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

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

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:

Predator-prey interactions shape the structure and function of ecosystems and their responses to external influences, including anthropogenic global change (CITE, McCann, others, extinction cascades). 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 (CITE). However, more generalizable predictions of feeding interactions can be made using non-specific species traits. Body size, for example, is a key trait that determines feeding interactions between predators and prey across ecosystems (Woodward, etc). 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 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 general rules. 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, 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 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 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 between these predators and their prey 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 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) 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) relate to 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 (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, Figure 1, 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.

*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).

*Dat 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 of size distribution) 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. Subsetting just these individuals ensured that the predators in the analysis were partitioning the same potential prey pool and thus removing the covariate of variation in prey availability across microhabitats. Finally, to explore 3) 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.

*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 (Versions, citations) 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 (versions). For the CCA analysis, we extracted the marginal significance of each predictor in the full model, refit the model without 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 is 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 2, SI Table 3). The best-fitting CCA model included the term of predator species identity, but not that of predator size (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 (Supplementary figure). A depiction of prey composition by predator species is given in Figure 3. Lastly, species that use webs either to capture or subdue prey (five arachnid species) had significantly larger prey relative to individual predator size 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).

± SE; p-value = 0.05, β = 1.25, Figure 4).

**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 predator mass, predator species identity, and predator feeding traits 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. In a shared environment, 2) predator individuals partition prey by predator species, but not by predator size, thus suggesting species niches dictated by traits 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 body size and predator traits that are conserved among phylogenetically-similar predator species, lending empirical evidence to an approach for building and predicting predator-prey interactions based on a combination of traits (including body size and hunting mode) and species-specific variation. Building comprehensive empirical datasets for interactions among small-bodied predators and their prey will help to build more robust empirical datasets with which to 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 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 one species may have the same prey size range available to it as a large-bodied individual of a different species with a more gape-limited hunting strategy (Nakazawa). While both sets of feeding interactions are still constrained by size, these constraints vary depending on predator identity, or, more broadly, predator traits. Our dataset demonstrates that some traits that limit prey size selection may be phylogenetically conserved (e.g. web use among Order: Araneae), suggesting that approaching these food webs with phylogenetic or trait-based filters (e.g. Grey et al., Brose et al. 2019, Laigle et al. 2017) could explain variation not explained by the trait of body size alone.

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

Our data demonstrate the influence of several traits, specifically body size and web use, in shaping predator-prey interactions for small-bodied terrestrial invertebrate predators. However, other traits are likely to influence prey selection criteria in this consumer groups, 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” seem like 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 clear that some species-specific or hunting trait drives variation in prey selection for small-bodied invertebrates, even when they have the same prey items available to them, regardless of prey size. In our dataset, a few promising traits and surprising variations from expected patterns may provide fruitful next steps for diet DNA metabarcoding in this consumer group. Specifically, while centipedes are generally thought to relax size constraints due to venom use (e.g. Digel et al.), the centipede Geophilomorpha sp., in our dataset, consumed only proportionally small prey (mean predator-prey mass ratio 247 ± 61:1), 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), 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.). 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).

[downsides/pluses/next directions]

Diet DNA metabarcoding data provide a promising validation and refinement of what we already know: species traits, including body size, hunting mode, and species identity, are important for structuring predator-prey interactions. Combined with other environmental datasets (e.g. prey densities), diet DNA metabarcoding can also take interactions beyond binary occurrence and provide a value for the strength of interactions, which, more than binary interaction occurrence, describes the dynamics of predator-prey interactions and how they relate to ecosystem function (Eitzinger et al. 2018, Preston et al. 2019, CITE FUNCTION ONE).

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] 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**

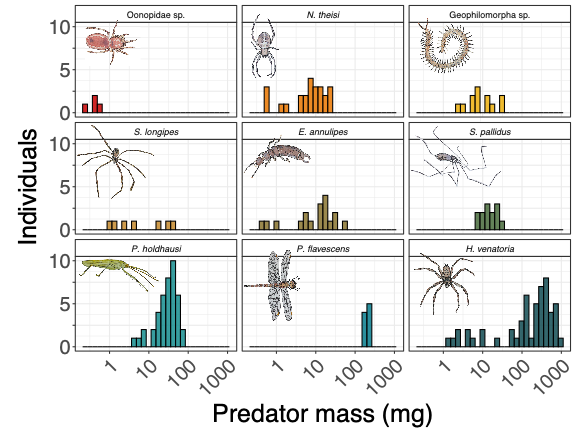
****

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.

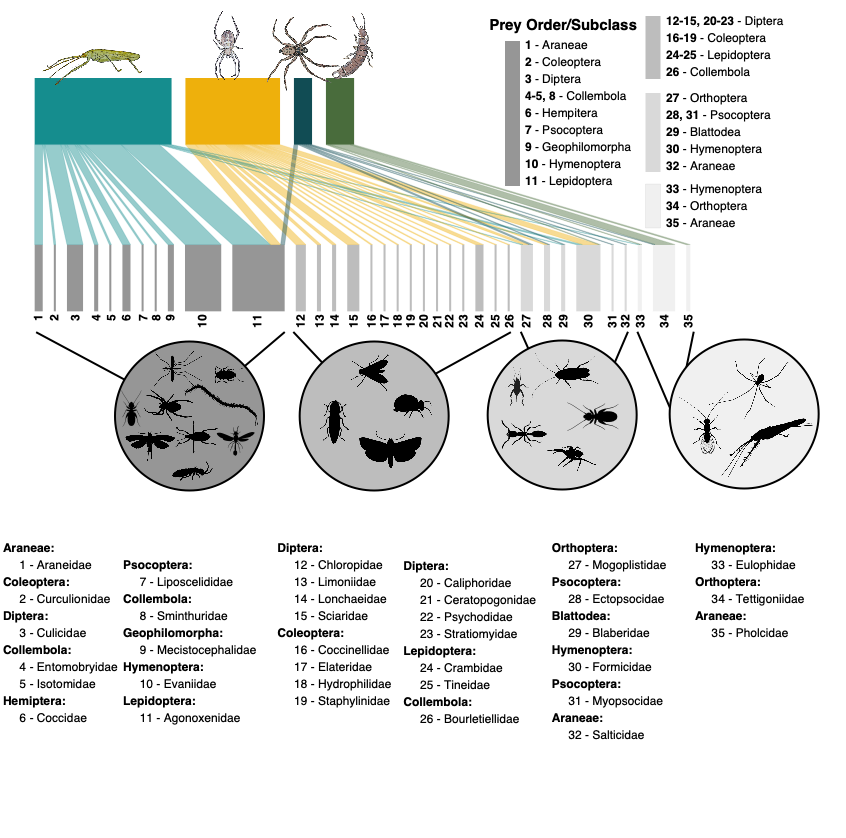


Figure 4: Predators in a shared environment partition resources by predator species, not by predator body size (CCA analysis R2 = 0.10, p-value < 0.001, Supplementary Figures). The top bars depict predators and are scaled by abundance of that predator species within the analysis (wider bar = more abundant). 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 as follows (and maybe just in supplement):

1. Araneae: Araneidae
2. Coleoptera: Curculionidae
3. Diptera: Culicidae
4. Collembola: Entomobryidae
5. Collembola: Isotomidae
6. Hemiptera: Coccidae
7. Psocoptera: Liposcelididae
8. Collembola: Sminthuridae
9. Geophilomoropha: Mecistocephalidae
10. Hymenoptera: Evaniidae
11. Lepidoptera: Agonoxenidae
12. Diptera: Chloropidae
13. Diptera: Limoniidae
14. Diptera: Lonchaeidae
15. Diptera: Sciaridae
16. Coleoptera: Coccinellidae
17. Coleoptera: Elateridae
18. Coleoptera: Hydrophilidae
19. Coleoptera: Staphylinidae
20. Diptera: Caliphoridae
21. Diptera: Ceratopohonidae
22. Diptera: Psychodidae
23. Diptera: Stratiomyidae
24. Lepidoptera: Crambidae
25. Lepidoptera: Tineidae
26. Collembola: Bourletiellidae
27. Orthoptera: Mogoplistidae
28. Psocoptera: Ectopsocidae
29. Blattodea: Blaberidae
30. Hymenoptera: Formicidae
31. Psocoptera: Myopsocidae
32. Araneae: Saliticidae
33. Hymenoptera: Eulophidae
34. Orthoptera: Tettigoniidae
35. Araneae: Pholcidae