

Cover page

Article type: Invited review

Author names and affiliations

- Tom Baden, <https://orcid.org/0000-0003-2808-4210>, t.baden@sussex.ac.uk
 - School of Life Sciences, University of Sussex, Brighton, UK
- John Briseño, <https://orcid.org/0000-0002-2702-5078>, john.briseno@uconn.edu
 - Molecular and Cell Biology Department, University of Connecticut, Storrs, CT 06269, USA
- Gabrielle Coffing <https://orcid.org/0000-0001-8546-6595> gcoffing@uoregon.edu
 - Biology Department: Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA
- Sophie Cohen-Bodénès, <https://orcid.org/0000-0001-7924-3942>, sophie.bodenes@gmail.com
 - Laboratoire des Systèmes Perceptifs, Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL University, CNRS, 75005 Paris, France
- Amy Courtney <https://orcid.org/0000-0003-0495-1711> acourtney@mrc-lmb.cam.ac.uk
 - MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, UK
- Dominick Dickerson dickerson.dominick@gmail.com
 - Friday Harbor Laboratory
- Gül Dölen (gul@jhu.edu)
 - Johns Hopkins University, Department of Neuroscience
- Graziano Fiorito, <https://orcid.org/0000-0003-2926-9479>, graziano.fiorito@szn.it
 - Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy
- Camino Gestal <https://orcid.org/0000-0003-1931-9567> cgestal@iim.csic.es
 - Laboratory of Marine Molecular Pathobiology, Institute of Marine Research (IIM), Spanish National Research Council (CSIC), Vigo, Spain
- Taryn Gustafson, taryngustafson@knights.ucf.edu
 - University of Central Florida, Orlando, FL, US
- Elizabeth Heath-Heckman, <https://orcid.org/0000-0001-5265-9032> each@msu.edu
 - Departments of Integrative Biology and Microbiology and Molecular Genetics, Michigan State University, East Lansing MI, USA
- Qiaz Hua <https://orcid.org/0000-0002-3128-7273>
 - Department of Ecology and Evolution, University of Adelaide, Adelaide, South Australia 5000, Australia
- Pamela Imperadore, <https://orcid.org/0000-0002-9947-0971>, pamela.imperadore@szn.it
 - Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy
- Ryosuke Kimbara, <https://orcid.org/0000-0003-3328-0638>, kimbara-ryosuke32@mmbs.s.u-tokyo.ac.jp
 - Misaki Marine Biological Station, School of Science, The University of Tokyo, Miura, Kanagawa, 238-0225, Japan.
- Mirela Król mirkro@st.amu.edu.pl
 - Adam Mickiewicz University in Poznań (Poznań, Poland)
- Zdeněk Lajbner <https://orcid.org/0000-0001-7528-5408>, lajbner@oist.jp
 - Physics and Biology Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Kunigami District, Okinawa 904-0495, Japan
- Nicolás Lichilín <https://orcid.org/0000-0002-3674-9561>
 - Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, 1010, Austria
- Filippo Macchi, <https://orcid.org/0000-0001-8226-4816>, fm1442@nyu.edu
 - Program in Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
- Matthew J. McCoy <https://orcid.org/0000-0002-8250-0564>

- Department of Pathology, Stanford University, Stanford, CA 94305, USA
- Michele K. Nishiguchi <https://orcid.org/0000-0002-8950-4621> nish@ucmerced.edu
 - Department of Molecular and Cell Biology, School of Natural Sciences, University of California, Merced, 5200 N. Lake Blvd., Merced, CA 95343 USA
- Spencer V. Nyholm spencer.nyholm@uconn.edu
 - Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA
- Eve Otjacques, <https://orcid.org/0000-0002-0880-9360> (ebotjacques@fc.ul.pt)
 - MARE—Marine and Environmental Sciences Centre & ARNET—Aquatic Research Network, Laboratório Marítimo da Guia, Faculdade de Ciências, Universidade de Lisboa, Av. Nossa Senhora do Cabo, 939, 2750-374 Cascais, Portugal
 - Division of Biosphere Sciences and Engineering, Carnegie Institution for Science, 1200 E. California Blvd, Pasadena, CA 91125, USA
- Pedro Antonio Pérez-Ferrer <https://orcid.org/0000-0002-5831-5078> pperez40@ucmerced.edu
 - Department of Molecular and Cell Biology, School of Natural Sciences, University of California, Merced, 5200 N. Lake Blvd., Merced, CA 95343 USA
- Giovanna Ponte, <https://orcid.org/0000-0002-1779-958x>, giovanna.ponte@szn.it
 - Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy
- Judit R. Pungor <https://orcid.org/0000-0002-3858-9442> jpungor@uoregon.edu
 - Biology Department: Institute of Neuroscience, University of Oregon, Eugene, OR, USA
- Thea F. Rogers <https://orcid.org/0000-0003-0198-0242> thea.rogers@univie.ac.at
 - Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, 1010, Austria
- Joshua J. C. Rosenthal [orcid] rosenthal.joshua@gmail.com
 - The Eugene Bell Center for Regenerative Biology and Tissue Engineering, Marine Biological Laboratory, Woods Hole, MA, USA.
- Lisa Rouressol <https://orcid.org/0000-0003-4918-0711>, lisa.rouressol@univie.ac.at
 - Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, 1010, Austria
- Noelle Rubas <https://orcid.org/0000-0002-1967-1352> nrubas@hawaii.edu
 - Department of Molecular Biosciences and Bioengineering, University of Hawaii Manoa, Honolulu HI, 96822
- Gustavo Sanchez, <https://orcid.org/0000-0002-6490-2842>, gus.sanchez.s89@gmail.com
 - Molecular Genetics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan.
- Catarina Pereira Santos cdcsantos@fc.ul.pt
 - University of Lisbon, Portugal
- Darrin T. Schultz <https://orcid.org/0000-0003-1190-1122> darrin.schultz@univie.ac.at
 - Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, 1010, Austria
- Eve Seuntjens, <https://orcid.org/0000-0002-0126-461X>, eve.seuntjens@kuleuven.be
 - Laboratory of Developmental Neurobiology, Department of Biology, KU Leuven, Leuven, Belgium
- Jeremea O. Songco-Casey, <https://orcid.org/0000-0003-3118-6111>, jsongco@uoregon.edu
 - Biology Department: Institute of Neuroscience, University of Oregon, Eugene, OR, USA
- Ian Erik Stewart, <https://orcid.org/0000-0001-5307-1564>, ian.stewart@mhc-berlin.de
 - Neural Circuits and Behaviour Lab, Max-Delbrück Ctr. for Mol. Med. (MDC), Berlin, Germany / Institute for Biology, Freie Universität Berlin
- Ruth Styfhals, <https://orcid.org/0000-0002-5287-8830>, ruth.styfhals@kuleuven.be
 - Laboratory of Developmental Neurobiology, Department of Biology, KU Leuven, Leuven, Belgium
- Surangkana Tuanapaya, <https://orcid.org/0000-0001-5004-5846>,

suangkana.tua@gmail.com

- Laboratory of genetics and applied breeding of molluscs, Fisheries College, Ocean University of China, China.
- Nidhi Vijayan <https://orcid.org/0000-0002-0615-9587>, nidhi.vijayan@uconn.edu
 - Molecular and Cell Biology Department, University of Connecticut, Storrs, CT 06269, USA
- Anton Weissenbacher a.weissenbacher@zoovienna.at
 - Vienna Zoo, Maxingstraße 13b, 1130 Vienna, Austria
- Lucia Zifcakova <https://orcid.org/0000-0002-8037-2116> lucia.zifcakova@oist.jp
 - Physics and Biology Unit (Jonathan Miller), Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Kunigami District, Okinawa 904-0495, Japan
- Grace Schulz NGSchulz@uchicago.edu
 - University of Chicago
- Willem Weertman willemlw@uw.edu
 - University of Washington, Friday Harbor Laboratories
- Oleg Simakov <https://orcid.org/0000-0002-3585-4511> oleg.simakov@univie.ac.at
 - Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, 1010, Austria
- Caroline Albertin, <https://orcid.org/0000-0003-0933-7665>, calbertin@mbl.edu
 - Eugene Bell Center for Tissue Engineering and Regenerative Biology, Marine Biological Laboratory, Woods Hole, MA 02543, USA

Correspondence:

Oleg Simakov (oleg.simakov@univie.ac.at)

Caroline Albertin (calbertin@mbl.edu)

Keywords:

cephalopods, omics, evolution, neurobiology, transgenics

Abstract:

Few animal groups can claim the level of wonder that cephalopods instill in the minds of researchers and the general public. Much of cephalopod biology, however, remains unexplored: the largest invertebrate brain, difficult husbandry conditions, complex (meta-)genomes, among many other things, have hindered progress in addressing key questions. However, recent technological advancements in sequencing, imaging, and genetic manipulation have opened new avenues for exploring the biology of these extraordinary animals. The cephalopod molecular biology community is thus experiencing a large influx of researchers, emerging from different fields, accelerating the pace of research in this clade. In the first post-pandemic event at the Cephalopod International Advisory Council (CIAC) conference in April 2022, over 40 participants from all over the world met and discussed key challenges and perspectives for current cephalopod molecular biology and evolution. Our particular focus was on the fields of comparative and regulatory genomics, gene manipulation, single cell transcriptomics, metagenomics and microbial interactions. This article is a result of this joint effort, summarizing the latest insights from these emerging fields, their bottlenecks and potential solutions. The article highlights the interdisciplinary nature of the cephalopod -omics community and provides an emphasis on continuous consolidation of efforts and collaboration in this rapidly evolving field.

Title:

Cephalopod-omics: emerging fields and technologies in cephalopod biology

Running title: The state of Cephalopod-omics

Author list: see submission forms

Corresponding authors: Oleg Simakov (oleg.simakov@univie.ac.at), Caroline Albertin (calbertin@mbl.edu)

Abstract

Few animal groups can claim the level of wonder that cephalopods instill in the minds of researchers and the general public. Much of cephalopod biology, however, remains unexplored: the largest invertebrate brain, difficult husbandry conditions, complex (meta-)genomes, among many other things, have hindered progress in addressing key questions. However, recent technological advancements in sequencing, imaging, and genetic manipulation have opened new avenues for exploring the biology of these extraordinary animals. The cephalopod molecular biology community is thus experiencing a large influx of researchers, emerging from different fields, accelerating the pace of research in this clade. In the first post-pandemic event at the Cephalopod International Advisory Council (CIAC) conference in April 2022, over 40 participants from all over the world met and discussed key challenges and perspectives for current cephalopod molecular biology and evolution. Our particular focus was on the fields of comparative and regulatory genomics, gene manipulation, single cell transcriptomics, metagenomics and microbial interactions. This article is a result of this joint effort, summarizing the latest insights from these emerging fields, their bottlenecks and potential solutions. The article highlights the interdisciplinary nature of the cephalopod -omics community and provides an emphasis on continuous consolidation of efforts and collaboration in this rapidly evolving field.

Introduction

Cephalopods (squid, octopus, cuttlefish, nautilus) are a molluscan class that is subdivided into the coleoid cephalopods (squid, octopus, cuttlefish) and nautiloids (e.g. *Nautilus*). Ten years have passed since the publication of the first cephalopod genomics white paper (Albertin *et al.* 2012). While it still took a few years until the first cephalopod genome was released (Albertin *et al.* 2015), cheap sequencing costs and improvement of assembly algorithms have since resulted in an accelerated pace of cephalopod genomics studies (Albertin and Simakov 2020). Furthermore, technological advances have allowed for the development of new fields in cephalopod molecular biology: from unraveling aspects of gene regulation (Schmidbaur *et al.* 2022) to single cell transcriptomics (Duruz *et al.* 2022; Gavriouchkina *et al.* 2022; Songco-Casey *et al.* 2022; Styhals *et al.* 2022) to targeted gene manipulation (Crawford *et al.* 2020). This summary of the state-of-the-art and future direction in cephalopod-omics is a result of a collaboration of more than 40 authors who met at a workshop at the Cephalopod International Advisory Council (CIAC) conference in Portugal in 2022, preceded by online meetings within the CephRes Virtual Event in 2020. Our goal was to identify major developments in the field and emerging bottlenecks, and outline major ongoing directions as well as possible solutions to arising problems. We focused on four major topics that were discussed at the workshop,

spanning comparative and regulatory genomics, transgenics, gene expression and single cell transcriptomics, and metagenomics with the results and views of each of the working groups presented below. These topics emerged from group discussions focused on outlining the current state of the field, as well as developing emerging and future approaches for use in studying cephalopod biology. Being a small subset of the whole community and due to space limitations we cannot cover all possible topics of cephalopod molecular research or go in too much depth even in the selected set of the research areas. We thus apologize to colleagues whose work we could not represent sufficiently here and for this we refer the reader to the more specialized review articles.

Resources and emerging new standards for cephalopod genomics

Sequencing strategy proposal

The diversity and quality of published cephalopod genome assemblies have increased considerably in the seven years since the publication of the first cephalopod genome, that of *Octopus bimaculoides* (Albertin *et al.* 2015). The accelerated pace of chromosomal-scale assembly and chromosomal-level genomic comparisons allows for new topics to be explored. Major emerging topics include: chromosomal evolution (Albertin *et al.* 2022), first insights into regulatory regions (Schmidbaur *et al.* 2022), gene architecture and non-coding element evolution (McCoy and Fire 2020; Heath-Heckman and Nishiguchi 2021; Marino *et al.* 2022), phylogenetics (Tanner *et al.* 2017; Sanchez *et al.* 2021), among others. These studies have so far shown that coleoid genomes are typically large, often spanning ~3-6Gb across 30 chromosomes for octopuses and ~46 for squids, while many molluscan genomes are less than 2Gb (genomesizes.com). While the number of genes in cephalopods are similar to many other invertebrates, the non-coding genome appears to be significantly larger (Albertin *et al.* 2015; McCoy and Fire 2020), with at least half of a typical cephalopod genome composed of repetitive elements (Marino *et al.* 2022) and increased number of miRNAs (Zolotarov *et al.* 2022). The large and highly repetitive nature of these genomes necessitates rigorous approaches in genome assembly and annotation similar to what is required for many vertebrate genomes.

Many genome sequencing consortia have developed standardized approaches to genome sequencing and assembly (Rhie *et al.* 2021). The current “gold standard” of cephalopod genomics is long read (PacBio, Oxford Nanopore) sequencing of around 25-40x coverage using HiFi (High-Fidelity) or at least 50x CLR (Continuous Long Reads) (McKenna *et al.* 2021). The first major community goal is thus to produce chromosome-scale and phased genome assemblies of cephalopods, which provides haplotype information to help untangle relationships between DNA sequence and phenotype. Resolving the genomes to chromosome-scale and phasing the haplotypes additionally requires chromosomal conformational data (Hi-C), or preferably the more recent derivative DNase Hi-C (Ma *et al.* 2015), from the same individual from which the long reads are derived. Recent algorithmic advancements integrate high-fidelity PacBio reads (HiFi), ultra-long nanopore reads, and phasing information from Hi-C libraries to generate telomere-to-telomere assemblies and should be considered for cephalopods (Cheng *et al.* 2021; Nurk *et al.* 2022). Alternatively, inbred individuals could be sequenced to decrease the assembly graph complexity introduced by the frequent heterozygous sites in cephalopod genomes (Albertin *et al.* 2015).

Towards high-quality genome annotations

The second major goal is producing high quality genome annotations. Standard annotation pipelines, such as MAKER (Cantarel *et al.* 2008), AUGUSTUS (Hoff and Stanke 2019) or BRAKER (Brůna *et al.* 2021), produce sufficient quality gene models for basic orthology and evolutionary studies. However, the use of standard Illumina RNA-seq reads can lead to several problems in genome annotations, including the lack of 3' and 5' UTRs, fragmented or fused gene models, and poor gene calling ability in repeat-rich regions. This can be particularly problematic for read mapping in single-cell RNA-seq data, which is commonly 3' biased (see below). Contemporary genome annotation pipelines, such as MAKER with ProtHint (Brůna, Lomsadze and Borodovsky 2020) incorporation, can facilitate the annotation process by using protein orthology information and full-length transcript sequences from PacBio Iso-Seq or Oxford Nanopore Technologies R2C2. However, we expect that the large cephalopod genomes will require further manual curation to correct inappropriately fused or fragmented gene models.

To facilitate gene annotation of the many emerging cephalopod systems, the community goal is to generate high quality annotations for a particular subset of species. There are several considerations for selecting among current coleoid cephalopods, including available resources, individual community sizes, and long-term outlook for transgenics. Ideally, these efforts should be distributed across the two major clades of coleoids: the Decapodiformes (the clade including bobtails, squid, and cuttlefish), such as *Euprymna scolopes* and *Euprymna berryi*, *Sepioteuthis lessona*, and the Octopodiformes (vampire squid, octopuses and their kin), such as: *Octopus vulgaris* and *O. bimaculoides* (Albertin and Simakov 2020). This will allow the community to produce a high quality set of curated orthologs - which we propose to term CephUSCO (single copy orthologs) for curation of follow-up models, similar to efforts completed in other clades, e.g., mollusks (Caurcel *et al.* 2021). Additionally, standardized gene naming is crucial for effective communication about genes (Bruford *et al.* 2020). Presently, we suggest that the cephalopod community follow the naming convention proposed for human genes (<https://vertebrate.genenames.org/about/>, the HUGO Gene Nomenclature Committee (HGNC), while the efforts to characterize the function of cephalopod genes are ongoing.

These recent advancements have made sequencing and assembling cephalopod genomes to chromosome-scale, then annotating them, easier for the growing cephalopod research community (e.g., Destanović *et al.* 2023). The increased amount of community genomic resources will help with the continued development of nascent model systems for molecular research (see below) as well as the investigation of poorly studied or rare cephalopod species (McKenna *et al.* 2021; Vecchione, Sweeney and Rothman 2022).

Towards cephalopod transgenics and regulatory genomics

The historical success and persisting popularity of the major invertebrate model organisms, *C. elegans* and *D. melanogaster*, is due, in part, to their genetic tractability. A genetically tractable organism is a research organism amenable to genetic modifications, i.e. the genetic makeup of the organism is altered to achieve a desired outcome. Genetic tractability is essential as it allows researchers to functionally test correlative genotype-phenotype observations and opens

up the possibility of creating tools, like reporter lines, that will enable new areas of study. Well established methods in other model organisms enable us to remove (knock-out) or insert (knock-in) specific genomic regions, or cause a downregulation (knock-down) or upregulation of expression (over-expression). Knocking out a gene of interest and observing phenotypic changes allows us to understand the role that gene plays. Knocking in fluorescent proteins under cell type specific promoters enables us to visualize expression patterns on a cellular or protein level. For neurobiological studies, cell type-specific knock-in of neural activity reporters (e.g. fluorescent calcium indicators like GCaMP) allows us to monitor the activation of specific neurons even while the animal is behaving (Broussard, Liang and Tian 2014). Progress applying these tools in other emerging model organisms such as the cnidarian *Hydra vulgaris*, the acoel *Hofstenia miamia*, the hemichordate *Saccoglossus kowalevskii* (Minor *et al.* 2019), and crustaceans like *Parhyale hawaiiensis*, was made possible after sequencing breakthroughs and development of transgenic lines (Pavlopoulos and Averof 2005; Badhiwala *et al.* 2021; Ricci and Srivastava 2021). Developing a genetically tractable organism requires a species that can be easily cultured in the lab, has high quality genomic and transcriptomic data, and has reliable transgene injection techniques. The following sections will discuss the progress the cephalopod community has made in these respects and what is still required to enable a fully genetic tractable cephalopod species.

Goals and challenges for gene manipulation in cephalopods

While gene manipulation tools have been developed in a number of emerging model systems in many animal clades, these methods are only recently becoming available in cephalopods. Thus far, CRISPR-mediated knockouts have only been demonstrated for *Doryteuthis pealeii* (Crawford *et al.* 2020), and transgenics have yet to be reported in any cephalopod. However, efforts are underway to develop transgenics in all of the genome-enabled systems (see above), in particular for the Japanese bobtail *Euprymna berryi* (Ahuja *et al.* 2023), the Hawaiian bobtail squid *Euprymna scolopes*, as well as the stumpy cuttlefish *Sepia bandensis*, although other species may become readily available in the future as well (see below). In general, major technical obstacles to achieve transgenesis include introducing reagents into a cell, identifying robust and reproducible methods for integrating exogenous DNA, and driving transgene expression (transiently or integrated). These barriers have been overcome in several other emerging animal models (Pavlopoulos and Averof 2005; Minor *et al.* 2019; Badhiwala *et al.* 2021; Ricci and Srivastava 2021), and incorporating approaches and reagents that have been successful in these other species may accelerate the development of these technologies in cephalopods.

Delivery timing and injection techniques are species-specific

A number of methods have been established in other model organisms to introduce reagents into a cell, including microinjection, electroporation, transfection, and viral vectors. Thus far, the only published method for introducing reagents into cephalopod cells is microinjection (Crawford *et al.* 2020). This method requires routine access to early cleavage stage embryos and means of accessing the early blastomeres. Delivery techniques for CRISPR guides vary depending on the species. Naturally laid decapodiform embryos typically have a number of extraembryonic structures protecting the developing embryo, including a chorion, jelly layers, and an outer tunic.

These structures differ between different groups, and each presents a barrier to accessing embryonic cells. For example, naturally laid *Doryteuthis* embryos have tough chorion surrounded by jelly layers and an outer tunic that joins embryos into “fingers”. However, *in vitro* fertilized *Doryteuthis* embryos lack the outer jelly layers and tunic (Crawford 2002; Crawford *et al.* 2020), allowing timed experiments. However, they do have a thick chorion that needs to be cut in order to pass a microinjection needle into the early blastomeres.

Injection techniques have also been developed for multiple *Euprymna* species using naturally laid embryos (Ahuja *et al.* 2023). The *Euprymna* chorion can be pierced directly with the micropipette once the jelly layers are manually removed. While *in vitro* fertilization is not currently established in *Euprymna*, first cleavage takes place 8 hours after fertilization, giving researchers a large window in which these embryos can be injected. Attempts to develop microinjection techniques have also been extended to several other cephalopod clades. While squid and cuttlefish embryos have similar extraembryonic structures (chorion, jelly layers, tunic) (Lee, Callaerts and de Couet 2009; Montague, Rieth and Axel 2021) suggesting that microinjection should be broadly possible in this group, the extraembryonic structures in octopuses are quite different (Shigeno *et al.* 2015). Octopus embryos lack the jelly layers and tunic; instead they have a very tough, leathery chorion, which may suggest that other approaches might be needed to deliver molecular reagents into cells in this group.

CRISPRopods: creating knockouts and transgenics

Recent advances in CRISPR-based approaches have been applied in two different cephalopods (*D. pealeii* and *E. berryi*) to create knockout mutations. Two to three guides are designed per target using freely available programs such as CRISPRscan (Moreno-Mateos *et al.* 2015) or CHOPCHOP (Labun *et al.* 2019), and injected into early blastomeres, as described above. When injected into the first cell, over 95% of alleles demonstrate a mutation in injected *D. pealeii* hatchlings, as shown by low-coverage amplicon sequencing (Crawford *et al.* 2020). G0 knockouts have been generated for species without a closed life cycle (e.g., *Doryteuthis pealeii*), which excludes the possibility of generating knockout lines. However, *Euprymna berryi* has been shown to have germline transmission and several knockout lines have been established (Ahuja *et al.* 2023).

An important advance for the study of many biological questions will be the generation of transgenic or “knock-in” cephalopods, where a gene is introduced into the genome, such as a fluorescent reporter. There are several methods for inserting DNA into the genome of a research organism, including Homology Directed Repair (HDR) mediated via CRISPR-Cas (Sander and Joung 2014), as well as random integration via restriction enzymes or transposases. While HDR via CRISPR offers precision in targeting the transgene to a particular locus, and can enable the exogenous gene to be expressed using the endogenous regulatory elements, the efficiency of integration is often low, and can be greatly influenced by the size of the insert. By contrast, integration via transposases, such as *tol2* (Kawakami and Shima 1999), or restriction enzymes such as I-SceI meganuclease (Thermes *et al.* 2002), can be fairly efficient, but, as the integration of these constructs is random, the sequence for regulatory elements to drive expression must also be included. The production of transgenic cephalopods

is an area of active research for several labs using different species and approaches, and will surely usher in exciting new insights into cephalopod biology.

Target selection

A key first step in these molecular approaches is to identify target sequences for genetic manipulation. Annotation of protein coding sequences in the genome assemblies have provided a crucial first step, with resources currently being developed to identify regulatory elements. Several resources are now available to the cephalopod community to aid in this task, including long-read transcriptome sequencing for UTR detection, as well as ATAC-seq (Buenrostro *et al.* 2015) and DNA methylation (Fraga and Esteller 2002; de Mendoza *et al.* 2021; Macchi, Edsinger and Sadler 2022) methods for putative regulatory region annotation. Recent studies also employed HiC methods to detect genome topology (Schmidbaur *et al.* 2022). Ongoing community efforts include development of ChIP-seq methods and micro-C high resolution HiC protocols. Additionally, several computational tools are available to predict regulatory elements based on databases of transcription factor (TF) binding profiles, such as Jaspár (Castro-Mondragon *et al.* 2022). These currently do not include mollusks or cephalopods, but comparative genomic data combined with published ATAC-seq and methylation data can be used to begin generating such resources. Further tools could be built to identify TF binding sites from gene expression lists for cephalopods (e.g. cistarget, Herrmann *et al.* 2012). Despite the availability of such tools and data, the bulk of the ongoing analyses are still performed on a gene-by-gene basis involving manual curation and annotation. Combined RNA and ATACseq allows for identification of specific enhancers (González-Blas *et al.* 2022) in specific tissues or cells and several methods, including scATAC-seq are being developed for cephalopods. Additionally, using high resolution topological information (micro-C) and single cell data resolution, we may be able to disentangle gene regulation and identify specific regulatory region - target gene associations.

Advances in gene expression and single cell transcriptomics

Transcriptomics and spatial expression profile methods have been successfully applied to numerous emerging model organisms and hold great promise for the identification of genomic determinants putatively involved (and testable via transgenics) in anatomical novelties. While these tools have proven to be incredibly powerful in some species, their transfer to cephalopods has often been nontrivial. Below we highlight the current state of cephalopod transcriptomics including both the use of expression profiling, such as bulk RNA sequencing, single-cell RNA-sequencing, and expression visualization, such as *in situ* hybridization and hybridization chain reaction (HCR). Together, these methods provide a more comprehensive view of cell types.

Expression profiling

With a growing number of resources to explore cephalopod gene expression and regulatory elements (Schmidbaur *et al.* 2022, Crawford *et al.* 2020), uncovering the genetic bases underlying a cephalopod's unique biological functions becomes increasingly feasible. Transcriptome sampling from diverse tissue types across cephalopod taxa ranges from neural systems (Albertin *et al.* 2015, 2022; Alon *et al.* 2015; Koenig *et al.* 2016; Liscovitch-Brauer *et al.*

2017) to microbe-associated, host-provided symbiotic organs (Castellanos-Martínez *et al.* 2014; Moriano-Gutierrez *et al.* 2019; Koch *et al.* 2020). Importantly, common RNA quality control metrics such as the RNA Integrity Numbers (RIN) cannot be used to assess cephalopod RNA quality. Mollusks (and most protostomes) possess a “hidden break”, where the 28S rRNA is cleaved into two similar-sized molecules, which leads to poor automatically assigned RIN values and a misinterpretation of the quality of the RNA (Natsidis *et al.* 2019). Sufficient cephalopod RNA integrity for next-generation sequencing should therefore be evaluated based on the presence of clear, narrow peak(s) and the absence of smearing.

Recent advances have enabled researchers to profile the transcriptomes of individual cells at a large scale. Single-cell RNA sequencing (e.g., 10x Genomics) has the potential to shed light on the molecular profiles of cephalopod cell types and to identify cell type-specific markers. The ability to distinguish cell types based on more than their morphology is an important step forward in the field and will facilitate the creation of specific transgenic reporter lines. Recently, cell type diversity has been described in the embryonic squid head (Duruz *et al.* 2022), the paralarval octopus brain (Styfhals *et al.* 2022), the octopus visual system (Songco-Casey *et al.* 2022), and bobtail squid visual system (Gavriouchkina *et al.* 2022). The commonly used 10X Genomics Chromium Single Cell 3' Kit, v3 chemistry is 3' biased such that the majority of reads are not counted when the genome annotation is suboptimal. Improving the annotation of the 3' untranslated regions of current and future cephalopod genomes will lead to better mapping statistics and higher estimated numbers of sequenced cells (Lawson *et al.* 2020; Songco-Casey *et al.* 2022; Styfhals *et al.* 2022). Optimizing dissociation and culturing conditions for cephalopod cells will be important to improve cell viability and potentially library preparation. This is especially important, as many single cell sequencing protocols have been developed for terrestrial and freshwater species, but do not account for the sensitivity of marine cells to changes in salinity. ScRNAseq studies in other marine organisms resuspend the cells in calcium/magnesium-free artificial seawater before proceeding to 10X Genomics (Cao *et al.* 2019; Sharma, Wang and Stolfi 2019; Paganos *et al.* 2021). To circumvent potential problems relating to dissociation, viability and sampling issues, single nuclei RNA sequencing provides a sound alternative.

Expression visualization

The ability to localize RNA and proteins in cephalopods is an important tool in cephalopod genomics. RNA and localization in cephalopods has been executed for decades, beginning with the use of radiographical *in situ* hybridization (ISH) (Capano *et al.* 1987), and later with widespread use of colorimetric ISH utilizing anti-digoxigenin antibodies linked to enzymes which cleave colorimetric substrates, e.g., (Shigeno *et al.* 2015; Tarazona *et al.* 2019). There has been a recent shift to the use of hybridization chain reaction fluorescent *in situ* hybridization (HCR-FISH) in cephalopod systems, including the localization of neuronal precursors (Deryckere *et al.* 2021; Elagoz *et al.* 2022) and localization of both host and symbiont factors in the bobtail squid light organ (Krasity *et al.* 2015; Nikolakakis *et al.* 2015), among other promising studies (Duruz *et al.* 2022; Gavriouchkina *et al.* 2022; Styfhals *et al.* 2022). As HCR-FISH utilizes self-amplifying fluorescent nucleotide hairpins, the technique allows for the cellular or subcellular localization of transcripts (Choi, Schwarzkopf and Pierce 2020).

Immunocytochemistry (ICC) and immunohistochemistry (IHC) have also long been powerful tools for the cephalopod community. Due to the divergence between mammalian and cephalopod protein sequences, many cephalopod antibody studies employ custom antibodies, which, while effective, are expensive to produce and require effort to ensure specificity (Troll *et al.* 2009; Heath-Heckman *et al.* 2014; Li *et al.* 2019; Murata *et al.* 2021). Therefore most studies use commercial antibodies on cephalopod tissue that cross-reacts to well conserved proteins, peptides, or molecules (Di Cosmo, Paolucci and Di Cristo 2004; Springer *et al.* 2005; Goodson *et al.* 2006; Wollesen, Degnan and Wanninger 2010; Altura *et al.* 2011; Shigeno and Ragsdale 2015; Baldascino *et al.* 2017). While these protocols have been well established, there are still frontiers such as the optimization of dual ICC/HCR protocols, either using established ICC protocols (Elagoz *et al.* 2022), or by new HCR-ICC techniques (Schwarzkopf *et al.* 2021). However, the lack of commercially available cephalopod-specific antibodies makes ICC out of reach for some research groups, and a centralized database of commercial antibodies that do work in cephalopod tissue could be an important resource for this burgeoning field.

Steps to integrate gene expression and morphology

A multi-modal approach to integrate gene expression visualization and single-cell -omics with other techniques promises to shed new light on cephalopod biology. By combining scRNAseq datasets with electron microscopy, we will be able to link the molecular information with their morphology and location, as was recently done for *Platynereis*, (Vergara *et al.* 2021). Spatial transcriptomics also holds promise to detangle the cell type composition of cephalopods. Multi-ome sequencing approaches that can capture both gene expression and chromatin accessibility of individual cells will lead to a comprehensive database of cell-type specific enhancers in cephalopods and characterize the epigenetic landscape.

Besides the interest in the developmental stages from embryos to adults in areas such as evolutionary developmental biology (evo-devo) or marine ecology, species such as *Octopus vulgaris* or *Sepioteuthis lessoniana* are highly attractive for human consumption and are a good candidate for aquaculture diversification. Transcriptome sequencing of early developmental stages cultured in different conditions are helping to the identification of biomarkers of growth, health and welfare in aquaculture (García-Fernández *et al.* 2019; Prado-Álvarez *et al.* 2022; Varó *et al.* 2022). The identification and selection of biomarker genes will lead to the construction of specific q-PCR arrays as tools to analyze and guarantee welfare in aquaculture practices.

Cephalopod metagenomics and microbial interactions

Yet another core aspect of cephalopod biology is how bacteria influence the development, evolution, and life history of the animal. Microbes have evolved along with their metazoan partners since their origin on Earth, and surveying their presence can inform us about host health (immunology, pathogen susceptibility), population dynamics, and the ecology of the association. Previous studies have demonstrated how bacteria influence specific developmental

and behavioral features of cephalopods, and the cross talk between the partners (Nyholm and McFall-Ngai 2004, 2021; Visick, Stabb and Ruby 2021). Over the past few decades, much of the experimental work on cephalopod microbiomes has been focused on monotypic associations, including the light organ symbiosis in sepiolid and loliginid squids, and complex microbiomes, such as the accessory nidamental glands (ANGs) in the Decapodiformes (Kaufman *et al.* 1998; Barbieri *et al.* 2001; Nyholm and McFall-Ngai 2021). Both of these microbial systems have been important models for understanding how bacteria influence development of the specific organs they are housed in, as well as their ability to produce antimicrobial/antifungal compounds found in newly laid egg masses (Barbieri *et al.* 1997; Kerwin *et al.* 2019). Microbial communities have also been characterized in the gut of cuttlefish (*Sepia*) and wild octopus (*O. vulgaris*), but very little is known about other communities of bacteria that reside on or inside cephalopods (Roura *et al.* 2017; Lutz *et al.* 2019). Lastly, host specificity, and the variation among environmentally transmitted bacteria throughout populations of cephalopods has illuminated our comprehension of microbial diversity, coevolution between the partners, and microbe-microbe interactions prior to and during the association (Jones *et al.* 2006; Guerrero-Ferreira *et al.* 2013; Coryell *et al.* 2018).

Modes of microbial associations in cephalopods

The symbiotic relationship of bioluminescent *Vibrio fischeri* and cephalopods have been studied for many decades (Nyholm and McFall-Ngai 2021). Different molecular tools are used to experimentally analyze the interaction of the light organ in *E. scolopes* and *V. fischeri*. Further, *Vibrio*-specific 16S rRNA primers are available to identify different *Vibrio* species in different cephalopods (Nishiguchi and Nair 2003; Zamborsky and Nishiguchi 2011; Howard *et al.* 2015). The interrogation of both the light organ and ANG in *E. scolopes* has the potential to reveal common and unique mechanisms by which hosts regulate different microbial niches (Nishiguchi, Lopez and Boletzky 2004). We can use 16S rRNA to identify bacteria in different species of cephalopods and in different organs. The ANG microbiome (Barbieri *et al.* 2001; Pichon *et al.* 2005; Kerwin and Nyholm 2017, 2018; Yang *et al.* 2021), along with the microbiota of gills, skin, and the digestive system (Roura *et al.* 2017; Lutz *et al.* 2019) have been characterized in some squid, cuttlefish and octopus (Roura *et al.* 2017; Lutz *et al.* 2019; Kang *et al.* 2022). A core microbiome based on taxonomy has not been identified, but there is potential to identify fundamental functional microbiomes in different organs.

Several methods have been employed to characterize microbiome interactions in the model cephalopod *E. scolopes* such as 16S rRNA gene diversity metagenomics, transcriptomics/RNA-seq, genomics, proteomics, and metabolomics (Collins *et al.* 2012a; Altura *et al.* 2013; Bongrand *et al.* 2016; Kerwin and Nyholm 2018; Belcaid *et al.* 2019; Moriano-Gutierrez *et al.* 2019; Bongrand *et al.* 2020; Koch *et al.* 2020) and techniques such as HCR-FISH, FISH, immunofluorescence and fluorescently tagged bacteria were used to visualize host-bacteria interactions in *E. scolopes* (Troll *et al.* 2010; Collins *et al.* 2012a; Nikolakakis *et al.* 2015). Macrophage cells (hemocytes) can be easily obtained and observed for their interactions with symbiotic and non-symbiotic bacteria with microscopy (Nyholm *et al.* 2009; Rader, McAnulty and Nyholm 2019) and at the molecular level with proteomics and transcriptomics (Collins *et al.* 2012b; Schleicher *et al.* 2014).

Recent technological advancements permit finer resolution studies of population genetic structures that will be relevant to symbiotic association analyses. Double digest restriction site-associated DNA sequencing (ddRADseq) is a low-cost, high-throughput technique in genotyping thousands of single nucleotide polymorphisms (SNPs) in any species including non-model organisms like cephalopods (Andrews *et al.* 2016). This reduced-representation sequencing approach could be powerful for deducing population genetic structures of the host cephalopods and obtain first indication of microbial species present. Since abiotic factors like temperature, current and salinity have been found to influence the movement and concentrations of *Vibrio* spp. in sepiolids (Nishiguchi 2000; Jones *et al.* 2006; Soto *et al.* 2009; Nourabadi and Nishiguchi 2021), understanding how the population genetic structures of the host relate to that of the symbionts could be crucial in discerning their population ecology and dynamics with important evolutionary outcomes (Pankey *et al.* 2014). SNP dataset obtained from ddRADseq and other metagenomic (*e.g.*, bulk sequencing) approaches on host and symbiont would allow further research into the associations between adaptive genetic markers (either unique or shared between hosts and symbionts) and environmental factors, thereby identifying potential selection pressures in particular populations (Lindgren *et al.* 2012).

Emerging directions in cephalopod symbiosis research

Experimentally tractable model systems have been critical for understanding the influence of microbiomes on eukaryotic biology (Douglas 2019; McFall-Ngai and Bosch 2021). Having the ability to generate axenic/germ-free hosts or hosts with a known microbiota (gnotobiotic) have led to groundbreaking discoveries in vertebrate host-microbe systems, *e.g.*, showing linkages between the microbiome and obesity, type 2 diabetes, autism, and behavior. Developing gnotobiotic and germ-free aquatic husbandry systems would allow for the further expansion of the use of cephalopods in microbiome studies. One challenge in microbiome research is studying direct interactions between host and microbial cells. Cell culture of cephalopod tissue types like the mucosal epithelia of organs that directly interact with symbionts and innate immune cells (hemocytes) would help reveal the molecular cross-talk between the partners. Single cell multi-omic methods could then also be applied to these systems to identify the role of cell-cell interactions in cephalopod symbioses.

Genomic and metagenomic analysis of microbial communities shed light on microbiome function *e.g.*, by revealing important metabolic pathways, signaling pathways, and effectors that influence host biology. In order to expand the study of cephalopod microbiomes it will be critical to increase the number of symbiont reference genomes that are available for metagenomic and metatranscriptomic datasets. Having better reference genomes will also allow for linking biosynthetic gene clusters (BGCs) to the production of specific compounds that can be identified from metabolomics studies.

Studies investigating how cephalopod microbiomes change over time, and whether these changes can give us a window to the temporal dynamics of the host would lead to a greater understanding of how host-microbe associations are initiated, persist, and evolve over time. This includes experimental evolution studies, where microbes are evolved in a novel host

species to determine which genetic cues are responsible for recognition and specificity (Soto, Punke and Nishiguchi 2012; Soto and Nishiguchi 2014; Soto, Rivera and Nishiguchi 2014). Interestingly, this type of work has already been shown to be important in human microbiome research (Bosch, Guillemin and McFall-Ngai 2019; Burnett and Ratcliff 2022; Garud 2022); having a temporal framework of these microbial populations would further our knowledge of the importance of bacteria in cephalopod health (Lynch and Hsiao 2019). Given that much of the applied work in cephalopod husbandry and aquaculture is focused on preventing disease, knowledge of the balance between beneficial and pathogenic microbial species will be important for the future success of expanding these important fisheries.

Conclusions and future directions: CephBase and Best Practices for Sharing Resources

We are at an exciting and critical time in the evolution of the cephalopod field, with many new resources and techniques becoming available that will allow us to address fundamental questions of cephalopod biology and evolution. As the pace and complexity of generated datasets for cephalopods is intensifying, an emerging community goal will be focused on data integration for cephalopod -omics, including but not limited to the creation of shared databases for genomes, annotations, expression data, as well as cultures. Current efforts are emerging from individual labs, e.g, Cephalopod Program at the Marine Biological Laboratory, which is a resource for genetically tractable lines and developing genetic tools for transgenics (<https://www.mbl.edu/research/resources-research-facilities/marine-resources-center/cephalopods/cephalopod-breeding-center>), *Euprymna* genomics portal (<https://metazoa.csb.univie.ac.at/>), and CephRes (www.cephalopodresearch.org), among others. While efforts are being made to make a repository of tractable cephalopod species, lab contacts, their techniques, as well as available resources (see below), a consolidated effort would require funding sources and is otherwise restricted by “good-will”, research grants to individual labs, or philanthropy.

Therefore, as the cephalopod community continues to grow, we propose several steps to ensure that the community can share resources effectively. The first resource that is important to share is genomic and transcriptomic data. Drawing on successful approaches from the *C. elegans* (wormbase) and the *D. melanogaster* (flybase) communities, we would like to establish a centralized database for cephalopod bioinformatic data. A simple user interface will allow researchers to compile the information they need more easily, including expression and verified functional annotation. This repository could also serve as a means to share techniques and protocols. To complement this, any plasmids used for these techniques should be deposited to Addgene, or other repositories. Importantly, the source of the genomic information (such as the geographic location of sampling), is crucial for traceability (Vecchione, Sweeney and Rothman 2022). The third resource that is important to share is animals, both wild type and, eventually, transgenics. Each cephalopod species has many unique characteristics that makes them ideal for answering different questions, and thus no one species has emerged as the main species within the cephalopod research community. This comes with drawbacks: for example a technique that works in squid may not work in octopus. To take advantage of the opportunities that will emerge from a field with a strong comparative element, our longer term vision includes

dedicated cephalopod rearing facilities and centers of excellence for specific techniques and species around the world. This will encourage collaboration and allow scientists to travel to learn/perform specialized experiments on specific species.

The first attempts have been made to molecularly characterize cephalopod cell types at different scales (organ, systems and whole organisms). A large collaborative effort is required to construct a “Tabula Cephalopodae” to obtain a complete taxonomy of cephalopod cell type diversity. By bringing together the expertise of different research groups, we hope to streamline future studies and attempt to combine these datasets in a single-cell cephalopod atlas. By including multiple key cephalopod species we can examine clade-specific novelties such as different brain lobes and the light organ and shed light on cell-type evolution across cephalopods.

Our community has also developed further momentum to go beyond single-organism focus. As outlined above, disruption of the microbiome integrity is a mechanism by which stress affects animal health and welfare (Uren Webster *et al.* 2020). Climate change increases ocean warming and acidification which constitutes a stressful condition (Aprill 2020; Bénard, Vavre and Kremer 2020). Additionally, the intensification of aquaculture is associated with challenges regarding its sustainability, including impact on animal health and welfare, the environment, food safety and economic viability. Intensive aquaculture can expose cephalopods to crowding, handling and social stressors, which can induce negative effects on health and fitness, such as impaired growth, depressed immunity, and increased susceptibility to disease. Besides developing juveniles and adults, early life stages of cephalopods are particularly sensitive to environmental stressors, and there is increasing recognition that microbial communities may contribute to harmful effects (Roura *et al.* 2017). To this end, the continued analysis of gut, skin, or fecal microbiomes can be of great interest to monitor the stress-response and therefore to assess the welfare and health status of cephalopods. Linking -omic datasets for cephalopods and their metagenomes, including data from bacteria, archaea, fungi and bacteriophages (Koskella and Meaden 2013; Guerin and Hill 2020), can thus also aid in disease identification in aquaculture and the natural environment.

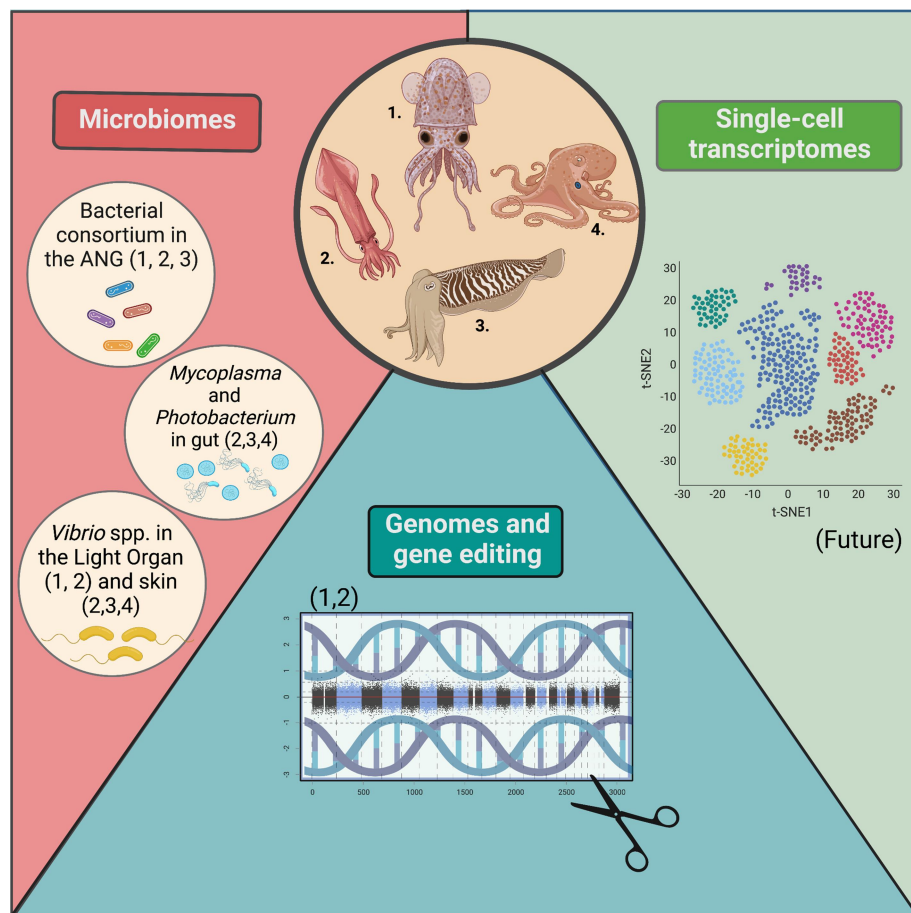


Figure 1. Research questions addressable with emerging technologies for cephalopods. Numbers in brackets refer to the numbers assigned to the different types of cephalopods. 1: Bobtail squid, 2. Loliginid squid, 3. Cuttlefish, 4. Octopus (Created with BioRender.com)

References

- Ahuja N, Hwaun E, Pungor JR *et al.* Creation of an albino squid line by CRISPR-Cas9 and its application for in vivo functional imaging of neural activity. *Curr Biol* 2023;**0**, DOI: 10.1016/j.cub.2023.05.066.
- Albertin CB, Bonnaud L, Brown CT *et al.* Cephalopod genomics: A plan of strategies and organization. *Stand Genomic Sci* 2012;**7**:175–88.
- Albertin CB, Medina-Ruiz S, Mitros T *et al.* Genome and transcriptome mechanisms driving cephalopod evolution. *Nat Commun* 2022;**13**:2427.
- Albertin CB, Simakov O. Cephalopod Biology: At the Intersection between Genomic and Organismal Novelties. *Annu Rev Anim Biosci* 2020;**8**:71–90.
- Albertin CB, Simakov O, Mitros T *et al.* The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* 2015;**524**:220–4.
- Alon S, Garrett SC, Levanon EY *et al.* The majority of transcripts in the squid nervous system are extensively recoded by A-to-I RNA editing. Guigó R (ed.). *eLife* 2015;**4**:e05198.
- Altura MA, Heath-Heckman EAC, Gillette A *et al.* The first engagement of partners in the *Euprymna scolopes*-*Vibrio fischeri* symbiosis is a two-step process initiated by a few environmental symbiont cells. *Environ Microbiol* 2013;**15**:2937–50.
- Altura MA, Stabb E, Goldman W *et al.* Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* 2011;**13**:527–37.
- Andrews KR, Good JM, Miller MR *et al.* Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat Rev Genet* 2016;**17**:81–92.
- Aprill A. The Role of Symbioses in the Adaptation and Stress Responses of Marine Organisms. *Annu Rev Mar Sci* 2020;**12**:291–314.
- Badhiwala KN, Primack AS, Juliano CE *et al.* Multiple neuronal networks coordinate Hydra mechanosensory behavior. *eLife* 2021;**10**:e64108.
- Baldascino E, Di Cristina G, Tedesco P *et al.* The Gastric Ganglion of *Octopus vulgaris*: Preliminary Characterization of Gene- and Putative Neurochemical-Complexity, and the Effect of *Aggregata octopiana* Digestive Tract Infection on Gene Expression. *Front Physiol* 2017;**8**:1001.
- Barbieri E, Barry K, Child A *et al.* Antimicrobial Activity in the Microbial Community of the Accessory Nidamental Gland and Egg Cases of *Loligo pealei* (Cephalopoda: Loliginidae). *Biol Bull* 1997;**193**:275–6.
- Barbieri E, Paster BJ, Hughes D *et al.* Phylogenetic characterization of epibiotic bacteria in the accessory nidamental gland and egg capsules of the squid *Loligo pealei* (Cephalopoda:Loliginidae). *Environ Microbiol* 2001;**3**:151–67.
- Belcaid M, Casaburi G, McAnulty SJ *et al.* Symbiotic organs shaped by distinct modes of genome evolution in cephalopods. *Proc Natl Acad Sci U S A* 2019;**116**:3030–5.
- Bénard A, Vavre F, Kremer N. Stress & Symbiosis: Heads or Tails? *Front Ecol Evol* 2020;**8**.
- Bongrand C, Koch EJ, Moriano-Gutierrez S *et al.* A genomic comparison of 13 symbiotic *Vibrio fischeri* isolates from the perspective of their host source and colonization behavior. *ISME J* 2016;**10**:2907–17.
- Bongrand C, Moriano-Gutierrez S, Arevalo P *et al.* Using Colonization Assays and Comparative Genomics To Discover Symbiosis Behaviors and Factors in *Vibrio fischeri*. *mBio* 2020;**11**:e03407-19.
- Bosch TCG, Guillemin K, McFall-Ngai M. Evolutionary “Experiments” in Symbiosis: The Study of Model Animals Provides Insights into the Mechanisms Underlying the Diversity of Host-Microbe Interactions. *BioEssays News Rev Mol Cell Dev Biol* 2019;**41**:e1800256.
- Broussard GJ, Liang R, Tian L. Monitoring activity in neural circuits with genetically encoded indicators. *Front Mol Neurosci* 2014;**7**.
- Bruford EA, Braschi B, Denny P *et al.* Guidelines for Human Gene Nomenclature. *Nat Genet*

- 2020;**52**:754–8.
- Brůna T, Hoff KJ, Lomsadze A *et al.* BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics Bioinforma* 2021;**3**:lqaa108.
- Brůna T, Lomsadze A, Borodovsky M. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genomics Bioinforma* 2020;**2**:lqaa026.
- Buenrostro JD, Wu B, Chang HY *et al.* ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 2015;**109**:21.29.1-21.29.9.
- Burnetti A, Ratcliff WC. Experimental evolution is not just for model organisms. *PLOS Biol* 2022;**20**:e3001587.
- Cantarel BL, Korf I, Robb SMC *et al.* MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* 2008;**18**:188–96.
- Cao C, Lemaire LA, Wang W *et al.* Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 2019;**571**:349–54.
- Capano CP, Giuditta A, Castigli E *et al.* Occurrence and sequence complexity of polyadenylated RNA in squid axoplasm. *J Neurochem* 1987;**49**:698–704.
- Castellanos-Martínez S, Arteta D, Catarino S *et al.* De Novo Transcriptome Sequencing of the Octopus vulgaris Hemocytes Using Illumina RNA-Seq Technology: Response to the Infection by the Gastrointestinal Parasite *Aggregata octopiana*. *PLOS ONE* 2014;**9**:e107873.
- Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2022;**50**:D165–73.
- Caurcel C, Laetsch DR, Challis R *et al.* MolluscDB: a genome and transcriptome database for molluscs. *Philos Trans R Soc Lond B Biol Sci* 2021;**376**:20200157.
- Cheng H, Concepcion GT, Feng X *et al.* Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* 2021;**18**:170–5.
- Choi HMT, Schwarzkopf M, Pierce NA. Multiplexed Quantitative In Situ Hybridization with Subcellular or Single-Molecule Resolution Within Whole-Mount Vertebrate Embryos: qHCR and dHCR Imaging (v3.0). *Methods Mol Biol Clifton NJ* 2020;**2148**:159–78.
- Collins AJ, LaBarre BA, Won BSW *et al.* Diversity and partitioning of bacterial populations within the accessory nidamental gland of the squid *Euprymna scolopes*. *Appl Environ Microbiol* 2012a;**78**:4200–8.
- Collins AJ, Schleicher TR, Rader BA *et al.* Understanding the role of host hemocytes in a squid/*Vibrio* symbiosis using transcriptomics and proteomics. *Front Immunol* 2012b;**3**:91.
- Coryell M, McAlpine M, Pinkham NV *et al.* The gut microbiome is required for full protection against acute arsenic toxicity in mouse models. *Nat Commun* 2018;**9**:5424.
- Crawford K. Culture Method for in vitro Fertilization to Hatching of the Squid, *Loligo pealeii*. *Biol Bull* 2002;**203**:216–7.
- Crawford K, Diaz Quiroz JF, Koenig KM *et al.* Highly Efficient Knockout of a Squid Pigmentation Gene. *Curr Biol CB* 2020;**30**:3484-3490.e4.
- Deryckere A, Styfhals R, Elagoz AM *et al.* Identification of neural progenitor cells and their progeny reveals long distance migration in the developing octopus brain. Sen S, Dulac C, Sen S, *et al.* (eds.). *eLife* 2021;**10**:e69161.
- Destanović D, Schultz DT, Styfhals R *et al.* A chromosome-level reference genome for the common octopus, *Octopus vulgaris* (Cuvier, 1797). 2023:2023.05.16.540928.
- Di Cosmo A, Paolucci M, Di Cristo C. N-methyl-D-aspartate receptor-like immunoreactivity in the brain of *Sepia* and *Octopus*. *J Comp Neurol* 2004;**477**:202–19.
- Douglas AE. Simple animal models for microbiome research. *Nat Rev Microbiol* 2019;**17**:764–75.
- Duruz J, Sprecher M, Kaldun JC *et al.* Molecular characterization of cell types in the squid

- Loligo vulgaris. 2022:2022.03.28.485983.
- Elagoz AM, Styfhals R, Maccuro S *et al.* Optimization of Whole Mount RNA Multiplexed in situ Hybridization Chain Reaction With Immunohistochemistry, Clearing and Imaging to Visualize Octopus Embryonic Neurogenesis. *Front Physiol* 2022;**13**:882413.
- Fraga MF, Esteller M. DNA methylation: a profile of methods and applications. *BioTechniques* 2002;**33**:632, 634, 636–49.
- García-Fernández P, Prado-Alvarez M, Nande M *et al.* Global impact of diet and temperature over aquaculture of Octopus vulgaris paralarvae from a transcriptomic approach. *Sci Rep* 2019;**9**:10312.
- Garud N. Microbial Evolution: An overlooked biomarker of host diet. *Cell Host Microbe* 2022;**30**:146–7.
- Gavriouchkina D, Tan Y, Ziadi-Künzli F *et al.* A single-cell atlas of bobtail squid visual and nervous system highlights molecular principles of convergent evolution. 2022:2022.05.26.490366.
- González-Blas CB, Winter SD, Hulselmans G *et al.* SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. 2022:2022.08.19.504505.
- Goodson MS, Crookes-Goodson WJ, Kimbell JR *et al.* Characterization and role of p53 family members in the symbiont-induced morphogenesis of the Euprymna scolopes light organ. *Biol Bull* 2006;**211**:7–17.
- Guerin E, Hill C. Shining Light on Human Gut Bacteriophages. *Front Cell Infect Microbiol* 2020;**10**.
- Guerrero-Ferreira R, Gorman C, Chavez AA *et al.* Characterization of the Bacterial Diversity in Indo-West Pacific Loliginid and Sepiolid Squid Light Organs. *Microb Ecol* 2013;**65**:214–26.
- Heath-Heckman E, Nishiguchi MK. Leveraging Short-Read Sequencing to Explore the Genomics of Sepiolid Squid. *Integr Comp Biol* 2021;**61**:1753–61.
- Heath-Heckman EAC, Gillette AA, Augustin R *et al.* Shaping the microenvironment: evidence for the influence of a host galaxin on symbiont acquisition and maintenance in the squid-Vibrio symbiosis. *Environ Microbiol* 2014;**16**:3669–82.
- Herrmann C, Van de Sande B, Potier D *et al.* i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. *Nucleic Acids Res* 2012;**40**:e114.
- Hoff KJ, Stanke M. Predicting Genes in Single Genomes with AUGUSTUS. *Curr Protoc Bioinforma* 2019;**65**:e57.
- Howard JH, Williams B, Breen M *et al.* Genomics of an environmentally-transmitted symbiosis: Newly sequenced Vibrio fischeri genomes from different host squids and geographic locations. 2015;**65**:187–200.
- Jones BW, Lopez JE, Huttenburg J *et al.* Population structure between environmentally transmitted vibrios and bobtail squids using nested clade analysis. *Mol Ecol* 2006;**15**:4317–29.
- Kang W, Kim PS, Tak EJ *et al.* Host phylogeny, habitat, and diet are main drivers of the cephalopod and mollusk gut microbiome. *Anim Microbiome* 2022;**4**:30.
- Kaufman MR, Ikeda Y, Patton C *et al.* Bacterial Symbionts Colonize the Accessory Nidamental Gland of the Squid Loligo opalescens via Horizontal Transmission. *Biol Bull* 1998;**194**:36–43.
- Kawakami K, Shima A. Identification of the Tol2 transposase of the medaka fish Oryzias latipes that catalyzes excision of a nonautonomous Tol2 element in zebrafish Danio rerio. *Gene* 1999;**240**:239–44.
- Kerwin AH, Gromek SM, Suria AM *et al.* Shielding the Next Generation: Symbiotic Bacteria from a Reproductive Organ Protect Bobtail Squid Eggs from Fungal Fouling. *mBio* 2019;**10**, DOI: 10.1128/MBIO.02376-19.

- Kerwin AH, Nyholm SV. Symbiotic bacteria associated with a bobtail squid reproductive system are detectable in the environment, and stable in the host and developing eggs. *Environ Microbiol* 2017;**19**:1463–75.
- Kerwin AH, Nyholm SV. Reproductive System Symbiotic Bacteria Are Conserved between Two Distinct Populations of *Euprymna scolopes* from Oahu, Hawaii. *mSphere* 2018;**3**:e00531-17.
- Koch EJ, Moriano-Gutierrez S, Ruby EG *et al*. The impact of persistent colonization by *Vibrio fischeri* on the metabolome of the host squid *Euprymna scolopes*. *J Exp Biol* 2020;**223**:jeb212860.
- Koenig KM, Sun P, Meyer E *et al*. Eye development and photoreceptor differentiation in the cephalopod *Doryteuthis pealeii*. *Development* 2016;**143**:3168–81.
- Koskella B, Meaden S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* 2013;**5**:806–23.
- Krasity BC, Troll JV, Lehnert EM *et al*. Structural and functional features of a developmentally regulated lipopolysaccharide-binding protein. *mBio* 2015;**6**:e01193-01115.
- Labun K, Montague TG, Krause M *et al*. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res* 2019;**47**:W171–4.
- Lawson ND, Li R, Shin M *et al*. An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes. *eLife* 2020;**9**:e55792.
- Lee PN, Callaerts P, de Couet HG. The embryonic development of the hawaiian bobtail squid (*Euprymna scolopes*). *Cold Spring Harb Protoc* 2009;**4**, DOI: 10.1101/pdb.ip77.
- Li H-W, Chen C, Kuo W-L *et al*. The Characteristics and Expression Profile of Transferrin in the Accessory Nidamental Gland of the Bigfin Reef Squid during Bacteria Transmission. *Sci Rep* 2019;**9**:20163.
- Lindgren AR, Pankey MS, Hochberg FG *et al*. A multi-gene phylogeny of Cephalopoda supports convergent morphological evolution in association with multiple habitat shifts in the marine environment. *BMC Evol Biol* 2012;**12**:129.
- Liscovitch-Brauer N, Alon S, Porath HT *et al*. Trade-off between Transcriptome Plasticity and Genome Evolution in Cephalopods. *Cell* 2017;**169**:191-202.e11.
- Lutz HL, Ramírez-Puebla ST, Abbo L *et al*. A Simple Microbiome in the European Common Cuttlefish, *Sepia officinalis*. *mSystems* 2019;**4**:e00177-19.
- Lynch JB, Hsiao EY. Microbiomes as sources of emergent host phenotypes. *Science* 2019;**365**:1405–9.
- Ma W, Ay F, Lee C *et al*. Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat Methods* 2015;**12**:71–8.
- Macchi F, Edsinger E, Sadler KC. Epigenetic machinery is functionally conserved in cephalopods. *BMC Biol* 2022;**20**:202.
- Marino A, Kizenko A, Wong WY *et al*. Repeat Age Decomposition Informs an Ancient Set of Repeats Associated With Coleoid Cephalopod Divergence. *Front Genet* 2022;**13**:793734.
- McCoy MJ, Fire AZ. Intron and gene size expansion during nervous system evolution. *BMC Genomics* 2020;**21**:360.
- McFall-Ngai M, Bosch TCG. Animal development in the microbial world: The power of experimental model systems. *Curr Top Dev Biol* 2021;**141**:371–97.
- McKenna V, Archibald JM, Beinart R *et al*. The Aquatic Symbiosis Genomics Project: probing the evolution of symbiosis across the tree of life. 2021, DOI: 10.12688/wellcomeopenres.17222.1.
- de Mendoza A, Poppe D, Buckberry S *et al*. The emergence of the brain non-CpG methylation system in vertebrates. *Nat Ecol Evol* 2021;**5**:369–78.
- Minor PJ, Clarke DN, Andrade López JM *et al*. I-SceI Meganuclease-mediated transgenesis in the acorn worm, *Saccoglossus kowalevskii*. *Dev Biol* 2019;**445**:8–15.

- Montague TG, Rieth IJ, Axel R. Embryonic development of the camouflaging dwarf cuttlefish, *Sepia bandensis*. *Dev Dyn Off Publ Am Assoc Anat* 2021;**250**:1688–703.
- Moreno-Mateos MA, Vejnar CE, Beaudoin J-D *et al*. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* 2015;**12**:982–8.
- Moriano-Gutierrez S, Koch EJ, Bussan H *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* 2019;**116**:7990–9.
- Murata R, Mushirobira Y, Tanaka Y *et al*. Expression profile of GnRH-like peptide during gonadal sex differentiation in the cephalopod kisslip cuttlefish, *Sepia lycidas*. *Gen Comp Endocrinol* 2021;**304**:113718.
- Natsidis P, Schiffer PH, Salvador-Martínez I *et al*. Computational discovery of hidden breaks in 28S ribosomal RNAs across eukaryotes and consequences for RNA Integrity Numbers. *Sci Rep* 2019;**9**:19477.
- Nikolakakis K, Lehnert E, McFall-Ngai MJ *et al*. Use of Hybridization Chain Reaction-Fluorescent In Situ Hybridization To Track Gene Expression by Both Partners during Initiation of Symbiosis. *Appl Environ Microbiol* 2015;**81**:4728–35.
- Nishiguchi MK. Temperature Affects Species Distribution in Symbiotic Populations of *Vibrio* spp. *Appl Environ Microbiol* 2000;**66**:3550–5.
- Nishiguchi MK, Lopez JE, Boletzky S v. Enlightenment of old ideas from new investigations: more questions regarding the evolution of bacteriogenic light organs in squids. *Evol Dev* 2004;**6**:41–9.
- Nishiguchi MK, Nair VS. Evolution of symbiosis in the Vibrionaceae: a combined approach using molecules and physiology. *Int J Syst Evol Microbiol* 2003;**53**:2019–26.
- Nourabadi N, Nishiguchi MK. pH Adaptation Drives Diverse Phenotypes in a Beneficial Bacterium-Host Mutualism. *Front Ecol Evol* 2021;**9**.
- Nurk S, Koren S, Rhie A *et al*. The complete sequence of a human genome. *Science* 2022;**376**:44–53.
- Nyholm SV, McFall-Ngai MJ. The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol* 2004;**2**:632–42.
- Nyholm SV, McFall-Ngai MJ. A lasting symbiosis: how the Hawaiian bobtail squid finds and keeps its bioluminescent bacterial partner. *Nat Rev Microbiol* 2021;**19**:666–79.
- Nyholm SV, Stewart JJ, Ruby EG *et al*. Recognition between symbiotic *Vibrio fischeri* and the haemocytes of *Euprymna scolopes*. *Environ Microbiol* 2009;**11**:483–93.
- Paganos P, Voronov D, Musser J *et al*. Single cell RNA sequencing of the *Strongylocentrotus purpuratus* larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. 2021:2021.03.16.435574.
- Pankey MS, Minin VN, Imholte GC *et al*. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proc Natl Acad Sci* 2014;**111**:E4736–42.
- Pavlopoulos A, Averof M. Establishing genetic transformation for comparative developmental studies in the crustacean *Parhyale hawaiiensis*. *Proc Natl Acad Sci* 2005;**102**:7888–93.
- Pichon D, Gaia V, Norman MD *et al*. Phylogenetic diversity of epibiotic bacteria in the accessory nidamental glands of squids (Cephalopoda: Loliginidae and Idiosepiidae). *Mar Biol* 2005;**147**:1323–32.
- Prado-Álvarez M, Dios S, García-Fernández P *et al*. De novo transcriptome reconstruction in aquacultured early life stages of the cephalopod *Octopus vulgaris*. *Sci Data* 2022;**9**:609.
- Rader B, McAnulty SJ, Nyholm SV. Persistent symbiont colonization leads to a maturation of hemocyte response in the *Euprymna scolopes/Vibrio fischeri* symbiosis. *MicrobiologyOpen* 2019;**8**, DOI: 10.1002/MBO3.858.
- Rhie A, McCarthy SA, Fedrigo O *et al*. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 2021;**592**:737–46.
- Ricci L, Srivastava M. Transgenesis in the acoele worm *Hofstenia miamia*. *Dev Cell*

- 2021;**56**:3160-3170.e4.
- Roura Á, Doyle SR, Nande M *et al.* You Are What You Eat: A Genomic Analysis of the Gut Microbiome of Captive and Wild Octopus vulgaris Paralarvae and Their Zooplankton Prey. *Front Physiol* 2017;**8**.
- Sanchez G, Fernández-Álvarez FÁ, Taite M *et al.* Phylogenomics illuminates the evolution of bobtail and bottletail squid (order Sepiolida). *Commun Biol* 2021;**4**:1–9.
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* 2014;**32**:347–55.
- Schleicher TR, VerBerkmoes NC, Shah M *et al.* Colonization state influences the hemocyte proteome in a beneficial squid-Vibrio symbiosis. *Mol Cell Proteomics MCP* 2014;**13**:2673–86.
- Schmidbaur H, Kawaguchi A, Clarence T *et al.* Emergence of novel cephalopod gene regulation and expression through large-scale genome reorganization. *Nat Commun* 2022;**13**:2172.
- Schwarzkopf M, Liu MC, Schulte SJ *et al.* Hybridization chain reaction enables a unified approach to multiplexed, quantitative, high-resolution immunohistochemistry and in situ hybridization. *Dev Camb Engl* 2021;**148**:dev199847.
- Sharma S, Wang W, Stolfi A. Single-cell transcriptome profiling of the Ciona larval brain. *Dev Biol* 2019;**448**:226–36.
- Shigeno S, Parnaik R, Albertin CB *et al.* Evidence for a cordal, not ganglionic, pattern of cephalopod brain neurogenesis. *Zool Lett* 2015;**1**:26.
- Shigeno S, Ragsdale CW. The gyri of the octopus vertical lobe have distinct neurochemical identities. *J Comp Neurol* 2015;**523**:1297–317.
- Songco-Casey JO, Coffing GC, Piscopo DM *et al.* Cell types and molecular architecture of the Octopus bimaculoides visual system. *Curr Biol* 2022;**32**:5031-5044.e4.
- Soto W, Gutierrez J, Remmenga MD *et al.* Salinity and Temperature Effects on Physiological Responses of Vibrio fischeri from Diverse Ecological Niches. *Microb Ecol* 2009;**57**:140–50.
- Soto W, Nishiguchi MK. Microbial experimental evolution as a novel research approach in the Vibrionaceae and squid-Vibrio symbiosis. *Front Microbiol* 2014;**5**:593.
- Soto W, Punke EB, Nishiguchi MK. Evolutionary perspectives in a mutualism of sepiolid squid and bioluminescent bacteria: combined usage of microbial experimental evolution and temporal population genetics. *Evol Int J Org Evol* 2012;**66**:1308–21.
- Soto W, Rivera FM, Nishiguchi MK. Ecological diversification of Vibrio fischeri serially passaged for 500 generations in novel squid host Euprymna tasmanica. *Microb Ecol* 2014;**67**:700–21.
- Springer J, Ruth P, Beuerlein K *et al.* Distribution and function of biogenic amines in the heart of Nautilus pompilius L. (Cephalopoda, Tetrabranchiata). *J Mol Histol* 2005;**36**:345–53.
- Styfals R, Zolotarov G, Hulsélmans G *et al.* Cell type diversity in a developing octopus brain. 2022:2022.01.24.477459.
- Tanner AR, Fuchs D, Winkelmann IE *et al.* Molecular clocks indicate turnover and diversification of modern coleoid cephalopods during the Mesozoic Marine Revolution. *Proc R Soc B Biol Sci* 2017;**284**:20162818.
- Tarazona OA, Lopez DH, Slota LA *et al.* Evolution of limb development in cephalopod mollusks. Prpic-Schäper N-M, Tautz D, Wagner G (eds.). *eLife* 2019;**8**:e43828.
- Thermes V, Grabher C, Ristoratore F *et al.* I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech Dev* 2002;**118**:91–8.
- Troll JV, Adin DM, Wier AM *et al.* Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. *Cell Microbiol* 2009;**11**:1114–27.
- Troll JV, Bent EH, Pacquette N *et al.* Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. *Environ Microbiol* 2010;**12**:2190–203.

- Uren Webster TM, Rodriguez-Barreto D, Consuegra S *et al.* Cortisol-Related Signatures of Stress in the Fish Microbiome. *Front Microbiol* 2020;**11**:1621.
- Varó I, Prado-Álvarez M, Ortea I *et al.* Proteogenomic Study of the Effect of an Improved Mixed Diet of Live Preys on the Aquaculture of Octopus vulgaris Paralarvae. *Front Mar Sci* 2022;**8**.
- Vecchione M, Sweeney MJ, Rothman PL. The geographic problem in cephalopod genomics. *Front Mar Sci* 2022;**9**.
- Vergara HM, Pape C, Meechan KI *et al.* Whole-body integration of gene expression and single-cell morphology. *Cell* 2021;**184**:4819-4837.e22.
- Visick KL, Stabb EV, Ruby EG. A lasting symbiosis: how *Vibrio fischeri* finds a squid partner and persists within its natural host. *Nat Rev Microbiol* 2021;**19**:654–65.
- Wollesen T, Degnan BM, Wanninger A. Expression of serotonin (5-HT) during CNS development of the cephalopod mollusk, *Idiosepius notoides*. *Cell Tissue Res* 2010;**342**:161–78.
- Yang S-H, Chen C, Hsieh YE *et al.* Bacterial Dynamics in the Accessory Nidamental Gland of *Sepioteuthis lessoniana* throughout Maturation. *Microbes Environ* 2021;**36**:ME21030.
- Zamborsky DJ, Nishiguchi MK. Phylogeographical Patterns among Mediterranean Sepioid Squids and Their *Vibrio* Symbionts: Environment Drives Specificity among Sympatric Species. *Appl Environ Microbiol* 2011;**77**:642–9.
- Zolotarov G, Fromm B, Legnini I *et al.* MicroRNAs are deeply linked to the emergence of the complex octopus brain. *Sci Adv* 2022;**8**:eadd9938.

ORIGINAL UNEDITED MANUSCRIPT