Title: Predator body size, species identity, and hunting traits determine predator-prey interactions in diet DNA metabarcoding data

Invited author list: Ana Miller-ter Kuile, Austen Apigo, An Bui, Bart DiFiore, Elizabeth Forbes, Daniel Preston, Taylor Bogar, Jasmine Childress, Rodolfo Dirzo, Magalay Espinoza, Cora Johnston, Maggie Klope, Kevin Lafferty, Michelle Lee, John McLaughlin, Marisa Morse, Carina Motta, Devyn Orr, Katherine Plummer, Chelsea Steele, David Weber, Hillary Young

Lab: anyone with 199?

ProcB: 6 MS pages, ~4000 words + 3 figures from abstract to end of literature cited. (Currently a ways over this limit).

**Abstract (200 words)**

**Keywords (3-6 words)**

**Introduction**

Baby:

Body size is a key trait that determines feeding interactions between predators and prey across ecosystems (Woodward, etc). Predator body size determines the range of prey a predator can consume, often with larger predators feeding on larger prey and a wider range of prey sizes due to optimal foraging (Stoufer 2015, Nakazawa). Because body size is integral to feeding interactions, it is one of the primary approaches for predicting the structure of feeding interactions for biological communities, or food webs (Woodward et al., Gravel et al. 2013, 2015). As such, body size is a useful species trait for determining feeding interactions when interactions have not or cannot be observed (e.g. Hines et al. 2019), or for generalizing patterns across environments and consumer groups (Brose et al. 2006, Brose et al. 2019). Furthermore, while body size alone predicts general patterns across food webs in multiple contexts, combining body size with other predator species characteristics, including species identity, and more broadly, species traits such as locomotion or metabolism, creates food web models that look even more similar to empirically-observed patterns (e.g. Grey et al., Pomeranz et al., Brose et al., Rudolf et al. 2014). Using general traits to describe food web patterns across ecosystems is not only important for the development of general predictable rules describing patterns in biological communities but could also be integral to predicting and mitigating species extinctions given the rate of anthropogenic species loss (Valiente-Banuet et al.).

Werewolf: While a few general rules accurately predict patterns in food webs that have been built fully or primarily from empirical data (e.g. Gravel et al PAPER), for some predators, in particular small-bodied invertebrate predators for which empirical diet methods (e.g. gut dissections) are impossible or often unfeasible (CITE From methods paper set of papers, McLaughlin et al. 2010), we have a dearth of observed interaction data with which to validate extrapolated approaches to predicting interactions based on species traits. For these consumers, species interactions are often inferred from literature reports of observed interactions from phylogenetically-related species (e.g. ) or based on body size feeding constraints (e.g. Hines et al. 2019). Because these are the very patterns we are aiming to extrapolate from these datasets, then, without empirical data it is unclear whether any patterns that occur for these consumers related to body size, species traits, or phylogeny are real ecological patterns or artefacts of the non-empirical diet assignment methods used to compile them. Beyond a better representation of food webs across environments and consumer groups, overlooking food web patterns that govern small-bodied terrestrial invertebrates is not inconsequential to global ecosystem processes; small-bodied terrestrial invertebrates comprise 50% of the earth’s animal biomass (Bar-on et al. 2018) and the majority of global animal species diversity (Mora et al. 2011, Costello et al 2013, Stork 2018). As such, understanding food web patterns for this consumers group is crucial for understanding the ecology of the world’s most dominant group of consumers shaping biomass and nutrient cycling dynamics (e.g. Bar-On et al. 2018).

Silver Bullet: In this study, we employ novel diet DNA metabarcoding data from 182 individuals of nine terrestrial invertebrate predator species to document predator-prey interactions between these predators and their prey in natural field conditions. Combined with an extensive dataset of body sizes for both predator individuals and the prey groups identified in their diets, we examine the predator-prey body size relationships for these predators across and within predator species. To understand how predator size, species identity, and feeding traits may drive empirical predator-prey interactions, we asked: 1) do larger individuals eat larger prey and is this mediated by predator species identity? 2) beyond species, do predator traits shared by multiple species related to hunting “tools” (web-use) relate to prey size selection? To understand how patterns we observed related to predator-prey body size compare to those used to extrapolate across systems, we then compared the predator-prey body size relationship we observed in this dataset to the predator-prey body size scaling reported from four terrestrial invertebrate food webs (from Brose et al. 2019, CITE OTHERS) to ask: 3) How does the predator-prey body size scaling relationship in empirical DNA-based diet data compare to the predator-prey body size scaling relationship across multiple studies in which a combination of primarily literature-based methods were used to assign interactions? Using empirical interaction data to understand how body size and species identity shape terrestrial invertebrate food webs will be key to building predictive food web models built on these parameters (Gravel et al. 2015, Pompanez et al. 2018).

**Methods:**

*Field site and collections*

We conducted this work on Palmyra Atoll National Wildlife Refuge, Northern Line Islands (5º53’ N, 162º05’W). Palmyra Atoll has a well-characterized species list, and like many atolls, is relatively species poor, allowing for characterization of potential diet items (Handler et al., 2007). Predator individuals were collected across habitat types representing likely loosely-connected compartments in the atoll metaweb (CITE), including different forest types and microhabitats, including understory vegetation, canopy vegetation, and soil habitats. For each of these habitats, we used a combination of methods, including individual collection during visual surveys for understory and soil collections and canopy fogging with insecticide (cite) onto collection sheets for canopy individuals. All individuals were collected individually with sterilized implements (ethanol-burned forceps) in sterilized collection containers containing 95% EtOH to avoid contamination (Greenstone et al., 2011). All individuals were stored in 95% EtOH a -20ºC before DNA extraction.

We identified all predators to morphospecies in a laboratory environment and later validated unique species by DNA metabarcoding sequence data. The predator species used in this study represent the most common predator species found in each food web compartment (habitat location) and span a body size range of 0.2 – 928 mg (wet mass, SI Table 2). These predators included five arachnid species (*Oonopidae* sp., *Neoscona theisi*, *Heteropoda venatoria*, *Smeringopus pallidus*, and *Scytodes longipes*), one dragonfly (*Pantala flavescens*), one predatory katydid (*Phisis holdhausi*), one earwig (*Euborellia annulipes*), and one soil-dwelling centipede species (*Geophilomorpha* sp.). These predators employ various hunting tools, including webs, venom, and grasping forearms and employ several different hunting strategies, including web-building, sit-and-wait, and active hunting. (DO I need a table?)

*DNA extraction, PCR amplification, library preparation, sequencing, and denoising*

Our full DNA extraction, PCR amplification, library preparation, sequencing, and denoising methods can be found in the Supplementary Information, though we provide an abridged version here.

We individually measured the length of each predator (mm) and separated the thorax, opisthosoma, or trunk (depending on predator species, Krehenwinkel, Macias Herndandez) for DNA extraction following a modified CTAB extraction protocol (Fulton et al., 1995). We followed methods in Krehenwinkel et al. (2017) for extracting diet DNA from terrestrial invertebrate predators, using PCR primers (mlCOIintf/Fol-degen-rev; Krehenwinkel et al. 2017, Leray et al. 2013, Yu et al. 2012) targeting the CO1 gene, which is well-represented in online databases (Porter and Hajibabaei 2018). Following amplification of the CO1 gene, we attached Illumina index primers (Nextera XT Index Kit v2) following the standard protocol for these primers (Illumina). To achieve ample sequencing depth (SI Figure 2), we ran samples across four separate sequencing runs (SI Table 1). All individuals within a predator species were sequenced on the same run and each run contained one to five predator species. We ran 19 samples of one predator species (*H. venatoria*) across all runs to quantify run-to-run variation in sequencing (SI Figure 3). We submitted multiplexed samples for sequencing at the University of California, Santa Barbara Biological Nanostructures Laboratory Genetics Core. Samples were run on an Illumina MiSeq platform (v2 chemistry, 500 cycles, paired-end reads) with a 15% spike-in of PhiX. Following sequencing, samples were demultiplexed using Illumina’s bcl2fastq conversion software (v2.20) at the Core facility. We merged, filtered, and denoised our sequences around amplicon sequence variants (ASVs) using the DADA2 algorithm in R (dada2 package version 1.1.14.0; Callahan et al., 2016). We removed samples from analysis that had not been sequenced to sufficient depth using iNEXT (CITE) and a lower quantile cutoff (SI Figures 6 & 7). We rarefied remaining samples (McKnight et al., 2019) based on the sample with the lowest sequencing depth which had been sequenced with 95%+ sampling completeness based on iNEXT (version 2.0.20) interpolation and extrapolation methods (Hsieh & Chao, 2017). We rarefied using the rrarefy() function in the vegan (version 2.5.6) package in R to 15,954 reads per sample.

*ASV taxonomic assignment with BLAST and BOLD*

We performed taxonomic assignment using BLAST and the blastn command (version 2.7.1) and using sequences in the GenBank and BOLD sequence databases (GenBank accessed: November 20, 2019 and BOLD accessed: May 21, 2020). BLAST analyses were run using the computing cluster at UC Santa Barbara. We exported taxonomic assignments using MEGAN Community Edition (version 6.18.0, Huson et al., 2016), using default settings and selecting the subtree with all possible diet items for this species (Kingdom: Animalia, Clade: Bilateria). We chose to combine prey taxonomies at the family level, similar to diet resolution in both metabarcoding and histological methods in this field (e.g. Brose et al. 2019, Kartzinel – find other metabarcoding ones) by summing the cumulative read abundances across the ASVs that corresponded to each diet family in each sample. All DNA matching any predator family present on an individual sequencing run was also removed as a conservative method to account for potential sequence jumping within sequencing runs which could alter prey identity or diversity in favor of predator species on a shared run (CITE SEQUENCE JUMPING).

*Predator and prey size determination*

We converted predator lengths to wet mass using mass-length scaling relationships for each predator species using existing datasets from the literature and the field site (Soehlstrom et al, Su et al. Yaninek et al. 1993, Miller-ter Kuile *unpublished data*, McLaughlin et al. *unpublished data*). Prey masses were taken as the average mass for individuals across species within each family (Supplementary Information, Figures).

*Statistical analyses*

To determine whether individual predator size, species, or both predicted prey size, we used model selection based on AICc to compare a set of nested linear models with the response variable of prey size. The full model included the interaction between log10 predator mass (in mg) and species identity, with a random effect of predator individual. We log10 transformed the response variable (log10 prey mass in wet mg). To explore how predator traits may influence predator-prey size ratios, we divided predators based on whether or not the predator species uses webs to either capture or subdue prey (n = 5 predator species, 101 individuals that build webs; n = 4 predator species, 81 individuals that do not build webs). We determined the ratio of predator to prey size for each of these interactions (raw predator mass/prey mass) and then built a linear mixed model of this ratio (log transformed for data normality) as the response variable, web-building (binary) as a predictor, and predator individual as a random effect. Finally, in order to compare our results to those from the literature, we extracted data from four invertebrate terrestrial food webs (compiled in Brose et al. 2019) on predator-prey body size relationships (not within species, so we could not examine the same species effects) to determine how the general scaling of predator-prey body size in our dataset compared to the relationship from previous food webs in which predator-prey interactions are built with a combination of literature documentation, phylogenetic relatedness, body size constraints, and some observation (Cattin Blandenier (2004), Digel et al. (2014), McLaughlin et al. (2010), Piechnik et al. (2008), Simberloff and Wilson (1969)). For this model, we created a full model of log10 prey mass predicted by the interaction between log10 predator mass and web type (DNA versus literature) with a random effect of food web identity, since data were drawn from a collection of food webs. Because our question was whether the scaling relationship was the same or different between our web and that from the literature webs (ANCOVA), we compared a model with and without the interaction between fixed effects (log10 predator mass and web type) to one without the interaction. For each question, we performed model selection, using the dredge() function in the MuMIn package in R (Versions, citations) to compare nested models and chose the model with the lowest AICc value. We verified model assumptions for best-fitting models using the DHARMa package in R (versions). All figures illustrating results were created in ggplot2 with color palettes from the calecopal package (CITE).

**Results**

*DNA extraction, PCR amplification, library preparation, sequencing, denoising, and ASV taxonomy*

Complete results and QC for each step of the DNA sequencing protocol can be found in the Supplementary Information.

Our final analyses were performed on a total of 182 predator individuals of nine species. Each predator had consumed 1 – 7 (average 1.86 ± 1.21 SD) prey families. Thirty percent (n = 524 of 1,738 total ASVs) of the total ASVs found in samples received taxonomic assignments from GenBank and BOLD corresponding to prey items at the family level or lower and so were used in analyses. These corresponded to 57 prey families (SUPP table).

*Statistical results*

The best model predicting prey size included the terms of predator mass and predator species identity, but not the interaction between these two terms. (log10(prey mass) = 0.41\*log10(predator mass), R2m = 0.30, R2c = 0.35, with significant variation in by-species intercepts, Figure, SI Table 3). Though we did not collect our dataset in order to elucidate what traits might drive patterns in terrestrial invertebrate food webs, we did observe that those species that use webs either to capture or subdue prey (five arachnid species) had a greater proportion of prey items closer to them in size than did predators without web-use traits. (give some means ± SE; p-value = 0.05, β = 1.25, Figure). The predator-prey mass scaling relationship for our dataset was significantly shallower than the relationship compiled for four terrestrial invertebrate food webs from the literature (log10(prey mass) = 0.26\*log10(predator mass) for our food web versus log10(prey mass) = 0.64\*log10(predator mass) for the four published food webs (ANCOVA p-value < 0.001, Figure).

**Discussion**

For terrestrial invertebrate predators like the ones in our study, comprehensive diet analyses have been nearly impossible or time-prohibitive (e.g. Polis 1991, McLaughlin et al. 2010) without genetic methods, so these data provide important empirical examination of food web patterns for these consumers. We found that both predator mass and predator species identity are important drivers of prey selection. Specifically, 1) we found that larger predator individuals do eat larger prey, however, individuals of some predator species eat proportionally smaller or larger prey than would be expected by one general cross-species relationship. Broadening this relationship to predict deviations from predator-prey feeding interactions based on species feeding traits, we demonstrate demonstrate that 2) the predator trait of web use, whether in capturing or subduing prey, enables some predator species to eat proportionally larger prey items. Finally, 3) when we compared the general predator-prey body size scaling relationship in our dataset to the patterns gathered from four literature food webs for similar consumers, we observed that our DNA diet metabarcoding data scaling relationship is significantly shallower than the relationship found in four food webs in the literature (CITE), suggesting that this general cross-food web relationship would overestimate the size of prey for larger individuals in our dataset and underestimate the size of prey for smaller individuals in our dataset. These results highlight that many food web patterns in small, terrestrial, invertebrate predator-prey interactions may be explained by a combination of body size and predator traits that are conserved among phylogenetically-similar predator species, lending empirical evidence to the approaches used in the field. Building comprehensive datasets for small-bodied predators will help to build more robust empirical datasets with which to continue to refine these general rules.

[predator-prey size relationships]

There is continued recognition in the field that food web patterns are likely the result of a combination of body size and species- or taxon-specific traits (Raffaieli 2007, Rudolf et al. 2014, Rall et al. 2011). Our results strengthen the need for combining body size with species identity or feeding traits for explaining and predicting food web patterns, since predator species significantly influenced predator-prey size scaling relationships. In our results, predator species that may be more gape-limited (e.g. *Pantala flavescens* and *Geophilomorpha* sp.) have smaller prey items on average compared to predators of similar, or even, smaller size of other species that may not be as gape limited due to “tools” such as webs (e.g. Oonopidae sp). Some traits that limit prey size selection may be phylogenetically conserved (e.g. mandible morphology), suggesting that approaching these food webs similarly to e.g. Grey at al. (DATE) with phylogenetic filters could explain variation not explained by the trait of body size alone.

[trait stuff]

Our results also highlight that it may be that approaching food web construction more generally than phylogeny, and rather using broad traits in addition to body size might be key to observed food web patterns (e.g. Laigle et al. 2017, Brose et al. 2019). For example, Brose et al. (2019) find that traits such as locomotion strategy or metabolism can explain variation in body size scaling patterns. In our study, we found that predator individuals of five arachnid species that use webs to either catch or subdue prey had significantly larger prey proportional to predator individual body size. This idea has already earned attention in terrestrial invertebrate communities (Laigle et al. 2017) and could be further explored by the use of diet DNA metabarcoding datasets to validate and add feeding interactions. While web-using seems to alter predator-prey size relationships, one predator, the centipede Geophilomorpha sp., which uses venom to subdue prey, consumed proportionally smaller prey (GET AVG SIZE HERE), suggesting that not all hunting tools always relax body size constraints. Conversely, one predator which likely uses neither of these tools, the dermaptera *E. annulipes,* fed on proportionally very large prey compared to predator body size (GIVE AVG SIZE HERE), either suggesting some other tool use in this predator or a reliance on scavenging (CITE scavenging). These patterns are promising for further studies on the traits that extend within and across groups of terrestrial invertebrate consumers and the use of combined genetic tools (e.g. RNA sequencing) to determine whether prey items detected in predator diets were consumed live or scavenged (e.g. Nielson et al.).

[Literature vs. empirical]

[WHAT WERE LITERATURE DIET METHODS to think about why ours are different]. Is this methodological or is it relative food web to food web? All of these food webs have a lot of spiders.

[1] "Cattin Blandenier (2004)" – literature with some niche processes and body size constraints

[2] "Digel et al. (2014)" - observed in feeding trials and/or single-target molecular, phylogenetic relatedness, body size ratios

[3] "McLaughlin et al. (2010)" – dissected 250 individuals but in no case could prey items be identified. Literature diet observations from other studies

[4] "Piechnik et al. (2008), Simberloff and Wilson (1969)" – literature. To order. Body size ratios.

While body size and species identity have been observed to be empirically important in some invertebrate food webs through empirical observation (e.g. aquatic invertebrate food webs, Woodward and Hildrew 2002, Rudolf et al. 2014), none of these food webs used empirical methods in natural environments.

[downsides/pluses/next directions]

New benefits: within-species data demonstrates cross-species differences in scaling (Nakazawa and other “food webs need individual observations” papers). Diet DNA metabarcoding dataset provides an opportunity to strengthen our understanding of predator-prey interactions for small-bodied organisms while examining these patterns at the level (individuals) at which predator-prey interactions occur, thus bridging the gap between species-level food webs and individual-level interactions (e.g. Nakazawa 2017, Stouffer 2010). These and other datasets that describe feeding interactions at the individual level (e.g. Woodward and Hildrew 2002) for small-bodied invertebrates will link patterns that emerge at the food web level with the mechanisms that drive predator-prey interactions at the individual level (Stouffer 2005, Ings et al. 2009).

[Somewhere in here contextualize Eitzinger et al. 2018 paper which is super cool about prey selection vs. prey availability using DNA metabarcoding]

Downsides: prey body size data aggregated at same level as in the past. Better data in databases online could refine these analyses in the future, suggesting they will just get better.

P: Concluding: not body size only, not species only, individuals acting on interactions at the individual level. A promising validation and refinement of what we already know: body size and species identity are important.

[Some rephrasing of this from intro] Small-bodied terrestrial invertebrates, including insects and spiders, comprise 50% of the earth’s animal biomass (Bar-On et al. 2018) and the majority of global animal species diversity (Mora et al. 2011 Costello et al. 2013), so understanding their trophic interactions is not only important for finding general food web patterns, but could be crucial for understanding the ecology of the world’s most dominant group of consumers shaping biomass and nutrient cycling dynamics (e.g. Bar-On et al. 2018).

**Acknowledgements**

This project was funded by the National Science Foundation (DEB #1457371), National Geographic Society, and a Faculty Research Grant from the UC Santa Barbara Academic Senate. We would like to thank field technicians [list non-authors] and laboratory technicians Emily Lutz and Tessa Chou [others!]. We would like to thank the U.S. Fish and Wildlife Service and Palmyra Atoll Research Consortium for supporting field work for this project. We would like to thank Dr. Ryoko Oono for use of her laboratory space and equipment. We acknowledge the use of the Biological Nanostructures Laboratory within the California NanoSystems Institute, supported by the University of California (UC) Santa Barbara and the University of California Office of the President. We especially thank Dr. Jennifer Smith, manager of the Biological Nanostructures Laboratory for her assistance in preparing and troubleshooting our samples. We acknowledge the use of computational facilities at the Center for Scientific Computing (CSC), which was purchased with funds from the National Science Foundation (CNS-1725797) and is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. We thank [EVERYONE WHO READS BUT NOT AUTHOR] H. Lowman, C. Jerde, M. Lee, and R. Ramiro for help in aspects of this manuscript, including statistics, bioinformatics, framing, and editing. We thank XX anonymous reviewers for help revising this manuscript. This is publication number PARC-XXX from the Palmyra Atoll Research Consortium.

**Literature Cited**

**Figures**

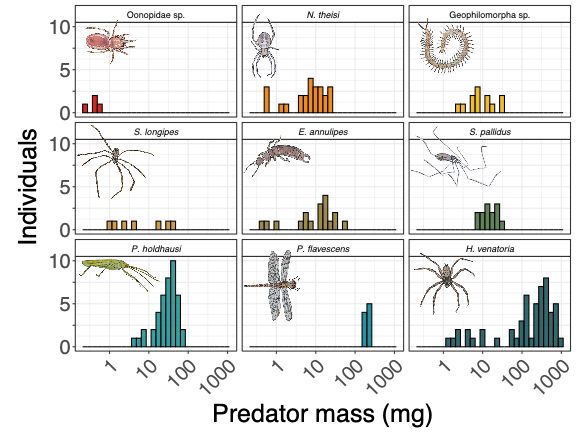
****

Figure: Predator size distributions across the nine predator species. The x-axis scale depicts absolute values but has been log10 transformed. Predator individuals span from 2x10-1 to 9.3x102 mg in wet weight and the facets in this figure have been ordered by increasing predator species mean size.



Figure: Larger predators eat larger prey based on a log10-log10 transformed relationship (panel (a), log10 prey size = 0.41\*log10 predator size), though the effect is mediated by species identity (b). Continuous axis labels represent absolute values but the scale between them has been log10 transformed. In panel (b), “+” and “-“ icons indicate species that either have higher (“+”) or lower (“-“) prey sizes relative to predator body size and the general predator-prey body size patterns. The dashed line in panel (a) represents the 1:1 relationship between predator and prey size.



Figure: Web-using traits increase the relative size of prey compared to predators (smaller ratios mean larger prey relative to predator individuals). Predators with web-using traits can relax traits related to gape limitation to access larger prey (p-value = 0.05, β = 1.25). The y-axis is presented with absolute values but displayed on a log10-transformed scale to demonstrate spread in the data. The dashed line indicates the 1:1 ratio where predators and prey are the same size; any interaction below this line indicates prey items that are larger than predator individuals.



Figure: comparison of our predator-prey size relationship to that from four food webs compiled from the literature. Our predator prey body size relationship had a significantly shallower slope than that of food webs in the literature (log10(prey mass) = 0.26\*log10(predator mass) vs. log10(prey mass) = 0.71\*log10(predator mass), p-value < 0.001).