PHILOSOPHICAL
TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

A highly resolved network reveals the role of terrestrial herbivory in structuring aboveground food webs

Kayla R. S. Hale[1,4], John David Curlis[1], Giorgia G. Auteri[5], Sasha Bishop[1], Rowan L. K. French[6], Lance E. Jones[7], Kirby L. Mills[1,2], Brian G. Scholtens[8], Meagan Simons[1], Cody Thompson[1,3], Jordon Tourville[9,10] and Fernanda S. Valdovinos[11]

[1]Department of Ecology and Evolutionary Biology, [2]School for Environment and Sustainability, and [3]Museum of Zoology, University of Michigan, Ann Arbor, MI, USA
[4]Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada
[5]Department of Biology, Missouri State University, Springfield, MO, USA
[6]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada
[7]Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana IL, USA
[8]Department of Biology, College of Charleston, Charleston, SC, 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

KRSH, 0000-0002-2627-9623; JDC, 0000-0002-2974-7330; GGA, 0000-0002-5579-8078; SB, 0000-0002-0877-3951; RLKF, 0000-0002-9178-0035; LEJ, 0000-0001-6368-9720; KLM, 0000-0001-7693-9629; BGS, 0009-0007-8413-4856; CT, 0000-0002-6686-6056; JT, 0000-0003-4005-5914; FSV, 0000-0002-5270-5286

Comparative studies suggest remarkable similarities among food webs across habitats, including systematic changes in their structure with diversity and complexity (scale-dependence). However, historic aboveground terrestrial food webs (ATFWs) have coarsely grouped plants and insects such that these webs are generally small, and herbivory is disproportionately under-represented compared to vertebrate predator–prey interactions. Furthermore, terrestrial herbivory is thought to be structured by unique processes compared to size-structured feeding in other systems. Here, we present the richest ATFW to date, including approximately 580 000 feeding links among approximately 3800 taxonomic species, sourced from approximately 27 000 expert-vetted interaction records annotated as feeding upon one of six different resource types: leaves, flowers, seeds, wood, prey and carrion. By comparison to historical ATFWs and null ecological hypotheses, we show that our temperate forest web displays a potentially unique structure characterized by two properties: (i) a large fraction of carnivory interactions dominated by a small number of hyper-generalist, opportunistic bird and bat predators; and (ii) a smaller fraction of herbivory interactions dominated by a hyper-rich community of insects with variably sized but highly specific diets. We attribute our findings to the large-scale, even resolution of vertebrate, insect and plant guilds in our food web.

This article is part of the theme issue ‘Connected interactions: enriching food web research by spatial and social interactions’.

Cite this article: Hale KRS *et al*. 2024 A highly resolved network reveals the role of terrestrial herbivory in structuring aboveground food webs. *Phil. Trans. R. Soc. B* **379**: 20230180.
https://doi.org/10.1098/rstb.2023.0180

Received: 3 October 2023
Accepted: 5 May 2024

One contribution of 19 to a theme issue ‘Connected interactions: enriching food web research by spatial and social interactions’.

Subject Areas:
ecology, environmental science, ecosystem

Keywords:
temperate forest ecosystem, multiplex ecological network, plant–insect interactions, niche model, scale-dependence, trophic species

Author for correspondence:
Kayla R. S. Hale
e-mail: kaylasal@umich.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7318729.

1. Introduction

Ecosystems contain immense biological complexity. Food webs represent part of this complexity by documenting the feeding interactions (links) between taxa (nodes). Comparative studies of food webs across habitats have revealed robust and non-random patterns suggestive of an underlying

THE ROYAL SOCIETY
PUBLISHING

architecture of life [1–7]. For example, the structure of food webs changes systematically with diversity and complexity (termed 'scale-dependence') but maintains hierarchy, demonstrated in aquatic and belowground food webs through modular, size-structured pathways of larger consumers feeding on smaller resources [8–13]. Aboveground terrestrial food webs (ATFWs) also exhibit size-structure in predator–prey interactions, but different mechanisms (e.g. chemical composition, trait matching) probably underlie the specialized feeding of insect herbivores on often-larger terrestrial plant resources [14]. However, there are few published ATFWs and even fewer that include high-resolution data for both plant–herbivore and predator–prey interactions across broad taxonomic groups (table 1). Therefore, whether and how the structure of ATFWs may fundamentally differ from those of other habitats remains unclear [32]. As a step towards answering this question, we construct the most extensive ATFW to date and study the mechanisms by which its increased taxonomic and trophic resolution lead to a unique structure relative to the scale-dependent pattern observed in previous webs.

Constructing food webs is fraught with methodological difficulties [33,34]. Observations of species and interactions depend on the boundaries of the system, the specific spatial (vertical versus horizontal transects, microhabitats) and temporal (seasonal, diurnal and duration) scales of sampling, as well as the taxonomic expertise of the investigators (including ability to detect and identify both consumer and resource species). Many organisms regularly cross ecosystem boundaries as part of their life cycles; for example some insect species spend larval stages underground or underwater, then move to aboveground habitats after maturation, after which they may migrate to a completely different region for breeding [35]. Species also exhibit adaptive foraging and defensive behaviours, effectively 'rewiring' interactions in response to changing biotic and abiotic conditions [14,36,37]. One approach to these problems is to construct an expert-vetted 'cumulative' or 'meta' food web that pools all species and interactions recorded across time and/or similar habitats [38]. This reduces the likelihood of missing cryptic or rare species and provides a more comprehensive accounting of all potential feeding interactions in the system. Additionally, as human activities alter species' distributions and habitats, 'rare' and novel interactions are increasing in frequency [36], making cumulative webs even more important.

Even cumulative food webs rely heavily on expertise and long-term and/or regional sampling. Perhaps for this reason, previous high-quality ATFWs have tended to focus either on taxonomic breadth or depth. Webs with taxonomic breadth (table 1, marked with asterisks) tend to resolve vertebrates most highly, while aggregating invertebrates and plants into coarse taxonomic or functional groups (e.g. into insect orders or plant tissue categories). This 'lumping' strategy *sensu* Briand [39] seeks to describe broad system-level behaviour but is largely a result of the technical difficulties associated with documenting and representing the pure volume of plant–insect associations [27]. Nevertheless, classic breadth webs—Coachella Valley [20], St Martin Island [21] and El Verde Rainforest [27]—have proved highly influential and remain perhaps our best description of ATFWs because they used cumulative approaches with known species lists from long-term fields sites. Indeed, these webs, with the Little Rock Lake web of Martinez [40], contributed to overturning 'empirical generalizations' (such as scale-invariance, low omnivory, etc.) derived from a catalogue of less-resolved webs [15,41,42].

By contrast, webs with greater depth of resolution tend to have narrower scope (table 1). These webs generally focus only on a single taxonomic group (e.g. only tetrapods) or on a single energetic pathway (e.g. 'source' or 'sink' webs) [38]. In the same vein, the explosion of ecological networks research in the last two decades has tended to focus on highly specific single interaction types such as frugivory or scavenging, demonstrating the unique structure and importance of these subnetworks for ecosystem dynamics and function [43–46]. However, different subnetworks are rarely recorded in the same system, and as such, it is unknown how they may connect with each other or to their broader food web [43]. An exception is the few 'multiplex' networks that report high-resolution interactions of different types (i.e. feeding on different resources or with different interaction outcomes) among non-disjoint sets of species [47,48]. These studies bring together interactions that otherwise rarely co-occur in food webs [44–46], especially mutualisms, such as pollination or seed dispersal, which can have a feeding component via consumption of nectar and pollen or seeds and fruits, with other forms of 'antagonistic' herbivory such as phloem-feeding by aphids [28,29,49].

Comparative studies attempt to standardize these diverse approaches to constructing food webs in three ways [38,50]. First, they aggregate empirical webs to 'trophic species' webs, where taxa with the same set of consumers and resources are grouped into the same node [41]. This reduces methodological biases within and between webs by retaining only functionally distinct units with unique trophic niches [2]. Second, they compare the properties of trophic species webs to null expectations provided by the well-known 'niche model' of Williams & Martinez [8], which embodies specific ecological hypotheses for the mechanisms structuring food webs [51]. This approach provides scale-dependent expectations for food web properties (i.e. given their richness and complexity), and deviations from null expectations (sometimes called 'errors') can be interpreted as rejecting the underlying hypotheses. However, errors are also scale-dependent, meaning that the properties of empirical webs increasingly deviate from niche model expectations with increasing richness [50]. Therefore, third, comparative studies extrapolate from scale-dependent errors to assess whether a focal web exhibits unique properties compared to other webs, given its scale [30,50].

In this study, we used more than a century of research at a biological research station to build the Michigan Temperate Forest (MTF) food web, the richest ATFW to date. We used a cumulative approach, incorporating public records and occurrence data, supplemented and vetted by experts for local plausibility given species' traits and behaviours. This resulted in approximately 580 000 feeding links among approximately 3800 taxonomic species, represented in a multiplex network according to feeding on different resource types ('prey', 'carrion', 'leaves', 'flowers', 'seeds', or 'wood'). Using comparative food web methods, we (i) characterized the properties of the MTF, (ii) studied whether the increased taxonomic and trophic resolution leads to unique structure, given its scale, compared to previous ATFWs, and, if so, (iii) identified potential mechanisms underlying the structure of more or less-resolved webs.

Table 1. Properties of ATFWs. (A collection of classic and more recent food webs used for comparison to our new Michigan Temperate Forest web for the University of Michigan Biological Station (UMBS). We include classic webs (published prior to 2000) traditionally considered 'highly resolved'; this excludes the many historical webs in the ECoWeB catalogue of low richness and variable resolution [15]. We exclude modern webs in 'container habitats' (e.g. under a log, in a tree hole) as well as newer parasitoid–host webs [6,16]. The High Arctic web was assembled from data in Appendix S2 of Wirta *et al.* [17], originally from Roslin *et al.* and Rasmussen *et al.* [18,19]. Column definitions: web, traditional name and reference; richness, number of 'species' in original publication (but see extent); S, number of trophic species; C, directed connectance; L, number of links; B, fraction of basal trophic species (with no consumers); taxonomic groups, taxa and resolution in the food web; basal nodes, types of trophic species or functional groups at the base of the food web; types of consumers, types of feeding interactions included in the food web; extent, notes on the space and time of food web construction. Unless otherwise noted, webs are cumulative metawebs, built from records pooled across time (including published literature) and similar habitats (usually contiguous field sites). Other definitions: 'lumped by taxa', grouped to order or family except potentially for key species; source-web, web recording the food chain(s) up from a set of resources; sink-web, web recording the food chain(s) down from a set of consumers; 'breadth' webs (marked with asterisk), webs including multiple taxonomic groups and energy pathways but with lower resolution; 'depth' webs, higher resolution webs missing key structural components, including some non-traditional food webs like the Pocock Farm multiplex network. Abbreviations: spp., species; incl., including.)

web	rich-ness	S	C	L	B	taxonomic groups	basal nodes	types of consumers	extent
Coachella Valley * [20]	30	29	0.312	262	0.103	vertebrates, arthropods, other invertebrates, microbes lumped by taxa, size and trophic role	plants and plant products; detritus; carrion	predators, parasitoids, detritivores, scavengers	lumps 138 vertebrate, 55 arach-nid, 174 vascular plant, and approx. 2–3000 insect spp.
St Martin Island* [21]	44	42	0.116	205	0.143	vertebrates resolved to spp.; fungi, arthropods, other invertebrates, lumped by taxa and lifestage	fruits and seeds; nectar and floral; leaves; roots; wood; detritus	predators, parasites, detritivores, herbivores (incl. nectar and fruit feeders)	'*Anolis*-centred'
UK Grassland [22]	87	67	0.027	119	0.149	grasses and wasps resolved to spp.	grass spp.	herbivores, parasitoids	source-web; direct observation
Scotch Broom [23]	154	83	0.032	220	0.012	broom, insects, fungi, bacteria, and arachnids resolved to spp.	broom (*Cytisus scoparius*)	predators, parasites, parasitoids, herbivores (incl. seed, sap and root feeders)	source-web; single field site
Serengeti – de Visser* [24]	350	85	0.075	545	0.071	vertebrates resolved to spp.; invertebrates incl. insects, gastropods, etc. lumped by taxa and size	detritus; plant juices; fruits and nectar; grains and seeds; grass and herbs; trees and shrubs	predators, scavengers, herbivores (incl. seed, nectar and fruit feeders)	lumps > 322 vertebrate spp., 28 invertebrate orders, 7 resource groups
Serengeti – Baskerville [25]	161	105	0.042	463	0.695	plants and mammals resolved to spp.	plant spp.	predators (carnivores), herbivores	
Shortgrass Prairie * [26]	133	105	0.034	379	0.657	plants and vertebrates resolved to spp.; arthropods lumped by taxa	plant spp.; un-identified seeds	predators, herbivores (incl. seed feeders)	sampling over 2 years at two sites
El Verde Rain-forest* [27]	156	155	0.063	1509	0.181	vertebrates and some invertebrates incl. insects, gastropods, etc. resolved to spp.; other invertebrates, lumped by taxa and lifestage	resource categories, e.g. leaves; dead wood; roots; bacteria; algae; slime molds; basal prey taxa	predators, parasites, scavengers, detritivores, herbivores (incl. leaf, seed, fruit, nectar, pollen, flower, wood, lichen, root and sap feeders), fungivores	lumps > 2601 known heterotroph and >214 autotroph plant spp.

(*Continued.*)

Table 1. (*Continued.*)

web	rich-ness	S	C	L	B	taxonomic groups	basal nodes	types of consumers	extent
High Arctic [17–19]	214	161	0.025	646	0.348	plants, spiders, insects, other invertebrates and birds resolved to spp.	plant and basal prey spp.	predators, parasitoids, herbivores, pollinators (nectar and pollen feeders)	Zackenberg Valley; source-webs coupled to sink-web following [28]
Norwood Farm [29]	560	327	0.011	1199	0.297	plants, insects, fleas, ticks, birds, rodents resolved to spp.	plant spp.	parasitoids, herbivores, flower visitors (nectar and pollen feeders), seed feeders, ectoparasites	source-web; recorded as multiplex network; sampled over 2 years at a 125 ha farm
Messel Forest* [30]	700	646	0.014	5758	0.307	plants, microbes, fungi, arthropods, vertebrates, other invertebrates; >50% resolved to genus or spp.	detritus, carrion, plant spp.	predators, herbivores (incl. seed, wood, root, pollen and leaf feeders)	taxa deposited in a lake basin during the Eocene
European Tetrapods [31]	1152	1054	0.043	47510	0.763	vertebrates resolved to spp.	basal prey spp.	predators (carnivores)	European continent; not all spp. co-occur locally; 17 diet categories provided as traits
Michigan Temperate Forest*	3802	2597	0.054	366385	0.245	vertebrates, plants, insects and other invertebrates resolved to spp.	plant and basal prey spp.	predators, parasitoids, ectoparasites, scavengers, herbivores (incl. leaf; seed and fruit; nectar and pollen; and wood feeders)	UMBS field station; recorded as multiplex network

2. Methods

(a) Site, species list and feeding records

The University of Michigan Biological Station (UMBS) was established in 1909 on approximately 10 000 acres of logged and burned land in northern lower Michigan, USA (45°35.5′ N, 84°43′ W) and represents a strongly seasonal system with historically cold, snowy winters and hot, humid summers. The site has, in recent years, been restored to predominantly dry-mesic, northern hardwood forests with patches of wooded wetlands (hardwood conifer swamp) [52]. A full description of the UMBS site and extended methods are available in the electronic supplementary material [53].

Briefly, experts (generally, the authors) vetted and approved species lists of mammals, amphibians, reptiles, vascular plants, birds, insects and non-insect arthropods from UMBS records. These were accumulated from resident biologists' personal observations, student projects, museum and herbarium specimens and semi-regular BioBlitz events, in which teams of biologists roamed the site and identified as many organisms as possible. Hereafter, we refer to all approved taxa as 'species', though a small fraction (4.5%) are genera.

The same experts vetted and annotated a list of potential feeding interactions, sourced from region-specific field guides and online databases [54]. 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). We included all records of direct interactions among species in our system with a bioenergetic flow (i.e. one species consuming another), regardless of lifestage or potential ecological effects (i.e. whether potentially 'mutualistic' or 'antagonistic' [14]). Experts approved recorded interactions between species as plausible if the species co-occur (with respect to phenology, activity patterns and microhabitat usage) and have no trait incompatibilities (with respect to acquisition, ingestion and assimilation). If a partner in a potential interaction was recorded at a coarser taxonomic level than species, the record was approved only if these conditions could also plausibly hold for all local species in that taxonomic unit.

Finally, experts categorized records by their focal resource type as animal tissues, either: (i) live tissues and as prey or (ii) scavenged as carrion, carcasses or other decaying animal remains, or as plant tissues, grouped as (iii) leaves and stems, (iv) flowers, nectar, pollen, etc., (v) seeds, fruits, etc., or (vi) wood and bark. Hereafter, we refer to these resource types simply as 'prey', 'carrion', 'leaves', 'flowers', 'seeds' and 'wood', respectively.

(b) Network representation

To translate our list of feeding records into a food web, we began with a 'multiplex network' approach (figure 1*a*) in which feeding on different resource types is represented by different types of links between the same set of nodes (taxonomic species). This allowed us to distinguish between the niches of animals feeding on different resources while also accounting for the fact that such resources are coupled together 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 l between nodes i, j if i consumes tissue type l of j or tissue type l 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. We characterized the complexity of the multiplex network by counting the number of links (L_l), consumer species (A_l) and resource species (P_l) involved in feeding of type l, and the fraction of the maximum possible links that were realized as either bipartite connectance = $L_l/(A_lP_l)$ for feeding on plant tissues (leaves, flowers, seeds or wood) or unipartite connectance = $L_l/(A_l + P_l)^2$ for feeding on animal tissues (prey or carrion).

As a direct comparison to previously published food webs (table 1), we used the conventional food web approach, where a binary link occurs between nodes i, j if i consumes j in the multiplex network (i.e. consumes any resource type of species j). Following convention, we aggregated all webs into 'trophic species' versions, wherein taxa with the same sets of consumer and resource species are grouped together (figure 1b(iii)).

To study the contribution of resolving feeding on different resource types to food web structure, we aggregated our multiplex network to a 'trophic species multiplex network' (figure 1b(i–ii)), wherein all taxa with the same set of consumers and resources, both in terms of taxonomic species and resource types, are grouped into a single node. The difference in resulting richness between this trophic species multiplex network (figure 1b(ii)) and that of our trophic species food web (figure 1b(iii)) indicates how many taxonomic species' trophic niches are differentiated only by feeding on specific resource types (e.g. on the leaves versus the flowers or the live prey versus the carrion of the same resource species).

Hereafter, we discuss the structure of a 'food web' as the trophic-species food web with richness denoted as S, number of links denoted as L and directed connectance (the fraction of observed links out of the maximum possible, L/S^2) denoted as C [40].

(c) Food web structure

To provide null expectations for the scale-dependent structure of food webs, we used the niche model of Williams & Martinez [8] to simulate $n = 1000$ matching webs using the S and C of our food web and each of the previous webs (table 1). Traditionally, the model assumes that all nodes are unique trophic species and webs that include nodes with the same consumer and resource set (i.e. duplicate trophic species) are rejected. For webs with low connectance, we relaxed that assumption and allowed niche model webs to be seeded with a slightly higher initial richness as long as (following another trophic species aggregation) S and C matched the empirical trophic-species web.

We calculated a suite of properties to characterize the composition, hierarchy and degree distribution of the empirical food webs [8,50,55–57]. See the electronic supplementary material, table S1 for a full list of properties and definitions. For each structural property, we assessed the significance of deviations from null expectations using normalized model errors (NMEs) [55,58]. NMEs are calculated as the difference between the median model and empirical values normalized by either the difference between the median model value and the 97.5 percentile of the model distribution if the empirical value is greater than the model median, or, if the empirical value is less than the model median, by the difference between the 2.5 percentile of the model distribution and the model median. Values greater than 1, or 1 or less indicate that the empirical value is significantly higher or lower, respectively, than the null expectation at the 95% confidence level. For the purposes of discussion, we follow previous works to summarize these as a composite *mean* |NME|, though the individual properties are not independent (see the electronic supplementary material, Supplementary Methods) [50].

To characterize species composition, we calculated the fraction of trophic species in each of the following categories: basal (*B*): with consumers but no resources; intermediate (*I*): with both consumers and resources; top (*T*): with resources but no consumers; *Herbiv* (TL2): eat only basal species (are strict herbivores, i.e. trophic level [TL] = 2); *Carniv*: eat only other consumers (strict carnivores); *Omniv*: eat both basal and consumer species (omnivores); *Cannib*: eat members of their own species (cannibals).

To characterize link composition, we calculated *HerbLink*, the fraction of total feeding links that are herbivorous (i.e. are on basal resources), and *TL2Link*, the fraction of feeding links from TL2 herbivores.

To characterize hierarchy, we calculated *meanTL* and *maxTL*, the mean and maximum short-weighted trophic level of consumers [57], and *meanTLTop*, the mean trophic level of top consumers.

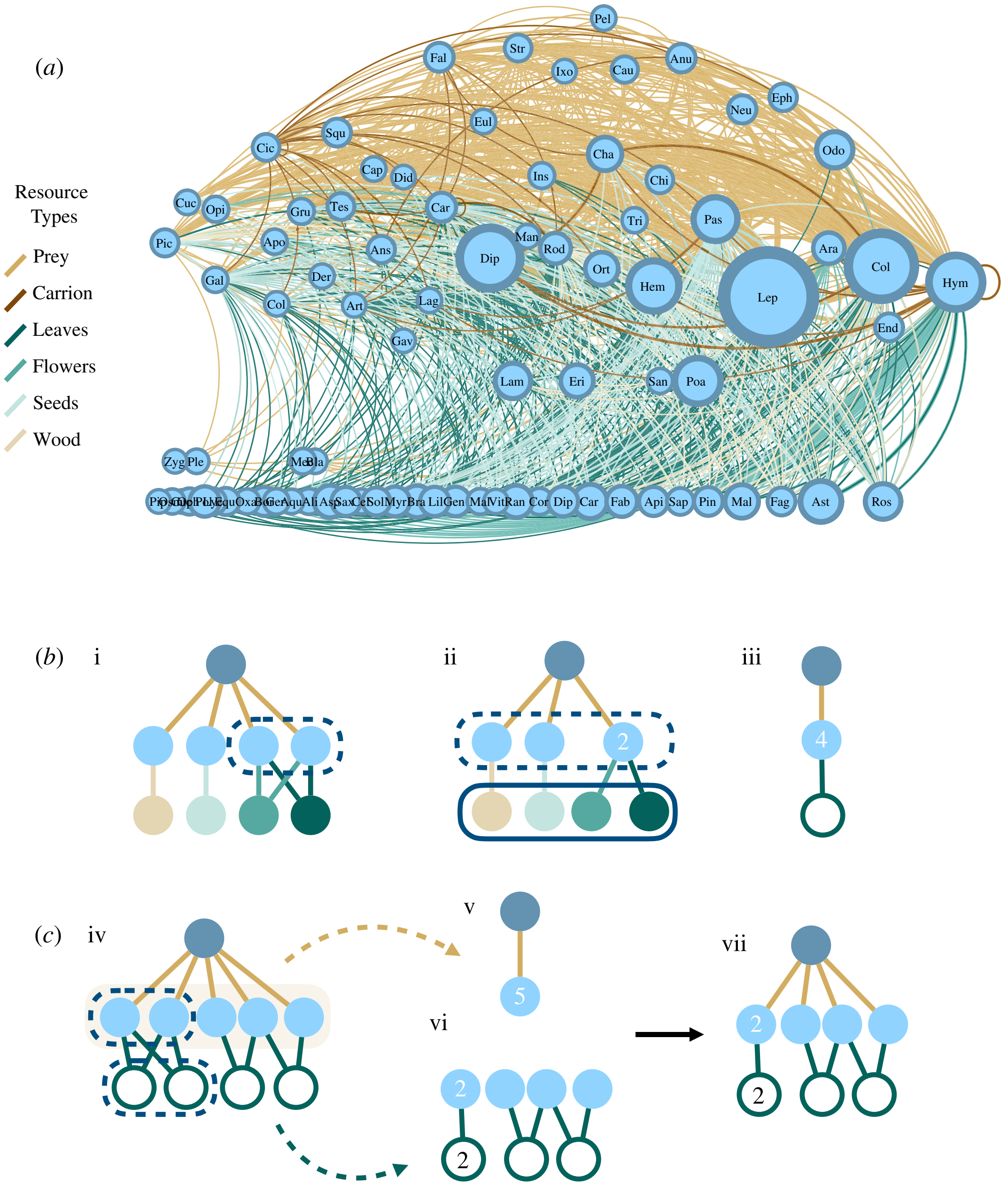

Figure 1. Representations of the MTF food web. (*a*) Simplified visualization of the multiplex food web. Each node represents a taxonomic order (labelled by its first three letters), with width scaled to the number of local species, ranging from one (e.g. order Gaviiformes, represented only by the common loon) to >1000 (order Lepidoptera). Each link is a feeding interaction between orders, with width scaled to the total number of feeding interactions between species in each order. Links are coloured by the resource type consumed as follows: tan, live prey and animal tissues; brown, scavenged carrion; dark green, leaves and stems; green, floral resources; light green, seeds and fruits; and beige, wood and bark. Self-links indicate feeding among species within the order, including cannibalism. Nodes are ordered horizontally by their number of consumers (in-degree), increasing from left to right, and vertically by increasing trophic level (TL) from basal resources on the bottom (TL = 1) to carnivores at the top. Three carnivorous/parasitic plant orders were assigned TL = 1.75 and four basal animal orders were assigned TL = 1.25 for visualization. The 85 orders shown here represent 3082 taxonomic species. (*b*) Illustrated effect of aggregating the multiplex network into 'trophic species'—species with same set of consumers and resources. Numbers indicate how many nodes were aggregated. Here, one generalist predator feeds on four herbivores of the same plant species, with its different tissue types (wood, seeds, flowers and leaves) illustrated as different nodes for clarity (colours follow panel (*a*)). There are six taxonomic multiplex species total in this example. The 'multiplex network' representation (i) can differentiate among the trophic niches of herbivores feeding on different plant tissues. Aggregating into a 'trophic-species multiplex network' (ii) groups the two herbivores that feed on the same plant tissues into one trophic species, resulting in five multiplex trophic species total. Further aggregating all four herbivores into one trophic species feeding on the plant species (white node), results in three trophic species total. This is the traditional representation of 'trophic-species food webs' (iii), where only taxonomic identity differentiates species' trophic niches. (*c*) Illustration of the disproportionate effect of herbivory on trophic species resolution in the MTF web. The generality of predators (especially birds and bats) in the MTF web (iv) means that prey species (especially insect herbivores) tend to be aggregated into fewer trophic species when considering carnivory interactions alone (v). The specificity of herbivores in the MTF web results in more trophic species resolved when considering herbivory interactions alone (vi). This leads to a pattern of numerous generalized carnivory interactions and fewer but more specific herbivory interactions among the more numerous plant and herbivore species (vii).

Degree describes the number of resources (in-degree) and consumers (out-degree) of a species. To characterize degree distribution, we calculated *meanGen*, the mean in-degree of consumers (i.e. their 'generality'), *GenSD*, the normalized variability of generality, *meanVul*, the mean out-degree of resources (i.e. their 'vulnerability'), and *VulSD*, the normalized variability of vulnerability. We also calculated these properties for specific subgroups of species to characterize their respective contribution to deviations observed from null expectations. Finally, we used two-sample Kolmogorov–Smirnov tests to directly test the null hypothesis that the observed and expected degree distributions were sampled from the same underlying distribution [58]. Full results are reported in the electronic supplementary material, table S2.

We observed that the niche model tended to generate webs with low B at high S, which may drive substantial NMEs simply owing to underlying differences in species composition. We directly tested whether this may be the case by simulating $n = 1000$ niche model webs with matching S, C, B values to our empirical food web, called the 'basal-matched' treatment. We seeded the niche model with $S = 2896$, $C = 0.057$, and set the 927 lowest niche-value species to have feeding ranges of zero (thus forcing them to be basal).

All data cleaning and network analyses were performed in MATLAB R2021b [59].

3. Results

(a) Species list and feeding records

Our final species list includes 3802 local species, representing 2073 genera in 451 families of 85 orders (figure 1*a*). Insects (2669 spp.) and vascular plants (781 spp.) numerically dominate the community, accounting for approximately 90% of the taxa, compared to vertebrates (313 spp.) and non-insect arthropods (39 spp.). The richest orders are insects, especially the Lepidoptera (butterflies and moths, 1168 spp.), Coleoptera (beetles, 512 spp.), Diptera (flies, 390 spp.), Hymenoptera (bees, wasps and ants, 265 spp.) and Hemiptera (true bugs, 211 spp.). Worldwide, there are more than twice as many named species of Coleoptera as Lepidoptera [60]; therefore, Lepidopterans are probably substantially over-represented in our list. Vascular plants represent the most taxonomically diverse group on our list, with 38 orders; however, most of these species belong either to Poales (grasses, sedges and rushes, 149 spp.) or Asterales (composite flowers, 96 spp.). Vertebrate species include birds (226 spp., including 127 passeriform birds), mammals (52 spp., including seven bats), amphibians (18 spp.) and reptiles (17 spp.). Finally, non-insect arthropods primarily include spiders and mites, but overall, this group is significantly under-represented in our list, both in terms of richness and taxonomic diversity.

In summary, we recorded 2541 species of consumers (including four carnivorous or parasitic plants) and 3782 species of resources. Approximately, two-thirds of species (69.7%) included in the records are resolved to taxonomic species level, while 13.7%, 13.1% and 3.0% of species (primarily insects) have records from genus-, family- or order-level records at best, respectively. We have no feeding records for 19 plants (0.5% of local species), including most of the Lycopodiales (clubmosses, five of six spp.) and the Polypodiales (ferns, 11 of 19 spp.), which represent 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, two spp.), Plecoptera (stoneflies, two spp.), Mecoptera (scorpionflies, four spp.) and Zygentoma (silverfish, one spp.). Some of these do not feed in aboveground terrestrial habits (or at all) during a certain life stage 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 expertise or the available natural history information for these species.

We constructed our multiplex network from 26 728 approved records of local taxa feeding on prey, carcass, leaf, flower, seed or wood resources, totalling 588 416 unique feeding interactions (links) between local species. These links primarily consist of feeding on prey (89.2%), especially insect prey, with the remaining links consisting of scavenging carcasses (0.54%) or feeding on plant leaves (6.7%), 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 interactions are realized among the 3023 predator and prey species in our carnivory subnetwork, with similarly low connectances for scavenging (0.2% among 1357 spp.) and the different types of herbivory (leaves: 3.1% among 2458 spp., flowers: 2.2% among 1533 spp., seeds: 5.1% among 796 spp., wood: 9.8% among 200 spp.).

Carnivory is the most numerically dominant interaction in our food web, but only approximately one-third (35.9%) of consumers feed on prey. In fact, the 120 most generalist species in our food web (4.7% of consumers) contribute over half (51.5%) of the unique carnivory links in our network, sourced from only 950 (3.6%) records of focal birds and bats thought to feed opportunistically upon entire insect orders. 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. In comparison, feeding between vertebrates accounts for only 1.3% of carnivory links.

Herbivory interactions are less numerous than carnivory, but most consumers (85.4%) 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 floral resources. Over half of herbivory interactions (52.4%) 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. 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 (39.6%) feed upon more than one type of resource. Around half of these feed on leaves and flowers (18.9% of consumers, primarily lepidopterans and coleopterans). A smaller fraction (11.6%) feed on more than two types of resources, but these represent a more diverse set of insects, mammals and birds feeding on prey, leaves and flowers or seeds, or, less frequently, leaves, flowers and wood. Ants (Formicidae, 48 spp.) uniquely feed on prey, leaves, flowers and carrion. Among consumers feeding on multiple resource types, we observed significant positive correlations between diet breadths when feeding on prey and plant leaves (Pearson correlation: $r = 0.26$, $p = 2.7 \times 10^{-6}$, $n = 319$), prey and seeds ($r = 0.14$, $p = 0.044$, $n = 202$), leaves and seeds ($r = 0.42$, $p = 7.2 \times 10^{-9}$, $n = 178$), and leaves and wood ($r = 0.24$, $p = 0.017$, $n = 98$; electronic supplementary material, S1). In other words, generalists on one resource type also tend to be generalists on others. 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 (81 species in 13 orders) sustaining feeding on all four recorded tissue types (electronic supplementary material, fignure S2).

(b) Trophic species composition

Our final food web consists of $S = 2595$ trophic species and $L = 365\,951$ links. An additional 29 trophic species and 8462 links would be distinguished by feeding only on different types of plant tissues (figure 1*b*). Nearly, all (92.3%, 2394) of the trophic species nodes correspond to taxonomic species, including all vertebrates, nearly all non-insect arthropods (94.9%), over two-thirds of plants (70.2%), and over half of insects (55.0%) in our species list. The remaining trophic species are each composed of taxonomic species from a single order. This is probably because entire insect orders, the primary resources in the food web, share the same opportunistic/generalist predators and are therefore only distinguished by their diets (figure 1*c*). Over half of these taxonomic species (54.1%, 761) are represented in only 27 trophic species groups; these comprise most (83.1%) of the 485 animal species without diet information (i.e. the species most likely lacking resolution).

Feeding on leaves contributed to distinguishing the most trophic species in our food web (924 trophic species), followed by feeding on prey, flowers, seeds, wood and then carrion (417, 322, 13, 5 and 0 respective additional trophic species relative to versions of the food web built with feeding only on the other resource types). Therefore, herbivory interactions, and particularly leaf herbivory, provide the key component for distinguishing species' trophic niches in our system.

(c) Scale-dependence of food web structure

The structure of our food web deviates significantly from the null expectation provided by the niche model in almost every metric. Beginning with trophic species composition (figure 2*a–e*), approximately one-tquarter of trophic species in our web are basal ($B_{\mathrm{obs}} = 0.245$) and the remaining three-quarters are intermediate ($I_{\mathrm{obs}} = 0.755$). Nearly half of trophic species are herbivores that eat only basal trophic species (i.e. are trophic level 2, $TL2TL2_{\mathrm{obs}} = 0.483$), and nearly one-quarter are omnivores that eat both basal and consumer trophic species ($OmnivOmniv_{\mathrm{obs}} = 0.235$), while <5% are carnivores that feed strictly on other consumers ($CarnivCarniv_{\mathrm{obs}} = 0.038$). This is in stark contrast to the null expectation that <2% of species are basal ($B_{\mathrm{null}} = 0.018$), and the remaining are intermediate ($I_{\mathrm{null}} = 0.977$) with <1% TL2 herbivores ($TL2TL2_{\mathrm{null}} = 0.006$). Moreover, the null model predicts one-third of species as carnivores ($Carniv_{\mathrm{null}} = 0.315$), and the remaining two-thirds as omnivores ($Omniv_{\mathrm{null}} = 0.662$). However, our observation of almost no top consumers was not significantly different from the null expectation of the niche model ($T_{\mathrm{obs}} = 3.85 \times 10^{-4}$, $T_{\mathrm{null}} = 0.006$), and the fraction of cannibals was also similar to the null expectation ($Cannib_{\mathrm{obs}} = 0.0312$, $Cannib_{\mathrm{null}} = 0.058$).

Our web contains a higher fraction of herbivorous links ($HerbLink_{\mathrm{obs}} = 0.135$, $HerbLink_{\mathrm{null}} = 0.018$)—especially a higher fraction by TL2 herbivores—than expected ($TL2Link_{\mathrm{obs}} = 0.071$, $TL2Link_{\mathrm{null}} = 4.87 \times 10^{-5}$; figure 2*f*) and tends to be shorter in terms of mean and max trophic level ($meanTL_{\mathrm{obs}} = 2.22$, $meanTL_{\mathrm{null}} = 4.37$; $maxTL_{\mathrm{obs}} = 4.62$, $maxTL_{\mathrm{null}} = 6.17$). Additionally, the mean trophic level of top species in our web is two (i.e. herbivores), significantly lower than expected ($meanTLTop_{\mathrm{obs}} = 2.0$, $meanTLTop_{\mathrm{null}} = 4.81$). This may be a limitation of the dataset rather than a true signal; our temperate forest system does not have megaherbivores, and we expect most insect herbivores to experience parasitoidy or natural enemies in aboveground terrestrial habitats.

Our web exhibits a higher average generality among consumers ($meanGen_{\mathrm{obs}} = 186.7$, $meanGen_{\mathrm{null}} = 143.5$; figure 2*h*) and a greater variability in generality overall ($GenSD_{\mathrm{obs}} = 2.82$, $GenSD_{\mathrm{null}} = 1.19$; figure 2*j*). Indeed, the distributions of generality between our web and the simulated niche models are significantly different, both visually (figure 3*b*) and statistically according to Kolmogorov–Smirnov tests (see the electronic supplementary material, table S2). These deviations can be attributed to two properties: first, a small core of hyper-generalist omnivores ($meanGenOmniv_{\mathrm{obs}} = 481.3$, $meanGenOmniv_{\mathrm{null}} = 144.4$; $GenSDOmniv_{\mathrm{obs}} = 1.35$, $GenSDOmniv_{\mathrm{null}} = 1.17$); second, a long tail of TL2 herbivores with greater generality and variability of generality than expected by the niche model ($meanGenTL2_{\mathrm{obs}} = 20.8$, $meanGenTL2_{\mathrm{null}} = 1.42$, figure 2*i*; $GenSDTL2_{\mathrm{obs}} = 1.30$, $GenSDTL2_{\mathrm{null}} = 0.40$; figure 2*k*). However, the TL2 herbivores produced by the niche model are far more specialized than the average consumers in our web (compare $meanGenTL2_{\mathrm{null}}$ to $meanGen_{\mathrm{obs}}$).

Our web does not exhibit significant deviations in average vulnerability ($meanVul_{\mathrm{obs}} = 141.1$, $meanVul_{\mathrm{null}} = 141.1$) or variability of vulnerability ($VulSD_{\mathrm{obs}} = 0.583$, $VulSD_{\mathrm{null}} = 0.583$; figure 2*l*) compared to the null expectation, despite significantly different degree distributions (figure 3*c*). However, our web specifically has a lower vulnerability of basal trophic species than expected ($meanVulBasal_{\mathrm{obs}} = 77.6$, $meanVulBasal_{\mathrm{null}} = 139.7$; figure 2*m*) and a higher variability of such vulnerability ($VulSDBasal_{\mathrm{obs}} = 1.052$, $VulSDBasal_{\mathrm{null}} = 0.580$; figure 2*n*).

To investigate the extent to which deviations in the structure of our food web from both previous webs and null expectations can be attributed simply to differences in species composition, we also compared our web to an ensemble of niche models with a matched fraction of basal trophic species, *B* (figure 2*a*, hollow triangle). This correction automatically fixes *I* (figure 2*b*) and *meanGen* (figure 2*h*) by restricting the available links to the correct number of consumers (figure 3*b*), thereby improving *Omniv*, *TL2* and *TL2Links* (figure 2*d–g, j*). However, this causes a compensatory overestimation of total herbivorous links, TL2 generality and basal vulnerability ($HerbLink_{\mathrm{nullB}} = 0.344$; $meanGenTL2_{\mathrm{nullB}} = 105.1$; figure 2*i*; $meanVulBasal_{\mathrm{nullB}} = 200.0$; figure 2*m*), as well

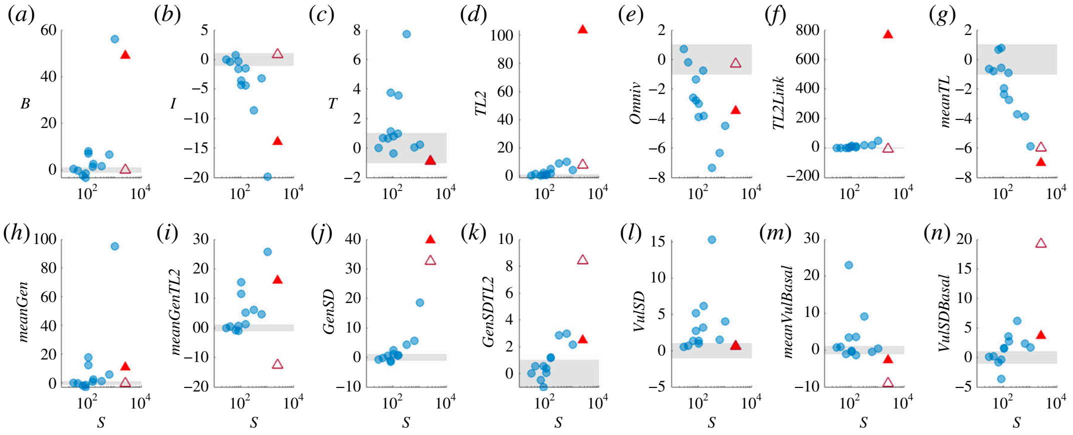

Figure 2. Structural properties of the MTF web deviate from previous ATFWs and null model expectations. Points are NMEs calculated for empirical food webs compared to null expectations from an ensemble of $n = 1000$ niche model food webs simulated with matched trophic species richness (S) and directed connectance (C). NMEs >0 or <0 indicate that observed properties are greater or less than expected, respectively. |NMEs| >1 (outside of the grey box) are significantly different from null expectations at the 95% confidence level. Blue dots represent previously published food webs (table 1). Red triangles are the MTF food web compared to the niche model with matched S, C (filled) or matched S, C and B, the fraction of species with no resources (i.e. basal species, achieved by setting species' feeding ranges to zero, hollow). Plots show that NMEs broadly scale with S, but the magnitudes of deviation associated with our web are greater than previously observed. See §2 for full definitions of food web properties. Fraction of (*a*) basal, (*b*) intermediate, (*c*) top, (*d*) strict herbivore (i.e. trophic level = 2) and (*e*) omnivorous species. (*f*) Fraction of feeding links by TL2 herbivores. (*g*) Average trophic level. Average generality of (*h*) consumers and (*i*) TL2 herbivores. Normalized standard deviation of generality of (*j*) consumer species and (*k*) TL2 herbivores. Normalized standard deviation of vulnerability of (*l*) resource species and (*n*) basal species. (*m*) Average vulnerability of basal species.

as a further underestimation of the variabilities of TL2 generality and basal vulnerability ($\mathrm{GenSDTL2}_{\mathrm{nullB}} = 0.92$; figure 2*k*; $\mathit{VulSDBasal}_{\mathrm{nullB}} = 0.31$; figure 2*n*). We also tested whether the higher B and herbivory interactions in our original trophic-species web were attributable to the presence of basal animals. After removing these species and links, we recovered similar properties to the original web ($B_{\mathrm{obsA}} = 0.232$, $I_{\mathrm{obsA}} = 0.768$, $\mathit{TL2}_{\mathrm{obsA}} = 0.491$, $\mathit{HerbLink}_{\mathrm{obsA}} = 0.117$).

This finding clarifies why our web deviates from niche model predictions (figure 3*d*). The strict hierarchy of the niche model causes decreasing B and T with increasing S and therefore becomes very unlikely to generate TL2 herbivores. However, even if forced to generate larger B, TL2 herbivores remain rare, with low variability in generality. This is because TL2 herbivores must have feeding ranges that only include basal species, which is most likely with narrow ranges (and therefore low niche values). By contrast, the TL2 herbivores in our web are numerous, have diets on average wider and more variable in size than those produced by the niche model, and their trophic species niches are distinguished by highly specific herbivory interactions rather than by their predator interactions.

4. Discussion

The food web for the MTF presented here is, to our knowledge, the largest yet published (table 1, [6,10]) and begins to shine light on the remarkable richness of feeding interactions between plants and animals in aboveground terrestrial systems. Our cumulative approach allowed us to evenly resolve the diets of both vertebrate and insect feeding guilds, revealing a clear pattern among herbivory and carnivory interactions. Herbivory interactions are far rarer (<15% of trophic links) than carnivory interactions, but the former were primarily sourced from the largest guild in the web (approx. 50% of trophic species)—the strict (i.e. trophic-level two) insect herbivores with variably sized but highly specific diets, both taxonomically and in terms of plant tissue types. In addition, carnivory interactions dominate the web (>85% of trophic links) but were primarily sourced from a small fraction of trophic species (<4.6%)—hyper-generalist birds and bats thought to feed opportunistically on entire insect orders. The combination of these two properties in our food web leads to a structure that qualitatively deviates from previous ATFWs of lower taxonomic and trophic resolution, as well as the scale-dependent null expectations of the niche model.

Given the richness and complexity of our trophic web, the niche model predicted an even larger fraction of carnivory interactions (>98%), stemming from the minimal fraction of basal trophic species with no consumers (B <2%) and herbivores ($\mathit{TL2}$ <1%), ultimately resulting in an overall NME far greater than for any previous webs studied here (MTF mean $|\mathrm{NME}| = 28.9$; figure 3*e*). ATFWs using the classic 'lumping' approach (Coachella Valley, St Martin Island, El Verde Rainforest, Serengeti de Visser [24]) are generally small, with low B and $\mathit{TL2}$ (owing to extremely coarse representation of plants and insects) and high fractions of carnivory links (owing to the disproportionate resolution of vertebrate predators). These properties align with the predictions of the niche model, as quantified by NMEs ($0.42 < \text{mean } |\mathrm{NME}| < 1.88$; figure 3*e*). High-resolution source and sink ATFWs (Scotch Broom [23], UK Grassland [22,61], High Arctic [17–19,28,62], Norwood Farm [29]) record primarily parasitoids and parasites as the higher trophic-level consumers but generally include herbivory links on multiple plant tissues. The relative lack of generalist or opportunistic predators leads to high fractions of top species and very low omnivory, both in contrast to our web and the expectations of the niche model ($0.97 < \text{mean } |\mathrm{NME}| < 6.21$). Most comparable to the MTF in terms of size and resolution is perhaps Messel Forest [30], though it is from the early Eocene and contains many now-extinct species.

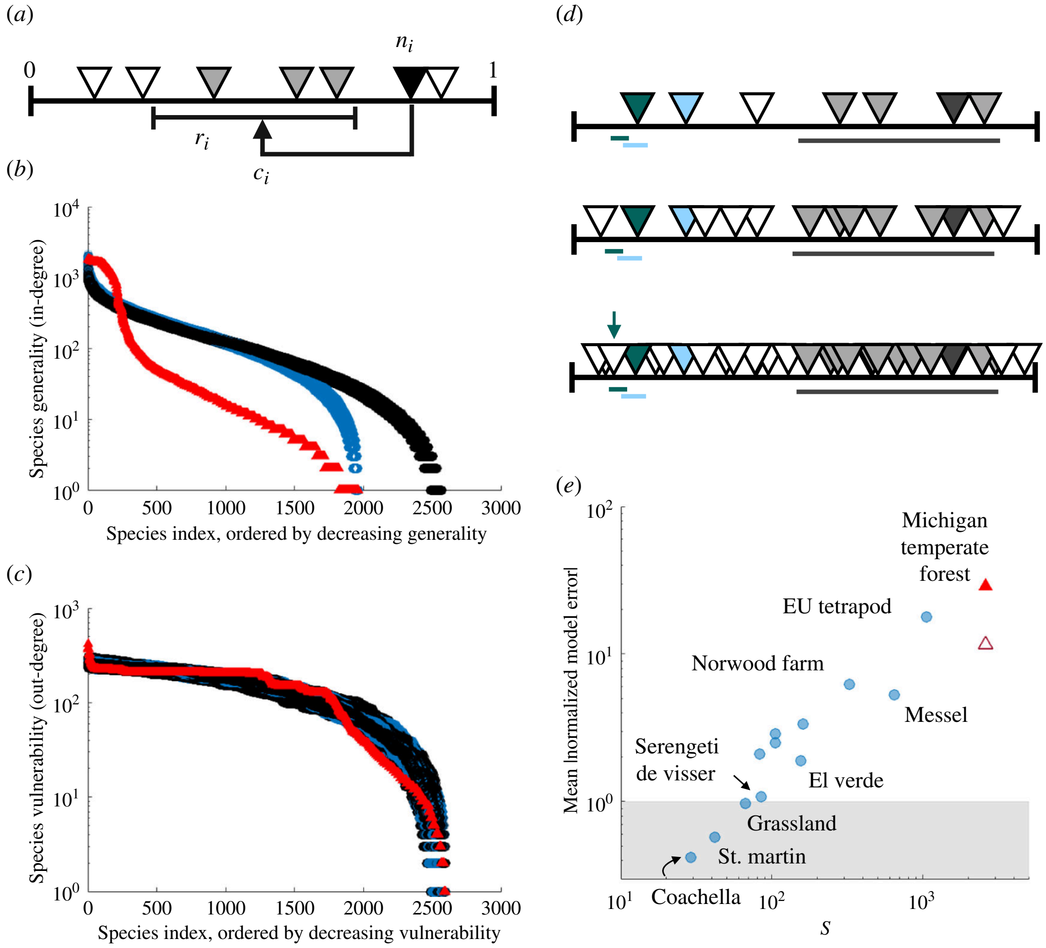

Figure 3. The MTF food web structure deviates from null expectations of the niche model owing to problems of niche packing at high richness. (*a*) We use Williams & Martinez' [8]) niche model to simulate an ensemble of null-model food webs with trophic species richness (S) and directed connectance (C) matched to the empirical food web. For each species i, the model randomly draws a niche value from a uniform distribution ($n_i \sim U(0,1)$) and assigns i a feeding range along the niche axis ($r_i = n_i x$, where x is β-distributed with $E[x] = 2C$) centered at a value (c_i) below its niche value ($c_i \in [r_i/2, min(n_i, 1 - r_i/2)]$). In this way, the higher the niche value of a species, the wider its feeding range. Then, i feeds on all species j with niche values in that range or is considered a basal species if the range is empty. Only trophic species with a unique set of consumers and resources are permitted. To guarantee a basal species, the feeding range of the species with the lowest niche value is set to zero ($r_i := 0$). (*b*) The observed in-degree distributions for the MTF web (red triangles) compared to in-degree distributions from $n = 1000$ simulated niche model webs with matched S, C (black circles) or matched S, C, and fraction of basal species, B, achieved by setting more species' feeding ranges to zero, blue circles). Degree is on a log scale (y-axis), with species (x-axis) ordered by decreasing degree. In-degree is generality, the richness of diet. The empirical web has a core of high-degree generalists (birds and bats) and a long tail of lower-degree trophic-level 2 (TL2) herbivores (insects), which leads to a significantly different distribution of generality than predicted by the niche model. (*c*) Observed and simulated out-degree distributions. Out-degree is vulnerability, the richness of predators. (*d*) The strict one-dimensional hierarchy of the niche model does not accommodate the rich community of TL2 herbivores observed in the empirical food web. As S increases, the niche axis becomes increasingly and uniformly packed with species, represented here as triangles. Horizontal bars represent the feeding ranges of each coloured species. At high S, the chance that species will have empty feeding ranges decreases, decreasing B (e.g. the green species). The fraction of TL2 herbivores also decreases with S because herbivores' feeding ranges must be perfectly placed to include a rare basal species but exclude any consumers. For example, the blue species feeds only on the green basal species but becomes a carnivore when the green species is switched to a consumer at high S. The fraction of top species (i.e. without predators, T) also decreases with S because the broad feeding ranges and tendency to cannibalism of high niche-value species increasingly cover the niche axis with predators and excludes those species from being top (e.g. the grey species). (*e*) Mean absolute NMEs across all traditional niche model properties. Formatting follows figure 2 but absolute rather than signed errors are shown, and points are on a log-log scale.

While the MTF has greater *TL2*, omnivores and carnivory links, as well as a lower fraction of herbivory links, the significantly different distributions of generality are probably the greatest contrast (electronic supplementary material, table S2). Messel Forest has fewer TL2 herbivores with narrower and less variable diets (deviating less from niche model predictions) and does not show an inflection in its (log) degree distribution caused by a group of hyper-generalists. As a result, the Messel Forest web shows an overall better correspondence to null expectations of the niche model (mean |NME| = 5.28), despite a substantially greater richness than classic webs.

The niche model is considered to successfully reiterate the properties of natural food webs and is even used to simulate network structures for studies of food web dynamics, though it is known to underestimate *TL2* and overestimate average trophic level [38,55]. This success can broadly be attributed to two mathematical properties, shared by other generating models of food web structure [1,3,63–65]: first, that species can be strictly ordered along a one-dimensional axis by their niche values ($n_1 < n_2 < \ldots n_S$), and second, that species' feeding ranges (r_i) are exponentially distributed, with a decaying probability of feeding on species of lower niche values. The superior performance of the niche model (given its simplicity) is explained by its third property [8]: it generates 'interval' webs, because species feed on all resources within intervals of the niche axis (i.e. they have contiguous diets). These properties are also ecological hypotheses for mechanisms that structure food webs [51]. For the MTF, the largest magnitudes of deviations ($|\text{NME}| > 95$) were in terms of species composition, which we attribute to issues with niche packing (figure 3*d*, but see [66]). This suggests that the MTF does not satisfy the first or third properties, since allowing more herbivores on a packed niche axis would require allowing an additional axis of food selection among herbivores of the same niche value or breaks in their diet contiguity. In fact, the MTF, like many others, is not strictly interval. As such, the more pertinent question is the level of intervality of the MTF relative to other webs of its scale and whether this could cause the large magnitude of deviations we observed. Unfortunately, this is combinatorically intractable for us to assess using current methods [3,67]. Finally, the MTF does not satisfy the second property of exponentially distributed feeding ranges (figure 3*b–c*). We hypothesize that for the niche model to better reproduce MTF properties, links would need to be preferentially allocated to a small group of generalist species and another axis associated with eating plants would need to be introduced to allow for rich communities of herbivores with variably sized diets [7].

Whether the structural patterns for the MTF could be general among aboveground terrestrial ecosystems remains to be tested. We observed an overall increase in NMEs with richness (figure 3*e*; figure 2), such that we cannot confidently reject the null hypothesis that our observed structure is attributable to scale-dependent errors associated with the niche model. However, the MTF appears to exhibit a qualitatively different structure than previous works, including a potentially novel degree distribution. Kolmogorov–Smirnov tests rejected the hypothesis that degree distributions of generality and vulnerability for our web were sampled from the same underlying process as the distributions of each of the other ATFWs studied here ($p < 1.0 \times 10^{-5}$ in all cases, electronic supplementary material, table S2). This could imply that terrestrial food webs do have unique structure—potentially driven by the opportunism of predators and a lack of hierarchy among strict herbivores—which may only be observed in the context of high taxonomic and trophic resolution among a rich community of plants, insects and vertebrates, such as in our temperate forest system.

As always, there are caveats to these conclusions, primarily associated with our methodological approach and general data limitations. It is possible that the critical hyper-generality of bats and birds in the MTF can be attributed simply to lack of taxonomic resolution regarding their specific foraging preferences (e.g. microlepidopterans, a paraphyletic group with <20 mm wingspans, may be too small to be eaten by vertebrate predators). Yet, studies examining species-specific diets of bats and birds have supported hyper-generalism (e.g. [68]), even to the extent that molecular identification of species in bat guano presents a roughly equivalent snapshot of insect biodiversity as traditional blacklight sampling of insects [69]. We also know that many plant and insect species are missing interactions because we were not able to find or verify species-specific data (owing to taxonomic or other data limitations) or because records were too vague (e.g. 'eating seeds' without further specificity). It is therefore possible that herbivory links were under-represented relative to carnivory in our observations. Interactions could be more thoroughly refined to account for species' traits and ecological habits, which would probably convert some links currently considered plausible to effectively 'forbidden links' *sensu* Jordano *et al*. [70,71]. Nevertheless, opportunistic carnivory links would still dominate the web and additional herbivory links would probably be taxonomically specific and trophically distinct, which may even serve to increase trophic species richness by distinguishing plants' or herbivores' trophic niches. In short, a further refinement of feeding interactions would most likely align with the food web structure described here.

Compiling taxonomic and feeding records for diverse groups into cumulative food webs potentially introduces diverse and compounding sampling biases, making it difficult to quantify overall uncertainty [33]. In the MTF, for example, the relative representation of some groups can be assessed directly through rarefaction curves over field seasons (e.g. lepidopteran leafminers), while some groups (e.g. non-insect arthropods) include diversity for which we were not able to source even regional lists as reference points (e.g. acariform mites). Likewise, feeding on some types of resources, especially by specific taxonomic groups, is more readily observed or charismatic than others, and therefore better documented. As food web research moves towards synthesizing bigger and more diverse data, an important theoretical question for future work is to develop quantitative methods for characterizing how uncertainty may affect observed network structure (see [33,72] for promising methods).

Like previous researchers, we chose to limit our scope to the aboveground portion of our food web. However, we recognize that all ATFWs are intimately and inextricably coupled with belowground soil food webs. A substantial fraction of plant biomass may exist belowground [73], creating habitat structure and providing food for organisms that consume roots, exudates and detritus. Decomposers return nutrients generated from aboveground waste to the soil for reuptake by plants, often facilitated by mutualistic symbionts (e.g. mycorrhizae) [74]. Moreover, many consumer species (including some in the MTF) live or feed belowground during certain lifestages or times of year. As such, complex food web dynamics belowground have considerable impacts on food web dynamics aboveground, and interactions between these two habitats can significantly influence ecosystem-level processes [75,76]. While this fact has long been recognized as a bias in ATFW research, rarely are above- and belowground food webs recorded in the same system (probably owing to the logistical challenges of sampling belowground). Doing so represents a critical research frontier for terrestrial food webs as we seek to understand their structure, dynamics and emergent ecosystem functions.

This study is only our first step towards documenting the immense taxonomic diversity and trophic complexity in the temperate forests of UMBS. Ecological networks have historically been published and analysed 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 ecologists and network scientists can perpetuate errors [34,77]. Our ability to create a large, highly resolved, expert-vetted ATFW was made possible through decades of research and observations at the UMBS and highlights the importance of local knowledge, taxon-specific expertise and collaboration among scientists, students and members of the public. As insights about the organisms present at UMBS will undoubtedly continue to grow, we consider the MTF web to be a living dataset that can be revised and expanded through time. To that end, our database is publicly available [54], and we are soliciting revisions, corrections and additions that will allow its continual improvement.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. Data and code are provided as supplementary material [53]. Species and interaction data are openly available via the Environmental Data Initiative at [78].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. K.R.S.H.: conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, visualization, writing—original draft, writing—review and editing; J.D.C.: data curation, writing—review and editing; G.G.A.: data curation, writing—review and editing; S.B.: data curation, writing—review and editing; R.L.K.F.: data curation, writing—review and editing; L.E.J.: data curation, writing—review and editing; K.L.M.: data curation, writing—review and editing; B.G.S.: data curation, writing—review and editing; M.S.: data curation, writing—review and editing; C.T.: data curation, writing—review and editing; J.T.: data curation, writing—review and editing; F.S.V.: conceptualization, funding acquisition, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was partially funded by NSF grants DEB−2129757 and DEB−2224915 to F.S.V. The University of Michigan Undergraduate Research Opportunities Program (UROP) paid the undergraduate researchers for their time.

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.

References

1. Camacho J, Guimerà R, Amaral LAN. 2002 Analytical solution of a model for complex food webs. *Phys. Rev. E* **65**, 030901. (doi:10.1103/PhysRevE.65.030901)
2. Dunne JA, Williams RJ, Martinez ND. 2002 Food-web structure and network theory: the role of connectance and size. *Proc. Natl Acad. Sci. USA* **99**, 12917–12922. (doi:10.1073/pnas.192407699)
3. Stouffer DB, Camacho J, Guimerà R, Ng CA, Amaral LAN. 2005 Quantitative patterns in the structure of model and empirical food webs. *Ecology* **86**, 1301–1311. (doi:10.1890/04-0957)
4. Stouffer DB, Camacho J, Jiang W, Amaral LAN. 2007 Evidence for the existence of a robust pattern of prey selection in food webs. *Proc. R. Soc. B* **274**, 1931–1940. (doi:10.1098/rspb.2007.0571)
5. Riede JO, Rall BC, Banasek-Richter C, Navarrete SA, Wieters EA, Emmerson MC, Jacob U, Brose U. 2010 Scaling of food-web properties with diversity and complexity across Ecosystems. In *Advances in ecological research* (ed. G Woodward), pp. 139–170, vol. **42**. Academic Press: Burlington. (doi:978-0-12-381363-3)
6. Cirtwill AR, Stouffer DB, Romanuk TN. 2015 Latitudinal gradients in biotic niche breadth vary across ecosystem types. *Proc. R. Soc. B* **282**, 20151589. (doi:10.1098/rspb.2015.1589)
7. Williams RJ, Purves DW. 2011 The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs. *Ecology* **92**, 1849–1857. (doi:10.1890/11-0200.1)
8. Williams RJ, Martinez ND. 2000 Simple rules yield complex food webs. *Nature* **404**, 180–183. (doi:10.1038/35004572)
9. Brose U *et al*. 2006 Consumer-resource body-size relationships in natural food webs. *Ecology* **87**, 2411–2417. (doi:10.1890/0012-9658(2006)87[2411:cbrinf]2.0.co;2)
10. Brose U *et al*. 2019 Predator traits determine food-web architecture across ecosystems. *Nat. Ecol. Evol.* **3**, 919–927. (doi:10.1038/s41559-019-0899-x)
11. Petchey OL, Beckerman AP, Riede JO, Warren PH. 2008 Size, foraging, and food web structure. *Proc. Natl Acad. Sci. USA* **105**, 4191–4196. (doi:10.1073/pnas.0710672105)
12. Stouffer DB, Bascompte J. 2011 Compartmentalization increases food-web persistence. *Proc. Natl Acad. Sci. USA* **108**, 3648–3652. (doi:10.1073/pnas.1014353108)
13. Potapov AM, Brose U, Scheu S, Tiunov AV. 2019 Trophic position of consumers and size structure of food webs across aquatic and terrestrial ecosystems. *Am. Nat.* **194**, 823–839. (doi:10.1086/705811)
14. Valdovinos FS *et al*. 2023 A bioenergetic framework for aboveground terrestrial food webs. *Trends Ecol. Evol. (Amst.)* **38**, 301–312. (doi:10.1016/j.tree.2022.11.004)
15. Cohen JE. 1989 Ecologists Co-operative Web Bank (ECOWeB) Version 1.0. Machine readable data base of food webs. New York, NY: Rockefeller University, New York. See https://digitalcommons.rockefeller.edu/cohen_joel_laboratory/1/.
16. Baiser B *et al*. 2019 Ecogeographical rules and the macroecology of food webs. *Glob. Ecol. Biogeogr.* **28**, 1204–1218. (doi:10.1111/geb.12925)
17. Wirta HK, Vesterinen EJ, Hambäck PA, Weingartner E, Rasmussen C, Reneerkens J, Schmidt NM, Gilg O, Roslin T. 2015 Exposing the structure of an Arctic food web. *Ecol. Evol.* **5**, 3842–3856. (doi:10.1002/ece3.1647)
18. Roslin T, Wirta H, Hopkins T, Hardwick B, Várkonyi G. 2013 Indirect interactions in the high Arctic. *PLoS One* **8**, e67367. (doi:10.1371/journal.pone.0067367)
19. Rasmussen C, Dupont YL, Mosbacher JB, Trøjelsgaard K, Olesen JM. 2013 Strong impact of temporal resolution on the structure of an ecological network. *PLoS One* **8**, e81694. (doi:10.1371/journal.pone.0081694)
20. Polis GA. 1991 Complex trophic interactions in deserts: an empirical critique of food-web theory. *Am. Nat.* **138**, 123–155. (doi:10.1086/285208)
21. Goldwasser L, Roughgarden J. 1993 Construction and analysis of a large Caribbean food web. *Ecology* **74**, 1216–1233. (doi:10.2307/1940492)
22. Martinez ND, Hawkins BA, Dawah HA, Feifarek BP. 1999 Effects of sampling effort on characterization of food-web structure. *Ecology* **80**, 1044–1055. (doi:10.1890/0012-9658(1999)080[1044:EOSEOC]2.0.CO;2)
23. Memmott J, Martinez ND, Cohen JE. 2000 Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web. *J. Anim. Ecol.* **69**, 1–15. (doi:10.1046/j.1365-2656.2000.00367.x)
24. De Visser SN, Freymann BP, Olff H. 2011 The Serengeti food web: empirical quantification and analysis of topological changes under increasing human impact. *J. Anim. Ecol.* **80**, 484–494. (doi:10.1111/j.1365-2656.2010.01787.x)
25. Baskerville EB, Dobson AP, Bedford T, Allesina S, Anderson TM, Pascual M. 2011 Spatial guilds in the Serengeti food web revealed by a Bayesian group model. *PLoS Comput. Biol.* **7**, e1002321. (doi:10.1371/journal.pcbi.1002321)

26. Harris LD, Paur Leonard Francis. 1972 *A quantitative food web analysis of a Shortgrass community. technical report no. 154, grassland Biome. U.S. International biological Program*. SPIRE. p. 17. See http://spire.umbc.edu/fwc/.
27. Reagan DP, Waide RB (eds). 1996 *The food web of a tropical rain forest*. Chicago, IL: University of Chicago Press.
28. Schmidt NM *et al*. 2017 Interaction webs in Arctic ecosystems: determinants of Arctic change? *Ambio* **46**, 12–25. (doi:10.1007/s13280-016-0862-x)
29. Pocock MJO, Evans DM, Memmott J. 2012 The robustness and restoration of a network of ecological networks. *Science* **335**, 973–977. (doi:10.1126/science.1214915)
30. Dunne JA, Labandeira CC, Williams RJ. 2014 Highly resolved early Eocene food webs show development of modern trophic structure after the end-Cretaceous extinction. *Proc. R. Soc. B* **281**, 20133280. (doi:10.1098/rspb.2013.3280)
31. O'Connor LMJ, Pollock LJ, Braga J, Ficetola GF, Maiorano L, Martinez-Almoyna C, Montemaggiori A, Ohlmann M, Thuiller W. 2020 Unveiling the food webs of tetrapods across Europe through the prism of the Eltonian niche. *J. Biogeogr.* **47**, 181–192. (doi:10.1111/jbi.13773)
32. Song C, Saavedra S. 2020 Telling ecological networks apart by their structure: an environment-dependent approach. *PLoS Comput. Biol.* **16**, e1007787. (doi:10.1371/journal.pcbi.1007787)
33. De Aguiar MAM *et al*. 2019 Revealing biases in the sampling of ecological interaction networks. *PeerJ* **7**, e7566. (doi:10.7717/peerj.7566)
34. Pringle RM, Hutchinson MC. 2020 Resolving food-web structure. *Annu. Rev. Ecol. Evol. Syst.* **51**, 55–80. (doi:10.1146/annurev-ecolsys-110218-024908)
35. Hu G, Lim KS, Horvitz N, Clark SJ, Reynolds DR, Sapir N, Chapman JW. 2016 Mass seasonal bioflows of high-flying insect migrants. *Science* **354**, 1584–1587. (doi:10.1126/science.aah4379)
36. Bartley TJ, McCann KS, Bieg C, Cazelles K, Granados M, Guzzo MM, MacDougall AS, Tunney TD, McMeans BC. 2019 Food web rewiring in a changing world. *Nat. Ecol. Evol.* **3**, 345–354. (doi:10.1038/s41559-018-0772-3)
37. Ceron K, Provete DB, Pires MM, Araujo AC, Blüthgen N, Santana DJ. 2022 Differences in prey availability across space and time lead to interaction rewiring and reshape a predator-prey metaweb. *Ecology* **103**, e3716. (doi:10.1002/ecy.3716)
38. Dunne JA. 2005 The network structure of food webs. In *Ecological networks* (eds M Pascual, JA Dunne), pp. 27–92. New York, NY: Oxford University Press. (doi:10.1093/oso/9780195188165.003.0002)
39. Briand F. 1983 Environmental control of food web structure. *Ecology* **64**, 253–263. (doi:10.2307/1937073)
40. Martinez ND. 1991 Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.* **61**, 367–392. (doi:10.2307/2937047)
41. Briand F, Cohen JE. 1984 Community food webs have scale-invariant structure. *Nature* **307**, 264–267. (doi:10.1038/307264a0)
42. Schoenly K, Beaver RA, Heumier TA. 1991 On the trophic relations of insects: a food-web approach. *Am. Nat.* **137**, 597–638. (doi:10.1086/285185)
43. Hale KRS, Thébault E, Valdovinos FS. 2023 A general trait-based model for multiplex ecological networks. *bioRxiv*. (doi:10.1101/2023.08.08.552546)
44. Kéfi S *et al*. 2012 More than a meal… integrating non-feeding interactions into food webs. *Ecol. Lett.* **15**, 291–300. (doi:10.1111/j.1461-0248.2011.01732.x)
45. Fontaine C, Guimarães PR, Kéfi S, Loeuille N, Memmott J, Van Der Putten WH, Veen FJF, Thébault E. 2011 The ecological and evolutionary implications of merging different types of networks: merging networks with different interaction types. *Ecol. Lett.* **14**, 1170–1181. (doi:10.1111/j.1461-0248.2011.01688.x)
46. Olff H, Alonso D, Berg MP, Eriksson BK, Loreau M, Piersma T, Rooney N. 2009 Parallel ecological networks in ecosystems. *Phil. Trans. R. Soc. B* **364**, 1755–1779. (doi:10.1098/rstb.2008.0222)
47. Pilosof S, Porter MA, Pascual M, Kéfi S. 2017 The multilayer nature of ecological networks. *Nat. Ecol. Evol.* **1**, 0101. (doi:10.1038/s41559-017-0101)
48. Kéfi S, Thébault E, Eklöf A, Lurgi M, Davis AJ, Kondoh M, Krumins JA. 2017 Toward multiplex ecological networks: accounting for multiple interaction types to understand community structure and dynamics. In *Adaptive food webs* (eds JC Moore, PC De Ruiter, KS McCann, V Wolters), pp. 73–87. New York, NY: Cambridge University Press. (doi:10.1017/9781316871867)
49. Timóteo S *et al*. 2023 Tripartite networks show that keystone species can multitask. *Funct. Ecol.* **37**, 274–286. (doi:10.1111/1365-2435.14206)
50. Dunne JA *et al*. 2013 Parasites affect food web structure primarily through increased diversity and complexity. *PLoS Biol.* **11**, e1001579. (doi:10.1371/journal.pbio.1001579)
51. Martinez ND. 2020 Allometric trophic networks from individuals to socio-ecosystems: consumer–resource theory of the ecological elephant in the room. *Front. Ecol. Evol.* **8**, 92. (doi:10.3389/fevo.2020.00092)
52. Barnett JW, Bergen KM, Currie WS. 2022 Assessing beech bark-diseased forest canopies over landscapes using high resolution open-source imagery in an ecological framework. *For. Ecol. Manage.* **523**, 120486. (doi:10.1016/j.foreco.2022.120486)
53. Hale KRS *et al*. 2024 Supplementary material from: A highly resolved network reveals the role of terrestrial Herbivory in structuring aboveground food webs. Figshare. (doi:10.6084/m9.figshare.c.7318729)
54. Hale KRS *et al*. 2023 University of Michigan biological station cumulative food web data for terrestrial habitats. (doi:10.6073/PASTA/840D70788BDE4692A7D6D45F8D04376F)
55. Williams RJ, Martinez ND. 2008 Success and its limits among structural models of complex food webs. *J. Anim. Ecol.* **77**, 512–519. (doi:10.1111/j.1365-2656.2008.01362.x)
56. Gauzens B, Rall BC, Mendonça V, Vinagre C, Brose U. 2020 Biodiversity of intertidal food webs in response to warming across latitudes. *Nat. Clim. Change* **10**, 264–269. (doi:10.1038/s41558-020-0698-z)
57. Williams RJ, Martinez ND. 2004 Limits to trophic levels and omnivory in complex food webs: theory and data. *Am. Nat.* **163**, 458–468. (doi:10.1086/381964)
58. Pires MM, Prado PI, Guimarães PR. 2011 Do food web models reproduce the structure of mutualistic networks? *PLoS One* **6**, e27280. (doi:10.1371/journal.pone.0027280)
59. 2021 MATLAB Vers. R2021b. Natick, MA: The MathWorks Inc.
60. Stork NE. 2018 How many species of insects and other terrestrial arthropods are there on earth? *Annu. Rev. Entomol.* **63**, 31–45. (doi:10.1146/annurev-ento-020117-043348)
61. Dawah HA, Hawkins BA, Claridge MF. 1995 Structure of the parasitoid communities of grass-feeding chalcid wasps. *J. Anim. Ecol.* **64**, 708. (doi:10.2307/5850)
62. Wirta HK, Hebert PDN, Kaartinen R, Prosser SW, Várkonyi G, Roslin T. 2014 Complementary molecular information changes our perception of food web structure. *Proc. Natl Acad. Sci. USA* **111**, 1885–1890. (doi:10.1073/pnas.1316990111)
63. Cattin MF, Bersier LF, Banasek-Richter C, Baltensperger R, Gabriel JP. 2004 Phylogenetic constraints and adaptation explain food-web structure. *Nat. New Biol.* **427**, 835–839. (doi:10.1038/nature02327)
64. Cohen JE, Palka ZJ. 1990 A stochastic theory of community food webs. V. Intervality and triangulation in the trophic-niche overlap graph. *Am. Nat.* **135**, 435–463. (doi:10.1086/285055)
65. Williams RJ, Anandanadesan A, Purves D. 2010 The probabilistic niche model reveals the niche structure and role of body size in a complex food web. *PLoS One* **5**, e12092. (doi:10.1371/journal.pone.0012092)
66. Martinez ND, Lawton JH. 1995 Scale and food-web structure: from local to global. *Oikos* **73**, 148. (doi:10.2307/3545903)
67. Eklöf A *et al*. 2013 The dimensionality of ecological networks. *Ecol. Lett.* **16**, 577–583. (doi:10.1111/ele.12081)

68. Clare EL, Fraser EE, Braid HE, Fenton MB, Hebert PDN. 2009 Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular approach to detect arthropod prey. *Mol. Ecol.* **18**, 2532–2542. (doi:10.1111/j.1365-294X.2009.04184.x)
69. Maslo B, Mau RL, Kerwin K, McDonough R, McHale E, Foster JT. 2022 Bats provide a critical ecosystem service by consuming a large diversity of agricultural pest insects. *Agric. Ecosyst. Environ.* **324**, 107722. (doi:10.1016/j.agee.2021.107722)
70. Jordano P, Bascompte J, Olesen JM. 2003 Invariant properties in coevolutionary networks of plant–animal interactions. *Ecol. Lett.* **6**, 69–81. (doi:10.1046/j.1461-0248.2003.00403.x)
71. Allesina S, Alonso D, Pascual M. 2008 A general model for food web structure. *Science* **320**, 658–661. (doi:10.1126/science.1156269)
72. Young JG, Cantwell GT, Newman MEJ. 2021 Bayesian inference of network structure from unreliable data. *J. Complex Netw.* **8**, cnaa046. (doi:10.1093/comnet/cnaa046)
73. Ma H, Mo L, Crowther TW, Maynard DS, Van Den Hoogen J, Stocker BD, Terrer C, Zohner CM. 2021 The global distribution and environmental drivers of aboveground versus belowground plant biomass. *Nat. Ecol. Evol.* **5**, 1110–1122. (doi:10.1038/s41559-021-01485-1)
74. Bardgett RD, Wardle DA. 2003 Herbivore-mediated linkages between aboveground and belowground communities. *Ecology* **84**, 2258–2268. (doi:10.1890/02-0274)
75. Wardle DA, Bardgett RD, Klironomos JN, Setälä H, Van Der Putten WH, Wall DH. 2004 Ecological linkages between aboveground and belowground biota. *Science* **304**, 1629–1633. (doi:10.1126/science.1094875)
76. Jochum M, Eisenhauer N. 2022 Out of the dark: using energy flux to connect above- and belowground communities and ecosystem functioning. *Eur. J. Soil Sci.* **73**, e13154. (doi:10.1111/ejss.13154)
77. Brimacombe C, Bodner K, Michalska-Smith M, Poisot T, Fortin M-J. 2023 Shortcomings of reusing species interaction networks created by different sets of researchers. *PLoS Biol.* **21**, e3002068. (doi:10.1371/journal.pbio.3002068)
78. Hale KRS *et al*. 2024 University of Michigan Biological Station cumulative food web data for terrestrial habitats, 1909-2023 ver 3. Environmental Data Initiative. See https://doi.org/10.6073/pasta/1b02ac7e71e55bcae3ee465a864c20ae (accessed 4 July 2024).