



## An assessment of sampling biases across studies of diel activity patterns in marine ray-finned fishes (Actinopterygii)

<sup>1</sup> North Carolina Museum of Natural Sciences, Raleigh, North Carolina 27601.

<sup>2</sup> Arnold Arboretum of Harvard University, Boston, Massachusetts 02131.

<sup>3</sup> Florida Atlantic University, Wilkes Honors College, Jupiter, Florida 33458.

<sup>4</sup> Florida Atlantic University, Harbor Branch Oceanographic Institution, Fort Pierce, Florida 34946.

<sup>5</sup> Macquarie University, North Ryde, New South Wales, Australia 2109.

<sup>6</sup> Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543.

\* Corresponding author  
email: <alex.dornburg@naturalsciences.org>, telephone: (919) 707-8866.

**Alex Dornburg**<sup>1\*</sup>  
**Elisabeth J Forrester**<sup>2</sup>  
**Jon A Moore**<sup>3,4</sup>  
**Teresa L Iglesias**<sup>5</sup>  
**Andrew Jones**<sup>6</sup>  
**Leela Rao**<sup>1</sup>  
**Dan L Warren**<sup>5</sup>

---

**ABSTRACT.**—Understanding the promotion and regulation of circadian rhythms in marine fishes is important for studies spanning conservation, evolutionary biology, and physiology. Given numerous challenges inherent to quantifying behavioral activity across the full spectrum of marine environments and fish biodiversity, case studies offer a tractable means of gaining insights or forecasting broad patterns of diel activity. As these studies continue to accumulate, assessing whether, and to what extent, the cumulatively collected data are biased in terms of geography, habitat, or taxa represents a fundamentally important step in the development of a broad overview of circadian rhythms in marine fish. As such investigations require a phylogenetic framework, general trends in the phylogenetic sampling of marine fishes should be simultaneously assessed for biases in the sampling of taxa and trait data. Here, we compile diel activity data for more than 800 marine species from more than five decades of scientific studies to assess general patterns of bias. We found significant geographic biases that largely reflect a preference toward sampling warm tropical waters. Additionally, taxonomic biases likewise reflect a tendency toward conspicuous reef associated clades. Placing these data into a phylogenetic framework that includes all known marine fishes revealed significant under-dispersion of behavioral data and taxon sampling across the whole tree, with a few subclades exhibiting significant over-dispersion. In total, our study illuminates substantial gaps in our understanding of diel activity patterns and highlights significant sampling biases that have the potential to mislead evolutionary or ecological analyses.

---

Date Submitted: 5 February, 2016.  
Date Accepted: 19 September, 2016.  
Available Online: 28 November, 2016.

Containing more than 30,000 described species (Eschmeyer and Fricke 2015), ray-finned fishes (Actinopterygii) represent what is arguably one of the most successful radiations in the evolutionary history of vertebrates. Half of this diversity, or one out of every four living vertebrate species, is found within the world's oceans (Carrete Vega and Wiens 2012) across a mosaic of habitats ranging from coral reefs to the polar seas (Nelson et al. 2016). This diversity and ubiquity of ray-finned fishes represents a wealth of information about the evolutionary process. Over the past several decades, continual advances in survey technology and efforts (Walker et al. 2000, Devine et al. 2006, Makris et al. 2006, Bollinger and Kline 2015, McIntyre et al. 2015) coupled with increased resolution of a ray-finned fish tree of life (Near et al. 2012b, 2013, Betancur-R et al. 2013, Miya et al. 2013, Rabosky et al. 2013, Sanciangco et al. 2016), suggest that the 21<sup>st</sup> century may be the time that some of the most vexing questions in ichthyology become answerable. However, the majority of data collected are often restricted to diurnal species. It is therefore unclear if this potential temporal bias is leaving our understanding of the ecology and evolution of nocturnal ray-finned fishes in the dark.

It is well known that day/light cycles act as zeitgebers for the circadian rhythms of marine fishes (Hobson 1965, 1975, Helfman 1986, Naylor 2005), lending support for the hypothesis that time (and its corresponding light cycle) acts as a primary component of the ecological niche in animals (Hut et al. 2012). Evidence for diversification dynamics corresponding to a "temporal niche" is readily apparent in the eyes of marine fishes, where adaptations to dim-light conditions have constrained optical and trophic diversity (Goatley et al. 2010, Schmitz and Wainwright 2011). True nocturnality has evolved independently within dozens of marine fish families (Pulcini et al. 2008, Goatley et al. 2010, Schmitz and Wainwright 2011, Brandl and Bellwood 2014), making ray-finned fishes an ideal clade with which to investigate the mechanisms underlying temporal niche transitions. Additionally, the local abundance of many species also makes them an excellent system for assessing how human-driven disturbances, such as light pollution or contaminants, affect circadian rhythms in wild species (Chepesiuk 2009, Gallaway et al. 2010). While the vast species-richness of marine ray-finned fishes presents a potential boon to research, it also presents the inordinate challenge of first quantifying activity patterns across a quarter of all living vertebrates.

Aside from the sheer number of species, challenges inherent to studying behavior in aquatic environments have presented major hurdles to quantifying diel activity patterns in marine fishes. These challenges range from safety limitations in scientific diving (Dardeau and McDonald 2007) to biases in species detection rates produced by differences in the efficiency of survey methods across taxa (Willis et al. 2000). Yet despite such challenges, the past 50 yrs have yielded a constant stream of insights into fish behavioral patterns (Hobson 1965, 1975, Lobel 1978, Horn 1980). In particular, technical improvements, such as the advent of cheaper and smaller acoustic transmitters for radio telemetry, have catalyzed a pulse of studies over the last two decades (Meyer et al. 2000, Arendt et al. 2001, Heupel et al. 2006, Fox and Bellwood 2011). As studies accumulate and we move toward a broader understanding of diel cycles in marine fishes, an overview of the sampling efforts underlying this research represents an important step in assessing what potential biases have accumulated along with these data.

Given the heterogeneity in logistical challenges associated with sampling different geographic regions, habitats, or taxa, research efforts to date should be expected to reflect a certain level of sampling bias. However, the extent of such potential biases remains unknown. This uncertainty raises numerous pressing questions. For example, asking “What regions or habitats are in the highest need of study?” is critical if we are to develop a broad overview of the circadian rhythms of marine fish. However, even within heavily studied habitats or regions, it is important to evaluate if there is a bias in which members of the community are included in the study. Likewise, there may be a global bias where either larger fishes or fishes feeding in higher trophic levels, such as the majority of commercially important species (Brodeur and Pearcy 1987, Brulé et al. 1994, Rooker 1995, Amundsen et al. 1999), are disproportionately studied, while smaller fishes that form the critical links in community food webs remain understudied. Answering such questions will not only provide insights into potential pitfalls that might occur when drawing generalities from existing data, but will also optimize the efficiency of ongoing and future work.

Concomitant with answers to such questions should be an assessment of whether available data depict any biases that may mislead comparative phylogenetic investigations. Are certain clades more represented than others? Are there biases correlated with ecology or geography within clades? With the rapid accumulation of phylogenetic studies resolving the backbone (Holcroft 2005, Holcroft and Wiley 2008, Near et al. 2012b, 2013, Betancur-R et al. 2013, Sanciangco et al. 2016), intra-ordinal (Miya et al. 2013, Chen et al. 2014, Dornburg et al. 2015a, Eytan et al. 2015, Near et al. 2015), and intra-familial relationships (Near et al. 2012a, Santini et al. 2013a,b, Dornburg et al. 2015b, Santini and Carnevale 2015) across most of the ray-finned fish tree of life, addressing such questions has become a possibility and represents an important first step toward understanding the evolution of circadian rhythms in vertebrates.

Here, we compile the results of 68 published studies of diel activity patterns across all marine ray-finned fishes to assess patterns of sampling bias. Specifically, we assess if body size, trophic ecology, geography, depth, or taxonomy are correlated with significant sampling biases. Integrating our results into a phylogenetic framework representing all epipelagic marine and brackish fishes, we further assess patterns of phylogenetic over- and under-dispersion in sampling both across the entire tree and between clades. In total, our results reveal several sampling biases across marine fishes that have not previously been acknowledged, highlighting several deficits in our knowledge of marine fish activity patterns.

## METHODS

**DATA ACQUISITION.**—We restricted our survey of the diel activity literature to species sampled in the 7822 taxa tree from Rabosky et al. (2013), which included 3703 marine fish lineages. This tree is currently the largest assembled phylogeny of ray-finned fishes to date and is based on a supermatrix of 13 genes constructed using a mega-phylogeny approach. Although this tree includes only a fraction of the total diversity of marine fishes, using this tree as a taxonomic guide makes attempting a survey of the published literature a tractable problem. For example, given the sheer number of marine ray-finned fishes, only 3 min of time spent querying the literature per species would require >800 hrs of total search time, yet obviously still not yield an adequate survey of published work. Additionally, this tree contains the majority

of marine fishes that have sequence data available on GenBank. Molecular phylogenetic studies often attempt to obtain as nearly complete taxon sampling as possible for major clades (e.g., “families” or “orders”), and make use of the wealth of genomic resources archived in natural history collections worldwide. This taxon sampling strategy makes it unlikely that a large number of diel activity pattern studies have been conducted on species not sampled in this phylogeny. Therefore, we believe the taxon sampling of this tree represents an opportune guide for our study.

Surveys of published studies were conducted using Google web search, Google Scholar, Google Books, FishBase, and the ISI Web of Knowledge portal. Combinations of the following keywords were entered as search terms: fish, diel, activity pattern, marine, ocean, teleost, foraging, feeding, nocturnal, diurnal, crepuscular, night, reef, sleep, active, and activity. In addition to these keywords, the above searches were repeated with combinations of higher taxonomic terms as keywords (e.g., Tetraodontiformes or Balistidae), as well as randomly selected individual species names from the topology of Rabosky et al. (2013). Our list of keywords was restricted to English and undoubtedly did not contain every possible keyword that could produce search engine hits for studies containing data on diel activity pattern. Regardless, the resulting survey spanned hundreds of fish species and a host of sampling methods ranging from telemetry to scuba-based visual census methods. The diversity of studies obtained reflects the use of the above search words in titles, abstracts, keywords, and other searchable portions of documents so that additional terms such as “telemetry” yielded no significant increase of hits that did not also include terms such as “diel” or “nocturnal.” As such, we argue that this approach to surveying the literature represents a reasonable approximation of data patterns in published studies concerning diel activity patterns.

Although many juvenile and mesopelagic fishes undergo daily vertical migrations (Roe and Badcock 1984, Benoit-Bird et al. 2001), we are primarily interested in changes occurring in taxa that experience shifts in light conditions over a 24-hr period. Therefore, we restrict our survey to adult fishes that occupy primarily the epipelagic and neritic realm (0–200 m). Taxa residing at depths >200 m were identified using trawl depth catch data, data from FishBase, and personal observations of one of the authors (JAM). These “deep-sea” fishes were “pruned” from the tree topology using functions in the APE (Paradis et al. 2004) and GEIGER (Harmon et al. 2008) packages in R.

**BIAS IN FAMILY REPRESENTATION.**—To assess which families were over- and underrepresented in sampling for diel activity data, we compared the frequency of appearance of each family in the activity data set to the overall representation of each family among epipelagic fishes on FishBase. First, we downloaded data for all epipelagic Actinopterygians using the *rfishbase* package (Boettiger et al. 2012), and used the family codes from that table to look up family names for each species. Species included in our study were all species within each family that have adult life stages in marine or brackish water, including anadromous and catadromous species. We then conducted a test for each family to assess whether that family’s representation in the diel activity data set (835 marine/brackish species) was significantly more or less than expected given their representation in the full data set containing 11,880 marine/brackish species. This was performed as a simple binomial test in R, with the number of “successes” representing the number of species from that family

that were represented in the activity data, the number of trials determined by total number of species in the activity data set, and the expected success rate for each family determined by the frequency of that family in the full data set from FishBase. For each family, we also estimated the standardized effect size using the following equation:

$$(F_o - F_e) / (\max\{F_o\} - \min\{F_o\}) \quad (\text{Eq. 1})$$

where  $F_o$  is the observed frequency,  $F_e$  is the expected frequency, and the  $\max\{F_o\} - \min\{F_o\}$  represents the width of confidence interval around observed frequency.

**BIAS AT THE SPECIES LEVEL.**—We also conducted a series of Monte Carlo simulations to look for statistically significant biases in data collection on diel activity patterns at the species level. As above, we restrict these analyses to epipelagic species, resulting in a total of 835 marine/brackish species for which we have activity data, and 3005 marine/brackish species for which we have FishBase and phylogenetic data. To examine geographic biases in data collection on activity patterns, we retrieved the contents of the “countries” and “ecosystems” tables linked to individual species’ description pages on FishBase. Despite the title, the areas listed in the “countries” table do not necessarily correspond to countries per se (e.g., “Alaska” and “Hawaii” are listed separately from the United States), instead corresponding to 283 geographically distinct areas of occurrence. Similarly, the FishBase definition of “ecosystem” is somewhat biologically unrealistic, with, for example, “USA” being defined as a single ecosystem. To a large extent, the 504 distinct FishBase “ecosystem” divisions primarily represent geographically (rather than biologically) distinct areas, albeit at a finer scale than do the “countries” designations.

For each of the geographic areas and ecosystems, we calculated the total number of species known to occur in an area for which diel activity patterns were known. To look for patterns of under- and overrepresentation, we compared these observed species counts to a null distribution constructed using Monte Carlo methods. For each of 1000 replicates, we sampled 835 species at random from the full data set and calculated the number of species observed that occur in each country or ecosystem. Using the species counts from these simulations, we were able to construct 95% confidence intervals on the expected number of species observed that occur in each country or ecosystem if species were sampled at random.

We also extracted data on size, depth distribution, maximum absolute latitude, and trophic level from FishBase pages for each species. We conducted Monte Carlo simulations to examine sampling biases based on these factors. First, each variable was split into 20 bins of equal width, and the frequency of species in each bin for which we had activity data was calculated. To estimate confidence intervals for the null distribution, we randomly selected the same number of species from the full data set and calculated the frequency in each bin, repeating this procedure 1000 times.

**PHYLOGENETIC BIASES IN SPECIES AND TRAIT SAMPLING.**—To test for phylogenetic biases in species sampling, we utilized the Rabosky et al. (2013) phylogeny and added species or tips to clades for which taxonomic membership is well-established, primarily representing families. Family-level diversity was restricted to brackish and marine species and their associated taxonomic information from FishBase. This

assignment of diversity prevented inflating diversity for clades spanning fresh and saltwater, such as Plotosidae, while also allowing us to capture lineages where adults or juveniles transitioning to adulthood migrate into brackish or marine water from fresh water. Tips were added to the tree with branch lengths of zero and polytomies were retained, as a fully bifurcating tree is unnecessary for analyses of phylogenetic over- and under-dispersion. This procedure yielded a completely sampled phylogeny of all 11,880 described marine epipelagic taxa. A second tree containing only reef fishes was also produced by “pruning” the tree of all marine epipelagic fishes to one with only species identified as reef associated on FishBase, resulting in a 4742 taxa tree. Tests of over- and under-dispersion for taxon and diel activity trait sampling were conducted for the complete epipelagic and reef fishes trees using the `mpd.ses` function in the `picante` library of R (Kembel et al. 2010), which measures the standardized mean pair-wise phylogenetic distances of species within a community, or in this case, those species which have been sampled (Webb et al. 2008). To assess significance, we generated a null distribution of 999 samples where the tip states were shuffled across the tree. Tests of node-based phylogenetic clustering of sampling were conducted using the clade significance test (JC Oliver et al. unpubl data), that has been used in recent studies of trait clustering (Forrestel et al. 2014, 2015). This test builds upon the `nodesig` test in `PhyloCom` (Webb et al. 2008) by calculating the clade density of a state for each node, allowing for the identification of clades within which descendant lineages are characterized by similar densities that differ substantially from the densities within their sister clade. This test was used to identify nodes and descendant taxa or clades with significant clustering of over or under sampling of taxa and diel activity. Significance was assessed using a null model where tip states were shuffled and a one-tailed test of significant ( $P < 0.05$ ) clustering was conducted independently for each state (i.e., sampled or unsampled).

## RESULTS

**BIAS AT THE FAMILY LEVEL.**—The results of the binomial tests for family representation are given in Table 1. Negative standardized effect sizes indicate families that are underrepresented in the diel activity data set compared to their prevalence in the epipelagic fauna at large, while positive effect sizes represent families that are relatively highly sampled for activity data. We found 29 families that were sampled significantly less often than expected, the majority of which represent difficult to detect lineages such as blennies, gobies, eels, and flatfishes. In contrast, we found 23 families that were sampled more often than expected, the majority of which are conspicuous members of the reef community, such as squirrelfishes, triggerfishes, and wrasses (Table 1).

**BIAS AT THE SPECIES LEVEL.**—Results of the simulation study of geographic sampling bias are given in Tables 2 and 3. Results for the two analyses are qualitatively similar; countries and ecosystems that are sampled at a significantly higher rate than expected tend to be small, tropical islands, particularly in the Indo-Pacific region. In fact, all overrepresented countries or ecosystems occurred in warm waters, with the tropical Indo-West Pacific biogeographic region (Briggs and Bowen 2012) accounting for the majority of areas (Tables 2, 3). Areas that are underrepresented include a mix of warm and cool regions that included North America, parts of Europe, as well

Table 1. Results of bias simulations based on family level taxonomy. Results are sorted by mean effect size with lower effect size values indicating less representation. Bold rows indicate families that are significantly over- or underrepresented. CI = confidence intervals.

Family	<i>P</i>	Lower CI	Upper CI	Effect Size
<b>Syngnathidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-3.3815</b>
<b>Tripterygiidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-3.2884</b>
<b>Callionymidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-3.2119</b>
<b>Ophichthidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-3.1907</b>
<b>Bythitidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-3.1354</b>
<b>Soleidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-3.0016</b>
<b>Gobiidae</b>	<b>&lt;0.0001</b>	<b>0.0263</b>	<b>0.0536</b>	<b>-2.9751</b>
<b>Pseudochromidae</b>	<b>&lt;0.0001</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-2.9060</b>
<b>Cyprinidae</b>	<b>0.0003</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-2.4089</b>
<b>Triglidae</b>	<b>0.0012</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-2.0074</b>
<b>Chaenopsidae</b>	<b>0.0028</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.8354</b>
<b>Gobiesocidae</b>	<b>0.0005</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.7276</b>
<b>Liparidae</b>	<b>0.0038</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.6824</b>
<b>Microdesmidae</b>	<b>0.0057</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.6060</b>
<b>Engraulidae</b>	<b>0.0015</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.5368</b>
<b>Platycephalidae</b>	<b>0.0087</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.5295</b>
<b>Batrachoididae</b>	<b>0.0083</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.4912</b>
<b>Bothidae</b>	<b>0.0022</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.4604</b>
<b>Sciaenidae</b>	<b>0.0004</b>	<b>0.0019</b>	<b>0.0139</b>	<b>-1.3513</b>
<b>Ariidae</b>	<b>0.0046</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.3459</b>
<b>Blenniidae</b>	<b>0.0002</b>	<b>0.0058</b>	<b>0.0219</b>	<b>-1.3192</b>
<b>Exocoetidae</b>	<b>0.0182</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.3001</b>
<b>Cynoglossidae</b>	<b>0.0044</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.2950</b>
<b>Nemipteridae</b>	<b>0.0176</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.2618</b>
<b>Zoarcidae</b>	<b>0.0176</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.2618</b>
<b>Congridae</b>	<b>0.0182</b>	<b>0.0000</b>	<b>0.0066</b>	<b>-1.0406</b>
<b>Salmonidae</b>	<b>0.0379</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-1.0324</b>
<b>Plesiopidae</b>	<b>0.0565</b>	<b>0.0000</b>	<b>0.0044</b>	<b>-0.9559</b>
Dactyloscopidae	0.0898	0.0000	0.0044	-0.9177
Opistognathidae	0.0521	0.0000	0.0066	-0.8370
Ophidiidae	0.0507	0.0000	0.0066	-0.7988
<b>Apogonidae</b>	<b>0.0118</b>	<b>0.0074</b>	<b>0.0249</b>	<b>-0.7911</b>
Malacanthidae	0.1280	0.0000	0.0044	-0.7839
Mugilidae	0.0723	0.0000	0.0066	-0.7480
<b>Cottidae</b>	<b>0.0419</b>	<b>0.0013</b>	<b>0.0122</b>	<b>-0.7194</b>
Stichaeidae	0.1054	0.0000	0.0066	-0.7098
Tetrarogidae	0.1227	0.0000	0.0044	-0.7074
Synanceiidae	0.1921	0.0000	0.0044	-0.6883
Sillaginidae	0.1857	0.0000	0.0044	-0.6500
Uranoscopidae	0.1857	0.0000	0.0044	-0.6500
Clinidae	0.1037	0.0003	0.0086	-0.6024
Pinguipedidae	0.1962	0.0003	0.0086	-0.5215
Cheilodactylidae	0.2743	0.0000	0.0044	-0.5162
Ogocephalidae	0.2743	0.0000	0.0044	-0.5162
Eleotridae	0.1900	0.0003	0.0086	-0.4811
Gerreidae	0.2832	0.0000	0.0066	-0.4808

Table 1. Continued.

Family	<i>P</i>	Lower CI	Upper CI	Effect Size
Nototheniidae	0.2832	0.0000	0.0066	-0.4808
Leiognathidae	0.2751	0.0000	0.0066	-0.4299
Achiridae	0.4156	0.0000	0.0044	-0.4206
Cyclopteridae	0.4133	0.0000	0.0044	-0.4015
Synodontidae	0.3470	0.0003	0.0086	-0.3901
Chlopsidae	0.6553	0.0000	0.0044	-0.3824
Samaridae	0.6487	0.0000	0.0044	-0.3633
Creediidae	0.6429	0.0000	0.0044	-0.3441
Mullidae	0.3964	0.0007	0.0105	-0.3413
Labrisomidae	0.3781	0.0019	0.0139	-0.3305
Cyprinodontidae	0.6381	0.0000	0.0044	-0.3250
Neosebastidae	0.6381	0.0000	0.0044	-0.3250
Plotosidae	0.6381	0.0000	0.0044	-0.3250
Polynemidae	0.5309	0.0000	0.0066	-0.2899
Fundulidae	0.6324	0.0000	0.0044	-0.2868
Moringuidae	0.6324	0.0000	0.0044	-0.2868
Pholidae	0.6324	0.0000	0.0044	-0.2868
Stromateidae	0.6324	0.0000	0.0044	-0.2868
Muraenidae	0.3441	0.0058	0.0219	-0.2863
Scorpaenidae	0.3719	0.0041	0.0188	-0.2838
Pristigasteridae	0.7381	0.0000	0.0066	-0.2518
Pentacerotidae	1.0000	0.0000	0.0044	-0.2485
Monacanthidae	0.4629	0.0019	0.0139	-0.2460
Aracanidae	1.0000	0.0000	0.0044	-0.2294
Bregmacerotidae	1.0000	0.0000	0.0044	-0.2294
Callanthiidae	1.0000	0.0000	0.0044	-0.2294
Centropomidae	1.0000	0.0000	0.0044	-0.2294
Paralichthyidae	0.5723	0.0019	0.0139	-0.2249
Galaxiidae	1.0000	0.0000	0.0044	-0.2103
Harpagiferidae	1.0000	0.0000	0.0044	-0.2103
Istiophoridae	1.0000	0.0000	0.0044	-0.2103
Phycidae	1.0000	0.0000	0.0044	-0.2103
Clupeidae	0.6373	0.0041	0.0188	-0.1975
Gasterosteidae	1.0000	0.0000	0.0044	-0.1912
Synphobranchidae	1.0000	0.0000	0.0044	-0.1912
Trachichthyidae	1.0000	0.0000	0.0044	-0.1912
Trichonotidae	1.0000	0.0000	0.0044	-0.1912
Atherinidae	0.7789	0.0003	0.0086	-0.1879
Champsodontidae	1.0000	0.0000	0.0044	-0.1721
Nomeidae	1.0000	0.0000	0.0044	-0.1721
Agonidae	0.7759	0.0003	0.0086	-0.1575
Hapalogenyidae	1.0000	0.0000	0.0044	-0.1529
Hemitripterae	1.0000	0.0000	0.0044	-0.1529
Kraemeriidae	1.0000	0.0000	0.0044	-0.1529
Trachinidae	1.0000	0.0000	0.0044	-0.1529
Hemiramphidae	0.8053	0.0007	0.0105	-0.1506
Ostraciidae	1.0000	0.0000	0.0066	-0.1373



Table 1. Continued.

Family	<i>P</i>	Lower CI	Upper CI	Effect Size
Paraulopidae	1.0000	0.0000	0.0044	-0.1338
Triacanthidae	1.0000	0.0000	0.0044	-0.1338
Bembridae	1.0000	0.0000	0.0044	-0.1147
Monodactylidae	1.0000	0.0000	0.0044	-0.1147
Retropinnidae	1.0000	0.0000	0.0044	-0.1147
Scophthalmidae	1.0000	0.0000	0.0044	-0.1147
Ariommatidae	1.0000	0.0000	0.0044	-0.0956
Gonorynchidae	1.0000	0.0000	0.0044	-0.0956
Latridae	1.0000	0.0000	0.0044	-0.0956
Leptoscopidae	1.0000	0.0000	0.0044	-0.0956
Pegasidae	1.0000	0.0000	0.0044	-0.0956
Scomberesocidae	1.0000	0.0000	0.0044	-0.0956
Scombridae	1.0000	0.0007	0.0105	-0.0899
Acropomatidae	1.0000	0.0000	0.0044	-0.0765
Bathydraconidae	1.0000	0.0000	0.0044	-0.0765
Citharidae	1.0000	0.0000	0.0044	-0.0765
Cryptacanthodidae	1.0000	0.0000	0.0044	-0.0765
Lophiidae	1.0000	0.0000	0.0044	-0.0765
Molidae	1.0000	0.0000	0.0044	-0.0765
Monocentridae	1.0000	0.0000	0.0044	-0.0765
Scatophagidae	1.0000	0.0000	0.0044	-0.0765
Anarhichadidae	1.0000	0.0000	0.0044	-0.0574
Drepaneidae	1.0000	0.0000	0.0044	-0.0574
Psettodidae	1.0000	0.0000	0.0044	-0.0574
Rhyacichthyidae	1.0000	0.0000	0.0044	-0.0574
Terapontidae	1.0000	0.0000	0.0066	-0.0482
Chirocentridae	1.0000	0.0000	0.0044	-0.0382
Dichistiidae	1.0000	0.0000	0.0044	-0.0382
Dinopercidae	1.0000	0.0000	0.0044	-0.0382
Lateolabracidae	1.0000	0.0000	0.0044	-0.0382
Lobotidae	1.0000	0.0000	0.0044	-0.0382
Pseudaphritidae	1.0000	0.0000	0.0044	-0.0382
Pseudotriconotidae	1.0000	0.0000	0.0044	-0.0382
Trichodontidae	1.0000	0.0000	0.0044	-0.0382
Veliferidae	1.0000	0.0000	0.0044	-0.0382
Antennariidae	1.0000	0.0007	0.0105	-0.0379
Pleuronectidae	1.0000	0.0019	0.0139	-0.0348
Tetraodontidae	1.0000	0.0066	0.0234	-0.0234
Artedidraconidae	1.0000	0.0000	0.0044	-0.0191
Centrogenyidae	1.0000	0.0000	0.0044	-0.0191
Cheimarrichthyidae	1.0000	0.0000	0.0044	-0.0191
Dinolestidae	1.0000	0.0000	0.0044	-0.0191
Eleginopsidae	1.0000	0.0000	0.0044	-0.0191
Enoplosidae	1.0000	0.0000	0.0044	-0.0191
Leptobramidae	1.0000	0.0000	0.0044	-0.0191
Menidae	1.0000	0.0000	0.0044	-0.0191
Normanichthyidae	1.0000	0.0000	0.0044	-0.0191

Table 1. Continued.

Family	<i>P</i>	Lower CI	Upper CI	Effect Size
Pomatomidae	1.0000	0.0000	0.0044	-0.0191
Ptilichthyidae	1.0000	0.0000	0.0044	-0.0191
Rachycentridae	1.0000	0.0000	0.0044	-0.0191
Rhamphocottidae	1.0000	0.0000	0.0044	-0.0191
Triodontidae	1.0000	0.0000	0.0044	-0.0191
Xiphiidae	1.0000	0.0000	0.0044	-0.0191
Zaproridae	1.0000	0.0000	0.0044	-0.0191
Atherinopsidae	1.0000	0.0007	0.0105	-0.0119
Osmeridae	0.6268	0.0000	0.0066	0.0027
Sphyraenidae	0.7158	0.0003	0.0086	0.0144
Grammatidae	0.5996	0.0000	0.0066	0.0154
Bovichtidae	0.5390	0.0000	0.0066	0.0408
Kuhliidae	0.5390	0.0000	0.0066	0.0408
Muraenesocidae	0.4693	0.0000	0.0066	0.0663
Echeneidae	0.4306	0.0000	0.0066	0.0790
Elopidae	0.3890	0.0000	0.0066	0.0917
Oplegnathidae	0.3890	0.0000	0.0066	0.0917
Sparidae	0.6247	0.0066	0.0234	0.0966
Bathymasteridae	0.3445	0.0000	0.0066	0.1045
Chironemidae	0.3445	0.0000	0.0066	0.1045
Latidae	0.3445	0.0000	0.0066	0.1045
Anomalopidae	0.2967	0.0000	0.0066	0.1172
Centriscidae	0.2967	0.0000	0.0066	0.1172
Fistulariidae	0.2454	0.0000	0.0066	0.1299
Glaucosomatidae	0.2454	0.0000	0.0066	0.1299
Coryphaenidae	0.1313	0.0000	0.0066	0.1554
Megalopidae	0.1313	0.0000	0.0066	0.1554
Pholidichthyidae	0.1313	0.0000	0.0066	0.1554
Hexagrammidae	0.2072	0.0003	0.0086	0.1660
Odacidae	0.2072	0.0003	0.0086	0.1660
Chanidae	0.0680	0.0000	0.0066	0.1681
Nematistiidae	0.0680	0.0000	0.0066	0.1681
Zanclidae	0.0680	0.0000	0.0066	0.1681
Albulidae	0.1819	0.0003	0.0086	0.1761
Acipenseridae	0.2032	0.0007	0.0105	0.1788
Pomacanthidae	0.3086	0.0049	0.0203	0.2125
Anguillidae	0.1352	0.0007	0.0105	0.2135
Dactylopteridae	0.0879	0.0003	0.0086	0.2166
<b>Aplodactylidae</b>	<b>0.0491</b>	<b>0.0003</b>	<b>0.0086</b>	<b>0.2368</b>
Ephippidae	0.0907	0.0007	0.0105	0.2395
Gadidae	0.0907	0.0007	0.0105	0.2395
<b>Arripidae</b>	<b>0.0329</b>	<b>0.0003</b>	<b>0.0086</b>	<b>0.2469</b>
<b>Moronidae</b>	<b>0.0329</b>	<b>0.0003</b>	<b>0.0086</b>	<b>0.2469</b>
Pomacentridae	0.2429	0.0273	0.0550	0.2493
<b>Aulostomidae</b>	<b>0.0194</b>	<b>0.0003</b>	<b>0.0086</b>	<b>0.2570</b>
Pempheridae	0.0780	0.0019	0.0139	0.2749
Kyphosidae	0.1126	0.0034	0.0172	0.2773

Table 1. Continued.

Family	<i>P</i>	Lower CI	Upper CI	Effect Size
<b>Diodontidae</b>	<b>0.0397</b>	<b>0.0013</b>	<b>0.0122</b>	<b>0.3000</b>
<b>Cirrhitidae</b>	<b>0.0097</b>	<b>0.0034</b>	<b>0.0172</b>	<b>0.4054</b>
<b>Belonidae</b>	<b>0.0062</b>	<b>0.0041</b>	<b>0.0188</b>	<b>0.4355</b>
<b>Lotidae</b>	<b>&lt;0.0001</b>	<b>0.0019</b>	<b>0.0139</b>	<b>0.4720</b>
<b>Priacanthidae</b>	<b>0.0005</b>	<b>0.0034</b>	<b>0.0172</b>	<b>0.4908</b>
<b>Caesionidae</b>	<b>0.0003</b>	<b>0.0041</b>	<b>0.0188</b>	<b>0.5218</b>
<b>Siganidae</b>	<b>0.0003</b>	<b>0.0049</b>	<b>0.0203</b>	<b>0.5404</b>
<b>Embiotocidae</b>	<b>&lt;0.0001</b>	<b>0.0066</b>	<b>0.0234</b>	<b>0.6566</b>
<b>Sebastidae</b>	<b>0.0001</b>	<b>0.0119</b>	<b>0.0324</b>	<b>0.6686</b>
<b>Lutjanidae</b>	<b>&lt;0.0001</b>	<b>0.0156</b>	<b>0.0381</b>	<b>0.7115</b>
<b>Haemulidae</b>	<b>&lt;0.0001</b>	<b>0.0194</b>	<b>0.0438</b>	<b>0.7672</b>
<b>Holocentridae</b>	<b>&lt;0.0001</b>	<b>0.0137</b>	<b>0.0353</b>	<b>0.7781</b>
<b>Carangidae</b>	<b>&lt;0.0001</b>	<b>0.0224</b>	<b>0.0480</b>	<b>0.8291</b>
<b>Lethrinidae</b>	<b>&lt;0.0001</b>	<b>0.0147</b>	<b>0.0367</b>	<b>0.9406</b>
<b>Balistidae</b>	<b>&lt;0.0001</b>	<b>0.0175</b>	<b>0.0410</b>	<b>1.0212</b>
<b>Serranidae</b>	<b>&lt;0.0001</b>	<b>0.0659</b>	<b>0.1046</b>	<b>1.1223</b>
<b>Scaridae</b>	<b>&lt;0.0001</b>	<b>0.0314</b>	<b>0.0605</b>	<b>1.2299</b>
<b>Acanthuridae</b>	<b>&lt;0.0001</b>	<b>0.0375</b>	<b>0.0687</b>	<b>1.4278</b>
<b>Labridae</b>	<b>&lt;0.0001</b>	<b>0.1105</b>	<b>0.1577</b>	<b>1.8904</b>
<b>Chaetodontidae</b>	<b>&lt;0.0001</b>	<b>0.0669</b>	<b>0.1059</b>	<b>1.8994</b>

as several areas that are generally considered difficult for travel purposes (e.g., North Korea, Iraq, Iran).

Results for the other Monte Carlo simulations are given in Figure 1A. Activity data availability does not show any readily discernible bias based on the size, trophic level, or depth distribution of the species. Maximum absolute latitude, however, shows a definitive pattern; species with very low maximum absolute latitudes (i.e., exclusively equatorial species) and species with very high maximum absolute latitudes (i.e., species extending into, if not necessarily exclusively found in, colder waters) tend to be under-sampled compared to expectations based on a random sampling of fish diversity (Fig. 1B, C). In contrast, we find a sharp peak in the frequency of sampled species with maximum absolute latitudes between approximately 28° and 33° (Fig. 1A). In part this represents the large number of species with latitudinal limits in this range (gray ribbon, Fig. 1A); however, the bias toward sampling these species exceeds even that expected given their higher representation in the full data set. Given this result and the results from the geographic bias test above, we find support for a general trend toward increased representation of widespread, tropical species.

**PHYLOGENETIC BIASES.**—Significant biases in taxon sampling were found across the entire tree of marine fishes (Fig. 2A). In terms of taxon sampling, these biases were not restricted to a certain region of the phylogeny and spanned clades such as notothenioids, damselfishes, parrotfishes, and moray eels. In total, 45 named clades exhibited significant clustering of sampled taxa relative to the taxon sampling present throughout the rest of the phylogeny (Fig. 2A). This high degree of bias was also reflected in a global calculation supporting significant under-dispersion across the phylogeny (Table 4). Global under-dispersion was also found for behavioral sampling patterns. These sampling patterns were similar to taxon sampling patterns, though

Table 2. Significantly over- or under-represented provinces based on boundaries defined on FishBase. Biogeographic regions and provinces that are included in the political boundaries of each "country" highlighted are based on Briggs and Bowen (2012). IWP = tropical Indo-West Pacific; WNP = western North Pacific; WP = western Pacific; ENP = eastern North Pacific; WA = western Atlantic; EP = eastern Pacific; SA = South American. Countries with multiple regions are listed in multiple rows. Numbers in parentheses indicate that a region had multiple provinces.

Country	Overrepresented countries			Underrepresented countries		
	<i>P</i>	Region	Province	<i>P</i>	Region	Province
Comoros	0.019	Warm: IWP	Western Indian Ocean	0.001	Cool: WNP	Oriental
Fiji	<0.001	Warm: IWP	Indo-Polynesian	0.001	Warm: WP	Sino-Japanese
Indonesia	0.004	Warm: IWP	Indo-Polynesian	0.009	Warm: IWP	Western Indian Ocean
Mauritius	0.001	Warm: IWP	Western Indian Ocean	0.013	Warm: IWP	Western Indian Ocean
Micronesia	<0.001	Warm: IWP	Indo-Polynesian	0.001	Cool: ENP	Aluetian
New Caledonia	0.001	Warm: IWP	Indo-Polynesian	0.001	Cool: WA	N/A
Palau	<0.001	Warm: IWP	Indo-Polynesian	0.001	Warm: WA (2)	Carolina, Caribbean
Papua New Guinea	0.011	Warm: IWP	Indo-Polynesian	0.005	Warm: IWP	Indo-Polynesian
Philippines	0.006	Warm: IWP	Indo-Polynesian	0.001	Cool: ENP (2)	Aluetian, Oregon
Reunion	0.005	Warm: IWP	Western Indian Ocean	0.001	Arctic	N/A
Tonga	0.000	Warm: IWP	Indo-Polynesian	0.001	Cool: WA	N/A
Madagascar	0.019	Warm: IWP	Western Indian Ocean	0.006	Cool: WNP (2)	Kurile, Okhotsk
Seychelles	<0.001	Warm: IWP	Western Indian Ocean	0.006	Arctic	N/A
American Samoa	0.003	Warm: IWP	Indo-Polynesian	0.007	Warm: IWP	Indo-Polynesian
Cargados Carajos	0.004	Warm: IWP	Western Indian Ocean	0.013	Warm: IWP	Western Indian Ocean
Chagos Is.	0.002	Warm: IWP	Indo-Polynesian	0.009	Warm: IWP	Western Indian Ocean
Cook Is.	0.011	Warm: IWP	Indo-Polynesian	0.001	Warm: IWP	Indo-Polynesian
French Polynesia	<0.001	Warm: IWP	Indo-Polynesian	0.018	Cool: WNP	Oriental
Guam	<0.001	Warm: IWP	Indo-Polynesian	0.008	Cool: WNP (2)	Kurile, Okhotsk
Hawaii	0.003	Warm: IWP	Hawaiian	0.008	Arctic	N/A
Maldives	0.001	Warm: IWP	Indo-Polynesian	0.021	Cool: EA	Eastern Atlantic Boreal

Table 2. Continued.

Country	Overrepresented countries			Underrepresented countries			
	<i>P</i>	Region	Province	Country	<i>P</i>	Region	Province
Marshall Is.	0.000	Warm: IWP	Indo-Polynesian	Chile	0.009	Warm: EP	Peru–Chilean
North Marianas	0.004	Warm: IWP	Indo-Polynesian	Chile	0.009	Cool: SA (2)	Southern Chile, Tierra del Feugo
Ogasawara Is.	0.020	Warm: IWP	Indo-Polynesian	Kuril Is.	0.015	Cool: WNP (2)	Kurile, Okhotsk
Panama	0.018	Warm: EP	Panamanian	Midway Is.	0.010	Warm: IWP	Indo-Polynesian
Panama	0.018	Warm: WA	Caribbean				
Samoa	<0.001	Warm: IWP	Indo-Polynesian				
Ryukyu Is.	<0.001	Warm: IWP	Indo-Polynesian				
Christmas Is.	<0.001	Warm: IWP	Indo-Polynesian				
Cocos Is. (Keel)	0.001	Warm: EP	Panamanian				
Kiribati	0.004	Warm: IWP	Indo-Polynesian				
Tuamotu Is.	0.001	Warm: IWP	Indo-Polynesian				
Pitcairn	0.004	Warm: IWP	Indo-Polynesian				
Revillagigedo	0.002	Warm: EP	Panamanian				
US Minor Is.	0.015	Warm: IWP	Indo-Polynesian				
Rodriguez	0.012	Warm: IWP	Western Indian Ocean				

Table 3. Significantly over- or under-represented ecosystems based on boundaries defined on FishBase. Biogeographic regions and provinces that are included within the boundaries of each ecosystem highlighted below are based on Briggs and Bowen (2012). IWP = tropical Indo-West Pacific; WP = western Pacific; WA = western Atlantic; EP = eastern Pacific; ENP = eastern North Pacific; WNP = western North Pacific. Countries with multiple regions are listed in multiple rows. Numbers in parentheses indicate that a region had multiple provinces.

Ecosystem	Overrepresented ecosystems			Underrepresented ecosystems			
	<i>P</i>	Region	Province	Ecosystem	<i>P</i>	Region	Province
Calamianes Islands	0.019	Warm: IWP	Indo-Polynesian	East China Sea	0.008	Warm: IWP	Indo-Polynesian
Coral Sea and GBR	0.001	Warm: IWP	Indo-Polynesian	East China Sea	0.008	Warm: WP	Sino-Japanese
Great Barrier Reef	0.001	Warm: IWP	Indo-Polynesian	NE US Continental Shelf	0.013	Cool: WA	N/A
Milne Bay	0.001	Warm: IWP	Indo-Polynesian	Sikao Creek	0.014	Warm: IWP	Indo-Polynesian
Verde Island Passage	0.009	Warm: IWP	Indo-Polynesian	California Current	0.001	Cool: ENP	Oregon
Gulf of California	0.001	Warm: EP	California	Sea of Japan	0.010	Cool: WNP	Kurile
Insular Pacific–Hawaiian	0.011	Warm: IWP	Hawaiian	Gulf of Alaska	0.003	Cool: ENP	Aluetian
Pacific Central–American Coastal	0.010	Warm: EP	Panamanian	Hecate Strait	0.010	Cool: ENP	Oregon
Polynesian Waters	0.002	Warm: IWP	Indo-Polynesian	Sakhalin Island	0.009	Cool: WNP	Okhotsk
Hsiao-liu-chiu	0.004	Warm: IWP	Indo-Polynesian	Salish Sea	0.001	Cool: ENP	Oregon
Lutao Island	0.008	Warm: WP	Sino-Japanese	Strait of Georgia	0.002	Cool: ENP	Oregon
Tahiti I.	0.001	Warm: IWP	Indo-Polynesian	Faroe Plateau	0.013	Cool: EA	N/A
Sogod Bay	0.009	Warm: IWP	Indo-Polynesian	Newfoundland-Labrador Shelf	0.001	Cool: WA	N/A
Tayabas Bay	0.006	Warm: IWP	Indo-Polynesian	Ionian Sea	0.003	Warm: EA	Lusitania
Spratly Islands	0.001	Warm: IWP	Indo-Polynesian				
Tawi tawi Bay	0.001	Warm: IWP	Indo-Polynesian				
Tubbataha Reefs	0.016	Warm: IWP	Indo-Polynesian				
Yun-gan	0.001	Warm: WP	Sino-Japanese				
Neotropical	0.014	Warm: EP	Panamanian				
Neotropical	0.014	Warm: WA (2)	Caribbean, Brazilian				

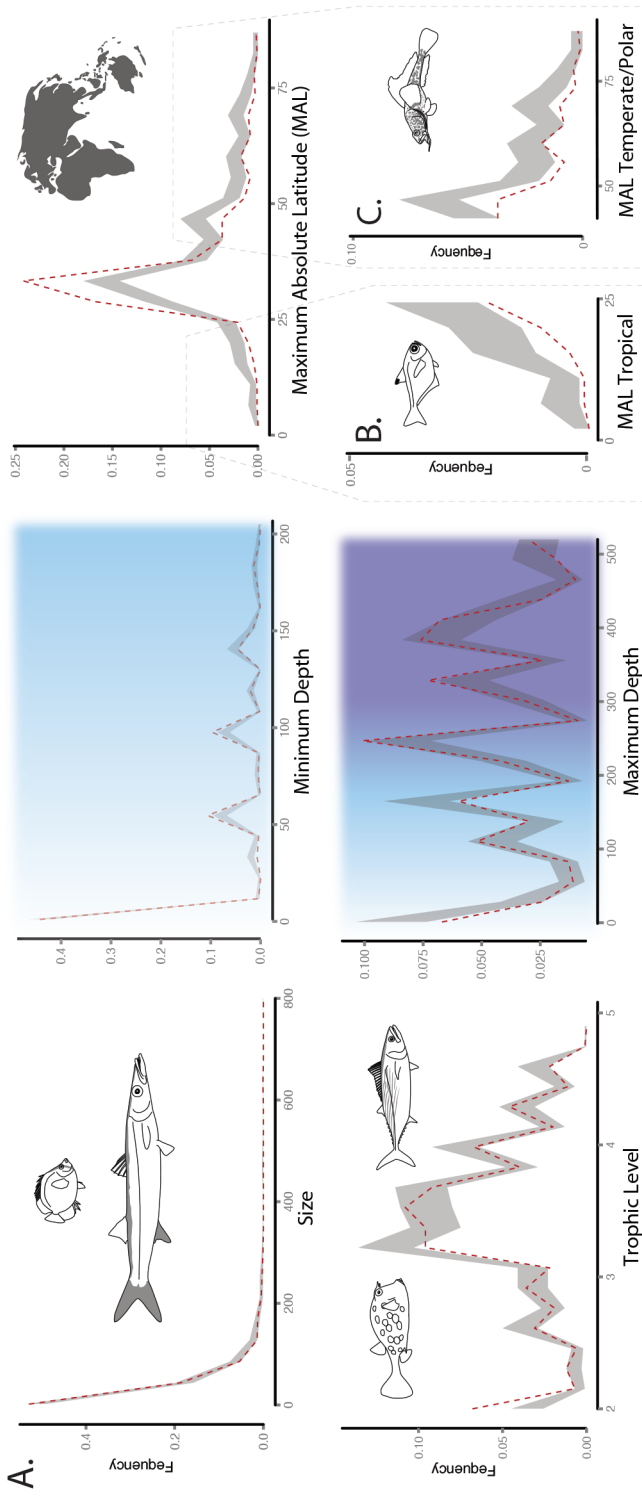


Figure 1. (A) Assessments of bias across potential categories. Gray bands represent null distributions, dotted lines represent empirical patterns. (B) Magnification of trends at low latitudes. (C) Magnification of trends at high latitudes.

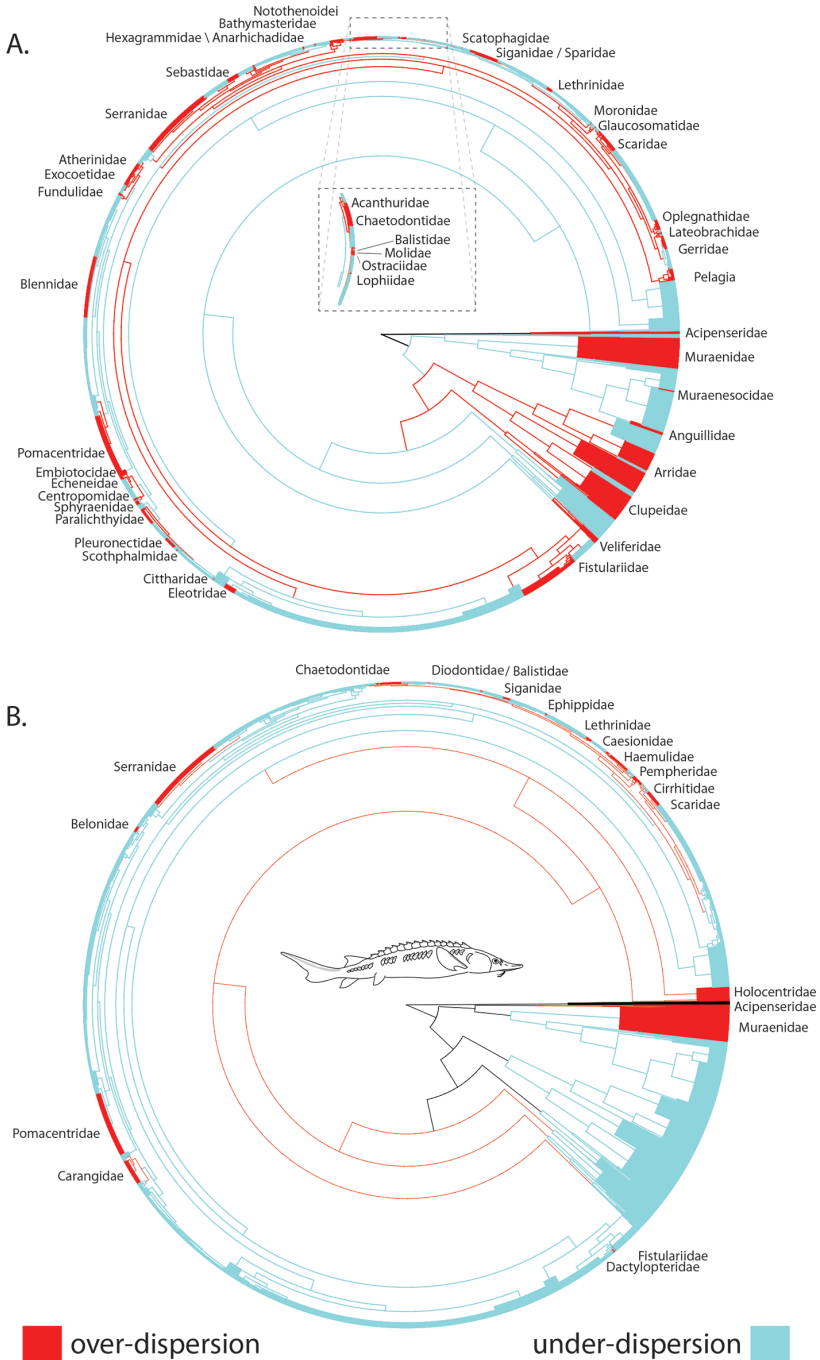


Figure 2. (A) Patterns of over (dark) and under-dispersion (light) in taxon sampling in the Rabosky et al. (2013) topology pruned to only marine species relative to a diversity tree containing all known marine fish species. (B) Patterns of over (dark) and under-dispersion (light) in activity data relative to a diversity tree containing all known marine fish species. Black lines in the phylogeny represent no deviation under or over dispersion of taxon sampling. Names represent significantly over-dispersed named clades. Illustration of a sturgeon (Acipenseridae) represents the earliest diverging lineage in our analyses.



Table 4. Patterns of phylogenetic over and under-sampling for all fishes and reef fishes, where MPD is the mean phylogenetic differences of all sampled taxa. obs = observations.

Tree	Bias	<i>n</i>	Observed MPD	Mean null model MPD	SD of null model MPD	Difference between obs	<i>P</i>
All fish	Activity	822	226.947	258.866	3.46764	-9.20485	<b>0.001</b>
All fish	Taxon	2,870	256.022	258.820	1.62375	-1.72312	<b>0.035</b>
Reef	Activity	671	214.174	235.555	2.75621	-7.75726	<b>0.001</b>
Reef	Taxon	1,346	227.726	235.549	1.72687	-4.53026	<b>0.001</b>

largely dominated by primarily tropically distributed clades. Clustering of behavioral trait sampling was also more restricted, with only 20 named clades exhibiting significant clustering of sampling while the majority of the nodes in the phylogeny exhibited significant under-sampling (Fig. 3B). Both taxon and behavioral data sampling patterns revealed the majority of internal nodes in the ray-finned fish phylogeny had clustering of under-sampled taxa, reflecting the concentration of bias in more tipward lineages and the overall low taxon and trait sampling.

Significant clustering of under-sampling in the phylogenetic and behavioral sampling of only reef-associated fishes was also detected (Table 4). In contrast to patterns across all marine fishes (Fig. 2), sampled taxa were significantly clustered for several deeply nested internal nodes in both the phylogenetic and behavioral sample sets (Fig. 3). This pattern reflects a larger clade-specific bias than that observed across all marine fishes. However, large portions of the reef fish topology exhibited significant clustering of under-sampling (Fig. 3). In terms of phylogenetic sampling, the majority of reef fish samples are distributed in moray eels, wrasses, parrotfishes, groupers, and damselfishes (Fig. 3A). Behavioral sampling biases of reef taxa closely mirror phylogenetic biases, though moray eels, wrasses, and damselfish are notably under-sampled (Fig. 3B).

## DISCUSSION

Quantification of phylogenetic and diel activity data sampling patterns revealed several significant and previously unacknowledged patterns of sampling bias that impose limits on our understanding of nocturnality and circadian rhythms in marine fishes. Diel activity data collected to date are dominated by wide-ranging tropical marine species, with a particular emphasis on species from the warm waters of the Pacific Ocean. This bias is largely driven by overrepresentation of conspicuous reef-associated fishes, such as squirrelfishes, triggerfishes, butterflyfishes, and grunts. In contrast, small cryptic reef lineages, such as blennies and gobies, are significantly underrepresented despite being among the most species-rich groups of coral reef fishes. Additional clades with clandestine habits, limited ranges, or those that reside in temperate and polar environments are also significantly under-studied. In contrast, phylogenetic sampling patterns are generally biased at higher taxonomic ranks across all latitudes, though clades composed of largely clandestine species, such as eels, gobies, and blennies, are again identified in need of increased sampling. In total, our analyses provide a roadmap to filling major gaps in our understanding of diel activity patterns and evolutionary history across the world's marine fishes.

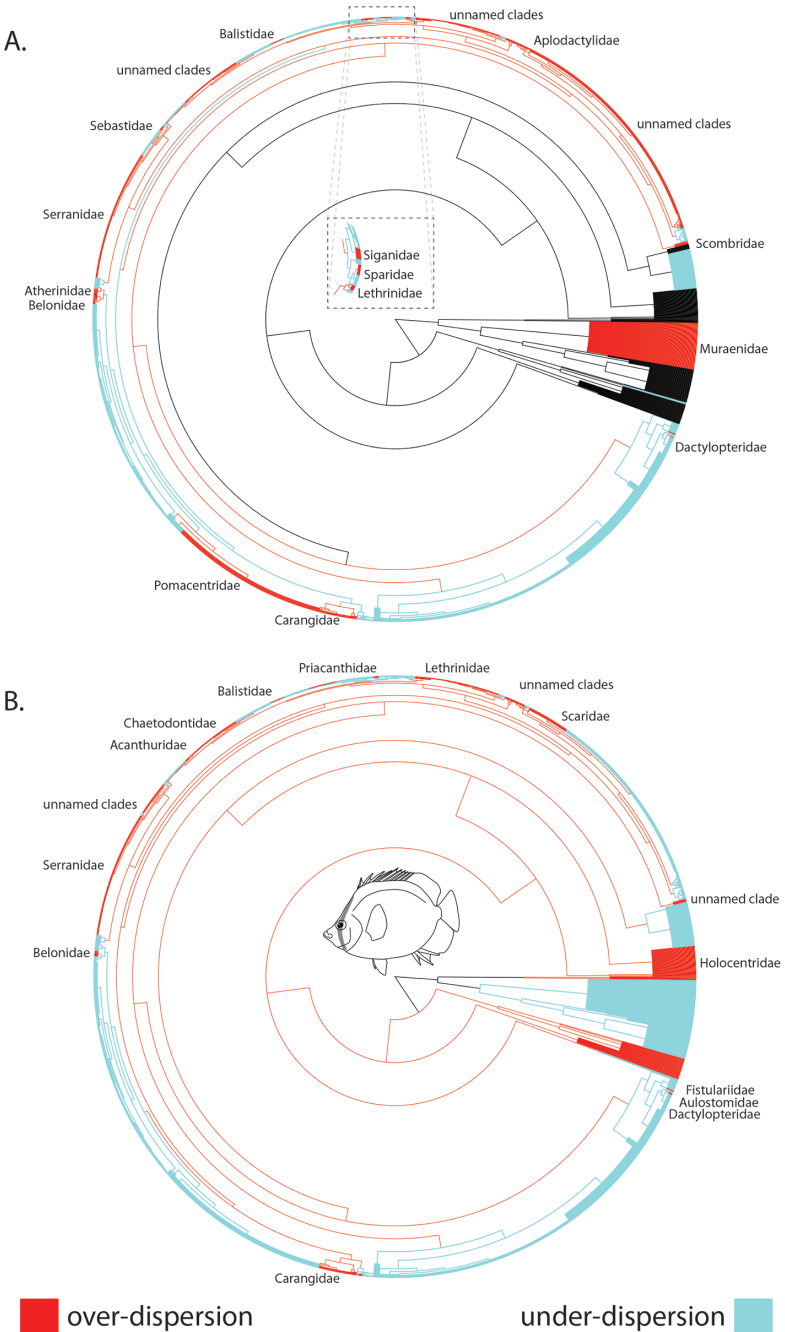


Figure 3. (A) Significant clustering of sampled (dark) and unsampled (light) in taxon sampling in the Rabosky et al. (2013) topology pruned to only reef associated species relative to a diversity tree containing all known reef associated fish species. (B) Significant clustering of sampled (dark) and unsampled (light) in activity data relative to a diversity tree containing all known reef associated fish species. Black lines in the phylogeny represent no deviation under or over dispersion of taxon sampling. Names represent significantly under-sampled named clades. Cartoon illustration of a butterflyfish (Chaetodontidae) represents one of the reef fish clades highlighted as highly sampled in our analyses.

**SAMPLING BIASES: WHAT ARE WE MISSING?**—Technical advances including remote underwater video, baited video, stereo-video, diving equipment, and sonic telemetry have catalyzed a previously unimaginable breadth of marine research (Willis et al. 2000, Harvey et al. 2007, Davis et al. 2015, Sopinka 2015, Lindfield et al. 2016). In the present study, we found diel activity data for 835 species of epipelagic fish from 68 studies (Online Appendix 1), reflecting the continual advancement of tools that enable underwater research. Additionally, many more studies were encountered; however, these included duplicate taxa from different localities and therefore represented redundant data for the purpose of our study. Cumulatively, the surveyed studies contain baseline data on diel patterns for 208 families of marine ray-finned fishes (Table 1). Across this sampling, we find no evidence for a sampling bias correlated with depth, maximum body size, or trophic ecology. However, several strongly supported spatial and taxonomic biases are present in these data.

Spatially, our analyses find strong support for an over-sampling of tropical species when assessing sampling patterns by latitude (Fig. 1), country (Table 2), or ecosystem (Table 3). Although there are a higher number of shallow marine fish species at low vs high latitudes globally (Tittensor et al. 2010), the number of species with diel activity data vastly exceeds expectations that account for uneven patterns of species richness (Fig. 1). This overrepresentation is not entirely unexpected. Large-scale, scuba-based behavioral and ecological studies of reef fish communities have been steadily accumulating for over 60 yrs (Hiatt and Strasburg 1960, Hobson 1965, Stephens et al. 1966, Albrecht 1969, Hobson 1975), providing baseline categorization of diel activity patterns across a wide spectrum of species. Based on our simulations, these sampling efforts have culminated in a much more complete profile of activity patterns for many entire families of reef fishes relative to other marine habitats (Table 1). While over half of the overrepresented families include reef fishes, such as butterflyfishes (Chaetodontidae), wrasses (Labridae), surgeonfishes (Acanthuridae), triggerfishes (Balistidae), and squirrelfishes (Holocentridae), more cryptic reef taxa, such as blennies and gobies, are highly underrepresented (Table 1). As blennies and gobies are among the most species-rich clades of reef fishes (Eschmeyer and Fricke 2015) with speciation rates that rival those of African rift lake cichlids (Near et al. 2013), this bias is not the result of rarity, but rather corresponds with another known bias: under-sampling small cryptic species during the collection of visual survey data (DeMartini and Roberts 1982, Ackerman and Bellwood 2000, Schmitt et al. 2002). This detection bias extends beyond blennies and gobies in our results. Over three quarters of the underrepresented families across all latitudes comprise difficult to detect, burrowing, or shelter-utilizing species, including several flatfish families, toadfishes, eels, and brotulas (Table 1).

In addition to hard to detect taxa, underrepresented species also included equatorial species with limited ranges (Fig. 1B). A potential explanation for this result may be found in the unusual patterns of geographic range size and endemism for tropical marine fishes. While tropical reefs represent biodiversity hotspots (Briggs 2003, Renema et al. 2008, Bellwood et al. 2012, Briggs and Bowen 2012, Dornburg et al. 2015b), many reef-associated species possess large geographic ranges (Hughes et al. 2002, Connolly et al. 2003, Bellwood and Meyer 2009). Therefore, unlike terrestrial biodiversity hotspots, fine scale patterns of endemism are less common while range overlap is common (Bellwood and Meyer 2009, Bellwood et al. 2012), so that diel data collected in one locality can represent species with ranges over thousands of miles.

The limited sampling of species with low maximum absolute latitude is therefore likely a reflection of range-restricted endemic species.

The interpretation of the data and results here is to a large degree reflective of the manner in which data are represented on FishBase; inclusion in the “countries” or “ecosystems” table means that a fish is known to occur in that area, not necessarily that activity data have been collected there. It is therefore possible for an area to appear as “overrepresented” in this sense if activity data for many of its fish species have been collected, regardless of whether that data have been collected within that specific geographic area or ecosystem. Some of the areas that are detected as “overrepresented” in these analyses may therefore represent localities containing a large number of widespread species (Table 2), rather than areas that have been sampled extensively. In this case, endemics restricted to small areas, such as archipelagos, will appear underrepresented, while other sympatric species with broad ranges may prevent the region from being detected as underrepresented in this analysis. This caveat notwithstanding, countries or ecosystems with disproportionately low representation should be considered good targets for increased sampling effort.

Outside of the tropics, we found strong evidence that several temperate and polar regions are underrepresented (Fig. 1C). While some of these regions lack data due to travel restrictions for some scientists (e.g., Iran, North Korea, Iraq; Table 2), other regions, including parts of the United States, Japan, and Russia, do not have such restrictions and are in need of more study (Table 2). The relative accessibility and close proximity to major research centers in these areas presents a unique opportunity to not only fill a critical gap in our understanding of diel activity patterns across fishes, but also enables the development of long-term studies and experiments to assess the impact of light and food regime changes on activity patterns and health. There is a growing recognition that changes in circadian rhythms affect organismal health (Stevens and Zhu 2015), and even minor changes in light regimes associated with urbanization can alter circadian rhythms in fishes (Brüning et al. 2015). Therefore, adding assessments of diel activity patterns alongside chemical (Myers et al. 1991, Lafferty et al. 2015) and hematological (Francesco et al. 2012, Fazio et al. 2013, Collins et al. 2016) health assessments could provide new insights into the health of wild populations in these areas.

**BIASES IN THE EVOLUTIONARY ANALYSES OF ACTIVITY DATA.**—It is hardly surprising that current attempts to infer the ray-finned fish tree of life contain taxon-sampling biases. Given the vast diversity of living fish species, detailed phylogenetic studies commonly focus on subclades that often correspond to taxonomic ranks, such as genera, families, or orders (Alfaro et al. 2007, Dornburg et al. 2008, 2012, Friedman et al. 2013, Miya et al. 2013, Santini et al. 2013b), while studies of deep evolutionary relationships contain fewer representative taxa (Miya et al. 2005, Near et al. 2013, Eytan et al. 2015). This approach to sampling the evolutionary history of marine ray-finned fishes is clearly reflected by our quantification of taxon-sampling biases (Fig. 2). Across all marine fishes, more than 30 clades were significantly overrepresented relative to other taxa (Fig. 2). Of these, the majority of significantly clustered nodes were found at the family level and dominated by conspicuous and well-studied reef-fish clades, such as butterflyfishes, damselfishes, triggerfishes, and parrotfishes, as well as scientifically prominent temperate and polar fishes, including Antarctic icefishes, porgies, and sturgeons (Fig. 2).

While the phylogenetic sampling biases align with the expectations of the field, sampling biases in diel activity data do not mirror the accumulation of phylogenetic data for marine fishes (Fig. 2B). Instead, the majority of overrepresented clades for which we were able to find diel activity data correspond to the biases we found without considering evolutionary history (Tables 1, 2, 3, Fig. 1). Reef fish clades such as squirrelfishes, bigeyes, hawkfishes, grunts, and emperors were highlighted as significantly over-dispersed, raising the strong caution that, at this point, meta-analyses of the evolution of circadian rhythms in marine fishes will be highly biased in their representation of the total diversity of marine fish ecology and life history.

In contrast to sampling bias patterns across all marine fishes, analyses of sampling biases for just reef-associated species revealed a tighter coupling of biases in diel activity and phylogenetic sampling (Fig. 3). Both in terms of activity data and phylogenetic sampling, numerous large, unnamed clades, representing the most recent common ancestors of large diversities of families, are over-dispersed (Fig. 3). These clades are largely consistent for activity data, with a few notable under-sampled clades, such as wrasses, damselfishes, and moray eels (Fig. 3B). In both cases, significantly under-sampled clades include cryptic, difficult-to-observe, or capture taxa, such as flatfishes, gobies, and blennies. This bias again warrants caution for not only meta-analyses of reef fish activity patterns, but also other comparative studies of reef fishes that could be misled. While gathering data for these under-sampled clades is certainly important, it should be noted that for clades highlighted as overrepresented in our study, we are not advocating that no further sampling is necessary. In fact, we would suggest the opposite.

Additional sampling of species already represented in studies of diel activity patterns in marine fishes allows for the testing of more complex hypotheses. For example, assessing diel activity patterns across different habitat classes within a species may reveal previously unrecognized patterns of behavioral plasticity (Fraser et al. 1993, Fox and Bellwood 2011). Likewise, testing for correlations between changes in diel activity patterns and artificial light conditions is a topic of considerable conservation concern that has tremendous conservation and management implications (Longcore and Rich 2004, Davies et al. 2014). Finally, greater taxon sampling increases statistical power in comparative studies (Smith et al. 2011, Beaulieu et al. 2012), highlighting already over-sampled large clades as a potential asset to researchers interested in macroevolutionary phenomena.

While we find significant biases in the sampling of both phylogenetic and diel activity data, our results are encouraging. Despite the logistical changes and challenges in sampling fishes that range from gear specific biases (Meekan et al. 2000, Latour et al. 2003) to behavioral differences between species or even individuals (Biro and Dingemanse 2009), our survey of the literature yielded compiled activity data for more than 800 species across more than 200 families of fishes. As ichthyology moves farther into the 21<sup>st</sup> century, these numbers will only grow. Likewise, the phylogenetic bias of underrepresented cryptic taxa is also sure to change with the continual accumulation of new studies sampling taxa highlighted here as underrepresented (Thacker et al. 2015, Yang et al. 2015, Miller et al. 2016). Given that tissue biopsies of the thousands of marine fish species sampled in phylogenetic analyses are often deposited in the world's natural history collections (Wandeler et al. 2007, Buerki and Baker 2016), researchers will soon be able to integrate next generation sequencing

techniques with this steady accumulation of phylogenetic and behavioral data to place shifts in diel activity patterns into a genomic perspective.

**CONCLUSIONS.**—Over the past several decades, technological innovations have enabled the collection of biological data across the world's oceans at a historically unprecedented pace and scale. As we look forward to the promises of the future, our study is a reminder that major advances in our understanding of marine biodiversity are contingent on the cumulative research efforts of ichthyologists and citizen scientists globally. Sampling bias is not a phenomenon restricted to studies of circadian rhythms. As studies addressing sampling biases across different data types continue to emerge, it is increasingly apparent that bias is a common problem in biological data. Factors such as variation in species ecology (Watson et al. 1995), behavior (Dingemanse et al. 2010), individual specialization/variation (Bolnick et al. 2003, Biro and Dingemanse 2009), and life history (Winemiller 1989, Gaillard et al. 2000) can systematically alter the intended outcomes of research studies in areas as diverse as niche modeling (Phillips et al. 2009, Warren et al. 2014), feeding ecology (Biro 2013), or conservation and management (Costa et al. 2010, Warren et al. 2014). Given the heterogeneous spatial distribution of species and resources/access to wild populations globally, sampling bias patterns should also be predicted to vary considerably at different geographic or taxonomic scales, as well as between different types of biological data. For example, simply shifting our focus from all marine fishes to only reef-associated fishes demonstrated the scale-specific aspect of phylogenetic biases (Figs. 2, 3). Although some areas and taxa are more represented than others, uneven data coverage should not be a cause for despair.

The 21<sup>st</sup> century represents the era of “big data,” a point in science where we are able to harness the cumulative efforts of centuries of scholarship and make use of data collection methods without historical parallels (Hampton et al. 2013, Marx 2013, Soranno and Schimel 2014). In the theme of the Fish at Night symposium, shining a light on fish at night is in some ways analogous to exploring biases lurking beneath the surface of an existing data structure. Focusing research efforts on one underrepresented component of marine vertebrates holds great potential for scientific discovery. Likewise, addressing emerging challenges associated with the accumulation of large data sets, such as sampling bias, also holds great promise. Assessing sampling bias patterns can increase the efficiency of experimental design, saving both research time and costs. Scrutinizing data for bias patterns can also spur the development of new methodologies while empowering new discoveries. As global biodiversity patterns continue to change in response to new pressures associated with the Anthropocene, such scrutiny of ecological or behavioral data will be essential if we are to accurately forecast and manage the future of the planet's biodiversity.

#### ACKNOWLEDGMENTS

We would like to thank the organizers of the Fish at Night symposium for motivating this work and conference participants for numerous wonderful discussions and proceedings. We would also like to thank two anonymous referees and WR Brooks for helpful comments on an earlier version of this manuscript. Partial funding was provided by the North Carolina Museum of Natural Sciences.

## LITERATURE CITED

- Ackerman JL, Bellwood DR. 2000. Reef fish assemblages: a re-evaluation using enclosed rotenone stations. *Mar Ecol Prog Ser.* 206:227–237. <http://dx.doi.org/10.3354/meps206227>
- Albrecht H. 1969. Behaviour of four species of Atlantic damselfishes from Columbia, South America (*Abudefduf saxatilis*, *A. taurus*, *Chromis multilineata*, *C. cyanea*; Pisces, Pomacentridae). *Z Tierpsychol.* 26:662–676. <http://dx.doi.org/10.1111/j.1439-0310.1969.tb01969.x>
- Alfaro ME, Santini F, Brock CD. 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfish and their allies (order Tetraodontiformes). *Evolution.* 61:2104–2126. <http://dx.doi.org/10.1111/j.1558-5646.2007.00182.x>
- Amundsen PA, Bergersen R, Huru H, Heggberget T. 1999. Diel feeding rhythms and daily food consumption of juvenile Atlantic salmon in the River Alta, northern Norway. *J Fish Biol.* 54:58–71. <http://dx.doi.org/10.1111/j.1095-8649.1999.tb00612.x>
- Arendt MD, Lucy JA, Evans DA. 2001. Diel and seasonal activity patterns of adult tautog, *Tautoga onitis*, in lower Chesapeake Bay, inferred from ultrasonic telemetry. *Environ Biol Fishes.* 62:379–391. <http://dx.doi.org/10.1023/A:1012266214143>
- Beaulieu JM, Jhweung DC, Boettiger C, O'Meara BC. 2012. Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution.* 66:2369–2383. <http://dx.doi.org/10.1111/j.1558-5646.2012.01619.x>
- Bellwood DR, Meyer CP. 2009. Searching for heat in a marine biodiversity hotspot. *J Biogeogr.* 36:569–576. <http://dx.doi.org/10.1111/j.1365-2699.2008.02029.x>
- Bellwood DR, Renema W, Rosen BR. 2012. Biodiversity hotspots, evolution and coral reef biogeography: a review. *In:* Gower, DJ, Johnson KG, Richardson JE, Rosen BR, Ruber L, Williams ST, editors. *Biotic evolution and environmental change in Southeast Asia*. Cambridge University Press. Asia:216. p. 216–245
- Benoit-Bird KJ, Au WWL, Brainard RE, Lammers MO. 2001. Diel horizontal migration of the Hawaiian mesopelagic boundary community observed acoustically. *Mar Ecol Prog Ser.* 217:1–14. <http://dx.doi.org/10.3354/meps217001>
- Betancur-R R, Broughton RE, Wiley EO, Carpenter K, López JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton JC 2nd. 2013. The tree of life and a new classification of bony fishes. *PLoS Curr.* 2013 April 18. Edition 1. <http://dx.doi.org/10.1371/currents.tol.53ba26640df0c caee75bb165c8c26288>
- Biro PA. 2013. Are most samples of animals systematically biased? Consistent individual trait differences bias samples despite random sampling. *Oecologia.* 171:339–345. <http://dx.doi.org/10.1007/s00442-012-2426-5>
- Biro PA, Dingemanse NJ. 2009. Sampling bias resulting from animal personality. *Trends Ecol Evol.* 24:66–67. <http://dx.doi.org/10.1016/j.tree.2008.11.001>
- Boettiger C, Lang D, Wainwright P. 2012. rfishbase: exploring, manipulating and visualizing FishBase data from R. *J Fish Biol.* 81:2030–2039. <http://dx.doi.org/10.1111/j.1095-8649.2012.03464.x>
- Bollinger M, Kline R. 2015. Validating side scan sonar as a fish survey tool over artificial reefs in the Gulf of Mexico. *J Acoust Soc Am.* 137:2334–2334. <http://dx.doi.org/10.1121/1.4920525>
- Bolnick DI, Svanbäck R, Fordyce JA, Yang LH, Davis JM, Hulsey CD, Forister ML. 2003. The ecology of individuals: incidence and implications of individual specialization. *Am Nat.* 161:1–28. <http://dx.doi.org/10.1086/343878>
- Brandl SJ, Bellwood DR. 2014. Pair-formation in coral reef fishes: an ecological perspective. *Oceanogr Mar Biol Annu Rev.* 52:1–80. <http://dx.doi.org/10.1201/b17143-2>
- Briggs JC. 2003. Marine centres of origin as evolutionary engines. *J Biogeogr.* 30:1–18. <http://dx.doi.org/10.1046/j.1365-2699.2003.00810.x>
- Briggs JC, Bowen BW. 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. *J Biogeogr.* 39:12–30. <http://dx.doi.org/10.1111/j.1365-2699.2011.02613.x>

- Brodeur R, Pearcy W. 1987. Diel feeding chronology, gastric evacuation and estimated daily ration of juvenile coho salmon, *Oncorhynchus kisutch* (Walbaum), in the coastal marine environment. *J Fish Biol.* 31:465–477. <http://dx.doi.org/10.1111/j.1095-8649.1987.tb05252.x>
- Brüning A, Hölker F, Franke S, Preuer I, Kloas W. 2015. Spotlight on fish: light pollution affects circadian rhythms of European perch but does not cause stress. *Sci Total Environ.* 511:516–522. <http://dx.doi.org/10.1016/j.scitotenv.2014.12.094>
- Brulé T, Avila DO, Crespo MS, Deniel C. 1994. Seasonal and diel changes in diet composition of juvenile red grouper (*Epinephelus morio*) from Campeche Bank. *Bull Mar Sci.* 55:255–262.
- Buerki S, Baker WJ. 2016. Collections-based research in the genomic era. *Biol J Linn Soc Lond.* 117:5–10. <http://dx.doi.org/10.1111/bij.12721>
- Carrete Vega G, Wiens JJ. 2012. Why are there so few fish in the sea? *Proc Biol Sci.* 279:2323–2329. <http://dx.doi.org/10.1098/rspb.2012.0075>
- Chen J-N, López JA, Lavoué S, Miya M, Chen W-J. 2014. Phylogeny of the Elopomorpha (Teleostei): evidence from six nuclear and mitochondrial markers. *Mol Phylogenet Evol.* 70:152–161. <http://dx.doi.org/10.1016/j.ympev.2013.09.002>
- Chepesiuk R. 2009. Missing the dark: health effects of light pollution. *Environ Health Perspect.* 117:A20–A27. <http://dx.doi.org/10.1289/ehp.117-a20>
- Collins S, Dornburg A, Flores JM, Dombrowski DS, Lewbart GA. 2016. A comparison of blood gases, biochemistry, and hematology to ecomorphology in a health assessment of pinfish (*Lagodon rhomboides*). *PeerJ.* 4:e2262. <http://dx.doi.org/10.7717/peerj.2262>
- Connolly SR, Bellwood DR, Hughes TP. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology.* 84:2178–2190. <http://dx.doi.org/10.1890/02-0254>
- Costa GC, Nogueira C, Machado RB, Colli GR. 2010. Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodivers Conserv.* 19:883–899. <http://dx.doi.org/10.1007/s10531-009-9746-8>
- Dardeau MR, McDonald CM. 2007. Pressure related incidence rates in scientific diving. *In: Pollock NW, Godfrey JM, editors. Diving for Science 2007. Proceedings of the American Academy of Underwater Sciences 26th Symposium. Dauphin Island, AL: AAUS.*
- Davies TW, Duffy JP, Bennie J, Gaston KJ. 2014. The nature, extent, and ecological implications of marine light pollution. *Front Ecol Environ.* 12:347–355. <http://dx.doi.org/10.1890/130281>
- Davis T, Harasti D, Smith SD. 2015. Compensating for length biases in underwater visual census of fishes using stereo video measurements. *Mar Freshw Res.* 66:286–291. <http://dx.doi.org/10.1071/MF14076>
- DeMartini E, Roberts D. 1982. An empirical test of biases in the rapid visual technique for species-time censuses of reef fish assemblages. *Mar Biol.* 70:129–134. <http://dx.doi.org/10.1007/BF00397676>
- Devine JA, Baker KD, Haedrich RL. 2006. Fisheries: deep-sea fishes qualify as endangered. *Nature.* 439:29–29. <http://dx.doi.org/10.1038/439029a>
- Dingemanse NJ, Kazem AJ, Réale D, Wright J. 2010. Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol Evol.* 25:81–89. <http://dx.doi.org/10.1016/j.tree.2009.07.013>
- Dornburg A, Friedman M, Near TJ. 2015. Phylogenetic analysis of molecular and morphological data highlights uncertainty in the relationships of fossil and living species of Elopomorpha (Actinopterygii: Teleostei). *Mol Phylogenet Evol.* 89:205–218. <http://dx.doi.org/10.1016/j.ympev.2015.04.004>
- Dornburg A, Moore AJ, Beaulieu JM, Eytan RI, Near TJ. 2015. The impact of shifts in marine biodiversity hotspots on patterns of range evolution: evidence from the Holocentridae (squirrelfishes and soldierfishes). *Evolution.* 69:146–161. <http://dx.doi.org/10.1111/evo.12562>
- Dornburg A, Moore JA, Webster R, Warren DL, Brandley MC, Iglesias TL, Wainwright PC, Near TJ. 2012. Molecular phylogenetics of squirrelfishes and soldierfishes (Teleostei: Beryciformes: Holocentridae): reconciling more than 100 years of taxonomic confusion. *Mol Phylogenet Evol.* 65:727–738. <http://dx.doi.org/10.1016/j.ympev.2012.07.020>



- Dornburg A, Santini F, Alfaro ME. 2008. The influence of model averaging on clade posteriors: an example using the triggerfishes (Family Balistidae). *Syst Biol.* 57:905–919. <http://dx.doi.org/10.1080/10635150802562392>
- Eschmeyer WN, Fricke R, editors. 2015. Catalog of fishes electronic version (9 Sept 2015). California Academy of Sciences, San Francisco. Available from: <http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>
- Eytan RI, Evans BR, Dornburg A, Lemmon AR, Lemmon EM, Wainwright PC, Near TJ. 2015. Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. *BMC Evol Biol.* 15:113. <http://dx.doi.org/10.1186/s12862-015-0415-0>
- Fazio F, Marafioti S, Arfuso F, Piccione G, Faggio C. 2013. Comparative study of the biochemical and haematological parameters of four wild Tyrrhenian fish species. *Vet Med.* 58:576–581.
- Forrestel EJ, Donoghue MJ, Smith MD. 2014. Convergent phylogenetic and functional responses to altered fire regimes in mesic savanna grasslands of North America and South Africa. *New Phytol.* 203:1000–1011. <http://dx.doi.org/10.1111/nph.12846>
- Forrestel EJ, Donoghue MJ, Smith MD. 2015. Functional differences between dominant grasses drive divergent responses to large herbivore loss in mesic savanna grasslands of North America and South Africa. *J Ecol.* 103:714–724. <http://dx.doi.org/10.1111/1365-2745.12376>
- Fox RJ, Bellwood DR. 2011. Unconstrained by the clock? Plasticity of diel activity rhythm in a tropical reef fish, *Siganus lineatus*. *Funct Ecol.* 25:1096–1105. <http://dx.doi.org/10.1111/j.1365-2435.2011.01874.x>
- Francesco F, Satheshkumar P, Kumar DS, Caterina F, Giuseppe P. 2012. A comparative study of hematological and blood chemistry of Indian and Italian grey mullet (*Mugil cephalus* Linnaeus 1758). *HOAJ Biology.* 1:5. <http://dx.doi.org/10.7243/2050-0874-1-5>
- Fraser NH, Metcalfe NB, Thorpe JE. 1993. Temperature-dependent switch between diurnal and nocturnal foraging in salmon. *Proc R Soc Lond B Biol Sci.* 252:135–139. <http://dx.doi.org/10.1098/rspb.1993.0057>
- Friedman M, Keck BP, Dornburg A, Eytan RI, Martin CH, Hulsey CD, Wainwright PC, Near TJ. 2013. Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proc Royal Soc B: Biol Sci.* 280:1–8. <http://dx.doi.org/10.1098/rspb.2013.1733>
- Gaillard J-M, Festa-Bianchet M, Yoccoz N, Loison A, Toigo C. 2000. Temporal variation in fitness components and population dynamics of large herbivores. *Annu Rev Ecol Syst.* 31:367–393. <http://dx.doi.org/10.1146/annurev.ecolsys.31.1.367>
- Gallaway T, Olsen RN, Mitchell DM. 2010. The economics of global light pollution. *Ecol Econ.* 69:658–665. <http://dx.doi.org/10.1016/j.ecolecon.2009.10.003>
- Goatley CH, Bellwood DR, Bellwood O. 2010. Fishes on coral reefs: changing roles over the past 240 million years. *Paleobiology.* 36:415–427. <http://dx.doi.org/10.1666/09035.1>
- Hampton SE, Strasser CA, Tewksbury JJ, Gram WK, Budden AE, Batcheller AL, Duke CS, Porter JH. 2013. Big data and the future of ecology. *Front Ecol Environ.* 11:156–162. <http://dx.doi.org/10.1890/120103>
- Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics.* 24:129–131. <http://dx.doi.org/10.1093/bioinformatics/btm538>
- Harvey ES, Cappo M, Butler JJ, Hall N, Kendrick GA. 2007. Bait attraction affects the performance of remote underwater video stations in assessment of demersal fish community structure. *Mar Ecol Prog Ser.* 350:245. <http://dx.doi.org/10.3354/meps07192>
- Helfman GS. 1986. Fish behaviour by day, night and twilight. *The behaviour of teleost fishes.* Springer. p. 366–387.
- Heupel M, Semmens J, Hobday A. 2006. Automated acoustic tracking of aquatic animals: scales, design and deployment of listening station arrays. *Mar Freshw Res.* 57:1–13. <http://dx.doi.org/10.1071/MF05091>
- Hiatt RW, Strasburg DW. 1960. Ecological relationships of the fish fauna on coral reefs of the Marshall Islands. *Ecol Monogr.* 30(1):65–127. <http://dx.doi.org/10.2307/1942181>

- Hobson ES. 1965. Diurnal-nocturnal activity of some inshore fishes in the Gulf of California. *Copeia*. 291–302. <http://dx.doi.org/10.2307/1440790>
- Hobson ES. 1975. Feeding patterns among tropical reef fishes. *Am Sci*. 63:382–392.
- Holcroft NI. 2005. A molecular analysis of the interrelationships of tetraodontiform fishes (Acanthomorpha: Tetraodontiformes). *Mol Phylogenet Evol*. 34:525–544. <http://dx.doi.org/10.1016/j.ympev.2004.11.003>
- Holcroft NI, Wiley EO. 2008. Acanthuroid relationships revisited: a new nuclear gene-based analysis that incorporates tetraodontiform representatives. *Ichthyol Res*. 55:274–283. <http://dx.doi.org/10.1007/s10228-007-0026-x>
- Horn MH. 1980. Diel and seasonal variation in abundance and diversity of shallow-water fish populations in Morro Bay, California. *Fish Bull*. 78:759–769.
- Hughes TP, Bellwood DR, Connolly SR. 2002. Biodiversity hotspots, centres of endemism, and the conservation of coral reefs. *Ecol Lett*. 5:775–784. <http://dx.doi.org/10.1046/j.1461-0248.2002.00383.x>
- Hut RA, Kronfeld-Schor N, Van der Vinne V, De la Iglesia HO. 2012. In search of a temporal niche: environmental factors. *Prog Brain Res*. 199:281–304. <http://dx.doi.org/10.1016/B978-0-444-59427-3.00017-4>
- Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*. 26:1463–1464. <http://dx.doi.org/10.1093/bioinformatics/btq166>
- Lafferty KD, Harvell CD, Conrad JM, Friedman CS, Kent ML, Kuris AM, Powell EN, Rondeau D, Saksida SM. 2015. Infectious diseases affect marine fisheries and aquaculture economics. *Annu Rev Mar Sci*. 7:471–496. <http://dx.doi.org/10.1146/annurev-marine-010814-015646>
- Latour RJ, Brush MJ, Bonzek CF. 2003. Toward ecosystem-based fisheries management: strategies for multispecies modeling and associated data requirements. *Fisheries*. 28:10–22. [http://dx.doi.org/10.1577/1548-8446\(2003\)28\[10:TEFM\]2.0.CO;2](http://dx.doi.org/10.1577/1548-8446(2003)28[10:TEFM]2.0.CO;2)
- Lindfield SJ, Harvey ES, Halford AR, McIlwain JL. 2016. Mesophotic depths as refuge areas for fishery-targeted species on coral reefs. *Coral Reefs*. 35:125. <http://dx.doi.org/10.1007/s00338-015-1386-8>
- Lobel PS. 1978. Diel, lunar, and seasonal periodicity in the reproductive behavior of the pomacanthid fish, *Centropyge potteri*, and some other reef fishes in Hawaii. *Pac Sci*. 32:193–207.
- Longcore T, Rich C. 2004. Ecological light pollution. *Front Ecol Environ*. 2:191–198. [http://dx.doi.org/10.1890/1540-9295\(2004\)002\[0191:ELP\]2.0.CO;2](http://dx.doi.org/10.1890/1540-9295(2004)002[0191:ELP]2.0.CO;2)
- Makris NC, Ratilal P, Symonds DT, Jagannathan S, Lee S, Nero RW. 2006. Fish population and behavior revealed by instantaneous continental shelf-scale imaging. *Science*. 311:660–663. <http://dx.doi.org/10.1126/science.1121756>
- Marx V. 2013. Biology: the big challenges of big data. *Nature*. 498:255–260. <http://dx.doi.org/10.1038/498255a>
- McIntyre F, Neat F, Collie N, Stewart M, Fernandes P. 2015. Visual surveys can reveal rather different ‘pictures’ of fish densities: comparison of trawl and video camera surveys in the Rockall Bank, NE Atlantic Ocean. *Deep Sea Res Part I Oceanogr Res Pap*. 95:67–74. <http://dx.doi.org/10.1016/j.dsr.2014.09.005>
- Meekan M, Doherty P, White L. 2000. Recapture experiments show the low sampling efficiency of light traps. *Bull Mar Sci*. 67:875–885.
- Meyer CG, Holland KN, Wetherbee BM, Lowe CG. 2000. Movement patterns, habitat utilization, home range size and site fidelity of whitesaddle goatfish, *Parupeneus porphyreus*, in a marine reserve. *Environ Biol Fishes*. 59:235–242. <http://dx.doi.org/10.1023/A:1007664813814>
- Miller EC, Lin H-C, Hastings PA. 2016. Improved resolution and a novel phylogeny for the Neotropical triplefin blennies (Teleostei: Tripterygiidae). *Mol Phylogenet Evol*. 96:70–78. <http://dx.doi.org/10.1016/j.ympev.2015.12.003>
- Miya M, Friedman M, Satoh TP, Takeshima H, Sado T, Iwasaki W, Yamanoue Y, Nakatani M, Mabuchi K, Inoue JG. 2013. Evolutionary origin of the Scombridae (tunas and mackerels):

- members of a Paleogene adaptive radiation with 14 other pelagic fish families. *PLoS One*. 8:e73535. <http://dx.doi.org/10.1371/journal.pone.0073535>
- Miya M, Satoh TR, Nishida M. 2005. The phylogenetic position of toadfishes (order Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. *Biol J Linn Soc Lond*. 85:289–306. <http://dx.doi.org/10.1111/j.1095-8312.2005.00483.x>
- Myers MS, Landahl JT, Krahn MM, McCain BB. 1991. Relationships between hepatic neoplasms and related lesions and exposure to toxic chemicals in marine fish from the US West Coast. *Environ Health Perspect*. 90:7. <http://dx.doi.org/10.2307/3430840>
- Naylor E. 2005. Chronobiology: implications for marine resource exploitation and management. *Sci Mar*. 69:157–167. <http://dx.doi.org/10.3989/scimar.2005.69s1157>
- Near TJ, Dornburg A, Eytan RI, Keck BP, Smith WL, Kuhn KL, Moore JA, Price SA, Burbrink FT, Friedman M. 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proc Natl Acad Sci USA*. 110:12,738–12,743. <http://dx.doi.org/10.1073/pnas.1304661110>
- Near TJ, Dornburg A, Harrington RC, Oliveira C, Pietsch TW, Thacker CE, Satoh TP, Katayama E, Wainwright PC, Eastman JT. 2015. Identification of the notothenioid sister lineage illuminates the biogeographic history of an Antarctic adaptive radiation. *BMC Evol Biol*. 15:109. <http://dx.doi.org/10.1186/s12862-015-0362-9>
- Near TJ, Dornburg A, Kuhn KL, Eastman JT, Pennington JN, Patarnello T, Zane L, Fernandez DA, Jones CD. 2012a. Ancient climate change, antifreeze, and the evolutionary diversification of Antarctic fishes. *Proc Natl Acad Sci USA*. 109:3434–3439. <http://dx.doi.org/10.1073/pnas.1115169109>
- Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. 2012b. Resolution of ray-finned fish phylogeny and timing of diversification. *Proc Natl Acad Sci USA*. 109:13,698–13,703. <http://dx.doi.org/10.1073/pnas.1206625109>
- Nelson JS, Grande TC, Wilson MV. 2016. *Fishes of the World*. John Wiley & Sons.
- Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*. 20:289–290. <http://dx.doi.org/10.1093/bioinformatics/btg412>
- Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl*. 19:181–197. <http://dx.doi.org/10.1890/07-2153.1>
- Pulcini D, Costa C, Aguzzi J, Cataudella S. 2008. Light and shape: a contribution to demonstrate morphological differences in diurnal and nocturnal teleosts. *J Morphol*. 269:375–385. <http://dx.doi.org/10.1002/jmor.10598>
- Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun*. 4:1958. <http://dx.doi.org/10.1038/ncomms2958>
- Renema W, Bellwood DR, Braga JC, Bromfield K, Hall R, Johnson KG, Lunt P, Meyer CP, McMonagle LB, Morley RJ, et al. 2008. Hopping hotspots: global shifts in marine biodiversity. *Science*. 321:654–657. <http://dx.doi.org/10.1126/science.1155674>
- Roe H, Badcock J. 1984. The diel migrations and distributions within a mesopelagic community in the north east Atlantic. 5. Vertical migrations and feeding of fish. *Prog Oceanogr*. 13:389–424. [http://dx.doi.org/10.1016/0079-6611\(84\)90014-4](http://dx.doi.org/10.1016/0079-6611(84)90014-4)
- Rooker JR. 1995. Feeding ecology of the schoolmaster snapper, *Lutjanus apodus* (Walbaum), from southwestern Puerto Rico. *Bull Mar Sci*. 56:881–894.
- Sanciango MD, Carpenter KE, Betancur-R R. 2016. Phylogenetic placement of enigmatic percormorph families (Teleostei: Percormorphaceae). *Mol Phylogenet Evol*. 94:565–576. <http://dx.doi.org/10.1016/j.ympev.2015.10.006>
- Santini F, Carnevale G. 2015. First multilocus and densely sampled timetree of trevallies, pompanos and allies (Carangoidei, Percormorpha) suggests a Cretaceous origin and Eocene radiation of a major clade of piscivores. *Mol Phylogenet Evol*. 83:33–39. <http://dx.doi.org/10.1016/j.ympev.2014.10.018>

- Santini F, Sorenson L, Alfaro ME. 2013a. A new multi-locus timescale reveals the evolutionary basis of diversity patterns in triggerfishes and filefishes (Balistidae, Monacanthidae; Tetraodontiformes). *Mol Phylogenet Evol.* 69:165–176. <http://dx.doi.org/10.1016/j.ympev.2013.05.015>
- Santini F, Sorenson L, Marcroft T, Dornburg A, Alfaro ME. 2013b. A multilocus molecular phylogeny of boxfishes (Aracanidae, Ostraciidae; Tetraodontiformes). *Mol Phylogenet Evol.* 66:153–160. <http://dx.doi.org/10.1016/j.ympev.2012.09.022>
- Schmitt E, Sluka R, Sullivan-Sealey K. 2002. Evaluating the use of roving diver and transect surveys to assess the coral reef fish assemblage off southeastern Hispaniola. *Coral Reefs.* 21:216–223.
- Schmitz L, Wainwright PC. 2011. Nocturnality constrains morphological and functional diversity in the eyes of reef fishes. *BMC Evol Biol.* 11:338. <http://dx.doi.org/10.1186/1471-2148-11-338>
- Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *Am J Bot.* 98:404–414. <http://dx.doi.org/10.3732/ajb.1000481>
- Sopinka N. 2015. Listening to fishes to see where they go. *Fisheries.* 40:524–524. <http://dx.doi.org/10.1080/03632415.2015.1082473>
- Soranno PA, Schimel DS. 2014. Macrosystems ecology: big data, big ecology. *Front Ecol Environ.* 12:3–3. <http://dx.doi.org/10.1890/1540-9295-12.1.3>
- Stephens JS Jr, Hobson ES, Johnson RK. 1966. Notes on the distribution, behavior, and morphological variation in some chaenopsis fishes from tropical eastern Pacific, with descriptions of two new species, *Acanthemblemaria castori* and *Coralliozetes springeri*. *Copeia.* 1966:424–438. <http://dx.doi.org/10.2307/1441061>
- Stevens RG, Zhu Y. 2015. Electric light, particularly at night, disrupts human circadian rhythmicity: is that a problem? *Philos Trans R Soc Lond B Biol Sci.* 370:20140120. <http://dx.doi.org/10.1098/rstb.2014.0120>
- Thacker CE, Satoh TP, Katayama E, Harrington RC, Eytan RI, Near TJ. 2015. Molecular phylogeny of Percomorpha resolves *Trichonotus* as the sister lineage to Gobioidi (Teleostei: Gobiiformes) and confirms the polyphyly of Trachinoidei. *Mol Phylogenet Evol.* 93:172–179. <http://dx.doi.org/10.1016/j.ympev.2015.08.001>
- Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Berghe EV, Worm B. 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature.* 466:1098–1101. <http://dx.doi.org/10.1038/nature09329>
- Walker RV, Myers KW, Davis ND, Aydin KY, Friedland KD, Carlson HR, Boehlert GW, Urawa S, Ueno Y, Anma G. 2000. Diurnal variation in thermal environment experienced by salmonids in the North Pacific as indicated by data storage tags. *Fish Oceanogr.* 9:171–186. <http://dx.doi.org/10.1046/j.1365-2419.2000.00131.x>
- Wandeler P, Hoeck PE, Keller LF. 2007. Back to the future: museum specimens in population genetics. *Trends Ecol Evol.* 22:634–642. <http://dx.doi.org/10.1016/j.tree.2007.08.017>
- Warren DL, Wright AN, Seifert SN, Shaffer HB. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Divers Distrib.* 20:334–343. <http://dx.doi.org/10.1111/ddi.12160>
- Watson R, Carlos G, Samoilys M. 1995. Bias introduced by the non-random movement of fish in visual transect surveys. *Ecol Modell.* 77:205–214. [http://dx.doi.org/10.1016/0304-3800\(93\)E0085-H](http://dx.doi.org/10.1016/0304-3800(93)E0085-H)
- Webb CO, Ackerly DD, Kembel SW. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics.* 24:2098–2100. <http://dx.doi.org/10.1093/bioinformatics/btn358>
- Willis TJ, Miller RB, Babcock RB. 2000. Detection of spatial variability in relative density of fishes: comparison of visual census, angling, and baited underwater video. *Mar Ecol Prog Ser.* 198:249–260. <http://dx.doi.org/10.3354/meps198249>

- Winemiller KO. 1989. Patterns of variation in life history among South American fishes in seasonal environments. *Oecologia*. 81:225–241. <http://dx.doi.org/10.1007/BF00379810>
- Yang Q-H, Lin Q, He L-B, Huang R-F, Lin K-B, Ge H, Wu J-S, Zhou C. 2015. The complete mitochondrial genome sequence of *Acentrogobius* sp.(Gobiiformes: Gobiidae) and phylogenetic studies of Gobiidae. *Mitochondrial DNA*. 27(4):2927–2928. <http://dx.doi.org/10.3109/19401736.2015>

