bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bacterial cross-feeding can promote gene retention by lowering gene expression costs

Short title: Cross-feeding can promote gene retention

Ying-Chih Chuang[1,2,†], Megan G. Behringer[3], Gillian Patton[3], Jordan T. Bird[4], Crystal E. Love[1], Ankur Dalia[1], and James B. McKinlay[1]*

[1]Department of Biology, Indiana University, Bloomington, IN, USA
[2]Biochemistry Program, Indiana University, Bloomington, IN, USA
[3]Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA
[4]Battelle Memorial Institute, Columbus, OH 43201

*Corresponding author: James B. McKinlay
Email: jmckinla@iu.edu

[†] Current address: Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA

Competing interests statement: The authors declare no conflicts of interest.

Keywords: cross-feeding; purine; adenine; siderophore, microbial interactions; Rhodopseudomonas; palustris; Black Queen Hypothesis; mutualism; excretion; microbial physiology, gene loss

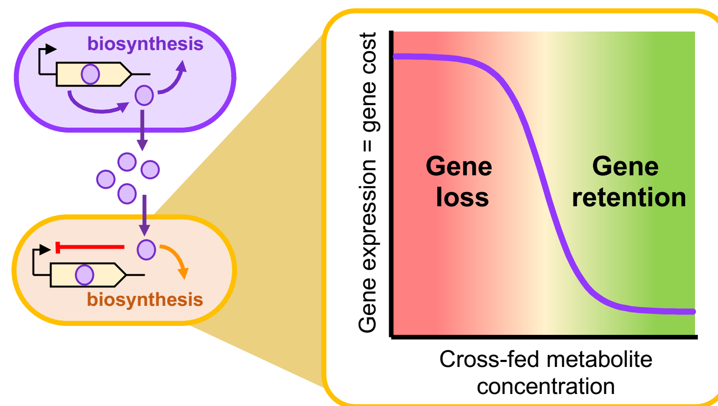
Study Funding information:

This work was supported in part by US Army Research Office grants W911NF-14-1-0411 and W911NF-17-1-0159, the National Science Foundation CAREER award MCB-1749489 to JBM, and National Institutes of Health grants R35GM128674 to ABD and R35GM150625 to MB. Supercomputing resources were supported in part by Lilly Endowment, Inc., through its support for the IU Pervasive Technology Institute.

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

33 **ABSTRACT**
34
35 Gene loss is expected in microbial communities when the benefit of obtaining a
36 biosynthetic precursor from a neighbor via cross-feeding outweighs the cost of retaining
37 a biosynthetic gene. However, gene cost primarily comes from expression, and many
38 biosynthetic genes are only expressed when needed. Thus, one can conversely expect
39 cross-feeding to repress biosynthetic gene expression and promote gene retention by
40 lowering gene cost. Here we examined long-term bacterial cocultures pairing
41 *Escherichia coli* and *Rhodopseudomonas palustris* for evidence of gene loss or
42 retention in response to cross-feeding of non-essential adenine. Although *R. palustris*
43 continued to externalize adenine in long-term cultures, *E. coli* did not accumulate
44 mutations in purine synthesis genes, even after 700 generations. *E. coli* purine
45 synthesis gene expression was low in coculture, suggesting that gene repression
46 removed selective pressure for gene loss. In support of this explanation, *R. palustris*
47 also had low transcript levels for iron-scavenging siderophore genes in coculture, likely
48 because *E. coli* facilitated iron acquisition by *R. palustris*. *R. palustris* siderophore gene
49 mutations were correspondingly rare in long-term cocultures but were prevalent in
50 monocultures where transcript levels were high. Our data suggests that cross-feeding
51 does not always drive gene loss, but can instead promote gene retention by repressing
52 costly expression.

53
54 **Graphical abstract**
55
56
57

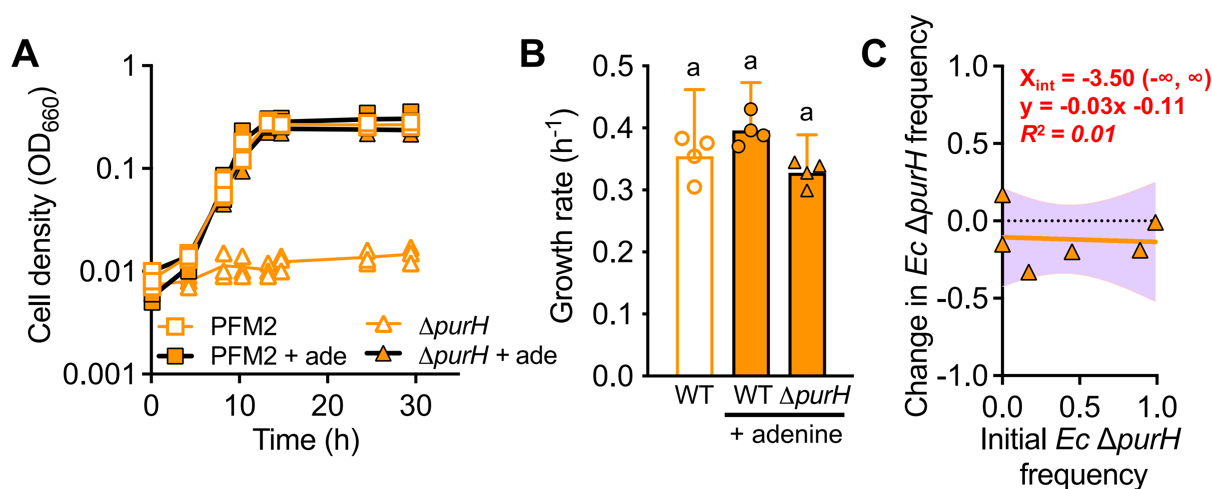


bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

58 **INTRODUCTION**
59
60 Changes to the genetic inventory of a given microbe can be influenced by living in a
61 community. An example of this influence is described in the Black Queen Hypothesis
62 (BQH). Based on the card game Hearts, where players avoid holding the costly Queen
63 of Spades, the BQH posits that gene loss (herein we do not distinguish loss-of-function
64 mutations from gene deletions for convenience) will occur when acquiring a resource
65 from neighbor is less costly than producing that resource (1). The BQH involves a
66 producer that generates a public resource, and a beneficiary that uses the resource and
67 is thus subject to beneficial gene loss. The BQH is analogous to the emergence of
68 cheaters that lose genes and benefit by exploiting a public resource. However, unlike a
69 cheater, a beneficiary is less likely to be the same species as the producer and is less
70 likely to harm the producer (1, 2). The BQH, or analogous processes, are now
71 commonly used to explain the prevalence of auxotrophs, which cannot synthesize one
72 or more essential nutrients (3, 4). Indeed, several groups have observed that
73 auxotrophs carrying biosynthetic gene deletions have a fitness advantage over wild-type
74 cells when corresponding biosynthetic precursors are provided (5-9).
75
76 A confounding aspect of BQH prediction of gene loss is that nutrient availability can also
77 repress gene expression and thus lower the cost of a gene (3, 10, 11); most of a gene's
78 cost comes from its expression (12-16), and thus repressing gene expression could
79 have similar cost-savings as a loss-of-function mutation. Thus, one can arrive at an
80 opposite prediction wherein cross-feeding could promote gene retention by repressing
81 costly gene expression. This prediction is well-documented from studies of cheaters.
82 For example, cheaters that had lost genes for iron-scavenging siderophores emerged
83 when iron was scare and siderophore genes were expressed, but gene loss was not
84 observed when iron was plentiful and siderophore genes were repressed (11, 17, 18).
85 Thus, nutrient availability alone might not be a good predictor for gene loss and the
86 emergence of beneficiaries.
87
88 Previously, we developed cocultures pairing fermentative *Escherichia coli* and
89 phototrophic *Rhodopseudomonas palustris*, wherein each species is dependent on the
90 other for essential nutrients; *E. coli* ferments glucose into organic acids that provide
91 essential carbon to *R. palustris* while *R. palustris* converts N_2 gas into NH_4^+, providing
92 essential nitrogen to *E. coli* (19). More recently, we discovered that *R. palustris* also
93 externalizes adenine at levels that can support an *E. coli* purine auxotroph (20, 21).
94 There is also suggestive evidence that *E. coli* facilitates *R. palustris* iron acquisition,
95 perhaps through cross-feeding of *E. coli* siderophores (22). Based on the BQH alone,
96 one might predict that the public availability of adenine and siderophores would enrich
97 for *E. coli* adenine auxotrophs and *R. palustris* siderophore mutants over time.
98
99 Here we examined long-term monocultures and cocultures for evidence of gene loss
100 due to the availability of non-essential adenine and possible siderophores from a
101 partner. In each case gene retention was correlated with low expression, suggesting
102 that repression of gene expression can decrease the frequency of gene loss.
103

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

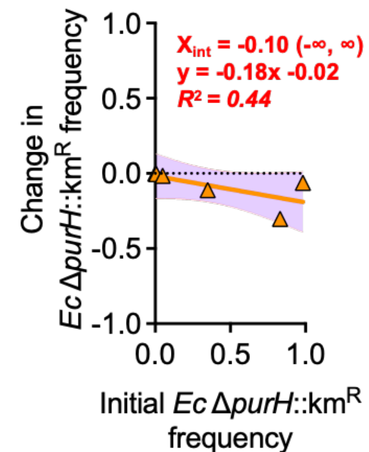
104 **RESULTS**
105
106 **Gene loss is not beneficial when adenine is provided to *E. coli* monocultures.** In
107 anaerobic cocultures where *R. palustris* provided essential nitrogen (NH_4^+) to *E. coli* and
108 *E. coli* provided essential carbon (organic acids) to *R. palustris* (19), we learned that *R.*
109 *palustris* also externalizes enough adenine to support an *E. coli ΔpurH* purine auxotroph
110 (20, 21). Herein, we recognized adenine availability as an opportunity to look for
111 evidence of the BQH; de novo purine synthesis costs more ATP than the purine salvage
112 pathway, and thus could lead to purine synthesis gene loss.
113
114 Fitness benefits have been observed for auxotrophic mutants, when the corresponding
115 nutrient is available either as a supplement or via cross-feeding (5-8). Thus, we first
116 tested whether loss of purine biosynthesis would be advantageous in the presence of
117 adenine by comparing the growth of *E. coli ΔpurH* versus wild-type (WT) PFM2 parent
118 cultures. We focused on *E. coli* PFM2 instead of the MG1655 (used previously by our
119 lab) because PFM2 does not carry the *rph* mutation that leads to slower growth in the
120 presence of adenine (23, 24), which can be alleviated by a *ΔpurH* mutant (Chuang and
121 McKinlay in prep). Despite the availability of adenine, the *ΔpurH* mutant did not grow
122 faster than the PFM2 parent in the presence of adenine (Fig 1A, B).
123



124
125 **Figure 1. Adenine availability does not provide a fitness advantage to an**
126 **engineered *E. coli* PFM2 purine auxotroph. A.** Growth of WT *E. coli* PFM2 and its
127 *ΔpurH* mutant in monoculture with and without 35 µM adenine (ade)**. B.** Monoculture
128 growth rates for WT PFM2 vs a corresponding *ΔpurH* mutant with and without 35 µM
129 adenine. 'a,' indicates statistically similar comparisons ($p > 0.01$) as determined using a
130 one-way ANOVA with Tukey's multiple comparisons test. **C.** Invasion-from-rare assay
131 competing a *ΔpurH* mutant against its PFM2 parent in coculture with *R. palustris*
132 CGA676 under N_2-fixing conditions where *R. palustris* excretes NH_4^+ and adenine. The
133 orange line is the best fit from a linear regression analysis with 95% CI shaded in
134 purple. Corresponding equations and x-intercepts (upper, lower bounds for 95% CI) are
135 shown. Change in frequency = (*E. coli* Δ*purH* / (*E. coli* WT + *E. coli* Δ*purH*))$_{final}$ – (*E. coli*
136 Δ*purH* / (*E. coli* WT + *E. coli* Δ*purH*))$_{initial}$. **A-C.** Each data point represents a
137 measurement from a single biological replicate.

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

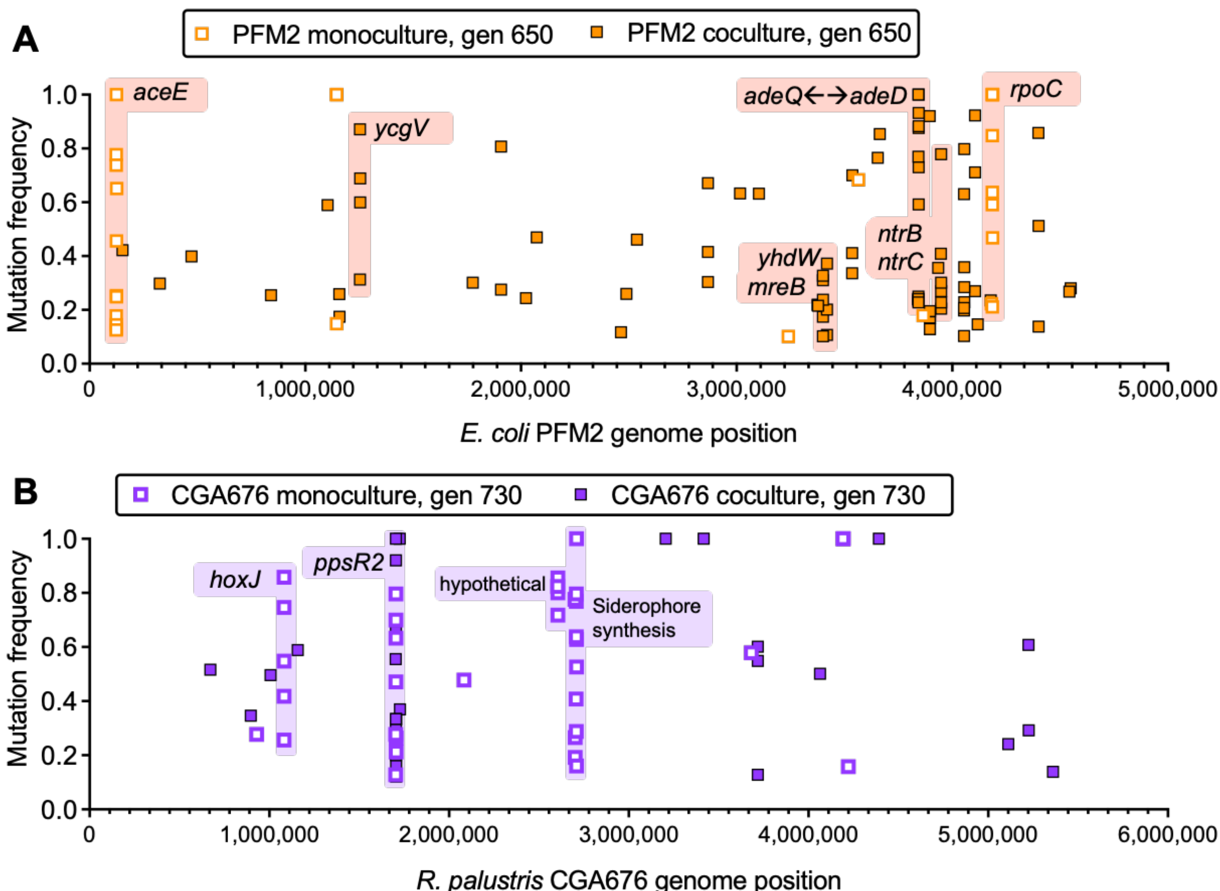
138 We also assessed the fitness impact of a *ΔpurH* mutation by competing the mutant
139 against the PFM2 parent in coculture with NH_4^+- and adenine-excreting *R. palustris*
140 NifA* (CGA676). We used a range of initial frequencies so that we could simultaneously
141 test for coexistence by mutual invasibility, where an equilibrium frequency can be
142 extrapolated from the x-intercept by linear regression analysis (25-27). However, there
143 was poor linear correlation in the invasion-from-rare assay (Fig 1C). The same was true
144 when a *ΔpurH*::km^R mutant was used (Fig S1). However, the *ΔpurH* mutant tended to
145 decrease in frequency relative to parent in cocultures, where adenine is available (Fig
146 1C and S1). Taken together, our results suggest that a *ΔpurH* mutant does not have a
147 competitive advantage over PFM2 in the presence of adenine. Thus, at least in this
148 engineered context, we do not expect that adenine availability to benefit *E. coli* PFM2
149 purine auxotrophs.
150



151 **Fig S1. Invasion-from-rare assay competing *E. coli***
152 ***ΔpurH*::km^R against its PFM2 parent in coculture with *R.***
153 ***palustris* CGA676.** The orange line is the best fit from a
154 linear regression analysis with 95% CI shaded in purple.
155 Corresponding equations and x-intercepts (upper, lower
156 bounds for 95% CI) are shown. Change in frequency = (*E.*
157 *coli* Δ*purH* / (*E. coli* WT + *E. coli* Δ*purH*))$_{final}$ – (*E. coli* Δ*purH*
158 / (*E. coli* WT + *E. coli* Δ*purH*))$_{initial}$. Each data point
159 represents a measurement from a single biological replicate.
160
161
162
163 **Adenine auxotrophs are rare in long-term cocultures.** Although the engineered
164 *ΔpurH* mutant did not have an obvious fitness advantage, it is possible that (i) there was
165 a subtle difference in growth rate that could be enriched over many generations and (ii)
166 other kinds of purine auxotrophy mutations could lead to a fitness advantage. To
167 account for these possibilities, we looked for the evidence of purine auxotrophs in long-
168 term monocultures and cocultures of *E. coli* PFM2 and NH_4^+- and adenine-excreting *R.*
169 *palustris* (CGA676; NifA*) that we had maintained for other reasons (Fig 2A; Table S1-
170 5). Adenine is amply available even in evolved cocultures (Chuang and McKinlay in
171 prep). Focusing on the latest common timepoint between monoculture vs coculture
172 treatments (generation 650), we did not observe obvious mutations in purine synthesis
173 genes in PFM2 in cocultures or monocultures (Fig 2B; Fig S2). When considering other
174 time points, only one purine biosynthesis gene mutation across was observed in one
175 monoculture line at generation 204 (*purT* at 54.3%, line 15; Table S2).
176
177 To account for non-obvious mutations that might lead to purine auxotrophy. We also
178 screened four *E. coli* isolates from each monoculture and coculture line (40 monoculture
179 isolates and 40 coculture isolates) for adenine auxotrophy in liquid cultures in 96-well
180 plates (Fig 2C). Although several monoculture isolates could not grow in minimal media,
181 suggesting that dependencies might have developed between *E. coli* subpopulations,
182 only two monoculture isolates were possibly adenine auxotrophs (line 20; Fig 2C). We
183 suspect that the other auxotrophs are due to mutations in *rpoC*, which were

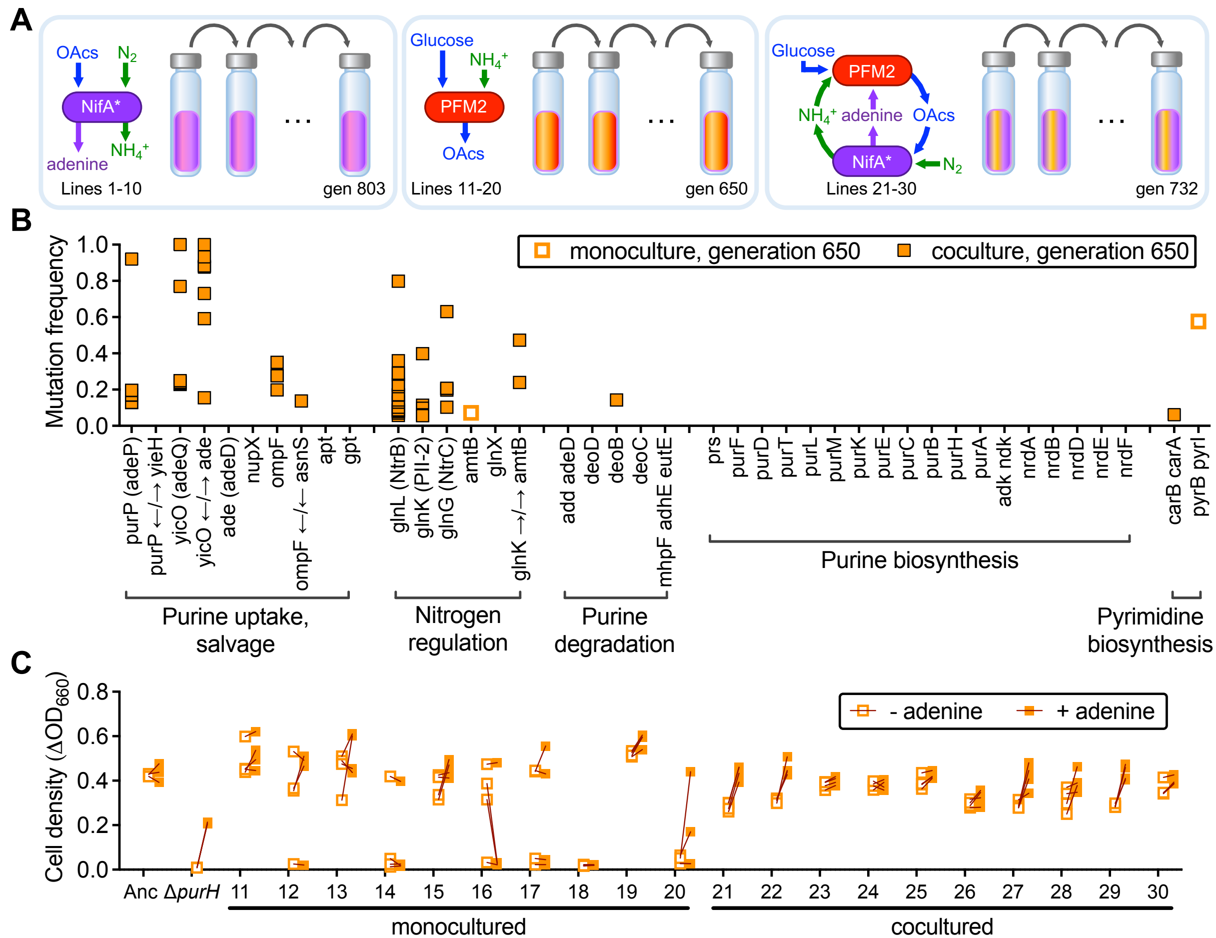
bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

184 conditionally enriched in monoculture (Fig S2A). Certain mutations in *E. coli rpoC* are
185 known to generate polyauxotrophies (28, 29).
186



187
188
189 **Fig S2. Mutations enriched in *E. coli* PFM2 (A) and *R. palustris* CGA676 (B) in**
190 **long-term monocultures and cocultures.** Each data point represents a mutation
191 observed at a frequency >0.1 in an experimental line. Generation: gen.
192
193 No adenine auxotrophs were isolated from cocultures and all coculture isolates grew in
194 defined media, suggesting that our coculture conditions might not favor emergence of
195 other dependencies between subpopulations. Overall, we conclude that *E. coli* adenine
196 auxotrophs are not a major subpopulation in any long term coculture despite the
197 availability of adenine.

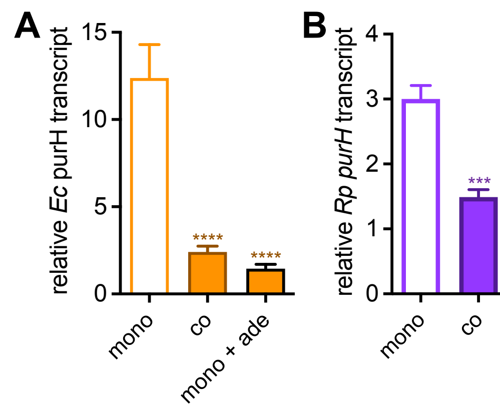
bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



198
199
200 **Fig 2. *E. coli* adenine auxotrophs are not prevalent in long-term cocultures. A.**
201 Long-term monoculture and coculture conditions. OAcs, organic acids. **B.** Enriched *E.*
202 *coli* mutations in genes for nucleobase metabolism and regulation of nitrogen
203 metabolism for long-term monocultures and cocultures. Each point represents a
204 mutation frequency in a given gene for a given evolutionary line. **C.** Random evolved *E.*
205 *coli* isolates (generation 650; n=4) were screened for adenine auxotrophy in liquid
206 monoculture conditions with and without 50 μM adenine. Ancestral strain (Anc) and
207 *ΔpurH* (*Δ*H) mutant values are also included (n=3). Each point represents a single
208 measurement for a single isolate (replicate measurements were not made for any
209 isolate). Each line connects measurements for the same isolate grown with and without
210 adenine.
211
212 **Cross-fed adenine represses purine synthesis gene expression.** One possible
213 explanation for why *E. coli* purine auxotrophs did not emerge in coculture is because
214 adenine availability repressed gene expression, thus lowering the purine biosynthesis
215 gene costs. In support of this notion, we previously saw that *E. coli* MG1655 down-
216 regulated several purine synthesis genes in coculture versus monoculture (22). To
217 determine if the same is true for PFM2, we quantified *purH* transcripts by RT-qPCR.
218 Coculture PFM2 *purH* levels were 20% of those in monoculture (Fig 3A). Adding
219 adenine to monocultures resulted in a *purH* transcript level that was 12% of that

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

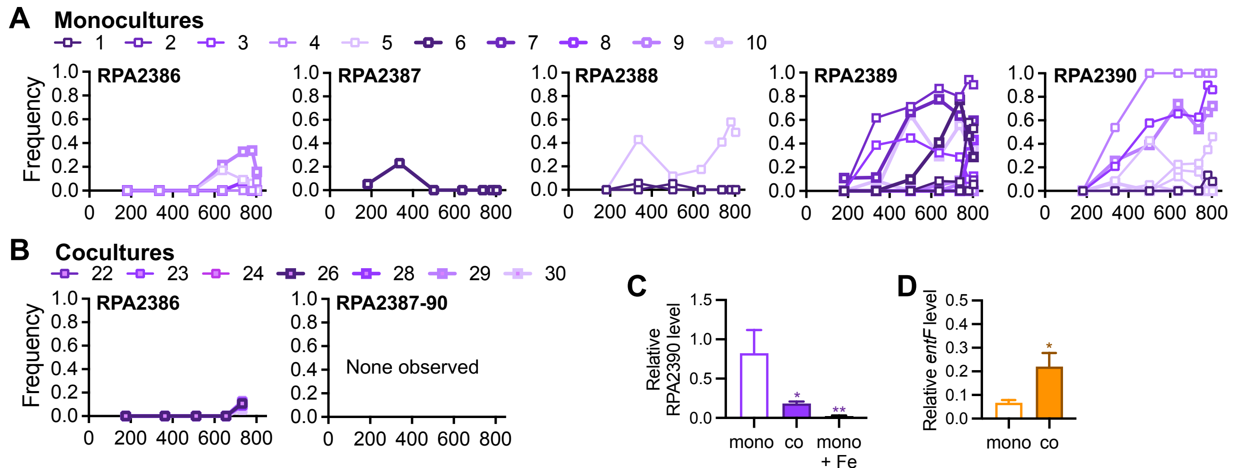
220 observed without adenine, suggesting that the low expression in coculture was due to
221 adenine availability (Fig 3A).
222
223 *R. palustris* CGA676 also showed less *purH* transcript in coculture, 50% of that
224 observed in monoculture (Fig 3B). It is unclear how *E. coli* would influence repression of
225 *R. palustris purH* expression. However, the lower expression supports previous findings
226 that *E. coli* does not stimulate *R. palustris* adenine production (20). Adenine availability
227 and expression of purine biosynthesis genes curiously did not lead to purine
228 biosynthesis gene mutations in *R. palustris* monocultures. It is possible that *R. palustris*
229 lacks effective adenine uptake mechanisms. In a separate study, we found that adenine
230 can be toxic to *R. palustris*, and purine externalizing strains, including the parent of
231 CGA676 are more resistant to exogenous adenine (Chuang and McKinlay, in prep).
232 Thus, CGA676 might be pre-dispositioned for adenine externalization rather than
233 uptake.
234



235 **Fig 3. Purine biosynthesis genes are down-**
236 **regulated in coculture.** RT-qPCR
237 quantification of *purH* transcript levels in *E. coli*
238 PFM2 (**A**; relative to *hcaT*) and *R. palustris*
239 CGA676 (**B**; relative to *fixJ*). Adenine (+ ade)
240 was added to a final concentration of 0.1 mM to
241 ensure that it was not used up during culture
242 growth. Values are the mean ± SD, n = 3-4.
243 Statistically significant differences from the
244 monoculture condition for each strain were
245 determined using an unpaired two-tail t-test;
246 ***, $p < 0.001$; ****, $p < 0.0001$.
247
248 ***R. palustris* siderophore gene loss is correlated with expression.** We looked for
249 other correlations of gene retention and low gene expression in our long-term cultures.
250 A standout was *R. palustris* siderophore genes which accumulated mutations in
251 monoculture but not coculture (Fig 4A, B; Fig S2). We previously saw that *R. palustris*
252 down-regulated siderophore synthesis genes in coculture with *E. coli* MG1655, relative
253 to monocultures (22). We verified that this trend was also true for *R. palustris* in
254 coculture with PFM2; RT-qPCR analysis showed that the transcript levels for the
255 siderophore synthesis gene RPA2390 was 20% in coculture compared to monoculture
256 (Fig 4C). Adding soluble iron (ammonium ferric citrate) to monocultures decreased
257 RPA2390 expression to 3% of that observed without added iron, suggesting that the low
258 expression in coculture was a response to enhanced iron availability in coculture (Fig
259 4C).
260
261 *R. palustris* might down-regulate siderophore production in coculture if *E. coli* facilitates
262 iron acquisition. Others have speculated that *R. palustris* can use siderophores from
263 other bacteria, because its genome encodes multiple transporters but only one
264 siderophore synthesis cluster for petrobactin-like siderophores (30, 31); *E. coli* produces
265 enterobactin siderophores (32). In agreement with this speculation, *E. coli* siderophore

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

266 gene expression seemed to respond to iron loss to *R. palustris*; the *entF* transcript level
267 was 3.3-fold higher in coculture than in monoculture (Fig. 4D). These observations
268 support a notion that *R. palustris* can use *E. coli* siderophores, however the actual
269 molecular nature of this relationship is pending a thorough investigation.
270
271 Despite the increase in *E. coli entF* expression in coculture, *E. coli* siderophore mutants
272 were rare and inconsistent across coculture lines and were not observed in long-term
273 monocultures (Table S2, S4). It is possible that this expression range is not enough to
274 pose a fitness cost, perhaps due to a low iron requirement under fermentative
275 conditions (e.g., little use for iron-containing cytochromes). In agreement with this
276 notion, transposon insertions in *E. coli* MG1655 siderophore genes had mainly neutral
277 effects on fitness in monocultures and cocultures (21).



278
279
280 **Fig 4. Siderophore gene loss is prevalent in *R. palustris* monocultures where**
281 **gene expression is high. A, B.** *R. palustris* CGA676 mutation frequencies in
282 siderophore synthesis genes for each monoculture (**A**) or coculture line (**B**) where
283 mutations were observed (no mutations were observed in lines 21, 25, and 27). Repeat
284 colors in a given graph indicate different mutations in the same gene. **C, D.** RT-qPCR
285 quantification of siderophore synthesis genes in *R. palustris* (**C**; RPA2390 relative to
286 *fixJ*) and in *E. coli* (**D**; *entF* relative to *hcaT*). Values are the mean ± SD, n = 3.
287 Statistically significant differences from the monoculture (mono) condition for each strain
288 were determined using an unpaired two-tail t-test; *, $p < 0.05$; **, $p < 0.01$; co, coculture,
289 Fe; ammonium ferric citrate.
290
291 **Discussion**
292
293 In long-term cocultures, the availability of cross-fed adenine, and possibly siderophores,
294 did not result in corresponding gene loss in a recipient. Our data suggest that the
295 externalized resource repressed recipient gene expression, thereby lowering gene cost
296 and selective pressure for gene loss. Conversely, siderophore loss-of-function
297 mutations accumulated in *R. palustris* monocultures where gene expression was high.
298 Although there are likely situations where cross-feeding promotes gene loss, there are

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

299 also likely situations where cross-feeding instead promotes gene retention through
300 repressing gene expression.
301
302 Our results do not counter the BQH, which posits that gene loss will occur if the benefit
303 of acquiring resources from a neighbor outweighs the cost of gene retention (1). If the
304 cost of gene retention is low, then gene loss would only have neutral fitness effects,
305 rather than beneficial effects. Others have suggested that low gene expression is not
306 enough to promote gene retention because the maintenance and activity of sensor and
307 regulatory proteins can still drive of gene loss (3). However, the fitness benefits of
308 mutants lacking biosynthetic genes suggests that benefits can be realized while
309 regulatory circuits remain intact (5-9). The association between gene cost and gene
310 expression is well-documented (10-16). Our data, and data from others, suggest that
311 there are cases where resource availability can lower gene cost simply by repressing
312 gene expression. For example, repression of an antibiotic resistance gene in *E. coli*
313 eliminated the fitness cost of that gene (10). In another example closer to the conditions
314 of the current study, iron-limitation that prompted high siderophore production by *P.*
315 *aeruginosa* led to a higher frequency of siderophore-deficient mutants compared to iron-
316 rich conditions where siderophore production was low (11). The *P. aeruginosa* results
317 were presented in the context of cheating, where gene loss is associated with
318 exploitation that harms the producer, unlike BQH beneficiaries that should not harm the
319 producer (1). Yet, the conditions leading to the emergence of cheaters and BQH
320 beneficiaries are analogous (1). As such, many lessons from literature on cheaters can
321 apply to BQH scenarios.
322
323 Insights into BQH gene-loss can also come from resource-rich monocultures, with the
324 caveat that nutrient access might be well above that from a cross-feeding partner. For
325 example, repression of gene expression in long-term nutrient-rich *E. coli* monocultures
326 could explain why the loss of some catabolic genes was not observed in parallel lines
327 and occurred without fitness effects (33). Repression of gene expression will shift the
328 cost of a gene towards neutral. Gene loss with neutral fitness effects can still occur, but
329 enrichment of those mutations would require hitchhiking with a beneficial mutation or
330 passage through a severe population bottleneck. Cost-neutral gene loss, as opposed to
331 beneficial gene loss, might also require time, and possibly spatial structure, that were
332 not part of our experimental design. Should cost-neutral auxotrophic mutants emerge in
333 nature, the chances of rescue via cross-feeding from a neighbor could be high (34, 35).
334
335 The extent to which gene repression can promote gene retention will also depend on
336 the level of expression strength of repression; as noted above, relatively low expression
337 might explain the lack of *E. coli* siderophore gene mutations. Gene expression can be
338 noisy, and that noise can be costly and subject to selective optimization in a stable
339 nutrient regime (36, 37). Noisy gene expression might explain why some engineered
340 auxotrophs have greater fitness over a prototrophic parent (5-9). In long-term *E. coli*
341 monocultures, loss-of-function mutations for maltose catabolism were arguably
342 associated with weak gene repression in the presence of glucose (38). During growth
343 on glucose, ancestral expression of the maltoporin was half that compared to during
344 growth on maltose. Several evolved isolates carried mutations that prevented growth on

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

345 maltose. However, in an evolved isolate that retained the ability to grow on maltose,
346 maltoporin expression with glucose was 5% of that with maltose (38), suggesting that
347 stronger repression contributed to gene retention.
348
349 The level of nutrient availability also likely affects the level of gene repression. In other
350 words, there might be a resource concentration range over which BQH gene loss would
351 occur. For example, *Acinetobacter baylyi* histidine autotrophs had a fitness advantage
352 over a prototrophic parent when there was $\leq$ 50 μM histidine, but the advantage was
353 reversed as histidine levels reached 100-200 μM (6), possibly due to stronger gene
354 repression in the parent, though this is speculation on our part. One might also consider
355 that the responses of a prototrophic strain versus an otherwise clonal auxotroph to an
356 extracellular nutrient can vary widely; whereas an auxotroph would benefit from the
357 nutrient, the prototroph might experience beneficial, neutral or even detrimental fitness
358 effects (6). The extent to which loss of a biosynthetic gene can have compounding
359 fitness effects by affecting expression of other genes deserves investigation.
360
361 Our work presents a caveat to the popular expectation of cross-feeding as a driver of
362 gene loss. Cross-feeding can instead promote gene retention when resource availability
363 is high enough to repress gene expression and thus lower the cost of a gene. The
364 likelihood of gene loss then becomes the same as any other gene with a neutral cost.
365

366 MATERIALS AND METHODS

367 **Bacterial strains.** *R. palustris* CGA676 is derived from CGA0092 (39) and carries a
368 *nifA** mutation that causes NH_4^+ excretion under N_2-fixing conditions (19, 40). *E. coli*
369 PFM2 (41, 42) Δ*purH*::km^R was made via lambda Red recombination (43) using
370 constructs amplified from KEIO mutants (44) using primers YCC29 (GCG CAA ACG
371 TTT TCG TTA CAA TGC) and YCC30 (TGC ATT ACC CGG AGC AAC). FLP-mediated
372 excision was used to remove the kanamycin resistance cassette (km^R) to generate the
373 Δ*purH* strain (43).
374
375 **Growth conditions.** Anoxic media in test tubes were prepared by bubbling N_2 through
376 10 ml of media in 27-ml anaerobic test tubes, then sealing with rubber stoppers and
377 aluminum crimps prior to autoclaving. Monocultures and cocultures were grown
378 horizontally at 30°C with light from a 45 W halogen bulb (430 lumens) and shaking at
379 150 rpm in minimal M9-derived coculture medium (MDC) (19) with either (i) *E. coli*
380 monocultures: 25 mM glucose, 10 mM NH_4Cl, and cation solution (100X stock: 100 mM
381 $MgSO_4$ and 10 mM $CaCl_2$); *R. palustris* monocultures: 20 mM sodium acetate and 10
382 mM NH_4Cl; or cocultures: 25 mM glucose and cation solution. Cultures with plasmid-
383 carrying strains were also supplemented with 100 μg/ml gentamycin or 25 μg/ml
384 chloramphenicol as appropriate. Starter cultures were inoculated with single colonies.
385 *R. palustris* starter cultures were grown in MDC with 20 mM acetate and 10 mM NH_4Cl.
386 *E. coli* starter cultures were grown aerobically in lysogeny broth, with 30 μg/ml
387 kanamycin (km) when appropriate. *E. coli* starter cultures were washed twice in 1 ml
388 MDC prior to inoculating test cultures or bioassays. Cocultures were inoculated with 0.1
389 ml each of *R. palustris* and *E. coli* to an initial optical density (OD_{660}) of ~0.003 each.
390

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

391 Cultures in 96-well plates were treated similarly except that oxic stock solutions were
392 used to prepare 0.2 ml volumes in each well. Anoxic conditions were achieved by
393 sealing plates inside a BD GasPak EZ large incubation container with 2 anaerobic
394 sachets.
395
396 **Experimental evolution.** Founder monocultures of *E. coli* PFM2 and *R. palustris* NifA*
397 CGA676 were grown from a single colony in anoxic MDC with either 25 mM glucose,
398 cation solution, and 3 mM NH_4Cl for PFM2 or 20 mM sodium acetate for CGA676. A
399 single founder monoculture was then used to inoculate 10 monocultures and 10
400 cocultures. All cultures were grown horizontally without shaking at 30°C with light in
401 MDC. PFM2 monocultures were supplemented with 10 mM glucose, cation solution, 25
402 mM NaCl, and 2.3 mM NH_4Cl. CGA676 monocultures were supplemented with 25 mM
403 glucose, cation solution, 8 mM disodium succinate, 7.3 mM sodium acetate, 0.25 mM
404 sodium formate, 1.4 mM sodium lactate, and 6.3 mM ethanol. Cocultures were
405 supplemented with 25 mM glucose, cation solution, and 25 mM NaCl. Every 7 days,
406 cultures were vortexed and 0.25 ml was transferred to fresh medium. About every 5
407 transfers, stocks were frozen in 25% glycerol at -80°C and separate cell pellets from 1
408 ml samples were frozen for gDNA extraction.
409
410 **Analytical procedures.** Cell densities were measured via turbidity (OD_{660}) using a
411 Genesys 20 visible spectrophotometer (Thermo-Fisher).
412
413 **Competition assays.** Competition assays were conducted in an invasion-from-rare
414 format to consider coexistence by the mutual invasion criterion, where each population
415 can increase when rare (25, 27). Cocultures were started from various initial
416 frequencies (targeting 0.01 – 0.99) of each *E. coli* strain (WT vs Δ*purH* or WT vs
417 Δ*purH*::km^R) for a total initial cell density of ~10^6 colony forming units (CFU) / ml. *R.*
418 *palustris* CGA676 was inoculated to an initial density of ~10^6 CFU / ml. Frequencies
419 were determined upon inoculation and after 5 days. WT and Δ*purH* mutants were
420 distinguished by counting CFUs on M9 agar with cations and 25 mM glucose and on LB
421 agar and then determining Δ*purH* populations from the difference. When Δ*purH*::km^R
422 mutants were used LB agar included km to allow for direct determination of population
423 size. Change in frequency = (*E. coli* Δ*purH* / (*E. coli* WT + *E. coli* Δ*purH*))$_{final}$ – (*E. coli*
424 Δ*purH* / (*E. coli* WT + *E. coli* Δ*purH*))$_{initial}$ (26).
425
426 **Genome sequencing and mutation analysis.** gDNA was purified from cells using a
427 Qiagen DNeasy Blood and Tissue kit following the manufacturer's instructions. Lysis
428 was facilitated after resuspension by adding proteinase K (50 µg/ml final), and
429 incubating at 56°C for 10 min. RNaseA (4 µl, Promega) was then added and the lysate
430 was incubated for 2 min before proceeding.
431
432 DNA fragment libraries were made using a NextFlex Bioo Rapid DNA kit and libraries
433 were sequenced using Illumina NextSeq 500 150x150 paired-end runs by the IU Center
434 for Genomics and Bioinformatics. Paired-end reads were pre-processed for quality with
435 cutadapt 3.4 (45) with the following options: -a AGATCGGAAGAGC -A
436 AGATCGGAAGAGC ; -q 15,10; -u 6. Mutations were called using breseq v. 0.32.0 on

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

437 polymorphism mode (46). *E. coli* monoculture population sequences were mapped to
438 the MG1655 genome (accession NC_000913); *R. palustris* monoculture population
439 sequences were mapped to a concatenated reference genome consisting of the
440 CGA009 chromosome (accession BX571963), and its plasmid pRPA (accession
441 BX571964). Co-culture sequences were mapped to a concatenation of the *E. coli* and
442 *R. palustris* reference genomes. Polymorphisms that co-occurred in both the
443 monoculture and co-culture datasets were filtered and maintained as a subset to enrich
444 for the most informative variants representing treatment differences (Table S1-4). This
445 filtering step also removed sequence differences between the reference sequences and
446 those of the experimental strains used. Variants were prioritized as mutations of interest
447 if they were detectable at the final two sequencing timepoints and co-occurred across
448 multiple populations in the same locus. All mutations can be found in Tables S1-4.
449 Locus tag conversions can be found in Table S5.
450
451 **Reverse transcription quantitative real-time PCR (RT-qPCR).** Cultures received 100
452 µM adenine or ammonium ferric [iron(III)] citrate as indicated. Cultures were harvested
453 in exponential phase at 0.6-0.8 OD_{660} except for *E. coli* monocultures (+/- adenine
454 experiment), which were harvested at 0.3-0.4 OD_{660} to avoid adenine depletion.
455 Cultures were chilled on ice and pelleted by centrifugation. Lysis, RNA purification, and
456 cDNA generation were performed exactly as described (20). Standard curves were
457 generated using gDNA. Transcripts were quantified as described (20) using the
458 appropriate primers (Table S6) with a Mastercycler ep *realplex* real-time PCR system
459 (Eppendorf). Data was analyzed by *realplex* software using Noiseband. Specificities
460 were validated by melting curves and by the presence of a single band on an agarose
461 gel.
462
463 **Statistical analyses.** Statistical analyses of growth and RT-qPCR data were performed
464 using Graphpad Prism (v10).
465
466 **ACKNOWLEDGEMENTS**
467
468 This work was supported in part by US Army Research Office grants W911NF-14-1-
469 0411 and W911NF-17-1-0159, the National Science Foundation CAREER award MCB-
470 1749489 to JBM, and National Institutes of Health grants R35GM128674 to ABD and
471 R35GM150625 to MB. Supercomputing resources were supported in part by Lilly
472 Endowment, Inc., through its support for the IU Pervasive Technology Institute.
473
474 We are grateful to J. Gliessman, A. Cairo, S. Kamaran, J. Mazny, and N. Ramli, for
475 maintaining long-term cultures. We thank P. Foster for PFM2 and helpful advice. We
476 also thank A. Drummond, J. Drummond, J.P. Gerdt, J. Lennon, M. Lynch, T. Romeo,
477 and the McKinlay lab for helpful advice.
478
479 **References**
480
481 1. J. J. Morris, R. E. Lenski, E. R. Zinser, The Black Queen Hypothesis: Evolution of
482 dependencies through adaptive gene loss. *mBio* **3**, e00036-00012 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

483 2. M. Ghoul, A. S. Griffin, S. A. West, Toward an evolutionary definition of cheating.
484 *Evolution* **68**, 318-331 (2014).
485 3. G. D'Souza *et al.*, Ecology and evolution of metabolic cross-feeding interactions in
486 bacteria. *Nat Prod Rep* **35**, 455-488 (2018).
487 4. S. Pande, C. Kost, Bacterial unculturability and the formation of intercellular metabolic
488 networks. *Trends Microbiol.* **25**, 349-361 (2017).
489 5. S. Zamenhof, H. H. Eichhorn, Study of microbial evolution through loss of biosynthetic
490 functions: establishment of "defective" mutants. *Nature* **216**, 456-458 (1967).
491 6. G. D'Souza *et al.*, Less is More: Selective advantages can explain the prevalent loss of
492 biosynthetic genes in bacteria. *Evolution* **68**, 2559-2570 (2014).
493 7. D. Dykhuizen, Selection for tryptophan auxotrophs of *Escherichia coli* in glucose-limited
494 chemostats as a test of the energy conservation hypothesis of evolution. *Evolution* **32**,
495 125-150 (1978).
496 8. S. Pande *et al.*, Fitness and stability of obligate cross-feeding interactions that emerge
497 upon gene loss in bacteria. *ISME J* **8**, 953-962 (2013).
498 9. G. D'Souza, C. Kost, Experimental evolution of metabolic dependency in bacteria. *PLoS*
499 *Genet.* **12**, e1006364 (2016).
500 10. T. N. Nguyen, Q. G. Phan, L. P. Duong, K. P. Bertrand, R. E. Lenski, Effects of carriage
501 and expression of the Tn10 tetracycline-resistance operon on the fitness of *Escherichia*
502 *coli* K12. *Mol. Biol. Evol.* **6**, 213-225 (1989).
503 11. Z. Dumas, R. Kümmerli, Cost of cooperation rules selection for cheats in bacterial
504 metapopulations. *J. Evol. Biol.* **25**, 473-484 (2012).
505 12. F. C. Neidhardt, J. L. Ingraham, M. Schaechter, *Physiology of the bacterial cell: a*
506 *molecular approach* (Sinauer Associates Inc, Sunderland, MA, 1990).
507 13. M. Lynch, G. K. Marinov, The bioenergetic costs of a gene. *Proc Natl Acad Sci U S A*
508 **112**, 15690-15695 (2015).
509 14. D. M. Stoebel, A. M. Dean, D. E. Dykhuizen, The cost of expression of Escherichia coli
510 lac operon proteins is in the process, not in the products. *Genetics* **178**, 1653-1660
511 (2008).
512 15. E. Dekel, U. Alon, Optimality and evolutionary tuning of the expression level of a
513 protein. *Nature* **436**, 588-592 (2005).
514 16. Z. Wu *et al.*, Expression level is a major modifier of the fitness landscape of a protein
515 coding gene. *Nat Ecol Evol* **6**, 103-115 (2022).
516 17. F. Harrison, J. Paul, R. C. Massey, A. Buckling, Interspecific competition and
517 siderophore-mediated cooperation in *Pseudomonas aeruginosa*. *ISME J* **2**, 49-55 (2008).
518 18. A. R. T. Figueiredo, A. Wagner, R. Kümmerli, Ecology drives the evolution of diverse
519 social strategies in *Pseudomonas aeruginosa*. *Mol. Ecol.* **30**, 5214-5228 (2021).
520 19. B. LaSarre, A. L. McCully, J. T. Lennon, J. B. McKinlay, Microbial mutualism dynamics
521 governed by dose-dependent toxicity of cross-fed nutrients. *ISME J* **11**, 337-348 (2017).
522 20. Y.-C. Chuang *et al.*, Bacterial adenine cross-feeding stems from a purine salvage
523 bottleneck *ISME J* **18**, wrae034 (2024).
524 21. B. LaSarre, A. M. Deutschbauer, C. E. Love, J. B. McKinlay, Covert cross-feeding
525 revealed by genome-wide analysis of fitness determinants in a synthetic bacterial
526 mutualism. *Appl Environ Microbiol* **86**, e00543-00520 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

527 22. A. L. McCully *et al.*, An *Escherichia coli* nitrogen starvation response is important for
528 mutualistic coexistence with *Rhodopseudomonas palustris*. *Appl Environ Microbiol* **84**,
529 e00404-00418 (2018).
530 23. K. F. Jensen, The *Escherichia coli* K-12 "wild types" W3110 and MG1655 have an rph
531 frameshift mutation that leads to pyrimidine starvation due to low pyrE expression levels.
532 *J. Bacteriol.* **175**, 3401-3407 (1993).
533 24. A. S. Bagnara, L. R. Finch, The effects of bases and nucleosides on the intracellular
534 contents of nucleotides and 5-phosphoribosyl 1-pyrophosphate in *Escherichia coli*. *Eur.*
535 *J. Biochem.* **41**, 421-430 (1974).
536 25. T. N. Grainger, J. M. Levine, B. Gilbert, The invasion criterion: a common currency for
537 ecological research. *Trends Ecol. Evol.* **34**, 925-935 (2019).
538 26. S. P. Hammarlund, J. M. Chacón, W. R. Harcombe, A shared limiting resource leads to
539 competitive exclusion in a cross-feeding system. *Environ. Microbiol.* **21**, 759-771 (2019).
540 27. P. Chesson, Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* **31**,
541 343-366 (2000).
542 28. D. Satory, J. A. Halliday, P. Sivaramakrishnan, R. C. Lua, C. Herman, Characterization
543 of a novel RNA polymerase mutant that alters DksA activity. *J. Bacteriol.* **195**, 4187-
544 4194 (2013).
545 29. L. Brown, D. Gentry, T. Elliott, M. Cashel, DksA affects ppGpp induction of RpoS at a
546 translational level. *J. Bacteriol.* **184**, 4455-4465 (2002).
547 30. F. W. Larimer *et al.*, Complete genome sequence of the metabolically versatile
548 photosynthetic bacterium *Rhodopseudomonas palustris*. *Nat. Biotechnol.* **22**, 55-61
549 (2004).
550 31. O. Baars, F. M. M. Morel, X. Zhang, The purple non-sulfur bacterium
551 *Rhodopseudomonas palustris* produces novel petrobactin-related siderophores under
552 aerobic and anaerobic conditions. *Environ. Microbiol.* **20**, 1667-1676 (2018).
553 32. L. Cavas, I. Kirkiz, Characterization of siderophores from E*scherichia coli* strains
554 through genome mining tools: an antiSMASH study. *AMB Express* **12**, 74 (2022).
555 33. N. Leiby, C. J. Marx, Metabolic erosion primarily through mutation accumulation, and
556 not tradeoffs, drives limited evolution of substrate specificity in *Escherichia coli*. *PLoS*
557 *Biol.* **12**, e1001789 (2014).
558 34. S. Giri, G. Yousif, S. Shitut, L. Oña, C. Kost, Prevalent emergence of reciprocity among
559 cross-feeding bacteria. *ISME Commun* **2**, 1-7 (2022).
560 35. S. Giri *et al.*, Metabolic dissimilarity determines the establishment of cross-feeding
561 interactions in bacteria. *Curr. Biol.* **31**, 5547-5557 e5546 (2021).
562 36. J. B. McKinlay, G. M. Cook, K. Hards, Microbial energy management-A product of three
563 broad tradeoffs. *Adv. Microb. Physiol.* **77**, 139-185 (2020).
564 37. E. J. O'Brien, J. Utrilla, B. O. Palsson, Quantification and classification of *E. coli*
565 proteome utilization and unused protein costs across environments. *PLoS Comput Biol*
566 **12**, e1004998 (2016).
567 38. L. Pelosi *et al.*, Parallel changes in global protein profiles during long-term experimental
568 evolution in *Escherichia coli*. *Genetics* **173**, 1851-1869 (2006).
569 39. B. E. Mazny, O. F. Sheff, B. LaSarre, A. McKinlay, J. B. McKinlay, Complete genome
570 sequence of *Rhodopseudomonas palustris* CGA0092 and corrections to the *R. palustris*
571 CGA009 genome sequence. *Microbiol Resour Announc* **12**, e0128522 (2023).

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

572 40. J. B. McKinlay, C. S. Harwood, Carbon dioxide fixation as a central redox cofactor
573 recycling mechanism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11669-11675 (2010).
574 41. P. L. Foster, H. Lee, E. Popodi, J. P. Townes, H. Tang, Determinants of spontaneous
575 mutation in the bacterium *Escherichia coli* as revealed by whole-genome sequencing.
576 *Proc Natl Acad Sci U S A* **112**, E5990-5999 (2015).
577 42. H. Lee, E. Popodi, H. Tang, P. L. Foster, Rate and molecular spectrum of spontaneous
578 mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing.
579 *Proc Natl Acad Sci U S A* **109**, E2774-2783 (2012).
580 43. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in
581 *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640-6645
582 (2000).
583 44. T. Baba *et al.*, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout
584 mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
585 45. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads.
586 *EMBnet j* **17** (2011).
587 46. D. E. Deatherage, J. E. Barrick, "Identification of Mutations in Laboratory-Evolved
588 Microbes from Next-Generation Sequencing Data Using breseq" in Engineering and
589 Analyzing Multicellular Systems: Methods and Protocols, L. Sun, W. Shou, Eds.
590 (Springer New York, New York, NY, 2014), 10.1007/978-1-4939-0554-6_12, pp. 165-
591 188.
592

bioRxiv preprint doi: https://doi.org/10.1101/2024.08.19.608702; this version posted August 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

593 **Table S6. Primers used in this study**

Primer	Sequence (5'-3')	Description
YCC29	GCG CAA ACG TTT TCG TTA CAA TGC G	5' of Δ*purH*::*Km* in JW3970
YCC30	TGC ATT ACC CGG AGC AAC	3' of Δ*purH*::*Km* in JW3970
YCC80	TAA GGG AAC CGT GCA TGT G	Forward qPCR primer for *Rp fixJ*
YCC81	GGA TTC GTA CAG CTT GAC CTC	Reverse qPCR primer for *Rp fixJ*
YCC97	ACG TCG TCG TGG TTC TTG	Forward qPCR primer for *Rp purH*
YCC98	CGA AGC CAC CGT CGA TAA A	Reverse qPCR primer for *Rp purH*
YCC99	GCA ACA CGT TCT GCT GAT G	Forward qPCR primer for *Rp RPA2390*
YCC100	CAT TGG TTC TCG GCC TAT CT	Reverse qPCR primer for *Rp RPA2390*
YCC91	AAC CGC ATG GCC CTT ATT	Forward qPCR primer for *Ec entF*
YCC92	GTA TCC AGC AAG CCA AGA AAT G	Reverse qPCR primer for *Ec entF*
YCC93	GTG TTC GAA GGC TTT GAT GG	Forward qPCR primer for *Ec purH*
YCC94	GTG AAG AGC AGC GAC TAT GA	Reverse qPCR primer for *Ec purH*
YCC95	GTC GCA CTT TGC CGT AAT C	Forward qPCR primer for *Ec hcaT*
YCC96	GCT GAT GCT GGT GAT GAT TG	Reverse qPCR primer for *Ec hcaT*

595