



HAL
open science

Microbiome diversity protects against pathogens by nutrient blocking

Frances Spragge, Erik Bakkeren, Martin Jahn, Elizete B. N. Araujo, Claire Pearson, Xuedan Wang, Louise Pankhurst, Olivier Cunrath, Kevin Foster

► **To cite this version:**

Frances Spragge, Erik Bakkeren, Martin Jahn, Elizete B. N. Araujo, Claire Pearson, et al.. Microbiome diversity protects against pathogens by nutrient blocking. *Science*, 2023, 382 (6676), pp.adj3502. 10.1126/science.adj3502 . hal-04737047

HAL Id: hal-04737047

<https://cnrs.hal.science/hal-04737047v1>

Submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Microbiome diversity protects against pathogens by nutrient blocking**

2 Frances Spragge^{1,2+} and Erik Bakkeren^{1,2+}, Martin T. Jahn^{1,2}, Elizete B. N. Araujo³, Claire F. Pearson³,

3 Xuedan Wang^{1,2}, Louise Pankhurst^{1,2}, Olivier Cunrath^{4*} and Kevin R. Foster^{1,2*}

4 ⁺These authors contributed equally to this work

5 *Corresponding authors: kevin.foster@biology.ox.ac.uk; olivier.cunrath@unistra.fr

6 1. Department of Biology, University of Oxford, Oxford, UK

7 2. Department of Biochemistry, University of Oxford, UK

8 3. Kennedy Institute of Rheumatology, University of Oxford, UK

9 4. CNRS, UMR7242, Biotechnology and cell signaling, University of Strasbourg, Illkirch, France

10 **Abstract**

11 The human gut microbiome plays an important role in resisting colonisation of the host by pathogens,
12 but we lack the ability to predict which communities will be protective. We studied how human gut
13 bacteria influence colonisation of two major bacterial pathogens, both *in vitro* and in gnotobiotic
14 mice. While single species alone had negligible effects, colonisation resistance greatly increased with
15 community diversity. Moreover, this community-level resistance rested critically upon certain species
16 being present. We explain these ecological patterns via the collective ability of resistant communities
17 to consume nutrients that overlap with those used by the pathogen. Further, we apply our findings to
18 successfully predict communities that resist a novel target strain. Our work provides a reason why
19 microbiome diversity is beneficial and suggests a route for the rational design of pathogen-resistant
20 communities.

21

22 **One sentence summary**

23 Diverse communities of gut bacteria collectively limit pathogen colonisation by blocking nutrient
24 access.

25 **Introduction**

26 The human gut is home to diverse bacterial species collectively known as the gut microbiota. A major
27 health benefit provided by the gut microbiota is protection against pathogen colonisation and
28 subsequent infection; a phenomenon known as colonisation resistance (1). The ability of the
29 microbiota to protect against numerous enteric pathogens is well-documented, with evidence that
30 particular species within the microbiota play a more important role than others (2-9). The ways that
31 colonisation resistance can arise include competition for nutrients and space, direct antagonism by
32 toxins and other harmful compounds, and promoting host immunity against pathogens (1, 10, 11).

33 While the importance of the microbiota for colonisation resistance is clear, however, we currently
34 lack the principles needed to predict, *a priori*, which microbiota species will be effective against a
35 given pathogen. A key challenge is the ecological complexity of the gut. The gut microbiome is a
36 diverse ecological system with many individual species that all have the potential to play a role in
37 colonisation resistance. Moreover, these constituent species can also affect each other and interact
38 ecologically in ways that are critical for colonisation resistance (12-16). This combination of species
39 diversity and the potential for ecological interactions makes colonisation resistance a challenging
40 phenotype to understand (17).

41 Here we approached the question of mechanisms of colonization resistance from the
42 perspective of the underlying ecological principles. To do this, we studied colonisation resistance
43 provided by a range of human gut bacteria, both alone and in combinations. We performed all
44 experiments in parallel using two species of pathogen, which are both on the WHO priority list:
45 *Klebsiella pneumoniae* and *Salmonella enterica* Serovar Typhimurium (18). Both are members of the
46 Enterobacteriaceae found in the human gut microbiome but they have very different lifestyles. *S.*
47 Typhimurium causes acute infection and gastroenteritis (19, 20). By contrast, *K. pneumoniae* is a
48 nosocomial, opportunistic pathogen that rarely causes disease in the gut itself, but gut colonisation is a
49 major risk factor for antimicrobial resistance associated infections elsewhere in the body (21).

50 Despite these differences, we have identified common principles that underlie colonisation resistance
51 to both species. Ecological diversity is important for colonisation resistance *in vitro* and in gnotobiotic
52 mice. Moreover, we found that colonisation resistance is an ecologically complex trait, whereby the
53 ability of one species to provide colonisation resistance can depend entirely upon the presence of
54 other species (22). Despite the complexity, we find that these ecological patterns are explained by a
55 simple underlying principle, the collective ability of certain communities to consume nutrients and
56 block pathogen growth. Further, we have shown that this principle offers a way to identify sets of
57 bacterial species that will collectively limit the growth of a particular pathogen.

58 **Results**

59 **Single species offer little protection in competition with pathogens**

60 Individual members of the microbiota can promote colonisation resistance in various contexts (2-8),
61 which suggests that some species are more important for colonisation resistance than others. To
62 systematically assess this variability, we screened a diverse set of 100 human gut symbionts (**Table**
63 **S1; also see Methods**) for their ability to limit pathogen growth. Competition in the gut occurs both at
64 the point a pathogen enters the gut and when a pathogen becomes established (23, 24). We designed
65 two co-culture assays to reflect these two aspects of competition in the mammalian gut (**Fig. 1a**). In
66 the first assay (ecological invasion assay), we pre-grew the symbiont alone in standard anaerobic
67 media (modified Gifu anaerobic media; mGAM) buffered to human colonic pH before adding the
68 pathogen. In the second assay (competition assay), we inoculated this media with an equal ratio of
69 symbiont to pathogen, which is designed to capture competition once a pathogen has established itself
70 in the gut.

71 To assess pathogen growth, we built luminescent strains of *K. pneumoniae* and *S.*
72 *Typhimurium* and compared luminescence when grown in monoculture and when grown in co-culture
73 with each symbiont. With this assay system, we could rank the strains based on their abilities to limit
74 pathogen growth in both the invasion and competition assays (**Fig. 1b-c, S1**). From this ranking, we
75 took the top ten best-performing non-pathogenic symbiont species in the screen (Methods; shown in

76 orange in **Fig. S1e-f; Table S1**) and subjected them to a more stringent test of colonisation resistance
77 designed to capture both phases of competition in the gut in one assay (extended competition assay,
78 **Fig. 1d**). Here, the pathogen is first introduced into a pre-grown culture of a given symbiont strain and
79 then, after 24 hours, the mixture is passaged into fresh media and allowed to grow for 24 hours,
80 whereafter pathogen abundance is assessed via flow cytometry (**Fig. 1d**). Despite choosing the best-
81 ranked species from the luminescence screen, all symbionts performed poorly under extended
82 competition, with the majority offering no discernible colonisation resistance (**Fig. 1e-f**). The best
83 performer was *Escherichia coli*, a known competitor of *S. Typhimurium*, and also a member of the
84 Enterobacteriaceae, but even here the protection offered was very limited with the pathogens still able
85 to reach $10^8 - 10^9$ cells/ml.

86 The outcome of the assay differed greatly when we pooled all ten strains together (**Fig. 1e-f**).
87 Now, the final abundance of both pathogens was strongly suppressed by over three orders of
88 magnitude for *K. pneumoniae* and about two orders of magnitude for *S. Typhimurium*. By contrast, a
89 community made up of the ten worst-performing strains from the luminescence screen (shown in blue
90 in **Fig. S1e-f**) provided little or no colonisation resistance (**Fig. 1e-f**). These results, therefore, suggest
91 that strain identity is important for colonisation resistance only in the context of a diverse community.

92

93 **Ecological diversity and complexity drive colonisation resistance *in vitro***

94 Our results indicated that microbiota diversity is important for colonisation resistance. This finding
95 fits well with the general idea that microbial diversity is beneficial for microbiome functioning,
96 whereas a loss of diversity, or dysbiosis, can be associated with poor health and disease (25-27).
97 While the potential benefits of diversity are clear, cause and effect can be confounded in observational
98 studies (28). To systematically test the role of diversity in colonisation resistance, we randomly
99 selected communities of increasing diversity from the best ranked 10 strains and competed them
100 against the pathogens in the extended competition assay. To further evaluate the importance of
101 diversity, we also assembled a community of 50 non-pathogenic symbiont species from the strains in

102 our initial luminescence screen (see Methods). These data indicated a relationship between diversity
103 and colonisation resistance. However, we also saw a large variation in colonisation resistance across
104 the communities that differed in their composition of two, three and five species. Visual inspection of
105 the data (**Fig. 2c-d**) suggested that a large component of this variability was driven by the composition
106 of the communities.

107 One species that appeared to be important for outcomes was *E. coli*. To explore this finding,
108 we randomly selected additional *E. coli*-containing communities and again evaluated colonisation
109 resistance (**Fig. 2c-d**). We also performed drop-out experiments, where we made up the 10 and 50
110 species communities without *E. coli* (**Fig. 2c-d**). These data revealed a strong and clear monotonic
111 increase in colonisation resistance as species diversity increases (**Fig. 2c-d, green circles**), but this
112 relationship is much weaker or disappears entirely in the absence of *E. coli* (**Fig. S2a-b**). In ecological
113 terms, these data show that colonisation resistance rests upon a strong higher-order effect involving
114 other community members and *E. coli* (22). By higher-order effects here, we mean cases where the
115 effects of species on one another are changed by the presence of a third-party species in a community
116 (22). That is, while *E. coli* alone or the rest of the community alone each have little impact on
117 pathogen growth, together they have a strong effect on pathogen growth. Such higher-order effects are
118 considered important in ecology as they imply context dependence, which can make a system difficult
119 to understand and predict (22, 29, 30).

120 Another way to illustrate the effect of diversity on colonisation resistance is to compare our
121 data to a simple null model. Consider, for example, a model where each additional species
122 proportionally reduces the amount of nutrients available to the pathogen. Specifically, we can
123 compare our experimental data for *E. coli* containing communities to a null model where pathogen
124 abundance scales according to $1/n$, where n is the number of species in the community. This analysis
125 shows that the deviation from the null model increases as diversity increases, where colonisation
126 resistance is again greater than expected for diverse communities (**Fig. S3**). We also asked whether
127 the role of *E. coli* within communities was a strain-specific effect. We replaced *E. coli* strain IA11,
128 identified in our screen, by each of four other *E. coli* strains historically isolated from the human gut

129 (31-33). The effect of *E. coli* was recreated when *E. coli* IAI1 was substituted by most *E. coli* strains
130 (Fig. 2e-f), which indicates that the higher-order effect involving *E. coli* is a general property of
131 closely-related strains.

132 Further inspection of the data pointed to other species that were important for colonisation
133 resistance in diverse communities. In *E. coli* containing communities, the presence of *Bifidobacterium*
134 *breve* appeared to be important in excluding *K. pneumoniae* (Fig. S4a), and *Lacrimispora*
135 *saccharolyticum* and *Phocaecicola vulgatus* in the exclusion of *S. Typhimurium* (Fig. S4b). We
136 confirmed these patterns through a series of systematic drop-out experiments (Fig. S4c-d). However,
137 it was still possible to achieve equivalent colonisation resistance in more diverse communities that
138 lack these species (Fig. S4c-d), which again points to the underlying benefits of a diverse microbiota.

139

140 **Ecological diversity and complexity also drive colonisation resistance *in vivo***

141 To validate our *in vitro* methods, we tested the ability of symbiont communities to resist pathogen
142 colonisation in gnotobiotic mice (Fig. 3a). Germ-free mice were colonised with symbiont
143 communities differing in diversity, and the presence or absence of *E. coli*. Successful colonisation by
144 *S. Typhimurium* causes an acute infection and massive gut inflammation, which is a major
145 confounding effect for studying the effects of community composition on pathogen growth. Animals
146 with a less protective microbiota can rapidly succumb to the infection, such that one cannot follow
147 ecological dynamics in a comparable way across treatments. We, therefore, chose to use an avirulent
148 variant of *S. Typhimurium* to eliminate the effect of gut inflammation on pathogen and host, where
149 one can use pathogen abundance as a measure of disease risk (19, 34, 35). We introduced
150 communities across the same range of diversities as before but, in contrast to *in vitro* assays, not all
151 symbiont species will reliably colonise germ-free mice (36). Therefore, we used metagenomic
152 sequencing to confirm that introducing a higher diversity community to the mice did indeed result in a
153 higher diversity of strains colonising the gut as measured by two metrics of alpha diversity, and to
154 identify the relative abundance of all members (Fig. 3b-c, S5).

155 These experiments revealed that, as observed *in vitro*, microbiome diversity is negatively
156 correlated with pathogen abundance in faeces for both pathogens (compare 10 vs 50-member
157 communities; **Fig. 3d-e**; **Fig. S6**). Moreover, drop-out experiments revealed again the importance of
158 the combination of *E. coli* and other community members for colonisation resistance (**Fig. 3d-e**). We
159 also observe that higher diversities are needed for efficient colonisation resistance in the mammalian
160 gut than in our *in vitro* assays, which is likely to be explained by the higher degree of environmental
161 and spatial heterogeneity in the gut compared to a test tube. Nevertheless, the key patterns remain the
162 same between the gnotobiotic mouse experiments and our *in vitro* assays. Both ecological diversity
163 and higher-order interactions are important for colonisation resistance to both pathogens. Moreover,
164 again as before, we saw a strong deviation from a simple null model of nutrient competition at high
165 levels of diversity (**Fig. S3**). In addition to showing the generality of these patterns, this fit between
166 the *in vitro* and *in vivo* methods validates our extended competition assay as an approach to
167 interrogate the ecology of colonisation resistance.

168

169 **A simple principle explains the roles of diversity and complexity in colonisation resistance**

170 The discovery of such higher-order effects in colonisation resistance indicates that colonisation
171 resistance is an ecologically complex trait (22), which can be challenging to work with owing to high
172 levels of context dependence (17, 22, 29, 30). Nevertheless, we sought to understand the mechanisms
173 underpinning colonisation resistance by returning to our *in vitro* data gathered from large numbers of
174 different communities. The genomic data enabled us to assess functional similarity between symbiont
175 communities and pathogens from overlap in protein compositions. Specifically, we calculated the
176 percentage of protein families carried by a pathogen that were also present in each community
177 investigated (see Methods). We reasoned that this measure of functional similarity may map to niche
178 overlap and, therefore, the strength of ecological competition between symbionts and pathogens. We
179 first confirmed that the number of encoded protein families covered by our experimental communities
180 increases proportionally with the number of added strains (**Fig. S7**). Permutation analyses also

181 confirmed that the randomly selected communities we have studied experimentally are a good
182 representation of all possible communities that we could have studied (**Fig. S8**).

183 The potential importance of protein family overlap was already clear from the effects of *E.*
184 *coli* in our experimental data (**Figs. 2-3**). *E. coli* is in the same family of bacteria as *K. pneumoniae*
185 and *S. Typhimurium* and can be seen to contribute greatly to the overlap in the protein families carried
186 by a given community and either of the pathogens in our experimental communities (**Fig. S9a-d**).
187 However, by taking only the communities that contain *E. coli* to control for this effect, we also see a
188 strong correlation between a community's protein family overlap with the pathogen and its
189 colonisation resistance in our *in vitro* assays (**Fig. 4a-b; Fig. S9**). In other words, if the symbiont
190 strain or community encoded many of the same (or similar) proteins as the pathogen, it provides better
191 colonisation resistance. The same analysis for communities that lack *E. coli* is not informative as
192 colonisation resistance is consistently so low across all communities (**Fig. 2**).

193 Altogether, our genomic analysis suggests that communities that overlap highly with the
194 pathogens in encoded functions provide the best colonisation resistance. These analyses support our
195 hypothesis that niche overlap is important for our observed ecological patterns in colonisation
196 resistance. One of the key drivers of niche overlap is resource competition (37, 38), which is a known
197 contributor to colonisation resistance to *K. pneumoniae* and *S. Typhimurium* (12, 13). We, therefore,
198 explored the role of nutrient competition by generating metabolic profiles for the two sets of 10 key
199 symbiont species identified in the original screen against each pathogen (**Fig. 1**) using AN Biolog
200 Microplates that profile the metabolic activity of each strain on 95 carbon sources (**Fig. S10**). Note
201 that to cover the two sets of 10 species, we only profiled 16 strains in practice because there were
202 some overlaps between the two sets of top-ranked strains in the luminescence screen. We first
203 established that there was a strong positive association between the protein family (genomic) and
204 metabolic (Biolog) overlap of communities with the pathogens (**Fig. S11**). We then assessed the
205 ability of metabolic overlap to predict colonisation resistance (**Fig. 4c-d**). Colonisation resistance was
206 only observed once communities shared sufficiently high overlap in their carbon source utilisation

207 profile with a pathogen. Moreover, communities with the greatest metabolic overlap with a pathogen
208 provided the greatest colonisation resistance.

209 Overall, these data point to the importance of nutrient competition, and specifically nutrient
210 use overlap between a community and a pathogen, as an explanation for the patterns we observe in
211 colonisation resistance. To further support this conclusion, we performed experiments where the
212 pathogens were grown in cell-free (spent) media collected from different communities, which
213 excluded cell-cell contact mechanisms as explanations for colonisation resistance. Growing the
214 pathogen in the spent media of *E. coli* and the 10 species communities recapitulated the patterns seen
215 in the competition experiments, consistent with the effect of nutrient competition (**Fig. S12**). As a
216 final test, we sought a nutrient that can be used by the pathogens only and used it to perform nutrient
217 supplementation experiments (**Fig. 4e-f**). We identified galactitol from the Biolog plates (**Fig. S10**).
218 The pathogens can use this sugar alcohol but it has the desirable property that it cannot be used by any
219 of the symbionts in our focal 10-species communities, except for *E. coli*. We engineered a strain of *E.*
220 *coli* that lacks the transporter for cell import (*E. coli gatABC* deletion mutant). By adding in galactitol
221 to our standard media, we found that colonisation resistance in a diverse community is lost if the
222 pathogens can use the nutrient but *E. coli* cannot (**Fig. 4e-f**). However, colonisation resistance is
223 restored when *E. coli* can use the nutrient. Further, if a pathogen is engineered so that it cannot use
224 galactitol (*S. Typhimurium gatABC* deletion mutant), colonisation resistance is restored. These
225 outcomes are exactly as expected if nutrient competition is the cause of colonisation resistance.

226 Our data show that the ability of a microbiota community to consume nutrients required by a
227 pathogen for growth underlies the colonisation resistance we observe. Importantly, the nutrient
228 blocking effect is a property of the whole community rather than any one species alone. That
229 colonisation resistance is a community-level trait explains the importance of the ecological diversity,
230 and complexity (22) we observed in our experiments. Despite considerable genomic and metabolic
231 overlap with the pathogens, a species like *E. coli* cannot alone block enough nutrients to provide
232 colonisation resistance. It is only in combination with other species, that *E. coli* becomes effective at
233 limiting pathogen growth.

234

235 **Nutrient blocking can identify protective communities**

236 Our experiments indicate that colonisation resistance is an ecologically complex trait, but that this
237 complexity can be understood and predicted via a simple underlying principle. As an additional test of
238 these findings, we used the nutrient blocking to predict community compositions that provide
239 colonisation resistance to a bacterial strain that was not present in our initial experiments. For this test,
240 we chose an antimicrobial resistant (AMR) clinical *E. coli* strain, which was isolated from the urine of
241 a patient. AMR *E. coli* strains are a major current target for alternatives to antibiotics because
242 members of this species have recently been found to be responsible for the most AMR-associated
243 deaths of any bacterial species (39).

244 We first analysed the AMR *E. coli* isolate on AN Biolog Microplates to assess its carbon
245 source utilisation and compared this to the top ranked strains from our initial luminescence screen
246 (**Fig. 1**). We reasoned that these top-ranked strains were a good place to start as *E. coli* is also a
247 member of the Enterobacteriaceae, like the two pathogens that were used to select the top-ranked
248 strains. As expected, the AMR *E. coli* had the greatest protein overlap with the symbiont *E. coli* in our
249 16 strains but, importantly, additional strains were predicted to be required to block nutrient
250 availability based on the overlap needed to suppress the two pathogens (**Fig. S13**). We next used the
251 Biolog data to computationally assemble all possible communities of one, two, three and five species
252 from the 16 strains and calculated their resource utilisation overlap with the AMR *E. coli* (**Fig. 5a**).
253 Again, in line with our findings, diversity improved the median resource utilisation overlap, but this
254 depended strongly on the presence of the symbiont *E. coli*.

255 The simplest test of the importance of nutrient blocking is to remove the symbiont *E. coli*
256 from a community and test the impact. Doing this for the community of all 16 strains confirmed the
257 importance of *E. coli* for colonisation resistance (**Fig. 5b**). However, we also tested our ideas on
258 communities that contain *E. coli*. In these experiments we identified communities predicted to have
259 the highest and lowest overlap with the target strain at each diversity level, where communities were

260 randomly chosen if there were ties in rank. We then used our extended competition assay (**Fig. 1d**) to
261 test the ability of the AMR *E. coli* to invade the communities. As predicted by nutrient blocking, this
262 revealed that increasing diversity leads to increased colonisation resistance and, critically, for each
263 diversity level, the community predicted to resist the AMR *E. coli* consistently performed better in
264 colonisation resistance than the community predicted to do poorly (**Fig. 5b**). This result was clearest
265 for the two and three species communities. For the five species community, the best performing
266 community was only marginally better than the worst. We reasoned that this was because, in these
267 experiments, we are limited to choosing from 16 strains that were preselected for being relatively
268 good competitors to Enterobacteriaceae (**Fig. 1b-c, S1**).

269 To test the nutrient blocking principle more robustly we selected from a wider range of
270 possible strains from our set of 50 strains that we used in our *in vitro* and *in vivo* experiments (**Fig. 2-**
271 **3**). Most of these strains had not been characterised for their functioning in community-level
272 colonisation resistance, other than in the 50 species treatment. We also used this set of experiments to
273 test the power of the nutrient blocking principle to predict colonisation resistance based upon genomic
274 data alone. Rather than using the Biolog phenotypic assay, therefore, we returned to our measure of
275 protein family overlap, which calculated the overlap in all protein types between an invading strain
276 and different communities. To do this we only had to sequence the AMR *E. coli* clinical isolate
277 because all 50 other strains were sequenced. Using the same approach as for the Biolog predictions,
278 we then assembled communities *in silico* that all contained the symbiont *E. coli* strain and, in each
279 case, calculated their protein family overlap with the AMR *E. coli* (**Fig. 5c-d**). As before, we chose
280 communities with the lowest and highest overlap to the AMR *E. coli* across a range of diversities
281 (randomly choosing communities if there were ties in rank) and experimentally assessed colonisation
282 resistance using the extended competition assay.

283 We again see the importance of community diversity in these experiments. Moreover, despite
284 using only genomic information and a much larger set of possible communities, we observed
285 improvement in colonisation resistance from the worst to best communities at each diversity level
286 (**Fig. 5e**). Finally, we evaluated our ability to select highly and poorly performing communities by

287 assessing colonisation resistance in additional five-species communities. Notably, at the five-species
288 level, more than 200,000 communities with *E. coli* can be assembled from 50 strains. We used our
289 algorithm to sample approximately 50,000 and, from these, we identified four additional community
290 compositions predicted to perform well and four predicted to perform poorly (to give five of each
291 class). In line with our predictions, the communities predicted to be colonisation resistant showed a
292 median 100-fold reduction in the abundance of the AMR *E. coli* compared to those predicted to be
293 permissive (**Fig. 5f**).

294

295 **Discussion**

296 A key benefit of the microbiome is its ability to reduce the probability of infection via colonisation
297 resistance (1, 2, 10, 40). Here we have used an ecological approach to understand the principles of
298 colonisation resistance in the gut microbiome. By screening a collection of human gut symbionts, we
299 found that individual strains were unable to provide effective resistance to pathogens (**Fig. 1**), but that
300 colonisation resistance increases monotonically with ecological diversity (**Fig. 2-3**). Our work,
301 therefore, supports the general hypothesis that a more diverse microbiome can carry health benefits
302 (28, 41-43). While much discussed, evidence for this hypothesis is typically based upon correlations
303 between microbiome diversity and health outcomes (28, 43, 44). Here, we provide experimental
304 evidence that microbiome diversity can provide health benefits via an increased ability to protect
305 against pathogens. Moreover, we explained this pattern in terms of the importance of the overlap
306 between the nutrient requirements of an invading pathogen and the community (**Fig. 4**).

307 We found that certain combinations of species display much greater colonisation resistance
308 together than when alone. These non-additive effects mean that colonisation resistance is formally a
309 complex ecological trait in the canon of ecology (22). Such effects are often assumed to imply a
310 complex network of interactions between species where, for example, one symbiont species affects a
311 second symbiont species and changes the way this second species interacts with a pathogen. However,
312 consideration of nutrient competition and particularly the level of overlap between a pathogen and

313 community revealed that much simpler processes explain the complexity we see. One species alone is
314 not sufficient to strongly impact pathogen growth, but rather a combination of species is required to
315 block nutrient access. Interestingly, the combinations of species that make colonisation-resistant
316 communities are often very phylogenetically diverse. Generating resistant communities does not rest
317 upon simply finding closely related species to a given pathogen, in our case against
318 Enterobacteriaceae members. Instead, a mixture of gram-positive and gram-negative species is often
319 what performs best (**Fig. 5f, Table S7**). While both of our pathogens are members of the same family,
320 they have different life histories (19-21). *S. Typhimurium* is a specialist gut pathogen while *K.*
321 *pneumoniae* is an opportunistic pathogen that typically causes no pathology in the gut itself, instead
322 causing infections in other parts of the body. Consistent with this, we find it is consistently easier to
323 generate colonisation resistance to *K. pneumoniae* than *S. Typhimurium*. Nevertheless, the importance
324 of ecological diversity and complexity is observed for both pathogens. We anticipate that the
325 importance of ecological diversity and the principle of nutrient blocking will apply generally, given
326 the widespread evidence that nutrient competition is important for diverse species in the microbiome,
327 including many other pathogens (9, 45-53).

328 By assembling a wide range of communities of defined compositions, we have been able to
329 establish links between ecological diversity, complexity and nutrient competition in colonisation
330 resistance. However, a limitation of this approach is that we have focused on relatively low diversity
331 communities. A question for the future is whether our findings will hold for the higher levels of
332 diversity that can naturally occur in the human microbiome. Consistent with our findings, colonisation
333 risk with species like *K. pneumoniae* is increased after antibiotic treatment that can lower species
334 diversity in the microbiome (54). Nutrient competition is central to the patterns that we have
335 described here. Colonisation resistance can develop via additional mechanisms, which include toxin-
336 mediated bacterial competition and effects via the host immune system (1, 10). Our work does not
337 exclude the potential importance of these additional mechanisms such as direct killing of invading
338 strains by members of the community, which can act in parallel to nutrient competition (12, 13, 15,
339 37). Moreover, our predictions of resistant and permissive communities are not perfect (e.g., high

340 overlap communities in **Fig. 4c**; the outlier in **Fig. 5f**). However, it is notable that these deviations are
341 in the direction of a community being more resistant than expected. This suggests that predictions
342 based upon nutrient blocking may often be conservative with errors resulting in communities
343 performing better than expected whenever other mechanisms of colonisation resistance are at play.

344 Our work shows that colonisation resistance is not the property of single microbiome species
345 but instead the collective property of multiple species. Specifically, we find that the effect of a given
346 symbiont species on a pathogen can be strongly dependent on whether other symbiont species are also
347 present to consume nutrients that the pathogen needs. This finding suggests that one can use the idea
348 of nutrient blocking to identify sets of microbiome species that will limit the growth of a target strain.
349 As a proof-of-principle, we tested this idea for an AMR *E. coli* strain, which revealed that species sets
350 can be successfully identified that collectively suppress an incoming strain (**Fig. 5**). Importantly, we
351 find that this can be done without specific information on available nutrients or the metabolism of the
352 species under study. Instead, a measure of genomic overlap can be used as a proxy for niche overlap
353 to assemble communities that perform nutrient blocking (**Fig. 5c-f**). The human microbiome is
354 dauntingly complex and has great potential for context-dependent effects. However, we found that
355 microbiome complexity can arise via simple underlying principles, which gives promise to the goal of
356 rationally designing microbiomes for better health.

357

358 *Please insert Materials and Methods here*

359 **Materials and methods**

360 **Bacterial strains and plasmids** A full list of bacterial strains used in this study is provided in **Table**
361 **S1**. The plasmids used in this study are listed in **Table S2**. *Klebsiella pneumoniae* subsp. *pneumoniae*
362 purchased from DSMZ (stock number 30104) was used in all experiments containing *K. pneumoniae*.
363 *Salmonella enterica* serovar Typhimurium (strain SL1344 (55)) was used in all experiments containing
364 *S. Typhimurium*. Symbiont strains were chosen based on those commonly isolated from the human gut
365 microbiota (56). Strains were kindly provided by Nassos Typas (EMBL Heidelberg, Germany), or were
366 ordered from the German Collection of Microorganisms, DSMZ or the American Type Culture
367 Collection (ATCC; *Staph epidermidis*). The *Lactobacillus plantarum* strain was provided by the

368 Department of Food and Nutritional Sciences, University of Reading. *Klebsiella pneumoniae* ATCC
369 700721 was provided by the Modernising Medical Microbiology research group, Nuffield Department
370 of Medicine, University of Oxford. *Escherichia coli* strains Z1331 and Z1269 were isolated from the
371 faeces of two healthy human donors in Switzerland (31). The ampicillin-resistant AMR *E. coli* strain is
372 a urine clinical isolate provided by the Pathogen Bank at Nottingham University Hospitals NHS Trust.

373 **Bacterial growth conditions** For engineering of *E. coli*, *K. pneumoniae*, or *S. Typhimurium*, strains
374 were grown aerobically in Lysogeny broth (LB; Fisher Scientific) with the appropriate antibiotics
375 (Table S1) at 37°C, shaking at 220 rpm. All symbionts were cultured under anaerobic conditions (5%
376 H₂, 5% CO₂, 90% N₂, <20ppm O₂) in modified Gifu Anaerobic Medium (mGAM; Nissui
377 Pharmaceuticals) broth buffered to pH 6.2 with 100mM 2-(N-morpholino)ethanesulfonic acid (MES;
378 Sigma-Aldrich). The redox indicator dye Resazurin (100µg/L media; Sigma-Aldrich) was added as a
379 quality control check for the presence of oxygen (turns red when conditions are not sufficiently
380 anaerobic). To prepare glycerol stocks, individual strains were first streaked onto mGAM agar (Nissui
381 Pharmaceuticals) and grown under anaerobic conditions (if possible, as not all strains can grow as single
382 colonies on agar-based media). The identity of each strain was confirmed using 16S rRNA sequencing
383 (Sanger sequencing; Source Biosciences) using the primers oOPC-953 and oOPC-954 for most species,
384 or g-Bifid-F and g-Bifid-R for Bifidobacteria (Table S3). Single colonies were then inoculated in
385 mGAM broth and stored at -70 degrees Celsius in mGAM with a final concentration of 25% glycerol.

386 **Genetic engineering of bacterial strains** Luminescence and fluorescence plasmids were transformed
387 into *K. pneumoniae* and *S. Typhimurium* using electroporation. Briefly, 5mL overnight culture was
388 washed 3x with cold Milli-Q water, before being concentrated in 500µl cold Milli-Q water. 2µl of
389 plasmid was mixed with 100µl of concentrated cells and electroporated (1.8kV; 0.1cm gap cuvettes)
390 before recovery in 1ml pre-warmed LB (1h at 37°C, shaking at 220rpm) and plating on LB with the
391 appropriate antibiotics.

392 Gene deletions were generated as described in (57). Briefly, 700 base pairs upstream and downstream
393 of region to be deleted were PCR amplified (Phusion[®], NEB) and inserted into the suicide vector pFOK,
394 which was linearised using the restriction enzymes BamHI and EcoRI, using the NEBuilder[®] HiFi DNA
395 Assembly (NEB). The plasmid was introduced into a diaminopimelic acid auxotroph *E. coli* strain
396 (JKe201). After 6h of mating between the plasmid-containing donor *E. coli* strain and the recipient *E.*
397 *coli*, *S. Typhimurium* or *K. pneumoniae* strain, trans-conjugants were selected on LB plates containing
398 50µg/mL kanamycin. Counter-selection was performed on no-salt LB plates supplemented with
399 0.5µg/mL of anhydrous tetracycline and 20% of sucrose at 30°C. Mutants were screened by colony
400 PCR and the sequence was verified using Sanger sequencing (Eurofins). Primers used for genetic
401 engineering in this study are listed in Table S3.

402 **Luminescence screen** *K. pneumoniae* (DSM 30104) and *S. Typhimurium* (SL1344) carrying a low
403 copy number plasmid with P_{nptII} promoter-driven expression of *luxCDBAE-frp* (pRSJ- $p_{nptII}::ilux$; also
404 known as the improved lux operon plasmid (58)) were used for the luminescence screen. Symbiont
405 strains were tested in pairwise co-culture with the pathogens. All strains tested were first grown in
406 monoculture anaerobically in static mGAM broth buffered to pH 6.2 with 100mM MES at 37°C. *K.*
407 *pneumoniae* WT and *S. Typhimurium* WT carrying the luminescence plasmid pRSJ- $p_{nptII}::ilux$ were
408 separately incubated in the same media overnight.

409 *Competition assay*: once all strains had reached stationary phase (12-72 hours of growth depending on
410 the species), 20µL of each symbiont strain was dispensed in technical triplicates in a 96-well plate
411 containing 160µL of mGAM broth. 20µL of *K. pneumoniae* pRSJ- $p_{nptII}::ilux$ overnight culture or *S.*
412 *Typhimurium* pRSJ- $p_{nptII}::ilux$ overnight culture was then added to each well.

413 *Ecological invasion assay*: 180µL of each symbiont strain was dispensed in technical triplicates in a
414 96-well plate and 20µL *K. pneumoniae* pRSJ- $p_{nptII}::ilux$ overnight culture or *S. Typhimurium* pRSJ-
415 $p_{nptII}::ilux$ overnight culture was added to each well.

416 96-well plates from both assays were incubated anaerobically for 6 hours at 37°C. Before luminescence
417 measurement, 96-well plates were brought out of the anaerobic chamber and exposed to O₂ for 10
418 minutes because, critically, O₂ is needed for light production catalysed by the enzyme luciferase.
419 Luminescence at 515-575nm and OD-600nm were measured using a CLARIOstar Plus spectrometer
420 (BMG Labtech). Each data point (**Fig. S1**) represents the median of at least 3 independent biological
421 replicates (that is, using different overnight cultures of the same strain in different 96-well plates).

422 **Selection of the top and bottom ten symbiont strains from the screen** For each of the ecological
423 invasion and competition assays, symbiont strains were ordered from best- to worst-performing.
424 Symbiont strains were given two rankings: one from the competition assay and one from the invasion
425 assay. The two rankings were summed and ordered from lowest sum (best overall performing) to highest
426 sum (worst overall performing). From these ranks, the top ten and bottom ten strains were chosen, with
427 the added selection criteria that hazard group 2 strains (European designation) and repeat strains of the
428 same species were excluded. We excluded hazard group 2 strains to focus on communities of species
429 that would be considered safe as therapeutics.

430 **Construction of phylogenetic trees** The phylogenomic trees of species in the luminescence screen
431 (n=100; **Fig. 1b-c; Table S1**) was inferred using default settings based on 67 single-copy core genes
432 using anvio v7.1 (59) and plotted using iTOL v6.7.5 (60).

433 **Extended competition assay** All culturing was performed in a shaking incubator at 225rpm and 37°C,
434 and under anaerobic conditions. Symbiont strains were grown to stationary phase (12-72 hours,
435 depending on the species), in Hungate tubes containing 5mL mGAM broth and 100mM MES and
436 buffered to pH 6.2. *K. pneumoniae* and *S. Typhimurium* carrying the fluorescence plasmid pBC11
437 (YPet) (61) were incubated overnight, with the appropriate antibiotics (**Table S1**).

438 Once in stationary phase, monocultures of symbionts were passaged (100µL culture into a new 5mL
439 tube of media) and grown for ~17 hours. Communities were assembled under anaerobic conditions into
440 new 5mL Hungate tubes (see section “community preparation”) and grown for 24 hours. Once grown,
441 communities were invaded with 100µL of the pathogen (10^6 cells/mL final concentration) in each
442 Hungate tube containing 5mL culture. Samples were taken immediately after addition of the pathogen
443 and prepared for measurement using flow cytometry (day 0). After 24 hours growth, the invaded
444 communities were sampled for flow cytometry (day 1) and 100µL was passaged into new tubes of 5mL
445 mGAM broth. After a further 24 hours, the end-point communities were sampled and measured a final
446 time (day 2).

447 For nutrient supplementation experiments (galactitol), 2.5mL mGAM buffered with 100mM MES to
448 pH 6.2 was prepared and autoclaved in Hungate tubes. A 2x stock solution of filter-sterilised sugar in
449 Milli-Q water (e.g. 0.2% or 2% galactitol) was prepared and de-oxygenated in the anaerobic chamber.
450 2.5mL of the stock solution was added into the Hungate tube to generate a 5ml total volume with the
451 appropriate concentration of the sugar. The extended competition experiment is performed as described
452 above, except samples were passaged again into tubes with the appropriate concentration of the sugar.
453 Samples were left to grow for 48h instead of the standard 24h to allow utilisation of low priority sugars,
454 such as galactitol.

455 For experiments with the AMR *E. coli* clinical isolate, the experiment was performed as described
456 above, except samples were analysed using selective plating instead of flow cytometry (LB + 100µg/mL
457 ampicillin to select for the AMR *E. coli* strain; MacConkey agar to enumerate total *E. coli* densities).

458 **Community preparation** All culturing was performed in a shaking incubator at 225rpm and 37°C, and
459 under anaerobic conditions. Constituent monocultures were grown to stationary phase, passaged into
460 new media and grown overnight for approximately ~17 hours, and then combined into communities.
461 The OD-600nm of the single strains was measured and the cultures were streaked anaerobically on
462 mGAM agar to check for contamination. Communities ranged from 2 to 50 strains in size and were
463 assembled under anaerobic conditions.

464 *In vitro*: For the communities containing 2-10 strains, equal volumes (100µL) of each overnight culture
465 were combined in a Hungate tube containing 5mL mGAM to form a community. For the 49- and 50-

466 strain communities, 20 μ L of each single strain was added, equating to a total volume of 1mL. The
467 communities were made by adding 1mL of each strain anaerobically to a 50mL falcon tube, mixing,
468 and removing 1mL to add to a fresh Hungate tube of 5mL mGAM broth. Communities were grown in
469 the shaking incubator (37°C, 225rpm) for 24 hours before being invaded with 10⁶ cells/mL of the
470 pathogen.

471 *In vivo*: Strains were grown separately, then OD-600nm of each culture was measured. For communities
472 containing 2-10 strains, the volume of culture containing 10⁹ cells was calculated for each individual
473 strain, based on a standard curve of OD-600nm and flow cytometry quantification. This was added to a
474 large falcon tube. The community mix was centrifuged (10 minutes, room temperature, 14000rpm) and
475 the resulting pellet resuspended in 2mL anaerobic PBS and kept on ice. 200 μ L were gavaged into each
476 mouse corresponding to 10⁸ cells for each strain. For the 49- or 50-strain communities, we used less
477 cells for each strain to avoid gavaging a large total cell density into mice (as done for the *in vitro*
478 experiments). The volume of culture containing 2*10⁸ cells was calculated for each constituent strain
479 and they were combined. The mixture was centrifuged and the pellet resuspended in 2mL PBS. Mice
480 were gavaged with 200 μ L of the inoculum. Therefore, the density of bacteria in the 200 μ L gavage was
481 equivalent for the 50-strain community as for the 10-strain community (10⁹ total cells). For all
482 inoculums, samples were taken for plating anaerobically on mGAM agar to confirm the cell density.

483 **Flow cytometry** Hungate tubes containing culture to be sampled were de-pressurised under anaerobic
484 conditions using needles. Samples were removed using needles and syringes. 10 μ L of each sample was
485 diluted in 90 μ L PBS + chloramphenicol 200 μ g/ml then 10 μ L added to a 96-well flat-bottomed plate
486 already containing 80 μ L PBS + chloramphenicol 200 μ g/ml. The 96-well plate was left to shake at room
487 temperature on the bench for 15 minutes to allow oxygen-dependent folding of YPet, while preventing
488 further growth of the sample with the bacteriostatic antibiotic chloramphenicol. Next, all bacterial cells
489 were fixed, permeabilised and fluorescently stained by adding 90 μ L PBS containing 4%
490 paraformaldehyde, 0.4% Triton X-100 and 1 μ g/mL 4',6-diamidino-2-phenylindole (DAPI) to each
491 sample well and incubating the plate in the dark for 1 hour at room temperature on a microplate shaker.
492 Prior to the measurement, 20 μ L AccuCheck counting beads (ThermoFisher- PBC100) was added to
493 every sample well. Two wells containing 200 μ L PBS were left between sample wells, in order to
494 prevent cross-contamination between samples. 96-well plates were run in a flow cytometer (Attune
495 NxT Autosampler with Attune NxT Flow Cytometer, ThermoFisher) set to: acquisition volume 50 μ L,
496 total draw volume 80 μ L, total sample volume 200 μ L, at a speed of 100 μ L/minute. Relevant spectral
497 parameters were recorded by the flow cytometer, equipped with 405nm, 488nm and 561nm lasers. The
498 following channels were used: DAPI staining, excitation 405nm, emission 415-465nm ('blue'); YPet
499 fluorescence, excitation 488nm, emission 505-515nm and 525-555nm ('green').

500 Flow cytometry data was processed using FlowJo v10.8.1 software. ‘Beads’ were gated on a linear
501 (FSC-H; SSC-H) axis. ‘Bacteria’ were gated on a logarithmic (FSC-H; SSC-H) axis. DAPI was gated
502 within the ‘Bacteria’ gate, on a logarithmic (Ex405-417; Ex405-495) axis. Within the DAPI gate, YPet
503 was gated on a logarithmic (Ex488-503; Ex488-555) axis. Count statistics for beads, DAPI and YPet
504 were exported to Microsoft Excel and pathogen density (cells/mL) was calculated.

505 **Mouse husbandry and experiments** Mouse experiments were performed with 6-8 weeks old germ-
506 free C57BL/6J female mice that were bred and maintained in germ-free isolators at the Kennedy
507 Institute, University of Oxford. During experiments, mice were housed as pairs or trios in sterile,
508 individually ventilated Sentry SPP cages (Allentown) with enrichment, irradiated food and autoclaved
509 water. Cages of mice were disinfected in TecCare Ultra Hydrogen peroxide, peracetic acid solution
510 before mice were handled using individually sterilised gloves (single use for each cage) in a biosafety
511 cabinet that was disinfected with the same acid solution in between each treatment group. Cages were
512 randomly assigned to a group, and cages were opened in the same order each time. Mouse experiments
513 were performed in accordance with the UK Animal Scientific Procedures Act (1986) under a Project
514 License authorised by the UK Home Office.

515 Germ-free mice were gavaged with a 200 μ L inoculum of symbionts (see “community preparation”) at
516 day -14 of the experiment and again on day -12 of the experiment (see scheme **Fig. 3a**). On day 0,
517 overnight cultures of the pathogens (*K. pneumoniae* and *S. Typhimurium*) grown aerobically in LB
518 containing the appropriate antibiotics were washed 3x with PBS and diluted. 10⁶ CFU of the pathogen
519 was given to each mouse as a 100 μ l gavage. Faeces were sampled daily after gavage of the pathogen.
520 Faecal pellets were weighed and homogenised with a 5mm stainless steel bead in 1ml PBS (samples
521 were shaken vigorously in 2ml Eppendorf tubes). CFUs were enumerated using selective plating
522 (MacConkey or LB agar for *E. coli*; MacConkey + 50 μ g/mL streptomycin for *S. Typhimurium*; LB +
523 50 μ g/mL carbenicillin for *K. pneumoniae*). *E. coli*, *K. pneumoniae* and *S. Typhimurium*, can be
524 differentiated by colour on MacConkey agar (*S. Typhimurium* cannot utilise lactose) or morphology
525 (*K. pneumoniae* produces a capsule). Mice were euthanised 4 days after infection with the pathogen.

526 *S. Typhimurium* triggers inflammation to bloom in the gut (19). To allow us to focus on early
527 colonisation events prior to triggering inflammation (i.e., colonisation resistance), we used an avirulent
528 strain (SL1344 $\Delta invG \Delta ssaV$) to avoid triggering inflammation, disease and mortality (62).

529 **DNA isolation from faeces and metagenomic sequencing** Aliquots of the faeces were frozen
530 immediately after collection (prior to homogenisation in PBS) and stored at -80°C until use. Faecal
531 samples were thawed and resuspended in nuclease-free water, before being transferred to a lysing matrix
532 B tube (MP Biomedicals). 3 rounds of bead beating were performed at 6 m/s for 40s. Samples were
533 centrifuged at high speed and DNA in the supernatant was precipitated by adding sodium acetate (1/10
534 volume) and ice-cold ethanol (96-100%; equal volume), and left at -20°C overnight. Samples were

535 centrifuged at high speed and the pellet was washed twice with 70% ethanol, before being dried and
536 resuspended in nuclease-free water.

537 Samples were further purified using an AMPure clean-up protocol. The samples were mixed with
538 AMPure XP beads (Beckman Coulter) and incubated at room temperature for 5 minutes, before being
539 washed twice with 70% ethanol, using a magnet to avoid removing the DNA bound to the AMPure
540 beads. After air drying, nuclease-free water was added to the beads, to allow collection of the purified
541 DNA in the supernatant.

542 CosmosID performed metagenomic sequencing on isolated DNA. DNA libraries were prepared using
543 the Nextera XT DNA Library Preparation Kit (Illumina) and IDT Unique Dual Indexes. Prepared
544 libraries were sequenced on the Illumina NovaSeq platform 2x150bp (3M reads per sample). CosmosID
545 performed bioinformatic analysis according to proprietary methods on the raw data to generate fine-
546 grained taxonomic and relative abundance estimates (**Fig. S5**).

547 **Spent media assay** Strains were anaerobically inoculated, grown, passaged and assembled into
548 communities as described above. Communities were grown anaerobically in the shaking incubator
549 (225rpm, 37°C) for 96 hours to maximise depletion of nutrients utilisable by the community.
550 Supernatants were prepared by centrifuging the 96-hour cultures (10 minutes, 4400rpm, room
551 temperature) and filter-sterilising the supernatant (0.2µm filter). All preparation of supernatants were
552 done under anaerobic conditions. Two of each of the communities were assembled so that, for each
553 community, one Hungate tube containing 5mL supernatant and another Hungate tube containing 2.5mL
554 supernatant and 2.5mL mGAM (the re-supplemented treatment) were prepared. The supernatants were
555 invaded with 10⁶ cells/mL of fluorescent pathogen, sampled and incubated for 24 hours as before.
556 Samples were taken on days 0 and 1.

557 **Nutrient utilisation overlap using biolog assays** Utilisation of carbon sources by symbionts and the
558 pathogens were assessed using AN Biolog MicroplatesTM (Biolog) according to the protocol of the
559 manufacturer. Briefly, strains were grown in mGAM and passaged (as detailed above). The cultures
560 were centrifuged anaerobically and washed twice in anaerobic PBS. The samples were concentrated
561 and an aliquot was taken to measure the OD-600nm. The cells were diluted or concentrated such that a
562 200µl aliquot of the cells would correspond to 65% transmittance (OD 0.187; the density outlined by
563 the manufacturer) in the 14ml volume of AN inoculating fluid provided. The concentrated cells were
564 added to the inoculating fluid aerobically (a small amount of O₂ is needed to oxidise the buffer) and
565 100µl was aliquoted into each well of the microplate. After 10min in aerobic conditions, the plates were
566 put in an air-tight container with a GasPak EZ Container System Sachet to generate hydrogen-free
567 anaerobic conditions. Samples were incubated at 35°C for 24h before measurement at 590nm. Since *S.*
568 Typhimurium SL1344 is a histidine auxotroph, a prototrophic strain was made using P22 transduction
569 (63) of the allele from a prototrophic strain of *S. Typhimurium* (ATCC 14028S). Positive clones were

570 selected for on M9 minimal media with 50µg/ml streptomycin and re-streaked 3 times on LB plates to
571 ensure phages are removed. Growth in the absence of exogenous histidine was validated by streaking
572 on M9 media.

573 For each plate, the absorbance reading at 590nm was subtracted from the blank (no carbon source
574 control in well A1). Each strain generates a different background signal so it is important to do this on
575 each plate individually. Each strain was measured as three independent biological replicates and the
576 median absorbance value was taken for each carbon source. We used a thresholding approach to
577 determine if a strain can metabolise a given carbon source (defined as Abs 590nm >0.1 median, after
578 blank subtraction; **Fig. S10**).

579 To predict overlap with the pathogens, whether a given strain can use a carbon source was compared to
580 the pathogen. We assessed which percentage of carbon sources that the pathogen can use can also be
581 used by the symbiont. To calculate overlap between pathogens and communities, we used an additive
582 calculation approach, where if a strain is contained in a community that can use a carbon source, the
583 entire community can use the carbon source. For simplicity, we did not treat cases where multiple
584 species use the same nutrients within the community differently than if a given nutrient is only covered
585 by one species.

586 **Genomic analysis using protein overlap** Genomic information from 50 symbiont strains and two
587 pathogens (*S. Typhimurium*, *K. pneumoniae*) were retrieved from the PATRIC database (64) (**Table**
588 **S1**). On this set, we applied a cluster-based analysis that groups proteins encoded in genomes into
589 PATRIC global protein families (65). A majority of symbiont strain proteins obtained a protein family
590 designation (clustering rate of 97.83%; 172 848 of 176 675) with protein families being populated by
591 proteins from on average 2.44 genomes (**Fig. S7**).

592 Based on this set, we drew protein family designations for focal symbiont communities and calculated
593 pathogen/community overlap using the same approach as for the Biolog plates above. That is, for each
594 protein family encoded in the pathogen's genome, we checked if it overlapped with any protein families
595 encoded in the community.

596 **Whole-genome sequencing** An overnight culture of *E. coli* 19Y000018 (the AMR clinical urine
597 isolate) was prepared in LB containing 100µg/ml ampicillin. A 1mL pellet was taken and DNA was
598 extracted using the same ethanol precipitation and AMPure clean-up protocol detailed above in the
599 section "DNA isolation from faeces". Source Biosciences performed whole-genome Illumina
600 sequencing on a NovaSeq 600 to generate 10M 150bp paired end reads. Genome assembly was
601 performed with Unicycler v0.4.8 using default settings and subjected to protein family annotation as
602 described in "Genomic analysis using protein overlap".

603 **Predictions based on biologists and protein families** *Biologist-based predictions:* The carbon source
604 utilisation profile of the AMR *E. coli* clinical isolate was assessed using AN Biolog Microplates™ as
605 detailed above. Next, all possible communities made of 1, 2, 3, and 5 species were identified from the
606 16 symbiont strains that were previously analysed for carbon source utilisation. For each community,
607 the carbon source utilisation overlap was calculated (as detailed above; **Fig. 5a**). At each diversity level,
608 communities were sorted by overlap to the AMR *E. coli* clinical isolate. Given that the symbiont *E. coli*
609 is needed for colonisation resistance in our experiments, we only included communities where the
610 symbiont *E. coli* is present. At each diversity level, we chose the highest ranked community (defined
611 “predicted best”) and lowest ranked community (defined “predicted worst”). If there were ties in rank,
612 one community was randomly chosen. These communities were experimentally assessed for
613 colonisation resistance using the extended competition *in vitro* assay as detailed above.

614 *Protein family-based predictions:* We used the whole-genome sequence of the AMR *E. coli* isolate to
615 calculate protein family overlap to communities of symbionts drawn from our 50-species community.
616 As for the biologist-based predictions, we restricted the analysis to communities that contained the
617 symbiont *E. coli* strain. Specifically, we generated 100,000 different communities randomly, each
618 containing the symbiont *E. coli* and a total of 2, 3, 5, or 10 species from the 50 species pool. We ranked
619 the communities according to their pathogen overlap at each diversity level and selected the highest and
620 lowest ranked communities for experimental validation. If ties occurred, we chose the communities at
621 random, as for the biologist-based predictions.

622 **Statistical analysis** All graphs were made and statistical analysis carried out in Prism v9.4.1
623 (GraphPad). Protein family overlap analyses and protein family-based predictions were computed in R
624 version v4.0.5 (66). Figure legends indicate the statistical test used and the sample sizes.

625 **Ethical statement** Mouse experiments were performed in accordance with the UK Animal Scientific
626 Procedures Act (1986) under a Project Licence authorised by the UK Home Office. Mice were
627 euthanised using cervical dislocation followed by exsanguination as a confirmation of death.

628 **Data and code availability** We will publish the code, sequences, and data along with publication of
629 the paper.

630 **References**

- 631 1. M. T. Sorbara, E. G. Pamer, Interbacterial mechanisms of colonization resistance and the
632 strategies pathogens use to overcome them. *Mucosal Immunol* **12**, 1-9 (2019).
- 633 2. A. Jacobson *et al.*, A Gut Commensal-Produced Metabolite Mediates Colonization Resistance
634 to Salmonella Infection. *Cell Host Microbe* **24**, 296-307 e297 (2018).
- 635 3. E. M. Velazquez *et al.*, Endogenous Enterobacteriaceae underlie variation in susceptibility to
636 Salmonella infection. *Nat Microbiol* **4**, 1057-1064 (2019).
- 637 4. S. Y. Wotzka *et al.*, Escherichia coli limits Salmonella Typhimurium infections after diet shifts
638 and fat-mediated microbiota perturbation in mice. *Nat Microbiol* **4**, 2164–2174 (2019).

- 639 5. R. P. Sequeira, J. A. K. McDonald, J. R. Marchesi, T. B. Clarke, Commensal Bacteroidetes protect
640 against *Klebsiella pneumoniae* colonization and transmission through IL-36 signalling. *Nat*
641 *Microbiol* **5**, 304-313 (2020).
- 642 6. C. G. Buffie *et al.*, Precision microbiome reconstitution restores bile acid mediated resistance
643 to *Clostridium difficile*. *Nature* **517**, 205-208 (2015).
- 644 7. S. G. Kim *et al.*, Microbiota-derived lantibiotic restores resistance against vancomycin-
645 resistant *Enterococcus*. *Nature* **572**, 665-669 (2019).
- 646 8. M. X. Byndloss *et al.*, Microbiota-activated PPAR-gamma signaling inhibits dysbiotic
647 *Enterobacteriaceae* expansion. *Science* **357**, 570-575 (2017).
- 648 9. R. A. Oliveira *et al.*, *Klebsiella michiganensis* transmission enhances resistance to
649 *Enterobacteriaceae* gut invasion by nutrition competition. *Nat Microbiol* **5**, 630-641 (2020).
- 650 10. G. Caballero-Flores, J. M. Pickard, G. Nunez, Microbiota-mediated colonization resistance:
651 mechanisms and regulation. *Nat Rev Microbiol* **21**, 347-360 (2023).
- 652 11. E. T. Granato, T. A. Meiller-Legrand, K. R. Foster, The Evolution and Ecology of Bacterial
653 Warfare. *Curr Biol* **29**, R521-R537 (2019).
- 654 12. C. Eberl *et al.*, *E. coli* enhance colonization resistance against *Salmonella Typhimurium* by
655 competing for galactitol, a context-dependent limiting carbon source. *Cell Host Microbe* **29**,
656 1680-1692 e1687 (2021).
- 657 13. L. Osbelt *et al.*, *Klebsiella oxytoca* causes colonization resistance against multidrug-resistant
658 *K. pneumoniae* in the gut via cooperative carbohydrate competition. *Cell Host Microbe* **29**,
659 1663-1679 e1667 (2021).
- 660 14. S. Caballero *et al.*, Cooperating Commensals Restore Colonization Resistance to Vancomycin-
661 Resistant *Enterococcus faecium*. *Cell Host Microbe* **21**, 592-602 e594 (2017).
- 662 15. S. Brugiroux *et al.*, Genome-guided design of a defined mouse microbiota that confers
663 colonization resistance against *Salmonella enterica* serovar *Typhimurium*. *Nat Microbiol* **2**,
664 16215 (2016).
- 665 16. A. G. Cheng *et al.*, Design, construction, and in vivo augmentation of a complex gut
666 microbiome. *Cell* **185**, 3617-3636 e3619 (2022).
- 667 17. S. Widder *et al.*, Challenges in microbial ecology: building predictive understanding of
668 community function and dynamics. *ISME J* **10**, 2557-2568 (2016).
- 669 18. E. Tacconelli *et al.*, Discovery, research, and development of new antibiotics: the WHO priority
670 list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**, 318-327 (2018).
- 671 19. B. Stecher *et al.*, *Salmonella enterica* serovar *typhimurium* exploits inflammation to compete
672 with the intestinal microbiota. *PLoS Biol* **5**, 2177-2189 (2007).
- 673 20. S. E. Majowicz *et al.*, The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin Infect*
674 *Dis* **50**, 882-889 (2010).
- 675 21. C. L. Gorrie *et al.*, Gastrointestinal Carriage Is a Major Reservoir of *Klebsiella pneumoniae*
676 Infection in Intensive Care Patients. *Clin Infect Dis* **65**, 208-215 (2017).
- 677 22. M. M. Mayfield, D. B. Stouffer, Higher-order interactions capture unexplained complexity in
678 diverse communities. *Nat Ecol Evol* **1**, 62 (2017).
- 679 23. Y. Litvak, A. J. Baumler, The founder hypothesis: A basis for microbiota resistance, diversity in
680 taxa carriage, and colonization resistance against pathogens. *PLoS Pathog* **15**, e1007563
681 (2019).
- 682 24. K. Z. Coyte, S. Rakoff-Nahoum, Understanding Competition and Cooperation within the
683 Mammalian Gut Microbiome. *Curr Biol* **29**, R538-R544 (2019).
- 684 25. O. Manor *et al.*, Health and disease markers correlate with gut microbiome composition
685 across thousands of people. *Nat Commun* **11**, 5206 (2020).
- 686 26. C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, R. Knight, Diversity, stability and
687 resilience of the human gut microbiota. *Nature* **489**, 220-230 (2012).
- 688 27. C. Tropini *et al.*, Transient Osmotic Perturbation Causes Long-Term Alteration to the Gut
689 Microbiota. *Cell* **173**, 1742-1754 e1717 (2018).

- 690 28. K. V. Johnson, P. W. Burnet, Microbiome: Should we diversify from diversity? *Gut Microbes* **7**,
691 455-458 (2016).
- 692 29. A. Baichman-Kass, T. Song, J. Friedman, Competitive interactions between culturable bacteria
693 are highly non-additive. *Elife* **12**, e83398 (2023).
- 694 30. T. J. C. Ian Billick, Higher Order Interactions in Ecological Communities: What Are They and
695 How Can They be Detected? *Ecology* **75**, 15 (1994).
- 696 31. S. Y. Wotzka *et al.*, Microbiota stability in healthy individuals after single-dose lactulose
697 challenge-A randomized controlled study. *PLoS One* **13**, e0206214 (2018).
- 698 32. S. B. Formal, G. J. Dammin, E. H. Labrec, H. Schneider, Experimental Shigella infections:
699 characteristics of a fatal infection produced in guinea pigs. *J Bacteriol* **75**, 604-610 (1958).
- 700 33. F. R. Blattner *et al.*, The complete genome sequence of Escherichia coli K-12. *Science* **277**,
701 1453-1462 (1997).
- 702 34. M. Ackermann *et al.*, Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**,
703 987-990 (2008).
- 704 35. E. Gul *et al.*, The microbiota conditions a gut milieu that selects for wild-type Salmonella
705 Typhimurium virulence. *PLoS Biol* **21**, e3002253 (2023).
- 706 36. R. J. Gibbons, S. S. Socransky, B. Kapsimalis, Establishment of Human Indigenous Bacteria in
707 Germ-Free Mice. *J Bacteriol* **88**, 1316-1323 (1964).
- 708 37. R. Freter, H. Brickner, M. Botney, D. Cleven, A. Aranki, Mechanisms that control bacterial
709 populations in continuous-flow culture models of mouse large intestinal flora. *Infect Immun*
710 **39**, 676-685 (1983).
- 711 38. A. Wagner, Competition for nutrients increases invasion resistance during assembly of
712 microbial communities. *Mol Ecol* **31**, 4188-4203 (2022).
- 713 39. C. Antimicrobial Resistance, Global burden of bacterial antimicrobial resistance in 2019: a
714 systematic analysis. *Lancet* **399**, 629-655 (2022).
- 715 40. M. R. McLaren, B. J. Callahan, Pathogen resistance may be the principal evolutionary
716 advantage provided by the microbiome. *Philos Trans R Soc Lond B Biol Sci* **375**, 20190592
717 (2020).
- 718 41. M. J. Blaser, The theory of disappearing microbiota and the epidemics of chronic diseases. *Nat*
719 *Rev Immunol* **17**, 461-463 (2017).
- 720 42. M. Fassarella *et al.*, Gut microbiome stability and resilience: elucidating the response to
721 perturbations in order to modulate gut health. *Gut* **70**, 595-605 (2021).
- 722 43. H. C. Wastyk *et al.*, Gut-microbiota-targeted diets modulate human immune status. *Cell* **184**,
723 4137-4153 e4114 (2021).
- 724 44. E. Le Chatelier *et al.*, Richness of human gut microbiome correlates with metabolic markers.
725 *Nature* **500**, 541-546 (2013).
- 726 45. S. E. Winter *et al.*, Gut inflammation provides a respiratory electron acceptor for Salmonella.
727 *Nature* **467**, 426-429 (2010).
- 728 46. K. M. Pruss, J. L. Sonnenburg, C. difficile exploits a host metabolite produced during toxin-
729 mediated disease. *Nature* **593**, 261-265 (2021).
- 730 47. L. Maier *et al.*, Microbiota-derived hydrogen fuels Salmonella typhimurium invasion of the gut
731 ecosystem. *Cell Host Microbe* **14**, 641-651 (2013).
- 732 48. M. L. Jenior, J. L. Leslie, V. B. Young, P. D. Schloss, Clostridium difficile Colonizes Alternative
733 Nutrient Niches during Infection across Distinct Murine Gut Microbiomes. *mSystems* **2**,
734 e00063-00017 (2017).
- 735 49. K. M. Ng *et al.*, Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric
736 pathogens. *Nature* **502**, 96-99 (2013).
- 737 50. A. W. Hudson, A. J. Barnes, A. S. Bray, D. A. Ornelles, M. A. Zafar, Klebsiella pneumoniae I-
738 Fucose Metabolism Promotes Gastrointestinal Colonization and Modulates Its Virulence
739 Determinants. *Infect Immun* **90**, e0020622 (2022).

- 740 51. A. G. Jimenez, M. Ellermann, W. Abbott, V. Sperandio, Diet-derived galacturonic acid regulates
741 virulence and intestinal colonization in enterohaemorrhagic Escherichia coli and Citrobacter
742 rodentium. *Nat Microbiol* **5**, 368-378 (2020).
- 743 52. F. C. Pereira *et al.*, Rational design of a microbial consortium of mucosal sugar utilizers reduces
744 Clostridiodes difficile colonization. *Nat Commun* **11**, 5104 (2020).
- 745 53. E. Gul *et al.*, Differences in carbon metabolic capacity fuel co-existence and plasmid transfer
746 between Salmonella strains in the mouse gut. *Cell Host Microbe* **31**, 1140-1153 e1143 (2023).
- 747 54. N. Raffelsberger *et al.*, Gastrointestinal carriage of Klebsiella pneumoniae in a general adult
748 population: a cross-sectional study of risk factors and bacterial genomic diversity. *Gut*
749 *Microbes* **13**, 1939599 (2021).
- 750 55. S. K. Hoiseth, B. A. Stocker, Aromatic-dependent Salmonella typhimurium are non-virulent
751 and effective as live vaccines. *Nature* **291**, 238-239 (1981).
- 752 56. L. Maier *et al.*, Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**,
753 623-628 (2018).
- 754 57. F. R. Cianfanelli, O. Cunrath, D. Bumann, Efficient dual-negative selection for bacterial genome
755 editing. *BMC Microbiol* **20**, 129 (2020).
- 756 58. R. Soldan *et al.*, From macro to micro: a combined bioluminescence-fluorescence approach to
757 monitor bacterial localization. *Environ Microbiol* **23**, 2070-2085 (2021).
- 758 59. A. M. Eren *et al.*, Community-led, integrated, reproducible multi-omics with anvi'o. *Nat*
759 *Microbiol* **6**, 3-6 (2021).
- 760 60. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display
761 and annotation. *Nucleic Acids Res* **49**, W293-W296 (2021).
- 762 61. O. Cunrath, D. Bumann, Host resistance factor SLC11A1 restricts Salmonella growth through
763 magnesium deprivation. *Science* **366**, 995-999 (2019).
- 764 62. S. Hapfelmeier *et al.*, The Salmonella pathogenicity island (SPI)-2 and SPI-1 type III secretion
765 systems allow Salmonella serovar typhimurium to trigger colitis via MyD88-dependent and
766 MyD88-independent mechanisms. *J Immunol* **174**, 1675-1685 (2005).
- 767 63. N. L. Sternberg, R. Maurer, Bacteriophage-mediated generalized transduction in Escherichia
768 coli and Salmonella typhimurium. *Methods Enzymol* **204**, 18-43 (1991).
- 769 64. A. R. Wattam *et al.*, Improvements to PATRIC, the all-bacterial Bioinformatics Database and
770 Analysis Resource Center. *Nucleic Acids Res* **45**, D535-d542 (2017).
- 771 65. J. J. Davis *et al.*, PATtyFams: Protein Families for the Microbial Genomes in the PATRIC
772 Database. *Front Microbiol* **7**, 118 (2016).
- 773 66. R Development Core Team, R. f. S. Computing, Ed. (Vienna, Austria, 2021).
- 774 67. B. Periaswamy *et al.*, Live attenuated S. Typhimurium vaccine with improved safety in
775 immuno-compromised mice. *PLoS One* **7**, e45433 (2012).
- 776 68. J. S. Johnson *et al.*, Evaluation of 16S rRNA gene sequencing for species and strain-level
777 microbiome analysis. *Nat Commun* **10**, 5029 (2019).
- 778 69. T. Matsuki *et al.*, Development of 16S rRNA-gene-targeted group-specific primers for the
779 detection and identification of predominant bacteria in human feces. *Appl Environ Microbiol*
780 **68**, 5445-5451 (2002).

781

782

783

784 **Acknowledgments**

785 We are indebted to members of the Foster lab for discussion, and to Emma Slack, Katharine Coyte,
786 Anna Weiss and Wolf-Dietrich Hardt for feedback on the manuscript. We thank Fiona Powrie and the
787 Oxford Centre for Microbiome Studies for germ-free mice and gnotobiotic mouse work. The
788 Heidelberg strains were a gift from Nassos Typas, EMBL Heidelberg. The AMR *E. coli* strain was
789 from Nottingham University Hospitals Pathogen Bank (<https://www.nuh.nhs.uk/pathogen-industry/>).

790 **Funding:** FS was supported by a BBSRC Studentship. EB was supported by SNSF postdoc mobility
791 fellowships (P2EZIP3_199916 and P500PB_210941). MTJ was supported by the Human Frontier
792 Science Program (LT000798/2020). This work was supported by Wellcome Trust Investigator award
793 209397/Z/17/Z and by European Research Council Grant 787932 to KRF. **Author contributions:**
794 Conceptualisation: KRF. Methodology: FS, EB, MTJ, EBNA, CFP, XW, LP, OC. Investigation: FS,
795 EB, MTJ, OC. Visualisation: FS, EB, MTJ. Funding acquisition: KRF. Supervision: EB, OC, KRF.
796 Writing – original draft: FS, EB, KRF. Writing – review and editing: FS, EB, MTJ, CFP, XW, LP,
797 OC and KRF. **Competing interests:** None. **Data and materials availability:** All data will be made
798 available in the main text or supplementary materials upon publication.

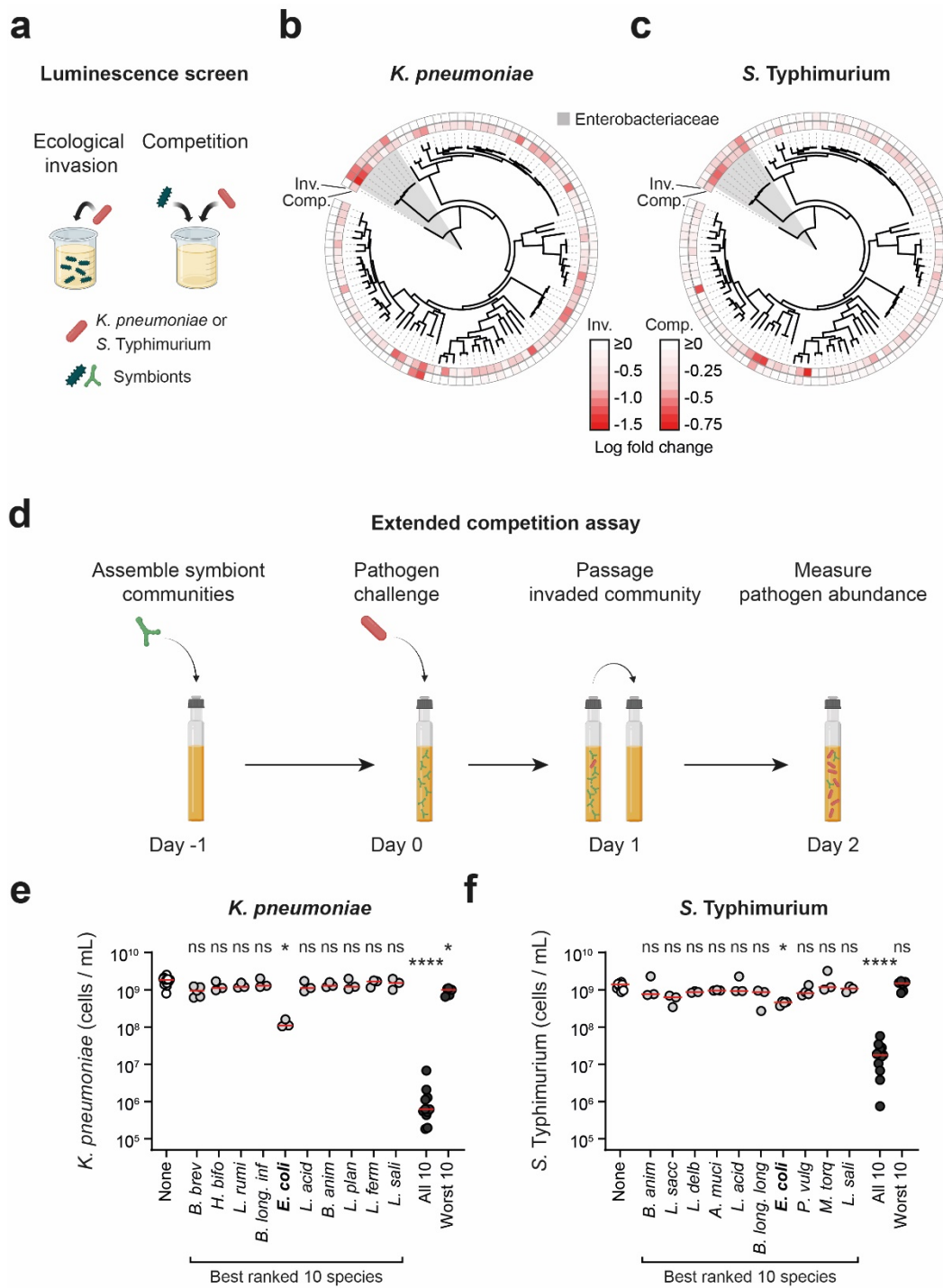
799 **License information:** Copyright © 20XX the authors, some rights reserved; exclusive licensee
800 American Association for the Advancement of Science. No claim to original US government works.
801 <https://www.science.org/about/science-licenses-journal-article-reuse>. This research was funded in
802 whole or in part by a Wellcome Trust Investigator award 209397/Z/17/Z, a cOAlition S organization.
803 The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public
804 copyright license.

805 **Supplementary Materials**

806 Figs. S1-S13

807 Tables S1-S7

808



810

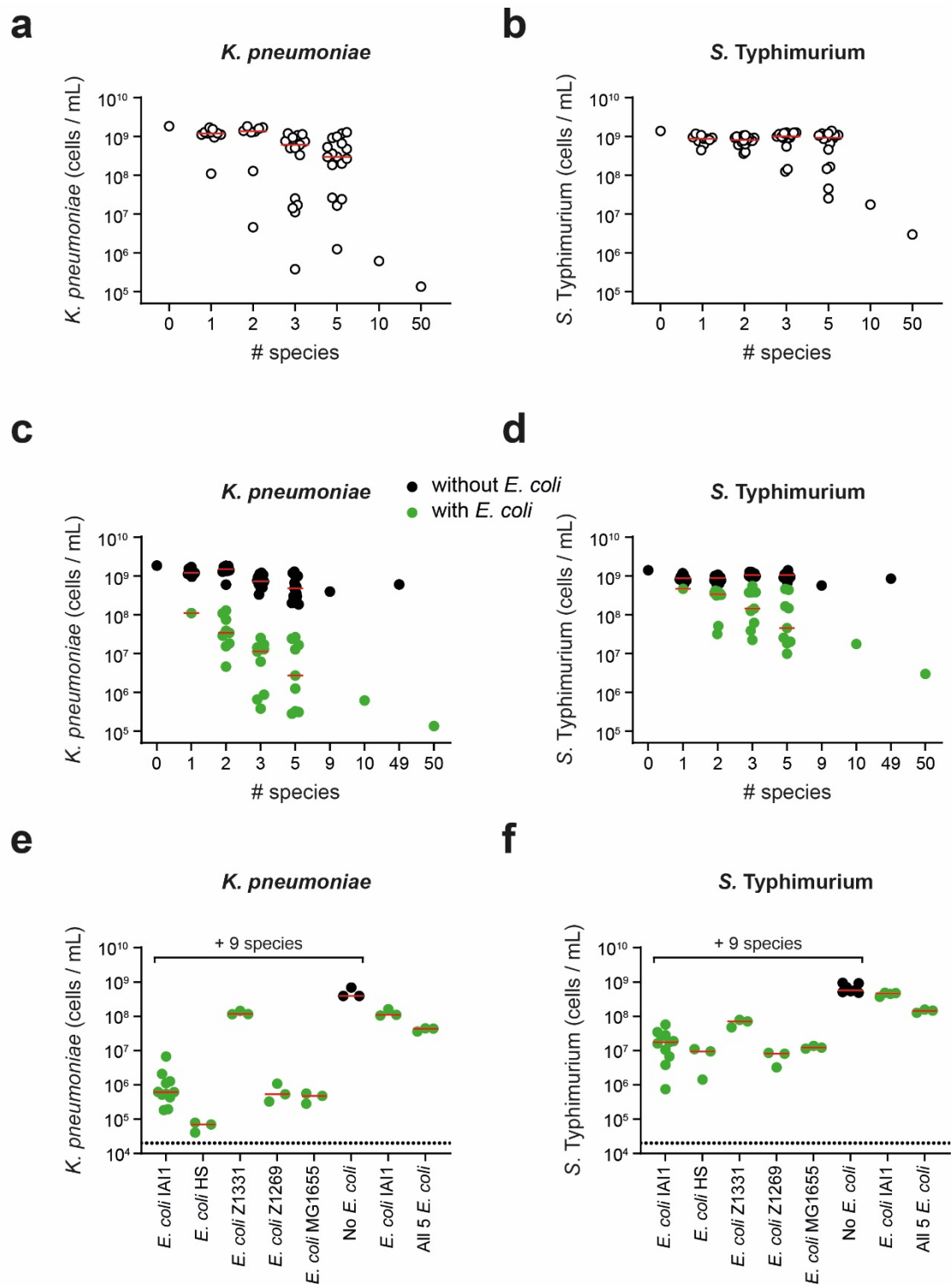
811 **Figure 1. Single strains do not provide robust colonisation resistance, but a diverse community**

812 **can, depending on its composition. a)** Overview of the luminescence co-culture assays. In the

813 ecological invasion assay, *K. pneumoniae* or *S. Typhimurium* (red) was inoculated in co-culture with

814 individual symbionts (different green symbols are used to represent the diversity of symbiont species

815 screened; 19:1 ratio of symbiont to pathogen). In the competition assay, the symbionts were
816 inoculated at an equal ratio to the pathogen to recapitulate competition between strains once the
817 pathogen is established. In both assays, luminescence produced by the pathogen was used as a proxy
818 for pathogen growth. Created with BioRender.com. **b-c)** Comparison of phylogenetic relatedness
819 between symbionts, and the ability of each symbiont to compete with the pathogen (inv=ecological
820 invasion assay; comp=competition assay). Data for *K. pneumoniae* shown in **(b)** and *S. Typhimurium*
821 shown in **(c)**. The family Enterobacteriaceae, which includes both *K. pneumoniae* and *S.*
822 *Typhimurium*, is shaded in grey. Luminescence fold change values are presented in **Fig. S1**. Data
823 presented as the median luminescence log fold change of n=3-10 independent experiments. Strains
824 with the most negative (most red) values inhibited growth of the pathogen most strongly. **d)** Overview
825 of the extended competition assay. Communities (or individual strains; green) of symbionts are pre-
826 grown in anaerobic rich media before addition of the pathogen (red). The community is passaged after
827 24h of growth, followed by another 24h of growth before quantification with flow cytometry. Created
828 with BioRender.com. **e-f)** The extended competition assay was performed for each individual species
829 identified in the best ranked 10 species, as well as for combinations of 10 species (both the best- and
830 worst-ranked 10 species; **Fig. S1**). Individual data points from n=3-15 independent experiments are
831 shown. Red lines indicate the median. A Kruskal-Wallis test with Dunn's multiple test correction
832 compares each group to the no-symbionts control ($p > 0.05 = \text{ns}$; $p < 0.05 = *$; $p < 0.0001 = ****$). Data for
833 *K. pneumoniae* shown in **(e)** and *S. Typhimurium* shown in **(f)**. See **Table S1** for species name
834 abbreviations.



835

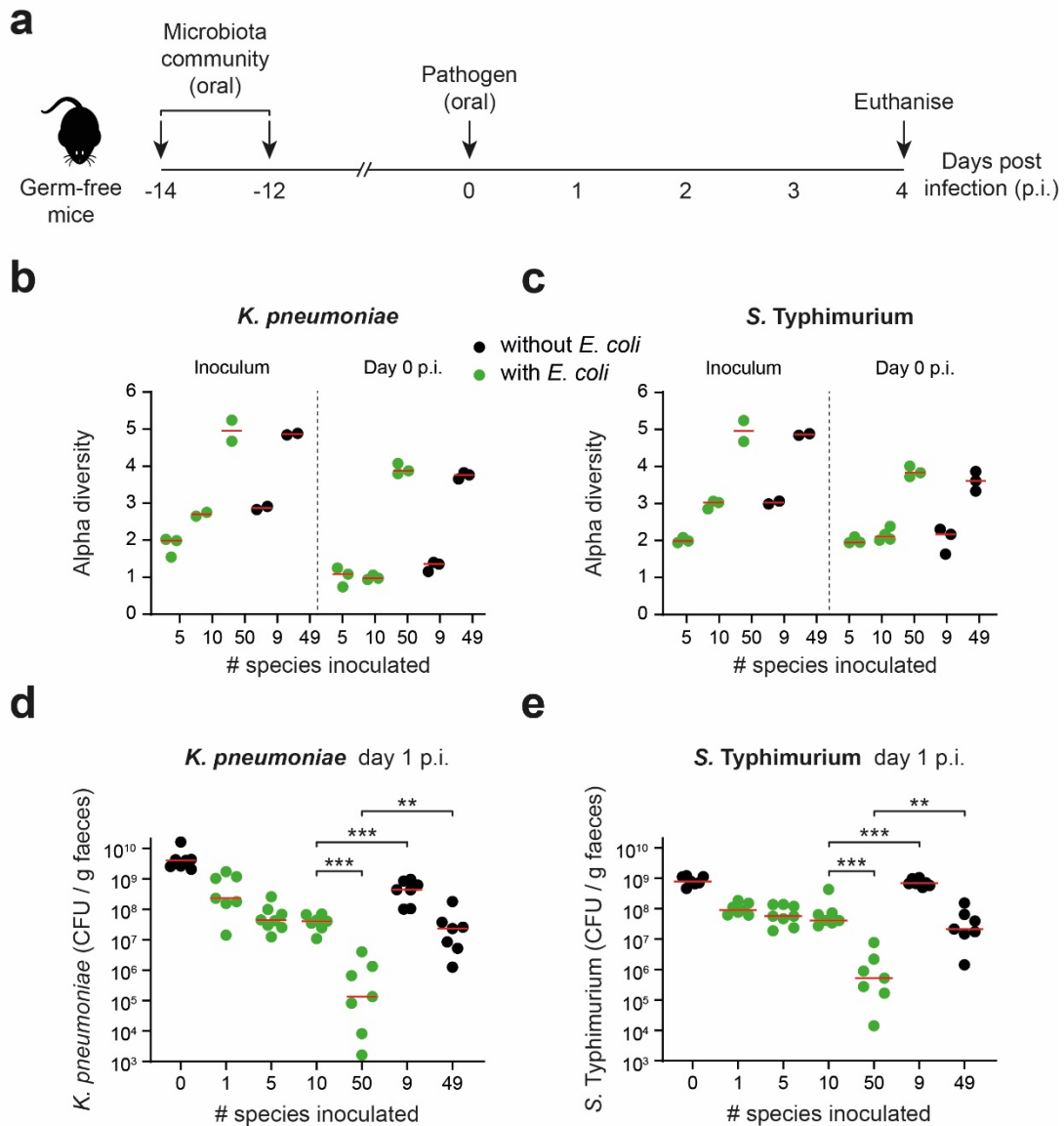
836 **Figure 2. Ecological diversity and key members are needed for efficient colonisation resistance**

837 *in vitro*. **a-d)** Extended competition assay on communities made up of an increasing number of

838 species. Each data point represents the median pathogen cells/mL value on day 2 of the extended

839 competition for a community (n=3-15 independent experiments for each community; up to 17

840 communities for each group). Communities with size ≤ 10 species were randomly selected from the
841 10 best ranked species for each pathogen. Community identities are shown in **Table S4-5**. Data for *K.*
842 *pneumoniae* shown in **(a)** and **(c)**, and for *S. Typhimurium* shown in **(b)** and **(d)**. Red lines indicate
843 the median value of communities at a given diversity level. In **c-d**), data from **a-b**) are replotted along
844 with additional communities that always contained *E. coli* but were otherwise randomly selected.
845 Communities without *E. coli* are depicted in black; communities with *E. coli* in green. Separate red
846 median lines shown for communities with and without *E. coli*. A linear regression is performed on
847 log-log transformed data in **Fig. S2a-b**, which shows that the association between diversity and
848 colonisation observed is statistically significant, and that this effect is greater for communities with *E.*
849 *coli* than those without (F tests, $p \leq 0.0001$). **e-f**) Extended competition assay testing *E. coli* strains
850 substituted into the best ranked 10 species community. Data for *K. pneumoniae* shown **(e)** and *S.*
851 *Typhimurium* in **(f)**. Red lines indicate median values.



852

853 **Figure 3. Ecological diversity and key members are needed for efficient colonisation resistance**

854 *in vivo*. **a)** Overview of gnotobiotic mouse experiments. Symbiont communities (or *E. coli* alone)

855 were given to germ-free mice by oral gavage twice (two days apart). 12 days later, the mice were

856 challenged with *K. pneumoniae* or *S. Typhimurium* by oral gavage. Faeces were collected from mice

857 daily before being euthanised on day 4 post infection (p.i.). **b-c)** Alpha diversity measured by

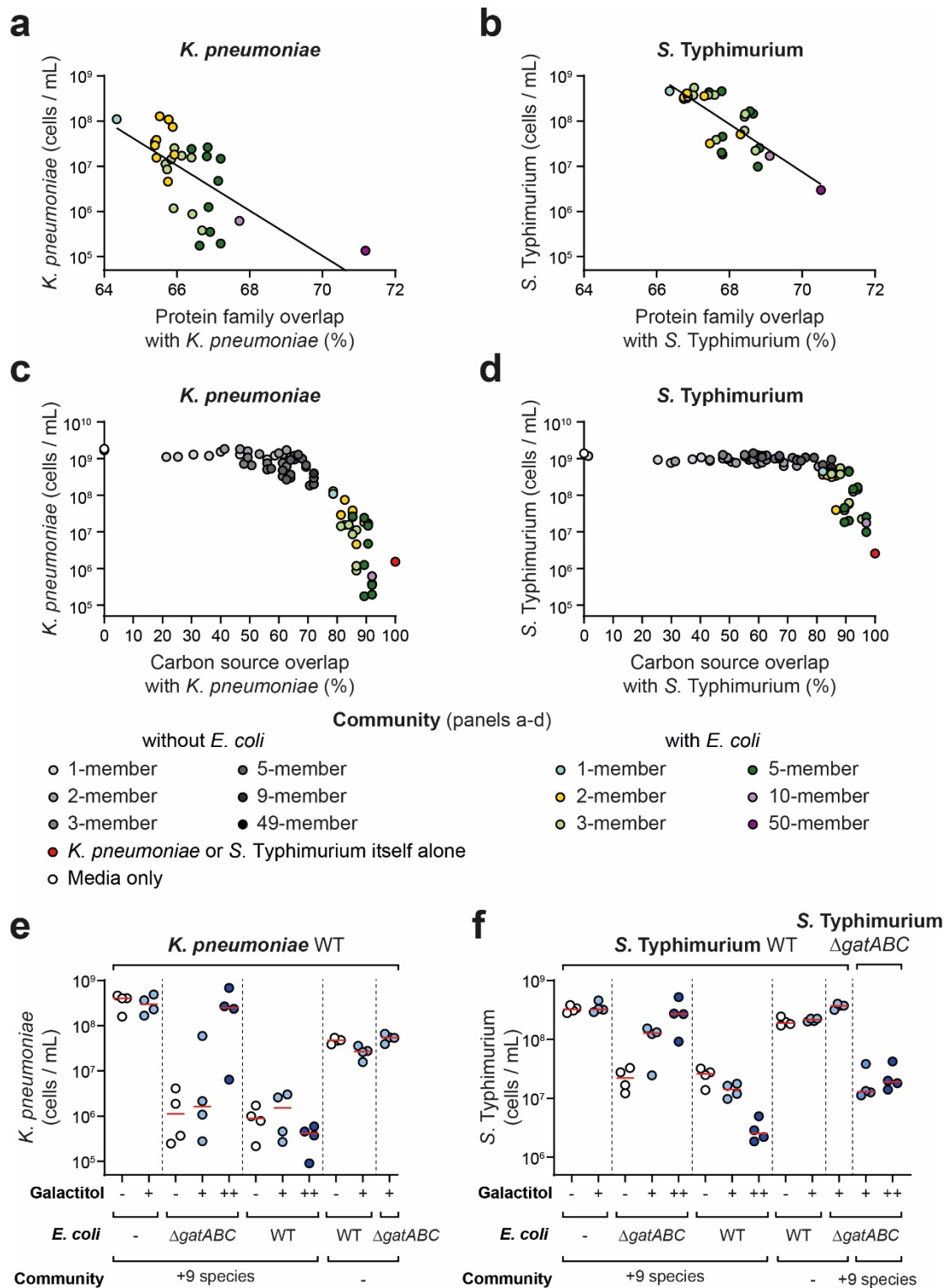
858 Shannon index of symbiont communities. Metagenomic sequencing was performed on the inoculum

859 and faecal samples at day 0 (when the pathogen is introduced) and used to calculate diversity. Data

860 for *K. pneumoniae* shown in **(b)** and *S. Typhimurium* in **(c)**. **d-e)** Pathogen abundances in the faeces

861 of gnotobiotic mice colonised with communities of increasing diversity (mice containing communities

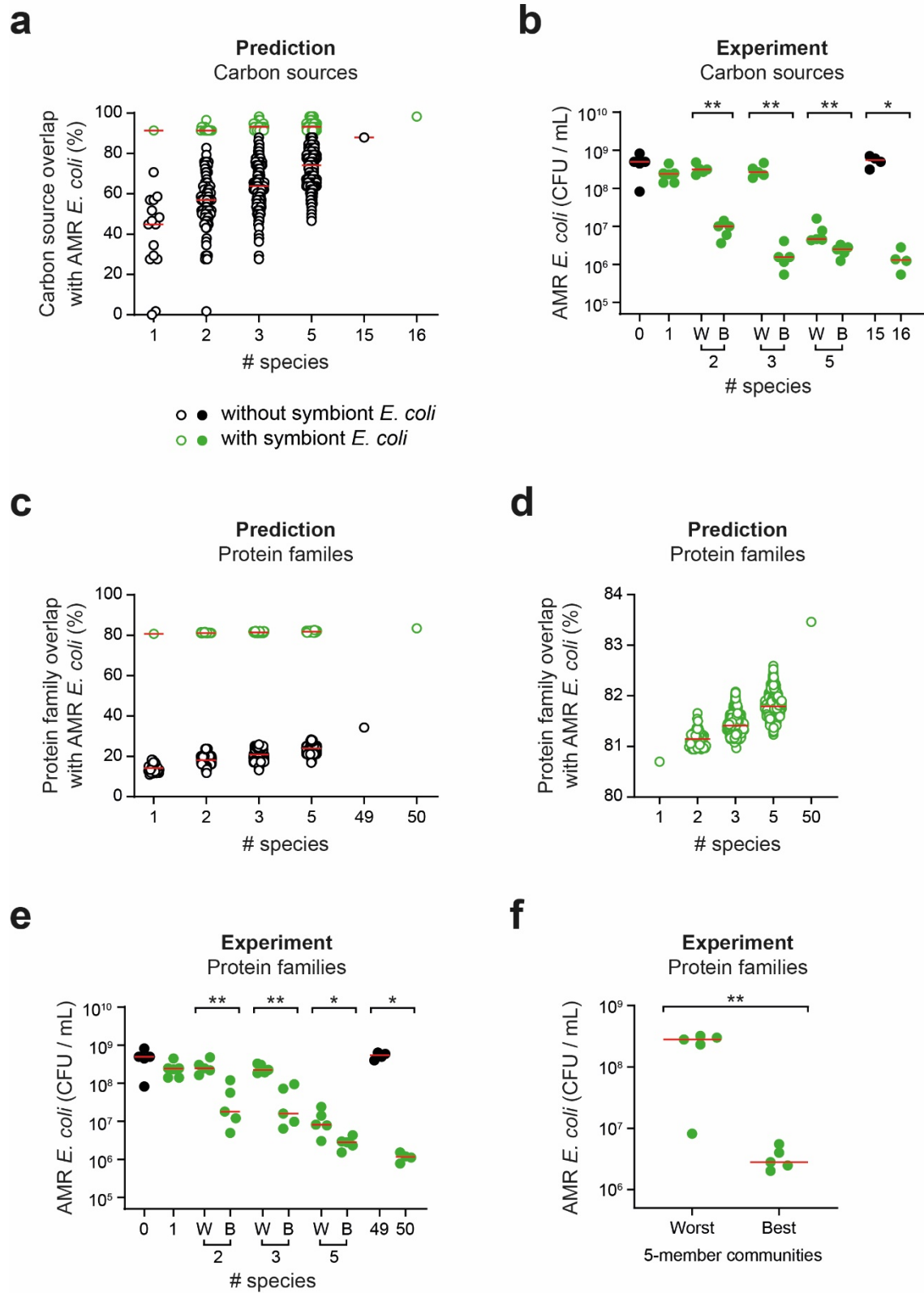
862 with *E. coli* shown in green; mice containing communities without *E. coli* shown in black; n=7-8 mice
863 per group in cages of 2-3 mice; 2-3 independent experiments). Red lines indicate the medians. Two-
864 tailed Mann-Whitney tests are used to compare the indicated groups ($p < 0.01 = **$; $p < 0.001 = ***$). Data
865 for *K. pneumoniae* shown in **(d)** and *S. Typhimurium* shown in **(e)**. Metagenomic analysis of strain
866 diversity and relative abundance is shown in **Fig. S5**. Pathogen abundance data from days 1-4 p.i. is
867 shown in **Fig. S6**. Community compositions are shown in **Table S6**.



868

869 **Figure 4. Nutrient overlap can explain the role of ecological diversity and the effect of *E. coli* in**
 870 **colonisation resistance. a-b)** Protein family overlap is compared to the median pathogen abundance
 871 values for each community containing *E. coli* from **Fig. 2c-d**. Diversity is visualised by a colour
 872 gradient. Data for *K. pneumoniae* shown in **(a)** and *Salmonella* shown in **(b)**. A line of best fit is

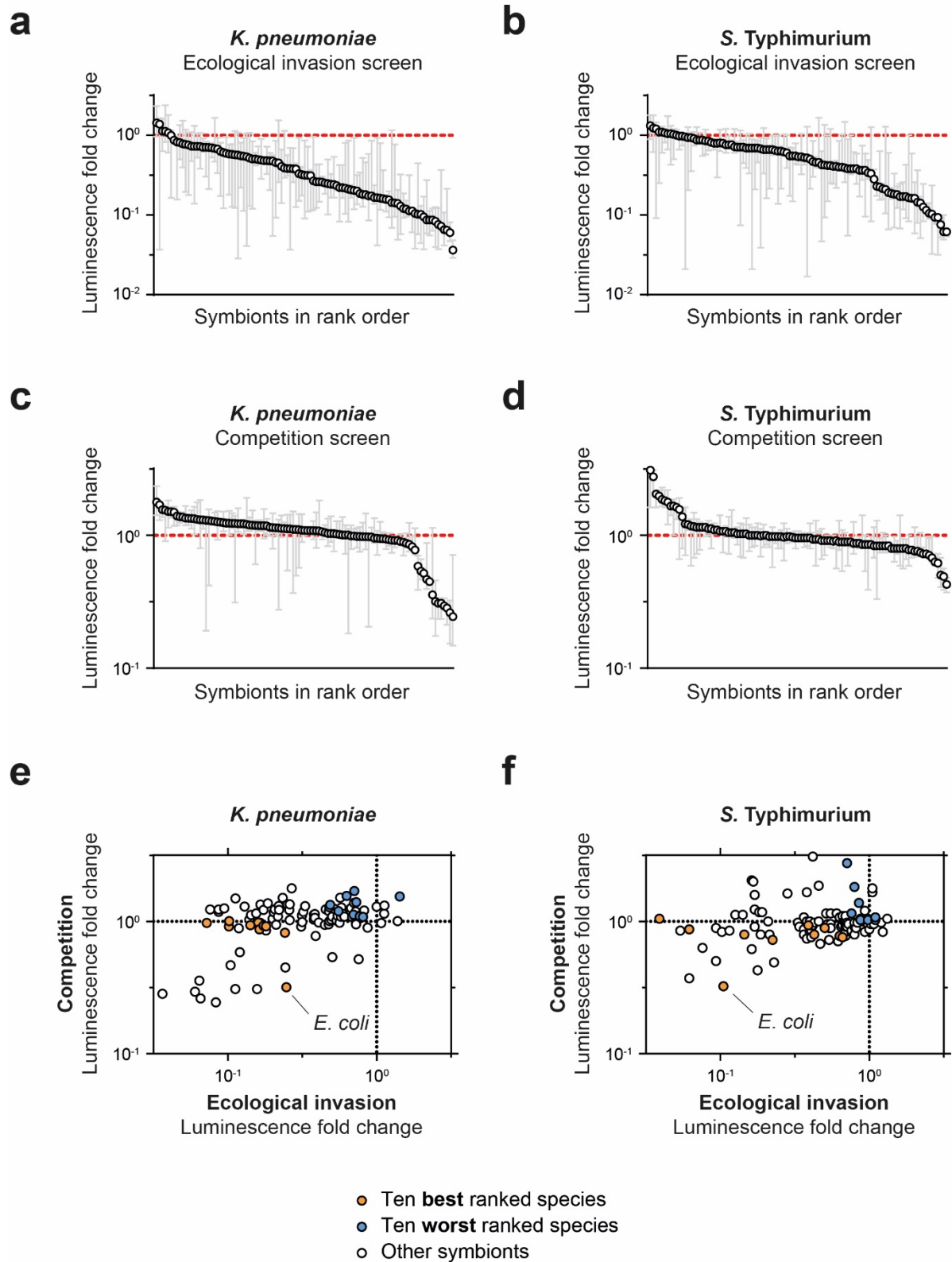
873 shown from a linear regression on log transformed data: $R^2 = 0.4255$ for *K. pneumoniae*; $R^2 = 0.603$
874 for *S. Typhimurium*; both slopes are significantly different from 0 using an F-test ($p < 0.0001$). Data
875 for communities without *E. coli* is presented in **Fig. S9e-f. c-d**. Overlap in carbon source utilisation
876 plotted against the median pathogen abundance measurements from experimental communities in **Fig.**
877 **2c-d**. Community carbon source overlap is calculated using an additive approach from carbon source
878 overlap of individual strains by measurement on AN Biolog Microplates (**Fig. S10**). Diversity is
879 visualised by a gradient of colour (for *E. coli*-containing communities) or greyscale (for communities
880 without *E. coli*). A control with the isogenic pathogen itself (100% overlap) is plotted in red. Data for
881 *K. pneumoniae* in **(c)** and *S. Typhimurium* shown in **(d). e-f)** A private nutrient, galactitol, that could
882 only be used by the WT *E. coli* strain and the pathogens but not by the other symbionts nor an *E. coli*
883 $\Delta gatABC$ mutant, was supplemented to the media and the extended competition assay performed as
884 before. In all treatments, pathogen abundance was measured by flow cytometry after 48h of growth
885 post-passage instead of the usual 24h. This change did not influence the control experiments without
886 galactitol, but proved informative because we found the growth impacts of galactitol were relatively
887 slow. Results for *K. pneumoniae* shown in **(e)** and *S. Typhimurium* in **(f)**. N=3-4 replicates per
888 treatment. Horizontal red lines show the median of the replicates. Light blue circles show results with
889 0.1% galactitol supplementation (+ symbol), dark blue circles show results with 1% galactitol
890 supplementation (++) symbol). White circles (control) show results with no nutrient supplementation
891 (- symbol). 9 species refers to the 9 additional species in the 10 best-performing species for each
892 respective pathogen (- symbol refers to when *E. coli* is added alone). In **(d)**, a $\Delta gatABC$ mutant of *S.*
893 *Typhimurium* was used in addition to the WT pathogen to verify the dependency of colonisation on a
894 private nutrient.



895

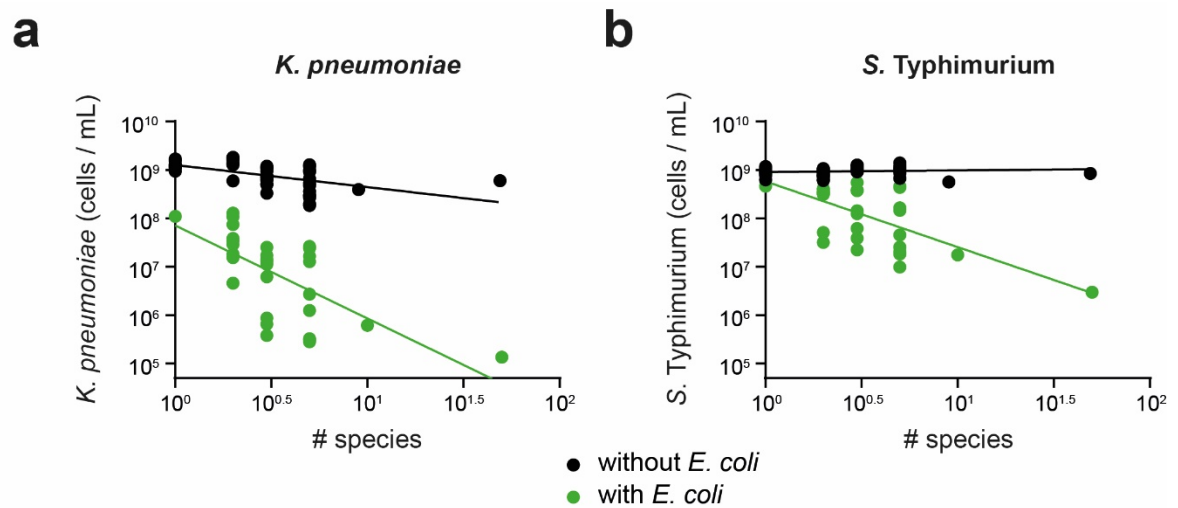
896 **Figure 5. Nutrient blocking predicts community colonisation resistance.** a) *In silico* prediction of
 897 carbon source overlap with the AMR *E. coli* strain for all possible combinations of symbiont
 898 communities at the indicated diversity levels. Each circle represents a different community.

899 Communities containing the symbiont *E. coli* IAI1 are shown in green and communities without *E.*
900 *coli* IAI1 are shown in black (predictions as hollow circles; experimental data as solid circles).
901 Predicted carbon source overlap calculated using an additive approach from carbon source use of
902 individual strains measured using AN Biolog Microplates (**Fig. S10, S13**). **b**) Experimental test of *in*
903 *silico* predictions in **(a)**. The two *E. coli* IAI1-containing communities predicted to have the best (B)
904 and worst (W) carbon source overlap were picked at each diversity level and competed against AMR
905 *E. coli* in the extended competition assay. A two-tailed Mann-Whitney U test was performed on
906 community pairs ($p > 0.05 = \text{ns}$; $p < 0.05 = *$; $p < 0.01 = **$) at the 2-, 3- and 5-strain diversity levels. Red
907 horizontal bars depict the median of each community tested. $N = 4-5$ replicates for each community. **c-**
908 **d**) *In silico* prediction of protein family overlap with the AMR *E. coli* strain for a random subset
909 ($n = 59,043$) of all possible symbiont communities at diversity levels 2-, 3-, and 5-strains, as well as all
910 individual strains and the 49- and 50-species communities. Each circle represents a different
911 community. Communities are selected from the strains comprising the 50-member community.
912 Communities containing *E. coli* IAI1 are shown in green and communities without *E. coli* IAI1 are
913 shown in black. **d**) Only the *E. coli*-containing communities are plotted to better visualise variation in
914 protein family overlap. **e**) Experimental test of *in silico* predictions based on protein cluster overlap in
915 **(c-d)**. The two *E. coli* IAI1-containing communities predicted to have the best (B) and worst (W)
916 protein family overlap were picked at each diversity level (randomly selected, for cases where there
917 were multiple communities with the same overlap), and competed against AMR *E. coli* in the
918 extended competition assay. Red horizontal bars depict the median of each community tested. $N = 5$
919 replicates for each community. A two-tailed Mann-Whitney U test was performed on community
920 pairs ($p > 0.05 = \text{ns}$; $p < 0.05 = *$; $p < 0.01 = **$) at the 2-, 3- and 5-strain diversity levels. **f**) Experimental
921 test of the predicted 5 best and 5 worst communities at the 5-strain diversity level, based on protein
922 family overlap with AMR *E. coli*. Each symbol represents the median of $N = 5$ replicates per
923 community. Red horizontal bars depict the median of the best and the worst predicted communities. A
924 two-tailed Mann-Whitney U test was performed ($p < 0.01 = **$). Community identities for **(b, e-f)** are
925 shown in **Table S7**.



928 **Figure S1. Human gut symbiont strains vary in their ability to inhibit growth of *K. pneumoniae***
 929 **and *S. Typhimurium* in the luminescence screen. a-d)** Waterfall plots show the luminescence median
 930 fold change (log ratio of the control luminescence divided by the treatment luminescence) of the

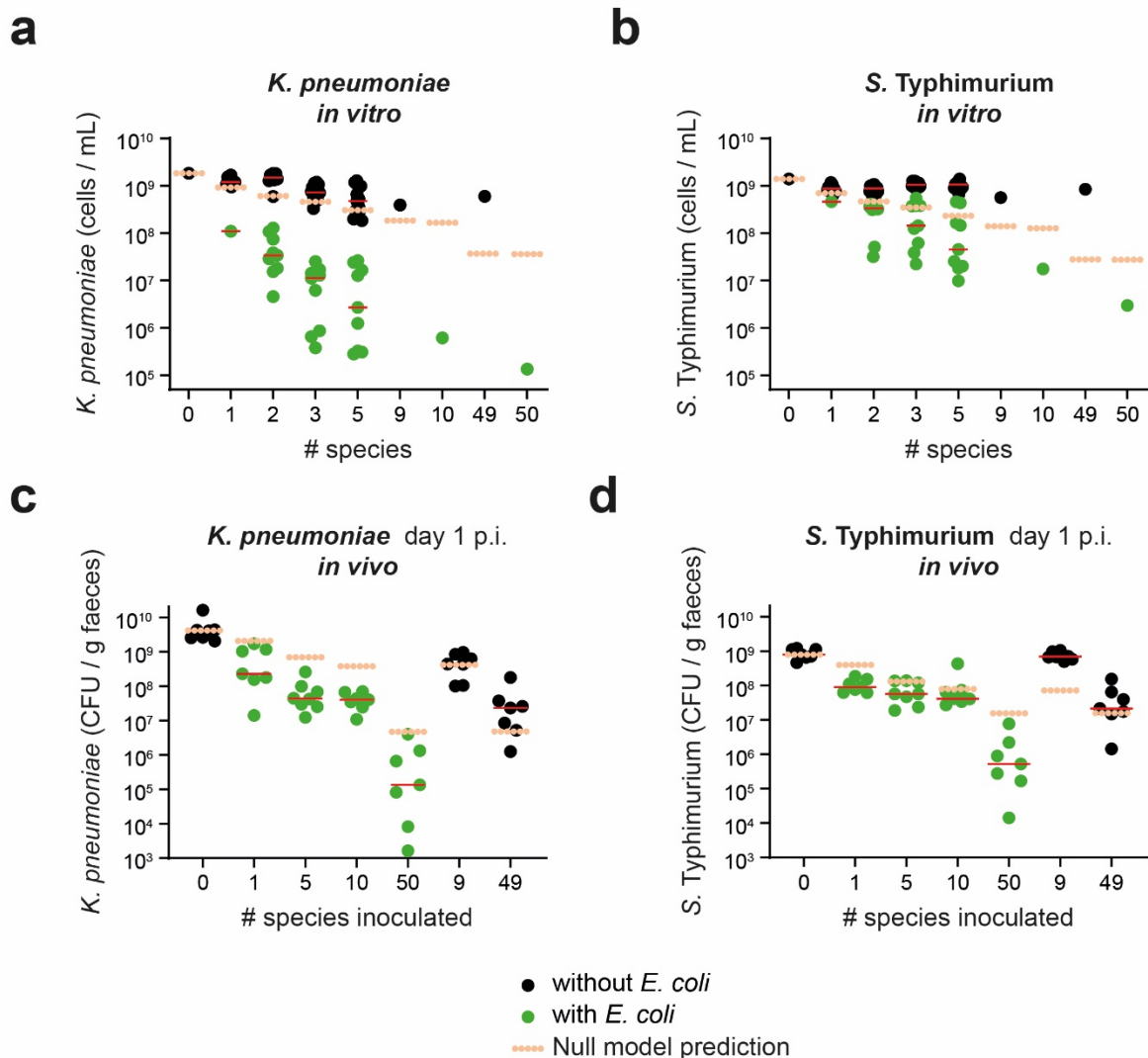
931 pathogen in combination with each human gut symbiont tested in the luminescence screen. Results for
932 *K. pneumoniae* are shown in **(a)** and **(c)** and for *S. Typhimurium* shown in **(b)** and **(d)**. Results of the
933 ecological invasion assay are shown in **(a)** and **(b)** and the competition assay in **(c)** and **(d)**. Black circles
934 represent the median value for each symbiont (N=3-10 replicates). Grey vertical lines represent range
935 bars. **e-f**) Correlation of the results of the invasion and competition assays of the luminescence screen.
936 Each circle represents the median value for a symbiont (N=3-10 replicates). Results for *K. pneumoniae*
937 shown in **(e)** and *S. Typhimurium* in **(f)**. The sum of the ranks of the competition and the invasion assays
938 were used to create an overall ranking of symbionts for each pathogen. The best 10 ranked non-
939 pathogenic species are shown in orange and the worst 10 shown in blue. Strains with the most negative
940 competition and invasion values inhibited the growth of the pathogen most strongly compared to the
941 media-only control.



942

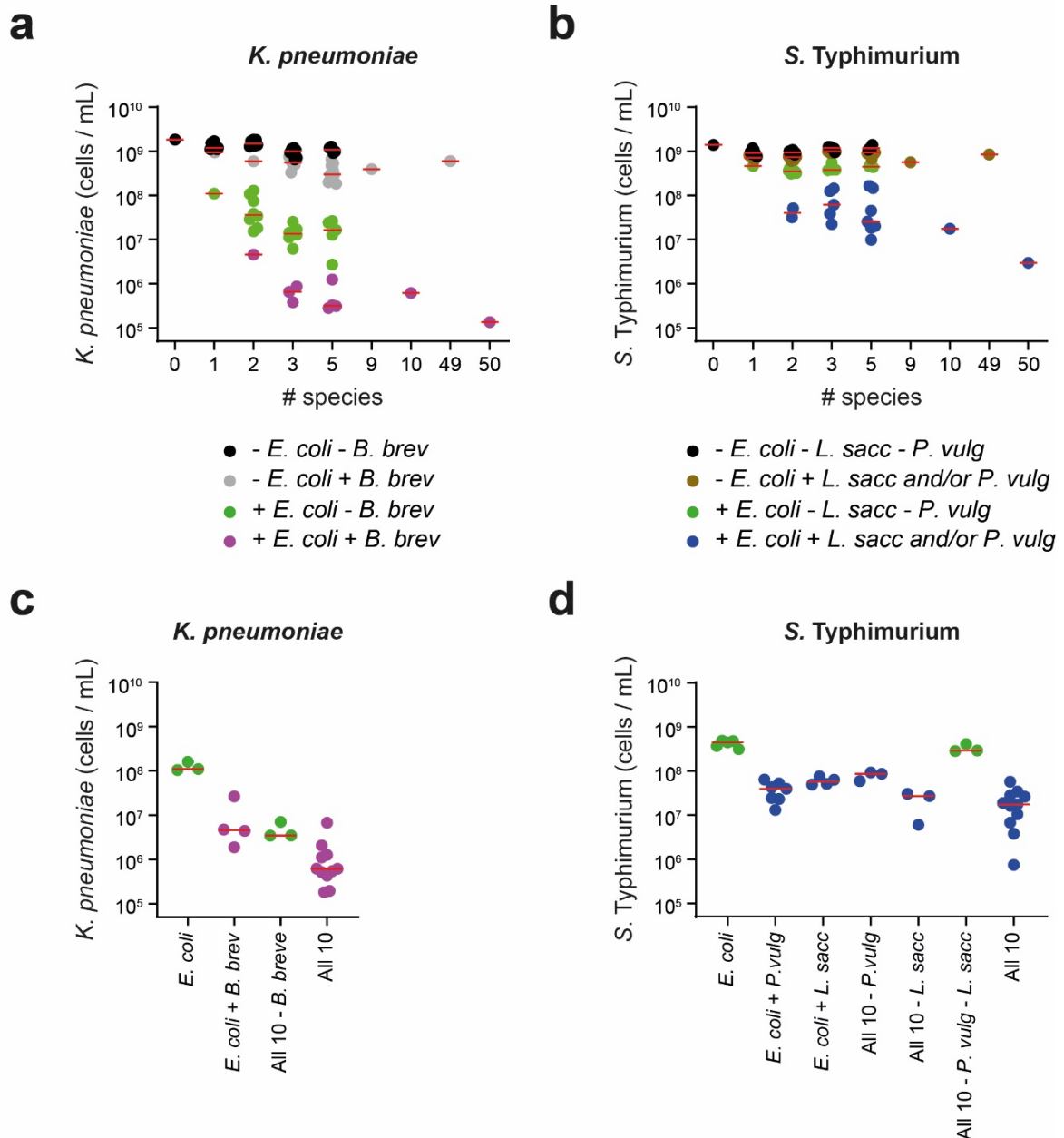
943 **Figure S2. Community diversity negatively correlates with pathogen abundance. a-b)** As the
 944 number of symbiont strains increases, pathogen density on day 2 of the extended competition decreases
 945 (x axis on a log scale). Data for *K. pneumoniae* shown in **(a)** and *S. Typhimurium* in **(b)**. Each circle
 946 represents the median value of a community tested in the extended competition assay (data from **Fig.**
 947 **2c-d**). Communities in green contain *E. coli*, communities in black do not contain *E. coli*. Linear
 948 regression of log-transformed data: **(a)** $R^2=0.4296$, non-zero slope for *E. coli* communities (F test,
 949 $p<0.0001$). $R^2=0.3103$, non-zero slope for communities without *E. coli* (F test, $p<0.0001$). Moreover,
 950 slopes of the two regressions are significantly different from each other (F test, $p=0.0001$). **(b)**
 951 $R^2=0.4234$, non-zero slope for *E. coli* communities (F test, $p<0.0001$). $R^2=0.01378$, slope not different
 952 to zero for communities without *E. coli* (F test, $p=0.4218$). Slopes of the two regressions are again
 953 significantly different from each other (F test, $p<0.0001$).

954



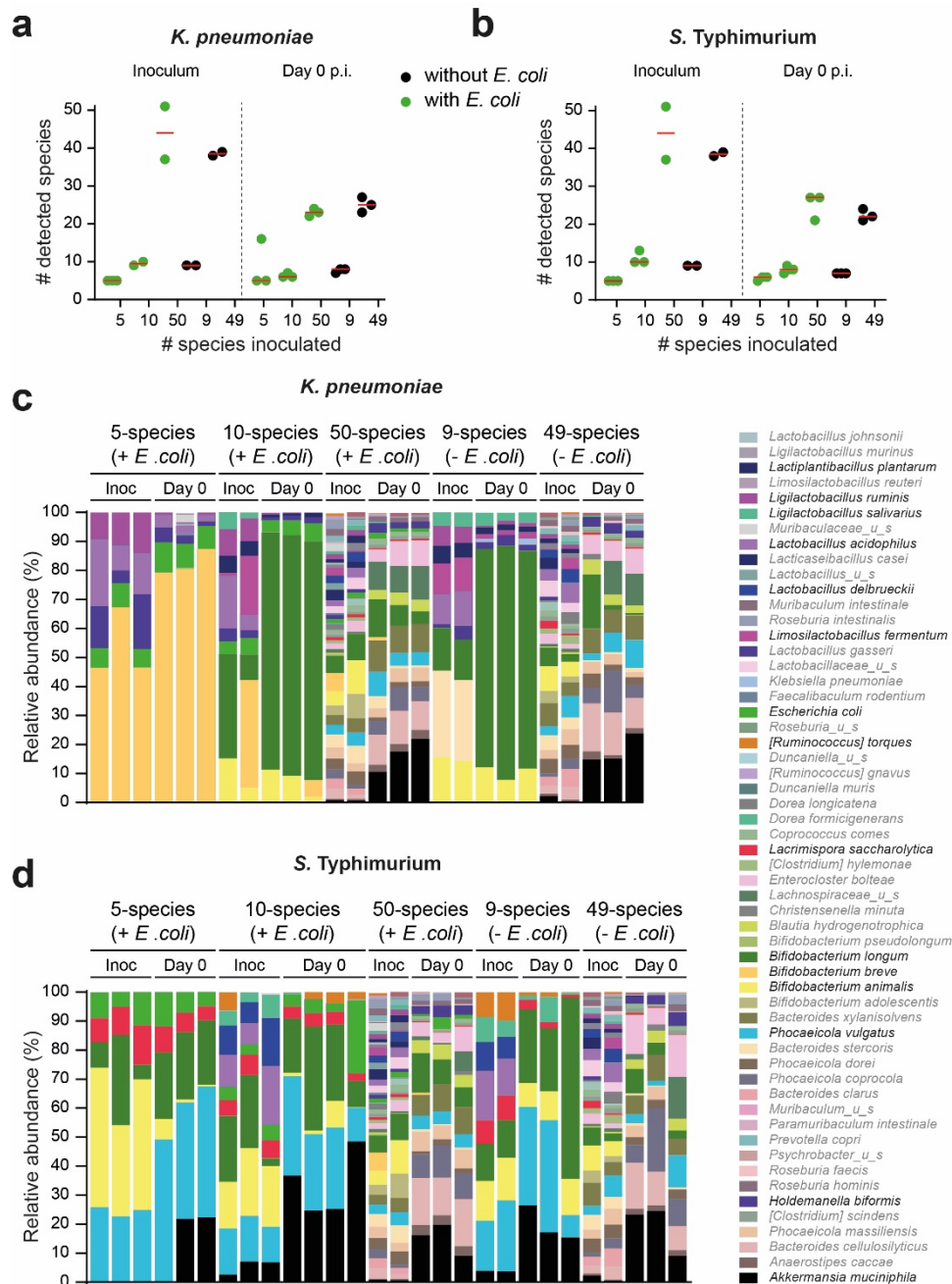
955
 956 **Figure S3. Experimental data deviates from a null model where the effect of each species is**
 957 **additive and restricts pathogen growth by affecting carrying capacity.** We compared our
 958 experimental data to a null model where the effect of colonisation resistance simply scales according to
 959 the number of species in the community. In this model, the expectation is that the effect of adding
 960 species will proportionally reduce the density of the pathogen simply by taking up a larger proportion
 961 of the carrying capacity. Therefore, we took the abundance of the pathogen at the end of the experiment
 962 when the pathogen is alone (ie, $n=0$) and multiplied it by $1/(n+1)$ where n refers to the number of species
 963 that contribute to the overall carrying capacity of the system. This value is plotted in as beige dotted
 964 lines for both *in vitro* data from Fig. 2c-d (panels a and b; dots indicate median values of each
 965 community) and *in vivo* data from Fig. 3d-e (panels c and d; dots indicate individual mice). a,c)
 966 Data for *K. pneumoniae*, b,d) Data for *S. Typhimurium*. Communities containing *E. coli* are shown in
 967 green and communities without *E. coli* are shown in black. Red lines indicate medians. In all cases, the
 968 ratio of observed results to the null model increased as diversity of the community increased for *E. coli*-
 969 containing community, indicating that our results are not solely driven by a reduction in the available
 970 carrying capacity of the pathogen (Two-tailed Mann-Whitney U tests are used to compare single species
 971 communities to 50-member communities; $p=0.0070$ for panel a; $p=0.0070$ for panel b; $p=0.0041$ for
 972 panel c; $p=0.0262$ for panel d).

973



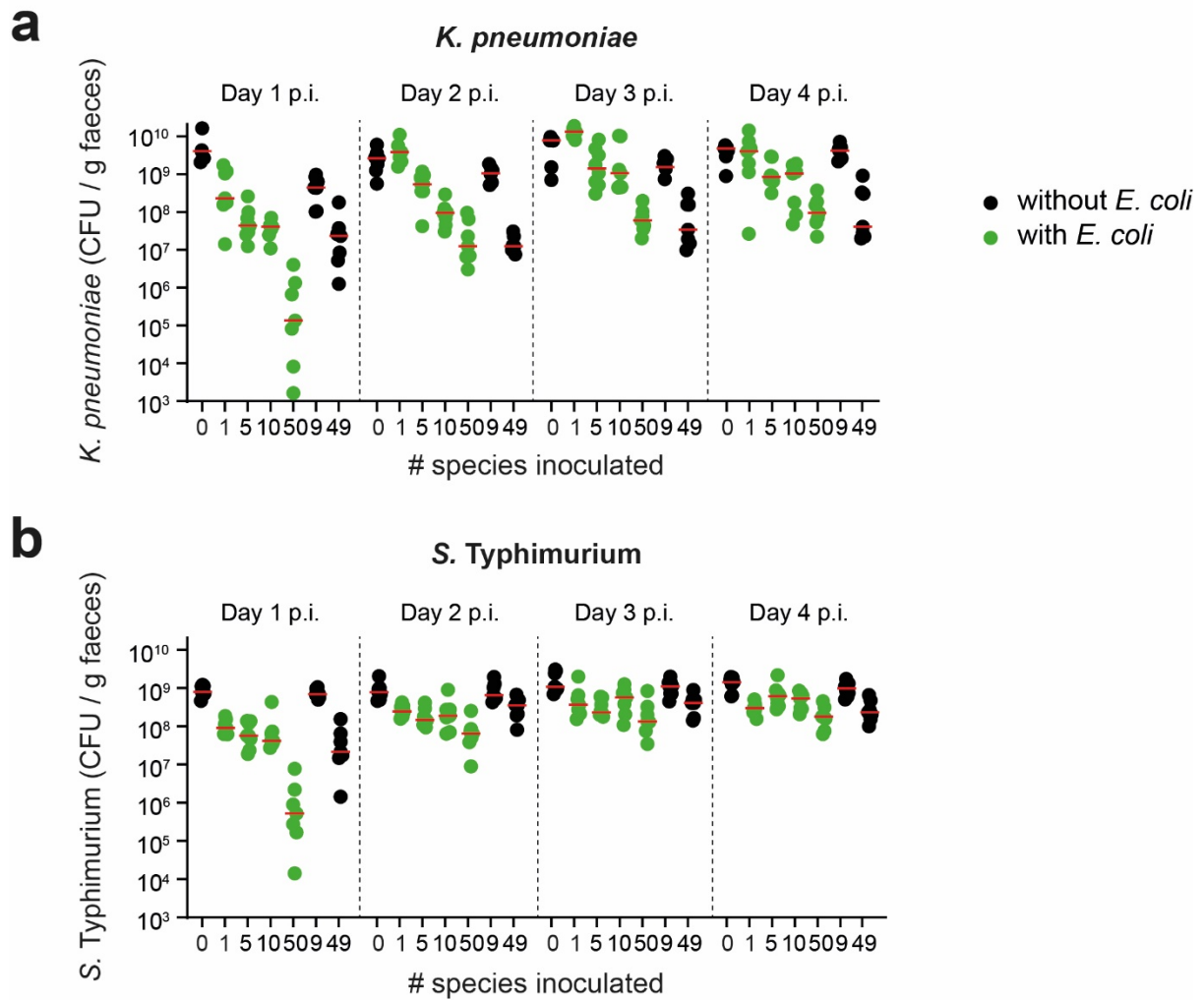
974

975 **Figure S4. Combinations of multiple species are important for colonisation resistance to each**
 976 **pathogen. a-b)** Equivalent figures to **Fig. 2c-d**, except the presence or absence of *B. breve* within
 977 communities are shown in **(a)** (grey for communities with *B. breve* but without *E. coli* and purple for
 978 communities with *B. breve* and *E. coli*), and the presence or absence of *L. saccharolyticum* and/or *P.*
 979 *vulgatus* shown in **(b)** (brown for communities with *L. saccharolyticum* and/or *P. vulgatus* but not *E.*
 980 *coli*, and blue for communities with *L. saccharolyticum* and/or *P. vulgatus* and *E. coli*). Results for *K.*
 981 *pneumoniae* are shown in **(a)** and for *S. Typhimurium* shown in **(b)**. In **(a-b)**, “+” and “-” refer to the
 982 presence or absence of species rather than the addition or subtraction of a species. Horizontal red lines
 983 depict the median of the communities at a particular diversity level containing the species indicated in
 984 the legend. Each circle represents the median pathogen abundance measured for a community on day 2
 985 of the extended competition assay. **c-d)** Drop-out experiments verify the context-dependent effect of
 986 key members on colonisation resistance. In **(c-d)**, “+” and “-” refer to the addition or subtraction of a
 987 species. Results for *K. pneumoniae* are shown in **(c)** and for *S. Typhimurium* shown in **(d)**. Horizontal
 988 red lines depict the median of the replicates for a particular community. See **Table S1** for species name
 989 abbreviations.



990

991 **Figure S5. Metagenomic sequencing shows that germ-free mice gavaged with more diverse**
 992 **communities were colonised with a higher number of bacterial strains. a-b)** Number of detected
 993 bacterial species (above a relative abundance threshold of 0.1%) in the inoculum given to the mice and
 994 in mouse faeces 14 days after the first inoculum gavage (Day 0 post infection; p.i.). Mice were given 2
 995 identical gavages containing symbiont communities 2 days apart; the first inoculum was sequenced as
 996 a representative. Each inoculum data point depicts an independent experiment and each day 0 data point
 997 indicates a representative mouse from each cage. Horizontal red lines represent median values of the
 998 replicates at each diversity level. (N=2-3 for the inoculum, N=3-4 for faecal samples). Green symbols
 999 represent communities that contain *E. coli*, black symbols represent communities without *E. coli*. **c-d)**
 1000 Relative abundance plot of symbiont strains in the inoculum and mouse faeces using metagenomic
 1001 sequencing data. Data for mice challenged with *K. pneumoniae* shown in **(c)** and *S. Typhimurium* in
 1002 **(d)**. The 10 best ranked strains from the luminescence screen for each pathogen are shown in black
 1003 writing (other detected species in grey writing).



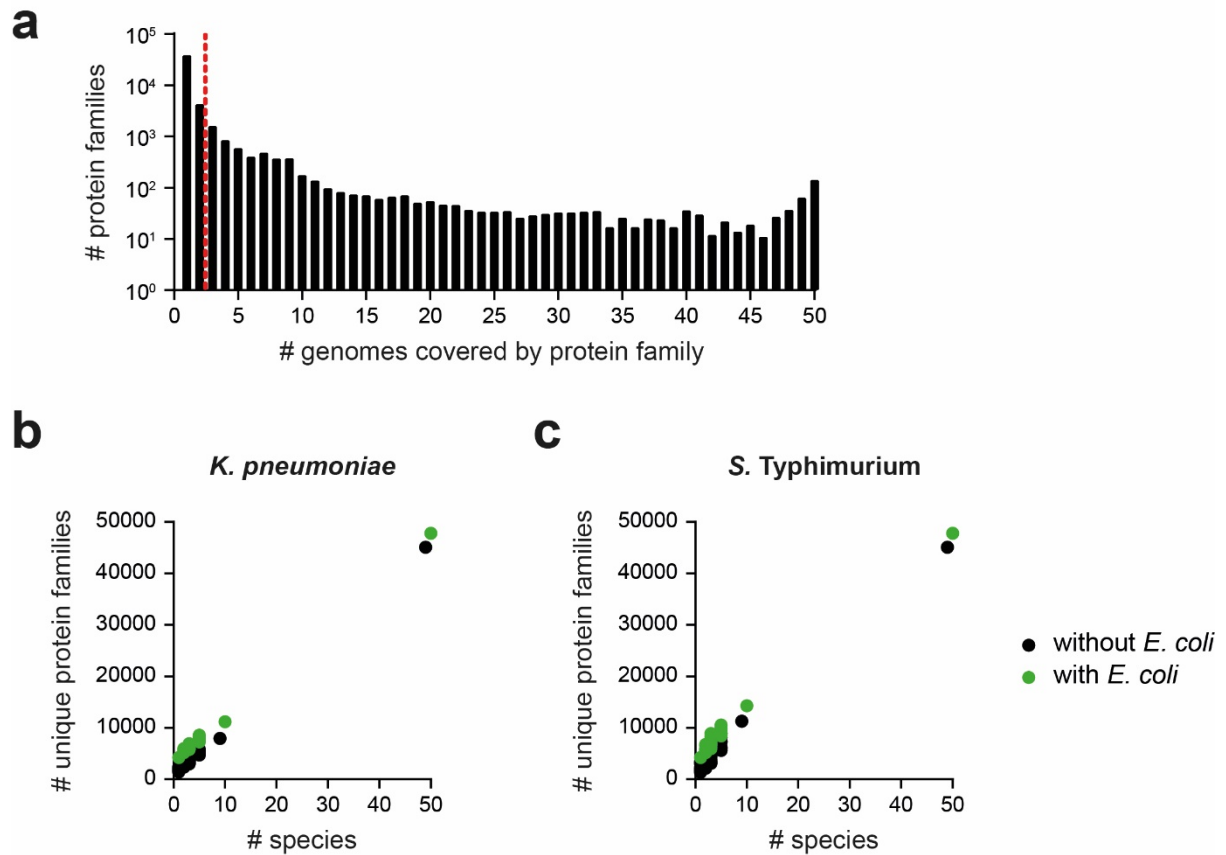
1005

1006

1007 **Figure S6. Pathogen abundance *in vivo* at later timepoints. a-b)** Mice were gavaged with *K.*
 1008 *pneumoniae* (a) or *S. Typhimurium* (b) on Day 0. Each symbol represents a faecal sample from 1 mouse.
 1009 Pathogen abundances were determined by selective plating aerobically on LB agar + carbenicillin (*K.*
 1010 *pneumoniae*) or LB agar + streptomycin (*S. Typhimurium*). Horizontal red lines represent median
 1011 values of the replicate mice tested at each diversity level. Communities containing *E. coli* are shown in
 1012 green whereas communities without *E. coli* are in black. N=7-8 mice per treatment. The day 1 post
 1013 infection (p.i.) data is the same as in Fig. 3d-e.

1014

1015

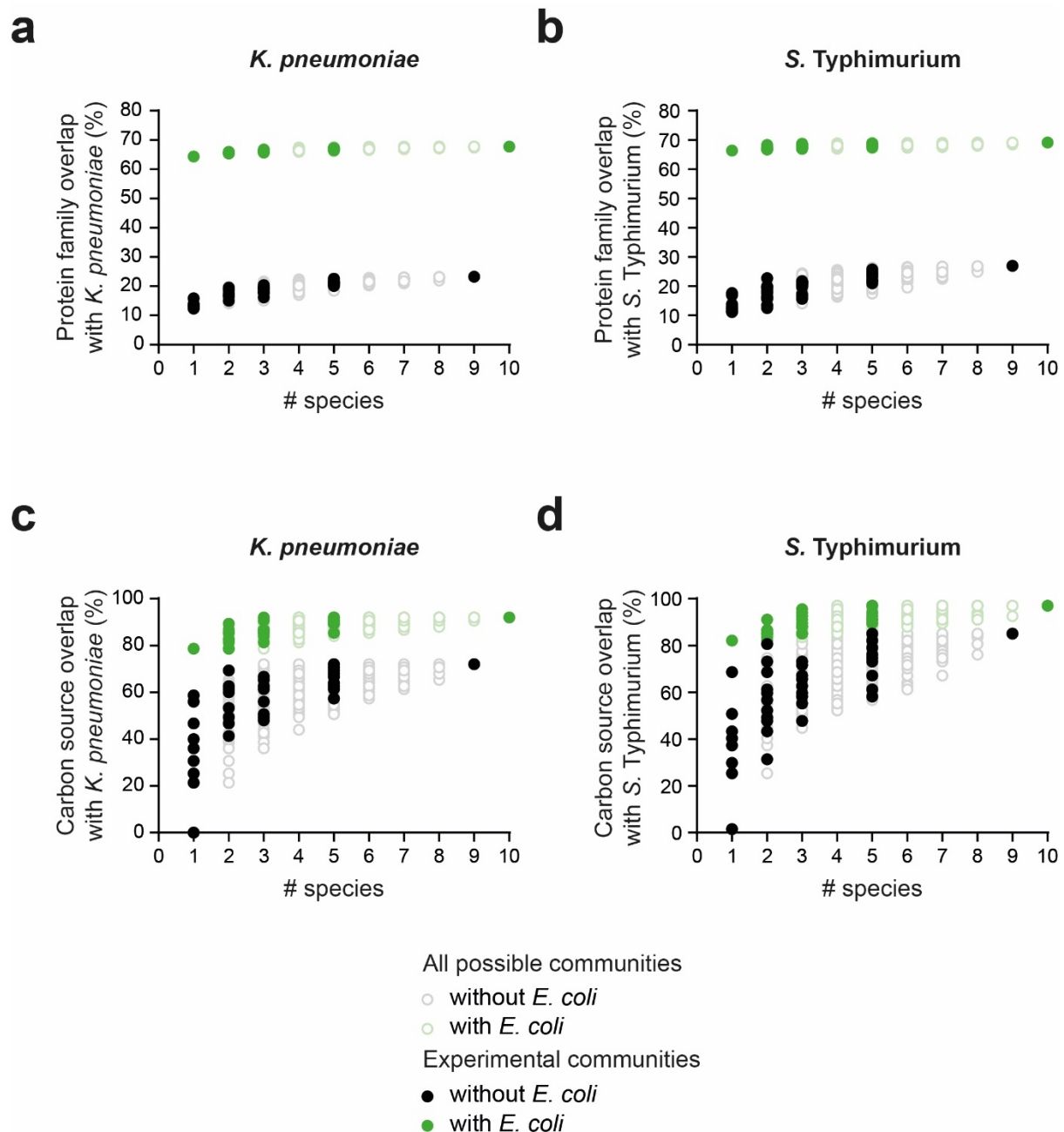


1016

1017

1018 **Figure S7. The number of protein families increase proportional to community diversity.** a)
1019 Histogram showing the distribution of protein families among the 50 species subset used for protein
1020 family analysis. The vertical red dashed line represents the average number of genomes out of the 50
1021 strains that share a particular protein family (2.44 genomes). There is an average of 3.22 proteins in
1022 each protein family. The histogram shows that many protein families are unique to a strain while others
1023 (141) are shared between all 50 strains. **b-c)** The number of protein families covered by a community
1024 increases as the number of strains in the community increases. Results for *K. pneumoniae* shown in (b)
1025 and for *S. Typhimurium* in (c). Community IDs taken from Fig. 2c-d. Each circle represents a different
1026 community. Green circles depict communities containing *E. coli*, while black circles are communities
1027 without *E. coli*.

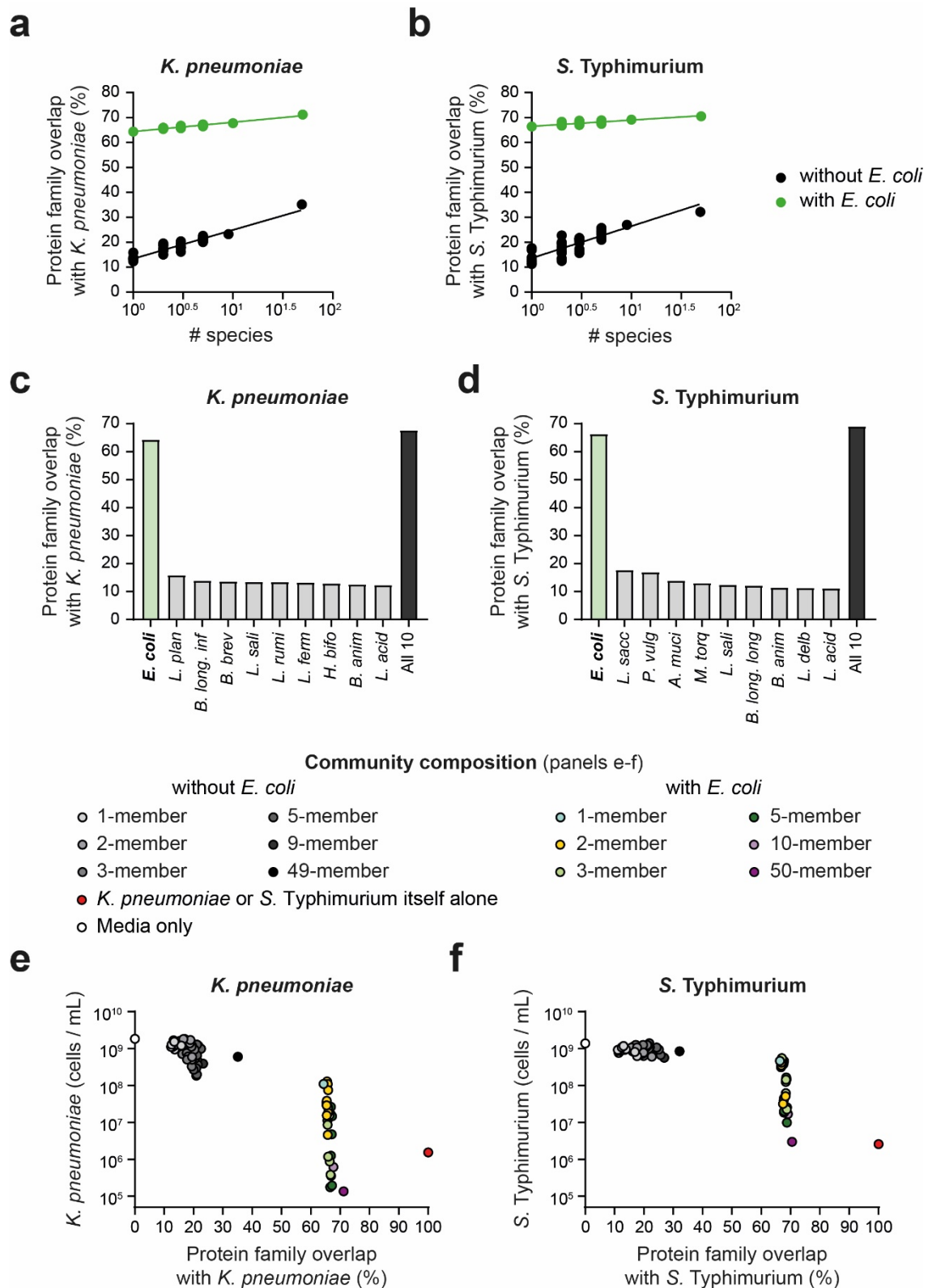
1028



1029

1030 **Figure S8. The randomly chosen communities used in the *in vitro* experiments are representative**
 1031 **of all possible combinations of the 10 best-ranked species. a-b)** Experimental communities contain
 1032 representative protein family overlap to the pathogens compared to all possible combinations of the 10
 1033 best-ranked species. All possible combinations of communities are depicted by unfilled circles,
 1034 experimentally tested communities are shown as filled circles. Communities in green contain *E. coli*
 1035 while communities in black do not contain *E. coli*. Data for *K. pneumoniae* shown in (a) and for *S.*
 1036 *Typhimurium* in (b). **c-d)** Experimental communities contain representative carbon source utilisation
 1037 overlap to the pathogens compared to all possible combinations of the 10 best-ranked species. All
 1038 possible combinations of communities are depicted by unfilled circles, experimentally tested
 1039 communities are shown as filled circles. Communities in green contain *E. coli* while communities in
 1040 black do not contain *E. coli*. Data for *K. pneumoniae* shown in (c) and for *S. Typhimurium* in (d).

1041



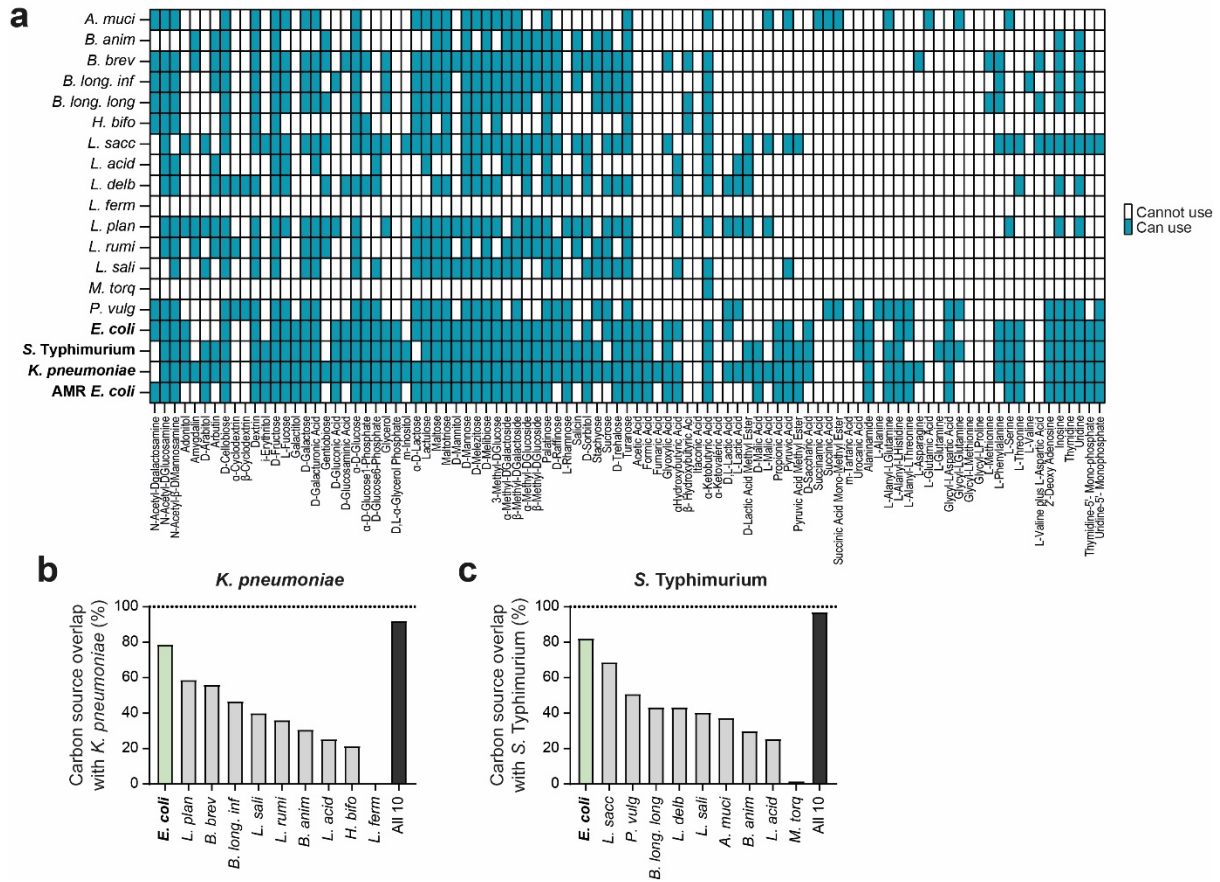
1042 **Figure S9. Protein family overlap between single strains or communities and the pathogen shows**
 1043 **that both diversity and key members (*E. coli*) are important in explaining predicted colonisation**
 1044 **resistance. a-b) As community diversity increases, protein family overlap with the pathogen increases.**
 1045 **Results for *K. pneumoniae* shown in (a) and for *S. Typhimurium* in (b). Each circle represents a**
 1046 **community (communities from Fig. 2c-d). Green circles depict communities containing *E. coli*, while**
 1047 **black circles are communities without *E. coli*. Linear regression of log-transformed data: (a) $R^2=0.9350$,**
 1048 **non-zero slope for *E. coli* communities (F test, $p<0.0001$). $R^2=0.9182$, non-zero slope for communities**

1049 without *E. coli* (F test, $p < 0.0001$). **(b)** $R^2 = 0.6825$, non-zero slope for *E. coli* communities (F test,
1050 $p < 0.0001$). $R^2 = 0.7484$, non-zero slope for communities without *E. coli* (F test, $p < 0.0001$). **c-d)** Bar
1051 chart showing the protein family overlap with the pathogen for the individual 10 best-ranked strains.
1052 Results for *K. pneumoniae* shown in **(c)** and for *S. Typhimurium* in **(d)**. The bar for *E. coli* is shown in
1053 green and the other strains in grey. The predicted protein family overlap for all 10 strains is shown in
1054 dark grey. See **Table S1** for species name abbreviations. **e-f)** As community cluster overlap with the
1055 pathogen increases, the pathogen abundance on day 2 of the extended competition decreases. Coloured
1056 circles depict communities containing *E. coli*, while black circles represent communities without *E. coli*
1057 (data from **Fig. 2c-d**). Colour or greyscale gradients indicate the diversity of the community. The red
1058 circles represent the isogenic wildtype pathogens. Results for *K. pneumoniae* shown in **(e)** and for *S.*
1059 *Typhimurium* in **(f)**.

1060

1061

1062



1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

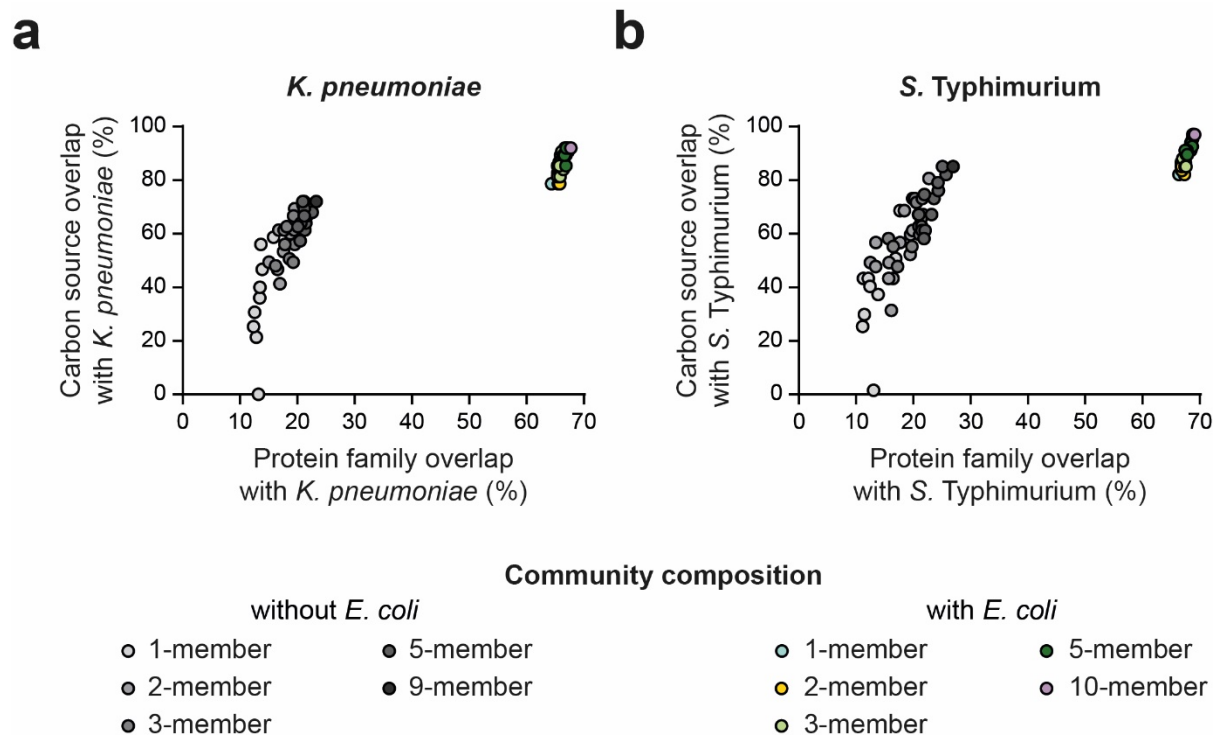
1073

1074

1075

1076

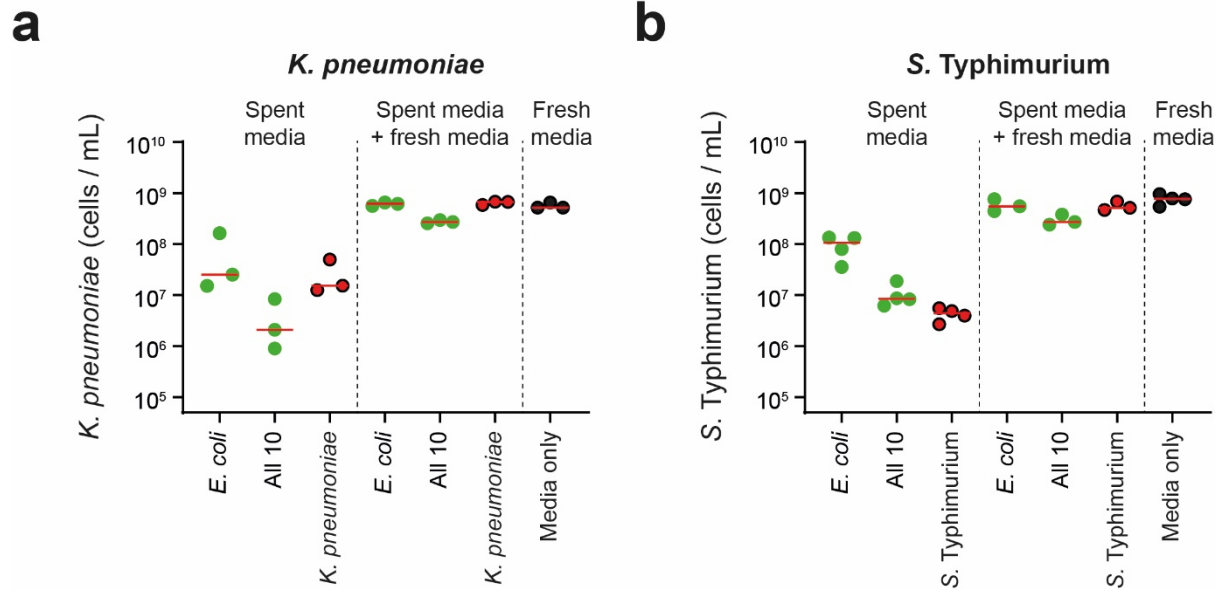
Figure S10. Individual carbon source utilisation profiles of the 10 best-ranked symbiont strains for *K. pneumoniae* and *S. Typhimurium* (16 symbiont strains total) and their overlap with the pathogens. (a) The x-axis lists the 95 individual nutrients tested in the AN Biolog microplates and the y-axis shows the 16 symbiont strains and the 3 pathogens (*K. pneumoniae*, *S. Typhimurium*, AMR *E. coli*). The pathogens are highlighted in bold on the y-axis. Nutrients shaded in blue could be used by a strain whereas those in white could not be used as defined by a threshold of background-subtracted Abs_{590nm} . **b-c)** Bar charts showing carbon source utilisation overlap (%) of the 10 best ranked individual strains with the pathogens. Results for *K. pneumoniae* in (b) and *S. Typhimurium* in (c). *E. coli* is shown in green and the other individual strains in grey. The predicted utilisation of all 10 best-ranked symbionts together is shown in dark grey. The percentage overlap with the pathogen is calculated as the proportion of the number of nutrients able to be used by the pathogen that can also be used by a particular symbiont strain or community. See **Table S1** for species name abbreviations.



1077

1078 **Figure S11. The protein family overlap and carbon source overlap prediction approaches are**
 1079 **positively correlated. a-b)** Correlation between protein cluster percentage overlap with the pathogen
 1080 and carbon source utilisation percentage overlap with the pathogen. Results for *K. pneumoniae* in **(a)**
 1081 and *S. Typhimurium* in **(b)**. Communities shown in colour contain *E. coli* and those in black do not.
 1082 Gradients of colour or greyscale intensity show community diversity. **(a)** $R^2=0.6119$, slope significantly
 1083 different than 0 by an F test for communities with *E. coli* ($p<0.0001$). $R^2=0.6838$, slope significantly
 1084 different than 0 by an F test for communities without *E. coli* ($p<0.0001$). **(b)** $R^2=0.8204$, slope
 1085 significantly different than 0 by an F test for communities with *E. coli* ($p<0.0001$). $R^2=0.7027$, slope
 1086 significantly different than 0 by an F test for communities without *E. coli* ($p<0.0001$). The communities
 1087 are the same as those in **Fig. 2c-d**. Percentage overlap calculated as the proportion of shared carbon
 1088 source use or shared protein families with the pathogen. Values for communities calculated in an
 1089 additive way based on the profiles of individual strains.

1090

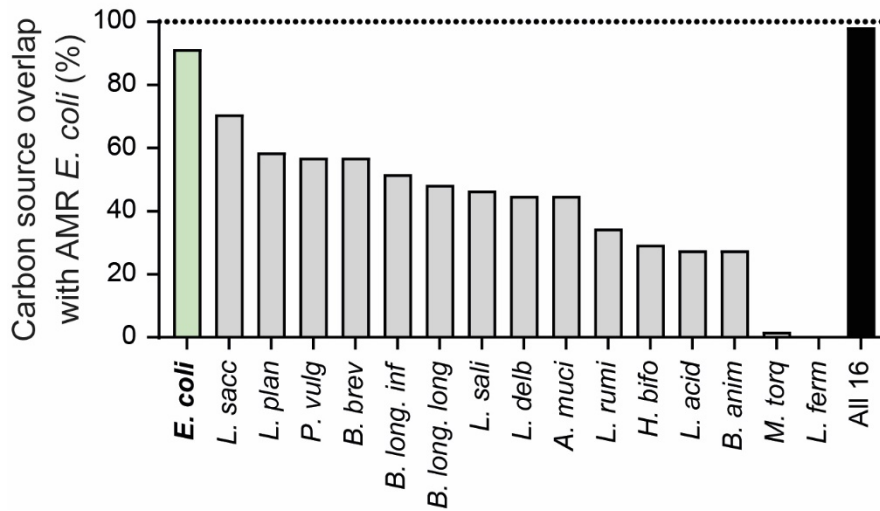


1091

1092 **Figure S12. Spent media experiment.** Communities were assembled and grown for 96 hours and the
 1093 pathogen invaded into the spent media or re-supplemented spent media (half volume spent media and
 1094 half volume nutrient media). Pathogen density measured by flow cytometry 24 hours after pathogen
 1095 invasion (day 1). Results for *K. pneumoniae* shown in (a) and *S. Typhimurium* in (b). N=3-4 replicates
 1096 per treatment. Horizontal red lines show the median of the replicates.

1097

AMR *E. coli*



1098

1099 **Figure S13. Carbon source utilisation overlap with AMR *E. coli*.** Bar chart showing carbon source
 1100 utilisation overlap (%) of the 16 best ranked individual strains (for *K. pneumoniae* and *S. Typhimurium*)
 1101 with the AMR *E. coli* strain. The *E. coli* symbiont is shown in green and the other individual strains in
 1102 grey. All 16 symbiont strains together are shown in black. The percentage overlap with the pathogen is
 1103 calculated as the proportion of the number of nutrients that can be used by AMR *E. coli* that can also
 1104 be used by a particular symbiont strain.

1105

1106

1107 **Supplementary Table S1. Strains.**

1108

Strain	Abbrev.	Relevant genotype *	Internal ID	PATRIC ID	Collection ID	Hazard group**	100-strain screen	50-strain community	<i>K. pneumoniae</i> 10 best-ranked	<i>S. Typhimurium</i> 10 best-ranked	Source
<i>Actinomyces odontolyticus</i>			NT5039	411466.7	DSM43331	2	+				Nassos Typas
<i>Akkermansia muciniphila</i>	<i>A.muci</i>		F3	349741.6	DSM22959	1	+	+		+	DSMZ***
<i>Anaerostipes caccae</i>			F2	105841.35	DSM14662	1	+	+			DSMZ
<i>Bacteroides caccae</i>			B1	411901.7	DSM19024	2	+				DSMZ
<i>Bacteroides cellulosilyticus</i>			B2	537012.5	DSM14838	1	+	+			DSMZ
<i>Bacteroides clarus</i>			NT5052	762984.10	DSM22519	1	+	+			Nassos Typas
<i>Bacteroides eggerthii</i>			B13	483216.6	DSM20697	2	+				DSMZ
<i>Bacteroides fragilis</i> enterotoxigenic 20656-2-1			NT5033	817.95	ATCC43860	2	+				Nassos Typas
<i>Bacteroides fragilis</i> 3_1_12			NT5057	457424.5	HM-20	2	+				Nassos Typas
<i>Bacteroides fragilis</i> CL07T12C05			NT5059	997883.3	HM-710	2	+				Nassos Typas
<i>Bacteroides fragilis</i> CL05T00C42			NT5060	997880.3	HM-711	2	+				Nassos Typas

<i>Bacteroides fragilis</i> CL05T12C13			NT5062	997881.3	HM-712	2	+				Nassos Typas
<i>Bacteroides fragilis</i> CL03T00C08			NT5063	997878.3	HM-713	2	+				Nassos Typas
<i>Bacteroides fragilis</i> CL03T12C07			NT5061	997879.7	HM-714	2	+				Nassos Typas
<i>Bacteroides fragilis</i> nontoxigenic			B4	272559.1 7	DSM2151	2	+				DSMZ
<i>Bacteroides ovatus</i>			B6	411476.1 1	DSM1896	2	+				DSMZ
<i>Bacteroides stercoris</i> VPI B5-21			B8	46506.15 62	DSM19555	1	+	+			DSMZ
<i>Bacteroides stercoris</i> CC31F			NT5055	1073351. 3	HM-1036	2	+				Nassos Typas
<i>Bacteroides thetaiotaomicro n</i>			B9	226186.1 2	DSM2079	2	+				DSMZ
<i>Bacteroides uniformis</i>			B10	820.37	DSM6597	2	+				DSMZ
<i>Bacteroides uniformis</i> CL03T12C37			NT5066	997890.6	HM-716	2	+				Nassos Typas
<i>Bacteroides xylanisolvens</i> XB1A			B12	657309.4	DSM18836	1	+	+			DSMZ
<i>Bacteroides xylanisolvens</i> CL03T12C04			NT5064	997892.5 0	DSM2079	2	+				Nassos Typas

<i>Bifidobacterium adolescentis</i>			NT5022	367928.6	DSM20083	1	+	+			Nassos Typas
<i>Bifidobacterium animalis</i>	<i>B. anim</i>		A2	555970.3	DSM10140	1	+	+	+	+	DSMZ
<i>Bifidobacterium animalis</i> subsp. lactis Bi-04			NT5043	580050.3		1	+				Nassos Typas
<i>Bifidobacterium animalis</i> subsp. lactis Bi-07			NT5044	742729.3	DGCC2908	1	+	+			Nassos Typas
<i>Bifidobacterium bifidum</i>			A1	500634.6	DSM20456	1	+	+			DSMZ
<i>Bifidobacterium breve</i>	<i>B. brev</i>		A3	518634.1 9	DSM20213	1	+	+	+		DSMZ
<i>Bifidobacterium longum</i> subsp. infantis	<i>B. long. inf</i>		A4	391904.8	DSM20088	1	+	+	+		DSMZ
<i>Bifidobacterium longum</i> subsp. longum	<i>B. long. long</i>		A5	565042.3	DSM20219	1	+	+		+	DSMZ
<i>Blautia hansenii</i>			NT5005	537007.6	DSM20583	1	+	+			Nassos Typas
<i>Blautia hydrogenotrophica</i>			F4	476272.2 1	DSM10507	1	+	+			DSMZ
<i>Christensenella minuta</i>			F7	626937.8	DSM22607	1	+	+			DSMZ
<i>Clostridium difficile</i> 630			NT23006	272563.8	DSM27543	2	+				Nassos Typas
<i>Clostridium leptum</i>			F12	428125.8	DSM753	1	+	+			DSMZ
<i>Clostridium perfringens</i>			NT5031	195103.1 0	DSM756	2	+				Nassos Typas

<i>Clostridium perfringens</i>			NT5032	451754.18	DSM11782	2	+				Nassos Typas
<i>Collinsella aerofaciens</i>			A6	411903.6	DSM3979	2	+				DSMZ
<i>Coprococcus comes</i>			NT5048	470146.3	ATCC27758	1	+	+			Nassos Typas
<i>Dorea formicigenerans</i>			NT5076	411461.20	DSM3992	1	+	+			Nassos Typas
<i>Dorea longicatena</i>			F13	411462.6	DSM13814	1	+	+			DSMZ
<i>Eggerthella lenta</i>			NT5024	479437.5	DSM2243	2	+				Nassos Typas
<i>Enterocloster bolteae</i>			NT5026	208479.10	DSM15670	1	+	+			Nassos Typas
<i>Erysipelatoclostridium ramosum</i>			NT5006	445974.19	DSM1402	2	+				Nassos Typas
AMR <i>Escherichia coli</i>		Wildtype (Ampicillin)	19Y000018		-	2					Nottingham University Hospital Pathogen Bank
<i>Escherichia coli</i> ED1a			NT5078	585397.9		1	+				Nassos Typas
<i>Escherichia coli</i> HS			e-OPC-323	331112.6		1	+				(32)
<i>Escherichia coli</i> IAI1	<i>E. coli</i>		NT5077	585034.5		1	+	+	+	+	Nassos Typas
<i>Escherichia coli</i> IAI1	<i>E. coli</i> Δ gatABC	Δ gatABC	eOPC-364			1					This study
<i>Escherichia coli</i> JKE201	+pOPC-231		eOPC-362			1					This study

<i>Escherichia coli</i> JKe201	+pOPC-232		eOPC-363			1					This study
<i>Escherichia coli</i> MG1655			e-OPC-292	511145.1 2		1	+				(33)
<i>Escherichia coli</i> Z1269			Z1269		-	1					(31)
<i>Escherichia coli</i> Z1331			Z1331		-	1					(31)
<i>Eubacterium rectale</i>			NT5009	657318.1 2	DSM17629	1	+	+			Nassos Typas
<i>Eubacterium siraeum</i>			NT5040	428128.1 9	DSM15702	1	+	+			Nassos Typas
<i>Faecalibacterium prausnitzii</i>			F25	411483.3	DSM17677	1		+			DSMZ
<i>Fusobacterium nucleatum</i> CTI-01			NT24006	1204474. 3		2	+				Nassos Typas
<i>Fusobacterium nucleatum</i> MJR7757B			NT24015	851.8		2	+				Nassos Typas
<i>Fusobacterium nucleatum</i> subsp. <i>nucleatum</i>			NT5025	190304.8	DSM15643	2	+				Nassos Typas
<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i>			NT24005	155615.5	DSM19508	2	+				Nassos Typas
<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i>			NT5030	209882.4	DSM19507	2	+				Nassos Typas
<i>Fusobacterium periodonticum</i> 1_A_54/D10			NT24011	546275.3	ATCC33693	1	+	+			Nassos Typas

<i>Fusobacterium periodonticum</i> 2_1_31			NT24012	469599.3		2	+				Nassos Typas
<i>Gemella morbillorum</i>			NT24013	562982.3		2	+				Nassos Typas
<i>Holdemanella bififormis</i>	<i>H. bifo</i>		F15	518637.5	DSM3989	1	+	+	+		DSMZ
<i>Intestinibacter bartletti</i>			NT5086	261299.1 18	DSM16795	1	+				Nassos Typas
<i>Klebsiella pneumoniae</i>			K1	272620.9	ATCC 700721	2	+				Modernising Medical Microbiology, Nuffield Department of Medicine
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>		Wildtype (Carbenicillin)	bFS-26	1162296. 3	DSM 30104	2	+				DSMZ
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>		+ pRSJ- p _{nptII} ::ilux	bFS-29		DSM 30104	2					This study
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>		+pBC11	bFS-34		DSM 30104	2					This study
<i>Lachnoclostridium symbiosum</i> WAL-14163			NT24007	742740.3		2	+				Nassos Typas
<i>Lachnoclostridium symbiosum</i> WAL-14673			NT24014	742741.3		2	+				Nassos Typas
<i>Lachnoclostridium hylemonae</i>			NT27002	553973.1 9	DSM15053	1	+	+			Nassos Typas

<i>Lachnoclostridium scindens</i>			F9	411468.4 1	DSM5676	1	+	+			DSMZ
<i>Lachnoclostridium symbiosum</i>			F8	411472.5	DSM934	2	+				DSMZ
<i>Lacrimispora saccharolytica</i>	<i>L. sacc</i>		NT5037	610130.3	DSM2544	1	+	+		+	Nassos Typas
<i>Lacticaseibacillus casei</i>			F16	219334.4	DSM20011	1	+	+			DSMZ
<i>Lacticaseibacillus paracasei</i>			NT5042	1226298. 3	ATCCSD5275	1	+	+			Nassos Typas
<i>Lactiplantibacillus plantarum</i> JDM1	<i>L. plan</i>		F14	644042.3	-	1	+	+	+		Department of Food and Nutritional Sciences, University of Reading
<i>Lactobacillus acidophilus</i>	<i>L. acid</i>		NT5041	272621.1 3	ATCC700936	1	+	+	+	+	Nassos Typas
<i>Lactobacillus delbrueckii</i> subsp. Delbrueckii	<i>L. delb</i>		NT14075	1423823. 4	DSM20074	1	+	+		+	Nassos Typas
<i>Lactobacillus gasseri</i>			F18	324831.1 3	DSM20243	1	+	+			DSMZ
<i>Ligilactobacillus ruminis</i>	<i>L. rumi</i>		F17	1423798. 5	DSM20403	1	+	+	+		DSMZ
<i>Ligilactobacillus salivarius</i>	<i>L. sali</i>		NT14072	1423799. 3	DSM20555	1	+	+	+	+	Nassos Typas

<i>Limosilactobacillus fermentum</i>	<i>L. ferm</i>		NT14076	1613.547	DSM20052	1	+	+	+		Nassos Typas
<i>Mediterraneibacter gnavus</i>			NT5046	411470.47	ATCC29149	1	+	+			Nassos Typas
<i>Mediterraneibacter torques</i>	<i>M. torq</i>		NT5047	411460.6	ATCC27756	1	+	+		+	Nassos Typas
<i>Odoribacter splanchnicus</i>			NT5081	709991.142	DSM20712	2	+				Nassos Typas
<i>Parabacteroides distasonis</i>			NT5074	435591.48	DSM20701	2	+				Nassos Typas
<i>Parabacteroides merdae</i>			NT5071	411477.88	DSM19495	1	+				Nassos Typas
<i>Peptostreptococcus stomatis</i>			NT24002	596315.3	DSM17678	2	+				Nassos Typas
<i>Phocaeicola coprocola</i>			B3	470145.69	DSM17136	1	+	+			DSMZ
<i>Phocaeicola dorei</i>			NT5049	357276.1035	DSM17855	1	+	+			Nassos Typas
<i>Phocaeicola massiliensis</i>			B5	1121098.3	DSM17679	1	+	+			DSMZ
<i>Phocaeicola vulgatus</i>	<i>P. vulg</i>		B11	435590.9	DSM1447	1	+	+		+	DSMZ
<i>Phocaeicola vulgatus</i> CL09T03C04			NT5056	997891.3		1	+				Nassos Typas
<i>Prevotella buccae</i>			B14	873513.3	DSM19025	2	+				DSMZ
<i>Prevotella copri</i>			B15	537011.439	DSM18205	1	+	+			DSMZ

<i>Roseburia faecis</i>			F24	301302.4	DSM16840	1	+	+			DSMZ
<i>Roseburia hominis</i>			F22	585394.18	DSM16839	1	+	+			DSMZ
<i>Roseburia intestinalis</i>			NT5011	536231.75	DSM14610	1	+	+			Nassos Typas
<i>Roseburia inulinivorans</i>			NT5012	622312.48	DSM16841	1	+	+			Nassos Typas
<i>Salmonella enterica</i> Typhimurium SL1344		Wildtype (Streptomycin)	SB300	216597.6	DSM24522	2	+				(55)
<i>Salmonella enterica</i> Typhimurium SL1344		+ pRSJ-p _{nptII} ::ilux	sOPC-406		DSM24522	2					This study
<i>Salmonella enterica</i> Typhimurium SL1344		+pBC11	sOPC-404		DSM24522	2					This study
<i>Salmonella enterica</i> Typhimurium SL1344		Avirulent	M2702		DSM24522	2					(67)
<i>Salmonella enterica</i> Typhimurium SL1344		<i>hisG</i> prototroph (Streptomycin)	EB199		DSM24522	2					This study
<i>Salmonella enterica</i> Typhimurium SL1344		Δ gatABC (Streptomycin)	sOPC-463		DSM24522	2					This study

<i>Salmonella enterica</i> Typhimurium SL1344		$\Delta gatABC$ +pBC11	EB149		DSM24522	2						This study
<i>Staphylococcus epidermis</i>			bOPC-105	176280.8 5	DSM1798	2	+					ATCC****
<i>Streptococcus parasanguinis</i>			NT5072	760570.3	DSM6778	2	+					Nassos Typas
<i>Streptococcus salivarius</i>			NT5038	1304.182 9	DSM20560	2	+					Nassos Typas
<i>Veillonella parvula</i>			NT5017	479436.6	DSM2008	2	+					Nassos Typas

1109

1110

1111

1112

1113 **Supplementary Table S2. Plasmids.**

1114

Plasmid name	Relevant genotype	Resistance*	Source
pRSJ-p _{nptII} ::ilux	<i>luxCDBAE-frp</i> expression	Tetracycline	(58)
pBC11	YPet expression	Kanamycin	(61)
pOPC-231	<i>SV-aphT-tetR-sceI-</i> <i>sacB-STm-ΔgatABC</i>	Kanamycin	This study
pOPC-232	<i>SV-aphT-tetR-sceI-</i> <i>sacB-EclAI-ΔgatABC</i>	Kanamycin	This study

1115

1116 *Relevant resistances only. Streptomycin= >50 ug/mL; carbenicillin=>50μg/mL;

1117 ampicillin=>100μg/mL; kanamycin=>50μg/mL; tetracycline=>50μg/mL.

1118 **According to European biosafety designation.

1119 *** Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Inhoffenstraße

1120 7, 38124 Braunschwe, Science Campus Braunschweig-Süd, Germany.

1121 **** ATCC (American Type Culture Collection), 10801 University Boulevard, Manassas, Virginia

1122 20110-2209, United States.

1123

1124

1125

1126

1127 **Supplementary Table S3. Primers.**

1128

Primer name	Sequence	Purpose	Source
oOPC-953	AGAGTTTGATCCTGGCTCAG	16S sequencing (27F)	(68)
oOPC-954	TACGGYTACCTTGTTACGACTT	16S sequencing (1492R)	(68)
g-Bifid-F	5'-CTCCTGGAAACGGGTGG-3'	16S sequencing for Bifidobacteria	(69)
g-Bifid-R	5'-GGTGTCTTCCCGATATCTACA-3'		(69)
oOPC-975	CCCAGTCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCaaatcgcgtttcgtgaatcagg	amplify 700 upstream <i>gatABC</i> for deletion - STm SL1344	This study
oOPC-976	taaaattaagaggcgatttgaatagttggctcataaattctccattattcagg		This study
oOPC-977	aatttatgagccaactatttcaaatcgctcttaattttaggggag	amplify 700 downstream <i>gatABC</i> for deletion - STm	This study
oOPC-978	CTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCctcccgctattacaggtatgcggttgcgc		This study
oOPC-979	agtcgatgctgcacgtacgc	check <i>gatABC</i> deletion - STm	This study
oOPC-980	atgtcggacaacgcggtctg		This study
oOPC-981	CCCAGTCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCtactgttaaatgttgcacgcacc	amplify 700 upstream <i>gatABC</i> for deletion - <i>E. coli</i> IA11	This study
oOPC-982	ggtatatgactaacctgtttgttctcgagaataattttacctgaggg		This study
oOPC-983	aaaaattattctgcgagaacaacaggttagtcatataccgtccttattccg	amplify 700 downstream <i>gatABC</i> for deletion - <i>E. coli</i> IA11	This study
oOPC-984	CTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCcttacgtacgcatcaaaggccttattgccc		This study

oOPC-985	aaacgctctgcatttcggc	check <i>gatABC</i> deletion - <i>E.coli</i> IAI1	This study
oOPC-986	cccatgttgaagatgccgcc		This study

1129

1130

1131 **Supplementary table S4.** Species compositions of communities used for *in*
 1132 *vitro* experiments with *Klebsiella pneumoniae* DSM 30104.

1133

Community ID	Community members	No. replicates	Median day 2 pathogen density (cells/mL)
No commensals	N/A	15	1.84E+09
KTS1	A3	5	9.45E+08
KTS2	F15	3	1.11E+09
KTS3	F17	3	1.19E+09
KTS4	A4	3	1.29E+09
KTS5	NT5077	3	1.10E+08
KTS6	NT5041	3	1.12E+09
KTS7	A2	3	1.28E+09
KTS8	F14	3	1.21E+09
KTS9	NT14076	3	1.68E+09
KTS10	NT14072	3	1.53E+09
KTD1	A4, NT5041	3	1.33E+09
KTD2	NT5041, F14	3	1.29E+09
KTD3	A3, NT5077	4	4.58E+06
KTD4	F15, F14	3	1.40E+09
KTD5	F15, A2	3	1.83E+09
KTD6	A2, F14	3	1.70E+09
KTD7	F17, NT14072	3	1.59E+09
KTD8	NT5077, NT14076	4	1.28E+08
KTD9	A3, F14	3	5.96E+08
KTD10*	NT5077, F14	3	1.82E+07
KTD11*	F15, NT5077	3	1.09E+08
KTD12*	F17, NT5077	3	3.42E+07
KTD13*	A4, NT5077	3	7.44E+07
KTD14*	NT5077, NT5041	3	1.55E+07
KTD15*	NT5077, A2	3	3.84E+07
KTD16*	NT5077, NT14072	3	2.90E+07
KTD17*	NT5041, A2	3	1.81E+09
KTT1	A3, F15, NT5041	3	3.30E+08
KTT2	A3, F15, A2	3	7.44E+08
KTT3	A2, NT14076, NT14072	3	6.56E+08
KTT4	F17, NT5077, NT14072	3	1.12E+07
KTT5	NT5041, NT14076, NT14072	3	7.11E+08
KTT6	F17, NT5077, A2	3	2.52E+07
KTT7	F17, A4, NT14076	3	1.07E+09
KTT8	A3, F15, F17	4	5.03E+08
KTT9	A3, NT5041, A2	3	7.59E+08
KTT10	A3, A4, NT5041	3	6.13E+08
KTT11	A2, F14, NT14072	3	1.19E+09
KTT12	A4, NT5041, F14	3	9.36E+08
KTT13	F15, F14, NT14072	3	1.13E+09
KTT14	A3, NT5077, F14	3	3.81E+05
KTT15	A3, F17, A2	3	5.07E+08
KTT16	NT5077, F14, NT14072	3	1.72E+07
KTT17	NT5077, NT14076, NT14072	3	1.43E+07
KTT23*	A3, F15, NT5077	2	8.73E+05

KTT24*	A4, NT5077, NT14072	3	1.29E+07
KTT25*	A3, NT5077, A2	3	6.53E+05
KTT26*	NT5077, NT5041, NT14072	3	6.20E+06
KTP1	A3, F15, F17, NT5041, NT14076	4	4.77E+08
KTP2	A3, F15, NT5041, A2, NT14072	4	2.98E+08
KTP3	NT5077, A2, F14, NT14076, NT14072	4	1.66E+07
KTP4	F17, NT5077, NT5041, A2, NT14072	4	2.42E+07
KTP5	A3, F15, NT5041, NT14076, NT14072	4	3.63E+08
KTP6	A3, F15, F17, A4, A2	4	5.33E+08
KTP7	F17, A4, NT5077, A2, NT14076	4	2.63E+07
KTP8	A3, F15, F17, NT5077, NT5041	4	1.25E+06
KTP9	A3, A4, NT5041, A2, F14	4	1.85E+08
KTP10	A3, A4, NT5041, F14, NT14072	4	1.99E+08
KTP11	F15, NT5041, A2, F14, NT14072	4	1.19E+09
KTP12	F17, NT5041, A2, F14, NT14076	4	9.19E+08
KTP13	F15, A4, F14, NT14076, NT14072	4	9.89E+08
KTP14	A3, A4, A2, F14, NT14076	4	6.60E+08
KTP15	F17, A4, F14, NT14076, NT14072	4	1.28E+09
KTP16	A3, F17, A4, F14, NT14072	4	2.95E+08
KTP17	A3, A4, A2, NT14076, NT14072	4	2.67E+08
KTP18*	F15, NT5077, NT5041, A2, F14	3	2.72E+06
KTP19*	F17, A4, NT5077, F14, NT14076	3	1.28E+07
KTP20*	A3, NT5077, NT5041, A2, NT14076	3	2.82E+05
KTP21*	A3, F15, NT5077, NT5041, F14	3	3.26E+05
KTP22*	A3, F17, NT5077, NT5041, F14	3	3.08E+05
KALLTOP	A3, F15, F17, A4, NT5077, NT5041, A2, F14, NT14076, NT14072	11	6.17E+05

1134

1135 *Communities that were additionally selected to contain *E. coli* IA11, but otherwise were selected at
1136 random.

1137

1138 **Supplementary table S5.** Species compositions of communities used for *in*
 1139 *vitro* experiments with *Salmonella enterica* serovar Typhimurium SL1344.

1140

Community ID	Community members	No. replicates	Median day 2 pathogen density (cells/mL)
No commensals	N/A	11	1.38E+09
STS1	A2	3	7.64E+08
STS2	NT5037	3	6.29E+08
STS3	NT14075	3	8.71E+08
STS4	F3	3	9.55E+08
STS5	NT5041	3	9.26E+08
STS6	A5	3	8.69E+08
STS7	NT5077	5	4.46E+08
STS8	B11	4	8.14E+08
STS9	NT5047	3	1.18E+09
STS10	NT14072	3	1.08E+09
STD1	NT5037, B11	3	6.04E+08
STD2	NT5077, NT5047	3	3.60E+08
STD3	NT5041, B11	3	7.16E+08
STD4	A5, NT5047	3	9.19E+08
STD5	NT14075, NT5077	3	3.98E+08
STD6	NT5037, NT14072	3	6.30E+08
STD7	B11, NT5047	3	7.64E+08
STD8	NT14075, NT14072	3	9.23E+08
STD9	NT5037, NT5047	3	7.74E+08
STD10	NT14075, NT5047	3	9.25E+08
STD11	NT14075, NT5041	3	9.62E+08
STD12	NT5041, NT14072	3	9.98E+08
STD13	A2, NT14072	3	1.06E+09
STD14	A2, NT5047	3	8.40E+08
STD15	B11, NT14072	3	7.79E+08
STD16	NT14075, F3	3	1.07E+09
STD17*	A2, NT5077	3	3.36E+08
STD18*	NT5037, NT5077	4	5.73E+07
STD19*	F3, NT5077	3	3.23E+08
STD20*	NT5041, NT5077	3	3.11E+08
STD21*	A5, NT5077	3	3.52E+08
STD22*	NT5077, B11	7	3.94E+07
STD23*	NT5077, NT14072	3	4.14E+08
STT1	A2, NT14075, A5	3	9.50E+08
STT2	NT14075, F3, B11	3	1.00E+09
STT3	F3, NT5041, NT5077	3	5.51E+08
STT4	A2, A5, NT14072	3	1.15E+09
STT5	A2, NT5037, A5	3	9.38E+08
STT6	NT5037, NT5077, NT14072	3	1.25E+08
STT7	F3, NT5041, A5	3	1.26E+09
STT8	NT5037, A5, NT5047	3	1.05E+09
STT9	NT5037, F3, NT5077	3	1.44E+08
STT10	NT5037, A5, NT14072	3	9.17E+08
STT11	NT14075, B11, NT5047	3	9.20E+08
STT12	A2, NT14075, B11	3	1.10E+09
STT13	A5, B11, NT14072	3	9.89E+08

STT14	NT5041, NT5047, NT14072	3	1.26E+09
STT15	A5, B11, NT5047	3	1.25E+09
STT16	A2, F3, B11	3	1.22E+09
STT17*	NT5037, NT5077, B11	5	2.23E+07
STT18*	A2, NT14075, NT5077	3	3.75E+08
STT19*	NT5041, NT5077, NT5047	3	3.80E+08
STT20*	NT5037, NT5077, NT5047	3	6.17E+07
STT21*	F3, NT5077, NT5047	3	3.79E+08
STT22*	NT14075, NT5077, B11	3	3.87E+07
STP1	NT14075, F3, NT5041, NT5047, NT14072	4	1.19E+09
STP2	A2, NT5037, NT14075, NT5077, NT14072	4	1.65E+08
STP3	NT14075, F3, A5, B11, NT14072	4	9.01E+08
STP4	A2, F3, A5, B11, NT5047	4	1.05E+09
STP5	F3, NT5041, A5, NT5077, B11	4	4.53E+07
STP6	NT5037, F3, B11, NT5047, NT14072	4	6.69E+08
STP7	NT14075, F3, NT5041, A5, NT5047	4	1.09E+09
STP8	A2, NT5037, F3, NT5041, NT5047	4	1.07E+09
STP9	NT5037, NT14075, A5, B11, NT14072	4	9.40E+08
STP10	A2, F3, NT5041, A5, NT5047	4	1.40E+09
STP11	A2, NT5037, A5, NT5077, B11	4	2.55E+07
STP12	NT5037, A5, NT5077, NT5047, NT14072	4	1.46E+08
STP13	A2, NT14075, A5, B11, NT14072	4	8.71E+08
STP14	F3, NT5041, NT5077, NT5047, NT14072	4	4.59E+08
STP15	A2, NT5037, NT14075, F3, A5	4	1.06E+09
STP16	A2, NT14075, F3, A5, NT14072	4	1.19E+09
STP17*	NT5041, A5, NT5077, B11, NT14072	3	2.01E+07
STP18*	NT5037, NT5041, A5, NT5077, B11	3	9.83E+06
STP19*	NT14075, F3, A5, NT5077, NT14072	3	4.38E+08
STP20*	A2, F3, A5, NT5077, B11	3	1.83E+07
SALLTOP	A2, NT5037, NT14075, F3, NT5041, A5, NT5077, B11, NT5047, NT14072	11	1.76E+07

1141

1142 *Communities that were additionally selected to contain *E. coli* IA11, but otherwise were selected at
1143 random.

1144

1145

1146

1147

1148

1149 **Supplementary table S6.** Species compositions of communities used for
 1150 gnotobiotic mouse experiments.

1151

Pathogen gavaged	No. symbionts in community	Community members	<i>E. coli</i> IA11 present	No. replicates	Median pathogen faecal density 24 hours p.i. (CFU/g)
<i>K. pneumoniae</i>	0	N/A	No	7	4.0375e+09
<i>K. pneumoniae</i>	1	NT5077	Yes	7	2.28846e+08
<i>K. pneumoniae</i>	5	KTP8 (see Table S4)	Yes	8	4.35148e+07
<i>K. pneumoniae</i>	10	KALLTOP (see Table S4)	Yes	7	4.07019e+07
<i>K. pneumoniae</i>	50	See Table S1	Yes	7	134667
<i>K. pneumoniae</i>	9	KALLTOP minus NT5077	No	7	4.4e+08
<i>K. pneumoniae</i>	49	50 minus NT5077	No	7	2.32e+07
<i>S. Typhimurium</i>	0	N/A	No	7	7.89e+08
<i>S. Typhimurium</i>	1	NT5077	Yes	7	9.01408e+07
<i>S. Typhimurium</i>	5	STP11 (see Table S5)	Yes	8	5.68115e+07
<i>S. Typhimurium</i>	10	SALLTOP (see Table S5)	Yes	7	4.1e+07
<i>S. Typhimurium</i>	50	See Table S1	Yes	7	521000
<i>S. Typhimurium</i>	9	SALLTOP minus NT5077	No	8	6.91e+08
<i>S. Typhimurium</i>	49	50 minus NT5077	No	7	2.15e+07

1152

1153

1154 **Supplementary table S7.** Species compositions of communities used for
 1155 prediction experiments with the AMR *E. coli* strain 19Y000018.

1156

Community ID	Community members	No. replicates	Median day 2 pathogen density (CFU/mL)
No commensals	N/A	5	4.96E+08
Best 2 Biolog	NT5077, NT5037	5	1.00E+07
Worst 2 Biolog	NT5077, A2	5	3.12E+08
Best 3 Biolog	NT5077, B11, NT5037	5	1.56E+06
Worst 3 Biolog	NT5077, F17, NT5041	5	2.64E+08
Best 5 Biolog	NT5077, NT5037, NT14075, NT5041, B11	5	2.48E+06
Worst 5 Biolog	NT5077, NT14075, A5, A3, F17	5	4.64E+06
Best 2 protein family	NT5077, NT5026	5	1.80E+07
Worst 2 protein family	NT5077, F3	5	2.48E+08
Best 3 protein family	NT5077, NT5026, F2	5	1.60E+07
Worst 3 protein family	NT5077, A2, NT5044	5	2.24E+08
Best 5 protein family	NT5077, NT5026, F2, NT5049, NT24011	5	2.80E+06
Worst 5 protein family	NT5077, NT14072, NT5041, F24, NT5040	5	8.16E+06
Best 10 protein family	NT5077, F2, NT14076, NT5026, NT5049, NT5047, F25, NT24011, NT5037, NT5012	5	1.60E+06
Worst 10 protein family	NT5077, F17, NT5041, NT5076, NT5048, A1, F15, A2, NT5009, NT5044	5	3.40E+06
Best 5 protein family #2	NT5077, NT5026, NT24011, NT5037, NT5052	5	4.00E+06
Best 5 protein family #3	NT5077, NT5026, NT24011, NT5037, NT5049	5	2.02E+06
Best 5 protein family #4	NT5077, F2, NT24011, NT5037, NT5052	5	5.52E+06
Best 5 protein family #5	NT5077, F2, NT5026, F16, NT5049	5	2.48E+06
Worst 5 protein family #2	NT5077, F17, F15, A2, NT5044	5	3.00E+08
Worst 5 protein family #3	NT5077, NT5048, F15, A2, NT5044	5	2.32E+08
Worst 5 protein family #4	NT5077, NT14072, NT5040, A2, NT5044	5	3.20E+08
Worst 5 protein family #5	NT5077, F17, A1, A2, NT5044	5	2.80E+08

1157

1158

1159