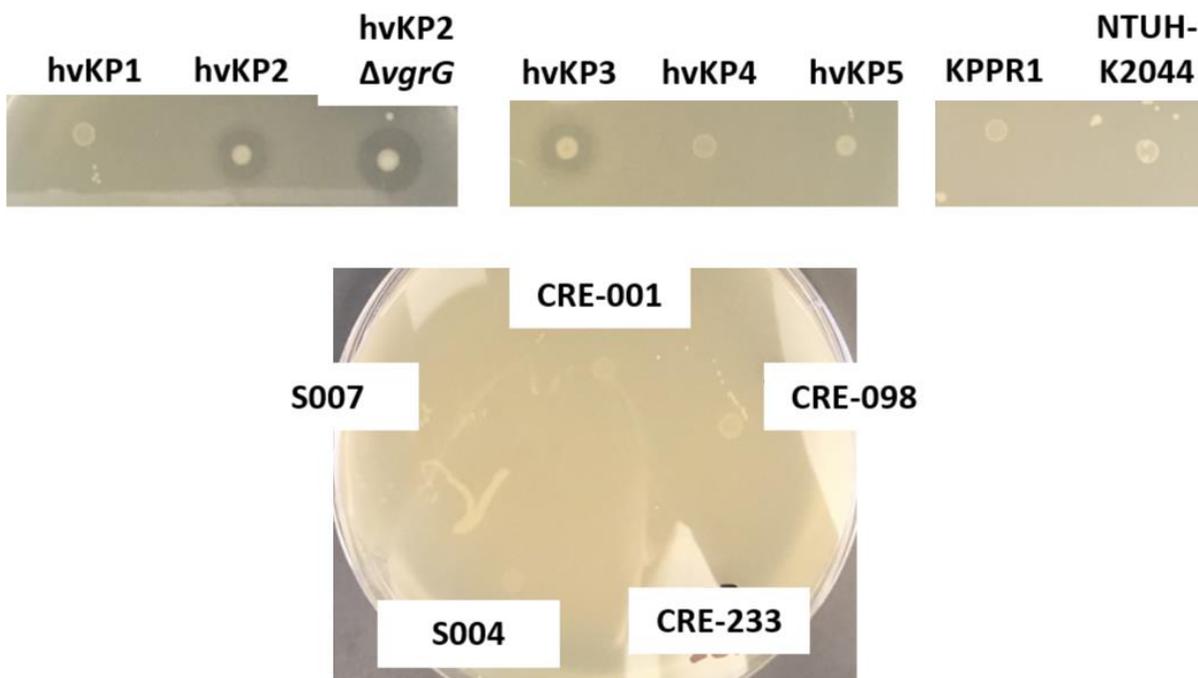
Figure 7 Growth inhibition of *E. coli* TOP10 by strains of *K. pneumoniae*
Strains of *K. pneumoniae* (labeled) were spotted onto LB agar plates, and 0.75% agar was spiked with *E. coli* TOP10 before being overlayed. Plates were incubated at 37ºC overnight. Images are representative of three biological replicates from different days.

Next, we used BLAST[117] to search for the previously described sequence of the microcin E492 biosynthetic cluster. We found that the genomes of hvKP2 and the other strains which produced zones of inhibition contained this cluster whereas most of the strains that could not inhibit the growth of TOP10 did not. However, we did find that hvKP5 did contain the cluster, but the sequence for its transporter contained a single-nucleotide polymorphism (SNP) leading to a premature stop codon. This indicated that hvKP5 might synthesize the microcin but not secrete

it, which would be supported by the overlay assay results. We then looked for the microcin E492

production genes in our *K. pneumoniae* clinical strain collections from Northwestern Memorial

Hospital (NMH, Chicago, USA) that we have whole genome sequenced. This included 182

carbapenem-resistant isolates and 20 non-carbapenem-resistant isolates, but we did not find any

with the microcin E492 gene cluster.

As we also had a collection of 141 non-sequenced bloodstream isolates of *K. pneumoniae*

from NMH, we performed a screen for microcins by arraying the isolates, stamping them on a

LB agar plate, and performing an overlay assay with 0.75% agar and TOP10. We detected a zone

of inhibition around one isolate, KPN46. Subsequent whole genome sequencing revealed that

this non-hypermucoviscous classical strain did indeed encode for microcin E492.

With both the classical (KPN46) and hypervirulent (hvKP2) strains, we created isogenic

deletion mutants of *mceAB* which encode for biosynthesis of the microcin. These deletion

mutants did not have the large zones of inhibition in overlay assays than the parent strains

(**Figure 8**). However, the zone of inhibition was restored when a plasmid with *mceAB* was used

to complement the KPN46Δ*mceAB* mutant, indicating the zone of inhibition was created by

expression of microcin E492.

Figure 8 Inhibition of *E. coli* by *K. pneumoniae* and microcin E492 mutants
Overlay assays were performed with wild-type hvKP2 and KPN46 as well as deletion mutants for microcin E492 (*mceAB*) for both, negative control hvKP5, and a plasmid-complemented strain KPN46Δ*mceAB* + pACYC184::*mceAB*. Spots of each of these strains were placed on LB agar, and 0.75% LB agar was spiked with *E. coli* TOP10 before being overlayed on the spots.

To test whether microcin E492 plays a role in GI colonization in the presence of an intact microbiome, we inoculated mice by orogastric gavage with the parent strains or deletion mutants separately (**Figure 9**). Again, inoculation with hvKP2 produced variable and non-dose-dependent shedding. When inoculated with 10^4 CFU, the hvKP2 parent strain exhibited greater colonization at a few timepoints (**Figure 9A**) whereas hvKP2Δ*mceAB* exhibited greater shedding at an inoculum of 10^5 CFU (**Figure 9C**). Additionally, mortality during the experiment (even in the hvKP2Δ*mceAB* group) created incomplete data sets and indicated that we were not achieving asymptomatic carriage.

Fecal shedding with KPN46 and KPN46Δ*mceAB* was higher (**Figure 8C**) following inoculation with 10^8 CFU of either strain, and no deaths occurred. However, there was no significant difference between colonization with the parent strain versus the microcin deletion mutant by two-way ANOVA.
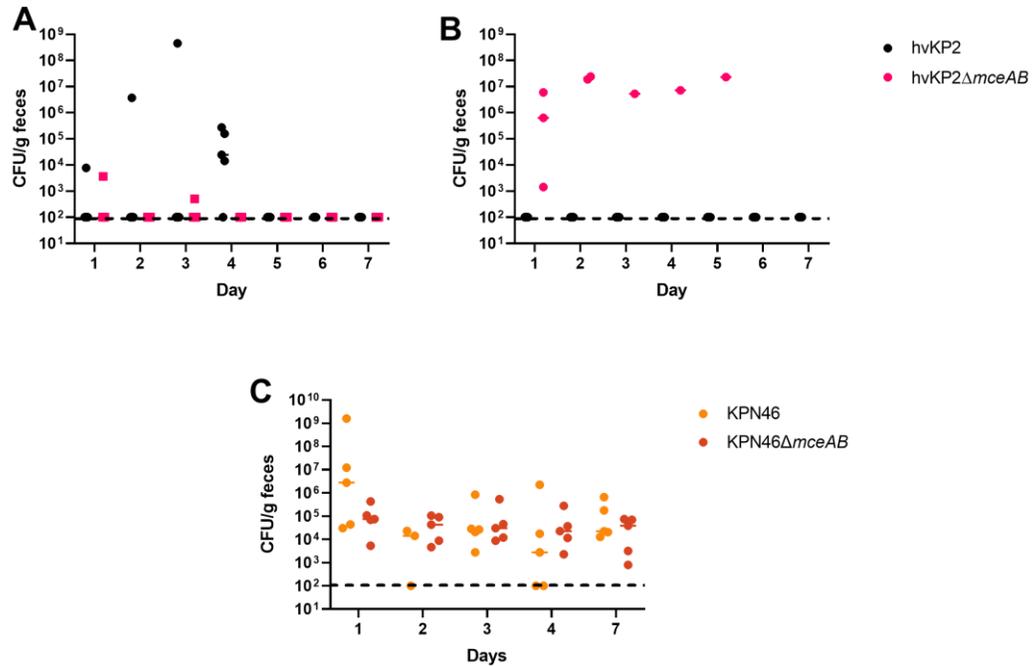
Figure 9 The role of microcin E492 in the colonization of the GI tract

Mice were inoculated with (A) 10^4 CFU or (B) 10^5 CFU of hvKP2 or hvKP2$\Delta mceAB$ (n = 3 for each group) or (C) 10^8 CFU KPN46 or KPN46 $\Delta mceAB$ (n = 5 for each group). Fecal pellets were collected and CFU enumerated. Dashed line indicates limit of detection.

Bottleneck detection in antibiotic-free GI colonization in mice

As one of the goals for colonization was to perform transposon mutant screens, we also performed the following studies to determine whether there were bottlenecks severe enough in our antibiotic-free mice to prevent the use of this type of screen. As one assumption of the screen is that mutants drop out of the output due to fitness defects, any other factors which cause drop out can confound results. If a severe bottleneck exists in the screen conditions, there would be stochastic elimination of mutants from the output pool.

To probe the bottleneck, we created a marked version of KPN46 with an apramycin-resistance cassette in the Tn7 site (KPN46 Tn7::Apr^R). We spiked this marked strain into an inoculum with the parent strain at a variety of ratios and determined whether we could still recover the marked strain. As our previous experiments above demonstrated that our detection of *K. pneumoniae* in the feces was not powerful enough to capture all colonization, we also determined whether bacterial burden in organs of the GI tract (small intestine, cecum, and colon) was high enough to support a screen at 24 hours post-gavage.

While the marked strain could be recovered from feces and organs when it made up a quarter of the initial inoculum, it could not be detected at any of the other ratios (**Figure 10**). As each transposon mutant makes up far less than a quarter of an input pool for a screen, we concluded that the bottleneck in antibiotic-free mice would be too severe to conduct a transposon mutant screen.
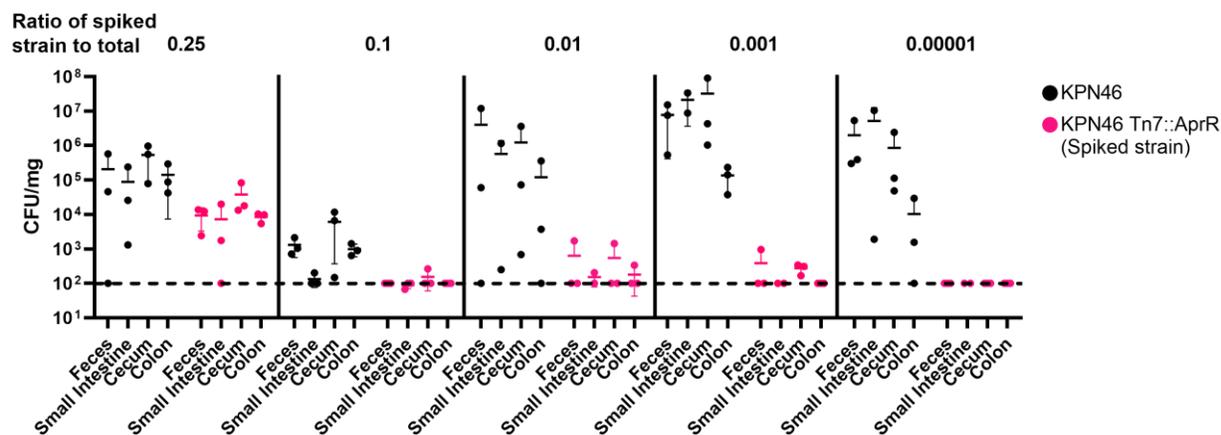
Figure 10 Recovery of a marked strain spiked into an inoculum for GI colonization

The marked strain KPN46 Tn7::Apr^R was spiked into an inoculum of 10^8 CFU of KPN46 and administered to mice by orogastric gavage. $n = 3$ for each group. At 24 hours post-gavage, feces and organs were collected. CFU were enumerated by plating on LB agar supplemented with carbenicillin (total *K. pneumoniae*) and apramycin (KPN46 Tn7::Apr^R). CFU of KPN46 were calculated by subtracting apramycin-resistant CFU from carbenicillin-resistant CFU.

2.3 Development of a clinically relevant mouse model of *K. pneumoniae* GI colonization

As attempts to establish GI colonization without antibiotics did not provide high levels of reliable fecal shedding, we then turned to administering an antibiotic regiment prior to oral gavage with *K. pneumoniae*. Since this would approximate hospitalized patients receiving antibiotics, we chose vancomycin, one of the most highly utilized antibiotics in the United States[118], administered through intraperitoneal injection to mimic the intravenous route of administration. Note that vancomycin lacks activity against gram-negative bacteria such as *K. pneumoniae*[119].

We conducted pilot experiments with two different doses of vancomycin: 20 mg/kg (which had been used in other studies)[120] and 350 mg/kg (a mouse equivalent of human dosing of 1 g/day)[121]. Mice received intraperitoneal injections of either dose for 3 days prior to gavage with 10^8 CFU of *K. pneumoniae* strain KPN46. While there were no significant differences between groups by multiple t-tests, fecal shedding of mice who received the higher dose (350 mg/kg) was much higher at Day 7 (10^8 CFU) than those who received the lower dose (20 mg/kg, 10^4 CFU) or PBS (10^4 CFU) (**Figure 11A**).

We next tested different durations of administration. Mice received intraperitoneal injections of vancomycin (350 mg/kg) for 3 or 5 days or PBS for 5 days (**Figure 11B**). While there were again no statistical differences between groups, the mice that received 5 days of vancomycin had a much higher fecal burden at Day 7 than the groups that received 3 days of vancomycin or PBS. As such, our optimized model consisted of 5 days of intraperitoneal injection of 350 mg/kg vancomycin followed by orogastric gavage with 10^8 CFU *K. pneumoniae*.
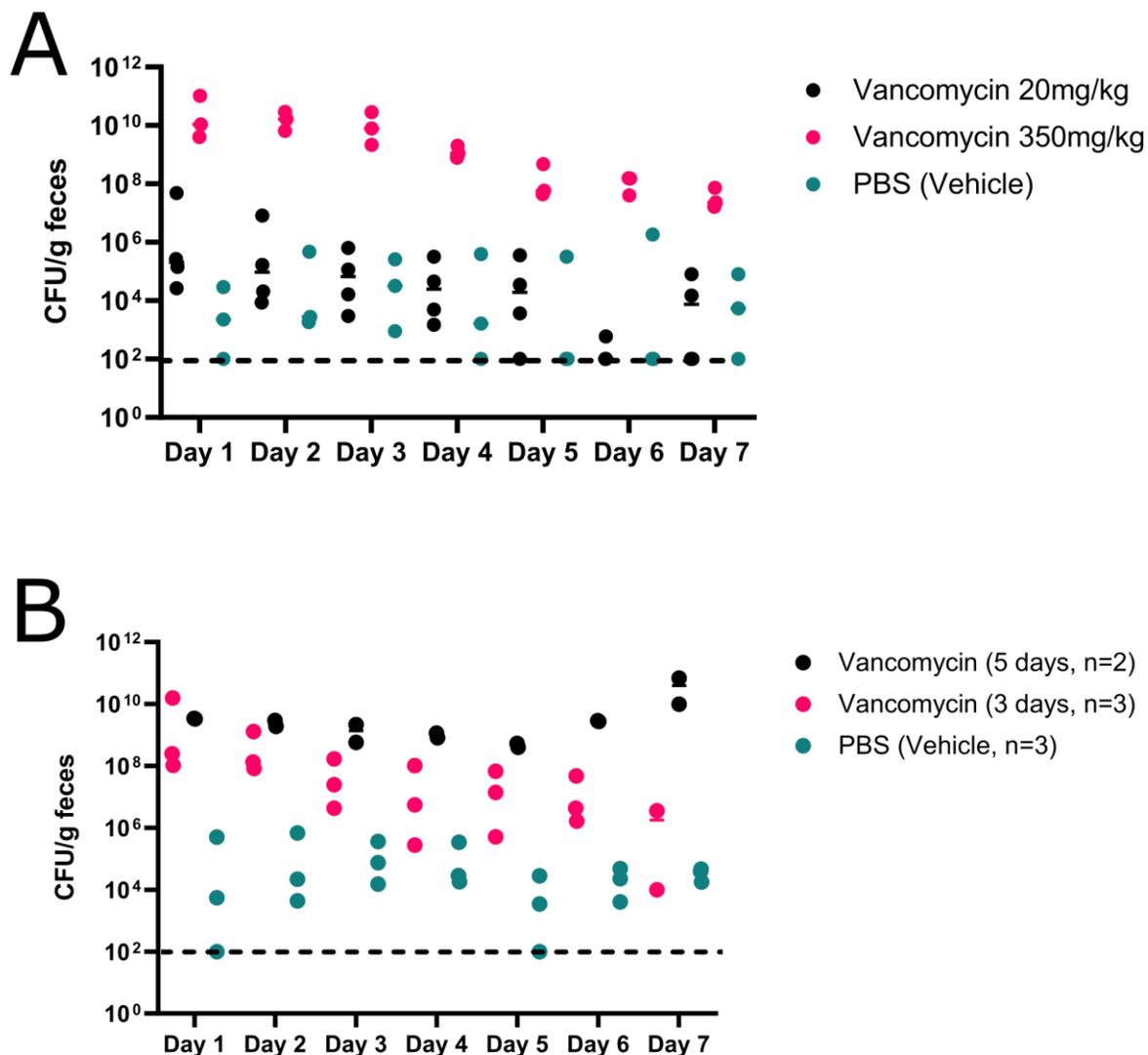
Figure 11 Optimization of vancomycin regimen for a model of GI colonization

(A) Mice were injected intraperitoneally with 20 mg/kg or 350 mg/kg vancomycin or PBS (vehicle) for three days prior to inoculation with *K. pneumoniae* strain KPN46. $n = 3$ for each group. (B) Mice received IP injections of 350 mg/kg vancomycin for three ($n = 3$) or five ($n = 2$) days or PBS for five days prior to inoculation with 10^8 CFU *K. pneumoniae* strain CRE-133. Fecal samples were collected and CFU were enumerated. Dashed line indicates limit of detection.

We confirmed the specificity of our method of detecting *K. pneumoniae* by selection with carbenicillin by culturing feces prior to inoculation with *K. pneumoniae*, which yielded no colonies (**Figure 12B**). Additionally, we verified that colonization was limited to the gut by quantifying organ burden in sterile sites (lung, liver, spleen). At Day 14, there was no recoverable *K. pneumoniae* at these sites even when there was 10^7 to 10^8 CFU/g in the feces (**Figure 12C**). Then, to test the duration of colonization (**Figure 12A**), we followed fecal burdens to Day 60 in both male and female mice. The mice shed 10^{10} CFU/g feces of *K. pneumoniae* in the first week followed by shedding of approximately 10^7 CFU/g for at least 60 days post-gavage in both male and female mice (**Figure 12D**). Furthermore, gavage with three different strains of *K. pneumoniae* (CRE-166, KPN46, and Z4160) led to high levels of fecal shedding up to 14 days post-gavage (**Figure 12E**). During these experiments, the mice did not exhibit signs of illness, suggesting that this method achieved colonization rather than systemic infection. Altogether, these results indicated that the vancomycin-treated mice can be used as a model for *K. pneumoniae* GI colonization.
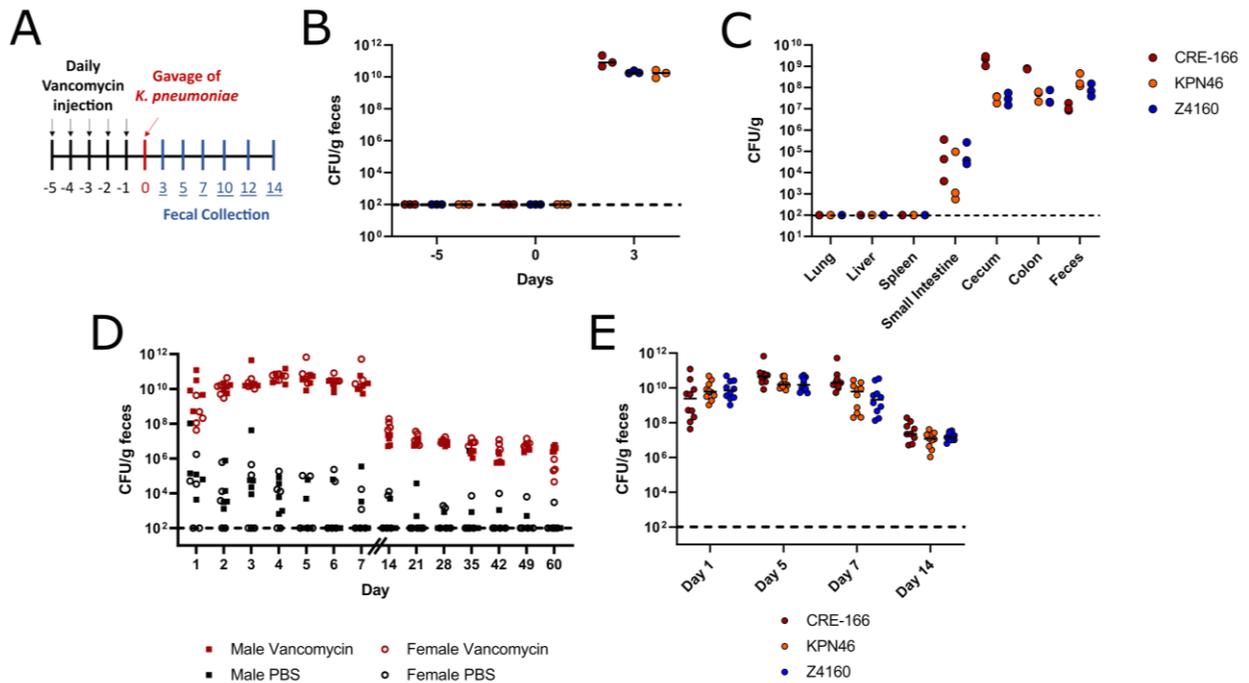
Figure 12 Longitudinal assessment of a mouse model of GI colonization with *K. pneumoniae* with antibiotic administration

(A) Schematic of *in vivo* model. Mice were administered 5 days of 350 mg/kg intraperitoneal vancomycin injections before orogastric gavage with 10^8 CFU *K. pneumoniae*. Fecal samples were collected after gavage and CFU enumerated. (B) Fecal burden of carbenicillin-resistant bacteria before (Days -5 and 0) and after (Day 5) inoculation with 10^8 CFU *K. pneumoniae* strains CRE-166, KPN46, and Z4160. $n = 3$ for each group. (C) Organ burden of CRE-166, KPN46, and Z4160 at Day 14. $n = 3$ for each strain. (D) Fecal burden of CRE-166 following gavage into male (square) or female (circle) mice with (red) or without (black) vancomycin treatment prior to gavage. $n = 5$ for each group. (E) Fecal burden of three strains of *K. pneumoniae* following gavage. $n = 10$ for each strain with inoculations performed on two separate days. Limit of detection was 10^2 CFU/g feces, denoted by a dotted line.

2.4 Colonization capacities of strains individually and in competition

Four strains of *K. pneumoniae* were selected to assess strain-dependent differences in GI colonization of mice. These clinical isolates from Northwestern Memorial Hospital had varying antibiotic resistance as well as epidemic status. However, when inoculated into antibiotic-treated mouse, each strain exhibited similar colonization patterns, with 10^{10} CFU/g feces recovered in the first week followed by 10^7 CFU/g feces in the second week (**Figure 13**). These results indicated that antibiotic-treated mice are able to support high levels of GI colonization by a range of classical *K. pneumoniae* strains.
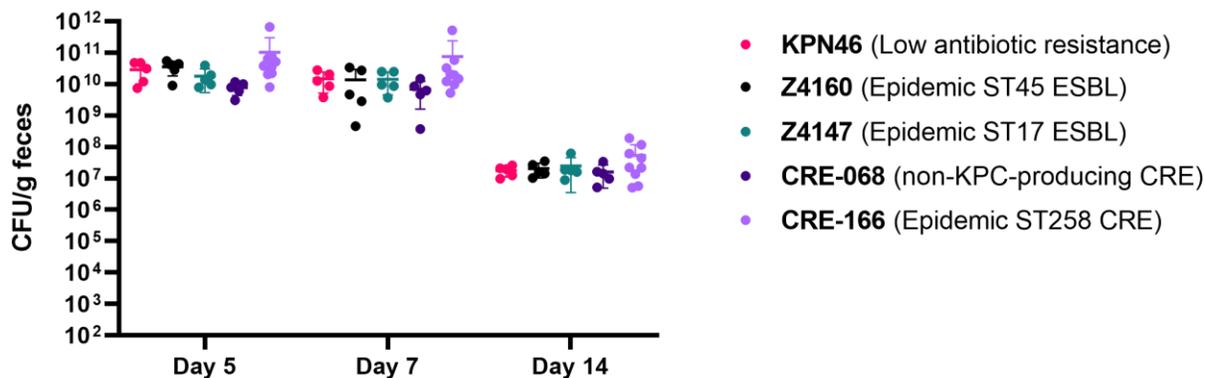
Figure 13 GI colonization by multiple strains of *K. pneumoniae*

Five classical strains of varying antibiotic resistance as well as global spread were inoculated by oral gavage (10^8 CFU). Feces were collected and CFU enumerated. $n \geq 5$ for each strain. There was no significant difference between each group by two-way ANOVA.

As we could not detect differences between the strains during individual inoculations, we turned to another method to detect more subtle differences in phenotypes: competition experiments. We performed *in vivo* competition experiments between pairs of three strains of *K. pneumoniae*: CRE-166, KPN46, and Z4160. In order to do so, we created two different marked strains for each parent strain by inserting an apramycin-resistance cassette (Apr^R) or hygromycin-resistance cassette (Hyg^R) into the Tn7 site. Then, 10^8 CFU of each strain was mixed in a pairwise manner and gavaged into mice, and each strain was enumerated by plating on apramycin or hygromycin. For instance, we combined CRE-166 Tn7::Apr^R and KPN46 Tn7::Hyg^R in an inoculum and were able to quantify them by plating on LB agar supplemented with apramycin or hygromycin, respectively. We calculated competitive indices by dividing CFU of one strain by the other (indicated in the y-axis title) and normalized to the ratio of CFU of each strain in the inoculum.

Each pair of strains exhibited a different pattern in competitive indices over two weeks (**Figure 14**). For CRE-166 and KPN46, KPN46 dominates in first week, but CRE-166 regains a competitive advantage over time, eventually coming to dominate the fecal burden at Day 14 (**Figure 14A**). However, when competed against Z4160, CRE-166 has a competitive defect in the first week that deteriorates even further in the second week (**Figure 14B**). In the final pair, neither Z4160 nor KPN46 showed a competitive advantage in the first week, but in the second week, Z4160 dominated the fecal burden (**Figure 14C**). As such, it appears that the ST258 epidemic strain (CRE-166) may do poorly against other epidemic strains (Z4160) but may colonize at higher levels long-term when competed against a non-epidemic strain (KPN46). In fact, KNP46 is outcompeted by both CRE-166 and Z4160 at Day 14.
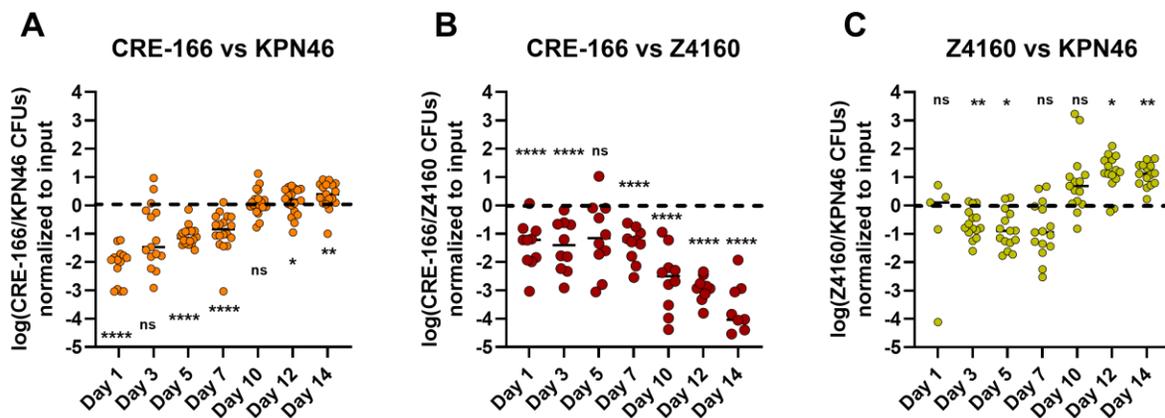
Figure 14 Competitive colonization between pairs of *K. pneumoniae* strains

After five days of intraperitoneal injections with vancomycin, mice were inoculated with pairs of strains (10^8 CFU per strain) (A) CRE-166 and KPN46 (n = 19) (B) CRE-166 and Z4160 (*n* = 10) and (C) Z4160 and KPN46 (*n* = 10). Inoculations were completed on at least two separate days. The first strain in each pair was marked with an apramycin-resistance cassette at the Tn7 site and the second strain was marked with a hygromycin-resistance cassette. Feces were collected and CFU were enumerated by plating on hygromycin and apramycin. Dashed line indicates equal CFU recovered of both strains. * indicates a significant difference from 0 by one-sample t-test with Dunn's correction; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

Next, we further investigated the change in strain domination over time during competition between CRE-166 and KPN46. The competitive index rose steadily over two weeks, and we hypothesized that this change in phenotype may have been due to the effects of vancomycin wearing off over that time period. To test this hypothesis, we injected mice already colonized with CRE-166 and KPN46 with vancomycin again for 5 days starting on Day 15.

During the injections, the upward trend of the competitive index began to reverse, and after this second course of vancomycin, the same initial dominance of KPN46 switching to dominance of CRE-166 over two weeks was observed (**Figure 15A**). Additionally, re-injection with vancomycin increased absolute CFU load back to levels similar to those observed after the initial vancomycin injection (**Figure 15B**).
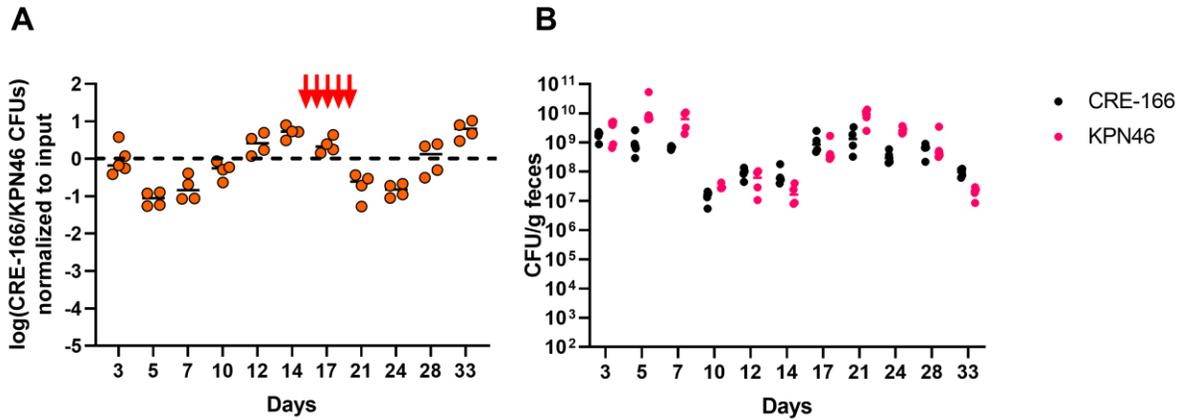
Figure 15 Competitive GI colonization between CRE-166 and KPN46 with re-injection of vancomycin after establishment

After 5 days of intraperitoneal injections with vancomycin, mice were inoculated with a mixture of 10^8 CFU each of CRE-166 Tn7::Apr^R and KPN46 Tn7::Hyg^R . Fecal pellets were collected and CFU enumerated by plating on selective media. Beginning on day 15 post-inoculation, mice received another 5 days of intraperitoneal injections with vancomycin (indicated by red arrows in panel A). (A) Competitive index over time. (B) CFU by strain over time. $n = 4$. Dashed line indicates equal CFU recovered of both strains.

To determine whether the patterns in competitive colonization could be generalized to other strains, we chose three additional strains for *in vivo* competition experiments (**Table 4**).

Strain	ST	High-risk clone?
CRE-166	ST258	Yes
KPN46	ST433	No
Z4160	ST45	Yes
KPN41	ST45	Yes
CRE-068	ST101	Yes
Z4147	ST17	Yes

Table 4 Characteristics of strains used in competitive GI colonization studies
MLST and capsule type were detected from whole genome sequences of each strain by the software Kleborate[23].

First, we tested whether our ST258 strain (CRE-166) would still be outcompeted by a strain of a high-risk sequence type (ST45) like Z4160 or if we would see strain dominance switch again from a low-resistance strain (KPN46) to CRE-166 in the second week. The strain for this comparison was KPN41, which belongs to a high-risk sequence type (ST45) but is antibiotic-susceptible strain. In competition with CRE-166, this pair was more similar to the phenotype from the CRE-166/Z4160 pairing, as KPN41 outcompeted CRE-166 throughout two weeks (**Figure 15A**).

Next, we asked whether another carbapenem-resistant strain (CRE-068) would also exhibit the same phenotype as the CRE-166/KPN46 pairing. However, KPN46 appeared to dominate throughout (**Figure 15B**).

Finally, we examined whether CRE-166 would still be outcompeted by a different ESBL-producing high-risk clone (Z4147, ST17). In this case, CRE-166 was outcompeted in the first week but appeared to recover to be equally as fit as Z4147 in the second week (**Figure 16C**).

In conclusion, strains of K. pneumoniae differ in their capacities to colonize the GI tract of mice, but fitness phenotypes are not tightly associated with other phenotypes such as antibiotic resistance or status as a high-risk clone. As such, we moved towards determining more specific genetic factors which different strains rely on to achieve GI colonization.
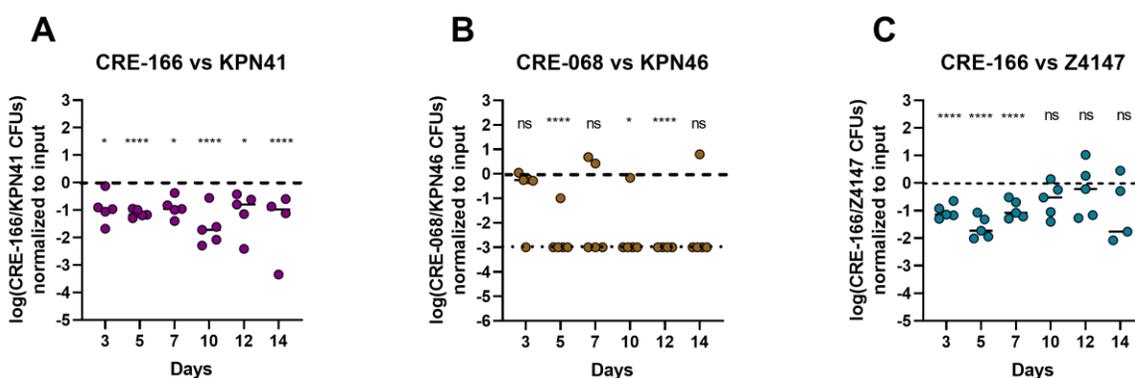
Figure 16 Competitive colonization between other strains of *K. pneumoniae*

After five days of intraperitoneal injections with vancomycin, mice were inoculated with 10^8 CFU of each strain in each pair (A) CRE-166 Tn7::Apr^R and KPN41 (B) CRE-068 and KPN46 Tn7::Apr^R and (C) CRE-166 Tn7::Apr^R and Z4147. $n = 5$ for each group. Feces were collected and CFU were enumerated by plating on carbenicillin and apramycin. CFU of the unmarked strain was calculated by subtracting apramycin-resistant CFU from carbenicillin-resistant CFU. Dashed line indicates equal CFU recovered of both strains. * indicates significant difference from 0 by one-sample *t*-test with Dunn's correction; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

2.4 Selection of three representative clinical *K. pneumoniae* strains

For our transposon mutant screen, we selected three clinical isolates of *K. pneumoniae* cultured at Northwestern Memorial Hospital that were representative of strains with varying levels of epidemic spread and antibiotic resistance. First, we chose CRE-166, a carbapenem-resistant strain of the ST258 high-risk clone that contained the *bla*$_{KPC}$ gene. Second, we selected an ESBL producer, Z4160, with both a widespread ESBL gene (*bla*$_{CTX-M-15}$) and an epidemic sequence type (ST45). For our third strain, KPN46, we chose a non-epidemic, antibiotic-susceptible strain (ST433) that we used to represent the many non-high-risk-clones that commonly infect hospitalized patients. CRE-166 was isolated from bronchiolar lavage (BAL) fluid while KPN46 and Z4160 were isolated from blood cultures. Samples isolated from feces fitting these patterns of epidemic sequence types and antimicrobial susceptibility were not available within our collection, but K. pneumoniae isolates from the lung and blood are thought to commonly originate from the GI tract.

The sizes of the CRE-166, Z4160, and KPN46 genomes were 6.00, 5.56, and 5.63 Mb, respectively. The core genome shared between them was 4.96 Mb (4,558 CDS), leaving CRE-166, Z4160, and KPN46 with 1.04, 0.63, 0.69 Mb (1115, 700, and 590 CDS), respectively, of accessory genetic content (**Figure 17**). Thus, CRE-166, Z4160, and KPN46 represented three clinical strains with phenotypic and genomic diversity suitable for subsequent studies of *K. pneumoniae* GI colonization.
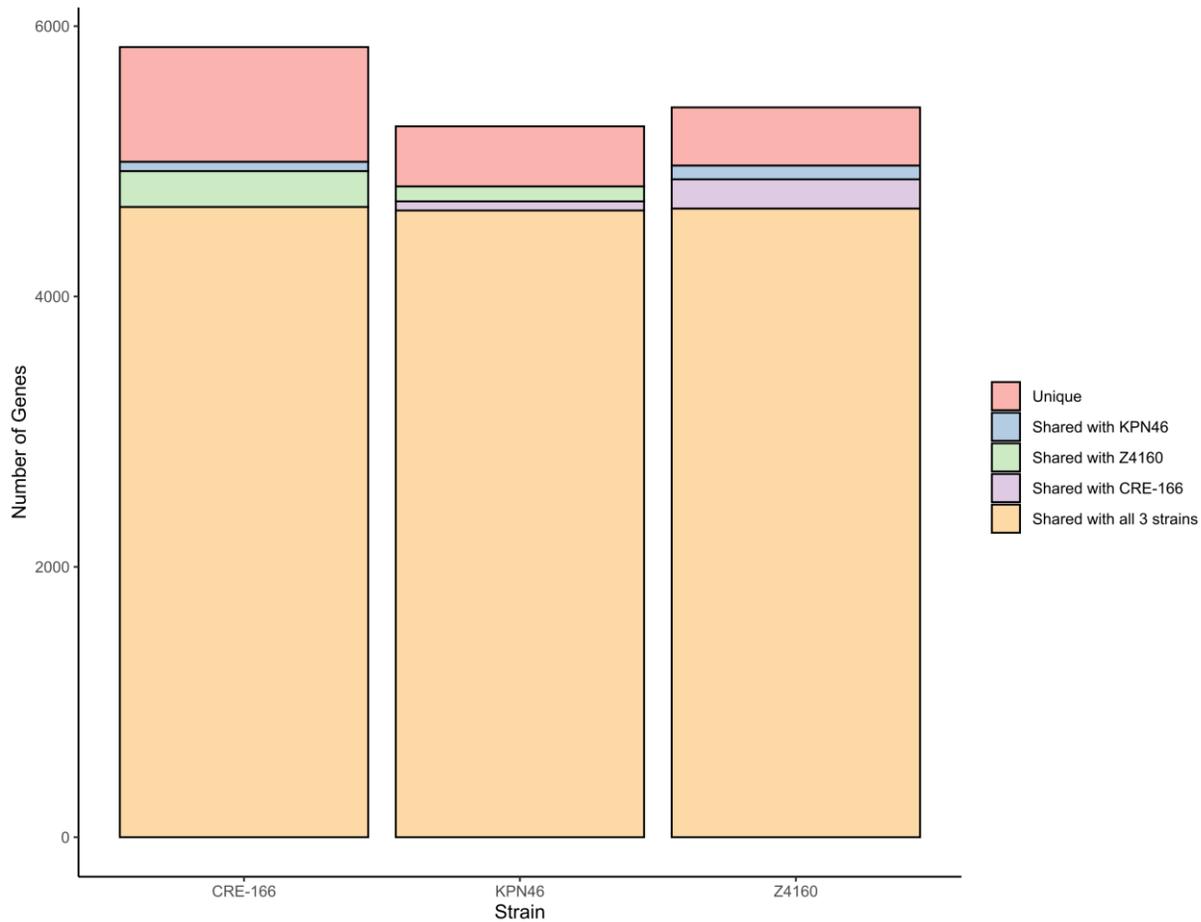
Figure 17 Unique and shared coding sequences between three representative strains of *K. pneumoniae*.

The software program Spine[110] was used to identify coding sequences with at least 85% homology in the strains indicated.

2.5 Evaluation of bottlenecks in the antibiotic-treated model of GI colonization

To determine whether transposon insertion sequencing experiments would be informative in our antibiotic-treated mcie, we next insured that mutants would not randomly drop out of the fecal output due to bottlenecks rather than to colonization defects. To this end, we constructed a marked CRE-166 strain by inserting an apramycin resistance cassette into the chromosomal Tn7 site. This marked strain did not have a growth defect in LB when compared to the parental strain (**Figure 18A**). To approximate the presence of a single transposon mutant within the pool of total mutants in an transposon mutant screen, we spiked this marked strain into an inoculum for gavage at a ratio of 1:100,000 with the parental strain. Next, we measured the ratio of the marked strain to total *K. pneumoniae* recovered from the feces of the mice. At Day 3 post-gavage, the marked strain was still detectable, suggesting the absence of a bottleneck significant enough to bias results (**Figure 18B**). However, at subsequent times post-gavage, we failed to recover the marked strain from some of the mice, indicating greater bottlenecks at later timepoints. We therefore chose Day 3 post-gavage as the timepoint for our screens.
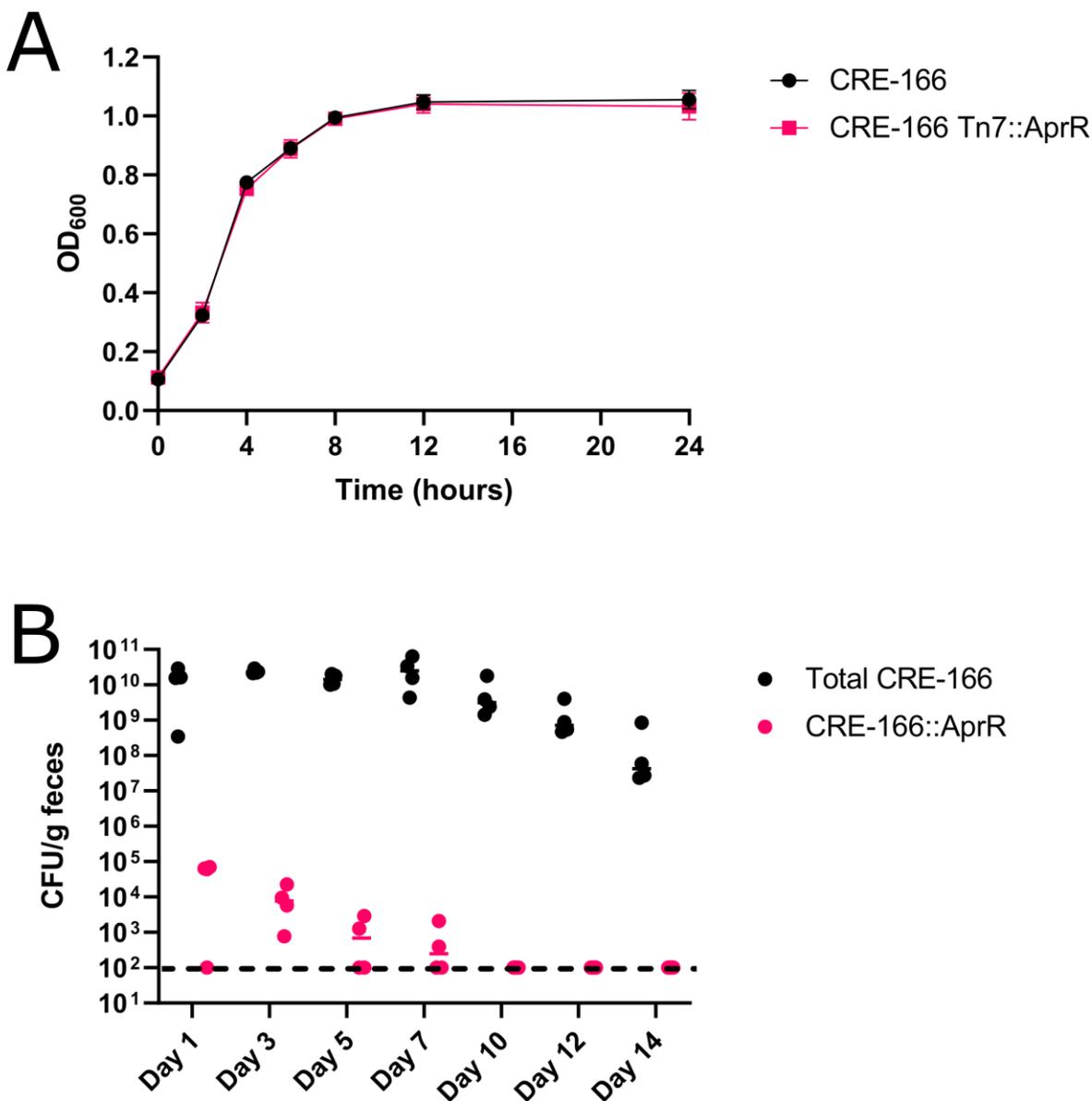
Figure 18 Validation and use of a marked *K. pneumoniae* strain to examine bottlenecks in the mouse model of GI colonization.

(A) Growth curve in LB for parent strain, CRE-166, and marked strain CRE-166 Tn7::Apr[R]. Points indicate an average of 3 technical replicates with a standard deviation marked in error bars. This is a representative curve from 3 biological replicates. (B) *In vivo* bottleneck detection experiment performed by spiking a marked strain, CRE-166 Tn7::Apr[R], into an inoculum of CRE-166 at a ratio of 1:100,000. Fecal samples were collected and CFU enumerated. $n = 4$.

2.6 Generation of three highly saturated transposon mutant libraries

To perform genome-wide screens for GI colonization factors, we generated transposon mutant libraries in three *K. pneumoniae* strains. The transposon vector pSAMerm was modified to express hygromycin resistance (pSAM*hygSDM*) to allow for selection of transposition in all three strains (including multidrug-resistant CRE-166). Libraries with over 145,000 CFU were generated for each strain (**Table 5**).

An initial assessment of library quality was performed by picking 32 colonies at random from each library and identifying transposon insertion sites with arbitrary PCR. Unique insertion sites for at least 26 colonies for each strain were successfully identified, and no colonies had more than one insertion site, indicating the libraries were of high quality.

	CRE-166 (ST258 CRE)	**Z4160** (ESBL producer)	**KPN46** (Antibiotic-susceptible)
Size (CFUs)	198,573	147,366	145,691
Unique insertion sites sequenced	29 (3 non-sequenceable)	26 (5 non-sequenceable, 1 with plasmid backbone)	26 (5 non-sequenceable, 1 with plasmid backbone)
Transposase present?	0/32	0/32	1/32

Table 5 Characteristics of Transposon Mutant Libraries Generated

Size of transposon mutant libraries and number of colonies (out of 32 picked for each library) with successful insertion site sequencing and detection of transposase by PCR are indicated.

2.7 Screening of mutant libraries for GI colonization factors *in vivo*

To specifically ensure the mice had no *K. pneumoniae* in their feces prior to gavage for this experiment, we collected feces before any experimental manipulation (Day -5) and then after vancomycin injections but before gavage (Day 0). Homogenized feces were serially diluted and plated on LB agar supplemented with carbenicillin to isolate *K. pneumoniae*, and no carbenicillin-resistant colonies were recovered on Day -5 or Day 0 (**Figure 19**).

On Day 0, we gavaged each of the three transposon mutant libraries into separate mice. A portion of the inoculum was saved and genomic DNA was extracted as the "input pool." At Days 1, 3, 5, 7, 10, 12, and 14, we collected fecal pellets for CFU enumeration and genomic DNA extraction. Although a bottleneck which may affect results past Day 3 was observed in pilot studies, we attempted to use a later timepoint in addition to Day 3 to determine if the data could still be interpretable. In the observed fecal burdens in **Figure 19**, we observed that the levels of CRE-166 and KPN46 at Days 12 and 14 appeared to be trending upwards rather than plateauing as expected from experiments with the parent strains (**Figure 12D**). As such, we opted to use the Day 10 timepoint which was more likely to represent a later timepoint which still followed similar colonization dynamics to the GI colonization experiments with the parent strains.

We used the method of Kazi et al (2020) to prepare sequencing libraries of the input (inoculum) and output (Day 3 and Day 10) pools. We prepared three biological replicates for each strain by selecting the samples for which there were the highest average DNA concentrations across both timepoints to increase the likelihood of successful library preparation.
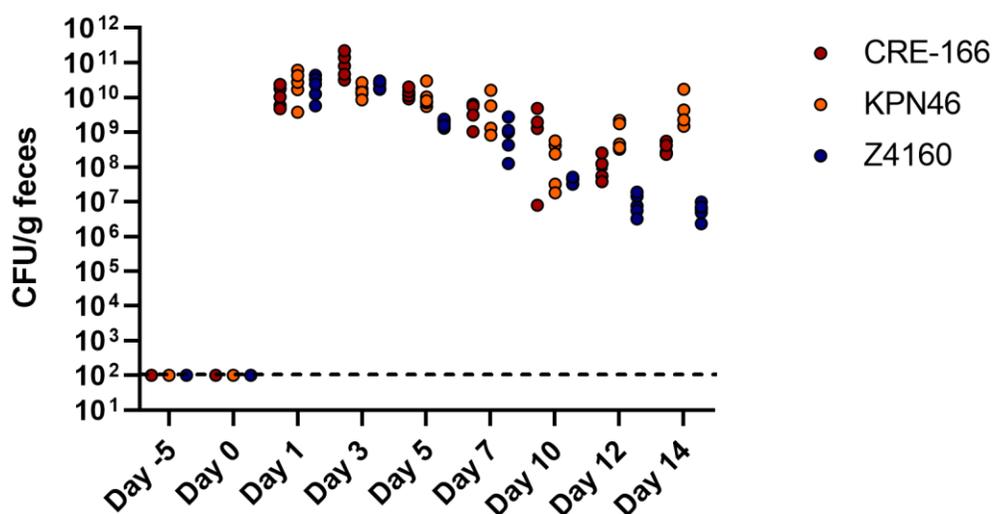
Figure 19 Fecal loads of transposon mutant libraries

Mice were injected with vancomycin for five days (Day -5 to -1) before gavage with 10^8 CFU of transposon mutant libraries (CRE-166, KPN46, or Z4160) on day 0. Fecal samples were collected and CFU enumerated. $n = 5$ for each strain. Dashed line indicates limit of detection.

However, following sequencing, the Day 10 output reads mapped to an extremely low number of loci in the genome (**Figure 20**, bottom track). In contrast, reads for the input replicates, and Day 3 outputs were well-distributed across the chromosome (**Figure 20**, top and middle tracks). The Day 10 results were not thought to be due to a technical error, as all of the Day 10 samples were sequenced to the same depth as the input and Day 3 samples.
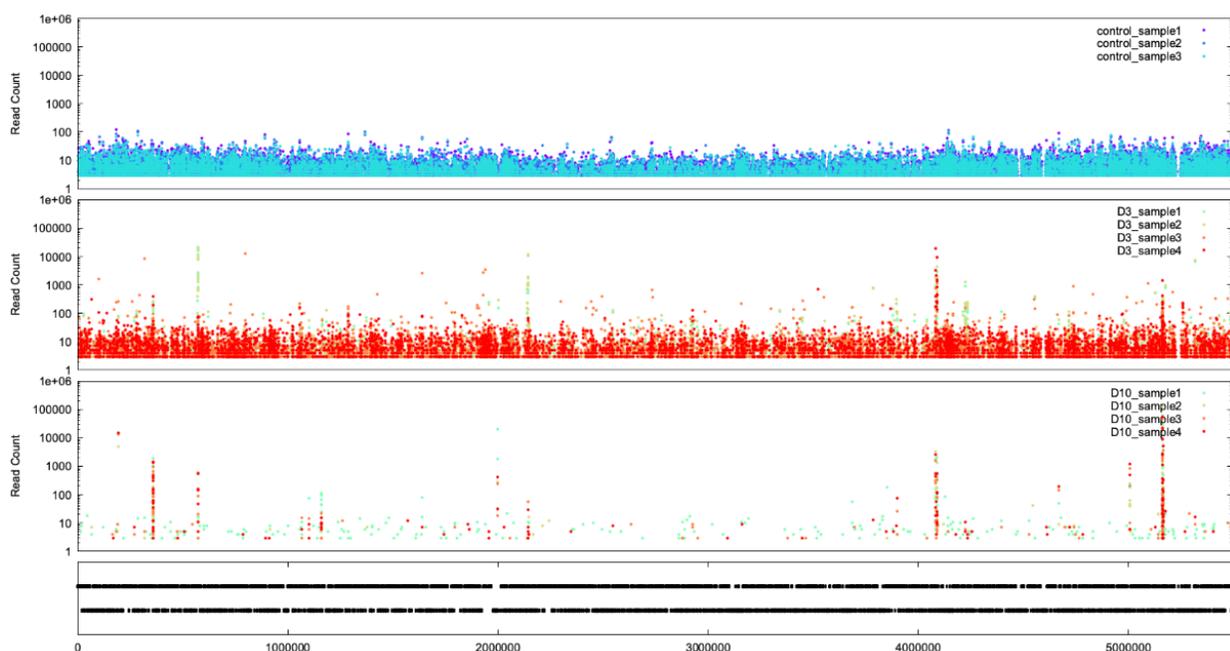
Figure 20 Mapped reads for CRE-166 transposon insertion sequencing (Run 1)
Transposon insertion sites from the (top track) input pool (middle track) Day 3 output and (bottom track) Day 10 output were sequenced and mapped to location on the genome. The x-axis represents position along the CRE-166 chromosome.

Closer inspection of the loci from the Day 10 output pool sequencing revealed that the loci with the largest numbers of reads were metabolic genes (**Table 6**). Of note, multiple genes involved in maltose transport (*malT*, *malE*, *malF*, and *malG*) were identified. As a previous study using similar methodologies had also observed outgrowth of mutants with transposon insertions in the *malT* gene[59], we concluded that these results represented the true biological state of the transposon mutant population at the timepoint rather than a technical error in library preparation. In other words, maltose transport mutants and several other metabolic mutants apparently had a strong selective advantage at later times, resulting in their overgrowth and predominance in the GI tract. However, these results were not informative, and we decided to proceed with analysis of Day 3 but not Day 10 output pools for all three strains.

Locus	*Gene*	*Product*
JCNGAGPE_00325	*malt_1*	HTH-type transcriptional regulator
JCNGAGPE_01889	*COQ5*	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial
JCNGAGPE_03887	*scrK*	Fructokinase
JCNGAGPE_03891	*cra_1*	Catabolite repressor/activator
JCNGAGPE_04777	*amiB*	N-acetylmuramoyl-L-alanine amidase AmiB
JCNGAGPE_04923	*malE*	Maltose/maltodextrin-binding periplasmic protein
JCNGAGPE_04924	*malF*	Maltose/maltodextrin transport system permease protein
JCNGAGPE_04925	*malG_2*	Maltose/maltodextrin transport system permease protein

Table 6 Transposon insertion loci overrepresented in Day 10 samples

Reads for sequencing libraries prepared from DNA extracted from feces 10 days after inoculation with a CRE-166 transposon mutant library were mapped to the CRE-166 genome. Loci with the highest number of reads are displayed along with their gene annotations.

Input pool sequencing demonstrated that over 82% of coding sequences had at least one insertion, and the average gene had 5 insertions. Coverage was distributed across chromosomes (**Figure 21**), confirming that all libraries were well-saturated.
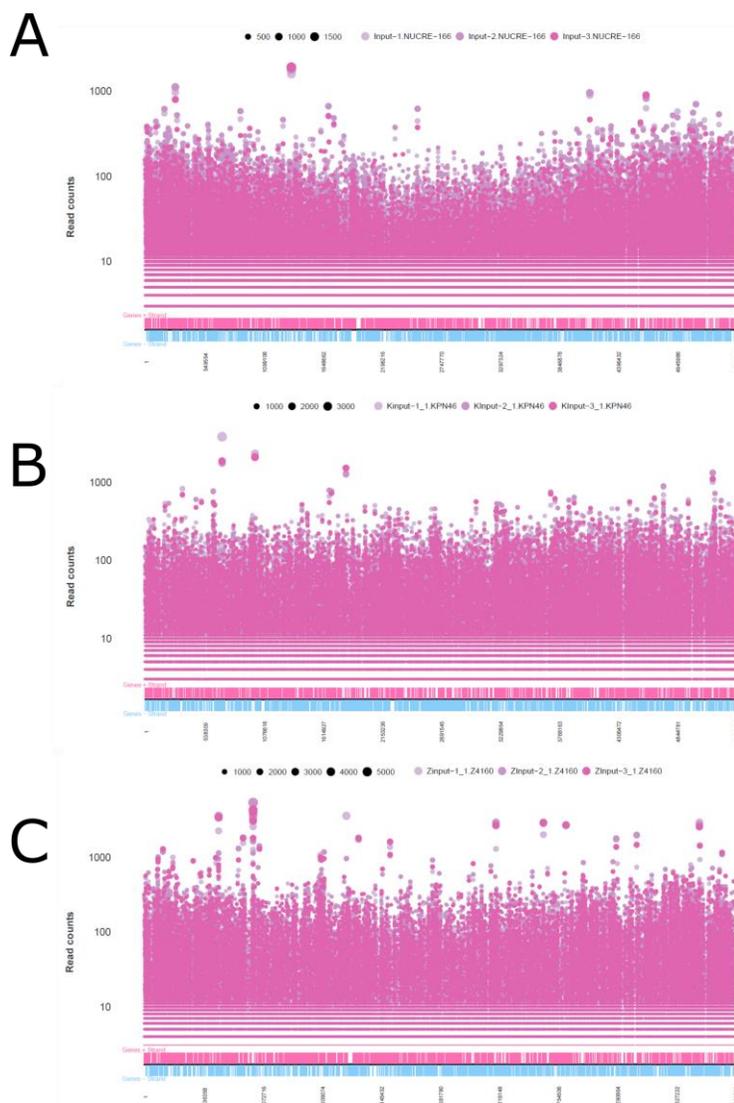
Figure 21 Distribution of transposon insertion sites and sequence read numbers across the *K. pneumoniae* chromosome.

Insertion sites are shown for (A) CRE-166, (B) KPN46, and (C) Z4160. Pink dots indicate reads of insertion sites, and dot size indicates number of reads. The first track below denotes coding sequences on the positive strand, and the lowest track indicates CDS on the negative strand.

Insertion site sequencing reads were then processed using a modified version of the previously described ESSENTIALS pipeline[122]. We first analyzed the input pools to identify genes required for the bacteria to grow in LB, denoted "essential genes." A total of 487 genes were identified as essential in all three strains, but a substantial number of genes were essential in only one or two strains (**Figure 22A**).
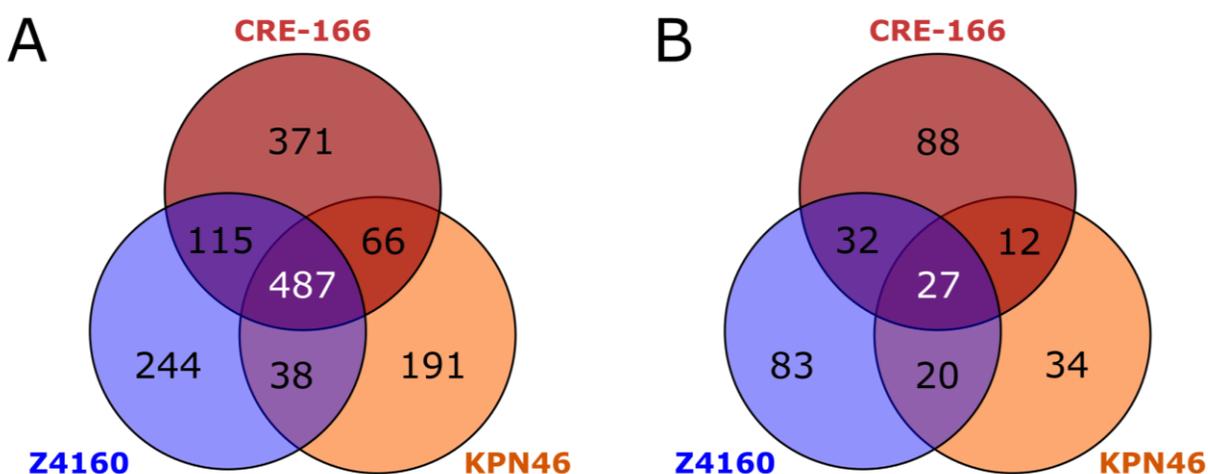
Figure 22 Genes required for growth in LB and for GI colonization.

Transposon mutant libraries were screened in a mouse model of GI colonization. n = 3 for inputs and outputs for each strain. (A) Shared and unique essential genes for each strain (B) Shared and unique genes for which transposon insertions resulted in reduced fitness in the gut. Genes had a $\log_2$ fold change < -2 and a false discovery rate < 0.05.

To determine which genes each strain utilized for GI colonization, we compared the total number of insertion reads per gene in the Day 3 output versus input pools. Bacteria containing transposon insertions that disrupted genes important for colonization are expected to be recovered in lower numbers in output versus input pools. We focused on genes that had a less than -2 $\log_2$(fold-change) (logFC) in output vs input insertion reads and a false discovery rate (FDR) less than 0.05 (**Figure 23**).
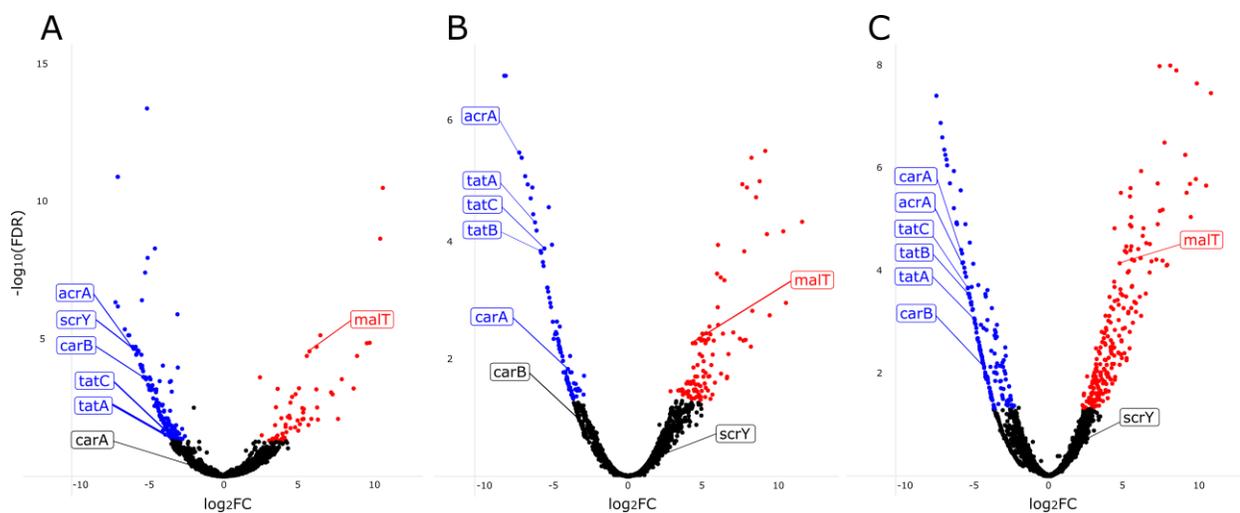
Figure 23 Volcano plots showing results of transposon insertion sequencing experiments in mice. Genes in which insertions were associated with decreased GI colonization (blue dots) and increased colonization (red dots) for (A) CRE-166, (B) KPN46, and (C) Z4160 are shown. Labeled points indicate targets which were chosen for the creation of isogenic mutants. FDR, false discovery rate; FC, fold change.

Multidimensional scaling (MDS) plots of input and Day 3 output pools demonstrated that the input pools were closely related and distinct from the Day 3 output pools for each strain (**Figure 24**).

Twenty-seven genes were used by all 3 strains for GI colonization (**Figure 22B**). However, many genes were used by only a single strain to establish colonization: 88 for CRE-166, 83 for Z4160, and 34 for KPN46. Intriguingly, most genes identified as important for colonization in at least one strain were present in all 3 strains. That is, only 22.6% of colonization genes for CRE-166, 3.3% for KPN46, and 4.9% for Z4160 were unique to each strain, suggesting that these three *K. pneumoniae* strains mostly rely on shared genes for GI colonization, but use different subsets of these shared genes for this purpose.

To determine whether these colonization genes were also found in most other *K. pneumoniae* strains, we calculated a core genome from a set of 323 previously described strains[109]. We defined the core genome as the genes shared by 95% of these strains and determined whether the genes identified as colonization factors in our experiments were in this broader core genome. Somewhat larger percentages of the colonization genes for each strain were now considered accessory genes: 25.8% for CRE-166, 16.1% for KPN46, and 15.4% for Z4160. However, most genes required for colonization by each strain were still genes shared across *K. pneumoniae* strains.

In addition to genes required for colonization, we also identified genes which, upon disruption with a transposon, conferred a colonization advantage. There were 7 genes found to confer an advantage when disrupted in all 3 strains (**APPENDIX 5**). Four were involved in maltose transport (*malT, malEFG*) while the other three had regulatory roles (*proQ, prc,* and *rspR*).
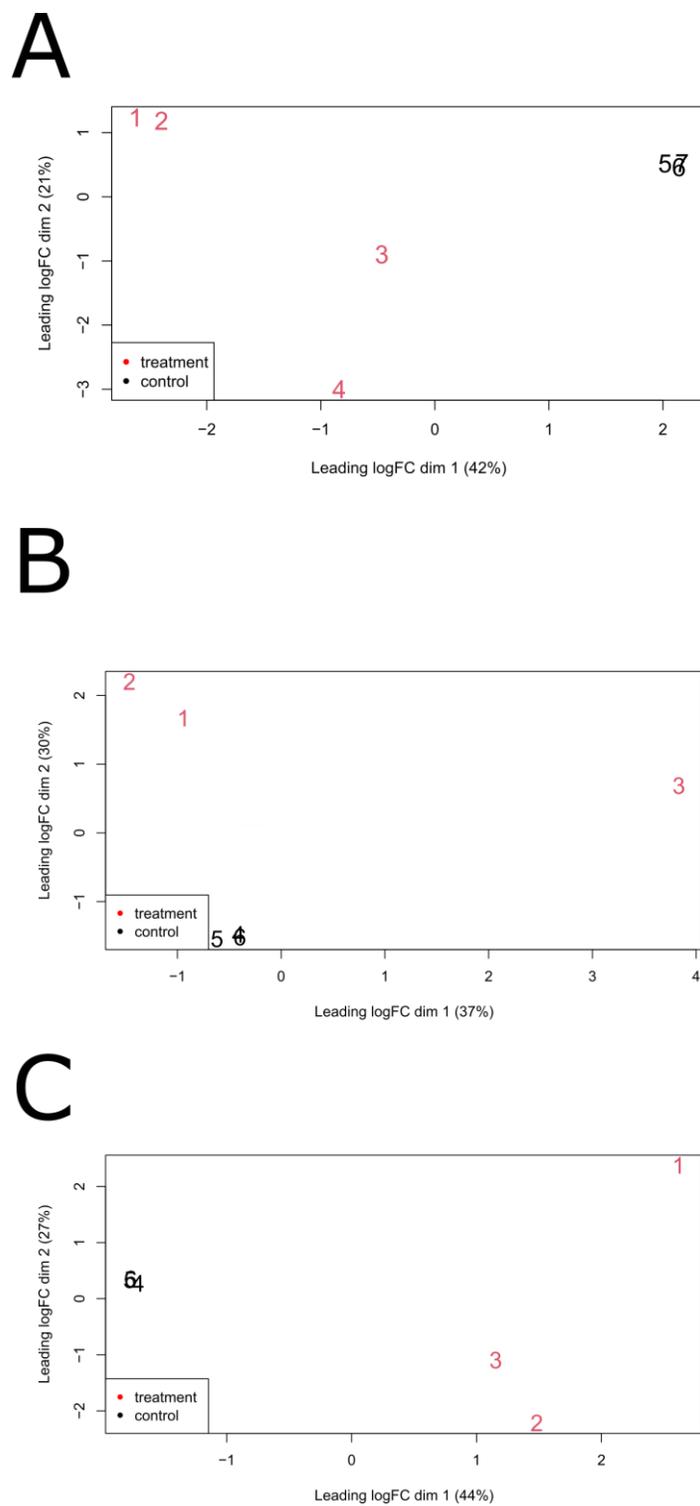
Figure 24 Multidimensional scaling (MDS) plots for transposon insertion results.
Input control pools (black font) and outputs from the treatment (GI colonization) (red font) for transposon mutant screens in (A) CRE-166, (B) KPN46, and (C) Z4160 are indicated.

2.8 Classification of GI colonization factors and pathways

To better understand the core colonization program in *K. pneumoniae*, we first focused on the 21 genes that contributed to colonization across all three strains (**Error! Reference source not found.**). As expected, we identified genes involved in anaerobic metabolism (e.g., *adhE, fnr, focA)*. We also found genes involved in other metabolic pathways, including *mtlD* (mannitol-1-phosphate dehydrogenase) and *carAB*, which encode the subunits of carbamoyl phosphate synthase that are responsible for the first committed step in synthesis of pyrimidine and arginine[123]. *carA* was identified in two strains and *carB* in the remaining strain. *tatA* and *tatC* (folded protein secretion apparatus[124]) and *acrA* (efflux pump[125]) were also identified.

| Gene | Annotation | log$_2$(Fold change) | | |
		CRE-166	KPN46	Z4160
aceE	Pyruvate dehydrogenase E1 component	-4.412	-5.504	-5.440
acrA	Multidrug efflux pump subunit	-5.906	-6.935	-5.512
adhE	Aldehyde-alcohol dehydrogenase	-3.353	-4.848	-7.339
arcB	Aerobic respiration control sensor protein	-5.342	-3.821	-6.018
bglY	Beta-galactosidase	-6.948	-2.896	-5.749
cvpA	Colicin V production protein	-4.611	-4.449	-6.031
cydA	Cytochrome bd-I ubiquinol oxidase subunit 1	-4.387	-6.933	-4.731
fnr	Fumarate and nitrate reduction regulatory protein	-4.248	-6.542	-5.723
focA	Formate Transporter	-6.486	-7.092	-6.679
glnA	Glutamine synthetase	-3.667	-3.755	-4.313
miaA	tRNA dimethylallyltransferase	-5.773	-3.856	-3.992
mtlD	Mannitol-1-phosphate 5-dehydrogenase	-5.711	-6.709	-5.990
ompC	Outer membrane porin C	-5.738	-3.129	-3.472
pal	Peptidoglycan-associated lipoprotein	-4.172	-6.183	-4.662
pflA	Pyruvate formate-lyase 1-activating enzyme	-5.617	-7.979	-6.821
pflB	Formate acetyltransferase 1	-4.766	-5.270	-4.077
pgi	Glucose-6-phosphate isomerase	-5.404	-5.665	-4.507
ptsI	Phosphoenolpyruvate-protein phosphotransferase	-3.088	-6.240	-7.061
purC	Phosphoribosylaminoimidazole-succinocarboxamide synthase	-3.470	-4.634	-4.273
purH	Bifunctional purine biosynthesis protein	-3.710	-5.209	-8.324
pykF	Pyruvate kinase I	-4.431	-4.701	-3.025
setA	Sugar efflux transporter A	-6.943	-5.168	-5.053
tatA	Sec-independent protein translocase protein	-3.384	-6.061	-4.829
tatC	Sec-independent protein translocase protein	-3.452	-5.452	-5.276
tolA	Tol-Pal system protein	-5.977	-5.965	-4.284
yeiE	HTH-type transcriptional activator	-5.445	-4.572	-4.522
-	Hypothetical polysaccharide deacetylase	-5.327	-4.906	-3.985
carA	Carbamoyl phosphate synthase small subunit	-	-3.917	-5.662
carB	Carbamoyl phosphate synthase large subunit	-5.267	-	-4.220

Table 7 Genes contributing to GI colonization in all 3 representative strains of *K. pneumoniae*.

For each strain, genes with log$_2$fold change < -2 and FDR < 0.05 were compared. For the *carAB* two-gene operon, *carB* met these criteria for CRE-166 and Z4160 whereas *carA* met these criteria for KPN46 and Z4160.

We were also interested in the pathways that each strain relied on for colonization. We assigned Kyoto Encyclopedia of Genes and Genomes (KEGG) identifiers to all genes in each strain and determined which pathways were enriched among colonization hits. These pathways fell into a few broad categories: metabolism, antimicrobial resistance, protein secretion, and environmental sensing (**Figure 25**). Our results confirm that metabolic capacities play an important role in the ability of bacteria to colonize the gut. However, there were differences in the metabolic pathways identified for each strain. For instance, colonization factors for Z4160 were enriched in the alanine, aspartate, and glutamate metabolism pathways. The colonization factors for CRE-166 and KPN46 were enriched for two-component systems, which may have played a role in metabolic adjustments caused by environmental sensing in the GI tract. In terms of antimicrobial resistance pathways, two strains (CRE-166 and KPN46) were reliant on genes that conferred resistance to cationic antimicrobial peptides (CAMPs), which are released by colonic epithelium and which are similar to microcins released by members of the microbiota. Thus, defense against host and microbiome factors is likely key to colonization by *K. pneumoniae*. Additionally, genes for resistance to beta-lactams were enriched in the screen results for KPN46. These included efflux pumps which likely play a role in the efflux of toxic compounds. Finally, protein export (the Tat secretion system) was enriched for KPN46 and Z4160, suggesting that secreted proteins may enhance the colonization of these strains.
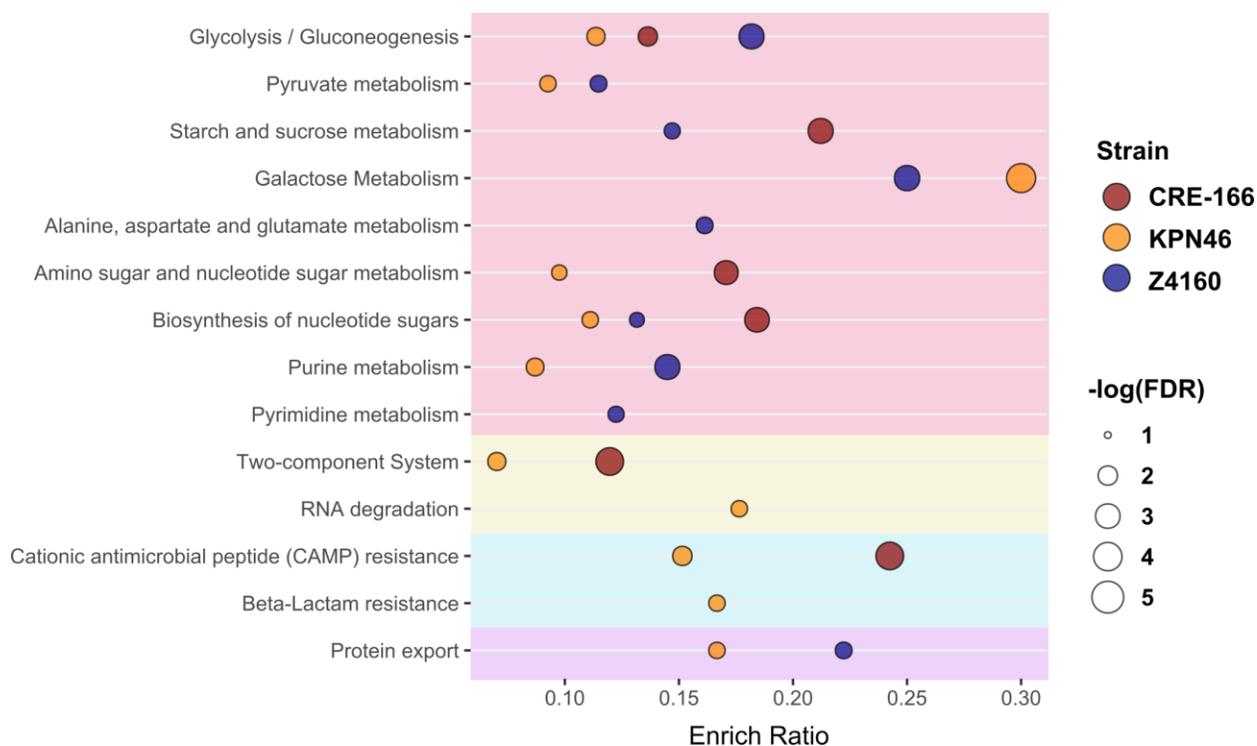
Figure 25 Pathways used for GI colonization in 3 strains of *K. pneumoniae*.
Kyoto Encyclopedia of Genes and Genomes (KEGG) identifiers were assigned to all genes in the genomes of CRE-166, KPN46, and Z4160. The enrichment ratio (Enrich Ratio) was calculated as the ratio of genes in the target list belonging to the specified KEGG pathway to the total number of genes in the pathway in the genome. A hypergeometric test was used to determine false discovery rate (FDR), and FDR < 0.05 was considered significant.

2.9 Validation of shared colonization factors

To validate our screen, we created isogenic mutants of 3 genetic loci—*acrA*, *carAB*, and *tatABCD*—required for GI colonization in all 3 strains (**Error! Reference source not found.**). These loci were chosen because they represent different functional groups: antimicrobial resistance, metabolism, and secretion. We generated isogenic mutants in which the coding sequence or operon of the target was replaced by an apramycin-resistance cassette. We verified that these mutants (and all others used in this study) did not have growth defects in LB compared to their marked parental strains (**Figure 26**).
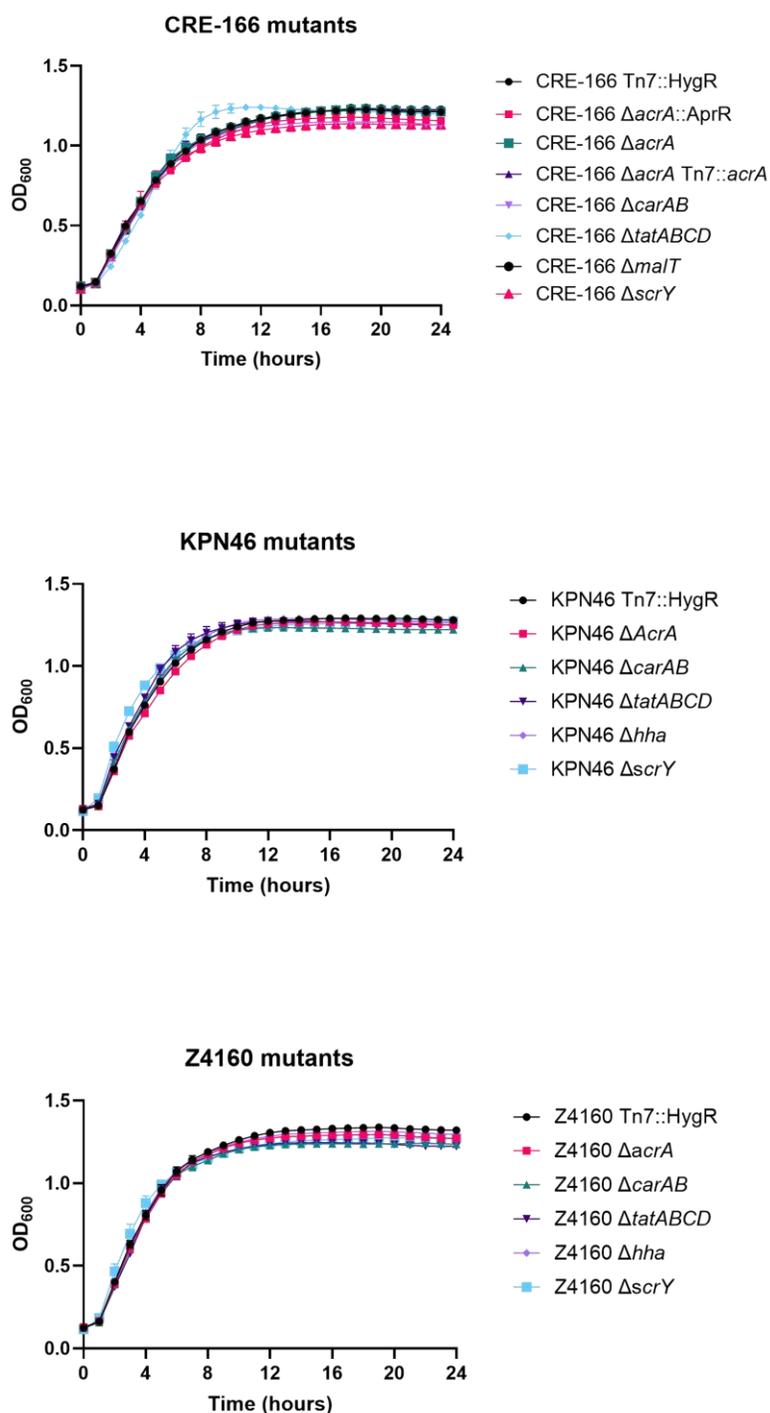
Figure 26 Growth curves in LB for marked parental strains and isogenic mutants of K. pneumoniae

Growth in LB at 37°C was measured by optical density (600 nm) for each individual mutant.

Then, we inoculated 1:1 mixtures of the marked parental strains and mutants into the mouse model of GI colonization and enumerated CFU in the feces at Day 3 (the screen timepoint) to calculate competitive indices (CI). To characterize the effects of these mutants at later timepoints, we also followed the fecal burdens to Day 14.

In all three strain backgrounds, *acrA* mutants displayed significant colonization defects at Day 3 (validating our screen) as well as beyond to Day 14 (**Figure 27A-C**). We constructed a complemented strain with an unmarked deletion of the gene locus, inserting *acrA* along with its upstream region and a downstream apramycin-resistance cassette into the chromosomal Tn7 site. This complement rescued the colonization defect fully at Day 3 and partially (**Figure 27D**) at subsequent timepoints.

The *carAB* deletion mutants were similarly tested in competition with their parental strains. At day 3 post-gavage, each *carAB* mutant exhibited colonization defects, continuing to Day 14 for CRE-166 and Z4160 (**Figure 27E-G**). For the KPN46 mutant, greater variability in CI was observed at later timepoints, suggesting the existence of a priority effect in the second week, during which mutants that initially established themselves tended to subsequently do very well while the others did progressively more poorly.

Finally, deletion of the *tatABCD* operon also significantly decreased colonization capacities, both at Day 3 and throughout subsequent days (**Figure 27H-J**). Insertion of the *tatABCD* operon at the Tn7 site fully rescued the colonization defect (**Figure 27K**). Additionally, insertion of *tatABC* at the Tn7 site also fully rescued colonization (**Figure 27L**), indicating the effect was not dependent on *tatD*. Thus, we verified our ability to detect shared factors essential for GI colonization.
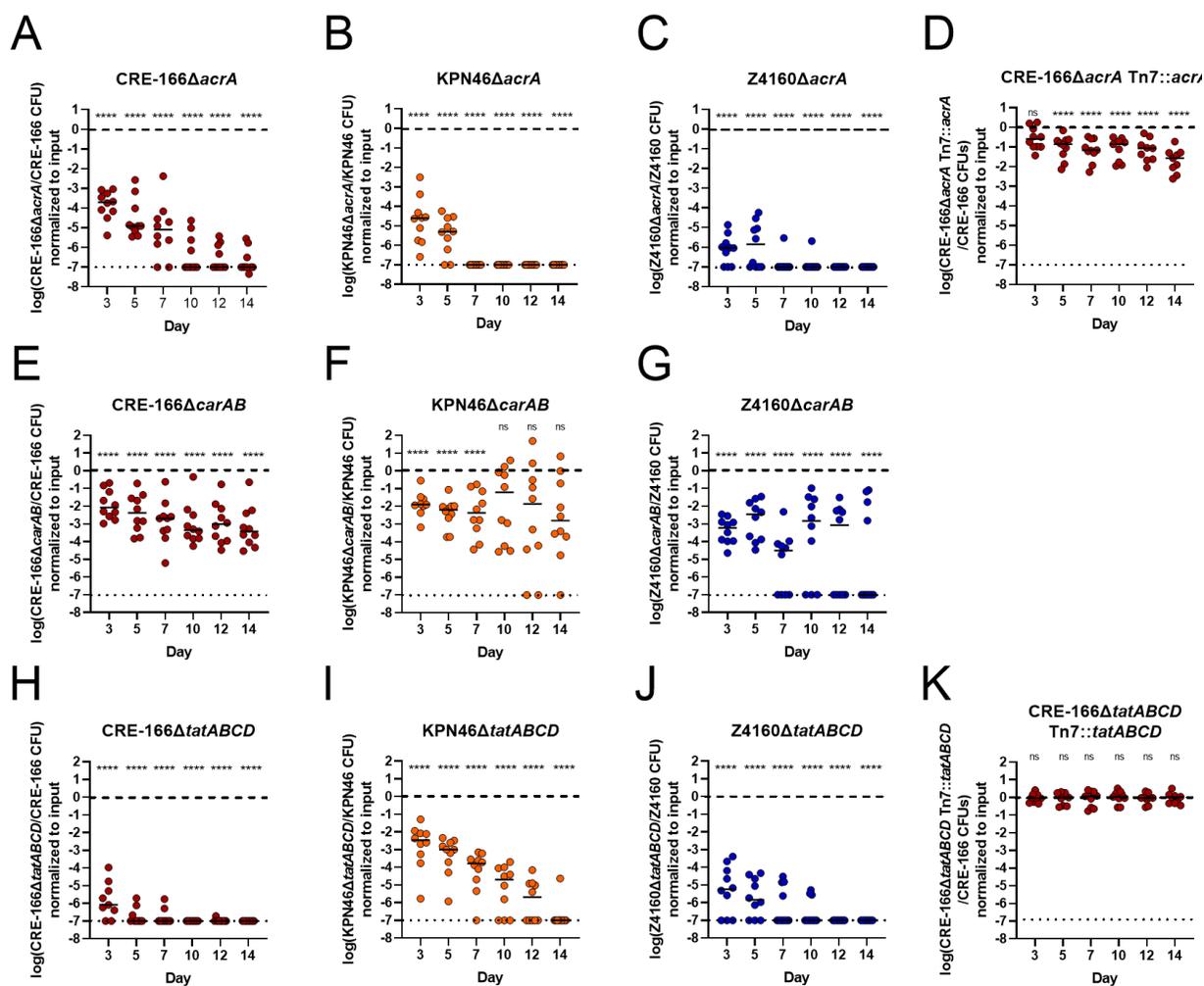
Figure 27 Competitive colonization between parent strains and isogenic mutants to validate genes identified as shared colonization factors in transposon mutant screens

Mice were treated with 5 days of vancomycin prior to gavage with 1:1 mixtures of marked parent strain (hygromycin-resistance cassette at the Tn7 site) and isogenic mutant (substitution of open reading frame with apramycin-resistance cassette) of target gene(s) or complemented mutant (insertion of gene(s) at the Tn7 site). $n = 10$ for A-G and K-M and $n \geq 9$ for H-J. Asterisks denote significance by one-sample t-tests with Dunn's correction where * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, and "ns" indicates not significant. Limit of detection was a competitive index of 10^{-7}, denoted with a dotted line. Log(competitive index) = 0, or equal recovered CFU of parental strain and mutant, is marked with a dashed line.

To determine whether the colonization defects were *in vivo*-specific, we also conducted *in vitro* competition experiments in LB. Isogenic mutants of CRE-166 were inoculated 1:1 with the marked parental strain in LB, and CFU of each strain were enumerated at 2 and 24 hours. Only the *tatABCD* mutant exhibited an *in vitro* competitive defect, indicating that the other mutants had *in vivo*-specific competitive defects (**Figure 28**).
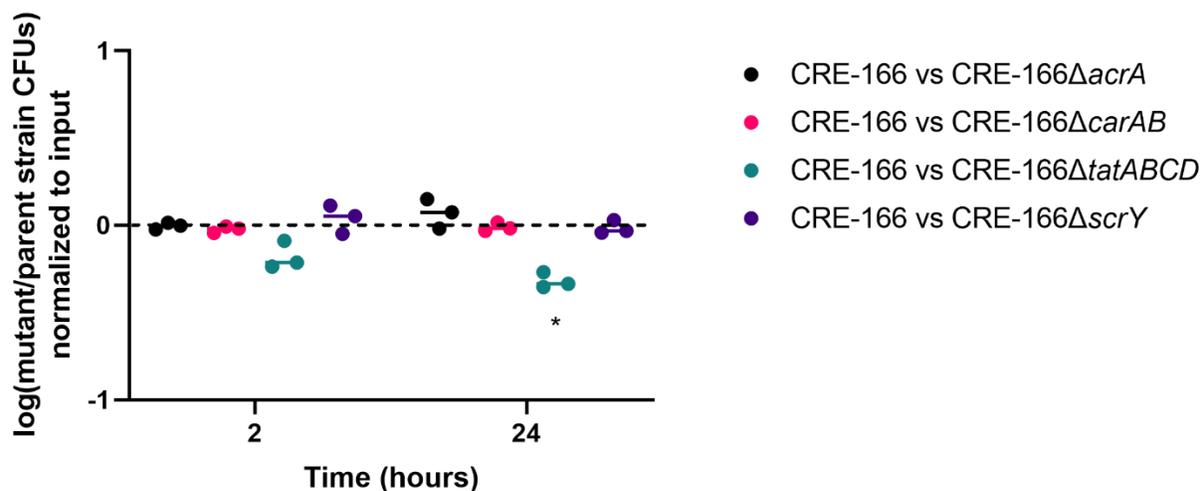
Figure 28 *In vitro* **competition experiments between marked parental strains and isogenic mutants of** *K. pneumoniae***.**

Strains were inoculated in a 1:1 mixture into LB, incubated, and CFU were plated for enumeration at the indicated timepoints. n = 3 biological replicates. Line denotes median. * indicates $p < 0.05$ in one-sample *t*-tests with Dunn's correction. Log(competitive index) = 0, or equal recovered CFU of parental strain and mutant, is marked with a dashed line.

We also selected one target that exhibited a colonization *advantage* upon disruption for validation. We chose *malT*, the transcriptional regulator for maltose uptake and metabolism. A *malT* deletion mutant in CRE-166 did not have a growth advantage in LB (**Figure 26**), but this deletion conferred a substantial colonization advantage over the parent strain at Day 3 and

beyond *in vivo* (**Figure 27M**), indicating our screen was also valid for detection of genes that confer colonization advantages.

While not identified in this screen, the maltose outer membrane porin, *lamB*, is the receptor for lambda phage. As such, we conducted plaque assays to determine whether the *malT* mutant was more successful in the GI tract due to increased resistance to phage. However, we were unable to observe any plaque formation from exposure to the supernatant of fecal homogenates from mice not colonized with *K. pneumoniae* and mice colonized with CRE-166. As such, we were unable to rule in the role of phage in this phenotype.

In addition to competition experiments *in vivo*, we also inoculated mice with just the deletion mutants of colonization factors to see if they would have a colonization defect on their own (in the absence of competition). However, when inoculated individually the Δ*acrA*, Δ*carAB*, Δ*tatABCD* mutants of CRE-166 still produced fecal burdens comparable to that of the parent strain (**Figure 29**). This may indicate that our model of GI colonization cannot distinguish between mutants and their parent strains unless they are in direct competition with each other.
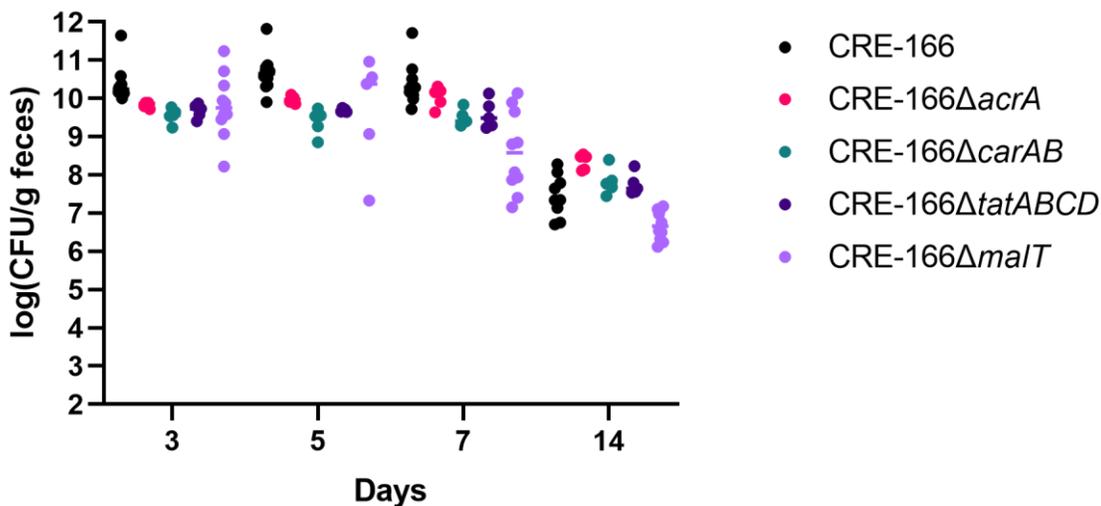
Figure 29 GI colonization with individual deletion mutants of colonization factors

Mice were treated with 5 days of vancomycin prior to gavage with individual deletion mutant strains. Fecal pellets were collected and CFU enumerated. $n \geq 5$ for all groups.

2.10 Deletion of *acrA* increases sensitivity to bile

We explored how one of our colonization genes, *acrA,* may contribute to GI colonization. As *acrA* encodes a component of an efflux pump that contributes to bile resistance in other GI pathogens[126], we tested whether our *acrA* mutant was more sensitive to bile. At 2, 4, and 24 hours after inoculation into 10% bile, the marked parental strain grew significantly better than the CRE-166Δ*acrA*::Apr^R mutant (**Figure 30**). We also generated a CRE-166Δ*acrA* mutant in which the apramycin-resistance cassette was removed from the Δ*acrA* allele. Like CRE-166Δ*acrA*::Apr^R, this mutant which also showed a growth defect in bile. A complemented strain (CRE-166Δ*acrA* Tn7::*acrA*) generated from this mutant showed that complementation rescued resistance to bile. Together, these data suggest that *acrA* supports *K. pneumoniae* GI colonization by providing resistance to bile.
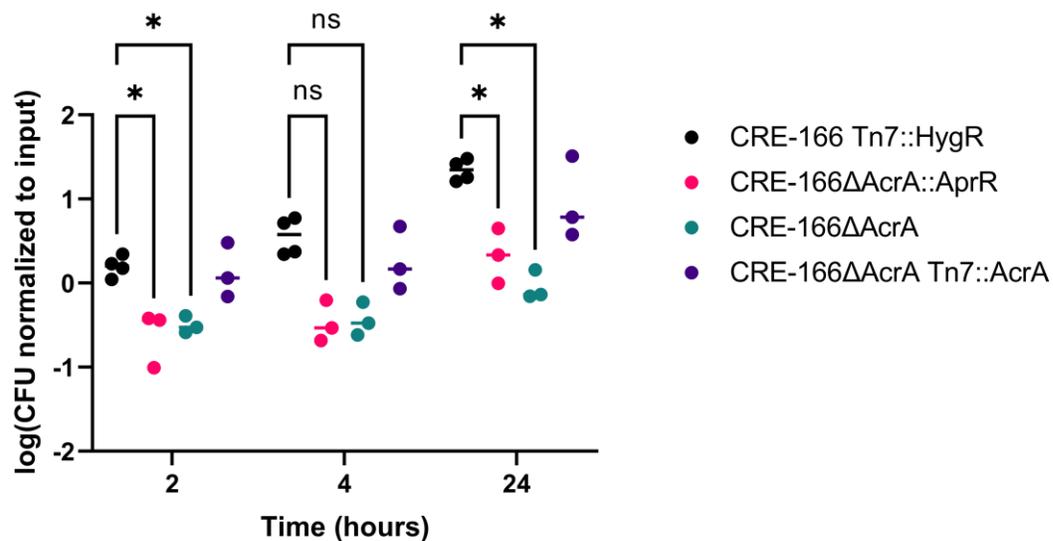
Figure 30 Resistance of CRE-166 and *acrA* mutants to 10% ox bile.

Strains were inoculated into 10% ox bile (w/v), incubated, and CFU were plated for enumeration at the indicated timepoints. $n = 3$ biological replicates. Line denotes median. * indicates $p < 0.05$ in two-way ANOVA with Tukey's HSD test.

2.11 Validation of strain-specific colonization factors

In addition to shared factors, we also wanted to confirm a strain-specific colonization factor. Hemolysin expression-modulating protein, encoded by *hha*, is a transcriptional regulator which scored as a hit in Z4160 but not other strains (**Table 8**). We generated *hha* deletion mutants in both Z4160 and KPN46 and confirmed that they did not have a growth defect in LB (**Figure 26**). At Day 3 in our mouse model of colonization, the day at which screen output pools were collected, the Δ*hha* mutant had a statistically significant colonization defect in Z4160 (**Figure 31A**). In KPN46, the Δ*hha* mutant had a slightly less severe colonization defect that was not statistically significant (**Figure 31B**). At subsequent timepoints, both *hha* mutants had colonization defects, but at Day 14 the median CI remained lower in Z4160Δ*hha* than in KPN46Δ*hha*.

Because the colonization difference of the *hha* mutant between two strains was small, we also investigated the role of *scrY*, a sucrose porin, which was identified as a hit for CRE-166 but not the other two strains (**Table 8**). As *scrY* occurs in an operon containing genes which had varying effects on colonization in the screen, we decided to create in-frame deletion mutants rather than marked mutants. As these deletion mutants did not encode for any antibiotic resistance, fecal burdens were estimated by recovering total *K. pneumoniae* present and subtracting marked parental strain CFU. This resulted in a narrower range of detection for competitive indices than for experiments where both strains were marked. Nevertheless, we were able to confirm that CRE-166Δ*scrY* had a colonization defect whereas KPN46Δ*scrY* and Z4160Δ*scrY* did not (**Figure 31C-E**). These data indicate that colonization factors may differ in their importance from strain to strain.

	hha		scrY	
	log$_2$FC	FDR	log$_2$FC	FDR
CRE-166	-1.83	0.362	**-5.39**	**0.000037**
KPN46	0.02	1	1.45	0.52
Z4160	**-4.95**	**0.0006**	1.68	0.25

Table 8 Genes identified as unique colonization factors

The log$_2$FC and FDR for each strain for *hha*, a colonization factor only identified for Z4160, and

scrY, one identified only for CRE-166 are displayed.
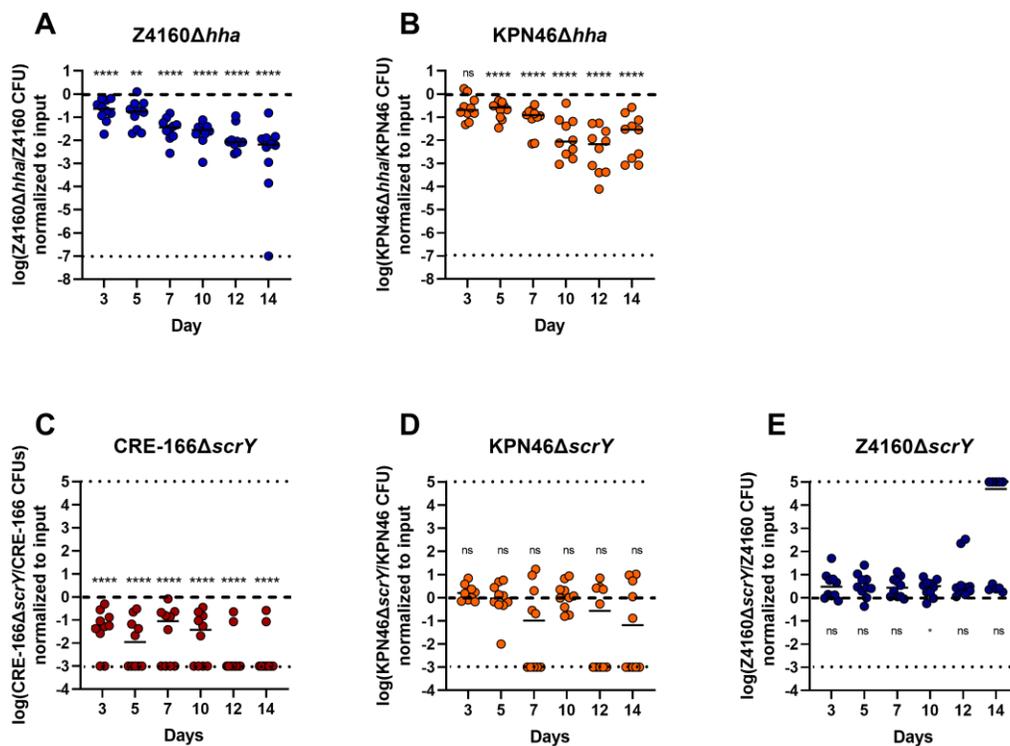
Figure 31 Competitive colonization between parent strains and isogenic mutants to validate genes identified as unique colonization factors in transposon mutant screens

Mice were treated with 5 days of vancomycin prior to gavage with 1:1 mixtures of marked parent strain (hygromycin-resistance cassette at the Tn7 site) and isogenic mutant. For A and B, the isogenic mutants consisted of substitution of open reading frame with apramycin-resistance cassette of *hha*. For C-E, the isogenic mutants consisted of an in-frame deletion of *scrY* without an apramycin-resistance cassette. $n = 10$ for each group, with two groups inoculated on separate days. Asterisks denote significance by one-sample *t*-tests with Dunn's correction where * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, and "ns" indicates not significant. Limit of detection was a competitive index of 10^{-7}, denoted with a dotted line. Log(competitive index) = 0, or equal recovered CFU of parental strain and mutant, is marked with a dashed line.

3. Discussion

Colonization of the GI tract by *K. pneumoniae* is a pivotal first step towards infection and transmission to other individuals[41]. The success of *K. pneumoniae* in causing many of the highly antibiotic resistant infections worldwide makes understanding the GI colonization process an important target for reducing morbidity and mortality from difficult-to-treat infections.

In this study, we aimed to answer the following questions. First, are there strain-dependent differences in GI colonization? To address this question, we developed a mouse model of colonization which mimics colonization of hospitalized patients and supports colonization by a wide range of classical *K. pneumoniae* strains. We also demonstrated that differences in colonization can be investigated with competition experiments in this model.

The second question we examined was how the wide genetic diversity of *K. pneumoniae* affects the colonization strategies of different strains. We investigated this question with three patient-derived strains with different phenotypic and genotypic features. We found that several sets of genes and pathways were used by all three strains to colonize the GI tract. However, other genes and pathways were unique to one or two strains, highlighting the need to examine multiple strains to more fully understand colonization in this genetically diverse species. In particular, we confirmed the importance of three different factors (*acrA*, *carAB, tatABCD*) in colonization by all three strains to validate our screens.

3.1 A clinically relevant murine model of *K. pneumoniae* GI colonization

We first attempted to study GI colonization without administration of antibiotics but found that fecal shedding was both low and variable. Similar attempts by other groups with classical strains has also demonstrated comparable fecal burdens[56]. Intriguingly, they did observe that after gavage, a fecal isolate was shed at more consistent, albeit still low, levels than the

hypervirulent lab strain KPPR1. This result suggests that passage through the gut may select for or upregulate features that promote fitness in the GI tract. To test this hypothesis, future studies could assess whether the fecal burden of a poorly colonizing classical strain could be increased by serially passaged through the mouse gut. Additionally, to date there have been no studies comparing colonization of multiple fecal isolates with that by strains isolated from different sources. In addition to observing similar results with classical strains, Young et al (2020) also reported the same difficulties that we experienced with establishing asymptomatic colonization with hypervirulent strains. Namely, gavage with hvKP1 induced mortality, just as we observed with hvKP2[56]. Thus, our results with antibiotic-free mice are similar to those of other studies.

But why does this lack of colonization occur? There are three likely reasons: colonization resistance from the microbiome, host factors, and technical limitations. First, the most obvious answer is that the native microbiome may prevent colonization and then, even if the *Klebiella* does manage to persist, the microbiome may suppress the *K. pneumoniae* population from blooming and causing a high fecal burden. This is likely the case, as we are able to induce high levels of fecal shedding using antibiotics which likely perturb the microbiome. Second, it is also possible that host factors are also important for colonization resistance. This may also be connected to the microbiome, as gut commensals are known to stimulate host-protective signaling and immune maturation[49]. Specifically, Bacteroidetes and host IL-36 signaling are known to exclude *K. pneumoniae* from the gut. It is possible that mice treated with vancomycin may also experience other gut changes, and future studies could be performed to examine inflammation or differential release of cytokines due to administration of vancomycin. Third, it is possible that some of the mice from which we did not recover any *K. pneumoniae* in the stool were actually colonized but shedding too little to be detectable. Our limit of detection was 10^2

CFU/g feces, and other studies with similar detection methods found occult colonization with *K. pneumoniae* can occur[57]. Additionally, another study demonstrated that fecal levels of *K. pneumoniae* in antibiotic-free mice might not be representative of colonic loads[66]. While our studies aimed to follow many colonization experiments over time (ruling out serial organ collection), organ loads or other tissue-based detection methods may be one option for investigating other questions concerning *K. pneumoniae* colonization without antibiotic treatment in the future.

As our study required stable, high levels of fecal shedding, we then turned to the use of antibiotics prior to bacterial inoculation. To mimic hospitalized patients, who are at risk for colonization and infection, we utilized the most commonly used antibiotic in the United States, vancomycin[118], and mimicked its most common route of administration (IV) through intraperitoneal injections. This antibiotic regimen differs from the ones used in previous screens in which antibiotics were administered, which may explain some of the differences in results. For instance, Jung, et al (2019) utilized a combination of vancomycin and metronidazole for 3 days prior to gavage while Struve, Forestier, and Krogfelt (2003) and Maroncle, et al (2002) administered streptomycin or ampicillin 1 day prior to gavage, respectively. Each of these 3 studies administered the antibiotics through drinking water and continued antibiotic treatment throughout the screen whereas our screen was executed without concurrent antibiotic administration.

Two factors in these different regimens may impact the results of these studies. First, the type of antibiotics chosen may have different effects on the microbiome. Unfortunately, metagenomic or 16S sequencing data is not available for these specific scenarios. However, the levels of fecal shedding which we observed in the first week of our colonization model (and our

screen timepoint at Day 3) was comparable to that reported by Jung, et al (2019). Second, it is possible that concurrent antibiotic stress may affect *K. pneumoniae* itself. While *Klebsiella* are resistant to all of the antibiotics used in the other three studies, previous studies have demonstrated that sub-inhibitory concentrations of antibiotics cause changes in bacterial gene expression[127]. As such, the results of the prior screens may also include signal related to antibiotic pressure in addition to colonization factors.

More broadly, our mouse model of GI colonization has other useful features. While the intraperitoneal injection of antibiotic may be more technically challenging than administration through drinking water, we are able to precisely dose the vancomycin and ensure each mouse receives the same amount of antibiotic. Additionally, this once daily regimen is still titratable by controlling the concentration and number of doses given. For administration of the *K. pneumoniae*, a one-time gavage is sufficient to produce high fecal burdens up to 60 days post-gavage, and multiple inoculations are not required. Our model also makes use of daily cage changes to minimize coprophagy, and therefore uncontrolled re-inoculation with fecally shed bacteria. Finally, our model has been validated in C56BL/6 mice. This background is widely used and readily available. Additionally, knock-out mouse lines can also be obtained allowing for future studies of host factors affecting *K. pneumoniae* GI colonization. In conclusion, our murine model of GI colonization is precise, clinically relevant, and reproducible.

3.2 Colonization capacities of different strains of *K. pneumoniae*

Our mouse model supported very high levels of fecal shedding for many *K. pneumoniae* strains, including antibiotic-resistant and susceptible strains of different sequence types. However, we were also unable to distinguish different colonization capacities of strains inoculated individually into the model. This may be because the antibiotic regimen may have

made the GI tract permissive to all colonization. In other words, 5 days of vancomycin may have removed all colonization resistance, and any strain of *K. pneumoniae* of any level of fitness may have been able to establish high levels of colonization in the gut. This may have also led to the finding that deletion mutants of colonization factors that exhibited competitive colonization defects against their parent strains were still able to colonize the GI tract readily when inoculated individually (**Figure 29**). Future studies could characterize colonization with lower or fewer doses of vancomycin to titrate to a level where differences in ability to colonize with more of an intact microbiome could be quantified.

Nevertheless, we were able to study differences in colonization fitness in this model by turning to competition experiments. We demonstrated that strain dominance could change over time, and that this dominance may be dependent on perturbations introduced by antibiotic administration. However, these patterns of strain dominance are not generalizable across strains with similar phenotypes. For instance, the carbapenem-resistant strain CRE-166 could outcompete the antibiotic-susceptible KPN46 after 2 weeks, but could not do the same to KPN41, another strain with low drug resistance. Additionally, CRE-166 was outcompeted by Z4160, an ESBL-producing high-risk clone, but not by Z4147, which also had those characteristics. Thus, global success of different strains may not be easily explainable by increased colonization fitness across the board. Despite this lack of overall generalizability, one of our sets of data suggests that colonization fitness may be associated with specific sequence types. CRE-166 was outcompeted by two different ST45 strains (Z4160 and KPN41). It is conceivable that lineage may be a better predictor of colonization ability than phenotypic characteristics and that ST45 strains are more likely to retain certain beneficial colonization

factors. This experiment (**Figure 16A**) was only completed with 5 mice and 2 ST45 strains, and more replicates and strains will be needed to test this hypothesis.

We move then to the question of how strains are able to outcompete others in the gut. Interbacterial competition can be mediated by direct methods like secretion of antibacterial effectors or indirect methods like increased metabolic fitness (ability to use different nutrients). *K. pneumoniae* strains can express Type VI Secretion Systems (T6SS) and microcins. The genomes of CRE-166, KPN46, and Z4160 all encode for the T6SS needle apparatus, and KPN46 is known to produce the microcin E492. Unfortunately, the T6SS effectors of *K. pneumoniae* have not been extensively characterized, and we do not know what these strains could be secreting to compete against one another and the microbiome. Overlay assays with CRE-166 and KPN46 were indeterminate on multiple occasions, so we are also unsure whether CRE-166 is susceptible to microcin E492 and whether direct killing contributes to KPN46 dominance during the first week. Future experiments could compare knockout mutants of the T6SS and microcin to see whether the same competitive index patterns still hold or if these systems contribute to dominance in the GI tract. In terms of metabolic contributions to fitness in the gut, a dominant strain might be able to utilize nutrients that the other cannot. Nutrient utilization in microbial communities is not only affected by what is available from the environment (gut lumen) but also what is available after other bacterial members of the community are able to utilize before their competitors. Additionally, recent work in bacterial communities indicates that timing or resource utilization also contributes to abundance of different members of the communities[128]. Strains may be able to utilize the same nutrient, but the one that switches on usage faster may be able to use up the nutrient before the other strain can. To probe these possibilities, metabolic phenotyping such growth on nutrient arrays or metabolic modeling from genome assemblies

could provide insight into the metabolic capabilities of different strains to correlate with competitive colonization phenotypes[129].

Finally, what is the biological relevance of these competitive colonization experiments? Namely, are patients ever colonized with multiple strains of *K. pneumoniae* at once? Rectal swabs are used for detection of carbapenem-resistant *Klebsiella*, but selective culture would exclude detection of non-resistant strains. Furthermore, it is unclear whether studies in which whole-genome sequencing is performed on isolates includes multiple colonies from the same swab identified as *K. pneumoniae* or not[41]. As a result, it is unlikely that most screening methods would pick up multiple strains of *K. pneumoniae* coexisting in the GI tract. However, a recent study in which *wzi* sequencing was performed on colonizing isolates demonstrated at least two patients where more than one *wzi* type was detected from banking up to 3 *K. pneumoniae* colonies cultured from rectal swabs[130]. This study also found instances in which strain replacement in the gut has occurred over time. In other words, at serial timepoints, *wzi* sequences different from what was previously isolated were in a single patient. Thus, colonization with two strains (as in our competition experiments) is a clinically occurring scenario. Differing colonization fitness may also play a role in strain replacement. One can envision a situation in which an already colonized patient is exposed to another strain of *K. pneumoniae* that is a better colonizer and so replaces the strain they were carrying previously.

3.3 Transposon Insertion Sequencing with multiple strains of the same species

Several studies in the past few years have highlighted the reality that though they may be members of the same species, individual strains can differ greatly in what genes are "essential[131,132]." One gene that is essential for growth in one strain may be dispensable in another. We can then extrapolate that this may be the case for scenarios of conditional

essentiality as well—scenarios like gut colonization. We hypothesized that factors that may be essential for GI colonization for one strain may not be so for another strain. Indeed, we found that most of the colonization factors that we identified in our screens were core genes, or genes that were shared by the other strains. However, many of these core genes were relied on by some strains but not others to establish colonization. It follows, then, that to appreciate the full range of mechanisms by which a species employs for pathogenesis, one must characterize the process in multiple strains.

Our study's novelty lies in the use of three different clinical strains to study one process: GI colonization. Our methods describe a procedure for generating saturating transposon mutant libraries in clinical and multidrug-resistant strains—features which have traditionally made genetic manipulation difficult. This method also allows for generation of libraries using rich media (LB), removing one step at which a library may be biased due to mutant drop out from culture in a minimal media even before being subjected to the experimental condition. In sum, our study demonstrates that additional information can be gleaned from including several strains in transposon screens as well as providing a method for generating the transposon libraries for such screens.

3.4 The core colonization program across multiple *K. pneumoniae* strains

Our transposon screen results defined a core colonization program of 27 genes utilized by all three strains. Most of these genes (16 of 21) were related to anaerobic metabolism and energy generation, as would be expected since metabolic adaption to the anaerobic colon is a prerequisite for successful colonization. In addition to these genes, three genetic loci involved in nucleotide synthesis were also necessary for all three strains to establish colonization: *carAB* and *purC/H*. Carbamoyl phosphate synthase, encoded by *carAB*, catalyzes a crucial step in the

biosynthesis of arginine and pyrimidines. *purC* and *purH* encode for enzymes in the biosynthetic

pathway for purines. Pyrimidines can also be obtained through the salvage pathway, which may

explain why deletion of *carAB* had no effect on growth in rich media but did impact colonization

of the GI tract, where these factors may be scarce.

Besides these metabolic genes, all three strains were reliant on the transcriptional

activator *yeiE*. The genes regulated by YeiE in *K. pneumoniae* are not well characterized, but in

Salmonella enterica, this activator controls expression of flagellar genes (which *K. pneumoniae*

do not possess) and a Δ*yeiE* mutant is defective in GI colonization[133].

Additionally, *tatA* and *tatC*, which encode components of the Tat folded-protein secretion

apparatus, were identified as elements of the core colonization program, also suggesting that

secreted factors may contribute to colonization. In particular, the Tat-secreted peptidoglycan

amidases, AmiA and AmiC, are necessary for colonization by *S. typhimurium*[134], and *amiC* was

identified in our CRE-166 screen. However, disruption of the Tat system also destabilizes the

cell envelope[135], decreasing resistance to bile acids[136]. Further studies will be necessary to

determine whether the contribution of the Tat system to GI colonization is through membrane

stabilization or through Tat-secreted effectors.

The porin *ompC* and components of the Tol-Pal system (*tolAB* and *pal*) were also

identified as critical for colonization for all 3 strains. In addition to allowing diffusion of small

solutes, OmpC is responsible for maintaining outer membrane leaflet asymmetry[137]. Although in

a different fashion, the Tol-Pal system also aids in maintaining the integrity of the outer

membrane[138]. Furthermore, deletion of *pal* in *K. pneumoniae* also increases sensitivity to bile,

one of the host-derived stresses bacteria encounter in the GI tract[84]. A few other genes that

contribute to bile resistance (*cvpA* and *acrA*) were critical for colonization. CvpA was first

characterized as a necessary component for the production of colicin V in *E. coli*[139]. However, its exact role in colicin production has remained unclear while more recent studies have found it is involved in colonization by multiple bacteria as well as in bile resistance in *E. coli*[140-143]. Finally, AcrA is a subunit of an efflux pump known to export a wide variety of molecules, including antibiotics and bile acids[125,126]. Thus, bile resistance may be an essential property for enteric colonization by all three *K. pneumoniae* strains used in this study. In summary, the core GI colonization program of *K. pneumoniae* is composed of genes involved in energy generation, nucleotide biosynthesis, protein folding and secretion, membrane homeostasis, and bile resistance.

We then validated three shared colonization factors *in vivo*. First, the efflux pump subunit *acrA* was identified in all three of our strains as a colonization factor, and levels of the isogenic mutant at Day 14 were undetectable in competition experiments for all biological replicates for KPN46 and Z4160 and most for CRE-166. AcrA is the periplasmic subunit of the tripartite efflux pumps which contain TolC and AcrB or AcrD[125]. These pumps export a large variety of substrates including multiple classes of antibiotics (beta-lactams, macrolides, fluoroquinolones, etc.) and bile acids in *E. coli* and *S. enterica*[125,144]. [125,144]. In addition, the AcrAB-TolC pump is required for *S. enterica* colonization in chickens and full virulence after oral inoculation of mice[144,145]. As the CRE-166Δ*acrA* mutant had decreased bile tolerance, the mechanism by which *acrA* supports *K. pneumoniae* GI colonization may be similar to its role in *S. enterica*. However, this would need further confirmation *in vivo*. In any case, due to the role of AcrA in increased antimicrobial resistance[94,146,147], its additional function in GI colonization makes it an attractive therapeutic target. Perhaps an AcrA inhibitor would not only increase efficacy of existing antibiotic treatments but also lower fecal burden of *K. pneumoniae* and risk of transmission to

other patients. As AcrA inhibitors are already in varying stages of pre-clinical development, further study with them or development of other inhibitors may prove fruitful[148].

The next shared colonization factor we confirmed was *carAB*, which encodes for the two subunits of carbamoyl phosphate synthase, which catalyzes the first committed step in arginine and pyrimidine synthesis[123]. Additionally, *carAB* was found to contribute to enteric colonization by *Salmonella typhimurium* along with other genes involved in purine and further pyrimidine synthesis[123]. As various KEGG pathways in the metabolism of nucleotides or nucleotide components were enriched for in both the CRE-166 and Z4160 screens, the production of nucleotides in the GI tract appears to be a crucial pathway for colonization. Thus, the mechanism by which *carAB* contributes to colonization may be through its role in *de novo* pyrimidine synthesis. As pyrimidines can also be obtained through the salvage pathway, this may explain why deletion of *carAB* had less of dramatic effect on colonization in competition experiments than deletion of the other targets (*acrA* and *tatABCD*). Outside the context of colonization, *carAB* contributes to inhibition of the oxidative burst in neutrophils by *Francisella tularensis* and virulence in the abscess model with *Porphyromonas gingivalis*[149,150]. Thus, it is also possible that *carAB* may also promote colonization by *K. pneumoniae* through a different mechanism, such as arginine biosynthesis and/or interruption of host immune processes, which could be a focus for future investigation.

Our final validated colonization factor in all 3 strains was *tatABCD*. The genes *tatABC* encode for the Tat folded protein secretion apparatus while *tatD* encodes for an exonuclease which is involved in DNA repair[124,151]. Although *tatA* and *tatC* were the genes identified in our screen, we opted to delete *tatABCD,* as they were arranged as an operon. Complementation of *tatABC* at the Tn7 site fully rescued the colonization defect of the *tatABCD* mutant, indicating

that the defect was not due to the loss of *tatD.* The Tat secretion pathway recognizes substrates with a twin-arginine motif and exports fully folded proteins, in contrast to the Sec pathway which translocates unfolded polypeptides[124]. The Tat pathway is thought to provide a solution for secreted proteins which cannot be folded correctly in environmental conditions without intracellular cofactors or chaperones. For *S. typhimurium*, *Vibrio cholerae*, *Yersinia pseudotuberculosis*, and *Campylobacter jejuni*, deletion of the Tat system results in diminished GI colonization[136,152-154]. The Tat-secreted peptidoglycan amidases, AmiA and AmiC*,* are known to be necessary for colonization by *S. typhimurium,* but the other Tat-secreted factors contributing to colonization for other bacterial species have not been described[134]. To determine whether the colonization defect of the *tatABCD* mutants is due to lack of secreted factors necessary for establishment of *K. pneumoniae* in the GI tract, future studies can be performed to determine what proteins are secreted by the Tat system in each strain and whether they have impacts on colonization ability. The Tat substrates in *K. pneumoniae* have not been studied extensively; only a beta-lactamase has been characterized in this species[155]. However, *amiC* was identified as a gene necessary for GI colonization by CRE-166 in our screen. Additionally, while it did not meet a logFC less than -2 or an FDR less than 0.05 for KPN46, the logFC and FDR for *amiC* in the Z4160 screen was -2.7 and 0.076 respectively. A second possibility for how deletion of the *tat* operon may impair colonization ability is through its pleiotropic effects on the cell envelope[135]. Additionally, deletion of *tatB* or *tatC* makes *E. coli* more susceptible to detergents due to the compromised permeability function of the outer membrane[135], and the *S. typhimurium* ΔtatC and ΔamiAamiC mutants were also more susceptible to deoxycholic acid[136]. Thus, the disruption of the *tat* operon in *K. pneumoniae* may also confer a colonization defect due to decreased bile resistance.

3.5 Comparison of identified colonization factors with results of published screens

In addition to validation by knockouts of individual targets, our results were also confirmed by the replication of several targets found in a previous screen of a different ST258 strain, MH258[59]. First, we also found that disruption of genes involved in maltose metabolism confers a fitness advantage in the GI tract. As other genes involved in metabolism were found in this category for all three strains, it is possible that these mutants were able to bypass a fitness cost to using those pathways in the gut. Specifically for the maltose-related genes, the lambda phage is known to use maltoporin *lamB* as its receptor[156]. The presence of this mutant in our list with increased fitness for both CRE-166 and Z4160 as well as *malEF* being common to all three strains may point to a role of phage in the dynamics of *K. pneumoniae* GI colonization. While we were unsuccessful in detecting phage in fecal homogenates, it is possible that another type of antibacterial factor such as a microcin may use the maltose transporter as a receptor.

In terms of other overlap with Jung, et al, ten of the genes identified as essential for GI colonization in their screen were also found to be essential in at least one of our strains (**Table 9**). Furthermore, the largest overlap was found with CRE-166 (9 of the 10 genes) whereas only 4 genes were shared with Z4160 and 2 with KPN46. As the previous screen was conducted with another ST258 strain (like CRE-166), this suggests that colonization strategies utilized by more closely related strains are more similar than those used by less related strains.

When we compared our results with those from Signature-Tagged Mutagenesis screens, we did not find any overlap with the 12 genes identified by Maroncle, et al (2002). However, out of the 19 genes from Struve, et al (2003), four were also identified as colonization factors for at least one of our strains. One of these genes, *arcB*, was also identified in the Jung, et al (2019) screen. This was not surprising, as it encodes for the aerobic respiration control sensor. As these

two STM screens identified only a small number of hits, it is not entirely unexpected that our results with a much larger transposon mutant library did not completely match.

Gene	Annotation	Struve, et al (2003)	Jung, et al (2019)	CRE-166	KPN46	Z4160
ackA	Acetate kinase		X	X	X	
arcB	Aerobic respiration control sensor protein	X	X	X	X	X
plsX	Phosphate acyltransferase	X				X
ptsI	Phosphoenolpyruvate-protein phosphotransferase		X	X	X	X
xylA	Xylose isomerase		X	X		X
csrD	RNase E specificity factor		X	X		
cyaA	Adenylate cyclase		X	X		
bamE	Outer membrane protein assembly factor		X	X	X	
ompA	Outer membrane protein A	X		X		X
ompC	Outer membrane porin C		X	X	X	X
bamB	Outer membrane protein assembly factor		X	X	X	
surA	Chaperone	X			X	X
tamB	Inner membrane component of TAM transport system		X			X

Table 9 Colonization factors found in this study and previous screens

Colonization factors from three screens in the GI tract with *K. pneumoniae* were compared from Maroncle, et al (2002)[103], Struve, et al (2003)[58], Jung, et al (2019)[59], and this study.

3.6 Strain-specificity of colonization genes and pathways

We also found a number of genes and pathways that contributed to colonization by one or two strains but not by all three strains. For instance, the KEGG pathway analysis revealed that unlike the other two strains, CRE-166 relies on amino sugar and nucleotide sugar metabolism and biosynthesis as well as two-component systems. This may indicate CRE-166 is especially reliant on metabolism of cell wall components as well as synthesis of carbohydrates for envelope decoration and employs more environmental sensing cues to achieve colonization. For Z4160, colonization factors were enriched for the pyruvate metabolism and purine and pyrimidine metabolism pathways. However, on further inspection, both CRE-166 and KPN46 colonization factors were enriched for genes in the KEGG module for *de novo* purine synthesis, and *de novo* pyrimidine biosynthesis genes were also overrepresented for KPN46. Thus, *de novo* nucleotide biosynthesis may be a key colonization process for *K. pneumoniae*.

In terms of unique colonization factors, CRE-166 had the most at 108. These included *nagAB*, two genes which encode for enzymes involved in the metabolism of N-acetylglucosamine (GlcNAc), a component of peptidoglycan. NagA is necessary for recycling of cell wall components, but NagB can catalyze a second step towards using GlcNAc as a carbon source[157]. Thus, CRE-166 may not only be dependent on recycling of cell wall components but also use of these amino sugars as a carbon source.

For our validation of unique colonization factors *in vivo*, we found that Z4160 relies more heavily than KPN46 on *hha* in early colonization. As Hha regulates virulence factors in *E. coli* and *S. enterica* in response to environmental cues, it may be responsible for transcriptional regulation of colonization factors in *K. pneumoniae*[158-160].

We also found that *scrY*, which encodes for a sucrose porin, was required for colonization by CRE-166 and was validated *in vivo*. However, KPN46Δ*scrY* and Z4160Δ*scrY* did not exhibit colonization defects. This porin is thought to be specific for uptake of sugar molecules[161]. However, whether other substrates can also be taken up through this porin is unknown. The genomes of all three strains contains several genes annotated as sucrose porins, and all three Δ*scrY* mutants are still able to grow in minimal media in which the only carbon source is sucrose. However, the differences observed in the validation experiments suggest that perhaps CRE-166 still relies on ScrY for another function. It is possible that KPN46 and Z4160 encode for other porins to uptake other nutrients that CRE-166 cannot make us of to compensate for the loss of ScrY. Metabolic phenotyping to determine if this is the case could be performed to probe the mechanism behind this finding.

In conclusion, pathways and genes which were colonization factors for individual strains were identified by our screen, and we were able to validate a large colonization defect associated with one of these factors. These findings support our hypothesis that colonization strategies differ between strains of *K. pneumoniae*.

3.7 A schematic for *K. pneumoniae* GI colonization

This study provides several insights into how *K. pneumoniae* establishes colonization in the GI tract. We found that high levels of GI colonization is difficult to establish in the absence of microbiome perturbation but that administration of vancomycin allows for dense colonization with many strains of *K. pneumoniae*. Additionally, we validated several shared colonization factors which likely aid *K. pneumoniae* by defending against host and microbiome factors (AcrA), possibly antagonizing competitors (Tat), and establishing a metabolic niche in the gut (CarAB).
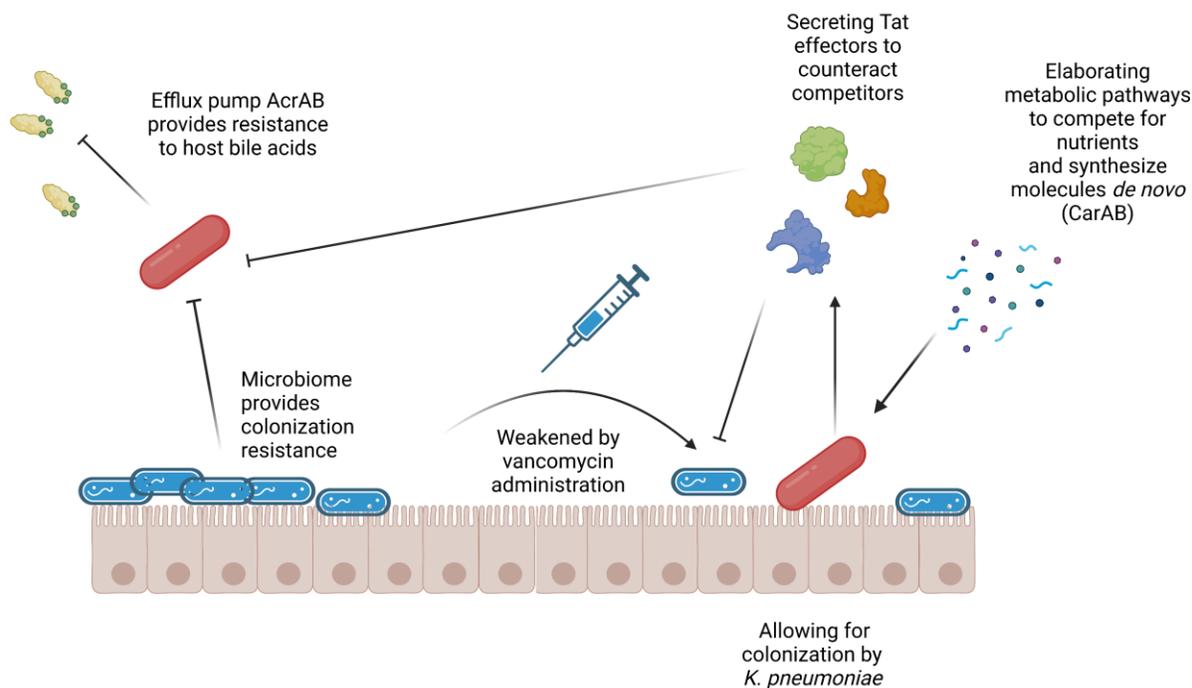
Figure 32 Schematic for validated shared colonization factors

3.8 Limitations of our approach

Our study had several limitations. All transposon insertion sequencing screens have a propensity to miss secreted factors and effectors because of trans-complementation. This may be one reason why we did not detect Type VI secretion systems in our screens even though they have been shown to contribute to colonization[66,102]. Interestingly, however, we did detect *amiC*, which encodes for a Tat-secreted substrate.

Additionally, our screen was performed on Day 3 of colonization, which was the latest timepoint at which we did not observe a substantial bottleneck. However, Day 3 may represent an early phase of colonization that may not be representative of longer term colonization. This would explain why our KPN46Δ*carAB* mutant did not exhibit a consistent colonization defect after Day 7 and why the subtle difference in colonization between Z4160Δ*hha* and KPN46Δ*hha* at Day 3 seems to resolve at later timepoints. However, for our other mutants, the phenotypes predicted by our screen are still consistent at Day 14, indicating that our results are still likely to represent many useful targets for further study.

Our use of -2 logFC as a strict threshold for identifying colonization genes may exaggerate strain differences when a gene falls slightly below this threshold in one strain but slightly above it in another. Finally, our study only used 3 strains. Additional studies will be necessary to determine whether our findings can be extrapolated to a larger number of *K. pneumoniae* strains.

As the *hha* mutants did not have as strong of a phenotype as the *scrY* mutants, we hypothesized that there may have been a technical limitation of the insertion sequencing analysis which could explain the differences. When we looked at the insertion sequencing reads which mapped to each locus, we found a comparatively large amount of reads (at least 60) for each *scrY*

input pools across all three strains (**Table 10**). Additionally, all output pool read counts were

either higher lower than input read counts in any particular strain. However, for *hha*, the CRE-

166 input pool counts were quite low, and the KPN46 output pool counts had one outlier of 7

(compared to 66 and 171 of the other two biological replicates (**Table 10**). As a result, the

average output read count was lower for KPN46 than it might have been without an outlier, and a

colonization defect may have been detected for CRE-166 if the overall read counts for the locus

had been higher. Thus, these two sets of mutants may demonstrate a false positive (*hha*) from

noise from the technique and analysis as well as a true positive hit (*scrY*). These findings suggest

that future transposon insertion sequencing studies should filter out genes for which input read

numbers are below a specific threshold. Ongoing work in the Hauser Laboratory is attempting to

generate analysis software that will allow for this adjustment.

		Input 1	Input 2	Input 3	Avg Input	Output 1	Output 2	Output 3	Avg Output
hha	CRE-166	10	17	6	**11**	2	3	1	**2**
	KPN46	96	210	114	**140**	66	171	7	**81**
	Z4160	114	96	70	**93**	5	1	1	**2**
scrY	CRE-166	201	301	280	**261**	2	2	10	**5**
	KPN46	188	208	217	**204**	817	439	324	**527**
	Z4160	60	71	63	**65**	142	88	144	**125**

Table 10 Read counts for genes identified as unique colonization factors
Number of insertion location reads in *hha* and *scrY* in each input pool technical replicate and
output pool biological replicate are displayed along with averages of each.

3.9 Future Directions

The many targets identified in this screen open up many possibilities for different avenues of future work such as determining the mechanisms by which the validated targets contribute to colonization. However, our murine model of colonization and competition experiments between strains also offers a powerful tool for phenotyping colonization abilities of a range of strains. Further investigation with a curated collection of strains may be able to help us answer whether certain sequence types of *K. pneumoniae* exhibit enhanced colonization abilities. This could direct further efforts into understanding the mechanisms used by these high-risk clones for colonization and explain why they have successfully spread across the globe. Additionally, microbiome sequencing such as metagenomics could shed light on how each *K. pneumoniae* strain interacts with the microbiome. These studies could also assess whether there are certain protective species in the microbiome which can prevent high levels of *K. pneumoniae* carriage and fecal shedding.

One of the most exciting possibilities for the results our transposon insertion screen is the development of therapeutics that may be able to reduce or eliminate colonization with *K. pneumoniae*. We were able to identify and validate several colonization factors shared by three clinical strains, suggesting that our screen was reliable. Our list of shared factors is likely to contain more genes which would validate *in vivo* and thus be good targets for therapeutics which could act on a range of *K. pneumoniae* strains.

The identification of *acrA* as necessary for GI colonization has translational implications. AcrA is the periplasmic subunit of the tripartite efflux pumps that contain TolC and AcrB or AcrD[125]. In *E. coli* and *S. enterica*, these pumps export a large variety of substrates including multiple classes of antibiotics and bile acids[125,144]. We demonstrated that the *acrA* mutant in

CRE-166 was more susceptible to bile (**Figure 30**), suggesting a similar function in *K. pneumoniae*. This efflux pump component was identified as a colonization factor in all three of our strains, and the *acrA* deletion mutants were undetectable at 14 days post-inoculation in most of our competition experiments. Because of its role in antibiotic resistance, the AcrAB efflux pump has been extensively studied, and several small molecule inhibitors are in varying stages of pre-clinical development[148]. We postulate that these inhibitors may have efficacy in preventing or eradicating *K. pneumoniae* GI colonization.

4. Materials & Methods

Bacterial Strains and Cultures

CRE-166, KPN46, and Z4160 are *K. pneumoniae* clinical isolates from Northwestern Memorial Hospital collected between 2013 and 2015. CRE-166 and KPN46 were previously described[113,162], whereas Z4160 was first used in the current study. *E. coli* strain PIR1 was used for cloning, and *E. coli* β3914 (diaminopimelic acid (DAP) auxotroph) was used to mate plasmids into *K. pneumoniae*[163,164].

Bacteria were grown in LB with shaking or on LB agar at 37°C unless otherwise stated. When appropriate, the following antibiotics were added: carbenicillin (100 μg/mL), hygromycin (100 μg/mL), apramycin (50 μg/mL) or tetracycline (10 μg/mL). Additionally, medium for the growth of β3914, was supplemented with 10 μg/mL of DAP.

Murine GI colonization with an intact microbiome

Six- to eight-week-old C57BL/6 mice (Jackson Laboratories) received 10^8 CFU *K. pneumoniae* by gavage unless otherwise indicated. Mice received daily cage changes to minimize coprophagy. CFU were enumerated by homogenization of fecal pellets in PBS with the Benchmark Bead Blaster 24 (Benchmark Scientific) followed by serial dilution and plating on LB agar with carbenicillin.

For treatment with sodium bicarbonate, mice received a sterile 8.5% solution of sodium bicarbonate by gavage prior to a second gavage with *K. pneumoniae*.

For sucrose feeding experiments, *K. pneumoniae* were suspended in 2% sterile sucrose solutions before being fed drop-wise to mice.

Murine GI colonization with vancomycin pre-treatment

Six- to eight-week-old C57BL/6 mice (Jackson Laboratories) received 5 daily intraperitoneal injections of vancomycin (350 mg/kg, Hospira) unless otherwise indicated. For gavage with individual strains, cultures at exponential phase were used to create inocula of 10^8 CFU in 50 μl of PBS. Mice received daily cage changes to minimize coprophagy. CFU were enumerated by homogenization of fecal pellets in PBS with the Benchmark Bead Blaster 24 (Benchmark Scientific) followed by serial dilution and plating on LB agar with carbenicillin.

Transposon mutant screen experiments were performed as above except that frozen aliquots of the transposon libraries were revived for 2 hours in 25 mL of LB at 37°C. CFU in fecal pellets were quantified as above, and DNA was extracted from the homogenates with the Maxwell 16 system.

For competitive colonization experiments, inocula of 1:1 mixtures of a hygromycin-resistant strain and an isogenic apramycin-resistant strain (10^8 CFU each) were created. Fecal CFU burdens were enumerated as above by plating on LB agar with hygromycin or apramycin. Competitive indices (CIs) were calculated as the ratio of apramycin-resistant CFU to hygromycin-resistant CFU, normalized to the input ratio. The Δ*scrY* mutants were not apramycin-resistant, and CFU of Δ*scrY* were calculated by subtracting hygromycin-resistant CFU (parent strain) from carbenicillin-resistant CFU (total *K. pneumoniae*).

Mice were housed in a containment ward of the Center for Comparative Medicine at Northwestern University. Experiments were approved by the Northwestern University Institutional Animal Care and Use Committee in compliance with ethical regulations.

Overlay Assays

Spots of 2 µl of *K. pneumoniae* strains were placed on LB agar plates and allowed to dry completely (at least 15 minutes). Then, 100 µl of *E. coli* TOP10 was spiked into 5 mL of 55°C molten 0.75% LB agar and immediately overlayed on the plates with the *K. pneumoniae* spots. Overlays were allowed to solidify for 15 minutes before incubation overnight at 37°C.

Preparation of Complete Genomes

Genomic DNA was purified using the Maxwell 16 system (Promega). Short-read genome assemblies for CRE-166 (accession number: SAMN17600204) and KPN46 (accession number: SAMN24040734) were published previously[113,162]. For Z4160, a Nextera XT kit (Illumina) was used for library preparation before sequencing on an Illumina MiSeq. Sequences were trimmed using Trimmomatic v0.32 [165], and *de novo* assembly was performed with SPAdes 3.9.1. Contigs were removed if they were shorter than 200 bp or had a mean fold coverage of <5x per base. Long-read nanopore sequencing libraries for CRE-166, KPN46, and Z4160 were prepared using the ligation sequencing kit (SQK-LSK109, Oxford Nanopore) and sequenced on a MinION using a FLO-MIN106 flow cell. Base calling and demultiplexing of sequence reads was performed using Guppy v3.4.5[166]. Hybrid assembly and circularization of Nanopore and Illumina reads were performed using Flye v2.9. Nanopore sequencing errors were corrected by aligning Illumina reads to the assembly using BWA v0.7.17 and using serial rounds of Pilon v1.23. Annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (CRE-166 and KPN46)[167] or Prokka (Z4160)[168].

Identification of shared genes

To identify genes shared between the three *K. pneumoniae* strains, we used the program Spine and defined shared coding sequences as those with >85% sequence homology between

strains[110]. To identify the core genome of *K. pneumoniae*, we examined the assemblies of 323

strains from a previous study[109]. We defined core genome sequences as those present in 95% of

strains and with >85% similarity between strains.

Construction of transposon mutant libraries

A suicide plasmid suitable for Himar1 mariner transposon mutagenesis in highly

antibiotic-resistant strains was generated from pSAM*erm* (gift from G. Pier[169]). pSAM*erm* was

modified by replacement of the erythromycin-resistance cassette with a hygromycin-resistance

cassette (HygR) amplified from the pFLP-hyg plasmid (Addgene plasmid #87831[170]) with

primers KEB306 and 307 (which also added MfeI and XbaI sites respectively) and ligated into

the cloning vector pCR®2.1-TOPO (Invitrogen). (All primers used in this study are listed in

Table 11) The hygromycin-resistance cassette had two internal MmeI sites, which had the

potential to interfere with amplification of transposon-flanking DNA by the Goodman

technique[107]. Thus, these sites were removed as follows. The Thermo Scientific Phusion Site-

Directed Mutagenesis Kit (Catalog number: F541) was used iteratively to perform site-directed

mutagenesis twice (round 1: primers KEB308, 309; round 2: KEB310, 311) to introduce

synonymous changes at amino acids serine 23 and serine 201 (cytosine to adenosine nucleotide

changes for both sites). These changes destroyed the internal MmeI recognition sites but

maintained protein sequence, yielding a modified hygromycin resistance cassette, *hygSDM*. The

functionality of *hygSDM* was confirmed by growth on hygromycin-supplemented LB agar. The

erythromycin cassette from pSAM*erm* was then excised with the restriction enzymes MfeI and

XbaI and replaced with the *hygSDM* resistance cassette from pCR®2.1-TOPO+*hygSDM* which

had the homologous ends added previously with KEB306 and 307. The plasmid was fully

sequenced and transformed into *E. coli* β3914 strain.

For conjugation of pSAM*hygSDM* into the *K. pneumoniae* strains, the recipient strains (CRE-166, Z4160, and KPN46) and the donor strain *E. coli* β3914(pSAM*hygSDM*) were mixed into a 1:2 ratio and resuspended in PBS. The mixture was spotted on LB in 10 μl droplets and allowed to conjugate at 37°C for 1 hour. The spots were resuspended in LB, pooled, and stored in 25% glycerol at -80°C. To select for transconjugants and eliminate β3914, the suspension was thawed and plated on LB agar supplemented with hygromycin (but lacking DAP) and incubated at 37°C overnight. Colonies were scraped and resuspended in LB with 25% glycerol. The resulting library was stored at -80°C.

Arbitrary PCR for library quality control

We performed arbitrary PCR to identify transposon insertion sites in 32 randomly selected colonies from each library for initial quality control. Genomic DNA was extracted with the Maxwell 16 system, and two rounds of nested PCR were performed to amplify the transposon insertion site for subsequent Sanger sequencing.

Round 1

pSAMseq1 (Transposon-specific primer): TGCGAGAGTAGGGAACTGCCAGG

Arbitrary primers components: sequence introduced for Round 2 – random nucleotides – sequence that occurs throughout genome

Arb1: GGCCACGCGTCGACTAGTAC-N_{10}-GATAT

Arb5: GGCCACGCGTCGACTAGTAC-N_{10}-CAAGG

Arb6: GGCCACGCGTCGACTAGTAC-N_{10}-ACGCC

Arb9: GGCCACGCGTCGACTAGTAC-N_{10}-CGACG

Note: Arb5 and Arb6 give bands for most of the mutants. You can start with those two and then perform the PCR with Arb1 or Arb9 for any that do not produce a unique band from Arb5 and Arb6.

Round 2

pSAMseq2 (Transposon-specific primer): CTGTTGTTTGTCGGTGAACGCTCTC

Arb2: GGCCACGCGTCGACTAGTAC

Arbitrary PCR Round 1

9.75ul	Water
1ul	DNA
1.25ul	Arb(1,5,6, or 9)
0.5ul	pSAMseq1
12.5ul	GoTaq Green Master Mix (Promega)
25ul	

Arbitrary PCR Round 1 Cycling Conditions
94°C x 3min
      *94°C x 45S
      30°C x 30s
      72°C x 1:30
      **Repeat 6x from *
      *94°C x 45s
      45°C x 30
      72°C x 2min
      **Repeat 30x from *
72°C x 10min

Arbitrary PCR Round 2

9ul	Water
2.5ul	Product from Round 1
0.5ul	Arb2
0.5ul	pSAMseq2
12.5ul	GoTaq Green Master Mix
25ul	

Cycling Conditions

94°C x 3min

    *94°C x 45s

    55°C x 40s

    72°C x 1:30

    **Repeat 30x from *

72°C x 10min

Gel

Run on 1.5% (w/v) agarose gel.

Extract unique bands for Sanger sequencing with pSAMseq2

Preparation of DNA for transposon insertion sequencing

DNA extracted from fecal pellets was prepared for insertion site sequencing using the method of Kazi and colleagues[171]. Three replicates for each input (technical) and output (biological) pool were prepared, plus one technical replicate of an output sample. DNA was sheared to ~250 bp fragments with the E220 ultrasonicator (Covaris), and poly-C tails were added with terminal deoxynucleotidyl transferase (Promega). A biotinylated primer annealing to the transposon and a second primer annealing to poly-C were used to amplify DNA at the transposon insertion site. After pulldown of biotinylated PCR products with streptavidin beads (New England BioLabs), a second PCR was performed with primers designed to add P5 and P7 capture/sequencing sites and library barcodes. Final library pool concentrations were quantified with Kapa library quantification (Roche) and sequenced on an Illumina MiSeq.

Transposon insertion sequencing data analysis

A modified version of the previously described ESSENTIALS pipeline[122] was used to identify genes necessary for growth in LB and genes that contributed to colonization[172]. Briefly, reads were first processed using Bowtie[173] to trim barcodes and transposon sequences and to

align the results to a reference genome[174]. Reads initially lacking a transposon sequence and reads aligning to the last 10% of a gene's sequence were discarded.

To identify genes that were necessary for growth in LB (referred to as "essential"), ESSENTIALS simulated a "perfect experiment" in which every TA site (the dinucleotide recognized by the Himar1 mariner transposon) received a transposon insertion and generated the same number of reads. Then, the logFC was calculated between these simulated number of reads per gene and those that were actually recovered from sequencing of the input pools. On the resulting density plot of logFC vs number of genes, two peaks were observed, reflecting genes with very few reads because of limited growth and genes with numerous reads because of abundant growth. The local minimum between the two peaks was calculated and designated as the threshold for considering a gene "essential" for growth in LB. These genes were removed from the analysis of the output pools from the colonization experiments. One limitation of the transoposon insertion sequencing approach and use of ESSENTIALS is that sequencing reads aligning to repetitive sequences in the genome cannot be assigned to a particular gene and are discarded. Thus, repetitive regions may be automatically sorted as "essential genes" since they have no assigned reads. For this reason, we removed all genes categorized as "essential" if they had multiple copies in the genome.

To identify colonization factors, ESSENTIALS compiled the ratio of the insertion site read numbers for each gene in the output pools to the insertion site read numbers in the input pool. Genes enriched and depleted in the output pools relative to the input pools were identified using a threshold of -2 or +2 logFC, respectively, and an FDR >0.05.

Pathway Analysis

KEGG identifiers were assigned to all genes in the 3 strains using BlastKOALA, and

KEGG pathways were assigned to each identifier using KEGG Mapper[175,176]. A hypergeometric

test was conducted in R (v4.2.2) to determine which KEGG pathways were enriched for in the

colonization factors identified for each strain.

Isogenic mutant construction

All isogenic mutants in this study except for the Δ*scrY* mutants were created with the

following method to replace the loci with an apramycin-resistance cassette. The 1000 bp regions

upstream and downstream of the target loci for deletion were amplified by PCR with the primers

listed in **Table 11**. These primers were designed to contain overlap sequences with primers used

to amplify the apramycin-resistance cassette from pIJ773[170]. These fragments were ligated

together and into the EcoRI site of pUC18R6K with Gibson assembly (New England BioLabs).

The resulting plasmids were propagated in *E. coli* PIR1. For electroporation into *K. pneumoniae*,

the fragment containing upstream and downstream regions flanking an apramycin-resistance

cassette were amplified by PCR.

For the Δ*scrY* mutants, Gibson assembly did not yield the correct clones. To create the

correct fragment containing the upstream region, apramycin-resistance cassette, and downstream

region, PCR splicing by overlap extension (SOE) was used to first assemble the upstream region

and apramycin-resistance cassette. The appropriate size band was gel purified and assembled

with the downstream region through a second PCR SOE. The appropriate size band was gel-

purified and PCR amplified for electroporation into *K. pneumoniae*.

Lambda red machinery was introduced into *K. pneumoniae* strains by electroporation

with pACBSR, which was maintained by growth at 30°C and supplementation with hygromycin.

The lambda red machinery was induced by growing the *K. pneumoniae* strain in 5mL LB supplemented with hygromycin and 475 µl of 1M arabinose for 3.5 hours at 30°C with shaking. Then, the *K. pneumoniae* were made electrically competent and electroporated with the fragment described above. Cells were recovered in LB for 2.5 hours at 37°C with shaking. Mutants were verified by Sanger sequencing, and pACBSR was removed by inoculating bacteria into LB, growing them overnight at 37°C, plating for single colonies, and screening for plasmid loss by patching onto LB agar supplemented with apramycin or hygromycin.

To create marked parental strains, homologous overhangs were amplified for the Tn7 insertion site in each strain using primers listed in **Table 11**. These were ligated using Gibson cloning to amplified apramycin- or hygromycin- resistance cassettes, as described above. For addition of the hygromycin cassette, a modified version of pACBSR was necessary because pACBSR also contains HygR. The original HygR in pACBSR was replaced with an apramycin-resistance cassette as follows. pACBSR was digested with XhoI and BglII to excise HygR. Primers AprR_F and AprR_R were then used to amplify the apramycin cassette from pIJ773 and add overlap sequences designed to hybridize to the ends of the digested pACBSR plasmid. After ligation with Gibson assembly, the resulting plasmid was named pACBSR*apr*. Transformation into *K. pneumoniae* and curing of pACBSR or pACBSR*apr* was achieved as described above.

For the creation of the Δ*acrA* complemented strain in CRE-166, designated CRE-166Δ*acrA* Tn7::*acrA*, we took advantage of FRT sites flanking the apramycin-resistance cassette within the Δ*acrA* allele. We transformed CRE-166Δ*acrA*::AprR with pFLP-hyg to excise this cassette. Hygromycin-resistant colonies were cultured in LB at 37°C with shaking overnight and plated for single colonies. The colonies were screened for excision of the apramycin-resistance cassette and curing of pFLP-hyg by patching onto LB agar with or without hygromycin or

apramycin. After the apramycin-resistance cassette was flipped out and pFLP-hyg was cured, the resulting strain (designated CRE-166Δ*acrA*) was then transformed with pACBSR. Then, the *acrA* gene along with 142 nucleotides of upstream were PCR amplified from CRE-166 and ligated with a downstream apramycin-resistance cassette between homologous overhangs to the Tn7 site as above. This fragment was transformed into electrocompetent CRE-166Δ*acrA*, and the screening and pACBSR curing process was completed as above. The Δ*tatABCD* complemented strain in CRE-166 was created in the same fashion. The in-frame Δ*scrY* mutants without an apramycin-resistance cassette were created by using and curing pFLP as above.

To create the plasmid pACYC184::*mceAB*, the entire *mceAB* coding sequence and 500 bp upstream were amplified with primers with homologous overhangs to pACYC184 digested with EcoRI and NcoI. Ligation of the fragment into the plasmid was achieved with T4 ligase, and the plasmid was propagated in *E. coli* TOP10. The plasmid was then purified by mini-prep and electroporated into KPN46Δ*mceAB*.

All mutants were confirmed by whole-genome sequencing using the seqWell 96 kit (seqWell) and Illumina NovaSeq 6000 or Illumina MiSeq.

Primers used in this study

Table 11 Primer Sequences

Name	Sequence	Purpose
KEB306	<u>CGTGCAATTG</u>CGCGGAACCCCTATTTG	Amplification of hygromycin-resistance cassette from pFLP-hyg for site-directed mutagenesis, underlined portions introduced MfeI site
KEB307	<u>CTAGTCTAGA</u>CTATTCCTTTGCCCTCGG	Amplification of hygromycin-resistance cassette from pFLP-hyg for site-directed mutagenesis, underlined portions introduced XbaI site
KEB308	5`Phos - TTCGACAGCGTCTC**a**GACCTGATGCAGC	Round 1 of site-directed mutagenesis of hygromycin-resistance cassette from pFLP-hyg, red "a" denotes site of introduction of SDM nucleotide
KEB309	5`Phos - CTTTTCGATCAGAAACTTCTCGACAGACGTCGC	Round 1 of site-directed mutagenesis of hygromycin-resistance cassette from pFLP-hyg
KEB310	5`Phos - GCGGATTTCGGCTC**a**AACAATGTCCTGAC	Round 2 of site-directed mutagenesis of hygromycin-resistance cassette from pFLP-hyg, red "a" denotes site of introduction of SDM nucleotide
KEB311	5`Phos - GTGCACGAGGTGCCGGACTTCGG	Round 2 of site-directed mutagenesis of hygromycin-resistance cassette from pFLP-hyg
pSAMseq1	TGCGAGAGTAGGGAACTGCCAGG	Round 1 of Arbitrary PCR
Arb1	GGCCACGCGTCGACTAGTAC-N$_{10}$-GATAT	Round 1 of Arbitrary PCR
Arb5	GGCCACGCGTCGACTAGTAC-N$_{10}$-CAAGG	Round 1 of Arbitrary PCR
Arb6	GGCCACGCGTCGACTAGTAC-N$_{10}$-ACGCC	Round 1 of Arbitrary PCR
Arb9	GGCCACGCGTCGACTAGTAC-N$_{10}$-CGACG	Round 1 of Arbitrary PCR
pSAMseq2	CTGTTGTTTGTCGGTGAACGCTCTC	Round 2 of Arbitrary PCR
Arb2	GGCCACGCGTCGACTAGTAC	Round 2 of Arbitrary PCR
olj510-Biotin	Biotin-GATGGCCTTTTTGCGTTTCTACCTGCAGGGCGCG	Transposon library sequencing preparation (anneals to transposon)
olj376	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGGGGGGGGGGGGG	Transposon library sequencing preparation (anneals to poly-C tail)

olj511	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTNNNNNGGGGA CTTATCATCCAACCTGTTAG	Transposon library sequencing preparation (nested to transposon and adds P5 adapter)
BC01	CAAGCAGAAGACGGCATACGAG AT**CGTGAT**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC02	CAAGCAGAAGACGGCATACGAG AT**ACATCG**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC03	CAAGCAGAAGACGGCATACGAG AT**GCCTAA**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC04	CAAGCAGAAGACGGCATACGAG AT**TGGTCA**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC05	CAAGCAGAAGACGGCATACGAG AT**CACTGT**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC06	CAAGCAGAAGACGGCATACGAG AT**ATTGGC**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
BC07	CAAGCAGAAGACGGCATACGAG AT**GATCTG**GTGACTGGAGTTCA GACGTGTG	Transposon library sequencing preparation (nested to olj376 and adds P7 adapter and **barcode**)
acrA_1	CCAAGCTTCTCGAGGACTGGGCT GAAAAGGCCAATAAAAC	Preparation of linear DNA fragment to generate ΔacrA
acrA_2	AGCCTACACATGATATGTAAACC TCGAGTGTCCAATTTC	Preparation of linear DNA fragment to generate ΔacrA
acrA_3	ACTCGAGGTTTACATATCATGTG TAGGCTGGGCTGCTTC	Preparation of linear DNA fragment to generate ΔacrA
acrA_4	CGGCTCCTGTTTAAGTTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate ΔacrA
acrA_5	GATCCCCGGAATTAACTTAAACA GGAGCCGTTAAGACATG	Preparation of linear DNA fragment to generate ΔacrA
acrA_6	CGGGCTGCAGGAATTTCATACGG GTAAACGATCTTCATCC	Preparation of linear DNA fragment to generate ΔacrA
carAB_1	CCAAGCTTCTCGAGGGGCCTGCG CCTGCAGCGTTTCATCA	Preparation of linear DNA fragment to generate ΔcarAB
carAB_2	CAGCCTACACATGATAAGCTTCC CGGCCTGGCCGGATAGC	Preparation of linear DNA fragment to generate ΔcarAB
carAB_3	CAGGCCGGGAAGCTTATCATGTG TAGGCTGGAGCTGCTTC	Preparation of linear DNA fragment to generate ΔcarAB
carAB_4	TCTCTGGAGGATGTTTTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate ΔcarAB

carAB_5	GAATCCCCGGAATTAAAACATCC TCCAGAGAATATCCACTC	Preparation of linear DNA fragment to generate ΔcarAB
carAB_6	CGGGCTGCAGGAATTATTTCGAC GTGTTGATCGATTTTAC	Preparation of linear DNA fragment to generate ΔcarAB
tatABCD_1	CCAAGCTTCTCGAGGTCGGCATA TTCAGCACTTGCTTGAA	Preparation of linear DNA fragment to generate ΔtatABCD
tatABCD_2	CAGCCTACACATGATTTAGATTT TCTGGAAGTCGGTATTT	Preparation of linear DNA fragment to generate ΔtatABCD
tatABCD_3	TTCCAGAAAATCTAAATCATGTG TAGGCTGGAGCTGCTTC	Preparation of linear DNA fragment to generate ΔtatABCD
tatABCD_4	CATAGGGGAACGTGTTTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate ΔtatABCD
tatABCD_5	GATCCCCGGAATTAAACACGTTC CCCTATGACAGATGATG	Preparation of linear DNA fragment to generate ΔtatABCD
tatABCD_6	CGGGCTGCAGGAATTATTTACGG TATCCCGGTGTCCGATG	Preparation of linear DNA fragment to generate ΔtatABCD
hha_1	CCAAGCTTCTCGAGGAAATCGG AAGATATAGAGCACAGCC	Preparation of linear DNA fragment to generate Δhha
hha_2	AGCCTACACATGATGAATTCCAC CTTTTGATTGTAATAAT	Preparation of linear DNA fragment to generate Δhha
hha_3	TCAAAAGGTGGAATTCATCATGT GTAGGCTGGAGCTGCTT	Preparation of linear DNA fragment to generate Δhha
hha_4	CGTAATACGCGTAAATTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate Δhha
hha_5	GATCCCCGGAATTAATTTACGCG TATTACGTGGTCTGCTG	Preparation of linear DNA fragment to generate Δhha
hha_6	CGGGCTGCAGGAATTTCTACCAG CGTTTCAGAGCGCAGCA	Preparation of linear DNA fragment to generate Δhha
malT_1	CCAAGCTTCTCGAGGGCGCGATC CGCAGAGCGGTATTTAT	Preparation of linear DNA fragment to generate ΔmalT
malT_2	AGCCTACACATGATGCGACGCC GTTGCCGTTTAGCACAGC	Preparation of linear DNA fragment to generate ΔmalT
malT_3	ACGGCAACGGCGTCGCATCATGT GTAGGCTGGAGCTGCTT	Preparation of linear DNA fragment to generate ΔmalT
malT_4	GCGAAATGTAGAACTTTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate ΔmalT
malT_5	GATCCCCGGAATTAAAGTTCTAC ATTTCGCTGTGCAAGAG	Preparation of linear DNA fragment to generate ΔmalT
malT_6	CGGGCTGCAGGAATTCGGTGAG GTTGACATCGTAGCCTTT	Preparation of linear DNA fragment to generate ΔmalT
scrY_1	CCAAGCTTCTCGAGGTATGAATG GAAAAATCTGGGTACTC	Preparation of linear DNA fragment to generate ΔscrY
scrY_2	GATCCCCGGAATTAACATGTTGG TGACATCCAAAGGTAAA	Preparation of linear DNA fragment to generate ΔscrY
scrY_3	GATGTCACCAACATGTTAATTCC GGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate ΔscrY

scrY_4	CCCCGGTGGCCGTCAATCATGTGTAGGCTGGAGCTGCTTC	Preparation of linear DNA fragment to generate ΔscrY
scrY_5	CAGCCTACACATGATTGACGGCCACCGGGGCGACAGGGTA	Preparation of linear DNA fragment to generate ΔscrY
scrY_6	CGGGCTGCAGGAATTTAGAGGCCGCCGAACAGCAGGCCCG	Preparation of linear DNA fragment to generate ΔscrY
Tn7AprR_1	CCAAGCTTCTCGAGGTAGCGAAGTTGTCTTCAGTCGGGCT	Preparation of linear DNA fragment to generate Tn7::AprR
Tn7AprR_2	GCTCCAGCCTACACATGATGGCACCTGCTGTTTTCCATCG	Preparation of linear DNA fragment to generate Tn7::AprR
Tn7AprR_3	GAAAACAGCAGGTGCCATCATGTGTAGGCTGGAGCTGCTT	Preparation of linear DNA fragment to generate Tn7::AprR
Tn7AprR_4	CGGCCCTTTTTATTTTTAATTCCGGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate Tn7::AprR
Tn7AprR_5	GATCCCCGGAATTAAAAATAAAAAGGGCCGCCGTCGGCAG	Preparation of linear DNA fragment to generate Tn7::AprR
Tn7AprR_6	CGGGCTGCAGGAATTTCAGATAGTTGCCTGCGGGACTTCC	Preparation of linear DNA fragment to generate Tn7::AprR
AprR_F	CTGCCTTAAAAAAACTTATGAGCTCAGCCAATCGACTGGC	Amplification of apramycin-resistance cassette from pIJ773 to insert into pACBSR digested with XhoI and BglII
AprR_R	GCACTTTGCAGATCCGTTGAGCACCGCCAGGTGCGAATAA	Amplification of apramycin-resistance cassette from pIJ773 to insert into pACBSR
Tn7HygR_1	same as Tn7AprR_1	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7HygR_2	AAACAAATAGGGGTTCCGCGGCACCTGCTGTTTTCCATCG	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7HygR_3	GAAAACAGCAGCTGCCGCGGAACCCCTATTTGTTTATTTT	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7HygR_4	CGGCCCTTTTTATTTCTATTCCTTTGCCCTCGGACGAGTG	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7HygR_5	AGGGCAAAGGAATAGAAATAAAAAGGGCCGCCGTCGGCAG	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7HygR_6	same as Tn7AprR_6	Preparation of linear DNA fragment to generate Tn7::HygR
Tn7acrA_1	same as Tn7AprR_1	Preparation of linear DNA fragment to generate Tn7::acrA
Tn7acrA_2	TAAAGTCATTAACCTTTAATTCCGGGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate Tn7::acrA
Tn7acrA_3	GATCCCCGGAATTAAAGGTTAATGACTTTACAGAGGTTACGTT	Preparation of linear DNA fragment to generate Tn7::acrA
Tn7acrA_4	CGGCCCTTTTTATTTTTAAGACTTGGTTTGTTCTGATGGC	Preparation of linear DNA fragment to generate Tn7::acrA
Tn7acrA_5	ACAAACCAAGTCTTAAAAATAAAAAGGGCCGCCGTCGGCAG	Preparation of linear DNA fragment to generate Tn7::acrA

Tn7acrA_6	same as Tn7AprR_6	Preparation of linear DNA fragment to generate Tn7::acrA
Tn7tat_1	same as Tn7AprR_1	Preparation of linear DNA fragment to generate Tn7::tatABCD
Tn7tat_2	AGAGGCGAACCGATTAATTCCG GGGATCCGTCGACCTG	Preparation of linear DNA fragment to generate Tn7::tatABCD
Tn7tat_3	GATCCCCGGAATTAATCGGTTCG CCTCTTCATGACGCAGG	Preparation of linear DNA fragment to generate Tn7::tatABCD
Tn7tat_4	CGGCCCTTTTTATTTCTAAACGC CGTTAATGTCAACGCCA	Preparation of linear DNA fragment to generate Tn7::tatABCD
Tn7tat_5	ATTAACGGCGTTTAGAAATAAA AAGGGCCGCCGTCGGCAG	Preparation of linear DNA fragment to generate Tn7::tatABCD
Tn7tat_6	same as Tn7AprR_6	Preparation of linear DNA fragment to generate Tn7::tatABCD

Growth curves

Overnight cultures of strains were diluted to an OD_{600} of 0.1 in fresh LB. Then, 200 µl

was inoculated into an optically clear 96 well plate. The SpectraMax iD3 was used to incubate

the plates at 37°C with shaking and OD_{600} readings every hour for 24 hours.

Bile assays

Dehydrated ox bile was resuspended in water (10% w/v, Sigma) and filtered through a

0.2µm filter. Strains were grown to an OD_{600} of 1.0, and 1 mL of each culture was pelleted and

resuspended in 1 mL PBS. Each strain was inoculated (100 µl) into 900 µl of 10% ox bile, PBS,

or LB and incubated with shaking at 37°C. At 0, 2, 4, and 24 hours, 20 µl aliquots were removed,

serially diluted, and plated on LB agar for CFU enumeration.

In vitro competition assays

Overnight cultures of strains were diluted to an OD_{600} of 1.0, and 50 µl of each strain was

inoculated into 900 µl of LB and incubated with shaking at 37°C. At 0, 2, and 24 hours, aliquots

were removed and plated on LB agar supplemented with the appropriate antibiotics for CFU

enumeration.

REFERENCES

Figures 1 and 32 were created using biorender.com.

1   Bagley, S. T. Habitat association of *Klebsiella* species. *Infection Control : IC* **6**, 52–58, doi:10.1017/s0195941700062603 (1985).

2   Podschun, R., Pietsch, S., Höller, C. & Ullmann, U. Incidence of *Klebsiella* Species in Surface Waters and Their Expression of Virulence Factors. *Applied and Environmental Microbiology* **67**, 3325, doi:10.1128/aem.67.7.3325-3327.2001 (2001).

3   da Silva, J. M., Menezes, J., Marques, C. & Pomba, C. F. Companion Animals—An Overlooked and Misdiagnosed Reservoir of Carbapenem Resistance. *Antibiotics* **11**, doi:10.3390/antibiotics11040533 (2022).

4   Davis, G. S. & Price, L. B. Recent Research Examining Links Among *Klebsiella pneumoniae* from Food, Food Animals, and Human Extraintestinal Infections. *Current Environmental Health Reports* **3**, 128–135, doi:10.1007/s40572-016-0089-9 (2016).

5   Rathnasingh, C., Park, J. M., Kim, D.-K., Song, H. & Chang, Y. K. Metabolic engineering of *Klebsiella pneumoniae* and *in silico* investigation for enhanced 2,3-butanediol production. *Biotechnology Letters* **38**, 975–982, doi:10.1007/s10529-016-2062-y (2016).

6   Chen, Z., Wu, Y., Huang, J. & Liu, D. Metabolic engineering of *Klebsiella pneumoniae* for the *de novo* production of 2-butanol as a potential biofuel. *Bioresource Technology* **197**, 260–265, doi:10.1016/j.biortech.2015.08.086 (2015).

7   Constantinides, B. *et al.* Genomic surveillance of *Escherichia coli* and *Klebsiella* spp. in hospital sink drains and patients. *Microbial Genomics* **6**, mgen000391., doi:10.1099/mgen.0.000391 (2020).

8   Córdova-Espinoza, M. G. *et al.* Isolation and Identification of Multidrug-Resistant *Klebsiella pneumoniae* Clones from the Hospital Environment. *Pathogens* **12**, doi:10.3390/pathogens12050634 (2023).

9   Magill, S. S. *et al.* Multistate Point-Prevalence Survey of Health Care–Associated Infections. *New England journal of medicine* **370**, 1198, doi:10.1056/NEJMoa1306801 (2014).

10   Gomez-Simmonds, A. *et al.* Population Structure of Klebsiella pneumoniae Causing Bloodstream Infections at a New York City Tertiary Care Hospital: Diversification of Multidrug-Resistant Isolates. *Journal of Clinical Microbiology* **53**, 2060, doi:10.1128/jcm.03455-14 (2015).

11   Chang, D., Sharma, L., Dela Cruz, C. S. & Zhang, D. Clinical Epidemiology, Risk Factors, and Control Strategies of Klebsiella pneumoniae Infection. *Frontiers in Microbiology* **12**, 750662, doi:10.3389/fmicb.2021.750662 (2021).

12   CDC.   (2020).

13   Braykov, N. P., Eber, M. R., Klein, E. Y., Morgan, D. J. & Laxminarayan, R. Trends in Resistance to Carbapenems and Third-Generation Cephalosporins among Clinical Isolates of Klebsiella pneumoniae in the United States, 1999–2010. *Infection Control & Hospital Epidemiology* **34**, 259–268, doi:10.1086/669523 (2013).

14   Karlowsky, J. A. *et al.* Prevalence of ESBL non-CRE Escherichia coli and Klebsiella pneumoniae among clinical isolates collected by the SMART global surveillance programme from 2015 to 2019. *International Journal of Antimicrobial Agents* **59**, 106535., doi:10.1016/j.ijantimicag.2022.106535 (2022).

15   Porreca, A. M., Sullivan, K. V. & Gallagher, J. C. The Epidemiology, Evolution, and Treatment of KPC-Producing Organisms. *Current Infectious Disease Reports* **20**, 13–12, doi:10.1007/s11908-018-0617-x (2018).

16   Elemam, A., Rahimian, J. & Mandell, W. Infection with Panresistant Klebsiella pneumoniae: A Report of 2 Cases and a Brief Review of the Literature. *Clinical Infectious Diseases* **49**, 271–274, doi:10.1086/600042 (2009).

17   Krapp, F., Ozer, E. A., Qi, C. & Hauser, A. R. Case Report of an Extensively Drug-Resistant Klebsiella pneumoniae Infection With Genomic Characterization of the Strain and Review of Similar Cases in the United States. *Open Forum Infectious Diseases* **5**, doi:10.1093/ofid/ofy074 (2018).

18   Hauck, C. *et al.* Spectrum of Excess Mortality due to Carbapenem-Resistant *Klebsiella pneumoniae* Infections. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **22**, 513, doi:10.1016/j.cmi.2016.01.023 (2016).

19   Mathers, A. J., Peirano, Gisele, Pitout, Johann D.D. in *Advances in Applied Microbiology* Vol. 90 (ed GEOFFREY MICHAEL GADD SIMA SARIASLANI) Ch. 4, 109–154 (Academic Press, 2015).

20   de Lagarde, M., Vanier, G., Arsenault, J. & Fairbrother, J. M. High Risk Clone: A Proposal of Criteria Adapted to the One Health Context with Application to Enterotoxigenic *Escherichia coli* in the Pig Population. *Antibiotics* **10**, doi:10.3390/antibiotics10030244 (2021).

21   Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements
        Associated with Antimicrobial Resistance. *Clinical Microbiology Reviews* **31**,
        doi:10.1128/cmr.00088-17 (2018).

22   D'Andrea, M. M., Arena, F., Pallecchi, L. & Rossolini, G. M. CTX-M-type β-lactamases:
        a successful story of antibiotic resistance. *International Journal of Medical Microbiology
        : IJMM* **303**, 305–317, doi:10.1016/j.ijmm.2013.02.008 (2013).

23   Lam, M. M. C. *et al.* A genomic surveillance framework and genotyping tool for
        Klebsiella pneumoniae and its related species complex. *Nature Communications* **12**, 1–
        16, doi:10.1038/s41467-021-24448-3 (2021).

24   Diancourt, L., Passet, V., Verhoef, J., Grimont, P. A. D. & Brisse, S. Multilocus
        Sequence Typing of *Klebsiella pneumoniae* Nosocomial Isolates. *Journal of Clinical
        Microbiology* **43**, 4178, doi:10.1128/jcm.43.8.4178-4182.2005 (2005).

25   Cantón, R. & Coque, T. M. The CTX-M beta-lactamase pandemic. *Current Opinion in
        Microbiology* **9**, 466–475, doi:10.1016/j.mib.2006.08.011 (2006).

26   Lascols, C. *et al.* Using nucleic acid microarrays to perform molecular epidemiology and
        detect novel β-lactamases: a snapshot of extended-spectrum β-lactamases throughout the
        world. *Journal of Clinical Microbiology* **50**, 1632–1639, doi:10.1128/jcm.06115-11
        (2012).

27   Cerqueira, G. C. *et al.* Multi-institute analysis of carbapenem resistance reveals
        remarkable diversity, unexplained mechanisms, and limited clonal outbreaks.
        Proceedings of the National Academy of Sciences of the United States of America **114**,
        1135, doi:10.1073/pnas.1616248114 (2017).

28   Pitout, J. D. D., Nordmann, P. & Poirel, L. Carbapenemase-Producing *Klebsiella pneumoniae*, a Key Pathogen Set for Global Nosocomial Dominance. *Antimicrobial Agents and Chemotherapy* **59**, 5873, doi:10.1128/aac.01019-15 (2015).

29   Marsh, J. W. *et al.* Evolution of Outbreak-Causing Carbapenem-Resistant *Klebsiella pneumoniae* ST258 at a Tertiary Care Hospital over 8 Years. *mBio* **10**, doi:10.1128/mBio.01945-19 (2019).

30   Wyres, K. L., Lam, M. M. C. & Holt, K. E. Population genomics of *Klebsiella pneumoniae*. *Nature Reviews Microbiology* **18**, 344–359, doi:10.1038/s41579-019-0315-1 (2020).

31   Struve, C. *et al.* Mapping the Evolution of Hypervirulent *Klebsiella pneumoniae*. *mBio* **6**, doi:10.1128/mBio.00630-15 (2015).

32   Yu, W.-L. *et al.* Association between rmpA and magA Genes and Clinical Syndromes Caused by Klebsiella pneumoniae in Taiwan. *Clinical Infectious Diseases* **42**, 1351–1358, doi:10.1086/503420 (2006).

33   Krapp, F., Morris, A. R., Ozer, E. A. & Hauser, A. R. Virulence Characteristics of Carbapenem-Resistant Klebsiella pneumoniae Strains from Patients with Necrotizing Skin and Soft Tissue Infections. *Scientific Reports* **7**, doi:10.1038/s41598-017-13524-8 (2017).

34   Fierer, J., Walls, L. & Chu, P. Recurring Klebsiella pneumoniae Pyogenic Liver Abscesses in a Resident of San Diego, California, Due to a K1 Strain Carrying the Virulence Plasmid. *Journal of Clinical Microbiology* **49**, 4371, doi:10.1128/jcm.05658-11 (2011).

35   Zhang, Y. *et al.* Emergence of a hypervirulent carbapenem-resistant Klebsiella pneumoniae isolate from clinical infections in China. *Journal of Infection* **71**, 553–560, doi:10.1016/j.jinf.2015.07.010 (2015).

36   Conlan, S., Kong, H. H. & Segre, J. A. Species-Level Analysis of DNA Sequence Data from the NIH Human Microbiome Project. *PLoS ONE* **7**, doi:10.1371/journal.pone.0047075 (2012).

37   Yasir, M. *et al.* Analysis of the nasopharyngeal microbiome and respiratory pathogens in COVID-19 patients from Saudi Arabia. *Journal of Infection and Public Health* **16**, 680, doi:10.1016/j.jiph.2023.03.001 (2023).

38   Ciptaningtyas, V. R. *et al.* Bacterial colonization of the upper airways of children positive and negative for SARS-CoV-2 during the COVID-19 pandemic. *BMC Infectious Diseases* **22**, doi:10.1186/s12879-022-07851-z (2022).

39   Farida, H. *et al.* Nasopharyngeal Carriage of *Klebsiella pneumoniae* and Other Gram-Negative Bacilli in Pneumonia-Prone Age Groups in Semarang, Indonesia. *Journal of Clinical Microbiology* **51**, 1614, doi:10.1128/jcm.00589-13 (2013).

40   Dao, T. T. *et al. Klebsiella pneumoniae* Oropharyngeal Carriage in Rural and Urban Vietnam and the Effect of Alcohol Consumption. *PLoS ONE* **9**, doi:10.1371/journal.pone.0091999 (2014).

41   Gorrie, C. L. *et al.* Gastrointestinal Carriage Is a Major Reservoir of *Klebsiella pneumoniae* Infection in Intensive Care Patients. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **65**, 208, doi:10.1093/cid/cix270 (2017).

42   Martin, R. M. *et al.* Molecular Epidemiology of Colonizing and Infecting Isolates of *Klebsiella pneumoniae*. *mSphere* **1**, doi:10.1128/mSphere.00261-16 (2016).

43   Tesfa, T., Mitiku, H., Edae, M. & Assefa, N. Prevalence and incidence of carbapenem-resistant *K. pneumoniae* colonization: systematic review and meta-analysis. *Systematic Reviews* **11**, doi:10.1186/s13643-022-02110-3 (2022).

44   Conlan, S. *et al.* Plasmid Dynamics in KPC-Positive *Klebsiella pneumoniae* during Long-Term Patient Colonization. *mBio* **7**, doi:10.1128/mBio.00742-16 (2016).

45   Giannella, M. *et al.* Risk factors for carbapenem-resistant *Klebsiella pneumoniae* bloodstream infection among rectal carriers: a prospective observational multicentre study. *Clinical Microbiology and Infection* **20**, 1357–1362, doi:10.1111/1469-0691.12747 (2014).

46   Jiao, Y. *et al.* Risk factors for carbapenem-resistant Klebsiella pneumoniae infection/colonization and predictors of mortality: a retrospective study. *Pathogens and Global Health* **109**, 68–74, doi:10.1179/2047773215y.0000000004 (2015).

47   Papadimitriou-Olivgeris, M. *et al.* Risk factors for KPC-producing Klebsiella pneumoniae enteric colonization upon ICU admission. *Journal of Antimicrobial Chemotherapy* **67**, 2976–2981, doi:10.1093/jac/dks316 (2012).

48   Gagliotti, C. *et al.* Risk factors for colonization with carbapenemase-producing Klebsiella pneumoniae in hospital: a matched case-control study. *American Journal of Infection Control* **42**, 1006–1008, doi:10.1016/j.ajic.2014.05.028 (2014).

49   Sequeira, R. P., McDonald, J. A. K., Marchesi, J. R. & Clarke, T. B. Commensal Bacteroidetes protect against *Klebsiella pneumoniae* colonization and transmission

through IL-36 signalling. *Nature Microbiology* **5**, 304–313, doi:10.1038/s41564-019-0640-1 (2020).

50   Tiri, B. *et al.* Antimicrobial Stewardship Program, COVID-19, and Infection Control: Spread of Carbapenem-Resistant Klebsiella Pneumoniae Colonization in ICU COVID-19 Patients. What Did Not Work? *Journal of Clinical Medicine* **9**, doi:10.3390/jcm9092744 (2020).

51   Sonnenburg, J. L., Chen, C. T. L. & Gordon, J. I. Genomic and Metabolic Studies of the Impact of Probiotics on a Model Gut Symbiont and Host. *PLoS Biology* **4**, doi:10.1371/journal.pbio.0040413 (2006).

52   de Sablet, T. *et al.* Human Microbiota-Secreted Factors Inhibit Shiga Toxin Synthesis by Enterohemorrhagic *Escherichia coli* O157:H7. *Infection and Immunity* **77**, 783, doi:10.1128/iai.01048-08 (2009).

53   Gantois, I. *et al.* Butyrate Specifically Down-Regulates *Salmonella* Pathogenicity Island 1 Gene Expression. *Applied and Environmental Microbiology* **72**, 946, doi:10.1128/aem.72.1.946-949.2006 (2006).

54   Sassone-Corsi, M. *et al.* Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* **540**, 280–283, doi:10.1038/nature20557 (2016).

55   Kamada, N. *et al.* Regulated Virulence Controls the Ability of a Pathogen to Compete with the Gut Microbiota. *Science (New York, N.Y.)* **336**, 1325, doi:10.1126/science.1222195 (2012).

56   Young, T. M. *et al.* Animal Model To Study *Klebsiella pneumoniae* Gastrointestinal Colonization and Host-to-Host Transmission. *Infection and Immunity* **88**, doi:10.1128/iai.00071-20 (2020).

57   Sim, C. K. *et al.* A mouse model of occult intestinal colonization demonstrating antibiotic-induced outgrowth of carbapenem-resistant Enterobacteriaceae. *Microbiome* **10**, doi:10.1186/s40168-021-01207-6 (2022).

58   Struve, C., Forestier, C. & Krogfelt, K. A. 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, doi:10.1099/mic.0.25833-0 (2003).

59   Jung, H.-J. *et al.* Genome-Wide Screening for Enteric Colonization Factors in Carbapenem-Resistant ST258 *Klebsiella pneumoniae*. *mBio* **10**, doi:10.1128/mBio.02663-18 (2019).

60   Lau, H. Y., Huffnagle, G. B. & Moore, T. A. Host and microbiota factors that control *Klebsiella pneumoniae* mucosal colonization in mice. *Microbes and infection / Institut Pasteur* **10**, 1283, doi:10.1016/j.micinf.2008.07.040 (2008).

61   Perez, F., Pultz, M. J., Endimiani, A., Bonomo, R. A. & Donskey, C. J. Effect of antibiotic treatment on establishment and elimination of intestinal colonization by KPC-producing *Klebsiella pneumoniae* in mice. *Antimicrobial Agents and Chemotherapy* **55**, 2585–2589, doi:10.1128/aac.00891-10 (2011).

62   Hoyen, C. K., Pultz, N. J., Paterson, D. L., Aron, D. C. & Donskey, C. J. Effect of parenteral antibiotic administration on establishment of intestinal colonization in mice by *Klebsiella pneumoniae* strains producing extended-spectrum beta-lactamases. *Antimicrobial Agents and Chemotherapy* **47**, 3610–3612, doi:10.1128/aac.47.11.3610-3612.2003 (2003).

63   Pultz, M. J., Nerandzic, M. M., Stiefel, U. & Donskey, C. J. Emergence and acquisition of fluoroquinolone-resistant gram-negative bacilli in the intestinal tracts of mice treated with fluoroquinolone antimicrobial agents. *Antimicrobial Agents and Chemotherapy* **52**, 3457–3460, doi:10.1128/aac.00117-08 (2008).

64   Pultz, N. J., Stiefel, U. & Donskey, C. J. Effects of daptomycin, linezolid, and vancomycin on establishment of intestinal colonization with vancomycin-resistant enterococci and extended-spectrum-beta-lactamase-producing *Klebsiella pneumoniae* in mice. *Antimicrobial Agents and Chemotherapy* **49**, 3513–3516, doi:10.1128/aac.49.8.3513-3516.2005 (2005).

65   Le Guern, R. *et al.* Impact of the Timing of Antibiotic Administration on Digestive Colonization with Carbapenemase-Producing Enterobacteriaceae in a Murine Model. *Antimicrobial Agents and Chemotherapy* **63**, 00360–00319, doi:10.1128/aac.00360-19 (2019).

66   Calderon-Gonzalez, R. *et al.* Modelling the Gastrointestinal Carriage of *Klebsiella pneumoniae* Infections. *MBio* **e0312122.**, doi:10.1128/mbio.03121-22 (2023).

67   Mike, L. A. *et al.* A systematic analysis of hypermucoviscosity and capsule reveals distinct and overlapping genes that impact *Klebsiella pneumoniae* fitness. *PLoS Pathogens* **17**, e1009376., doi:10.1371/journal.ppat.1009376 (2021).

68   Snitkin, E. S. *et al.* Tracking a Hospital Outbreak of Carbapenem-Resistant *Klebsiella pneumoniae* with Whole-Genome Sequencing. *Science translational medicine* **4**, 148ra116, doi:10.1126/scitranslmed.3004129 (2012).

69   Ramage, B. *et al.* Comprehensive Arrayed Transposon Mutant Library of *Klebsiella pneumoniae* Outbreak Strain KPNIH1. *J Bacteriol* **199**, doi:10.1128/jb.00352-17 (2017).

70   Bleul, T. *et al.* Different Innate Immune Responses in BALB/c and C57BL/6 Strains following Corneal Transplantation. *Journal of Innate Immunity* **13**, 49–59, doi:10.1159/000509716 (2021).

71   Sellers, R. S., Clifford, C. B., Treuting, P. M. & Brayton, C. Immunological Variation Between Inbred Laboratory Mouse Strains: Points to Consider in Phenotyping Genetically Immunomodified Mice. *Veterinary Pathology* **49**, 32–43, doi:10.1177/0300985811429314 (2011).

72   Shu, H.-Y. *et al.* Genetic diversity of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* clinical isolates. *Microbiology* **155**, 4170–4183, doi:10.1099/mic.0.029017-0 (2009).

73   Domenico, P., Salo, R. J., Cross, A. S. & Cunha, B. A. Polysaccharide capsule-mediated resistance to opsonophagocytosis in Klebsiella pneumoniae. *Infection and Immunity* **62**, 4495 (1994).

74   Alvarez, D., Merino, S., Tomás, J. M., Benedí, V. J. & Albertí, S. Capsular polysaccharide is a major complement resistance factor in lipopolysaccharide O side chain-deficient Klebsiella pneumoniae clinical isolates. *Infection and Immunity* **68**, 953–955, doi:10.1128/iai.68.2.953-955.2000 (2000).

75   Campos, M. A. *et al.* Capsule polysaccharide mediates bacterial resistance to antimicrobial peptides. *Infection and Immunity* **72**, 7107–7114, doi:10.1128/iai.72.12.7107-7114.2004 (2004).

76   Lawlor, M. S., Handley, S. A. & Miller, V. L. Comparison of the Host Responses to Wild-Type and cpsB Mutant *Klebsiella pneumoniae* Infections. *Infection and Immunity* **74**, 5402, doi:10.1128/iai.00244-06 (2006).

77   Fang, C.-T., Chuang, Y.-P., Shun, C.-T., Chang, S.-C. & Wang, J.-T. A novel virulence gene in Klebsiella pneumoniae strains causing primary liver abscess and septic metastatic complications. *Journal of Experimental Medicine* **199**, 697–705, doi:10.1084/jem.20030857 (2004).

78   Favre-Bonté, S., Licht, T. R., Forestier, C. & Krogfelt, K. A. Klebsiella pneumoniae capsule expression is necessary for colonization of large intestines of streptomycin-treated mice. *Infection and Immunity* **67**, 6152–6156, doi:10.1128/iai.67.11.6152-6156.1999 (1999).

79   Tan, Y. H., Chen, Y., Chu, W. H. W., Sham, L.-T. & Gan, Y.-H. Cell envelope defects of different capsule-null mutants in K1 hypervirulent *Klebsiella pneumoniae* can affect bacterial pathogenesis. *Molecular Microbiology* **113**, 889–905, doi:10.1111/mmi.14447 (2020).

80   Brisse, S. *et al. wzi* Gene Sequencing, a Rapid Method for Determination of Capsular Type for *Klebsiella* Strains. *Journal of Clinical Microbiology* **51**, 4073, doi:10.1128/jcm.01924-13 (2013).

81   Rahn, A., Beis, K., Naismith, J. H. & Whitfield, C. A Novel Outer Membrane Protein, Wzi, Is Involved in Surface Assembly of the *Escherichia coli* K30 Group 1 Capsule. *J Bacteriol* **185**, 5882, doi:10.1128/jb.185.19.5882-5890.2003 (2003).

82   Wyres, K. L. *et al.* Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microbial Genomics* **2**, doi:10.1099/mgen.0.000102 (2016).

83   Lam, M. M. C., Wick, R. R., Judd, L. M., Holt, K. E. & Wyres, K. L. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. *Microbial Genomics* **8**, doi:10.1099/mgen.0.000800 (2022).

84   Hsieh, P.-F. *et al. Klebsiella pneumoniae* Peptidoglycan-Associated Lipoprotein and
      Murein Lipoprotein Contribute to Serum Resistance, Antiphagocytosis, and
      Proinflammatory Cytokine Stimulation. *Journal of Infectious Diseases* **208**, 1580–1589,
      doi:10.1093/infdis/jit384 (2013).

85   Klemm, P. & Schembri, M. A. Fimbrial surface display systems in bacteria: from
      vaccines to random libraries. *Microbiology (Reading, England)* **146**, 12,
      doi:10.1099/00221287-146-12-3025 (2000).

86   Tarkkanen, A.-M., Westerlund-Wikström, B., Erkkilä, L. & Korhonen, T. K.
      Immunohistological Localization of the MrkD Adhesin in the Type 3 Fimbriae of
      Klebsiella pneumoniae. *Infection and Immunity* **66**, 2356, doi:10.1128/iai.66.5.2356-
      2361.1998 (1998).

87   Struve, C., Bojer, M. & Krogfelt, K. A. Characterization of *Klebsiella pneumoniae* Type
      1 Fimbriae by Detection of Phase Variation during Colonization and Infection and Impact
      on Virulence. *Infection and Immunity* **76**, 4055, doi:10.1128/iai.00494-08 (2008).

88   Jagnow, J. & Clegg, S. *Klebsiella pneumoniae* MrkD-mediated biofilm formation on
      extracellular matrix- and collagen-coated surfaces. *Microbiology (Reading, England)*
      149, 9, doi:10.1099/mic.0.26434-0 (2003).

89   Bachman, M. A. *et al. Klebsiella pneumoniae* Yersiniabactin Promotes Respiratory Tract
      Infection through Evasion of Lipocalin 2. *Infection and Immunity* **79**, 3309,
      doi:10.1128/iai.05114-11 (2011).

90   Fischbach, M. A. *et al.* The pathogen-associated *iroA* gene cluster mediates bacterial
      evasion of lipocalin 2. *Proceedings of the National Academy of Sciences of the United
      States of America* **103**, 16502, doi:10.1073/pnas.0604636103 (2006).

91   Bachman, M. A., Miller, V. L. & Weiser, J. N. Mucosal Lipocalin 2 Has Pro-Inflammatory and Iron-Sequestering Effects in Response to Bacterial Enterobactin. *PLoS Pathogens* **5**, doi:10.1371/journal.ppat.1000622 (2009).

92   Nielsen, D. W. *et al.* Outer membrane protein A (OmpA) of extraintestinal pathogenic Escherichia coli. *BMC Research Notes* **13**, doi:10.1186/s13104-020-4917-5 (2020).

93   March, C. *et al. Klebsiella pneumoniae* Outer Membrane Protein A Is Required to Prevent the Activation of Airway Epithelial Cells. *Journal of Biological Chemistry* **286**, 9956, doi:10.1074/jbc.M110.181008 (2011).

94   Padilla, E. *et al. Klebsiella pneumoniae* AcrAB efflux pump contributes to antimicrobial resistance and virulence. *Antimicrobial Agents and Chemotherapy* **54**, 177–183, doi:10.1128/aac.00715-09 (2010).

95   Masi, M., Pagès, J.-M., Villard, C. & Pradel, E. The eefABC Multidrug Efflux Pump Operon Is Repressed by H-NS in Enterobacter aerogenes. *J Bacteriol* **187**, 3894, doi:10.1128/jb.187.11.3894-3897.2005 (2005).

96   Coudeyras, S., Nakusi, L., Charbonnel, N. & Forestier, C. A Tripartite Efflux Pump Involved in Gastrointestinal Colonization by *Klebsiella pneumoniae* Confers a Tolerance Response to Inorganic Acid. *Infection and Immunity* **76**, 4633, doi:10.1128/iai.00356-08 (2008).

97   Cherrak, Y., Flaugnatti, N., Durand, E., Journet, L. & Cascales, E. Structure and Activity of the Type VI Secretion System. *Microbiology Spectrum* **7**, . doi:10.1128/microbiolspec.PSIB-0031-2019 (2019).

98   Sá-Pessoa, J. *et al.* A trans-kingdom T6SS effector induces the fragmentation of the mitochondrial network and activates innate immune receptor NLRX1 to promote infection. *Nature Communications* **14**, 871., doi:10.1038/s41467-023-36629-3 (2023).

99   Lung, T. W. F. *et al. Klebsiella pneumoniae* induces host metabolic stress that promotes tolerance to pulmonary infection. *Cell metabolism* **34**, 761, doi:10.1016/j.cmet.2022.03.009 (2022).

100   Hsieh, P.-F., Lu, Y.-R., Lin, T.-L., Lai, L.-Y. & Wang, J.-T. *Klebsiella pneumoniae* Type VI Secretion System Contributes to Bacterial Competition, Cell Invasion, Type-1 Fimbriae Expression, and *In Vivo* Colonization. *Journal of Infectious Diseases* **219**, 637, doi:10.1093/infdis/jiy534 (2019).

101   Wang, H., Guo, Y., Liu, Z. & Chang, Z. The type VI secretion system contributes to the invasiveness of liver abscess caused by *Klebsiella pneumoniae*. *Journal of Infectious Diseases* **jiad166.**, doi:10.1093/infdis/jiad166 (2023).

102   Merciecca, T. *et al.* Role of *Klebsiella pneumoniae* Type VI secretion system (T6SS) in long-term gastrointestinal colonization. *Scientific Reports* **12**, 16968., doi:10.1038/s41598-022-21396-w (2022).

103   Maroncle, N., Balestrino, D., Rich, C. & Forestier, C. Identification of *Klebsiella pneumoniae* genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis. *Infection and Immunity* **70**, 4729–4734, doi:10.1128/iai.70.8.4729-4734.2002 (2002).

104   Hudson, A. W., Barnes, A. J., Bray, A. S., Ornelles, D. A. & Zafar, M. A. *Klebsiella pneumoniae* l-Fucose Metabolism Promotes Gastrointestinal Colonization and Modulates

Its Virulence Determinants. *Infection and Immunity* **90**, e0020622.,
doi:10.1128/iai.00206-22 (2022).

105   Hennequin, C. & Forestier, C. oxyR, a LysR-type regulator involved in *Klebsiella
pneumoniae* mucosal and abiotic colonization. *Infection and Immunity* **77**, 5449–5457,
doi:10.1128/iai.00837-09 (2009).

106   Mazurkiewicz, P., Tang, C. M., Boone, C. & Holden, D. W. Signature-tagged
mutagenesis: barcoding mutants for genome-wide screens. *Nature Reviews Genetics* **7**,
929–939, doi:10.1038/nrg1984 (2006).

107   Goodman, A. L., Wu, M. & Gordon, J. I. Identifying microbial fitness determinants by
insertion sequencing using genome-wide transposon mutant libraries. *Nature Protocols* **6**,
1969–1980, doi:10.1038/nprot.2011.417 (2011).

108   van Opijnen, T. & Camilli, A. Transposon insertion sequencing: a new tool for systems-
level analysis of microorganisms. *Nature Reviews Microbiology* **11**, 435–442,
doi:10.1038/nrmicro3033 (2013).

109   Holt, K. E. *et al.* Genomic analysis of diversity, population structure, virulence, and
antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health.
Proceedings of the National Academy of Sciences of the United States of America **112**,
E3574, doi:10.1073/pnas.1501049112 (2015).

110   Ozer, E. A., Allen, J. P. & Hauser, A. R. Characterization of the core and accessory
genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. *BMC
Genomics* **15**, 737., doi:10.1186/1471-2164-15-737 (2014).

111   Vernikos, G., Medini, D., Riley, D. R. & Tettelin, H. Ten years of pan-genome analyses.
Current Opinion in Microbiology **23**, 148–154, doi:10.1016/j.mib.2014.11.016 (2015).

112   Medini, D., Donati, C., Tettelin, H., Masignani, V. & Rappuoli, R. The microbial pan-genome. *Current Opinion in Genetics & Development* **15**, 589–594, doi:10.1016/j.gde.2005.09.006 (2005).

113   Kochan, T. J. *et al.* Genomic surveillance for multidrug-resistant or hypervirulent *Klebsiella pneumoniae* among United States bloodstream isolates. *BMC Infectious Diseases* **22**, doi:10.1186/s12879-022-07558-1 (2022).

114   Olivier, V., Queen, J. & Satchell, K. J. F. Successful Small Intestine Colonization of Adult Mice by *Vibrio cholerae* Requires Ketamine Anesthesia and Accessory Toxins. *PLoS ONE* **4**, doi:10.1371/journal.pone.0007352 (2009).

115   Lagos, R. *et al.* Structure, organization and characterization of the gene cluster involved in the production of microcin E492, a channel-forming bacteriocin. *Molecular Microbiology* **42**, 229–243, doi:10.1046/j.1365-2958.2001.02630.x (2001).

116   Mercado, G., Tello, M., Marín, M., Monasterio, O. & Lagos, R. The production *in vivo* of microcin E492 with antibacterial activity depends on salmochelin and EntF. *J Bacteriol* **190**, 5464–5471, doi:10.1128/jb.00351-08 (2008).

117   Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410, doi:10.1016/s0022-2836(05)80360-2 (1990).

118   Baggs, J., Fridkin, S. K., Pollack, L. A., Srinivasan, A. & Jernigan, J. A. Estimating National Trends in Inpatient Antibiotic Use Among US Hospitals From 2006 to 2012. *JAMA Internal Medicine* **176**, 1639–1648, doi:10.1001/jamainternmed.2016.5651 (2016).

119   Hauser, A. R. *Antibiotic Basics for Clinicians: The ABCs of Choosing the Right Antibacterial Agent*. 3rd edn, (Lippincott Williams & Wilkins, 2018).

120   Hawk, C. T., Leary, S. & Morris, T. *Formulary for Laboratory Animals*. (Wiley, 2005).

121   Administration., U. F. D.   (2018).

122   Zomer, A., Burghout, P., Bootsma, H. J., Hermans, P. W. M. & van Hijum, S. A. F. T. ESSENTIALS: Software for Rapid Analysis of High Throughput Transposon Insertion Sequencing Data. *PLOS ONE* **7**, e43012, doi:10.1371/journal.pone.0043012 (2012).

123   Charlier, D., Nguyen Le Minh, P. & Roovers, M. Regulation of carbamoylphosphate synthesis in *Escherichia coli*: an amazing metabolite at the crossroad of arginine and pyrimidine biosynthesis. *Amino Acids* **50**, 1647–1661, doi:10.1007/s00726-018-2654-z (2018).

124   Palmer, T. & Berks, B. C. The twin-arginine translocation (Tat) protein export pathway. *Nature Reviews. Microbiology* **10**, 483–496, doi:10.1038/nrmicro2814 (2012).

125   Li, X.-Z., Plésiat, P. & Nikaido, H. The Challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria. *Clinical Microbiology Reviews* **28**, 337, doi:10.1128/cmr.00117-14 (2015).

126   Thanassi, D. G., Cheng, L. W. & Nikaido, H. Active efflux of bile salts by *Escherichia coli*. *J Bacteriol* **179**, 2512, doi:10.1128/jb.179.8.2512-2518.1997 (1997).

127   Goh, E.-B. *et al.* Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 17025, doi:10.1073/pnas.252607699 (2002).

128   Aranda-Díaz, A. *et al.* Assembly of gut-derived bacterial communities follows "early-bird" resource utilization dynamics. *bioRxiv*, doi:10.1101/2023.01.13.523996 (2023).

129   Vezina, B. *et al.*   (eLife Sciences Publications, Ltd, 2023).

130   Vornhagen, J. *et al.* Combined comparative genomics and clinical modeling reveals plasmid-encoded genes are independently associated with *Klebsiella* infection. *Nature Communications* **13**, doi:10.1038/s41467-022-31990-1 (2022).

131   Rousset, F. *et al.* The impact of genetic diversity on gene essentiality within the *Escherichia coli* species. *Nature Microbiology* **6**, 301–312, doi:10.1038/s41564-020-00839-y (2021).

132   Rosconi, F. *et al.* A bacterial pan-genome makes gene essentiality strain-dependent and evolvable. *Nature Microbiology* **7**, 1580–1592, doi:10.1038/s41564-022-01208-7 (2022).

133   Westerman, T. L., McClelland, M. & Elfenbein, J. R. YeiE Regulates Motility and Gut Colonization in *Salmonella enterica* Serotype *Typhimurium*. *mBio* **12**, doi:10.1128/mBio.03680-20 (2021).

134   Fujimoto, M. *et al.* Tat-exported peptidoglycan amidase-dependent cell division contributes to *Salmonella Typhimurium* fitness in the inflamed gut. *PLoS Pathogens* **14**, doi:10.1371/journal.ppat.1007391 (2018).

135   Stanley, N. R., Findlay, K., Berks, B. C. & Palmer, T. *Escherichia coli* Strains Blocked in Tat-Dependent Protein Export Exhibit Pleiotropic Defects in the Cell Envelope. *J Bacteriol* **183**, 139, doi:10.1128/jb.183.1.139-144.2001 (2001).

136   Reynolds, M. M. *et al.* Abrogation of the twin arginine transport system in *Salmonella enterica* serovar *Typhimurium* leads to colonization defects during infection. *PloS One* **6**, e15800., doi:10.1371/journal.pone.0015800 (2011).

137   Chong, Z.-S., Woo, W.-F. & Chng, S.-S. Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in *Escherichia coli*. *Molecular Microbiology* **98**, 1133–1146, doi:10.1111/mmi.13202 (2015).

138   Szczepaniak, J., Press, C. & Kleanthous, C. The multifarious roles of Tol-Pal in Gram-negative bacteria. *FEMS Microbiology Reviews* **44**, 490, doi:10.1093/femsre/fuaa018 (2020).

139   Fath, M. J., Mahanty, H. K. & Kolter, R. Characterization of a purF operon mutation which affects colicin V production. *J Bacteriol* **171**, 3158–3161, doi:10.1128/jb.171.6.3158-3161.1989 (1989).

140   Warr, A. R. *et al.* Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization. *PLoS Pathogens* **15**, doi:10.1371/journal.ppat.1007652 (2019).

141   Warr, A. R., Giorgio, R. T. & Waldor, M. K. Genetic analysis of the role of the conserved inner membrane protein CvpA in EHEC resistance to deoxycholate. *J Bacteriol* **203**, 00661–00620, doi:10.1128/jb.00661-20 (2020).

142   Hubbard, T. P. *et al.* Genetic analysis of *Vibrio parahaemolyticus* intestinal colonization. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 6283, doi:10.1073/pnas.1601718113 (2016).

143   Chaudhuri, R. R. *et al.* Comprehensive Assignment of Roles for *Salmonella Typhimurium* Genes in Intestinal Colonization of Food-Producing Animals. *PLoS Genetics* **9**, doi:10.1371/journal.pgen.1003456 (2013).

144   Nishino, K., Latifi, T. & Groisman, E. A. Virulence and drug resistance roles of multidrug efflux systems of *Salmonella enterica* serovar *Typhimurium*. *Molecular Microbiology* **59**, 126–141, doi:10.1111/j.1365-2958.2005.04940.x (2006).

145   Buckley, A. M. *et al.* The AcrAB-TolC efflux system of Salmonella enterica serovar Typhimurium plays a role in pathogenesis. *Cellular Microbiology* **8**, 847–856, doi:10.1111/j.1462-5822.2005.00671.x (2006).

146   Ashwath, P. *et al.* Sequence-Specific Gene Silencing of acrA in the Multi-drug Efflux System AcrAB Induces Sensitivity in Drug-Resistant Klebsiella pneumoniae. *Molecular Biotechnology*, 1–8, doi:10.1007/s12033-022-00585-y (2022).

147   Zhong, X. *et al.* First emergence of acrAB and oqxAB mediated tigecycline resistance in clinical isolates of *Klebsiella pneumoniae* pre-dating the use of tigecycline in a Chinese hospital. *PloS One* **9**, e115185., doi:10.1371/journal.pone.0115185 (2014).

148   Sharma, A., Gupta, V. K. & Pathania, R. Efflux pump inhibitors for bacterial pathogens: From bench to bedside. *Indian Journal of Medical Research* **149**, 129, doi:10.4103/ijmr.IJMR_2079_17 (2019).

149   Schulert, G. S. *et al. Francisella tularensis* Genes Required for Inhibition of the Neutrophil Respiratory Burst and Intramacrophage Growth Identified by Random Transposon Mutagenesis of Strain LVS. *Infection and Immunity* **77**, 1324, doi:10.1128/iai.01318-08 (2009).

150   Miller, D. P. *et al.* Genes Contributing to *Porphyromonas gingivalis* Fitness in Abscess and Epithelial Cell Colonization Environments. *Frontiers in Cellular and Infection Microbiology* **7**, doi:10.3389/fcimb.2017.00378 (2017).

151   Wexler, M. *et al.* TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export. *Journal of Biological Chemistry* **275**, 16717–16722, doi:10.1074/jbc.M000800200 (2000).

152   Zhang, L. *et al.* Pleiotropic effects of the twin-arginine translocation system on biofilm formation, colonization, and virulence in *Vibrio cholerae*. *BMC Microbiology* **9**, 114, doi:10.1186/1471-2180-9-114 (2009).

153   Lavander, M., Ericsson, S. K., Bröms, J. E. & Forsberg, Å. The Twin Arginine Translocation System Is Essential for Virulence of *Yersinia pseudotuberculosis*. *Infection and Immunity* **74**, 1768, doi:10.1128/iai.74.3.1768-1776.2006 (2006).

154   Rajashekara, G. *et al.* Functional characterization of the twin-arginine translocation system in *Campylobacter jejuni*. *Foodborne Pathogens and Disease* **6**, 935–945, doi:10.1089/fpd.2009.0298 (2009).

155   Bharathwaj, M. *et al.* The Carbapenemase BKC-1 from Klebsiella pneumoniae Is Adapted for Translocation by Both the Tat and Sec Translocons. *mBio* **12**, doi:10.1128/mBio.01302-21 (2021).

156   Randall-Hazelbauer, L. & Schwartz, M. Isolation of the Bacteriophage Lambda Receptor from *Escherichia coli*. *J Bacteriol*, doi:10.1128/jb.116.3.1436-1446.1973 (1973).

157   Park, J. T. & Uehara, T. How Bacteria Consume Their Own Exoskeletons (Turnover and Recycling of Cell Wall Peptidoglycan). *Microbiology and Molecular Biology Reviews : MMBR* **72**, 211, doi:10.1128/mmbr.00027-07 (2008).

158   Mouriño, M. *et al.* The Hha protein as a modulator of expression of virulence factors in *Escherichia coli*. *Infection and Immunity* **64**, 2881, doi:10.1128/iai.64.7.2881-2884.1996 (1996).

159   Silphaduang, U., Mascarenhas, M., Karmali, M. & Coombes, B. K. Repression of Intracellular Virulence Factors in *Salmonella* by the Hha and YdgT Nucleoid-Associated Proteins. *J Bacteriol* **189**, 3669, doi:10.1128/jb.00002-07 (2007).

160   Madrid, C., Nieto, J. M. & Juárez, A. Role of the Hha/YmoA family of proteins in the thermoregulation of the expression of virulence factors. *International Journal of Medical Microbiology : IJMM* **291**, 425–432, doi:10.1078/1438-4221-00149 (2002).

161   Schmid, K., Ebner, R., Jahreis, K., Lengeler, J. W. & Titgemeyer, F. A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria. *Molecular Microbiology* **5**, 941–950, doi:10.1111/j.1365-2958.1991.tb00769.x (1991).

162   Bulman, Z. P. *et al.* Genomic Features Associated with the Degree of Phenotypic Resistance to Carbapenems in Carbapenem-Resistant *Klebsiella pneumoniae*. *mSystems* **6**, doi:10.1128/mSystems.00194-21 (2021).

163   Wang, N., Ozer, E. A., Mandel, M. J. & Hauser, A. R. Genome-Wide Identification of *Acinetobacter baumannii* Genes Necessary for Persistence in the Lung. *mBio* **5**, doi:10.1128/mBio.01163-14 (2014).

164   Le Roux, F., Binesse, J., Saulnier, D. & Mazel, D. Construction of a *Vibrio splendidus* Mutant Lacking the Metalloprotease Gene *vsm* by Use of a Novel Counterselectable Suicide Vector. *Applied and Environmental Microbiology* **73**, 777, doi:10.1128/aem.02147-06 (2007).

165   Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114, doi:10.1093/bioinformatics/btu170 (2014).

166   Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology* **20**, doi:10.1186/s13059-019-1727-y (2019).

167   Tatusova, T. *et al.* NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research* **44**, 6614, doi:10.1093/nar/gkw569 (2016).

168   Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069, doi:10.1093/bioinformatics/btu153 (2014).

169   Skurnik, D. *et al.* A comprehensive analysis of *in vitro* and *in vivo* genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. *PLoS Pathogens* **9**, e1003582., doi:10.1371/journal.ppat.1003582 (2013).

170   Huang, T.-W. *et al.* Capsule deletion via a λ-Red knockout system perturbs biofilm formation and fimbriae expression in *Klebsiella pneumoniae* MGH 78578. *BMC Research Notes* **7**, 13, doi:10.1186/1756-0500-7-13 (2014).

171   Kazi, M. I., Schargel, R. D. & Boll, J. M. Generating Transposon Insertion Libraries in Gram-Negative Bacteria for High-Throughput Sequencing. *JoVE (Journal of Visualized Experiments)*, e61612, doi:10.3791/61612 (2020).

172   egonozer/essentials_local: Version 2.1 v. 2.1 (Zenodo, 2023).

173   Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, 1–10, doi:10.1186/gb-2009-10-3-r25 (2009).

174   Wang, N. & Ozer, E. A. A Method for Bioinformatic Analysis of Transposon Insertion Sequencing (INSeq) Results for Identification of Microbial Fitness Determinants. *Methods in Molecular Biology (Clifton, N.J.)* **1498**, ;, doi:10.1007/978-1-4939-6472-7_16 (2017).

175   Kanehisa, M., Sato, Y. & Morishima, K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology* **428**, 726–731, doi:10.1016/j.jmb.2015.11.006 (2016).

176   Kanehisa, M., Sato, Y. & Kawashima, M. KEGG mapping tools for uncovering hidden features in biological data. *Protein Science : a Publication of the Protein Society* **31**, 47–53, doi:10.1002/pro.4172 (2022).

APPENDIX 1: Colonization-associated transposon sequencing results for CRE-166

Gene Identifier	Gene	Product	log2(fold change)	False Discovery Rate
JCNGAGPE_00682	tolC	Outer membrane protein TolC	-7.104	4.56E-07
ECJKFODA_00215	sopB	Protein SopB	-6.965	1.23E-11
JCNGAGPE_04042	bglY	Beta-galactosidase BglY	-6.948	6.47E-07
ECJKFODA_00216	-	hypothetical protein	-6.943	1.23E-11
JCNGAGPE_03410	focA	putative formate transporter 1	-6.486	4.39E-06
JCNGAGPE_04043	ganB	Arabinogalactan endo-beta-1,4-galactanase	-6.251	7.25E-06
JCNGAGPE_03639	pal_2	Peptidoglycan-associated lipoprotein	-6.232	7.25E-06
JCNGAGPE_03360	ompA_3	Outer membrane protein A	-6.224	7.25E-06
JCNGAGPE_01434	bepA_2	Beta-barrel assembly-enhancing protease	-6.222	7.25E-06
JCNGAGPE_03640	tolB	Tol-Pal system protein TolB	-6.196	7.25E-06
JCNGAGPE_03641	-	hypothetical protein	-5.977	1.90E-05
JCNGAGPE_04029	acrA	Multidrug efflux pump subunit AcrA	-5.906	2.24E-05
JCNGAGPE_00499	sspA	Stringent starvation protein A	-5.880	2.24E-05
JCNGAGPE_04775	miaA	tRNA dimethylallyltransferase	-5.773	1.90E-05
JCNGAGPE_01614	ompC	Outer membrane porin C	-5.738	3.71E-05
JCNGAGPE_00167	mtlD	Mannitol-1-phosphate 5-dehydrogenase	-5.711	3.50E-05
JCNGAGPE_03411	pflB	Formate acetyltransferase 1	-5.617	2.57E-05
JCNGAGPE_04134	phoB	Phosphate regulon transcriptional regulatory protein PhoB	-5.583	2.79E-05
JCNGAGPE_05180	dsbA_2	Thiol:disulfide interchange protein DsbA	-5.461	8.58E-05
JCNGAGPE_01655	yeiE	Transcriptional activator yeiE	-5.445	8.71E-05
JCNGAGPE_00153	envC	Murein hydrolase activator EnvC	-5.405	9.31E-05
JCNGAGPE_01487	ptsI	Phosphoenolpyruvate-protein phosphotransferase	-5.404	0.000108568
JCNGAGPE_03888	scrY_3	Sucrose porin	-5.390	3.71E-05
JCNGAGPE_01261	bamE	Outer membrane protein assembly factor BamE	-5.361	0.000115648
ECJKFODA_00181	higA1	Antitoxin HigA1	-5.360	3.84E-07
JCNGAGPE_04116	tgt	Queuine tRNA-ribosyltransferase	-5.359	0.000108568

JCNGAGPE_00508	arcB	Aerobic respiration control sensor protein ArcB	-5.342	9.31E-05
JCNGAGPE_00142	-	hypothetical protein	-5.327	0.00011935
JCNGAGPE_00370	crp	cAMP-activated global transcriptional regulator CRP	-5.277	0.000148301
JCNGAGPE_04445	carB	Carbamoyl-phosphate synthase large chain	-5.267	0.000239124
JCNGAGPE_02792	mlc	Protein mlc	-5.161	0.00025692
CAHBADIM_00011	-	hypothetical protein	-5.144	3.78E-08
JCNGAGPE_00144	-	hypothetical protein	-5.121	0.000247372
JCNGAGPE_00145	-	hypothetical protein	-5.083	0.000292079
MHOCDJNC_00056	traA	Pilin	-4.983	1.10E-08
JCNGAGPE_00476	csrD	RNase E specificity factor CsrD	-4.947	0.000465009
JCNGAGPE_05052	wecG	UDP-N-acetyl-D-mannosaminuronic acid transferase	-4.895	0.000496939
JCNGAGPE_03643	tolQ	Tol-Pal system protein TolQ	-4.871	0.000569615
JCNGAGPE_05084	hdfR_5	HTH-type transcriptional regulator HdfR	-4.863	0.000705837
JCNGAGPE_05203	mnmG	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	-4.849	0.000248062
JCNGAGPE_02122	ychF	Ribosome-binding ATPase YchF	-4.804	0.000616926
JCNGAGPE_04117	queA	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	-4.789	0.000292835
JCNGAGPE_04934	pgi	Glucose-6-phosphate isomerase	-4.766	0.000705837
JCNGAGPE_04337	dksA_2	RNA polymerase-binding transcription factor DksA	-4.752	0.000718109
JCNGAGPE_04773	hflX	GTPase HflX	-4.661	0.000480278
GHJMACIB_00082	-	hypothetical protein	-4.645	0.000676705
JCNGAGPE_01547	cvpA	Colicin V production protein	-4.611	0.00143582
JCNGAGPE_03890	scrB	Sucrose-6-phosphate hydrolase	-4.530	0.002718259
MHOCDJNC_00070	-	hypothetical protein	-4.499	5.00E-09
JCNGAGPE_04415	setA	Sugar efflux transporter A	-4.431	0.000859679
JCNGAGPE_03647	cydB_3	Cytochrome bd-I ubiquinol oxidase subunit 2	-4.421	0.000749449
GHJMACIB_00012	-	hypothetical protein	-4.416	0.000676705

JCNGAGPE_04367	aceE	Pyruvate dehydrogenase E1 component	-4.412	0.00213906
JCNGAGPE_03648	cydA_2	Cytochrome bd-I ubiquinol oxidase subunit 1	-4.387	0.001610871
JCNGAGPE_03686	nagB	Glucosamine-6-phosphate deaminase	-4.355	0.002527976
JCNGAGPE_01281	raiA	Ribosome-associated inhibitor A	-4.308	0.002704814
JCNGAGPE_05123	menA	1,4-dihydroxy-2-naphthoate octaprenyltransferase	-4.306	0.003186716
JCNGAGPE_00109	recG	ATP-dependent DNA helicase RecG	-4.286	0.003286977
JCNGAGPE_02837	fnr	Fumarate and nitrate reduction regulatory protein	-4.248	0.004192053
JCNGAGPE_00327	malQ	4-alpha-glucanotransferase	-4.239	0.002155363
JCNGAGPE_01258	smpB	SsrA-binding protein	-4.188	0.005814795
ECJKFODA_00161	-	hypothetical protein	-4.186	0.000339475
JCNGAGPE_03889	sacX	Negative regulator of SacY activity	-4.180	0.004492147
JCNGAGPE_03412	pflA	Pyruvate formate-lyase 1-activating enzyme	-4.172	0.006550888
JCNGAGPE_02451	mnaT	L-amino acid N-acyltransferase MnaT	-4.145	0.005923768
JCNGAGPE_04603	-	hypothetical protein	-4.126	0.006277821
JCNGAGPE_01654	-	hypothetical protein	-4.106	0.006186136
JCNGAGPE_01259	ratA	Ribosome association toxin RatA	-4.081	0.008539878
JCNGAGPE_05129	cpxA	Sensor histidine kinase CpxA	-4.080	0.007063744
JCNGAGPE_00811	speB	Agmatinase	-4.049	0.007010349
JCNGAGPE_04416	-	hypothetical protein	-4.028	0.002634736
JCNGAGPE_00182	xylA_1	Xylose isomerase	-3.989	0.002434318
JCNGAGPE_03621	galE	UDP-glucose 4-epimerase	-3.987	0.008077597
JCNGAGPE_03054	hipO_2	Hippurate hydrolase	-3.975	0.021592144
ECJKFODA_00038	-	hypothetical protein	-3.975	9.22E-05
JCNGAGPE_01740	-	hypothetical protein	-3.953	0.009254878
JCNGAGPE_03080	topA_2	DNA topoisomerase 1	-3.934	0.011181324
JCNGAGPE_00310	gntR_1	HTH-type transcriptional regulator GntR	-3.930	0.008077597
JCNGAGPE_00126	rph	Ribonuclease PH	-3.911	0.010384778
GHJMACIB_00097	-	hypothetical protein	-3.904	0.004353576
JCNGAGPE_01542	truA	tRNA pseudouridine synthase A	-3.894	0.004021724
JCNGAGPE_03622	galT	Galactose-1-phosphate uridylyltransferase	-3.827	0.014499305

JCNGAGPE_00311	gntK	Thermoresistant gluconokinase	-3.821	0.014731811
JCNGAGPE_02641	-	hypothetical protein	-3.811	0.030145252
JCNGAGPE_02605	glnQ_4	Glutamine transport ATP-binding protein GlnQ	-3.800	0.013866994
JCNGAGPE_04618	blc_1	Outer membrane lipoprotein Blc	-3.795	0.014455064
JCNGAGPE_04268	gmhB	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase	-3.780	0.007639943
JCNGAGPE_04676	pepA	Cytosol aminopeptidase	-3.761	0.015550294
JCNGAGPE_01563	ackA	Acetate kinase	-3.753	0.015902147
JCNGAGPE_01357	bamB	Outer membrane protein assembly factor BamB	-3.741	0.017754304
JCNGAGPE_02235	pykF	Pyruvate kinase I	-3.710	0.017479495
JCNGAGPE_01271	rplS	50S ribosomal protein L19	-3.696	0.018879938
JCNGAGPE_01082	hcpA_1	Major exported protein	-3.680	0.018699995
JCNGAGPE_05172	glnA	Glutamine synthetase	-3.667	0.021592144
JCNGAGPE_03710	-	hypothetical protein	-3.663	0.01359159
JCNGAGPE_03687	nagA	N-acetylglucosamine-6-phosphate deacetylase	-3.663	0.010943149
JCNGAGPE_03646	cydX	Cytochrome bd-I ubiquinol oxidase subunit X	-3.659	0.007684072
JCNGAGPE_04306	degP	Periplasmic serine endoprotease DegP	-3.658	0.021104562
JCNGAGPE_03423	cydD	ATP-binding/permease protein CydD	-3.629	0.021592144
JCNGAGPE_02255	-	3-mercaptopropionate dioxygenase	-3.598	0.013648617
JCNGAGPE_05122	hslU	ATP-dependent protease ATPase subunit HslU	-3.595	0.009254878
JCNGAGPE_02383	slyA_1	Transcriptional regulator SlyA	-3.591	0.010371149
JCNGAGPE_00067	bcr_1	Bicyclomycin resistance protein	-3.591	0.023602038
JCNGAGPE_05128	cpxR	Transcriptional regulatory protein CpxR	-3.589	0.024360058
JCNGAGPE_03406	aroA	3-phosphoshikimate 1-carboxyvinyltransferase	-3.588	0.013648617
JCNGAGPE_00338	ompR_1	Transcriptional regulatory protein OmpR	-3.587	0.02330413
JCNGAGPE_03526	gsiC_3	Glutathione transport system permease protein GsiC	-3.581	0.010371149
JCNGAGPE_01628	glcR_1	HTH-type transcriptional repressor GlcR	-3.566	0.024473139

JCNGAGPE_00339	envZ	Osmolarity sensor protein EnvZ	-3.543	0.012482087
JCNGAGPE_03352	hspQ	Heat shock protein HspQ	-3.522	0.030145252
MHOCDJNC_00075	-	hypothetical protein	-3.517	0.000365103
JCNGAGPE_00136	-	hypothetical protein	-3.512	0.030145252
JCNGAGPE_03675	pgm	Phosphoglucomutase	-3.500	0.036040283
ECJKFODA_00162	-	hypothetical protein	-3.496	0.00407046
JCNGAGPE_01541	usg	USG-1 protein	-3.496	0.030145252
JCNGAGPE_04960	purH	Bifunctional purine biosynthesis protein PurH	-3.470	0.013648617
JCNGAGPE_00959	amiC_1	N-acetylmuramoyl-L-alanine amidase AmiC	-3.462	0.021592144
JCNGAGPE_05013	tatC	Sec-independent protein translocase protein TatC	-3.452	0.032262482
MHOCDJNC_00054	traM	Relaxosome protein TraM	-3.441	0.000337688
JCNGAGPE_01828	-	hypothetical protein	-3.423	0.040243327
JCNGAGPE_01625	yejM	Inner membrane protein YejM	-3.414	0.02037644
JCNGAGPE_04790	epmA	Elongation factor P--(R)-beta-lysine ligase	-3.411	0.037648215
JCNGAGPE_00836	argP_1	HTH-type transcriptional regulator ArgP	-3.401	0.015120695
GHJMACIB_00081	vapC	tRNA(fMet)-specific endonuclease VapC	-3.389	0.000820042
JCNGAGPE_05015	tatA	Sec-independent protein translocase protein TatA	-3.384	0.042481688
JCNGAGPE_01512	glk	Glucokinase	-3.374	0.014404063
JCNGAGPE_05083	-	hypothetical protein	-3.357	0.042215911
JCNGAGPE_02166	adhE_2	Aldehyde-alcohol dehydrogenase	-3.353	0.027830042
JCNGAGPE_03388	-	hypothetical protein	-3.344	0.017479495
JCNGAGPE_05058	rffH_2	Glucose-1-phosphate thymidylyltransferase 2	-3.334	0.016042824
JCNGAGPE_03642	tolR	Tol-Pal system protein TolR	-3.294	0.048223587
JCNGAGPE_03407	serC	Phosphoserine aminotransferase	-3.293	0.015120695
JCNGAGPE_04145	proC	Pyrroline-5-carboxylate reductase	-3.265	0.028934082
JCNGAGPE_03760	mtnA	Methylthioribose-1-phosphate isomerase	-3.209	0.027087996
JCNGAGPE_03389	-	hypothetical protein	-3.184	0.018997376
JCNGAGPE_03580	moaC	Cyclic pyranopterin monophosphate synthase	-3.170	0.036962216

JCNGAGPE_01425	purN	Phosphoribosylglycinamide formyltransferase	-3.170	0.025199472
JCNGAGPE_05042	cyaA	Adenylate cyclase	-3.167	0.037648215
JCNGAGPE_03107	-	hypothetical protein	-3.164	0.021592144
JCNGAGPE_04187	-	hypothetical protein	-3.148	0.042916525
JCNGAGPE_03378	pncB	Nicotinate phosphoribosyltransferase	-3.140	0.039581763
JCNGAGPE_00100	zapB	Cell division protein ZapB	-3.133	0.035704137
JCNGAGPE_03754	-	hypothetical protein	-3.132	0.02330413
JCNGAGPE_03833	-	hypothetical protein	-3.118	0.048223587
JCNGAGPE_00481	rng	Ribonuclease G	-3.098	0.013648617
JCNGAGPE_01440	purC	Phosphoribosylaminoimidazole-succinocarboxamide synthase	-3.088	0.038960074
JCNGAGPE_03792	-	hypothetical protein	-3.078	0.035704137
JCNGAGPE_01318	purL	Phosphoribosylformylglycinamidine synthase	-3.044	0.031260192
CAHBADIM_00023	-	hypothetical protein	-3.010	1.25E-06
JCNGAGPE_00482	-	hypothetical protein	-2.999	0.024624001
CAHBADIM_00022	-	hypothetical protein	-2.986	0.000108494
JCNGAGPE_04044	malG_1	Maltose/maltodextrin transport system permease protein MalG	-2.966	0.042175111
JCNGAGPE_05096	fabR	HTH-type transcriptional repressor FabR	-2.939	0.049308648
JCNGAGPE_03062	sapF	Peptide transport system ATP-binding protein SapF	-2.922	0.045460155
JCNGAGPE_00326	malP_4	Maltodextrin phosphorylase	-2.901	0.027830042
JCNGAGPE_00884	mdtO	Multidrug resistance protein MdtO	-2.823	0.041626346
ECJKFODA_00057	-	hypothetical protein	-2.704	0.018591331
GHJMACIB_00163	-	hypothetical protein	-2.684	0.043459194
ECJKFODA_00167	traV	Protein TraV	-2.525	0.03418481
CAHBADIM_00010	ccdA	Antitoxin CcdA	-1.931	0.003144391
CAHBADIM_00018	-	hypothetical protein	2.437	0.000246025
GHJMACIB_00121	-	hypothetical protein	2.553	0.031747248
JCNGAGPE_00904	yadV_1	putative fimbrial chaperone YadV	3.045	0.049234457
JCNGAGPE_02320	-	hypothetical protein	3.207	0.049819258
JCNGAGPE_01043	gudD	Glucarate dehydratase	3.259	0.04950649

JCNGAGPE_02782	speG	Spermidine N(1)-acetyltransferase	3.268	0.042481688
JCNGAGPE_00997	cbiM	Cobalt transport protein CbiM	3.305	0.045460155
JCNGAGPE_00909	-	hypothetical protein	3.439	0.013648617
JCNGAGPE_03186	nimR_5	HTH-type transcriptional regulator NimR	3.456	0.049819258
ECJKFODA_00144	ylpA	Lipoprotein YlpA	3.493	0.003028696
JCNGAGPE_01945	dcyD	D-cysteine desulfhydrase	3.507	0.037850898
JCNGAGPE_04357	-	hypothetical protein	3.521	0.041626346
JCNGAGPE_01218	-	hypothetical protein	3.537	0.046612591
JCNGAGPE_04150	-	hypothetical protein	3.590	0.03247458
MHOCDJNC_00082	-	hypothetical protein	3.601	0.000649103
JCNGAGPE_02827	pcaI	3-oxoadipate CoA-transferase subunit A	3.801	0.047676592
JCNGAGPE_02007	-	hypothetical protein	3.844	0.04415016
JCNGAGPE_01759	-	hypothetical protein	3.848	0.007582352
JCNGAGPE_02770	-	hypothetical protein	3.856	0.048444515
JCNGAGPE_05223	gamA	putative glucosamine-6-phosphate deaminase 2	3.918	0.02786328
JCNGAGPE_03705	gltJ_1	Glutamate/aspartate import permease protein GltJ	3.978	0.037364781
JCNGAGPE_01291	-	hypothetical protein	4.058	0.039205871
JCNGAGPE_04513	rimI	[Ribosomal protein S18]-alanine N-acetyltransferase	4.076	0.048787584
JCNGAGPE_01790	-	hypothetical protein	4.128	0.015728881
JCNGAGPE_05097	sthA	Soluble pyridine nucleotide transhydrogenase	4.132	0.012980429
JCNGAGPE_01729	dnaK_1	Chaperone protein DnaK	4.148	0.049996439
GHJMACIB_00164	-	hypothetical protein	4.177	0.024525602
JCNGAGPE_03514	ybjI	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YbjI	4.277	0.006862519
JCNGAGPE_01889	COQ5	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial	4.365	0.006186136
JCNGAGPE_02899	paaZ	Bifunctional protein PaaZ	4.414	0.007538534
JCNGAGPE_02020	kdgR_2	Transcriptional regulator KdgR	4.449	0.002054845
JCNGAGPE_05011	rfaH	Transcription antitermination protein RfaH	4.458	0.017479495
JCNGAGPE_05146	rafR	HTH-type transcriptional regulator RafR	4.557	0.001271868

JCNGAGPE_01796	yeaR_1	putative protein YeaR	4.759	0.016358181
JCNGAGPE_03688	nagC_2	N-acetylglucosamine repressor	4.788	0.001012333
JCNGAGPE_02831	iaaH	Indole-3-acetyl-aspartic acid hydrolase	4.919	0.009167914
JCNGAGPE_02347	rhtC_1	Threonine efflux protein	4.965	0.021592144
JCNGAGPE_02776	rspR	HTH-type transcriptional repressor RspR	5.014	0.000616926
JCNGAGPE_03395	-	hypothetical protein	5.078	0.012482087
JCNGAGPE_04495	slt	Soluble lytic murein transglycosylase	5.232	0.00326483
JCNGAGPE_04047	malK_1	Maltose/maltodextrin import ATP-binding protein MalK	5.233	0.008297827
GHJMACIB_00061	-	hypothetical protein	5.287	0.00469881
JCNGAGPE_01555	-	hypothetical protein	5.315	0.008077597
JCNGAGPE_05099	argH	Argininosuccinate lyase	5.330	0.003512415
JCNGAGPE_04923	malE	Maltose/maltodextrin-binding periplasmic protein	5.529	4.10E-05
JCNGAGPE_00097	gltS	Sodium/glutamate symporter	5.636	0.013417991
JCNGAGPE_00325	malT_1	HTH-type transcriptional regulator MalT	5.707	2.79E-05
JCNGAGPE_00784	xerC_1	Tyrosine recombinase XerC	5.829	0.008077597
GHJMACIB_00025	-	hypothetical protein	6.160	0.000676705
JCNGAGPE_04925	malG_2	Maltose/maltodextrin transport system permease protein MalG	6.178	1.91E-05
JCNGAGPE_03605	-	hypothetical protein	6.214	0.003093365
JCNGAGPE_00996	cbiL	Cobalt-precorrin-2 C(20)-methyltransferase	6.264	0.008539878
JCNGAGPE_04924	malF	Maltose/maltodextrin transport system permease protein MalF	6.426	7.25E-06
JCNGAGPE_03721	dacA	D-alanyl-D-alanine carboxypeptidase DacA	7.134	0.000870513
JCNGAGPE_04040	lamB_2	Maltoporin	7.232	0.001023655
JCNGAGPE_03387	gloC	Hydroxyacylglutathione hydrolase GloC	7.593	0.008077597
JCNGAGPE_00066	murQ_1	N-acetylmuramic acid 6-phosphate etherase	7.839	0.000289173
JCNGAGPE_02016	proQ	RNA chaperone ProQ	8.623	0.000628077
JCNGAGPE_05070	-	hypothetical protein	8.848	4.10E-05
JCNGAGPE_00556	nlpI	Lipoprotein NlpI	9.496	1.41E-05
JCNGAGPE_02017	prc	Tail-specific protease	9.683	1.36E-05
JCNGAGPE_03887	scrK	Fructokinase	10.376	2.20E-09
JCNGAGPE_03891	cra_1	Catabolite repressor/activator	10.549	3.10E-11

APPENDIX 2: **Colonization-associated transposon sequencing results for KPN46**

Gene Identifier	Gene	Product	log2(fold change)	False Discovery Rate
GDOLJONB_05051	acrB_2	Multidrug efflux pump subunit AcrB	-8.075	1.55E-07
GDOLJONB_00499	pflB	Formate acetyltransferase 1	-7.979	1.55E-07
GDOLJONB_00500	focA	putative formate transporter 1	-7.092	3.13E-06
GDOLJONB_05052	acrA	Multidrug efflux pump subunit AcrA	-6.935	3.84E-06
GDOLJONB_00305	cydA_1	Cytochrome bd-I ubiquinol oxidase subunit 1	-6.933	3.84E-06
GDOLJONB_03700	mtlD	Mannitol-1-phosphate 5-dehydrogenase	-6.709	7.90E-06
GDOLJONB_01119	fnr	Fumarate and nitrate reduction regulatory protein	-6.542	1.09E-05
GDOLJONB_02602	purM	Phosphoribosylformylglycinamidine cyclo-ligase	-6.347	1.88E-05
GDOLJONB_02588	purC	Phosphoribosylaminoimidazole-succinocarboxamide synthase	-6.240	1.23E-05
GDOLJONB_00498	pflA	Pyruvate formate-lyase 1-activating enzyme	-6.183	3.50E-05
GDOLJONB_04085	tatA	Sec-independent protein translocase protein TatA	-6.061	4.81E-05
GDOLJONB_00312	-	hypothetical protein	-5.965	6.63E-05
GDOLJONB_04086	tatB	Sec-independent protein translocase protein TatB	-5.697	0.0001476
GDOLJONB_00650	pyrC	Dihydroorotase	-5.667	0.000158
GDOLJONB_02541	ptsI	Phosphoenolpyruvate-protein phosphotransferase	-5.665	0.000157
GDOLJONB_00712	phoQ	Sensor protein PhoQ	-5.549	0.0002275
GDOLJONB_04710	aceE	Pyruvate dehydrogenase E1 component	-5.504	0.0002649
GDOLJONB_04087	tatC	Sec-independent protein translocase protein TatC	-5.452	0.0001337
GDOLJONB_04159	pgi	Glucose-6-phosphate isomerase	-5.270	0.0006189

GDOLJONB_04005	fabR	HTH-type transcriptional repressor FabR	-5.247	0.0006189
GDOLJONB_01768	pykF	Pyruvate kinase I	-5.209	0.0007231
GDOLJONB_03973	cpxR	Transcriptional regulatory protein CpxR	-5.127	0.0009178
GDOLJONB_04638	apaH	Bis(5'-nucleosyl)-tetraphosphatase, symmetrical	-5.069	0.0011247
GDOLJONB_04642	surA	Chaperone SurA	-5.025	0.0013208
GDOLJONB_03725	-	hypothetical protein	-4.906	0.0023538
GDOLJONB_01837	adhE_1	Aldehyde-alcohol dehydrogenase	-4.848	0.0045968
GDOLJONB_03881	pstB	Phosphate import ATP-binding protein PstB	-4.759	0.0038511
GDOLJONB_01271	hipB	Antitoxin HipB	-4.707	0.0036875
GDOLJONB_04662	setA	Sugar efflux transporter A	-4.701	0.0035264
GDOLJONB_04405	pepA	Cytosol aminopeptidase	-4.653	0.0037563
GDOLJONB_04140	purH	Bifunctional purine biosynthesis protein PurH	-4.634	0.0023416
GDOLJONB_02322	cmpR_1	HTH-type transcriptional activator CmpR	-4.572	0.0045968
GDOLJONB_03742	rph	Ribonuclease PH	-4.527	0.0049199
GDOLJONB_01163	mlc	Protein mlc	-4.515	0.0059157
GDOLJONB_00277	seqA	Negative modulator of initiation of replication	-4.515	0.005205
GDOLJONB_04176	dgkA	Diacylglycerol kinase	-4.477	0.0068082
GDOLJONB_03897	mnmG	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	-4.451	0.0063067
GDOLJONB_02429	cvpA	Colicin V production protein	-4.449	0.0028416
GDOLJONB_02635	iscA	Iron-binding protein IscA	-4.410	0.0069066
GDOLJONB_02168	-	hypothetical protein	-4.406	0.0077084
GDOLJONB_00669	fabF_1	3-oxoacyl-[acyl-carrier-protein] synthase 2	-4.370	0.0078313
GDOLJONB_00330	galK	Galactokinase	-4.278	0.0090615
GDOLJONB_00959	-	hypothetical protein	-4.235	0.0108495
GDOLJONB_00937	sapB	Putrescine export system permease protein SapB	-4.208	0.0170306
GDOLJONB_04189	-	hypothetical protein	-4.115	0.0168593

GDOLJONB_03370	rpoN	RNA polymerase sigma-54 factor	-4.077	0.0168593
GDOLJONB_03956	rafA	Alpha-galactosidase	-4.074	0.0170306
GDOLJONB_02613	guaB	Inosine-5'-monophosphate dehydrogenase	-4.066	0.0174198
GDOLJONB_01443	fieF_1	Ferrous-iron efflux pump FieF	-4.019	0.0223394
GDOLJONB_00507	ihfB	Integration host factor subunit beta	-3.988	0.0204243
GDOLJONB_02973	recC	RecBCD enzyme subunit RecC	-3.984	0.0234755
GDOLJONB_02507	-	tRNA-Arg	-3.982	0.0223394
GDOLJONB_04459	-	hypothetical protein	-3.958	0.0216354
GDOLJONB_02618	bamB	Outer membrane protein assembly factor BamB	-3.931	0.0242634
GDOLJONB_02413	ackA	Acetate kinase	-3.924	0.0262738
GDOLJONB_04632	carA_2	Carbamoyl-phosphate synthase small chain	-3.917	0.0149549
GDOLJONB_00309	ybgC	Acyl-CoA thioester hydrolase YbgC	-3.904	0.0260496
GDOLJONB_03985	metJ	Met repressor	-3.881	0.0264385
GDOLJONB_02323	-	hypothetical protein	-3.870	0.0264385
GDOLJONB_00866	pyrF	Orotidine 5'-phosphate decarboxylase	-3.863	0.0268054
GDOLJONB_04311	miaA	tRNA dimethylallyltransferase	-3.856	0.0264536
GDOLJONB_03167	mltC	-	-3.853	0.0265944
GDOLJONB_04467	-	hypothetical protein	-3.849	0.0291801
GDOLJONB_01840	galU	UTP--glucose-1-phosphate uridylyltransferase	-3.842	0.0331472
GDOLJONB_04772	glnD	Bifunctional uridylyltransferase/uridylyl-removing enzyme	-3.828	0.030414
GDOLJONB_03377	arcB	Aerobic respiration control sensor protein ArcB	-3.821	0.0331872
GDOLJONB_03392	degS	Serine endoprotease DegS	-3.783	0.0311319
GDOLJONB_03347	rlmE	Ribosomal RNA large subunit methyltransferase E	-3.783	0.0312166
GDOLJONB_01841	rssB_2	Regulator of RpoS	-3.775	0.0089887
GDOLJONB_03701	mtlR	Mannitol operon repressor	-3.767	0.0309513

GDOLJONB_03929	glnA	Glutamine synthetase	-3.755	0.0170306
GDOLJONB_01147	mdcG	Phosphoribosyl-dephospho-CoA transferase	-3.752	0.0338847
GDOLJONB_03509	damX	Cell division protein DamX	-3.701	0.0408887
GDOLJONB_03976	pfkA	ATP-dependent 6-phosphofructokinase isozyme 1	-3.694	0.035289
GDOLJONB_04312	hfq	RNA-binding protein Hfq	-3.679	0.0170306
GDOLJONB_04970	-	hypothetical protein	-3.655	0.0400702
GDOLJONB_04820	-	hypothetical protein	-3.630	0.0406796
GDOLJONB_00955	tpx	Thiol peroxidase	-3.630	0.0408887
GDOLJONB_04318	purA	Adenylosuccinate synthetase	-3.624	0.0406796
GDOLJONB_00812	cho	Excinuclease cho	-3.600	0.0457291
GDOLJONB_00682	ndh	NADH dehydrogenase	-3.586	0.04481
GDOLJONB_00331	galT	Galactose-1-phosphate uridylyltransferase	-3.574	0.0476859
GDOLJONB_03758	recG	ATP-dependent DNA helicase RecG	-3.526	0.0260496
GDOLJONB_00463	artI	Putative ABC transporter arginine-binding protein 2	-3.434	0.0331484
GDOLJONB_03957	melY	Melibiose permease	-3.407	0.0268054
GDOLJONB_01628	-	hypothetical protein	-3.361	0.0436981
GDOLJONB_02981	rppH_1	RNA pyrophosphohydrolase	-3.315	0.0348922
GDOLJONB_03139	epd	D-erythrose-4-phosphate dehydrogenase	-3.245	0.0321827
GDOLJONB_02363	ompC	Outer membrane protein C	-3.129	0.0346591
GDOLJONB_05039	bglY	Beta-galactosidase BglY	-2.896	0.0406796
GDOLJONB_01520	ansP2	L-asparagine permease 2	3.286	0.0346591
GDOLJONB_00003	-	hypothetical protein	3.387	0.0338847
GDOLJONB_01003	-	hypothetical protein	3.485	0.035289
GDOLJONB_00016	ushA	-	3.519	0.035289
GDOLJONB_03572	livJ_2	Leu/Ile/Val-binding protein	3.543	0.0331484
GDOLJONB_02870	cysD	Sulfate adenylyltransferase subunit 2	3.581	0.0338847
GDOLJONB_03564	-	hypothetical protein	3.592	0.0408887
GDOLJONB_00975	-	hypothetical protein	3.715	0.0385665
GDOLJONB_02843	-	hypothetical protein	3.743	0.0382322
GDOLJONB_02831	licC_3	Lichenan permease IIC component	3.824	0.0264536
GDOLJONB_02916	-	hypothetical protein	3.855	0.0365449

GDOLJONB_00897	-	hypothetical protein	3.897	0.0247484
GDOLJONB_01717	-	hypothetical protein	3.937	0.0406796
GDOLJONB_02771	emrB_2	Multidrug export protein EmrB	3.967	0.0445364
GDOLJONB_02826	gmuD_1	6-phospho-beta-glucosidase GmuD	4.032	0.0264536
GDOLJONB_02829	licC_2	Lichenan permease IIC component	4.068	0.0436887
GDOLJONB_00970	-	hypothetical protein	4.070	0.0248228
GDOLJONB_05047	ylaC	Inner membrane protein YlaC	4.075	0.0265944
GDOLJONB_02892	syd	Protein Syd	4.151	0.0473365
GDOLJONB_02837	mntB_2	Manganese transport system membrane protein MntB	4.162	0.0264536
GDOLJONB_02770	emrA_2	Multidrug export protein EmrA	4.168	0.0481205
GDOLJONB_02682	pat_2	Protein lysine acetyltransferase Pat	4.182	0.0472751
GDOLJONB_03537	malT_2	HTH-type transcriptional regulator MalT	4.234	0.0054619
GDOLJONB_01988	exoX	Exodeoxyribonuclease 10	4.273	0.0240064
GDOLJONB_02239	-	hypothetical protein	4.278	0.0141029
GDOLJONB_04535	hpcE_2	Homoprotocatechuate catabolism bifunctional isomerase/decarboxylase	4.315	0.0334204
GDOLJONB_02778	-	hypothetical protein	4.337	0.035289
GDOLJONB_02740	fimD_2	Outer membrane usher protein FimD	4.343	0.0471218
GDOLJONB_04801	trmO	tRNA (adenine(37)-N6)-methyltransferase	4.394	0.0200997
GDOLJONB_02875	cysJ	Sulfite reductase [NADPH] flavoprotein alpha-component	4.402	0.0331484
GDOLJONB_02910	fucP_2	L-fucose-proton symporter	4.418	0.0054619
GDOLJONB_02915	rlmM	Ribosomal RNA large subunit methyltransferase M	4.425	0.0274699
GDOLJONB_02928	cbiQ	Cobalt transport protein CbiQ	4.479	0.0445364
GDOLJONB_00477	aqpZ_1	Aquaporin Z	4.518	0.020108
GDOLJONB_02752	-	hypothetical protein	4.526	0.0359852
GDOLJONB_04497	lsrB_3	-	4.545	0.012818

GDOLJONB_00026	-	hypothetical protein	4.595	0.0274699
GDOLJONB_04374	-	hypothetical protein	4.635	0.0406796
GDOLJONB_02789	mntA	Manganese-binding lipoprotein MntA	4.681	0.0457291
GDOLJONB_04169	malE	Maltose-binding periplasmic protein	4.686	0.0083504
GDOLJONB_00164	-	hypothetical protein	4.735	0.0291801
GDOLJONB_04167	malG_1	Maltose transport system permease protein MalG	4.748	0.0046349
GDOLJONB_02836	fhuC_2	Iron(3+)-hydroxamate import ATP-binding protein FhuC	4.749	0.049666
GDOLJONB_02886	gudD	Glucarate dehydratase	4.761	0.0445364
GDOLJONB_00161	-	ISNCY family transposase ISPlu15	4.791	0.0049199
GDOLJONB_02855	dmlR_14	HTH-type transcriptional regulator DmlR	4.821	0.035289
GDOLJONB_02283	adeQ	Adenine permease AdeQ	4.840	0.0090535
GDOLJONB_01179	rspR	HTH-type transcriptional repressor RspR	4.861	0.0037517
GDOLJONB_02919	csdA	Cysteine desulfurase CsdA	4.917	0.0160562
GDOLJONB_04359	ytfP	Gamma-glutamylcyclotransferase family protein ytfP	4.930	0.0047769
GDOLJONB_02802	norV	Anaerobic nitric oxide reductase flavorubredoxin	4.931	0.0202044
GDOLJONB_04022	ilvE	Branched-chain-amino-acid aminotransferase	4.947	0.0110638
GDOLJONB_00652	mdtH_1	Multidrug resistance protein MdtH	5.012	0.0264536
GDOLJONB_02795	srlE	PTS system glucitol/sorbitol-specific EIIB component	5.013	0.0262738
GDOLJONB_01601	ydhF	Oxidoreductase YdhF	5.051	0.04481
GDOLJONB_03181	-	hypothetical protein	5.066	0.0054619
GDOLJONB_00559	prsE	Type I secretion system membrane fusion protein PrsE	5.074	0.0174198
GDOLJONB_02747	acoR	Acetoin dehydrogenase operon transcriptional activator AcoR	5.084	0.0312166
GDOLJONB_04997	-	hypothetical protein	5.118	0.0036875

GDOLJONB_02238	-	Putative tyrosine-protein kinase in cps region	5.126	0.0049199
GDOLJONB_02960	rsxC_2	Electron transport complex subunit RsxC	5.226	0.0346591
GDOLJONB_01974	proQ	RNA chaperone ProQ	5.257	0.0200997
GDOLJONB_00060	iolU_1	scyllo-inositol 2-dehydrogenase (NADP(+)) IolU	5.262	0.0045968
GDOLJONB_02820	hypA	Hydrogenase maturation factor HypA	5.321	0.0197861
GDOLJONB_01898	-	hypothetical protein	5.323	0.0028416
GDOLJONB_02240	-	hypothetical protein	5.384	0.0038511
GDOLJONB_02862	surE_2	5'/3'-nucleotidase SurE	5.391	0.0481205
GDOLJONB_03906	rbsC_7	Ribose import permease protein RbsC	5.446	0.012818
GDOLJONB_02709	ypjD	Inner membrane protein YpjD	5.520	0.0406796
GDOLJONB_03978	menA	1,4-dihydroxy-2-naphthoate octaprenyltransferase	5.538	0.0085287
GDOLJONB_04664	gltC_4	HTH-type transcriptional regulator GltC	5.607	0.0049199
GDOLJONB_02975	-	hypothetical protein	5.682	0.0255274
GDOLJONB_03553	yhhW	Quercetin 2,3-dioxygenase	5.821	0.0003581
GDOLJONB_00176	entC	Isochorismate synthase EntC	5.865	0.001334
GDOLJONB_03955	rafR	HTH-type transcriptional regulator RafR	5.882	0.0001164
GDOLJONB_04826	yafV	Omega-amidase YafV	5.889	0.0026677
GDOLJONB_04168	malF	Maltose transport system permease protein MalF	6.063	0.0004153
GDOLJONB_02796	srlB_2	PTS system glucitol/sorbitol-specific EIIA component	6.081	0.0179635
GDOLJONB_01820	-	hypothetical protein	6.204	0.0037517
GDOLJONB_04280	cvaA_2	Colicin V secretion protein CvaA	6.301	0.0004637
GDOLJONB_02816	hycD	Formate hydrogenlyase subunit 4	6.463	0.0214804
GDOLJONB_02790	mntB_1	Manganese transport system membrane protein MntB	6.479	0.0202044
GDOLJONB_02036	-	hypothetical protein	6.565	0.0083389
GDOLJONB_03386	sspA	Stringent starvation protein A	6.881	0.0035763

GDOLJONB_01653	acrF_1	Multidrug export protein AcrF	7.313	0.0045676
GDOLJONB_04880	-	hypothetical protein	7.470	1.08E-05
GDOLJONB_02264	ddpC_3	putative D,D-dipeptide transport system permease protein DdpC	7.543	0.0049199
GDOLJONB_05000	cyoC	Cytochrome bo(3) ubiquinol oxidase subunit 3	7.595	0.0001494
GDOLJONB_00345	bluR	HTH-type transcriptional repressor BluR	7.687	0.0047769
GDOLJONB_03145	galP	Galactose-proton symporter	7.771	1.23E-05
GDOLJONB_02739	yadV_2	-	8.030	0.0062394
GDOLJONB_03081	btuD_9	Vitamin B12 import ATP-binding protein BtuD	8.069	3.84E-06
GDOLJONB_01973	prc	Tail-specific protease	8.124	0.0015372
GDOLJONB_04628	yhaI	Inner membrane protein YhaI	8.370	1.80E-05
GDOLJONB_01105	fumA	Fumarate hydratase class I, aerobic	8.605	9.62E-06
GDOLJONB_01139	apbE_1	FAD:protein FMN transferase	8.967	2.94E-06
GDOLJONB_02229	wcaJ	UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase	9.077	7.63E-05
GDOLJONB_02861	pcm	Protein-L-isoaspartate O-methyltransferase	9.257	0.00183
GDOLJONB_04668	-	NADH oxidase	10.145	6.84E-05
GDOLJONB_04927	-	-	10.318	0.0011247
GDOLJONB_00580	-	hypothetical protein	11.376	4.71E-05
GDOLJONB_05217	-	hypothetical protein	-5.168	2.66E-05
GDOLJONB_05218	sopB	Protein SopB	-4.960	0.0001156
GDOLJONB_05220	-	IS3 family transposase ISSen4	-2.862	0.0195054
GDOLJONB_05195	silP	Silver exporting P-type ATPase	2.785	0.0356532
GDOLJONB_05095	-	hypothetical protein	3.950	0.0356532
GDOLJONB_05150	livF_7	High-affinity branched-chain amino acid transport ATP-binding protein LivF	4.655	0.0045861

APPENDIX 3: Colonization-associated transposon sequencing results for Z4160

Gene Identifier	Gene	Product	log2(fold change)	False Discovery Rate
QLD26_10890	pykF	pyruvate kinase PykF	-8.324	2.14E-09
QLD26_10530	adhE	bifunctional acetaldehyde-CoA/alcohol dehydrogenase	-7.339	4.10E-08
QLD26_06745	-	phosphoribosylaminoimidazolesuccinocarb oxamide synthase	-7.061	1.38E-07
QLD26_17960	ybgC	tol-pal system-associated acyl-CoA thioesterase	-6.954	2.64E-07
QLD26_16955	pflB	formate C-acetyltransferase	-6.821	4.57E-07
QLD26_19930	acrB	multidrug efflux RND transporter permease subunit AcrB	-6.754	5.77E-07
QLD26_16950	focA	formate transporter FocA	-6.679	7.11E-07
QLD26_06385	purL	phosphoribosylformylglycinamidine synthase	-6.634	9.21E-07
QLD26_07310	purF	amidophosphoribosyltransferase	-6.454	2.06E-06
QLD26_03570	cpdA	3',5'-cyclic-AMP phosphodiesterase	-6.195	6.26E-06
QLD26_25750	-	MFS transporter	-6.194	1.19E-06
QLD26_07305	cvpA	colicin V production protein	-6.031	1.22E-05
QLD26_02620	arcB	aerobic respiration two-component sensor histidine kinase ArcB	-6.018	1.18E-05
QLD26_00895	mtlD	mannitol-1-phosphate 5-dehydrogenase	-5.990	1.28E-05
QLD26_19995	-	beta-galactosidase	-5.749	2.82E-06
QLD26_14050	fnr	fumarate/nitrate reduction transcriptional regulator Fnr	-5.723	4.04E-05
QLD26_22140	carA	glutamine-hydrolyzing carbamoyl-phosphate synthase small subunit	-5.662	4.74E-05
QLD26_08270	sanA	outer membrane permeability protein SanA	-5.632	7.28E-05
QLD26_20000	-	arabinogalactan endo-beta-1,4-galactanase	-5.595	7.04E-05
QLD26_19925	acrA	multidrug efflux RND transporter periplasmic adaptor subunit AcrA	-5.512	8.99E-05
QLD26_02650	hpf	ribosome hibernation promoting factor	-5.473	0.000111
QLD26_21745	aceE	pyruvate dehydrogenase (acetyl-transferring), homodimeric type	-5.440	1.28E-05

QLD26_16760	pyrD	quinone-dependent dihydroorotate dehydrogenase	-5.397	0.000134
QLD26_02640	rapZ	RNase adapter RapZ	-5.291	0.000225
QLD26_25095	tatC	Sec-independent protein translocase subunit TatC	-5.276	0.00022
QLD26_21240	gmhB	D-glycero-beta-D-manno-heptose 1,7-bisphosphate 7-phosphatase	-5.256	0.000287
QLD26_17940	tolB	Tol-Pal system beta propeller repeat protein TolB	-5.207	0.000287
QLD26_04475	-	glycoside hydrolase family 32 protein	-5.205	0.000287
QLD26_21600	dksA	RNA polymerase-binding protein DksA	-5.201	0.000291
QLD26_01615	gntR	gluconate operon transcriptional repressor GntR	-5.198	0.000292
QLD26_25100	tatB	Sec-independent protein translocase protein TatB	-5.170	0.000331
QLD26_16700	ompA	porin OmpA	-5.159	2.92E-05
QLD26_00695	pyrE	orotate phosphoribosyltransferase	-5.154	0.000344
QLD26_20150	lon	endopeptidase La	-5.097	0.000427
QLD26_18120	pgm	phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent)	-5.054	0.000418
QLD26_27660	sopA	plasmid-partitioning protein SopA	-5.053	0.000472
QLD26_16315	mdoG	glucans biosynthesis protein MdoG	-5.034	0.000212
QLD26_27275	-	fertility inhibition protein FinO	-4.984	0.000595
QLD26_19940	-	HHA domain-containing protein	-4.955	0.000572
QLD26_08225	-	amino acid permease	-4.931	7.53E-05
QLD26_05360	nlpD	murein hydrolase activator NlpD	-4.866	0.000868
QLD26_08070	yejM	LPS biosynthesis-modulating metalloenzyme YejM	-4.835	0.001039
QLD26_25105	tatA	Sec-independent protein translocase subunit TatA	-4.829	0.000914
QLD26_02770	rlmE	23S rRNA (uridine(2552)-2'-O)-methyltransferase RlmE	-4.777	0.001129
QLD26_17980	cydA	cytochrome ubiquinol oxidase subunit I	-4.731	0.001334
QLD26_12370	-	general stress protein	-4.722	0.001629
QLD26_16960	pflA	pyruvate formate lyase 1-activating protein	-4.662	0.002127
QLD26_26050	rsmG	16S rRNA (guanine(527)-N(7))-methyltransferase RsmG	-4.641	0.002256

QLD26_25655	pfkA	6-phosphofructokinase	-4.631	0.002095
QLD26_07115	glk	glucokinase	-4.601	0.000195
QLD26_04465	-	hypothetical protein	-4.561	0.0025
QLD26_08220	yieE	DNA-binding transcriptional regulator YeiE	-4.522	0.00293
QLD26_00785	-	sugar glycosyltransferase	-4.514	0.003207
QLD26_06990	ptsI	phosphoenolpyruvate-protein phosphotransferase PtsI	-4.507	0.003515
QLD26_17935	pal	peptidoglycan-associated lipoprotein Pal	-4.501	0.003366
QLD26_06670	purM	phosphoribosylformylglycinamidine cyclo-ligase	-4.456	0.003816
QLD26_10535	-	hypothetical protein	-4.451	0.003585
QLD26_26190	yidC	membrane protein insertase YidC	-4.429	0.001156
QLD26_16165	fabF	beta-ketoacyl-ACP synthase II	-4.414	0.004333
QLD26_22090	surA	peptidylprolyl isomerase SurA	-4.377	0.004676
QLD26_09895	fadR	fatty acid metabolism transcriptional regulator FadR	-4.357	0.004783
QLD26_25355	trxA	thioredoxin TrxA	-4.354	0.005042
QLD26_16260	pyrC	dihydroorotase	-4.325	0.005838
QLD26_16200	yceD	23S rRNA accumulation protein YceD	-4.322	0.005681
QLD26_25885	glnA	glutamate--ammonia ligase	-4.313	0.006199
QLD26_17945	tolA	cell envelope integrity protein TolA	-4.284	0.006474
QLD26_24830	purH	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	-4.273	0.006633
QLD26_23910	hflX	GTPase HflX	-4.259	0.001334
QLD26_04030	serA	phosphoglycerate dehydrogenase	-4.227	0.007943
QLD26_22135	carB	carbamoyl-phosphate synthase large subunit	-4.220	0.00754
QLD26_01050	ghrB	glyoxylate/hydroxypyruvate reductase GhrB	-4.151	0.009705
QLD26_16190	plsX	phosphate acyltransferase PlsX	-4.142	0.011871
QLD26_03990	epd	erythrose-4-phosphate dehydrogenase	-4.137	0.000291
QLD26_02550	zapG	Z-ring associated protein ZapG	-4.132	0.010658
QLD26_05165	-	flavodoxin	-4.128	0.01073
QLD26_11465	-	winged helix-turn-helix domain-containing protein	-4.115	0.000326
QLD26_21915	cra	catabolite repressor/activator	-4.112	0.010974
QLD26_02655	rpoN	RNA polymerase factor sigma-54	-4.078	0.012262

QLD26_24740	pgi	glucose-6-phosphate isomerase	-4.077	0.013054
QLD26_20160	clpP	ATP-dependent Clp endopeptidase proteolytic subunit ClpP	-4.067	0.001453
QLD26_25340	wzzE	ECA polysaccharide chain length modulation protein	-4.017	0.011731
QLD26_25445	-	DUF413 domain-containing protein	-4.006	0.000932
QLD26_04470	-	MFS transporter	-4.003	0.000244
QLD26_27395	-	hypothetical protein	-3.997	0.012016
QLD26_23920	miaA	tRNA (adenosine(37)-N6)-dimethylallyltransferase MiaA	-3.992	0.013054
QLD26_00770	-	polysaccharide deacetylase family protein	-3.985	0.013521
QLD26_03895	speB	agmatinase	-3.951	0.015468
QLD26_11650	slyB	outer membrane lipoprotein SlyB	-3.888	0.019049
QLD26_19750	purE	5-(carboxyamino)imidazole ribonucleotide mutase	-3.859	0.019647
QLD26_10230	-	DUF6392 family protein	-3.830	0.022852
QLD26_25645	menA	1,4-dihydroxy-2-naphthoate polyprenyltransferase	-3.812	0.007561
QLD26_20790	-	tRNA-Thr	-3.811	0.022852
QLD26_06615	guaB	IMP dehydrogenase	-3.805	0.023899
QLD26_24375	-	type IV toxin-antitoxin system AbiEi family antitoxin	-3.802	0.024944
QLD26_16210	rluC	23S rRNA pseudouridine(955/2504/2580) synthase RluC	-3.800	0.025305
QLD26_06085	-	type II toxin-antitoxin system RatA family toxin	-3.784	0.029475
QLD26_01750	ompR	two-component system response regulator OmpR	-3.770	0.025849
QLD26_06710	bepA	beta-barrel assembly-enhancing protease	-3.766	0.025849
QLD26_01340	pitA	inorganic phosphate transporter PitA	-3.755	0.003692
QLD26_12795	-	NADH:flavin oxidoreductase/NADH oxidase	-3.742	0.028991
QLD26_15000	topA	type I DNA topoisomerase	-3.737	0.031958
QLD26_11595	gloA	lactoylglutathione lyase	-3.728	0.030458
QLD26_01690	malP	maltodextrin phosphorylase	-3.718	0.000547
QLD26_08665	-	glycosyltransferase family 4 protein	-3.691	0.033112

QLD26_01045	-	DUF3053 domain-containing protein	-3.679	0.033601
QLD26_16920	rpsA	30S ribosomal protein S1	-3.604	0.041512
QLD26_07175	-	tRNA-Arg	-3.597	0.049584
QLD26_18125	seqA	replication initiation negative regulator SeqA	-3.597	0.042531
QLD26_22110	apaH	bis(5'-nucleosyl)-tetraphosphatase (symmetrical) ApaH	-3.586	0.042521
QLD26_25635	hslV	ATP-dependent protease subunit HslV	-3.581	0.045586
QLD26_02765	yhbY	ribosome assembly RNA-binding protein YhbY	-3.580	0.0473
QLD26_23470	pyrB	aspartate carbamoyltransferase	-3.568	0.046101
QLD26_23360	-	MurR/RpiR family transcriptional regulator	-3.568	0.009705
QLD26_25310	rffA	dTDP-4-amino-4,6-dideoxygalactose transaminase	-3.552	0.0473
QLD26_17005	lrp	leucine-responsive transcriptional regulator Lrp	-3.540	0.001858
QLD26_25930	dsbA	thiol:disulfide interchange protein DsbA	-3.539	0.048918
QLD26_13640	-	DeoR/GlpR family DNA-binding transcription regulator	-3.538	0.004167
QLD26_23685	tamB	autotransporter assembly complex protein TamB	-3.516	0.00146
QLD26_00830	gpmM	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	-3.494	0.014167
QLD26_06080	smpB	SsrA-binding protein SmpB	-3.479	0.009705
QLD26_07645	-	porin OmpC	-3.472	0.002095
QLD26_05020	gcvA	glycine cleavage system transcriptional regulator GcvA	-3.463	0.001621
QLD26_16665	mgsA	methylglyoxal synthase	-3.443	0.007561
QLD26_07760	-	hypothetical protein	-3.400	0.010238
QLD26_08745	rfbB	O-antigen export ABC transporter ATP-binding protein RfbB	-3.396	0.001964
QLD26_04020	argP	DNA-binding transcriptional regulator ArgP	-3.384	0.001156
QLD26_25290	wecG	lipopolysaccharide N-acetylmannosaminouronosyltransferase	-3.371	0.006479
QLD26_25755	-	alpha-galactosidase	-3.291	0.039885

QLD26_05675	-	DNA-binding transcriptional repressor	-3.183	0.01447
QLD26_21490	mrcB	bifunctional glycosyl transferase/transpeptidase	-3.166	0.007273
QLD26_25400	ilvC	ketol-acid reductoisomerase	-3.127	0.014174
QLD26_04595	rppH	RNA pyrophosphohydrolase	-3.121	0.019647
QLD26_27220	-	thermonuclease family protein	-3.044	0.026846
QLD26_21990	-	sugar efflux transporter	-3.025	0.02825
QLD26_25545	ppc	phosphoenolpyruvate carboxylase	-3.022	0.006255
QLD26_13470	-	D-threonate 4-phosphate dehydrogenase	-2.994	0.031698
QLD26_00975	xylA	xylose isomerase	-2.985	0.007561
QLD26_25305	wzxE	lipid III flippase WzxE	-2.955	0.02036
QLD26_03245	-	beta-galactosidase subunit beta	-2.941	0.005607
QLD26_02685	-	calcium/sodium antiporter	-2.890	0.009105
QLD26_27485	-	helix-turn-helix transcriptional regulator	-2.839	0.00318
QLD26_08535	yegQ	tRNA 5-hydroxyuridine modification protein YegQ	-2.833	0.01027
QLD26_26665	-	transposase domain-containing protein	-2.808	0.004625
QLD26_03255	ebgR	transcriptional regulator EbgR	-2.807	0.013287
QLD26_25890	glnL	nitrogen regulation protein NR(II)	-2.764	0.040895
QLD26_04665	amiC	N-acetylmuramoyl-L-alanine amidase AmiC	-2.738	0.035056
QLD26_08530	yegS	lipid kinase YegS	-2.722	0.040224
QLD26_12025	pqqU	TonB-dependent receptor PqqU	-2.687	0.037067
QLD26_00540	-	YiiQ family protein	-2.682	0.042521
QLD26_08685	wcaJ	undecaprenyl-phosphate glucose phosphotransferase	-2.669	0.037067
QLD26_25905	hemN	oxygen-independent coproporphyrinogen III oxidase	-2.646	0.035767
QLD26_20370	tgt	tRNA guanosine(34) transglycosylase Tgt	-2.634	0.033112
QLD26_12720	-	GNAT family N-acetyltransferase	-2.593	0.018005
QLD26_26600	-	GNAT family N-acetyltransferase	-2.591	0.018227
QLD26_20485	rdgC	recombination-associated protein RdgC	-2.569	0.034089
QLD26_20005	-	sugar ABC transporter permease	-2.503	0.027664
QLD26_13420	-	tagaturonate reductase	-2.493	0.031231
QLD26_03885	yqgB	acid stress response protein YqgB	-2.468	0.033085
QLD26_13230	-	hypothetical protein	-2.429	0.041566
QLD26_25785	-	PTS sugar transporter subunit IIA	-2.350	0.045586

QLD26_09780	manX	PTS mannose transporter subunit IIAB	-2.323	0.042804
QLD26_25935	-	serine/threonine protein kinase	-2.266	0.042804
QLD26_26230	-	IS5-like element IS903B family transposase	2.200	0.041949
QLD26_27320	-	conjugal transfer protein TrbF	2.270	0.047356
QLD26_26265	-	IS3 family transposase	2.282	0.026846
QLD26_27530	-	hypothetical protein	2.284	0.027022
QLD26_27260	-	DUF2726 domain-containing protein	2.302	0.024963
QLD26_15600	-	phage tail protein	2.365	0.045884
QLD26_05140	ppnN	nucleotide 5'-monophosphate nucleosidase PpnN	2.471	0.048252
QLD26_23135	traD	conjugative transfer system coupling protein TraD	2.477	0.035444
QLD26_26540	-	SDR family oxidoreductase	2.504	0.011612
QLD26_19955	-	YlaC family protein	2.627	0.030323
QLD26_26375	pcoR	copper response regulator transcription factor PcoR	2.630	0.031157
QLD26_19490	-	terminase	2.647	0.037891
QLD26_14505	-	Lrp/AsnC family transcriptional regulator	2.653	0.036785
QLD26_20570	sbmA	peptide antibiotic transporter SbmA	2.688	0.03322
QLD26_11735	-	GH1 family beta-glucosidase	2.696	0.018642
QLD26_18945	-	DUF2157 domain-containing protein	2.735	0.03144
QLD26_06625	-	sulfatase-like hydrolase/transferase	2.758	0.037784
QLD26_25590	-	bifunctional UDP-sugar hydrolase/5'-nucleotidase	2.768	0.024944
QLD26_18440	-	LysR family transcriptional regulator	2.783	0.036785
QLD26_17280	-	oligosaccharide MFS transporter	2.783	0.039885
QLD26_15945	pepT	peptidase T	2.800	0.026471
QLD26_16330	ymdB	O-acetyl-ADP-ribose deacetylase	2.815	0.042519
QLD26_14730	-	HD domain-containing protein	2.831	0.016036
QLD26_11850	-	DUF3313 domain-containing protein	2.834	0.033112
QLD26_26595	-	Tn3 family transposase	2.856	0.014538
QLD26_18085	kdpF	K(+)-transporting ATPase subunit F	2.859	0.023417
QLD26_19545	-	recombination protein NinG	2.871	0.040873
QLD26_08760	-	glycosyltransferase	2.871	0.041913

QLD26_27060	-	IS6-like element IS6100 family transposase	2.890	0.019333
QLD26_15975	-	long-chain fatty acid--CoA ligase	2.893	0.03322
QLD26_26215	-	tyrosine-type recombinase/integrase	2.903	0.011612
QLD26_24140	cutA	divalent cation tolerance protein CutA	2.923	0.01376
QLD26_04200	-	LysR family transcriptional regulator	2.924	0.045884
QLD26_25595	-	cytoplasmic protein	2.941	0.034107
QLD26_15370	-	phosphatase PAP2 family protein	2.942	0.019628
QLD26_22380	deoB	phosphopentomutase	2.952	0.010656
QLD26_07965	-	DUF3168 domain-containing protein	2.962	0.045468
QLD26_14985	acnA	aconitate hydratase AcnA	2.966	0.033112
QLD26_24360	-	ABC transporter substrate-binding protein	2.982	0.0334
QLD26_16220	-	LysR family transcriptional regulator	2.985	0.009705
QLD26_21190	-	endonuclease/exonuclease/phosphatase family protein	3.009	0.035767
QLD26_26615	-	IS3 family transposase	3.034	0.014538
QLD26_26135	-	carbohydrate porin	3.038	0.015295
QLD26_21295	-	YaeP family protein	3.058	0.024664
QLD26_24595	tyrB	aromatic amino acid transaminase	3.059	0.008432
QLD26_27550	-	ParB/RepB/Spo0J family partition protein	3.061	0.011612
QLD26_20190	cyoB	cytochrome o ubiquinol oxidase subunit I	3.081	0.00673
QLD26_26390	-	type II toxin-antitoxin system RelE/ParE family toxin	3.084	0.0033
QLD26_25835	fdoI	formate dehydrogenase cytochrome b556 subunit	3.091	0.013054
QLD26_21760	-	MFS transporter	3.112	0.024944
QLD26_26425	-	L-lactate permease	3.116	0.007799
QLD26_07745	-	hypothetical protein	3.126	0.005332
QLD26_22600	-	2-dehydro-3-deoxygalactonokinase	3.128	0.025033
QLD26_15115	-	MarR family winged helix-turn-helix transcriptional regulator	3.131	0.042915
QLD26_01595	yhhY	N-acetyltransferase	3.139	0.014086
QLD26_25070	fadA	acetyl-CoA C-acyltransferase FadA	3.150	0.016662
QLD26_19165	-	HlyD family secretion protein	3.151	0.0331

QLD26_06540	-	alpha-2-macroglobulin	3.158	0.00918
QLD26_00380	-	MDR family MFS transporter	3.171	0.027885
QLD26_20515	-	YaiI/YqxD family protein	3.177	0.030323
QLD26_14075	-	hypothetical protein	3.213	0.028863
QLD26_20580	ampH	D-alanyl-D-alanine-carboxypeptidase/endopeptidase AmpH	3.221	0.018503
QLD26_06520	pepB	aminopeptidase PepB	3.225	0.024365
QLD26_04225	-	LysR family transcriptional regulator	3.248	0.019906
QLD26_24470	-	glutathione S-transferase	3.257	0.031307
QLD26_16365	-	arginase family protein	3.262	0.028335
QLD26_09350	uvrY	UvrY/SirA/GacA family response regulator transcription factor	3.273	0.003816
QLD26_11710	rsxC	electron transport complex subunit RsxC	3.283	0.042804
QLD26_20935	accC	acetyl-CoA carboxylase biotin carboxylase subunit	3.283	0.014167
QLD26_20300	-	aldo/keto reductase	3.336	0.004182
QLD26_21100	-	NADH:ubiquinone reductase (Na(+)-transporting) subunit D	3.343	0.033112
QLD26_26505	arsB	arsenite efflux transporter membrane subunit ArsB	3.355	0.006722
QLD26_00110	-	DUF202 domain-containing protein	3.378	0.024944
QLD26_27475	-	DUF932 domain-containing protein	3.379	0.003165
QLD26_15800	-	UbiD family decarboxylase	3.379	0.010525
QLD26_23120	-	IS3 family transposase	3.382	0.021984
QLD26_18970	-	NUDIX domain-containing protein	3.394	0.011303
QLD26_14175	yedE	selenium metabolism membrane protein YedE/FdhT	3.402	0.009705
QLD26_18160	chiP	chitoporin	3.458	0.001585
QLD26_22425	yjjG	pyrimidine 5'-nucleotidase	3.490	0.013244
QLD26_05460	-	thiamine pyrophosphate-requiring protein	3.491	0.0025
QLD26_13160	-	LysR family transcriptional regulator	3.525	0.025219
QLD26_22265	thrC	threonine synthase	3.537	0.035002
QLD26_16930	aroA	3-phosphoshikimate 1-carboxyvinyltransferase	3.548	0.015874
QLD26_12030	-	YncE family protein	3.591	0.002108
QLD26_12040	ansP	L-asparagine permease	3.606	0.008278

QLD26_21755	aroP	aromatic amino acid transporter AroP	3.634	0.001684
QLD26_25180	recQ	ATP-dependent DNA helicase RecQ	3.643	0.008703
QLD26_26855	lacI	DNA-binding transcriptional repressor LacI	3.646	0.014538
QLD26_03875	-	sugar porter family MFS transporter	3.655	0.005512
QLD26_11370	-	LysR family transcriptional regulator	3.664	0.002631
QLD26_24685	malK	maltose/maltodextrin ABC transporter ATP-binding protein MalK	3.702	0.00077
QLD26_03630	-	YgiQ family radical SAM protein	3.703	0.002816
QLD26_15540	-	DUF488 family protein	3.709	0.021368
QLD26_16690	-	hypothetical protein	3.717	0.023211
QLD26_13930	mqo	malate dehydrogenase (quinone)	3.717	0.00558
QLD26_20835	-	metallophosphoesterase	3.735	0.001328
QLD26_13750	-	GntR family transcriptional regulator	3.753	0.00146
QLD26_11790	-	membrane-bound PQQ-dependent dehydrogenase, glucose/quinate/shikimate family	3.755	0.016272
QLD26_08020	-	hypothetical protein	3.756	0.010332
QLD26_08980	-	transposase	3.783	0.006927
QLD26_20715	-	CS1-pili formation C-terminal domain-containing protein	3.804	0.009705
QLD26_25125	rmuC	DNA recombination protein RmuC	3.847	0.01185
QLD26_19060	-	APC family permease	3.849	0.030729
QLD26_11575	-	AraC family transcriptional regulator	3.853	0.001935
QLD26_10455	-	DMT family transporter	3.857	0.010252
QLD26_18925	-	PTS system mannose/fructose/sorbose family transporter subunit IID	3.889	0.000964
QLD26_11420	-	ABC transporter permease	3.908	0.022456
QLD26_22910	-	HlyD family efflux transporter periplasmic adaptor subunit	3.910	0.005042
QLD26_14535	-	aminotransferase class I/II-fold pyridoxal phosphate-dependent enzyme	3.920	0.011871
QLD26_25450	hdfR	HTH-type transcriptional regulator HdfR	3.949	0.000775
QLD26_16710	-	Lon protease family protein	3.967	0.000331

QLD26_12280	-	ABC transporter ATP-binding protein	4.029	0.003136
QLD26_21280	nlpE	envelope stress response activation lipoprotein NlpE	4.040	0.001039
QLD26_02455	csrD	RNase E specificity factor CsrD	4.047	0.013287
QLD26_22270	thrB	homoserine kinase	4.054	0.003585
QLD26_16980	dmsA	dimethylsulfoxide reductase subunit A	4.054	0.002307
QLD26_23410	-	GNAT family N-acetyltransferase	4.062	0.003816
QLD26_13055	-	alpha/beta hydrolase	4.066	0.000524
QLD26_18850	-	DMT family transporter	4.099	0.003136
QLD26_11815	-	carboxylesterase/lipase family protein	4.103	0.007965
QLD26_23065	-	MFS transporter	4.103	0.004182
QLD26_24910	thiG	thiazole synthase	4.139	0.006633
QLD26_24265	-	alpha-glucosidase/alpha-galactosidase	4.145	0.016036
QLD26_02440	msrQ	protein-methionine-sulfoxide reductase heme-binding subunit MsrQ	4.170	0.002647
QLD26_22340	sltY	murein transglycosylase	4.193	0.003276
QLD26_10700	-	MetQ/NlpA family lipoprotein	4.205	0.000572
QLD26_25520	oxyR	DNA-binding transcriptional regulator OxyR	4.251	0.006633
QLD26_17035	clpA	ATP-dependent Clp protease ATP-binding subunit ClpA	4.270	0.000168
QLD26_08310	ascB	6-phospho-beta-glucosidase	4.275	0.004574
QLD26_26420	-	IS3-like element ISEc52 family transposase	4.277	0.00318
QLD26_24145	-	protein-disulfide reductase DsbD	4.291	0.000212
QLD26_26605	-	ISNCY family transposase	4.294	1.54E-05
QLD26_07640	rcsD	phosphotransferase RcsD	4.300	0.000399
QLD26_20505	aroL	shikimate kinase AroL	4.311	0.001453
QLD26_14820	-	6-phospho-beta-glucosidase	4.311	0.013521
QLD26_25210	yigB	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YigB	4.330	0.011871
QLD26_20785	proA	glutamate-5-semialdehyde dehydrogenase	4.496	0.00038
QLD26_13140	-	MFS transporter	4.542	0.003988
QLD26_14695	hcp	type VI secretion system effector Hcp	4.563	0.000932
QLD26_27240	repA	plasmid replication initiator RepA	4.565	0.000472

QLD26_04105	-	protein-disulfide reductase DsbD family protein	4.574	0.004117
QLD26_25375	-	amidohydrolase family protein	4.599	0.001878
QLD26_18280	gltK	glutamate/aspartate ABC transporter permease GltK	4.628	0.0069
QLD26_01685	malT	HTH-type transcriptional regulator MalT	4.629	7.36E-05
QLD26_21970	-	MFS transporter	4.676	0.000235
QLD26_13150	-	sugar diacid recognition domain-containing protein	4.705	0.000285
QLD26_27195	-	IS66 family transposase	4.707	3.15E-06
QLD26_22990	-	CusA/CzcA family heavy metal efflux RND transporter	4.716	0.004312
QLD26_20720	-	fimbrial chaperone EcpB	4.744	0.002307
QLD26_23915	hfq	RNA chaperone Hfq	4.752	0.00243
QLD26_04565	lplT	lysophospholipid transporter LplT	4.771	0.005103
QLD26_25540	argE	acetylornithine deacetylase	4.817	0.000308
QLD26_22245	tal	transaldolase	4.836	0.001382
QLD26_18995	-	fimbrial protein	4.883	0.000547
QLD26_20215	-	IclR family transcriptional regulator C-terminal domain-containing protein	4.938	0.004105
QLD26_24700	malG	maltose ABC transporter permease MalG	4.962	4.23E-05
QLD26_25345	wecA	UDP-N-acetylglucosamine--undecaprenyl-phosphate N-acetylglucosaminephosphotransferase	4.968	0.001323
QLD26_16250	mdtH	multidrug efflux MFS transporter MdtH	5.052	0.000326
QLD26_24695	malF	maltose ABC transporter permease MalF	5.060	3.47E-05
QLD26_18840	-	PLP-dependent aminotransferase family protein	5.067	4.53E-05
QLD26_07730	-	ead/Ea22-like family protein	5.077	0.005633
QLD26_15350	gdhA	NADP-specific glutamate dehydrogenase	5.083	0.005145
QLD26_19310	-	fimbrial protein	5.156	0.000111
QLD26_20420	fba	class II fructose-1,6-bisphosphate aldolase	5.185	0.000689
QLD26_22695	hpaB	4-hydroxyphenylacetate 3-monooxygenase, oxygenase component	5.199	0.000646

QLD26_06795	nudK	GDP-mannose pyrophosphatase NudK	5.230	0.000104
QLD26_24345	-	ATP-binding protein	5.279	0.003305
QLD26_10580	cls	cardiolipin synthase	5.289	0.001585
QLD26_24690	malE	maltose/maltodextrin ABC transporter substrate-binding protein MalE	5.302	3.70E-06
QLD26_18045	-	DUF969 domain-containing protein	5.303	1.32E-05
QLD26_06845	eutM	ethanolamine utilization microcompartment protein EutM	5.335	4.12E-05
QLD26_07190	fadL	long-chain fatty acid transporter FadL	5.337	2.55E-06
QLD26_13380	tam	trans-aconitate 2-methyltransferase	5.383	0.000545
QLD26_16410	efeB	iron uptake transporter deferrochelatase/peroxidase subunit	5.384	9.29E-06
QLD26_06725	bcp	thioredoxin-dependent thiol peroxidase	5.391	1.42E-05
QLD26_01160	-	cellulose synthase operon protein YhjQ/BcsQ	5.393	3.45E-05
QLD26_18865	-	efflux RND transporter permease subunit	5.421	0.000212
QLD26_15705	-	Dam family site-specific DNA-(adenine-N6)-methyltransferase	5.428	6.22E-05
QLD26_22755	-	DUF445 domain-containing protein	5.440	0.000203
QLD26_21085	nqrM	(Na+)-NQR maturation NqrM	5.492	0.001052
QLD26_14165	-	hypothetical protein	5.578	6.54E-05
QLD26_19665	-	hypothetical protein	5.585	0.000412
QLD26_24580	-	MmcQ/YjbR family DNA-binding protein	5.688	0.000438
QLD26_21790	gspE	type II secretion system protein GspE	5.690	0.000646
QLD26_19765	-	porin	5.843	3.89E-05
QLD26_03890	speA	biosynthetic arginine decarboxylase	5.940	0.000329
QLD26_19985	-	maltoporin	5.986	5.85E-05
QLD26_22350	nadR	multifunctional transcriptional regulator/nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase NadR	6.002	5.81E-05

QLD26_20480	mak	fructokinase	6.029	0.000773
QLD26_07635	rcsB	response regulator transcription factor RcsB	6.030	1.19E-06
QLD26_21770	ampE	beta-lactamase regulator AmpE	6.134	2.16E-05
QLD26_01645	glgX	glycogen debranching protein GlgX	6.183	1.54E-05
QLD26_19535	-	bacteriophage antitermination protein Q	6.336	0.000286
QLD26_24845	zraS	two-component system sensor histidine kinase ZraS	6.366	2.98E-05
QLD26_14125	fumA	class I fumarate hydratase FumA	6.588	0.000111
QLD26_25145	metR	HTH-type transcriptional regulator MetR	6.599	3.11E-05
QLD26_16965	-	MFS transporter	6.705	6.74E-05
QLD26_26155	-	NCS2 family permease	7.023	6.22E-05
QLD26_19170	-	DUF3302 domain-containing protein	7.110	2.07E-06
QLD26_23580	-	DUF4311 domain-containing protein	7.116	1.28E-05
QLD26_13050	-	DUF2255 family protein	7.233	1.09E-08
QLD26_13760	-	fructuronate reductase	7.252	7.10E-06
QLD26_17880	pnuC	nicotinamide riboside transporter PnuC	7.427	6.54E-05
QLD26_20020	ugpC	sn-glycerol-3-phosphate ABC transporter ATP-binding protein UgpC	7.435	6.70E-06
QLD26_11130	-	iron ABC transporter permease	7.556	3.33E-07
QLD26_18080	-	DUF2517 family protein	7.696	8.13E-05
QLD26_23350	iolD	3D-(3,5/4)-trihydroxycyclohexane-1,2-dione acylhydrolase (decyclizing)	7.738	7.89E-05
QLD26_24190	-	LuxR C-terminal-related transcriptional regulator	7.933	1.06E-08
QLD26_24200	pgaA	poly-beta-1,6 N-acetyl-D-glucosamine export porin PgaA	7.965	4.64E-09
QLD26_24215	pgaD	poly-beta-1,6-N-acetyl-D-glucosamine biosynthesis protein PgaD	8.330	1.32E-08
QLD26_24210	pgaC	poly-beta-1,6-N-acetyl-D-glucosamine synthase	8.488	2.52E-09
QLD26_24205	pgaB	poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase PgaB	8.810	3.61E-09
QLD26_11480	-	hypothetical protein	8.908	5.79E-07

QLD26_10670	ihfA	integration host factor subunit alpha	8.916	5.77E-07
QLD26_19485	-	DUF1073 domain-containing protein	8.999	3.14E-06
QLD26_18355	dacA	D-alanyl-D-alanine carboxypeptidase DacA	9.222	2.11E-06
QLD26_20920	-	ABC transporter substrate-binding protein	9.273	9.29E-06
QLD26_17770	modB	molybdate ABC transporter permease subunit	9.603	1.70E-06
QLD26_02860	nlpI	lipoprotein NlpI	9.660	2.52E-09
QLD26_14305	-	GFA family protein	9.668	2.36E-08
QLD26_21055	frsA	esterase FrsA	10.275	2.28E-06
QLD26_16915	ihfB	integration host factor subunit beta	10.505	4.35E-09
QLD26_09700	proQ	RNA chaperone ProQ	10.588	3.64E-08
QLD26_23215	-	DUF2157 domain-containing protein	10.950	2.26E-09
QLD26_09705	prc	carboxy terminal-processing peptidase	10.953	2.52E-09

APPENDIX 4: **Shared colonization factors between strains**

CRE-166 and KPN46	CRE-166 and Z4160	KPN46 and Z4160
ackA	amiC	acrB
bamB	argP	apaH
cpxR	bepA	carA
fabR	carB	epd
galT	dksA	fabF
hypothetical protein	dsbA	guaB
mlc	ganB	hypothetical protein
mnmG	glk	hypothetical protein
pepA	gmhB	melY
recG	gntR	pfkA
rph	hflX	purM
sopB	higA1	pyrC
	hypothetical protein	rafA
	hypothetical protein	rlmE
	malG	rpoN
	malP	rppH
	menA	seqA
	ompA	surA
	ompR	tatB
	pal	ybgC
	pgm	
	purL	
	ratA	
	smpB	
	speB	
	tgt	
	tolB	
	topA	
	wecG	
	xylA	
	yejM	

APPENDIX 5: Shared genes in which transposon insertion confers a colonization advantage

CRE-166 and KPN46	CRE-166 and Z4160	KPN46 and Z4160	CRE-166, KPN46, and Z4160
gudD	nlpI	mdtH	malE
rafR	hypothetical protein - glycosyltransferase	fumA	malF
	dacA	ansP2	malG
	lamB_2	galP	malT
	malK_1	ylaC	prc
	slt		proQ
	hypothetical protein - amidohydrolase family protein		rspR

VITA

EDUCATION

MD-PhD Student

Northwestern University Feinberg School of Medicine (entry: 2017)

B.S. Biomedical Engineering *cum laude*, with distinction in the major

Yale University (2012-2016)

RESEARCH TRAINING

Doctoral Research: Alan Hauser Laboratory, Northwestern University. 2019-2023

Genomic and murine model approaches to identify genes involved in three strains of *Klebsiella pneumoniae*

- Developed mouse model of robust *K. pneumoniae* GI Colonization with precise antibiotic dosing
- Created transposon insertion mutant libraries covering the genomes for three phylogenetically distinct *K. pneumoniae* strains
- Performed Insertion Sequencing in an *in vivo* model of GI Colonization and validated both shared and strain-specific colonization factors *in vivo*

Research Associate: Chirag Parikh Laboratory, Yale School of Medicine. 2016-2017

Clinical coordinator for recruitment of participants to identify biomarkers for acute kidney injury

- Obtained informed consent and enrolled over 100 participants in a study to identify biomarkers for diagnosis of Acute Interstitial Nephritis
- Developed and implemented a phone survey which revealed that participants had positive attitudes towards acquisition of extra kidney biopsy specimens in procedure for research

Research Assistant: Themis Kyriakides Laboratory, Yale University. 2015-2016

Engineering of bulk metallic glass nanorod arrays to modulate the foreign body response

- Created bulk metallic glass (BMG) nanorod arrays and quantified the effects of substrate topography and stiffness on primary macrophage fusion in the foreign body response
- Implemented high-throughput screening of BMG alloys for bacterial growth inhibition

Research Assistant: Heidi Noels Laboratory, RWTH Aachen. Summer 2014.

Isolation of platelets and conducting luminescence assays for enzyme activity

- Isolated and tested human platelets for DPP4 enzyme activity to elucidate mechanism of improved cardiovascular outcome following treatment with DPP4 inhibitors

Research Assistant: Choukri Ben Mamoun Laboratory, Yale University. Summer 2013.

Cloning of a biosensor and enzyme purification and kinetics

- Cloned membrane phospholipid biosensor for transfection into malarial parasites
- Enzyme production, purification, and kinetics characterizations

PUBLICATIONS

Kochan, TJ, Nozick SH, **Cheung BH**, Gatesy SWM, Lebrun Corbin M, Mitra SD, Khalatyan N, Krapp F, Zi C, Ozer EA, Hauser AR. (2022). Genomic surveillance for multidrug-resistant or hypervirulent *Klebsiella pneumoniae* among United States bloodstream isolates. *BMC Infect Dis, 22*: 603.

Bachta KER, Allen JP, **Cheung BH**, Chiu CH, Hauser AR. (2020). Systemic infection facilitates transmission of Pseudomonas aeruginosa in mice. *Nat Commun, 11*:543.

Shayan M, Padmanabhan J, Morris AH, **Cheung B**, et al. (2018). Nanopatterned bulk metallic glass-based biomaterials modulate macrophage polarization. *Acta Biomaterialia, 75*, 427-38.

Moledina DG, **Cheung B**, et al. (2017). A Survey of Patient Attitudes Toward Participation in Biopsy-Based Kidney Research. *Kidney International Reports, 3*(2), 412-6.

Padmanabhan J, Augelli MJ, **Cheung B**, et al. (2016). Regulation of cell-cell fusion by nanotopography. *Scientific Reports, 6*, 33277.

Liu Y, Padmanabhan J, **Cheung B**, et al. (2016). Combinatorial development of antibacterial Zr-Cu-Al-Ag thin film metallic glasses. *Scientific Reports, 6*, 26950.

PRESENTATIONS

4. "Identification of factors contributing to gastrointestinal colonization by *Klebsiella pneumoniae*." **Poster presentation** at the American Society of Microbiology: Microbe meeting, 2023 (Houston, TX).

3. "Genome-wide screens reveal shared and strain-specific genes necessary for enteric colonization by *Klebsiella pneumoniae*." **Oral presentation** at KlebClub, 2023. (Webinar Series)

2. "Identification of factors contributing to gastrointestinal colonization by *Klebsiella pneumoniae*." **Poster presentation and NIH Abstract Award** at the Midwest Microbial Pathogenesis Conference, 2022 (Madison, WI)

1. "A murine model of *Klebsiella pneumoniae* gastrointestinal colonization with parenteral vancomycin administration." **Poster presentation** at ID Week, 2021 (Virtual)

HONORS/AWARDS

2018	**5T32GM008152-33**, NIH/NIGMS
2015	**George J. Schulz Fellowship in the Physical Sciences**, Yale University
2014	**Silliman Sherwood Fellowship**, Yale University
2013	**Freshman Summer Research Fellowship in the Sciences and Engineering**, Yale University
2012-2016	**National Merit Scholarship Finalist**, Yale University

TEACHING EXPERIENCE

2020	**Teaching Assistant**, Intro to Life Sciences Research
2019 – 2020	**Clinical microbiology review session developer and leader**, Undergraduate Medical Education
2019 – 2020	**Question writer**, Undergraduate Medical Education Study Bank

LEADERSHIP

2018 – present	**President and Co-Founder**, Northwestern Medical Orchestra
2014 – 2015	**President and Publicity Officer**, Coup de Brass Horn Choir, Yale University

2013 – 2015	**Co-President and Librarian**, Davenport Pops Orchestra, Yale University
2013 – 2016	**Work Leader**, Yale University Stacks Department

PRIOR WORK EXPERIENCE

2016	**Research Assistant**, Linguistics Department, Yale University, New Haven CT
2012 – 2016	**Page and Work Leader**, Yale University Stacks Department, New Haven, CT
2009 – 2012	**Page,** Ardsley Public Library, Ardsley NY