

1 **Genome-wide screens reveal shared and strain-specific genes that facilitate**
2 **enteric colonization by *Klebsiella pneumoniae***

3 Bettina H Cheung¹, Arghavan Alisoltani^{1,2,3}, Travis J Kochan¹, Marine Lebrun-Corbin¹,
4 Sophia H Nozick¹, Christopher MR Axline¹, Kelly ER Bachta², Egon A Ozer^{2,3}, and Alan
5 R Hauser^{1,2#}

6 ¹Department of Microbiology-Immunology, Northwestern University, Feinberg
7 School of Medicine, Chicago, IL, USA

8 ²Division of Infectious Diseases, Department of Medicine, Northwestern
9 University, Feinberg School of Medicine, Chicago, IL, USA

10 ³Center for Pathogen Genomics and Microbial Evolution, Havey Institute for
11 Global Health, Northwestern University Feinberg School of Medicine, Chicago,
12 IL, USA

13 # Corresponding author: Alan R Hauser (ahauser@northwestern.edu)

14 **RUNNING TITLE** Colonization factors across 3 *Klebsiella* strains

15 **ABSTRACT**

16 Gastrointestinal (GI) colonization by *Klebsiella pneumoniae* is a risk factor for
17 subsequent infection as well as transmission to other patients. Additionally, colonization
18 is achieved by many strain types that exhibit high diversity in genetic content. Thus, we
19 aimed to study strain-specific requirements for *K. pneumoniae* GI colonization by
20 applying transposon insertion sequencing to three classical clinical strains: a
21 carbapenem-resistant strain, an extended-spectrum beta-lactamase producing strain,

22 and a non-epidemic antibiotic-susceptible strain. The transposon insertion libraries were
23 screened in a murine model of GI colonization. At three days post-inoculation, 27 genes
24 were required by all three strains for colonization. Isogenic deletion mutants for three
25 genes/operons (*acrA*, *carAB*, *tatABCD*) confirmed colonization defects in each of the
26 three strains. Additionally, deletion of *acrA* reduced bile tolerance *in vitro*, while
27 complementation restored both bile tolerance *in vitro* and colonization ability *in vivo*.
28 Transposon insertion sequencing suggested that some genes were more important for
29 colonization of one strain than the others. For example, deletion of the sucrose porin-
30 encoding gene *scrY* resulted in a colonization defect in the carbapenemase-producing
31 strain but not in the extended-spectrum beta-lactamase producer or the antibiotic-
32 susceptible strain. These findings demonstrate that classical *K. pneumoniae* strains use
33 both shared and strain-specific strategies to colonize the mouse GI tract.

34 **IMPORTANCE** *Klebsiella pneumoniae* is a common cause of difficult-to-treat infections
35 due to its propensity to express resistance to many antibiotics. For example,
36 carbapenem-resistant *K. pneumoniae* (CR-Kp) has been named an urgent threat by the
37 United States Centers for Disease Control and Prevention. Gastrointestinal colonization
38 of patients with *K. pneumoniae* has been linked to subsequent infection, making it a key
39 process to control in prevention of multidrug-resistant infections. However, the bacterial
40 factors which contribute to *K. pneumoniae* colonization are not well understood.
41 Additionally, individual strains exhibit large amounts of genetic diversity, begging the
42 question of whether some colonization factors are strain-dependent. This study
43 identifies the enteric colonization factors of 3 classical strains using transposon mutant

44 screens to define a core colonization program for *K. pneumoniae* as well as detecting
45 strain-to-strain differences in colonization strategies.

46 INTRODUCTION

47 As multidrug-resistant bacteria continue to pose a looming threat to our ability to
48 treat infections, alternative methods for controlling disease burden—such as infection
49 prevention—are increasingly important. For *Klebsiella pneumoniae*, a highly multidrug-
50 resistant bacteria, gastrointestinal (GI) colonization is a risk factor for subsequent
51 infection. Patients are prone to infections caused by the strains they carry and may also
52 spread them to other hospitalized patients^{1,2}. Thus, GI colonization by *K. pneumoniae* is
53 an attractive target for infection prevention.

54 The genes required for *K. pneumoniae* gut colonization are not well understood.
55 Studies of intestinal colonization are complicated by the fact that this species exhibits
56 substantial genetic diversity. While the average *K. pneumoniae* genome contains about
57 5000-6000 genes, only ~1700 are shared by most strains^{3,4}. Studies have identified
58 colonization factors for individual strains⁵⁻⁷, but whether *K. pneumoniae*'s genetic
59 diversity gives rise to strain-to-strain differences in GI colonization strategies remains
60 unclear.

61 Two broad categories of *K. pneumoniae* strains have been described: classical
62 and hypervirulent. Of the classical strains, several have achieved global dominance and
63 are referred to as high-risk clones. Two such high-risk clones are the ST45 and ST258
64 sequence types. ST45 strains frequently produce extended-spectrum beta-lactamases
65 (ESBLs), making them resistant to most beta-lactams except carbapenems. Many

66 ST258 strains produce carbapenemases, causing a significant portion of carbapenem-
67 resistant infections worldwide⁸. While it is conceivable that the global success of these
68 lineages has resulted in part from their capacity to better colonize the GI tract, how their
69 colonization strategies might differ from those of antibiotic-susceptible strains has not
70 been investigated.

71 Here, we selected three representative clinical classical strains of *K. pneumoniae*
72 to investigate GI colonization by lineages of varying epidemicity and antibiotic
73 resistance. We used transposon insertion sequencing to identify genes in each strain
74 that were required for GI colonization in a mouse model. We found that a subset of
75 genes was required by all three strains, but each strain relied on a much larger set of
76 additional genes that were dispensable for the other strains. Thus, the set of genes, and
77 therefore presumed GI colonization strategies, vary substantially across
78 phylogenetically distinct *K. pneumoniae* strains.

79 RESULTS

80 Three clinical strains of *K. pneumoniae* with distinct genetic and phenotypic 81 characteristics

82 We selected three clinical isolates of *K. pneumoniae* representative of strains
83 with varying levels of epidemic spread and antibiotic resistance. First, we chose CRE-
84 166, a carbapenem-resistant strain of the ST258 high-risk clone with the *bla*_{KPC} gene,
85 which was isolated from bronchioalveolar lavage. Second, we selected Z4160, an ESBL
86 producer isolated from the bloodstream with both a widespread ESBL gene (*bla*_{CTX-M-15})
87 and an epidemic sequence type (ST45). Third, we chose KPN46, a non-epidemic,

88 antibiotic-susceptible bloodstream isolate (ST433). The sizes of the CRE-166, Z4160,
89 and KPN46 genomes were 6.00, 5.56, and 5.63 Mb, respectively. The core genome
90 shared between them was 4.96 Mb (4,662 coding sequences [CDS]), leaving CRE-166,
91 Z4160, and KPN46 with 1.04, 0.63, 0.69 Mb (1183, 623, and 749 CDS) of accessory
92 genetic content (Figure 1). Thus, CRE-166, Z4160, and KPN46 represented three
93 clinical strains with phenotypic and genomic diversity suitable for subsequent studies of
94 *K. pneumoniae* GI colonization.

95 **A clinically relevant murine model of GI colonization**

96 To model hospitalized patients receiving antibiotics and to induce robust fecal
97 shedding of *K. pneumoniae* in C57BL/6 mice, we administered an antibiotic regimen
98 prior to oral gavage with bacteria. Vancomycin, one of the most highly utilized antibiotics
99 in the United States⁹, was administered through daily intraperitoneal injection for 5 days
100 (350 mg/kg, equivalent to a human dose of 1 g/day)¹⁰ (Figure 2A). In contrast to other
101 screens for GI colonization factors in which antibiotics were administered through
102 drinking water and throughout the screen⁵⁻⁷, we opted for injections to control dosage to
103 each mouse and to allow for precise adjustments in future studies. We also ceased
104 administration prior to *K. pneumoniae* inoculation to investigate colonization after
105 antibiotic exposure. Following the last dose of vancomycin, 10⁸ colony forming units
106 (CFU) of CRE-166 were inoculated by orogastric gavage. *K. pneumoniae* was then
107 selectively cultured from feces by plating on lysogeny broth (LB) agar supplemented
108 with carbenicillin, an antibiotic to which all *K. pneumoniae* are resistant⁴. Culture of
109 feces on this medium prior to inoculation with *K. pneumoniae* yielded no colonies,
110 confirming specificity for experimentally introduced *K. pneumoniae* (Supplemental

111 Figure 1A). These conditions supported 10^{10} CFU/g fecal shedding of *K. pneumoniae* in
112 the first week followed by shedding of 10^7 CFU/g for at least 60 days post-gavage in
113 both male and female mice (Figure 2B). At Day 14, no CFU could be detected in the
114 lung, liver, or spleen, indicating there was no dissemination from the gut (Supplemental
115 Figure 1B). Additionally, throughout the infection, mice did not exhibit signs of illness,
116 suggesting that they were colonized rather than infected.

117 To ensure these conditions would produce similar levels of GI colonization with
118 KPN46 and Z4160, we individually inoculated CRE-166, KPN46, and Z4160 into mice.
119 Fecal burdens were similar across 14 days (Figure 2C). These results indicated that the
120 vancomycin-treated mice were a robust model of *K. pneumoniae* GI colonization.

121 To determine whether transposon insertion sequencing experiments would be
122 informative in our model, we next ensured that mutants would not randomly drop out of
123 the fecal output due to bottlenecks rather than colonization defects. To this end, we
124 constructed a marked CRE-166 strain by inserting an apramycin-resistance cassette
125 into the chromosomal Tn7 site. This marked strain did not have a growth defect in LB
126 when compared to the parental strain (Supplemental Figure 2A). To mimic the presence
127 of a single transposon mutant within the pool of total mutants in a screen, we spiked this
128 marked strain into an inoculum at a ratio of 1:100,000 with the parental strain. Next, we
129 measured the ratio of the marked strain to total *K. pneumoniae* recovered from the
130 feces of the mice. At Day 3 post-gavage, the marked strain was still detectable,
131 suggesting the absence of a bottleneck significant enough to bias transposon
132 sequencing results (Supplemental Figures 2B & 2C). However, at subsequent
133 timepoints, we failed to recover the marked strain from some of the mice, indicating

134 greater bottlenecks. We therefore chose Day 3 post-gavage as the timepoint for
135 transposon sequencing experiments.

136 **Generation of highly saturated transposon mutant libraries**

137 To perform genome-wide screens for GI colonization factors, we generated
138 transposon mutant libraries in all three *K. pneumoniae* strains. The transposon vector
139 pSAMerm was modified to express hygromycin resistance (pSAMhygSDM) to allow for
140 selection of transposition in all three strains. Libraries with over 145,000 mutants were
141 generated. An initial assessment of library quality was performed by randomly selecting
142 32 colonies from each library and identifying transposon insertion sites with arbitrary
143 PCR. Unique insertion sites for at least 26 colonies for each strain were successfully
144 identified, and no colonies had more than one insertion site, indicating the libraries were
145 of high quality.

146 **Screening the mutant libraries for GI colonization factors *in vivo***

147 We gavaged each of the three transposon mutant libraries into mice pre-treated
148 with vancomycin. A portion of the inoculum was saved as the “input pool.” At Day 3, we
149 collected fecal pellets, or the “output pools.” Input pool sequencing demonstrated that
150 over 82% of coding sequences had at least one insertion, and on average, each gene
151 had 5 insertions. Coverage was distributed across chromosomes (Supplemental Figure
152 3), confirming that all libraries were well-saturated.

153 Insertion site sequencing reads were then processed using a modified version of
154 the previously described ESSENTIALS pipeline¹¹. We first analyzed the input pools to
155 identify “essential genes” required for the bacteria to grow in LB. A total of 487 genes

156 were identified as essential in all three strains, but a substantial number of genes were
157 essential in only one or two strains (Figure 3A, Supplemental Table 1). Between 14-
158 18% of CDS in each strain were found to be essential, comparable to the estimates of
159 11-17% reported for other strains of *K. pneumoniae*^{5,12-14}.

160 To determine which genes each strain utilized for establishing GI colonization,
161 we compared the total number of insertion reads per gene in the Day 3 output pool to
162 those of the input pools. We focused on genes that had a less than $-2 \log_2(\text{fold-change})$
163 (logFC) in output vs. input insertion reads and a false discovery rate (FDR) less than
164 0.05 (Figure 4, Supplemental Table 2). Multidimensional scaling (MDS) plots of input
165 and Day 3 output pools demonstrated that the input pools were closely related and
166 distinct from the Day 3 output pools for each strain (Supplemental Figure 4), indicating
167 our approach was technically robust.

168 Twenty-seven genes were used by all three strains for GI colonization (Figure
169 3B). However, many genes were used by only one strain to establish colonization: 88
170 for CRE-166, 83 for Z4160, and 34 for KPN46. Intriguingly, most genes identified as
171 important for colonization in at least one strain were present in all 3 strains. That is, only
172 22.6% of colonization genes for CRE-166, 3.3% for KPN46, and 6.8% for Z4160 were
173 absent from the genomes of the other strains, suggesting that these strains mostly rely
174 on shared genes to establish GI colonization, but use different sets of these genes for
175 this purpose.

176 To determine whether these colonization genes were also found in broader
177 populations of *K. pneumoniae* strains, we calculated a core genome (genes shared by
178 95% of strains) from a set of 323 previously described strains⁴. Upon comparison with

179 this broader core genome, somewhat larger percentages of the colonization genes for
180 each strain were now considered accessory genes: 25.8% for CRE-166, 16.1% for
181 KPN46, and 15.4% for Z4160. However, most genes required for colonization by each
182 strain were still genes shared across *K. pneumoniae* strains.

183 In addition to genes required for colonization, we also identified genes which,
184 upon disruption with a transposon, conferred a colonization advantage (red points in
185 Figure 4). There were 7 genes found to confer an advantage when disrupted in all 3
186 strains (Supplemental Table 2). Four were involved in maltose transport (*malt*,
187 *maltEFG*) while the other three had regulatory roles (*proQ*, *prc*, and *rspR*).

188 **Classification of GI colonization factors**

189 To better understand the core colonization program in *K. pneumoniae*, we
190 focused on the 27 genes important for colonization across all three strains (Table 1). As
191 expected, we identified genes involved in anaerobic metabolism (e.g., *adhE*, *fnr*, *focA*).
192 We also found genes involved in other metabolic pathways, including *mtlD* (mannitol-1-
193 phosphate dehydrogenase) and *carAB*, which encode the subunits of carbamoyl
194 phosphate synthase that are responsible for the first committed step in synthesis of
195 pyrimidine and arginine¹⁵. *carA* was identified in two strains and *carB* in the remaining
196 strain. *tatA* and *tatC* (folded protein secretion apparatus¹⁶) and *acrA* (efflux pump¹⁷)
197 were also identified.

198 To examine the pathways that strains relied on for colonization, we assigned
199 Kyoto Encyclopedia of Genes and Genomes (KEGG) identifiers to all genes in each
200 strain and determined which pathways were enriched among colonization hits. These

201 pathways fell into a few broad categories: metabolism, antimicrobial resistance, protein
202 secretion, and environmental sensing (Figure 5). Metabolic capacities played an
203 important role in the ability of bacteria to colonize the gut, but individual pathways
204 identified differed by strain. Additionally, the colonization factors for CRE-166 and
205 KPN46 were enriched for two-component systems, which may have played a role in
206 metabolic adjustments caused by environmental sensing in the GI tract. In terms of
207 antimicrobial resistance pathways, two strains (CRE-166 and KPN46) were reliant on
208 genes that conferred resistance to cationic antimicrobial peptides (CAMPs), which are
209 released by colonic epithelium and are similar to microcins released by the microbiota.
210 Thus, defense against host and microbiome factors is likely key to colonization by *K.*
211 *pneumoniae*. KPN46 colonization factors were enriched for genes for resistance to
212 beta-lactams, including efflux pumps, which play a role in the efflux of toxic compounds.
213 Finally, protein export (the Tat secretion system) was enriched for KPN46 and Z4160,
214 suggesting that secreted proteins may enhance colonization.

215 **Validation of colonization genes using isogenic mutants**

216 To validate our screen, we created isogenic mutants of 3 genetic loci—*acrA*,
217 *carAB*, and *tatABCD*—required for GI colonization in all 3 strains (Table 1). These loci
218 were chosen because they represent different functional groups: antimicrobial
219 resistance, metabolism, and secretion. We generated isogenic mutants in which the
220 coding sequence of the target was replaced by an apramycin-resistance cassette. We
221 verified that these apramycin-resistant mutants did not have growth defects in LB when
222 grown individually (Supplemental Figure 5). Of these mutants, only the *tatABCD* mutant
223 had a slight defect in LB when competed against their marked (hygromycin-resistant)

224 parental strains (Supplemental Figure 6). Then, we inoculated 1:1 mixtures of the
225 marked parental strains and mutants into the mouse model of GI colonization and
226 enumerated CFU in the feces at Day 3 (the screen timepoint) to calculate competitive
227 indices (CI). To characterize the effects of these mutants at later timepoints, we also
228 followed the fecal burdens to Day 14.

229 In all three strain backgrounds, *acrA* mutants displayed significant colonization
230 defects at Day 3 (validating our screen) as well as beyond to Day 14 (Figure 6A-C). We
231 constructed a complemented strain with an unmarked deletion of the *acrA* locus,
232 inserting *acrA* along with its upstream region and a downstream apramycin-resistance
233 cassette into the chromosomal Tn7 site. This complement rescued the colonization
234 defect (Figure 6D).

235 The *carAB* deletion mutants were similarly tested in competition with their
236 parental strains. At Day 3, each *carAB* mutant exhibited colonization defect, continuing
237 to Day 14 for CRE-166 and Z4160 (Figure 6E-G). For the KPN46 mutant, greater
238 variability in CI was observed at later timepoints, suggesting the existence of a priority
239 effect in the second week, during which mutants that initially established themselves
240 tended to do very well while the others did progressively more poorly. Due to technical
241 limitations, a *carAB* complement could not be constructed. However, we performed
242 whole genome sequencing on all mutants, confirming that the CRE-166 and Z4160
243 mutants did not have off-site mutations likely to be responsible for observed
244 phenotypes. For KPN46, sequencing indicated that two nonsynonymous mutations
245 emerged during the course of mutant generation. We performed an *in vivo* competition
246 experiment between two marked KPN46 parental strains—one with these off-site

247 mutations and one without—to show that the mutations did not confer a colonization
248 defect (Supplemental Figure 7). These data indicate that the *carAB* deletions are
249 responsible for the colonization defects observed in Figure 6E-G.

250 Finally, deletion of the *tatABCD* operon also significantly decreased colonization
251 capacities, both at Day 3 and throughout subsequent days (Figure 6H-J). Insertion of
252 the *tatABCD* operon at the Tn7 site fully rescued the colonization defect (Figure 6K).
253 Thus, we verified our ability to detect shared factors essential for GI colonization.

254 In addition to shared factors, we wanted to confirm strain-specific colonization
255 factors. Hemolysin expression-modulating protein, encoded by *hha*, is a transcriptional
256 regulator which scored as a hit in Z4160 but not in KPN46 and CRE-166. We generated
257 *hha* deletion mutants in both Z4160 and KPN46, neither of which had *in vitro* growth
258 defects (Supplemental Figure 5). At Day 3, the Δhha mutant had a statistically
259 significant colonization defect in Z4160 (Figure 6L). In KPN46, the Δhha mutant had a
260 slightly less severe colonization defect that was not statistically significant (Figure 6M).

261 As a second strain-specific factor, we selected *scrY*, which encodes for a
262 sucrose porin, and which was identified in our screen as a colonization factor for CRE-
263 166 but not KPN46 or Z4160. We created in-frame deletions to preserve the remainder
264 of the operon downstream from *scrY* and found that CRE-166 $\Delta scrY$ but not
265 KPN46 $\Delta scrY$ or Z4160 $\Delta scrY$ had a colonization defect (Figure 6O-Q). Together, these
266 data indicate that colonization factors may differ in their importance from strain to strain.

267 Finally, we also selected one target that exhibited a colonization *advantage* upon
268 disruption. We chose *maltT*, the transcriptional regulator for maltose uptake and

269 metabolism. A *malT* deletion mutant in CRE-166 did not have a growth advantage in LB
270 (Supplemental Figure 5), but this deletion conferred a substantial colonization
271 advantage over the parent strain *in vivo* (Figure 6N), indicating our screen was also
272 valid for detection of genes that confer colonization advantages.

273 **Deletion of *acrA* reduces resistance to ox bile**

274 We further explored how one of our colonization genes, *acrA*, may contribute to
275 GI colonization. As *acrA* encodes a component of an efflux pump that contributes to bile
276 resistance in other GI pathogens¹⁸, we tested whether our *acrA* mutant was more
277 sensitive to bile. At 2 and 24 hours after inoculation into 10% bile, the marked parental
278 strain grew significantly better than the CRE-166 Δ *acrA*::AprR mutant (Figure 7). We
279 also generated a CRE-166 Δ *acrA* mutant in which the apramycin-resistance cassette
280 was removed from the Δ *acrA* allele; this mutant also showed a growth defect in bile. A
281 complemented strain (CRE-166 Δ *acrA* Tn7::*acrA*) generated from this mutant showed
282 that complementation rescued resistance to bile. Together, these data suggest that
283 *acrA* supports *K. pneumoniae* GI colonization by providing resistance to bile.

284 **DISCUSSION**

285 In this study, we aimed to answer the following question: does the genetic
286 diversity of *K. pneumoniae* affect the colonization strategies of different strains? We
287 compared three clinically-relevant strains (a globally distributed ST258 strain with a
288 carbapenemase gene, an epidemic ST45 strain with an ESBL gene, and a non-
289 epidemic antibiotic-susceptible strain) and identified a core set of genes used by all
290 three strains to colonize the GI tract. However, we found other genes and pathways

291 unique to one or two strains, highlighting the diversity of colonization strategies in this
292 genetically diverse species. In particular, we confirmed the importance of three different
293 factors (*acrA*, *carAB*, *tatABCD*) in colonization for all three strains to validate our
294 screens.

295 Our results defined a core colonization program of 27 genes utilized by all 3
296 strains. Most of these genes (16 of 27) were related to metabolism, as expected since
297 metabolic adaption to the anaerobic colon is a prerequisite for successful colonization.
298 In addition, we also identified three genetic loci involved in pyrimidine and purine
299 synthesis (*carAB* and *purC/purH*, respectively). Beyond metabolic genes, all strains
300 relied on *yeiE*, which encodes a transcriptional activator. The genes regulated by YeiE
301 in *K. pneumoniae* are not well characterized, but in *Salmonella enterica* this activator
302 also controls GI colonization¹⁹, targeting flagellar genes (which *K. pneumoniae* does not
303 possess). In addition, *tatA* and *tatC*, which encode components of the Tat folded-protein
304 secretion apparatus, were identified as elements of the core colonization program,
305 suggesting secreted factors may contribute to colonization, as they do for several gut
306 pathogens²⁰⁻²³. In particular, the Tat-secreted peptidoglycan amidases, AmiA and AmiC,
307 are necessary for colonization by *S. typhimurium*²⁴, and *amiC* was identified in our CRE-
308 166 screen (Supplemental Table 2). However, disruption of the Tat system also
309 destabilizes the cell envelope²⁵, decreasing resistance to bile acids²⁰. Genes encoding
310 the porin OmpC (*ompC*) and components of the Tol-Pal system (*tolA* and *pal*) were also
311 identified as critical for colonization for all 3 strains. In addition to allowing diffusion of
312 small solutes, OmpC is responsible for maintaining outer membrane leaflet
313 asymmetry²⁶. In a different fashion, the Tol-Pal system also aids in maintaining the

314 integrity of the outer membrane²⁷. Deletion of *pal* in *K. pneumoniae* increases sensitivity
315 to bile, one of the host-derived stresses encountered in the GI tract²⁸. A few other genes
316 implicated in bile resistance (*cvpA* and *acrA*) were critical for colonization. In summary,
317 the core GI colonization program of *K. pneumoniae* is composed of genes involved in
318 energy generation, nucleotide biosynthesis, protein secretion, membrane homeostasis,
319 and bile resistance.

320 We found many genes that contributed to colonization by one or two strains but
321 not by all three strains. For instance, pathway analysis revealed that unlike the other
322 two strains, Z4160 colonization factors were enriched for alanine, aspartate, and
323 glutamate metabolism pathways. In addition, we found that CRE-166 depends on the
324 sucrose porin *scrY* for colonization whereas Z4160 and KPN46 do not. It is possible that
325 redundancy in either sucrose uptake or functional redundancy in other metabolic
326 pathways can compensate for *scrY* deletion in Z4160 and KPN46 but not CRE-166.
327 These findings support our hypothesis that colonization strategies differ between strains
328 of *K. pneumoniae*.

329 The identification of *acrA* as necessary for GI colonization in all three strains has
330 translational implications. AcrA is the periplasmic subunit of the tripartite efflux pumps
331 that contain TolC and AcrB or AcrD¹⁷. In *E. coli* and *S. enterica*, these pumps export a
332 large variety of substrates, including multiple classes of antibiotics and bile acids^{17,29}.
333 We demonstrated that the *acrA* mutant in CRE-166 was more susceptible to bile (Figure
334 7), suggesting a similar function in *K. pneumoniae*. Furthermore, the *acrA* deletion
335 mutants were undetectable at 14 days post-inoculation in most of our competition
336 experiments (Figure 6A-C). Because of its role in antibiotic resistance, the AcrAB efflux

337 pump has been extensively studied, and several small molecule inhibitors are in varying
338 stages of pre-clinical development³⁰. We postulate that these inhibitors may have
339 efficacy in preventing or eradicating *K. pneumoniae* GI colonization.

340 CRE-166 belongs to the globally disseminated ST258 group of carbapenemase-
341 producing strains, and there is substantial interest in understanding how these strains
342 successfully colonize patients. Jung and colleagues previously performed transposon
343 insertion sequencing on a different ST258 strain, MH258, and our results have
344 similarities. First, both studies found that disruption of genes involved in maltose
345 metabolism conferred a fitness advantage in the GI tract. The lambda phage uses
346 maltoporin LamB as its receptor³¹, so it is possible that the maltose system serves as a
347 receptor for a lytic phage, bacteriocin, or other bacterial competition systems. Second, 9
348 of 35 MH258 colonization factors were also important for CRE-166. However, we also
349 found 150 genes to be important for GI colonization with CRE-166 that were not
350 identified for MH258. These differences may reflect genetic variation between the
351 ST258 strains or differences in screen analysis and antibiotic regimens.

352 Our study had several limitations. All transposon insertion sequencing screens
353 have a propensity to miss secreted factors because of trans-complementation. This may
354 be one reason why we did not detect Type VI secretion systems in our screens even
355 though they have been shown to contribute to colonization^{32,33}. Interestingly, we did
356 detect *amiC*, which encodes for a Tat-secreted substrate, but this substrate may remain
357 localized within bacteria. Additionally, our screen was performed on Day 3 of
358 colonization—the latest timepoint at which we did not observe a substantial bottleneck.
359 Factors necessary for colonization on Day 3 may be different from those necessary for

360 longer term colonization. However, these factors would likely be relevant for the
361 development of prophylactic therapies against colonization, and our mutant studies
362 show that deletion of several genes required for Day 3 colonization were also necessary
363 at Day 14 (Figure 6). Additionally, we found that one strain-specific factor, *scrY*, had a
364 stronger phenotype in targeted deletion and validation than another factor, *hha*. This
365 may be explained by outlier read counts that have an inflated effect on the averages
366 used to calculate log fold change. This type of technical limitation of the analysis may
367 explain why *toIC*, which is usually complexed to *acrA*, was identified as a colonization
368 factor for CRE-166 but not the two other strains. Another limitation is that we used as
369 our input pool the mutant library that was inoculated into mice rather than a mutant
370 library passaged in LB. As a result, some of the genes we identified as necessary for GI
371 colonization were likely necessary for normal bacterial growth in rich medium. Such
372 genes, although mechanistically less interesting, are in the strictest sense still
373 necessary for colonization. Our study used only 3 strains, all of which were classical
374 strains. The conserved genome between these clinical strains was 4558 CDS, which is
375 larger than the core content between larger collections of *K. pneumoniae* strains
376 (around 1700 CDS). Additional studies will be necessary to determine whether our
377 findings can be extrapolated to a larger number of *K. pneumoniae* strains, including
378 hypervirulent strains. Finally, our studies were done with C57BL/6 mice; experiments
379 with different mouse strains are necessary to determine whether mouse background
380 affects the genes necessary for colonization.

381 In conclusion, our study found that while different strains of *K. pneumoniae* rely
382 on different genes and pathways to colonize the GI tract, there is a core set of

383 colonization factors used by multiple strains. Inhibition of the proteins encoded by
384 shared genes could theoretically block colonization by nearly all *K. pneumoniae* strains,
385 whereas inhibition of lineage-specific factors could selectively block colonization by
386 specific genotypes.

387 **MATERIALS AND METHODS**

388 **Bacterial strains and cultures**

389 CRE-166, KPN46, and Z4160 are *K. pneumoniae* clinical isolates from
390 Northwestern Memorial Hospital in Chicago collected between 2014 and 2015. CRE-
391 166 and KPN46 were previously described^{34,35}, whereas Z4160 was first used in the
392 current study. *E. coli* strain PIR1 was used for cloning, and *E. coli* β 3914
393 (diaminopimelic acid auxotroph) was used to mate plasmids into *K. pneumoniae*^{36,37}.

394 Bacteria were grown in LB 37°C unless otherwise stated. When appropriate, the
395 following antibiotics were added: carbenicillin (100 μ g/mL), hygromycin (100 μ g/mL), or
396 apramycin (50 μ g/mL). Medium for β 3914 was supplemented with 10 μ g/mL of DAP.

397 **Preparation of complete genomes**

398 Genomic DNA was extracted from CRE-166, KPN46, and Z4160 and sequenced
399 on Illumina and Nanopore platforms to create complete genomes. Annotation was
400 performed using the NCBI Prokaryotic Genome Annotation Pipeline (Z4160)³⁸ or Prokka
401 (CRE-166 and KPN46)³⁹.

402 **Identification of shared genes**

403 To identify genes shared between the three *K. pneumoniae* strains, we used the
404 program Spine and defined shared coding sequences as those with >85% sequence
405 homology between strains⁴⁰. We defined the core genome of *K. pneumoniae* as
406 sequences present in 95% of 323 strains from a previous study⁴ (with homologous
407 genes defined as those with >85% similarity between strains).

408 **Murine model of GI Colonization**

409 Six- to eight-week-old C57BL/6 mice (Jackson Laboratories) received 5 daily
410 intraperitoneal injections of vancomycin (350 mg/kg, Hospira) unless otherwise
411 indicated. For gavage with individual strains, inocula of 10⁸ CFU in 50 µl of PBS were
412 used. Mice received daily cage changes to minimize coprophagy. CFU were
413 enumerated by homogenization of fecal pellets in PBS with the Benchmark Bead
414 Blaster 24 (Benchmark Scientific) followed by serial dilution and plating on LB agar with
415 carbenicillin.

416 Transposon insertion sequencing experiments were performed as above. Frozen
417 aliquots of the transposon libraries were revived for 2 hours in 25 mL of LB. CFU in
418 fecal pellets were quantified as above, and DNA was extracted from the homogenates
419 with the Maxwell 16 system.

420 For competitive colonization experiments, inocula of 1:1 mixtures of a
421 hygromycin-resistant parental strain and an isogenic apramycin-resistant mutant (10⁸
422 CFU each) were created. Fecal CFU burdens were enumerated as above by plating on
423 LB agar with hygromycin or apramycin. Competitive indices (CIs) were calculated as the
424 ratio of mutant CFU to parent strain CFU, normalized to the input ratio.

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

428 **Construction of transposon mutant libraries**

429 A suicide plasmid suitable for Himar1 mariner transposon mutagenesis in highly
430 antibiotic-resistant strains was generated from pSAMerm (gift from G. Pier⁴¹). Briefly,
431 pSAMerm was modified by replacement of the erythromycin-resistance cassette with a
432 hygromycin-resistance cassette (HygR), and the Mmel sites in HygR were removed by
433 site-directed mutagenesis. The resulting plasmid, pSAMhygSDM, was transformed into
434 *E. coli* β 3914 to generate a donor strain for conjugation with *K. pneumoniae* recipient
435 strains. To select for transconjugants and eliminate β 3914, we plated on LB agar
436 supplemented with hygromycin (but lacking DAP) and incubated at 37°C overnight.
437 Colonies were scraped and resuspended in LB with 25% glycerol. The resulting library
438 was stored at -80°C.

439 **Arbitrary PCR for library quality control**

440 Genomic DNA was extracted from 32 randomly selected colonies from each
441 library, and two rounds of nested PCR were performed to amplify the transposon
442 insertion site for subsequent Sanger sequencing. Primer sequences are listed in
443 Supplemental Table 3.

444 **Preparation of DNA for transposon insertion sequencing**

445 DNA extracted from fecal pellets was prepared for insertion site sequencing
446 using the method of Kazi and colleagues⁴². Three replicates for each input (technical)

447 and output (biological) pool were prepared, plus one technical replicate of an output
448 sample. DNA was sheared to ~250 bp fragments with the E220 ultrasonicator (Covaris),
449 and insertion sites were amplified with addition of capture/sequencing sites and
450 barcodes. Final library pool concentrations were quantified with Kapa library
451 quantification (Roche) and sequenced on an Illumina MiSeq.

452 **Transposon insertion sequencing data analysis**

453 A modified version of the previously described ESSENTIALS pipeline¹¹ was used
454 to identify genes necessary for growth in LB and genes that contributed to
455 colonization⁴³.

456 **Pathway analysis**

457 KEGG identifiers were assigned to all genes using BlastKOALA, and KEGG
458 pathways were assigned using KEGG Mapper^{44,45}. A hypergeometric test was
459 conducted in R (v4.2.2) to determine which KEGG pathways were enriched for
460 colonization factors.

461 **Isogenic mutant construction**

462 Deletion mutants were created using lambda red recombination as previously
463 described⁴⁶. All mutants were confirmed by whole-genome sequencing on Illumina
464 platforms.

465 **Bile assays**

466 Dehydrated ox bile was resuspended in water (10% w/v, Sigma) and filtered
467 through a 0.2 µm filter. Strains were grown to an OD₆₀₀ of 1.0, and 1 mL of each culture

468 was pelleted and resuspended in 1 mL PBS. Each strain was inoculated (100 μ l) into
469 900 μ l of 10% ox bile, PBS, or LB and incubated with shaking at 37°C. At 0, 2, 4, and 24
470 hours, aliquots were removed and plated on LB agar for CFU enumeration.

471 **Data availability**

472 The complete genomes of CRE-166, KPN46, and Z4160 have been deposited,
473 and respective accession numbers are: GCA_016797255.2, GCA_021272285.2,
474 GCA_030019695.1.

475 **ACKNOWLEDGEMENTS**

476 We thank Mark Mandel for the β 3914 strain and David Amici for comments on
477 the manuscript. Support for this work was provided by the National Institutes of Health
478 grants R01 AI118257, U19 AI135964, K24 AI04831, R21 AI164254, R21 AI153953
479 (A.R.H) and T32GM008152 (B.H.C), American Cancer Society grant 134251-CSDG-20-
480 053-01-MPC (K.E.R.B), and American Heart Association grant 837089 (T.K.).

481 **BIBLIOGRAPHY**

- 482 1. Giannella M, Trearichi EM, De Rosa FG, et al. Risk factors for carbapenem-resistant
483 *Klebsiella pneumoniae* bloodstream infection among rectal carriers: a prospective observational
484 multicentre study. *Clinical Microbiology and Infection* 2014;20:1357–62.
- 485 2. Gorrie CL, Mirčeta M, Wick RR, et al. Gastrointestinal Carriage Is a Major Reservoir of
486 *Klebsiella pneumoniae* Infection in Intensive Care Patients. *Clinical Infectious Diseases: An*
487 *Official Publication of the Infectious Diseases Society of America* 2017;65:208.
- 488 3. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic surveillance
489 framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex.
490 *Nature Communications* 2021;12:1–16.
- 491 4. Holt KE, Wertheim H, Zadoks RN, et al. Genomic analysis of diversity, population
492 structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to
493 public health. *Proceedings of the National Academy of Sciences of the United States of America*
494 2015;112:E3574.

- 495 5. Jung H-J, Littmann ER, Seok R, et al. Genome-Wide Screening for Enteric Colonization
496 Factors in Carbapenem-Resistant ST258 *Klebsiella pneumoniae*. *mBio* 2019;10.
- 497 6. Maroncle N, Balestrino D, Rich C, Forestier C. Identification of *Klebsiella pneumoniae*
498 genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis.
499 *Infection and Immunity* 2002;70:4729–34.
- 500 7. Struve C, Forestier C, Krogfelt KA. Application of a novel multi-screening signature-
501 tagged mutagenesis assay for identification of *Klebsiella pneumoniae* genes essential in
502 colonization and infection. *Microbiology* 2003;149:167–76.
- 503 8. Cerqueira GC, Earl AM, Ernst CM, et al. Multi-institute analysis of carbapenem resistance
504 reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks.
505 *Proceedings of the National Academy of Sciences of the United States of America*
506 2017;114:1135.
- 507 9. Baggs J, Fridkin SK, Pollack LA, Srinivasan A, Jernigan JA. Estimating National Trends in
508 Inpatient Antibiotic Use Among US Hospitals From 2006 to 2012. *JAMA Internal Medicine*
509 2016;176:1639–48.
- 510 10. Administration. UFD. Estimating the Maximum Safe Starting Dose in Initial Clinical Trials
511 f. 2018.
- 512 11. Zomer A, Burghout P, Bootsma HJ, Hermans PWM, van Hijum SAFT. ESSENTIALS:
513 Software for Rapid Analysis of High Throughput Transposon Insertion Sequencing Data. *PLOS*
514 *ONE* 2012;7:e43012.
- 515 12. Jana B, Cain AK, Doerrler WT, et al. The secondary resistome of multidrug-resistant
516 *Klebsiella pneumoniae*. *Scientific Reports* 2017;7.
- 517 13. Ramage B, Erolin R, Held K, et al. Comprehensive Arrayed Transposon Mutant Library of
518 *Klebsiella pneumoniae* Outbreak Strain KPNH1. *J Bacteriol* 2017;199.
- 519 14. Bachman MA, Breen P, Deornellas V, et al. Genome-Wide Identification of *Klebsiella*
520 *pneumoniae* Fitness Genes during Lung Infection. *MBio* 2015;6:e00775.
- 521 15. Charlier D, Nguyen Le Minh P, Roovers M. Regulation of carbamoylphosphate synthesis
522 in *Escherichia coli*: an amazing metabolite at the crossroad of arginine and pyrimidine
523 biosynthesis. *Amino Acids* 2018;50:1647–61.
- 524 16. Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway.
525 *Nature Reviews Microbiology* 2012;10:483–96.
- 526 17. Li X-Z, Plésiat P, Nikaido H. The Challenge of Efflux-Mediated Antibiotic Resistance in
527 Gram-Negative Bacteria. *Clinical Microbiology Reviews* 2015;28:337.
- 528 18. Thanassi DG, Cheng LW, Nikaido H. Active efflux of bile salts by *Escherichia coli*. *J*
529 *Bacteriol* 1997;179:2512.
- 530 19. Westerman TL, McClelland M, Effenbein JR. YeiE Regulates Motility and Gut Colonization
531 in *Salmonella enterica* Serotype *Typhimurium*. *mBio* 2021;12.
- 532 20. Reynolds MM, Bogomolnaya L, Guo J, et al. Abrogation of the twin arginine transport
533 system in *Salmonella enterica* serovar *Typhimurium* leads to colonization defects during
534 infection. *PloS One* 2011;6:e15800.
- 535 21. Zhang L, Zhu Z, Jing H, et al. Pleiotropic effects of the twin-arginine translocation system
536 on biofilm formation, colonization, and virulence in *Vibrio cholerae*. *BMC Microbiology*
537 2009;9:114.

- 538 22. Lavander M, Ericsson SK, Bröms JE, Forsberg Å. The Twin Arginine Translocation System
539 Is Essential for Virulence of *Yersinia pseudotuberculosis*. *Infection and Immunity* 2006;74:1768.
- 540 23. Rajashekara G, Drozd M, Gangaiyah D, Jeon B, Liu Z, Zhang Q. Functional characterization
541 of the twin-arginine translocation system in *Campylobacter jejuni*. *Foodborne Pathogens and*
542 *Disease* 2009;6:935–45.
- 543 24. Fujimoto M, Goto R, Hirota R, et al. Tat-exported peptidoglycan amidase-dependent cell
544 division contributes to *Salmonella Typhimurium* fitness in the inflamed gut. *PLoS Pathogens*
545 2018;14.
- 546 25. Stanley NR, Findlay K, Berks BC, Palmer T. *Escherichia coli* Strains Blocked in Tat-
547 Dependent Protein Export Exhibit Pleiotropic Defects in the Cell Envelope. *J Bacteriol*
548 2001;183:139.
- 549 26. Chong Z-S, Woo W-F, Chng S-S. Osmoporin OmpC forms a complex with MlaA to
550 maintain outer membrane lipid asymmetry in *Escherichia coli*. *Molecular Microbiology*
551 2015;98:1133–46.
- 552 27. Szczepaniak J, Press C, Kleanthous C. The multifarious roles of Tol-Pal in Gram-negative
553 bacteria. *FEMS Microbiology Reviews* 2020;44:490.
- 554 28. Hsieh P-F, Liu J-Y, Pan Y-J, et al. *Klebsiella pneumoniae* Peptidoglycan-Associated
555 Lipoprotein and Murein Lipoprotein Contribute to Serum Resistance, Antiphagocytosis, and
556 Proinflammatory Cytokine Stimulation. *Journal of Infectious Diseases* 2013;208:1580–9.
- 557 29. Nishino K, Latifi T, Groisman EA. Virulence and drug resistance roles of multidrug efflux
558 systems of *Salmonella enterica* serovar *Typhimurium*. *Molecular Microbiology* 2006;59:126–41.
- 559 30. Sharma A, Gupta VK, Pathania R. Efflux pump inhibitors for bacterial pathogens: From
560 bench to bedside. *Indian Journal of Medical Research* 2019;149:129.
- 561 31. Randall-Hazelbauer L, Schwartz M. Isolation of the Bacteriophage Lambda Receptor
562 from *Escherichia coli*. *J Bacteriol* 1973.
- 563 32. Calderon-Gonzalez R, Lee A, Lopez-Campos G, et al. Modelling the Gastrointestinal
564 Carriage of *Klebsiella pneumoniae* Infections. *MBio* 2023;e0312122.
- 565 33. Merciecca T, Bornes S, Nakusi L, et al. Role of *Klebsiella pneumoniae* Type VI secretion
566 system (T6SS) in long-term gastrointestinal colonization. *Scientific Reports* 2022;12:16968.
- 567 34. Bulman ZP, Krapp F, Pincus NB, et al. Genomic Features Associated with the Degree of
568 Phenotypic Resistance to Carbapenems in Carbapenem-Resistant *Klebsiella pneumoniae*.
569 *mSystems* 2021;6.
- 570 35. Kochan TJ, Nozick SH, Medernach RL, et al. Genomic surveillance for multidrug-resistant
571 or hypervirulent *Klebsiella pneumoniae* among United States bloodstream isolates. *BMC*
572 *Infectious Diseases* 2022;22.
- 573 36. Wang N, Ozer EA, Mandel MJ, Hauser AR. Genome-Wide Identification of *Acinetobacter*
574 *baumannii* Genes Necessary for Persistence in the Lung. *mBio* 2014;5.
- 575 37. Le Roux F, Binesse J, Saulnier D, Mazel D. Construction of a *Vibrio splendidus* Mutant
576 Lacking the Metalloprotease Gene *vsm* by Use of a Novel Counterselectable Suicide Vector.
577 *Applied and Environmental Microbiology* 2007;73:777.
- 578 38. Tatusova T, DiCuccio M, Badretdin A, et al. NCBI prokaryotic genome annotation
579 pipeline. *Nucleic Acids Research* 2016;44:6614.
- 580 39. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*
581 2014;30:2068–9.

- 582 40. Ozer EA, Allen JP, Hauser AR. Characterization of the core and accessory genomes of
583 *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. BMC Genomics
584 2014;15:737.
- 585 41. Skurnik D, Roux D, Aschard H, et al. A comprehensive analysis of *in vitro* and *in vivo*
586 genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon
587 libraries. PLoS Pathogens 2013;9:e1003582.
- 588 42. Kazi MI, Schargel RD, Boll JM. Generating Transposon Insertion Libraries in Gram-
589 Negative Bacteria for High-Throughput Sequencing. JoVE (Journal of Visualized Experiments)
590 2020:e61612.
- 591 43. Ozer EA. *egonozer/essentials_local*: Version 2.1. 2.1 ed: Zenodo; 2023.
- 592 44. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for
593 Functional Characterization of Genome and Metagenome Sequences. Journal of Molecular
594 Biology 2016;428:726–31.
- 595 45. Kanehisa M, Sato Y, Kawashima M. KEGG mapping tools for uncovering hidden features
596 in biological data. Protein Science : a Publication of the Protein Society 2022;31:47–53.
- 597 46. Huang T-W, Lam I, Chang H-Y, Tsai S-F, Palsson BO, Charusanti P. Capsule deletion via a
598 λ -Red knockout system perturbs biofilm formation and fimbriae expression in *Klebsiella*
599 *pneumoniae* MGH 78578. BMC Research Notes 2014;7:13.

600

601

Table 1 Genes contributing to GI colonization in all 3 strains of *K. pneumoniae*.

Gene	Annotation	log ₂ (Fold change)		
		CRE-166	KPN46	Z4160
<i>aceE</i>	Pyruvate dehydrogenase E1 component	-4.412	-5.504	-5.440
<i>acrA</i>	Multidrug efflux pump subunit	-5.906	-6.935	-5.512
<i>adhE</i>	Aldehyde-alcohol dehydrogenase	-3.353	-4.848	-7.339
<i>arcB</i>	Aerobic respiration control sensor protein	-5.342	-3.821	-6.018
<i>bglY</i>	Beta-galactosidase	-6.948	-2.896	-5.749
<i>cvpA</i>	Colicin V production protein	-4.611	-4.449	-6.031
<i>cydA</i>	Cytochrome bd-I ubiquinol oxidase subunit 1	-4.387	-6.933	-4.731
<i>fnr</i>	Fumarate and nitrate reduction regulatory protein	-4.248	-6.542	-5.723
<i>focA</i>	Formate Transporter	-6.486	-7.092	-6.679
<i>glnA</i>	Glutamine synthetase	-3.667	-3.755	-4.313
<i>miaA</i>	tRNA dimethylallyltransferase	-5.773	-3.856	-3.992
<i>mtlD</i>	Mannitol-1-phosphate 5-dehydrogenase	-5.711	-6.709	-5.990
<i>ompC</i>	Outer membrane porin C	-5.738	-3.129	-3.472
<i>pal</i>	Peptidoglycan-associated lipoprotein	-4.172	-6.183	-4.662
<i>pflA</i>	Pyruvate formate-lyase 1-activating enzyme	-5.617	-7.979	-6.821
<i>pflB</i>	Formate acetyltransferase 1	-4.766	-5.270	-4.077
<i>pgi</i>	Glucose-6-phosphate isomerase	-5.404	-5.665	-4.507
<i>ptsI</i>	Phosphoenolpyruvate-protein phosphotransferase	-3.088	-6.240	-7.061
<i>purC</i>	Phosphoribosylaminoimidazole-succinocarboxamide synthase	-3.470	-4.634	-4.273
<i>purH</i>	Bifunctional purine biosynthesis protein	-3.710	-5.209	-8.324
<i>pykF</i>	Pyruvate kinase I	-4.431	-4.701	-3.025
<i>setA</i>	Sugar efflux transporter A	-6.943	-5.168	-5.053
<i>tatA</i>	Sec-independent protein translocase protein	-3.384	-6.061	-4.829
<i>tatC</i>	Sec-independent protein translocase protein	-3.452	-5.452	-5.276
<i>tolA</i>	Tol-Pal system protein	-5.977	-5.965	-4.284
<i>yeiE</i>	HTH-type transcriptional activator	-5.445	-4.572	-4.522
-	Hypothetical polysaccharide deacetylase	-5.327	-4.906	-3.985
<i>carA</i>	Carbamoyl phosphate synthase small subunit	-	-3.917	-5.662
<i>carB</i>	Carbamoyl phosphate synthase large subunit	-5.267	-	-4.220

*For each strain, genes with log₂fold change < -2 and FDR < 0.05 were compared. For the *carAB* two-gene operon, *carB* met these criteria for CRE-166 whereas *carA* met these criteria for KPN46 and Z4160.

602

603 FIGURE LEGENDS

604

605 **Figure 1** Unique and shared coding sequences between 3 strains of *K. pneumoniae*.
 606 The software program Spine was used to identify coding sequences with at least 85%
 607 homology in the strains indicated.

608 **Figure 2** Mouse model of GI colonization with *K. pneumoniae*. (A) Schematic of the *in*
 609 *vivo* model. Mice were administered 5 days of 350 mg/kg intraperitoneal vancomycin
 610 injections before orogastric gavage with 10⁸ CFU *K. pneumoniae*. Fecal samples were
 611 collected after gavage and CFU enumerated. (B) Fecal burden of CRE-166 following
 612 gavage into male (square) or female (circle) mice with (red) or without (black)
 613 vancomycin treatment prior to gavage. n = 5 for each group. (C) Fecal burdens of
 614 strains CRE-166, KPN46, and Z4160 in mouse model of GI colonization. n = 10 for each

615 group. Line indicates median. Limit of detection was 10^2 CFU/g feces, denoted by a
616 dotted line.

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

622 **Figure 4** Volcano plots showing results of transposon insertion sequencing experiments
623 in mice. Genes in which insertions were associated with decreased GI colonization
624 (blue dots) and increased colonization (red dots) for (A) CRE-166, (B) KPN46, and (C)
625 Z4160 are shown. Labeled points indicate targets that were chosen for the creation of
626 isogenic mutants. FDR, false discovery rate; FC, fold change.

627 **Figure 5** Pathways used for GI colonization in 3 strains of *K. pneumoniae*. Kyoto
628 Encyclopedia of Genes and Genomes (KEGG) identifiers were assigned to all genes in
629 the genomes of CRE-166, KPN46, and Z4160. The enrichment ratio (Enrich Ratio) was
630 calculated as the ratio of genes in the target list belonging to the specified KEGG
631 pathway to the total number of genes in the pathway in the genome. A hypergeometric
632 test was used to determine false discovery rate (FDR), and $FDR < 0.05$ was considered
633 significant.

634 **Figure 6** Competitive colonization between parent strains and isogenic mutants to
635 validate genes identified in transposon insertion screens. Mice were treated with 5 days
636 of vancomycin prior to gavage with 1:1 mixtures of marked parent strain (hygromycin-
637 resistance cassette at the Tn7 site) and isogenic mutant (substitution of open reading
638 frame with apramycin-resistance cassette [A-N] or an unmarked in-frame deletion [O-
639 Q]) of the target gene(s). $n = 10$ for A-G and L-N. $n \geq 9$ for H-K and O-Q. Asterisks
640 denote significance by one-sample t-tests with Dunn correction where * $p < 0.05$, ** $p <$
641 0.01 , *** $p < 0.001$, **** $p < 0.0001$, and “ns” indicates not significant. Limit of detection
642 was a competitive index of 10^{-7} for A-N and 10^{-3} for O-Q, denoted with a dotted line.
643 $\text{Log}(\text{competitive index}) = 0$, or equal recovered CFU of parental strain and mutant, is
644 marked with a dashed line.

645 **Figure 7** Resistance of CRE-166 and *acrA* mutants to 10% ox bile. Strains were
646 inoculated into 10% ox bile (w/v), incubated, and CFU were plated for enumeration at
647 the indicated timepoints. $n = 3$ biological replicates. Line denotes median. * indicates p
648 < 0.05 in two-way ANOVA with Tukey’s HSD test.













