

# 1 **Single mutation makes *Escherichia coli* an insect mutualist**

2

3 Ryuichi Koga<sup>1\*</sup>, Minoru Moriyama<sup>1</sup>, Naoko Onodera-Tanifuji<sup>1</sup>, Yoshiko Ishii<sup>1</sup>, Hiroki Takai<sup>1</sup>,  
4 Masaki Mizutani<sup>1</sup>, Kohei Oguchi<sup>1</sup>, Reiko Okura<sup>2</sup>, Shingo Suzuki<sup>3</sup>, Yasuhiro Goto<sup>4</sup>, Tetsuya  
5 Hayashi<sup>4</sup>, Masahide Seki<sup>5</sup>, Yutaka Suzuki<sup>5</sup>, Yudai Nishide<sup>1,6</sup>, Takahiro Hosokawa<sup>7</sup>, Yuichi  
6 Wakamoto<sup>2,8</sup>, Chikara Furusawa<sup>3,8</sup>, Takema Fukatsu<sup>1,9,10\*</sup>

7

8 <sup>1</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and  
9 Technology (AIST), Tsukuba, Japan.

10 <sup>2</sup>Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo,  
11 Tokyo, Japan.

12 <sup>3</sup>Center for Biosystem Dynamics Research, RIKEN, Osaka, Japan.

13 <sup>4</sup>Department of Bacteriology, Faculty of Medical Sciences, Kyushu University, Fukuoka, Japan.

14 <sup>5</sup>Laboratory of Systems Genomics, Department of Computational Biology and Medical  
15 Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan.

16 <sup>6</sup>National Agriculture and Food Research Organization (NARO), Institute of Agrobiological  
17 Sciences, Tsukuba, Japan.

18 <sup>7</sup>Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan.

19 <sup>8</sup>Universal Biology Institute, The University of Tokyo, Tokyo, Japan.

20 <sup>9</sup>Department of Biological Sciences, The University of Tokyo, Tokyo, Japan.

21 <sup>10</sup>Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan.

22

23 \*Corresponding authors. Email: [r-koga@aist.go.jp](mailto:r-koga@aist.go.jp), [t-fukatsu@aist.go.jp](mailto:t-fukatsu@aist.go.jp)

24

25

26 **Abstract**

27 We report an experimental system in which *Escherichia coli* evolves into an insect mutualist.  
28 When the essential gut symbiont of the stinkbug *Plautia stali* was replaced by *E. coli*, a few  
29 survivor insects exhibited specific localization and vertical transmission of *E. coli*. Through  
30 trans-generational maintenance with *P. stali*, several hyper-mutating *E. coli* lines independently  
31 evolved host's high adult emergence and improved body color. Such "mutualistic" *E. coli* lines  
32 exhibited independent mutations disrupting the carbon catabolite repression (CCR) global  
33 transcriptional regulator. Each of the mutations reproduced the mutualistic phenotypes when  
34 introduced into wild-type *E. coli*, confirming that the single CCR mutations instantly make *E.*  
35 *coli* an insect mutualist. Our discovery uncovers that evolution of elaborate mutualism can  
36 proceed more easily and rapidly than conventionally envisaged.

37

38

39 Microbial symbioses are among the major evolutionary drivers underpinning the biodiversity  
40 (1,2). How ordinary free-living microbes have become sophisticated mutualists is an important  
41 but unanswered question in understanding the evolution of symbiosis. To address this  
42 fundamental issue, experimental evolutionary approaches may provide valuable insights (3-8).  
43 Here we report a novel experimental system in which the model bacterium *Escherichia coli*  
44 evolves into an insect mutualist, thereby demonstrating that evolution of mutualism can proceed  
45 very easily and quickly via disruption of a global transcriptional regulator system.

46

#### 47 ***E. coli* potentially capable of symbiosis with *P. stali***

48 Plant-sucking heteropteran bugs generally possess specific symbiotic bacteria in the midgut,  
49 which contribute to their growth and survival via provisioning of essential amino acids and/or  
50 vitamins (9,10). The brown-winged green stinkbug *Plautia stali* (Hemiptera: Pentatomidae)  
51 (Fig. 1a) develops a specialized symbiotic organ consisting of numerous crypts in a posterior  
52 region of the midgut (Fig. 1b). The crypt cavities are densely populated by a specific bacterial  
53 symbiont of the genus *Pantoea* (Fig. 1c, d). The symbiont is essential for growth and survival  
54 of the host insect. Normal insects infected with the uncultivable obligatory symbiont, *Pantoea*  
55 sp. A (11,12), attained over 70% adult emergence rates (Fig. 1e), smeared the symbiont cells  
56 onto the eggs upon oviposition (Fig. 1f), and transmitted the symbiont vertically to the offspring  
57 via nymphal probing of the eggshell (Fig. 1g). Aposymbiotic insects generated by egg surface  
58 sterilization died out with no adult emergence (Fig. 1e). Non-symbiotic bacteria, such as  
59 *Bacillus subtilis* and *Burkholderia insecticola*, cannot establish infection and symbiosis with *P.*  
60 *stali* (11). Meanwhile, when *E. coli* was inoculated to sterilized newborn nymphs, the insects  
61 certainly exhibited retarded growth and high mortality, but a small number of adult insects  
62 emerged, attaining 5-10% adult emergence rates (Fig. 1e; fig. S1) (11). Such adult insects,  
63 which were dwarf in size and dark in color (Fig. 1h), tended to die early, but some insects  
64 managed to survive, mate, and produce a small number of eggs. We dissected and inspected  
65 these insects, and found that, surprisingly, although the symbiotic organ was atrophied (Fig. 1i),  
66 *E. coli* localized to the midgut crypts just like the original symbiont, although the infection  
67 patterns were often patchy (Fig. 1j, k; fig. S2). Furthermore, *E. coli* cells were smeared on the  
68 eggshell and vertically transmitted to the offspring (Fig. 1l, m), although the transmission rates  
69 and the infection titers were unstable in comparison with those of the original symbiont (Fig.  
70 1l). These results suggested that, though incipiently, *E. coli* is capable of localized infection,  
71 vertical transmission, and supporting host survival in *P. stali*. Considering that *E. coli* belongs  
72 to the same Enterobacteriaceae as the original *Pantoea* symbiont, *E. coli* may be able to co-opt  
73 the mechanisms for infection and localization of the symbiont to establish the incipient

74 symbiosis (11). In this context, it seems relevant that, in the stinkbug family Pentatomidae, the  
75 gut symbiotic bacteria have evolved repeatedly from the Enterobacteriaceae through recurrent  
76 acquisitions and replacements (13,14).

77

### 78 **Experimental evolution using hyper-mutating *E. coli***

79 This finding prompted us to apply experimental evolutionary approaches to the *P. stali*-*E. coli*  
80 relationship. By continuously inoculated to and maintained with *P. stali*, would *E. coli* improve  
81 the symbiosis-related traits and finally evolve into a symbiont-like entity? Considering the  
82 expected difficulty in observing the evolution of elaborate symbiosis in a realistic time frame,  
83 we adopted the hyper-mutating *E. coli* strain,  $\Delta$ mutS, in which the DNA mismatch repair  
84 enzyme gene mutS is disrupted and the molecular evolutionary rate is elevated by two orders  
85 of magnitude (15). The *E. coli* strain of the same genetic background,  $\Delta$ intS, in which the phage  
86 integrase gene is disrupted, was used as control. Two selection schemes, growth selection and  
87 color selection, were conducted (fig. S3). In growth selection lines (GmL for hyper-mutating  
88  $\Delta$ mutS lines; GiL for non-mutating  $\Delta$ intS lines), the first-emerged adult insect was subjected to  
89 dissection of the symbiotic organ for inoculation to the next generation as well as freeze storing.  
90 In color selection lines (CmL for  $\Delta$ mutS lines; CiL for  $\Delta$ intS lines), the most greenish adult  
91 insect was subjected to dissection of the symbiotic organ for inoculation to the next generation  
92 as well as freeze storing. Throughout the evolutionary experiments, the host insects were  
93 supplied from a mass-reared inbred population of *P. stali*, thereby homogenizing the host  
94 genetic background and focusing on the evolutionary changes of the *E. coli* side. Since it takes  
95 around 1 month for newborn nymphs of *P. stali* to become adults under the rearing condition,  
96 it was expected that, ideally, we would be able to run 12 host generations per year. Actually,  
97 however, it took almost for two years because (i) the *E. coli*-inoculated insects generally  
98 exhibited high mortality and retarded growth, (ii) for keeping the insects under a good condition,  
99 frequent care without overcrowding was essential, which limited the manageable number of  
100 insects per evolutionary line ranging from 50 to 100, and (iii) consequently, extended generation  
101 time and stochastic extinction of the evolutionary lines frequently occurred, which had to be  
102 restarted from the frozen *E. coli* stocks.

103

### 104 **Evolution of mutualistic *E. coli***

105 We established and maintained 12 CmL color selection lines with 11 CiL control lines, and 7  
106 GmL growth selection lines with 7 CiL control lines (Fig. 2a, b). While the control  $\Delta$ intS-  
107 infected lines almost constantly exhibited low adult emergence rates, some of the hyper-  
108 mutating  $\Delta$ mutS-infected lines started to produce more adult insects. Notably, in a color

109 selection line CmL05, the adult emergence rate jumped up at generation 7, and the high  
110 emergence rates were maintained thereafter (Fig. 2a). In a growth selection line GmL07, the  
111 adult emergence rate improved as early as at generation 2, which was maintained thereafter  
112 (Fig. 2b). In CmL05 and GmL07, coincident with the improvement of the adult emergence rate,  
113 body color of the adult insects improved from dark to greenish (Fig. 2a-c; fig. S4), and  
114 furthermore, the colony morphology of *E. coli* changed from large and flat with rich  
115 extracellular matrix to small and convex with little extracellular matrix (Fig. 2c). When the  
116 frozen stocks of CmL05 and GmL07 were inoculated to *P. stali*, the improved adult emergence  
117 rate, the greenish body color, and the small and convex colony shape were reproducibly  
118 observed (Fig. 2d, e; fig. S5). These results indicated that some evolutionary lines of hyper-  
119 mutating *E. coli* have developed mutualistic traits for the host insect and that the phenotypic  
120 effects are attributable to genetic changes in the evolutionary *E. coli* lines.

121

### 122 **Microbial traits of mutualistic *E. coli***

123 In addition to the colony size, shape and extracellular matrix on agar plates (Fig. 2c), the  
124 mutualistic *E. coli* lines CmL05 and GmL07 in culture exhibited distinct microbial traits in  
125 comparison with the original *E. coli* strains: slower growth rate, smaller cell size, loss of  
126 flagellar motility, and unstable cell shape (fig. S6a-g). Within the host insect, the evolutionary  
127 *E. coli* lines CmL05 and GmL07 showed significantly higher infection densities than the  
128 original *E. coli* strains (fig. S6h). These observations revealed that the mutualistic *E. coli* lines  
129 certainly have developed a variety of “symbiont-like” microbial traits.

130

### 131 **Transcriptomics and genomics of mutualistic *E. coli***

132 An aliquot of the dissected symbiotic organ from each generation of the color selection line  
133 CmL05 was subjected to RNA sequencing, from which *E. coli*-derived reads were extracted  
134 and analyzed (tables S1 and S2). Interestingly, the gene expression patterns of *E. coli* at  
135 generations 7-14 after the improvement of host phenotypes were separately clustered in contrast  
136 to those at generations 1-6 before the improvement (Fig. 3a). In the growth selection line  
137 GmL07, similarly, the gene expression patterns of *E. coli* at generations 2-12 after the  
138 improvement were distinct from that at generation 1 before the improvement and also from  
139 those of the other growth selection lines GmL02 and GmL04 in which the improvement of host  
140 phenotypes did not occur (Fig. 3b; tables S1 and S3). These results suggested that the evolution  
141 of the mutualistic *E. coli* lines entails specific and global change of gene expression patterns.

142 In the growth selection line GmL07 and the color selection line CmL05, we surveyed  
143 differentially expressed genes before and after the improvement of host phenotypes (table S4),

144 which identified 193 commonly down-regulated genes and 95 commonly up-regulated genes  
145 across GmL07 and CmL05 (fig. S7a, b). The commonly down-regulated genes contained a  
146 number of metabolism-related genes such as transporter genes for sugars and other nutrients  
147 like maltose, ribose, galactitol, trehalose, mannose, branched chain amino acids, etc.,  
148 glyoxylate bypass genes, fatty acid degradation genes, and others. Notably, core genes involved  
149 in extracellular matrix (= Curli fimbriae) production were significantly down-regulated after  
150 the improvement (fig. S7c), which accounted for the altered colony morphology of *E. coli*  
151 associated with the improvement of host phenotypes (see Fig. 2c).

152 The improved lines CmL05 and GmL07 and the non-improved lines GmL02 and GmL04  
153 were subjected to genome sequencing throughout the evolutionary course (table S5), which  
154 identified many mutations accumulated in the hyper-mutating *E. coli* lines (tables S6 and S7;  
155 fig. S8). In an attempt to identify candidate mutations that are correlated with the improvement  
156 of the host phenotypes, we surveyed the mutations that appeared at generation 7 of CmL05 and  
157 then fixed, which yielded 7 candidate genes, and also the mutations that appeared at generation  
158 2 of GmL07 and then fixed, which yielded 9 candidate genes (Fig. 3c).

159

### 160 **Disrupted CCR pathway in mutualistic *E. coli***

161 Of these candidates, we focused on a frame shift mutation that disrupted adenylate cyclase  
162 (CyaA) in CmL05, and a non-synonymous mutation that changed a functionally important  
163 cAMP binding site of cAMP receptor protein (Crp) from leucine to proline in GmL07 (Fig. 3d).  
164 Despite their independent origins in distinct evolutionary lines, CyaA and Crp are pivotal  
165 components of the same global metabolic regulator system, the carbon catabolite repression  
166 (CCR) pathway, operating in diverse bacteria including *E. coli* (16,17) (Fig. 3e). With sufficient  
167 availability of glucose as the primary carbon source for *E. coli*, the CCR components are  
168 subjected to glucose-mediated suppression, being in an unphosphorylated form incapable of  
169 activating CyaA, by which the intracellular cAMP is maintained at a low level (fig. S9a). When  
170 glucose is used up, the glucose-mediated suppression is released, by which the CCR  
171 components are phosphorylated and activate CyaA, which results in an elevated intracellular  
172 cAMP level and promotes cAMP binding to Crp. The resultant global transcriptional regulator  
173 Crp-cAMP activates and/or represses several hundreds of operons throughout the bacterial  
174 genome, referred to as the Crp-cAMP regulon, by which the bacterial metabolic pathways are  
175 switched to exploit other carbon sources for adaptation to nutrient-deficient and/or high  
176 bacterial density conditions (fig. S9b) (18,19). According to RegulonDB (20), the Crp-cAMP  
177 regulon of *E. coli* consists of some 390 up-regulated genes and 80 down-regulated genes (fig.  
178 S9c), which are involved in, for example, up-regulation of transporters and catabolic enzymes



179 for non-glucose sugars (21), quorum sensing induction (22), and production of extracellular  
180 matrix (23).

181 Both the CyaA mutation in CmL05 and the Crp mutation in GmL07 are disruptive of the  
182 CCR pathway. Considering that *E. coli* cells are packed in the host symbiotic organ very densely  
183 (see Fig. 1k; fig. S2i, k), it seems likely that the symbiotic *E. coli* may be under a nutrient-  
184 limited condition within the host insect, at least locally. If so, it is expected that, in the  
185 evolutionary *E. coli* lines, while the Crp-cAMP transcriptional regulator was activated before  
186 the mutations occurred, the activation was disabled after the mutations occurred. Notably, of  
187 193 genes commonly down-regulated after the CyaA mutation in CmL05 and the Crp mutation  
188 in GmL07, 55 genes were reported as activated by Crp-cAMP (fig. S10a). These genes, which  
189 are expected to be silenced upon disruption of the CCR system, were significantly down-  
190 regulated in CmL05 and GmL07, which represented many transporter genes for non-glucose  
191 sugars, carbohydrate metabolism genes, quorum sensing genes, extracellular matrix production  
192 genes, transcription factor genes, and others (fig. S10b-i).

193

#### 194 **Disrupted CCR genes make *E. coli* an insect mutualist**

195 In order to test whether these mutations are involved in the mutualistic traits of the evolutionary  
196 *E. coli* lines, we prepared *E. coli* strains that carry the mutations under the wild-type genetic  
197 background: the strain  $\Delta$ cyaA in which cyaA gene is disrupted; and the strain crp<sup>221T>C</sup> whose  
198 crp gene was engineered to carry the leucine-proline replacement at the cAMP binding site.  
199 Both the mutant *E. coli* strains exhibited small and convex colonies with little extracellular  
200 matrix, somewhat slower growth rate, smaller cell size, and loss of flagellar motility (Fig. 4a;  
201 fig. S11a-e), which were generally reminiscent of the characteristic traits of the improved  
202 evolutionary *E. coli* lines CmL05 and GmL07 (Fig. 2c; fig. S6a-e). When the mutant *E. coli*  
203 strains were inoculated to sterilized newborn nymphs of *P. stali*, both the  $\Delta$ cyaA-infected  
204 insects and the crp<sup>221T>C</sup>-infected insects exhibited remarkably high adult emergence rates,  
205 which were comparable to the insects infected with the improved evolutionary *E. coli* lines and  
206 were significantly higher than the insects infected with the control *E. coli* strains (Fig. 4b).  
207 Moreover, the  $\Delta$ cyaA-infected insects and the crp<sup>221T>C</sup>-infected insects were greenish in color,  
208 which were comparable to the greenish insects infected with the improved evolutionary *E. coli*  
209 lines and distinct from the dwarf brown insects infected with the control *E. coli* strains (Fig.  
210 4c). On the other hand, infection densities of crp<sup>221T>C</sup> and  $\Delta$ cyaA were not comparable to those  
211 of the improved evolutionary *E. coli* lines (fig. S11f). These results demonstrated that, strikingly,  
212 the single mutations that disrupt the CCR global regulator system make *E. coli* mutualistic to  
213 the host insect *P. stali*.

214

## 215 **Discussion**

216 We established an experimental insect-*E. coli* symbiotic system in which the model bacterium  
217 is localized to host symbiotic organ, transmissible to host offspring vertically, and supportive  
218 of host survival, though incompletely. By infecting and passaging a hyper-mutating *E. coli*  
219 strain with the host insect trans-generationally, several evolutionary lines rapidly developed  
220 improved adult emergence and body color, realizing recurrent evolution of mutualism in the  
221 laboratory. Strikingly, the *E. coli*'s evolution into the insect mutualist was ascribed to single  
222 mutations that convergently disrupted the bacterial CCR pathway, uncovering unexpected  
223 involvement of the nutrient-responsive global transcriptional regulator in the establishment of  
224 symbiosis.

225 Our finding sheds new light on the evolvability of symbiosis – elaborate mutualistic  
226 symbiosis can evolve much more easily and rapidly than conventionally envisaged. We suggest  
227 the possibility that the inactivation of the CCR global regulator may represent a pivotal  
228 evolutionary step at an early stage of symbiosis. Densely packed in the symbiotic organ, the  
229 symbiotic bacteria are expected to constantly suffer nutritional shortage and activate the CCR  
230 pathway in vain, which may incur substantial metabolic cost and destabilize the symbiotic  
231 association. In this context, the disruption of the CCR pathway should benefit and stabilize the  
232 symbiosis. Our finding may be also relevant to the general evolutionary trend of symbiont  
233 genomes toward size reduction (24) and lack of transcription factors (25). The disruption of the  
234 CCR pathway causes silencing of otherwise activated about 400 genes under the Crp-cAMP  
235 regulon (20), which accounts for about 10% of the whole *E. coli* genome and provides potential  
236 targets for gene disruption, IS bombardment, intragenomic recombination, and reductive  
237 genome evolution. We propose that, although speculative, inactivation of transcriptional  
238 regulators and genome size reduction might have concurrently proceeded in this way during  
239 the symbiont genome evolution.

240 The *P. stali-E. coli* experimental symbiotic system will open a new window to directly  
241 observe and analyze the evolutionary processes and mechanisms of mutualistic symbiosis in  
242 real-time. *E. coli* is among the best understood cellular organisms, whose 4.5-5.5 Mb genome  
243 encodes over 4,000 genes and around 70% of them are with functional information (26,27).  
244 Laboratory evolution of mutualism using such a model bacterium with ample technological and  
245 genetic resources will lead to an ultimate understanding of the symbiotic evolution. Considering  
246 that *E. coli* represents a universal component of the gut microbiome of human, mouse, and other  
247 vertebrates (28), the insect-*E. coli* system in combination with the germfree mouse-*E. coli*  
248 experimental evolution systems (29,30) would enable us to pursue not only the differences but



249 also the commonality underpinning the mechanisms of gut symbiosis across vertebrates and  
250 invertebrates.

251

## 252 **Methods**

### 253 Insect and bacterial strains used in this study

254 An inbred laboratory strain of the brown-winged green stinkbug *P. stali* was established from  
255 several adult insects collected at Tsukuba, Ibaraki, Japan in September 2012, and has been  
256 maintained in the laboratory for years. This strain is associated with an essential and  
257 uncultivable gut symbiont *Pantoea* sp. A (II) in a posterior midgut region specialized as the  
258 symbiotic organ (Fig. 1; fig. S2). The insects were reared on raw peanuts, soybeans and water  
259 containing 0.05% ascorbic acid (Merck, Germany) at  $25 \pm 1$  °C and  $50 \pm 5\%$  relative humidity  
260 under a long-day regime of 16 h light and 8 h dark.

261 *E. coli* strains and mutants used in this study are listed below. The mutants  $\Delta$ intS,  $\Delta$ mutS and  
262  $crp^{221T>C}$  were generated as described later.

Name	Description	Reference
BW25113	Obtained from National Bio Resource Project (NBRP)	31
$\Delta$ intS	The intS gene was replaced with kanamycin resistance gene.	This study
$\Delta$ mutS	The mutS gene was removed from $\Delta$ intS.	This study
$\Delta$ cyxA	Obtained from Keio single-gene knock-out mutant library via NBRP	27
$Crp^{221T>C}$	A nonsynonymous single nucleotide substitution was introduced in the <i>crp</i> gene	This study
CmL05G13a	An <i>E. coli</i> clone isolated from the glycerol stock of CmL05G13.	This study
GmL07G12a	An <i>E. coli</i> clone isolated from the glycerol stock of GmL07G12.	This study
EPI300	Obtained from Epicentre	N.A.
DH5 $\alpha$	Common laboratory strain	N.A.
JM109	Common laboratory strain	N.A.
BL21	Common laboratory strain	N.A.

263

### 264 Construction of *E. coli* mutants

265 The *E. coli* mutant  $\Delta$ intS was established by replacing intS gene of *E. coli* BW25113 with nptII  
266 gene that confers kanamycin resistance by  $\lambda$ -Red homologous recombination using the  
267 pRed/ET plasmid (Gene Bridges, Germany). The *E. coli* mutant  $\Delta$ mutS was established by  
268 replacing mutS gene of  $\Delta$ intS with the FRT-Cm-FRT cassette (Gene Bridges, Germany) by  $\lambda$ -  
269 Red homologous recombination, and then Cm<sup>R</sup> was eliminated by Flp-FRT recombination. The  
270 *E. coli* mutant  $crp^{221T>C}$  was established from  $\Delta$ intS by replacing the 221<sup>st</sup> nucleotide T of the  
271 wild type *crp* gene with C, which changed 74<sup>th</sup> amino acid leucine of the Crp protein to proline.

272 This replacement was introduced by MAGE method (32) with a 90-mer DNA oligo (5'-  
273 taaagaaatg-atcctctct-atctgaatca-gggtgatttt-attggcgaac-Cgggcctgtt-tgaagagggc-caggaacgta-  
274 gcgcatgggt) whose 1<sup>st</sup> to 4<sup>th</sup> nucleotides were phosphothioated.

275

#### 276 Preparation of symbiont-free nymphs by surface sterilization of eggs

277 Egg clutches produced by the stock culture of *P. stali* were soaked in 4% formaldehyde for 10  
278 min, rinsed with sterilized water several times, and kept in sterilized plastic boxes until use.  
279 While this treatment does not affect hatchability and survival of the eggs, newborn nymphs fail  
280 to acquire the symbiotic bacteria and become symbiont-free (33).

281

#### 282 Experimental evolution of *P. stali*-*E. coli* artificial symbiotic system

283 Evolutionary experiments in this study consisted of, for each evolutionary *P. stali* line, (i)  
284 preparation of an inoculum either from *E. coli* culture of  $\Delta$ mutS or  $\Delta$ intS (only G1) or from an  
285 adult female of the previous generation (from G2 and on), (ii) oral administration of the  
286 inoculum to symbiont-free nymphs, (iii) rearing of the nymphs either to their adulthood or death,  
287 (iv) selection of an adult female for inoculation to the next generation, (iv) contamination check  
288 of the selected adult female, (v) preparation of an inoculum and a glycerol stock from the  
289 symbiotic organ dissected from the selected female, and (vi) morphological measurements of  
290 all adult insects obtained.

291 Either diluted *E. coli* culture (2.5 ml adjusted to OD<sub>600</sub> = 0.1) or homogenate of the  
292 symbiotic organ dissected from a selected female of the previous generation (2.5 ml containing  
293 1/2 organ equivalent) was soaked in a cotton pad and orally administered to around 84  
294 symbiont-free hatchlings derived from six surface-sterilized egg masses, by making use of the  
295 nymphal behavior that, after egg surface probing for about 30 min and resting for around a day,  
296 they take water without feeding and molt to second instar in a few days (11,33). These nymphs  
297 were reared on sterilized peanuts, soybeans and ascorbic acid water as described previously  
298 (33). In the evolutionary experiments, two selection schemes, growth selection and color  
299 selection, were conducted (fig. S3). In growth selection lines (GmL for hyper-mutating  $\Delta$ mutS  
300 lines; GiL for non-mutating  $\Delta$ intS lines), the first-emerged adult female was subjected to  
301 dissection of the symbiotic organ for inoculation to the next generation as well as freeze storing.  
302 In color selection lines (CmL for  $\Delta$ mutS lines; CiL for  $\Delta$ intS lines), adult females were collected  
303 for 35 days after hatching or until at least one adult female emerged. These adult females were  
304 anesthetized on ice and photographed from the ventral side using a digital camera. Their body  
305 color was measured using the image analyzing software Natsumushi ver. 1.10 (34), and the  
306 adult female that exhibited the highest hue angle (= greenness) was subjected to dissection of

307 the symbiotic organ for inoculation to the next generation as well as freeze storing.

308 The symbiotic organ of the selected female was dissected in PBS (0.8% NaCl, 0.02% KCl,  
309 0.115% Na<sub>2</sub>HPO<sub>4</sub>, 0.02% KH<sub>2</sub>PO<sub>4</sub>, pH 7.4), rinsed with 70% ethanol, and homogenized in 200  
310 μL sterile water. Of the 200 μL homogenate, 5 μL was used for contamination check by  
311 quantitative PCR. The number of *E. coli* genome copies was evaluated in terms of kanamycin  
312 resistance gene copies, which is present in the ΔintS and ΔmutS mutants but absent in wild-  
313 type *E. coli* and other bacteria. The number of total bacterial genome copies was evaluated  
314 based on bacterial 16S rRNA gene copies. When the former *E. coli* genome copy number was  
315 approximately the same as the latter bacterial genome copy number, the specimen was  
316 diagnosed as free of contamination. When the specimen was diagnosed as contaminated, the  
317 next best female was used. For quantitative PCR, the primers Tn5-1789F (5'-TGC TCG ACG  
318 TTG TCA CTG AA-3') and Tn5-1879R (5'-GCA GGA GCA AGG TGA GAT GA-3') were used  
319 for kanamycin resistance gene, while the primers 16S-967F (5'-CAA CGC GAA GAA CCT  
320 TAC C-3') and 16S-1046R (5'-CGA CAG CCA TGC ANC ACC T-3') were used for bacterial  
321 16S rRNA gene. The PCR reaction was performed using Brilliant PCR mix (Agilent  
322 Technologies, USA). The standard curve was drawn using serially diluted ΔintS genomic DNA,  
323 which contains one kanamycin gene copy and seven 16S rRNA gene copies per genome. The  
324 thermal profile was the initial denaturation at 95°C for 3 min followed by 40 cycles of  
325 incubation at 95°C for 5 sec and at 60°C for 10 sec. To confirm specific amplification, melting  
326 curve analysis was also included. The reaction was conducted on Mx3000p (Agilent  
327 Technologies, USA). While 100 μL of the homogenate of the female symbiotic organ diagnosed  
328 as free of contamination was used as the inoculum to the next generation, the remaining  
329 homogenate (~95 μL) was mixed with an equal volume of 20% glycerol and stored at -80°C.

330

### 331 Inoculation of *E. coli* frozen stocks to *P. stali*

332 The frozen glycerol stocks were thawed, of which 50 mL was taken and diluted with sterile  
333 water to 3 mL. Each of three replicates of around 84 symbiont-free hatchlings from six surface-  
334 sterilized egg masses was fed with 1 ml inoculum soaked in a cotton pad as described above.  
335 The symbiont A and ΔmutS were included in the evaluation as positive and negative controls,  
336 respectively. Adult emergence of the insects was monitored for 50 days after hatching. All the  
337 adult insects were photographed from the dorsal side with a digital camera, and hue angle (=   
338 greenness) of the scutellum and thorax width were measured using ImageJ (35). For the  
339 subsequent RNA sequencing analyses and resequencing of *E. coli* genomes, the symbiotic  
340 organs were isolated from the adult insects and homogenized in 100 μL PBS. Of 100 μL  
341 homogenate, 50 μL was subjected to RNA sequencing and the remaining 50 μL was used for

342 genome resequencing.

343

#### 344 RNA sequencing analyses

345 The homogenate of the symbiotic organ was subjected to total RNA extraction using RNAiso  
346 (Takara Bio, Japan) and RNeasy Mini Kit (Qiagen, Nederland). Then, ribosomal RNAs of both  
347 insect and bacterial origins were removed from the total RNA samples using Ribo-Zero Gold  
348 rRNA Removal Kit (Epidemiology) (Illumina, USA). The rRNA-depleted RNAs were  
349 converted to paired end libraries using Sure Select Strand Specific RNA Kit (Agilent  
350 Technologies, USA) or TruSeq RNA Library Prep Kit v2 (Illumina, USA) (see [table S1](#)). The  
351 libraries were sequenced with Hiseq 3000 or Hiseq X (Illumina, USA).

352 The obtained sequences were trimmed, mapped to *E. coli* BW25113 genome sequence  
353 (Accession number NZ\_CP009273), and read-counted with CLC Genomics Workbench 10.0  
354 (Qiagen, Germany). Normalizations and differential expression analyses were conducted with  
355 EdgeR ver. 3.32.1 (36). Complex Heatmap ver. 2.10.0 (37) was used for clustering analyses and  
356 drawing heatmaps of the RNA sequencing libraries.

357

#### 358 Genome resequencing and detection of structural changes

359 DNA samples were extracted from the homogenates of the symbiotic organ using QIAamp  
360 DNA Mini Kit (Qiagen, Germany). The extracted DNAs were converted to paired end libraries  
361 using Nextera XT DNA Library Prep Kit (Illumina, USA) and the libraries were sequenced  
362 with Miseq (Illumina, USA). CLC Genomic Workbench ver. 10.0 was used for detection of *E.*  
363 *coli* genome variants that emerged during the evolutionary experiments. The heatmaps of the  
364 variant frequency data were drawn using Complex Heatmap (37).

365

#### 366 Fluorescence in situ hybridization

367 Fluorescence in situ hybridization (FISH) analyses were performed essentially as described  
368 (38). The whole insect bodies or isolated digestive tracts were fixed with PBS containing 4%  
369 formaldehyde (Fujifilm, Japan). The fixed samples were embedded in Technovit 8100 (Kulzer,  
370 Germany) and processed into 2  $\mu$ m tissue sections using a rotary microtome RM2255 (Leica,  
371 Germany). The AlexaFluor555-labeled oligonucleotide probes Eco934 (5'-CAT GCT CCA  
372 CCG CTT GTG-3') and SymAC89R (5'-GCA AGC TCT TCT GTG CTG CC-3') were used to  
373 detect *E. coli* and the symbiont A, respectively (12). Host nuclei were counterstained with 4',  
374 6-diamidino-2-phenylindole (DAPI) (Dojindo, Japan). The hybridized specimens were  
375 observed using a fluorescence dissection microscope M165FC (Leica, Germany), an  
376 epifluorescence microscope DM6B (Leica, Germany), and a laser confocal microscope

377 LSM700 (Zeiss, Germany).

378

379 Infection of *E. coli* mutants and effects on host phenotypes

380 *E. coli* mutants were cultured, diluted, and orally administrated to symbiont-free newborn  
381 nymphs of *P. stali* as described above. The insects were reared to monitor their adult emergence  
382 for 42 days after hatching. The dorsal images of the adults were taken with an image scanner  
383 GT-X830 (Epson, Japan), and the hue angle of the scutellum and thorax width were measured  
384 and analyzed using the software Natsumushi (34). *P. stali* harboring the original symbiont  
385 *Pantoea* sp. A was also included as a reference. As for the adult females infected with *E. coli*,  
386 bacterial titers in the symbiotic organs were measured by quantitative PCR. KAPA SYBR Fast  
387 qPCR Kit (Roche, USA), Tn5-1789F and Tn5-1879R primer sets were used for quantification.  
388 The standard curves were drawn using serially diluted pT7Blue (Takara Bio, Japan) plasmid  
389 carrying a kanamycin resistance gene fragment. The quantitative PCR reactions were conducted  
390 on Light Cycler 96 (Roche, Switzerland).

391

392 Measurement of *E. coli* phenotypes

393 For inspection of colony morphology and extracellular matrix production, *E. coli* cultures were  
394 spread onto LB agar plates containing 80 µg/mL Congo Red (Merck, USA) and incubated at  
395 25°C for 3 days. Colonies formed on the plate were photographed by using a scanner GT-X850  
396 and/or dissection microscope S9i (Leica, Germany).

397 For growth curve measurements, each glycerol stock of *E. coli* was inoculated to 2 mL LB  
398 broth (Becton Dickinson, USA) and incubated at 25°C for 16 h with shaking at 200 rpm. The  
399 cell culture was diluted to OD<sub>600</sub> = 0.005 in 25 mL LB broth, and incubated at 25°C with shaking  
400 at 200 rpm. From the bacterial culture, 120 µL of cell suspension was sampled every hour, and  
401 the samples were subjected to measurement of OD<sub>600</sub> using a spectrometer UV-1800 (Shimadzu,  
402 Japan).

403 For time-lapse analyses of growth and morphology of individual *E. coli* cells, two types  
404 of microfluidic devices were used. One type was a microfluidic device in which bacterial cells  
405 were enclosed in microchambers etched on a glass coverslip. A cellulose membrane was  
406 attached to a coverslip via biotin-streptavidin binding, on which the microchambers were  
407 fabricated as described previously (39,40). Another type was a microfluidic device made of  
408 polydimethylsiloxane (PDMS) with a channel structure similar to Mother Machine 3 as described  
409 previously (41) (see [fig. S6f](#)). The width of the cell observation channels in this device was 9  
410 µm, which was broader than that of the Mother Machine and thus each cell observation channel  
411 could harbor 30-70 individual *E. coli* cells depending on cell sizes. *E. coli* cells in exponential

412 phase were introduced into both types of the microfluidic devices and observed under a Nikon  
413 Ti-E microscope (Nikon, Japan) equipped with ORCA-fusion camera (Hamamatsu Photonics,  
414 Japan). In the time-lapse measurements, phase-contrast images were acquired with a 100 × oil  
415 immersion objective lens (Plan Apol, NA 1.45) at an interval of 3 min, in which 50-100 XY  
416 positions were simultaneously observed. The microscope was controlled from a computer using  
417 Micromanager 4. In the microchamber device measurements, LB broth was supplemented with  
418 0.1% bovine serum albumin and 0.02% Tween-80 to suppress cell adhesion, and introduced  
419 into the devices at a flow rate of 2 mL/h.

420 For measurements of size and flagellar motility, *E. coli* cells were grown in LB medium  
421 with shaking at 25°C to around OD<sub>600</sub> = 2.0, observed under a phase-contrast microscope IX71  
422 (Olympus, Japan), recorded by a CCD camera DMK33UP5000.WG (The Imaging Source,  
423 Germany) at 30 frames per second, and analyzed using ImageJ v1.53 (35) and IGOR Pro 8.02  
424 J (WaveMetrics, USA). The cell size data were measured for individual six cultures. The  
425 swimming ratio data were obtained as the number of swimming cells in 100 cells from  
426 individual eight cultures.

427

#### 428 Statistics

429 Statistical analyses were conducted by using R ver. 4.1.2 (42) and RStudio (43). R was also  
430 used to plot the data.

431

#### 432 **References**

- 433 1. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences.  
434 *Proc. Natl. Acad. Sci. USA* **110**, 3229-3236 (2013).
- 435 2. Gilbert, S. F., Bosch, T. C. G. & Ledón-Rettig, C. Eco-Evo-Devo: developmental symbiosis  
436 and developmental plasticity as evolutionary agents. *Nat. Rev. Genet.* **16**, 611-622 (2015).
- 437 3. Hoang, K. L., Morran, L. T. & Gerardo, N. M. Experimental evolution as an underutilized  
438 tool for studying beneficial animal–microbe interactions. *Front. Microbiol.* **7**, 1444 (2016).
- 439 4. King, K. C. *et al.* Rapid evolution of microbe-mediated protection against pathogens in a  
440 worm host. *ISME J.* **10**, 1915-1924 (2016).
- 441 5. Tso, G. H. W. *et al.* Experimental evolution of a fungal pathogen into a gut symbiont.  
442 *Science* **362**, 589-595 (2018).
- 443 6. Robinson, C. D. *et al.* Experimental bacterial adaptation to the zebrafish gut reveals a  
444 primary role for immigration *PLoS Biol.* **16**, e2006893 (2018).
- 445 7. Mehta, A. P. *et al.* Engineering yeast endosymbionts as a step toward the evolution of  
446 mitochondria. *Proc. Natl. Acad. Sci. USA* **115**, 11769-11801 (2018).



- 447 8. Drew, G. C., EJ Stevens, E. J. & King, K. C. Microbial evolution and transitions along the  
448 parasite–mutualist continuum. *Nat. Rev. Microbiol.* **19**, 623-638 (2021).
- 449 9. Nikoh, N. *et al.* Reductive evolution of bacterial genome in insect gut environment. *Genome*  
450 *Biol. Evol.* **3**, 702-714 (2011).
- 451 10. Salem, H. *et al.* Vitamin supplementation by gut symbionts ensures metabolic homeostasis  
452 in an insect host. *Proc. R. Soc. B* **281**, 20141838 (2014).
- 453 11. Hosokawa, T. *et al.* Obligate bacterial mutualists evolving from environmental bacteria in  
454 natural insect populations. *Nat. Microbiol.* **1**, 15011 (2016).
- 455 12. Oishi, S. *et al.* Morphogenesis and development of midgut symbiotic organ of the stinkbug  
456 *Plautia stali* (Hemiptera: Pentatomidae). *Zool. Let.* **5**, 16 (2019).
- 457 13. Hosokawa T. *et al.* Recurrent evolution of gut symbiotic bacteria in pentatomid stinkbugs.  
458 *Zool. Let.* **2**, 34 (2016).
- 459 14. Otero-Bravo, A. & Sabree, Z. L. Multiple concurrent and convergent stages of genome  
460 reduction in bacterial symbionts across a stink bug family. *Sci. Rep.* **11**, 7731 (2021).
- 461 15. Giraud, A. *et al.* Costs and benefits of high mutation rates: adaptive evolution of bacteria in  
462 the mouse gut. *Science* **291**, 2606-2608 (2001).
- 463 16. Deutscher, J. *et al.* How phosphotransferase system-related protein phosphorylation  
464 regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 939-1031  
465 (2006).
- 466 17. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the  
467 most out of nutrients *Nat. Rev. Microbiol.* **6**, 613-624 (2008).
- 468 18. Martinez-Antonio, A. & Collado-Vides, J. Identifying global regulators in transcriptional  
469 regulatory networks in bacteria. *Curr. Opin. Microbiol.* **6**, 482-489 (2003).
- 470 19. Zheng, D. *et al.* Identification of the CRP regulon using in vitro and in vivo transcriptional  
471 profiling. *Nucleic Acids Res.* **32**, 5874-5893 (2004).
- 472 20. Santos-Zavaleta, A. *et al.* RegulonDB v 10.5: tackling challenges to unify classic and high  
473 throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* **47**, D212-  
474 D220 (2019).
- 475 21. Shimada, T. *et al.* Novel roles of cAMP receptor protein (CRP) in regulation of transport  
476 and metabolism of carbon sources. *PLoS One* **6**, e20081 (2011).
- 477 22. Xavier, K. B. & Bassler, B. L. Regulation of uptake and processing of the quorum-sensing  
478 autoinducer AI-2 in *Escherichia coli*. *J. Bacteriol.* **187**, 238-248 (2005).
- 479 23. Müller, C. M. *et al.* Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia*  
480 *coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathog.* **5**, 1000303 (2009).
- 481 24. McCutcheon, J. P. & Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nat.*

- 482 *Rev. Microbiol.* **10**, 13-26 (2012).
- 483 25. Moran, N. A. *et al.* Regulation of transcription in a reduced bacterial genome: nutrient-  
484 provisioning genes of the obligate symbiont *Buchnera aphidicola*. *J. Bacteriol.* **187**, 4229-  
485 4237 (2005).
- 486 26. Blount, Z. D. The unexhausted potential of *E. coli*. *eLife* **4**, e05826 (2015).
- 487 27. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout  
488 mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
- 489 28. Tenaillon, O. *et al.* The population genetics of commensal *Escherichia coli*. *Nat. Rev.*  
490 *Microbiol.* **8**, 207-217 (2010).
- 491 29. Giraud, A. *et al.* Dissecting the genetic components of adaptation of *Escherichia coli* to the  
492 mouse gut. *PLoS Genet.* **4**, e2 (2019).
- 493 30. Barroso-Batista, J. *et al.* Specific eco-evolutionary contexts in the mouse gut reveal  
494 *Escherichia coli* metabolic versatility. *Curr. Biol.* **30**, 1049-1062.e7 (2020).
- 495 31. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in  
496 *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645  
497 (2000).
- 498 32. Nyerges Á., *et al.* A highly precise and portable genome engineering method allows  
499 comparison of mutational effects across bacterial species. *Proc. Natl. Acad. Sci. USA* **113**,  
500 2502–2507 (2016).
- 501 33. Nishide, Y., *et al.* Aseptic rearing procedure for the stinkbug *Plautia stali* (Hemiptera:  
502 Pentatomidae) by sterilizing food-derived bacterial contaminants. *Appl. Entomol. Zool.* **53**,  
503 407-415 (2017).
- 504 34. Tanahashi, M. & Fukatsu, T. *Natsumushi* - Image measuring software for entomological  
505 studies. *Entomol. Sci.* **21**, 347-360 (2018).
- 506 35. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image  
507 analysis. *Nat. Methods* **9**, 671–675 (2012).
- 508 36. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. EdgeR: a Bioconductor package for  
509 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140  
510 (2010).
- 511 37. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in  
512 multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
- 513 38. Koga, R., Tsuchida, T. & Fukatsu, T. Quenching autofluorescence of insect tissues for in  
514 situ detection of endosymbionts. *Appl. Entomol. Zool.* **44**, 281–291 (2009).
- 515 39. Inoue, I., Wakamoto, Y., Moriguchi, H., Okano, K. & Yasuda, K. On-chip culture system  
516 for observation of isolated individual cells. *Lab. Chip* **1**, 50-55 (2001).

- 517 40. Hashimoto, M., *et al.* Noise-driven growth rate gain in clonal cellular populations. *Proc.*  
518 *Natl. Acad. Sci. USA* **113**, 3251-3256 (2016).
- 519 41. Wang, P., *et al.*, Robust growth of *Escherichia coli*. *Curr. Biol.* **20**, 1099–1103 (2010).
- 520 42. R Core Team. R: A language and environment for statistical computing. R Foundation for  
521 Statistical Computing, Vienna, Austria (2021). <https://www.R-project.org/>
- 522 43. RStudio Team. RStudio: Integrated Development Environment for R (2020).  
523 <http://www.rstudio.com/>
- 524 44. Hayashi, T., *et al.* Female-specific specialization of a posterior end region of the midgut  
525 symbiotic organ in *Plautia splendens* and allied stinkbugs. *Appl. Environ. Microbiol.* **81**,  
526 2603-2611 (2015).

527

## 528 **Acknowledgments**

529 We thank U. Asaga, S. Kimura, J. Makino and T. Matsushita for insect rearing and technical  
530 assistance. This study was supported by the JST ERATO grants JPMJER1803 and  
531 JPMJER1902 (TF, CF, YW, RK) and the JSPS KAKENHI grant JP25221107 (TF, RK).  
532 Genome sequencing and analyses were supported by the JSPS KAKENHI grant JP16H06279.

533

534 **Author contributions:** RK and TF conceived the project and designed the experiments. RK,  
535 MMo, NOT and YN performed insect-*E. coli* evolutionary experiments, RK, MMo, NOT, YI,  
536 HT, YN and THo analyzed insect phenotypes, RK, MMi, KO, RO and YW analyzed *E. coli*  
537 phenotypes, RK, MMo, NOT, YG and THa performed genome sequencing and analyses, MMo,  
538 RK, NOT, MS and YS conducted RNA sequencing and analyses, RK, HT, SS and CF designed  
539 and generated hyper-mutating and other *E. coli* strains, and TF wrote the manuscript with input  
540 from all the other authors.

541

## 542 **Competing interests**

543 The authors declare no competing interests.

544

## 545 **Data and materials availability**

546 All RNA sequencing and DNA sequencing data produced in this study were deposited in DDBJ  
547 Sequence Read Archive (DRA) (see [tables S1 and S5](#)). All data are available in the manuscript  
548 or the supplementary materials.

549

550

551 **Fig. 1. Infection, localization and vertical transmission of *E. coli* in the gut symbiotic**  
552 **system of *P. stali*.** (a) Normal symbiotic adult female, large in size and green in color. (b)  
553 Dissected alimentary tract, in which symbiotic organ is well developed and yellow in color. (c)  
554 Fluorescence in situ hybridization (FISH) localization of symbiont cells to the symbiotic organ.  
555 (d) Magnified FISH image showing symbiont localization to crypt cavities of the symbiotic  
556 organ. (e) Adult emergence rates of newborn nymphs inoculated with normal symbiont (Sym,  
557 *Pantoea* sp. A), no bacteria (Apo, aposymbiotic), and *E. coli*. (f) Symbiont cells smeared on  
558 egg surface. (g) Newborn nymphs sucking symbiont cells from eggshell. (h) *E. coli*-infected  
559 adult female, dwarf in size and brown in color. (i) Dissected alimentary tract, in which  
560 symbiotic organ is atrophied. (j) FISH localization of *E. coli* to the symbiotic organ. (k)  
561 Magnified FISH image visualizing *E. coli* localization to crypt cavities of the symbiotic organ.  
562 (l) Bacterial titers in symbiont-inoculated and *E. coli*-inoculated nymphs one day after second  
563 instar molt in terms of groEL and nptII gene copies per insect, respectively. (m) *E. coli* cells  
564 smeared on egg surface.

565

566

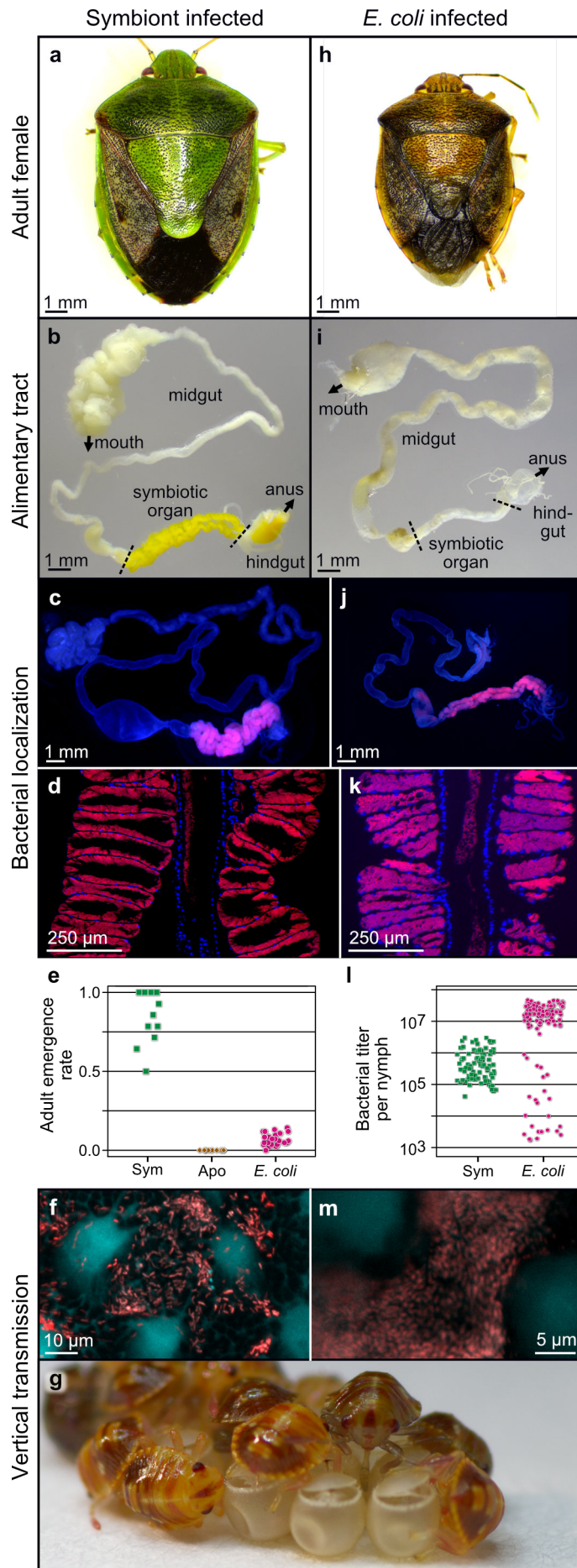


Fig. 1

570 **Fig. 2. Evolution of mutualistic traits for *P. stali* in hyper-mutating *E. coli* lines. (a)**  
571 Evolutionary *E. coli* lines subjected to host's body color selection. Data of adult emergence rate  
572 and body color are displayed by heat maps. White asterisks indicate missing data of body color  
573 measurement. **(b)** Evolutionary *E. coli* lines subjected to host's growth speed selection. Data of  
574 adult emergence rate and days to the first adult emergence are displayed by heat maps. Note  
575 that in **(a)** and **(b)**, when an evolutionary line produced no adult insect and recovery from the  
576 freeze stock failed twice consecutively, the evolutionary line was terminated due to shortage of  
577 inoculum. From generation 10 and on, selected evolutionary lines were maintained. **(c)** Host's  
578 body color and colony morphology of evolutionary *E. coli* lines. Red colonies are due to rich  
579 extracellular matrix produced on the agar plates containing Congo red. **(d, e)** Adult emergence  
580 patterns of *P. stali* infected with the representative *E. coli* lines, CmL05, GmL07, GmL02 and  
581 GmL04, in the original evolutionary experiments **(d)** and those in the confirmation experiments  
582 using frozen *E. coli* stocks **(e)**. In **(c)**-**(e)**, magenta lines and blue lines highlight "improved"  
583 generations and "non-improved" generations, respectively.  
584  
585  
586



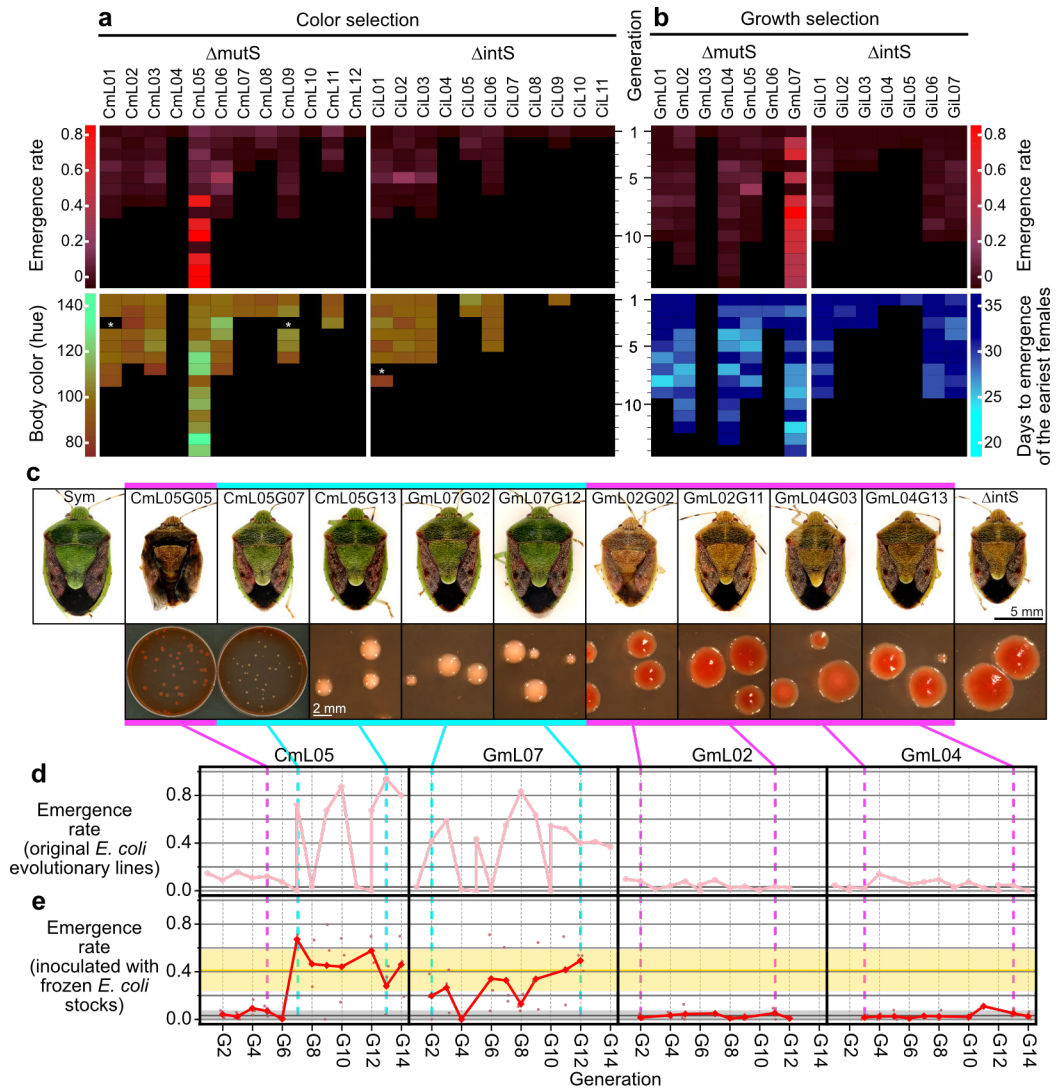


Fig. 2

590 **Fig. 3. Transcriptomics and genomics of evolutionary *E. coli* lines.** (a, b) Clustering  
591 dendrograms and heatmaps based on gene expression levels across generations of evolutionary  
592 *E. coli* lines subjected to color selection (right, 3,401 genes) (a) and growth selection (left,  
593 3,360 genes) (b). Gray and colored areas depict non-improved and improved generations,  
594 respectively. (c) Mutations identified in the genomes of CmL05 and GmL07 as coincident with  
595 the improvement of host phenotypes. (d) Candidate mutations disrupting the catabolite  
596 repression pathway: a frame shift mutation in *cyaA* of CmL05 (top) and a non-synonymous  
597 mutation causing change from leucine to proline at a functionally important cAMP binding  
598 domain in *crp* of GmL07 (bottom). (e) Schematic presentation as to how CRP pathway is  
599 disrupted by the *cyaA* and *crp* mutations.  
600  
601

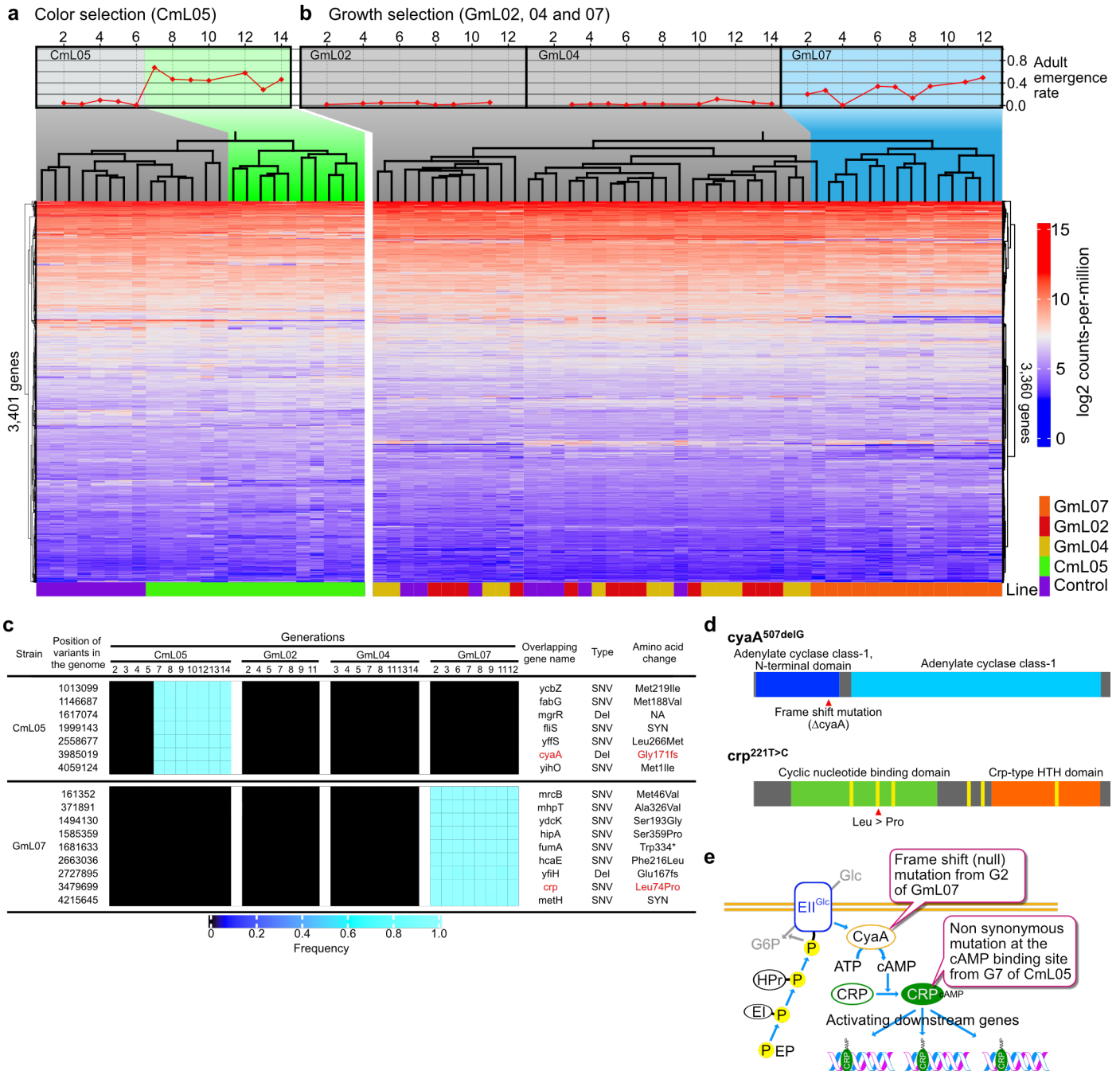


Fig. 3

605 **Fig. 4. Single mutations disrupting carbon catabolite repression pathway make *E. coli***  
606 **mutualistic to *P. stali*.** (a) Small, convex and white colonies of  $\Delta\text{cyaA}$  and  $\text{crp}^{221\text{T}>\text{C}}$ . (b) Adult  
607 emergence rates of *P. stali* infected with  $\Delta\text{cyaA}$  and  $\text{crp}^{221\text{T}>\text{C}}$ . Different alphabetical letters  
608 indicate statistically significant differences (pairwise Wilcoxon rank sum test with Bonferroni  
609 correction:  $P < 0.05$ ). (c) Adult insects infected with  $\Delta\text{cyaA}$  and  $\text{crp}^{221\text{T}>\text{C}}$ , which are larger in  
610 size and green in color in comparison with those infected with control  $\Delta\text{intS}$ .  
611  
612

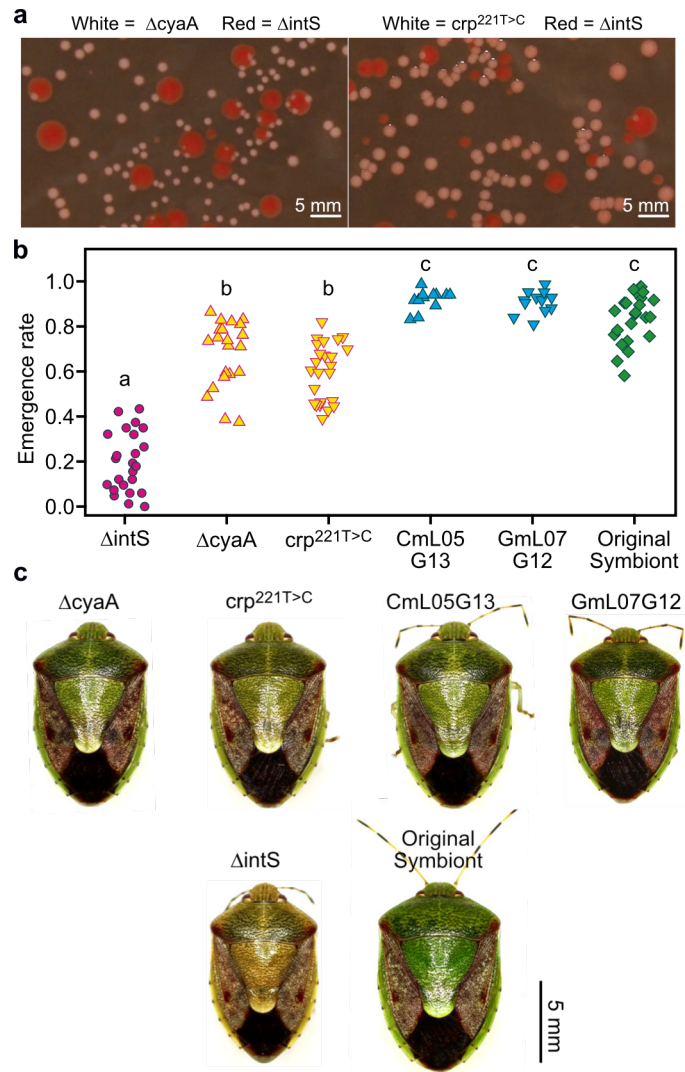
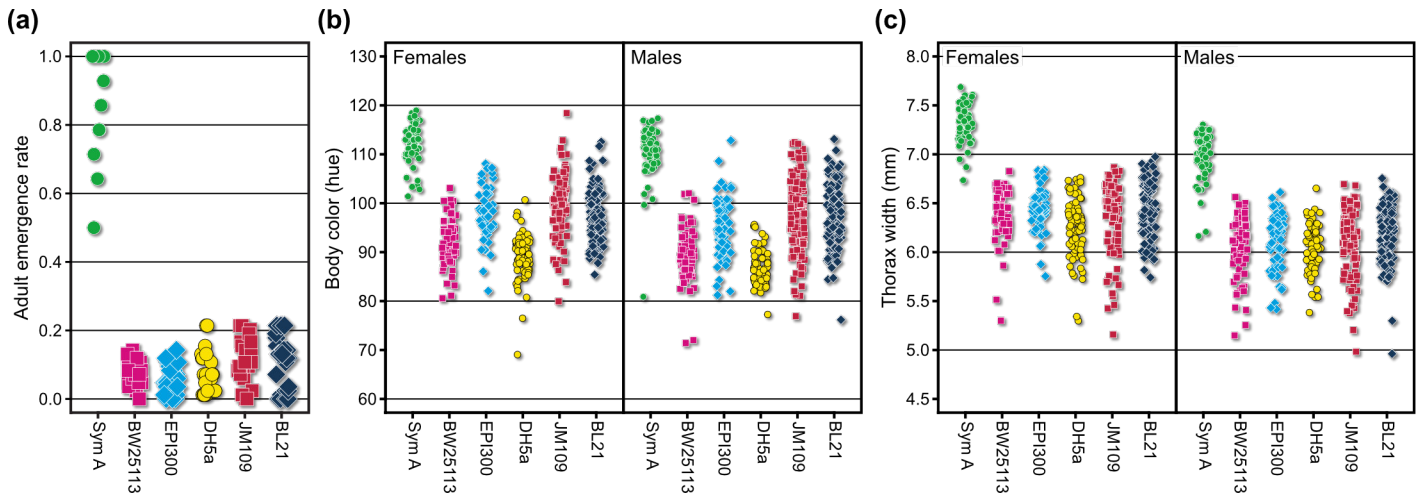


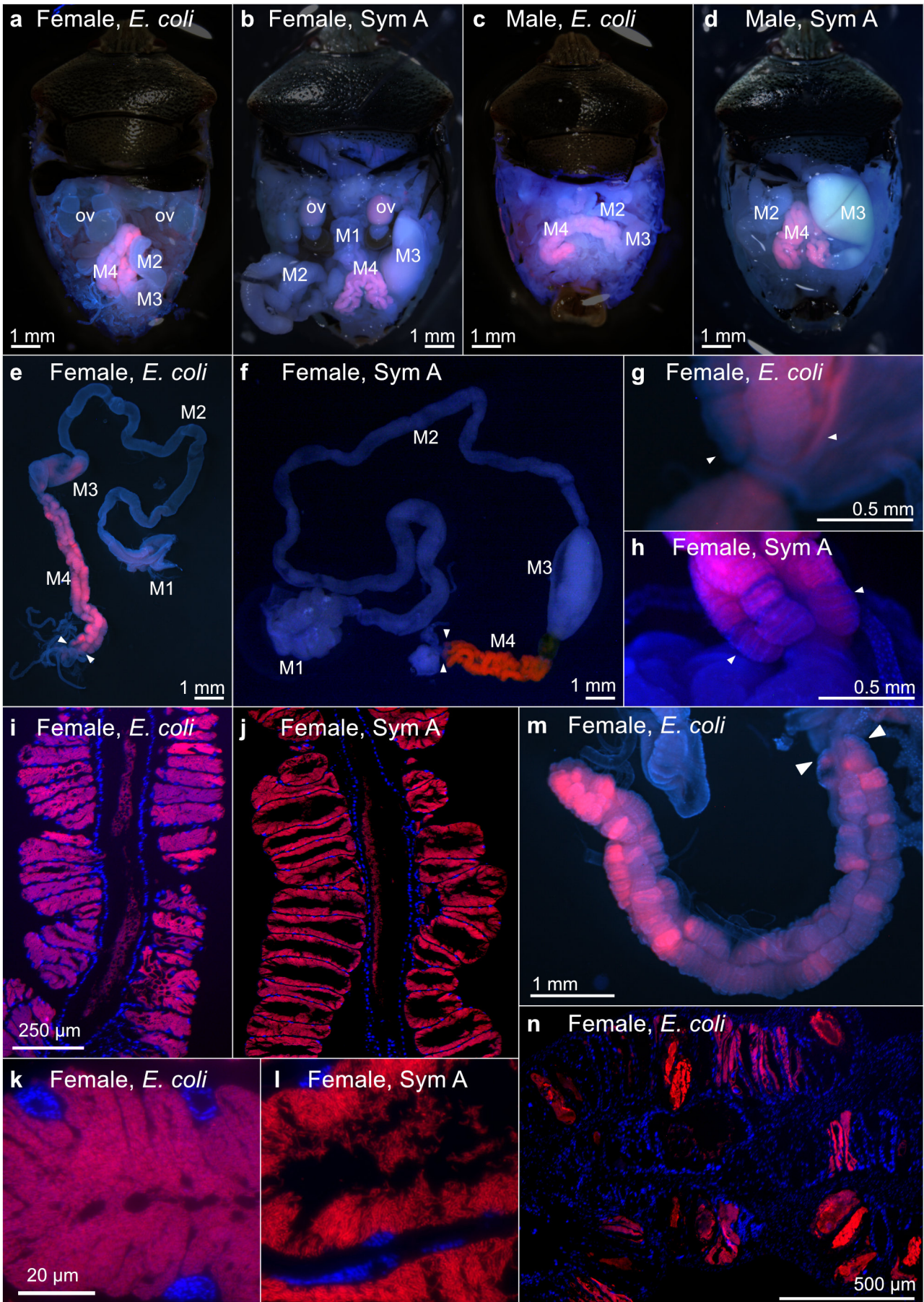
Fig. 4

616 **Fig. S1. Phenotypes of *P. stali* adults infected with laboratory strains of *E. coli*.** (a) Adult  
617 emergence rate. (b) Body color (greenish hue) of females (left) and males (right). (c) Body size  
618 (thorax width) of females (left) and males (right). Sym A is *Pantoea* sp. A, the original,  
619 uncultivable and essential gut symbiont of *P. stali* (11). BW25113, EPI300, DH5a, JM109 and  
620 BL21 are commonly used laboratory strains of *E. coli*.  
621  
622





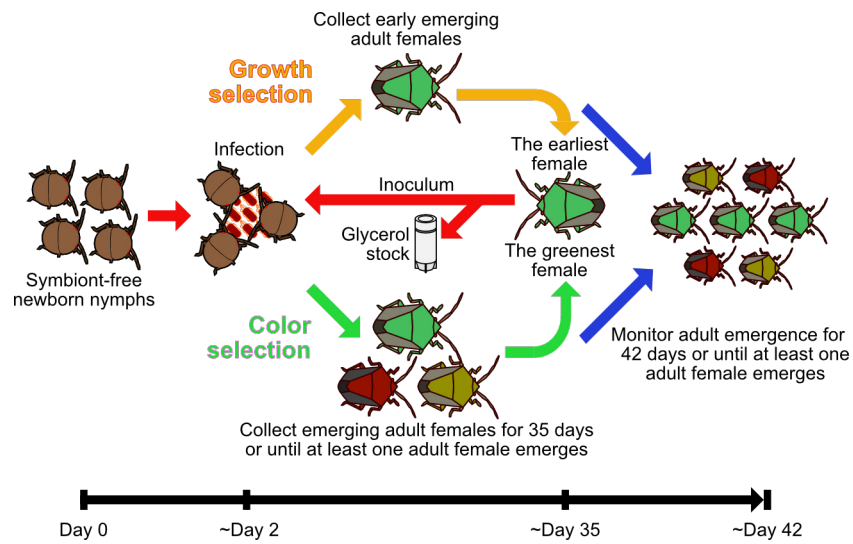
626 **Fig. S2. FISH localization of *E. coli* and original symbiont *Pantoea* sp. A (= Sym A) in *P.***  
627 ***stali*. (a-d)** Localization in abdominal body cavity of adult insects: **(a)** *E. coli* in adult female,  
628 **(b)** Sym A in adult female, **(c)** *E. coli* in adult male, and **(d)** Sym A in adult male. FISH signals  
629 are localized to the midgut M4 region. Signals in oocytes are due to autofluorescence.  
630 Abbreviations: M1, M2, M3, and M4, midgut regions M1, M2, M3, and M4 (= symbiotic  
631 organ); ov, ovary. **(e, f)** Localization of *E. coli* **(e)** and Sym A **(f)** in dissected alimentary tract  
632 of adult females. Arrowheads indicate female-specific enlarged end crypts at the posterior end  
633 of the symbiotic organ, which are presumably involved in vertical symbiont transmission by  
634 storing bacteria-containing secretion (44). **(g, h)** Magnified images of the end crypts infected  
635 with *E. coli* **(g)** and Sym A **(h)**. Note that *E. coli*-infected end crypts are atrophied in comparison  
636 with Sym A-infected ones. **(i, j)** Localization of *E. coli* **(i)** and Sym A **(j)** in the crypt cavities  
637 of the symbiotic organ. **(k, l)** Magnified images of *E. coli* cells **(k)** and Sym A cells **(l)** packed  
638 in the crypt cavity. **(m, n)** Patchy localization patterns of *E. coli* in the symbiotic organ, which  
639 are often found with *E. coli* but seldom observed with Sym A.  
640  
641



645 **Fig. S3. Experimental scheme for evolution of mutualistic *E. coli* with *P. stali*.**

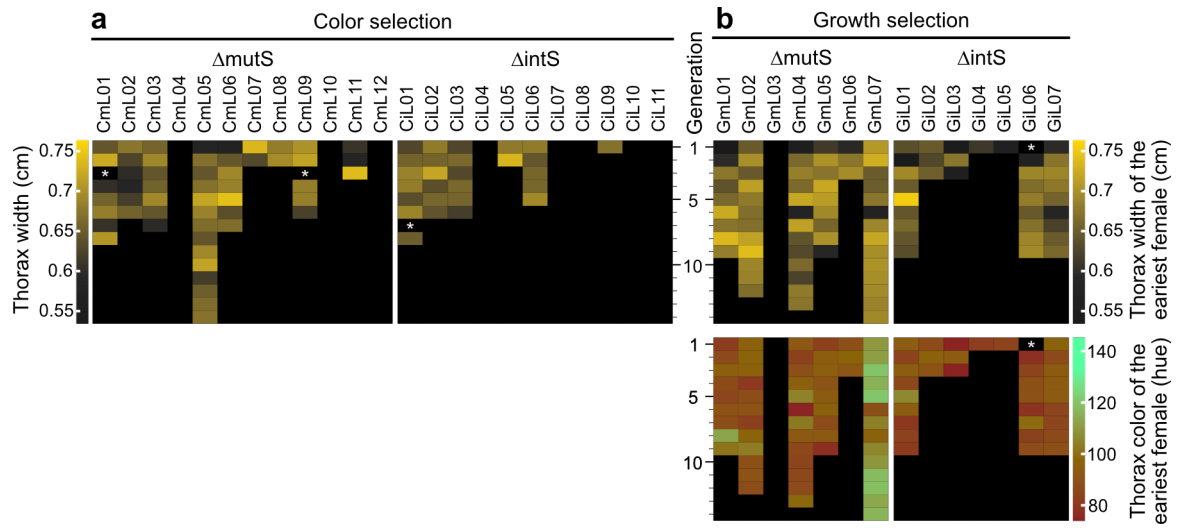
646

647

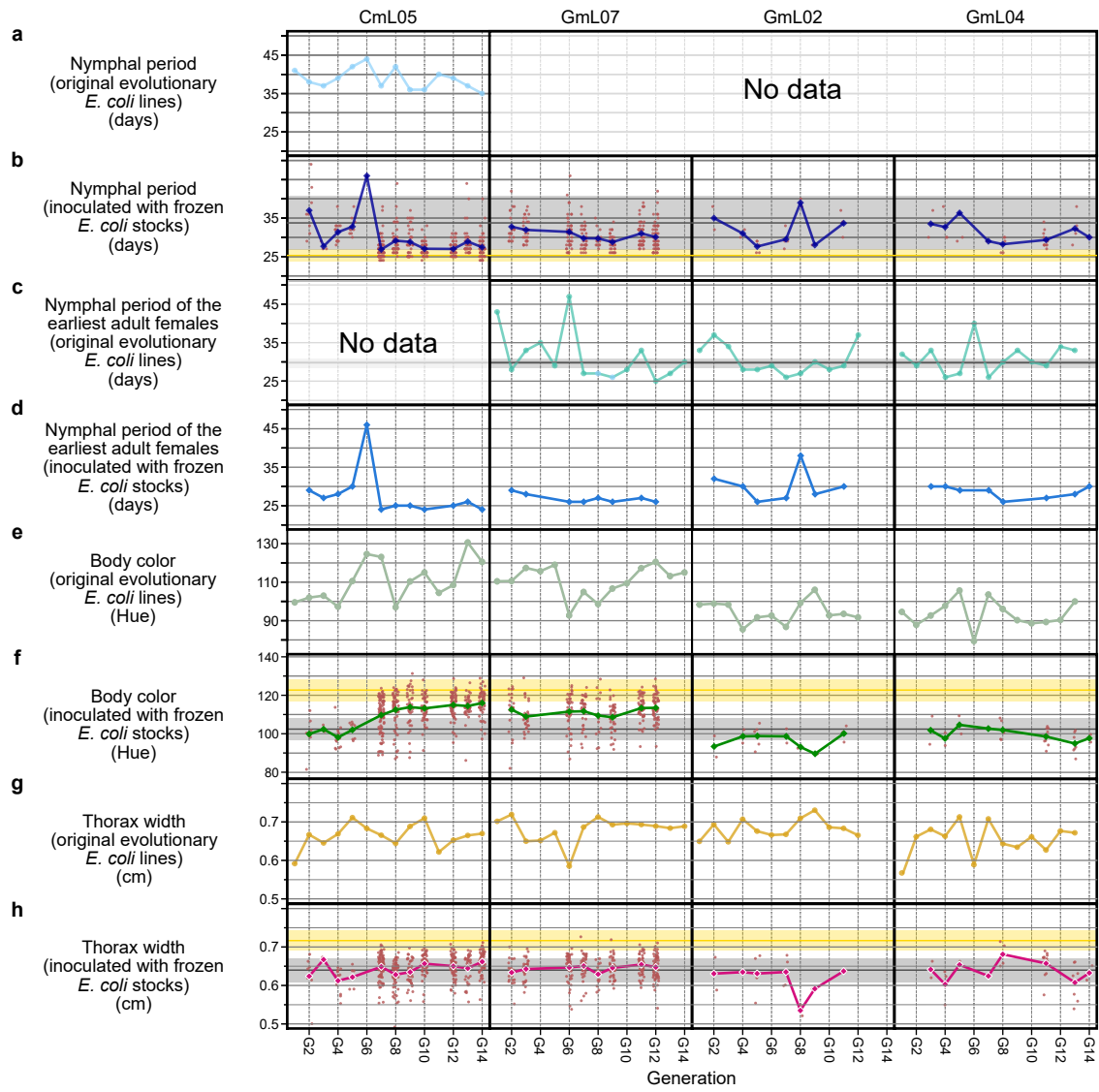


651 **Fig. S4. Effects of evolutionary *E. coli* lines on body size and color of *P. stali*.** (a)  
652 Evolutionary *E. coli* lines subjected to host's body color selection. Data of host's body width  
653 are displayed by heat maps. Also see [Fig. 2a](#). (b) Evolutionary *E. coli* lines subjected to host's  
654 growth speed selection. Data of host's body width and color are displayed by heat maps. Also  
655 see [Fig. 2b](#).  
656  
657

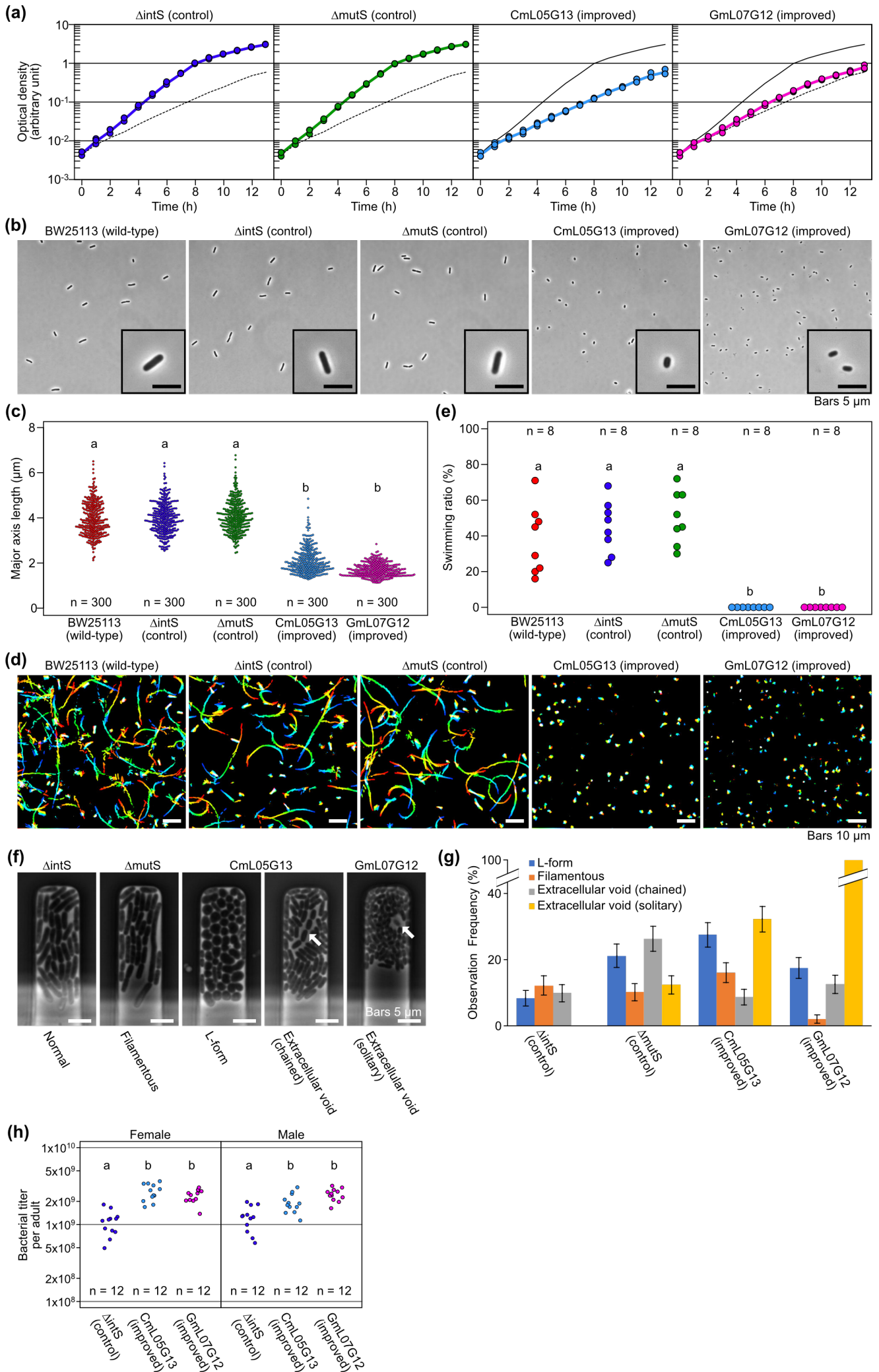




661 **Fig. S5. Adult phenotypes of *P. stali* infected with the evolutionary *E. coli* lines CmL05,**  
662 **GmL07, GmL02 and GmL04. (a, b) Nymphal period. (c, d) Nymphal period of the earliest**  
663 **adult females. (e, f) Body color. (g, h) Thorax width. (a, c, e, g) Phenotypes of adult insects**  
664 **infected with the original evolutionary *E. coli* lines. (b, d, f, h) Phenotypes of adult insects**  
665 **inoculated with the frozen *E. coli* stocks. Line charts show mean values while dots indicate**  
666 **individual data points. Note that, corresponding to each original evolutionary *E. coli* line, three**  
667 **insect groups were inoculated with the frozen *E. coli* stock.**  
668  
669  
670

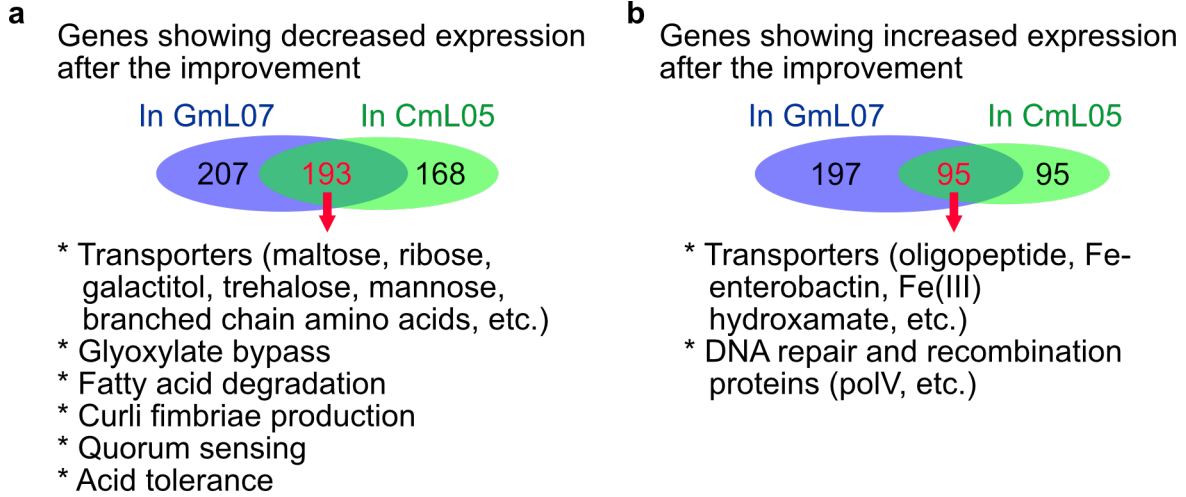


674 **Fig. S6. Microbial traits of evolutionary *E. coli* lines CmL05 and GmL07 in comparison**  
675 **with original *E. coli* strains BW25113,  $\Delta$ intS and  $\Delta$ mutS cultured in liquid medium. (a)**  
676 Growth curves (3 replicates each). Upper solid line is the trace of  $\Delta$ intS growth curve, whereas  
677 lower dotted line is the trace of CmL05 growth curve. (b) Morphology of bacterial cells. (c)  
678 Quantification of cell size in terms of major axis length. (d) Motility of bacterial cells visualized  
679 by rainbow plot for 2 sec. (e) Quantification of bacterial motility in terms of number of  
680 swimming cells per 100 cells observed. (f) Characteristic cellular shape and growth mode in  
681 microfluidic channels. From left to right, the micrographs show the microchannels harboring  
682 *E. coli* cells with normal rod-like shape ( $\Delta$ intS), filamentation shape ( $\Delta$ mutS), L-form-like  
683 round shape (CmL05), extracellular void space and chained growth (CmL05), and extracellular  
684 void space and solitary growth (GmL07). Arrows indicate the cells showing the extracellular  
685 void space. (g) Frequency of the microchannels in which *E. coli* cells exhibited characteristic  
686 cell shape and growth mode. The total numbers of microchannels observed in the time-lapse  
687 measurements were 131 ( $\Delta$ intS), 137 ( $\Delta$ mutS), 149 (CmL05G13), and 143 (GmL07G12). (h)  
688 Bacterial titers in adult females 35 days after emergence in terms of *ntpII* gene copies per insect.  
689 In (c), (e) and (h), different alphabetical letters indicate statistically significant differences  
690 (pairwise Wilcoxon rank sum test:  $P < 0.05$ ).  
691  
692

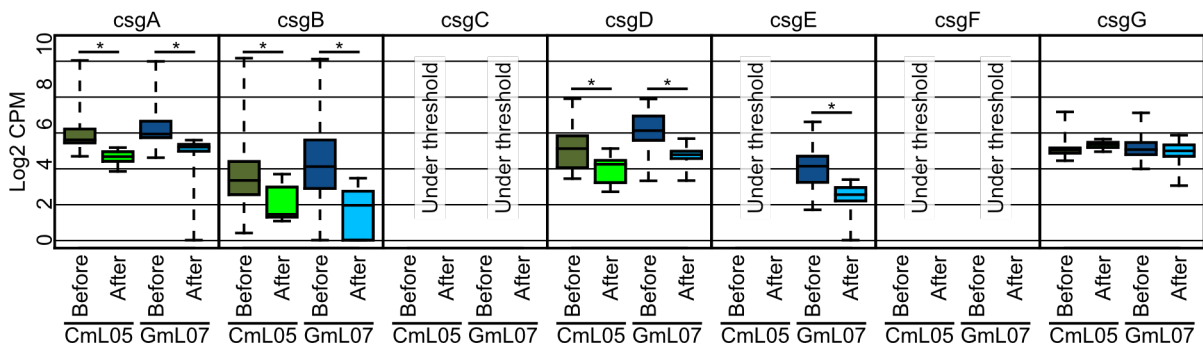


696 **Fig. S7. Gene expression changes of evolutionary *E. coli* lines GmL07 and CmL05 before**  
697 **and after improvement of host phenotypes. (a, b)** Venn diagrams showing down-regulated  
698 genes **(a)** and up-regulated genes **(b)** after the improvement of host phenotypes. **(c)** Expression  
699 levels of genes involved in extracellular matrix (Curli fimbriae) production before and after the  
700 improvement of host phenotypes. Asterisks indicate statistically significant differences (FDR  
701 q-value < 0.01).  
702  
703  
704

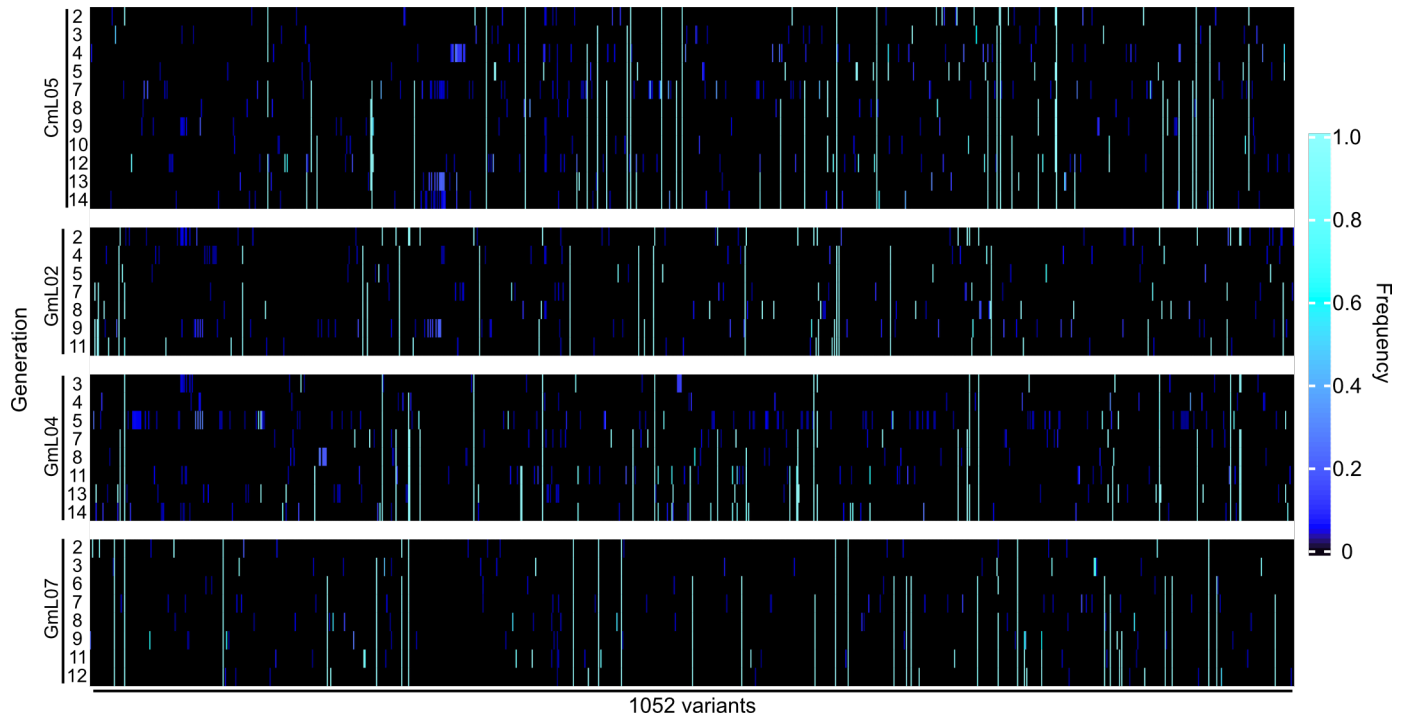




**c** Curli fimbriae production



708 **Fig. S8. Mutations in the genomes of evolutionary *E. coli* lines CmL05, GmL02, GmL04**  
709 **and GmL07 in the experimental evolutionary course.** Frequencies of 1,052 variants  
710 identified in the experimental evolution lines and generations are color-coded. Vertical axis  
711 represents the generations of the experimental evolution lines whereas horizontal axis  
712 represents an array of 1,052 variants. This figure is the graphical representation of [table S7](#).  
713  
714  
715



719 **Fig. S9. Carbon catabolite repression (CCR) pathway and Crp-cAMP regulon of *E. coli*.**

720 **(a)** CCR pathway repressed in the presence of glucose. **(b)** CCR pathway activated in the

721 absence of glucose. **(c)** Number of genes constituting the Crp-cAMP regulon of *E. coli*

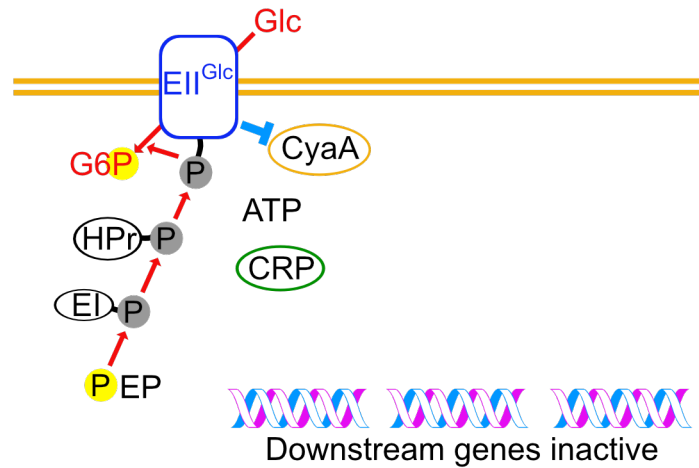
722 estimated by RegulonDB (20).

723

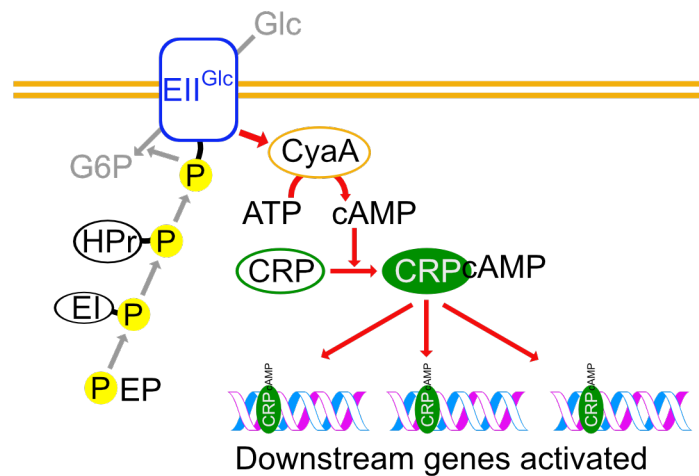
724

725

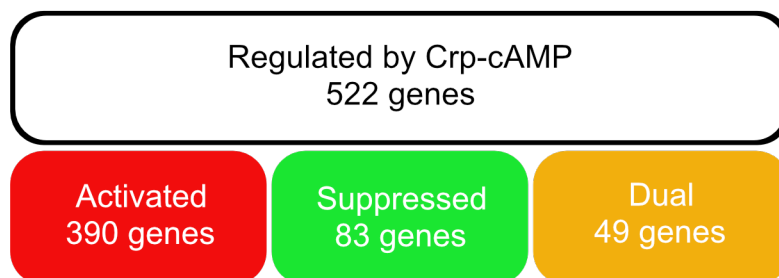
### a CCR pathway repressed in the presence of glucose



### b CCR pathway activated in the absence of glucose

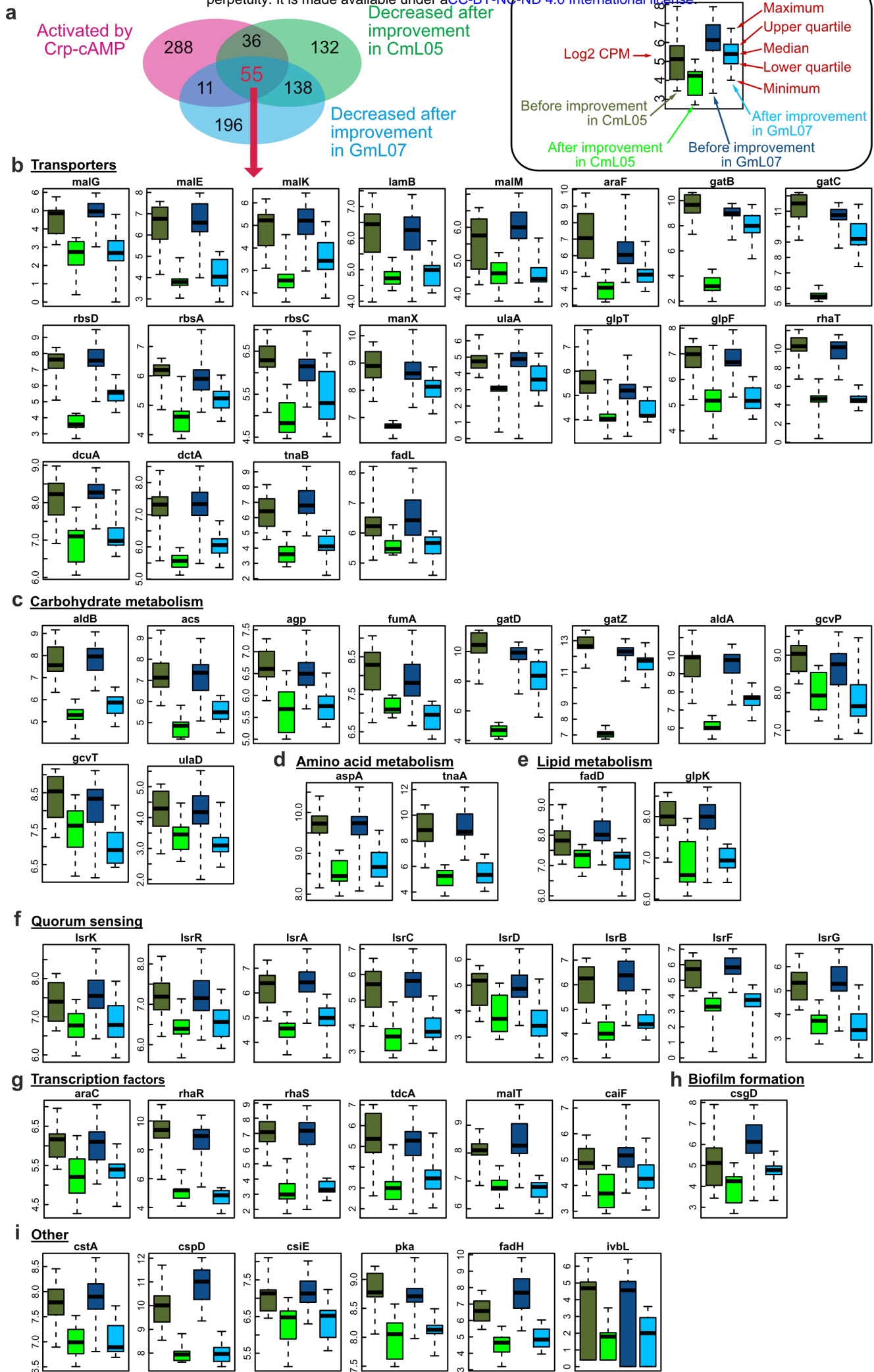


### c Crp-cAMP regulon genes in *E. coli*



729 **Fig. S10. Genes commonly down-regulated in GmL07 and CmL05 after the**  
730 **improvement of host phenotypes, and also down-regulated by disruption of Crp-cAMP**  
731 **in *E. coli*.** (a) Venn diagram showing the commonly down-regulated genes. (b-i) Expression  
732 levels of the commonly down-regulated genes in GmL07 and CmL05 after the improvement  
733 of host phenotypes. (b) Transporter genes. (c) Carbohydrate metabolism genes. (d) Amino  
734 acid metabolism genes. (e) Lipid metabolism genes. (f) Quorum sensing genes. (g)  
735 Transcription factor genes. (h) Biofilm (= Curli fimbriae) formation genes. (i) Other genes.  
736  
737  
738





742 **Fig. S11. Phenotypic traits of  $\Delta$ cyaA and crp<sup>221T>C</sup> mutants of *E. coli*.** (a) Growth curves  
743 (3 replicates each). Upper solid line is the trace of  $\Delta$ intS growth curve, whereas lower dotted  
744 line is the trace of CmL05 growth curve. (b) Morphology of bacterial cells. (c) Quantification  
745 of cell size in terms of major axis length. (d) Motility of bacterial cells visualized by rainbow  
746 plot for 2 sec. (e) Quantification of bacterial motility in terms of number of swimming cells  
747 per 100 cells observed. (f) Bacterial titers in adult females 35 days after emergence in terms  
748 of ntpII gene copies per insect. In (c), (e) and (f), different alphabetical letters indicate  
749 statistically significant differences (pairwise Wilcoxon rank sum test:  $P < 0.05$ ).

750

751

752

