PALEOBIOLOGY
A PUBLICATION OF THE

Paleontological
S O C I E T Y

Featured Article

Disentangling ecological and taphonomic signals in ancient food webs

Jack O. Shaw ⓘ*, Emily Coco, Kate Wootton, Dries Daems* ⓘ*, Andrew Gillreath-Brown* ⓘ*, Anshuman Swain, and Jennifer A. Dunne*

Abstract.—Analyses of ancient food webs reveal important paleoecological processes and responses to a range of perturbations throughout Earth's history, such as climate change. These responses can inform our forecasts of future biotic responses to similar perturbations. However, previous analyses of ancient food webs rarely accounted for key differences between modern and ancient community data, particularly selective loss of soft-bodied taxa during fossilization. To consider how fossilization impacts inferences of ancient community structure, we (1) analyzed node-level attributes to identify correlations between ecological roles and fossilization potential and (2) applied selective information loss procedures to food web data for extant systems. We found that selective loss of soft-bodied organisms has predictable effects on the trophic structure of "artificially fossilized" food webs because these organisms occupy unique, consistent food web positions. Fossilized food webs misleadingly appear less stable (i.e., more prone to trophic cascades), with less predation and an overrepresentation of generalist consumers. We also found that ecological differences between soft- and hard-bodied taxa—indicated by distinct positions in modern food webs—are recorded in an early Eocene web, but not in Cambrian webs. This suggests that ecological differences between the groups have existed for ≥48 Myr. Our results indicate that accounting for soft-bodied taxa is vital for accurate depictions of ancient food webs. However, the consistency of information loss trends across the analyzed food webs means it is possible to predict how the selective loss of soft-bodied taxa affects food web metrics, which can permit better modeling of ancient communities.

Jack O. Shaw. Department of Earth and Planetary Sciences, Yale University, New Haven, Connecticut 06511, U.S.A. E-mail: jack.shaw@yale.edu

Emily Coco. Center for the Study of Human Origins, Department of Anthropology, New York University, New York, New York 10003, U.S.A. E-mail: ec3307@nyu.edu

Kate Wootton. Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 75651 Uppsala, Sweden. E-mail: kate.wootton@slu.se

Dries Daems. Sagalassos Archaeological Research Project, Department of Archaeology, University of Leuven, Blijde-Inkomststraat 21/3314, 3000 Leuven, Belgium. E-mail: dries.daems@kuleuven.be

Andrew Gillreath-Brown. Department of Anthropology, Washington State University, Pullman, Washington 99163, U.S.A. E-mail: andrew.d.brown@wsu.edu

Anshuman Swain. Department of Biology, University of Maryland, College Park, Maryland 20742, U.S.A. E-mail: answain@terpmail.umd.edu

Jennifer A. Dunne. Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, U.S.A. E-mail: jdunne@santafe.edu

Introduction

Communities of interacting animals, plants, and other organisms can be represented as food webs, networks of nodes (taxa) connected by links (consumer–resource interactions). This network-based approach is vital for community ecology studies, as it connects interspecific interactions to community-level responses, which are unpredictable if taxa are considered in isolation (Schmitz 1997). Comparative studies of ecological network properties, also referred to as metrics (e.g., Table 1), across habitats and species types have revealed similarities in community structure (e.g., Stouffer et al. 2007; Dunne et al. 2013). Combined analyses of community composition (types of organisms present), community structure (whole-network

0094-8

CrossMark

TABLE 1. Descriptions of terms, metrics, and acronyms used throughout this paper.

	Term	Description	Abbreviation
Misc.	Fossilization probability	An individual taxon's probability of fossilizing, not accounting for the environmental filter. Measured as a value between 0 and 1, where 1 means that the taxon always fossilizes, and 0 means the species never fossilizes. We used values of 0.25 for soft-bodied species, 0.5 for intermediate species, and 0.75 for hard-bodied species, unless noted otherwise.	FP
	Environmental filter distribution	Values drawn from this beta distribution are compared with fossilization probabilities to determine whether a taxon successfully fossilizes in the "artificial fossilization" process. The alpha and beta components of the beta distribution sum to 10. Higher alpha parameters indicate a stronger environmental filter, which removes more taxa during the artificial fossilization process.	EF distribution
	Model error	We use model error to describe differences between the network-level metrics of a target web and a null model (here either niche model webs or randomly fossilized webs). Values exceeding ±1 indicate significant differences from the null model.	ME
Network-level metrics	Clustering	Probability that adjacent nodes of a node are connected to each other (also known as transitivity).	
	Connectance	Proportion of possible links in a network that actually occur. Calculated as the number of links (L) divided by number of species (S) squared (i.e., L/S^2).	C
	Diameter	Shortest distance between the two most distant nodes of a network, (i.e., the longest path of all shortest paths between each pair of nodes).	
	Mean normalized degree	Mean of all normalized node degrees within a network. The degree of a node the number of links it has to other nodes. Normalized degree is the degree normalized by the size of the network minus 1 (e.g., L/S-1).	Mean degree
	Characteristic path length	Average number of links in the shortest paths between all node pairs.	CPL
	Mean short-weighted trophic level	Average of short-weighted trophic level across all taxa in a network, where short-weighted trophic level is calculated as the average of the shortest TL of a consumer (1 + shortest chain length from the consumer to a basal node) and prey-averaged TL (1 + mean TL of all the consumer's resources).	Mean TL
	Systems omnivory index (Libralato 2008)	Distribution of feeding interactions among trophic levels through the weighted average of omnivory of the consumers. Uses short-weighted trophic level values.	SOI
	Size	Number of nodes in a graph.	S/taxon richness
Node-level metrics	Normalized betweenness centrality	Extent to which a node lies on paths between other nodes. Calculated as the number of shortest paths that pass through a given vertex and normalized to size of network excluding the node in question (S - 1).	Btw.
	Normalized in-degree	How many resource nodes a consumer node has. Normalized to the average number of links per node for that network.	Generality
	Normalized out-degree	How many consumer nodes a resource node has. Normalized to the average number of links per node for that network.	Vulnerability
	Short-weighted trophic level (Williams and Martinez 2004)	The average of the shortest TL of a consumer (1 + shortest chain length from the consumer to a basal node) and prey averaged TL (1 + mean TL of all the consumer's resources).	TL
	Omnivory index (Christensen and Pauly 1992)	Variance of short-weighted trophic levels of resource taxa.	OI

properties), and organismal interactions (node-level structural properties and the distribution of interactions) have revealed various aspects of community-level dynamics, function, and stability (e.g., McCann 2000; Montoya et al. 2006; Cardinale et al. 2009; Thompson et al. 2012; Gauzens et al. 2020).

The application of network approaches to ecological interaction data has supported ecologically specific interpretations of various network metrics (Lau et al. 2017). For example, in-degree (Table 1) has been shown to have implications for how vulnerable a food web is likely to be to biodiversity loss (Dunne et al. 2002b), and path lengths have been used to understand how quickly perturbations can spread through ecosystems (Williams et al. 2002; Dunne et al. 2016; Delmas et al. 2019). Although there are questions about how variable species interactions are across space and time (Poisot et al. 2015) and how universal ecological network metrics are across systems (i.e., Blüthgen 2010), research evaluating how network metrics can reflect ecological processes, dynamics, and responses is extensive (Delmas et al. 2019). These network metrics are often used to evaluate biotic responses to future perturbations (e.g., Albouy et al. 2019), but inferences are limited by recent ecosystems having not experienced the range or intensity of climatic and other perturbations forecast for the future.

The fossil record, given its greater range of taxonomic diversity, community types, and disturbance types and intensities, has the potential to greatly enhance our understanding of community responses to significant current and future perturbations (Williams and Jackson 2007). The analysis of fossil food webs holds particular promise for these types of questions, although only a few such studies have been conducted thus far. Two food web analyses of fossil data have revealed similarities of community structure among ancient and modern ecosystems and examined the implications for the stability of food web structure over deep time (Dunne et al. 2008, 2014). Other studies have assessed probabilistic community responses to perturbations such as mass extinctions (Roopnarine 2009; Roopnarine and Angielczyk 2015). However, these studies did not fully

account for the preservation biases inherent in fossil data sets. Unlike modern data, records of ancient communities are subject to information loss caused by geological processes—such as the selective loss of soft-bodied taxa—which affects some taxa and environments more than others. This reduced fidelity may bias food web properties and result in inaccurate inferences of ancient community structure and associated processes. Such biases need to be accounted for to draw robust ecological inferences about the past and future.

Specifically, the faunal composition of fossil deposits is affected by (1) environmental attributes, such as time averaging, environmental conditions, and changes in seawater chemistry; and (2) organismal attributes, such as biomineralization, size, and life habit (for a review of fossil record biases, see Kidwell and Holland 2002). One study estimated that only 40% of taxa in intertidal communities were likely to leave any fossil evidence due to organismal attributes alone (Schopf 1978). We refer to the percentage of taxa likely to leave fossil evidence as an assemblage's "community fossilization potential." If an organismal attribute is correlated with both fossilization potential and community role, then the selective loss of certain taxa will lead to fossil assemblages lacking taxa representing particular community roles. As an extension of this, food webs based on these depauperate assemblages will depict misleading structures and can result in unsound inference of relevant processes and dynamics. For community paleoecology studies, a key source of information loss—and the primary focus of this paper—is the selective loss of weakly and non-biomineralizing organisms (Kidwell and Flessa 1995). However, how biomineralization, fossilization potential, and community role covary—and thus how the selective removal of soft-bodied taxa influences analyses of food web structure—is unknown.

Sampling biases affect both modern and fossil data. A variety of studies have analyzed the impacts of different modes of information loss on apparent food web structure for extant systems, including analyses focused on sampling effort, spatial extent of sampling, resolution, and aggregation of data (e.g., Martinez 1991; Goldwasser and Roughgarden 1997;

Wood et al. 2015). Fossilization adds an extra layer of complexity to interpretations of ancient systems, given that one must identify whether something was truly absent from a record or whether it was lost due to taphonomy—processes affecting organic remains as they become fossilized, such as burial and decay. The few studies of fossil food webs have considered a limited range of information loss issues related to these processes. Analyses of Cambrian (Dunne et al. 2008) and early Eocene (Dunne et al. 2014) food webs based on Lagerstätten—exceptionally preserved fossil deposits containing both soft- and hard-bodied organisms—found that food web structure was relatively insensitive to the exclusion of low-certainty or random links. Studies of less well-preserved, more typical fossil assemblages (i.e., those containing only hard-bodied components) grouped taxa with similar functional properties and likely trophic habits into guilds to account for limited information about specific interspecies interactions (Roopnarine et al. 2007, 2019; Roopnarine and Angielczyk 2015).

One recent study (Roopnarine and Dineen 2018) used a modern coral reef food web to explore the effects of (1) selective information loss due to taphonomy and (2) aggregation of species into guilds on potential analyses of fossil food webs. The authors used an indirect proxy for fossilization potential, a taxon's appearance in the Paleobiology Database (PBDB), to simulate differences between modern and ancient community data. Although the distribution of dietary breadths and modularity of webs subjected to information loss were indistinguishable from the original webs, the removal of taxa not represented in the PBDB—inferred to have low fossilization potentials—significantly impacted inferences of trophic structure. These taxa were generally positioned at lower trophic levels, such that their selective removal resulted in decreased trophic-level assignments for remaining taxa. While the authors helped to constrain the broad impacts of information loss on a single trophic system, it is less clear how their findings apply to other systems or how they are influenced by the various organismal and environmental factors influencing preservation. Given

that taphonomic information loss affects different taxa and environments to varying degrees (Kidwell and Holland 2002), generalizable conclusions about the effect of information loss on food webs cannot be drawn from one system.

Here, we identify correlations between fossilization potential and food web features across eight highly resolved food webs to elucidate the implications of taphonomic information loss on the interpretation of ancient community structure. We first evaluate how differences in community structure and community fossilization potential could influence how food webs respond to taphonomic information loss. We then test whether the anatomical fossilization potential of a taxon—based on the presence of biomineralizing elements—varies with its ecological position in a community by comparing node-level attributes between taxa of differing fossilization potentials. Next, we test whether the selective loss of low fossilization-potential taxa has systematic effects on network-level food web structure by applying information loss models to modern systems. Finally, we conclude by considering what ecological inferences can reliably be drawn from fossil data and whether systematic effects could be accounted for in future analyses of ancient food webs. Our focus is on aquatic systems, given that they are best documented in the fossil record.

Methods

Data

We analyzed food webs for four modern marine systems, one modern lake system, two ancient marine systems, and one ancient lake system from previous publications (Martinez 1991; Dunne et al. 2008, 2014, 2016; Jacob et al. 2011; Roopnarine and Hertog 2013; Kéfi et al. 2016; Table 2, Supplementary Text). All webs have similar, broad higher-rank taxonomic compositions, are mostly resolved to the genus and species level (Supplementary Fig. 1), and contain at least 85 nodes (the size of the smallest ancient network considered). Nodes in the Cayman Islands web represent guilds rather than individual taxa (see Supplementary Text for more information on data processing). The three ancient food webs are based

TABLE 2.   Summary of webs studied. Connectance, clustering, characteristic path length (CPL), mean degree, diameter, mean trophic level (mean TL), and system omnivory index (SOI) are calculated using trophic species webs. Ancient webs are ordered by age and modern webs by the number of taxa. Int., intermediate-group taxa (see text); C, connectance.

Web	% Soft-bodied	% Int.	% Hard-bodied	No. of taxa	No. of trophic species	C	Clustering	CPL	Mean degree	Diameter	Mean TL	SOI	% Primary producers	Environment
Chengjiang Shale (Dunne et al. 2008)	29	32	39	85	26	0.11	0.68	1.77	0.23	4	2.85	2.83	5	Marine–offshore
Burgess Shale (Dunne et al. 2008)	20	24	56	143	43	0.11	0.43	1.64	0.22	3	2.57	3.56	17	Marine–basinal
Messel Shale (Dunne et al. 2014)	21	52	27	106	104	0.07	0.30	2.25	0.14	6	2.30	11.20	27	Lake
Chilean (Kéfi et al. 2016)	7	64	29	106	69	0.13	0.18	1.30	0.27	3	1.82	2.49	45	Marine–intertidal
Little Rock (Martinez 1991)	12	82	6	178	76	0.12	0.41	1.89	0.24	5	2.39	4.96	35	Lake
Cayman Islands (Roopnarine and Hertog 2013)	7	11	82	234	229	0.07	0.11	1.82	0.13	6	2.86	30.80	6	Marine–reef
Weddell Sea (Jacob et al. 2011)	24	31	45	492	275	0.09	0.38	2.27	0.18	8	2.97	49.20	13	Marine–shelf
Sanak archipelago (Dunne et al. 2016)	17	21	62	513	510	0.03	0.18	3.41	0.05	13	2.71	46.30	11	Marine–nearshore

on Lagerstätten where specific trophic links are inferred based on various lines of evidence, such as functional morphology, damage patterns, and gut contents (Dunne et al. 2008) (different from stochastic modeling procedures used to generate other ancient food webs; e.g., Roopnarine 2009). Although many taxa have been discovered from the Chengjiang, Burgess, and Messel deposits since the publication of the three food webs, updating the webs is outside of the scope of this paper.

We delimit three "preservation groups": (1) "hard-bodied" taxa are those with biomineralized skeletons, heavily sclerotized parts, or decay-resistant organic cuticle; (2) "intermediate-group" taxa are those with light sclerotization or unsclerotized cuticle; and (3) "soft-bodied" taxa are those with soft cellular outer layers and soft internal tissues (Fig. 1). Hard-bodied taxa are by far the most abundantly represented in the fossil record. Although we distinguish between soft-bodied and intermediate group taxa—given that the latter display some degree of biomineralization—both are severely underrepresented in the fossil record. We refer to the combination of the two as "poorly preserved" taxa. We assigned preservation groups to metazoans at the ordinal level (Supplementary Data) based on preservation group assignment for the majority of taxa within the clade. Assignments were primarily based on descriptions from standard sources (Kardong 1995; Guralnick et al. 2000; Brusca and Brusca 2003; Benton 2014), but additionally consulted references are listed in Supplementary Data 2. Enigmatic metazoans, bacteria, fungi, nonmineralizing plankton, and detritus were assigned to the soft-bodied group. Plants and mineralizing plankton were assigned to the intermediate group. We individually assigned preservation groups to taxa without order-level taxonomic information (e.g., many Cambrian taxa) based on assessments of fossilization potential at the lowest taxonomic rank possible (usually species). All data and code used herein are available in the corresponding Dryad repository.

Evaluating Food Web Structure

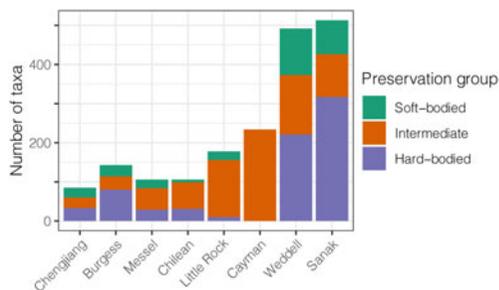*Node-Level Attributes.*—We used five node-level metrics (Table 1) to assess whether taxa

FIGURE 1. Breakdown of food web preservation group composition. Ancient webs (Chengjiang, Burgess, and Messel) ordered by age followed by modern webs ordered by taxon richness.

in different preservation groups occupied distinct network positions spanning modern and ancient webs, indicative of ecological differences between the groups. These metrics included short-weighted trophic level (TL), omnivory index (OI), normalized in-degree (generality), normalized out-degree (vulnerability), and betweenness centrality. We assessed the significance of differences between preservation groups using pairwise *t*-tests of mean differences and two-sided 95% confidence intervals (CIs) assuming unequal variance.

In addition, we calculated trophic overlap to consider dietary similarity between taxa in the same or different preservation groups, which reveals aspects of ecological similarity. Species with high trophic overlap are frequently aggregated into trophic species or guilds in ecological network studies. We calculated resource and consumer trophic overlap between all pairs of taxa in a web as the number of shared resources (or consumers) divided by the minimum number of resources (or consumers) held by either of the two taxa. The pairwise resource and consumer overlap values were multiplied together to generate a single combined "trophic overlap" value for each pair of taxa. A pairwise trophic overlap value of 1 indicated that two taxa had exactly the same resources and consumers. To calculate mean trophic overlap within and between preservation groups, we averaged trophic overlap values for all pairs of taxa within a web by intra- or inter-preservation group.

Network-Level Attributes.—We used seven network-level metrics (Table 1) to assess the community structure of the food webs before

and after information loss: mean short-weighted trophic level (mean TL), system omnivory index (SOI), diameter, connectance, mean normalized degree (degree), clustering, and characteristic path length (CPL). Unless otherwise noted, for network-level metrics we analyzed "trophic species" versions of webs, in which taxa with identical resources (prey) and consumers (predators) were aggregated into single trophic species; this is a common practice among ecological network analyses, as it reduces biases generated by uneven sampling and resolution (Williams and Martinez 2000; Dunne et al. 2002a, 2008; Beckerman et al. 2006) (see Supplementary Figs. 2–5 for analyses of *species* webs rather than *trophic species* webs). The concept of trophic species is different from trophic overlap; the former describes the aggregation of taxa with similar dietary habits, whereas the latter quantifies similarity in dietary habits between two taxa.

Comparing Food Webs.—The empirical food webs we use here vary in richness (i.e., number of taxa) and connectance (i.e., density of links)—both of which are known to vary systematically with other network metrics. Thus, to compare metrics across uniquely structured food webs, we normalized webs by comparing them with simulated null model food webs generated with the "niche model" (Williams and Martinez 2000). The niche model uses taxon richness and connectance to stochastically generate simulated food webs with properties similar to real food webs (Dunne et al. 2008). Normalization was achieved by calculating model error (ME) for each network-level metric, where values exceeding ±1 are considered significantly different from the niche model. The ME between the metric value of the target empirical web and that of the null model (in this case replicated niche model–generated webs) is the normalized difference between the null model's median metric value (here based on 1000 replicates) and the target web metric value. ME is normalized to the difference between the model's median metric value and the model's value at the upper or lower bound of the 95% CI, depending on whether the target web metric value falls above or below the model's median metric value (e.g., Dunne et al. 2008).

Quantifying the Effects of Information Loss

To assess the impacts of information loss on whole-network structure, we applied an artificial fossilization procedure to modern webs. The procedure consisted of two variables: a fossilization probability (FP) unique to each taxon and an environmental filter (EF) applied equally across the food web (Fig. 2).

Taxa were assigned fossilization probabilities based upon their preservation groups (Fig. 2B). Compared with previous studies estimating the fossilization potential of well-preserved taxa (Foote and Sepkoski 1999), our methodology did not require knowledge of the fossil record and could be applied across multiple systems and taxonomic groups, including those without fossil data. We utilized a tiered system in which the least fossilizable preservation group (soft-bodied) was assigned an arbitrary FP of 0.25, the intermediate group (weakly biomineralizing) was assigned 0.5, and the most fossilizable (hard-bodied) was assigned 0.75. Given the arbitrary nature of these values and the heterogeneity of fossilization potential among taxa in reality, we also tested alternative FP values, including different static values and variable values drawn from uniform distributions (see Supplementary Text, Supplementary Table 1, Supplementary Figs. 6–17). In both cases, we found no differences in the resulting trends, so we used the abovementioned static values to reduce computational complexity. See the Supplementary Text for a comparison between using preservation groups (e.g., herein) and known fossil occurrences (e.g., Roopnarine and Dineen 2018) as proxies for fossilization potential.

We defined an EF based upon beta distributions (notation = 'beta [alpha parameter, beta parameter]'; Fig. 2C) to mimic randomness in local environmental fossilization potential conditions and to incorporate variation in fossilization potential within preservation groups. The environment most conducive to preservation was modeled using a beta distribution with alpha equal to 0.1 and beta equal to 9.9 (right-skewed distribution), while the least conducive used a distribution with alpha equal to 9.9 and beta equal to 0.1 (left-skewed distribution). In all beta distributions, alpha plus beta equaled 10. For ease of interpretation,
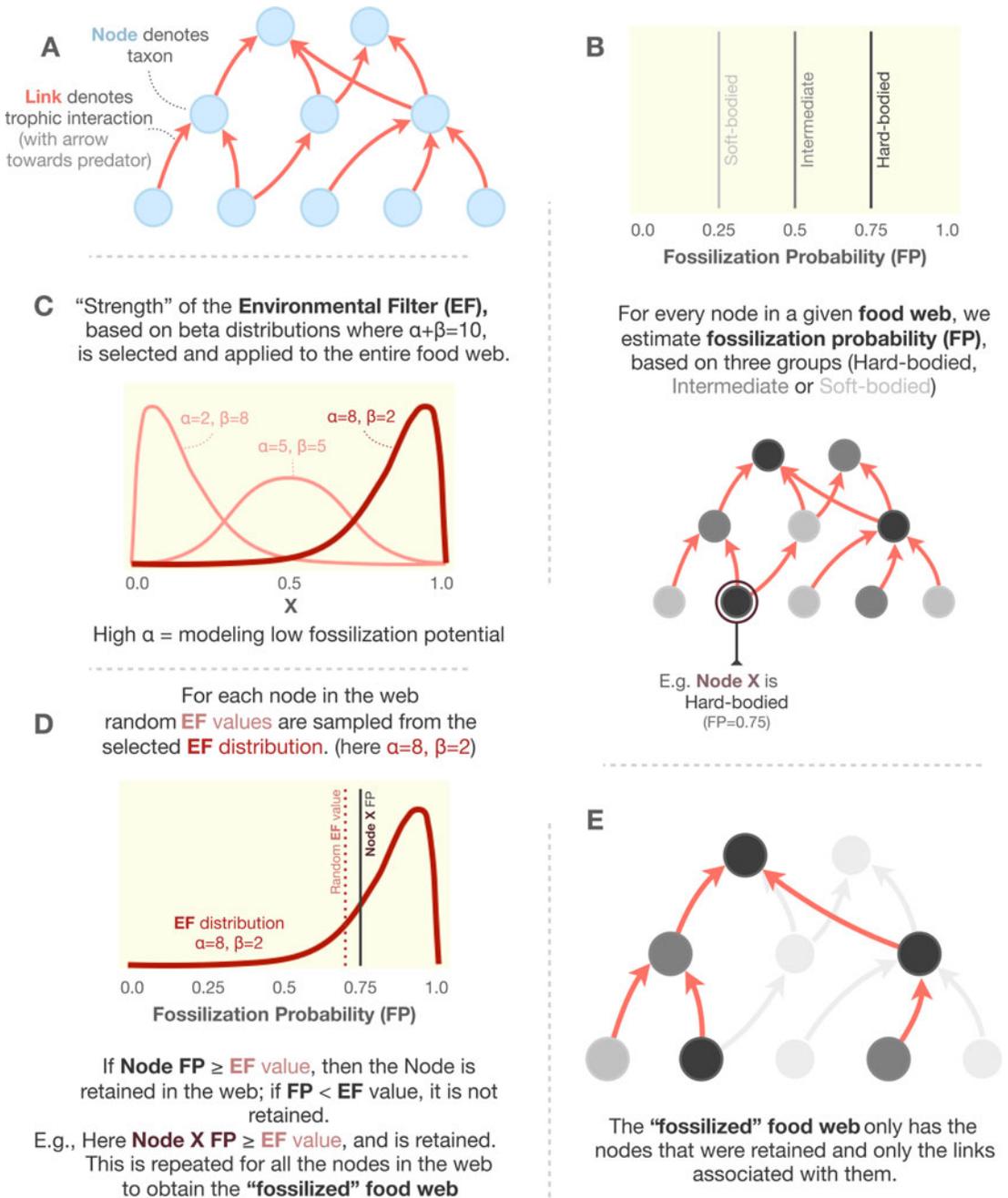
FIGURE 2. Schematic describing the selective information loss procedure. A, How network nodes and links correspond to trophic interactions in food webs; B, how fossilization probabilities were assigned to nodes; C, how environmental filters were defined using beta distributions and how distributions correspond to "environments" with different fossilization potentials; D, how an environmental filter is compared with node fossilization potential values to "artificially fossilize" a food web; E, how a food web looks after selective information loss.

we present data versus percent node loss, rather than the alpha parameter (see Supplementary Fig. 18 for comparisons of the two).

To "fossilize" a taxon within a web, a number was drawn from a specified EF distribution and compared with the taxon's FP (Fig. 2D). If the
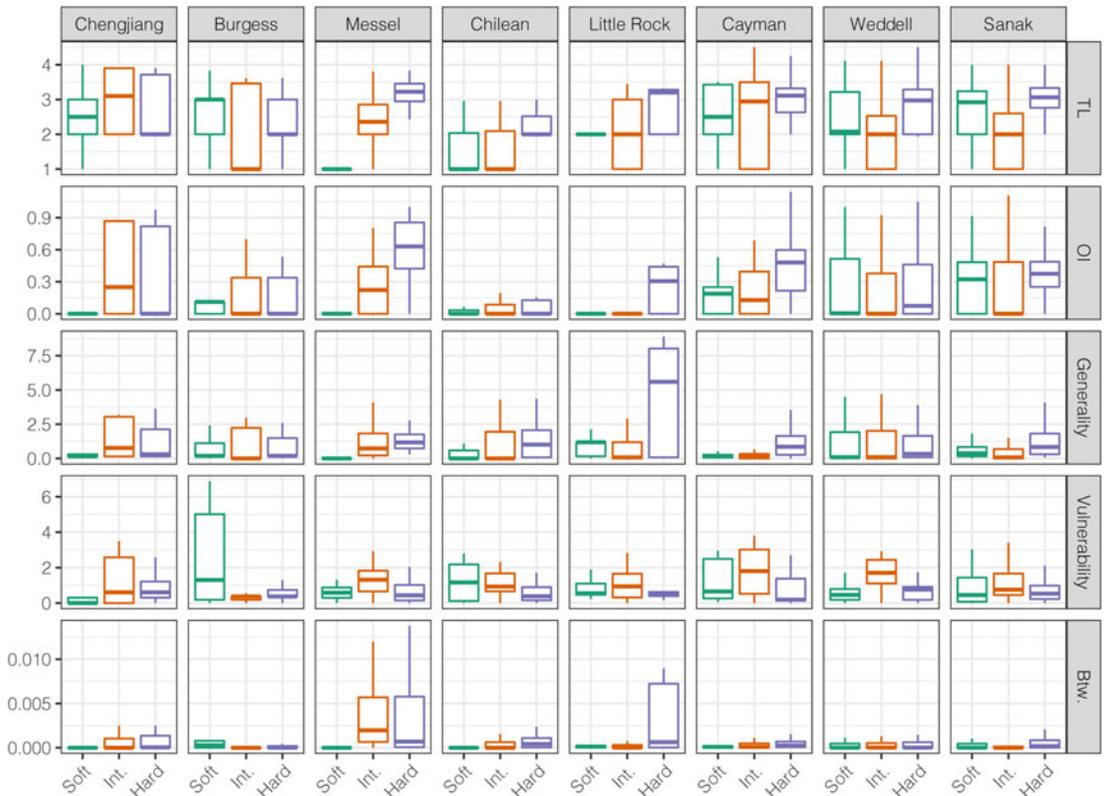
FIGURE 3. Comparison of node-level network metrics for the preservation groups. Outliers removed. TL, trophic level; OI, omnivory index; Btw, normalized betweenness centrality. See Supplementary Fig. 22 for outliers included.

FP value was greater than the drawn EF value, then the taxon "fossilized" and remained part of the food web; otherwise, the taxon was removed. We removed consumers without resources and primary producers without consumers after fossilization to ensure that we were analyzing a connected trophic system without consumers appearing as primary producers. For each instance of fossilizing a food web, a single EF distribution was used (i.e., the alpha and beta parameters of the beta distribution did not vary within a run). We "artificially fossilized" each web 200 times for each 0.1 increment increase of the alpha component of the EF beta distribution (0 < alpha < 10) to show a range of environmental conditions varying in conduciveness to fossilization.

To test the hypothesis that selective loss of poorly preserved taxa results in systematic changes to network structure, we compared the effects of selective and random information loss. Although information loss in the fossil

record is selective, we wanted to determine whether increasing random information loss—acting as a null model—had distinctive effects on network structure, given that this form of loss is easier to quantify. Nodes were removed either (1) selectively—mimicking the selective loss of soft-bodied taxa—by comparing FPs with values drawn from the EF; or (2) randomly, by shuffling FPs among taxa in a web and comparing them with the EF. Conversion to trophic species webs followed node removal. We compare these forms of information loss in two ways: comparisons of selectively and randomly fossilized webs to niche model webs with identical richness and connectance values as the fossilized webs, and direct comparisons between selectively and randomly fossilized webs. Comparisons between the niche model webs and the fossilized webs allowed us to account for the richness and connectance dependence of network metrics when considering the effects of
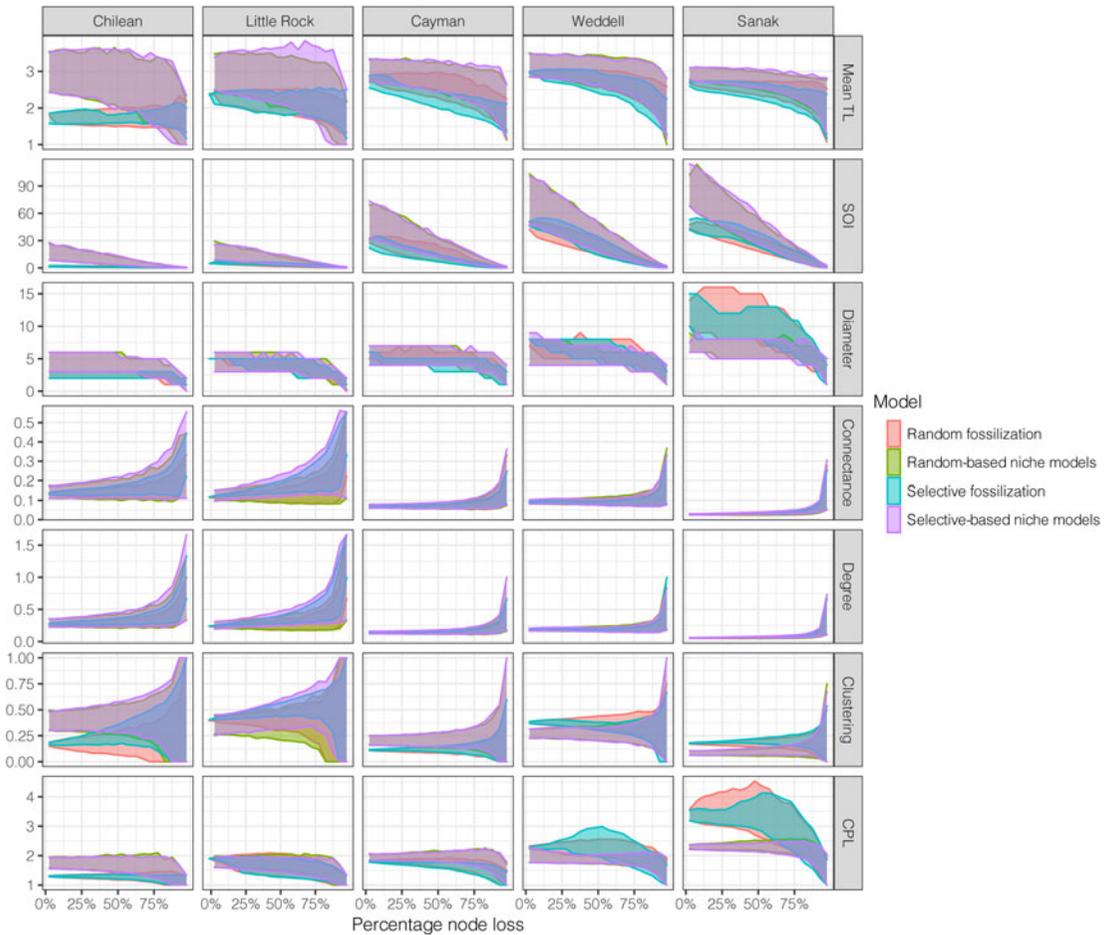
FIGURE 4. Comparison of connected 95% confidence intervals (CIs) of network-level metrics for trophic species webs exposed to fossilization procedures mimicking random and selective information loss. Replicates generated by repeating the fossilization procedure 200 times for each 0.1 increase in the alpha value of the environmental filter beta distribution. CIs calculated in 5% node-loss increment bins. As the "environmental filter" is increased, the percentage of nodes removed increases. Environmental filter ranges from beta [0.1,9.9] to beta [9.9,0.1]. These webs are compared with those generated by a niche model, based on identical taxon richness and connectance values to the webs with which they are being compared. Mean TL, mean short-weighted trophic level; SOI, system omnivory index; CPL, characteristic path length.

information loss across webs with different structures. Direct comparisons between the randomly and selectively fossilized webs allowed us to assess whether the selective information loss based on preservation group has nonrandom effects on the structure of a food web. We quantified differences using ME.

Results

Baseline Food Web Structure and Fossilization Potential

Before modeling information loss, we evaluated the network structure of unaltered webs,

as differences could affect the webs' responses to information loss. The webs presented in this study ranged widely in richness, connectance, percentage of primary producers, and taxonomic resolution (Table 2, Supplementary Figs. 1, 19). However, structure varied predictably with richness and connectance, as described in previous studies (Supplementary Fig. 20), supporting the use of the niche model as a tool for ecological network normalization when comparing structure across webs (Dunne et al. 2013). Network-level structure metrics (Tables 1, 2) and degree distributions (Supplementary Fig. 21) indicated overall
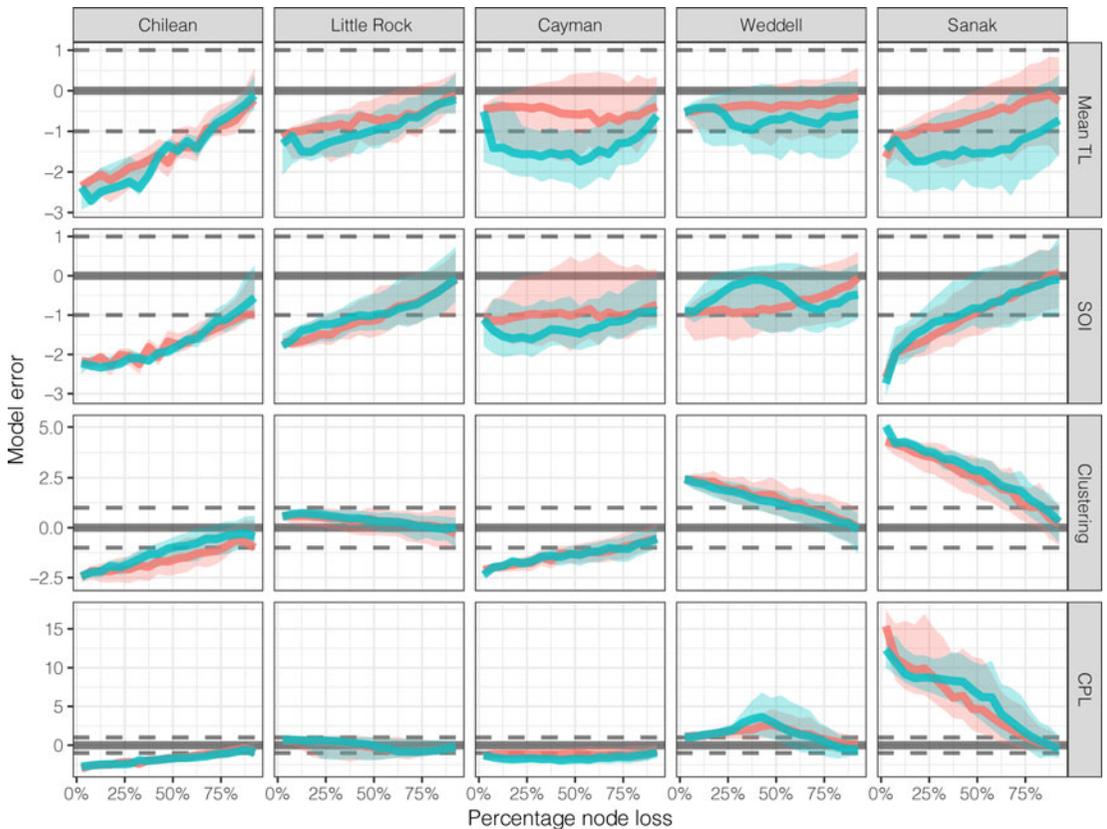
FIGURE 5. Selectively and randomly fossilized trophic species webs compared with those generated by respective niche models (i.e., selectively fossilized webs compared with niche models based on selectively fossilized webs, random compared with niche model webs based on randomly fossilized webs). Plotted as model errors relative to the niche model. Values within ±1 are not significantly different from the niche model. Connected 95% confidence intervals and mean values (calculated in 5% bins) of randomly fossilized webs given in red, and of selectively fossilized webs in blue. Mean TL, mean short-weighted trophic level; SOI, system omnivory index; CPL, characteristic path length.

structural similarity of ancient and modern webs, with the exception of clustering coefficients, for which ancient webs had higher values.

Previous research has shown systematic differences in fossilization potential (e.g., Schopf 1978; Foote and Raup 1996; Foote and Sepkoski 1999), but recent work has started to consider how this scales to and varies at the community level (Shaw et al. 2020). We tested for differences in the baseline fossilization potential of different food webs by estimating community fossilization potential as the number of hard-bodied taxa in a web compared with the total number of taxa (Table 2). Little Rock had the lowest community fossilization potential (6%), Cayman Islands had the highest (81%), and mean community fossilization potential across

the webs was 43%. This estimate of mean community fossilization potential is similar to Shaw et al.'s (2020) observation that an average of 38% of genera in marine communities would be preserved in the fossil record. The Sanak and Weddell webs had taxa more evenly distributed across the three preservation groups than other webs. Differences in baseline community fossilization potential mean that some webs are more susceptible to the removal of poorly preserved taxa than others and are thus likely to contain less information when represented as fossil assemblages.

Ecological Differences between Soft- and Hard-bodied Organisms

If there are ecological differences among preservation groups, fossilization is likely to

unevenly affect certain functional groups. Such uneven preservation will, in turn, systematically bias inferences of community structure. To assess ecological differences, we examined node positions and dietary similarity among taxa within and between preservation groups.

Across modern and ancient webs, hard-bodied and intermediate group taxa exhibited distinct node-level metric distributions, as did soft- and hard-bodied taxa (Fig. 3, Supplementary Figs. 22, 23). Distributions of node-level metrics were more similar between soft-bodied and intermediate group taxa, suggesting similar ecological roles, although TL and vulnerability distributions were distinct between the two groups in all three ancient webs. When compared with the intermediate group, hard-bodied taxa displayed significantly lower vulnerability values (six of eight webs) and significantly greater TL values (seven of eight webs), but not significantly different betweenness, generality, and OI value distributions. When compared with the soft-bodied group, hard-bodied taxa displayed significantly greater generality values (six of eight webs) but similar betweenness and vulnerability value distributions (five of eight webs). Thus, hard-bodied taxa generally display higher OI, higher TL, higher generality, lower vulnerability, and higher betweenness values. These patterns were evident in the Messel Shale web but not in the Burgess Shale or Chengjiang Shale webs.

Trophic overlap was greater between taxa within preservation groups than between taxa in different groups (Supplementary Fig. 24). However, within-group overlap for soft-bodied taxa was relatively lower in ancient webs and the small Chilean web. High within-group trophic overlap means that trophic species will generally be composed of taxa belonging to a single preservation group. Consequently, trophic species in a food web based on shelly fossil deposits will not account for the lack of non-biomineralizing taxa.

Systematic Effects of Random Information Loss on Food Web Structure

Each network-level metric responded to random node loss similarly across all webs (Fig. 4). In general, mean TL, SOI, and CPL decreased with increased random information loss. For these metrics, deviations from the original values were largest when node loss surpassed ~75%. Connectance and clustering showed increased variability with node loss, although the averages remained similar to that of the unaltered web with no node loss.

We then compared the randomly fossilized webs with corresponding niche model webs, which highlighted that structural changes resulting from artificial fossilization were not simply due to the fossilized network being smaller than the original network. Mean TL and SOI of randomly fossilized webs were lower than those of niche model webs at low node-loss values (e.g., when few taxa are removed). As more species were lost from the food web, randomly fossilized webs appeared to be more similarly structured to niche model webs due to increased variance in metrics at higher levels of node loss. The variance of mean TL values in randomly fossilized webs remained mostly stable across increasing node loss. In contrast, the variance of SOI values increased. The mean TL values of randomly fossilized webs mostly fell within ME, except for the Chilean web. The SOI values of these randomly fossilized webs were more distinct from niche webs, except for the Weddell web. The clustering coefficients of randomly fossilized webs were variably greater or lesser than predicted by the niche model—as opposed to being the same—depending on the food web being fossilized, but clustering coefficient values converged upon those of niche model webs with increased node loss as variance increased. CPLs of randomly fossilized webs were also variably greater or less than those for niche model webs, dependent on the food web. Whereas CPLs of the Chilean and Sanak webs responded to node loss similarly to clustering coefficient, other webs showed unique responses.

Systematic Effects of the Removal of Soft-bodied Organisms on Food Web Structure

To further understand the dynamics of information loss, we compared selectively fossilized webs via model errors to (1) niche model webs and to (2) randomly fossilized webs. Comparisons of selectively fossilized food webs with the niche model webs broadly show the same

trends as comparisons between randomly fossilized webs and corresponding niche model webs (Fig. 5). Comparisons between selectively and randomly fossilized webs indicate systematic differences in how metrics responded to information loss. In general, increased selective node loss led to changes in metrics with the same directionality as random node loss (Fig. 4). Mean ME (averaged in 5% node-loss bins) often displayed *S*-shaped trajectories across increasing node loss (Figs. 6, 7). This was caused by the initial selective removal of soft-bodied taxa with a distinct distribution of node attributes, followed by the removal of intermediate group taxa, then hard-bodied taxa as the EF increased. Randomly and selectively fossilized species webs show greater ME values compared with aggregated trophic species webs (Supplementary Figs. 3–5), in line with prior work showing that aggregating to trophic species reduces variability within the data (e.g., Williams and Martinez 2000; Dunne et al. 2002a, 2008; Beckerman et al. 2006).

The responses of metrics to selective information loss were generally within model error of responses to random information loss. However, some metrics and webs showed significant differences (ME > ±1) from randomly fossilized webs, such as clustering in the Little Rock web and mean TL in the Cayman Islands web (Fig. 7). As an indicator of overall difference between randomly and selectively fossilized webs, the mean absolute error (MAE) values were calculated for each variable (Supplementary Table 2). The least sensitive variable was diameter (MAE = 0.44) and the most sensitive was mean TL (MAE = 0.70).

Random information loss alone generates much of the change to apparent network structure, yet selective information loss does have a distinct effect. While MEs comparing random and selectively fossilized webs generally fall within ±1—indicating a lack of statistically significant difference between the structures of the two—the consistency of ME trends across webs, metrics, and increased node loss confirms that selective information loss has distinct effects. Along with ecological differences indicated by consistent node positions and trophic overlap trends, differences in whole-network structure suggest an important effect of the selective removal of soft-bodied taxa.

Discussion

Our analyses of node- and network-level food web features reveal that the selective removal of poorly preserved organisms generates predictable changes to food webs, biasing inferences of community ecology. This suggests that ancient food webs based on typical fossil deposits (i.e., assemblages composed of only biomineralizing taxa) are likely to be missing aspects of trophic function performed by poorly preserved taxa.

Baseline Food Web Structure and Fossilization Potential

Although the food webs considered here vary in structure, age, and aquatic type (i.e., lake versus marine; see Shurin et al. [2006] for a review of differences between terrestrial, marine, and lake food webs), all are comparable after accounting for systematic variations with richness and link distributions, corroborating previous evidence (Dunne et al. 2008, 2013). These similarities permitted comparisons of information loss trends across systems. However, webs varied substantially in community fossilization potential, meaning that the removal of soft-bodied taxa generated more information loss in some systems than others.

Ecological Differences between Soft- and Hard-bodied Organisms

We show that node positions vary systematically among preservation groups—indicating that taxa with differing fossilization potentials tend to play different trophic roles in aquatic communities—such that the selective removal of a group will result in a biased distribution of node positions. Across the modern webs studied, the greatest differences in node positions are between hard-bodied and intermediate group taxa, and the node-level attributes of soft-bodied taxa generally fit between the two groups. Webs in which soft-bodied and intermediate group taxa are selectively lost will overestimate mean TL, SOI, mean generality, and betweenness centrality, but underestimate mean vulnerability.
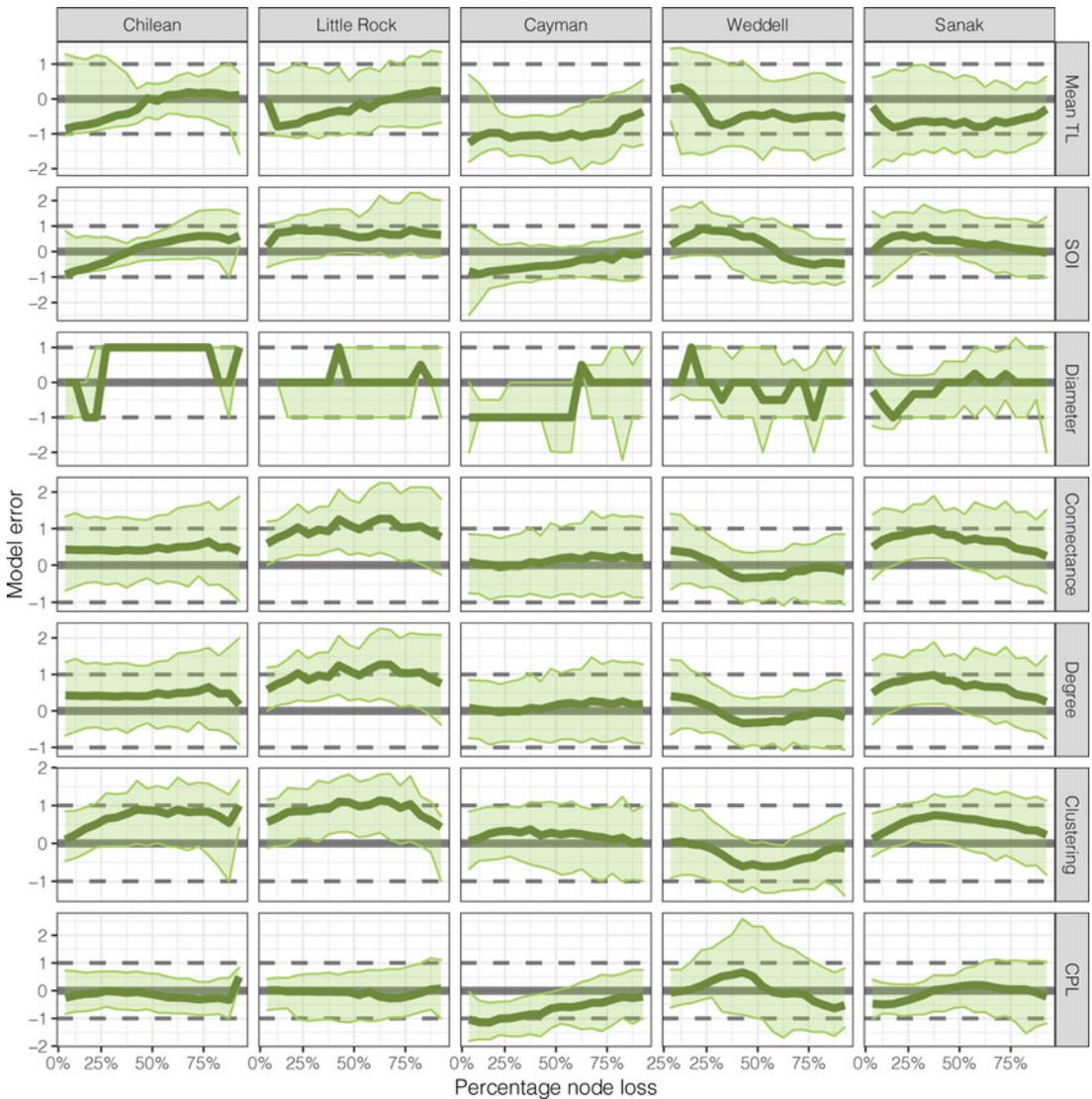
FIGURE 6. Comparison of connected 95% confidence intervals for model error of selectively fossilized trophic species webs relative to the randomly fossilized trophic species webs. Values within ±1 are not significantly different from the niche model. Mean model error indicated by bold line. Mean TL, mean short-weighted trophic level; SOI, system omnivory index; CPL, characteristic path length.

Ecological metrics—both network and node level—have real-world implications for community structure and function, although they are imperfect predictors and must be interpreted with caution. Research shows that trophic cascades can be generated by the removal of high- or low-TL taxa (Thompson et al. 2007), such that the selective removal of poorly preserved taxa—which generally have lower TL values, as shown here and in

Roopnarine and Dineen (2018)—would generate food webs that may be incorrectly interpreted as less stable. Additionally, the higher OI values of hard-bodied taxa are associated with more resistant and resilient communities, given the ability to feed across trophic levels (Christensen and Pauly 1992; Libralato 2008), such that the selective removal of poorly preserved taxa—with lower OI values—would generate networks that are considered more
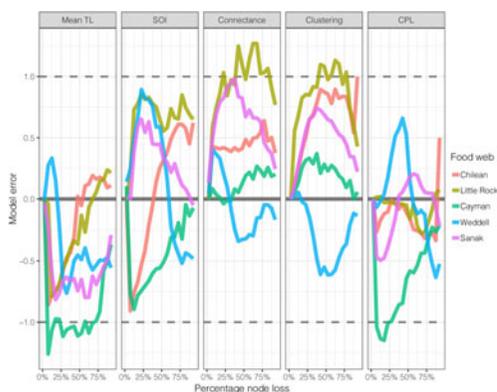
FIGURE 7. Mean model error of selectively fossilized trophic species webs relative to randomly fossilized trophic species webs. Mean values calculated in 5% bins. Mean model error lines are identical to those depicted in Fig. 6, except here multiple webs are overlain to compare the relative impacts of information loss. Mean TL, mean short-weighted trophic level; SOI, system omnivory index; CPL, characteristic path length.

stable. Higher generality (i.e., number of resources) values imply that hard-bodied taxa have a wider range of resources and are more resistant to the loss of individual taxa. Thus, the selective loss of soft-bodied and intermediate group taxa would generate webs dominated by highly generalized consumers. Conversely, higher vulnerability (i.e., number of consumers) values imply that intermediate group taxa are consumed more than both soft- and hard-bodied taxa. Finally, greater betweenness centrality values indicate that hard-bodied taxa take part in more food chains and may exert more control over energy flows. Thus, the selective loss of poorly preserved taxa generates a web that could appear more fragile, given that all remaining taxa would have large control over energy flows.

The lack of significant differences between the node positions of soft- and hard-bodied groups may be the result of distinct life habits. For instance, most soft-bodied taxa in the Weddell web are infaunal worms and thus may be isolated from consumption by highly predatory hard-bodied surficial and pelagic taxa—here, tiering may explain the low vulnerability of soft-bodied taxa compared with the intermediate group.

Notably, consistent differences in node positions between preservation groups are not common in the two Cambrian webs—the Chengjiang Shale and Burgess Shale—but are present in the early Eocene Messel lake web, where network positions of soft-bodied, intermediate group, and hard-bodied taxa are consistent with modern webs. Although hard-bodied metazoan life proliferated during the Cambrian, it diversified dramatically during the Paleozoic (e.g., the great Ordovician biodiversification event) and its greatest relative expansion occurred during the Mesozoic marine revolution (MMR) (Knoll 2003). The MMR—which predates the early Eocene Messel Shale—is synchronous with the appearance of modern ecological structure (Hull 2017) and is thought to have been a response to the diversification of durophagous and boring organisms. Analyses of webs postdating the Cambrian and predating the MMR would help to confirm whether that event—or another diversification event—cemented the ecological differences between soft- and hard-bodied taxa indicated by their unique network positions. Furthermore, reassessments of the fossil webs used here, in light of discoveries of taxa postdating the original publications, may elucidate some of the trends we identified.

Systematic Effects of Information Loss on Food Web Structure

We show that the selective removal of poorly preserved taxa affects whole-network structure, although random information loss accounts for much of the bias. Selective information loss is most distinct from random loss in the cases of mean TL, connectance, and clustering, such that we caution interpretations of these metrics in fossil assemblages without both soft- and hard-bodied diversity. These consistencies indicate that the impacts of information loss could be incorporated into future analyses of ancient food webs. New techniques quantifying sample completeness (Hsieh et al. 2016) and taphonomic bias (Saleh et al. 2020) may permit accurate predictions of the magnitude of change to a network metric caused by taphonomic processes. Furthermore, the integration of data on life habits, differential preservation among environments (Foote et al. 2015; Shaw et al. 2020), differential preservation among specific taxa (Foote and Sepkoski

1999), and how these components covary would illuminate the relative importance of different taphonomic processes biasing community compositions and interpretations of structure.

Conclusions

Fossil data offer tantalizing glimpses into ancient communities and permit consideration of how systems respond to large perturbations. However, the deposits used to describe ancient communities are impacted by taphonomic processes, reducing both the amount and fidelity of information that can be gleaned. As seen with our sample of aquatic food webs, community fossilization potential varies widely (Table 2), but the general effects of information loss based on preservation group are consistent. This suggests that ancient community data can be used to examine ecological network structure across deep time, as long as all preservation groups are considered (e.g., in the case of Lagerstätten) or if biases in metrics are accounted for (as described here). It is also important to remember that ecological network metrics are not perfect predictors of community structure and response—they should be considered in tandem with other features of assemblages, such as taxonomic composition, functional diversity, and population sizes. Future studies could consider the effects of information loss based on other organismal and environmental components of fossilization potential—such as taxon size or habitat type—and how these components interact.

Acknowledgments

References

Albouy, C., P. Archambault, W. Appeltans, M. B. Araújo, D. Beauchesne, K. Cazelles, A. R. Cirtwill, M. J. Fortin, N. Galiana, S. J. Leroux, L. Pellissier, T. Poisot, D. B. Stouffer, S. A. Wood, and D. Gravel. 2019. The marine fish food web is globally connected. Nature Ecology and Evolution 3:1153–1161.

Beckerman, A. P., O. L. Petchey, and P. H. Warren. 2006. Foraging biology predicts food web complexity. Proceedings of the National Academy of Sciences USA 103:13745–13749.

Benton, M. 2014. Vertebrate palaeontology. Wiley, Hoboken, N.J.

Blüthgen, N. 2010. Why network analysis is often disconnected from community ecology: a critique and an ecologist's guide. Basic and Applied Ecology 11:185–195.

Brusca, R. C., and G. J. Brusca. 2003. Invertebrates. Sinauer Associates, Sunderland, Mass.

Cardinale, B., E. Duffy, D. Srivastava, M. Loreau, M. Thomas, and M. Emmerson. 2009. Towards a food web perspective on biodiversity and ecosystem functioning. Pp. 105–120 in S. Naeem, D. E. Bunker, A. Hector, M. Loreau, and C. Perrings, eds. Biodiversity, ecosystem functioning, and human wellbeing: an ecological and economic perspective. Oxford University Press, Oxford, United Kingdom.

Christensen, V., and D. Pauly. 1992. ECOPATH II—a software for balancing steady-state ecosystem models and calculating network characteristics. Ecological Modelling 61:169–185.

Delmas, E., M. Besson, M. H. Brice, L. A. Burkle, G. V. Dalla Riva, M. J. Fortin, D. Gravel, P. R. Guimarães, D. H. Hembry, E. A. Newman, J. M. Olesen, M. M. Pires, J. D. Yeakel, and T. Poisot. 2019. Analysing ecological networks of species interactions. Biological Reviews 94:16–36.

Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002a. Food-web structure and network theory: the role of connectance and size. Proceedings of the National Academy of Sciences USA 99:12917–12922.

Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002b. Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology Letters 5:558–567.

Dunne, J. A., R. J. Williams, N. D. Martinez, R. A. Wood, and D. H. Erwin. 2008. Compilation and network analyses of Cambrian food webs. PLoS Biology 6:e102.

Dunne, J. A., K. D. Lafferty, A. P. Dobson, R. F. Hechinger, A. M. Kuris, N. D. Martinez, J. P. McLaughlin, K. N. Mouritsen, R. Poulin, K. Reise, D. B. Stouffer, D. W. Thieltges, R. J. Williams, and C. D. Zander. 2013. Parasites affect food web structure primarily through increased diversity and complexity. PLoS Biology 11:e1001579.

Dunne, J. A., C. C. Labandeira, and R. J. Williams. 2014. Highly resolved early Eocene food webs show development of modern trophic structure after the end-Cretaceous extinction. Proceedings of the Royal Society of London B 281:20133280.

Dunne, J. A., H. Maschner, M. W. Betts, N. Huntly, R. Russell, R. J. Williams, and S. A. Wood. 2016. The roles and impacts of human hunter-gatherers in North Pacific marine food webs. Scientific Reports 6:21179.

Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.

Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record. Nature 398:415–417.

Foote, M., J. S. Crampton, A. G. Beu, and C. S. Nelson. 2015. Aragonite bias, and lack of bias, in the fossil record: lithological, environmental, and ecological controls. Paleobiology 41:245–265.

Gauzens, B., B. C. Rall, V. Mendonça, C. Vinagre, and U. Brose. 2020. Biodiversity of intertidal food webs in response to warming across latitudes. Nature Climate Change 10:264–269.

Goldwasser, L., and J. Roughgarden. 1997. Sampling effects and the estimation of food-web properties. Ecology 78:41–54.

Guralnick, R., A. Collins, B. Waggoner, B. R. Speer, C. Whitney, and D. Smith. 2000. Tree of Life. University of California Museum of Paleontology. https://ucmp.berkeley.edu, accessed 1 August 2019.

Hsieh, T. C., K. H. Ma, and A. Chao. 2016. iNEXT: interpolation and extrapolation for species diversity, R package version 2.0.8. R-project, https://cran.r-project.org/web/packages/iNEXT/index.html, accessed 1 August 2019.

Hull, P. M. 2017. Emergence of modern marine ecosystems. Current Biology 27:R466–R469.

Jacob, U., A. Thierry, U. Brose, W. E. Arntz, S. Berg, T. Brey, I. Fetzer, T. Jonsson, K. Mintenbeck, C. Möllmann, O. L. Petchey, J. O. Riede, and J. A. Dunne. 2011. The role of body size in complex food webs: a cold case. Advances in Ecological Research 45:181–223.

Kardong, K. 1995. Vertebrates: comparative anatomy, function, evolution. McGraw-Hill, New York.

Kéfi, S., V. Miele, E. A. Wieters, S. A. Navarrete, and E. L. Berlow. 2016. How structured is the entangled bank? The surprisingly simple organization of multiplex ecological networks leads to increased persistence and resilience. PLoS Biology 14:e1002527.

Kidwell, S. M., and K. W. Flessa. 1995. The quality of the fossil record: populations, species, and communities. Annual Review of Ecology and Systematics 26:269–299.

Kidwell, S. M., and S. M. Holland. 2002. The quality of the fossil record: implications for evolutionary analyses. Annual Review of Ecology and Systematics 33:561–588.

Knoll, A. H. 2003. Biomineralization and evolutionary history. Reviews in Mineralogy and Geochemistry 54:329–356.

Lau, M. K., S. R. Borrett, B. Baiser, N. J. Gotelli, and A. M. Ellison. 2017. Ecological network metrics: opportunities for synthesis. Ecosphere 8:e01900.

Libralato, S. 2008. System omnivory index. Pp. 3472–3477 in S. E. Jørgensen and B. D. Fath, eds. Encyclopedia of ecology. Elsevier, Amsterdam.

Martinez, N. D. 1991. Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. Ecological Monographs 61:367–392.

McCann, K. S. 2000. The diversity–stability debate. Nature 405:228–233.

Montoya, J. M., S. L. Pimm, and R. V. Solé. 2006. Ecological networks and their fragility. Nature 442:259–264.

Poisot, T., D. B. Stouffer, and D. Gravel. 2015. Beyond species: why ecological interaction networks vary through space and time. Oikos 124:243–251.

Roopnarine, P. D. 2009. Ecological modeling of paleocommunity food webs. Paleontological Society Papers 15:195–220.

Roopnarine, P. D., and K. D. Angielczyk. 2015. Community stability and selective extinction during the Permian–Triassic mass extinction. Science 350:90–93.

Roopnarine, P. D., and A. A. Dineen. 2018. Coral reefs in crisis: the reliability of deep-time food web reconstructions as analogs for the present. Pp. 105–141 in C. L. Tyler and C. L. Schneider, eds. Marine conservation paleobiology. Springer, Cham, Switzerland.

Roopnarine, P. D., and R. Hertog. 2013. Detailed food web networks of three Greater Antillean coral reef systems: the Cayman Islands, Cuba, and Jamaica. Dataset Papers in Ecology 2013:857470.

Roopnarine, P. D., K. D. Angielczyk, S. C. Wang, and R. Hertog. 2007. Trophic network models explain instability of Early Triassic terrestrial communities. Proceedings of the Royal Society of London B 274:2077–2086.

Roopnarine, P. D., K. D. Angielczyk, A. Weik, and A. Dineen. 2019. Ecological persistence, incumbency and reorganization in the Karoo Basin during the Permian–Triassic transition. Earth-Science Reviews 189:244–263.

Saleh, F., J. B. Antcliffe, B. Lefebvre, B. Pittet, L. Laibl, F. P. Peris, L. Lustri, P. Gueriau, and A. C. Daley. 2020. Taphonomic bias in exceptionally preserved biotas. Earth and Planetary Science Letters 529:115873.

Schmitz, O. J. 1997. Press perturbations and the predictability of ecological interactions in a food web. Ecology 78:55–69.

Schopf, T. J. M. 1978. Fossilization potential of an inter tidal fauna: Friday Harbor, Washington. Paleobiology 4:261–270.

Shaw, J. O., D. E. G. Briggs, and P. M. Hull. 2020. Fossilization potential of marine assemblages and environments. Geology. doi: 10.1130/G47907.1.

Shurin, J. B., D. S. Gruner, and H. Hillebrand. 2006. All wet or dried up? Real differences between aquatic and terrestrial food webs. Proceedings of the Royal Society of London B 273:1–9.

Stouffer, D. B., J. Camacho, W. Jiang, and L. A. Nunes Amaral. 2007. Evidence for the existence of a robust pattern of prey selection in food webs. Proceedings of the Royal Society of London B 274:1931–1940.

Thompson, R. M., M. Hemberg, B. M. Starzomski, and J. B. Shurin. 2007. Trophic levels and trophic tangles: the prevalence of omnivory in real food webs. Ecology 88:612–617.

Thompson, R. M., U. Brose, J. A. Dunne, R. O. Hall, S. Hladyz, R. L. Kitching, N. D. Martinez, H. Rantala, T. N. Romanuk, D. B. Stouffer, and J. M. Tylianakis. 2012. Food webs: reconciling the structure and function of biodiversity. Trends in Ecology and Evolution 27:689–697.

Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Frontiers in Ecology and the Environment 5:475–482.

Williams, R. J., and N. D. Martinez. 2000. Simple rules yield complex food webs. Nature 404:180–183.

Williams, R. J., and N. D. Martinez. 2004. Limits to trophic levels and omnivory in complex food webs: theory and data. American Naturalist 163:458–468.

Williams, R. J., E. L. Berlow, J. A. Dunne, A. L. Barabási, and N. D. Martinez. 2002. Two degrees of separation in complex food webs. Proceedings of the National Academy of Sciences USA 99:12913–12916.

Wood, S. A., R. Russell, D. Hanson, R. J. Williams, and J. A. Dunne. 2015. Effects of spatial scale of sampling on food web structure. Ecology and Evolution 5:3769–3782.