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

Hector L. Burgos[1], Emanuel F. Burgos[1], Andrew J. Steinberger[2], Garret Suen[2], and Mark J. Mandel[1,*]

[1]Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI USA

[2]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI USA

Short title: *Vibrio fischeri* barcoded deletions and BarSeq

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

* Correspondence to:
Mark J. Mandel
University of Wisconsin-Madison
Department of Medical Microbiology and Immunology
1550 Linden Drive
Madison, WI 53706
Phone: (608) 261-1170
Fax: (608) 262-8418
Email: mmandel@wisc.edu

ABSTRACT

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

IMPORTANCE

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

INTRODUCTION

Beneficial symbioses are ubiquitous in the environment and have substantial impacts on the health and development of animal hosts. In animals, symbionts can affect host organ morphogenesis, immune system development, reproduction, susceptibility to disease, and even behavior (1–4). In humans, the gut, skin, lungs, and urogenital tract all have specific microbiomes for which their dysbiosis has been associated with disease (5–8). The gut microbiome, in particular, contains a dense community of diverse bacteria that influences the immune system and host behavior, and is increasingly being linked to healthy and disease states, including cardiovascular disease, irritable bowel syndrome, obesity, colon cancer, and susceptibility to *Clostridium difficile* infection (8–11). It is clear that molecular communication between animal hosts and their microbial partners leads to selection and retention of the cognate microbiome: while many microbes are obtained from the environment, the composition of mature microbiomes is often largely stable and resilient within members of a host species (12, 13). Metagenomics, transcriptomics, and metabolomics have all been applied to study the microbiome in various animal model systems (14). While these approaches can characterize aspects of microbial communities, the complexity of animal-associated microbiomes and the inability to culture and genetically-manipulate many symbionts make it difficult to study the precise molecular mechanisms that establish specific relationships.

The binary symbiosis between genetically-tractable *Vibrio fischeri* and the Hawaiian bobtail squid *Euprymna scolopes* serves as a model system to study the molecular interactions underlying microbiome assembly (14–19). The squid hatch aposymbiotically (without symbiont)

and are colonized by *V. fischeri* in a multi-step process that leads to the specific recruitment of the symbiont from a marine environment in which the bacteria are < 0.1% of the bacterioplankton (17, 20). The symbionts are housed in the dedicated light organ (LO) within the squid's mantle cavity, where they generate light that the host uses for counterillumination to hide its shadow while it hunts at night (21). The host provides the symbionts with a protected niche, nutrients, and oxygen (18). Once the symbiosis is irreversibly established in juvenile squid, a daily cycle proceeds where 90-95% of the bacteria are expelled from the LO at dawn. The remaining symbionts grow during the day until they fill the niche, and at night the dense population of symbionts provide light (22). Because the aposymbiotic hatchlings can be cultured in the lab and infected with genetically-tractable *V. fischeri*, colonization experiments can be performed to study the molecular factors that play a role during this process (15, 19, 23). In addition, the translucent nature of the squid hatchlings allows for visualization of the colonization process by microscopy (24–26).

Microbe-host signaling mechanisms and developmental transitions ensure specificity during colonization (17, 20). Upon detection of bacterial-derived peptidoglycan, the ciliated appendages on the surface of the LO secrete mucus that traps bacteria circulating within the mantle cavity (16, 27, 28). In the mucus field, *V. fischeri* bacteria bind to cilia and form aggregates by expressing symbiosis polysaccharide (*syp*) genes, a locus of 18 structural and regulatory genes that induce biofilm formation (29, 30). Approximately 3-4 hr post-inoculation, bacterial aggregates migrate through the host mucus toward the pores that lead into the LO ducts (28). While the initial migration is independent of flagellar motility (31), at the pores squid-

101 produced chitin oligosaccharides serve as a chemoattractant to direct the symbiotic bacteria into
102 the host crypts (24). Motility and chemotaxis are required for colonization, and mutants of genes
103 required for these processes—such as *cheA*, *flrA*, and *rpoN*—are unable to successfully
104 colonize the squid LO (24, 31). The squid also secretes nitric oxide (NO) into the mucus and
105 throughout the ducts, which likely plays a role in colonization specificity, since the symbiont
106 contains specialized mechanisms to resist and respond to NO production (32). Once within the
107 LO *V. fischeri* generates light through expression of the *lux* operon in a quorum sensing-
108 dependent process (33). Symbionts that fail to produce luminescence, such as strains with
109 mutations in the autoinducer synthases *ainS* or *luxI*, or deletions of the *lux* operon, are
110 preferentially eliminated during daily venting of the LO so that they do not persist in the
111 symbiosis (34, 35).

112 Synchronous colonization studies with the squid host have demonstrated that establishment
113 of the symbiosis occurs in stages. Three broad stages have been defined (36): initiation
114 (approximately 1-4 h post inoculation), in which bacteria aggregate in the host and use biofilm
115 formation and chemotaxis/motility to enter the squid (37, 38); accommodation (4-36 h), where
116 bacteria grow to fill the LO (39); and persistence (36+ h), where the population is maintained in
117 the host through the daily cycles of expulsion and regrowth (34, 38, 40). Aspects of this process
118 are well understood, including the biofilm formation and luminescence required for initiation and
119 persistence, respectively. However, symbiont and host factors that regulate accommodation and
120 key transitions between stages are not well understood.

While traditional genetic engineering techniques have been instrumental in characterizing molecular factors that play a role during squid colonization, forward genetic screens based on random mutagenesis and deep sequencing have accelerated the discovery of a large number of genes that modulate symbiotic phenotypes (31, 37, 41–44). To identify novel colonization factors in *V. fischeri* our lab previously used a global transposon insertion sequencing approach (INSeq) to identify bacterial mutants that were depleted after 48 hours in the squid host (44). This approach successfully identified previously-known colonization factors, such as *rscS*, *rpoN, ompU*, various motility factors, and the *syp* biofilm locus, and in addition revealed 344 novel putative colonization factors. Twenty candidates were tested in competitive colonization assays of WT vs mutant strains and the results showed that nine factors had colonization defects, suggesting that about 45% of INSeq hits are true colonization factors. Some of the validated factors encompass roles in protein quality control (DnaJ and DegS) and copper detoxification (CopA and CusC), inner membrane proteins predicted to play a role in secretion of autotransporters (TamB/YtfN), and other poorly characterized factors (YdhC, YafD, and YhcB). This global approach was crucial in identifying putative colonization factors. However, further study is required to address which genes are true colonization factors, when they act during colonization, and how their products modulate the interaction with the host.

Approximately 32% of putative colonization factors identified did not fall into a curated Clusters of Orthologous Groups (COGs) category, suggesting that the ability to interrogate the function of these colonization factors will reveal novel biology. Traditional genetic engineering techniques in *V. fischeri* are either random (transposon mutagenesis) or labor-intensive

(plasmid-based allelic exchange) (43–45). We therefore considered approaches by which we could isolate mutants and examine phenotypes in a multiplexed fashion. One possible approach was to retrieve transposon insertions of interest from an arrayed library (46–48). A second approach we considered was to adapt a newly-developed method for transformation-mediated mutagenesis using linear DNA (49) with an in-frame barcoding strategy to facilitate precise mutations. This latter option was attractive in that we hoped that it would limit the effects of polar mutations and provide a set of defined deletions that can be characterized by amplicon PCR. Barcode sequencing (BarSeq), in which each strain is labeled with a unique DNA sequence and sequenced using high-throughput next-generation sequencing, has been used successfully to track population dynamics in multiple systems, including in yeast genomic libraries, during *Vibrio cholerae* infection, and to track and phenotype laboratory-evolved *Escherichia coli* (50–54). Here, we describe an approach to generate barcode-tagged gene deletions in *V. fischeri* and perform high-throughput colonization experiments using BarSeq. We also describe the barseq python computational package used to analyze the results.

RESULTS

Generation of barcoded gene deletions. To generate barcoded deletions of specific *V. fischeri* genes, we designed an approach that takes advantage of splicing-by-overlap extension PCR and *tfoX*-based natural transformation (Fig. 1) (49, 55–57). The first step uses PCR to amplify DNA fragments upstream and downstream of the gene targeted for deletion, fused to the left and right linker sequences, respectively (Fig. 1A). A separate PCR is performed with pHB1 as a template to generate the central fragment of DNA containing the linker sequences, a selectable marker (*erm*, conferring erythromycin resistance) surrounded by FRT sites, and the semi-random barcode sequence. The barcode is provided by the reverse primer, which contains a region of semi-randomized-sequence. The three resulting DNA fragments—upstream, central, and downstream—are then fused into one fragment via their overlapping linker sequences by SOE-PCR (57) and transformed into *V. fischeri* upon *tfoX* induction (55). Finally, the *erm* cassette is removed via FLP recombinase (56).

The resulting 138-bp deletion scar (*barcode* scar, or "*bar*" scar, Fig 1B) lies between the deleted gene's first codon and last seven codons (i.e., the final six amino acid-encoding codons plus the stop codon). The scar is designed to be in-frame to prevent polar effects on gene expression when targeting genes within operons. The terminal codons were retained in case they contain a ribosomal binding site for downstream gene(s) (58). In addition to the barcode, the additional sequence in the scar includes left and right linker sequences that are shared among all of the mutants, which allows us to identify and quantify the abundance of each

barcoded-strain using amplicon sequencing, while minimizing amplification bias by using common primers that amplify the same size product.

To test this new approach, we investigated the *copA* gene. Among Gammaproteobacteria, CopA is the main exporter of cytoplasmic copper and is the most widely conserved copper detoxification factor (59, 60). Although our laboratory previously demonstrated that *copA* is a squid colonization factor, its role in copper resistance has not been examined (44). We therefore targeted *copA* for deletion using our mutagenesis approach as a proof-of-concept and subsequently tested its role in copper resistance in *V. fischeri*. To ensure that the deletion process worked as intended, we used four sets of diagnostic PCR primers that would report on correct *erm* insertion, subsequent removal of the *erm* cassette by FLP recombinase, and absence of the targeted gene from the bacterial chromosome. PCR with various pairs of oligonucleotides that target the *copA* gene and its deletion constructs produced amplicons of the expected size in each strain (Fig. 2A). These results show that the *erm* cassette was successfully inserted into *copA* generating Δ*copA*::*erm-bar*, and subsequently removed by FLP recombination to generate the in-frame deletion scar in Δ*copA*::*bar*. Furthermore, sequencing of the deletion scar for several *ΔcopA*::*erm-bar* candidates showed that after a single round of mutagenesis, multiple uniquely barcoded deletion strains were generated (Fig. 2B). These results demonstrate that our deletion method is successful in generating uniquely barcoded mutant strains of *V. fischeri*.

The presence of a barcode within a gene deletion does not alter mutant phenotypes. To test that the barcoded scar does not affect the mutant phenotypes, we measured the copper

197 sensitivity of strains deleted for *copA* using various methods. In addition to the mutants
198 generated using our deletion approach (Δ*copA*::*erm-bar* and Δ*copA*::*bar*) we constructed a
199 deletion of *copA* using plasmid-based allelic exchange (Δ*copA*) (45) and obtained a *copA*
200 transposon mutant isolated from our previous study (Δ*copA*::Tn*erm*) (44). We then tested the
201 growth of these various *copA* mutants in the presence of varying amounts of copper. Our results
202 show that, regardless of the mutagenesis method, the growth of *copA* mutants is similarly
203 impeded in the presence of copper, with the severity of the growth defect increasing in
204 proportion to the concentration of copper: at 0.2 μM Cu^{2+} the *copA* mutants were able to grow
205 slightly, whereas at 20 μM Cu^{2+} these strains were unable to grow (Fig. 3A). This is in contrast
206 to the WT strain that achieved the same growth yield regardless of the concentration of copper
207 present. The Δ*copA*::*erm-bar* and Δ*copA*::*bar* mutants showed the same degree of copper
208 sensitivity (Fig. 3A).
209 To corroborate that the observed growth defects were due specifically to excess copper, we
210 measured the growth of Δ*copA*::*bar* in the presence of copper, with and without the copper
211 chelator bathocuproinedisulfonic acid (BCS). As expected, Δ*copA*::*bar* was unable to grow in
212 the presence of 20 μM Cu^{2+}, whereas the WT is unaffected (Fig. 3B). However, growth of
213 Δ*copA*::*bar* in the presence of copper was rescued by addition of 80 μM of BCS (Fig. 3B),
214 suggesting that free copper is indeed responsible for the observed lack of growth in the mutant.
215 To verify that the absence of CopA was responsible for susceptibility to copper toxicity, we
216 complemented *copA* at the chromosomal *att*Tn*7* site in the Δ*copA*::*bar* strain and observed that
217 growth was rescued in the presence of copper (Fig. 3B). Based on these results, we conclude

that CopA is required for copper detoxification in *V. fischeri*, consistent with its function in other Gammaproteobacteria (59).

To further test our deletion method, we generated mutants in multiple genes required for *V. fischeri* motility—*rpoN*, *flrA*, and *flaA*—and tested the resulting strains' motility phenotypes on soft agar plates (31, 61, 62). While motility of WT *V. fischeri* resulted in a migration disc with a diameter of 21 mm from the inoculation point on soft agar plates, deleting *flaA* resulted in a drastic reduction in migration (5 mm), while deleting either *flrA* or *rpoN* resulted in no motility (1 mm) (Fig. 4). Again, we observed that both the *erm-bar* and *bar* versions of the deletions displayed equivalent phenotypes, showing that the strains behave as null alleles regardless of whether the scar contains the *erm* cassette (Fig. 4).

Removing the erythromycin-resistance cassette minimizes polar effects of the barcoded deletions. While the presence or absence of the *erm* cassette does not prevent deletion strains from manifesting the corresponding phenotypes (Fig. 3A and 4), we were concerned about polar effects on downstream gene expression upon insertion of the 1,049 bp heterologous *erm* cassette (63–65). To test the effect of the *erm* cassette on downstream gene expression we used reverse transcriptase quantitative PCR (RT-qPCR) to measure expression of genes immediately upstream and downstream of a targeted gene deletion for three different operons. In each case, we measured the ratio in expression levels of the downstream vs upstream genes in the mutant, normalized to the ratio in WT (defining this normalized value as the "polarity ratio"). For both *rpoN* and *cheA*, deletion scars of either *erm-bar* or *bar* resulted in negligible changes in the polarity ratio (Fig. 5); the slight decrease in the polarity ratio for the

Δ*rpoN*::*bar* construct compared to WT is not statistically significant. In contrast, the polarity ratio of Δ*cusA*::*erm-bar* was 26-fold higher than WT, whereas removal of the *erm* cassette to form the in-frame deletion scar resulted in the polarity ratio reverting to basal levels (no statistically significant differences are observed in the polarity ratio between Δ*cusA*::*bar* and WT; Fig. 5). We conclude that, in at least some cases, *gene*::*bar* deletion scars can alleviate collateral effects on flanking genes that are caused by inserting an antibiotic-resistance cassette.

Development of a computational pipeline to analyze *V. fischeri* BarSeq data. With the ability to quickly generate precise barcoded deletions, we next sought to compete the Δ*gene*::*bar* deletions en masse during host colonization. We therefore developed a BarSeq sample preparation protocol, and an accompanying computational package to analyze the data (Fig. 6). To accomplish this, we mixed barcoded strains to generate an input library (i.e., a synthetic microbiome). This library was then used to inoculate media and/or squid hatchlings, which were then sampled at the desired time points. Samples were then processed to extract gDNA, and PCR was performed with dual-index Illumina sequencing primers to obtain dsDNA fragments containing the barcoded deletion scars. The resulting library was then sequenced on an Illumina MiSeq and demultiplexed based on the unique dual indexes (66). The resulting sequence data was then analyzed using the barseq package, which identifies and counts the barcodes present in the samples, assigns strain identity, normalizes strain counts, and calculates relative frequency and the competitive index (CI) for each strain within the samples. The BarSeq protocol provides a streamlined and effective way to measure population dynamics throughout squid colonization.

BarSeq enables sensitive multiplex competition experiments. To test our BarSeq protocol in tracking individual strains within a population, we performed an *in vitro* competition and a competitive colonization experiment using an input library of 7 different barcoded strains mixed in an equivalent ratio. In addition to several mutant strains, we included three WT strains that had the *bar* scar inserted at the Tn7 site that could be similarly tracked using amplicon sequencing, but without affecting the phenotypes of the strains. After 15 generations of growth *in vitro*, the proportion of most strains relative to WT remained stable except for *flrA* and *rpoN*, which were 4-fold higher and lower, respectively, when compared to WT (Fig. 7A). In contrast, following 48 h of squid colonization—which corresponds to approximately 15 bacterial generations (44)—we observed reduced levels of the *flaA* flagellin mutant and severely reduced levels of the *flrA*, *rpoN*, and *cheA* strains, all of which were near the limit of detection (Fig. 7B). This result is consistent with their previously known role as necessary factors for squid colonization, although we did observe higher levels of *flaA* in the competitive colonization that are observed when a transposon insertion is competed against wild-type (31, 61, 62, 67). We note that there was relatively little variability among the tagged wild-type strains in the analysis, whereas the 4-5 log scale in which to identify colonization defects provided a substantially greater range over which to identify and refine colonization phenotypes *in vivo* (Fig. 7B). Taken together, these results show that our method for targeted barcoded deletions, multiplex squid colonization, and analysis by BarSeq allows for reproducible competition experiments *in vitro* and *in vivo* with high sensitivity.

DISCUSSION

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

Generation of targeted, barcoded gene deletions that minimize effects on neighboring genes. Our approach builds upon previous SOE-PCR/*tfoX* mutagenesis techniques to incorporate a unique barcode in each deletion strain, which enables high-throughput experiments via barcode sequencing (BarSeq). Since BarSeq relies on amplicon sequencing, library preparation is straightforward and allows for a large number of samples to be processed in parallel. The method we have employed to design the barcode and flanking sequences was planned to minimize disruption on flanking genes. Expression of bacterial genes is frequently organized by their genetic arrangement into operons where expression of operon members is driven by a common promoter (68). However, given that some regulatory regions overlap in neighboring genes, deletion of one gene can alter the expression level of another nearby cistron. These off-target effects on gene expression could obfuscate the analysis of experimental results. Similar to the approach used by Baba et al. (64), our deletion approach reduces off-target effects on gene expression by ensuring the formation of an in-frame open reading frame within the deletion scar and including several codons at the end of the deletion

target where the ribosome binding site of downstream genes is frequently located (Fig. 1B and Fig. 5). There are some cases where the *bar* scar affects gene expression of the downstream gene, but these effects are small (2-4 fold decrease) and are less than those introduced by the insertion of the *erm-bar* scar (Fig. 5).

BarSeq enables detailed studies of the molecular mechanisms that result in establishment of the *Vibrio*-squid symbiosis. Using an INSeq screen, our lab previously identified 344 putative novel squid colonization factors in *V. fischeri* (44). Our deletion approach, combined with BarSeq, will enable the high-throughput characterization of these factors during squid colonization by allowing multiplexing of colonization factor mutants and tracking of individual strains. By enabling a more precise study of colonization factors BarSeq has several potential applications.

BarSeq can be applied to gain a high temporal resolution picture of the colonization process. Traditional colonization experiments are usually performed either before 3 hr to probe initiation of colonization and aggregate formation, or at 24 and 48 hr post-infection after bacteria have reached and colonized the LO. However, various transitions occur during colonization within 3-24 hr, including entrance to the pores, travel through the ducts into the antechamber, travel through the constricted bottleneck before entry into the crypts, induction of luminescence, and cell growth (16, 17, 25). Confocal microscopy experiments have been used successfully to assay these early stages of colonization but are low-throughput (25, 26, 28). Using BarSeq, we are able to sample many timepoints during colonization and accurately define the role of colonization factors at different stages. One change that occurs within the 3-24 hr window is

constriction of the bottleneck within the LO in response to colonization by *V. fischeri* (25); this phenotype varies depending on the symbiont species, suggesting that differing molecular interactions can affect bottleneck constriction. BarSeq can be used to quickly identify the factors involved in bottleneck constriction by sampling multiple timepoints within the colonization stage of interest using a high number of multiplexed deletion mutants.

BarSeq can also be applied to the study of strain variation and evolution of colonization mechanisms in the *Vibrio*-squid symbiosis. *V. fischeri* strain variation is an important consideration when studying the mechanisms of colonization of the squid LO (69). Previous studies have shown that multiple strains can co-colonize the squid LO and that they do so at different rates (70–72). More recent studies have focused on deciphering the specific mechanisms that result in differing colonization behavior (73, 74). However, the large number of *V. fischeri* strains make the study of the evolution of colonization mechanisms labor-intensive and prohibitive. The barcode-tagged mutagenesis method presented here can be applied to generate uniquely-tagged WT or mutant strains of the various phylogenetically-distinct *V. fischeri* strains and assayed in multiplexed format during squid colonization using BarSeq. We have already successfully used our mutagenesis approach to make deletions in the ancestral strain SR5, showing that this method is applicable to *V. fischeri* strains that are evolutionarily distant to the frequently used ES114 strain (73).

BarSeq can also be used in directed evolution experiments to test the function of genetic variation of colonization factors. Directed evolution experiments have recently been applied to study colonization factors in *V. fischeri* (75). Barcode tagging also has applications in tracking

specific lineages in long-term evolution experiments and analyzing microbial population dynamics during infection (50, 53, 54). The ease of tracking large numbers of individual *V. fischeri* strains using BarSeq enables long-term evolution experiments that track lineages through multiple rounds of colonization, growth, expulsion, and recolonization. This will allow the study of large numbers of phylogenetically-distinct strains at a single time and help measure the evolutionary dynamics of colonization factors across *V. fischeri* species.

Phenotypes of *rpoN* and *flrA* mutants during competitive growth in media. Both *rpoN* and *flrA* deletion strains showed a statistically significant 4-fold decrease and increase, respectively, during competitive growth in media compared to WT (Fig. 7A). Due to the nature of the factors they encode, the observed growth defects are likely due to changes in energetic and nutritional requirements when *rpoN* and *flrA* are deleted. The *rpoN* gene encodes the alternative σ^{54} factor that is responsible for recognition and expression of various genes involved in squid colonization, including Syp biofilm formation, flagellar motility, and luminescence (62, 76, 77). Therefore, it is not surprising that deleting the gene encoding σ^{54} has pleiotropic effects on gene expression due to misregulation of the RpoN regulon and could reduce the ability of the mutant strain to effectively compete for growth *in vitro*, though further experiments are necessary to define the precise mechanism for the defect. FlrA is the σ^{54}-dependent transcription factor that activates expression of the flagellar biosynthesis cascade and is required for motility and squid colonization (61). The high energetic cost of expressing all genes related to flagellar biosynthesis (78), which in *V. cholerae* requires FlrA-dependent regulation of 52 genes (79) and in *V. fischeri* between 39 and 131 genes (31), is consistent with the observed increase in growth

of the *flrA* deletion strain compared to WT during competitive growth in media (Fig. 7A). Nonetheless, even though the changes observed in competitive growth of the *rpoN* and *flrA* mutants *in vitro* are in opposing directions (less vs more growth, respectively), both are severely defective in squid colonization (Fig. 7B). Future experiments using BarSeq to probe bacterial growth during colonization will elucidate the relationship between bacterial growth rates *in vitro* and *in vivo*, and efficient squid colonization.

Discrepancy in *flaA* colonization efficiency measured by BarSeq vs traditional competitive squid colonization experiments. In our BarSeq experiment, the known colonization factor *flaA* only shows a small (~2-fold) colonization defect after 48 hr post-squid infection (Fig. 7B). In contrast, Millikan and Ruby showed that a *flaA* deletion made by insertion of a Kan^R-cassette is severely defective during competitive colonization against WT *V. fischeri* (67). Using confocal microscopy, their work showed that LO colonization by *flaA* is delayed compared to WT by ~8 hr. However, because our competitive colonization experiment using BarSeq was done at 48 hpi, this delayed colonization is not enough to explain the observed discrepancy. Previous work has shown that the concentration of *V fischeri* in the inoculum can affect the number of different strains that can co-colonize the squid LO (80). This raises the possibility that the inoculum amount or the ratio of strains within the synthetic microbiome might affect the observed colonization defect. To address this, future experiments should examine how inoculum amount and the ratio of mutant strains to WT within a synthetic barcode-tagged population affects colonization efficiency for the different strains in the population.

Insights into the role of copper during colonization. CopA is a copper resistance factor widely-distributed throughout the Gammaproteobacteria (59), but its function had not been tested in *V. fischeri*. Here we show that CopA, which is required for robust colonization of the squid host, is required for *V. fischeri* to avoid intoxication at high copper concentrations (Fig. 3A and 3B). Copper is an essential metal but is toxic due to its ability to produce reactive oxygen species and outcompete other metals for binding to essential metalloproteins (81). Copper is used in innate immunity as a bactericidal compound at host-microbe interfaces (82); e.g., macrophages pump copper into phagosomes to kill ingested pathogens (83). Together with the identification of *copA*, *cusA*, and *cusC* mutants as defective during colonization (44), these results suggest that appropriately managing intracellular copper is paramount for *V. fischeri* to successfully colonize its squid host. It is intriguing to speculate that squid hemocytes, the invertebrate equivalent of macrophages that accumulate within the ciliated appendages on the surface of the LO during initiation of colonization, may use copper to select for the correct symbiont (84).

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

MATERIALS AND METHODS

Bacterial strains, growth conditions, plasmids, and primers. Bacterial strains used in this study are listed in Table 1, with Table S1 containing an Extended Table 1 showing the oligos used to generate the specified barcode-tagged gene deletions. Plasmids are listed in Table 2, and DNA oligonucleotides are listed in Table S2. DNA oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA), and Sanger DNA sequencing was performed through the University of Wisconsin–Madison Biotechnology Center. *Escherichia coli* strains were grown in Luria-Bertani (LB) medium [per liter, 25 g Difco LB broth (BD), in distilled water] at 37°C with aeration. Unless otherwise indicated, *V. fischeri* strains were grown in Luria-Bertani salt (LBS) medium [per liter, 25 g Difco LB broth (BD), 10 g NaCl, and 50 ml 1 M Tris buffer, pH 7.0, in distilled water] at 25°C with aeration. When necessary, growth media was solidified by adding 15 g Bacto agar (BD) per liter. For growth of *V. fischeri*, antibiotics (Gold Biotechnology) were added at the following concentrations: 5 µg/ml erythromycin, 5 µg/ml or 2.5 µg/ml chloramphenicol as indicated, and 100 µg/ml kanamycin. For *E. coli* the antibiotic concentrations used were 100 µg/ml carbenicillin, 25 µg/ml chloramphenicol, and 50 µg/ml kanamycin. The *E. coli* strain π3813 containing pKV496 is a thymidine auxotroph and was grown in LB with 50 µg/ml kanamycin supplemented with 0.3 mM thymidine (49, 85).

The unmarked deletion of *copA* in MJM1100 was made by allelic exchange as described previously (73). Briefly, 1.6 kb upstream (US) and 1.6 kb downstream (DS) sequences of *copA* were amplified by PCR using oligos HB44 and HB45, and HB46 and HB47, respectively, and were cloned using Gibson Assembly (NEBuilder HiFi DNA assembly cloning kit) into the

linearized vector pEVS79 (linearized using oligos HB52 and HB53) (Table S2). The Gibson mix was transformed into NEB5α chemically-competent cells and selected on chloramphenicol. The resulting pEVS79-Δ*copA* candidates were screened using PCR with oligos HB54 and HB55 and confirmed by sequencing, generating pHB3, which was conjugated into *V. fischeri* MJM1100 (ES114) via triparental mating with MJM534, which contains the helper plasmid pEVS104 (45). Single recombinants of pHB3 into the chromosome were screened and selected by growth on chloramphenicol (MJM3400), and double recombinants by loss of the antibiotic resistance cassette and *copA* (MJM3401). The resulting constructs were verified by PCR and sequencing (Table S2).

The *copA* gene was inserted into the *att*Tn*7* site in the chromosome using pEVS107 (80). The *copA* gene including 191 bp US and 321 bp DS sequences was amplified by PCR using oligos HB27 and HB34, the product was digested with AscI, and cloned into the AscI site of pEVS107. The resulting plasmid, pHB2 (pEVS107-*copA*), was transformed into and maintained in *E. coli* DH5α λ*pir* cells and verified by PCR (Table S2) and sequencing. pHB2 was conjugated into *ΔcopA* (MJM3401) via tetraparental mating with donor MJM3288 (DH5α λ*pir*/pHB2), helper strains MJM637 (S17-1 λ*pir*/pUX-BF13) (86, 87) and MJM534 (CC118 λ*pir*/pEVS104) (45), and the recipient MJM3543 (Δ*copA*::*bar*), resulting in MJM3790 (Δ*copA*::*bar att*Tn7::*copA*). Candidates were confirmed by PCR (Table S2) and sequencing.

Construction of barcode-tagged gene deletions. The deletion protocol demonstrated in Fig 1A is based on splicing by overlap extension PCR (SOE-PCR) and *tfoX* transformation (49, 55–57) to directly delete and tag targeted genes with a randomized sequence (barcode). Our

protocol was in development prior to publication of the previous method (49), so while it is conceptually similar, the sequences of the linkers and primers are distinct. First, several oligos were designed specific to the targeted genes to amplify 1 kb of US (F1 and R1-LL) and DS (F2-RL and R2) DNA tagged with the left linker (LL) and right linker (RL) sequences, respectively, to screen the deletion scar via PCR (FO and RO), and to assay for the absence of the targeted gene (FW and RW) (Fig. 1A, Table S1, and Table S2). FW and RW were designed to amplify a fragment of 500-1,000 bp, depending on the size of the gene. The F1 and R2 oligos were designed to anneal 1 Kb US and DS, respectively, of the targeted gene. The R1 oligo was designed to anneal starting at the start codon of the targeted ORF going upstream, then the reverse complement of the LL sequence (LL reverse complement: 5'-CTGGCGAAGCATATATAAGAAGCTCGTCTCGT-3') was attached to the 5'-end of the R1 oligo, resulting in R1-LL. The F2 oligo was designed to anneal at the last 7 codons (6 aa and stop codon) on the 3'-end of the targeted ORF going downstream, then the RL sequence (RL: 5'-GACTTGACCTGGATGTCTCTACCCACAAGATCG-3') was attached to the 5'-end of the F2 oligo, resulting in F2-RL. The FO and RO oligos (forward outside and reverse outside, respectively) were designed to anneal 500 bp away from the annealing sites of F1 and R2, respectively, and were used to probe the targeted genomic region for insertion of the desired deletion scar.

The middle dsDNA fragment containing the *erm* cassette flanked by FRT sites and the randomized barcode was obtained by PCR with Phusion Hot Start Flex 2X master mix (NEB; M0536L) and pHB1 as template, which contains the LL-FRT-*erm*-FRT-spacer sequences and

465 was built as described previously (73), and oligos HB42 and HB154. Oligo HB154 is a reverse
466 primer and contains the RL sequence, 18 bp of randomized sequence composed of 6 trimers of
467 'NNB' to prevent formation of stop codons (results in 'VNN' codons in the forward direction), and
468 the spacer sequence (Table S2). The resulting 1,049 bp product containing LL-FRT-*erm*-FRT-
469 spacer-random barcode-RL was purified by gel extraction using a QIAquick Gel Extraction Kit
470 (Qiagen; 28706). The flanking 1 kb US and DS fragments for each targeted gene were then
471 fused to this middle DNA fragment via the homology between the LL and RL sequences and
472 using SOE-PCR with the F1 and R2 oligos, resulting in the 3 kb mutagenic dsDNA. The reaction
473 mixture contained 10 ng of each the middle, US, and DS DNA fragments, 200 nM of the
474 corresponding F1 and R2 oligos (Table S2), 1X Phusion Hot Start Flex Master Mix (NEB;
475 M0536L) and H_2O up to a total volume of 25 µl. SOE-PCR conditions were 98°C for 30 sec,
476 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
477 for 5 min.

478 The 3 Kb mutagenic DNA fragments were purified using a QIAquick PCR Purification Kit
479 (Qiagen; 28106) and transformed into *V. fischeri* ES114 via natural transformation with pLostfoX
480 (MJM1538) (44, 55) where the flanking sequences guide the barcoded *erm* cassette to
481 substitute the targeted gene. Mutant candidates were selected on LBS-Erm5 and screened by
482 PCR with oligo pairs F1/R2, FO/HB8, and FW/RW (as shown in Fig. 2A). The insertion of the
483 *erm-bar* scar was confirmed by Sanger sequencing with primers HB8, HB9, HB42, and HB146,
484 and the unique barcode sequence was recorded for each strain. We checked that the barcode

485 sequences were unique by calculating the hamming distance between the new barcodes and
486 those recorded for other deletion strains in our strain collection.
487 The final *bar* scars were made by triparental mating of donor MJM3478 (π3813/pKV496)
488 (49) and helper strain MJM534 (CC118 λ*pir*/pEVS104) into recipient *V. fischeri* strains
489 containing the *erm-bar* scar and selection on LBS-Kan100. Plasmid pKV496 contains the FLP
490 recombinase that removes the *erm* cassette and fuses the two surrounding FRT sites into one,
491 resulting in the final *bar* scar as shown in Fig. 1B. The plasmid was eliminated by growing the
492 candidates on LBS without selection twice and selecting colonies that were Erm^S and Kan^S. The
493 *gene*::*bar* candidates were screened by PCR using oligo pairs F1/R2, FO/HB146 (RL), and
494 FW/RW, and the deletion scar verified by Sanger sequencing using oligos HB42 and HB146.
495 The barcode sequences were verified to match the barcode within the parental strains
496 containing the *gene*::*erm-bar* scar.
497 The barcoded WT *V. fischeri* strains (WT::*bar*) were constructed using the same procedure
498 as outlined above for the gene deletions but targeting a site next to the *att*Tn*7* site in the
499 intergenic region of *yeiR* and *glmS*. The 1 kb US and DS arms were amplified using PCR with
500 ES114 gDNA and oligo pairs HB239/HB240 and HB241/HB242. After SOE-PCR to form the
501 mutagenic DNA and *tfoX* transformation, the WT::*erm-bar* candidates were screened by PCR
502 with oligo pairs HB243/HB244 and HB243/HB8. Sanger sequencing was used to confirm
503 insertion of the *erm-bar* scar and record the unique barcode sequences. Triparental mating as
504 described above was performed to remove the *erm* cassette using pKV496. The *bar* scar was
505 confirmed by PCR with HB243/HB146 and Sanger sequencing.

506 **Growth assays in the presence of copper.** Colonies from freshly streaked LBS plates of
507 the indicated *V. fischeri* strains were inoculated into 3 ml LBS with the appropriate antibiotics
508 and grown for 8 hr at 25°C with shaking. Three microliters of the LBS cultures were subcultured
509 into 3 ml Tris minimal medium [per liter, 500 ml DSW (2X), 50 ml 1 M Tris base, pH 7.5, 1 ml
510 5.8% K_2HPO_4, 1 ml 10 mM $FeSO_4$, and 20 ml 10% *N*-acetylglucosamine (GlcNAc), in distilled
511 water; DSW (2X) = 100 mM $MgSO_4$, 20 mM $CaCl_2$, 600 mM NaCl, and 20 mM KCl] and
512 incubated at 25°C overnight for ≤ 16 hr. Overnight cultures were diluted to an OD_{600} of 0.5 in
513 200 µl, then 2 µl of 0.5 OD_{600} were transferred into 198 µl of fresh Tris minimal medium
514 containing the appropriate amounts of copper and/or BCS in a 96-well plate. The plate was then
515 incubated in a Synergy Neo2 Multi-Mode Microplate Reader (BioTek) at 25°C with OD_{600}
516 measurements every 15 min for 20 hr. Copper stock solutions (100 mM) were prepared from
517 copper (II) sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$; Sigma-Aldrich; 203165), and BCS stock
518 solutions (50 mM) from bathocuproinedisulfonic acid disodium salt (Sigma-Aldrich; B1125).
519 **Motility assays.** The indicated bacterial strains were streaked onto fresh LBS plates with
520 the appropriate antibiotics and grown overnight at 25°C. Single colonies were picked with a
521 sterile toothpick and deposited onto TBS Agar plates [per liter, 10 g Gibco Bacto Tryptone
522 (Thermo Fisher Scientific; 211705), 50 ml 1 M Tris buffer, pH 7.0, 20 g NaCl, 8.63 g $MgSO_4$,
523 and 3 g Agar, in distilled water] by stabbing the toothpick into the media at a single spot. Plates
524 were incubated at 28°C for 4 hr and the outer diameter of swimming cells was measured.
525 **Measuring polarity ratio via RT-qPCR.** The indicated bacterial strains were grown in 3 ml
526 LBS with the appropriate antibiotics and grown at 25°C overnight. On the day of the experiment,

527 15 µl of the overnight cultures were transferred into 3 ml of fresh LBS and growth was continued
528 at 25°C with aeration. Samples were harvested at an OD_{600} of 0.2-0.4 (mid-log phase) by
529 transferring 800 µl of culture into a 2 ml screw-cap tube containing 100 µl of a cold 95% EtOH–
530 5% Phenol solution that inactivates RNases (88). RNA extraction and RT-qPCR were performed
531 as described previously (89). Briefly, cells were lysed in Tris-EDTA (TE) buffer (10 mM Tris-Cl,
532 pH 8.0, 1 mM EDTA) containing lysozyme (Epicentre; R1804M) and 1% SDS. RNA was
533 extracted using the hot phenol method (88) and digested with DNase I (NEB; M0303S).
534 cDNA was synthesized from 0.5 µg of total RNA using the iScript Advanced cDNA synthesis
535 kit (Bio-Rad; 1725037) following the protocol 25°C for 5 min, 46°C for 20 min, and 95°C for 1
536 min. Quantitative PCR was performed using 1:10 dilutions of cDNA synthesis products with the
537 iTaq Universal SYBR green supermix (Bio-Rad; 1725121) on a CFX Connect real-time PCR
538 detection system (Bio-Rad). The qPCR protocol was 95°C for 30 sec, 95°C for 5 sec, 58°C for
539 30 sec (40 cycles), with a final melt curve analysis to ensure specificity in the reaction. The
540 mRNA levels of *rpoD*, *lptB*, *hpf*, *cheZ*, *cheB*, *cusB* and *cusF* were measured using the oligo
541 pairs listed in Table S2. Expression levels for each gene were normalized to *rpoD* and the
542 mutants were normalized to WT using the $2^{-\Delta\Delta CT}$ method (90). The polarity ratio of *rpoN*, *cheA*,
543 and *cusA* was calculated as "expression of the downstream gene / expression of the upstream
544 gene" using the respective flanking genes in each putative operon–*lptB-rpoN-hpf*, *cheZ-cheA-*
545 *cheB*, and *cusB-cusA-cusF* (91).
546 **Barseq bioinformatic tool.** To quantify barcodes within each sequenced sample, we
547 developed barseq (https://github.com/mjmlab/barseq), a python package that identifies putative

548 barcodes in the sequenced reads and matches them to a user provided barcode library. The
549 program iterates through each sample and uses regular expressions to search within the reads
550 for flanking sequences on the left (GCTCATGCACTTGATTCC; spacer sequence) and the right
551 (GACTTGACCTGGATGTCT; right linker sequence) of the barcode (Fig. 1B), while also allowing
552 for 18 random nucleotides that represent a candidate barcode. The putative barcode sequence
553 is then mapped against the reference barcode library and increases the count for the matched
554 strain. Barseq outputs a tab-delimited table with the barcode/strain counts for each of the
555 samples analyzed.

556 **Barcode Sequencing and multiplexed competitive experiments.** Cells of the indicated
557 strains (Fig. 7) were grown in 3 ml LBS at 25°C overnight with aeration. The cultures were then
558 diluted (1:80) into 3 ml fresh LBS and grown to mid-log phase (0.2 OD_{600}). Equivalent ODs of
559 cells from each strain [volume to mix calculated as Vol. (µl) = (1.25/OD_{600}) X 50] were mixed,
560 resulting in a multiplexed population with each strain present at a 1 to 1 ratio. A sample from this
561 input library was harvested by collecting cells from 700 µl by centrifugation and storing the cell
562 pellet at -80°C. The input library was then used to inoculate hatchling *Euprymna scolopes* squid
563 at 5-9 x 10^3 CFU/ml for 3 hr in FSIO (filter sterilized Instant Ocean) as previously described (23).
564 Squid samples (n = 24, per replicate) were harvested at 48 hpi and surface sterilized by storing
565 at -80°C. Concurrently to squid colonization, the input library was competed for growth *in vitro*
566 for 15 generations by diluting the library 1:181 into LBS, growing at 25°C with aeration back to
567 the starting OD_{600}, repeating this process once more, and harvesting samples as described
568 above. Individual squid were homogenized in 700 µl of FSIO, 500 µl of each homogenate was

mixed in a 50 ml conical tube, diluted 1:20 in 70% IO (Instant Ocean), and 50 µl plated onto LBS plates in triplicate. After a 17 hr overnight incubation at 25°C the bacterial colonies from each plate were scraped with a sterile cell scraper into 1 ml of 70% IO and collected by centrifugation. Cell pellets were stored at -20°C prior to DNA extraction.

Genomic DNA from the cell pellets was extracted and purified using the Qiagen DNeasy Blood and Tissue Kit (Qiagen; 69506) following the Gram-negative bacteria protocol, and was quantified using a NanoDrop spectrophotometer (Thermo Scientific). The barcoded scars were amplified together with dual-index Illumina sequencing primers (66). The reaction mixtures contained 50 ng of gDNA, 200 nM of each oligo (Table S2), 1X Phusion Hot Start Flex Master Mix (NEB; M0536L) and H_2O up to a total volume of 50 µl. PCR conditions were 98°C for 30 sec, 98°C for 10 sec, 60°C for 10 sec, 72°C for 10 sec (20 cycles), with a final extension step at 72°C for 5 min. PCR products were visualized using a 2% agarose gel to confirm the dual-indexed amplicon of 231 bp and purified using a QIAquick PCR Purification Kit (Qiagen; 28106). Purified PCR products were quantified using a Qubit 3 fluorometer (Life Technologies) and pooled in equimolar amounts, and diluted to 4 nM. The pool was sequenced on an Illumina MiSeq using the 2 X 250 bp v2 kit with a 10% PhiX control following the manufacturer's protocol (Illumina, Inc., San Diego, CA) and using custom primers developed from (66). The sequencing data was processed using the barseq python package to obtain strain counts per sample, and mutants that were in the input library but still being validated were removed from the dataset. The relative frequency (RF) for each strain in a sample was calculated, normalized to the RF in

589 the input library and the average RF in the sample, and the competitive index (CI) was then

590 calculated using the formula: $CI = Log_{10} [(RF_{mutant}/Avg.\ RF_{WT})_{Sample}/(RF_{mutant}/Avg.\ RF_{WT})_{Input}]$.

FIGURE LEGENDS

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

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

FIG 3 Assaying the phenotype of strains deleted for copper-resistance factors. (A and B) Bar graphs showing the average OD_{600} of the indicated Δ*copA* mutants after 20 hr of growth in the presence of the indicated amounts of copper and/or bathocuproinedisulfonic acid (BCS). (A)

Error bars represent the standard deviation of the mean (*n* = 3). (B) Data are from two independent replicates (*n* = 2). Statistical analysis was performed using a Two-way ANOVA test. **** *P* < 0.0001.

FIG 4 Assaying the phenotype of strains deleted for motility factors. Representative TBS agar plates showing the migration of strains from the inoculation point after incubation at 28°C for 4 hr. Bar graph shows the quantified data from three independent replicates with error bars showing the standard deviation of the mean (*n* = 3). Statistical analysis was performed using a Two-way ANOVA test. **** *P* < 0.0001.

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

FIG 6 Overview of BarSeq experiments and computational package. Methodology and software for performing BarSeq experiments as described in the main text. An input population is used to inoculate squid or media, and samples are taken at the times of interest for gDNA extraction and processing to be sequenced by Illumina sequencing. The i5 and i7 segments are the index sequences in the dual indexed DNA fragments, whereas the P5 and P7 sequences are

sequencing adapters for the MiSeq flow cell (details in Table S2). Sequencing reads are analyzed by the barseq package to obtain counts for individual strains in a sample based on their unique barcodes. Those counts are then used to calculate the relative frequencies of individual strains at each timepoint and the competitive index (CI) as described in Materials and Methods.

FIG 7 BarSeq enables high-throughput competition experiments. (A and B) Graphs show the mean competitive index (CI) on a Log_{10}-scale for each barcoded strain in the population using the WT strains as controls as described in Brooks et al. (92) after (A) 15 generations *in vitro* in LBS and (B) 48 hr post-squid infection. LOD = limit of detection for the experiment (3.39×10^{-5}). Each symbol represents one biological replicate. Statistical analysis was performed using a One-way ANOVA test comparing each strain to WT-1. *** $P < 0.001$, **** $P < 0.0001$.

Table 1. Bacterial strains. [a]Thymidine auxotroph, growth conditions in Materials and Methods.

N/A = Not applicable.

Strain	Alias	Genotype/Description	Source
V. fischeri			
MJM1100	ES114 (WT)	ES114	91, 93
MJM1538	ES114/pLostfoX	MJM1100/pLostfoX	44
MJM1902	Δ*copA*::Tn*erm*	MJM1100 Δ*copA*::Tn*erm*	44
MJM3400	Δ*copA*::pEVS79-Δ*copA*	MJM1100 Δ*copA*::pEVS79-Δ*copA*	This work
MJM3401	Δ*copA*	MJM1100 Δ*copA*	This work
MJM3529	Δ*copA*::*erm-bar*	MJM1100 Δ*copA*::*erm-bar*	This work
MJM3534	Δ*cusA*::*erm-bar*	MJM1100 Δ*cusA*::*erm-bar*	This work
MJM3543	Δ*copA*::*bar*	MJM1100 Δ*copA*::*bar*	This work
MJM3565	Δ*cusA*::*bar*	MJM1100 Δ*cusA*::*bar*	This work
MJM3620	WT::*erm-bar*1	MJM1100 IG(*yeiR-glmS*)::*erm-bar*1	This work
MJM3621	WT::*erm-bar2*	MJM1100 IG(*yeiR-glmS*)::*erm-bar*2	This work
MJM3622	WT::*erm-bar3*	MJM1100 IG(*yeiR-glmS*)::*erm-bar*3	This work
MJM3629	WT::*bar*1	MJM1100 IG(*yeiR-glmS*)::*bar*1	This work
MJM3630	WT::*bar2*	MJM1100 IG(*yeiR-glmS*)::*bar*2	This work
MJM3631	WT::*bar3*	MJM1100 IG(*yeiR-glmS*)::*bar*3	This work
MJM3785	Δ*flrA*::*erm-bar*	MJM1100 Δ*flrA*::*erm-bar*	This work
MJM3785	Δ*flaA*::*erm-bar*	MJM1100 Δ*flaA*::*erm-bar*	This work
MJM3786	Δ*rpoN*::*erm-bar*	MJM1100 Δ*rpoN*::*erm-bar*	This work
MJM3788	Δ*cheA*::*erm-bar*	MJM1100 Δ*cheA*::*erm-bar*	This work
MJM3790	Δ*copA*::*bar att*Tn7::*copA*	MJM1100 Δ*copA*::*bar att*Tn7::*copA*	This work
MJM3792	Δ*flrA*::*bar*	MJM1100 Δ*flrA*::*bar*	This work
MJM3795	Δ*flaA*::*bar*	MJM1100 Δ*flaA*::*bar*	This work
MJM3796	Δ*rpoN*::*bar*	MJM1100 Δ*rpoN*::*bar*	This work
MJM3798	Δ*cheA*::*bar*	MJM1100 Δ*cheA*::*bar*	This work
E. coli			
MJM534	CC118 λ*pir*/pEVS104	Δ(*ara-leu*) *araD* Δ*lacX74 galE galK phoA20 thi-1 rpsE rpoB argE*(Am)	45

		recA1, lysogenized with λ*pir*/pEVS104	
MJM537	DH5α λ*pir*	F- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *supE44 hsdR17* (r_K^-, m_K^+) *endA1 recA1 gyrA96 thi-1 relA1 uidA*::*pir*$^+$	Laboratory stock
MJM570	DH5α/pEVS79	F- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *supE44 hsdR17* (r_K^-, m_K^+) *endA1 recA1 gyrA96 thi-1 relA1*/pEVS79	45
MJM637	S17-1 λ*pir*/pUX-BF13	*pro res− hsdR17* (r_K^- m_K^+) *recA−* with an integrated *RP4-2-Tc*::*Mu-Km*::Tn*7* λ*pir*/pUX-BF13	86, 87
MJM658	BW23474/pEVS107	Δ*lac-169 robA1 creC510 hsdR514 uidA*(Δ*Mlu*I)::*pir116 endA*(*BT33*) *recA1*/pEVS107	80
MJM3287	NEB5α/pHB1	F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *glnV44 hsdR17* (r_K^-, m_K^+) *endA1 recA1 gyrA96 thi-1 relA1 fhuA2 phoA*/pHB1	73
MJM3288	DH5α λ*pir/*pHB2	MJM537/pHB2	This work
MJM3383	NEB5α/pHB3	F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *glnV44 hsdR17* (r_K^-, m_K^+) *endA1 recA1 gyrA96 thi-1 relA1 fhuA2 phoA*/pHB3	This work
MJM3478	KV8052: π3813[a]/pKV496	*lacI*^{q} *thi-1 supE44 endA1 recA1 hsdR17 gyrA462 zei-298*::Tn*10* Δ*thyA*::(*erm-pir-116*)/pKV496	49, 85

649
650

Table 2. Plasmids.

Plasmid	Relevant properties	Source
pEVS79	Vector backbone for deletion construct via allelic-exchange, Cam^R	45
pEVS104	Conjugation helper plasmid, Kan^R	45
pEVS107	mini-Tn7 mobilizable vector, Erm^R (transposon), Kan^R	80
pKV496	pEVS79 containing the FLP recombinase, Kan^R	49
pLostfoX	*tfoX* overexpression vector, Cam^R	55
pUC19	Cloning vector, $Carb^R$	Laboratory stock
pUX-BF13	Tn7 transposase helper plasmid (*tns* genes), $Carb^R$	86
pHB1	pUC19 containing the LL-FRT-*erm*-FRT-spacer sequence in the HindIII/BamHI site	73
pHB2	pEVS107 containing *copA* (including 191 bp upstream and 321 bp downstream of the *copA* ORF) at the AscI site	This work
pHB3	pEVS79 containing 1.6-kb upstream/1.6-kb downstream of *copA*	This work

Supplementary Tables S1-S3:

- **Table S1. Expanded bacterial strains.** [a]Thymidine auxotroph, growth conditions in Materials and Methods. N/A = Not applicable.
- **Table S2. DNA oligonucleotides.**
- **Table S3. Strain counts from barseq output.** Strain counts for individual strains within each sample was obtained using the barseq package. The samples for competitive squid colonization (48 hpi) were processed in triplicate as technical replicates. The counts in '_other' represent sequence reads that contain the appropriate sequences flanking the barcode region but the barcode sequence does not match any present in the reference barcode library (as described in Materials and Methods).

ACKNOWLEDGMENTS

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

FUNDING SOURCES

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

REFERENCES

1. Taschuk R, Griebel PJ. 2012. Commensal microbiome effects on mucosal immune system development in the ruminant gastrointestinal tract. Anim Health Res Rev 13:129–141.

2. Montgomery MK, McFall-Ngai M. 1994. Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. Development 120:1719–1729.

3. Shropshire JD, Bordenstein SR. 2016. Speciation by symbiosis: the microbiome and behavior. MBio 7:e01785.

4. Honda K, Littman DR. 2012. The microbiome in infectious disease and inflammation. Annu Rev Immunol 30:759–795.

5. Byrd AL, Belkaid Y, Segre JA. 2018. The human skin microbiome. Nat Rev Microbiol 16:143–155.

6. O'Dwyer DN, Dickson RP, Moore BB. 2016. The lung microbiome, immunity, and the pathogenesis of chronic lung disease. J Immunol 196:4839–4847.

7. Aragón IM, Herrera-Imbroda B, Queipo-Ortuño MI, Castillo E, Del Moral JS-G, Gómez-Millán J, Yucel G, Lara MF. 2018. The urinary tract microbiome in health and disease. Eur Urol Focus 4:128–138.

8. Kinross JM, Darzi AW, Nicholson JK. 2011. Gut microbiome-host interactions in health and

disease. Genome Med 3:14.

9. Shreiner AB, Kao JY, Young VB. 2015. The gut microbiome in health and in disease. Curr Opin Gastroenterol 31:69–75.

10. Lazar V, Ditu L-M, Pircalabioru GG, Gheorghe I, Curutiu C, Holban AM, Picu A, Petcu L, Chifiriuc MC. 2018. Aspects of gut microbiota and immune system interactions in infectious diseases, immunopathology, and cancer. Front Immunol 9:1830.

11. Vuong HE, Yano JM, Fung TC, Hsiao EY. 2017. The microbiome and host behavior. Annu Rev Neurosci 40:21–49.

12. Parfrey LW, Knight R. 2012. Spatial and temporal variability of the human microbiota. Clin Microbiol Infect 18 Suppl 4:8–11.

13. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. Science 326:1694–1697.

14. Kostic AD, Howitt MR, Garrett WS. 2013. Exploring host-microbiota interactions in animal models and humans. Genes Dev 27:701–718.

15. Ruby EG. 1999. The *Euprymna scolopes*-*Vibrio fischeri* symbiosis: a biomedical model for the study of bacterial colonization of animal tissue. J Mol Microbiol Biotechnol 1:13–21.

16. McFall-Ngai M. 2008. The squid-vibrio association- a naturally occurring, experimental

model of animal/bacterial partnerships, p. 102–112. *In* Gary B. Huffnagle, ANCN (ed.), GI Microbiota and Regulation of the Immune System. Landes Bioscience and Springer Science + Business Media.

17. McFall-Ngai MJ. 2014. The importance of microbes in animal development: lessons from the squid-vibrio symbiosis. Annu Rev Microbiol 68:177–194.

18. Alice H Tischler, Kelsey M Hodge-Hanson, and Karen L Visick. 2019. *Vibrio fischeri*–squid symbiosis. Elsevier Oceanogr Ser 267:1–9.

19. McFall-Ngai M. 2014. Divining the essence of symbiosis: insights from the squid-vibrio model. PLoS Biol 12:e1001783.

20. Nyholm SV, McFall-Ngai M. 2004. The winnowing: establishing the squid–vibrio symbiosis. Nat Rev Microbiol 2:632–642.

21. Jones BW, Nishiguchi MK. 2004. Counterillumination in the Hawaiian bobtail squid, *Euprymna scolopes* Berry (Mollusca: Cephalopoda). Mar Biol 144:1151–1155.

22. Lee KH, Ruby EG. 1994. Effect of the squid host on the abundance and distribution of symbiotic *Vibrio fischeri* in nature. Appl Environ Microbiol 60:1565–1571.

23. Naughton LM, Mandel MJ. 2012. Colonization of *Euprymna scolopes* squid by *Vibrio fischeri*. J Vis Exp e3758.

24. Mandel MJ, Schaefer AL, Brennan CA, Heath-Heckman EAC, DeLoney-Marino CR,

McFall-Ngai MJ, Ruby EG. 2012. Squid-derived chitin oligosaccharides are a chemotactic signal during colonization by *Vibrio fischeri*. Appl Environ Microbiol 78:4620–4626.

25. Essock-Burns T, Bongrand C, Goldman WE, Ruby EG, McFall-Ngai MJ. 2020. Interactions of symbiotic partners drive the development of a complex biogeography in the squid-vibrio symbiosis. MBio 11.

26. Sycuro LK, Ruby EG, McFall-Ngai M. 2006. Confocal microscopy of the light organ crypts in juvenile *Euprymna scolopes* reveals their morphological complexity and dynamic function in symbiosis. J Morphol 267:555–568.

27. Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ. 2002. Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. Appl Environ Microbiol 68:5113–5122.

28. Nyholm SV, Stabb EV, Ruby EG, McFall-Ngai MJ. 2000. Establishment of an animal–bacterial association: recruiting symbiotic vibrios from the environment. Proc Natl Acad Sci U S A 97:10231–10235.

29. Altura MA, Heath-Heckman EAC, Gillette A, Kremer N, Krachler A-M, Brennan C, Ruby EG, Orth K, McFall-Ngai MJ. 2013. The first engagement of partners in the *Euprymna scolopes-Vibrio fischeri* symbiosis is a two-step process initiated by a few environmental symbiont cells. Environ Microbiol 15:2937–2950.

30. Visick KL. 2009. An intricate network of regulators controls biofilm formation and

colonization by *Vibrio fischeri*. Mol Microbiol 74:782–789.

31. Brennan CA, Mandel MJ, Gyllborg MC, Thomasgard KA, Ruby EG. 2013. Genetic determinants of swimming motility in the squid light-organ symbiont *Vibrio fischeri*. Microbiologyopen 2:576–594.

32. Wang Y, Ruby EG. 2011. The roles of NO in microbial symbioses. Cell Microbiol 13:518–526.

33. Stabb E.V., Schaefer A., Bose J.L., Ruby E.G. 2008. Quorum signaling and symbiosis in the marine luminous bacterium *Vibrio fischeri*, p. 233–250. *In* Winans, Bassler (eds.), Chemical Communication among Bacteria. American Society of Microbiology.

34. Visick KL, Foster J, Doino J, McFall-Ngai M, Ruby EG. 2000. *Vibrio fischeri lux* genes play an important role in colonization and development of the host light organ. J Bacteriol 182:4578–4586.

35. Lupp C, Urbanowski M, Greenberg EP, Ruby EG. 2003. The *Vibrio fischeri* quorum-sensing systems *ain* and *lux* sequentially induce luminescence gene expression and are important for persistence in the squid host. Mol Microbiol 50:319–331.

36. Visick KL, Ruby EG. 2006. *Vibrio fischeri* and its host: it takes two to tango. Curr Opin Microbiol 9:632–638.

37. Visick KL, Skoufos LM. 2001. Two-component sensor required for normal symbiotic colonization of *Euprymna scolopes* by *Vibrio fischeri*. J Bacteriol 183:835–842.

38. Lupp C, Ruby EG. 2005. *Vibrio fischeri* uses two quorum-sensing systems for the regulation of early and late colonization factors. J Bacteriol 187:3620–3629.

39. Whistler CA, Ruby EG. 2003. GacA regulates symbiotic colonization traits of *Vibrio fischeri* and facilitates a beneficial association with an animal host. J Bacteriol 185:7202–7212.

40. Schwartzman JA, Koch E, Heath-Heckman EAC, Zhou L, Kremer N, McFall-Ngai MJ, Ruby EG. 2015. The chemistry of negotiation: rhythmic, glycan-driven acidification in a symbiotic conversation. Proc Natl Acad Sci U S A 112:566–571.

41. Singh P, Brooks JF 2nd, Ray VA, Mandel MJ, Visick KL. 2015. CysK plays a role in biofilm formation and colonization by *Vibrio fischeri*. Appl Environ Microbiol 81:5223–5234.

42. O'Shea TM, Klein AH, Geszvain K, Wolfe AJ, Visick KL. 2006. Diguanylate cyclases control magnesium-dependent motility of *Vibrio fischeri*. J Bacteriol 188:8196–8205.

43. Lyell NL, Dunn AK, Bose JL, Stabb EV. 2010. Bright mutants of *Vibrio fischeri* ES114 reveal conditions and regulators that control bioluminescence and expression of the *lux* operon. J Bacteriol 192:5103–5114.

44. Brooks JF 2nd, Gyllborg MC, Cronin DC, Quillin SJ, Mallama CA, Foxall R, Whistler C, Goodman AL, Mandel MJ. 2014. Global discovery of colonization determinants in the squid symbiont *Vibrio fischeri*. Proc Natl Acad Sci U S A 111:17284–17289.

45. Stabb EV, Ruby EG. 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods Enzymol 358:413–426.

46. Baym M, Shaket L, Anzai IA, Adesina O, Barstow B. 2016. Rapid construction of a whole-genome transposon insertion collection for *Shewanella oneidensis* by Knockout Sudoku. Nat Commun 7:13270.

47. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. 2009. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 6:279–289.

48. Studer SV, Mandel MJ, Ruby EG. 2008. AinS quorum sensing regulates the *Vibrio fischeri* acetate switch. J Bacteriol 190:5915–5923.

49. Visick KL, Hodge-Hanson KM, Tischler AH, Bennett AK, Mastrodomenico V. 2018. Tools for rapid genetic engineering of *Vibrio fischeri*. Appl Environ Microbiol 84:e00850–18.

50. Jahn LJ, Porse A, Munck C, Simon D, Volkova S, Sommer MOA. 2018. Chromosomal barcoding as a tool for multiplexed phenotypic characterization of laboratory evolved lineages. Sci Rep 8:6961.

51. Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, Chee M, Roth FP, Giaever G, Nislow C. 2009. Quantitative phenotyping via deep barcode sequencing. Genome Res 19:1836–1842.

52. Robinson DG, Chen W, Storey JD, Gresham D. 2014. Design and analysis of Bar-seq experiments. G3 4:11–18.

53. Blundell JR, Levy SF. 2014. Beyond genome sequencing: lineage tracking with barcodes to

study the dynamics of evolution, infection, and cancer. Genomics 104:417–430.

54. Abel S, Abel zur Wiesch P, Chang H-H, Davis BM, Lipsitch M, Waldor MK. 2015. Sequence tag-based analysis of microbial population dynamics. Nat Methods 12:223–226.

55. Pollack-Berti A, Wollenberg MS, Ruby EG. 2010. Natural transformation of *Vibrio fischeri* requires *tfoX* and *tfoY*. Environ Microbiol 12:2302–2311.

56. De Souza Silva O, Blokesch M. 2010. Genetic manipulation of *Vibrio cholerae* by combining natural transformation with FLP recombination. Plasmid 64:186–195.

57. Dalia AB, Lazinski DW, Camilli A. 2013. Characterization of undermethylated sites in *Vibrio cholerae*. J Bacteriol 195:2389–2399.

58. Baba T, Mori H. 2008. The construction of systematic in-frame, single-gene knockout mutant collection in *Escherichia coli* K-12, p. 171–181. *In* Osterman, AL, Gerdes, SY (eds.), Microbial Gene Essentiality: Protocols and Bioinformatics. Humana Press, Totowa, NJ.

59. Hernández-Montes G, Argüello JM, Valderrama B. 2012. Evolution and diversity of periplasmic proteins involved in copper homeostasis in gamma proteobacteria. BMC Microbiol 12:249.

60. Rensing C, Fan B, Sharma R, Mitra B, Rosen BP. 2000. CopA: An *Escherichia coli* Cu(I)-translocating P-type ATPase. Proc Natl Acad Sci U S A 97:652–656.

61. Millikan DS, Ruby EG. 2003. FlrA, a σ^{54}-dependent transcriptional activator in *Vibrio*

fischeri, is required for motility and symbiotic light-organ colonization. J Bacteriol 185:3547–3557.

62. Wolfe AJ, Millikan DS, Campbell JM, Visick KL. 2004. *Vibrio fischeri* σ^{54} controls motility, biofilm formation, luminescence, and colonization. Appl Environ Microbiol 70:2520–2524.

63. Raynal A, Karray F, Tuphile K, Darbon-Rongère E, Pernodet J-L. 2006. Excisable cassettes: new tools for functional analysis of *Streptomyces* genomes. Appl Environ Microbiol 72:4839–4844.

64. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008.

65. Oppenheim DS, Yanofsky C. 1980. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. Genetics 95:785–795.

66. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112–5120.

67. Millikan DS, Ruby EG. 2004. *Vibrio fischeri* flagellin A is essential for normal motility and for symbiotic competence during initial squid light organ colonization. J Bacteriol 186:4315–4325.

68. Koonin EV. 2009. Evolution of genome architecture. Int J Biochem Cell Biol 41:298–306.

69. Bongrand C, Ruby EG. 2019. The impact of *Vibrio fischeri* strain variation on host colonization. Curr Opin Microbiol 50:15–19.

70. Wollenberg MS, Ruby EG. 2009. Population structure of *Vibrio fischeri* within the light organs of *Euprymna scolopes* squid from two Oahu (Hawaii) populations. Appl Environ Microbiol 75:193–202.

71. Bongrand C, Ruby EG. 2018. Achieving a multi-strain symbiosis: strain behavior and infection dynamics. ISME J 13:698–706.

72. Wollenberg MS, Ruby EG. 2012. Phylogeny and fitness of *Vibrio fischeri* from the light organs of *Euprymna scolopes* in two Oahu, Hawaii populations. ISME J 6:352–362.

73. Rotman ER, Bultman KM, Brooks JF 2nd, Gyllborg MC, Burgos HL, Wollenberg MS, Mandel MJ. 2019. Natural strain variation reveals diverse biofilm regulation in squid-colonizing *Vibrio fischeri*. J Bacteriol 201.

74. Bongrand C, Moriano-Gutierrez S, Arevalo P, McFall-Ngai M, Visick KL, Polz M, Ruby EG. 2020. Using colonization assays and comparative genomics to discover symbiosis behaviors and factors in *Vibrio fischeri*. MBio 11.

75. Pankey MS, Foxall RL, Ster IM, Perry LA, Schuster BM, Donner RA, Coyle M, Cooper VS, Whistler CA. 2017. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. Elife 6:e24414.

76. Yip ES, Grublesky BT, Hussa EA, Visick KL. 2005. A novel, conserved cluster of genes

promotes symbiotic colonization and σ^{54}-dependent biofilm formation by *Vibrio fischeri*. Mol Microbiol 57:1485–1498.

77. Hunt TP, Magasanik B. 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc Natl Acad Sci U S A 82:8453–8457.

78. Soutourina OA, Bertin PN. 2003. Regulation cascade of flagellar expression in Gram-negative bacteria. FEMS Microbiol Rev 27:505–523.

79. Prouty MG, Correa NE, Klose KE. 2001. The novel σ^{54}- and σ^{28}-dependent flagellar gene transcription hierarchy of *Vibrio cholerae*. Mol Microbiol 39:1595–1609.

80. McCann J, Stabb EV, Millikan DS, Ruby EG. 2003. Population dynamics of *Vibrio fischeri* during infection of *Euprymna scolopes*. Appl Environ Microbiol 69:5928–5934.

81. Becker KW, Skaar EP. 2014. Metal limitation and toxicity at the interface between host and pathogen. FEMS Microbiol Rev 38:1235–1249.

82. Fu Y, Chang F-MJ, Giedroc DP. 2014. Copper transport and trafficking at the host-bacterial pathogen interface. Acc Chem Res 47:3605–3613.

83. Hodgkinson V, Petris MJ. 2012. Copper homeostasis at the host-pathogen interface. J Biol Chem 287:13549–13555.

84. McAnulty SJ, Nyholm SV. 2016. The role of hemocytes in the Hawaiian bobtail squid,

Euprymna scolopes: a model organism for studying beneficial host-microbe interactions. Front Microbiol 7:2013.

85. Le Roux F, Binesse J, Saulnier D, Mazel D. 2007. Construction of a *Vibrio splendidus* mutant lacking the metalloprotease gene *vsm* by use of a novel counterselectable suicide vector. Appl Environ Microbiol 73:777–784.

86. Bao Y, Lies DP, Fu H, Roberts GP. 1991. An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. Gene 109:167–168.

87. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Biotechnology 1:784–791.

88. Khodursky AB, Bernstein JA, Peter BJ, Rhodius V, Wendisch VF, Zimmer DP. 2003. *Escherichia coli* spotted double-strand DNA microarrays: RNA extraction, labeling, hybridization, quality control, and data management. Methods Mol Biol 224:61–78.

89. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. 2016. ppGpp binding to a site at the RNAP-DksA interface accounts for its dramatic effects on transcription initiation during the stringent response. Mol Cell 62:811–823.

90. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. Methods 25:402–408.

91. Mandel MJ, Stabb EV, Ruby EG. 2008. Comparative genomics-based investigation of resequencing targets in *Vibrio fischeri*: focus on point miscalls and artefactual expansions. BMC Genomics 9:138.

92. Brooks JF 2nd, Mandel MJ. 2016. The histidine kinase BinK is a negative regulator of biofilm formation and squid colonization. J Bacteriol 198:2596–2607.

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

Figure 1

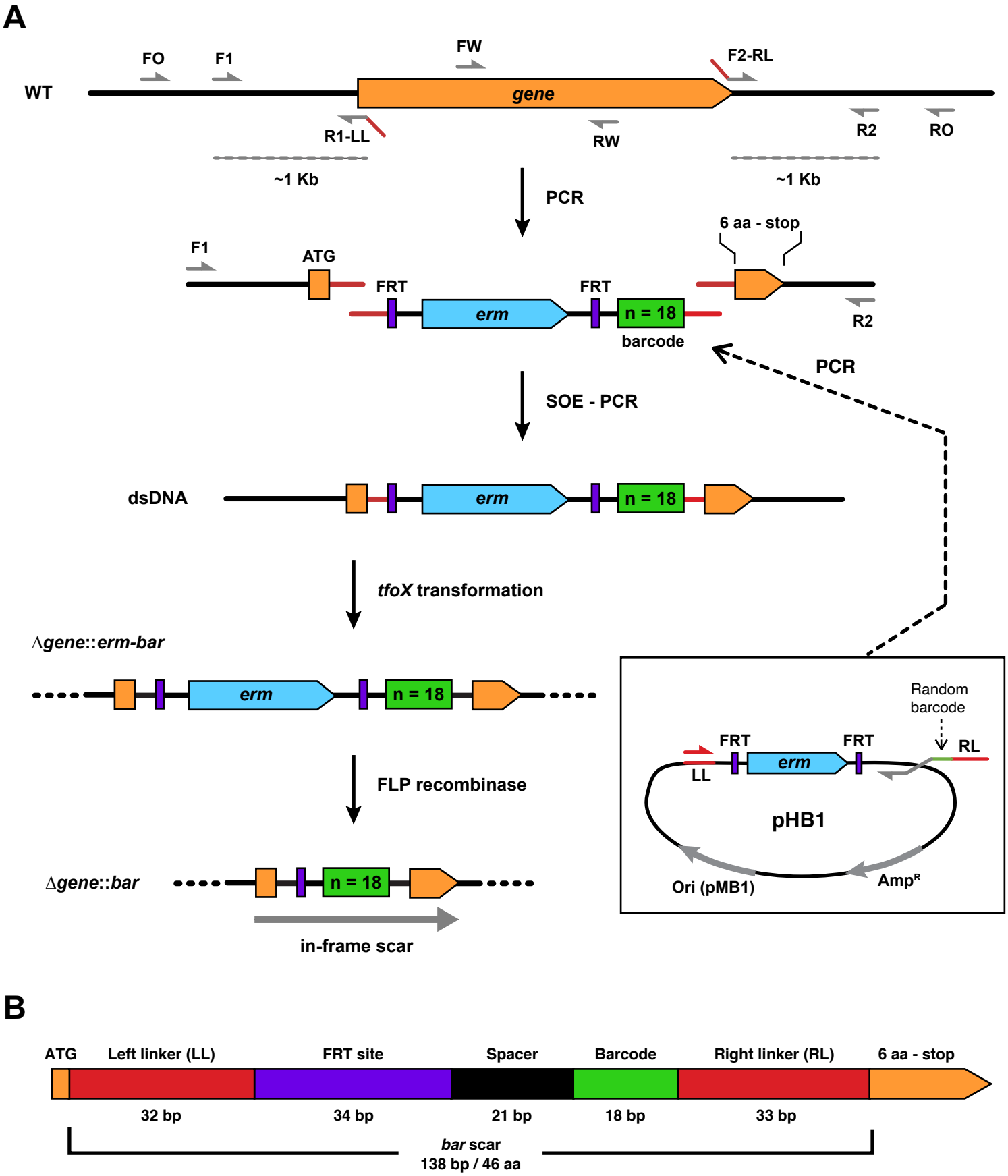
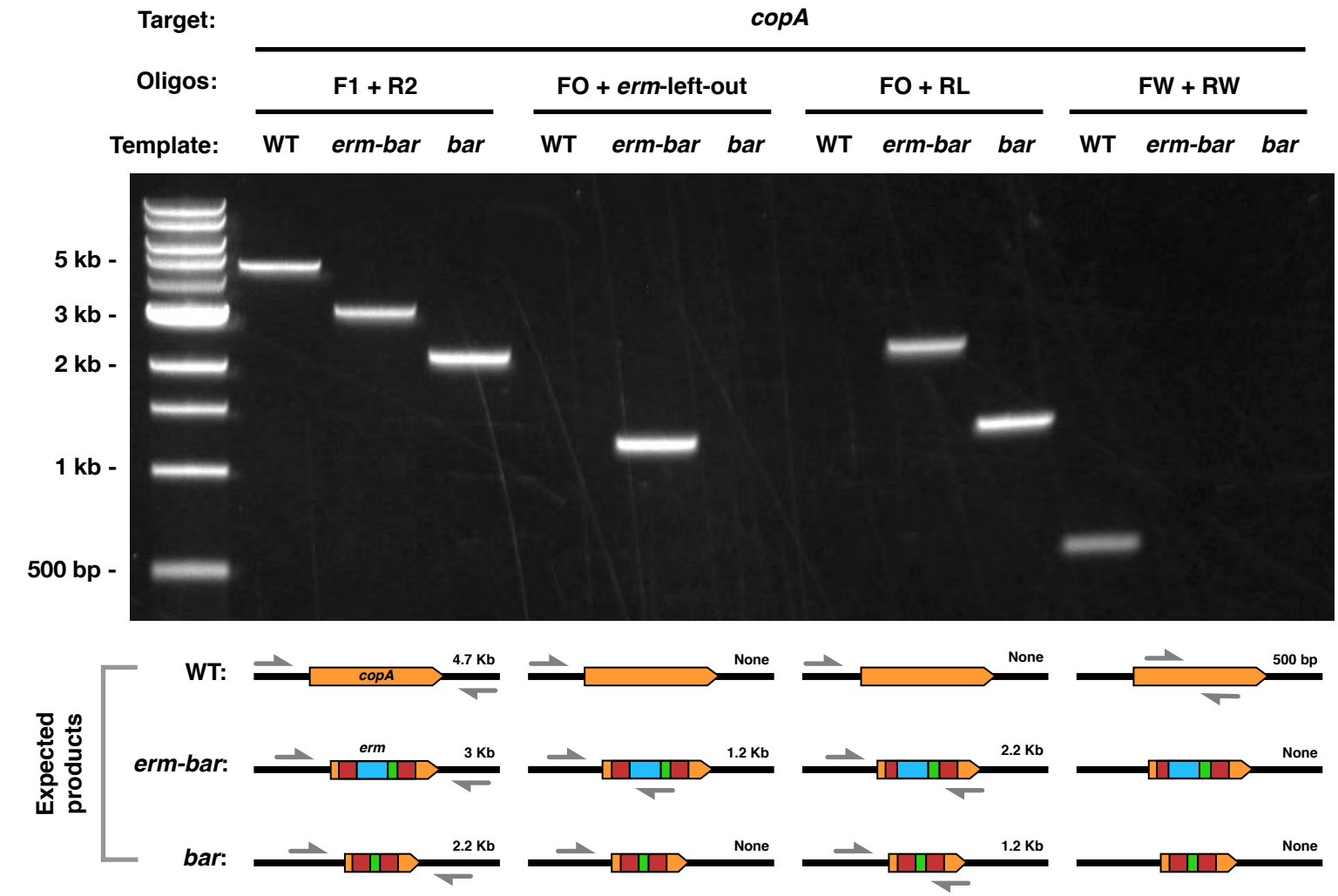

A
WT
FO
F1
FW
F2-RL
gene
R1-LL
RW
R2
RO
~1 Kb
~1 Kb
PCR
F1
ATG
FRT
erm
FRT
n = 18
barcode
6 aa - stop
R2
PCR
SOE - PCR
dsDNA
erm
n = 18
tfoX transformation
Δgene::erm-bar
erm
n = 18
FLP recombinase
Δgene::bar
n = 18
in-frame scar
Random barcode
FRT
FRT
RL
LL
erm
pHB1
Ori (pMB1)
AmpR
B
ATG
Left linker (LL)
FRT site
Spacer
Barcode
Right linker (RL)
6 aa - stop
32 bp
34 bp
21 bp
18 bp
33 bp
bar scar
138 bp / 46 aa

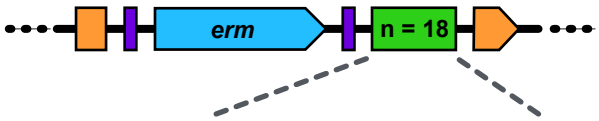
Figure 2

A



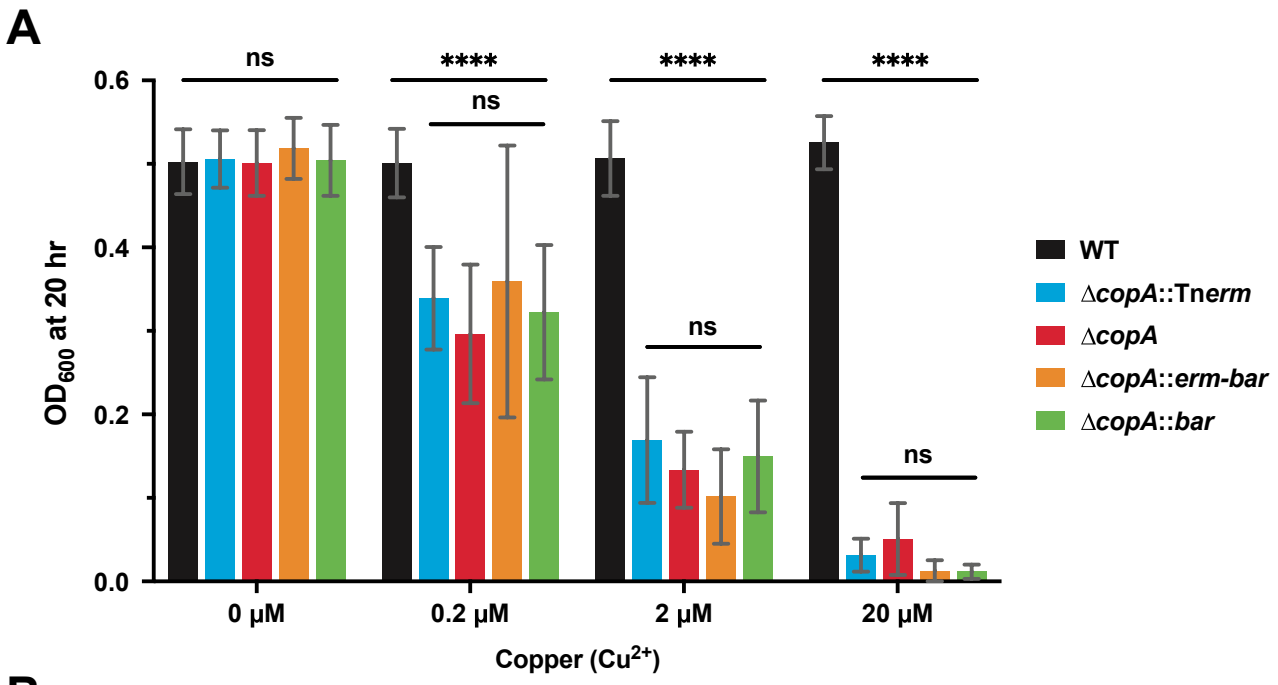
B

Δ*copA*::*erm-bar* deletion scar



Barcode	VNNVNNVNNVNNVNNVNN
Isolate A	GGCCCCCCGCCCGTCCCC
Isolate B	ATGAAGACTGTTGCCGTA
Isolate C	CACGACGCCCTCCGCGGA
Isolate D	ACTATTACGCAAAATAAT

Figure 3


A
ns
****
****
****
ns
ns
ns
0.6
0.4
0.2
0.0
OD600 at 20 hr
0 μM
0.2 μM
2 μM
20 μM
Copper (Cu2+)
WT
ΔcopA::Tnerm
ΔcopA
ΔcopA::erm-bar
ΔcopA::bar

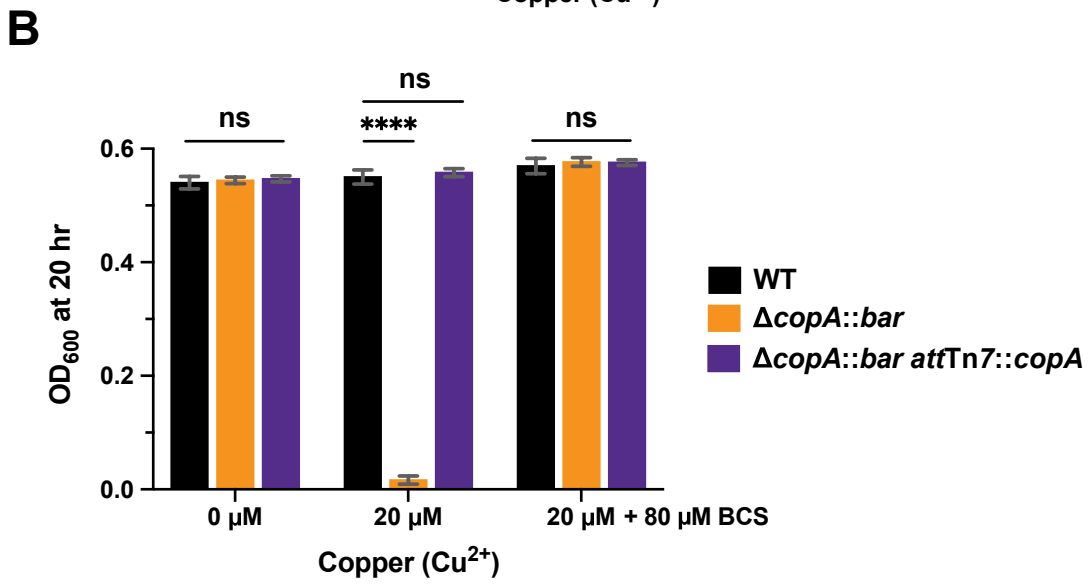

B
ns
ns
****
ns
0.6
0.4
0.2
0.0
OD600 at 20 hr
0 μM
20 μM
20 μM + 80 μM BCS
Copper (Cu2+)
WT
ΔcopA::bar
ΔcopA::bar attTn7::copA

Figure 4

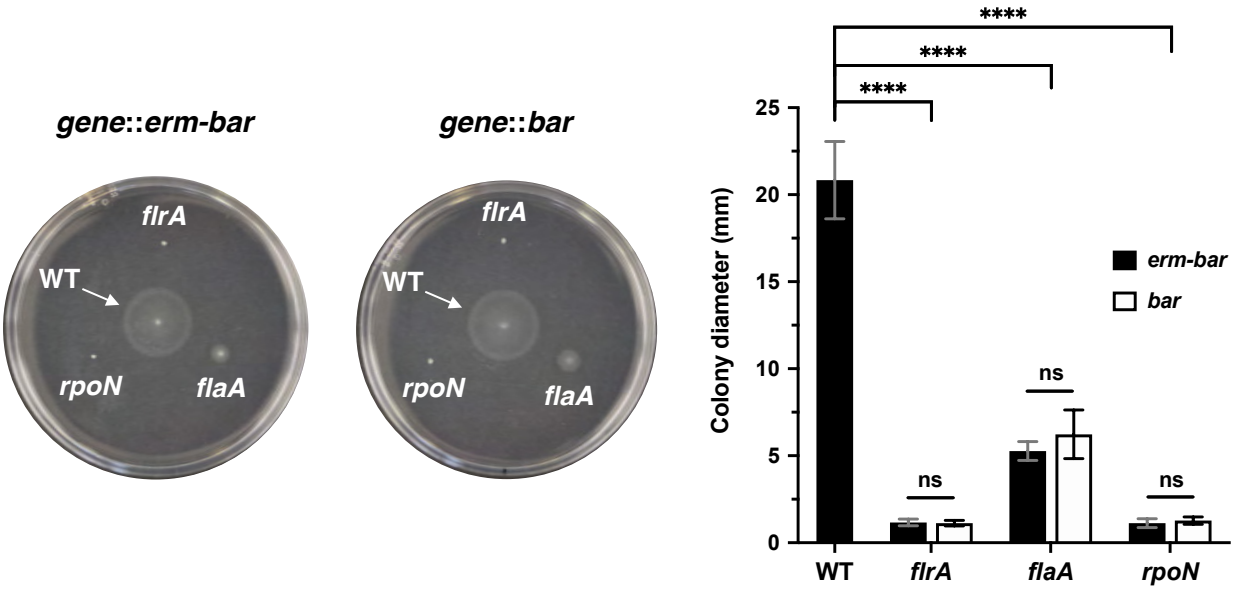

gene::erm-bar
gene::bar
flrA
WT
rpoN
flaA
flrA
WT
rpoN
flaA
****
****
****
ns
ns
ns
Colony diameter (mm)
25
20
15
10
5
0
erm-bar
bar
WT
flrA
flaA
rpoN

Figure 5

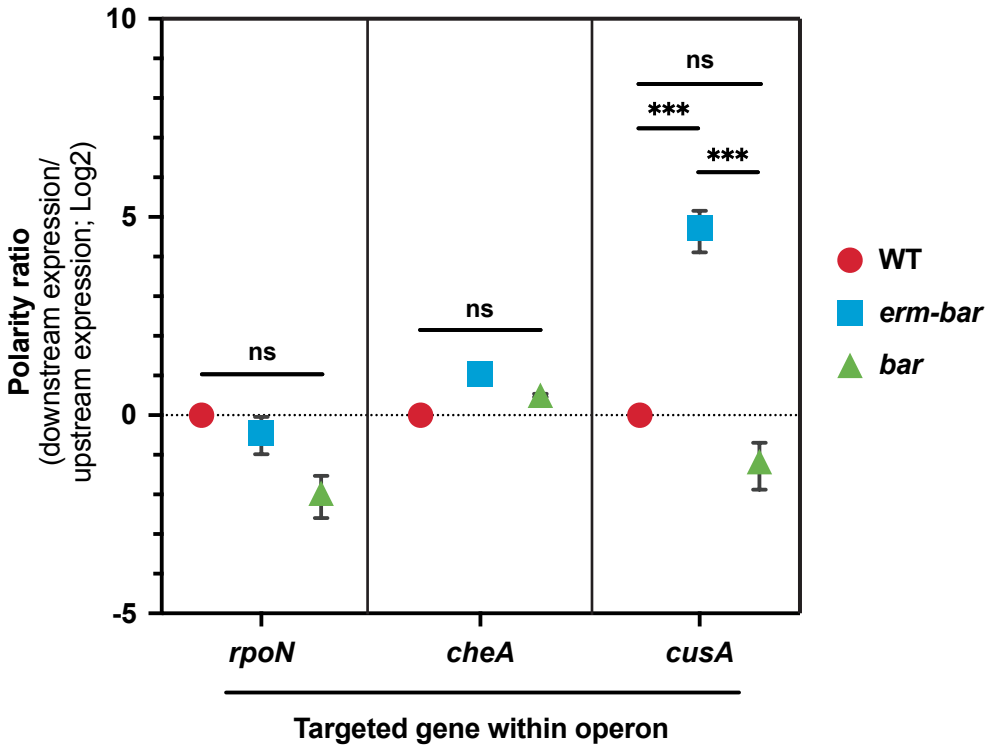
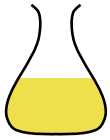

10
5
0
-5
Polarity ratio
(downstream expression/
upstream expression; Log2)
ns
ns
ns
***
***
rpoN
cheA
cusA
Targeted gene within operon
WT
erm-bar
bar

Figure 6

Input population (synthetic microbiome)

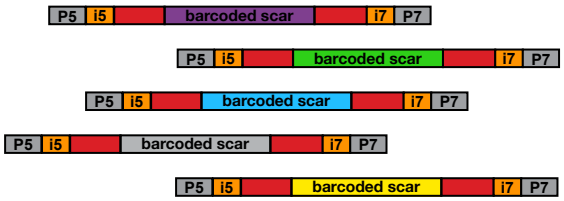


Inoculate culture medium and/or squid hatchlings

Grow in vitro for the desired number of generations and/or allow colonization to proced 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
...	...	...

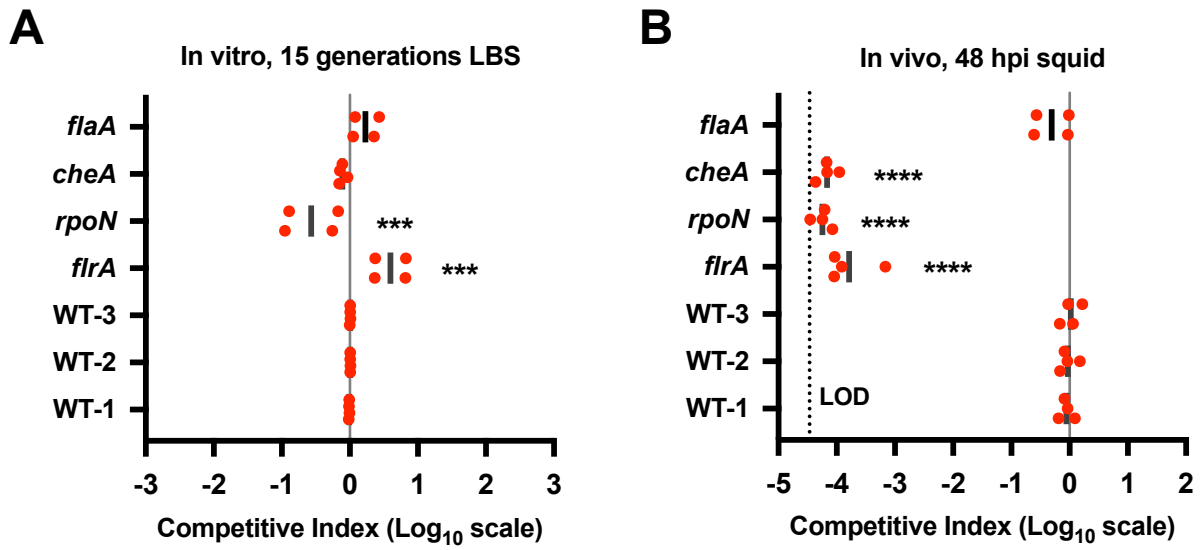
Barseq package: identifies barcodes and assigns strain identity

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

$$CI = \mathrm{Log}_{10}\left(\frac{\left(\frac{RF_{Mut}}{\overline{RF}_{WT}}\right)_{Sample}}{\left(\frac{RF_{Mut}}{\overline{RF}_{WT}}\right)_{Input}}\right)$$

Normalize counts to input and calculate competitive index (CI)

Figure 7


A
In vitro, 15 generations LBS
flaA
cheA
rpoN
flrA
WT-3
WT-2
WT-1
***
***
-3
-2
-1
0
1
2
3
Competitive Index (Log10 scale)
B
In vivo, 48 hpi squid
flaA
cheA
rpoN
flrA
WT-3
WT-2
WT-1
****
****
****
LOD
-5
-4
-3
-2
-1
0
1
2
Competitive Index (Log10 scale)