Phylogeny of siphonophore specimens collected from multiple year fieldwork expeditions

Website: https://www.bco-dmo.org/dataset/907525
Data Type: Synthesis
Version: 1
Version Date: 2023-09-01

Project
» Collaborative research: The effects of predator traits on the structure of oceanic food webs (SiphWeb)

Contributors	Affiliation	Role
Haddock, Steven H. D.	Monterey Bay Aquarium Research Institute (MBARI)	Principal Investigator, Contact
Damian-Serrano, Alejandro	Yale University	Co-Principal Investigator
Dunn, Casey W.	Yale University	Co-Principal Investigator
Heyl, Taylor	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Predator specialization has often been considered an evolutionary ‘dead-end’ due to the constraints associated with the evolution of morphological and functional optimizations throughout the organism. However, in some predators, these changes are localized in separate structures dedicated to prey capture. One of the most extreme cases of this modularity can be observed in siphonophores, a clade of pelagic colonial cnidarians that use tentilla (tentacle side branches armed with nematocysts) exclusively for prey capture. Here we study how siphonophore specialists and generalists evolve, and what morphological changes are associated with these transitions. To answer these questions, we: (1) measured 29 morphological characters of tentacles from 45 siphonophore species, (2) mapped these data to a phylogenetic tree, and (3) analyzed the evolutionary associations between morphological characters and prey type data from the literature. Instead of a dead-end, we found that siphonophore specialists can evolve into generalists, and that specialists on one prey type have directly evolved into specialists on other prey types. Our results show that siphonophore tentillum morphology has strong evolutionary associations with prey type, and suggest that shifts between prey types are linked to shifts in the morphology, mode of evolution, and genetic correlations of tentilla and their nematocysts. The evolutionary history of siphonophore specialization helps build a broader perspective on predatory niche diversification via morphological innovation and evolution. These findings contribute to understanding how specialization and morphological evolution have shaped present-day food webs.

Table of Contents

- Dataset Description
  - Methods & Sampling
  - Data Processing Description
- Data Files
- Supplemental Files
- Related Publications
- Parameters
- Project Information
- Funding

Methods & Sampling

While the main goal of this work is not to elucidate a novel phylogeny for Siphonophora, we did expand on the most recent transcriptome based phylogeny [@munro2018improved] to accommodate a larger taxon sampling. In order to do this, we ran a constrained analysis on an extensive 18S+16S dataset. The phylogenetic analysis included 55 siphonophore species and 6 outgroup cnidarian species (*Clytia

hemisphaerica*, *Hydra circumcincta*, *Ectopleura dumortieri*, *Porpita porpita*, *Velella velella*, *Staurocladia wellingtoni*). The gene sequences we used in this study are available online (accession numbers in Dryad repository). Some of the sequences we used were accessioned in [@dunn2005molecular], and others we extracted from the transcriptomes in [@munro2018improved]. Two new 16S sequences for *Frillagalma vityazi* (MK958598) and *Thermopalia* sp. (MK958599) sequenced by Lynne Christianson using the primers from [@cunningham1993molecular] (read 3' to 5' F: TCGACTGTTTACCAAAAACATAGC , R: ACGGAATGAACTCAAATCATGTAAG) were included and accessioned to NCBI. Additional details on the phylogenetic inference methods can be found in the Supplementary Methods.

Unconstrained ML and Bayesian phylogenies were congruent (S2,S5). Given the broader sequence sampling of the transcriptome phylogeny, we ran constrained inferences (using both ML and Bayesian approaches, which produced fully congruent topologies (S4,S6)) after clamping the 5 nodes (S3, blue circles in Fig. \@ref(figure4)) that were incongruent with the topology of the consensus tree in [@munro2018improved]. This topology was then used to inform a Bayesian relaxed molecular clock time-tree in RevBayes, using a birth-death process (sampling probability calculated from the known number of described siphonophore species) to generate ultrametric branch lengths (S7-8). Scripts and tree files are available in the Dryad repository.

Published May 07, 2021 on Dryad. https://doi.org/10.5061/dryad.p2ngf1vp2

Data Processing Description

Statistical analyses
We used a series of phylogenetic comparative methods to test the evolutionary hypotheses presented in this study. We reconstructed ancestral states using ML (R phytools::anc.ML [@revell2012phytools]), and stochastic character mapping (R phytools::make.simmap) for categorical characters. For more details on the data wrangling prior to these analyses, please see the Supplementary Methods. R scripts are available in the Dryad repository.

In order to study the evolution of predatory specialization, we reconstructed components of the diet and prey selectivity on the phylogeny using ML (R phytools::anc.ML). To identify evolutionary associations of diet with tentillum and nematocyst characters, we compared the performance of a neutral evolution model to that of a diet-driven directional selection model. First, we collapsed the diet data into the five feeding guilds mentioned above (fish specialist, small crustacean specialist, large crustacean specialist, gelatinous specialist, generalist), based on which prey types they were observed consuming most frequently. Then, we reconstructed the feeding guild ancestral states using the ML function ace (package ape [@paradis2019package]), removing tips with no feeding data. The ML reconstruction was congruent with the consensus stochastic character mapping (S15). Then, using the package *OUwie* [@beaulieu2012ouwie], we fitted an OU model with multiple optima and rates of evolution (OUm) matched to the reconstructed ancestral diet regimes, a single optimum OU model, and a BM null model, inspired by the analyses in [@cressler2015detecting]. We then ranked the models in order of increasing parametric complexity (BM, OU, OUm), and compared the corrected Akaike Information Criterion (AICc) support scores [@sugiura1978further] to the lowest (best) score, using a cutoff of 2 units to determine significantly better support. When the best fitting model was not significantly better than a less complex alternative, we selected the least complex model (S9). In addition, we calculated and reported the model adequacy scores using the R package *arbutus* [@pennell2015model].

In order to study correlations between the rates of evolution between different characters, we fitted a set of evolutionary variance-covariance matrices [@revell2009phylogenetic] (R phytools::evol.vcv). For more details on the data wrangling preceding these analyses, please see Supplementary Methods. To test whether phenotypic integration changed across selective regimes determined by the reconstructed feeding guilds, we carried out character-pairwise variance-covariance analysis comparing alternative models (R phytools::evolvcv.lite), including those where correlations are the same across the whole tree and models where correlations differ between selective regimes (S19). Number of taxa used in each pairwise comparison is reported in S20. Finally, we compared regime-specific variance-covariance matrices to the general matrix and to their preceding regime matrix to identify the changes in character dependences unique to each regime (S21-22).

We carried out a linear discriminant analysis of principal components (DAPC) using the dapc function (R adegenet::dapc) [@jombart2010discriminant]. This function allowed us to incorporate more predictors than individuals. We generated discriminant functions for feeding guild, and for the presence of copepods, fish, and shrimp (large crustaceans) in the diet (S10-13). From these DAPCs we obtained the highest contributing morphological characters to the discrimination (characters in the top quartile of the weighted sum of the linear

discriminant loadings controlling for the eigenvalue of each discriminant). In order to identify the sign of the relationship between the predictor characters and prey type presence in the diet, we then generated generalized logistic regression models (as a type of generalized linear model, or GLM using R stats::glm) and phylogenetic generalized linear models (R phylolm::phyloglm) with the top contributing characters (from the corresponding DAPC) as predictors (S14). We also carried out these GLMs on the Ivlev's selectivity indices for each prey type calculated from [@purcell1981dietary]. In addition, we ran a series of comparative analyses to address hypotheses of diet-tentillum relationships posed in the literature. Additional details on the DAPC optimization are available in the Supplementary Methods.

Supplementary Methods

Phylogenetic inference:

We aligned the sequences using MAFFT [@katoh2002mafft] (alignments available in Dryad). We inferred a Maximum Likelihood (ML) phylogeny (S2) from 16S and 18S ribosomal rRNA genes using IQTree [@nguyen2014iq] with 1000 bootstrap replicates (iqtree -s alignment.fa -nt AUTO -bb 1000). We used ModelFinder [@kalyaanamoorthy2017modelfinder] implemented in IQTree v1.5.5. to assess the relative model fit. ModelFinder selected GTR+R4 for having the lowest Bayesian Information Criterion score. Additionally, we inferred a Bayesian tree with each gene as an independent partition in RevBayes [@hohna2016revbayes] (S5), which was topologically congruent with the unconstrained ML tree. The *alpha* priors were selected to minimize prior load in site variation.

Data wrangling for comparative analyses:

For comparative analyses, we removed species present in the tree but not represented in the morphology data, and *vice versa*. Although we measured specimens labeled as *Nanomia bijuga* and *Nanomia cara*, we are not confident in some of the species-level identifications, and some specimens were missing diagnostic zooids. Thus, we decided to collapse these into a single taxonomic concept (*Nanomia* sp.). All *Nanomia* sp. observations were matched to the phylogenetic position of *Nanomia bijuga* in the tree. We carried out all phylogenetic comparative statistical analyses in the programming environment R [@team2017r], using the Bayesian ultrametric species tree (S8), and incorporating intraspecific variation estimated from the specimen data as standard error whenever the analysis tool allowed it. R scripts and summarized species-collapsed data available in the Dryad repository. For each character (or character pair) analyzed, we removed species with missing data and reported the number of taxa included. We tested each character for normality using the Shapiro-Wilk test [@shapiro1965analysis] and log-transformed those that were non-normal.

Data wrangling for the variance-covariance analyses:

When fitting all variance-covariance terms simultaneously (S16-18), we selected the largest set of characters that would allow the analysis to run without computational singularities. This excluded many of the morphometric characters which are linearly dependent on other characters. Since the functions do not tolerate missing data, we ran the analyses in two ways: One including all taxa but transforming absent states to zeroes, and another removing the taxa with absent states. These analyses could only be carried out on the subset of taxa for which diet data is available, and only among character pairs that are not computationally singular for that taxonomic subset. Gelatinous specialist correlations could only be estimated for a small subset of characters present in *Apolemia* (S21F, S22E, S23D) and should be interpreted with care.

Comparative tools used to test character associations:

To test for correlated evolution among binary characters, we used Pagel's test [@pagel1994detecting]. To characterize and evaluate the relationship between continuous characters, we used phylogenetic generalized least squares regressions (PGLS) [@grafen1989phylogenetic]. To compare the evolution of continuous characters with categorical aspects of the diet, we carried out a phylogenetic logistic regression (R nlme::gls using the 'corBrownian' function for the argument 'correlation').

DAPC optimization:

Some taxa have inapplicable states for certain absent characters (such as the length of a nematocyst subtype that is not present in a species), which are problematic for DAPC analyses. We tackled this by transforming the absent states to zeroes. This approach allows us to incorporate all the data, but creates an attraction bias between small character states (*e.g.* small tentilla) and absent states (*e.g.* no tentilla). Absent characters are likely to be very biologically relevant to prey capture and we believe they should be accounted for in a predictive approach. We limited the number of linear discriminant functions retained to the number of groupings in each case. We selected the number of principal components retained using the a-score optimization function (R adegenet::optim.a.score) [@jombart2010discriminant] with 100 iterations, which yielded more stable results than the cross validation function (R adegenet::xval). This optimization aims to find the compromise value with highest discrimination power with the least overfitting. The discriminant analysis for feeding guild (7 principal components, 4 discriminants) produced 100% discrimination, and the highest loading contributions were found for the characters (ordered from highest to lowest): Involucrum length, heteroneme

volume, heteroneme number, total heteroneme volume, tentacle width, heteroneme length, total nematocyst volume, and heteroneme width (S10).

[table of contents | back to top]

Data Files

File	
16S_alignment.nexus	(Octet Stream, 178.79 KB) MD5:b14c64e20b36858bfd17e6debd9d1403
18S_alignment.nexus	(Octet Stream, 325.24 KB) MD5:30648beba60240864eb5ec17d180c420
concatenate_18S16S_alignment.nexus	(Octet Stream, 503.92 KB) MD5:4e3c94d2cc578b650331578c93c698bc
raw_categorical_data.csv	(Comma Separated Values (.csv), 6.69 KB) MD5:083988091e28be94eecde239bb690d38

[table of contents | back to top]

Supplemental Files

File	
accession_numbers.pdf	(Portable Document Format (.pdf), 94.21 KB) MD5:73cec5bf026cecfc9e8ad60078d71a63
character_definitions.pdf	(Portable Document Format (.pdf), 112.56 KB) MD5:a815274ffc59f52b99d37ed78d11c6cc
IQtrees.pdf	(Portable Document Format (.pdf), 210.59 KB) MD5:2b0d24fee876f1cf9729d79d8b89fe57
R_Code.R	(R Script, 56.82 KB) MD5:f7a46e2fc405a954e4c0753bd4a6251e
RevBayes_topology_scripts.pdf	(Portable Document Format (.pdf), 43.35 KB) MD5:088cd82f1cf35906e9a563857fb585c0
RevBayestimetrees.pdf	(Portable Document Format (.pdf), 181.04 KB) MD5:a5dee3bf174c7c61e308fd575d180ea1

[table of contents | back to top]

Related Publications

Damian-Serrano, A., Haddock, S., & Dunn, C. (2020). *Data and code for: The evolution of siphonophore tentilla for specialized prey capture in the open ocean* (Version 7) [Data set]. Dryad. https://doi.org/10.5061/DRYAD.P2NGF1VP2 https://doi.org/10.5061/dryad.p2ngf1vp2
Results

[table of contents | back to top]

Parameters

Parameters for this dataset have not yet been identified

[table of contents | back to top]

Project Information

Collaborative research: The effects of predator traits on the structure of oceanic food webs (SiphWeb)

Coverage: North Pacific

Food webs describe who eats whom, tracing the flow of energy from plants up to large animals. While many connections in food webs on land are quite familiar (lions eat antelope and antelope eat grass, for example), there are large gaps in our understanding of ocean food webs. Closing these gaps is critical to understanding how nutrients and energy move through ocean ecosystems, how organisms interact in the ocean, and how best to manage ocean resources. This project will study ocean food web structure with a focus on siphonophores, an abundant group of predators in the open ocean that range in length from less than an inch to more than one hundred feet. Siphonophores are closely related to corals and many jellyfish. They are known to be important predators within ocean food webs, but they are difficult to study because they live across great ocean depths and are gelatinous and fragile. The details of what they eat, as well as many other features of their biology, remain poorly known. This project will combine direct observations of feeding, genetic analysis of siphonophore gut contents, and stable isotope analyses to identify what different species of siphonophores eat. The team will also examine why they eat what they do. This will provide a new understanding of how the structure of food webs arise, aiding in our ability to predict future changes to food webs as the global climate shifts. Siphonophores feed in a very unique manner--they have highly specialized tentacles that are used solely for capturing prey--thus, the prey captured is determined largely by the anatomy and function of these tentacles. The project will describe these tentacles, reconstruct their evolutionary history, and investigate how evolutionary shifts in tentacle structure have led to changes in diet. This project will train one PhD student, one Master's student, a postdoc, and undergraduate students, including individuals of underrepresented groups. This project will support the production of scientifically rigorous yet engaging videos, foster the expansion of a citizen-science program, and create K-12 teaching modules.

This project will advance three scientific aims: First, it will identify the diet of a diverse range of siphonophores using DNA metabarcoding of gut contents and prey field, remotely operated vehicle (ROV) video of prey encounters, and stable isotope analysis. These approaches are highly complementary and allow for extensive cross validation. Second, the project will characterize the selectivity of siphonophore diets by comparing them to the relative prey abundances in the habitats of each of these species. Third, the project will characterize the structure of the siphonophore prey capture apparatus across species through detailed morphological analysis of their tentacles and nematocysts. These data will be integrated in an ecological and evolutionary framework to identify predator features associated with prey specialization. In a larger context, addressing these questions will advance our understanding of oceanic predation by revealing how evolutionary changes in predator selectivity correspond to evolutionary changes in habitat and feeding apparatus and how these changes shape current food web structure in the open ocean. We will test and refine an integrated approach to describing the structure and origin of food web topology, and evaluate the potential for phylogenetic relationships to explain prey selectivity.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1829835

[table of contents | back to top]