

1 **Multiplexed competition in a synthetic squid light organ microbiome using barcode-**
2 **tagged gene deletions**

3

4 Hector L. Burgos¹, Emanuel F. Burgos¹, Andrew J. Steinberger², Garret Suen², and Mark J.
5 Mandel^{1,*}

6

7 ¹Department of Medical Microbiology and Immunology, University of Wisconsin-Madison,
8 Madison, WI USA

9 ²Department of Bacteriology, University of Wisconsin-Madison, Madison, WI USA

10

11 Short title: *Vibrio fischeri* barcoded deletions and BarSeq

12

13 Keywords: Barcode sequencing, amplicon sequencing, sequence-tagged gene deletions, *Vibrio*
14 *fischeri*, *Aliivibrio fischeri*

15

16 * Correspondence to:

17 Mark J. Mandel

18 University of Wisconsin-Madison

19 Department of Medical Microbiology and Immunology

20 1550 Linden Drive

21 Madison, WI 53706

22 Phone: (608) 261-1170

23 Fax: (608) 262-8418

24 Email: mmandel@wisc.edu

25 **ABSTRACT**

26 Beneficial symbioses between microbes and their eukaryotic hosts are ubiquitous and have
27 widespread impacts on host health and development. The binary symbiosis between the
28 bioluminescent bacterium *Vibrio fischeri* and its squid host *Euprymna scolopes* serves as a
29 model system to study molecular mechanisms at the microbe-animal interface. To identify
30 colonization factors in this system, our lab previously conducted a global transposon insertion
31 sequencing (INSeq) screen and identified over 300 putative novel squid colonization factors in
32 *V. fischeri*. To pursue mechanistic studies on these candidate genes, we present an approach
33 to quickly generate barcode-tagged gene deletions and perform high-throughput squid
34 competition experiments with detection of the proportion of each strain in the mixture by
35 barcode sequencing (BarSeq). Our deletion approach improves on previous techniques based
36 on splicing-by-overlap extension PCR (SOE-PCR) and *tfoX*-based natural transformation by
37 incorporating a randomized barcode that results in unique DNA sequences within each deletion
38 scar. Amplicon sequencing of the pool of barcoded strains before and after colonization faithfully
39 reports on known colonization factors and provides increased sensitivity over colony counting
40 methods. BarSeq enables rapid and sensitive characterization of the molecular factors involved
41 in establishing the *Vibrio*-squid symbiosis and provides a valuable tool to interrogate the
42 molecular dialogue at microbe-animal host interfaces.

43

44 **IMPORTANCE**

45 Beneficial microbes play essential roles in the health and development of their hosts.
46 However, the complexity of animal microbiomes and general genetic intractability of their
47 symbionts have made it difficult to study the coevolved mechanisms for establishing and
48 maintaining specificity at the microbe-animal host interface. Model symbioses are therefore
49 invaluable for studying the mechanisms of beneficial microbe-host interactions. Here we present
50 a combined barcode-tagged deletion and BarSeq approach to interrogate the molecular
51 dialogue that ensures specific and reproducible colonization of the Hawaiian bobtail squid by
52 *Vibrio fischeri*. The ability to precisely manipulate the bacterial genome, combined with multiplex
53 colonization assays, will accelerate the use of this valuable model system for mechanistic
54 studies of how environmental microbes—both beneficial and pathogenic—colonize specific
55 animal hosts.

56

57

58 INTRODUCTION

59 Beneficial symbioses are ubiquitous in the environment and have substantial impacts on the
60 health and development of animal hosts. In animals, symbionts can affect host organ
61 morphogenesis, immune system development, reproduction, susceptibility to disease, and even
62 behavior (1–4). In humans, the gut, skin, lungs, and urogenital tract all have specific
63 microbiomes for which their dysbiosis has been associated with disease (5–8). It is clear that
64 molecular communication between animal hosts and their microbial partners leads to selection
65 and retention of the cognate microbiome: while many microbes are obtained from the
66 environment, the composition of mature microbiomes is often largely stable and resilient within
67 members of a host species (9, 10). While microbial communities have been characterized using
68 metagenomic, transcriptomic, and metabolomic approaches (11), the complexity of animal-
69 associated microbiomes and the inability to culture and genetically-manipulate many symbionts
70 make it difficult to study the precise molecular mechanisms that establish specific relationships.

71 The binary symbiosis between genetically-tractable *Vibrio fischeri* and the Hawaiian bobtail
72 squid *Euprymna scolopes* serves as a model system to study the molecular interactions
73 underlying microbiome assembly (11–16). The squid hatch aposymbiotically (without symbiont)
74 and are colonized by *V. fischeri* in a multi-step process that leads to the specific recruitment of
75 the symbiont from a marine environment in which the bacteria are < 0.1% of the
76 bacterioplankton (14, 17). The symbionts are housed in the dedicated light organ (LO) within the
77 squid's mantle cavity, where they generate light that the host uses for counterillumination to hide
78 its shadow while hunting at night (18). The host provides the symbionts with a protected niche,
79 nutrients, and oxygen (15). Once the symbiosis is irreversibly established in juvenile squid, a
80 daily cycle proceeds where 90-95% of the bacteria are expelled from the LO at dawn. The
81 remaining symbionts grow during the day until they fill the LO, and at night the dense population
82 of symbionts provides light (19). Because the aposymbiotic hatchlings can be cultured in the lab
83 and infected with genetically-tractable *V. fischeri*, colonization experiments can be performed to

84 study the molecular factors that play a role during this process (12, 16, 20). In addition, the
85 translucent nature of the LO in squid hatchlings allows for visualization of the colonization
86 process by microscopy (21–23).

87 Microbe-host signaling mechanisms and developmental transitions ensure specificity during
88 colonization (14, 17). Upon detection of bacterial-derived peptidoglycan, the ciliated
89 appendages on the surface of the LO secrete mucus that traps bacteria circulating within the
90 mantle cavity (13, 24, 25). In the mucus field, *V. fischeri* bacteria bind to cilia and form
91 aggregates by expressing symbiosis polysaccharide (*syp*) genes, a locus of 18 structural and
92 regulatory genes whose products contribute to biofilm formation (26, 27). Approximately 3-4
93 hours post-inoculation, bacterial aggregates migrate through the host mucus toward the pores
94 that lead into the LO ducts (25). While the initial migration is independent of flagellar motility
95 (28), at the pores squid-produced chitin oligosaccharides serve as a chemoattractant to direct
96 the symbiotic bacteria into the host crypts (21). Motility and chemotaxis are required for
97 colonization, and strains with mutations in genes required for these processes—such as *cheA*,
98 *flrA*, and *rpoN*—are unable to successfully colonize the squid LO (21, 28). Once within the LO
99 *V. fischeri* generates light through expression of the *lux* operon in a quorum sensing-dependent
100 process (29). Symbionts that fail to produce luminescence, such as strains with mutations in the
101 autoinducer synthases *ainS* or *luxI*, or deletions of the *lux* operon, are unable to persist in the
102 symbiosis (30, 31).

103 To identify novel colonization factors in *V. fischeri* our lab previously used a global
104 transposon insertion sequencing approach (INSeq) to identify bacterial mutants that were
105 depleted after 48 hours in the squid host (32). This approach successfully identified previously-
106 known colonization factors, such as *rscS*, *rpoN*, *ompU*, various motility factors, and the *syp*
107 biofilm locus, and in addition revealed 344 putative novel colonization factors. Twenty
108 candidates were tested in competitive colonization assays of wild-type (WT) vs mutant strains
109 and the results showed that nine factors had colonization defects. Some of the validated factors

110 encompass roles in protein quality control (DnaJ and DegS) and copper detoxification (CopA
111 and CusC), inner membrane proteins predicted to play a role in secretion of autotransporters
112 (TamB/YtfN), and other poorly characterized factors (YdhC, YafD, and YhcB). This global
113 approach was crucial in identifying putative colonization factors. However, further study is
114 required to address which genes are true colonization factors, when they act during
115 colonization, and how their products modulate the interaction with the host.

116 Approximately 32% of putative colonization factors identified by INSeq did not fall into a
117 curated Clusters of Orthologous Groups (COGs) category, suggesting that the ability to
118 interrogate the function of these colonization factors will reveal novel biology. Traditional genetic
119 engineering techniques in *V. fischeri* are either random (transposon mutagenesis) or labor-
120 intensive (plasmid-based allelic exchange) (32–34). We therefore considered approaches by
121 which we could isolate mutants and examine phenotypes in a multiplexed fashion. One possible
122 approach was to retrieve transposon insertions of interest from an arrayed library (35–37). A
123 second approach we considered was to adapt a newly-developed method for transformation-
124 mediated mutagenesis using linear DNA (38) with an in-frame barcoding strategy to facilitate
125 precise mutations. This latter option was attractive in that we hoped that it would limit the effects
126 of polar mutations and provide a set of defined deletions that can be characterized by amplicon
127 PCR. Barcode sequencing (BarSeq), in which each strain is uniquely labeled and identified
128 using high-throughput next-generation sequencing, has been used successfully to track
129 population dynamics in multiple systems, including in yeast genomic libraries, during *Vibrio*
130 *cholerae* infection, and to track and phenotype laboratory-evolved *Escherichia coli* (39–43).
131 Here, we describe an approach to generate barcode-tagged gene deletions in *V. fischeri* and
132 perform high-throughput colonization experiments using BarSeq. We also describe the barseq
133 python computational package used to analyze the results.

134 RESULTS

135 **Generation of barcoded gene deletions.** To generate barcoded deletions of specific *V.*
136 *fischeri* genes, we designed an approach that takes advantage of splicing-by-overlap extension
137 PCR and *tfoX*-based natural transformation (Fig. 1) (38, 44–46). The first step uses PCR to
138 amplify DNA fragments upstream and downstream of the gene targeted for deletion, fused to
139 the left and right linker sequences, respectively (Fig. 1A). A separate PCR is performed with
140 pHB1 as a template to generate the central fragment of DNA containing the linker sequences, a
141 selectable marker (*erm*, conferring erythromycin resistance) surrounded by FRT sites, and the
142 semi-random barcode sequence. The barcode is provided by the reverse primer, which contains
143 a region of semi-randomized-sequence. The three resulting DNA fragments—upstream, central,
144 and downstream—are then fused into one fragment via their overlapping linker sequences by
145 SOE-PCR (46) and transformed into *V. fischeri* upon *tfoX* induction (44). Finally, the *erm*
146 cassette is removed via FLP recombinase (45).

147 The resulting 138-bp deletion scar (*barcode* scar, or “*bar*” scar, Fig 1B) lies between the
148 deleted gene’s first codon and last seven codons (i.e., the final six amino acid-encoding codons
149 plus the stop codon). The scar is designed to be in-frame to prevent polar effects on gene
150 expression when targeting genes within operons. The terminal codons were retained in case
151 they contain a ribosomal binding site for downstream gene(s) (47). In addition to the barcode,
152 the additional sequence in the scar includes left and right linker sequences that are shared
153 among all of the mutants, which allows us to identify and quantify the abundance of each
154 barcoded-strain using amplicon sequencing, while minimizing amplification bias by using
155 common primers that amplify the same size product.

156 To test this new approach, we investigated the *copA* gene. Among Gammaproteobacteria,
157 CopA is the main exporter of cytoplasmic copper and is the most widely conserved copper
158 detoxification factor (48, 49). Although our laboratory previously demonstrated that *copA* is a
159 squid colonization factor, its role in copper resistance has not been examined (32). We therefore
160 targeted *copA* for deletion using our mutagenesis approach as a proof-of-concept and

161 subsequently tested its role in copper resistance in *V. fischeri*. To ensure that the deletion
162 process worked as intended, we used four sets of diagnostic PCR primers that would report on
163 correct *erm* insertion, subsequent removal of the *erm* cassette by FLP recombinase, and
164 absence of the targeted gene from the bacterial chromosome. PCR with various pairs of
165 oligonucleotides that target the *copA* gene and its deletion constructs produced amplicons of the
166 expected size in each strain (Fig. 2A). These results show that the *erm* cassette was
167 successfully inserted into *copA* generating $\Delta copA::erm\text{-}bar$, and subsequently removed by FLP
168 recombination to generate the in-frame deletion scar in $\Delta copA::bar$. Furthermore, sequencing of
169 the deletion scar for several $\Delta copA::erm\text{-}bar$ candidates showed that after a single round of
170 mutagenesis, multiple uniquely barcoded deletion strains were generated (Fig. 2B). These
171 results demonstrate that our deletion method is successful in generating uniquely barcoded
172 mutant strains of *V. fischeri*.

173 **The presence of a barcode within a gene deletion does not alter mutant phenotypes.**

174 To test that the barcoded scar does not affect the mutant phenotypes, we measured the copper
175 sensitivity of strains deleted for *copA* using various methods. In addition to the mutants
176 generated using our deletion approach ($\Delta copA::erm\text{-}bar$ and $\Delta copA::bar$) we constructed a
177 deletion of *copA* using plasmid-based allelic exchange ($\Delta copA$) (33) and obtained a *copA*
178 transposon mutant isolated from our previous study ($\Delta copA::Tnerm$) (32). We then tested the
179 growth of these various *copA* mutants in the presence of varying amounts of copper. Our results
180 show that, regardless of the mutagenesis method, the growth of *copA* mutants is similarly
181 impeded in the presence of copper, with the severity of the growth defect increasing in
182 proportion to the concentration of copper: at 0.2 μM Cu^{2+} the *copA* mutants were able to grow
183 slightly, whereas at 20 μM Cu^{2+} these strains were unable to grow (Fig. 3A). This is in contrast
184 to the WT strain that achieved the same growth yield regardless of the concentration of copper
185 present. The $\Delta copA::erm\text{-}bar$ and $\Delta copA::bar$ mutants showed the same degree of copper
186 sensitivity (Fig. 3A).

187 To corroborate that the observed growth defects were due specifically to excess copper, we
188 measured the growth of $\Delta copA::bar$ in the presence of copper, with and without the copper
189 chelator bathocuproinedisulfonic acid (BCS). As expected, $\Delta copA::bar$ was unable to grow in
190 the presence of 20 μM Cu^{2+} , whereas the WT is unaffected (Fig. 3B). However, growth of
191 $\Delta copA::bar$ in the presence of copper was rescued by addition of 80 μM of BCS (Fig. 3B),
192 suggesting that free copper is indeed responsible for the observed lack of growth in the mutant.
193 To verify that the absence of CopA was responsible for susceptibility to copper toxicity, we
194 complemented *copA* at the chromosomal *attTn7* site in the $\Delta copA::bar$ strain and observed that
195 growth was rescued in the presence of copper (Fig. 3B). Based on these results, we conclude
196 that CopA is required for resistance to copper in *V. fischeri*, consistent with its function in other
197 Gammaproteobacteria (48).

198 To further test our deletion method, we generated mutants in multiple genes required for *V.*
199 *fischeri* motility—*rpoN*, *flrA*, and *flaA*—and tested the resulting strains' motility phenotypes on
200 soft agar plates (28, 50, 51). We also included a WT strain tagged with the deletion scars at the
201 *attTn7* site (WT-1) and *copA* mutants as controls. While motility of WT *V. fischeri* resulted in a
202 migration disc with a diameter of 26 mm from the inoculation point on soft agar plates, deleting
203 *flaA* resulted in a drastic reduction in migration (9 mm), while deleting either *flrA* or *rpoN*
204 resulted in no motility (1.5 mm) (Fig. 4). We observed that both the *erm-bar* and *bar* versions of
205 the gene deletions displayed equivalent phenotypes, showing that the strains behave as null
206 alleles regardless of whether the scar contains the *erm* cassette (Fig. 4). Both WT-1 and *copA*
207 strains have motility comparable to WT, showing that motility defects are due to the deleted loci
208 and not to the insertion of the deletion scars.

209 **Removing the erythromycin-resistance cassette minimizes polar effects of the**
210 **barcoded deletions.** While the presence or absence of the *erm* cassette does not prevent
211 deletion strains from manifesting the corresponding phenotypes (Fig. 3A and 4), we were
212 concerned about polar effects on downstream gene expression upon insertion of the 1,049 bp
213 heterologous *erm* cassette (52–54). To test the effect of the *erm* cassette on downstream gene

214 expression we used reverse transcriptase quantitative PCR (RT-qPCR) to measure expression
215 of genes immediately upstream and downstream of a targeted gene deletion for three different
216 predicted operons. In each case, we measured the ratio in expression levels of the downstream
217 vs upstream genes in the mutant, normalized to the ratio in WT *Vibrio fischeri* (defining this
218 normalized value as the “polarity ratio”). For both *rpoN* and *cheA*, deletion scars of either *erm-*
219 *bar* or *bar* resulted in negligible changes in the polarity ratio (Fig. 5). In contrast, the polarity
220 ratio of $\Delta cusA::erm\text{-}bar$ was 26-fold higher than WT, whereas removal of the *erm* cassette to
221 form the in-frame deletion scar restored the polarity ratio to basal levels (Fig. 5). We conclude
222 that, in at least some cases, *gene::bar* deletion scars can alleviate collateral effects on flanking
223 genes that are caused by inserting an antibiotic-resistance cassette.

224 **Development of a computational pipeline to analyze *V. fischeri* BarSeq data.** With the
225 ability to quickly generate precise barcoded deletions, we next sought to compete the
226 $\Delta gene::bar$ deletions en masse during host colonization. We therefore developed a BarSeq
227 sample preparation protocol, and an accompanying computational package to analyze the data
228 (Fig. 6). To accomplish this, we mixed barcoded strains to generate an input library (i.e., a
229 synthetic microbiome). This library was then used to inoculate media and/or squid hatchlings,
230 which were then sampled at the desired time points. Samples were then processed to extract
231 gDNA, and PCR was performed with dual-index Illumina sequencing primers to obtain dsDNA
232 fragments containing the barcoded deletion scars. The resulting library was then sequenced on
233 an Illumina MiSeq and demultiplexed based on the unique dual indexes (55). The resulting
234 sequence data was then analyzed using the barseq package, which identifies and counts the
235 barcodes present in the samples, assigns strain identity, normalizes strain counts, and
236 calculates relative frequency and the competitive index (CI) for each strain within the samples.
237 The BarSeq protocol provides a streamlined and effective way to measure population dynamics
238 throughout squid colonization.

239 **BarSeq enables sensitive multiplex competition experiments.** To test our BarSeq
240 protocol in tracking individual strains within a population, we performed an *in vitro* competition

241 and a competitive colonization experiment using an input library of seven different barcoded
242 strains mixed in an equivalent ratio. In addition to several mutant strains, we included three
243 WT::*bar* strains that had the *bar* scar inserted at the Tn7 site that could be similarly tracked
244 using amplicon sequencing but without affecting the phenotypes of the strains (WT-1, WT-2,
245 and WT-3 in Fig. 7). After 15 generations of growth *in vitro*, the proportion of most strains
246 relative to the WT::*bar* strains remained stable except for *flrA* and *rpoN*, which were 4-fold
247 higher and lower, respectively, when compared to WT::*bar* (Fig. 7A). In contrast, following 48 h
248 of squid colonization—which corresponds to approximately 15 bacterial generations (32)—we
249 observed reduced levels of the *flaA* flagellin mutant and severely reduced levels of the *flrA*,
250 *rpoN*, and *cheA* strains, all of which were near the limit of detection (Fig. 7B). This result is
251 consistent with their previously known roles as necessary factors for squid colonization,
252 although we did observe higher levels of *flaA* in the competitive colonization than are observed
253 when a transposon insertion is competed against wild-type (28, 50, 51, 56). We note that there
254 was relatively little variability among the WT::*bar* strains in the analysis, whereas the 4-5 log
255 scale in which to identify colonization defects provided a substantially greater range over which
256 to identify and refine colonization phenotypes *in vivo* (Fig. 7B). Taken together, these results
257 show that our method for targeted barcoded deletions, multiplex squid colonization, and
258 analysis by BarSeq allows for reproducible competition experiments *in vitro* and *in vivo* with high
259 sensitivity.

260 **DISCUSSION**

261 In this study, we developed a method to quickly generate gene deletions where the resulting
262 strains are tagged with unique DNA barcodes. We demonstrated the utility of these strains in
263 performing BarSeq high-throughput competitive colonization experiments and introduced a
264 software package to analyze the resulting sequencing data. BarSeq provides a sensitive
265 method to track population dynamics of squid colonization by *V. fischeri*.

266 **Generation of targeted, barcoded gene deletions that minimize effects on neighboring**
267 **genes.** Our approach builds upon previous SOE-PCR/*tfoX* mutagenesis techniques to
268 incorporate a unique barcode in each deletion strain, which enables high-throughput
269 experiments via barcode sequencing (BarSeq). Since BarSeq relies on amplicon sequencing,
270 library preparation is straightforward and allows for a large number of samples to be processed
271 in parallel. The method we have employed to design the barcode and flanking sequences was
272 planned to minimize disruption on flanking genes. Expression of bacterial genes is frequently
273 organized by their genetic arrangement into operons where expression of operon members is
274 driven by a common promoter (57). However, given that some regulatory regions overlap in
275 neighboring genes, deletion of one gene can alter the expression level of another nearby
276 cistron. These off-target effects on gene expression could obfuscate the analysis of
277 experimental results. Similar to the approach used by Baba et al. (53), our deletion approach
278 reduces off-target effects on gene expression by ensuring the formation of an in-frame open
279 reading frame within the deletion scar and including several codons at the end of the deletion
280 target where the ribosome binding site of downstream genes is frequently located (Fig. 1B and
281 Fig. 5).

282 **BarSeq enables detailed studies of the molecular mechanisms that result in**
283 **establishment of the *Vibrio*-squid symbiosis.** Using an INSeq screen, our lab previously
284 identified 344 putative novel squid colonization factors in *V. fischeri* (32). Our deletion approach,
285 combined with BarSeq, will enable the high-throughput characterization of these factors during
286 squid colonization by allowing multiplexing of colonization factor mutants and tracking of

287 individual strains. By enabling a more precise study of colonization factors, BarSeq has several
288 potential applications.

289 BarSeq can be applied to the study of strain variation and evolution of colonization
290 mechanisms in the *Vibrio*-squid symbiosis. *V. fischeri* strain variation is an important
291 consideration when studying the mechanisms of colonization of the squid LO (58, 59). Previous
292 studies have shown that multiple strains can co-colonize the squid LO and that they do so at
293 different rates (60–62). More recent studies have focused on deciphering the specific
294 mechanisms that result in differing colonization behavior (63, 64). The barcode-tagged
295 mutagenesis method presented here can be applied to generate uniquely-tagged WT or mutant
296 strains of the various phylogenetically-distinct *V. fischeri* strains and assayed in multiplexed
297 format during squid colonization using BarSeq. We have already successfully used our SOE-
298 PCR/*tfoX* mutagenesis approach to make targeted deletions in the ancestral strain SR5,
299 showing that this method is applicable to *V. fischeri* strains that are evolutionarily distant to the
300 frequently used ES114 strain (63).

301 BarSeq can also be used in directed evolution experiments to examine the functional
302 evolution of colonization factors. Directed evolution experiments have recently been applied to
303 study colonization factors in *V. fischeri* (65). The ease of tracking large numbers of individual *V.*
304 *fischeri* strains using BarSeq could enable tracking of strain lineages in long-term evolution
305 experiments, as has been conducted in other organisms (39, 42, 43).

306 **Phenotypes of *rpoN* and *flrA* mutants during competitive growth in media.** Both *rpoN*
307 and *flrA* deletion strains showed a statistically significant 4-fold decrease and increase,
308 respectively, during competitive growth in media compared to WT (Fig. 7A). Due to the nature of
309 the factors they encode, the observed growth defects are likely due to changes in energetic and
310 nutritional requirements when *rpoN* and *flrA* are deleted. The *rpoN* gene encodes the alternative
311 σ^{54} factor that is responsible for expression of various systems involved in squid colonization,
312 including Syp biofilm formation, flagellar motility, and luminescence (51, 66, 67). Therefore, it is
313 not surprising that deleting the gene encoding σ^{54} has pleiotropic effects on gene expression

314 due to misregulation of the RpoN regulon and could reduce the ability of the mutant strain to
315 effectively compete for growth *in vitro*, though further experiments are necessary to define the
316 precise mechanism for the defect. FlrA is the σ^{54} -dependent transcription factor that activates
317 expression of the flagellar biosynthesis cascade and is required for motility and squid
318 colonization (50). The high energetic cost of expressing all genes related to flagellar
319 biosynthesis (68), which in *V. cholerae* requires FlrA-dependent regulation of 52 genes (69) and
320 in *V. fischeri* between 39 and 131 genes (28), is consistent with the observed increase in growth
321 of the *flrA* deletion strain compared to WT during competitive growth in media (Fig. 7A).
322 Nonetheless, even though the changes observed in competitive growth of the *rpoN* and *flrA*
323 mutants *in vitro* are in opposing directions (less vs more growth, respectively), both are severely
324 defective in squid colonization (Fig. 7B). Future experiments using BarSeq to probe bacterial
325 growth *in vitro* and during colonization have the potential to elucidate heretofore hidden
326 phenotypes.

327 **Discrepancy in *flaA* colonization efficiency measured by BarSeq vs traditional**
328 **competitive squid colonization experiments.** In our BarSeq experiment, the known
329 colonization factor *flaA* only shows a small (~2-fold) colonization defect after 48 hr post-squid
330 inoculation (Fig. 7B). In contrast, Millikan and Ruby showed that a *flaA* deletion made by
331 insertion of a Kan^R-cassette is severely defective during competitive colonization against WT *V.*
332 *fischeri* (56). Using confocal microscopy, their work showed that LO colonization by *flaA* is
333 delayed compared to WT by ~8 hr. However, because our competitive colonization experiment
334 using BarSeq was done at 48 hours post-inoculation, this delayed colonization is not enough to
335 explain the observed discrepancy. Previous work has shown that the concentration of *V. fischeri*
336 in the inoculum can affect the number of different strains that can co-colonize the squid LO (70).
337 This raises the possibility that the inoculum amount or the ratio of strains within the synthetic
338 microbiome might affect the observed colonization defect. To address this, future experiments
339 should examine how inoculum amount and the ratio of mutant strains to WT within a synthetic

340 barcode-tagged population affects colonization efficiency for the different strains in the
341 population.

342 In summary, we provide a new method for constructing barcoded deletions of *V. fischeri*
343 genes; demonstrate the utility of this method for generating in-frame deletions and discovering
344 new functions of squid colonization factors; and combine this approach with a computational
345 tool to conduct multiplex animal colonization assays using barcode sequencing (BarSeq).

346 **MATERIALS AND METHODS**

347 **Bacterial strains, growth conditions, plasmids, and primers.** Bacterial strains used in
348 this study are listed in Table 1, with Table S1 containing an Extended Table 1 showing the
349 oligos used to generate the specified barcode-tagged gene deletions. Plasmids are listed in
350 Table 2, and DNA oligonucleotides are listed in Table S2. DNA oligonucleotides were
351 synthesized by Integrated DNA Technologies (Coralville, IA), and Sanger DNA sequencing was
352 performed through the University of Wisconsin–Madison Biotechnology Center. *Escherichia coli*
353 strains were grown in Luria-Bertani (LB) medium [per liter, 25 g Difco LB broth (BD), in distilled
354 water] at 37°C with aeration. Unless otherwise indicated, *V. fischeri* strains were grown in Luria-
355 Bertani salt (LBS) medium [per liter, 25 g Difco LB broth (BD), 10 g NaCl, and 50 ml 1 M Tris
356 buffer, pH 7.0, in distilled water] at 25°C with aeration. When necessary, growth media was
357 solidified by adding 15 g Bacto agar (BD) per liter. For growth of *V. fischeri*, antibiotics (Gold
358 Biotechnology) were added at the following concentrations: 5 µg/ml erythromycin, 5 µg/ml or 2.5
359 µg/ml chloramphenicol as indicated, and 100 µg/ml kanamycin. For *E. coli* the antibiotic
360 concentrations used were 100 µg/ml carbenicillin, 25 µg/ml chloramphenicol, and 50 µg/ml
361 kanamycin. The *E. coli* strain π3813 containing pKV496 is a thymidine auxotroph and was
362 grown in LB with 50 µg/ml kanamycin supplemented with 0.3 mM thymidine (38, 71).

363 The unmarked deletion of *copA* in MJM1100 was made by allelic exchange as described
364 previously (63). Briefly, 1.6 kb upstream (US) and 1.6 kb downstream (DS) sequences of *copA*
365 were amplified by PCR using oligos HB44 and HB45, and HB46 and HB47, respectively, and
366 were cloned using Gibson Assembly (NEBuilder HiFi DNA assembly cloning kit) into the
367 linearized vector pEVS79 (linearized using oligos HB52 and HB53) (Table S2). The Gibson mix
368 was transformed into NEB5α chemically-competent cells and selected on chloramphenicol. The
369 resulting pEVS79-Δ*copA* candidates were screened using PCR with oligos HB54 and HB55 and
370 confirmed by sequencing, generating pHB3, which was conjugated into *V. fischeri* MJM1100
371 (ES114) via triparental mating with MJM534, which contains the helper plasmid pEVS104 (33).
372 Single recombinants of pHB3 into the chromosome were screened and selected by growth on

373 chloramphenicol (MJM3400), and double recombinants by loss of the antibiotic resistance
374 cassette and *copA* (MJM3401). The resulting constructs were verified by PCR and sequencing
375 (Table S2).

376 The *copA* gene was inserted into the *attTn7* site in the chromosome using pEVS107 (70).
377 The *copA* gene including 191 bp US and 321 bp DS sequences was amplified by PCR using
378 oligos HB27 and HB34, the product was digested with *Ascl*, and cloned into the *Ascl* site of
379 pEVS107. The resulting plasmid, pHB2 (pEVS107-*copA*), was transformed into and maintained
380 in *E. coli* DH5 α λ *pir* cells and verified by PCR (Table S2) and sequencing. pHB2 was conjugated
381 into Δ *copA* (MJM3401) via tetraparental mating with donor MJM3288 (DH5 α λ *pir*/pHB2), helper
382 strains MJM637 (S17-1 λ *pir*/pUX-BF13) (72, 73) and MJM534 (CC118 λ *pir*/pEVS104) (33), and
383 the recipient MJM3543 (Δ *copA*::*bar*), resulting in MJM3790 (Δ *copA*::*bar attTn7::copA*).
384 Candidates were confirmed by PCR (Table S2) and sequencing.

385 **Construction of barcode-tagged gene deletions.** The deletion protocol demonstrated in
386 Fig 1A is based on splicing by overlap extension PCR (SOE-PCR) and *tfoX* transformation (38,
387 44–46) to directly delete and tag targeted genes with a randomized sequence (barcode). Our
388 protocol was in development prior to publication of the previous method (38), so while it is
389 conceptually similar, the sequences of the linkers and primers are distinct. First, several oligos
390 were designed specific to the targeted genes to amplify 1 kb of US (F1 and R1-LL) and DS (F2-
391 RL and R2) DNA tagged with the left linker (LL) and right linker (RL) sequences, respectively, to
392 screen the deletion scar via PCR (FO and RO), and to assay for the absence of the targeted
393 gene (FW and RW) (Fig. 1A, Table S1, and Table S2). FW and RW were designed to amplify a
394 fragment of 500-1,000 bp, depending on the size of the gene. The F1 and R2 oligos were
395 designed to anneal 1 Kb US and DS, respectively, of the targeted gene. The R1 oligo was
396 designed to anneal starting at the start codon of the targeted ORF going upstream, then the
397 reverse complement of the LL sequence (LL reverse complement: 5'-
398 CTGGCGAAGCATATATAAGAAGCTCGTCTCGT-3') was attached to the 5'-end of the R1
399 oligo, resulting in R1-LL. The F2 oligo was designed to anneal at the last 7 codons (6 aa and

400 stop codon) on the 3'-end of the targeted ORF going downstream, then the RL sequence (RL:
401 5'-GACTTGACCTGGATGTCTCTACCCACAAGATCG-3') was attached to the 5'-end of the F2
402 oligo, resulting in F2-RL. The FO and RO oligos (forward outside and reverse outside,
403 respectively) were designed to anneal 500 bp away from the annealing sites of F1 and R2,
404 respectively, and were used to probe the targeted genomic region for insertion of the desired
405 deletion scar.

406 The middle dsDNA fragment containing the *erm* cassette flanked by FRT sites and the
407 randomized barcode was obtained by PCR with Phusion Hot Start Flex 2X master mix (NEB;
408 M0536L) and pHB1 as template, which contains the LL-FRT-*erm*-FRT-spacer sequences and
409 was built as described previously (63), and oligos HB42 and HB154. Oligo HB154 is a reverse
410 primer and contains the RL sequence, 18 bp of randomized sequence composed of 6 trimers of
411 'NNB' to prevent formation of stop codons (results in 'VNN' codons in the forward direction), and
412 the spacer sequence (Table S2). The resulting 1,049 bp product containing LL-FRT-*erm*-FRT-
413 spacer-random barcode-RL was purified by gel extraction using a QIAquick Gel Extraction Kit
414 (Qiagen; 28706). The flanking 1 kb US and DS fragments for each targeted gene were then
415 fused to this middle DNA fragment via the homology between the LL and RL sequences and
416 using SOE-PCR with the F1 and R2 oligos, resulting in the 3 kb mutagenic dsDNA. The reaction
417 mixture contained 10 ng of each the middle, US, and DS DNA fragments, 200 nM of the
418 corresponding F1 and R2 oligos (Table S2), 1X Phusion Hot Start Flex Master Mix (NEB;
419 M0536L) and H₂O up to a total volume of 25 µl. SOE-PCR conditions were 98°C for 30 sec,
420 98°C for 5 sec, 60°C for 20 sec, 72°C for 1 min (30 cycles), with a final extension step at 72°C
421 for 5 min.

422 The 3 Kb mutagenic DNA fragments were purified using a QIAquick PCR Purification Kit
423 (Qiagen; 28106) and transformed into *V. fischeri* ES114 via natural transformation with pLostfoX
424 (MJM1538) (32, 44) where the flanking sequences guide the barcoded *erm* cassette to
425 substitute the targeted gene. Mutant candidates were selected on LBS-Erm5 and screened by
426 PCR with oligo pairs F1/R2, FO/HB8, and FW/RW (as shown in Fig. 2A). The insertion of the

427 *erm-bar* scar was confirmed by Sanger sequencing with primers HB8, HB9, HB42, and HB146,
428 and the unique barcode sequence was recorded for each strain.

429 The final *bar* scars were made by triparental mating of donor MJM3478 (π 3813/pKV496)
430 (38) and helper strain MJM534 (CC118 λ *pir*/pEVS104) into recipient *V. fischeri* strains
431 containing the *erm-bar* scar and selection on LBS-Kan100. Plasmid pKV496 contains the FLP
432 recombinase that removes the *erm* cassette and fuses the two surrounding FRT sites into one,
433 resulting in the final *bar* scar as shown in Fig. 1B. The plasmid was eliminated by growing the
434 candidates on LBS without selection twice and selecting colonies that were Erm^S and Kan^S. The
435 *gene::bar* candidates were screened by PCR using oligo pairs F1/R2, FO/HB146 (RL), and
436 FW/RW, and the deletion scar verified by Sanger sequencing using oligos HB42 and HB146.
437 The barcode sequences were verified to match the barcode within the parental strains
438 containing the *gene::erm-bar* scar.

439 The barcoded WT *V. fischeri* strains (WT::*bar*) were constructed using the same procedure
440 as outlined above for the gene deletions but targeting a site next to the *attTn7* site in the
441 intergenic region of *yeiR* and *glmS*. The 1 kb US and DS arms were amplified using PCR with
442 ES114 gDNA and oligo pairs HB239/HB240 and HB241/HB242. After SOE-PCR to form the
443 mutagenic DNA and *tfoX* transformation, the WT::*erm-bar* candidates were screened by PCR
444 with oligo pairs HB243/HB244 and HB243/HB8. Sanger sequencing was used to confirm
445 insertion of the *erm-bar* scar and record the unique barcode sequences. Triparental mating as
446 described above was performed to remove the *erm* cassette using pKV496. The *bar* scar was
447 confirmed by PCR with HB243/HB146 and Sanger sequencing.

448 **Growth assays in the presence of copper.** Colonies from freshly streaked LBS plates of
449 the indicated *V. fischeri* strains were inoculated into 3 ml LBS with the appropriate antibiotics
450 and grown for 8 hr at 25°C with shaking. Three microliters of the LBS cultures were subcultured
451 into 3 ml Tris minimal medium [per liter, 500 ml DSW (2X), 50 ml 1 M Tris base, pH 7.5, 1 ml
452 5.8% K₂HPO₄, 1 ml 10 mM FeSO₄, and 20 ml 10% *N*-acetylglucosamine (GlcNAc), in distilled
453 water; DSW (2X) = 100 mM MgSO₄, 20 mM CaCl₂, 600 mM NaCl, and 20 mM KCl] and

454 incubated at 25°C overnight for \leq 16 hr. Overnight cultures were diluted to an OD₆₀₀ of 0.5 in
455 200 μ l, then 2 μ l of 0.5 OD₆₀₀ were transferred into 198 μ l of fresh Tris minimal medium
456 containing the appropriate amounts of copper and/or BCS in a 96-well plate. The plate was then
457 incubated in a Synergy Neo2 Multi-Mode Microplate Reader (BioTek) at 25°C with OD₆₀₀
458 measurements every 15 min for 20 hr. Copper stock solutions (100 mM) were prepared from
459 copper (II) sulfate pentahydrate (CuSO₄·5H₂O; Sigma-Aldrich; 203165), and BCS stock
460 solutions (50 mM) from bathocuproinedisulfonic acid disodium salt (Sigma-Aldrich; B1125).

461 **Motility assays.** The indicated bacterial strains were streaked onto fresh LBS plates with
462 the appropriate antibiotics and grown overnight at 25°C. Single colonies were picked with a
463 sterile toothpick and deposited onto OmniTrays (Thermo Fisher Scientific; 242811) containing
464 TBS Agar [per liter, 10 g Gibco Bacto Tryptone (Thermo Fisher Scientific; 211705), 50 ml 1 M
465 Tris buffer, pH 7.0, 20 g NaCl, 8.63 g MgSO₄, and 3 g Agar, in distilled water] by stabbing the
466 toothpick into the media at a single spot. Trays were incubated at 28°C for 4 hr and the outer
467 diameter of swimming cells was measured.

468 **Measuring polarity ratio via RT-qPCR.** The indicated bacterial strains were grown in 3 ml
469 LBS with the appropriate antibiotics and grown at 25°C overnight. On the day of the experiment,
470 15 μ l of the overnight cultures were transferred into 3 ml of fresh LBS and growth was continued
471 at 25°C with aeration. Samples were harvested at an OD₆₀₀ of 0.2-0.4 (mid-log phase) by
472 transferring 800 μ l of culture into a 2 ml screw-cap tube containing 100 μ l of a cold 95% EtOH–
473 5% Phenol solution that inactivates RNases (74). RNA extraction and RT-qPCR were performed
474 as described previously (75). Briefly, cells were lysed in Tris-EDTA (TE) buffer (10 mM Tris-Cl,
475 pH 8.0, 1 mM EDTA) containing lysozyme (Epicentre; R1804M) and 1% SDS. RNA was
476 extracted using the hot phenol method (74) and digested with DNase I (NEB; M0303S).

477 cDNA was synthesized from 0.5 μ g of total RNA using the iScript Advanced cDNA synthesis
478 kit (Bio-Rad; 1725037) following the protocol 25°C for 5 min, 46°C for 20 min, and 95°C for 1
479 min. Quantitative PCR was performed using 1:10 dilutions of cDNA synthesis products with the
480 iTaq Universal SYBR green supermix (Bio-Rad; 1725121) on a CFX Connect real-time PCR

481 detection system (Bio-Rad). The qPCR protocol was 95°C for 30 sec, 95°C for 5 sec, 58°C for
482 30 sec (40 cycles), with a final melt curve analysis to ensure specificity in the reaction. The
483 mRNA levels of *rpoD*, *lptB*, *hpf*, *cheZ*, *cheB*, *cusB* and *cusF* were measured using the oligo
484 pairs listed in Table S2. Expression levels for each gene were normalized to *rpoD* and the
485 mutants were normalized to WT using the $2^{-\Delta\Delta CT}$ method (76). The polarity ratio of *rpoN*, *cheA*,
486 and *cusA* was calculated as “expression of the downstream gene / expression of the upstream
487 gene” using the respective flanking genes in each putative operon—*lptB-rpoN-hpf*, *cheZ-cheA-*
488 *cheB*, and *cusB-cusA-cusF*. Operons were predicted using the BioCyc database for “*Aliivibrio*
489 *fischeri*”, Strain ES114, version 24.1, which is based on the sequenced genome in Mandel et al.
490 (77, 78).

491 **Barseq bioinformatic tool.** To quantify barcodes within each sequenced sample, we
492 developed barseq (<https://github.com/mjmlab/barseq>), a python package that identifies putative
493 barcodes in the sequenced reads and matches them to a user provided barcode library. The
494 program iterates through each sample and uses regular expressions to search within the reads
495 for flanking sequences on the left (GCTCATGCACTTGATTCC; spacer sequence) and the right
496 (GACTTGACCTGGATGTCT; right linker sequence) of the barcode (Fig. 1B), while also allowing
497 for 18 random nucleotides that represent a candidate barcode. The putative barcode sequence
498 is then mapped against the reference barcode library and increases the count for the matched
499 strain. Barseq outputs a tab-delimited table with the barcode/strain counts for each of the
500 samples analyzed.

501 **Barcode Sequencing and multiplexed competitive experiments.** Cells of the indicated
502 strains (Fig. 7) were grown in 3 ml LBS at 25°C overnight with aeration. The cultures were then
503 diluted (1:80) into 3 ml fresh LBS and grown to mid-log phase (0.2 OD₆₀₀). Equivalent ODs of
504 cells from each strain [volume to mix calculated as Vol. (μl) = (1.25/OD₆₀₀) X 50] were mixed,
505 resulting in a multiplexed population with each strain present at a 1 to 1 ratio. A sample from this
506 input library was harvested by collecting cells from 700 μl by centrifugation and storing the cell
507 pellet at -80°C. The input library was then used to inoculate hatchling *Euprymna scolopes* squid

508 at $5-9 \times 10^3$ CFU/ml for 3 hr in FSIO (filter sterilized Instant Ocean) as previously described (20).
509 Squid samples (n = 24, per replicate) were harvested at 48 hr post-inoculation and surface
510 sterilized by storing at -80°C . Concurrently to squid colonization, the input library was competed
511 for growth *in vitro* for 15 generations by diluting the library 1:181 into LBS, growing at 25°C with
512 aeration back to the starting OD_{600} , repeating this process once more, and harvesting samples
513 as described above. Individual squid were homogenized in 700 μl of FSIO, 500 μl of each
514 homogenate was mixed in a 50 ml conical tube, diluted 1:20 in 70% IO (Instant Ocean), and 50
515 μl plated onto LBS plates in triplicate. After a 17 hr overnight incubation at 25°C the bacterial
516 colonies from each plate were scraped with a sterile cell scraper into 1 ml of 70% IO and
517 collected by centrifugation. Cell pellets were stored at -20°C prior to DNA extraction.

518 Genomic DNA from the cell pellets was extracted and purified using the Qiagen DNeasy
519 Blood and Tissue Kit (Qiagen; 69506) following the Gram-negative bacteria protocol, and was
520 quantified using a NanoDrop spectrophotometer (Thermo Scientific). The barcoded scars were
521 amplified together with dual-index Illumina sequencing primers (55). The reaction mixtures
522 contained 50 ng of gDNA, 200 nM of each oligo (Table S2), 1X Phusion Hot Start Flex Master
523 Mix (NEB; M0536L) and H_2O up to a total volume of 50 μl . PCR conditions were 98°C for 30
524 sec, 98°C for 10 sec, 60°C for 10 sec, 72°C for 10 sec (20 cycles), with a final extension step at
525 72°C for 5 min. PCR products were visualized using a 2% agarose gel to confirm the dual-
526 indexed amplicon of 231 bp and purified using a QIAquick PCR Purification Kit (Qiagen; 28106).
527 Purified PCR products were quantified using a Qubit 3 fluorometer (Life Technologies) and
528 pooled in equimolar amounts, and diluted to 4 nM. The pool was sequenced on an Illumina
529 MiSeq using the 2 X 250 bp v2 kit with a 10% PhiX control following the manufacturer's protocol
530 (Illumina, Inc., San Diego, CA) and using custom primers developed from (55). The sequencing
531 data was processed using the barseq python package to obtain strain counts per sample, and
532 mutants that were in the input library but still being validated were removed from the dataset.
533 The relative frequency (RF) for each strain in a sample was calculated, normalized to the RF in

534 the input library and the average RF in the sample, and the competitive index (CI) was then
535 calculated using the formula: $CI = \text{Log}_{10} [(RF_{\text{mutant}}/\text{Avg. RF}_{\text{WT}})_{\text{Sample}}/(RF_{\text{mutant}}/\text{Avg. RF}_{\text{WT}})_{\text{Input}}]$.

536 **FIGURE LEGENDS**

537 **FIG 1** Approach for quickly generating barcode-tagged gene deletions in *V. fischeri*. (A)
538 Schematic diagram (not to scale) of the process used to generate the barcoded deletions as
539 described in the main text. Multiple primers are designed for use in PCR to generate the desired
540 DNA molecules and screen/sequence for the correct deletion mutants as described in the
541 Materials and Methods. (B) To scale schematic of the resulting *bar* scar containing the start
542 codon, the left and right linker sequences (LL and RL), the FRT site that remains after removal
543 of the *erm* cassette, a spacer sequence, the unique barcode, and the last 7 codons of the
544 targeted open-reading frame (ORF). The barcode sequence is designed to lack in-frame stop
545 codons, which results in an in-frame ORF together with the start codon and the last 7 codons of
546 the targeted gene.

547
548 **FIG 2** Evaluating the genotype of a *copA* deletion strain. (A) Representative 1% agarose gel
549 showing the products generated by PCR when using the specified primer pairs and templates.
550 DNA ladder is the 1 kb Plus DNA Ladder (New England BioLabs). (B) Table showing several
551 unique barcode sequences within the $\Delta copA::erm-bar$ deletion scar of various deletion
552 candidates that were generated from a single round of mutagenesis. Diagram is not to scale.

553
554 **FIG 3** Assaying the phenotype of strains deleted for copper-resistance factors. (A and B) Bar
555 graphs showing the average OD₆₀₀ of the indicated $\Delta copA$ mutants after 20 hr of growth in the
556 presence of the indicated amounts of copper and/or bathocuproinedisulfonic acid (BCS). (A)
557 Error bars represent the standard deviation of the mean ($n = 3$). (B) Data are from two
558 independent replicates ($n = 2$). Statistical analysis was performed using a Two-way ANOVA
559 test. **** $P < 0.0001$.

560
561 **FIG 4** Assaying the phenotype of strains deleted for motility factors. Representative TBS agar
562 trays showing the migration of strains from the inoculation point after incubation at 28°C for 4 hr.

563 WT is MJM1100 (ES114), while WT-1 represents the *attTn7*-marked MJM1100 strain with
564 barcode 1 (either WT::*erm-bar1* or WT::*bar1*). Bar graph shows the quantified data from five
565 independent replicates with error bars showing the standard deviation of the mean ($n = 5$).
566 Statistical analysis was performed using a Two-way ANOVA test. **** $P < 0.0001$.

567

568 **FIG 5** The *gene::bar* deletion scar reduces polar effects on gene expression introduced by the
569 *erm* cassette. Graph showing the polarity ratio (expression of the downstream gene / expression
570 of the upstream gene; relative to the indicated gene deletion target) for the indicated gene
571 constructs within their respective predicted operons as measured by RT-qPCR. Statistical
572 analysis was performed using a Two-way ANOVA test. *** $P < 0.001$.

573

574 **FIG 6** Overview of BarSeq experiments and computational package. Methodology and software
575 for performing BarSeq experiments as described in the main text. An input population is used to
576 inoculate squid or media, and samples are taken at the times of interest for gDNA extraction
577 and processing to be sequenced by Illumina sequencing. The i5 and i7 segments are the index
578 sequences in the dual indexed DNA fragments, whereas the P5 and P7 sequences are
579 sequencing adapters for the MiSeq flow cell (details in Table S2). Sequencing reads are
580 analyzed by the barseq package to obtain counts for individual strains in a sample based on
581 their unique barcodes. Those counts are then used to calculate the relative frequencies of
582 individual strains at each timepoint and the competitive index (CI) as described in Materials and
583 Methods.

584

585 **FIG 7** BarSeq enables high-throughput competition experiments. (A and B) Graphs show the
586 mean competitive index (CI) on a Log_{10} -scale for each barcoded strain in the population using
587 the WT strains as controls as described in Brooks et al. (79) after (A) 15 generations *in vitro* in
588 LBS and (B) 48 hr post-squid inoculation (hpi). WT is MJM1100 (ES114). WT-1 represents the
589 *attTn7*-marked MJM1100 strain with barcode 1 (WT::*bar1*), and similarly for WT-2 and WT-3 for

590 barcodes 2 and 3, respectively. LOD = limit of detection for the experiment (3.39×10^{-5}). Each
591 symbol represents one biological replicate. Statistical analysis was performed using a One-way
592 ANOVA test comparing each strain to WT-1. *** $P < 0.001$, **** $P < 0.0001$.

593

594

595 **Table 1. Bacterial strains.** ^aThymidine auxotroph, growth conditions in Materials and Methods.

596 N/A = Not applicable.

Strain	Alias	Genotype/Description	Source
<i>V. fischeri</i>			
MJM1100	ES114 (WT)	ES114	78, 80
MJM1538	ES114/pLostfoX	MJM1100/pLostfoX	32
MJM1902	$\Delta copA::Tnerm$	MJM1100 $\Delta copA::Tnerm$	32
MJM3400	$\Delta copA::pEVS79-$ $\Delta copA$	MJM1100 $\Delta copA::pEVS79-\Delta copA$	This work
MJM3401	$\Delta copA$	MJM1100 $\Delta copA$	This work
MJM3529	$\Delta copA::erm-bar$	MJM1100 $\Delta copA::erm-bar$	This work
MJM3534	$\Delta cusA::erm-bar$	MJM1100 $\Delta cusA::erm-bar$	This work
MJM3543	$\Delta copA::bar$	MJM1100 $\Delta copA::bar$	This work
MJM3565	$\Delta cusA::bar$	MJM1100 $\Delta cusA::bar$	This work
MJM3620	WT:: <i>erm-bar1</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>erm-bar1</i>	This work
MJM3621	WT:: <i>erm-bar2</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>erm-bar2</i>	This work
MJM3622	WT:: <i>erm-bar3</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>erm-bar3</i>	This work
MJM3629	WT:: <i>bar1</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>bar1</i>	This work
MJM3630	WT:: <i>bar2</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>bar2</i>	This work
MJM3631	WT:: <i>bar3</i>	MJM1100 IG(<i>yeiR-glmS</i>):: <i>bar3</i>	This work
MJM3785	$\Delta flrA::erm-bar$	MJM1100 $\Delta flrA::erm-bar$	This work
MJM3785	$\Delta flaA::erm-bar$	MJM1100 $\Delta flaA::erm-bar$	This work
MJM3786	$\Delta rpoN::erm-bar$	MJM1100 $\Delta rpoN::erm-bar$	This work
MJM3788	$\Delta cheA::erm-bar$	MJM1100 $\Delta cheA::erm-bar$	This work
MJM3790	$\Delta copA::bar$ <i>attTn7::copA</i>	MJM1100 $\Delta copA::bar attTn7::copA$	This work

MJM3792	$\Delta flrA::bar$	MJM1100 $\Delta flrA::bar$	This work
MJM3795	$\Delta flaA::bar$	MJM1100 $\Delta flaA::bar$	This work
MJM3796	$\Delta rpoN::bar$	MJM1100 $\Delta rpoN::bar$	This work
MJM3798	$\Delta cheA::bar$	MJM1100 $\Delta cheA::bar$	This work
<i>E. coli</i>			
MJM534	CC118 $\lambda pir/pEVS104$	$\Delta(ara-leu) araD \Delta lacX74 galE galK$ <i>phoA20 thi-1 rpsE rpoB argE(Am)</i> <i>recA1</i> , lysogenized with $\lambda pir/pEVS104$	33
MJM537	DH5 α λpir	F- $\Phi 80 lacZ \Delta M15 \Delta(lacZYA-$ <i>argF)U169 supE44 hsdR17 (r_K⁻, m_K⁺)</i> <i>endA1 recA1 gyrA96 thi-1 relA1</i> <i>uidA::pir⁺</i>	Laboratory stock
MJM570	DH5 α /pEVS79	F- $\Phi 80 lacZ \Delta M15 \Delta(lacZYA-$ <i>argF)U169 supE44 hsdR17 (r_K⁻, m_K⁺)</i> <i>endA1 recA1 gyrA96 thi-1</i> <i>relA1/pEVS79</i>	33
MJM637	S17-1 $\lambda pir/pUX-BF13$	<i>pro res- hsdR17 (r_K⁻ m_K⁺) recA-</i> with an integrated <i>RP4-2-Tc::Mu-Km::Tn7</i> $\lambda pir/pUX-BF13$	72, 73
MJM658	BW23474/pEVS107	$\Delta lac-169 robA1 creC510 hsdR514$ <i>uidA(\Delta Mlul)::pir116</i> <i>endA(BT33) recA1/pEVS107</i>	70
MJM3287	NEB5 α /pHB1	F ⁻ $\Phi 80 lacZ \Delta M15 \Delta(lacZYA-$ <i>argF)U169 glnV44 hsdR17 (r_K⁻, m_K⁺)</i> <i>endA1 recA1 gyrA96 thi-1 relA1</i> <i>fhuA2 phoA/pHB1</i>	63

MJM3288	DH5α <i>λpir</i> /pHB2	MJM537/pHB2	This work
MJM3383	NEB5α/pHB3	F ⁻ Φ80/ <i>lacZ</i> ΔM15 Δ(<i>lacZYA-argF</i>)U169 <i>glnV44 hsdR17</i> (<i>r_K⁻, m_K⁺</i>) <i>endA1 recA1 gyrA96 thi-1 relA1 fhuA2 phoA</i> /pHB3	This work
MJM3478	KV8052: π3813 ^a /pKV496	<i>lacI^a thi-1 supE44 endA1 recA1 hsdR17 gyrA462 zei-298::Tn10 ΔthyA::(erm-pir-116)</i> /pKV496	38, 71

597

598

599 **Table 2. Plasmids.**

Plasmid	Relevant properties	Source
pEVS79	Vector backbone for deletion construct via allelic-exchange, Cam ^R	33
pEVS104	Conjugation helper plasmid, Kan ^R	33
pEVS107	mini-Tn7 mobilizable vector, Erm ^R (transposon), Kan ^R	70
pKV496	pEVS79 containing the FLP recombinase, Kan ^R	38
pLostfoX	<i>tfoX</i> overexpression vector, Cam ^R	44
pUC19	Cloning vector, Carb ^R	Laboratory stock
pUX-BF13	Tn7 transposase helper plasmid (<i>tns</i> genes), Carb ^R	72
pHB1	pUC19 containing the LL-FRT- <i>erm</i> -FRT-spacer sequence in the HindIII/BamHI site	63
pHB2	pEVS107 containing <i>copA</i> (including 191 bp upstream and 321 bp downstream of the <i>copA</i> ORF) at the <i>Ascl</i> site	This work
pHB3	pEVS79 containing 1.6-kb upstream/1.6-kb downstream of <i>copA</i>	This work

600

601 **Supplementary Tables S1-S3:**

- 602 • **Table S1. Expanded bacterial strains.** ^aThymidine auxotroph, growth conditions in
603 Materials and Methods. N/A = Not applicable.
604 • **Table S2. DNA oligonucleotides.**

605 ● **Table S3. Strain counts from barseq output.** Strain counts for individual strains within
606 each sample was obtained using the barseq package. The samples for competitive
607 squid colonization (48 hpi) were processed in triplicate as technical replicates. The
608 counts in ‘_other’ represent sequence reads that contain the appropriate sequences
609 flanking the barcode region but the barcode sequence does not match any present in the
610 reference barcode library (as described in Materials and Methods).

611 **ACKNOWLEDGMENTS**

612 We thank Karen L. Visick for plasmid pKV496, and members of the Mandel lab for
613 comments on the mutagenesis protocol and manuscript.

614

615 **FUNDING SOURCES**

616 This work was funded by NIGMS grant R35 GM119627 to M.J.M., an American Society for
617 Microbiology Undergraduate Research Fellowship to E.B., and the McNair Scholars Program
618 (E.B.). G.S. acknowledges support from the USDA National Institute of Food and Agriculture
619 (NIFA), Agricultural and Food Research Initiative (AFRI) Foundation grant no. 2020-67015-
620 31576, and USDA NIFA HATCH grant no. WIS02007.

621

622

623 **REFERENCES**

- 624 1. Taschuk R, Griebel PJ. 2012. Commensal microbiome effects on mucosal immune system
625 development in the ruminant gastrointestinal tract. *Anim Health Res Rev* 13:129–141.
- 626 2. Montgomery MK, McFall-Ngai M. 1994. Bacterial symbionts induce host organ
627 morphogenesis during early postembryonic development of the squid *Euprymna scolopes*.
628 *Development* 120:1719–1729.
- 629 3. Shropshire JD, Bordenstein SR. 2016. Speciation by symbiosis: the microbiome and
630 behavior. *MBio* 7:e01785.
- 631 4. Honda K, Littman DR. 2012. The microbiome in infectious disease and inflammation. *Annu*
632 *Rev Immunol* 30:759–795.
- 633 5. Byrd AL, Belkaid Y, Segre JA. 2018. The human skin microbiome. *Nat Rev Microbiol*
634 16:143–155.
- 635 6. O’Dwyer DN, Dickson RP, Moore BB. 2016. The lung microbiome, immunity, and the
636 pathogenesis of chronic lung disease. *J Immunol* 196:4839–4847.
- 637 7. Aragón IM, Herrera-Imbroda B, Queipo-Ortuño MI, Castillo E, Del Moral JS-G, Gómez-
638 Millán J, Yucel G, Lara MF. 2018. The urinary tract microbiome in health and disease. *Eur*
639 *Urol Focus* 4:128–138.
- 640 8. Kinross JM, Darzi AW, Nicholson JK. 2011. Gut microbiome-host interactions in health and
641 disease. *Genome Med* 3:14.
- 642 9. Parfrey LW, Knight R. 2012. Spatial and temporal variability of the human microbiota. *Clin*
643 *Microbiol Infect* 18 Suppl 4:8–11.
- 644 10. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial

- 645 community variation in human body habitats across space and time. *Science* 326:1694–
646 1697.
- 647 11. Kostic AD, Howitt MR, Garrett WS. 2013. Exploring host-microbiota interactions in animal
648 models and humans. *Genes Dev* 27:701–718.
- 649 12. Ruby EG. 1999. The *Euprymna scolopes-Vibrio fischeri* symbiosis: a biomedical model for
650 the study of bacterial colonization of animal tissue. *J Mol Microbiol Biotechnol* 1:13–21.
- 651 13. McFall-Ngai M. 2008. The squid-vibrio association- a naturally occurring, experimental
652 model of animal/bacterial partnerships, p. 102–112. *In* Gary B. Huffnagle, ANCN (ed.), *GI*
653 *Microbiota and Regulation of the Immune System*. Landes Bioscience and Springer
654 Science + Business Media.
- 655 14. McFall-Ngai MJ. 2014. The importance of microbes in animal development: lessons from
656 the squid-vibrio symbiosis. *Annu Rev Microbiol* 68:177–194.
- 657 15. Alice H Tischler, Kelsey M Hodge-Hanson, and Karen L Visick. 2019. *Vibrio fischeri*–squid
658 symbiosis. *Elsevier Oceanogr Ser* 267:1–9.
- 659 16. McFall-Ngai M. 2014. Divining the essence of symbiosis: insights from the squid-vibrio
660 model. *PLoS Biol* 12:e1001783.
- 661 17. Nyholm SV, McFall-Ngai M. 2004. The winnowing: establishing the squid–vibrio symbiosis.
662 *Nat Rev Microbiol* 2:632–642.
- 663 18. Jones BW, Nishiguchi MK. 2004. Counterillumination in the Hawaiian bobtail squid,
664 *Euprymna scolopes* Berry (Mollusca: Cephalopoda). *Mar Biol* 144:1151–1155.
- 665 19. Lee KH, Ruby EG. 1994. Effect of the squid host on the abundance and distribution of
666 symbiotic *Vibrio fischeri* in nature. *Appl Environ Microbiol* 60:1565–1571.

- 667 20. Naughton LM, Mandel MJ. 2012. Colonization of *Euprymna scolopes* squid by *Vibrio*
668 *fischeri*. J Vis Exp e3758.
- 669 21. Mandel MJ, Schaefer AL, Brennan CA, Heath-Heckman EAC, DeLoney-Marino CR,
670 McFall-Ngai MJ, Ruby EG. 2012. Squid-derived chitin oligosaccharides are a chemotactic
671 signal during colonization by *Vibrio fischeri*. Appl Environ Microbiol 78:4620–4626.
- 672 22. Essock-Burns T, Bongrand C, Goldman WE, Ruby EG, McFall-Ngai MJ. 2020. Interactions
673 of symbiotic partners drive the development of a complex biogeography in the squid-vibrio
674 symbiosis. MBio 11.
- 675 23. Sycuro LK, Ruby EG, McFall-Ngai M. 2006. Confocal microscopy of the light organ crypts
676 in juvenile *Euprymna scolopes* reveals their morphological complexity and dynamic function
677 in symbiosis. J Morphol 267:555–568.
- 678 24. Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ. 2002. Roles of *Vibrio*
679 *fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont
680 colonization of the *Euprymna scolopes* light organ. Appl Environ Microbiol 68:5113–5122.
- 681 25. Nyholm SV, Stabb EV, Ruby EG, McFall-Ngai MJ. 2000. Establishment of an animal–
682 bacterial association: recruiting symbiotic vibrios from the environment. Proc Natl Acad Sci
683 U S A 97:10231–10235.
- 684 26. Altura MA, Heath-Heckman EAC, Gillette A, Kremer N, Krachler A-M, Brennan C, Ruby
685 EG, Orth K, McFall-Ngai MJ. 2013. The first engagement of partners in the *Euprymna*
686 *scolopes*-*Vibrio fischeri* symbiosis is a two-step process initiated by a few environmental
687 symbiont cells. Environ Microbiol 15:2937–2950.
- 688 27. Visick KL. 2009. An intricate network of regulators controls biofilm formation and
689 colonization by *Vibrio fischeri*. Mol Microbiol 74:782–789.

- 690 28. Brennan CA, Mandel MJ, Gyllborg MC, Thomasgard KA, Ruby EG. 2013. Genetic
691 determinants of swimming motility in the squid light-organ symbiont *Vibrio fischeri*.
692 *Microbiologyopen* 2:576–594.
- 693 29. Stabb E.V., Schaefer A., Bose J.L., Ruby E.G. 2008. Quorum signaling and symbiosis in
694 the marine luminous bacterium *Vibrio fischeri*, p. 233–250. *In* Winans, Bassler (eds.),
695 *Chemical Communication among Bacteria*. American Society of Microbiology.
- 696 30. Visick KL, Foster J, Doino J, McFall-Ngai M, Ruby EG. 2000. *Vibrio fischeri lux* genes play
697 an important role in colonization and development of the host light organ. *J Bacteriol*
698 182:4578–4586.
- 699 31. Lupp C, Urbanowski M, Greenberg EP, Ruby EG. 2003. The *Vibrio fischeri* quorum-sensing
700 systems *ain* and *lux* sequentially induce luminescence gene expression and are important
701 for persistence in the squid host. *Mol Microbiol* 50:319–331.
- 702 32. Brooks JF 2nd, Gyllborg MC, Cronin DC, Quillin SJ, Mallama CA, Foxall R, Whistler C,
703 Goodman AL, Mandel MJ. 2014. Global discovery of colonization determinants in the squid
704 symbiont *Vibrio fischeri*. *Proc Natl Acad Sci U S A* 111:17284–17289.
- 705 33. Stabb EV, Ruby EG. 2002. RP4-based plasmids for conjugation between *Escherichia coli*
706 and members of the Vibrionaceae. *Methods Enzymol* 358:413–426.
- 707 34. Lyell NL, Dunn AK, Bose JL, Stabb EV. 2010. Bright mutants of *Vibrio fischeri* ES114
708 reveal conditions and regulators that control bioluminescence and expression of the *lux*
709 operon. *J Bacteriol* 192:5103–5114.
- 710 35. Baym M, Shaket L, Anzai IA, Adesina O, Barstow B. 2016. Rapid construction of a whole-
711 genome transposon insertion collection for *Shewanella oneidensis* by Knockout Sudoku.
712 *Nat Commun* 7:13270.

- 713 36. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI.
714 2009. Identifying genetic determinants needed to establish a human gut symbiont in its
715 habitat. *Cell Host Microbe* 6:279–289.
- 716 37. Studer SV, Mandel MJ, Ruby EG. 2008. AinS quorum sensing regulates the *Vibrio fischeri*
717 acetate switch. *J Bacteriol* 190:5915–5923.
- 718 38. Visick KL, Hodge-Hanson KM, Tischler AH, Bennett AK, Mastrodomenico V. 2018. Tools
719 for rapid genetic engineering of *Vibrio fischeri*. *Appl Environ Microbiol* 84:e00850–18.
- 720 39. Jahn LJ, Porse A, Munck C, Simon D, Volkova S, Sommer MOA. 2018. Chromosomal
721 barcoding as a tool for multiplexed phenotypic characterization of laboratory evolved
722 lineages. *Sci Rep* 8:6961.
- 723 40. Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, Chee M, Roth FP, Giaever G,
724 Nislow C. 2009. Quantitative phenotyping via deep barcode sequencing. *Genome Res*
725 19:1836–1842.
- 726 41. Robinson DG, Chen W, Storey JD, Gresham D. 2014. Design and analysis of Bar-seq
727 experiments. *G3* 4:11–18.
- 728 42. Blundell JR, Levy SF. 2014. Beyond genome sequencing: lineage tracking with barcodes to
729 study the dynamics of evolution, infection, and cancer. *Genomics* 104:417–430.
- 730 43. Abel S, Abel zur Wiesch P, Chang H-H, Davis BM, Lipsitch M, Waldor MK. 2015. Sequence
731 tag-based analysis of microbial population dynamics. *Nat Methods* 12:223–226.
- 732 44. Pollack-Berti A, Wollenberg MS, Ruby EG. 2010. Natural transformation of *Vibrio fischeri*
733 requires *tfoX* and *tfoY*. *Environ Microbiol* 12:2302–2311.
- 734 45. De Souza Silva O, Blokesch M. 2010. Genetic manipulation of *Vibrio cholerae* by

- 735 combining natural transformation with FLP recombination. *Plasmid* 64:186–195.
- 736 46. Dalia AB, Lazinski DW, Camilli A. 2013. Characterization of undermethylated sites in *Vibrio*
737 *cholerae*. *J Bacteriol* 195:2389–2399.
- 738 47. Baba T, Mori H. 2008. The construction of systematic in-frame, single-gene knockout
739 mutant collection in *Escherichia coli* K-12, p. 171–181. In Osterman, AL, Gerdes, SY (eds.),
740 *Microbial Gene Essentiality: Protocols and Bioinformatics*. Humana Press, Totowa, NJ.
- 741 48. Hernández-Montes G, Argüello JM, Valderrama B. 2012. Evolution and diversity of
742 periplasmic proteins involved in copper homeostasis in gamma proteobacteria. *BMC*
743 *Microbiol* 12:249.
- 744 49. Rensing C, Fan B, Sharma R, Mitra B, Rosen BP. 2000. CopA: An *Escherichia coli* Cu(I)-
745 translocating P-type ATPase. *Proc Natl Acad Sci U S A* 97:652–656.
- 746 50. Millikan DS, Ruby EG. 2003. FlrA, a σ^{54} -dependent transcriptional activator in *Vibrio*
747 *fischeri*, is required for motility and symbiotic light-organ colonization. *J Bacteriol* 185:3547–
748 3557.
- 749 51. Wolfe AJ, Millikan DS, Campbell JM, Visick KL. 2004. *Vibrio fischeri* σ^{54} controls motility,
750 biofilm formation, luminescence, and colonization. *Appl Environ Microbiol* 70:2520–2524.
- 751 52. Raynal A, Karray F, Tuphile K, Darbon-Rongère E, Pernodet J-L. 2006. Excisable
752 cassettes: new tools for functional analysis of *Streptomyces* genomes. *Appl Environ*
753 *Microbiol* 72:4839–4844.
- 754 53. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M,
755 Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene
756 knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008.

- 757 54. Oppenheim DS, Yanofsky C. 1980. Translational coupling during expression of the
758 tryptophan operon of *Escherichia coli*. *Genetics* 95:785–795.
- 759 55. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a
760 dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data
761 on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120.
- 762 56. Millikan DS, Ruby EG. 2004. *Vibrio fischeri* flagellin A is essential for normal motility and for
763 symbiotic competence during initial squid light organ colonization. *J Bacteriol* 186:4315–
764 4325.
- 765 57. Koonin EV. 2009. Evolution of genome architecture. *Int J Biochem Cell Biol* 41:298–306.
- 766 58. Bongrand C, Ruby EG. 2019. The impact of *Vibrio fischeri* strain variation on host
767 colonization. *Curr Opin Microbiol* 50:15–19.
- 768 59. Mandel MJ. 2010. Models and approaches to dissect host-symbiont specificity. *Trends*
769 *Microbiol* 18:504–511.
- 770 60. Wollenberg MS, Ruby EG. 2009. Population structure of *Vibrio fischeri* within the light
771 organs of *Euprymna scolopes* squid from two Oahu (Hawaii) populations. *Appl Environ*
772 *Microbiol* 75:193–202.
- 773 61. Bongrand C, Ruby EG. 2018. Achieving a multi-strain symbiosis: strain behavior and
774 infection dynamics. *ISME J* 13:698–706.
- 775 62. Wollenberg MS, Ruby EG. 2012. Phylogeny and fitness of *Vibrio fischeri* from the light
776 organs of *Euprymna scolopes* in two Oahu, Hawaii populations. *ISME J* 6:352–362.
- 777 63. Rotman ER, Bultman KM, Brooks JF 2nd, Gyllborg MC, Burgos HL, Wollenberg MS,
778 Mandel MJ. 2019. Natural strain variation reveals diverse biofilm regulation in squid-

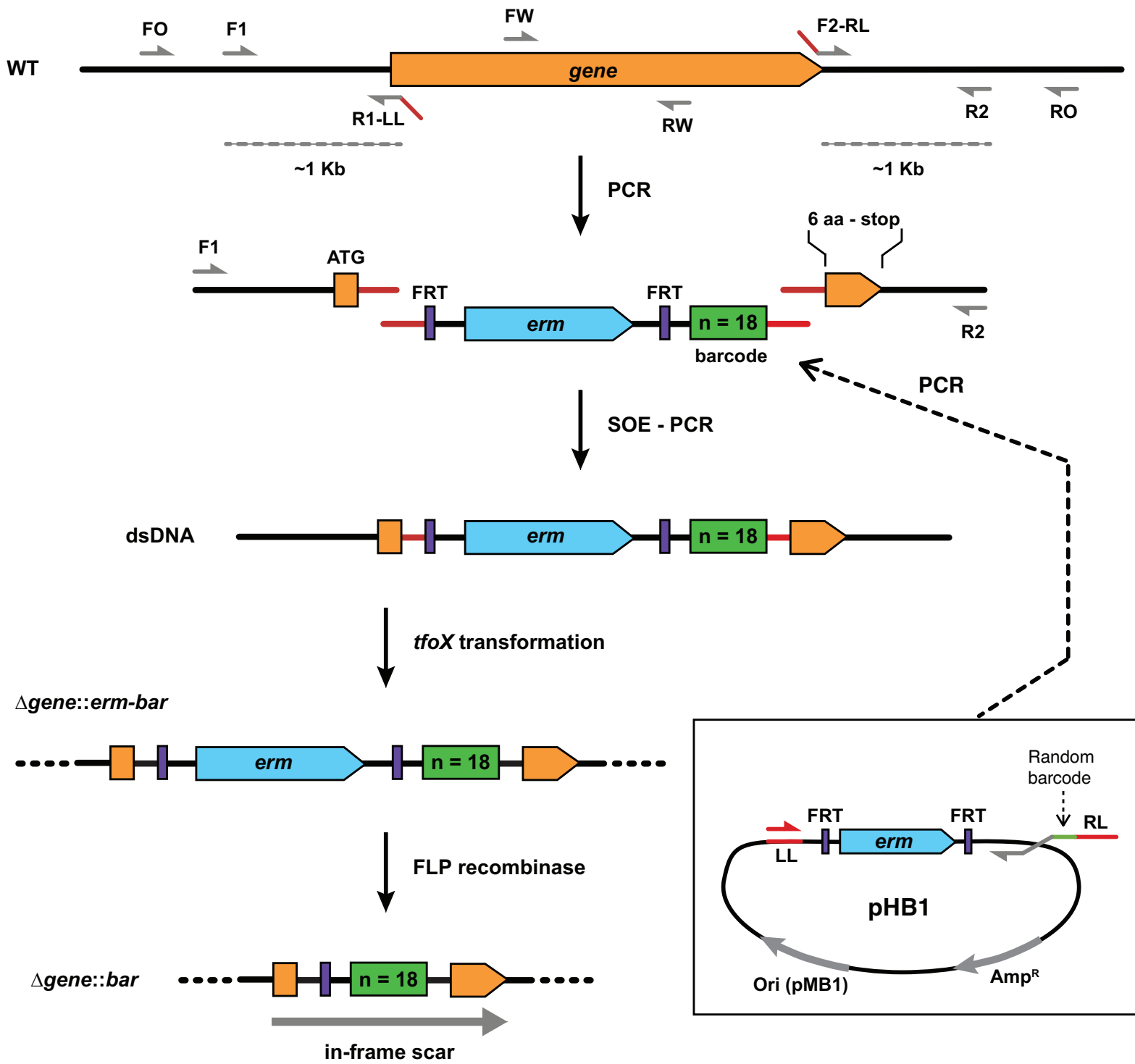
- 779 colonizing *Vibrio fischeri*. J Bacteriol 201.
- 780 64. Bongrand C, Moriano-Gutierrez S, Arevalo P, McFall-Ngai M, Visick KL, Polz M, Ruby EG.
781 2020. Using colonization assays and comparative genomics to discover symbiosis
782 behaviors and factors in *Vibrio fischeri*. MBio 11.
- 783 65. Pankey MS, Foxall RL, Ster IM, Perry LA, Schuster BM, Donner RA, Coyle M, Cooper VS,
784 Whistler CA. 2017. Host-selected mutations converging on a global regulator drive an
785 adaptive leap towards symbiosis in bacteria. Elife 6:e24414.
- 786 66. Yip ES, Grublesky BT, Husa EA, Visick KL. 2005. A novel, conserved cluster of genes
787 promotes symbiotic colonization and σ^{54} -dependent biofilm formation by *Vibrio fischeri*. Mol
788 Microbiol 57:1485–1498.
- 789 67. Hunt TP, Magasanik B. 1985. Transcription of *glnA* by purified *Escherichia coli*
790 components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc Natl
791 Acad Sci U S A 82:8453–8457.
- 792 68. Soutourina OA, Bertin PN. 2003. Regulation cascade of flagellar expression in Gram-
793 negative bacteria. FEMS Microbiol Rev 27:505–523.
- 794 69. Prouty MG, Correa NE, Klose KE. 2001. The novel σ^{54} - and σ^{28} -dependent flagellar gene
795 transcription hierarchy of *Vibrio cholerae*. Mol Microbiol 39:1595–1609.
- 796 70. McCann J, Stabb EV, Millikan DS, Ruby EG. 2003. Population dynamics of *Vibrio fischeri*
797 during infection of *Euprymna scolopes*. Appl Environ Microbiol 69:5928–5934.
- 798 71. Le Roux F, Binesse J, Saulnier D, Mazel D. 2007. Construction of a *Vibrio splendidus*
799 mutant lacking the metalloprotease gene *vsm* by use of a novel counterselectable suicide
800 vector. Appl Environ Microbiol 73:777–784.

- 801 72. Bao Y, Lies DP, Fu H, Roberts GP. 1991. An improved Tn7-based system for the single-
802 copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene*
803 109:167–168.
- 804 73. Simon R, Prierer U, Pühler A. 1983. A broad host range mobilization system for *in vivo*
805 genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnology*
806 1:784–791.
- 807 74. Khodursky AB, Bernstein JA, Peter BJ, Rhodius V, Wendisch VF, Zimmer DP. 2003.
808 *Escherichia coli* spotted double-strand DNA microarrays: RNA extraction, labeling,
809 hybridization, quality control, and data management. *Methods Mol Biol* 224:61–78.
- 810 75. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. 2016. ppGpp
811 binding to a site at the RNAP-DksA interface accounts for its dramatic effects on
812 transcription initiation during the stringent response. *Mol Cell* 62:811–823.
- 813 76. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
814 quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25:402–408.
- 815 77. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM,
816 Krummenacker M, Midford PE, Ong Q, Ong WK, Paley SM, Subhraveti P. 2019. The
817 BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform* 20:1085–
818 1093.
- 819 78. Mandel MJ, Stabb EV, Ruby EG. 2008. Comparative genomics-based investigation of
820 resequencing targets in *Vibrio fischeri*: focus on point miscalls and artefactual expansions.
821 *BMC Genomics* 9:138.
- 822 79. Brooks JF 2nd, Mandel MJ. 2016. The histidine kinase BinK is a negative regulator of
823 biofilm formation and squid colonization. *J Bacteriol* 198:2596–2607.

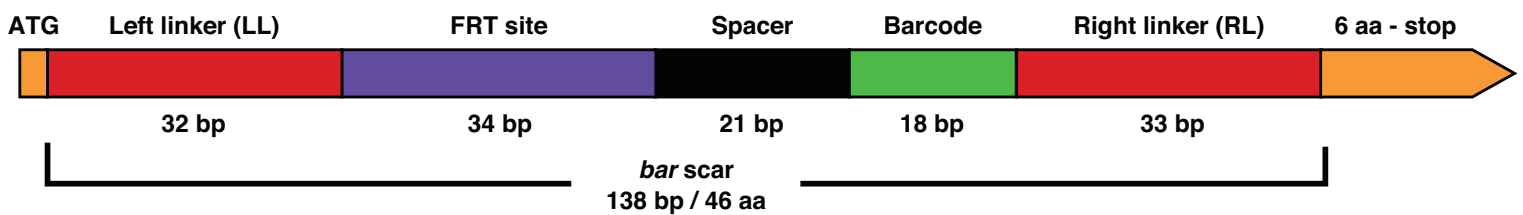
824 80. Boettcher KJ, Ruby EG. 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the
825 sepiolid squid *Euprymna scolopes*. J Bacteriol 172:3701–3706.

826

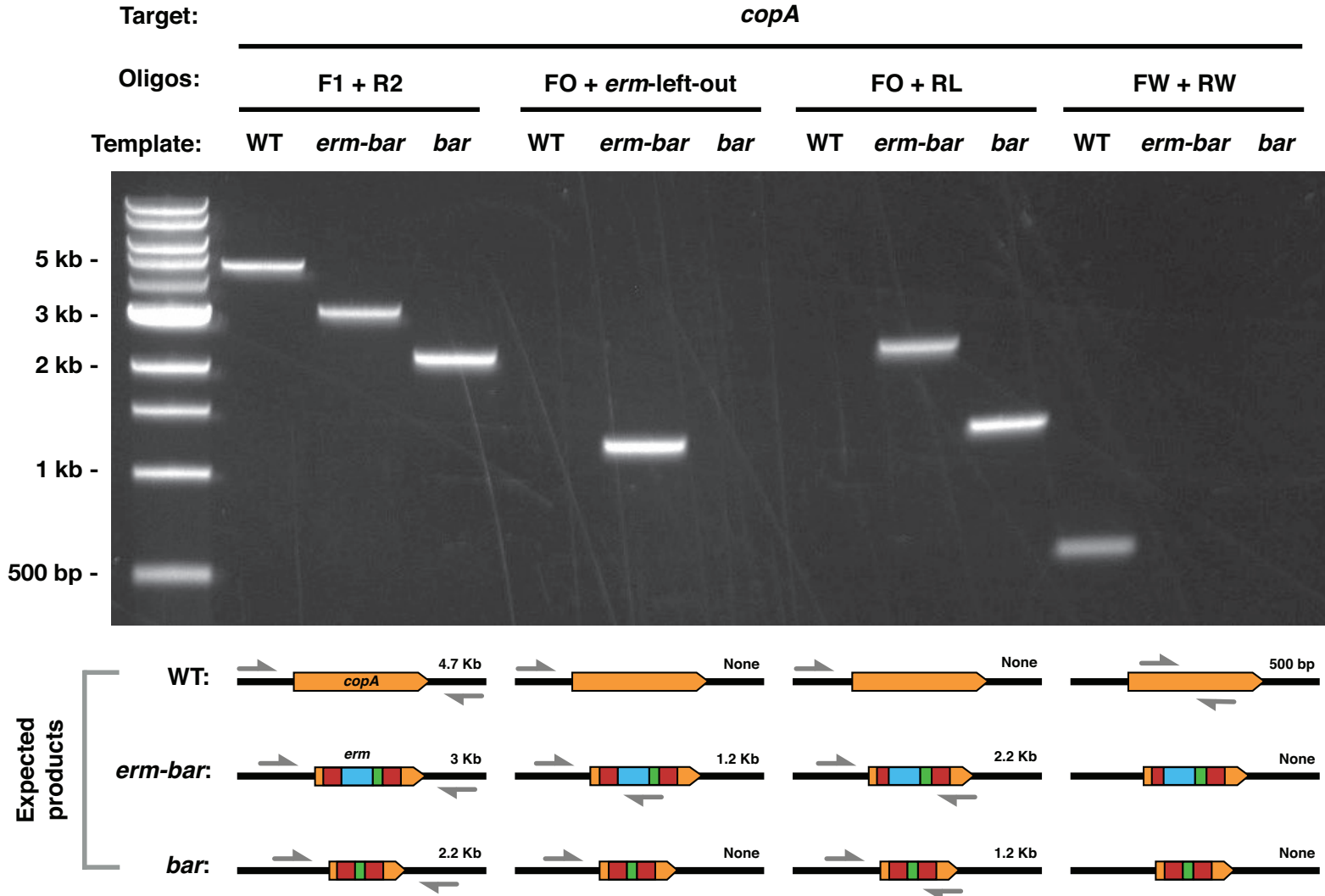
A



B

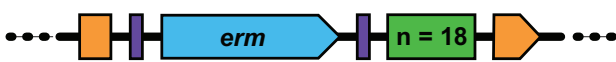


A

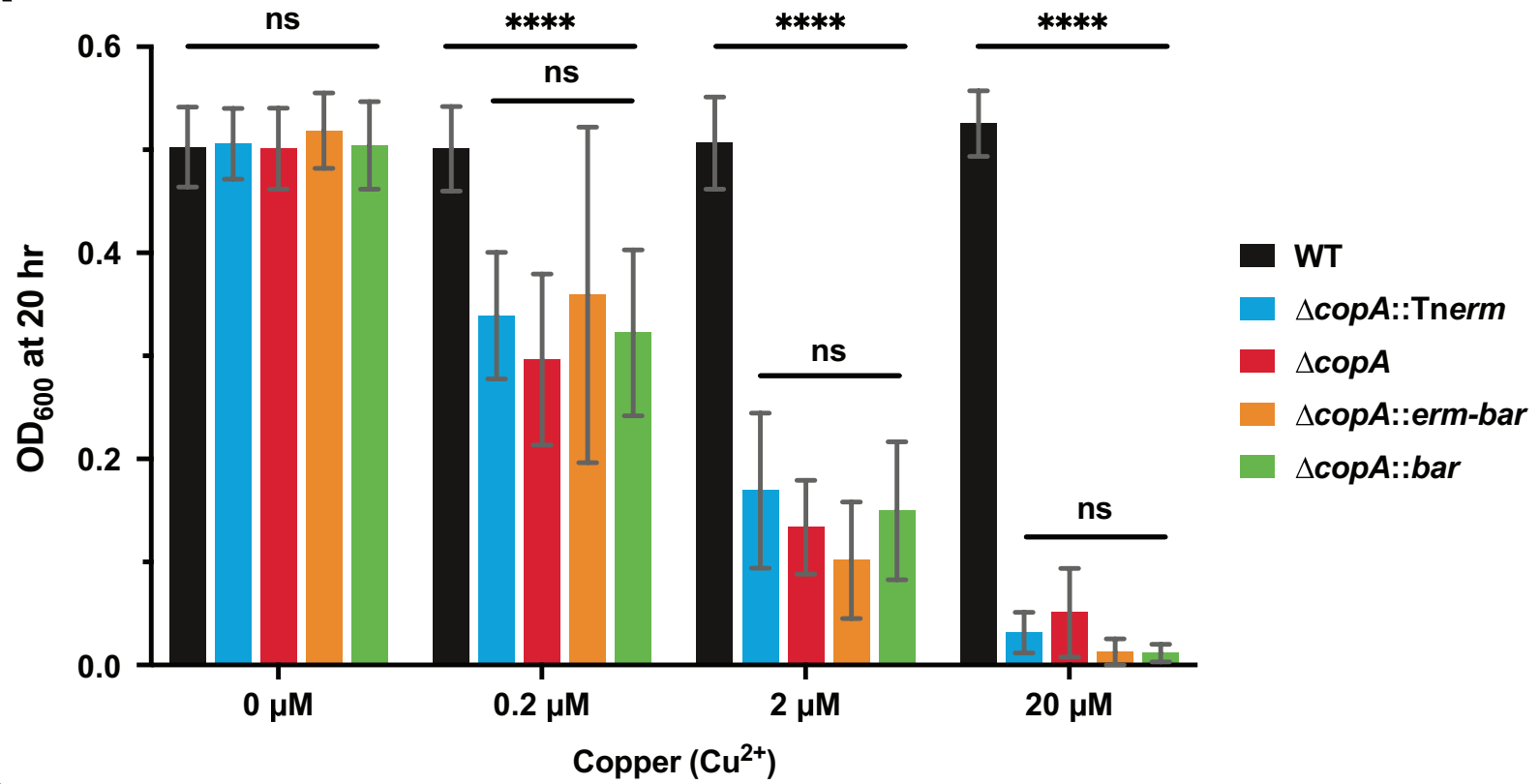
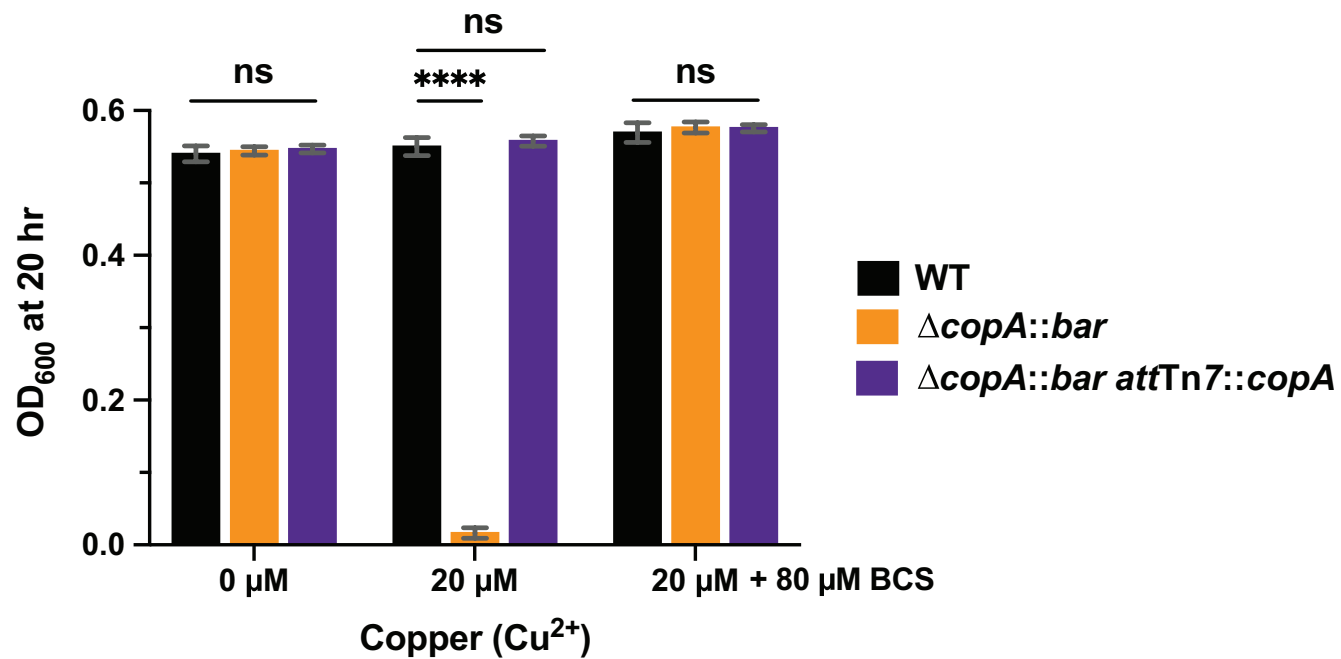


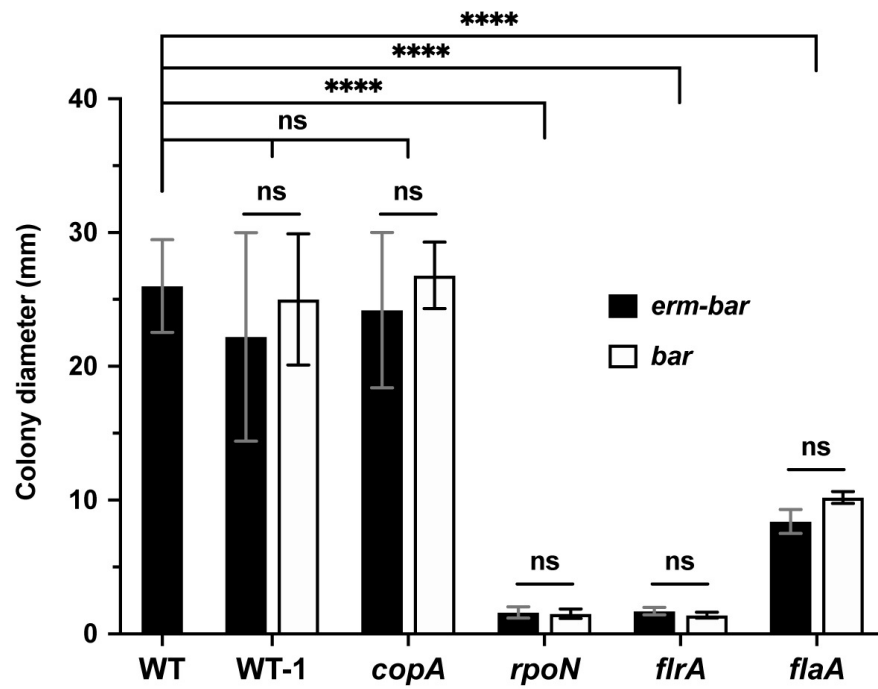
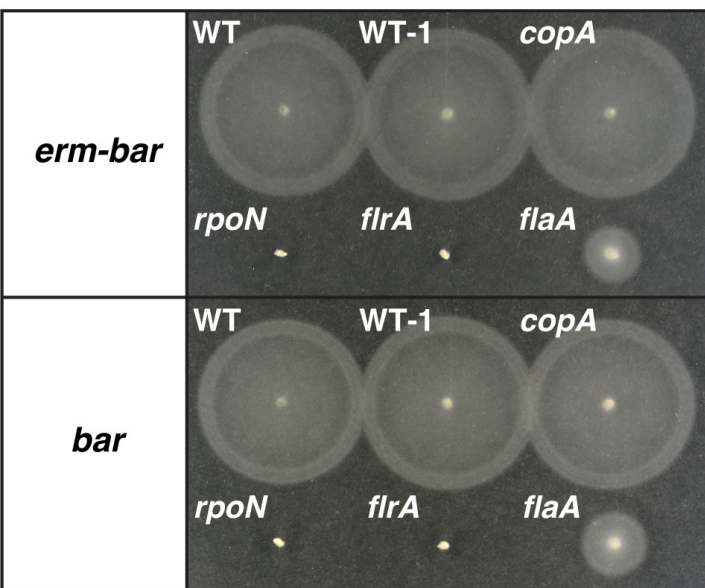
B

$\Delta copA::erm$ -bar deletion scar



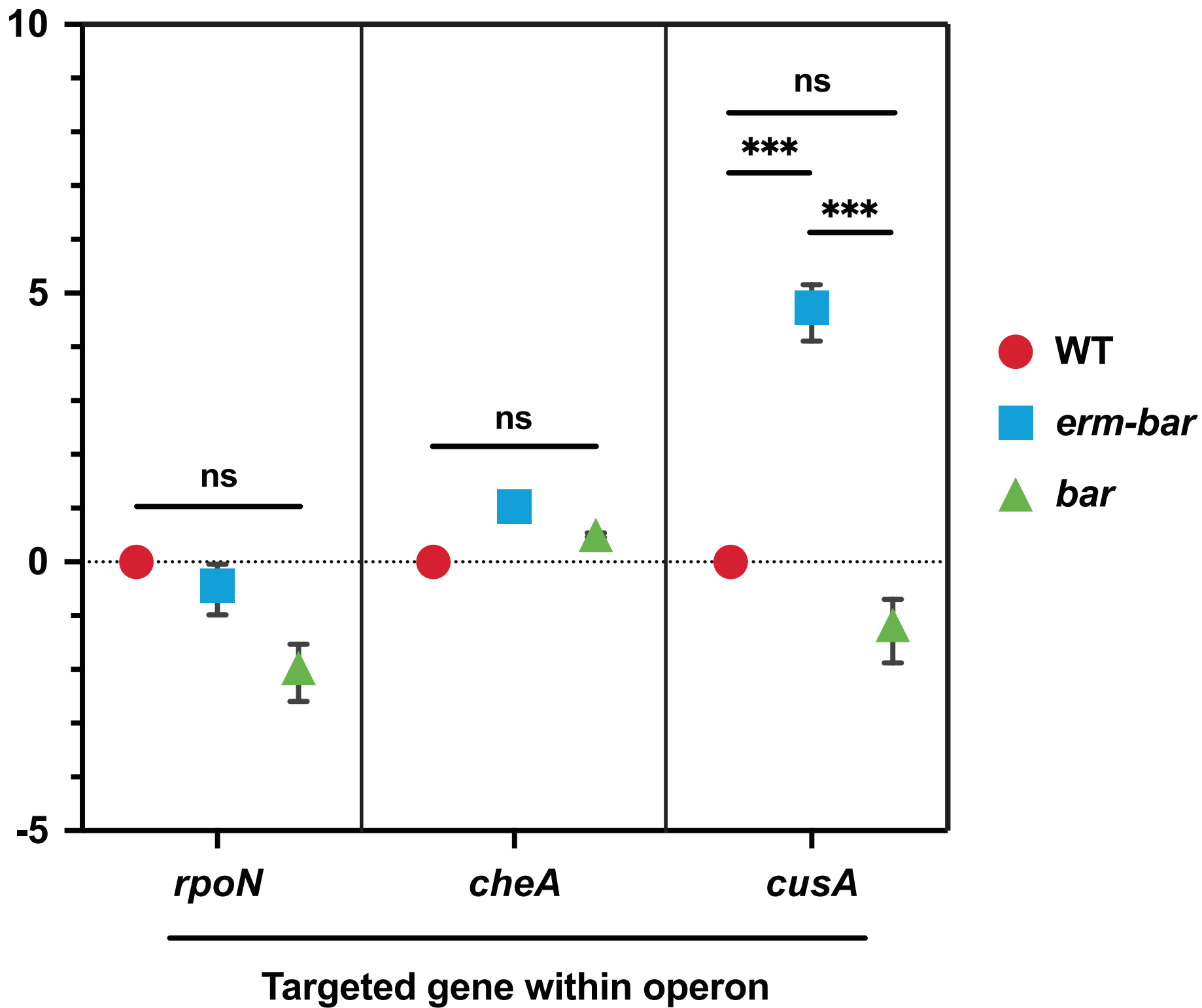
Barcode	VNNVNNVNNVNNVNNVNN
Isolate A	GGCCCCCGCCCGTCCCC
Isolate B	ATGAAGACTGTTGCCGTA
Isolate C	CACGACGCCCTCCGCGGA
Isolate D	ACTATTACGCAAATAAT

A**B**



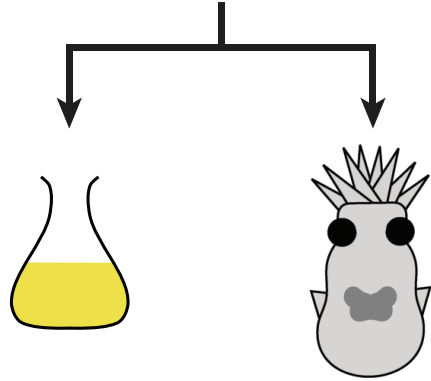
Polarity ratio

(downstream expression/
upstream expression; Log2)





Input population (synthetic microbiome)

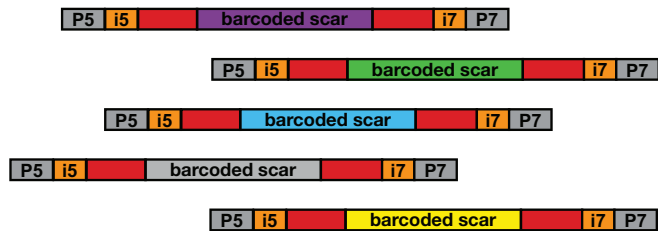


Inoculate culture medium and/or squid hatchlings

Grow in vitro for the desired number of generations and/or allow colonization to proceed for the desired time

Take samples at the desired time points, extract gDNA, amplicon/index PCR

Pool of dual indexed DNA fragments carrying the barcoded deletion scars



Illumina multiplex sequencing (e.g., MiSeq system)

Output fastq file:

```
read 1      TGACGACTGATCG...
read 2      AGCTACGCGATAC...
read 3      TAACTGCTAGTAC...
...         ...
```

Demultiplex based on unique dual indexes (i5, i7)

```
Sample 1   Sample 2   Sample 3
read 1     read 1     read 1
read 2     read 2     read 2
...        ...        ...
```

Barseq package: identifies barcodes and assigns strain identity

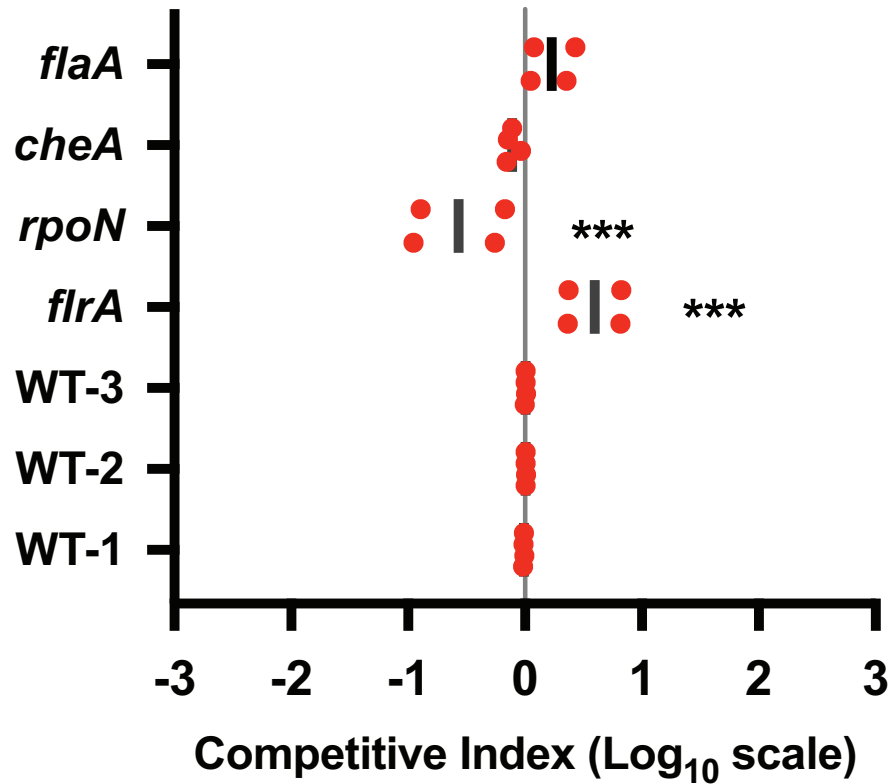
Strain	Barcode	Sample 1	Sample 2	...
WT::bar1	ATGAAGACTGTTGCCGTA	23694	15692	...
cheA::bar	AATGCCCATATTGAGGTG	3	4	...
flaA::bar	ATAATACGTCATACAGCT	10043	8098	...
...

$$CI = \log_{10} \left(\frac{\left(\frac{RF_{Mut}}{RF_{WT}} \right)_{Sample}}{\left(\frac{RF_{Mut}}{RF_{WT}} \right)_{Input}} \right)$$

Normalize counts to input and calculate competitive index (CI)

A

In vitro, 15 generations LBS

**B**

In vivo, 48 hpi squid

