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

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

ProcB: 6 MS pages, ~4000 words + 3 figures from abstract to end of literature cited.

**Abstract (200 words)**

Predator-prey interactions are shaped by a combination of predator traits, including body size, hunting mode, and phylogenetic identity. However, while these traits can predict predator-prey interactions across systems where predator-prey interactions are empirically observed, less is known about how these traits mediate predator-prey interactions that are challenging to observe, such as for small-bodied terrestrial invertebrates. In this study, we combined diet DNA metabarcoding data of 182 individual invertebrate predators from nine species with predator and prey sizes to explore how predator traits shape interactions. We found that 1) predator size and species identity shape prey size selection, with species-specific variation to a general size scaling relationship, 2) predator individuals with a shared prey pool partition prey by predator species but not predator size, and 3) the use of hunting “tools” (webs) allows predator individuals to select significantly larger prey than predators that do not use webs. Our findings indicate that predator-prey traits are important in shaping trophic interactions in invertebrate food webs and could help predict how anthropogenic biodiversity change will influence terrestrial invertebrates, the earth’s most diverse and biomass-dominant taxonomic group.

**Keywords (3-6 words)**

Food web, arthropod, allometry, phylogenetic relationships, predator traits

**Introduction**

Baby:

Predator-prey interactions shape the structure and function of ecosystems and their responses to external influences, including anthropogenic global change (McCann 2000, Brodie et al. 2014). Because predator-prey interactions are integral to ecosystem function, predicting these interactions is key for extrapolating general rules of predator-prey interactions in new contexts and finding general patterns that describe biological communities. Traditionally, predator-prey interactions have been approached from a species-specific framework; specifically, species identity or phylogenetic relatedness shapes feeding interactions (Ings et al. 2009). However, more generalizable predictions of feeding interactions can be made using non-specific traits. Body size, for example, is a key trait that determines feeding interactions between predators and prey across ecosystems (Woodward et al. 2005). Because body size is integral to feeding interactions, both dictating the rate and range of prey a predator can consume, it is one of the primary approaches for predicting the structure of feeding interactions for biological communities, or food webs (Stoufer 2015, Nakazawa, Woodward et al., Gravel et al. 2013, 2015). While body size alone predicts general patterns across food webs in multiple contexts, combining body size with more species-specific 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 generalizable 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 sometimes predict patterns in empirical food webs (e.g. Gravel et al. 2013), for some predators, in particular small-bodied invertebrate predators for which empirical diet methods (e.g. gut dissections) are impossible or often unfeasible (Sheppard and Harwood 2005, McLaughlin et al. 2010), we have a dearth of observed interaction data with which to validate extrapolated approaches to predicting interactions based on general rules. For these consumers, species interactions are often inferred from literature reports of observed interactions from phylogenetically-related species (e.g. Laigle et al, Cattin Bannuet, Simberloff and Wilson, Piechnek et al.) or based on body size feeding constraints (e.g. Laigle et al, Digel et al., 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, other 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 consequential 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 consumer 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 in natural field conditions. We combined these data with an extensive dataset of body sizes for both predator individuals and the prey groups identified in their diets. To understand how predator size, species identity, and cross-species feeding traits may drive empirical predator-prey interactions, we asked: 1) do larger predator individuals eat larger prey and is this mediated by predator species identity? 2) do predator individuals in a shared environment partition prey resources based on predator species identity, predator size, or both? and 3) beyond species, do predator traits shared by multiple species related to hunting “tools” (web-use) change prey size selection? 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, including different forest types and microhabitats representing 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 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, Miller-ter Kuile et al. *in revision*). All individuals were stored in 95% EtOH at -20ºC before DNA extraction.

We identified all predators to morphospecies and later validated unique species by DNA metabarcoding sequence data. The predators sampled represent the most common predator species found in each food web compartment (microhabitat location) and span a body size range of 0.2 – 928 mg (wet mass, Figure 1, SI Table 1). These predators included five arachnid species (Oonopidaesp., *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 (Geophilomorphasp.). 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.

*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 1), we ran samples across four separate sequencing runs (SI Table 2). 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 2). 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, SI Figures 3 & 4). We removed samples from analysis that had not been sequenced to sufficient depth using iNEXT (Hsieh and Chao 2016) and a lower quantile cutoff (SI Figures 5 & 6). 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 et al., Eitzinger et al.) 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 (van der Valk et al. 2020).

*Predator and prey size determination*

We converted predator lengths to wet mass using mass-length scaling relationships for each predator species from existing datasets (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 (SI Figure 9).

*Data analyses*

To determine 1) whether individual predator size, species, or both predicted prey size, we fit a linear mixed effects model with the response variable of log10 prey mass (in mg) and predictor variables of log10 predator mass (in mg), species identity, and their interaction, with a random effect of predator individual. Then, to examine 2) whether predators in a shared environment partition prey based on predator species identity, size, or both, we subset a set of predators collected within the same collection site (the canopy of *Pisonia grandis* on Sand Island n = 73 individuals of four species, including *P. holdhausi*, *E. annulipes*, *H. venatoria*, and *N. theisi*, SI Figure 10) and performed a canonical correspondence analysis (CCA) of the matrix of the feeding interactions of these individuals and the predictor variables of predator species identity and log10 predator mass using the vegan package in R (R version 4.0.2, package version 2.5-7, Oksanen et al. 2020). Subsetting just these individuals ensured that the predators in the analysis were partitioning the same potential prey pool and thus removed the covariate of variation in prey availability across microhabitats. Finally, to explore 3) how other predator hunting 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.

*Statistical model selection*

For the linear mixed effects models examining size, species identity, and species traits (1) and 3)), we performed model selection using the dredge() function in the MuMIn package in R (package version 1.43.17, Barton 2020) to compare nested models and chose the model with the lowest AICc value. For all models, we verified model assumptions for best-fitting models using the DHARMa package in R (version 0.3.3.0, Hartig 2020). For the CCA analysis, we extracted the marginal significance of each predictor in the full model, refit the model removing the least significant (at p-value > 0.05) variables and chose the best fitting model comparing both the full and reduced models with the anova() function in R. The color palette in our figures is from the calecopal package (version 0.1.0, Bui et al. 2020).

**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 (SI Table 3, SI Figures 8 & 9).

*Statistical results*

The best performing model predicting prey size included the terms of predator mass and predator species identity, but not their interaction (log10(prey mass) = 0.41\*log10(predator mass), R2m = 0.30, R2c = 0.35, with significant variation in by-species intercepts, Figure 2, SI Table 4). The best-fitting CCA model included the term of predator species identity, but not that of predator size (Figure 3, adjusted R2 = 0.10, p-value < 0.001) and explained 15% of the variation in the data, with 5.7 and 5.0% being explained by the first two axes, and loadings being driven by *E. annulipes* on CCA axis 1 and *P. holdhausi* and *N. theisi* on CCA axis 2 (SI Tables 5 & 6, SI Figures 10-12). Lastly, species that use webs either to capture or subdue prey (five arachnid species) had significantly smaller predator-prey size ratios than predators without web-use traits. (median ratios of predator to prey size of 9.71 for non web-builders versus 2.57 for web-builders, with larger numbers corresponding to larger predator in relation to prey; p-value = 0.05, β = 1.25, Figure 4, SI Table 7).

**Discussion**

For terrestrial invertebrate predators, 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 predator mass, species identity, and feeding traits are important drivers of prey selection and resulting food web patterns. 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. In a shared environment, 2) predator individuals partition prey by predator species, but not by predator size, thus suggesting species niches dictated by attributes beyond body size particular to those predator species. Broadening beyond species-specific analyses, 3) we demonstrate that more general species traits, in particular, the hunting strategy of web use, whether in capturing or subduing prey, enables some predator species to eat proportionally larger prey items. These results highlight that many food web patterns in small, terrestrial, invertebrate predator-prey interactions may be explained by a combination of species traits and that not one predator attribute alone predicts all interactions (e.g. approach in Pomeranz et al. ). Building comprehensive empirical datasets for interactions among small-bodied predators and their prey will continue to refine these multi-trait approaches.

[Not **just** size, but also traits and species]

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. In our results, individuals of predator species that may be more limited in diet based on prey that can fit in their mouths (“gape-limited”, e.g. *Pantala flavescens* and *Geophilomorpha* sp.) have smaller prey items on average compared to predator individuals 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). What this means is that a small-bodied individual of a non-gape-limited species may have the same prey size range available to it as a large-bodied individual of a gape-limited species (Nakazawa). While both sets of feeding interactions are still constrained by predator and prey size, these constraints vary depending on predator identity, or, more broadly, predator traits related to gape limitation. Our dataset demonstrates that some traits that limit prey size selection are phylogenetically conserved (e.g. web use among Order: Araneae), suggesting approaching these food webs with phylogenetic or trait-based filters (e.g. Grey et al., Brose et al. 2019, Laigle et al. 2017, Pomeranz) in addition to body size constraints.

[What other traits, because sometimes size doesn’t predict things??]

Other predator traits are likely to influence prey selection criteria in this consumer group as well, an idea that has already gained attention both within invertebrate food webs (Laigle et al. 2017) and scaled across food webs in many environmental contexts (Brose et al. 2019). For invertebrate-only food webs, hunting strategy and hunting “tools” are promising traits in terms of generality and their role in shaping predator niche (e.g. Schmitz spider papers). Indeed, given that the predators in our analysis of a single environment partitioned prey items by predator species without a significant effect of predator body size (Figure 4), it seems likely that some species-specific hunting trait may drive variation in prey selection for small-bodied invertebrates, even when they have the same available prey items. In our dataset, a few promising traits and surprising variations from expected body size patterns may provide fruitful next steps for diet DNA metabarcoding in these consumers. Specifically, while centipedes are generally thought to relax size constraints due to venom use (e.g. Digel et al.), the centipede, Geophilomorpha sp., consumed only proportionally small prey (mean predator-prey mass ratio 247 ± 61:1, Figure 2B), suggesting not all hunting tools always relax body size constraints. Conversely, the earwig *E. annulipes,* for which neither venom or web use is reported as a hunting strategy, fed on proportionally very large prey compared to predator body size (mean ± standard error of predator-prey mass ratio 4.35 ± 1.99:1, Figure 2B), either suggesting some other tool use in this predator or a reliance on scavenging (Wilson-Rankin et al. papers). 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.). How important these traits may be in structuring a particular food web will depend on the species composition and the relative proportion of predator diversity and biomass with each of these traits, but differences among these communities in structure and function may be linked to the different constraints faced by predators with different hunting traits (e.g. Os Schmitz papers).

[DNA data is cool]

Diet DNA metabarcoding will continue to be an important tool in understanding the biology of small-bodied invertebrate consumers because it allows us to examine invertebrate diets at the same resolution as that of the diets of larger-bodied species (Baker et al. 2014, Duffy and Jackson 1986, Hyslop 1980). As databases of DNA sequences continue to grow (Porter and… 2018), these analyses will likely get more specific, potentially surpassing the resolution of other methods (e.g. McElroy et al. 2020). Going beyond species-level traits, diet DNA metabarcoding data can provide data at the individual level, thus illuminating mechanisms describing patterns in predator-prey interactions at the scale at which they occur, thus bridging the gap between species- or size-based food webs and individual interactions (Nakazawa 2017, Ings et al., Stouffer 2010). Combined with other environmental datasets (e.g. prey densities), diet DNA metabarcoding data can be used to describe interaction strengths, which can predict dynamics in biological communities through time and space more accurately than binary interaction occurrence data (Eitzinger et al. 2018, Preston et al. 2019, Tang et al. 2014, Emmerson et al. 2004).

[mysteries revealed and an appeal for empirical data for the most abundant organisms on earth]

Small-bodied invertebrate predators are the most diverse and abundant predators on earth (Bar-on et al.2018, Mora et al. 2011, Costello et al. 2013) and until now, the predation interactions of these consumers have been largely unknown. Like other predators in multiple other ecosystem contexts (Brose et al 2019), the predation interactions of small-bodied predators are driven by a combination of measurable and generalizable traits, including body size, phylogeny, and hunting strategy. Using empirical datasets, such as those built by diet DNA metabarcoding data, will be key to determining which traits shape and mediate species interactions. Not only will this build a deeper understanding of the generality of feeding interactions and food webs across environmental contexts and consumer groups, but could be key to predicting and mitigating ongoing biodiversity loss (e.g. Valiente-Banuet, Borvvall et al., Donohue et al. 2017).

**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] C. Owens, R. Harris-Gavin, S. Copeland, G. Titcomb, 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**

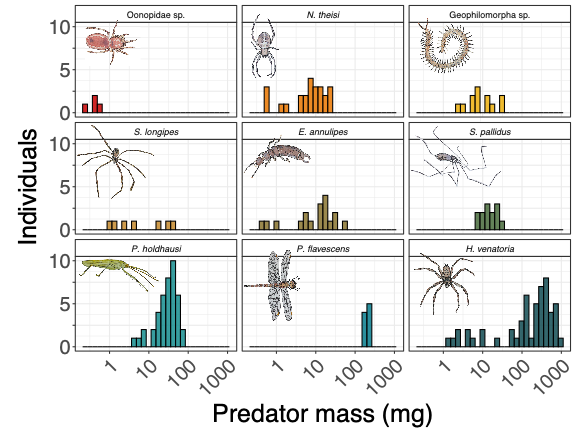
****

Figure 1: 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.



Figure2: 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 3: 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.

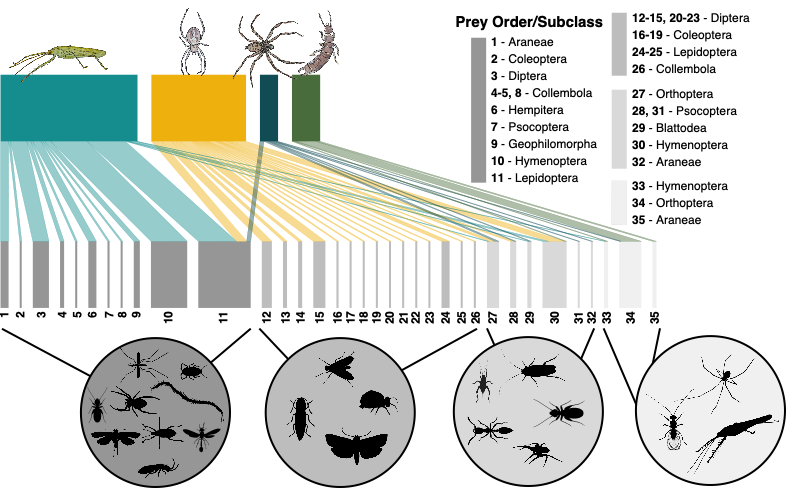


Figure 4: Predators in a shared environment partition resources by predator species, not by predator body size (CCA analysis R2 = 0.10, p-value < 0.001, Supplementary Figures). The top bars depict predators and are scaled by abundance of that predator species within the analysis (wider bar = more abundant). The bottom bars depicting prey are scaled by the frequency of that prey in an interaction with a predator individual (wider bar = more frequently consumed across predator individuals). Each prey family is identified by a number with the order or subclass (for Collembola) indicated in the legend. The particular prey families are included in an alternative version of this figure in the supplement.