A highly resolved multiplex network reveals the structural role of insects and plants in terrestrial food webs

Kayla R. S. Hale1,2*, John David Curlis1, Giorgia G. Auteri3, Sasha Bishop1, Rowan L. K. French4, Lance E. Jones5, Kirby L. Mills1,6, Brian G. Scholtens7, Meagan Simons1, Cody Thompson8, Jordon Tourville9,10, Fernanda S. Valdovinos11

1 Department of Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA
2 Department of Integrative Biology, University of Guelph, Guelph, ON, CA
3 Department of Biology, Missouri State University, Springfield, MO, USA
4 Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, ON, CA
5 Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
6 School for Environment & Sustainability, University of Michigan, Ann Arbor, MI, USA
7 Department of Biology, College of Charleston, Charleston, SC, USA
8 Museum of Zoology, University of Michigan, Ann Arbor, MI, USA
9 Department of Environmental Biology, SUNY College of Environmental Science and Forestry, Syracuse, NY, USA
10 Research Department, Appalachian Mountain Club, Boston, MA, USA
11 Department of Environmental Science and Policy, University of California, Davis, CA, USA

* Corresponding author; e-mail: kaylasal@umich.edu

ORCIDs

Hale, https://orcid.org/0000-0002-2627-9623; Curlis, https://orcid.org/0000-0002-2974-7330; Auteri, https://orcid.org/0000-0002-5579-8078; Bishop, https://orcid.org/0000-0002-0877-3951; Jones, https://orcid.org/0000-0001-6368-9720; Mills, https://orcid.org/0000-0001-7693-9629; Thompson, https://orcid.org/0000-0002-6686-6056; Tourville, https://orcid.org/0000-0003-4005-5914; Valdovinos, https://orcid.org/0000-0002-5270-5286

Keywords

Ecological network, aboveground terrestrial ecosystems, temperate forest, trophic structure, plant-insect interactions, taxonomic resolution
Abstract

Terrestrial ecosystems support rich communities of species feeding on each other in different ways. Insects and plants comprise much of this species richness, but the structure of their feeding interactions in aboveground terrestrial food webs is not well known. Historically, food web research has coarsely or unevenly grouped insects and plants, excluded “mutualistic” feeding interactions (e.g., pollinators eating nectar), and focused only on subsets of species or feeding interactions, especially “antagonistic” interactions involving tetrapods. Here, we combine public data and records from a biological research station into a cumulative food web of ~580,000 interactions among ~3,800 species in a temperate hardwood forest ecosystem. We include all feeding interactions, subdivided by whether animal tissues or plant leaves, flowers, fruits, or wood are being eaten. We represent these different types of feeding interactions in a multiplex food web and study the effect of taxonomic resolution on our understanding of food web structure. Our results indicate that insect herbivores engage in highly species- and tissue-specific feeding that is inaccurately represented by coarse taxonomic groups. The food web presented here is the richest and most evenly-resolved representation of feeding interactions yet reported and provides a tool for future management and research on terrestrial ecosystems.

1) Introduction

Empirical food webs are resources both for advancing ecological research and informing management practices. Food webs represent the potential biomass flows through an ecosystem by documenting the feeding interactions (links) between taxa (nodes). Studying these representations has allowed researchers to identify fundamental patterns in the structure of ecological systems. For example, aquatic and belowground food webs have been shown to be

Nevertheless, a lack of systematically collected data has limited even a basic account of the structure of aboveground terrestrial food webs (ATFWs). Most of the few published ATFWs (Brose et al. 2019) have focused on resolving only a subset of taxa (e.g., tetrapods), interactions (e.g., carnivory), energy pathways (e.g., a partial food web on a single resource), or potential habitats (e.g., under a log) in a system. Structural patterns in some subnetworks have been identified: larger predators eat smaller prey (Brose et al. 2006), while some types of insect herbivores form nested, specialized associations with plants (Thébault & Fontaine 2010). Yet it is still unknown how these subnetworks interconnect into ATFWs, or if these patterns are relevant in the context of a single system (Hale et al. 2023). Here, we present and study the structure of a food web that evenly resolves feeding interactions among all plants, insects, and tetrapods across aboveground terrestrial habitats at a long-term biological research site.

In building our empirical food web, we identified three problems that are especially prominent in aboveground terrestrial ecosystems. The first problem is uneven and coarse taxonomic resolution. In many cases, tetrapods are recorded to species-level, while insects are recorded to order or even coarser taxonomic levels, even though they represent a taxonomically richer part of the system (Harris & Paur 1972, Dunne 2006, O’Connor et al. 2020, Pringle &
As a result, potentially thousands of plant-insect interactions are reduced to a single trophic link of equal weight in analyses to a species-specific predator-prey interaction between tetrapods. The coarseness of this approach may undermine the usefulness of food webs for management, e.g., by missing critical interactions that affect the abundance of emerging insect pests. In other cases, species with the same or similar sets of consumers and resources are grouped into a single node and labeled a “trophic species” (Williams & Martinez 2000, Delmas et al. 2019). This more tractable representation includes only the unique biomass flows in the food web and is especially appropriate when all species within a taxonomic group have the same trophic interactions, rendering the resolution of interactions to species level unnecessary.

Nevertheless, even within such a group, species often vary in their resource usage (e.g., by consuming vegetation versus seeds), with different nutritional and demographic consequences (Pringle & Hutchinson 2020). A food web that evenly resolves both resource- and species-specific interactions could be a powerful tool to study the correspondence between taxonomic groups and trophic species.

The second problem in building an ATFW is the sensitivity of food webs to sampling effort. A single set of observations provides only a “snapshot” of an ecosystem (Dunne et al. 2006, de Aguiar et al. 2019), biased by the specific spatial (vertical versus horizontal transects, microhabitats) and temporal (seasonal, diurnal, duration) scales of sampling, as well as the taxonomic expertise of the investigators (including ability to catch and identify organisms).

Species also exhibit adaptive foraging and defensive behaviors, effectively “rewiring” trophic interactions in response to changing biotic and abiotic conditions (Bartley et al. 2019, Ceron et al. 2022, Valdovinos et al. 2023). As such, limited surveys miss rare or cryptic species and interactions, even though such species and interactions likely provide critical contributions to
ecosystem dynamics and function (Dee et al. 2019, Simpson et al. 2022). One solution to this
sampling problem is the concept of a cumulative food web or “metaweb” (Dunne 2006), in
which all species and interactions recorded over years and/or across similar habitats are pooled
into one network of all possible interactions between local species. Such an approach risks
missing system-specific interactions or swamping out the common species and links that
comprise the bulk of the bioenergetic flows in the system. However, as human actions change
species’ distributions and habitats, “rare” and novel interactions are increasing in frequency
(Bartley et al. 2019), making cumulative webs all the more important.

The third problem affecting the construction of empirical food webs is the conceptual
inconsistency in the criteria for including different types of feeding interactions. Ideally, a food
web maps all bioenergetic flows between organisms in an ecosystem. Historical ATFWs have
typically only included feeding interactions with ostensibly antagonistic (+/−) effects, such as
predation, herbivory, and sometimes parasitism/parasitoidy. Other feeding interactions with
ostensibly mutualistic (+/+ effects, such as nectarivory (pollination) and frugivory (seed
dispersal), are excluded. However, mutualistic feeding interactions constitute dynamically
important bioenergetic flows in aboveground terrestrial ecosystems (Baude et al. 2016, Hale et
al. 2020), and ignoring them is more an historical artifact of disciplinary subdivisions than a
useful approach to studying biotic interactions or ecosystem functions (Hale & Valdovinos
2021). Moreover, categorizing interactions by effect is problematic because effects can change
over time, space, and according to biotic and abiotic conditions (Chamberlain et al. 2014). For
example, mutualistic interactions can be antagonistic when nectarivory is not coupled with
sufficiently effective pollination behavior (Bronstein 2001). Similarly, antagonistic interactions
can be mutualistic, such as when herbivory stimulates greater plant production in a phenomenon
called overcompensation (Ramula et al. 2019). Therefore, there is growing consensus (Abrams 1987, McCall & Irwin 2006, Holland & DeAngelis 2009, Valdovinos et al. 2023) that feeding interactions should be differentiated by their mediating traits (phenology, ontogeny, morphology), tissues (nectar, fruit, leaves), and bioenergetic or demographic consequences (yields and production costs), rather than their presumptive net effects.

We began to address these problems by building a food web for a temperate hardwood forest based upon ~100 years of study by diverse specialists at a biological research station. In constructing the food web, we emphasized the even resolution of plant, insect, and tetrapod taxa, resulting in ~580,000 feeding interactions among ~3,800 species. Interactions were collected from public records and online databases, supplemented by experts and vetted for local plausibility given species’ spatial and temporal co-occurrences, traits, and behaviors. We included all types of trophic interactions in a “multiplex” food web subdivided by resource type (i.e., animal and plant tissue types) rather than presumptive effect (i.e., antagonistic versus mutualistic). Using this food web, we discuss the effects of taxonomic resolution on our understanding of food web structure.

2) Methods

a) Conceptual framework

To determine criteria for including interactions in our food web, we applied a recently-proposed framework for studying aboveground terrestrial food webs (ATFWs, Valdovinos et al. 2023). This framework asserts that the complexity of plant tissues fundamentally distinguishes aboveground terrestrial ecosystems from their aquatic and belowground counterparts, shaping the structure and dynamics of ATFWs through distinctive patterns of growth and turnover. Different plant tissue types are often only expressed at certain times phenologically and
ontogenetically and often require consumers to possess specific morphological and/or physiological traits to access them. The availability and nutritional yield to the consumer, as well as the costs of production, maintenance, and loss to the resource, vary substantially between tissue types. Such differences are highly consequential; tissue-specific feeding may underlie bioenergetic flows, such as primary production and consumption, in aboveground terrestrial ecosystems and explain how species coexist even when their partners (consumers or resources) are otherwise similar. Accounting for this tissue-specific complexity therefore provides more information than traditional static food web structures and may be key for moving food webs from simple graphical representations to useful tools for research and management.

With this in mind, we chose to include all direct interactions among species in our system with a bioenergetic flow (i.e., one species consuming another, regardless of their potential ecological effects), differentiated by their focal resource. We broadly categorized the resources as (1) animal tissues (live and scavenged) and plant tissues, grouped into (2) leaves and stems, (3) flowers, nectar, pollen, etc., (4) seeds, fruits, etc., and (5) wood and bark. Hereafter, we refer to these resource types simply as “animals,” “leaves,” “flowers,” “seeds,” and “wood.” Due to our focus on aboveground terrestrial plants and insects, we excluded fungi and detritus resources. The tissues in each resource category do not have uniform nutritional content, but our framework provides a start towards representing the different types of resource use and associated foraging behaviors in food webs.

b) Site description

The University of Michigan Biological Station (UMBS) includes ~10,000 acres of land that is used for teaching and research in northern lower Michigan, USA (45°35.5′ N, 84°43′ W). The property is composed predominantly of dry-mesic, northern hardwood forests with patches
of wooded wetlands (hardwood conifer swamp) flanked by two lakes (Barnett et al. 2022). This
is a strongly seasonal, temperate system with historically cold, snowy winters (average minimum
temperature of \(-12.1\, ^\circ C\) in January) and hot, humid summers (average maximum temperature of
26.2\, ^\circ C\) in July; Wang & Ibanez 2022). Soil composition varies across microhabitat patches from
sandy outwash plain to moraine. Prior to acquisition by the University of Michigan in 1909, the
landscape was almost completely logged and burned, with only a few old-growth forest tracts
remaining. Therefore, the vegetative landscape is relatively uniform in age, with variation in
stand structure and composition attributable to glacial landforms and differences in soil

UMBS is contiguous to other forested habitats, allowing free movement in and out for
mobile species, including those that seasonally migrate. Such transient organisms, coupled with
those that are only active or present aboveground during certain times of year, lead to strong
effects of seasonality on species composition. Though their impacts may only be temporary, we
included these species in our analyses because they may represent critical consumers or
resources for local species at certain times of year. In particular, most species of ectotherms at
UMBS are inactive during winter months, and many migratory bird species are present only
during the spring and summer. By pooling observations into a cumulative web, we avoid
sensitivity of food web structure to this spatial and temporal transience.

c) Species list

To construct the food web, we began from taxonomic lists of mammals, amphibians,
reptiles, vascular plants, birds, insects, and non-insect arthropods. These lists represent an
accumulation of decades of records at UMBS, from resident biologists’ personal observations,
student papers and research projects, regional lists, museum specimens, online databases (such as
eBird and iNaturalist), and semi-regular BioBlitz events, in which teams of biologists roamed the site and identified as many organisms as possible.

Where possible, we updated records to the most recently accepted species names according to the Integrated Taxonomic Information System (ITIS) using the “taxize” package in R (Chamberlain et al. 2013). We excluded taxa that could not be resolved to at least genus-level (e.g., parasitoid wasp family Diapriidae). For taxa that could only be resolved to genus, we excluded those that are highly speciose in the Nearctic (> 20 species, e.g., the bloodworm genus *Chironomus*) and those that had congenerics already included in the food web (e.g., six species of masked bee *Hylaeus* were included, but another unidentified species was excluded). Finally, we treated variants and subspecies (e.g., deer mouse *Peromyscus maniculatus gracilis*) as their binomial species name. Hereafter, we refer to all taxa occurring at UMBS as “species,” though a small fraction (4.5%) are genera.

The species lists were vetted and approved by experts (generally, the authors) that have extensive knowledge of the communities and natural histories of the organisms that occur in the region. In determining which species to include or exclude, we excluded any species that do not occur at UMBS or do not have a significant lifestage or feeding behavior in aboveground terrestrial habitats (O’Connor et al. 2020). We defined “aboveground terrestrial habitats” as land above the soil layer (including leaf litter and above), on terrestrial plants (growing potentially over but not exclusively in water), or in air (above other terrestrial habitats). Thus, we excluded species that live and feed primarily on the surface of aquatic habitats, such as waterlilies (Nymphaceae) and water bugs (Belostomatidae). Additionally, given our primary focus on even resolution among plants, arthropods, and tetrapods, we chose to exclude some major groups, such as bryophytes, fungi, lichens, and molluscs.
d) Feeding interactions

To assemble a list of potential feeding interactions between species, we referenced region-specific field guides, online databases (e.g., Animal Diversity Web, Birds of the World), and central repositories that automatically scrape data from museum records and the web (e.g., Global Biotic Interactions). Trained undergraduate researchers searched sources by species name and recorded predators, diet, and other potential feeding interactions with as much resolution as possible. Each focal taxon was resolved to species-level, but their interaction partners could be recorded at any taxonomic level (e.g., species $x$ eats Family $y$). If we could not find information on a given focal species, we did not extrapolate from closely related species. Our records were supplemented, revised, and annotated by the same experts who vetted the species lists. Where possible, records were assigned to experts for both the focal species and its partner’s group (that is, two separate experts). We provide the final list of annotated interactions and their experts online (Hale 2023).

Experts were asked to assess the plausibility of the recorded interactions occurring at UMBS. Interactions between species $x$ and $y$ were considered plausible if the two species co-occur (with respect to phenology, activity patterns, microhabitat usage) and have no trait incompatibilities (it would not be possible for $x$ to acquire, ingest, or assimilate $y$). Interactions were considered plausible even if they could be considered inefficient ($x$ can consume $y$, but this is rare or unlikely, especially if more rewarding and easily acquired foods than $y$ are available). Interactions that are non-consummptive (e.g., nesting or hunting locations) or occur outside of aboveground terrestrial habitats were excluded even if they occur between local species (e.g., least sandpiper *Calidris minutilla* preys upon toad *Anaxyrus americanus* tadpoles, but this occurs exclusively in aquatic habitats).
Experts were also asked to assess the plausibility of the taxonomic level at which interactions were recorded. For example, when Wilson’s warbler (*Cardellina pusilla*) is recorded to eat beetles (order Coleoptera), it is unclear if: (case 1) *C. pusilla* eats every species of Coleoptera, potentially or opportunistically, including all local Coleopterans, or (case 2) *C. pusilla* eats (at least) one species of Coleoptera that was not identified in the original record and which may or may not be local. Either case may be possible, depending on the biology of the species. At one extreme (case 1), the lack of taxonomic resolution in the record could reflect an ecologically relevant lack of discrimination by the focal species (e.g., opportunism due to sensory capabilities in distinguishing between predators or resources), with the probability of interactions occurring between any given partner species determined more by interaction efficiency or abundance than by taxonomy. At the other extreme (case 2), the lack of taxonomic resolution could simply reflect a lack of knowledge, representing a summary of potentially highly specific interactions. In the first case, experts accepted the interaction, while in the second, experts rejected the interaction unless it could be plausibly revised for local taxa.

In the original sources, interactions were often recorded following traditional disciplinary categorizations. In the absence of other information, we assumed that animals recorded as “eating,” “feeding on,” “consuming,” “parasitizing,” or “hosted by” plants were consuming plant leaves. We assumed that animals recorded as “pollinator of,” “visiting flowers of,” “visiting,” or “mutalist of” plants were consuming flowers. Other mutualisms involving potential feeding such as seed dispersal (through frugivory, scatterhoarding, etc.) and ant protection (of hemipterans or plants) were noticeably underrepresented in our records. We assumed animals recorded as “disperser of” plants were consuming seeds. We assumed that interactions between animals recorded as “kills,” “predates upon,” “scavenges,” “host of,” “parasite of,” or “parasitoid
of” involved consumption of animal tissue. When necessary, these default assumptions were revised by experts to reflect true feeding on another tissue or a lack of feeding altogether.

e) Network representation and analysis

To translate our list of interactions into a food web, we used a “multiplex” network approach in which feeding on different types of tissues is represented by different types of links between the same set of nodes (species). This allows us to distinguish between the niches of animals feeding on different tissues of the same plant, while also representing that these tissues are coupled to each other in the same organism. Specifically, we defined a node for each focal species $i$ in our list. Then, we defined a directed link of type $k$ between nodes $i, j$ if $i$ consumes tissue type $k$ of $j$ or tissue type $k$ of a broader taxonomic group including $j$. Links are binary, indicating the presence or absence of potential feeding, not its frequency, probability, rate, or strength. We retained only unique links, but tracked the most resolved taxonomic level from which each link was sourced.

To assess the extent to which taxonomic groupings represent the unique trophic functions in food webs, we calculated trophic similarities across taxonomic resolutions. We used the Jaccard index, which ranges from 1, when a pair of species have all consumers and resources (both species and tissue types) in common, to 0, when species have completely unique sets of consumers and resources. For each genus, family, and order containing $> 1$ species, we calculated the average of the minimum trophic similarities between each species $i$ with all other species $j$ in its taxonomic group. We also calculated trophic similarity between each pair of species $(i, j)$ in our food web and created a “trophic species web” by grouping species with trophic similarities of 1 into single nodes. This grouping removes functionally redundant nodes.
and links. All data cleaning and network analyses were performed in MATLAB R2021b (Mathworks).

3) **Results**

a) **Species composition and database**

Our final species list includes 3,802 local species, representing 2,073 genera in 451 families of 86 orders (Fig. 1). Insects (2,669 species) and vascular plants (781 species) numerically dominate the food web, accounting for ~90% of the taxa, compared to tetrapods (313 species) and non-insect arthropods (39 species). The richest orders in our list are insects, especially Lepidoptera (butterflies and moths, 1,168 species), Coleoptera (beetles, 512 species), Diptera (flies, 390 species), Hymenoptera (bees, wasps, and ants, 265 species), and Hemiptera (true bugs, 211 species). Worldwide, there are more than twice as many named species of Coleoptera as Lepidoptera; therefore, Lepidopterans are likely substantially overrepresented in our list. The vascular plants are our most taxonomically diverse group, with 38 orders represented primarily by two groups of flowering plants, Poales (grasses, sedges, and rushes, 149 species) and Asterales (composite flowers, 96 species). Our tetrapod species are birds (226 species, including 127 passeriform birds), mammals (52 species, including 7 bats), amphibians (18 species), and reptiles (17 species). Finally, our non-insect arthropods primarily include spiders and mites, but overall, this group is significantly underrepresented in our list, both in terms of richness and taxonomic diversity.

Our database of feeding interactions between local species amounted to 26,747 approved records. In sum, our final food web totals 2,541 species of consumers (including 4 carnivorous or parasitic plants) and 3,782 species of resources. We have no records for 19 plants (~0.5% of local species), including most of the Lycopodiales (clubmosses, 5 of 6 species) and the
Polypodiales (ferns, 11 of 19 species), two of the major groups of non-seed plants in our system. Additionally, our records include no diet information for 485 species of insects (~16.3% of local animals), primarily from the richest orders (183 Lepidopterans, 145 Dipterans, 109 Coleopterans), but also including all Blattodea (cockroaches, 2 species), Plecoptera (stoneflies, 2 species), Mecoptera (scorpionflies, 4 species), and Zygentoma (silverfish, 1 species). Some of these do not feed in aboveground terrestrial habits (or at all) during a certain lifestage or feed entirely upon resources we excluded (fungi, detritus, lichens, etc.), limiting their potential diet in our food web. However, these gaps in our dataset may also indicate broader gaps in our natural history knowledge for these species.

b) Feeding interactions

Our final food web totals 586,601 unique feeding interactions between local species. These interactions consist primarily of feeding on animals (89.7%), especially insects, with the remaining interactions consisting of feeding on plant leaves (6.8%), flowers (2.2%), seeds (1.2%), or wood (0.17%). Though numerous, these interactions are only a small fraction of the possible links. Separating feeding on each type of resource, we calculate that only 5.8% of the carnivory interactions are realized among the 3,023 participating species, with similarly low connectance for the different types of herbivory (leaves: 3.1% among 2,480 species, flowers: 2.2% among 1,537 species, seeds: 5.0% among 804 species, wood: 9.8% among 200 species).

Comparisons to similarly rich networks are not available in the literature. However, our carnivory subnetwork is both more connected given its richness than expected from a scaling law fit to previous food webs (1.5%, Cohen & Briand 1984, Montoya & Solé 2003, Ings et al. 2009) and more connected than a recently published metaweb of only tetrapods (3.9% with 1,136 species, O’Connor et al. 2020).
Carnivory is the numerically dominant interaction in our food web, but only approximately one-third (35.9%) of consumers feed on animal tissues. These consumers exhibit a heavily skewed diet breadth distribution, where most feed on a small number of animal species but a few (primarily the bats and birds) are hypergeneralists (Fig. 2A). In fact, the 127 passeriform bird species (5.0% of consumers) contribute over half (51.4%) of the unique carnivory links in our network, sourced from only 711 (2.7%) records of focal bird species thought to feed opportunistically upon entire insect orders (especially Coleoptera and Lepidoptera). Records of feeding on insect orders by any taxon contribute 87.1% of unique carnivory links overall, meaning that they are not otherwise included by records at lower taxonomic levels (Supplementary Fig. 1A). In comparison, feeding between tetrapods accounts for only 1.3% of carnivory links.

Herbivory interactions are less numerous than carnivory, but most consumers (86.0%) in our food web feed on plants, with nearly half (46.0%) feeding on a single plant tissue. These are primarily insects, dominated by Lepidopterans eating leaves (as caterpillars), but also including Hymenopterans and Dipterans eating flowers (nectar). Like carnivores, herbivores across resource types have skewed diet breadth distributions (Fig. 2B-E). Considering specialization in terms of number of taxonomic resource families rather than resource species richness, we also see a skewed distribution, with many insect families feeding on a single plant family and a long tail of a few taxonomically generalized insect families (Fig. 2F). Perhaps reflecting this taxonomic specialization, over half of our herbivory interactions (66.4%) stem from records of feeding between insects and plants at the genus- and species-level, with only 4.1% of unique herbivory interactions contributed by order-level records across all taxonomic groups (Supplementary Fig. 1B). Therefore, in contrast to carnivory, our herbivory records at coarser
taxonomic levels do not include or are redundant to interactions from more taxonomically-
resolved records.

Over one-third of consumers (40.1%) feed upon more than one type of resource. Around
half of these consumers feed on leaves and flowers (19.1% of consumers, primarily
Lepidopterans and Coleopterans). A smaller fraction (11.3%) feed on > 2 types of resources, but
these are a more diverse set of insects, mammals, and birds feeding on animals, leaves, and
flowers or seeds, or, less frequently, leaves, flowers, and wood. Among these consumers, we
observed significant positive correlations between diet breadths when feeding on animals and
plant leaves (Pearson correlation: $r = 0.28$, $p = 2.0 \times 10^{-7}$, $N = 340$), animals and plant seeds ($r =
0.15$, $p = 0.044$, $N = 202$), plant leaves and seeds ($r = 0.42$, $p = 7.2 \times 10^{-9}$, $N = 178$), and plant
leaves and wood ($r = 0.23$, $p = 0.017$, $N = 98$; also see Supplementary Fig. 2). In other words,
among animals that access multiple types of resources, generalists on one resource type also tend
to be generalists on others. Interestingly, though many consumers access multiple types of
resources, only a very small fraction of interactions (0.65%) are between the same pair of
consumer and resource species, indicating feeding on multiple plant tissues by the same
consumer.

In contrast to the tissue specialization by most animals, most plants (91.3% of 781
species) support consumers on more than one of their tissues, with a small set of diverse plants
(85 species in 13 orders) sustaining feeding on all four recorded tissue types. We observed
significant positive correlations between the richnesses of consumers feeding on plants’ leaves
and flowers (Pearson correlation: $r = 0.16$, $p = 3.4 \times 10^{-5}$, $N = 653$), leaves and seeds ($r = 0.35$, $p
= 8.1 \times 10^{-17}$, $N = 548$), and leaves and wood ($r = 0.61$, $p = 1.7 \times 10^{-11}$, $N = 98$; also see
Supplementary Fig. 3). In other words, the plants most consumed by leaf-herbivores are also
most likely to host rich communities of consumers eating their other tissues. However, the richnesses of animals eating plants’ flowers and seeds was negatively correlated ($r = -0.11$, $p = 0.013$, $N = 514$). This may suggest a trade-off in plants’ ability to support flower and seed eaters (potential pollinator and seed dispersal mutualists, respectively) or a deterrence effect between seed- and flower-eaters in our system (primarily birds and insects, respectively).

c) **Trophic similarity**

Two thirds of our taxonomic species (2,396 of 3,802, 63.0%) are unique trophic species (Fig. 3A), meaning that they have unique sets of consumers and resources. Of these, 29 species are distinguished only by feeding on different plant tissue types. In total, all tetrapods, nearly all non-insect arthropods (97.4%), over three-quarters of plants (77.1%), and almost two-thirds of insects (62.7%) are trophically unique species. The remaining 30.9% of taxonomic species are pooled into 231 trophic species groups, each composed of species from a single taxonomic order. This may be because entire orders of insects share the same opportunistic/generalist predators and are therefore only distinguished by their diets. Indeed, over half of these species (54.1%, 760) are represented in only 27 trophic species groups; these comprise most (82.9%) of the 485 animal species without diet information (that is, the species most likely lacking resolution in our food web). The remaining trophic species groups include closely related taxonomic species, but do not perfectly correspond to lower taxonomic levels. For example, the 10 species of *Zale* moths in our food web are present in 8 different trophic species groups with other Lepidopterans across 8 families.

Even if they are technically unique, closely related taxonomic species tend to be trophically similar. Most genera and families have an average minimum trophic similarity close to one (median among genera: 0.884, families: 0.745, Fig. 3B-C), indicating that species share
most (but not strictly all) trophic interactions with their congenerics and confamiliars. On the other hand, taxonomic orders have a very low trophic similarity (< 0.05, Fig. 3D), indicating the presence of species that share very few or no trophic links with other species in their order. This result is primarily driven by plants, which generally show low trophic similarity even within families (median among plant orders: 0.023, families: 0.148, genera: 0.688).

Relative to the original food web, our trophic species web includes more richness (81.5%) and more unique links (75.1%) among herbivory interactions than among carnivory interactions (67.1% and 62.5% of the original richness and links). In other words, the herbivory interactions are more frequently functionally distinct than carnivory interactions in our food web and therefore provide the key source of niche differentiation (decreased trophic similarity) between species.

4) Discussion

Aboveground terrestrial ecosystems contain immense taxonomic diversity, especially among insect and plant species. Studying these interactions in aboveground terrestrial food webs (ATFWs) has been limited by three problems: taxonomic resolution of species, sensitivity to sampling effort, and conceptual inconsistency in the definition of feeding interactions. To overcome the problem of taxonomic resolution, we built our food web with taxa evenly resolved to species level and investigated the effect of resolution on unique trophic functions. To overcome the problem of sampling, we pooled public records and decades of biological research at a single site into an expert-approved, cumulative food web. This reduces the likelihood that we missed cryptic or rare species in our system and provides a more comprehensive accounting of the potential feeding interactions. Finally, to overcome the problem of conceptual inconsistencies, we extended a framework for studying ATFWs in which we included all types
of feeding interactions represented in a multiplex network as different types of feeding links by
the type of resource (animals or plant tissues) being consumed.

Our ATFW is the largest food web yet published and begins to shine light on the
remarkable richness of feeding interactions between plants and animals in temperate forests. Our
approach allowed us to evenly resolve the diets of both tetrapod and insect feeding guilds,
revealing a clear division between herbivory and carnivory interactions in our food web.
Carnivory interactions were numerous and non-specific, primarily sourced from interaction
records resolved to the order-level. Herbivory interactions were less numerous but more specific,
both taxonomically and in terms of different resource types. Most unique herbivory interactions
were sourced from records resolved to the genus- or species-level, and this specificity
corresponded to decreased trophic similarity (i.e., increased niche differentiation) among species
in our food web. The lack of taxonomic resolution in our carnivory records may simply reflect
limited knowledge about dietary preferences. However, our results are consistent with previous
food web research (primarily in aquatic systems) showing that larger species tend to have higher
trophic levels and more generalized diets in terms of number of resources (Brose et al. 2006).

Species in our food web shared many of their consumers and resources with other species
in their genus or family (i.e., were trophically similar). This is not surprising – close taxa share
most of their evolutionary histories and therefore many of the traits (e.g., body size, shape of
mouthparts, bio/phytochemistry) that govern their interactions (Ibanez et al. 2016). Insect
herbivores are particularly specialized to sense and tolerate the chemical and physical cues and
defenses of plant hosts with which they have coevolved (Loxdale & Harvey 2016). Indeed, insect
families in our food web not only share the same bird predators but also most frequently are
specialists on a single plant family, limiting their potential diet differences. A global study of diet
breadth among insect herbivores (Forister et al. 2015) showed even more frequent taxonomic specialization than we observed, but we also expect that our temperate site, inclusion of non-angiosperm host plants, and likely under-sampling of leaf-gallers leads to less specialization in our system than globally. Given this and the immense richness of insects, it is surprising that we resolved most insects in our food web as unique trophic species. We expect that further study would reveal nearly all taxonomic species to be trophically unique (Hutchinson & Pringle 2020). Niche theory suggests that distinct species (even closely related ones) should vary in their resource usage, though potentially in dimensions we did not record, including spatially, temporally, within our resource groupings, on resources we excluded, or outside of our system boundaries (i.e., in aquatic or belowground habitats). Regardless, the higher resolution herbivory records in our system differentiated insects by the species or tissue types of their plant resources. If we had focused primarily on tetrapod interactions as in previous work, coarse insect groupings would appear to be justified. Instead, our findings suggest that coarse taxonomic groupings (order- and family-level) are not an appropriate proxy for trophic species groupings in ATFWs.

a) Limitations

Though we sought to document all species and feeding interactions at UMBS, our approach still had methodological limits. As in all food webs, our system boundaries were largely arbitrary. Many of our species live or feed in aquatic or belowground habitats during a certain lifestage, and interactions coupling these different habitats can significantly affect ecosystem-level processes (e.g., through plant growth). Our lists excluded fungi, non-vascular plants, most non-insect invertebrates, and detritus, though these likely account for a substantial fraction of the species richness and feeding interactions in our system. Furthermore, regional lists suggest that there should be many more insect species (especially beetles, wasps, and true bugs) than have
been directly recorded at UMBS. We also are aware that many of our species are missing interactions because we were not able to find or verify species-specific data (due to taxonomic or other data limitations) or because records were too vague (e.g., “eating seeds” without further specificity). Together, this meant that 3.0% of our species had interactions only from order-level records, while 14.1% and 20.6% of our species (primarily insects) had interactions only from family- or genus-level records, respectively. In contrast, interactions with common, charismatic, or economically important species are likely more comprehensively documented in our food web. Lastly, just as we designated different plant tissues (see Conceptual Framework), feeding by or on different animal tissues (e.g., blood) or lifestages (e.g., eggs) likely has different implications for the relevant consumer and resource species, as well as anticipated biomass flows in the food web. Future work should investigate the effect of separating these interaction types in a multiplex network.

b) A path forward

We are in a new era of ecological big data, which provides an opportunity for a new era of food web ecology. Building databases of potential species interactions at research stations is a scalable process that could facilitate management decisions and stimulate ecological networks research. In this way, our approach complements recent advances in DNA barcoding, text-mining, and remote sensing that could ultimately lead to automated approaches to collecting, analyzing, and interpreting ecological networks using artificial intelligence and machine learning.

Nevertheless, the construction of our food web demonstrated that the role of ecological expertise (natural history and taxonomic) remains invaluable for at least three reasons: (1) identifying taxonomic conflicts, (2) clarifying nomenclature across ecological subdisciplines,
and (3) interpreting the scale of interactions. First, because taxonomic names are consistently updated, experts were critical for clarifying the relevant interaction records and the most up-to-date species nomenclature. Second, when interaction records listed species as “host,” “parasite,” “mutualist,” etc., experts could clarify whether feeding occurred and on what type of resource. Third, and most critically for the network presented here, experts assessed whether an interaction could feasibly occur in our system given species’ temporal and spatial co-occurrence and whether the taxonomic level of the record (genus, family, etc.) was representative of true potential interactions between species given their biologies (sensory capabilities, size, mobility, etc.). Some of these issues could be resolved with better data management (e.g., clearly defining the types of identifiers used to report species interactions), but local and organism-specific knowledge will likely remain an important part of building and using ecological networks (also see Poisot et al. 2023). Our work therefore highlights both the opportunity and the challenge to develop large collaborations among specialists, students, and the public as we seek to understand the complexity of natural ecosystems.

Finally, ecological networks have historically been published and analyzed as static structures, encapsulating the biases and practical limitations of their collection. As such, publication in online databases and consistent re-use in meta-analyses by network scientists can perpetuate errors (Hutchinson & Pringle 2020). Our ATFW represents a first attempt at documenting the immense complexity of aboveground terrestrial ecosystems. We advocate a shift in mindset when it comes to ecological networks, considering them as living datasets that can be revised and expanded as collaborations and knowledge grows. To that end, our database is publicly available (Hale 2023), and we are soliciting revisions, corrections, and additions that will allow its continual improvement.
Acknowledgements

We thank Adam Schubel, Jason Tallant, Aimee Classen, Knute Nadelhoffer, and other current and former University of Michigan BioStation (UMBS) staff for providing the original species lists and hosting the living version of the dataset. We are immensely grateful to Teresa Pegan, Eric Gulson, Simone Oliphant, Nate Sanders, Daniel Swanson, Anton Reznicek, Erika Tucker, and undergraduate research team Kathrine Northman, Taylor Brubaker, John Kelly, Lynnae Gilman, Matthew Palumbo, Lex Newman, and Stephan Verral for contributions to data acquisition and vetting. We acknowledge that the Indian Point Reserve (gifted to UMBS in 1987) includes lands of the Burt Lake Band of Ottawa and Chippewa Native Americans from which they were brutally and illegally evicted in 1900. The University of Michigan Undergraduate Research Opportunities Program (UROP) paid the undergraduate researchers for their time. This work was partially funded by NSF grants DEB-2129757 and DEB-2224915 to F.S.V.

Data Availability Statement

Species and interaction data are openly available via the Environmental Data Initiative at [doi forthcoming, also attached here for review]. Supplementary figures are available in the online version of this article.

References


Ávila-Thieme MI et al. 2021 Alteration of coastal productivity and artisanal fisheries interact to affect a marine food web. Scientific Reports, 11(1). (doi:10.1038/s41598-021-81392-4)


5) Figures

Figure 1 Visualization of a temperate forest food web in northern Michigan
Multiplex food web for the aboveground habitats in the University of Michigan Biological Station, grouped to taxonomic order. Each node is a taxonomic order, sized according to the number of species in each order and colored according to taxonomic group (plants in dark green, insects orange, birds in pink, mammals in yellow, amphibians in green, reptiles in blue, non-insect arthropods in brown). Each link is a feeding interaction between orders, with width scaled to the total number of feeding interactions by participating species, and colored according to the type of resource being fed upon (both live and scavenged animals in gray, leaves and stems in green, floral resources in pink, seeds and fruits in yellow, wood and bark in brown). Self-links indicate feeding among species within the order, including cannibalism. Nodes are ordered horizontally by their number of consumers, from most on the left to least on the right, and vertically by their trophic level (TL), from basal resources on the bottom (TL = 1) to carnivores at the top. Three carnivorous or parasitic plant orders and four basal animal orders with no resources (see Main Text) were assigned trophic levels between 1 and 2. The 86 orders here (labeled by their first three letters) represent 3,082 taxonomic species.
Figure 2 Distribution of diet breadths across resource types

(A-E) Histograms showing the frequency of consumers’ diet breadths in terms of number of resource species (in-degree) when feeding on (A) animals, both live and scavenged (N = 913), (B) leaves and stems (N = 1,738), (C) flowers, including nectar and pollen (N = 866), (D) seeds and fruit (N = 243), and (E) wood and bark (N = 102). (F) Histogram showing the counts of diet breadths for insect families in terms of number of plant resource families they use as hosts, pooled across plant tissue types (N = 264). Histograms include only consumer species (A-E) or (F) insect families that feed on the focal resource.
**Figure 3 Trophic similarity across taxonomic groupings**

(A-C) Histograms showing the frequency of the average minimum trophic similarities among species in each (A) genus (N = 635), (B) family (N = 296), and (C) order (N = 78). Trophic similarity is the similarity (quantified by the Jaccard index) between a pair of species’ consumers and resources. (D) Histogram showing the frequency of taxonomic species richness in each trophic species grouping. Trophic species are groups of taxonomic species that have the same sets of consumers and resources (that is, have a trophic similarity = 1). For visual clarity, the y-axis is scaled to log10(frequency) + 1.
Supplementary Figure 1 Taxonomic source and level of food web links

The taxonomic resolution of feeding links on (A) animals (N = 520,851 links) and (B), including leaf, flower, seed, and wood tissues (N = 59,898 links total). Bar graphs represent the fraction of unique links between taxonomic groups of consumer (left sub-panel) and resources (right sub-panel), colored by the coarsest level of taxonomic resolution in each group of feeding links. For example, panel (A) shows that among carnivory links, most are birds consuming insects, and furthermore that in most of these links one partner is most coarsely resolved to taxonomic order.

Consumer and resource link percentages both total to 100% because all feeding links are directed between a consumer and resource taxa. N = 586,601 total links are represented, including 526,206 carnivory links and 60,395 herbivory links.
Supplementary Figure 2 Correlations between consumer diet breadths feeding on different resources

Scatter plots showing the relationship between consumers’ diet breadths in terms of number of resource species when feeding on (A) animals and plant leaves ($r = 0.28$, $p = 2.0 \times 10^{-7}$, $N = 340$), (B) animals and plant seeds ($r = 0.15$, $p = 0.044$, $N = 202$), (C) plant leaves and seeds ($r = 0.42$, $p = 7.2 \times 10^{-9}$, $N = 178$), and (D) plant leaves and wood ($r = 0.23$, $p = 0.017$, $N = 98$). Each point is a species that feeds on both focal resources. Lines are least squares fits. Only relationships with significant Pearson’s correlations ($r$) are shown.
Supplementary Figure 3 Correlations between animal richnesses hosted by plants on different tissues

Scatter plots showing the relationship between the richness in terms of number of consumers species when feeding on plant species’ (A) leaves and flowers (r = 0.16, p = 3.4 x 10^{-5}, N = 653), (B) leaves and seeds (r = 0.35, p = 8.1 x 10^{-17}, N = 548), (C) leaves and wood (r = 0.61, p = 1.7 x 10^{-11}, N = 98), and (D) flowers and seeds (r = -0.11, p = 0.013, N = 514). Each point is a plant species with consumers that feed on both focal resources. Lines are least squares fits. Only relationships with significant Pearson’s correlations (r) are shown.