1 **Costs of antibiotic resistance genes depend on host strain and environment and can influence**
2 **community composition.**

3 Huei-Yi Lai[1] and Tim F. Cooper[1]

4 [1] School of Natural Sciences, Massey University, Auckland 0634, New Zealand

5 Running title: Resistance gene costs

6 Keywords: Antibiotic, antibiotic resistance genes, fitness, beta lactamase, plasmids

Abstract

The prevalence of antibiotic resistance genes (ARGs) is a major contributor to bacterial antibiotic resistance. ARGs are strongly selected in environments containing corresponding antibiotics, but it is less clear how ARGs are maintained in environments where antibiotic selection might be weak or sporadic. In particular, few studies have directly estimated the effect of ARGs on host fitness in the absence of direct selection. To the extent that ARGs impose costs, it is not clear if these are fixed or might depend on the host strain, perhaps marking some ARG-host combinations as low-cost reservoirs that can act to maintain ARGs in the absence of antibiotic selection. We quantified the fitness effects of six ARGs in each of 12 diverse *Escherichia* spp. strains in two environments. Only one ARG *(bla*$_{\text{TEM-116}}$, encoding resistance to some β-lactam antibiotics) imposed an overall cost. However, all but one ARG imposed a cost in at least one host strain, reflecting a significant ARG cost-by-strain interaction effect. A simulation model predicts that these interactions cause the ecological success of ARGs to depend on available host strains, and, to a lesser extent, for successful host strains to depend on the ARGs present in a community. Together, these results indicate the need to consider the potential of reservoir strains for maintenance of ARGs.

22 **Introduction**

23 Widespread use of antibiotics has led to selection for antibiotic-resistant bacteria, which has been
24 identified as a major global health concern (1). Resistance can derive from spontaneous mutations
25 that generally reduce affinity or access of antibiotics to cellular targets and from antibiotic resistance
26 genes (ARGs) that typically act to detoxify or export antibiotics from cells (2). The link between
27 antibiotic exposure and selection of resistance is clear: ARGs tend to rise in frequency shortly after
28 widespread use of the cognate antibiotics and ARGs of recent origin are associated with host bacteria
29 that can cause infection in humans or domestic animals (3–6). Even exposure to low levels of
30 antibiotics can confer a significant advantage to resistant cells, potentially allowing selection in a
31 wide range of environments(7).

32 The benefit conferred by ARGs to host cells during antibiotic exposure is, however, only one factor
33 determining their overall success (8–11). Of 22 ARGs for which a recent origin could be inferred,
34 most were from species with some human or domestic animal association (5). This link supports the
35 idea that direct selection for resistance is important for the success of ARGs, many of these species
36 also transit through environments where antibiotic concentrations are low (12). For other ARGs, no
37 originating host species can be inferred, consistent with them originating in poorly characterized
38 environmental species (5). These findings underlie the view that costs associated with encoding
39 ARGs will play an important role in determining the overall success of those genes (10, 13, 14).

40 Costs of antibiotic resistance have been examined in some detail for spontaneously occurring
41 resistance alleles (15). In one experiment, costs associated with rifampicin resistance were sufficient
42 to drive selection for frequent reversion to sensitivity in a mouse infection model (16). In other cases,
43 costs drive the selection of secondary mutations that compensate for the costs of resistance allowing
44 the resistance phenotype to persist with relaxed antibiotic-driven selection (reviewed in (13, 15).
45 Spontaneous resistance often occurs through mutations in essential genes encoding products targeted
46 by the relevant antibiotic (17). Changes to these products will generally alter their structure in some
47 way that reduces the strength of antibiotic interaction and are likely to have some deleterious effect
48 on the original function (14). By contrast, less work has been done examining effects to host cells of
49 expressing introduced ARGs.

50 Studies have shown fitness costs to hosts that carry ARG-encoding plasmids but estimates of the
51 effect of ARGs themselves are often indirect. A meta-analysis found that the costs associated with
52 plasmids generally increase with the number of encoded ARGs (14). Specific examples include

53 analysis of a *tetAR* operon, encoding tetracycline resistance, that was estimated to impose a cost of
54 around 1% in rich medium (18, 19). Similarly, β-lactamases, a large class of resistance genes that
55 confer resistance to one or more of the β-lactam antibiotics, have been shown to decrease host fitness,
56 though different variants have different effects in different hosts (20–23). In at least some cases,
57 however, ARGs, including *tetAR*, not only do not confer a cost but can provide a benefit to host cells
58 (24–26).

59 A factor that makes it especially important to understand the potential for costs of ARGs is their
60 association with horizontally mobile elements that allow them to be transmitted between different
61 hosts (reviewed in (12)). Studies of mutation fitness effects, including some conferring antibiotic
62 resistance, often find significant differences depending on the host background (27–32). For example,
63 although not isolating the effect of the ARG itself, a plasmid encoding the $bla_{\text{OXA-48}}$ resistance gene
64 conferred an overall cost across a range of host strains but with a range between ~20% cost and ~20%
65 benefit (33). Simulations predict that this high variation can increase plasmid, and, therefore, ARG,
66 persistence because hosts in which it has low or no cost can serve as reservoirs during periods when
67 the plasmid is not directly selected (33).

68 We cloned ARGs into a control plasmid and estimated their fitness effect in the absence of a cognate
69 antibiotic. We estimated effects of six ARGs conferring resistance to four different classes of
70 antibiotic in 12 genetically diverse *Escherichia* spp.. We found that costs were diverse across ARGs
71 and strains, and that they interact so that different strains have the potential to act as no-cost
72 reservoirs for different ARGs. Simulations predict that the magnitude of these effects is sufficient to
73 cause both the presence of different ARGs to influence which strains can persist in a community and
74 the availability of strains to determine which ARGs can be maintained. Our study reveals the
75 importance of variation in fitness effects of ARGs in bacterial populations on the prevalence of ARGs
76 in the absence of antibiotics.

77 **Materials and Methods**

78 *Selection of candidate ARGs.* The ARGs used in this study were chosen based on their prevalence in
79 accessible *E. coli* genomes and their availability in our laboratory (34). The 2022 release of the
80 CARD Prevalence, Genomes and Variants dataset gives the prevalence of our chosen ARGs among
81 sequenced *E. coli* genomes as: *aadA* (ARG type = 15.3%; specific ARG allele = 1.73%), *cat* (5.3%,
82 4.1%), *dfrA5* (19.4%, 1.8%), $bla_{TEM\text{-}116}$ (22.7%, 0.07%), $bla_{CTX\text{-}M\text{-}15}$ (17.1%, 5.83%) and bla_{SHV12}
83 (0.59%, 0.04%).

84 *Construction of ARG-plasmids and bacterial strains.* The focal ARGs were originally cloned into
85 pUA66, which contains a promoterless *gfp* (35). However, we found leaky expression of GFP in this
86 vector (data not shown). To avoid any influence of GFP expression on fitness measurements, we
87 removed the GFP by amplifying the pUA66 backbone (using pUA_minusFP_F:
88 5'atgtccagacctgcaggcatg; pUA_minusFP_R: 5'ggatccatcgaggtgaagacg) and circularizing the product
89 with a bridging oligonucleotide (5'cgtcttcacctcgatggatccatgtccagacctgcaggcatg) using the
90 NEBuilder® HiFi DNA Assembly (NEB) kit. The resulting plasmid, pmFP, retains the kanamycin
91 resistance gene originally present on pUA66. The six ARGs were amplified from template sources as
92 listed in Table S1 and separately added into the pmFP vector (amplified using primers: pUA66_F:
93 aataggcgtatcacgagg and pUA66_EcoRI_R: gaattcatggtttcttagacgtcgg) using NEBuilder® HiFi DNA
94 Assembly. All PCR reactions used Q5® High-Fidelity polymerase (NEB) unless otherwise specified.
95 When sequencing cloned ARGs we found that the $bla_{TEM-116}$ gene in pTarget, which was used as a
96 template, has a non-synonymous mutation (Q274R) relative to the reference $bla_{TEM-116}$ sequence
97 (NCBI accession: U36911). To distinguish between these genes we identify the variant used here as
98 $bla_{TEM-116}$*.

99 The six constructed pmFP::ARG plasmids, and the reference pmFP backbone vector, were each used
100 to transform one *E. albertii* and 11 genetically divergent *E. coli* strains (Fig. S1) followed the method
101 outlined in ref. 36. After transformation, cells were recovered at 37 °C for one hour before being
102 plated on lysogeny broth (LB) supplemented with 50 μg/ml kanamycin to select plasmid containing
103 cells. One strain, H305, was resistant to ampicillin, chloramphenicol, and trimethoprim, antibiotics.
104 Examining an annotated sequence of this strain we found that it contains bla_{TEM} and *cat* genes,
105 conferring resistance to B-lactams and chloramphenicol, respectively. Omitting H305 makes no
106 qualitative difference to the interpretation of any analysis we present, though we note that H305 was
107 one of four strains in which no ARG conferred a significant cost. We leave it in the presented
108 analyses because already resistant strains can clearly serve as hosts to ARG-encoding plasmids in
109 natural communities and because it is reasonable to think that additional copies of a given resistance
110 gene can impose some additional fitness effect.

111 *Fitness competitions.* To measure the fitness effect of each ARG, we performed a series of
112 competition assays. The fitness effect of an ARG was estimated as the fitness difference between
113 cells carrying an ARG plasmid (W_{ARG}) and cells carrying the control pmFP plasmid (W_{Vec}). The
114 fitness of ARG and vector plasmid-carrying cells was estimated separately by competition with
115 reference cells carrying a GFP expressing plasmid (pUA66-$_{PrpsL}$GFP).

116 Briefly, cells carrying an ARG plasmid, pmFP, or pUA66-$_{PrpsL}$GFP, were inoculated in 200 µL of
117 Davis-Mignioli broth supplemented with 250 µg/ml glucose (DM250) and 50 µg/ml kanamycin from
118 frozen stocks and cultured at 37ºC overnight. The overnight cultures were diluted 1:100 in 200 µL of
119 fresh DM250 and incubated for a 24-hour growth cycle. Cultures were grown over two additional
120 daily growth cycles with 1:100 dilution between each cycle to condition them to the antibiotic-free
121 medium, then used to set up competition assays. Five microliters of culture of each competitor –
122 reference cells with pUA66-$_{PrpsL}$GFP and either cells with an ARG plasmid or with the control
123 plasmid – were mixed in 40 µL of fresh DM250. Ten microliters of each cell mixture was added to
124 190 µL of fresh DM250 and incubated for 24 hours. The remaining 40 µL of each cell mix was added
125 to phosphate buffered saline (PBS) and placed on ice or fixed with formaldehyde before being
126 assayed by flow cytometry to determine the proportion of reference (GFP+) and ARG or control
127 plasmid-containing cells (Day 0). Overnight cultures of the cell mixtures were diluted 20-fold in PBS
128 and assayed by flow cytometry (Day 1). This protocol was repeated to obtain ARG effect estimates in
129 a competition environment supplemented with kanamycin, to which all plasmids confer resistance to.

130 The fitness effect of an ARG was estimated as W_{ARG} / W_{Vec}, where $W_{ARG} = \ln(ARG_{Day1} \times 100 /$
131 $ARG_{Day0}) / \ln(GFP_{Day1} \times 100 / GFP_{Day0})$ and $W_{Vec} = \ln(Vec_{Day1} / Vec_{Day0}) / \ln(GFP_{Day1} / GFP_{Day0})$.
132 ARG_{Day1}, ARG_{Day0}, Vec_{Day1} and Vec_{Day0} are the proportion of flow cytometry events without a GFP
133 signal, and GFP_{Day1} and GFP_{Day0} are the proportion of events with a GFP signal. Day 1 proportions
134 are multiplied by 100 to account for growth occurring during the competition. We note that a number
135 of flow cytometry runs resulted in total event counts or counts of one competitor that were
136 substantially different to mean estimates over all replicates of any given competition type. We
137 consider that the most likely reason for this is that the natural isolate strains used here are subject to
138 clumping that introduces variability in passage of cells through the cytometer flow cell. To reduce the
139 effect of this variation we excluded competitions where the starting frequency of either competitor
140 was estimated as more than 15% different than the target of 50%. This step removed a total of 911 of
141 4563 competitions. We emphasize that this filtering was performed without regard to the fitness
142 estimated considering the change in frequency over the course of the competition. We also omitted a
143 further 16 fitness estimates of one host strain collected in one experimental block. These estimates
144 were more than 50% higher than a set of 205 estimates collected for the same strain in other
145 experimental blocks, likely indicating that they were compromised by some experimental error.

146 In the antibiotic-free assay environment it is possible that plasmid loss could occur, which could
147 influence fitness estimates through production of relatively higher fitness plasmid free sub-

148 populations. To estimate the influence of plasmid-losing cells during the competition assay,
149 monocultures of cells with the pUA66-$_{PrpsL}$GFP plasmid were grown following the competition
150 protocol and plated on non-selective media to allow identification of the proportion of plasmid
151 containing cells. We found plasmid-containing cells remained at a frequency of 0.996 (± 0.004
152 95%CI, n = 4) after three days of transfer. If we conservatively take this fraction to represent the
153 fraction of cells incorrectly identified as being the ARG-plasmid containing competitor instead of
154 cells that lost the reference pUA66-$_{PrpsL}$GFP plasmid during a single day of competition, this would
155 represent a fitness inflation of the ARG of only <0.1% over the true value. Moreover, plasmid-loss
156 cannot easily explain differences in costs of different ARGs in the same host. Finally, we repeated
157 fitness estimates in an environment supplemented with kanamycin, which ensures that only plasmid-
158 containing cells are present. The overall magnitude of fitness effects in this environment were similar
159 to those in a kanamycin-free environment, suggesting that plasmid loss is not influencing fitness
160 estimates in the conditions used in our experiment.

161 *Determination of copy number of the* bla$_{TEM-1}$ *plasmid.* The copy number of plasmids encoding the
162 *bla*$_{TEM-116}$* ARG was measured by quantitative PCR (qPCR) using the comparative Ct (ΔΔCt)
163 method (37). Plasmid carrying cells were cultured overnight in DM250 supplemented with 50 µg/ml
164 kanamycin and DNA isolated using a Wizard Genomic DNA Purification Kit (Promega).
165 Amplification of target genes was carried out using a SYBR Green based qPCR mix consisting of
166 Q5® High-Fidelity 2× Master Mix (NEB), SYBR Green, and relevant primers (at a final
167 concentration of 0.25 µM). Between one and two ng of template DNA was used per 10 µl reaction.
168 Reactions were performed using a Thermo Scientific PikoReal Real-Time PCR System with default
169 settings. Primers were designed using PrimerQuest (Integrated DNA Technologies, Inc.) to amplify
170 the chromosomal *dxs* gene, in order to determine the absolute Ct value of the bacterial chromosome
171 (Ct_{ch}; xs_qpcr_F1:cgagaaactggcgatcctta; dxs_qpcr_R1: cttcatcaagcggtttcaca), and the pUA66 plasmid
172 encoded *aph(3')-IIa* gene, to determine the Ct value of the plasmid (Ct_p; pUA_Kan_qpcr_F1:
173 ctcgtcaagaaggcgatagaag; pUA_Kan_qpcr_R1:cgttggctacccgtgatatt). The relative copy number of the
174 *bla*$_{TEM-116}$* plasmid to the chromosome (ΔCt) was calculated as $2^{-(Ct_p - Ct_{ch})}$. The ΔCt of different
175 bacterial strains was then normalized to the lab strain, REL606, to obtain the ΔΔCt value.

176 *Statistics, phylogeny construction and phylogenetic signal testing.* The R statistical computing
177 platform was used for all analysis and visualization (38). The flowCore package (39) was used to
178 analyze flow cytometry data and the lme4 and lmerTest (40, 41) packages were used to run and
179 evaluate mixed models. Environment, ARG and host strain were included as random effects in these

180 models as noted in the text. Core and accessory genomes of a collection of 96 environmental *E. coli*
181 isolates were determined and used previously to build phylogenies (42). These phylogenies were used
182 to test for a phylogenetic signal of ARG effects using Pagel's lambda and Blomberg's K metrics as
183 implemented in the phylosig function of the phytools package 1.0-3 (43).

184 *Simulation of host-ARG communities.* Competitive fitness assays allow fitness of derived strains to be
185 measured and compared to a corresponding ancestral strain. It is not possible to extend this approach
186 to compare the fitness of our distinct progenitor strains because competition between these strains
187 may be affected by complex interactions, for example mediated by colicins or phage, that lead to non-
188 transitivity in their relative fitness. To estimate fitness of all starting and evolved strains on a
189 comparable scale, we analyzed growth curves of progenitor strains to estimate their fitness relative to
190 one another as determined only by differences in resource use (44). The fitness of each strain-ARG
191 combination was determined by adding the estimated baseline strain fitness and the relative fitness
192 effect of each ARG measured in that strain. These estimates were used to parameterize an individual
193 based model that simulated strain-ARG dynamics in a community with ARG effects drawn form a
194 normal distribution with mean and standard deviation taken from experimental estimates. Simulations
195 were initiated with equal numbers of each strain-ARG combination and a population of size 10^4
196 propagated using a Wright-Fisher model for 200 generations. Strains and ARGs remaining above a
197 threshold frequency of 0.01 in at least 20% of 10 replicate simulations were considered to have
198 persisted in the community. Simulations were repeated separately omitting each strain and each ARG
199 from the starting community.

200 **Results**

201 *ARGs can impose costs and benefits in the absence of direct selection.* We separately cloned six
202 ARGs belonging to four distinct mechanistic classes into a low-copy number plasmid vector (Table
203 1). The fitness effects of each ARG in the absence of cognate antibiotics was estimated in each of 12
204 divergent *Escherichia* spp. host strains (Fig. 1). We found significant differences in the mean effect
205 of the different ARGs ($\chi^2 = 40.4$, $P < 0.001$). Of the six tested ARGs, only two conferred a significant
206 overall change in fitness measured across all host strains: $bla_{TEM116*}$ conferred a 2.8% (±1.0% 95CI)
207 cost and $bla_{CTX\text{-}M\text{-}15}$ conferred a 1.5% (±1.1% 95CI) benefit (Dunnett's test: $bla_{TEM116*}$ $P < 0.01$;
208 $bla_{CTX15\text{-}M\text{-}15}$ $P = 0.04$). Two other ARGs, *cat* and *dfrA5,* conferred marginally non-significant costs
209 (Dunnett's test: *cat,* cost = 1.6%, $P = 0.06$; *dfrA5* cost = 1.6%, $P = 0.06$). Together these results
210 clearly indicate that ARGs can confer significant effects on fitness in the absence of direct selection
211 for their resistance phenotypes. Moreover, the three β-lactamases had significantly different effects

212 on host fitness, demonstrating that effects can differ even between ARGs with similar resistance
213 profiles ($\chi^2 = 37.4$, $P < 0.001$).

214 *Fitness effects of ARGs depend on the host strain.* The results presented above consider the overall
215 effect of each tested ARG over 12 host strains. To determine if effects differed between strains, we
216 extended our model to include an ARG effect-by-strain interaction term. This extended model
217 provided a significantly improved fit compared to the previous model, indicating a dependence of
218 ARG effects on host genotype ($\chi^2 = 102.6$, $P<0.001$). We found that all ARGs conferred a significant
219 fitness effect in at least one strain (Dunnett's test $P < 0.05$: *aadA* - 4 strains; $bla_{CTX\text{-}M\text{-}15}$ - 2 strains,
220 bla_{SHV12} - 1 strain, $bla_{TEM116*}$ - 3 strains; *cat* - 3 strains; *dfr*A5 - 2 strain2) (Fig. 2). In 11 of these 15
221 cases effects were negative. Of the 12 tested strains, all except H305, M605, and B354 had a
222 significant change in fitness caused by at least one ARG (Dunnett's test $P < 0.05$). Mean fitness
223 difference between different ARG-strain combinations was substantially larger than mean differences
224 between ARG effects, indicating a potential for some strains to act as reservoirs for ARGs in
225 antibiotic-free environments (ARG-Strain combinations mean difference = 3.1% (2.4 – 3.9% 95CI);
226 ARG mean difference = 1.6% (0.4 – 3.5% 95CI)). Indeed, four of the six tested ARGs were estimated
227 as having a neutral or beneficial effect in at least one host strain.

228 *Fitness effects of ARGs depend on the environment.* Fitness estimates reported above were carried out
229 in a minimal medium environment in the absence of any antibiotic. The focal ARGs were cloned into
230 a vector that also included an additional ARG gene (*aph3'-II*, encoding resistance to kanamycin)
231 giving us the opportunity to assess how the presence of an antibiotic might influence the effect of our
232 non-selected ARGs. We note that kanamycin is an aminoglycoside antibiotic with some similarity to
233 spectinomycin and streptomycin, substrates for the *aadA* ARG, but there is no detectable cross-
234 activity (45). When we measured the fitness effects of each ARG-strain combination in the same
235 environment as used previously, but supplemented with kanamycin, we again found a significant
236 interaction between ARG and host strain in determining fitness ($\chi^2 = 59.26$, $P<0.001$). The fitness of
237 all strains except for M605 was significantly affected by at least one ARG and all ARGs except *cat*
238 and bla_{SHV12} significantly affected fitness of at least one strain. There was a significant effect of
239 environment comparing fitness estimates made in the minimal medium and kanamycin supplemented
240 environments ($\chi^2 = 84.2$, $P<0.001$). Evidently, a general pattern of ARG fitness effects depending on
241 the strain background is common across environments, although the exact nature of that dependence
242 is different.

243 *Fitness effects of ARGs are not explained by plasmid copy number.* A possible explanation for
244 different effects of ARGs across host strains is that the plasmid they are encoded on has a different
245 copy number in different strains. Different copy number can affect expression levels of plasmid-
246 borne genes and, thus, any phenotypic effect of gene products. One well known example is the effect
247 of bla_{TEM1} copy number on the degree of ampicillin resistance (46). We used qPCR to determine the
248 copy number of the $bla_{TEM116*}$ plasmid across our set of 12 host strains. This ARG showed the highest
249 variation in fitness cost so represents a good test of the contribution of copy number differences to
250 differences in fitness effects. We found that the relationship between $bla_{TEM116*}$ copy number and
251 fitness cost was positive, but non-significant, in both the minimal medium and kanamycin-
252 supplemented environments, indicating that differences in plasmid copy number is unlikely to
253 represent a general mechanism of strain-specific ARG fitness effects (minimal: $r = 0.28$, $P = 0.36$;
254 kanamycin-supplemented: $r = 0.26$, $P = 0.41$).

255 *ARG effects are not predicted by evolutionary relationships between host strains.* Variation in ARG
256 effects across host strains depends on the genotype of those strains. To test if the dependence of ARG
257 effect on host strain can be predicted through knowledge of the genetic relationship of strains we
258 tested for a phylogenetic signal in the effect of each ARG. We estimated Pagel's λ, a measure of the
259 phylogenetic signal in a response variable, considering the fitness effect of each ARG across
260 phylogenies derived from the core and accessory genomes of our host strains (47). In no case was any
261 significant signal found. An alternative measure of phylogenetic signal, Blomberg's K (48), gave a
262 qualitatively consistent result (Table S2). With only 12 host strains, the power of individual tests is
263 not high. However, the lack of signal given that our host strains cover a wide range of *E. coli* species
264 genetic diversity and that tests were repeated for each of our six ARGs, indicate, at least, that that
265 fitness effects of ARGs are not well predicted by the overall genetic similarity of diverse host strains.

266 *Dependence of ARG fitness effect on host strain predicts interdependence between strain and ARG*
267 *success.* The above results demonstrate variation between ARGs in their mean effect on host fitness
268 and dependence on the host strain for the effect of any single ARG. These factors combine to cause
269 the success of host strains and ARGs to be codependent: the overall fitness of each ARG depends on
270 available host strains and the overall fitness of each host strain depends on which ARGs it contains. A
271 consequence of this dependence is a potential for some strains to act as low-cost reservoirs of ARGs,
272 allowing these ARGs to persist when they would not otherwise be able to. It is also possible that the
273 presence of some ARGs will affect the relative fitness of some strains allowing them to be maintained
274 in a community when they would otherwise be outcompeted.

275 As a first step to examine implications of host-ARG fitness codependence we simulated 200
276 generations of growth of a community started with equal numbers of all 72 host-ARGs combinations
277 considered in our study. At the end of the simulation strains and ARGs remaining above a threshold
278 frequency in at least 20% of replicate simulated communities were counted as having persisted (see
279 Materials and Methods for details). We found that the final community typically contained of two
280 strains, B354 and H504, and a single ARG, bla_{SHV1} (Fig. 3A and B, top right squares). This reflects
281 that these strain-ARG combinations had the highest fitness of any that we measured (Fig. 2 combined
282 with baseline strain fitness estimates, see Materials and Methods for details). To determine the effect
283 of strain dependent ARG costs on this result, we repeated simulations separately omitting each strain
284 and ARG. In some cases, quite different final communities were produced (Fig. 3). For example,
285 omission of the H504 strain from the starting community led to the B354 strain dominating. In this
286 strain as many as five of the six focal ARGs remained present at the end of the simulation period,
287 reflecting the relatively small and consistent effects of ARGs in the B354 host strain.

288 **Discussion**

289 We quantified the effect conferred by six different ARGs in each of 12 host strains. Selection for
290 genes depends on the net effect of positive selection for the function they encode and costs to the host
291 cell of expressing that function (49, 50). ARGs clearly benefit through protection of host cells in the
292 presence of antibiotics but many ARGs have been isolated from environments where antibiotics are
293 not present at high concentrations and where costs might be influential in determining their fate (10,
294 12–14). We found that costs of carrying ARGs were common (Figs 1 & 2). All tested ARGs except
295 bla_{SHV12} conferred a cost in at least one host strain. These costs differed over host strains, but all of
296 the 12 hosts had reduced fitness conferred by at least one ARG in at least one of the two assay
297 environments. Conversely, all ARGs were neutral or beneficial in at least one host strain, indicating
298 the potential for specific strains to provide a 'selective refuge' for ARGs in the absence of antibiotic
299 selection.

300 Within the set of six ARGs we considered, three—$bla_{\mathrm{TEM116*}}$, $bla_{\mathrm{CTX\text{-}M\text{-}15}}$ and bla_{SHV12}—were β-
301 lactamases. We expected the similar function of the encoded gene products to lead to similar host
302 dependent effects, but in fact these genes showed a significant host strain interaction effect (51).
303 Indeed, one, $bla_{\mathrm{TEM116*}}$, conferred the highest cost of the genes we considered while another, $bla_{\mathrm{CTX\text{-}M\text{-}}}$
304 $_{15}$, conferred the highest benefit (Fig. 1). While we do not focus on the basis of costs, we note that
305 previous work on the effects of β-lactamases have proposed several mechanisms that may be relevant
306 to our findings. β-lactamases need to be transported to the periplasmic region of cells to function and

307 some signal peptides that direct this secretion have been associated with a fitness cost (20). β-
308 lactamases also have similarity with the penicillin binding proteins (PBPs) that process peptidoglycan
309 components of cell walls. Residual activity of β-lactamases for PBP substrates might interfere with
310 the dynamic process of peptidoglycan processing required to accommodate cell growth (21, 22). The
311 host strains we used are genetically distinct, likely leading to different metabolic, physiological, and
312 regulatory backgrounds on which effects of ARGs are tested. Whatever the mechanism of ARG costs,
313 differences among our host strains evidently provide opportunity for them to play out differently.

314 Co-existence theory has been developed to understand the role of specific interactions in determining
315 broader community dynamics (52–54). This theory has recently been applied to consider ARG
316 success given the availability of a range of environmental niches in which host strains might be
317 differently affected (11). One key finding was that ARGs that reduce host fitness in fewer
318 environmental niches will generally be more successful than those that impose costs generally over
319 most environments. Because ARGs are frequently encoded by horizontally mobile elements that can
320 transfer between hosts, different host strains can be thought of as distinct niches with transfer
321 facilitating fluctuating selection. From the perspective of the ARG, strains in which they confer low
322 costs represent refuges allowing them to minimize niche overlap with competitors. The more of these
323 kind of strains are available to an ARG, the more successful an ARG will be. We have not yet
324 experimentally tested this prediction, but, even in the absence of any horizontal transfer, our
325 simulations indicate that measured ARG fitness effects are of a magnitude where strains acting as a
326 refuge for a particular ARG can influence the outcome of competition (Fig. 3). We also note that
327 ARG-strain interactions depended on the assay environment, a finding consistent with other studies
328 measuring effects of antibiotic resistance (16, 55–58) . This dependence increases the chance that
329 refuge strains will be available for any given ARG but may also select for transfer between strains if
330 costs change following environmental shifts.

331 To our knowledge the influence of ARG refuge strains has not been experimentally examined. There
332 are, however, some observations that are consistent with the idea that different costs of ARGs across
333 host strains are relevant for determining community dynamics. In one study trimethoprim use was
334 decreased by 85% for two years in one Swedish county (59). This change was correlated with a small
335 overall decrease in frequency of the corresponding *dfrA* resistance gene that masked larger changes
336 within different *E. coli* subpopulations. Trimethoprim resistance was both unevenly distributed across
337 subpopulations at the beginning of the intervention and changed differently over subpopulations
338 during the intervention, decreasing from being present in 67% to 41% of isolates from one

339 subpopulations, but increasing from 17% to 31% in another (59).The authors note that explanations
340 other than strain-specific differential fitness costs are possible, including clonal expansions of
341 sequence types within subpopulations may influence ARG frequencies, though the frequent
342 horizontal transfer of resistance genes means that they are unlikely to be the sole explanation.

343 One complication of predictions based on co-existence theory is that costs of ARGs can evolve.
344 Indeed, costs of resistance phenotypes, especially those due to spontaneous mutations, are often
345 reduced during growth in antibiotic-free environments (reviewed in (13, 15). Several studies have
346 also demonstrated changes in the fitness effect of ARGs due to chromosomal mutations (30, 60, 61).
347 A consequence of evolved changes in ARG costs is that ARG effects will not only differ across host
348 stains and environments, but that they will also differ through time as host strains adapt and
349 compensate for any costs they confer. How this effect might influence long-term ARG fitness will be
350 an important area for future study.

351 Investigation of the influence of ARGs on microbial communities has been dominated by
352 consideration of their effects in the presence of antibiotics. It was generally thought that ARGs would
353 impose costs that would select against them in environments without this selection pressure, but it
354 now seems that this assumption was optimistic (62). While many ARGs can impose costs, a
355 dependence on environment and host strain may create refuges that keep ARG frequencies high even
356 in the absence of direct selection.

357 **References**
358

359 1. Collaborators AR, Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, Han C, Bisignano C, Rao P,
360 Wool E, Johnson SC, Browne AJ, Chipeta MG, Fell F, Hackett S, Haines-Woodhouse G, Hamadani BHK, Kumaran
361 EAP, McManigal B, Agarwal R, Akech S, Albertson S, Amuasi J, Andrews J, Aravkin A, Ashley E, Bailey F, Baker S,
362 Basnyat B, Bekker A, Bender R, Bethou A, Bielicki J, Boonkasidecha S, Bukosia J, Carvalheiro C, Castañeda-Orjuela C,
363 Chansamouth V, Chaurasia S, Chiurchiù S, Chowdhury F, Cook AJ, Cooper B, Cressey TR, Criollo-Mora E,
364 Cunningham M, Darboe S, Day NPJ, Luca MD, Dokova K, Dramowski A, Dunachie SJ, Eckmanns T, Eibach D, Emami
365 A, Feasey N, Fisher-Pearson N, Forrest K, Garrett D, Gastmeier P, Giref AZ, Greer RC, Gupta V, Haller S, Haselbeck A,
366 Hay SI, Holm M, Hopkins S, Iregbu KC, Jacobs J, Jarovsky D, Javanmardi F, Khorana M, Kissoon N, Kobeissi E,
367 Kostyanev T, Krapp F, Krumkamp R, Kumar A, Kyu HH, Lim C, Limmathurotsakul D, Loftus MJ, Lunn M, Ma J, Mturi
368 N, Munera-Huertas T, Musicha P, Mussi-Pinhata MM, Nakamura T, Nanavati R, Nangia S, Newton P, Ngoun C,
369 Novotney A, Nwakanma D, Obiero CW, Olivas-Martinez A, Olliaro P, Ooko E, Ortiz-Brizuela E, Peleg AY, Perrone C,
370 Plakkal N, Ponce-de-Leon A, Raad M, Ramdin T, Riddell A, Roberts T, Robotham JV, Roca A, Rudd KE, Russell N,
371 Schnall J, Scott JAG, Shivamallappa M, Sifuentes-Osornio J, Steenkeste N, Stewardson AJ, Stoeva T, Tasak N,
372 Thaiprakong A, Thwaites G, Turner C, Turner P, Doorn HR van, Velaphi S, Vongpradith A, Vu H, Walsh T, Waner S,
373 Wangrangsimakul T, Wozniak T, Zheng P, Sartorius B, Lopez AD, Stergachis A, Moore C, Dolecek C, Naghavi M. 2022.
374 Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399:629–655.

375 2. Darby EM, Trampari E, Siasat P, Gaya MS, Alav I, Webber MA, Blair JMA. 2022. Molecular mechanisms of
376 antibiotic resistance revisited. Nat Rev Microbiol 21:1–16.

377 3. Rahman S, Kesselheim AS, Hollis A. 2023. Persistence of resistance: a panel data analysis of the effect of antibiotic
378 usage on the prevalence of resistance. J Antibiot 76:270–278.

379 4. Olesen SW, Barnett ML, MacFadden DR, Brownstein JS, Hernández-Díaz S, Lipsitch M, Grad YH. 2018. The
380 distribution of antibiotic use and its association with antibiotic resistance. eLife 7:e39435.

381 5. Ebmeyer S, Kristiansson E, Larsson DGJ. 2021. A framework for identifying the recent origins of mobile antibiotic
382 resistance genes. Commun Biology 4:8.

383 6. Poku E, Cooper K, Cantrell A, Harnan S, Sin MA, Zanuzdana A, Hoffmann A. 2023. Systematic review of time lag
384 between antibiotic use and rise of resistant pathogens among hospitalized adults in Europe. JAC-Antimicrob Resist
385 5:dlad001.

386 7. Gullberg E, Cao S, Berg OG, Ilbäck C, Sandegren L, Hughes D, Andersson DI. 2011. Selection of resistant bacteria at
387 very low antibiotic concentrations. Plos Pathog 7:e1002158.

388 8. Huijben S, Chan BHK, Nelson WA, Read AF. 2018. The impact of within-host ecology on the fitness of a drug
389 resistant parasite. Evol Medicine Public Heal 2018 1:127-137.

390 9. Sundqvist M. 2014. Reversibility of antibiotic resistance. Upsala J Med Sci 119:142–148.

391 10. Andersson D, Hughes D. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev
392 Microbiol 8:260 271.

393 11. Letten AD, Hall AR, Levine JM. 2021. Using ecological coexistence theory to understand antibiotic resistance and
394 microbial competition. Nat Ecol Evol 5:431–441.

395 12. Larsson DGJ, Flach C-F. 2021. Antibiotic resistance in the environment. Nat Rev Microbiol 20:1–13.

396 13. Hernando-Amado S, Sanz-García F, Blanco P, Martínez JL. 2017. Fitness costs associated with the acquisition of
397 antibiotic resistance. Essays Biochem 61:37–48.

398 14. Vogwill T, MacLean RC. 2014. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis
399 approach. Evol Appl 8:284 295.

400 15. Hall AR, Angst DC, Schiessl KT, Ackermann M. 2015. Costs of antibiotic resistance – separating trait effects and
401 selective effects. Evol Appl 8:261–272.

402 16. Björkman J, Nagaev I, Berg OG, Hughes D, Andersson DI. 2000. Effects of environment on compensatory mutations
403 to ameliorate costs of antibiotic resistance. Science 287:1479–1482.

404 17. Alekshun MN, Levy SB. 2007. Molecular mechanisms of antibacterial multidrug resistance. Cell 128:1037–1050.

405 18. Gelder LD, Ponciano JM, Abdo Z, Joyce P, Forney LJ, Top EM. 2004. Combining mathematical models and
406 statistical methods to understand and predict the dynamics of antibiotic-sensitive mutants in a population of resistant
407 bacteria during experimental evolution. Genetics 168:1131–1144.

408 19. Nguyen TN, Phan QG, Duong LP, Bertrand KP, Lenski RE. 1989. Effects of carriage and expression of the Tn*10*
409 tetracycline-resistance operon on the fitness of *Escherichia coli* K12. Mol Biol Evol 6:213–25.

410 20. Rajer F, Sandegren L. 2022. The role of antibiotic resistance genes in the fitness cost of multiresistance plasmids.
411 Mbio 13:e03552-21.

412 21. Fernández A, Pérez A, Ayala JA, Mallo S, Rumbo-Feal S, Tomás M, Poza M, Bou G. 2012. Expression of OXA-type
413 and SFO-1 β-lactamases induces changes in peptidoglycan composition and affects bacterial fitness. Antimicrob Agents
414 Chem 56:1877–1884.

415 22. Colquhoun JM, Farokhyfar M, Hutcheson AR, Anderson A, Bethel CR, Bonomo RA, Clarke AJ, Rather PN. 2021.
416 OXA-23 β-Lactamase overexpression in *Acinetobacter baumannii* drives physiological changes resulting in new genetic
417 vulnerabilities. Mbio 12:e03137-21.

418 23. Marciano DC, Karkouti OY, Palzkill T. 2007. Afitness cost associated with the antibiotic resistance enzyme SME-1 β-
419 lactamase. Genetics 176:2381–2392.

420 24. Enne VI, Delsol AA, Davis GR, Hayward SL, Roe JM, Bennett PM. 2005. Assessment of the fitness impacts on
421 *Escherichia coli* of acquisition of antibiotic resistance genes encoded by different types of genetic element. J Antimicrob
422 Chemoth 56:544–551.

423 25. Kassen R, Bataillon T. 2006. Distribution of fitness effects among beneficial mutations before selection in
424 experimental populations of bacteria. Nat Genet 38:484 488.

425 26. Lenski RE, Simpson SC, Nguyen TT. 1994. Genetic analysis of a plasmid-encoded, host genotype-specific
426 enhancement of bacterial fitness. J Bacteriol 176:3140–3147.

427 27. Knopp M, Andersson DI. 2018. Predictable phenotypes of antibiotic resistance mutations. Mbio 9:552 14.

428 28. Peng F, Widmann S, Wünsche A, Duan K, Donovan KA, Dobson RCJ, Lenski RE, Cooper TF. 2018. Effects of
429 beneficial mutations in *pykF* gene vary over time and across replicate populations in a long-term experiment with
430 bacteria. Mol Biol Evol 35:202 210.

431 29. Trindade S, Sousa A, Xavier KB, Dionisio F, Ferreira MG, Gordo I. 2009. Positive epistasis drives the acquisition of
432 multidrug resistance. Plos Genet 5:e1000578.

433 30. Silva RF, Mendonça SCM, Carvalho LM, Reis AM, Gordo I, Trindade S, Dionisio F. 2011. Pervasive sign epistasis
434 between conjugative plasmids and drug-resistance chromosomal mutations 7:e1002181.

435 31. Sousa JM de, Balbontín R, Durão P, Gordo I. 2017. Multidrug-resistant bacteria compensate for the epistasis between
436 resistances. Plos Biol 15:e2001741 24.

437 32. Rozen DE, McGee L, Levin BR, Klugman KP. 2007. Fitness costs of fluoroquinolone resistance in *Streptococcus*
438 *pneumoniae*. Antimicrob Agents Chem 51:412–416.

33. Valle AA, León-Sampedro R, Rodríguez-Beltrán J, DelaFuente J, Hernández-García M, Ruiz-Garbajosa P, Cantón R, Peña-Miller R, Millán AS. 2021. Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. Nat Commun 12:2653.

34. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, McArthur AG. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

35. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U. 2006. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat Methods 3:623 628.

36. Lessard JC. 2013. Transformation of *E. coli* via electroporation. Methods Enzym 529:321–327.

37. Lee C, Kim J, Shin SG, Hwang S. 2006. Absolute and relative qPCR quantification of plasmid copy number in *Escherichia coli*. J Biotechnol 123:273–280.

38. Team RC. 2021. R: A Language and Environment for Statistical Computing.

39. Ellis B, Haaland P, Hahne F, Le Meur N, Gopalakrishnan N, Spidlen J, Jiang M, Finak G. 2022. flowCore: Basic structures for flow cytometry data. R package version 2.8.0.

40. Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw 67. 1-48.

41. Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in linear mixed effects models. J Stat Softw 82. 1-26.

42. Wang Y, Arenas CD, Stoebel DM, Flynn K, Knapp E, Dillon MM, Wünsche A, Hatcher PJ, Moore FB-G, Cooper VS, Cooper TF. 2016. Benefit of transferred mutations is better predicted by the fitness of recipients than by their ecological or genetic relatedness. Proc National Acad Sci 113:5047 5052.

43. Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

44. Ram Y, Dellus-Gur E, Bibi M, Karkare K, Obolski U, Feldman MW, Cooper TF, Berman J, Hadany L. 2019. Predicting microbial growth in a mixed culture from growth curve data. Proc National Acad Sci 116:14698–14707.

45. Dhar R, Bergmiller T, Wagner A. 2014. Increased gene dosage plays a predominant role in the initial stages of duplicate TEM-1 beta lactamase genes. Evolution 68:1775–1791.

46. Papadovasilaki M, Oberthür D, Gessmann R, Sarrou I, Betzel C, Scoulica E, Petratos K. 2015. Biophysical and enzymatic properties of aminoglycoside adenylyltransferase AadA6 from *Pseudomonas aeruginosa*. Biochem Biophysics Reports 4:152–157.

47. Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877 884.

48. Blomberg SP, Garland T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717 745.

49. Dekel E, Alon U. 2005. Optimality and evolutionary tuning of the expression level of a protein. Nature 436:588 592.

50. Eames M, Kortemme T. 2012. Cost-benefit tradeoffs in engineered *lac* operons. Science 336:911 915.

51. Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, Takebayashi Y, Spencer J. 2019. β-Lactamases and β-Lactamase Inhibitors in the 21st Century. J Mol Biol 431:3472–3500.

477 52. HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM. 2012. Rethinking community assembly
478 through the lens of coexistence theory. Ecol Evol Syst 43:227–248.

479 53. Barabás G, D'Andrea R, Stump SM. 2018. Chesson's coexistence theory. Ecol Monogr 88:277–303.

480 54. Ellner SP, Snyder RE, Adler PB, Hooker G. 2019. An expanded modern coexistence theory for empirical applications.
481 Ecol Lett 22:3–18.

482 55. Petersen A, Aarestrup FM, Olsen JE. 2009. The in vitro fitness cost of antimicrobial resistance in *Escherichia coli*
483 varies with the growth conditions. FEMS Microbiol Lett 299:53–59.

484 56. Paulander W, Maisnier-Patin S, Andersson DI. 2009. The fitness cost of streptomycin resistance depends on *rpsL*
485 mutation, carbon source and RpoS (σS). Genetics 183:539–546.

486 57. Qi Q, Preston GM, MacLean RC. 2014. Linking system-wide impacts of RNA polymerase mutations to the fitness
487 cost of rifampin resistance in *Pseudomonas aeruginosa*. mBio 5:e01562-14.

488 58. Trindade S, Sousa A, Gordo I. 2012. Antibiotic resistance and stress in the light of Fisher's model. Evolution 66:3815
489 3824.

490 59. Sundqvist M, Granholm S, Naseer U, Rydén P, Brolund A, Sundsfjord A, Kahlmeter G, Johansson A. 2014. Within-
491 population distribution of trimethoprim resistance in *Escherichia coli* before and after a community-wide intervention on
492 trimethoprim use. Antimicrob Agents Ch 58:7492–7500.

493 60. Turner PE, Cooper VS, Lenski RE. 1998. Tradeoff between horizontal and vertical modes of transmission in bacterial
494 plasmids. Evolution 52:315–329.

495 61. Hellweger FL. 2013. *Escherichia coli* adapts to tetracycline resistance plasmid (pBR322) by mutating endogenous
496 potassium transport: in silico hypothesis testing. FEMS Microbiol Ecol 83:622–631.

497 62. Andersson DI, Nicoloff H, Hjort K. 2019. Mechanisms and clinical relevance of bacterial heteroresistance. Nat Rev
498 Microbiol 17:479–496.

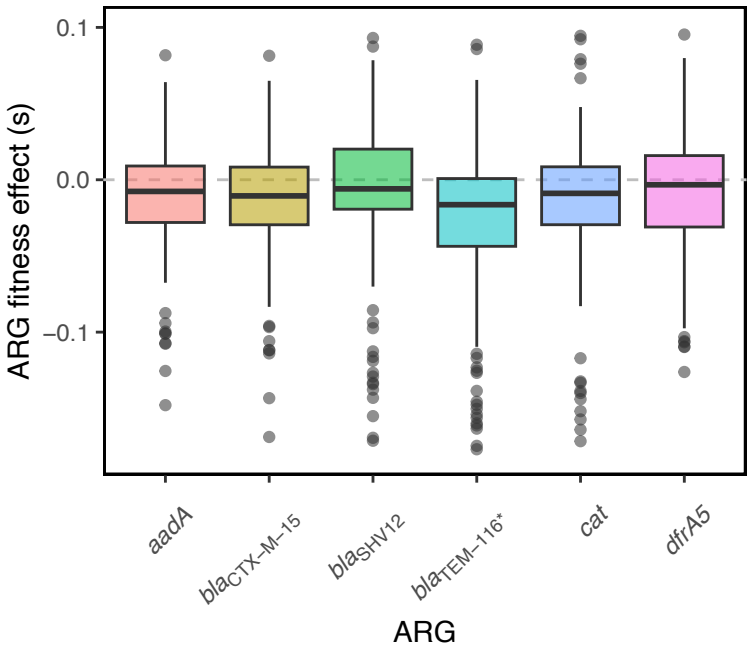
499
500
501

502 **Table 1. ARGs used in this study.**

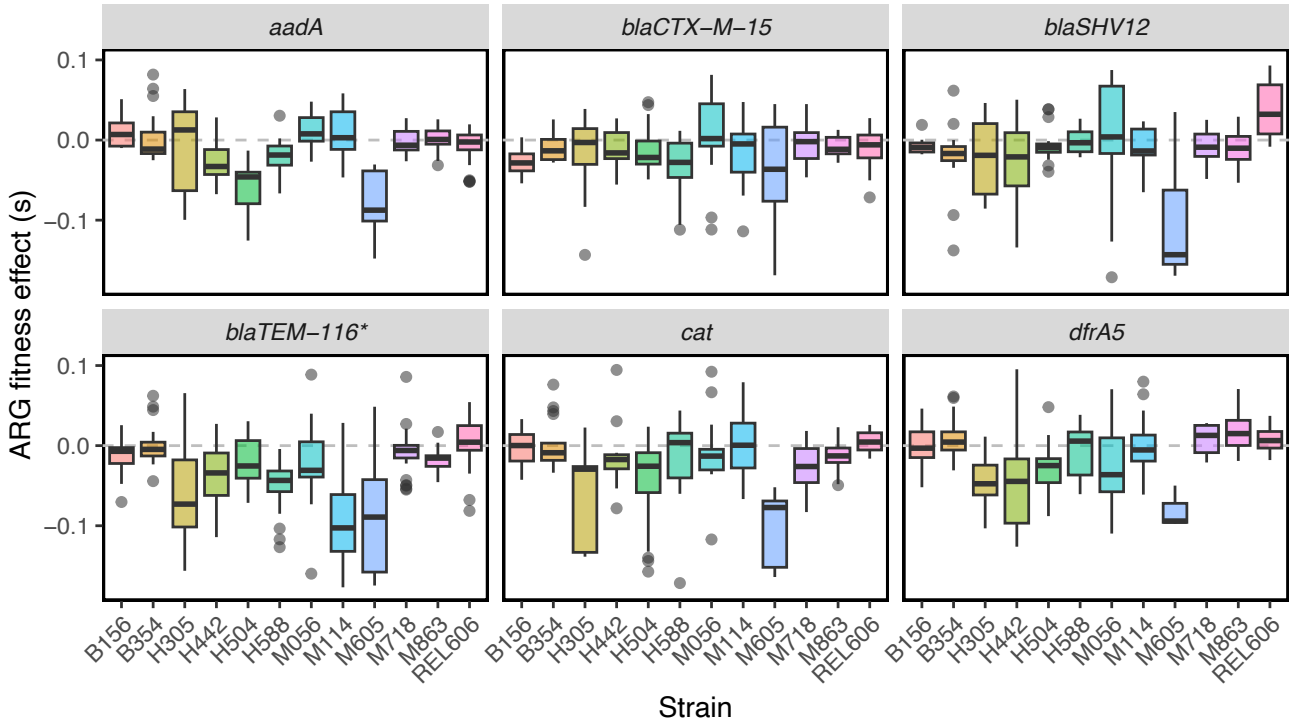
Gene	Enzyme type	Conferring resistance to:	Mechanism
aadA	Aminoglycoside adenyltransferase	Aminoglycosides such as spectinomycin and streptomycin	Drug-modification
$bla_{TEM-116}$*	Beta lactamase	β-lactam such as ampicillin	Drug-modification
bla_{SHV12}	Beta lactamase	β-lactam such as ampicillin and cefotaxime	Drug-modification
$bla_{CTX-M-15}$	Beta lactamase	β-lactam such as ampicillin and cefotaxime	Drug-modification
cat	Chloramphenicol acetyl transferase	Chloramphenicol	Drug-modification
dfrA5	Dihydrofolate reductase	Trimethoprim	Target replacement

503

504



505 **Figure 1. Fitness effect of each ARG across 12 *Escherichia* spp. host strains.** Boxes indicate median (line), 25th and
506 75th percentiles, and whiskers indicate 1.5× the inter-quartile range. Outliers are indicated as individual points. The
507 horizontal dashed line indicates no fitness effect. The fitness effect of each ARG effect was estimated in each of the 12
508 host strains. Estimates over these strains are pooled in this analysis ($n \geq 5$ for each host-ARG combination).
509



510
511 **Figure 2. Fitness effect of ARGs in each host strain.** Boxes indicate median (line), 25th and 75th percentiles, and
512 whiskers indicate 1.5× the inter-quartile range. Outliers are indicated as individual points. The horizontal dashed line
513 indicates no fitness effect. n ≥ 5 for each host-ARG combination.

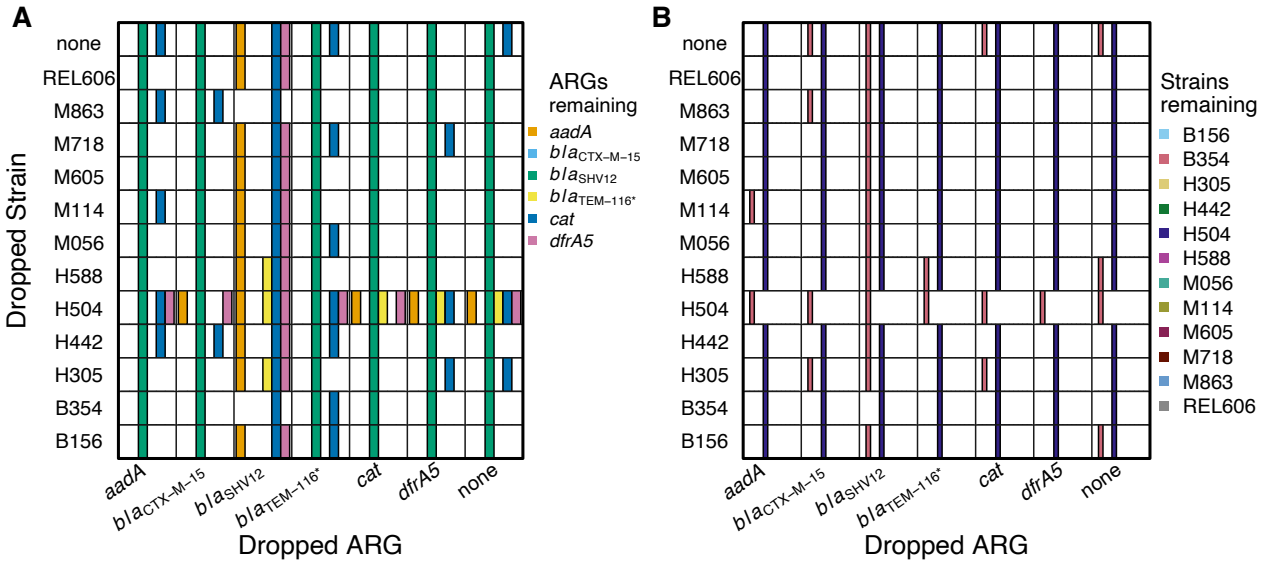


Figure 3 Host strain and ARG persistence during community growth in the absence of direct ARG selection. Communities consisting of all 72 host strain-ARG combinations and with omission of one strain and/or ARG as indicated were simulated over 200 generations with fitness effects estimated as described in Materials and Methods. Coloured bars indicate the ARGs (panel A) and host strains (panel B) remaining at greater than 1% in at least two of ten replicate communities simulated in each condition.