

Multi-scale systems
analysis of the squid-vibrio
symbiosis

Thesis by
Vera Michelle Beilinson

In Partial Fulfillment of the Requirements for
the degree of
Biology

The Caltech logo, featuring the word "Caltech" in a bold, orange, sans-serif font, centered within a light orange rectangular background.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2027
Defended June 22, 2026

© 2027

Vera Michelle Beilinson
ORCID: 0000-0002-1259-733X

ACKNOWLEDGEMENTS

First and foremost, I want to state how fortunate I have been to be at Caltech, getting to meet and work with such amazing scientists. I am thankful for every PI, postdoc, graduate student, and undergraduate I have had the privilege of interacting with during my time here.

I would like to express my deepest gratitude to my primary advisor, Dr. Margaret McFall-Ngai, for her support and mentorship over the past years. Margaret's passion, curiosity, and positive science attitude is truly inspiring and has kept me going when projects were at a low. I will especially appreciate her encouragement for her students and postdocs to explore their own scientific interest. My time in the McFall-Ngai lab has taught me that passion for your work is essential for success and can bring a lifetime of happiness.

I would like to also acknowledge my second advisor, Dr. Lior Pachter, who jumped into the role of my second advisor without knowing what he got himself into- the world of squid. Having studied mathematics as an undergraduate before and transitioning into biology, I deeply respect Lior's rigor and care for data, as well as his enthusiasm for learning across disciplines in support of his students.

I also thank for the rest of my thesis committee: Drs. Edward Ruby, Sarkis Mazmanian, and Smruthi Karthikeyan. Not only are they exceptional scientists, but amazing people who have supported me as a scientist and as a person.

I also thank all the members of the squid-vibrio community, especially the McFall-Ngai and Ruby labs. We have been together through many transitions building the lab at Caltech, welcoming new members, and exploring new kinds of experiments. I especially want to thank the lab technicians, Brandon Jatico, Nicolina Ranieri, and Amber Castro, for always keeping the lab running and vibrant. You have such amazing futures. I want to thank the postdocs, Drs. Jill Smith, Liu Yang, Xiao-Meng Hu, Viridiana Avila-Magaña, and Drew Honson, who have taught me so much in the time we overlapped. Thank you to Randall Scarborough and Grischa Chen for keeping the lab running smoothly and dealing with my antics, such as constant requests for a private jet. I also want to thank all of the Pachter lab for accepting me with open arms.

Dr. Kohta Ikegami – thank you for taking a chance on me as an undergraduate. I remember telling you I had zero biology experience, yet you took me in with open arms. You even believed that one day I would become an astronaut; while I am still far from that, I think about your encouragement

when the science gets tough. Thank you for introducing me to the world of bioinformatics, I have found my passion.

Dr. Tatyana Golvokina- I am the scientist I am today because of you. You taught me how to be a scientist: always think critically and design experiments well (controls!).

Alessandra Veinbach- my friend and colleague. I cannot express my gratitude to you for being my friend. I will always remember days in the Golvokina lab where we would gossip about boys while you taught me how to use a pipette, run a PCR, and so much more. You are an inspiration to all scientists; you grab challenges and make them look easy and with elegance and grace!

To Joclyn Mendes and every other single friend who has carried me throughout my PhD: AVM, BM, SF, AA, BG, DC, SS, GH, EK (and many more)- thank you for all the yapping.

Mom and dad – “It’s more fun to talk with someone who doesn’t use long, difficult words but rather short easy words like “What about lunch”” – Winnie-the-Pooh

Dr. Helen Beilinson- I cannot explain how lucky I am to have you as my sister and my colleague. You have bulldozed your way through life and let me take the easy road. I am incredibly thankful for all you have done for me in this life and am in awe of everything you have built.

To Lemon Beilinson-Kuang – woof woof woof.

And finally..

To Ryan Kuang – The best thing that has happened to me is falling so madly in love with you. As a kid, I always wished on shooting stars to find true love, like the ones you read about in books and see in TV shows. You have fulfilled my childhood dreams. I found true love and more in you. You have not only opened my heart to you and to others, but also to myself. I do not think I would be the person I am today without you. You are the reason I left the lab each day- to come home to you- and the reason I returned each morning- to make you proud. All I can say is - thank you.

“Me not working hard? Yeah, right, picture that with a Kodak. Or better yet, go to Times Square. take a picture of me with a Kodak.” – Pitbull

ABSTRACT

Animal–microbe symbioses are fundamental to the biology of nearly all multicellular organisms, yet many questions remain regarding how hosts recognize beneficial microbes and establish stable associations. The symbiosis between the Hawaiian bobtail squid, *Euprymna scolopes*, and the bioluminescent bacterium *Vibrio fischeri* provides a powerful model for investigating these processes.

In this dissertation, I examined host–microbe interactions across multiple biological scales, from bacterial genetic variation and tissue-level responses to cellular mechanisms of symbiosis. I demonstrate that bacterial strain identity influences host transcriptional responses. I further identify the skin as an early site of host–microbe interaction, revealing that hosts respond to symbiont exposure prior to stable colonization of the light organ. Finally, I establish methodologies for single-cell and single-nucleus transcriptomics in *E. scolopes*, providing a foundation for investigating symbiosis at cellular resolution.

Together, these findings show that host responses to symbiotic bacteria emerge from interconnected processes spanning bacterial genetics, host tissues, and individual cell types. This work advances our understanding of how beneficial host–microbe partnerships are established and maintained and provides new tools for studying symbiosis in cephalopods and other animal systems.

PUBLISHED CONTENT AND CONTRIBUTIONS

V. Beilinson, G.Y. Chen, A.C. Hargadon, E.G. Ruby, M.J. McFall-Ngai. 2025. Strain matters: host responses reflect symbiont origin in the squid-vibrio symbiosis. *mSystems* 10:e00498-25. <https://doi.org/10.1128/msystems.00498-25>

V.B. participated in the conception of the project, gathered and analyzed data, prepared the data, and participated in the writing of the manuscript.

J.T. Kuwabara, V. Beilinson, A.C. Hargadon, G.Y. Chen, X. Hu, M.S. Ladinsky, K.T. Hackett, J.P. Dillard, K.L. Visick, E.G. Ruby, & M. McFall-Ngai. 2026. SypC, a symbiont outer membrane vesicle protein, impacts the development of the squid–vibrio partnership. *Proc. Natl. Acad. Sci. U.S.A.* 123 (11) e2524648123. <https://doi.org/10.1073/pnas.2524648123>

V.B. participated in the data analysis and participated in the writing of the manuscript.

C. Oakes, V. Beilinson, M.J. McFall-Ngai, L. Pachter. 2025. Uniform pre-processing of bacterial single-cell RNA-seq. *bioRxiv* 2025.12.04.692398. <https://doi.org/10.64898/2025.12.04.692398>

V.B. participated in the conception of the project, gathered and analyzed data, and participated in the writing of the manuscript.

TABLE OF CONTENTS

Acknowledgements.....	iii
Abstract	v
Published Content and Contributions.....	vi
Table of Contents.....	vii
List of Illustrations and/or Tables.....	ix
Nomenclature.....	xii
Chapter I: Introduction.....	1
1.1 The squid-vibrio system as a model	2
1.2 Dynamic communication.....	4
1.3 Specificity and immune perception of symbiosis	5
1.4 Mutli-scale nature of host-microbe interactions.....	6
1.5 Overarching goals of dissertation.....	7
Chapter II: Bacterial genetics shapes host response during symbiosis	11
2.1 Published: Strain-specific effects on symbiosis.....	12
Abstract	
Importance	
Materials and methods	
Results and discussion	
Conclusions and future directions	
Acknowledgments	
References	
2.2 Published: The effect of SypC on the host.....	45
Abstract	
Importance	
Materials and methods	
Results and discussion	
Conclusions and future directions	
Acknowledgments	
References	
2.3 The effect of bacterial T6SS2 on host.....	76
2.3.1 T6SS2 influences host cellular homeostasis.....	77
2.3.2 T6SS2 influences host lipid remodeling and bacterial density, but not early phenotypes.....	82
2.4 The effect of bacterial <i>acuC</i> on the host	85
2.4.1 <i>AcuC</i> influences host metabolism.....	85
2.4.2 <i>AcuC</i> influences host size	92
2.4.3 <i>AcuC</i> influences bacteria inside the host	93
2.5 Summary and discussion	99
2.6 Contributions and methods for sections 2.3 and 2.4	101
Chapter III: Skin as an early detection organ	109

Abstract	
Importance	
Materials and methods	
Results and discussion	
Conclusions and future directions	
Acknowledgements	
References	
Chapter IV: Host-transcriptomic responses across scales	137
4.1 Bulk and meta-transcriptomic analysis of host responses	141
4.1.1 Light organ: APO vs WT meta RNA-seq	142
4.1.2 Light organ: WT vs Δ lux	147
4.1.3 Eye: APO vs WT vs Δ lux	150
4.1.4 Brain: APO vs WT vs Δ lux	153
4.2 Single-cell and single-nuclei optimization for <i>E. scolopes</i>	155
4.2.1 Single-cell	157
4.2.2 Single-nuclei	162
4.3 snRNA-seq analysis for <i>E. scolopes</i>	169
4.3.1 snRNA-seq analysis for light organ	171
4.3.2 snRNA-seq analysis for brain	177
4.4 Summary and future directions	185
4.5 Contributions and methods	186
Chapter V: Bacterial single-cell RNA sequencing in the squid-vibrio system	202
5.1 Development of bacterial scRNA-seq for <i>V. fischeri</i>	204
5.1.1 MATQ-seq for <i>V. fischeri</i>	205
5.1.1.1 Notes for MATQ-seq	210
5.1.2 Isolation of bacteria from host for MATQ-seq	210
5.1.2.1 Isolation method	215
5.1.2.2 Methods for flow cytometry on bacteria	217
5.2 MATQ-seq on bacteria isolated from the host	218
5.3 Published: Analysis development for bacterial scRNA-seq	222
Abstract	
Introduction	
Methods	
Results	
Discussion	
References	
5.4 Summary and future directions	240
5.5 Methods	241
Chapter VI: Discussion/Future directions	243
Appendix A: New tools for genome annotation	246
Appendix B: Role of host exosomes in symbioses	256

LIST OF ILLUSTRATIONS AND/OR TABLES

Unpublished figures only

<i>Chapter Figures</i>	<i>Page</i>
1.1 The light organ.....	3
2.1 Strains used in T6SS.....	76
2.2 Host genetics strongly influences transcriptional variation in the squid light organ in T6SS	77
2.3 Identification of host genes associated with T6SS2-dependent transcriptional responses	79
2.4 Tomograms of light organs of squid colonized with ES401 (T6SS2+) or Δ vasA (T6SS2-).	83
2.5 Host genetics strongly influences transcriptional variation in the squid light organ in acuC	85
2.6 Identification of host genes associated with acuC dependent transcriptional responses	87
2.7 Crypt regions of colonized squid.....	89
2.8 Long-term colonization with the acuC* strain is associated with reduced squid mantle length	91
2.9 Tomographs of bacteria inside crypt spaces of the squid	93
Supp. Fig 2.1 Disruption of acuC or T6SS2 does not produce major early symbiotic phenotypes in the squid light organ	94
Supp. Fig 2.2 PCA of squid LO colonized either with APO, ES114 (T6SS1+T6SS2-), ES401 (T6SS1+T6SS2+), ES114 acuC* (T6SS1+T6SS2-), or ES401 Δ vasA (T6SS1+T6SS2-).	96
Supp. Fig 2.3 Long-term colonization dynamics of ES114 and acuC* during persistent squid symbiosis.....	97
3.1 Skin transcriptomic profiles are altered in the presence of symbiont within a short time frame	118
Supp. Fig 3.1	120
3.2 Rapid repression and epithelial localization of vitellogenin-like LLTP family proteins during early symbiont exposure	123
Supp Fig 3.2 Expansion and expression dynamics of large lipid transfer protein (LLTP) genes in E. scolopes.	124
Supp Fig 3.3 HCR of LLTPs at 1 hr.	126
3..3 Quantification of LLTP-associated signal intensity within the mantle ciliary tuft cells.	128
3.4 Early transcriptional and spatial remodeling of mantle epithelial mucus-associated factors during symbiont exposure	130

4.1 Overview of host tissues examined by single-cell and single-nucleus RNA sequencing	138
4.2 Variance partitioning analysis for the sources J, S, and T light organ RNA-seq datasets	141
4.3 Principal component analysis (PCA) of combined light organ RNA-seq datasets under different modeling approaches	142
4.4 Overlap of DEG ($\text{padj} < 0.05$; $ \log_2\text{FC} \geq 1$) between three independent RNA-seq datasets (J, S, T)	143
4.5 UpSet plot showing overlap of differentially expressed genes (DEGs) identified across individual RNA-seq datasets and meta-analysis approaches	144
4.6 Host genetic background is a major source of transcriptional variation in the WT, Δlux , and APO light organ RNA-seq dataset	147
4.7 Host genetic background is a major source of transcriptional variation in the WT, Δlux , and APO eye RNA-seq dataset	149
4.8 Host genetic background is a major source of transcriptional variation in the WT, Δlux , and APO brain RNA-seq dataset	152
4.9 Dissociation of the LO into single cells	156
Sup Fig 4.1. Whole juvenile isolation and fixation of cells with ACME	157
4.10 Live/dead staining of cells dissociated in 1% pronase in CMF-ASW	158
4.11 CMF-ASW vs PBS+Mannitaol	159
4.12 Nuclei from brain	161
4.13 Nuclei isolation of light organ	161
4.14 LO nuclei isolation after improved isolation method	163
4.15 Flow cytometry gating strategy for isolation of single nuclei from the squid light organ	164
4.16 Final experimental design for single-nucleus RNA-seq analysis	166
4.17 Preprocessing pipeline	169
4.18 BioAnalyzer electropherogram of amplified cDNA generated from 10x Genomics single-nucleus RNA-seq libraries	170
4.19 Quality-control metrics for the pilot light organ single-nucleus RNA-seq dataset	171
4.20 Heatmap of marker gene expression across metacells identified in the pilot light organ snRNA-seq dataset	172
4.21 Cephalopod brain	177
4.22 Cell clusters	178
4.23 DEGs identified between APO and WT colonized animals for each Leiden cluster	181
4.24 DEGs identified between Δlux and WT colonized animals for each Leiden cluster	182

5.2 Lysis optimization for <i>V. fischeri</i>	205
5.3 Overall sequencing results of pre-DASH samples.....	206
5.4 Overall sequencing results of post-DASH sample.....	207
5.5 Number of coding genes detected after DASH.....	207
5.6 GFP loses signal due to RNALater	210
5.7 Live bacterial cells (from host tissue) were sorted directly onto LBS agar plates arranged to mimic the layout of a 96-well plate	211
5.8 Effect of counterstains on <i>Vibrio fischeri</i> fluorescence and viability.....	212
5.9 Growth curves and RNA content measurements for <i>Vibrio fischeri</i>	217
5.10 Quality control of RNA extraction and library preparation from host-isolated <i>V. fischeri</i> samples	219
B.1 mRNA in crypt spaces	256
B.2 Exosomes in crypt spaces	258
B.3 Evidence for extracellular vesicle-associated RNA sorting and candidate RNA cargo in <i>Euprymna scolopes</i>	259

NOMENCLATURE

APO. Aposymbiotic, animals not exposed to their symbiont, but to other species of the bacterioplankton.

DE. Differentially expressed.

DEG. Differentially expressed genes.

HCR. Hybridization chain reaction *in situ*. Imaging method to localize RNA.

LO. Light organ.

PCA. Principle component analysis.

WT. Wild-type strain (of *V. fischeri* (strain ES114) unless otherwise noted).

scRNA-seq. Single-cell RNA sequencing.

SYM squid. Symbiotic squid (colonized with *V. fischeri* (strain ES114) unless otherwise noted).

Chapter 1

INTRODUCTION

Animals are constantly exposed to microbial communities that can profoundly influence host health and development. Historically, symbiotic organisms have been categorized into three broad categories: pathogens, which are harmful on the host; mutualists, which are beneficial to the host; and commensals, which are neutral to the host. However, these classifications are highly context dependent. The outcome of a host-microbe interactions can vary depending on the host species, tissue, environmental condition, and even the specific microbial strain involved. As a result, many symbionts possess characteristics associated with multiple categories. For example, microbes who fall under “mutualists” have “pathogenic” traits, such as suppressing the host immune system¹. Such shared features highlight that symbiosis has overlapping molecular toolkits, raising important questions about how hosts distinguish between different types of microbes.

The structure and complexity of these microbial communities vary widely across organisms. Some animals, like humans, have a diverse microbiome. This diversity is thought, in part, to be supported by the adaptive immune system, which enables flexible and targeted responses to a broad range of microbes while still maintaining tolerance^{2,3}. In contrast, other animals, including many insects, have highly selective and simplified microbiomes. In these systems, the innate immune system plays a dominate role in controlling microbial composition, often enforcing strict specificity³. Despite the differences in microbiome complexity, all hosts must distinguish between wanted and unwanted microbes. This ability to establish and maintain this specificity is critical for stable symbioses and host health, yet the mechanism underlying these processes remain incompletely understood.

To maintain stable symbiosis, the host exerts tight spatial and physiological control. In many tissues, such as the skin or gastrointestinal tract, microbes are restricted to epithelial surfaces and prevented from entering deeper tissues. In contrast, some symbiosis involves the

controlled internalization of microbes, as seen in corals, where symbionts reside within specialized host cell. In many cases, hosts evolve dedicated structures or organs that create a regulated environment for the microbial partners, highlighting the evolutionary importance of symbiosis. On such example is the highly specialized symbiosis between the Hawaiian bobtail squid, *Euprymna scolopes*, and the bioluminescent bacterium *Vibrio fischeri*. Despite encountering a diverse microbial environment, newly hatched squid selectively recruit *V. fischeri* from seawater and establish a stable association within a specialized light organ. This experimentally tractable partnership has become a powerful model for understanding how hosts recognize, select, and maintain beneficial microbial symbionts.

In this dissertation, I use the squid-vibrio symbiosis to investigate host responses across multiple biological scales, from the earliest interactions between environmental bacteria and host epithelial surfaces to transcriptional responses within individual host cells. By combining bulk and single-cell approaches, this work seeks to provide a deeper understanding of how beneficial symbioses are established, maintained, and integrated into host biology.

1.1 The squid-vibrio system as a model

The Hawaiian bobtail squid, *Euprymna scolopes*, forms a highly specific symbiosis with the bioluminescent bacterium *Vibrio fischeri*. The bacteria only reside in the light organ (LO). It does not provide nutrients to host but rather provides the host with protection via counterillumination. In turn the squid provide the bacteria with nutrients and a niche where it can survive. This naturally simplified association, in which a single bacterial species colonizes the host, provides a powerful model for studying host-microbe interactions without the complexity of a diverse microbiome⁴. As a result, it enables detailed investigation of the molecular and biochemical conversation that occurs between host and microbe. Additionally, because cephalopods lack an adaptive immune system, this system provides the opportunity to examine how innate immune mechanisms, in all animals, contribute to symbiotic specificity. Together with their ease of keeping them in the lab, makes the squid-vibrio a good model for studying host-microbe interactions.

In this horizontally acquired symbiosis, the squid hatch asymbiotically. Within minutes of exposure to environmental microbes, the host releases mucus and increases ciliary beating of the light organ to attract bacterium to the light organ⁵. Host-derived biochemicals within the mucus help inhibit many non-symbiotic bacteria while promoting the enrichment of *V. fischeri* near the light organ⁴. Within three hours, *V. fischeri* cells enter the pores, three on each side of the organ, and colonize three crypts where they will persist (Figure 1)⁶. Despite the fact that *V. fischeri* makes up a ~~mere~~ less than 0.01% of the surrounding bacteria, the system achieves remarkable specificity⁴. The animal now is closed off to new bacteria from entering. The animal begins its profound day-night cycle. During the night, the crypt spaces are filled with the bacteria. With the dawn light cue, 70-95% of the bacteria are expelled⁴. Thus, many datasets explore these first 24 hours. At 24 hours the squid has accepted its symbiont, allowed it to grow filling the crypt spaces, and has undergone one day-night cycle.

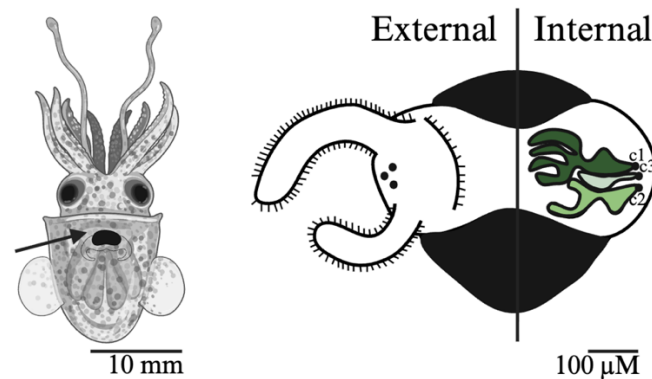


Figure 1 The light organ. The light organ site inside the mantle of *E. scolopes* (arrow). Exposed to the water are ciliated appendages and pores. *V. fischeri* enters the light organ via the pores and colonized three internal crypts on either side of the organ: crypt 1 (c1) which is the most mature, crypt 2 (c2), and crypt 3 (c3) which is the least mature/developed.

1.2 Dynamic communication

From the first interaction with the light organ the symbiont traverses a series of microenvironments, that enforce symbiont specificity. When the symbiont is present in the water, the exterior ciliated appendages rapidly beat⁵, promoting the entrainment of the symbiont to the region above the pores, which connect to the internal part of the light organ. A chemical gradient of chitin dimer or chitobiose is used to move the symbiont to the pores and through the migration pathway⁷. After entering the light organ via the three pores in the light organ, the bacteria enter the duct, then the antechamber, and then the bottleneck before reaching the crypt space. Successful colonization requires the symbiont to survive and navigate each of these environments. Only 1-3 *V. fischeri* cells enter the crypt spaces. Once this happens, the bottleneck narrows, restricting the bacteria on either side⁴. Throughout this process the host changes gene expression with contact with the migrating symbiont^{8,9}. Within the crypts, the host provides nutrients that support bacterial growth; however, the three-dimensional structure of the crypts likely generates gradients in oxygen, nutrients, and bacterial density, creating additional heterogeneity within the symbiotic habitat. Together, these observations suggest that establishment of the squid–vibrio symbiosis requires continuous adaptation by both partners. The bacterium must successfully navigate and survive multiple host-generated microenvironments, while the host undergoes physiological and transcriptional changes in response to bacterial colonization.

These changing conditions do not end once the symbionts reach the crypt spaces. Even within the crypt spaces, bacteria occupy distinct microenvironments that vary in developmental state and host architecture. The three crypt spaces of the juvenile light organ differ substantially from one another. Crypt 1 is the most mature and is the first to develop during embryogenesis, whereas Crypt 3 is the least mature and develops last¹⁰. Crypts 1 and 2 undergo daily venting, while Crypt 3 initially does not because of its immature state⁶. Correspondingly, bacterial populations within Crypt 3 display a less organized association with the epithelial surface and are less densely packed than those found in Crypts 1 and 2⁶. These observations suggest that each crypt provides a distinct local environment for the

symbiont. As a result, differences in epithelial physiology, nutrient availability, host signaling, or immune activity are likely present among the crypts, potentially shaping bacterial behavior and host responses at a fine spatial scale.

The existence of these distinct microenvironments raises an important question: how are bacterial populations regulated within the light organ, and how does the host maintain specificity throughout colonization? Despite occupying different regions of the organ and experiencing different local conditions, *V. fischeri* populations remain tightly controlled by the host. This regulation requires continuous communication between host and symbiont, allowing the host to detect the presence of microbes, distinguish beneficial partners from potential competitors or pathogens, and modulate physiological responses accordingly. Understanding the molecular mechanisms underlying this communication is therefore central to understanding how symbiosis is established and maintained.

1.3 Specificity and immune perception of symbiosis

Hosts and symbionts use biochemical language to establish and maintain association. This communication often begins with host pattern-recognition receptors (PRRs), which detect conserved microbe-associated molecular patterns (MAMPs) and signal the presence of microbes. PRRs are highly diverse across animal lineages and include receptors such as Toll-like receptors (TLRs), peptidoglycan recognition proteins (PGRPs), and NOD-like receptors. When these PRRs are triggered, the innate immune system is activated to initiate inflammatory responses or other pathways to promote the symbiosis¹¹.

In *Hydra* spp., TLR signaling has been shown to promote bacterial colonization and contribute to microbiome composition¹¹. Similarly, PRR signaling pathways help maintain the squid-vibrio symbiosis, where host recognition of bacterial molecules contributes to developmental and physiological changes within the light organ. In addition, the bacteria have the ability to turn down the immune system, preventing inflammation pathways¹². In vertebrates such as zebrafish and mammals, innate immune signaling also contributes to microbial tolerance, allowing hosts to maintain beneficial microbial communities while

limiting harmful colonization¹¹. Together, these studies highlight that microbial sensing is central not only to immune defense, but also to symbiotic specificity and the selective maintenance of beneficial microbial partners.

While molecular recognition is essential for establishing specificity, the consequences of these signaling events extend far beyond individual receptors and pathways. Recognition of microbial molecules can alter host gene expression, tissue development, cellular behavior, and organismal physiology. Likewise, bacterial responses to host-derived signals can influence colonization dynamics, metabolism, and population structure. These interconnected responses occur across multiple levels of biological organization, making symbiosis fundamentally a multi-scale process.

1.4 Multi-scale nature of host-microbe interactions

Host–microbe interactions occur across multiple biological scales, ranging from organism-level physiology to tissue-specific responses, cellular heterogeneity, and microbial population dynamics. As a result, understanding symbiosis requires approaches that integrate information across these different levels of organization. Changes observed at one biological scale may emerge from highly localized or cell-type-specific processes that are not apparent when examining the organism as a whole.

Historically, many studies of symbiosis have relied on bulk tissue-level approaches, including transcriptomics, imaging, and physiological measurements. These methods have provided important insights into how hosts respond to microbial colonization and how symbionts influence host development. In the squid–vibrio system, bulk transcriptomic studies have identified large-scale changes in host gene expression associated with colonization, including pathways related to immunity, tissue remodeling, metabolism, and development¹². However, bulk approaches average signals across many cell types and tissues, potentially obscuring spatially restricted or cell-specific responses that contribute to symbiotic specificity.

Recent advances in single-cell and single-nuclei sequencing technologies have begun to reveal the cellular complexity underlying host–microbe interactions. These approaches have demonstrated that microbial exposure can produce highly heterogeneous responses across individual cell populations, even within the same tissue¹³. Similarly, microbial populations themselves are often heterogeneous, with distinct bacterial subpopulations exhibiting different physiological states and metabolic activities¹⁴. Understanding symbiosis therefore requires consideration of both host and microbial heterogeneity.

Importantly, biological processes occurring at one scale can propagate across others. Early microbial sensing at epithelial surfaces may trigger organism-wide physiological responses, while strain-specific bacterial traits can influence tissue development, cellular signaling, and microbial community structure. Together, these observations highlight that symbiosis is not governed by a single interaction, but instead emerges from dynamic processes operating across multiple interconnected biological scales.

1.5 Overarching goals of dissertation thesis

In this dissertation, I investigate how animal hosts detect, respond to, and maintain associations with beneficial microbial symbionts. Using the squid-vibrio symbiosis as a model system, I examine how host–microbe interactions are initiated, how hosts distinguish symbiotic bacteria from other environmental microbes, and how these interactions propagate across multiple biological scales. By integrating organismal, tissue-level, cellular, and microbial perspectives, this work seeks to provide a more comprehensive understanding of how symbiosis is established and maintained.

In Chapter 2, I examine how host transcriptional and physiological responses vary depending on the strain of *V. fischeri*. I demonstrate that host gene expression and aspects of host development differ with exposure to different bacterial strains, indicating that host responses are sensitive to strain-level variation. To better understand the basis of these differences, I also investigate variation in bacterial functional potential across strains. Together, these

findings demonstrate that symbiotic outcomes are shaped by both host responses and bacterial genetic diversity.

In Chapter 3, I identify the skin as an early site of symbiont detection that responds prior to colonization of the light organ. I demonstrate that host responses to symbiotic bacteria are initiated in non-symbiotic tissues, revealing that early microbial sensing is not restricted to the nascent organ itself. Importantly, these responses are specific to symbiotic bacteria, suggesting that the host is not merely responding to microbial presence, but instead detecting symbiont-associated signals. These findings indicate that peripheral tissues may play an active role in initiating symbiosis and represent the earliest stage of host–microbe communication.

In Chapter 4, I extend these studies to the cellular level by developing and applying single-cell and single-nuclei transcriptomic approaches for *E. scolopes*. Using these methods, I investigate how individual host cell populations respond to microbial colonization and explore the cellular heterogeneity underlying tissue-level responses observed in previous studies. This work provides a framework for understanding symbiosis at single-cell resolution in a cephalopod model system.

In Chapter 5, I further expand the investigation of cellular heterogeneity by developing approaches for bacterial single-cell transcriptomics. This chapter focuses on methodological advances and computational frameworks designed to characterize bacterial transcriptional states within host-associated environments. These approaches provide a foundation for examining how individual bacterial cells respond to host-derived conditions and contribute to symbiotic function.

Together, these studies establish a multi-scale framework for understanding how beneficial symbioses emerge from the interplay between host recognition mechanisms and symbiont-specific bacterial traits. By linking early microbial detection, strain-level variation, host cellular responses, and bacterial heterogeneity, this dissertation provides new insights into

the mechanisms that govern host–microbe specificity and the establishment of stable symbiotic associations.

REFERENCES

1. Wiesmann, C. L., Wang, N. R., Zhang, Y., Liu, Z. & Haney, C. H. Origins of symbiosis: shared mechanisms underlying microbial pathogenesis, commensalism and mutualism of plants and animals. *FEMS Microbiol. Rev.* **47**, fuac048 (2023).
2. Lee, Y. K. & Mazmanian, S. K. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **330**, 1768–1773 (2010).
3. Woodhams, D. C. *et al.* Host-associated microbiomes are predicted by immune system complexity and climate. *Genome Biol.* **21**, 23 (2020).
4. Nyholm, S. V. & McFall-Ngai, M. J. A lasting symbiosis: how the Hawaiian bobtail squid finds and keeps its bioluminescent bacterial partner. *Nat. Rev. Microbiol.* **19**, 666–679 (2021).
5. Gundlach, K. A. *et al.* Ciliated epithelia are key elements in the recruitment of bacterial partners in the squid-vibrio symbiosis. *Front. Cell Dev. Biol.* **10**, 974213 (2022).
6. Essock-Burns, T., Bongrand, C., Goldman, W. E., Ruby, E. G. & McFall-Ngai, M. J. Interactions of Symbiotic Partners Drive the Development of a Complex Biogeography in the Squid-Vibrio Symbiosis. *mBio* **11**, e00853-20 (2020).
7. Visick, K. L., Stabb, E. V. & Ruby, E. G. A lasting symbiosis: how *Vibrio fischeri* finds a squid partner and persists within its natural host. *Nat. Rev. Microbiol.* **19**, 654–665 (2021).
8. Kremer, N. *et al.* Initial Symbiont Contact Orchestrates Host-Organ-wide Transcriptional Changes that Prime Tissue Colonization. *Cell Host Microbe* **14**, 183–194 (2013).
9. Nikolakakis, K., Lehnert, E., McFall-Ngai, M. J. & Ruby, E. G. Use of Hybridization Chain Reaction-Fluorescent *In Situ* Hybridization To Track Gene Expression by Both Partners during Initiation of Symbiosis. *Appl. Environ. Microbiol.* **81**, 4728–4735 (2015).
10. Montgomery, M. K. & McFall-Ngai, M. Embryonic Development of the Light Organ of the Sepiolid Squid *Euprymna scolopes* Berry. *Biol. Bull.* **184**, 296–308 (1993).
11. Chu, H. & Mazmanian, S. K. Innate immune recognition of the microbiota promotes host-microbial symbiosis. *Nat. Immunol.* **14**, 668–675 (2013).
12. Moriano-Gutierrez, S. *et al.* Critical symbiont signals drive both local and systemic changes in diel and developmental host gene expression. *Proc. Natl. Acad. Sci.* **116**, 7990–7999 (2019).
13. Sharma, P. V. & Thaiss, C. A. Host-Microbiome Interactions in the Era of Single-Cell Biology. *Front. Cell. Infect. Microbiol.* **10**, 569070 (2020).
14. Homberger, C., Hayward, R. J., Barquist, L. & Vogel, J. Improved Bacterial Single-Cell RNA-Seq through Automated MATQ-Seq and Cas9-Based Removal of rRNA Reads. *mBio* **14**, e03557-22 (2023).

Chapter 2

. BACTERIAL GENETICS SHAPES HOST RESPONSE DURING SYMBIOSIS

The symbiosis between the Hawaiian bobtail squid, *Euprymna scolopes*, and its bacterial partner *Vibrio fischeri* has long served as a powerful model for understanding how animals establish and maintain symbiosis with microbes. This interaction is highly specific, as only *V. fischeri* can successfully colonize the squid light organ¹. However, not all strains of *V. fischeri* are able to colonize, while some are more competitive colonizers^{2,3}. This suggests that the relationship extends beyond simple species-level recognition; rather, bacterial strain background and physiological function can strongly influence symbiotic interactions.

This dissertation examines host-microbe symbiosis across multiple biological scales, beginning here with variation at the bacterial strain level. By integrating comparative colonization experiments, host transcriptomics, and bacterial pangenome analyses, this chapter investigates how microbial genetic diversity influences both bacterial behavior and host responses during symbiosis. These studies establish strain-level variation as an important driver of symbiotic outcomes and provide a conceptual foundation for subsequent chapters examining host-microbe interactions across tissues, developmental time points, and individual cells.

In this chapter, I explore how distinct bacterial strains, and specific genetic factors modulate the squid-vibrio symbiosis. Section 2.1 characterizes strain-specific effects on colonization and host transcriptional responses, followed by investigations bacterial gene *sypC* (Section 2.2), the Type VI secretion system (T6SS2) (Section 2.3), and the metabolic regulator *acuC* (Section 2.4). Together, these studies highlight that symbiosis is shaped by coordinated interactions between both host and microbe, requiring specific bacterial traits and corresponding host responses for successful

establishment and maintenance. Furthermore, these interactions can be altered by bacterial strain background and genetic variation, demonstrating that symbiosis is not a fixed biological state, but a dynamic process influenced by microbial functional diversity.

2.1 Strain specific effects of symbiosis

This section is modified from “Strain matters: host responses reflect origin in the squid-vibrio symbiosis.” Published in mSystems, November 2025.

Vera Beilinson¹, Grischa Y. Chen^{1,2}, Alexis C. Hargadon^{1,2}, Edward G. Ruby¹, Margaret J. McFall-Ngai^{1,2}

1. Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, USA
2. Division of Biosphere Sciences and Engineering, Carnegie Institution for Science, Pasadena, California, USA

V. Beilinson participated in the conception of the project, gathered and analyzed data, prepared the data, confocal imaging and analysis, and participated in the writing of the manuscript.

Supplemental figures (S1-S4) and tables (S1-S4) can be found online at: <https://doi.org/10.1128/msystems.00498-25>

Dissertation Note

This study demonstrates that bacterial strain identity is a major determinant of symbiotic outcome. Distinct *Vibrio fischeri* strains elicited different host transcriptional responses, with native *E. scolopes* isolates producing more similar host expression profiles and stronger suppression of immune-associated pathways than non-native strains. These findings suggest that long-term host-symbiont

coevolution shapes both bacterial colonization strategies and host responses, establishing strain-level variation as an important source of functional diversity within an otherwise species-specific symbiosis.

Abstract

Understanding the cause and consequences of bacterial strain variation remains a challenge in the study of symbioses. While the diverse reactions of the host immune system to strain variants have been well studied in pathogenesis, much less is known about how strain variation influences beneficial associations. From the complex vertebrate gut microbiome to the more tractable invertebrate models of symbiosis, the host's cellular and molecular responses to this diversity remain largely a mystery. Here, we explore strain diversity in *Vibrio fischeri*, the bioluminescent bacterial symbiont of the Hawaiian bobtail squid, *Euprymna scolopes*. Phylogenetic analyses of the genomes of 62 *V. fischeri* strains, including 50 light organ-associated and 12 planktonic isolates, revealed several genes that were absent in planktonic strains, but uniformly present in symbiotic ones. To better understand the consequences of this diversity to the host, we selected five light-organ associated strains: three from *E. scolopes* but having different combinations of colonization factors, one from a congeneric squid host, and one from a marine fish. We colonized juvenile *E. scolopes* with these strains and, using RNAseq, found that (i) the most similar host transcriptomic responses occurred among the native *E. scolopes* strains, (ii) intermediate was the strain from the related squid, and (iii) least similar was the fish strain. Importantly, native strains downregulated immune-related genes more than non-native ones. Finally, host development was atypical or delayed when colonized

by non-native strains. These experiments point the way to more targeted studies of the mechanisms underlying host responses to symbiont strain diversity.

Importance

Variation among strains of a bacterial species is a powerful factor underlying the intensity of host responses during pathogenic infections. Less is known about the cellular and molecular responses of host tissues to differences between the strains present in an animal's normal microbiome. We use a natural, species-specific, symbiosis to explore the influence of strain-level differences on host gene expression and morphogenesis. Analysis of symbiotic strains from squids and fishes, as well as free-living strains, shows that the carriage of colonization determinants, while critical to competitive success among strains of a species, has a minimal effect on the transcriptional response of the host. We provide evidence that a more important driver of normal gene expression during the development of symbiosis is the history of a strain's co-diversification with its host species. Such studies, using simple invertebrate models, allow the recognition of otherwise obscured interactions underlying the more complex microbiomes of vertebrates.

Introduction

Bacteria exhibit remarkable genome plasticity, enabling the diversification of species into genetically distinct strains that collectively can occupy a wider range of environments and ecological niches. This natural strain-level variation often has profound functional consequences. In host–microbe symbioses of plants and animals, strain-level differences in the bacterial symbiont can shape the success of microbial colonization and persistence, as well as the nature of the host responses to its bacterial partners. For example, differing outcomes of bacterial pathogenesis associated with strain variation has long been known (1), and in some diseases, an individual can carry multiple strains of a pathogen, where such diversity can modulate disease intensity and progression (2). Recent studies have demonstrated that in a mutually

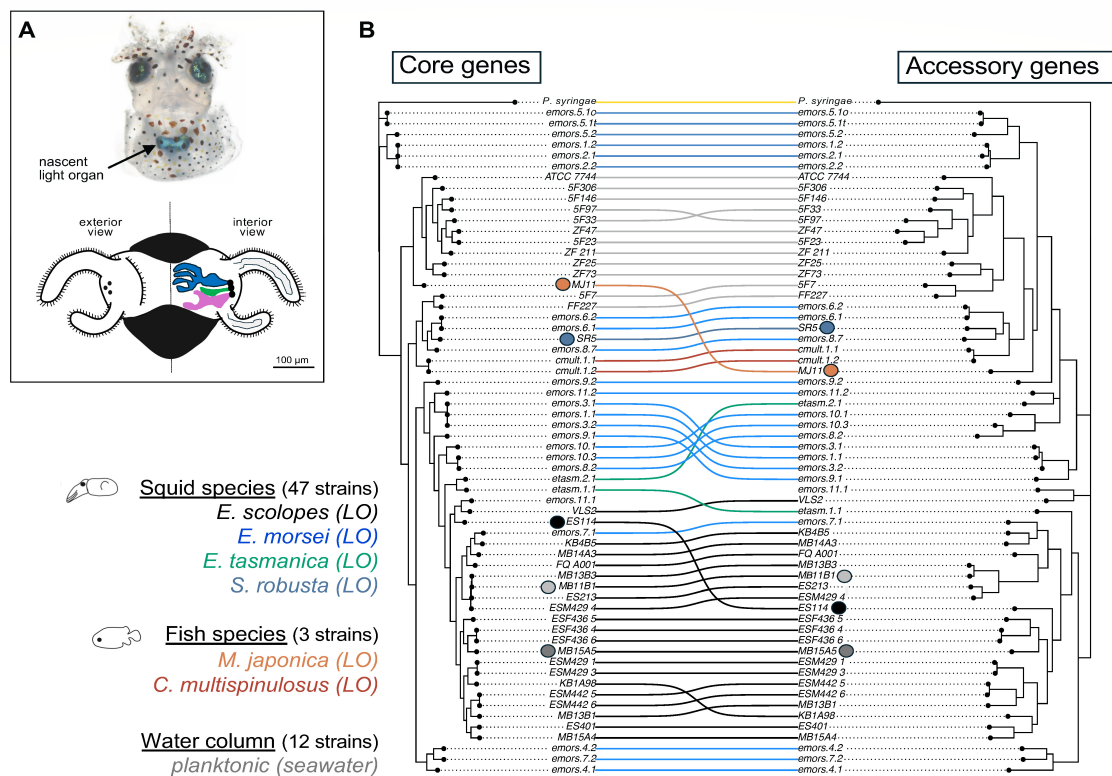
beneficial context, natural strain variation appears to be a key feature of the vertebrate microbiome (for review, see references 3, 4). However, because of the bacterial complexity of many microbiomes and the inability to culture many of their component species, defining strain-level characteristics has been challenging although great strides in analysis of this complexity are being made (5).

Several studies using invertebrate models, which typically have a microbiome with fewer species, have demonstrated the importance of strain variation, but few have examined the molecular and cellular impact on the host (see, e.g., 6–9). Naturally occurring monospecific bacterial symbioses offer the opportunity to experimentally manipulate strain variation and study its impact with high resolution. The best studied of such binary associations are the legume–rhizobium symbioses (10), where specificity is dictated by genetic factors from both partners. For example, the bacterial *nodE* gene influences host range by modifying nodulation (Nod)-factor structure, while the host, *Medicago truncatula*, produces nodule-specific cysteine-rich (NCR) peptides that act selectively against incompatible strains (11). Beyond molecular compatibility, functional outcomes also reflect co-evolutionary history. Studies have shown that native rhizobial strains often have a competitive advantage when colonizing the roots of their cognate host, outperforming non-native strains when both are present (12). These findings illustrate a key principle: host-microbe compatibility and potential co-evolution can develop both at the species level and the strain level.

To explore these phenomena within an animal-microbe symbiosis, we have used the species-specific light-organ association between the Hawaiian bobtail squid *Euprymna scolopes* and its bioluminescent bacterial symbiont *Vibrio fischeri* (13, 14). Males of this host have only the *V. fischeri* light-organ (LO) symbiosis throughout their lifespan, whereas females have two sites of symbiosis: the LO, which is present throughout life in both sexes, as well as a consortial accessory nidamental gland (ANG) (Fig. S1), which develops only in mature females

(13) and is typical of the females of many squid species. Like the plant-rhizobium association, the binary system of the juvenile *E. scolopes* LO provides a tractable model to study how colonization by different symbiont strains influences the host response.

Juvenile squid hatch into seawater, where *V. fischeri* strains make up <0.1% of the microbial community (15, 16). Nonetheless, it is the only bacterial species capable of entering and colonizing the host's nascent LO, reflecting the host's remarkable selectivity for its symbiont. While *V. fischeri* defines the species-level symbiont identity, the LO of the Hawaiian host can be colonized by a number of different *V. fischeri* strains found in the bacterioplankton (17). However, at the level of the individual host, the number of colonizing strains is affected by the biogeography of the organ, i.e., each juvenile host organ has six independent interior crypts and typically entertains between one and six *V. fischeri* strains (Fig. 1A). This observation and other studies have provided evidence that each crypt is usually colonized by a single bacterial strain (16, 18).



Protein description	Light-organ associated (%)	Bacterioplankton associated (%)	Protein description	Squid associated (%)	Fish associated (%)
Iron-citrate transporter (FecB)	81	0	FadB-like enoyl-CoA hydratase	100	0
Catechol receptor (Fiu)	81	0	Carnitine metabolism (CaiD)	100	0
Peptide transporter (SalY)	46	0	Glycerol-3-phosphate transporter (UgpBE)	0	67
CRISPR helicase (Cas3)	2	38	Capsular colonic acid synthesis (ManC)	0	67
CRISPR endoribonuclease (Cas6f)	2	38	Mg ²⁺ efflux exchanger (CorC)	0	67
CRISPR integrase (Cas1)	6	44	Ribosome acetylation (RimL)	4	100

Figure 2 Patterns of gene content in *Vibrio fischeri* isolates diverge across niches and hosts. (A) Cartoon of *E. scolopes* nascent light-emitting organ (LO). Left side: external view showing ciliated appendages and three surface pores leading to the internal crypts in which symbionts will be housed. Right side: cut-away view of LO interior showing three of the organ's six crypts. (B) Co-phylogenetic analysis. Left tree: Core genome comparison (nucleotide level) based on orthologous genes. Right tree: Accessory genome

clustering based on presence/absence of gene content. Analysis includes 50 light organ-associated strains (from squids or fishes) and 12 planktonic (i.e., seawater) isolates. Trees are rooted on the outgroup species *Pseudomonas syringae*. Branch lengths are log-transformed for visualization. The branch length of *P. syringae* was shortened to fit within the figure. (C) Functional enrichment analysis (percent of strains encoding these symbiosis-associated proteins). Left: Proteins enriched in *V. fischeri* LO isolates (50 strains) compared to *V. fischeri* isolates not associated with a LO (12 strains). Right: Proteins enriched in strains associated with squid LOs (47 strains) compared to fish LOs (3 strains).

Some strains of *V. fischeri* are not as effective at colonizing the Hawaiian squid host as native strains from *E. scolopes* LOs. For example, MJ11, a *V. fischeri* strain isolated from the LO of the Japanese pinecone fish, *Monocentris japonica*, is unable to colonize *E. scolopes* under typical environmental conditions of low *V. fischeri* abundance (19). MJ11 colonizes poorly because it lacks the gene encoding the Syp-biofilm-promoting regulator, RscS (Fig. S2), which induces the production of an exopolysaccharide that facilitates the colonization of juvenile squid (20, 21). Introducing *rscS* into the MJ11 genome conferred colonization ability under typical conditions, illustrating how specific genetic elements shape host compatibility. These examples highlight how different strains can compete for colonization (14), but how do strain differences affect host responses?

The carriage of several colonization factors differs among native *E. scolopes*-associated strains (14, 22, 23) (Fig. S2). For example, dominant “D” strains are faster colonizers and more competitive than sharing “S” strains, at times excluding them from the crypt spaces (24). Competition within the host crypts, where the symbionts permanently reside, is also reflected in features such as the presence or absence of a type VI secretion system (T6SS) that mediates inter-strain antagonism (25, 26).

However, little is known about how carriage of these colonization factors influences host responses.

While much is known about the host's reactions to colonization by the best-studied *V. fischeri* strain, ES114 (13), the transcriptional and developmental responses to other symbiont strains of different hosts have remained poorly examined. In this study, we compare the effects of colonization by several *V. fischeri* strains, isolated from the LOs of several species of squids and fish, on relative LO gene expression (as based on transcript levels) and developmental phenotypes in juvenile *E. scolopes*. We find that the host's transcriptional signature varies with the evolutionary and environmental history of the symbiont strain. Specifically, gene-expression responses are most similar when newly hatched juvenile squid are exposed to *E. scolopes* LO isolates, intermediate for strains derived from another species of squid, and most divergent for a fish LO isolate. In addition, several differences in the trajectory of juvenile light-organ development were observed. These findings reveal that the host can discriminate among conspecific strains, suggesting that host-symbiont compatibility and communication can evolve through fine-scale, strain-level co-adaptation. The results add to growing evidence for the importance of considering bacterial diversity below the species level in studies of beneficial symbiosis (27, 28).

Materials and methods

Phylogenetic trees, the pangenome, and functional enrichment

Publicly available *V. fischeri* strains with clearly defined isolation sources were downloaded from NCBI in May 2025. RefSeq (GCF) genome assemblies were used. Pangenome analysis was performed using Anvi'o (v8), and functional enrichment

between pairwise comparisons was assessed using anvi-compute-functional-enrichment-in-pan.

Orthologs were identified using OrthoFinder (v2.5.5). A binary matrix of gene presence/absence was generated using default parameters, and a nucleotide-based phylogenetic tree was constructed using the -n flag. Multiple-sequence alignment was performed with MAFFT, and gene trees were inferred using IQ-TREE. Co-phylogeny trees were visualized in R using ggtree (v3.12.0) and ape (v5.8-1).

Sources of *V. fischeri* strains and construction of an *rscS* expression vector

A total of 62 strains of *V. fischeri* were isolated from a number of light organs and seawater samples (Table S1). Particular emphasis was placed on five strains of symbionts (Fig. 1B): ES114, MB11B1, and MB15A5, which were isolated from the light organs of adult *E. scolopes* collected in nearshore waters of Oahu, Hawaii, as previously described (29, 30). Strain SR5 was obtained from the light organ of an adult specimen of the related sepiolid squid *Sepiola robusta*, collected off the French Mediterranean coast (31). Strain MJ11 was isolated from the light organ of an adult Japanese pinecone fish, *Monocentris japonica* (20). Construction of the *rscS* expression plasmid, pLMS33, and its empty control vector, pKV69, and their conjugation into MJ11 have been previously described (20, 32).

Animal colonization by *V. fischeri* strains

Juvenile squid were exposed to bacterial strains within 4 h of hatching in offshore seawater obtained from Scripps Institution of Oceanography, La Jolla, CA. The inoculum was 50,000 CFU/mL, and exposure was for 3 h, unless otherwise noted. This unnaturally high concentration of *V. fischeri* cells (19) was used to limit the possibility that bacterial availability was limiting the inoculation process. Following exposure, animals were rinsed twice with offshore seawater and transferred to individual vials. At 24 h post-inoculation, luminescence was measured, and

individuals were classified as colonized when the relative light unit (RLU) level had exceeded the detection limit of the luminometer (>100 RLU). The animals were transferred every 24 h to fresh offshore seawater as described previously (33). Only colonized animals were maintained for analysis at 72 h.

Squid were considered part of the same clutch if their eggs were laid by the same parental pair at the same time and the hatchlings emerged synchronously. Clutch identity was tracked in downstream analyses to account for genetic background effects.

Determination of symbiont colonization level

At designated time points, individual squid were transferred to 1.5 mL Eppendorf tubes with a minimal volume seawater and anesthetized on ice. Samples were then frozen at -80°C to reduce surface contamination from external bacteria (34). After thawing, squid were homogenized and plated to determine CFU/mL per individual as described previously (33).

Calculation of light-organ appendage length

At 72 h post-inoculation, squid were preserved on ice in 4% paraformaldehyde (PFA) in marine PBS (mPBS) consisting of 50 mM Na_2PO_4 , 0.45 M NaCl, pH 7.4 (35). Samples were incubated overnight at 4°C on a rotating platform and then rinsed four times for 15 min with 4 mL mPBS. Squid tissue was stained with the nuclear stain DAPI (Thermo Fisher Scientific, Irwindale, CA) diluted 1:50 in mPBS for 24 h, followed by two rinses in mPBS to remove excess stain. Samples were mounted on glass depression slides with coverslips. Imaging was performed on an upright Zeiss (Dublin, CA) LSM 900 microscope. Arm length was estimated in FIJI (ImageJ) using

the measurement tool, from a point at the center of the three pores to the tip of each arm.

Measurement of bottleneck closure

Squid were colonized overnight (18 h) with a single *V. fischeri* strain at an inoculum of 10,000 to 50,000 CFU/mL. Colonization was confirmed using a luminometer; animals with luminescence below ~100 relative light units (RLU) were excluded from analysis to ensure consistent colonization. Squid were anesthetized on ice to prevent premature venting of their symbionts and fixed at 4°C in 4% PFA in mPBS for 24 h with rotation. Samples were rinsed four times for 15 min in mPBS. Light organs were dissected and incubated for 24 h in 400 µL mPBS containing 1% Triton X-100, DAPI (1:50), and rhodamine phalloidin (Biotium, Fremont, CA) diluted 1:50 in mPBS. Samples were rinsed two times in mPBS and mounted on glass depression slides with coverslips. Imaging was performed on a Zeiss LSM 900 confocal microscope. DAPI was used to visualize host nuclei, confirming colonization, and rhodamine phalloidin was used to stain filamentous actin to easily visualize and determine bottleneck diameter, as previously described (36). On both sides of the organ, crypt 1 (Fig. 1A), the most mature crypt with the least variation in structure (37), was imaged, and bottleneck diameters were measured using FIJI.

Transcriptome assembly, annotation, and differential gene-expression analysis

A previous study had shown that *V. fischeri* colonization by strains isolated from the *E. scolopes* LO significantly changed host gene expression by 24 h post inoculation (hpi) (38). In this study, based on transcript levels, we characterized symbiosis-induced changes in gene expression at 72 hpi to minimize any effects of delayed colonization or morphological development by non-native strains. The CFU and luminescence differences among strains were negligible. At 72 hpi, squid were preserved in RNAlater (Thermo Fisher) overnight at 4°C and stored at 80°C until RNA extraction was performed. Light-organ total RNA was extracted using the

Qiagen RNeasy Mini Kit with QIAshredder columns (Qiagen, Valencia, CA) per manufacturer's instructions. Samples were treated with TURBO DNase (Thermo Fisher) to remove DNA contamination. RNA integrity was assessed using RNA 6000 Pico Kit for Bioanalyzer (Agilent Technologies, Santa Clara, CA). mRNA was separated from total RNA using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, Ipswich, MA), and RNA-seq libraries were constructed using the NEBNext Ultra II RNA Library Prep Kit for Illumina following the manufacturer's protocols. The libraries were sequenced as paired-end reads (2×100 bp) on an Illumina NextSeq2000 at the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech.

Sequencing reads were quality-trimmed using fastp (v0.24.0) with default parameters. Cleaned reads were aligned to the *E. scolopes* reference genome (v2.2) (39) using STAR (v2.7.10b) in two-pass mode with a read overhang of 99. The resulting coordinate-sorted BAM file was used for genome-guided transcriptome assembly with Trinity (v2.1.1), using the `--genome_guided_max_intron 500000` flag.

Assembled transcripts were clustered using CD-HIT-EST (v4.8.1) at a 95% similarity threshold (`-c 0.95`) and then mapped to the genome using GMAP (gmapl, v2024-08-14) with output format `gff3_match_cdna`. The resulting GFF3 file was converted to BED format using AGAT (v1.4.1), and overlapping features were identified with bedtools intersect (v2.31.0). Novel transcripts were retained and merged with published annotations using AGAT. As is the case for essentially all RNAseq analyses, the annotations reported here are inferred from sequence homologies and are not supported by biochemical confirmation of their functionality. With that understanding, the descriptor “putative” has not been appended in every reference to a gene's annotation.

Reads were quantified using featureCounts (v2.0.1) with the updated genome annotation using flags `-g gene_id -t exon`. Differential expression analysis was performed in R (v4.4.0) using DESeq2 (v1.44.0), with the design formula `~clutch +`

condition to account for clutch-specific variation. Differential expression between conditions was assessed using contrasts in DESeq2 with Benjamini-Hochberg correction for multiple testing and log2 fold-change shrinkage.

Result and discussion

Phylogenetic comparisons of free-living and host-associated *V. fischeri* strains reveal a strong clustering by ecological origin

Whole-genome analyses of 62 strains of *V. fischeri* isolated from a variety of sources (Table S1) indicated a strong clustering according to whether they were obtained from seawater (i.e., planktonic) or from the tissues of one of six animal host species (Fig. 1B). Strains obtained from the LOs of *Euprymna* spp. formed distinct clades in the core-genome-based phylogeny (Fig. 1B, left). In general, fish strains, and non-*Euprymna* squid strains clustered separately, suggesting the presence of both host-specific and tissue-specific lineages. This pattern demonstrated that the phylogeny of the *V. fischeri* strains reflects the source of the isolate and supports a possible effect of host-symbiont co-adaptation; however, an influence of geographic distribution is likely also present. Interestingly, analysis using the accessory (or “adaptive”) genome (Fig. 1B, right) revealed a similar pattern, but with higher variation (40), possibly capturing additional strain-level adaptations involving genes that are enriched in isolates from a particular niche. Such a genomic comparison further showed that, while these 62 *V. fischeri* strains share a conserved core genome, each strain also encodes a considerable number of singleton genes, i.e., ones that are found in only one strain. An example of such a comparison of singleton gene carriage was depicted for 5 *V. fischeri* strains (Fig. S3), showing between 150–400 singletons among those LO isolates. The apparently limited distribution of these singleton genes suggested

they have been acquired relatively recently and may be an important class of genetic elements in the diversification of symbiotic *V. fischeri* (17, 41).

Another pattern arising from a genome-level comparison between all host-associated *V. fischeri* strains (Fig. 1C, left) was the high frequency of the carriage of the genes *fecB*, encoding an iron-citrate transporter, and *fiu*, encoding an Fe³⁺-catechol receptor. This pattern was true of strains isolated from both fish and squid LOs. These two Fe-related genes were missing from the genomes of non-LO-associated planktonic strains (Table S1). One interpretation of these results is that strains isolated from host animal tissues encode pathways that support survival in a low-iron environment. Previous research has provided evidence that symbiotic *V. fischeri* is adapted to growth under limiting iron-availability (42, 43), and analysis of the bacterium's response to colonizing the LO has indicated that, as is true for many animal tissues, the LO is an iron-restricted environment (43–45). However, it should be noted that in some host-associated bacteria, an iron-citrate transporter is used in the opposite direction to efflux excess Fe³⁺ out of the bacterial cell. Reducing the level of free-iron limits the toxic effects of certain host-produced antibiotics and nitric oxide resulting from the Fenton reaction (46). The usefulness of encoding a similar efflux capacity (e.g., *salY*) is also reported for antimicrobial peptides (47, 48). Conditions such as nitric oxide and antimicrobials have been reported in the squid light organ and other host tissues (13, 49), suggesting that the *FecB* and *SalY* efflux transporters of *V. fischeri* have a role in the association (Fig. 1C left).

In contrast to LO symbionts, strains isolated as free-living planktonic cells in seawater had a higher likelihood of encoding genes annotated as anti-viral CRISPR-Cas systems compared to host-associated ones (Fig. 1C, left) (Table S2). This difference may reflect the observation that host animals are able to protect their symbiont populations from phage infection while they are in tissues like the LO (50) as opposed to when *V. fischeri* cells are free-living in seawater (51); thus, CRISPR-Cas may be less of a fitness factor for strains that do most of their replication within

a symbiotic organ (23). This finding was consistent with the absence of CRISPR-Cas genes in 99% of the sequenced genomes of symbiotic rhizobia, as well as a comprehensive analysis of the patterns of other bacteria lacking CRISPR-Cas (52). Together, these studies, as well as those presented here, have found that symbionts whose growth occurs mainly in their host seem to be more likely to lack CRISPR-Cas. More symbiotic systems should be studied to determine whether loss of CRISPR genes is a general trend among symbiotic species. It is possible that other anti-phage systems (53) may compensate for the loss of CRISPR-Cas in symbiotic *Vibrio* populations. Another consideration is that the loss of CRISPR genes within symbionts may reflect differences in physiological adaptations in the host since CRISPR systems may also function in gene regulation or DNA repair in other bacteria (54, 55).

We also found that the genomes of strains associated with different host species were more likely to encode proteins conferring distinct physiological capabilities (Table S3). For example, as previously reported (22), only isolates from sepiolid squid shared an FadB-like enoyl-CoA hydratase used in lipid degradation and catabolism. Similarly, squid LO-derived *V. fischeri* strains were found to carry the gene annotated as carnitine racemase (*caiD*), used to synthesize osmoprotectants like glycine betaine (56, 57) that could be valuable in conditions of squid tissues, which are typically isosmotic or hyperosmotic to seawater, i.e., at or above 1,000 milliosmoles/liter (mOsm/L) (58) (Fig. 1C, right); however, producing such additional osmolytes would be detrimental to strains growing in the hypoosmotic tissues of marine fishes, which are typically 300–400 mOsm/L (59, 60). Not surprisingly, it has also been reported that the bioluminescence of luminous symbionts isolated from squids requires an osmolarity that is high relative to the optimal conditions of fish LO symbionts (61, 62). Finally, similar to animal pathogens, LO-derived strains of fishes (Fig. 1C, right) more often encoded genes whose annotations are associated with the production of cell-surface modifications (e.g., *manC* and *spsG*) (63, 64) or that lower Mg²⁺ concentration (*corC*) and modify

ribosomal components (rimL) as a strategy to resist protein-synthesis inhibiting antimicrobials (65, 66). Such immunity-evading mechanisms may provide a higher fitness value for symbionts of a vertebrate host with adaptive immunity. Also notable was a putative high-affinity glycerol-3-phosphate transporter locus (ugpBE) that was encoded in fish isolates, but not squid isolates (Fig. 1C). It has been hypothesized that in the squid LO the host provides this metabolite as a nutrient (67) during its daily period of rapid growth (33); however, glycerol-3-phosphate may be less available in fishes, thereby requiring its symbionts to have a higher affinity transporter.

It should be noted that associating certain genetic capabilities with symbiont success in fish LOs (Fig. 1C, right) must be tempered by the low number of isolates from fishes that were analyzed in this study (Table S1). Nonetheless, taken together, the genomic and phylogenetic results presented here provide evidence that *V. fischeri* strains have evolved in response to their ecological context, free-living or host-associated, and that host identity may further drive strain diversification.

Colonization by different *V. fischeri* strains leads to different transcriptional responses in the host

In the squid-vibrio field, most research focusing on the host transcriptomic and developmental responses has been done with strains isolated from *E. scolopes*, principally strain ES114 (13, 14). To understand how strain variation influences LO gene expression and development in squid, we colonized newly hatched juvenile *E. scolopes* with strains isolated from LOs across a phylogenetic gradient: (i) ES114, MB15A5, MB11B1 from *E. scolopes*; (ii) SR5 from the Mediterranean sepiolid squid *Sepiolo robusta*; and (iii) MJ11 from the fish *M. japonica* (Table S1). Four replicate sets of host colonization were obtained for each strain (Fig. 2C). During a quality control check of the samples, it was determined that one of the four ES114 replicates was an outlier (Fig. S4A). PCA plots of the change in expression of the top

3,000 most variable host genes supported this conclusion (Fig. S4B) so that replicate was eliminated from further analyses.

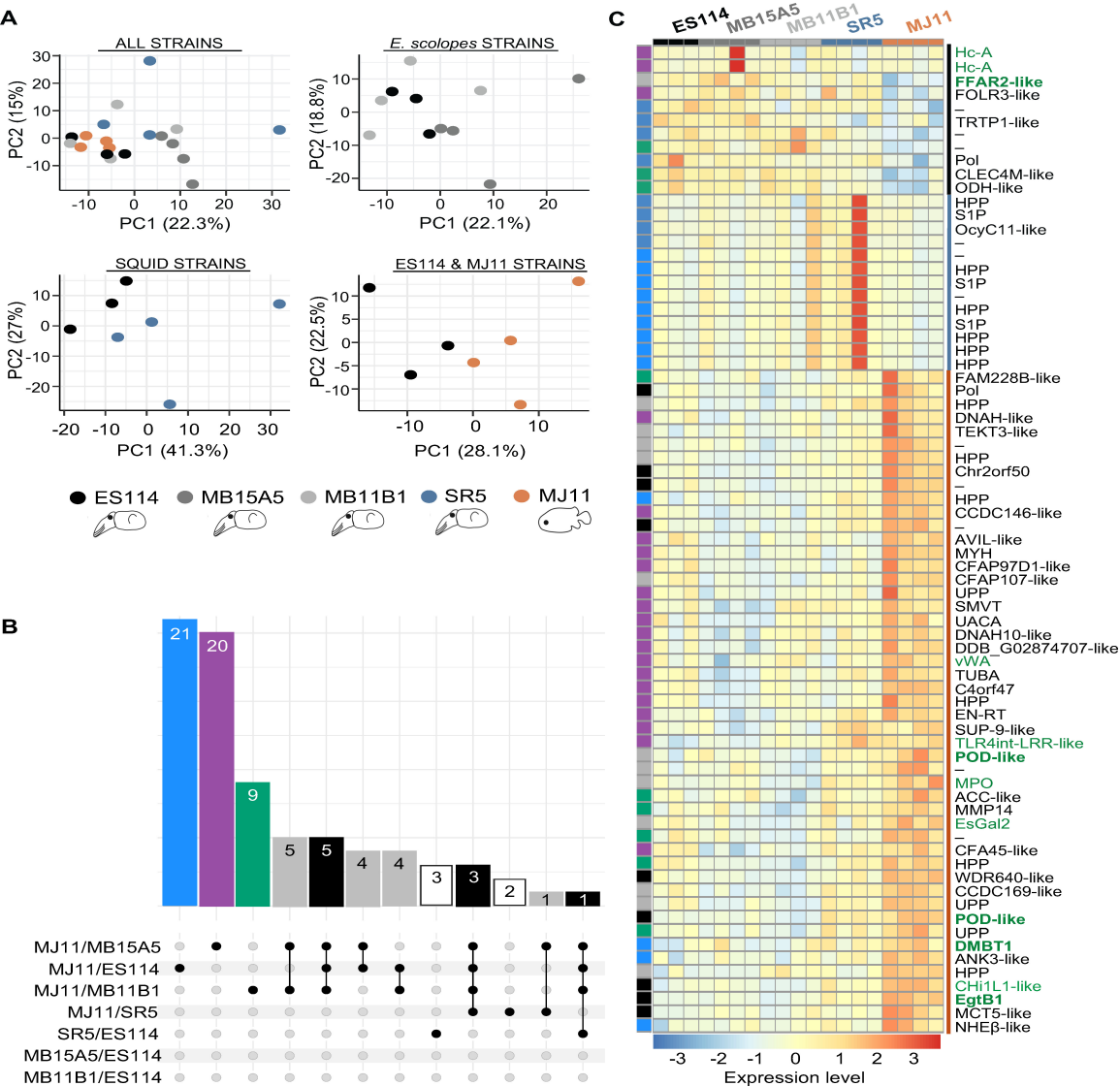


Figure 3 Colonization by non-native *V. fischeri* strains differentially impacts host gene expression at 72 hpi. (A) Principal component analysis (PCA) plots representing expression of the top 3,000 most variable host genes 72 h after colonization by one of 5 strains (color-coded). The top-left PCA plot compares all strains to show global variation. The additional three plots highlight differences between specific strain combinations. (B) UpSet plot, capturing both shared and strain-specific transcriptional responses, and showing overlap in differentially expressed genes (DEGs) of the host from pairwise comparisons between strains. There were no DEGs

between ES114 and the other two *E. scolopes*-derived strains (bottom two rows). DEGs were defined as those with adjusted P-value (padj) < 0.05 and absolute log2-fold change ($|\log_2FC|$) ≥ 0.58 , with log2FC values shrunk using the ashR method. In the MJ11 vs. ES114 comparison, 18 genes had $|\log_2FC| \geq 1$ and 10 had $|\log_2FC| \geq 2$. The 0.58 threshold was selected to capture biologically meaningful changes influenced by bacterial colonization and bioluminescence. Color legend: Blue bar: DEGs unique to MJ11 vs. ES114; Purple bar: DEGs unique to MJ11 vs. MB15A5; Green bar: DEGs unique to MJ11 vs. MB11B1; Gray bar: DEGs in MJ11 vs. one or two other strains (but not all four); Black bar: DEGs shared in MJ11 vs. all *E. scolopes* strains (used in heatmap); White bar: DEGs not included in the heatmap. Color-group memberships of individual DEGs are identified on left column of heat map. (C) Heatmap of selected DEGs. Gene annotations are based on BLAST nucleotide hits; annotated gene IDs abbreviated (Table S4); (–) =no BLAST match found; (HPP) = hypothetical/predicted protein; (UPP) = uncharacterized/predicted protein. Functional groupings of GO-term classes are indicated by a colored line on the far right; Black = Transposase and oxidoreductase activity; Teal = Membrane associated proteolysis; Orange = Cytoskeleton, transport and stress response. Gene names with immune-related annotations are in dark green. The five genes examined in Fig. 3 are in bold dark green. Expression values are variance-stabilizing transformed (VST) and row-scaled for visualization. Rows (genes) are clustered by a tree, which is not represented.

By 72 hpi, well after colonization had been established and all strains had triggered morphological changes in the LO (13), the overall host RNA profiles remained broadly comparable among strains (Fig. 2A, upper left). However, certain host genes exhibited differences in strain-dependent expression. In subgroup comparisons, the gene expression patterns of the three isolates from *E. scolopes* LOs were similar, forming an overlapping group (Fig. 2A, upper right). In contrast, pairwise comparisons between *E. scolopes* strain ES114 and the non-*E. scolopes* LO strains showed distinct groupings (Fig. 2A, lower).

We then performed pairwise comparisons between strains with the most distinctive patterns and displayed them as an UpSet plot (Fig. 2B). A major conclusion from this analysis was that the transcriptional response of the host to colonization by any of the 3 *E. scolopes*-derived *V. fischeri* strains was the same (i.e., no DEGs in the bottom

two comparisons). The transcription data were also used to develop a heat map displaying differences in transcript level of the 73 most highly regulated genes (Fig. 2C; Table S4). For example, colonization of *E. scolopes* by the fish symbiont MJ11 led to a markedly higher transcript level of a gene predicted to encode glycoprotein DMBT1 than native squid symbionts (Fig. 3A). This protein has a variety of roles in innate immunity, such as the recognition of MAMPs and epithelial-cell differentiation. Several studies have indicated that DMBT1 is a key player in maintaining homeostasis of innate immunity along mucosal surfaces, where it often occurs in high abundance (for review, see reference 68). In a recent contribution, the DMBT1 gene was shown to be regulated in bacteria-associated epithelial tumorigenesis in the mammalian colon (69).

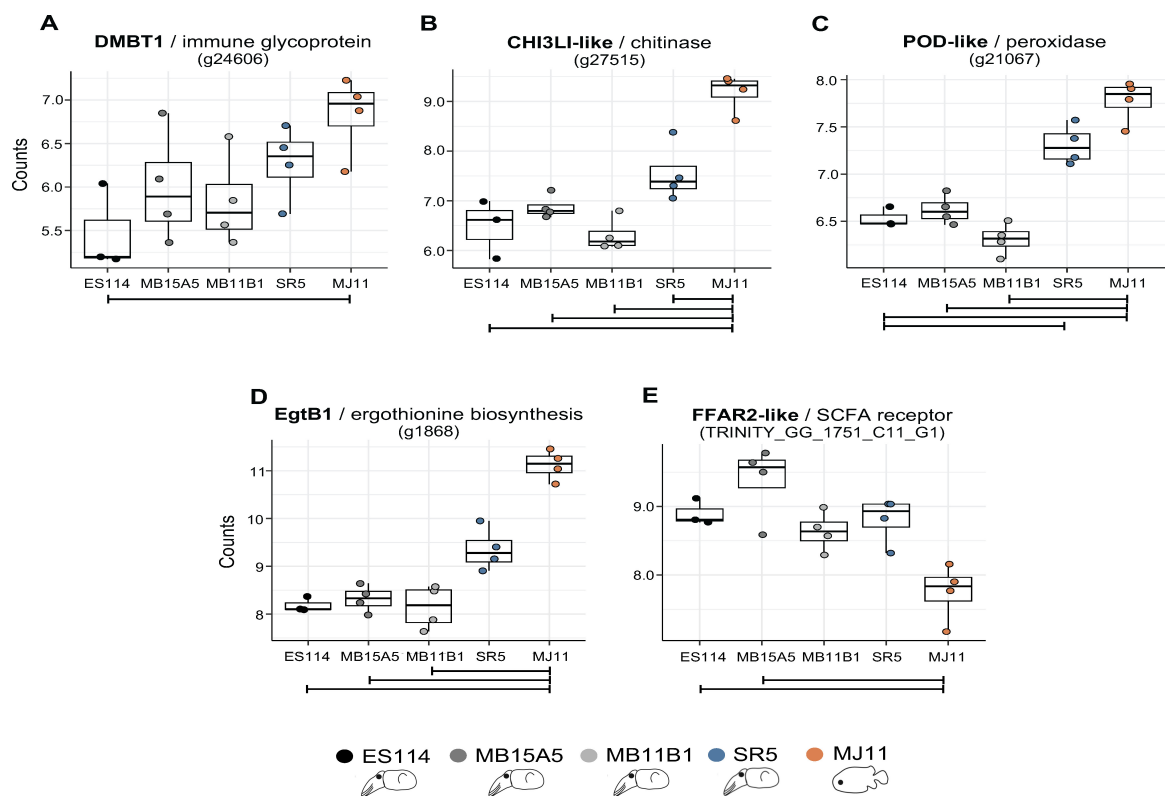


Figure 4 Strain colonization impacts host genes associated with symbiotic interactions. Patterns of expression at 72 hpi of 5 host genes (A–E) of interest from Fig. 2C (names in bold). Counts are VST transformed. Expression levels that were significantly different between pairs of strains are indicated by the line

below them; $\text{padj} < 0.05$, $|\log_2\text{FC}| \geq 0.58$. For simplicity, SR5 was only compared with ES114 or MJ11. Data are also present in heatmap, Fig. 2C.

The same pattern of differential regulation was observed with most of the other immune-related proteins (Fig. 2C). For example, a putative antimicrobial chitinase also had higher expression in the non-*E. scolopes* isolates (Fig. 3B). Similarly, the genes encoding two proteins associated with the production of antioxidants were differentially regulated among the strains: POD-like, a peroxidase, and EgtB1-like (annotated as “ergothioneine biosynthesis protein 1 isoform 1”). The presence of the gene encoding EgtB1 was unexpected; ergothioneine, a sulfur-containing amino-acid derivative of histidine, is a powerful antioxidant believed to be produced only by certain fungi, actinomycetota, and cyanobacteria (70). When BLASTed against the genomes of the phylum Mollusca, the EgtB1-like mRNA sequence returned a best hit to a sequence encoding a domain predicted to have oxidoreductase activity in the oyster *Crassostrea gigas*. Thus, these data suggest that the EgtB1-like antioxidant protein that, together with the antioxidant POD-like, had significantly lower levels of their encoding transcripts in LOs colonized by strains isolated from *E. scolopes* than by other host-associated strains (Fig. 3C and D). Taken together, these patterns were consistent with previous reports of the native symbiont transcriptionally “calming” the immune response of *E. scolopes* (71, 72), a strategy that non-native *V. fischeri* strains, adapted to other host species, may be unable to deploy.

Finally, some putative immune-related host genes that were differentially regulated when colonized by *E. scolopes*-derived strains, responded less to SR5 or MJ11 (Fig. 2C). For example, the gene encoding FFAR2 (free fatty acid receptor 2) was more highly induced by the native squid strains than by MJ11, the fish LO symbiont (Fig. 3E). In mammals, FFAR2 is a 7-membrane-spanning G-protein-coupled receptor with the highest expression in the gut epithelium, where it senses microbially produced short-chain fatty acids (SCFAs) and acts in concert with other immune system responses to the normal microbiota (73, 74). In *E. scolopes*, the LO symbionts

are believed to produce SCFAs that are used by the host (67). It is intriguing to note that the expression pattern seen for the putative FFAR2 is suggestive of a host receptor that is induced only when the native symbiont and its resulting metabolic activity are detected.

Across all comparisons of gene expression, colonization by SR5 usually resulted in an intermediate transcriptional profile, falling between the fish LO symbiont and the three strains isolated from *E. scolopes* LOs (Fig. 3). This trend reinforces the correlation between the nature of host responses to the strain and the relative phylogenetic distance of the hosts from which the symbionts were isolated.

Colonization by non-native strains alters or delays host phenotypes

As mentioned above, in the squid-vibrio field, similar to the study of host transcriptomic responses, most analyses of the host phenotypes have been performed using strains isolated from *E. scolopes*, principally strain ES114 (13, 14). When a host is colonized by strains isolated from squid hosts, both sides of the bilaterally symmetric organ were typically colonized by 24 h and, as such, the symbiont-driven developmental program behaved similarly on either side, as previously reported (13) (Fig. 4A). We found that the fish strain MJ11 initiated symbiosis more slowly and to a lesser extent (Fig. 4B) than any squid-derived strain (23). Specifically, we observed that MJ11 colonized only one side of the LO in 25% of the juvenile hosts, even when exposed to an unnaturally high inoculum, leaving the other half uncolonized for at least 72 hpi (Fig. 4C). This differential colonization was reflected in the morphological development of the organ. While the colonized sides of the organs showed the signature regression of the ciliated superficial epithelium, the sides that were uncolonized had a phenotype similar to the light-organ surfaces of newly hatched (Fig. 1A) or 72 h aposymbiotic animals (75). The appendages of juveniles

that were bilaterally colonized by MJ11 or by one of the other strains all regressed to the same degree (Fig. 4C).

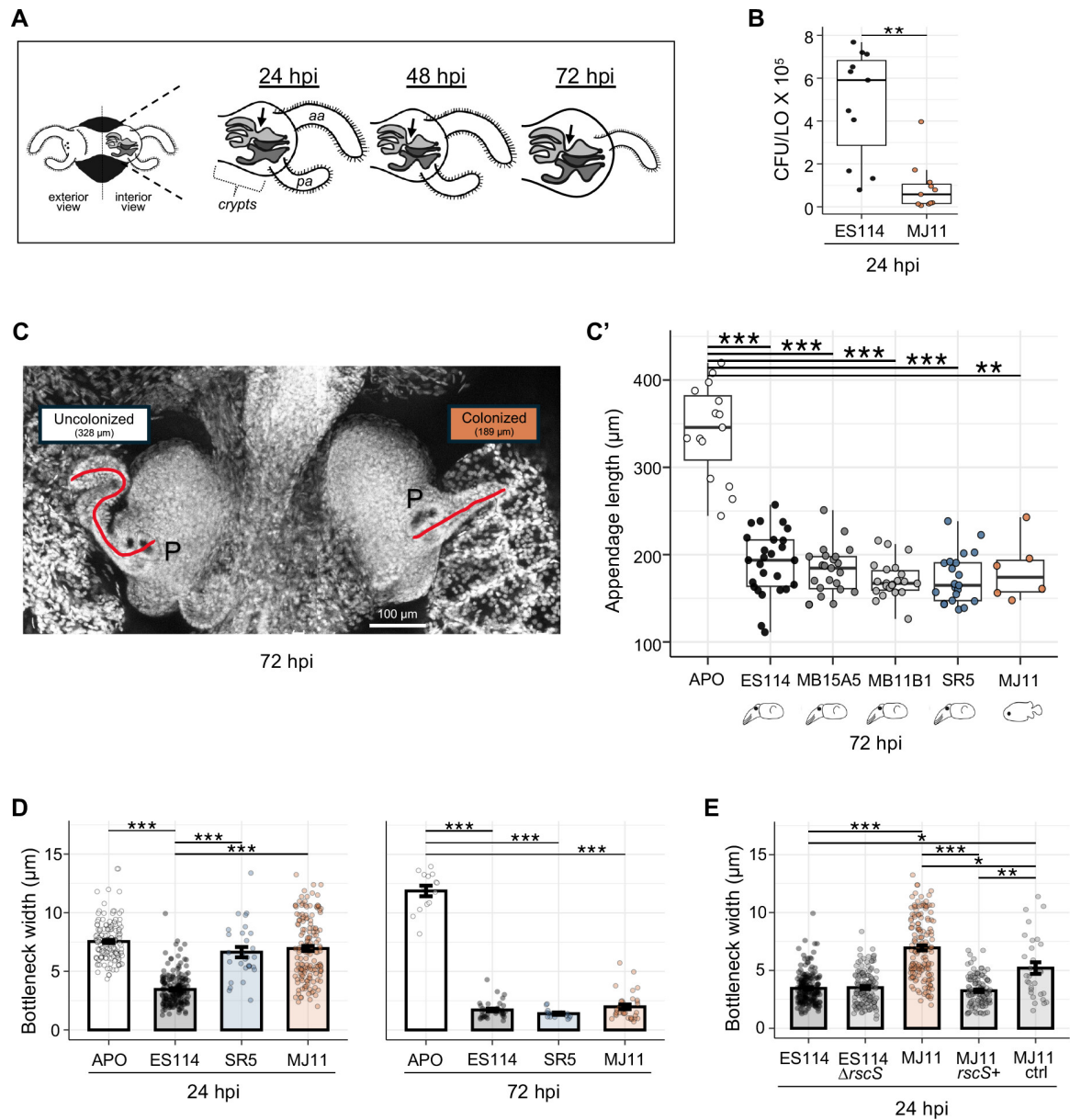


Figure 5 Colonization by non-native strains delays host development. (A) Cartoon illustrating some of the developmental events typically occurring in the juvenile LO during the first 72 hpi. The leftmost image shows the entire organ of a hatchling, while the other three depict an interior view of only one side at three times post inoculation. During the first 3 days after colonization, the

ciliated appendages begin their trajectory toward full regression, and the bottleneck (arrows) connecting the migration path to the crypts reduces in diameter. aa, anterior appendage; pa, posterior appendage. (B) Colony-forming units per light organ (CFU/LO) recovered from individual squids at 24 hpi. Statistical comparisons were performed using a Kruskal–Wallis test followed by Dunn’s post hoc tests with Bonferroni correction. Asterisks indicate significance: ** $P < 0.01$. (C) Image of the LO of a juvenile squid at (72 hpi) colonized on only one side with strain MJ11. This condition was noted in about 25% of MJ11-colonized light organs but was never observed in the several hundred animals colonized by the other four strains. (C’) Average length of the ciliated appendages of the LO at 72 hpi with one of five strains. Measurements were made only on animals in which both sides of the organ were colonized. Data shown are for the length of the left anterior appendage. Statistical comparisons were performed using a Kruskal–Wallis test followed by Dunn’s post hoc tests with Bonferroni correction. Asterisks indicate significance: *** $P < 0.001$. ** $P < 0.01$, * $P < 0.05$. APO (aposymbiotic, not colonized). (D) LO bottleneck width at 24 hpi (left) and 72 hpi (right) of colonized LOs compared to uncolonized LOs (APO). Statistical comparisons were performed at each time point, using a Kruskal–Wallis test followed by Dunn’s post hoc tests with Bonferroni correction. Asterisks indicate significance: *** $P < 0.001$. (E) Bottleneck closure response to the carriage of the *rscS* gene. At 24 hpi, bottleneck closure induced by ES114 and ES114 Δ rscS was not statistically different; however, MJ11 closed less than both of them. MJ11 rscS+ has *rscS* present on a plasmid. MJ11 ctrl carries only the backbone plasmid. Statistical comparisons were performed using a Kruskal–Wallis test followed by Dunn’s post hoc tests with Bonferroni correction. Asterisks indicate significance: *** $P < 0.001$. ** $P < 0.01$, * $P < 0.05$. Only some statistical significances are shown for plot clarity.

We also examined the impact of strain variation on another developmental phenotype (76), the symbiont-induced narrowing of a “bottleneck” that restricts the diameter of the migration path just before the LO crypt entrance (Fig. 4A). This constriction serves to spatially restrict the symbiont population to the crypts, opening only during the crypt venting at dawn (36, 37). By 24-hpi, colonization by *E. scolopes*-derived strains had triggered a significant bottleneck narrowing; however, the non-native strains SR5 and MJ11 had not (Fig. 4D, left). Because bottleneck constriction begins when symbiont numbers are as low as 10 cells (36), the reduced population sizes of

these strains at 24 hpi (Fig. 4B) are not responsible for the lack of bottleneck closure. Instead, intrinsic differences between the behavior or biochemistry of native and non-native strains are likely accountable. By 72 hpi, the bottleneck was fully narrowed by all strains (Fig. 4D, right).

To better understand the basis of the strain-level difference in the timing of bottleneck narrowing, we explored how the carriage of *rscS*, which encodes a regulator of biofilm formation in *V. fischeri* (21), may be related to the narrowing of the host's LO bottleneck (Fig. 4E). The wild-type MJ11 strain, which does not carry *rscS* and colonizes poorly at normal inoculum (20), is delayed in bottleneck constriction at 24 hpi (Fig. 4D, left); however, when strain MJ11 was genetically modified to carry the *rscS* allele, colonization resulted in a normal bottleneck narrowing by 24 hpi. Interestingly, the wild-type squid symbiont ES114 typically carries *rscS*, but colonization by a deletion mutant of *rscS* still causes a narrowing of the bottleneck. This result suggested that the carriage of *rscS* is sufficient for bottleneck closure for some strains but is not necessary for others. By 72 hpi, the bottlenecks of all LOs colonized by wild-type and genetically modified strains were closed and not significantly different from strain ES114.

The bottleneck phenotype is complex, apparently requiring multiple inputs from the symbiont to activate. Our previous work showed that bottleneck closure requires a bioluminescence-independent, quorum sensing-dependent pathway (77). It is also known through multiple studies that quorum sensing and RscS/biofilm formation are transcriptionally and post-transcriptionally co-regulated (78–81). It is possible that in ES114, overlapping signals from both the quorum sensing and biofilm pathways are detected by the host to quickly trigger bottleneck closure. In the absence of any one of these signals, closure may lag or not complete. MJ11 constitutively expressing

RscS appears to be primed to express the necessary signals upon colonization of the crypts, whereas wild-type MJ11 takes longer to turn on these signals.

Taken together, the developmental phenotypes described above show that, over time, *E. scolopes* ultimately remodels its LO in response to *V. fischeri* colonization, regardless of the strain's origin. However, the length of time required for the developmental trajectory to result in normal maturation of the LO varies among *V. fischeri* strains. Similar differences in development have also been found to occur with strain variation in legume-rhizobia associations (10, 11, 82), demonstrating how a comparative strain-level approach can provide clues to symbiotic mechanisms.

Conclusions and future directions

This study represents a broad view of how symbiont strain-level variation influences the establishment and development of a symbiotic association and serves as a discovery tool for further analysis of this and other symbioses. Specifically, we have shown how host responses to different strains that typically colonize the LO of one host species, *E. scolopes*, differ from the responses of that host to strains from the LOs of other squid and fish species. We anticipate that the patterns in host responses to strain-level differences described here will help generate hypotheses that can be tested in this and other host-microbe symbioses.

The results revealed an array of possible areas in which to more deeply pursue the mechanisms underlying the phenomena discovered. One conspicuous area is the basis for or consequences of changes in gene expression. Specifically, this study did not reveal whether the differences in transcript levels detected during colonization by different strains resulted from changes in the stability of the transcript, or whether the differences were reflected in the levels of transcript-encoded protein abundance and/or activity. Similarly, it remains to be determined whether any of the hypotheses

proposed above for underlying functions of the differences noted between strains isolated from squid or fish tissue are valid ([Fig. 1C](#)).

Another area of possible exploration would be a comparison with transcriptional responses by other sepiolids species that have symbiotic associations with *V. fischeri* strains. Such an approach would provide insight into whether the responses were common among, or specific to, hosts beyond *E. scolopes*. For example, several *Euprymna* spp. occur in Indo-Pacific and Australian waters, including *E. berryi*, which is being developed as another model sepiolid species ([83](#)). In addition, several *Sepiola* spp. with symbiotic organs containing *V. fischeri* occur in the Northeastern Atlantic Ocean and Mediterranean. Such an approach may also allow a refined analysis of the relationship between bacteria genomic differences, which are extensive, and host responses.

Acknowledgements

The authors thank David Romero (Universidad Nacional Autónoma de México) for his insights concerning the encoding of CRISPR-Cas in symbiotic rhizobia genomes. We also acknowledge the McFall-Ngai and Ruby lab members for providing ongoing advice and support throughout this project, in particular, Viridiana Avila Magana.

This work was supported by the National Institute of Allergy and Infectious Diseases (R37 AI050661 to Margaret J. McFall-Ngai and Edward G. Ruby), the National Institute of General Medical Sciences (R01 GM135254 to Edward G. Ruby and Margaret J. McFall-Ngai), and the Caltech 2023 Center for Evolutionary Science Award (Vera Beilinson).

V.B., Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing—original draft, Writing—review and editing. G.Y.C., Conceptualization, Data curation, Investigation, Supervision, Validation, Writing—review and editing. A.C.H., Data curation, Investigation, Methodology,

Validation, Writing—review and editing. E.G.R., Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—original draft, Writing—review and editing. M.J.M.-N., Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources,

References

1. Robertson BD, Meyer TF. 1992. Genetic variation in pathogenic bacteria. *Trends Genet* 8:422–427. [https://doi.org/10.1016/0168-9525\(92\)90325-x](https://doi.org/10.1016/0168-9525(92)90325-x)
2. Acosta N, Waddell B, Heirali A, Somayaji R, Surette MG, Workentine ML, Rabin HR, Parkins MD. 2020. Cystic fibrosis patients infected with epidemic *Pseudomonas aeruginosa* strains have unique microbial communities. *Front Cell Infect Microbiol* 10:173. <https://doi.org/10.3389/fcimb.2020.00173>
3. Van Rossum T, Ferretti P, Maistrenko OM, Bork P. 2020. Diversity within species: interpreting strains in microbiomes. *Nat Rev Microbiol* 18:491–506. <https://doi.org/10.1038/s41579-020-0368-1>
4. Yan Y, Nguyen LH, Franzosa EA, Huttenhower C. 2020. Strain-level epidemiology of microbial communities and the human microbiome. *Genome Med* 12:71. <https://doi.org/10.1186/s13073-020-00765-y>
5. Ghazi AR, Thompson KN, Bhosle A, Mei Z, Yan Y, Wang F, Wang K, Franzosa EA, Huttenhower C. 2025. Quantifying metagenomic strain associations from microbiomes with Anpan. *bioRxiv*:2025.01.06.631550. <https://doi.org/10.1101/2025.01.06.631550>
6. Jehrke L, Stewart FA, Droste A, Beller M. 2018. The impact of genome variation and diet on the metabolic phenotype and microbiome composition of *Drosophila melanogaster*. *Sci Rep* 8:6215. <https://doi.org/10.1038/s41598-018-24542-5>
7. Motta EVS, Moran NA. 2024. The honeybee microbiota and its impact on health and disease. *Nat Rev Microbiol* 22:122–137. <https://doi.org/10.1038/s41579-023-00990-3>
8. Murfin KE, Lee MM, Klassen JL, McDonald BR, Larget B, Forst S, Stock SP, Currie CR, Goodrich-Blair H. 2015. *Xenorhabdus bovienii* strain diversity impacts coevolution and symbiotic maintenance with *Steinernema* spp. nematode hosts. *mBio* 6:e00076. <https://doi.org/10.1128/mBio.00076-15>
9. Pérez-Ferrer PA, Ashraf M, Rodrigues M, Troncoso J, Nishiguchi MK. 2025. Genetic variation in the Atlantic bobtail squid-*Vibrio* symbiosis from the Galician rias. *Mol Ecol* 34:e17596. <https://doi.org/10.1111/mec.17596>
10. Riaz MR, Sosa Marquez I, Lindgren H, Levin G, Doyle R, Romero MC, Paoli JC, Drnevich J, Fields CJ, Geddes BA, Marshall-Colón A, Heath KD. 2025. Mobile gene clusters and coexpressed plant-rhizobium pathways drive partner quality variation in symbiosis. *Proc Natl Acad Sci USA* 122:e2411831122. <https://doi.org/10.1073/pnas.2411831122>

11. Yang S, Wang Q, Fedorova E, Liu J, Qin Q, Zheng Q, Price PA, Pan H, Wang D, Griffiths JS, Bisseling T, Zhu H. 2017. Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc Natl Acad Sci USA* 114:6848–6853. <https://doi.org/10.1073/pnas.1700460114>
12. Wippel K, Tao K, Niu Y, Zgadzaj R, Kiel N, Guan R, Dahms E, Zhang P, Jensen DB, Logemann E, Radutoiu S, Schulze-Lefert P, Garrido-Oter R. 2021. Host preference and invasiveness of commensal bacteria in the *Lotus* and *Arabidopsis* root microbiota. *Nat Microbiol* 6:1150–1162. <https://doi.org/10.1038/s41564-021-00941-9>
13. Nyholm SV, McFall-Ngai MJ. 2021. A lasting symbiosis: how the Hawaiian bobtail squid finds and keeps its bioluminescent bacterial partner. *Nat Rev Microbiol* 19:666–679. <https://doi.org/10.1038/s41579-021-00567-y>
14. Visick KL, Stabb EV, Ruby EG. 2021. A lasting symbiosis: how *Vibrio fischeri* finds a squid partner and persists within its natural host. *Nat Rev Microbiol* 19:654–665. <https://doi.org/10.1038/s41579-021-00557-0>
15. Lee KH, Ruby EG. 1992. Detection of the light organ symbiont, *Vibrio fischeri*, in Hawaiian seawater by using lux gene probes. *Appl Environ Microbiol* 58:942–947. <https://doi.org/10.1128/aem.58.3.942-947.1992>
16. 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. <https://doi.org/10.1038/ismej.2011.92>
17. Bongrand C, Koch E, Mende D, Romano A, Lawhorn S, McFall-Ngai M, DeLong EF, Ruby EG. 2022. Evidence of genomic diversification in a natural symbiotic population within its host. *Front Microbiol* 13:854355. <https://doi.org/10.3389/fmicb.2022.854355>
18. Sun Y, LaSota ED, Cecere AG, LaPenna KB, Larios-Valencia J, Wollenberg MS, Miyashiro T. 2016. Intraspecific competition impacts *Vibrio fischeri* strain diversity during initial colonization of the squid light organ. *Appl Environ Microbiol* 82:3082–3091. <https://doi.org/10.1128/AEM.04143-15>
19. 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. <https://doi.org/10.1128/aem.60.5.1565-1571.1994>
20. Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG. 2009. A single regulatory gene is sufficient to alter bacterial host range. *Nature* 458:215–218. <https://doi.org/10.1038/nature07660>
21. Shibata S, Yip ES, Quirke KP, Ondrey JM, Visick KL. 2012. Roles of the structural symbiosis polysaccharide (syp) genes in host colonization, biofilm formation, and polysaccharide biosynthesis in *Vibrio fischeri*. *J Bacteriol* 194:6736–6747. <https://doi.org/10.1128/JB.00707-12>
22. Bongrand C, Ruby EG. 2019. The impact of *Vibrio fischeri* strain variation on host colonization. *Curr Opin Microbiol* 50:15–19. <https://doi.org/10.1016/j.mib.2019.09.002>
23. Ruby EG, Lee KH. 1998. The *Vibrio fischeri*-*Euprymna scolopes* light organ association: current ecological paradigms. *Appl Environ Microbiol* 64:805–812. <https://doi.org/10.1128/AEM.64.3.805-812.1998>

24. Koehler S, Gaedeke R, Thompson C, Bongrand C, Visick KL, Ruby E, McFall-Ngai M. 2018. The model squid–vibrio symbiosis provides a window into the impact of strain- and species-level differences during the initial stages of symbiont engagement. *Environ Microbiol*. <https://doi.org/10.1111/1462-2920.14392> 2
25. Guckes KR, Miyashiro TI. 2023. The type-VI secretion system of the beneficial symbiont *Vibrio fischeri*. *Microbiology (Reading, Engl)* 169. <https://doi.org/10.1099/mic.0.001302>
26. Suria AM, Smith S, Speare L, Chen Y, Chien I, Clark EG, Krueger M, Warwick AM, Wilkins H, Septer AN. 2022. Prevalence and diversity of type VI secretion systems in a model beneficial symbiosis. *Front Microbiol* 13:988044. <https://doi.org/10.3389/fmicb.2022.988044>
27. Jones KR, Song Y, Rinaldi SS, Moran NA. 2025. Effects of priority on strain-level composition of the honey bee gut community. *Appl Environ Microbiol* 91:e0082825. <https://doi.org/10.1128/aem.00828-25>
28. Li Z, Xia J, Wang J. 2025. Unveiling strain-level dynamics in the human skin microbiome. *Cell Host Microbe* 33:615–617. <https://doi.org/10.1016/j.chom.2025.03.014>
29. Boettcher KJ, Ruby EG. 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. *J Bacteriol* 172:3701–3706. <https://doi.org/10.1128/jb.172.7.3701-3706.1990>
30. 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. <https://doi.org/10.1128/AEM.01792-08>
31. Fidopiastis PM, von Boletzky S, Ruby EG. 1998. A new niche for *Vibrio logei*, the predominant light organ symbiont of squids in the genus *Sepiola*. *J Bacteriol* 180:59–64. <https://doi.org/10.1128/JB.180.1.59-64.1998>
32. Visick KL, Skoufos LM. 2001. Two-component sensor required for normal symbiotic colonization of *Euprymna scolopes* by *Vibrio fischeri*. *J Bacteriol* 183:835–842. <https://doi.org/10.1128/JB.183.3.835-842.2001>
33. Yang L, Lawhorn S, Bongrand C, Kosmopoulos JC, Kuwabara J, VanNieuwenhze M, Mandel MJ, McFall-Ngai M, Ruby E. 2024. Bacterial growth dynamics in a rhythmic symbiosis. *Mol Biol Cell* 35:ar79. <https://doi.org/10.1091/mbc.E24-01-0044>
34. Naughton LM, Mandel MJ. 2012. Colonization of *Euprymna scolopes* squid by *Vibrio fischeri*. *J Vis Exp*:e3758. <https://doi.org/10.3791/3758>
35. Hargadon AC, Viliunas JW, Koehler S, Thies AB, Chen GY, Ladinsky MS, Kuwabara J, Avila-Magana V, Ruby EG, Tresguerres M, McFall-Ngai MJ. 2024. An acidic microenvironment produced by the V-type ATPase of *Euprymna scolopes* promotes specificity during *Vibrio fischeri* recruitment. *Commun Biol* 7:1642. <https://doi.org/10.1038/s42003-024-07348-2>
36. 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:e00853-20. <https://doi.org/10.1128/mBio.00853-20>
37. 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. <https://doi.org/10.1002/jmor.10422>
38. Moriano-Gutierrez S, Koch EJ, Bussan H, Romano K, Belcaid M, Rey FE, Ruby EG, McFall-Ngai MJ. 2019. Critical symbiont signals drive both local and systemic changes in diel and developmental host gene expression. *Proc Natl Acad Sci USA* 116:7990–7999. <https://doi.org/10.1073/pnas.1819897116>
39. Rogers TF, Yalçın G, Briseno J, Vijayan N, Nyholm SV, Simakov O. 2024. Gene modelling and annotation for the Hawaiian bobtail squid, *Euprymna scolopes*. *Sci Data* 11:40. <https://doi.org/10.1038/s41597-023-02903-8>
40. Domingo-Sananes MR, McInerney JO. 2021. Mechanisms that shape microbial pangenomes. *Trends Microbiol* 29:493–503. <https://doi.org/10.1016/j.tim.2020.12.004>
41. Soto W, Travisano M, Tolleson AR, Nishiguchi MK. 2019. Symbiont evolution during the free-living phase can improve host colonization. *Microbiology (Reading)* 165:174–187. <https://doi.org/10.1099/mic.0.000756>
42. Eickhoff MJ, Bassler BL. 2020. *Vibrio fischeri* siderophore production drives competitive exclusion during dual-species growth. *Mol Microbiol* 114:244–261. <https://doi.org/10.1111/mmi.14509>
43. Graf J, Ruby EG. 2000. Novel effects of a transposon insertion in the *Vibrio fischeri* *glnD* gene: defects in iron uptake and symbiotic persistence in addition to nitrogen utilization. *Mol Microbiol* 37:168–179. <https://doi.org/10.1046/j.1365-2958.2000.01984.x>
44. Septer AN, Wang Y, Ruby EG, Stabb EV, Dunn AK. 2011. The haemuptake gene cluster in *Vibrio fischeri* is regulated by Fur and contributes to symbiotic colonization. *Environ Microbiol* 13:2855–2864. <https://doi.org/10.1111/j.1462-2920.2011.02558.x>
45. 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. <https://doi.org/10.1128/JB.185.24.7202-7212.2003>
46. Frawley ER, Crouch M-LV, Bingham-Ramos LK, Robbins HF, Wang W, Wright GD, Fang FC. 2013. Iron and citrate export by a major facilitator superfamily pump regulates metabolism and stress resistance in *Salmonella Typhimurium*. *Proc Natl Acad Sci USA* 110:12054–12059. <https://doi.org/10.1073/pnas.1218274110>
47. Ma J, Feng J, Shan Y, Zhao Y, Qiao H, Xie L, Lin X, Wang C, Chuai X. 2020. Characteristic antimicrobial resistance of clinically isolated *Stenotrophomonas maltophilia* CYZ via complete genome sequence. *J Glob Antimicrob Resist* 23:186–193. <https://doi.org/10.1016/j.jgar.2020.09.008>
48. Phelps HA, Neely MN. 2007. *SalY* of the *Streptococcus pyogenes* lantibiotic locus is required for full virulence and intracellular survival in macrophages. *Infect Immun* 75:4541–4551. <https://doi.org/10.1128/IAI.00518-07>

49. Wang Y, Ruby EG. 2011. The roles of NO in microbial symbioses. *Cell Microbiol* 13:518–526. <https://doi.org/10.1111/j.1462-5822.2011.01576.x> 5
50. Lynch JB, Bennett BD, Merrill BD, Ruby EG, Hryckowian AJ. 2022. Independent host- and bacterium-based determinants protect a model symbiosis from phage predation. *Cell Rep* 38:110376. <https://doi.org/10.1016/j.celrep.2022.110376>
51. Xiao W, Weissman JL, Johnson PLF. 2024. Ecological drivers of CRISPR immune systems. *mSystems* 9:e0056824. <https://doi.org/10.1128/msystems.00568-24>
52. Burstein D, Sun CL, Brown CT, Sharon I, Anantharaman K, Probst AJ, Thomas BC, Banfield JF. 2016. Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nat Commun* 7:10613. <https://doi.org/10.1038/ncomms10613>
53. Millman A, Melamed S, Leavitt A, Doron S, Bernheim A, Hör J, Garb J, Bechon N, Brandis A, Lopatina A, Ofir G, Hochhauser D, Stokar-Avihail A, Tal N, Sharir S, Voichek M, Erez Z, Ferrer JLM, Dar D, Kacen A, Amitai G, Sorek R. 2022. An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* 30:1556–1569. <https://doi.org/10.1016/j.chom.2022.09.017>
54. Newsom S, Parameshwaran HP, Martin L, Rajan R. 2020. The CRISPR-Cas mechanism for adaptive immunity and alternate bacterial functions fuels diverse biotechnologies. *Front Cell Infect Microbiol* 10:619763. <https://doi.org/10.3389/fcimb.2020.619763>
55. Shabbir MAB, Shabbir MZ, Wu Q, Mahmood S, Sajid A, Maan MK, Ahmed S, Naveed U, Hao H, Yuan Z. 2019. CRISPR-cas system: biological function in microbes and its use to treat antimicrobial resistant pathogens. *Ann Clin Microbiol Antimicrob* 18:21. <https://doi.org/10.1186/s12941-019-0317-x>
56. Meadows JA, Wargo MJ. 2015. Carnitine in bacterial physiology and metabolism. *Microbiology (Reading)* 161:1161–1174. <https://doi.org/10.1099/mic.0.000080>
57. Gregory GJ, Dutta A, Parashar V, Boyd EF. 2020. Investigations of dimethylglycine, glycine betaine, and ectoine uptake by a betainecarnitine-choline transporter family transporter with diverse substrate specificity in *Vibrio* species. *J Bacteriol* 202:e00314-20. <https://doi.org/10.1128/JB.00314-20>
58. Wells MJ, Wells J. 1989. Water uptake in a cephalopod and the function of the so-called ‘pancreas’. *J exp Biol* 145:215–226. <https://doi.org/10.1242/jeb.145.1.215>
59. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. 1982. Living with water stress: evolution of osmolyte systems. *Science* 217:1214–1222. <https://doi.org/10.1126/science.7112124>
60. Kültz D. 2015. Physiological mechanisms used by fish to cope with salinity stress. *J Exp Biol* 218:1907–1914. <https://doi.org/10.1242/jeb.118695>
61. Dunlap PV. 1985. Osmotic control of luminescence and growth in *Photobacterium leiognathi* from ponyfish light organs. *Arch Microbiol* 141:44–50. <https://doi.org/10.1007/BF00446738>

62. Stabb EV, Butler MS, Adin DM. 2004. Correlation between osmolarity and luminescence of symbiotic *Vibrio fischeri* strain ES114. *J Bacteriol* 186:2906–2908. <https://doi.org/10.1128/JB.186.9.2906-2908.2004>
63. Dubois T, Krzewinski F, Yamakawa N, Lemy C, Hamiot A, Brunet L, Lacoste AS, Knirel Y, Guerardel Y, Faille C. 2020. The *sps* genes encode an original legionaminic acid pathway required for crust assembly in *Bacillus subtilis*. *mBio* 11:e01153-20. <https://doi.org/10.1128/mBio.01153-20>
64. Post DMB, Yu L, Krasity BC, Choudhury B, Mandel MJ, Brennan CA, Ruby EG, McFall-Ngai MJ, Gibson BW, Apicella MA. 2012. O-antigen and core carbohydrate of *Vibrio fischeri* lipopolysaccharide: composition and analysis of their role in *Euprymna scolopes* light organ colonization. *J Biol Chem* 287:8515–8530. <https://doi.org/10.1074/jbc.M111.324012>
65. Kazakov T, Kuznedelov K, Semenova E, Mukhamedyarov D, Datsenko KA, Metlitskaya A, Vondenhoff GH, Tikhonov A, Agarwal V, Nair S, Van Aerschot A, Severinov K. 2014. The RimL transacetylase provides resistance to translation inhibitor microcin C. *J Bacteriol* 196:3377–3385. <https://doi.org/10.1128/JB.01584-14>
66. Trachsel E, Redder P, Linder P, Armitano J. 2019. Genetic screens reveal novel major and minor players in magnesium homeostasis of *Staphylococcus aureus*. *PLoS Genet* 15:e1008336. <https://doi.org/10.1371/journal.pgen.1008336>
67. Wier AM, Nyholm SV, Mandel MJ, Massengo-Tiassé RP, Schaefer AL, Koroleva I, Splinter-Bondurant S, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, Bonaldo M de F, Casavant TL, Soares MB, Cronan JE, Reed JL, Ruby EG, McFall-Ngai MJ. 2010. Transcriptional patterns in both host and bacterium underlie a daily rhythm of anatomical and metabolic change in a beneficial symbiosis. *Proc Natl Acad Sci USA* 107:2259–2264. <https://doi.org/10.1073/pnas.0909712107>
68. Reichhardt MP, Holmskov U, Meri S. 2017. SALSA-A dance on a slippery floor with changing partners. *Mol Immunol* 89:100–110. <https://doi.org/10.1016/j.molimm.2017.05.029>
69. Green EH, Kotrannavar SR, Rutherford ME, Lunnemann HM, Kaur H, Heiser CN, Ding H, Simmons AJ, Liu X, Lacy DB, Washington MK, Shrubsole MJ, Liu Q, Lau KS, Sears CL, Coffey RJ, Drewes JL, Markham NO. 2025. Multiomic spatial atlas shows deleted in malignant brain tumors 1 (DMBT1) glycoprotein is lost in colonic dysplasia. *J Pathol* 266:51–65. <https://doi.org/10.1002/path.6406>
70. Borodina I, Kenny LC, McCarthy CM, Paramasivan K, Pretorius E, Roberts TJ, van der Hoek SA, Kell DB. 2020. The biology of ergothioneine, an antioxidant nutraceutical. *Nutr Res Rev* 33:190–217. <https://doi.org/10.1017/S0954422419000301>
71. Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ. 2011. Attenuation of host NO production by MAMPs potentiates development of the host in the squid–vibrio symbiosis. *Cell Microbiol* 13:527–537. <http://doi.org/10.1111/j.1462-5822.2010.01552.x>

72. Moriano-Gutierrez S, Bongrand C, Essock-Burns T, Wu L, McFall-Ngai MJ, Ruby EG. 2020. The noncoding small RNA SsrA is released by *Vibrio fischeri* and modulates critical host responses. *PLoS Biol* 18:e3000934. <https://doi.org/10.1371/journal.pbio.3000934>
73. Corrêa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MAR. 2016. Regulation of immune cell function by short-chain fatty acids. *Clin Transl Immunology* 5:e73. <https://doi.org/10.1038/cti.2016.17>
74. Mohammad S. 2015. Role of free fatty acid receptor 2 (FFAR2) in the regulation of metabolic homeostasis. *Curr Drug Targets* 16:771–775. <https://doi.org/10.2174/1389450116666150408103557>
75. Doino JA, McFall-Ngai MJ. 1995. A transient exposure to symbiosis-competent bacteria induces light organ morphogenesis in the host squid. *Biol Bull* 189:347–355. <https://doi.org/10.2307/1542152>
76. Essock-Burns T, Lawhorn S, Wu L, McClosky S, Moriano-Gutierrez S, Ruby EG, McFall-Ngai MJ. 2023. Maturation state of colonization sites promotes symbiotic resiliency in the *Euprymna scolopes*-*Vibrio fischeri* partnership. *Microbiome* 11:68. <https://doi.org/10.1186/s40168-023-01509-x>
77. Essock-Burns T, Bennett BD, Arencibia D, Moriano-Gutierrez S, Medeiros M, McFall-Ngai MJ, Ruby EG. 2021. Bacterial quorum-sensing regulation induces morphological change in a key host tissue during the *Euprymna scolopes*-*Vibrio fischeri* symbiosis. *mBio* 12:e0240221. <https://doi.org/10.1128/mBio.02402-21>
78. Fung BL, Visick KL. 2025. LitR and its quorum-sensing regulators modulate biofilm formation by *Vibrio fischeri*. *J Bacteriol* 207:e0047624. <https://doi.org/10.1128/jb.00476-24>
79. Miyashiro T, Oehlert D, Ray VA, Visick KL, Ruby EG. 2014. The putative oligosaccharide translocase SypK connects biofilm formation with quorum signaling in *Vibrio fischeri*. *Microbiologyopen* 3:836–848. <https://doi.org/10.1002/mbo3.199>
80. Ray VA, Visick KL. 2012. LuxU connects quorum sensing to biofilm formation in *Vibrio fischeri*. *Mol Microbiol* 86:954–970. <https://doi.org/10.1111/mmi.12035>
81. Tepavčević J, Yarrington K, Fung B, Lin X, Visick KL. 2022. sRNA chaperone Hfq controls bioluminescence and other phenotypes through Qrr1-dependent and -independent mechanisms in *Vibrio fischeri*. *Gene* 809:146048. <https://doi.org/10.1016/j.gene.2021.146048>
82. Rahman A, Mancini M, Nadon C, Perez IA, Farsamin WF, Lampe MT, Le TH, Torres Martínez L, Weisberg AJ, Chang JH, Sachs JL. 2023. Competitive interference among rhizobia reduces benefits to hosts. *Curr Biol* 33:2988–3001. <https://doi.org/10.1016/j.cub.2023.06.081>
83. Imes AM, Pavelsky MN, Badal K, Kamp DL, Briseno JL, Sakmar T, Vogt MA, Nyholm SV, Heath-Heckman EAC, Grasse B, Septer AN, Mandel MJ. 2025. *Euprymna berryi* as a comparative model host for *Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* 91:e0000125. <https://doi.org/10.1128/aem.00001-25>

2.2 The effect of SypC on the host

This section is modified from “SypC, a symbiont outer membrane vesicle protein, impacts the development of the squid-vibrio partnership .” Published in PNAS, March 2026.

Jill T. Kuwabara^{1,2}, Vera Beilinson³, Alexis C. Hargadon², Grischa Y. Chen^{2,3}, Xiao-Meng Hu^{2,3}, Mark S. Ladinsky³, Kathleen T. Hacksett⁴, Joseph R. Dillard⁴, Karen L. Visick⁵, Edward G. Ruby^{1,2}, Margaret McFall-Ngai^{1,2,3}

1. Pacific Biosciences Research Center, Kewalo Marine Laboratory, University of Hawaii at Manoa, Honolulu, HI 96813
2. Division of Biosphere Sciences and Engineering, Carnegie Science, Pasadena, CA 91125
3. Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125
4. Department of Medical Microbiology and Immunology, University of Wisconsin- Madison, School of Medicine and Public Health, Madison, WI 53705
5. Department of Microbiology and Immunology, Loyola University Medical Center, Maywood, IL 6015

V.Beilinson participated in RNA-seq analysis, NanoString analysis, and writing.

Supplemental figure (S1-S8) and tables (S1-S2) can be found online at:
<https://doi.org/10.1073/pnas.2524648123>

Dissertation Note

This study moves from observing strain-level differences to identifying the genetic mechanisms that produce them. The characterization of *sypC* demonstrates how variation in individual bacterial genes can alter symbiotic performance, providing a mechanistic framework for understanding how microbial genetic diversity influences host colonization and the establishment of stable symbiotic relationships.

Abstract

Bacterial outer membrane vesicles (OMVs) and the cargo they carry are increasingly recognized as a means of communication between microbial symbionts and the cells of their host. However, few studies have focused on the biochemical and molecular mechanisms underlying OMV signaling during symbiosis onset and development. We show here that SypC, an OMV protein of the bioluminescent symbiont *Vibrio fischeri*, is taken up by cells of the squid host *Euprymna scolopes* where it assumes a new function, i.e., the facilitation of symbiont-induced light-organ morphogenesis. SypC is a Wza-like outer membrane protein found in host-associated *Vibrionaceae* and is essential for *V. fischeri* biofilm formation. Colonization or direct treatment with *V. fischeri* OMVs triggers host development, which was reduced or delayed if the host is instead exposed to a Δ sypC mutant or Δ sypC OMVs. RNA-seq analyses comparing light organs colonized by either the mutant or its parent revealed differential expression of host genes associated with immune responses and tissue morphogenesis. In immunocytochemical imaging, SypC-bearing OMVs were taken up by the host's macrophage-like cells near the light-organ crypts, revealing the mechanism by which SypC travels through tissue to trigger morphogenesis. Taken together, the data provide evidence that in addition to its role in biofilm formation and colonization, SypC has a second function promoting the induction of symbiotic-tissue development. These findings provide a critical piece of a puzzle whereby a rich array of host and symbiont molecules work in

concert to orchestrate normal symbiont colonization and host development within the first hours to days of symbiosis.

Importance

Bacterial symbionts can profoundly influence host development. Using the *Euprymna scolopes-Vibrio fischeri* model symbiosis, this study provides evidence for recruitment of the outer membrane protein SypC, which is necessary for *V. fischeri* biofilm formation, into the process of host morphogenesis that occurs early in the symbiosis. This divergent dual-function role that SypC plays requires it to be shed from the cell surface of *V. fischeri* in outer membrane vesicles (OMVs). Host macrophage-like cells phagocytose and traffic the OMVs to a distant site where they are essential for induction of light-organ development. Because many species of the *Vibrionaceae* share unusual structural features of SypC, this protein may be critical for host–vibrio interactions in both nonpathogenic and pathogenic association

Introduction

Most, if not all, Gram-negative bacteria produce outer membrane vesicles (OMVs), which have recently been recognized to be involved in communication with host cells (for review, see refs. 1–3). For many species, OMVs are released from the bacterial cell surface by a nonlytic budding mechanism of the outer membrane (2, 4) that can be recognized by the almost solely periplasmic and outer membrane derivation of their proteome (e.g., refs. 5, 6). The cargo present in OMVs also includes nucleic acids and components of the outer envelope, such as lipopolysaccharide (LPS) and peptidoglycan (PGN) derivatives. Following release from the cell surface of bacterial symbionts, these vesicles interact with surface receptors on host epithelial or blood cells, or can even enter the cytoplasm, where they are often recognized by intracellular receptors. Increasing evidence suggests that OMVs internalized in host cells can release their cargo into the host-cell cytoplasm and influence the function and viability of those cells (6–8). OMVs have been principally characterized in bacterial pathogenesis (9, 10), but recent studies have shown them to be critical

elements of beneficial symbioses of both vertebrates (6) and invertebrates (see, e.g., refs. 11–13). For example, OMVs of the microbiota have been shown to enter the bloodstream of mammals, where they influence host immunity and disease response (12, 14), as well as to cross the blood–brain barrier, where they influence neurological health and disease (15).

Studies of the squid–vibrio model symbiosis have provided insights into the involvement of symbiont OMVs in modulating host–microbe interactions (16, 17). Host exposure to either ^{15}N -labeled *V. fischeri* cells or OMVs derived from these cells results in indistinguishable patterns of heavy-isotope labeling of the host tissues (7). In this system, the luminous bacterium *Vibrio fischeri* resides in crypt spaces along the apical epithelium of the symbiotic light-emitting organ of the Hawaiian bobtail squid, *Euprymna scolopes* (18, 19). Each generation, the aposymbiotic (APO) (i.e., symbiont-free) light organ of the newly hatched juvenile host specifically recruits *V. fischeri* cells from the bacterioplankton within hours; these founders and their progeny remain the symbionts throughout the squid’s lifespan. The host hatches with a light-organ morphology that potentiates symbiont recruitment (Fig. 1A). Ciliated fields on the nascent symbiotic organ entrain environmental *V. fischeri* cells above pores on the organ surface (Fig. 1 A’, Left). These cells aggregate in host-derived mucus (Fig. 1A’'). This behavior requires the production of an exopolysaccharide on the surface of *V. fischeri* cells, a function that is encoded by the symbiont’s *syp* locus (20). Thus, the biofilm consists of two components: host mucus and the bacterial exopolysaccharide. After aggregating, the cells migrate through host tissues to colonize the crypt spaces (Fig. 1 A’, Right), where they induce the loss of the superficial ciliated fields over the first 4 to 5 d following colonization (18). This gradual process involves two conspicuous phenotypes: i) widespread cell death of the superficial ciliated cells (Fig. 1B); and ii) the migration of squid-host macrophage-like blood cells, or hemocytes, into the blood sinuses of the ciliated-appendage epithelia (Fig. 1C). Following an upregulation of matrix metalloprotease activity in the basement membrane (21), the dying cells are sloughed from the

surface, i.e., they are not phagocytosed by the hemocytes in the blood sinus. The microbe-associated molecular patterns, or MAMPs, present in the OMVs of *V. fischeri* cells, specifically LPS and PGN derivatives, trigger most of this cell death and hemocyte migration, with symbiont light production being required for the full developmental program (22). While LPS and PGN are common to many bacterial species, this response is specific to *V. fischeri* because the signals originate in light-organ crypts, where only this species can persist (16, 23). Whereas a great deal of this process has been described, the mechanism by which the inductive message is transferred from the bacteria in the crypt spaces to the superficial ciliated fields has remained unresolved.

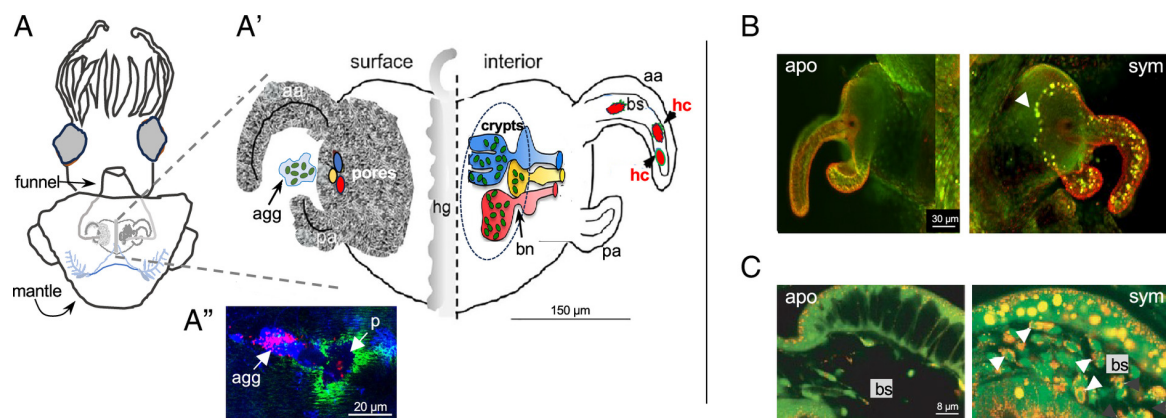


Figure 6 Juvenile light organ features and colonization developmental events. (A) The morphology and anatomy of the symbiotic organ. Ventral view of an *E. scolopes* showing the location of the light organ beneath the mantle and funnel. (A') Expanded diagram of the light organ. Left half: surface features of the organ, showing the superficial ciliated fields (gray) and three pores, relative to the position of the hindgut (hg). *V. fischeri* cells (green) in the ambient seawater aggregate in mucus outside of the pores (agg). Right half: internal regions of the organ including 3 crypts and the migration-pathway bottlenecks at their entrances (one indicated, bn). *V. fischeri* cells (green) are located in the crypts; the blood sinuses (bs) of the superficial ciliated appendages (aa, anterior appendage; pa, posterior appendage) contain migrating hemocytes (hc, red). (A'') Confocal image of aggregated *V. fischeri* cells (magenta) in host mucus (blue) outside of a pore (p) surrounded by cilia (green). (B) Once the light organ is colonized, the superficial ciliated fields regress by programmed cell death; dying cells (yellow puncta) were visualized by acridine orange

staining, characteristic of early stage cell death of symbiotic (sym) but not aposymbiotic (apo) light organs. (C) Close-up images of tissues of the blood sinuses (bs) of the ciliated appendages. After colonization by *V. fischeri* cells, the blood sinuses fill with hemocytes (white arrowheads).

Investigations of *V. fischeri* OMVs have shown that LPS, as well as hundreds of proteins and several RNA types (tRNA, rRNA, and small noncoding RNAs or sRNAs), are trafficked into host cells. Thus far, three of these biomolecules in the OMVs have been deeply characterized, i.e., the LPS (24), the sRNA SsrA, also called tmRNA (17), and the outer membrane protein OmpU (5). The data showed that these molecules each impact a specific feature of early symbiosis development.

In this study, we present data from an examination of the importance of SypC [See SI Appendix, Fig. S1A; UniProt (25)], a *V. fischeri* protein present in the OMVs (26) (SI Appendix, Fig. S1B), on host light-organ gene expression and development. SypC was selected for study because there is a deep understanding of components of the *V. fischeri* syp locus (20), which encodes proteins that produce an exopolysaccharide critical in symbiosis competency. SypC is an intrinsic outer membrane-spanning protein that is a paralog of the *V. fischeri* GfcE and its ortholog *Escherichia coli* Wza (SI Appendix, Fig. S1A) (27). It is essential for the export of the bacterium's exopolysaccharide (EPS) or Syp polysaccharide (28) and is the only protein encoded by the syp locus that is detected in OMVs (5). The data presented here provide evidence that SypC is critical for conveying the information that potentiates host light-organ development and promotes colonization of the crypts by *V. fischeri*.

Materials and methods

Animal Collection and Care.

All animal protocols and experiments were approved by the California Institute of Technology Institutional Animal Care and Use Committee. Adult *Euprymna*

scolopes Hawaiian bobtail squid were collected from Maunalua Bay, Oahu, HI, and either maintained in flow-through seawater tanks at Kewalo Marine Laboratory (Honolulu, HI) or transferred to recirculating tanks at Caltech (Pasadena, CA) and maintained as previously described (57). Juveniles bred from wild-caught adults were used for all experiments. After embryonic day 10, egg clutches were moved to a natural 12:12-hour light–dark cycle. Upon hatching, juvenile squid were collected, washed, and placed in i) 0.2 μ m-pore size filter-sterilized artificial seawater (Instant Ocean, Spectrum Brands, Blacksburg, CA) (FSIO = filter-sterilized Instant Ocean), ii) natural ocean seawater collected from the pier pump of Scripps Institution of Oceanography, La Jolla, CA (SIO seawater), or iii) filter-sterilized ocean seawater (FSOW), as indicated in the figure legends of a given experiment.

Bacterial Strains and Host Colonization Assay.

V. fischeri strains and plasmids used in this study are listed in SI Appendix, Table S1. The ability of different strains to colonize newly hatched juvenile *E. scolopes* was determined as described elsewhere (58, 59) (SI Appendix, Supplemental Methods).

Imaging of Cell Death in the Appendages.

Juvenile squid were collected at 24 h postinoculation, placed in 0.0001% acridine orange (Thermo Fisher Scientific, Irwindale, CA) in FSOW for 1 min, and washed with FSOW. Animals were anesthetized in 2% ethanol and their mantles and funnels were removed to expose the light organ. Light organ appendages were visualized to determine the number of acridine orange positive (early-stage cell death) nuclei using fluorescence confocal microscopy (60) (See SI Appendix, Table S2 for details). Late-

stage cell death, characterized by DNA fragmentation, was detected by TUNEL staining, according to the manufacturer's instructions (Promega, Madison, WI).

Construction of Protein Fusions in *V. fischeri*.

Because an antibody raised to the SypC protein itself was not easily detected by ICC, N-terminal mCherry translational fusions were constructed by amplification of mCherry from pSMV3-mCherry-zapA (61) and either sypC or ompU from the *V. fischeri* ES114 genome by standard Gibson assembly into pVSV105 (62). All fusions include an internal linker sequence (5'-CTCGAGGGATCCGGA-3') between mCherry and the gene of interest (See SI Appendix, Table S1). pCMC37 was cloned with the native signal peptide sequence of ompU on the 5' end of the fusion sequence. Transfer of plasmids into *V. fischeri* strains was performed by standard biparental conjugation protocol (63) using the *E. coli* donor strain MFDpir (64).

Immunocytochemistry of SypC and OmpU.

Juvenile squid were colonized within 3 h of hatching, at 1 h postdusk with 10,000 CFU of bacteria per ml (either WT+ mCherry, WT + mCherry-SypC conjugate, or WT + mCherry-OmpU conjugate). Colonization was confirmed via luminometer prior to fixation. Squid were fixed at 4 and 10 h postdawn (15 and 21 h postcolonization) in 4% PFA in marine PBS (mPBS) for 12 h at 4 °C with agitation, then rinsed four times for 15 min in mPBS. Light organs were dissected out of the squid and permeabilized for two days in mPBS/1% Triton X-100 (mPBST) at 4 °C with agitation. Blocking solution (0.5% BSA, 1% goat serum in mPBST) was added and samples were incubated in the dark overnight at 4 °C with agitation. Either an anti-mCherry monoclonal antibody (Thermo Fisher Scientific) or the corresponding rat IgG isotype control (Thermo Fisher Scientific) was added at a concentration of 1:100 in blocking solution and incubated in the dark for 14 d at 4 °C with agitation; fresh antibody was replaced at 7 d. Samples were rinsed 4 × 1 h with mPBST and

goat anti-rat IgG secondary antibody (Alexa Fluor 647 conjugated, Thermo Fisher Scientific) was added at 1:1000 concentration in blocking solution overnight. Samples were rinsed 5×5 min in mPBST, and 1×5 min in $5\times$ SSCT, then counterstained with 1:1,000 CellMask Orange Plasma Membrane Stain, and 1:50 DAPI (Thermo Fisher Scientific), in mPBST for two days at 4 °C with agitation.

Light organs were mounted in Vectashield (Fisher Scientific, Hampton, NH) on slides for confocal microscopy imaging. Lasers were set to the negative control to eliminate background autofluorescence and identical imaging parameters were used across all samples. Specific confocal microscope and imaging acquisition parameters are listed in SI Appendix, Table S2. Three-dimensional reconstructions of 0.3 μm thick Z-stacks were generated using Imaris 10.1 (Bitplane, Zurich, Switzerland).

Electron Tomography.

Samples of the light organ were prepared for electron microscopy. For sample preparation and details of the electron tomography and scanning electron microscopy methods, see SI Appendix, Table S2 and Supplemental Methods.

TCT Production by *V. fischeri*.

Analyses of the production of this MAMP by cultures of *V. fischeri* have been reported elsewhere and can be found in SI Appendix, Supplemental Methods.

OMV Isolation.

OMVs were isolated from *V. fischeri* cultures after 24 h of growth in LBS at 28 °C as previously described (5) (see SI Appendix, Supplemental Methods for details).

RNA Extraction for RNAseq and NanoString Experiments.

Juvenile squid were hatched into *V. fischeri*-free natural ocean water. Animals hatched on the same day and from the same clutch (i.e., with a related genetic

background) were separated into 3 treatments groups: APO (no exposure to *V. fischeri*), WT (exposed to wild-type *V. fischeri* strain ES114 at ~5,000 CFU/mL), and Δ sypC (exposed to a *V. fischeri* Δ sypC derivative at ~5,000 CFU/mL); n = 20 animals per sample, with 4 biological replicates. After 24 h of exposure, total light-organ RNA was isolated using an RNeasy Mini Kit (Qiagen, Germantown, MD) and Qiagen QIAshredder according to the manufacturer's instructions. Samples were treated with TURBO DNase (Thermo Fisher Scientific) to remove DNA contamination. RNA quality and concentration were determined by spectrophotometry using a NanoPhotometer (Implen, Westlake City, CA), and the concentration was confirmed using Qubit RNA BR assay kit (Thermo Fisher Scientific). For RNAseq analysis, the Illumina TruSeq Stranded mRNA Sample Prep with polyA selection v4.0 protocol was used for library preparation. Sequencing of light-organ tissue samples was performed at the NYU Genome Center on an Illumina HiSeq 4000 using a paired-end, 100-nucleotide-length run mode that produced ~35 to 85 million reads/sample.

For NanoString codeset analysis, RNA was prepared from APO, WT, and Δ sypC animals as described above. Raw nCounter data were passed into Rs NanoStringDiff (v1), which computed log2FC and qvalue to identify differentially expressed genes (DEGs); two host genes, serine hydroxymethyltransferase and REL_NF-kappaB, were chosen for normalization. Clutch number (as a proxy for genetic background) was included as a covariate in the model. nSolver (v4.0) was used to normalize raw NanoString data for visualization.

Determination of Differential Gene Expression.

Raw PE reads were quality- and adapter-trimmed using fastp (v0.24.0) with default parameters (65). Cleaned reads were then pseudoaligned, to the most recently published *E. scolopes* reference genome and annotations (v2.2) (66), using kb-python with bulk RNA-seq and reverse stranded tags (kb-python: v0.28.2, kallisto v0.50.0, bustools: v0.43.2) (67). On average, 64.4% of reads successfully

pseudoaligned. Gene-level counts were then passed into R's DESeq2 (v1.44.0) for differential expression analysis (68). Genetic background, or clutch, was included as a covariate in the model. DEG were identified for each pairwise comparison, due to high within sample variation, with P-value adjusted globally (DE: global padj <0.05, log2FC $\pm$ 0.5). Plots were made with ggplot2 (v3.5.1), ComplexUpsetR (v1.3.3), and variancePartition (v1.34.0).

Statistical Analyses.

All statistical analyses were conducted using Prism 8 (Graphpad Software, La Jolla, CA). Statistical analyses between two groups were analyzed using unpaired Student's t test. Comparisons between multiple groups were analyzed using two-way ANOVA with Tukey test. Statistical variability is expressed as mean $\pm$ 1 SD; ns: not significant, $P > 0.05$; * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$. N values for experiments are indicated in figure legends as appropriate.

Results and discussion

Δ sypC-Colonized Light Organs Have Altered Transcriptional Responses.

To understand how colonization by a Δ sypC symbiont might differentially influence host gene expression, we performed an RNA-seq experiment at 24 h postinoculation (hpi). Relative to APO animals, *V. fischeri* colonization, regardless of the presence of sypC in the symbionts (i.e., WT or Δ sypC), modulated the expression of 396 genes, clearly distinguishing uncolonized squid from symbiotic ones. (Fig. 2 A, Left, and B). This result is not unexpected, as multiple factors from *V. fischeri* (e.g., MAMPs and light production) are known to modulate host gene expression (22). Nevertheless, the presence of sypC did have a clear transcriptional impact, specifically affecting a distinct subset of host genes. Differential expression (DE) analysis (see PERMANOVA, in SI Appendix, Supplemental Results) comparing colonization by either the WT or the Δ sypC strain identified 138 sypC-dependent regulated genes (green and yellow, Fig. 2B). Hierarchical clustering showed that

these genes had similar expression levels in APO and Δ sypC treatments, indicating that to be regulated, these genes require the presence of sypC in the symbionts (Fig. 2 C, Left). Additionally, we identified 345 genes that were not significantly different after WT or Δ sypC colonization but instead exhibited an intermediate level of expression in Δ sypC relative to WT and APO (Fig. 2 C, Right). These genes were statistically DE between WT/APO, but not between either Δ sypC/APO or Δ sypC/WT (brown, Fig. 2B). Such a pattern suggests that sypC contributed to achieving full expression of these genes.

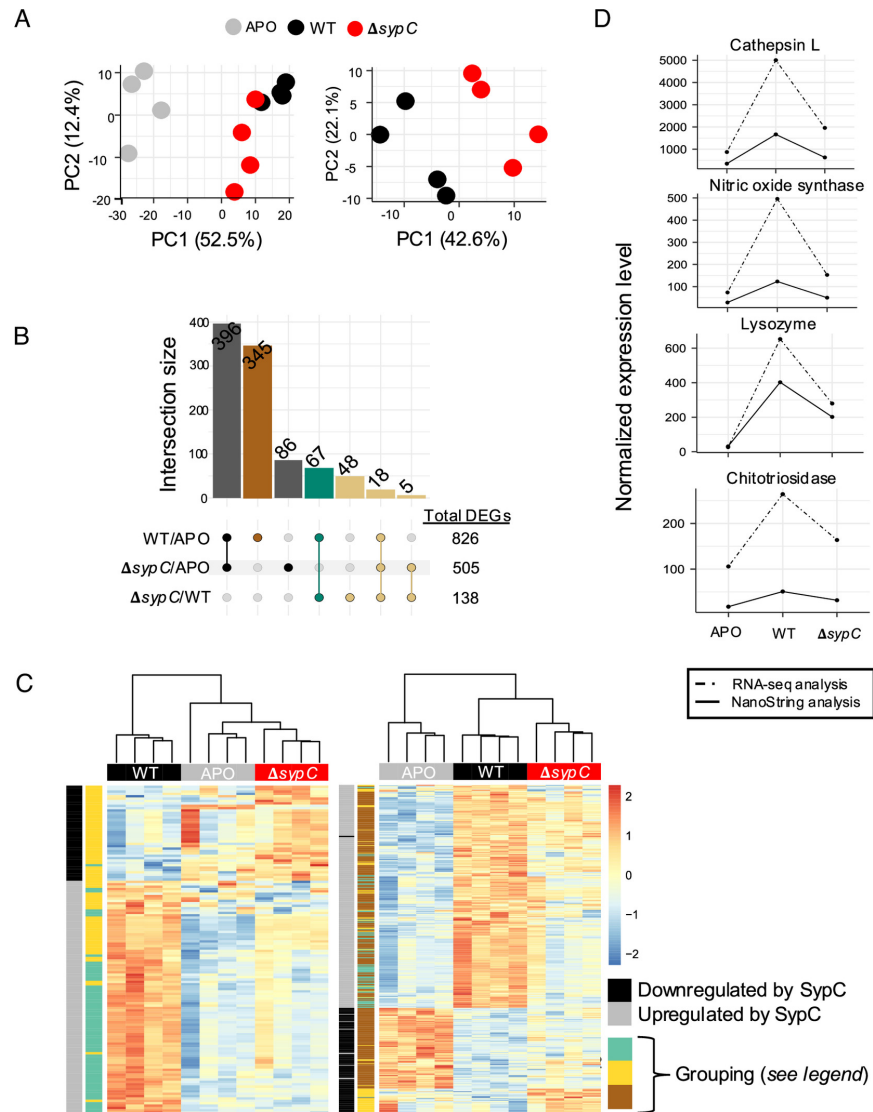


Figure 7 Relative impact of the $\Delta sypC$ mutant on light-organ gene expression. (A) Principal component analysis (PCA) of host transcriptomes. Left: three conditions are shown: uncolonized (APO, gray), colonized with wild-type bacteria (WT, black), or colonized with the $\Delta sypC$ mutant ($\Delta sypC$, red). Right: Same data as Left, but with APO excluded to better highlight the separation between WT and $\Delta sypC$. Percent variance explained by each principal component is indicated on the axes. (B) UpSet plot of differentially expressed genes (DEGs). Comparisons include WT vs. APO and $\Delta sypC$ vs. APO ($\log_2 FC \geq |1|$, adjusted $P < 0.05$), and $\Delta sypC$ vs. WT ($\log_2 FC \geq |0.58|$, adjusted $P < 0.05$). (C) Expression patterns of selected DEG categories. Left: Green and yellow genes highlight that $\Delta sypC$ colonized animals resemble APO more closely for this subset. Right: Green, yellow, and brown gene groupings illustrate that

Δ sypC represents an intermediate expression state between APO and WT. Color coding: Black: Downregulated by SypC (Δ sypC > WT). Gray: Upregulated by SypC (Δ sypC < WT). Green: DE in both WT vs. APO and Δ sypC vs. WT. Yellow: DE only in Δ sypC vs. WT. Brown: DE only in WT vs. APO. (D) Expression validation of selected genes comparing RNA-seq and NanoString analyses.

Examining all genes of interest (Fig. 2; brown, yellow, and green), GO-term enrichment analysis highlights the possibility that symbiont carriage of sypC suppresses pathways related to cytoskeletal organization and cellular motility, and upregulates host responses involved in metabolism, immune modulation, cell adhesion, and extracellular remodeling (SI Appendix, Fig. S2), including production of a matrix metalloprotease by the ciliated epithelium, which is increased 24 h after colonization and is involved in light-organ morphogenesis (29). Twenty-five other upregulated host genes of interest were associated with cell death and proteolysis (SI Appendix, Fig. S3 A and B), supporting the idea that the presence of the SypC protein promotes host-cell death in the ciliated fields (Fig. 3A).

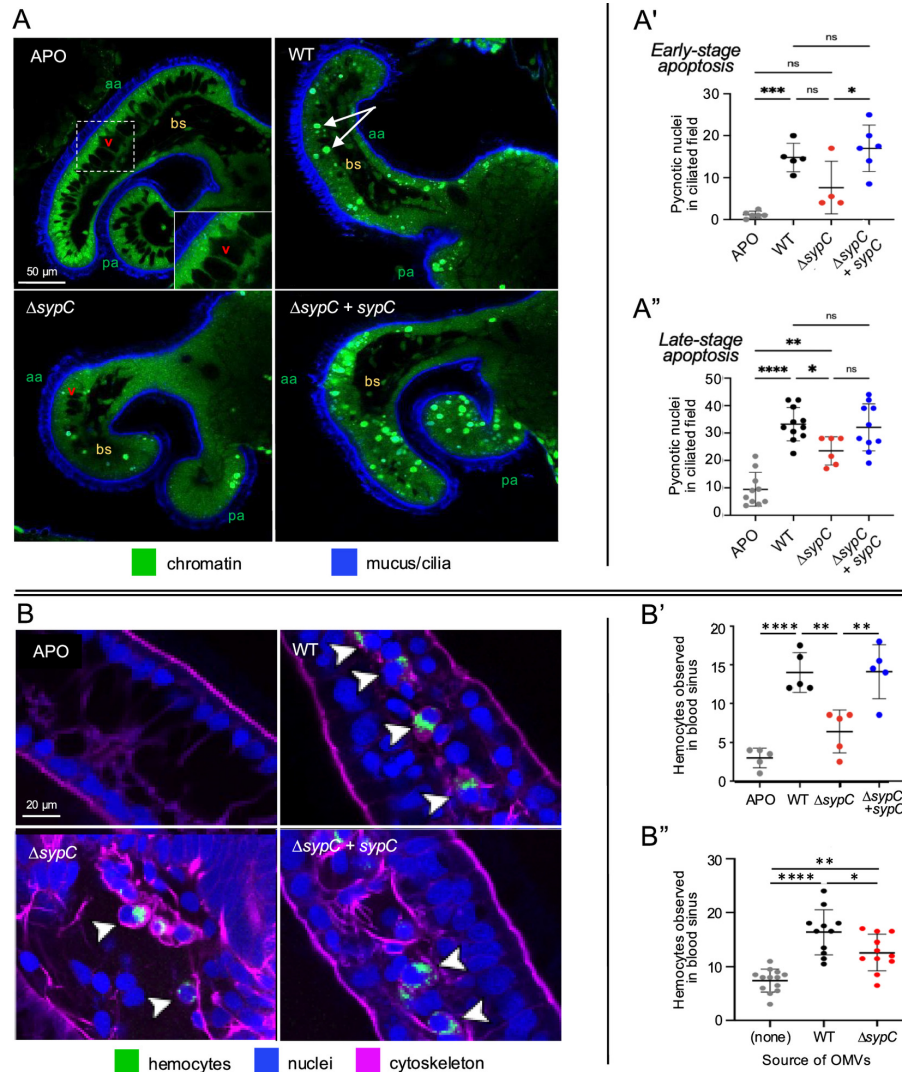


Figure 8 Colonization by the *V. fischeri* $\Delta sypC$ mutant attenuates normal cellular phenotypes of the host. (A) Representative images of the superficial ciliated fields of light organs that were APO or colonized with wild-type *V. fischeri* (WT), $\Delta sypC$, or the complemented $\Delta sypC + sypC$ strains. At 24 h postcolonization, levels of symbiosis-induced early-stage apoptosis were determined in the ciliated field with each condition, as indicated by pycnotic nuclei (condensed chromatin, e.g., white arrows) in these regions. Fluorochromes used: cell death was indicated by acridine-orange stained condensed chromatin (green); mucus was stained with wheat-germ agglutinin (blue). aa, anterior appendage; bs, blood sinus; pa, posterior appendage; v, vacuoles within the lateral epithelial cells of the ciliated field. (A') Quantification of nuclei indicating early-stage cell death. (A'') Quantification of late-stage apoptosis was determined using the fluorochrome TUNEL (terminal deoxynucleotidyl transferase dUTP nicked-end labeling).

(B) Representative images of hemocytes (blood cells, white arrowheads) trafficking into the blood sinus of the anterior appendage under the four colonization conditions. Fluorochromes used: hemocytes were identified by DNase-1 staining (green); F-actin cytoskeleton was stained with phalloidin (magenta); nuclei, TO-PRO-3 (blue). (B') The number of hemocytes present in the blood sinus of the anterior appendage was quantified at 24 h after colonization by different *V. fischeri* strains. (B'') Hemocyte trafficking was also determined in animals exposed only to OMVs that had been isolated from either WT or Δ sypC strains. For all conditions in A' and A'', and B' and B'' each data point represents a biological replicate. Each experiment was repeated 3 times with the same outcome. Results are presented as mean $\pm$ 1 SD. Comparisons between multiple groups were analyzed using one-way ANOVA with Tukey test. ns: not significant, $P > 0.05$; * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$.

To validate the RNA-seq results, NanoString analysis was performed on selected targets (Fig. 2D and SI Appendix, Fig. S3C). Previously studied symbiosis-associated genes such as cathepsin L, nitric oxide synthase, lysozyme, and chitotriosidase (reviewed in ref. 18) were upregulated by symbionts encoding SypC, consistent with both cell-death enhancement and immune-response modulation. Most of these crucial developmental genes were downregulated in a colonization by a Δ sypC mutant compared to WT, further highlighting the importance of SypC during tissue morphogenesis.

Colonization by the Δ sypC Mutant Affects Certain Features of Host Development.

In response to colonization by the Δ sypC mutant, several typical wild-type-induced developmental phenotypes were significantly reduced (Fig. 3). Both early-stage (characterized by condensed chromatin; Fig. 3 A and A') and late-stage (characterized by fragmented chromatin; Fig. 3A'') cell death were often statistically different; however, due to natural genetic variation of juveniles during development, some clutches only showed significant differences during late-stage cell death. Similarly, the number of hemocytes migrating into the blood sinuses of the anterior appendages of the ciliated fields was consistently statistically reduced in Δ sypC-colonized animals (Fig. 3 B and B'). Treatment of juveniles with *V. fischeri* OMVs

was sufficient to induce the migration of hemocytes to the blood sinus (23), even in the absence of the bacteria themselves (Fig. 3B''), indicating that OMVs alone were adequate to trigger this key developmental event. Further, the reduced level of hemocyte migration that resulted from a Δ sypC colonization (Fig. 3B') was phenocopied when animals were exposed to OMVs produced by the mutant strain relative to wild-type OMVs (Fig. 3B'').

SypK, another protein encoded by the syp locus and essential for biofilm formation (28), contains the conserved RfbX flippase domain found in Wzx (30). It likely transports subunits of the Syp polysaccharide from the cytoplasm into the periplasm where they are assembled. SypK is predicted to be associated with the cytoplasmic membrane and thus, unlike the outer membrane SypC, was not detected in OMVs (5). In contrast to the Δ sypC mutant, colonization by the Δ sypK mutant did not significantly affect the level of hemocyte migration (SI Appendix, Fig. S4). These results indicate that it is the absence of the SypC protein itself, not a disruption of exopolysaccharide production, that is responsible for the inability of Δ sypC to induce normal hemocyte migration.

Certain other morphological events in the developmental program are also absent or delayed in Δ sypC colonized animals, such as the loss of the ciliated ridge on the light organ's surface (31) (SI Appendix, Fig. S5A) at 24 hpi. However, squids colonized by Δ sypC did retain several other normal developmental phenotypes, such as the typical regression of the superficial ciliated fields, which did not differ significantly between WT- and Δ sypC-colonized animals at 72 hpi (SI Appendix, Fig. S5 B–D). In addition, a phenotype occurring more deeply in the symbiont's migration path, i.e., the symbiont-induced closure of the migration path "bottleneck" (32, 33), was similar in animals colonized by either the WT or Δ sypC strain (SI Appendix, Fig. S5E).

Given that the peptidoglycan monomer TCT (tracheal cytotoxin) induces hemocyte migration into the blood sinuses of the appendages (26), we hypothesized that the Δ sypC would be deficient in TCT production. To test this notion, we measured the

amount of TCT produced by WT and Δ sypC cells in culture and determined that the mutant produced roughly the same level of TCT as the WT strain (See SI Appendix, Fig. S6 A and B). This result suggests that the developmental defect caused by the Δ sypC symbiont is not in the production of TCT but instead in the ability of this TCT to induce hemocyte migration normally. Interestingly, the experimental addition of high levels of TCT directly to seawater can itself induce hemocyte migration in both uncolonized and Δ sypC-colonized animals (SI Appendix, Fig. S6C), indicating that such a pharmacological treatment with TCT alone can be sufficient to overcome the typical requirement that SypC be present to enable induction of hemocyte migration.

Symbiont SypC Protein Is Found within Host Hemocytes.

Because the SypC protein itself is required for normal host responses to either colonization by symbionts or treatment with OMVs, we investigated whether this protein's effect was due to accumulation within host cells. Prior studies have shown that components of symbiont-produced OMVs can enter host cells, including LPS, which induces early-stage cell death (24), and small RNAs (17). These findings suggest a mechanism by which OMV-associated SypC (5) could also be delivered into host cells. To determine whether SypC is, in fact, translocated within the light organ, we introduced into wild-type *V. fischeri* a gene encoding an mCherry-SypC fusion protein that was constitutively expressed from its plasmid (SI Appendix, Table S1). Thus, this SypC fusion protein was well expressed, functional, and could easily be detected in bacterial cell extracts using a fluorescent antibody made against mCherry (SI Appendix, Fig. S7). When juvenile animals were examined at 24 hpi, the SypC fusion protein was observed within the colonized light organ by immunocytochemistry using an anti-mCherry antibody (Fig. 4). As expected, the symbionts themselves were strongly labeled. The SypC fusion protein was also observed in the cytoplasm of the epithelial cells lining the crypts (Fig. 4A); however, when colonized by wild-type symbionts carrying mCherry alone, no antibody reaction was observed in hemocytes (SI Appendix, Fig. S8), indicating that these

cells internalize OMVs, but not entire *V. fischeri* symbionts. Further, the SypC-fusion labeling was in proximity to the nucleus, but excluded from the nuclear region (Fig. 4 B and C). Examination of tissues distant from the crypts revealed labeled SypC (Fig. 4 D–G'') appearing even as remotely as the blood sinus of the external ciliated appendages (Fig. 5A). In colonized animals, hemocytes migrate to the appendages prior to the onset of cell death (34); thus, we asked whether SypC might be associated with these tissue-targeting cells and, if so, whether the presence of SypC might be connected to cell-death-driven morphogenesis.

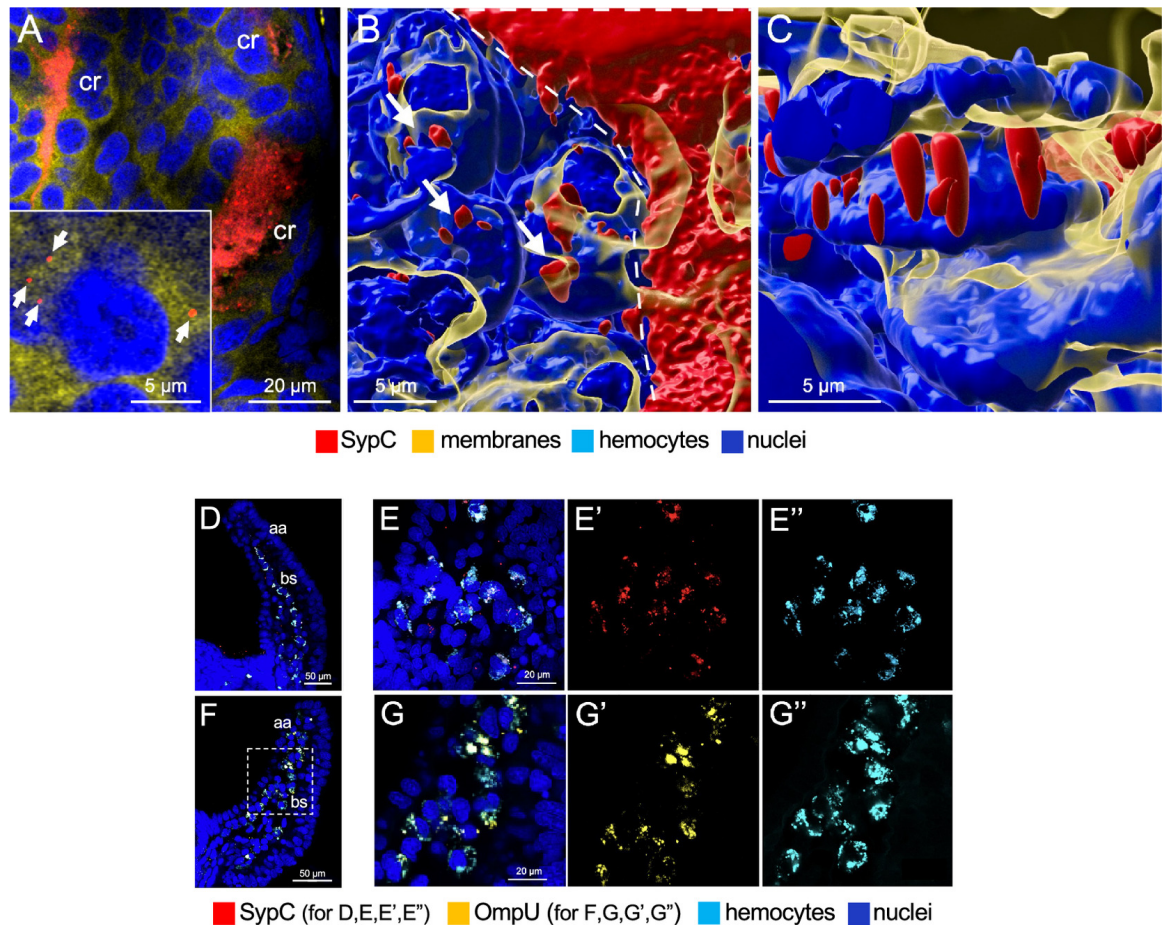


Figure 9 Components and activities of migrating hemocytes. Localization of SypC within and near the symbionts at 24 hpi. (A) Abundance of antibody labeling of *V. fischeri* SypC in the light-organ crypt spaces (cr). Insert, mCherry-SypC fusion protein detected by fluorescent mCherry antibody (white arrows) within the

cytoplasm of host cells lining the crypts. (B) Imaris rendering of higher-magnification confocal images showing SypC (white arrows) within host cells in regions near nuclei. The dashed line indicates the boundary between the host cells (Left) and the symbiont-containing crypts. CellMask Orange-stained membranes in and surrounding the cytoplasm. (C) Imaris rendering demonstrating that SypC is extranuclear. (D) Labeling of the *V. fischeri* SypC-fusion protein within the blood sinus (bs) of an anterior appendage (aa) of the superficial ciliated field of the organ; low magnification image showing the general location of higher magnification area. (E–E'') Higher magnification images showing the association of SypC with hemocytes in the blood sinus. DAPI-stained nuclei, (blue); fluorescent antibody to mCherry portion of mCherry-SypC fusion protein (red); CellMask Orange-stained hemocytes (turquoise). (F) Labeling of the *V. fischeri* OmpU protein within the blood sinus; low magnification image as in D. (G–G'') Higher magnification images showing the association of OmpU with hemocytes in the blood sinus. DAPI-stained nuclei, (blue); fluorescent antibody to OmpU (yellow); CellMask Orange-stained hemocytes (turquoise). Images of experimental controls can be found in SI Appendix, Fig. S8.

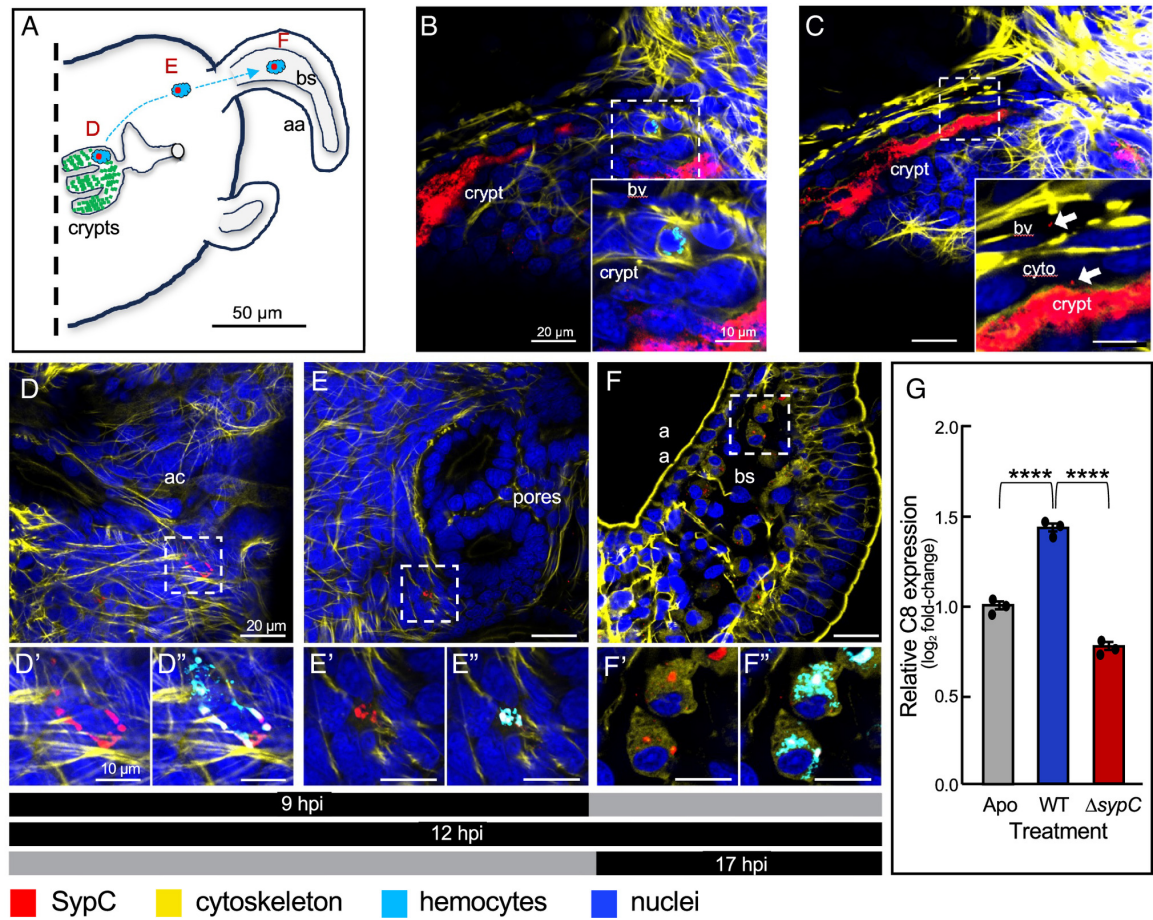


Figure 10 Migration of SypC-bearing hemocytes to the blood sinus. (A) Tracking of host hemocytes during the first 17 h postinfection (hpi) as they migrate from the crypts to the blood sinus (bs) of the anterior appendage (aa). The presence of SypC-bearing hemocytes was indicated by the colocalization of a mCherry-SypC fusion protein (red) detected by a fluorescent mCherry antibody, with the CellMask Orange-stained hemocytes (turquoise). The hemocytes were observed near the medial crypt tissues at 9 hpi (B–D), progressed through the lateral tissues of the light organ by 12 hpi (E), and became situated in the blood sinus of the anterior appendage as soon as 17 hpi (F). The areas indicated by the dashed white boxes are enlarged in the insets and lower images. SypC could also occasionally be seen free in the blood vessels (bv) and in the cytoplasm of the crypt epithelium (white arrows) (C). The black timeline indicates at which image locations (D, E, and/or F) the SypC-carrying hemocytes could be observed when viewed at 9, 12, or 17 hpi. Nuclei were stained with DAPI (blue) and cytoskeleton was labeled with rhodamine phalloidin (yellow). ac, antechamber; cyto, host cytoplasm. (G) qRT-PCR determination of the expression level of the C8 subunit of the proteasome.

Relative expression of the C8 gene in hemocytes in uncolonized juveniles (Apo) or 24 h postcolonization by either wild-type (WT) or Δ sypC symbionts. Differences are indicated by a one-way ANOVA with Tukey's multiple comparison test ($n = 3$). Data are presented as the mean $\pm$ 1 SEM. **** $P < 0.0001$.

Colocalization of the SypC label with a hemocyte-specific marker confirmed that this symbiont protein was internalized and carried into the sinuses of the appendages by these host cells (Fig. 4 D and E–E’’). We next hypothesized that the SypC protein entered the phagocytic hemocytes through the uptake of symbiont OMVs they encountered while traversing the crypts (18). We tested this by asking whether OmpU, another symbiont protein present in OMVs (5), also colocalized with the hemocytes in the appendages. As predicted, an antibody to an mCherry-OmpU fusion protein labeled these host cells (Fig. 4 F and G–G’’), indicating that the uptake of SypC was not specific to that protein, and likely occurred during the phagocytosis of OMVs.

The Progression of SypC Transported to the Blood Sinuses by Migrating Hemocytes Is Associated with the Induction of a Proteolytic Gene.

The path of SypC-containing hemocytes was tracked from the light-organ crypts to the blood sinuses of the superficial ciliated appendages. At 9 hpi, SypC was first observed in the hemocytes near the crypt spaces, presumably after they had phagocytosed symbiont OMVs (Fig. 5 A–D). By 12 hpi, labeled hemocytes were most easily detected distally near the region of the pores (Fig. 5E). Finally, by 17 hpi, hemocytes containing SypC were primarily in the blood sinuses of the appendages (Fig. 5F). Hemocytes collected from symbiotic animals at 24 hpi had increased levels of the C8 subunit of the proteasome, a proteolytic complex that is induced in hemocytes of the wild-type colonized light organ during cell death of the appendages

(35). In contrast, no such induction was seen in hemocytes collected from Δ sypC colonized squid (Fig. 5G).

Discussion

SypC is a putative outer membrane transport protein involved in Syp-polysaccharide secretion and biofilm formation, activities that are critical for the initiation of symbiosis by planktonic *V. fischeri* cells (for review, see ref. 20), as well as for their colonization of the light organ (36). Recent studies have shown that, of the 18 proteins encoded by the well-studied syp locus, SypC is the only one that is present in *V. fischeri* OMVs (5), likely due to its being embedded within the outer membrane. Here, we present data on host phenotypes that rely on the presence of SypC for normal symbiosis onset. Expression of hundreds of juvenile-host genes is dependent on colonization by symbionts that encode SypC (Fig. 2). Further, in the early hours of the host–symbiont partnership, SypC-bearing OMVs are taken up by phagocytic blood cells, or hemocytes, of the juvenile host and trafficked through its tissues to the outer surface of the light organ. The SypC delivered to the surface is essential for inducing, either directly or indirectly, normal development of the symbiosis.

Prior to the work described here, the presence of symbiont OMVs in the bloodstream of the squid host (17) had suggested that bacterial proteins like SypC, present in phagocytosed OMVs, might impact not only the light organ, but other tissues as well (37). Our findings add to the growing recognition that OMV trafficking through the body of a host is an evolutionarily conserved mechanism of host–symbiont communication (38, 39). Such long-range trafficking of OMVs has been documented in a variety of eukaryotic hosts spanning from plants, choanoflagellates, and sponges to squids and vertebrates (5, 6, 40, 41). In mammals, not only do OMVs of the gut microbiome have either beneficial or pathologic influences locally on the gastrointestinal epithelium, they also cross into the bloodstream where they are transported to distant tissues with wide-ranging effects (for reviews, see refs. (42, 43)). Recent studies of pregnant women have provided evidence that this

phenomenon even occurs with the delivery of OMVs from the mother's microbiota to the developing fetus (44). Once the OMVs are delivered, they can have an array of effects (38). For example, notable here is the role of the *V. fischeri* OMVs in inducing normal cell death in the light organ's superficial ciliated epithelium (45); such OMV-induced cell death has also been reported in a diversity of other host animals, from sponges to mammals (40, 46).

The activity of SypC in bacterial-induced animal development is only one of a variety of bacterial and host features that work in concert to promote a normal and productive symbiosis. SypC is unusual in that its role in inducing host development is so distinct from its role in biofilm secretion. Proteins that have taken on a divergent role have been termed as "dual-function" (see, e.g., ref. 47). These proteins are not uncommon within a bacterial species, or between bacterial species, but it is rare to have trans-Domain functions. While this study of the Syp protein may be the first to describe a bacterial-animal dual-function protein, such proteins have previously been reported between plants and their rhizobial symbionts (48, 49).

Studies of the OMVs have demonstrated that they contain cargo other than SypC that impacts light-organ morphogenesis, such as another outer membrane protein, OmpU (5). In addition, nonprotein cargo of the OMVs influences organ development, including i) MAMPs such as LPS (24) and peptidoglycan derivatives (23, 26) that directly induce the developmental program; and, ii) the small RNA, SsrA (also called tmRNA), which influences host development and symbiont persistence by modulating the host's immune system (17). Several other non-OMV elements are critical for morphogenesis including the symbiont's production of light (22) and modulation of crypt pH (5). The role of the host in this dialogue is similarly complex, coordinating hemocyte migration from the light-organ crypts where OMVs are engulfed, to the blood sinuses of the superficial epithelium where their contents are delivered (Fig. 5). The host responds to this delivery by upregulating expression of genes, such as those encoding the proteasome in the hemocytes as well as

metalloproteinases at the base of the developing epithelium. These events enable the normal cell-death process that underlies the morphogenetic program (for review, see ref. 18). Overall, the level of complexity characterizing host–symbiont codevelopment is reminiscent of that evident during embryogenesis: the migration of cells and diffusion of morphogens from one tissue to another, delivering signals and triggering developmental programs. In both progressions, the specification of cell fates is required for a normal, healthy outcome, demanding coordination of several factors both across time and space, and in balanced proportions. The principal difference is that, with a few exceptions (50, 51), embryonic development typically requires coordination only within a single genome; as described here, in a beneficial symbiosis the two evolutionarily distinct genomes of the bacterium and its host must communicate.

A common feature of animal associations with *Vibrio* species, whether pathogenic or beneficial, may be the functional diversification of SypC, reflected in both the protein's structure and its role in symbiosis. A simulation of SypC using AlphaFold2 (52) suggests the protein may have evolved from another outer membrane transporter, Wza (53), with an increase in domain type (54) and number (SI Appendix, Fig. S1). The UniProt database (25) currently lists 14 SypC homologs in conspecifics or congeners of *V. fischeri*. These proteins all have subunits consisting of ~703 amino acids, have similar secondary/tertiary structures, and most have been reported as proteins of mutualistic or pathogenic bacteria of aquatic animals. At the 50+% level of amino acid identity, dozens of proteins are listed, ~2/3 of which are in the *Vibrionaceae* and have a similar structure, with 700 (± 5) amino acids. For example, a SypC homolog has been found in the hydrothermal-vent bacterium, *Vibrio diabolicus* (55). In addition, and relevant here, a sypC gene encoding a protein of 699 aa, with 68% sequence identity to *V. fischeri* SypC has been reported as cargo in the OMVs of *Vibrio tasmaniensis*, a pathogen of oysters

(56). These phylogenetic patterns of SypC-like proteins suggest a widespread, essential function in host–symbiont interactions of the *Vibrionaceae*.

The results presented here provide another piece of the puzzle of how a host and its species-specific bacterial symbiont communicate at each new interaction during their 5 My-old partnership. In particular, they address the challenge of how symbiont signals produced deep in the light-organ crypts are transmitted to, and drive host development in, the surface epithelium. The data provide evidence that the transmission is not indirect, i.e., broadcast from the crypt spaces where the symbionts reside by signal transduction cascades passing through the tissues to the surface. Instead, signal delivery appears to be direct, i.e., by the hemocyte-trafficking of OMVs and their cargo to distant tissues that undergo morphogenesis. Many more questions concerning the coordination of this process remain to be addressed. For example: What are the molecular interactions of the SypC protein with the host macrophages that trigger host morphogenesis? Are these interactions direct, or are they indirect, perhaps by affecting the composition or activity of other outer membrane proteins in the OMVs? Do the mechanisms underlying the described behaviors involve the host's immune system or other symbiont MAMPs? Finally, is the structure of the SypC molecule, which appears to be unique to *Vibrio* spp., critical for interactions with the host?

Acknowledgements

We thank Zachary A. Yahiku and Hannah Osland for their technical help. Dr. William Goldman (University of North Carolina at Chapel Hill) provided purified TCT. Dr. Joram Piatigorsky provided valuable advice. This work was supported by funding from the NIH Grants R37-AI50661 to M.M.-N. and E.G.R. and R01-GM135254 to E.G.R. and M.M.-N.

J.T.K., V.B., G.Y.C., J.P.D., K.L.V., E.G.R., and M.M.-N. designed research; J.T.K., V.B., A.C.H., G.Y.C., X.-M.H., M.S.L., K.T.H., and K.L.V. performed research; J.P.D. and K.L.V. contributed new reagents/analytic tools; J.T.K., V.B., A.C.H.,

G.Y.C., X.-M.H., M.S.L., K.T.H., J.P.D., K.L.V., E.G.R., and M.M.-N. analyzed data; and J.T.K., V.B., A.C.H., G.Y.C., M.S.L., J.P.D., K.L.V., E.G.R., and M.M.-N. wrote the paper.

References

1. V. Tiku, M. W. Tan, Host immunity and cellular responses to bacterial outer membrane vesicles. *Trends Immunol.* 42, 1024–1036 (2021).
2. M. Toyofuku, S. Schild, M. Kaparakis-Liaskos, L. Eberl, Composition and functions of bacterial membrane vesicles. *Nat. Rev. Microbiol.* 21, 415–430 (2023).
3. L. Turnbull et al., Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* 7, 11220 (2016).
4. L. Zavan et al., The mechanism of *Pseudomonas aeruginosa* outer membrane vesicle biogenesis determines their protein composition. *Proteomics* 23, e2200464 (2023).
5. J. B. Lynch et al., Ambient pH alters the protein content of outer membrane vesicles, driving host development in a beneficial symbiosis. *J. Bacteriol.* 201, e00319-19 (2019).
6. M. G. Sartorio, E. J. Pardue, N. E. Scott, M. F. Feldman, Human gut bacteria tailor extracellular vesicle cargo for the breakdown of diet- and host-derived glycans. *Proc. Natl. Acad. Sci. U.S.A.* 120, e2306314120 (2023).
7. S. K. Cohen et al., Tracking the cargo of extracellular symbionts into host tissues with correlate delectron microscopy and nanoscale secondary ion mass spectrometry imaging. *Cell Microbiol.* 22, e13177 (2020).
8. D. N. Villageliu, D. R. Samuelson, The role of bacterial membrane vesicles in human health and disease. *Front. Microbiol.* 13, 828704 (2022).
9. F. S. Abolhasani et al., Outer membrane vesicles in gram-negative bacteria and its correlation with pathogenesis. *Front. Immunol.* 16, 1541636 (2025).
10. J. Li et al., Outer membrane vesicles produced by coral-associated *Vibrio coralliilyticus* inhibit bacteriophage infection and its ecological implications. *Microbiol. Res.* 281, 127607 (2024).
11. H. Gao et al., Outer membrane vesicles from a mosquito commensal mediate targeted killing of *Plasmodium* parasites via the phosphatidylcholine scavenging pathway. *Nat. Commun.* 14, 5157(2023).
12. J. B. Lynch, R. A. Alegado, Spheres of hope, packets of doom: The good and bad of outer membrane vesicles in interspecies and ecological dynamics. *J. Bacteriol.* 199, e00012-17 (2017).
13. G. Wu et al., Outer membrane vesicles induce the mussel plantigrade settlement via regulation of c-di-GMP. *Biofouling* 39, 359–370 (2023).
14. Y. Shen et al., Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell Host Microbe* 12, 509–520 (2012).

15. A. Kaisanlahti et al., Bacterial extracellular vesicles - brain invaders? A systematic review. *Front. Mol. Neurosci.* 16, 1227655 (2023).
16. M. S. Aschtgen et al., Rotation of *Vibrio fischeri* flagella produces outer membrane vesicles that induce host development. *J. Bacteriol.* 198, 2156–2165 (2016).
17. S. Moriano-Gutierrez et al., The noncoding small RNA SsrA is released by *Vibrio fischeri* and modulates critical host responses. *PLoS Biol.* 18, e3000934 (2020).
18. S. V. Nyholm, M. J. McFall-Ngai, A lasting symbiosis: How the Hawaiian bobtail squid finds and keeps its bioluminescent bacterial partner. *Nat. Rev. Microbiol.* 19, 666–679 (2021).
19. K. L. Visick, E. V. Stabb, E. G. Ruby, A lasting symbiosis: How *Vibrio fischeri* finds a squid partner and persists within its natural host. *Nat. Rev. Microbiol.* 19, 654–665 (2021).
20. B. L. Fung, J. J. Esin, K. L. Visick, *Vibrio fischeri*: A model for host-associated biofilm formation. *J. Bacteriol.* 206, e0037023 (2024).
21. T. Koropatnick, M. S. Goodson, E. A. Heath-Heckman, M. McFall-Ngai, Identifying the cellular mechanisms of symbiont-induced epithelial morphogenesis in the squid- association. *Biol. Bull.* 226, 56–68 (2014).
22. M. McFall-Ngai, E. A. Heath-Heckman, A. A. Gillette, S. M. Peyer, E. A. Harvie, The secret languages of coevolved symbioses: Insights from the *Euprymna scolopes-Vibrio fischeri* symbiosis. *Semin. Immunol.* 24, 3–8 (2012).
23. M. S. Aschtgen, K. Wetzel, W. Goldman, M. McFall-Ngai, E. Ruby, *Vibrio fischeri*-derived outer membrane vesicles trigger host development. *Cell Microbiol.* 18, 488–499 (2016).
24. J. S. Foster, M. A. Apicella, M. J. McFall-Ngai, *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev. Biol.* 226, 242–254 (2000).
25. C. UniProt, UniProt: The universal protein knowledgebase in 2025. *Nucleic Acids Res.* 53, D609–D617 (2025).
26. T. A. Koropatnick et al., Microbial factor-mediated development in a host-bacterial mutualism. *Science* 306, 1186–1188 (2004).
27. E. S. Yip, B. T. Grublesky, E. A. Hussa, K. L. Visick, A novel, conserved cluster of genes promotes symbiotic colonization and sigma-dependent biofilm formation by *Vibrio fischeri*. *Mol. Microbiol.* 57, 1485–1498 (2005).
28. S. Shibata, E. S. Yip, K. P. Quirke, J. M. Ondrey, K. L. Visick, Roles of the structural symbiosis polysaccharide (*syp*) genes in host colonization, biofilm formation, and polysaccharide biosynthesis in *Vibrio fischeri*. *J. Bacteriol.* 194, 6736–6747 (2012).
29. T. A. Koropatnick, M. S. Goodson, E. A. Heath-Heckman, M. McFall-Ngai, Identifying the cellular mechanisms of symbiont-induced epithelial morphogenesis in the squid-vibrio association. *Biol. Bull.* 226, 56–68 (2014).

30. Y. Hong, M. A. Liu, P. R. Reeves, Progress in our understanding of Wzx flippase for translocation of bacterial membrane lipid-linked oligosaccharide. *J. Bacteriol.* 200, e00154-17 (2018).
31. A. Doino, M. J. McFall-Ngai, A transient exposure to symbiosis-competent bacteria induces light organ morphogenesis in the host squid. *Biol. Bull.* 189, 347–355 (1995).
32. T. Essock-Burns, C. Bongrand, W. E. Goldman, E. G. Ruby, M. J. McFall-Ngai, Interactions of symbiotic partners drive the development of a complex biogeography in the squid-vibrio symbiosis. *mBio* 11,e00853-20 (2020).
33. L. K. Sycuro, E. G. Ruby, M. McFall-Ngai, 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 (2006).
34. T. A. Koropatnick, J. R. Kimbell, M. J. McFall-Ngai, Responses of host hemocytes during the initiation of the squid-vibrio symbiosis. *Biol. Bull.* 212, 29–39 (2007).
35. J. R. Kimbell, T. A. Koropatnick, M. J. McFall-Ngai, Evidence for the participation of the proteasome in symbiont-induced tissue morphogenesis. *Biol. Bull.* 211, 1–6 (2006).
36. K. E. Zink et al., A small molecule coordinates symbiotic behaviors in a host organ. *mBio* 12,e03637-20 (2021).
27. S. Moriano-Gutierrez et al., Critical symbiont signals drive both local and systemic changes in diel and developmental host gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 116, 7990–7999 (2019).
38. J. D. Cecil, N. Sirisaengtaksin, N. M. O'Brien-Simpson, A. M. Krachler, Outer membrane vesicle-host cell interactions. *Microbiol. Spectr.* 7, 7.1.06 (2019).
39. J. M. Bomberger et al., Long-distance delivery of bacterial virulence factors by *Pseudomonas aeruginosa* outer membrane vesicles. *PLoS Pathog.* 5, e1000382 (2009).
40. K. Wang, C. Jia, B. Zhang, J. Chen, J. Zhao, Outer membrane vesicles from commensal microbes contribute to the sponge *Tedania* sp. development by regulating the expression level of apoptosis-inducing factor (AIF). *Commun. Biol.* 7, 952 (2024).
41. A. Woznica et al., Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7894–7899 (2016).
42. S. Chen, Q. Lei, X. Zou, D. Ma, The role and mechanisms of gram-negative bacterial outer membrane vesicles in inflammatory diseases. *Front. Immunol.* 14, 1157813 (2023).
43. R. Stentz, A. L. Carvalho, E. J. Jones, S. R. Carding, Fantastic voyage: The journey of intestinal microbiota-derived microvesicles through the body. *Biochem. Soc. Trans.* 46, 1021–1027 (2018).
44. A. Kaisanlahti et al., Maternal microbiota communicates with the fetus through microbiota-derived extracellular vesicles. *Microbiome* 11, 249 (2023).

45. M. S. Aschtgen et al., Insights into flagellar function and mechanism from the squid-Vibriosisymbiosis. *NPJ Biofilms Microbiomes* 5, 32 (2019).
46. S. Dhital, P. Deo, I. Stuart, T. Naderer, Bacterial outer membrane vesicles and host cell death signaling. *Trends Microbiol.* 29, 1106–1116 (2021).
47. B. M. Coffey, S. S. Akhand, G. G. Anderson, MgtE is a dual-function protein in *Pseudomonas aeruginosa*. *Microbiology (Reading, Engl.)*. 160, 1200–1213 (2014).
48. D. Liu, A. K. Bhunia, Anchorless bacterial moonlighting metabolic enzymes modulate the immune system and contribute to pathogenesis. *ACS Infect. Dis.* 10, 2551–2566 (2024).
49. F. Sorroche et al., The ex planta signal activity of a *Medicago* ribosomal uL2 protein suggests a moon lighting role in controlling secondary rhizobial infection. *PLoS One* 15, e0235446 (2020).
50. K. E. Malter et al., A bacterial membrane-disrupting protein stimulates animal metamorphosis. *mBio* 16, e0357324 (2025).
51. B. Qi, M. Han, Microbial siderophore enterobactin promotes mitochondrial ion uptake and development of the host via interaction with ATP synthase. *Cell* 175, 571–582.e511 (2018).
52. J. Jumper et al., Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589(2021).
53. R. F. Collins, J. P. Derrick, Wza: A new structural paradigm for outer membrane secretory proteins?. *Trends Microbiol.* 15, 96–100 (2007).
54. A. M. Burroughs, S. Balaji, L. M. Iyer, L. Aravind, A novel superfamily containing the beta-grasp foldinvolved in binding diverse soluble ligands. *Biol. Direct* 2, 4 (2007).
55. D. Goudenege et al., Genome sequence of *Vibrio diabolus* and identification of the exopolysaccharide HE800 biosynthesis locus. *Appl. Microbiol. Biotechnol.* 98, 10165–10176(2014).
56. A. S. Vanhove et al., Outer membrane vesicles are vehicles for the delivery of *Vibrio tasmaniensis* virulence factors to oyster immune cells. *Environ. Microbiol.* 17, 1152–1165 (2015).
57. A. C. Hargadon et al., An acidic microenvironment produced by the V-type ATPase of *Euprymna scolopes* promotes specificity during *Vibrio fischeri* recruitment. *Commun. Biol.* 7, 1642 (2024).
58. E. J. Koch, T. Miyashiro, M. J. McFall-Ngai, E. G. Ruby, Features governing symbiont persistence in the squid-vibrio association. *Mol. Ecol.* 23, 1624–1634 (2014).
59. L. M. Naughton, M. J. Mandel, Colonization of *Euprymna scolopes* squid by *Vibrio fischeri*. *J. Vis.Exp.* 10, e3758 (2012).
60. M. K. Montgomery, M. McFall-Ngai, Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. *Development* 120, 1719–1729(1994).
61. L. Yang et al., Bacterial growth dynamics in a rhythmic symbiosis. *Mol. Biol. Cell* 35, ar79 (2024).

62. A. K. Dunn, D. S. Millikan, D. M. Adin, J. L. Bose, E. V. Stabb, New rfp- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and lux expression in situ. *Appl. Environ. Microbiol.* 72, 802–810 (2006).
63. K. L. Visick, K. M. Hodge-Hanson, A. H. Tischler, A. K. Bennett, V. Mastrodomenico, Tools for rapid genetic engineering of *Vibrio fischeri*. *Appl. Environ. Microbiol.* 84, e00850-18 (2018).
64. L. Ferrieres et al., Silent mischief: Bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J. Bacteriol.* 192, 6418–6427 (2010).
65. S. Chen, Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta* 2, e107 (2023).
66. T. F. Rogers et al., Gene modelling and annotation for the Hawaiian bobtail squid. *Euprymna scolopes*. *Sci. Data* 11, 40 (2024).
67. P. Melsted et al., Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat. Biotechnol.* 39, 813–818 (2021).
68. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
69. J. T. Kuwabara, V. Beilinson, Data from “SypC, a symbiont outer membrane vesicle protein, impacts the development of the squid-vibrio partnership”. *Gene Expression Omnibus*. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE307160>. Deposited 3 September 2025.
70. J. T. Kuwabara, V. Beilinson, E.G. Ruby, Data from “SypC, a symbiont outer membrane vesicle protein, impacts the development of the squid-vibrio partnership”. *Gene Expression Omnibus*. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL36208>. Deposited 3 September 2025.
71. V. Beilinson, Data from “SypC, a symbiont outer membrane vesicle protein, impacts the development of the squid-vibrio partnership”. *GitHub*. https://github.com/verabeil/escolopes_vfischeri_sypc. Deposited 25 February 2026.

[Intentionally redacted.]

Chapter 3

SKIN AS AN EARLY DETECTION ORGAN

[This chapter is temporarily embargoed.]

Chapter 4

HOST RESPONSE TO SYMBIONT ON THE SINGLE-CELL LEVEL

[This chapter is temporarily embargoed.]

Chapter 5

BACTERIAL SINGLE-CELL RNA SEQUENCING IN THE SQUID-VBIRIO SYSTEM

[This chapter is temporarily embargoed.]

Chapter 6

CONCLUSIONS

Animal–microbe symbioses are fundamental to the biology of nearly all multicellular organisms. Understanding how these relationships evolve, are established, and are maintained is critical for understanding the health and function of animals, plants, and their associated microbial communities. Disruptions to these interactions can have profound consequences for host development, physiology, and disease susceptibility. The squid–vibrio symbiosis provides a uniquely tractable system for investigating these processes because colonization occurs reproducibly, involves a single dominant symbiont, and elicits well-characterized developmental and physiological changes in the host. In this dissertation, I examined host–microbe interactions across multiple biological scales, from bacterial strain-level variation and tissue-specific responses to the development of single-cell approaches capable of resolving cellular responses to colonization. Examining symbiosis across these biological scales is essential because stable host–microbe partnerships emerge from interconnected processes spanning microbial genetic variation, host tissues, and individual cell types. Only by integrating these levels of organization can we develop a comprehensive understanding of how symbioses are established and maintained.

Collectively, these studies in this thesis demonstrate that host responses to symbiosis are dynamic, distributed, and highly context dependent. In Chapter 2, I showed that different *Vibrio fischeri* strains elicit distinct host transcriptional responses, indicating that symbiotic outcomes are influenced not only by the presence of a symbiont but also by variation within the symbiont species itself. By knocking out genes in the symbiont, we reveal how those specific genes influence the host. Together this highlights that the symbionts' potential function plays a critical role in symbiosis.

[Intentionally redacted.]

Together, these findings support a model in which symbiosis emerges from interactions occurring across multiple levels of biological organization. Hosts respond not only to the presence of microbes but also to variation among strains, signals encountered at external barriers, and the coordinated activities of diverse cell types. Understanding these processes requires integrating perspectives from microbiology, genomics, physiology, and cell biology.

More broadly, the principles revealed by the squid–vibrio symbiosis are likely relevant beyond this specific partnership. Most animals acquire microbial associates from complex environmental communities and must distinguish beneficial organisms from the many microbes they encounter. The mechanisms identified here may represent general features of how hosts establish and maintain beneficial microbial associations. As single-cell and spatial approaches continue to advance, future studies will be able to connect microbial colonization to specific cellular behaviors and intercellular communication networks, ultimately providing a more complete understanding of how animals and microbes build stable partnerships.

A p p e n d i x A

NEW TOOLS FOR GENOME ANNOTATION

A.1 Challenges of annotation in a non-model organism

For transcriptomics, genome annotation files (GTF/GFF files) hold critical information about gene, including location, gene ids, and gene name/description. In non-model organisms, these annotations are often incomplete and continuously evolving. As a result, genes may be absent from annotations, described as “uncharacterized,” or assigned different identifiers across genome versions and transcriptome assemblies. As research progresses in non-model organisms, genome annotations require being updated and refined.

These annotation challenges become even more significant when integrating transcriptomic datasets across studies. Transcriptomic analyses in non-model organisms are often generated using different reference genomes, annotation releases, or de novo transcriptome assemblies. As genome resources improve over time, gene identifiers, transcript structures, and annotation completeness may change substantially between versions. Consequently, the same gene may be represented by different identifiers, renamed between annotation releases, or only exist in certain annotation frameworks. This creates substantial difficulties when comparing differential expression results or performing cross-study analyses, requiring researchers to reconcile multiple annotation systems simultaneously.

To address these challenges, I developed lightweight computational tools designed to facilitate annotation refinement and identifier translation in non-model

transcriptomic datasets. These tools are intended to complement, rather than replace, traditional genome annotation pipelines and RNA-seq re-analysis workflows.

A.2 GenePatch

Can be found here: <https://github.com/verabeil/GenePatch>

A.2.1 Motivation

As research progresses in non-model organisms, new genes might be found or existing genes may need an update to their description. In many cases, genes are labeled as “uncharacterized” or “hypothetical” within genome annotation files. Upon additional and experimental work, these genes may eventually need to be updated.

For example, in *Euprymna scolopes*, the galaxin gene family has been studied extensively in the context of the squid–vibrio symbiosis. However, when locating these genes within the published genome annotation from NCBI, these genes are not present or not labeled with their names from the literature (EsGal 1 – EsGal 11). For researchers specifically studying these galaxin genes, it would be advantageous for the annotation itself to contain the biologically relevant gene names (EsGal 1 – EsGal 11) so that this information is propagated through downstream transcriptomic pipelines and differential expression analyses.

To facilitate incorporation of external sequence evidence into genome annotations, as well as refinement of existing gene names, I developed GenePatch, a easy-to-use tool for mapping user-provided sequences directly to a reference genome and

updating existing annotations. GenePatch was designed to allow researchers to provide sequences of interest and update annotations.

A.2.2 Overview of the workflow

GenePatch is a command-line tool that curates genome annotations using known gene sequences. The tool takes three main input files: a reference genome FASTA file, an existing genome annotation file in GTF format, and a query FASTA file containing one or more gene sequences. For each query sequence, GenePatch aligns the sequence to the reference genome using minimap2, displays the top genomic alignment hits, and reports whether each hit overlaps an existing annotated gene.

After alignment, the user can select which genomic hit to use. If the selected hit overlaps an existing gene, GenePatch allows the user to update the gene description while preserving the original gene identifier. For example, a gene previously labeled as “uncharacterized” can be updated to include a biologically meaningful name, such as a galaxin gene name. If the selected hit does not overlap an existing gene,

GenePatch can add a new gene annotation to the GTF file. Newly added annotations consist of a simple gene model containing a gene, transcript, and exon entry.

The general workflow consists of the following steps:

1. Input a reference genome FASTA file, genome annotation GTF file, and query FASTA file containing known gene sequences.
2. Align each query sequence to the reference genome using minimap2¹.
3. Display the top genomic alignment hits, including alignment location and overlap with existing annotated genes.
4. Allow the user to select the appropriate genomic hit.
5. If the selected hit overlaps an existing gene, optionally update the gene description.
6. If the selected hit does not overlap an existing gene, optionally add a new gene annotation to the GTF file.

Export an updated GTF file and a report summarizing the changes.

This approach provides a user-friendly method for improving genome annotations by either refining existing gene descriptions or incorporating previously missing genes into the annotation. By integrating biologically relevant gene names directly into the annotation framework, downstream transcriptomic analyses become easier to interpret and more biologically informative.

A.2.3 Applications and limitations

GenePatch was used to facilitate annotation refinement in *E. scolopes*, including identification and renaming of galaxin-related genes. More broadly, the tool can be

applied to incorporate candidate genes, improve interpretability of transcriptomic datasets, and refine evolving annotations in non-model systems.

Importantly, GenePatch is not intended to replace comprehensive genome annotation pipelines. Instead, it provides a modular approach for incorporating external biological evidence into existing annotations. Likewise, updating annotation names does not substitute for remapping RNA-seq reads to an improved reference genome when major structural changes occur.

The accuracy of the approach also depends on sequence quality, alignment specificity, and the quality of the underlying reference genome assembly.

A.3 map_query_to_target_genes

Can be found here: https://github.com/verabeil/map_query_to_target_genes

A.3.1 Motivation

Another major challenge in non-model transcriptomics is the instability of gene identifiers across genome versions, annotation releases, and transcriptome assemblies. Differential expression analyses performed using different references may produce gene identifiers that are not directly comparable, making integration of datasets difficult.

As genome and annotations evolve, datasets for publication may have been generated using different annotation frameworks or transcriptome assemblies. In the squid–vibrio system, transcriptomic datasets were generated using multiple genome

versions as well as independently generated transcriptome assemblies, complicating direct comparison of gene-level results.

To facilitate cross-study integration and reproducibility, I developed `map_query_to_target_genes`, a practical and easy to use tool for translating identifiers between annotation frameworks.

A.3.2 Mapping strategy

`map_query_to_target_genes` is a command-line tool designed to facilitate translation of gene or transcript identifiers between different annotation frameworks. The tool was developed to improve interoperability between transcriptomic datasets generated using different genome annotations, transcriptome assemblies, or genome versions.

The workflow takes as input a query annotation set and a target annotation set, then identifies corresponding genes or transcripts between the two references. Query transcript sequences are first extracted from the query genome and annotation using `gffread`². These transcript sequences are then aligned to the target genome using `minimap2`¹ with splice-aware alignment settings.

Following alignment, the genomic overlap between aligned query transcripts and target genome annotations is evaluated. Exon-level overlap between the aligned query transcript and annotated target genes is calculated to identify the most likely

corresponding target gene. The software then generates mapping tables linking query identifiers to corresponding target identifiers.

The general workflow consists of the following steps:

1. Extract transcript sequences from the query genome and annotation using gffread².
2. Align query transcripts to the target genome using minimap2¹ splice-aware alignment.
3. Evaluate overlap between aligned query transcripts and target gene annotations.
4. Calculate exon overlap statistics between query transcripts and candidate target genes.
5. Assign each query transcript to the best matching target gene based on overlap metrics.
6. Generate mapping tables linking query identifiers to corresponding target identifiers.
7. Export mapping statistics and reports in tabular formats suitable for downstream transcriptomic analyses. The output is a table with the query gene id and the best matching target gene id if present.

To improve mapping confidence, the workflow incorporates multiple filtering criteria, including minimum alignment quality, minimum query transcript coverage, minimum exon overlap thresholds, and comparison between the best and second-best

overlapping target genes. Query transcripts with weak or ambiguous overlap can therefore be identified and flagged separately.

This approach provides a user-friendly method for reconciling transcriptomic datasets generated using evolving annotation resources. For example, the tool can be used to map identifiers between older and newer genome versions, de novo transcriptome assemblies and genome annotations, or independently generated transcriptomic datasets. This enables researchers to identify genes of interest from older studies and determine whether corresponding genes are present in newer datasets without requiring complete re-analysis or remapping of the original sequencing data. By facilitating identifier translation across annotation frameworks, the tool improves reproducibility and simplifies cross-study comparison in non-model transcriptomic analyses.

A.3.3 Applications and limitations

`map_query_to_target_genes` was used to facilitate comparison of transcriptomic datasets generated using different *E. scolopes* genome annotations and transcriptome assemblies. A gene of interest from an early genome was easily found in a newer genome version which has a new gene id.

The tool enabled translation of identifiers across annotation frameworks, improving reproducibility and facilitating comparison of conserved transcriptional responses across independent studies.

However, identifier mapping should not be considered equivalent to full RNA-seq re-analysis using a unified reference genome. Mapping approaches cannot fully account for differences in transcript structure, annotation completeness, or genome assembly quality. Instead, the tool is intended to provide a practical strategy for

improving interoperability between datasets generated using different annotation resources.

A.4 Broader implications for non-model transcriptomics

Together, these tools reflect a broader reality of transcriptomic analysis in non-model organisms: genome annotations are dynamic resources that evolve alongside improved assemblies, experimental evidence, and community-driven annotation efforts.

Unlike annotations in highly curated model systems, non-model annotations should not be viewed as static or definitive. Instead, transcriptomic interpretation in non-model organisms often requires iterative refinement, integration of external biological evidence, and reconciliation across multiple annotation frameworks.

As sequencing technologies continue to improve and additional transcriptomic datasets become available, annotation resources for non-model organisms will continue to evolve. Flexible computational approaches that facilitate annotation refinement and cross-study integration will therefore remain important components of reproducible transcriptomic analysis.

REFERENCES

1. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572–4574 (2021).
2. Perte, G. & Perte, M. GFF Utilities: GffRead and GffCompare. *F1000Research* **9**, 304 (2020)

Appendix B

ROLE OF HOST EXOSOMES IN SYMBIOSIS

[Intentionally redacted.]

