

**SPECIAL ISSUE: SPECIES INTERACTIONS,
ECOLOGICAL NETWORKS AND
COMMUNITY DYNAMICS****Approaches to integrating genetic data into ecological networks**

Elizabeth L. Clare^{1,2}  | Aron J. Fazekas³ | Natalia V. Ivanova² | Robin M. Floyd⁴ |
Paul D. N. Hebert² | Amanda M. Adams⁵ | Juliet Nagel⁶ | Rebecca Girton¹ |
Steven G. Newmaster⁷ | M. Brock Fenton⁸

¹School of Biological and Chemical Sciences, Queen Mary University of London, London, UK

²Centre for Biodiversity Genomics, University of Guelph, Guelph, Ontario, Canada

³The Arboretum, University of Guelph, Guelph, Ontario, Canada

⁴Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK

⁵Department of Biology, Texas A&M University, College Station, Texas

⁶Center for Environmental Science, University of Maryland, Frostburg, Maryland

⁷Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

⁸Department of Biology, University of Western Ontario, London, Ontario, Canada

Correspondence

Elizabeth L. Clare, School of Biological and Chemical Sciences, Queen Mary University of London, London, UK.
Email: e.clare@qmul.ac.uk

Abstract

As molecular tools for assessing trophic interactions become common, research is increasingly focused on the construction of interaction networks. Here, we demonstrate three key methods for incorporating DNA data into network ecology and discuss analytical considerations using a model consisting of plants, insects, bats and their parasites from the Costa Rica dry forest. The simplest method involves the use of Sanger sequencing to acquire long sequences to validate or refine field identifications, for example of bats and their parasites, where one specimen yields one sequence and one identification. This method can be fully quantified and resolved and these data resemble traditional ecological networks. For more complex taxonomic identifications, we target multiple DNA loci, for example from a seed or fruit pulp sample in faeces. These networks are also well resolved but gene targets vary in resolution and quantification is difficult. Finally, for mixed templates such as faecal contents of insectivorous bats, we use DNA metabarcoding targeting two sequence lengths (157 and 407 bp) of one gene region and a MOTU, BLAST and BIN association approach to resolve nodes. This network type is complex to generate and analyse, and we discuss the implications of this type of resolution on network analysis. Using these data, we construct the first molecular-based network of networks containing 3,304 interactions between 762 nodes of eight trophic functions and involving parasitic, mutualistic and predatory interactions. We provide a comparison of the relative strengths and weaknesses of these data types in network ecology.

KEYWORDS

bats, DNA barcoding, food webs, high-throughput sequencing, interaction networks, metabarcoding

1 | INTRODUCTION**1.1 | Ecological networks, DNA and opportunities**

Ecosystem functioning is driven by a network of interactions among species affected by diverse abiotic and biotic variables such as climate, habitat and resource distribution (McCann, 2007) with global economic (Costanza et al., 1997) and conservation (Worm et al.,

2006) impacts. The analysis of interaction networks is of increasing interest across many disciplines, spurring the development of new mathematical and statistical tools (Poisot, Stouffer, & Kéfi, 2016). In ecology, visual representations provide a synoptic view of complex interactions and are primarily displayed as bipartite networks, where trophic levels are depicted as layers (upper and lower) composed of species as nodes connected by links representing interactions

(Dormann, Fründ, Blüthgen, & Gruber, 2008). When multiple networks are combined, it is possible to conceptualize multiple trophic levels simultaneously (e.g., Pocock, Evans, & Memmott, 2012) clarifying ecosystem assembly and structure (Milo et al., 2002), functional roles, and mechanisms of stability (McCann, 2000; Thébault & Fontaine, 2010). Comparisons between networks can assess natural or anthropogenic impacts (McCann, 2007), the evolution of networks (Guimarães et al., 2011; Nuismer, Jordano, & Bascompte, 2013) and the role and response of specific nodes (Martín González, Dalsgaard, & Olesen, 2010; McDonald-Madden et al., 2016; Strona & Lafferty, 2016).

Many networks are incredibly complex with multiple trophic levels and high taxonomic diversity (e.g., Pocock et al., 2012) and are therefore time-consuming to construct, often requiring years of ecological observations and considerable taxonomic expertise (Evans, Kitson, Lunt, Straw, & Pocock, 2016). Consequently, they are not readily scalable to rapid biomonitoring or, if they can scale, they routinely suffer from problems of network completeness and poor or uneven resolution of taxa (Hemprich-Bennett, Oliveira, Le Comber, Rossiter, & Clare, 2018; Ings et al., 2009). Incorporating dietary tracers such as fatty acids, isotopes and genetic tools such as DNA sequencing is a growing trend for measuring species interactions, though each has advantages and disadvantages (reviewed in Nielsen, Clare, Hayden, Brett, & Kratina, 2018). Genetic analyses are expanding at a remarkable rate and have evolved from enzyme-linked immunosorbent assay (ELISA) and targeted species detection (Symondson, 2002) to the use of high-throughput sequencing (HTS) for the analysis of target markers or “metabarcoding” (reviewed in Pompanon et al., 2012). While these techniques are quickly becoming common for the dietary analysis of single species with many proposed applications (Clare, 2014), they have not been widely incorporated into network analysis (but see González-Varo, Arroyo, & Jordano, 2014; Hemprich-Bennett et al., 2018; Wirta et al., 2014) despite strong advocates (Evans et al., 2016; Roslin & Majaneva, 2016).

Many reviews, authors and developers of these techniques have discussed the challenges in DNA-based analyses of species interactions including the impacts of primer choice on taxonomic coverage and resolution, the completeness of reference databases (Pompanon et al., 2012), bioinformatics methods (Clare, Chain, Littlefair, & Cristescu, 2016) and the role of quantification (Deagle et al., 2018) but the specific implications for constructing networks vary with data type. In traditional DNA barcoding, a specimen's ID is delimited by generating one sequence per specimen and comparing it to a reference dataset to confirm its identity. These data are not fundamentally different from traditional observations for generating an interaction matrix. However, at the other end of the complexity spectrum, metabarcoding represents a novel data type for network ecology. HTS generates millions of sequences from each sample of mixed template. While the process can uniquely deal with otherwise intractable sources such as trace material and liquid feeders, it poses novel problems for ecological analysis and network ecology. First, the data require complex bioinformatics handling to remove

unwanted (often error prone) data, but in many cases the impacts of these parameter choices on ecological analysis are unknown (Clare et al., 2016). Second, the ability to quantify DNA within a sample is highly controversial and while, in some cases, proportions of recovered sequence correspond to biomass, in other cases the connection is not clear (Deagle et al., 2018; Deagle, Thomas, Shaffer, Trites, & Jarman, 2013; Thomas, Deagle, Eveson, Harsch, & Trites, 2016). Finally, in an ideal situation, the recovered sequences are matched to a complete reference dataset to identify taxa, but in most cases the reference library is incomplete. In these cases, either an incomplete network is created biased towards the contents of the reference collection (often larger, more charismatic or economically important taxa) or the recovered pool of DNA is converted into molecular operational taxonomic units—MOTUs (Floyd, Abebe, Papert, & Blaxter, 2002)—which are best viewed as pools of equivalent genetic diversity that can be compared, whether we know their identity or not (Clare et al., 2016; Floyd et al., 2002). In this case, each MOTU becomes a node in the network and this level of the network is entirely resolved to a common point of reference with both known and unknown items included, a distinct advantage when mixed resolution presents an analytical problem (Hemprich-Bennett et al., 2018; Ings et al., 2009). However, the level of this resolution is arbitrarily defined by the bioinformatics assessment (see an analysis of parameters for MOTU definition in ecological analysis (Clare et al., 2016; Flynn, Brown, Chain, MacIsaac, & Cristescu, 2015) which may collapse trophic levels and thus generate fundamentally different structures. This is of importance when selecting what metrics can or should be measured. Network metrics can be divided into several broad classes. Network level metrics (e.g., connectance, nestedness, generality) are measured across the entire network. Node level metrics (e.g., centrality, species strength, partner diversity) are specific to the interactions of a given node. Motif measurements are subnetworks of a particular pattern which may define specific ecological interaction types or functions (Milo et al., 2002). Each metric type needs to be considered separately in light of the new data type. Many have concluded that the molecular approach will transform the discipline of ecological biomonitoring and ecological network analysis permitting rapid consistent assessments in systems that are otherwise intractable (Gibson et al., 2015; Roslin & Majaneva, 2016; Toju, 2015; Wirta et al., 2014) while others have advocated adoption but raised serious analytical concerns (e.g., Evans et al., 2016). How then might we proceed?

Here, we address this challenge by focussing on a single complex assemblage of interacting species to demonstrate three approaches to the use of DNA data to resolve interactions and measure several network and node level metrics. While these are not without controversy, our objective is to present an example of methods of data integration into a “network of networks” and we include the most commonly analysed interaction types (antagonistic, mutualistic, parasitic) and the three key methods that have been discussed for DNA and network integration. First, we use standard single-gene DNA barcoding to resolve taxonomy in cryptic organisms and to validate field identifications. Second, we use multigene DNA barcoding

to resolve more complex taxa and single-sourced trace material that cannot be identified by morphological methods. Third, we use metabarcoding to resolve mixed material and then discuss the advantages and challenges of applying these approaches. While these have been used previously, our analysis provides the first example of integrating these data types to form a multitrophic level assemblage resolved entirely with DNA and the first to contrast these data. We hope to provoke discussion about the appropriate use of these data types.

2 | MATERIALS AND METHODS

2.1 | A case study from Costa Rica: Plants, bats, insects and parasites

From May to July 2009, a field team visited Sector Santa Rosa of the Area de Conservación Guanacaste (ACG). The present analysis relies on material collected during that period and a preliminary ecological analysis of this case is presented (Box 1). A total of 801 bats were captured representing 26 species morphologically identified using available field keys and checklists (Reid, 2009; Simmons, 2005). From these individuals, we analysed 466 parasites that were sampled from 18 host species and 260 guano samples from 21 species of which visual inspection led to 132 samples being classified as containing plant materials (seeds or fruit pulp) and the rest insect material. Some species are integrated into all trophic levels while others are only loosely associated, for example the sanguivore *Desmodus rotundus* did not produce a faecal sample so was retained in the dataset as a parasite host only.

2.2 | Method one: Sanger sequencing to resolve species ID of bats and parasites

The simplest way to integrate DNA data involves the use of Sanger sequencing to clarify species boundaries or to confirm and improve upon identifications made in the field. For bats and parasites, we targeted the 5' end of the mitochondrial cytochrome *c* oxidase subunit 1 gene (COI) as described by Hebert, Ratnasingham, and DeWaard (2003) using full length (658 bp) DNA barcodes which provide taxonomic discrimination for most animal groups (Hebert, Cywinska, Ball, & DeWaard, 2003). For bats, we used small tissue fragments from each individual captured and the "routine" method of DNA barcoding described in Ivanova, Clare, and Borisenko (2012) and the mammal primer cocktail, PCR reagent mix and the thermocycler conditions "MamCOI" described in Supporting Information Tables S1, S2 and S3 of that publication. We edited sequences in CodonCode Aligner (CodonCode Corporation, Centerville, MA, USA) and compared the resulting DNA barcodes to existing reference databases (Clare, Lim, Fenton, & Hebert, 2011) using a neighbour-joining (NJ) tree to confirm they clustered with other representatives of their species assignment based on morphological inspection in the field (Supporting Information Figure S1). Sequences, collection information and primer

names are available in the Barcode of Life Data System (BOLD; www.barcodinglife.org) (Ratnasingham & Hebert, 2007) project BCCR for each recovered sequence.

For parasitic flies and mites, we extracted DNA from whole specimens using voucher-retention procedures (Porco, Rougerie, Deharveng, & Hebert, 2010). Our subsequent PCRs used a variety of primer combinations which are associated with individual records available in the project BCPB available in the BOLD website with corresponding primer sequences online at http://www.boldsystems.org/index.php/Public_Primer_PrimerSearch. Our PCR protocols followed Hebert et al., (2008) with postsequence analysis employing CodonCode Aligner (CodonCode Corporation). Unlike bats, field taxonomic designations were minimal and a full reference database of voucher-linked barcodes was not available. As a consequence, we employed the Barcode Identification Number (BIN; Ratnasingham & Hebert, 2013) method of delimiting MOTUs in BOLD to identify species and compared this to terminal clusters in an NJ tree generated in BOLD. Three clusters were unassigned to any BIN because their sequence lengths were insufficient to provide a sequence match with high confidence; therefore, we designate these as taxa based on reciprocal monophyly of their sequences in an NJ tree (Supporting Information Figure S2).

2.3 | Method two: Sanger sequencing with multiple targets

A more complex problem involves the analysis of material from one source when that material is degraded, making DNA analysis a preferred option, but where the taxa involved are difficult to resolve using this approach. In this case, the seeds defecated by bats may be identifiable from morphology, but fragmented seeds and digested fruit pulp are almost never identifiable morphologically. Consequently, plants whose seeds are too large to be swallowed are often excluded from food webs and dietary analyses unless direct observation confirms their consumption. Plants represent an additional hurdle as species delimitation by DNA often requires multiple genetic markers (CBOL plant working group, 2009).

For all guano samples containing seeds, we separated three to five intact morphologically identical seeds from each sample. For samples containing only pulp or pollen, we separated approximately 10 mg of dried guano for DNA extraction. DNA isolation employed the NucleoSpin® 96 Plant II DNA isolation kit (Macherey-Nagel) following the manufacturer's protocol with an extension of lysis to 1 hr.

We amplified the *rbcL* and *trnH-psbA* regions using primers *rbcLa_F/rbcLajf_634R* and *trnH/psbA* (Fazekas et al., 2008; Kress, Wurdack, Zimmer, Weigt, & Janzen, 2005). We amplified *matK* using primers *1R_KIM/3F_KIM* (Fazekas et al., 2008) and repeated the PCR for failed reactions using alternate primers: *XF/3F_KIM* (Fazekas et al., 2008; Ford et al., 2009). PCRs were carried out in 10 µl volumes containing 2 µl of 5× Phire® reaction buffer (Finnzymes), 0.05 µl of 10 mM dNTPs, 0.1 µl of each 10 µM primer and 0.125 µl of Phire® Hot Start II polymerase (Finnzymes) using the following protocol: initial denaturation at 98°C for 90 s, 35 cycles of 98°C for 5 s, 55–66°C

Box 1 A preliminary analysis of a tropical bat community

Field Methods

All materials were acquired from past research at the field location and/or held in personal collections. All bats were caught over a six-week period from late May – early July of 2009 using mistnets or harptraps in Sector Santa Rosa of the Area de Conservación Guanacaste, Costa Rica. Net locations were alternated nightly between the Bosque Humeda, La Casona and the Picnic area with an extra netting night at the Playa Naranjo targeting *Noctilio*. Each bat was identified and placed in a cloth bag for approximately 1 hr. Any guano produced was collected for taxonomic identification of prey items and the bats were released at the point of capture. Ectoparasites and wing biopsies were preserved in isopropyl alcohol; guano was frozen. Morphological identification of the ectoparasites to fly or mite was made in order to separate functional groups. Guano samples were screened for seed fragments and insect remains, and classified as containing plant material or insect material. Two genera, *Glossophaga* and *Micronycteris*, were assigned provisional species-level ID. *Artibeus lituratus* and *A. intermedius* could not be distinguished in the field and are referred to as *A. sp.*

Network Analysis

We visualized the interaction networks using Bipartite (Dormann et al., 2008) and Cheddar (Hudson et al., 2013) as implemented in R (R Development Core Team, 2015). We compared the structural metrics (links per species, asymmetry, connectance, nestedness, generality and vulnerability) of each traditional bipartite subnetwork (bat–parasite, bat–plant, bat–insect). We evaluated the robustness of each network and modelled the effects of species loss and restoration within the networks. We employed three extinction models: species removed randomly (null model), species removed from most to least connected (Rd–worst-case scenario; Kaiser-Bunbury et al., 2010) and species removed from least to most frequently detected (Ra–best case scenario). Species lose connections within the network when their hosts, prey, predators or mutualists are eliminated. From each of our three component networks (parasitism, mutualism and predation), we measured network robustness (Kaiser-Bunbury et al., 2010; Memmott et al., 2004). We then modelled a restoration scenario where bat species are re-introduced from greatest to least connected (best case) and assessed the proportion of links restored to the structure. To pinpoint possible keystone species, we examined the role of each bat species within the entire network of network using betweenness and closeness centrality scores (Martín González et al., 2010) in igraph (Csardi & Nepusz, 2006). For simplicity, when individual networks are depicted, we present bats on the top rather than arranging these by trophic level (bats occupy multiple trophic levels making any other display exceedingly complex). Finally, we analysed the impact of OTU clustering thresholds of insects in the bat–prey network on the measurement of these metrics considering clustering thresholds from 90% to 99%.

A Network of Networks

Using these data, we present the first “network of networks” where all underlying data have been generated using a molecular approach (Figure 1a). We evaluated the structural metrics (Table 1) and robustness (Figure 2) of traditional subnetworks (Figures 1b–d) and modelled the effects of bat species loss on parasite persistence (Figure 2a), plant mutualism (Figure 2b), predation (Figure 2c) and secondary extinction of bats from prey loss (Figure 2d). Under all models, parasite networks were less robust ($R_a = 0.69/R_d = 0.36$) and mutualistic relationships were only slightly more robust ($R_a = 0.74/R_d = 0.4$). However, arthropods responded differently to models of extinction: a high proportion of prey face predation, even when the highest-ranking bat species by abundance are eliminated ($R_a = 0.86$), while, conversely, arthropods experience a tremendous release from predation when bat species go extinct by connectance ($R_d = 0.28$). Insectivorous bats appear robust to the loss of prey species ($R_a = 0.998/R_d = 0.85$). Even under the worst-case scenario, the first bat species is not lost until 32% of arthropods are extinct, and even when >90% of arthropods are lost, >70% of bat species remain in the network if prey biomass was sufficient (Figure 2d). Only *G. soricina* showed significant trophic flexibility operating in both a mutualistic and strong predatory role (high centrality scores, Supporting Information Table S7). This is also evident in our restoration ecology model (Figure 3) where the third bat returned based on connectance is *G. soricina* introducing parasites, insects and plants at the same time.

A Snap Shot of a Bat Community

Even considering the variability of metrics across multiple MOTU resolutions (Figure 4), the generality of bat–prey networks is extreme compared to the bat–parasite and bat–plant networks. This significantly impacts on our understanding of robustness in this system and may provide evidence in the diversity vs. stability debate (McCann, 2000). The data suggest extraordinary behavioural flexibility of insectivorous bats and their lack of reliance on specific prey. While there is evidence for resource specialization (e.g., the preference for beetles in *Eptesicus* (Clare, Symondson, & Fenton, 2014) or moths in sibling rhinolophids (Arrizabalaga-Escudero et al., 2018), most studies that have employed molecular techniques have observed very generalist flexible behaviour in bat foraging (Salinas-Ramos et al., 2015; Sedlock,

Box 1 (Continued)

Krüger, & Clare, 2014) though none have examined a community on this scale. Second, perhaps the most interesting observation is the position of the bat *Glossophaga soricina* in the network. Clare, Goerlitz, et al. (2014) used a molecular dietary analysis to identify a novel hunting strategy that permits this supposed “nectar bat” to sneak up on insects. The bats’ echolocation calls are low enough in intensity that prey with ears do not detect the approaching threat in time to evade it (Clare, Goerlitz, et al., 2014). Our network analysis suggests that insectivory in *Glossophaga* is not a rare behaviour but rather, during the period of this study, *G. soricina* was the third best-connected insectivore in the community in addition to its role in pollination and, seed dispersal and as a parasite host. Its diverse functional roles make it a probable keystone species with very high betweenness and closeness centrality (Supporting Information Table S7). This distinguishes it as the only bat occupying all these functional roles in the network and thus a species of special conservation interest. In contrast, some species are only very tangentially associated with this network. For example, *Desmodus rotundus*, the vampire bat, is connected to only one parasite and thus forms its own module of interactions unconnected to the rest of the community.

for 10 s (depending on primer set), 72°C for 7–10 s (depending on region), followed by a final extension at 72°C for 60 s and hold at 4°C (see primer references for additional details).

We sequenced each amplicon bidirectionally with the same primers used for amplification in 11 µl reaction volumes containing 0.5 µl of BigDye terminator mix (ABI), 2 µl of 5× sequencing buffer, 1 µl 10 µM primer and 0.5 µl of undiluted PCR product using the following protocol: initial denaturation at 96°C for 2 min, 30 cycles of 96°C for 30 s, 55°C for 15 s, 60°C for 4 min, followed by a hold at 4°C.

We assembled contigs and edited all sequences using SEQUENCHER 4.8 (Gene Codes Corp, Ann Arbor, MI, USA). We then ascertained the percentage similarity of all recovered sequences to available reference sequences in GenBank and BOLD, with the exception of the *trnH-psbA* region which was not searchable within BOLD.

Identification to known taxa is more complex as the different regions provide resolution at different taxonomic depths in different taxa. For example, *rbcl* typically provides generic level resolution (CBOL plant working group, 2009) (occasionally to species level), whereas the *matK* and *trnH-psbA* regions can provide resolution to species in ~60%–90% of cases (depending on the taxa and geographic scope) (Braukmann, Kuzmina, Sills, Zakharov, & Hebert, 2017; Burgess et al., 2011; Lahaye et al., 2008). Due to incompleteness of the reference sequence databases for the flora, many sequences did not show 100% identity to any species in the reference database. We therefore assigned sequences to family, genus or species depending on the region and per cent identity using the following criteria. For *rbcl*, sequence matches with 99.75%–100% identity were assigned to a genus, while matches with 99%–99.75% identity were only placed to a family. For *matK*, matches with 100% identity were assigned to a species or a species cluster when more than one species in the reference set matched with 100% identity; matches with 99.0%–99.9% identity were assigned to genus, while matches with 98%–99% identity were only assigned to a family. For the *trnH-psbA* region, most matched sequences ranged from 98% to 99% identity (no 100% matches were observed). The variable length of the region, the presence of repeated sequence motifs and the

small number of reference sequences complicated the interpretation of BLAST analysis with the GenBank dataset so most assignments were only made to a genus. For two genera, however, the *trnH-psbA* data corroborated the *matK* designation and enabled an increased level of resolution. Unique sequences for these samples were therefore designated with a number (in addition to genus) and treated as putative species. Species-level designation was also accepted for sequences that matched a monotypic genus, and where sequences matched a genus of three species, two of which occur outside the study area.

2.4 | Method three: Metabarcoding of mixed unknowns

When the material to be analysed is a mixed sample of unknown taxa (in this case arthropods), the entire assemblage must be targeted, followed by the use of bioinformatics tools to process the sequences and ascertain the number of taxa in each sample. In this case, we used DNA metabarcoding that targeted two segments of the COI DNA barcode region and processed these data using a series of bioinformatics tools in a well-established analytical pipeline (e.g., Alberdi, Aizpurua, Gilbert, & Bohmann, 2018; Clare, Symondson, et al., 2014; Salinas-Ramos, Herrera Montalvo, León-Regagnon, Arrizabalaga-Escudero, & Clare, 2015). In brief, DNA was extracted using Qiagen Stool mini-kits (Qiagen, CA, USA) with modifications (Clare, Goerlitz, et al., 2014; Zeale, Butlin, Barker, Lees, & Jones, 2011) and eluted in 35 µl of molecular grade water. We targeted 157 and 407 bp amplicons of the DNA barcode region. PCRs were conducted in 20 µl reactions containing 10 µl of Qiagen multiplex master mix (Qiagen), 6 µl of water, 1 µl of each 10 µM primer and 2 µl of DNA. PCRs were as follows: 95°C – 15 min; 50 cycles of 95°C – 30 s, 52°C – 30 s, 72°C – 30 s (1 min for the 407 bp region); 72°C – 10 min. Amplicons were visualized on 96-well 2% agarose precast E-gels (Invitrogen, Life Technologies). For the 157 bp region, we used the Zeale primers (Zeale et al., 2011) which do not amplify bat DNA well, modified with the two adaptor molecular identification tags (MIDs) system to identify individual samples (Clare, Symondson, et al., 2014) without pooling. For the 407 bp region,

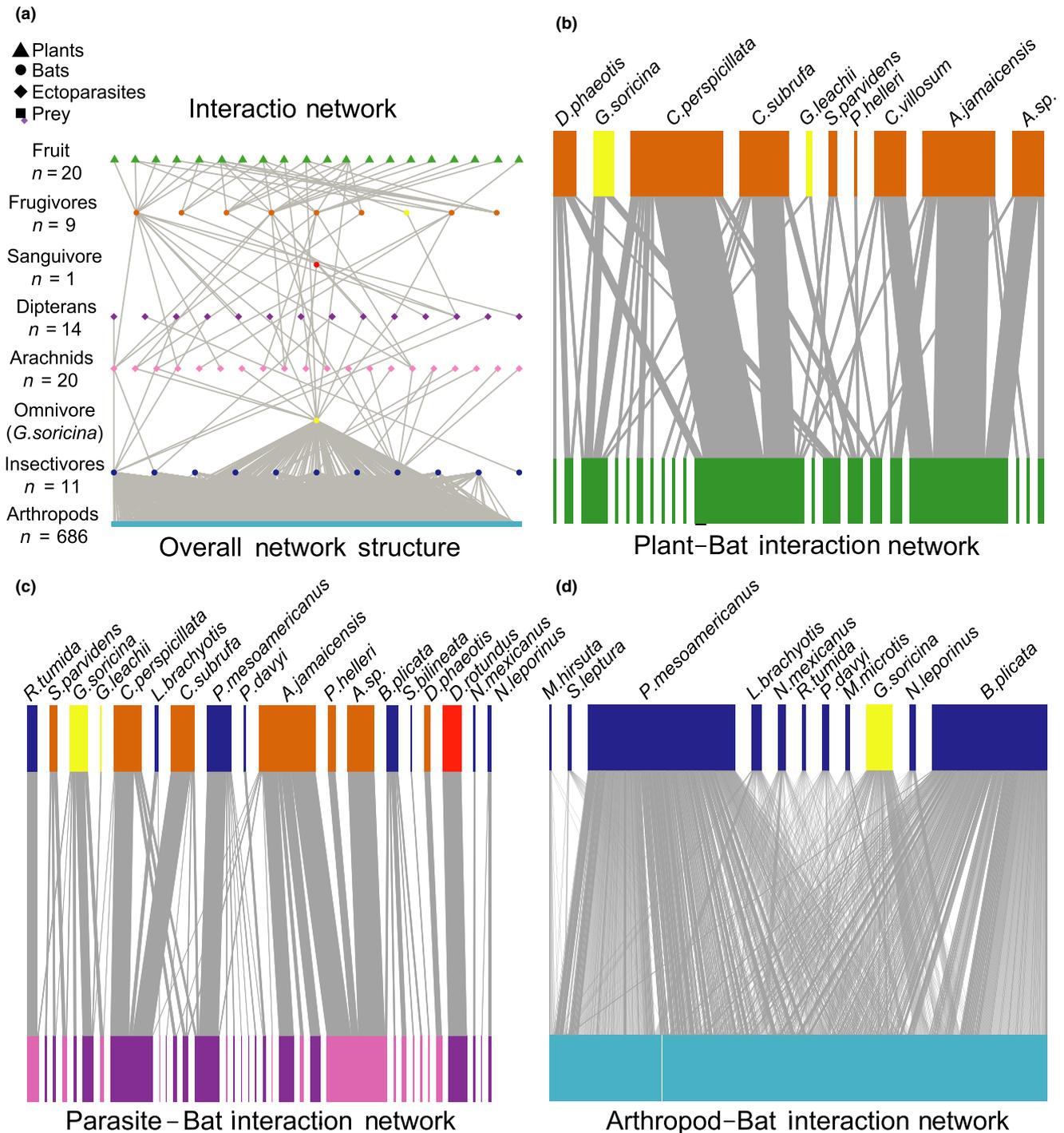


FIGURE 1 Species' interaction networks. The network of networks (a) displays interaction structure organized by behavioural ecology (rather than traditional trophic structure). The visualization of this network is not presented as standard trophic levels for two reasons. First, the arthropod prey represents multiple trophic levels themselves which cannot be differentiated. Second, the density of connections makes links to plants impossible to distinguish if the plants are presented as the lowest trophic level. In this case, the network has been structured to depict function rather than trophic levels. For example, arachnid mites of bats are parasites that spend their entire life cycle on their host (Christe, Arlettaz, & Vogel, 2000) which restricts their dispersal so horizontal transmission primarily occurs via host-to-host contact. Therefore, mites and their hosts are usually regarded as the product of long co-evolution. In comparison, parasitic Diptera (flies) can be highly mobile and often spend part of their life cycle apart from their host (Fritz, 1983). Because of such distinct life history differences, alternative hypotheses of function can be advanced. We depict them as separate functional groups (a) and in their traditional parasite role (c). *N*-values represent the number of taxa detected. Semi-quantified individual trophic networks (b–d) display traditional trophic organizations (though for simplicity of comparison bats are always presented on top). Detection frequency data for each species is given by the width of the block proportional to species' frequency in the network. Colours indicate behavioural role from (a). See Supporting Information for a discussion of visualization orientation and Supporting Information Tables S4 and S8 for matrices of parasite, plant and bat taxonomic identifications

TABLE 1 Structure of the subnetworks

	Links per species	Asymmetry (A)	Connectance (C)	Nestedness (N)	Generality (G) Ψ	Vulnerability (V) Ψ
Bat-Parasite	1.02	-0.31 ^a	0.09	12.60	1.95 ^a	1.32 ^a
Bat-Plant	1.37	-0.33	0.21	26.25	2.74	1.80
Bat-Prey	1.70	-0.97	0.16	14.11	76.72	1.52
Bat-BIN	1.17	-0.91	0.12	14.79	41.25	1.37

Note. Ψ unweighted but see Supporting Information for the appropriateness of unweighted measures.

^aSee Supporting Information for an interpretation of positive vs. negative values and structural arrangement.

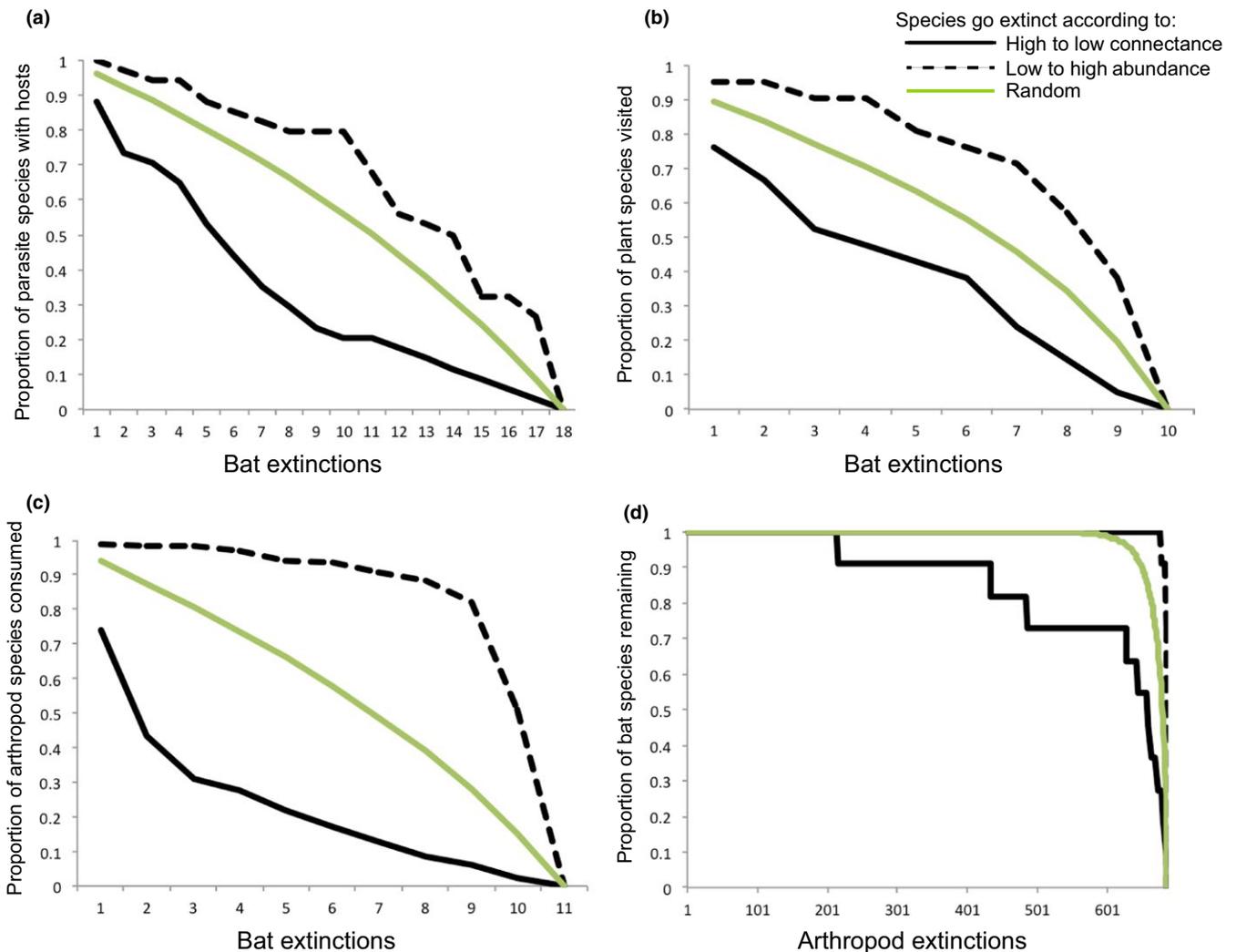


FIGURE 2 The robustness of interaction network structure to the sequential removal of species under three extinction models. The number of bat species removed has an extreme impact on the loss of parasites (a), while plants are slightly more resilient (b). The proportion of arthropods released from predation (c) is strongly dependent on the model of extinction, while insectivorous bats are extremely resilient to the loss of prey under any model (d)

we used primers MLepF1 (GCTTTCCACGAATAAATAATA) and RonMWASPdeg (GGWTCWCCWGATATAKCWTTTCC) combined in equal quantities with LepR1 (TAAACTTCTGGATGTCCAAAAAATCA) and HCO2198 (TAAACTTCAGGGTGACCAAAAAATCA). Sequence recovery is predicted to be lower with longer amplicons due to DNA degradation in digested material but longer reads maximize

identification. For this region, we extracted and PCR-amplified all samples independently but unlike the Zeale region we did not use MID. This does not impact MOTU estimates, but we cannot assign individual sequences to their source bat so they were analysed as a pool and we do not generate networks from these data, just compare MOTU estimation from alternative regions.

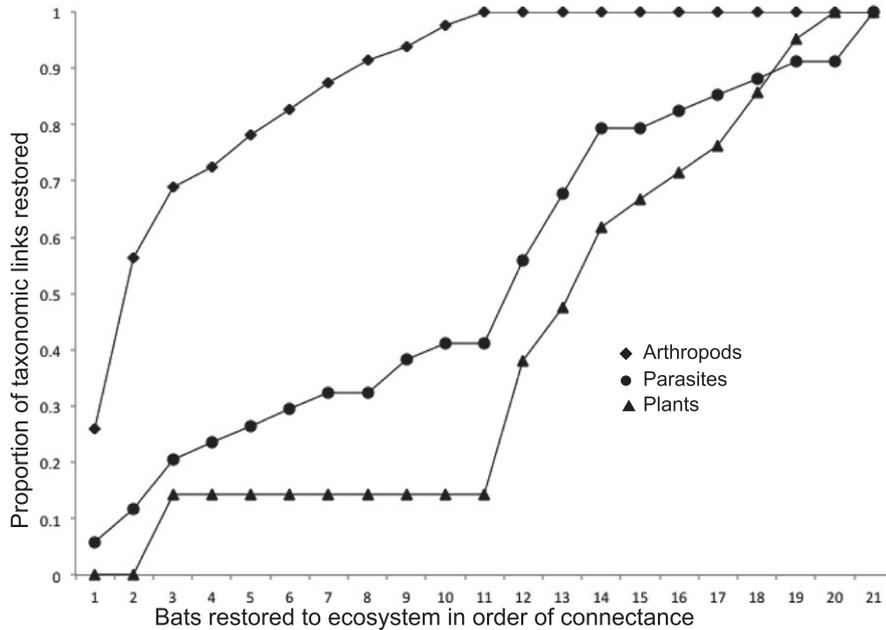


FIGURE 3 A restoration ecology model showing the proportion of links restored if bats are introduced to the ecosystem in order of connectance (best case scenario). With the restoration of only the three most strongly connected species (*Pteronotus mesoamericanus*, *Balantiopteryx plicata* and *Glossophaga soricina*), 72% of arthropod species are under predation, 24% of parasite species have a host, and 14% of plant species are visited. See Clare, Goerlitz, et al. (2014) for a discussion of trophic roles of *Glossophaga*

PCR products were pooled without normalization and 70 μ l of the pooled product was cleaned using the PCR Clean DX kit (Aline Biosciences) for a double size selection purification protocol (Supporting Information Table S1). Purified PCR products were eluted in 36 μ l of water. The concentration was measured on the Qubit 2.0 spectrophotometer using a Qubit dsDNA HS Assay Kit (Invitrogen, Life Technologies). All products were normalized to 1 ng/ μ l prior to final library dilution. Sequencing was performed using the Ion PGM Template OT2 400 kit for template preparation according to manufacturer's instructions, except for a ~ 2–3 \times recommended dilution with water (Supporting Information Table S2), and a 316 chip. After the chip check (prior to loading), the chip was flushed once with 100 μ l of 100% isopropanol and three times with 100 μ l of annealing buffer.

2.5 | Bioinformatics analysis

The sequences were processed using two analytical pipelines for comparison. First, we employed well-established tools in the Galaxy platform (Afgan et al., 2016). Reads from the 157 bp Zeale region were separated by MID allowing two indels and two mismatches using the barcode splitter tool. For both the 157 and 407 bp datasets, primers (MIDs for the 157 bp region) and adaptors were removed using the clip tool (both tools from the FASTX tool kit (Assaf Gordon [2010]. FASTQ/A short reads preprocessing tools. http://hannonlab.cshl.edu/fastx_toolkit/). The resulting amplicons were filtered for length (157 or 407 bp \pm 10 bp) and dereplicated (Supporting Information Figure S3) using the collapse tool (FASTX tool kit). We used custom scripts to remove singletons (Supporting Information Table S3).

For the 157 bp dataset, we clustered the remaining haplotypes into MOTUs at 90%–97% similarity in QIIME using the pick_otu and uclust methods (<http://qiime.sourceforge.net/>). See Clare et

al., (2016) and Flynn et al., (2015) for a discussion of MOTU thresholds. For each dataset, we used a BLAST analysis interpreted in MEGAN (Huson, Mitra, Ruscheweyh, Weber, & Schuster, 2011) to filter out MOTUs that could not be reliably assigned to an arthropod order using a reference database of >600,000 COI sequences extracted from GenBank. Parameters in Megan were as follows: Max Expected = 0.01, Top Percent = 10, Min Support Percent (Off), Min Support = 1, Min Complexity = 0.2, Min Score = 250. We tested a representative sequence from each MOTU in UCHIME as implemented in MOTHUR (Schloss et al., 2009) to filter out MOTUs that were likely to be chimeras.

For each MOTU dataset (90%–97% clustering), we examined a BLAST assignment for MOTU representatives in MEGAN. If two or more reads were assigned to the same species, we considered MOTUs oversplit, rejected that threshold and tested the next most conservative option. We particularly considered assignments in the Lepidoptera because this order was heavily represented in the reference database. A QIIME threshold of 92 minimized MOTU oversplitting and this dataset was used for network construction. The same analysis was performed for the 407 bp dataset (MEGAN Min Score = 500) but without network construction (Supporting Information Figure S4).

We further queried the 157 and 407 bp datasets by comparing all sequences to the same reference sequence library and to a reference library provided by D. Janzen and W. Hallwachs generated from specimens (primarily Lepidoptera) from the study area visualized in MEGAN (Supporting Information Figures S5 and S6) and with custom BLAST parsing scripts. This analysis extracts species-level identifications, but is biased towards identification of Lepidoptera, which dominate the reference database from the study location, and the accuracy of database curation (e.g., databases generally provide better resolution of larger, more charismatic and economically important species).

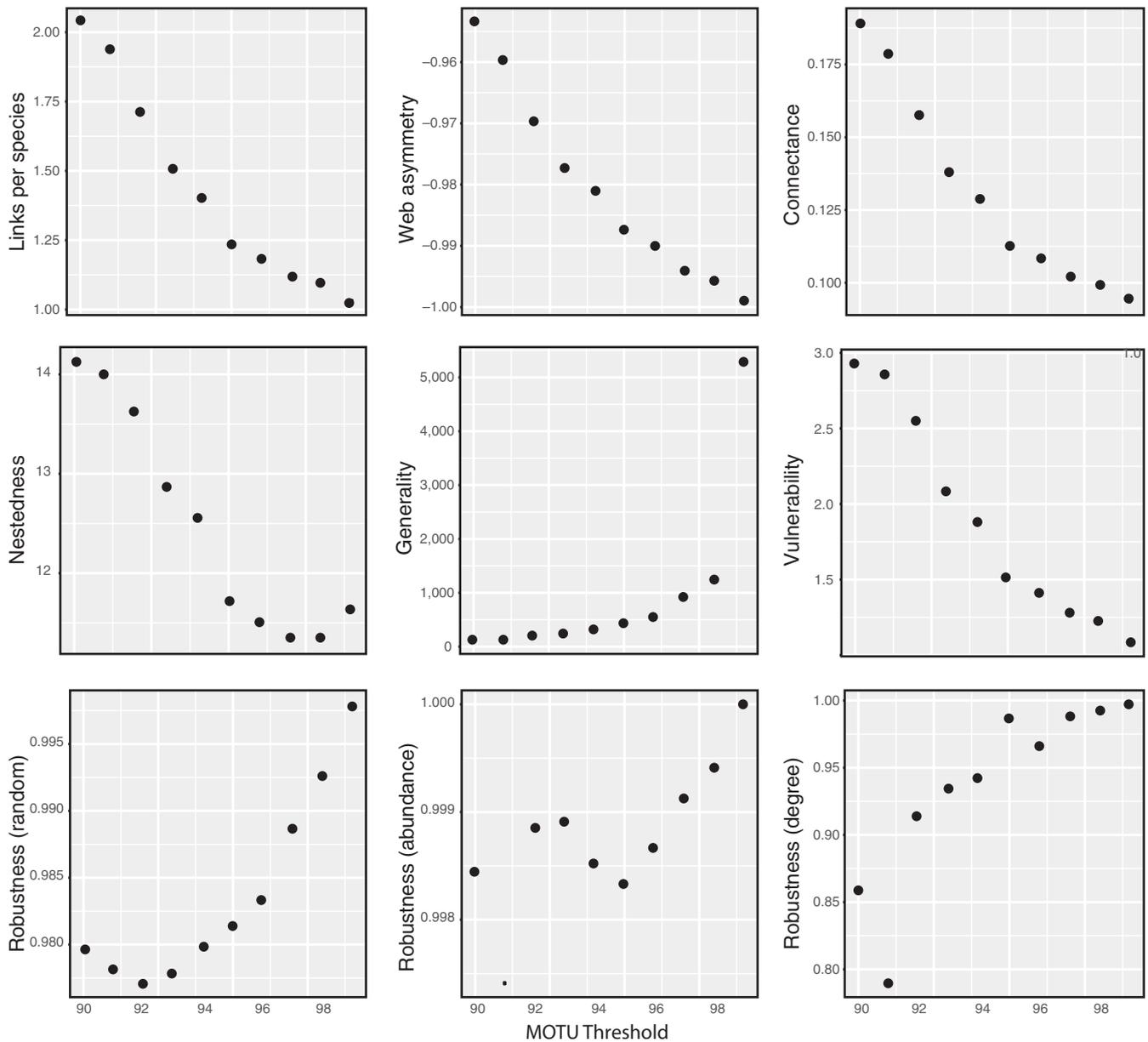


FIGURE 4 Metabarcoding data are a fundamentally new type of data for network ecology. Nodes in metabarcoding normally do not represent a specific taxon or taxonomic level, but are molecular operational taxonomic units (MOTUs) best described as taxa that are defined by being genetically pools of equal diversity. They are defined by a series of bioinformatics steps with the ultimate decision dependent on the threshold employed for splitting versus lumping sequences into a MOTU (nodes in our networks). As the MOTU threshold changes, taxa are lumped or split to a greater or lesser extent. For most network metrics (top two rows), this has a predictable effect as the resource level in our networks increases in richness relative to the consumers. The same pattern is evident but less predictable in measures of robustness (bottom row). For a complete analysis, see Hemprich-Bennett et al. (2018)

For a second comparative analytical approach, we used a non-MOTU-based method. Initial steps were similar with reads processed in Galaxy to split by MID and remove primers using cutadapt (<https://cutadapt.readthedocs.io/en/stable/guide.html>). FASTQ files were then transferred to the mBRAVE platform (<http://www.mbrave.net/>) and processed using the parameters trim front 0 bp, trim end 0 bp (primers and adaptors had already been removed via cutadapt) trim length 500 bp and filtering of MINQV 0QV, min length 147 bp, and max bases with low or ultra low QV of 100% (to avoid specific quality filter parameters). We set a preclustering threshold

of none, and ID distance threshold of 2% and left OTU thresholds as preset standards as we ignored MOTU analysis for this comparison. Chimera screening and dereplication were performed automatically by the mBRAVE platform. The resulting data are automatically compared to the BOLD system library for insects using the BIN approach to attempt to associate the reads with known BINs (this library contained 580, 824 reference sequences from 434, 878 known BINs, last updated 21 October 2018). The resulting dataset was then converted to a matrix of bat species vs. associated prey BINs for further analysis.

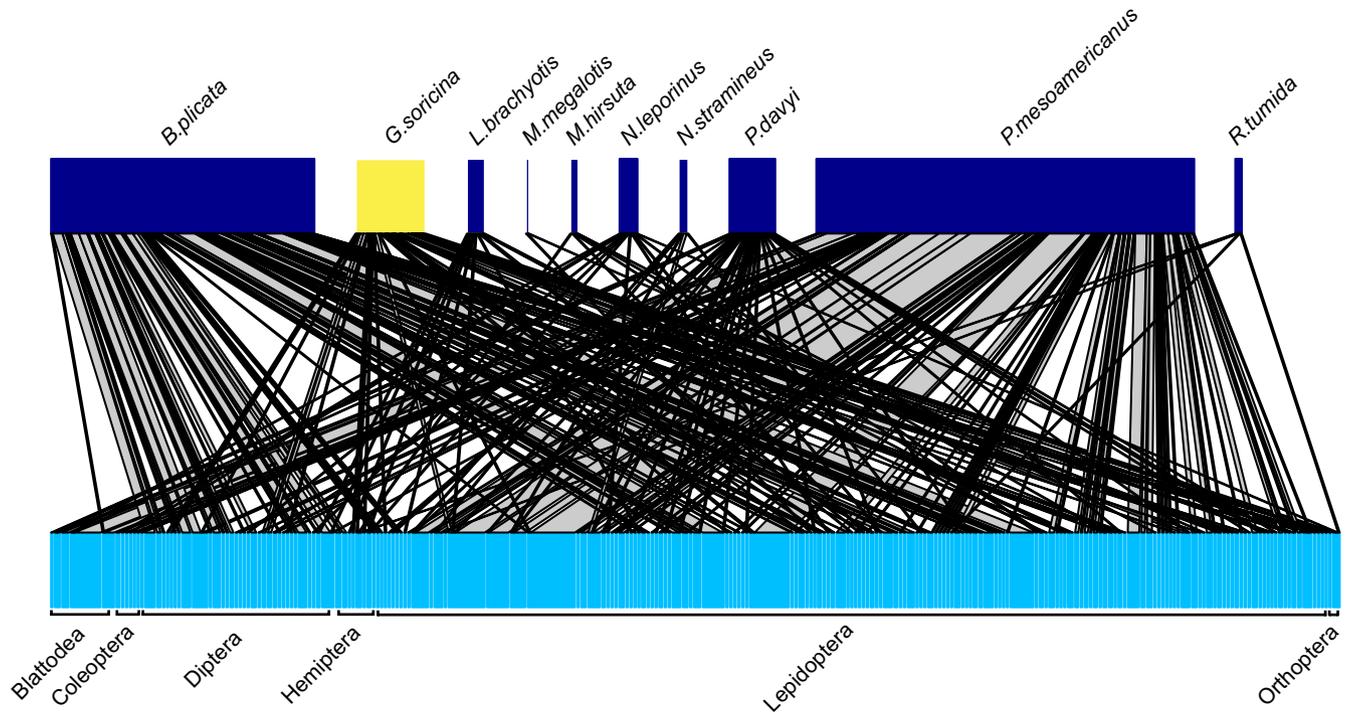


FIGURE 5 Species' interaction network for bats and prey identified using the BIN association method employed on the mBRAVE platform. See Supporting Information Table S9 for a matrix of bats and BIN-based nodes with full taxonomic identifications [Colour figure can be viewed at wileyonlinelibrary.com]

2.6 | Network analysis

Using the data produced by all three approaches, we constructed a “network of networks” in Cheddar (Hudson et al., 2013) and Bipartite (Dormann et al., 2008) in R (R Development Core Team, 2008) which represents all identified taxa or MOTUs. This network has differing levels of resolution based on the trophic level or taxonomic group. As the bats are, with a single exception, resolved to species level, they are fully quantified. The mites and flies that feed on them are identified by a Barcode Identification Number (BIN; Ratnasingham & Hebert, 2013) as a proxy for species. This trophic level is also well resolved, but individual taxa are only partially quantified from each bat (finding all individual parasites is not practically possible). Similarly, the plant and arthropod prey levels are frequency-based as it is not possible to assess ingested plant biomass from seeds (plant ID) and metabarcoding data are poorly suited to quantify the biomass or abundance of species represented in the data (Deagle et al., 2018). The arthropods are represented by MOTUs (Floyd et al., 2002). In addition, we produced a separate network of bats and prey employing the non-OTU-based BIN association matrix.

3 | RESULTS

3.1 | Molecular analysis

We recovered DNA barcodes from 698 bats representing 24 species belonging to 17 genera. The barcode results generally confirmed

the field IDs, but could not distinguish *Artibeus lituratus* and *Artibeus intermedius* (Clare, Lim, et al., 2011) leaving this node unresolved. However, other cases of taxonomic uncertainty were resolved. An unknown species of *Carollia* was identified as *C. sowelli* and members of two genera (*Glossophaga* and *Micronycteris*) gained species assignments. We suspect one genetically divergent specimen of *Sturnira parvidens* may be a sister taxon, but since this outcome could not be confirmed, it was retained as a single node (Supporting Information Figure S1).

We recovered DNA barcodes from 445 of the 466 mites and flies found on 18 host species. Parasite diversity varied from a single ectoparasite species per bat species (nine cases) to nine ectoparasite species on *A. jamaicensis*. Among the seven bat families, the Emballonuridae, Mormoopidae and Phyllostomidae hosted the greatest diversity of parasite species whereas individuals of the Noctilionidae were only associated with two ectoparasite species and individual Molossididae, Natalidae and Vespertilionidae were only parasitized by one ectoparasite species at a time. Two-thirds of the 34 ectoparasite species, nine mites and 13 flies were only collected from one host species. The maximum number of host species inhabited by a mite or fly species was four.

We recovered plant DNA from 112 guano samples from 12 species of bat. Guano from seven bats contained two seed morphotypes analysed separately, producing 119 sequenced seed samples. We recovered *rbcl* from 102, *matK* from 81 and *trnH-psbA* from 106 samples. Through comparison to GenBank and BOLD, 103 samples had sequences assigned to eight genera based on at least two of

the three loci. Of these, 97 seed samples had sequences assigned to a putative species and 16 samples had sequences placed to a genus based on a single-gene region (Supporting Information Table S4). Comparison of *rbcL* sequences to GenBank often returned multiple BLAST hits with equivalent best scores. For example, top BLAST matches to *Ficus* or *Solanum* matched (100% or 99% identity, respectively) multiple species within these genera. Although some sequences did not have an identical match on GenBank, all *rbcL* sequences matched with 100% identity to a sequence on BOLD, presumably reflecting the greater diversity of taxa present in the latter database. Similarly, all *matK* sequences matched with 100% identity to sequences on BOLD vs. lower values on GenBank (94%–100%). In some cases, this allowed a more precise taxonomic assignment on BOLD, either to a species (e.g., *Guazuma ulmifolia*), or species cluster (e.g., *Cecropia obtusifolia/peltata/insignis*) vs. assignment to a higher taxonomic rank (e.g., Urticaceae or *Cecropia* sp.; *Ficus* sp.).

The GenBank BLAST of *trnH-psbA* sequences corroborated results obtained with *rbcL* and *matK*. In all cases, samples that yielded unique sequences for *matK* also had unique sequences for *trnH-psbA*. Although the limited taxonomic coverage for the latter gene region on GenBank often prevented an assignment to a known species, these sequence variants were treated as putative species. We also detected a probable case of taxonomic error in GenBank. Two *trnH-psbA* sequences from our samples showed high similarity (98% identity) to *Cecropia obtusifolia*, an unexpected result as several other sequences of almost twice the length showed nearly 100% similarity to several other species of *Cecropia* on GenBank. Further investigation revealed that these sequences probably belong to a species of *Vismia* (Hypericaceae).

We used HTS to recover two regions of the mitochondrial COI gene (157 and 407 bp, Supporting Information Tables S2–S3). The 157 bp region has been used extensively (Alberdi et al., 2018; Zeale et al., 2011) and generated high recovery rates in the present study; it is fully analysed and generated 686 MOTU at the given parameters. Surprisingly, given the degradation induced by digestion, the 407 bp region also showed high sequencing success. These two regions (Supporting Information Tables S5 and S6) identified a similar number of species (118 vs. 109 taxa for the 157 bp and 407 bp regions, respectively) from all the same classes and orders of arthropods (except one mantid). Many of the same species, for example, 32 species of Lepidoptera, were common in the two lists. However, there were also different species identified and in a number of cases identifications were improved using the longer target region. For example, sequences assigned to the genus *Culex* by the short region could be identified as *Culex nigripalpus* by the 407 bp region. This outcome suggests these two regions may be complementary, adding confidence to the general diversity recovered and the specific taxa identified. However, the 407 bp region pushes the current limits of amplicon size recovery on most HTS platforms, creating constraints on quality and recovery rate. Analysis with the BIN association method in mBRAVE identified 212 potential prey in the 157 bp dataset.

3.2 | The impact of OTUs on network metrics

The most novel data type generated is the metabarcoded data that underlie the bat–prey network because the prey nodes do not represent a particular taxonomic level or taxon but a measure of prey genetic diversity. As a result, we investigated the impact of the key bioinformatics step—that of generating MOTUs—on the measurement of common network variables. Our data suggest that MOTU thresholds have a significant impact on standard network metrics as taxa are lumped or split to a greater or lesser extent. For most network metrics (Figure 4), an increase in the MOTU threshold (e.g., from 90% to 99%) split taxa so the resource level in our network increased in richness relative to the consumers with expected outcomes for each metric. In the case of links between species, connectance, nestedness and vulnerability, this variation can result in different relative rankings of these metrics between network types. For a complete analysis, see Hemprich-Bennett et al., (2018). The effect is consistent but less predictable in measures of robustness (Figure 4), but in all cases we would have drawn the same conclusion. The BIN association network (Figure 5) contained substantially fewer prey nodes, which is to be expected, as the reference database for the area is minimal. Of these, 75% were Lepidoptera reflecting the substantial effort to create a Lepidoptera reference library for the site (see below). Interestingly, the actual measurements of network properties did not differ substantially (Table 1) which reflects the tremendous prey diversity represented by any method.

4 | DISCUSSION

We have demonstrated that three types of molecular data can be incorporated into network analysis. DNA can be used to confirm field identities (e.g., bats) or differentiate cryptic taxa (e.g., parasites) and to identify morphologically compromised material (e.g., plant pulp). DNA can also be used to generate complex and fundamentally novel data via metabarcoding of mixed material (e.g., faeces) that can be analysed using MOTU or association with taxa in reference collections (e.g., the BIN method). While these data types can effectively generate rapid, scalable analyses of entire communities, there are challenges in both generating data and in the interpretation of network metrics to ensure biologically meaningful results (Table 2, Figure 4).

The incorporation of DNA analysis into networks presents both straightforward use cases and challenges. Confirming field IDs is a common molecular procedure (Borisenko, Lim, Ivanova, Hanner, & Hebert, 2008) and differentiating cryptic or taxonomically complex species is now routine (Smith, Woodley, Janzen, Hallwachs, & Hebert, 2006). These approaches have successfully been incorporated into network analysis (e.g., Wirta et al., 2014). However, the inclusion of metabarcoding results is more challenging and requires special consideration to integrate with network analysis. Metabarcoding is best applied to mixed faecal samples, gut contents (particularly liquid

TABLE 2 A comparison of three data types for network analysis

	Data type	Taxonomic resolution	Application	Required references	Advantages	Disadvantages	Network implications
Resolving ID with DNA The plant network	A few sequences per individual	Species	Identification of fragments	Complete database	Can deal with forensic trace material Produces traditional networks Minimal taxonomic expertise	Requires molecular expertise Requires a well-resolved database	Produces a network equivalent to a traditional food web
Resolving species when taxonomy is not known The parasite network	A few sequences per individual	Species but without names	Identification of taxa where taxonomy may be incomplete or cryptic	Incomplete database	Can deal with forensic trace material Can include species with incomplete taxonomic investigation Resolves cryptic taxa	Requires molecular expertise May not encourage taxonomic work	Similar to a traditional food web, but may collapse trophic levels where ID is not clear
Using MOTUs without a taxonomic unit The arthropod prey network	Millions of sequences per sample	Arbitrary but comparable units	Rapid surveys where identification is not possible	No database required	Can deal with forensic trace material May include both known and unknown data (MOTUs) Can be used in any context regardless of taxonomic knowledge Rapid and MOTUs are mathematically and genetically identical yielding perfect resolution Quantification is controversial ^a	Requires molecular expertise Does not represent real taxa May include error-prone data Actual MOTU numbers may be meaningless May be biased by primers or other protocol choices	Fundamentally different MOTUs are not species and probably collapse trophic levels but allow rapid structural comparisons Node numbers are likely meaningless

Note. ^aSee a review by Deagle et al. (2018).

feeders, e.g., Piñol, San Andrés, Clare, Mir, & Symondson, 2014) or pollen carried by generalists. However, it is challenging to generate reliable metabarcoded data (Alberdi et al., 2018; Arrizabalaga-Escudero et al., 2018; Atwell et al., 2010; King, Read, Traugott, & Symondson, 2008). The methods of interpreting individual dietary analyses using these data have been studied in several contexts (Clare et al., 2016; Flynn et al., 2015; Pompanon et al., 2012; Symondson, 2002). However, certain challenges are unique to the interpretation of food webs. Debate about the quantification of metabarcoding data centres largely around whether sequence recovery is linked to original biomass (Deagle et al., 2013; Nielsen et al., 2018; Pompanon et al., 2012). While this is possible in restricted scenarios (Bowles, Schulte, Tollit, Deagle, & Trites, 2011; Thomas et al., 2016), in many cases frequency-based measures of interactions are more appropriate (Nielsen et al., 2018). Frequency-based approaches are common in network ecology, for example, visitation frequency to specific flowers is a standard metric of the strength of mutualistic interactions (Memmott et al., 2004; Vázquez, Morris, & Jordano, 2005). However, incomplete quantification needs to be considered when weighted metrics are used (Kaiser-Bunbury, Muff, Memmott, Müller, & Caflisch, 2010), as rare and common interactions may be equally weighted (Clare, 2014).

We suggest two alternative ways to incorporate metabarcoding data: using MOTUs and screening for taxonomic identities (e.g., BINs, similarity searches). The advantage of MOTUs is that all data are incorporated, both known and unknown taxa. However, by incorporating unknowns, one may inadvertently include nontarget taxa (e.g., intestinal parasites or bacteria that are not screened out bioinformatically) that may generate nodes in networks unrelated to the behaviour under study or even false nodes from sequencing error (Clare et al., 2016; Flynn et al., 2015). In many systems, MOTUs collapse all prey levels into one “resource” level rather than revealing the complexity among trophic levels. For example, in our case some insects were primary consumers while others were predators, but all were treated as MOTU “prey” of an undifferentiated consumer level. Our data further suggest that the protocols used to differentiate MOTUs will themselves impact network metrics (Figure 4). The effect of node resolution has been discussed for decades with analyses showing that the impact of resolution on node, chain length and trophic levels significantly alters the observations of network properties (Brose, Ostling, Harrison, & Martinez, 2004). The situation is similar but not identical to the node resolution issue of employing MOTU. The impacts of the informatic steps used to generate MOTU are only starting to be considered in ecological (Clare et al., 2016) or network analysis (Hemprich-Bennett et al., 2018). Any network that incorporates taxa with different levels of resolution (e.g., mixing genus and species designations) faces the same challenge (Hemprich-Bennett et al., 2018; Martinez, 1991). However, MOTUs make it possible to easily re-analyse any dataset to empirically estimate that impact (Figure 4) and one potential advantage is that MOTUs generate a uniform level of resolution in a network. By their nature, MOTUs represent equal and repeatable measures of biodiversity (Floyd et al., 2002), even if that level does not equate

to a standard level of taxonomy. This may represent a powerful advantage in comparing network structure across systems, but presents a challenge in interpretation. For example, unnamed MOTUs of unknown life cycle and unknown affinity to each other provide limited information on the nature of the ecological interaction being measured beyond the general structure of the community. Similarly, if the numbers of nodes and their connections vary with analysis parameters (e.g., MOTU threshold), a network on its own holds little biological meaning. However, if the same methods are replicated a biological picture can emerge. For example, if the prey level undergoes a population crash, the genetic diversity and the MOTU number would similarly decline relative to the consumer level and fluctuations in parameters such as generality or nestedness would be measurable. The key then is to compare only analyses that employ the same methods from sequencing platform and field and laboratory protocol to informatics choices, just as sampling protocols and node resolution should be maintained in traditional networks being compared. This would be required to avoid context specificity. It is also necessary to pick specific metrics; for example network level metrics may be more reliable than motif measurements (though see Hemprich-Bennett et al., 2018).

In contrast, similarity-based searches and BIN association type data provide better ability to determine exactly what is being included as a node (e.g., Figure 5) and yield greater ecological information about the type of interaction being measured, but will be biased by the contents of the reference library being used. In this case, the network metrics were similar enough that comparative conclusions about bat-prey/BIN, bat-parasite and bat-plant networks would remain the same but some specific indicators change. For example, generality of the bat-BIN network was much lower reflecting the substantial reduction in prey nodes when relying exclusively on reference collections for the inclusion of a prey node. As reference collections improve this effect will diminish but it is a very important factor in relatively unexplored faunas.

4.1 | Three distinct data types

This paper has considered three distinct types of molecular data. The bat and plant identifications provide by DNA deliver nearly perfect resolution of the network. Such analysis generates data similar to that employed in traditional network ecology, the only major difference being the need for multilocus data to obtain species-level resolution for plants. The parasite identifications were generated in a similar fashion to the bat data (one sequence per specimen), but with the crucial difference that current reference databases are very incomplete. As a consequence, we employed an alternative taxonomic system, the Barcode Index Number (BIN). The performance of the BIN system has been extensively tested (Ratnasingham & Hebert, 2013) and these studies have shown that it delivers taxonomic resolution that is very close to traditional taxonomy. This data type (Table 2) has the advantage of making it possible to incorporate taxa which are apparently different species but where the current taxonomic system is incomplete. BIN analysis avoids unresolved

nodes in network construction, but imposes a constraint that the identification is based on a measure of sequence differentiation observed in related taxa. Unlike other MOTU-generating methods, the BIN system is not based on strict *a priori* threshold delimitation but has been trained specifically using the large Sanger datasets for the DNA barcode region. In a test of 1,400 species spanning birds, bees, fishes and Lepidoptera, the correspondence between species counts based on traditional taxonomy and BINs was very high ($r^2 = 0.99$) and the actual mapping of species to BIN was approximately 90% (varying from 79% to 97% between taxonomic group) (Ratnasingham & Hebert, 2013). Thus, when viewed from the context of DNA barcoding, BINs are a strong proxy for species. Because the definition of new BINs requires at least 500 bp of sequence information from the COI barcode region, the short reads generated by most current HTS platforms cannot be used to delineate new BINs although they can be matched to existing BINs. Reflecting this constraint, there is a need for other methods of MOTU generation. For this third data type, data are most often analysed using MOTU without much (if any) taxonomic identification (e.g., Figure 1). This has the advantage of making it possible to analyse mixed sources (e.g., stomach contents) but imposes unique problems for network ecology as it compresses trophic levels and dispenses with traditional taxonomy. While such analysis can generate data for a comparable interaction network model, it may not represent a trophic food web. In the study location, most arthropods remain undescribed despite decades of intense taxonomic work. For example, some 400,000 arthropod species are estimated to be present in the ACG, but just 43,000 of these species have been barcoded over 14 years, revealing the scope of the taxonomic challenge (D.H. Janzen Pers. Comm.). In such locations, a complete food web or interaction network is impossible and restricting analysis to those species which can gain a full taxonomic designation (either by morphology or DNA) would introduce a substantial bias (Supporting Information Table S6). As a consequence, our bat–BIN network contained substantially fewer nodes than our bat–prey network based on MOTU and would be less comparable to a network generated in an area with a different/reduced reference database. In such cases, a BIN or MOTU approach to generating a reference collection and then some sort of association or matching system is the only means of developing an ecosystem network model. The use of reference databases can provide a familiar binomial designation on some nodes by similarity searchers or BIN association but imposes a significant bias on the data, which is then composed of “things found in databases” while novel BINs, and MOTUs do not impose this bias. On the other hand, novel BINs and MOTUs may include nontarget taxa such as parasites, parasitoids or taxa acquired via secondary predation.

4.2 | Comparison of 157 and 407 bp datasets

Current consensus suggests that short reads are required to maximize MOTU and taxonomic ID recovery in digested material because of DNA degradation. Contrary to this expectation, the 407 bp region had higher MOTU estimates and broader taxonomic coverage

when evaluated using BLAST, suggesting it has less amplification bias and hence a complementary region for arthropod diversity analysis. However, these conclusions need to be considered with caution. Longer reads should generate better taxon identification scores (more information) but will also generate high rates of low-quality BLAST scores (local alignments). We modified the MEGAN Min Support value to partially compensate for this and to maximize assignment with most scores >98% similar to references. Taxonomic assignments of MOTUs (e.g., Supporting Information Tables S5 and S6) should be considered for interest's sake only in this dataset, particularly when a reference databases contains errors or skewed coverage. For this reason, we considered only MOTUs for network analysis. Despite the promise of the 407 bp region, we used the 157 bp region MOTUs for network analysis for two reasons. First, unlike the 157 bp region, the 407 bp region is new to NGS analysis and has not been evaluated for this purpose before. We consider it an interesting and potentially important tool but are hesitant to rely on it until further testing has been completed. Second, the 407 bp region is long compared to the capacity of most high-throughput sequencing platforms which limits its use and prevented us from employing MIDTs to separate samples. Platform read length has generally fallen since the first highly popular Roche454 platform capable of 1,000 bp reads to the now standard 250 bp paired end reads of the MiSeq, thus while promising, the 407 bp read will be analytically challenging. Newer platforms such as the SMRT sequencing platform (PACBIO, Pacific Biosciences) can overcome this problem allowing longer reads and thus higher taxonomic resolution assuming that digestion has not substantially compromised the DNA.

4.3 | DNA integration into network ecology

Despite challenges, incorporating DNA into networks has significant advantages. First, the technique does not rely on the need to observe interactions or the time-consuming rearing practices used to establish many cases of parasitism (Wirta et al., 2014). It can be applied to broken seeds, fruit pulp (e.g., Lim et al., 2017), single grains of pollen or morphologically destroyed material (e.g., digested remains) as well as entire specimens. Even traces of DNA (eDNA) with no observable material are amenable (Bohmann et al., 2014; Drinkwater et al., 2018); for example, seeds dropped on the forest floor will have DNA of the plant, but also of the animal that dropped them from either saliva or cells from a digestive tract. González-Varo et al. (2014) have spectacularly demonstrated this method to capture bird DNA on the surface of dispersed seeds. Similarly, the detection of cryptic species and relationships represents a huge shift in the resolution of interaction networks. This was demonstrated by Wirta et al., (2014) who observed that DNA dramatically increased the number of identified interaction types and altered the perceived host specificity of host-parasitoid networks.

A rapid DNA-based network biomonitoring tool will require us to understand: first, which data can be quantified (Deagle et al., 2018) and second, which metrics are reliable, in relative or absolute terms,

to ensure we produce biologically meaningful outputs (Clare et al., 2016; Hemprich-Bennett et al., 2018; Ings et al., 2009). However, these datasets are already being demonstrated as powerful tools to resolve complex interaction networks quickly and in exquisite detail. Here, we have generated a detailed network of networks in a complex tropical ecosystem incorporating different molecular data types as a case study. Ecologically, our data suggest a hitherto unrecognized keystone species and behavioural flexibility that may be critical to the success of insectivores (Box 1). Methodologically, our findings support the approach, but also highlight the need for rigorous testing of methods. The rapidly advancing technology of this field means that such analyses will soon become a common and relatively inexpensive tool for understanding biostructure (McCann, 2007). While a fully resolved and taxonomically identified network will always be the goal, our analysis demonstrates the utility of these tools for network ecology and produces the first full network of networks resolved entirely by DNA.

ACKNOWLEDGEMENTS

We are indebted to collaborators who have supported this research including staff at the Canadian Centre for DNA Barcoding and the Biodiversity Institute of Ontario and the staff of Sector Santa Rosa in the Area de Conservación Guanacaste (ACG) Costa Rica particularly Daniel Janzen, Winnie Hallwachs, Roger Blanco and Maria Marta Chavarria. We also appreciate funding from Bat Conservation International to ELC which supported original field work.

AUTHOR CONTRIBUTIONS

E.L.C., A.M.A. and J.N. performed fieldwork. E.L.C., A.J.F., N.V.I. and R.M.F. performed laboratory work. E.L.C. and R.G. performed analysis. P.D.N.H., M.B.F. and S.G.N. helped design the study. All authors contributed to the writing of the manuscript.

DATA ACCESSIBILITY

All molecular data can be found in Dryad <https://doi.org/10.5061/dryad.0k90c0v> and BOLD projects (BCCR Bats of Costa Rica ACG & BCPB Parasites of tropical bats) also contain sequences and collection metadata and associated GenBank accessions.

ORCID

Elizabeth L. Clare  <https://orcid.org/0000-0002-6563-3365>

REFERENCES

- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., ... Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, 44(W1), W3–W10. <https://doi.org/10.1093/nar/gkw343>
- Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, 9, 134–147. <https://doi.org/10.1111/2041-210X.12849>
- Arrizabalaga-Escudero, A., Clare, E. L., Salsamendi, E., Alberdi, A., Garin, I., Aihartza, J., & Goiti, U. (2018). Assessing niche partitioning of co-occurring sibling bat species by DNA metabarcoding. *Molecular Ecology*, 27(5), 1273–1283. <https://doi.org/10.1111/mec.14508>
- Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., ... Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, 465(7298), 627–631. <https://doi.org/10.1038/nature08800>
- Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., ... de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, 29(6), 358–367. <https://doi.org/10.1016/j.tree.2014.04.003>
- Borisenko, A. V., Lim, B. K., Ivanova, N. V., Hanner, R. H., & Hebert, P. D. N. (2008). DNA barcoding in surveys of small mammal communities: A field study in Suriname. *Molecular Ecology Resources*, 8(3), 471–479. <https://doi.org/10.1111/j.1471-8286.2007.01998.x>
- Bowles, E., Schulte, P. M., Tollit, D. J., Deagle, B. E., & Trites, A. W. (2011). Proportion of prey consumed can be determined from faecal DNA using real-time PCR. *Molecular Ecology Resources*, 11(3), 530–540. <https://doi.org/10.1111/j.1755-0998.2010.02974.x>
- Braukmann, T. W. A., Kuzmina, M. L., Sills, J., Zakharov, E. V., & Hebert, P. D. N. (2017). Testing the efficacy of DNA barcodes for identifying the vascular plants of Canada. *PLoS One*, 12(1), e0169515. <https://doi.org/10.1371/journal.pone.0169515>
- Brose, U., Ostling, A., Harrison, K., & Martinez, N. D. (2004). Unified spatial scaling of species and their trophic interactions. *Nature*, 428(6979), 167–171. <https://doi.org/10.1038/nature02297>
- Burgess, K. S., Fazekas, A. J., Kesanakurti, P. R., Graham, S. W., Husband, B. C., Newmaster, S. G., ... Barrett, S. C. H. (2011). Discriminating plant species in a local temperate flora using the rbcL+matK DNA barcode. *Methods in Ecology and Evolution*, 2(4), 333–340. <https://doi.org/10.1111/j.2041-210X.2011.00092.x>
- CBOL plant working group (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), 12794–12797. <https://doi.org/10.1073/pnas.0905845106>
- Christe, P., Arlettaz, R., & Vogel, P. (2000). Variation in intensity of a parasitic mite (*Spinturnix myoti*) in relation to the reproductive cycle and immunocompetence of its bat host (*Myotis myotis*). *Ecology Letters*, 3(3), 207–212. <https://doi.org/10.1046/j.1461-0248.2000.00142.x>
- Clare, E. L. (2014). Molecular detection of trophic interactions: Emerging trends, distinct advantages, significant considerations and conservation applications. *Evolutionary Applications*, 7(9), 1144–1157. <https://doi.org/10.1111/eva.12225>
- Clare, E. L., Barber, B. R., Sweeney, B. W., Hebert, P. D. N., & Fenton, M. B. (2011). Eating local: Influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, 20(8), 1772–1780. <https://doi.org/10.1111/j.1365-294X.2011.05040.x>
- Clare, E. L., Chain, F. J. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome*, 59(11), 981–990. <https://doi.org/10.1139/gen-2015-0184>
- Clare, E. L., Lim, B. K., Fenton, M. B., & Hebert, P. D. N. (2011). Neotropical bats: Estimating species diversity with DNA barcodes. *PLoS ONE*, 6(7), e22648. <https://doi.org/10.1371/journal.pone.0022648>
- Clare, E. L., Symondson, W. O. C., & Fenton, M. B. (2014). An inordinate fondness for beetles? Variation in seasonal dietary preferences of night roosting big brown bats (*Eptesicus fuscus*). *Molecular Ecology*, 23(15), 2333–2647. <https://doi.org/10.1111/mec.12519>

- Clare, E. L., Goerlitz, H. R., Drapeau, V. A., Holderied, M. W., Adams, A. M., Nagel, J., ... Brock Fenton, M. (2014). Trophic niche flexibility in *Glossophaga soricina*: How a nectar seeker sneaks an insect snack. *Functional Ecology*, 28(3), <https://doi.org/10.1111/1365-2435.12192>
- Costanza, R., D'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., ... van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, 387(6630), 253–260. <https://doi.org/10.1038/387253a0>
- Csardi, G., & Nepusz, T. (2006). The Igraph Software Package for Complex Network Research. InterJournal 2006, Complex Systems, 1695.
- Deagle, B. E., Thomas, A. C., Shaffer, A. K., Trites, A. W., & Jarman, S. N. (2013). Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: Which counts count? *Molecular Ecology Resources*, 13(4), 620–633. <https://doi.org/10.1111/1755-0998.12103>
- Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., ... Eveson, J. P. (2018). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, 28(2), 391–406. <https://doi.org/10.1111/mec.14734>
- R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Development Core Team.
- R Development Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Dormann, C. F., Fründ, J., Blüthgen, N., & Gruber, B. (2009). Indices, graphs and null models: Analyzing bipartite ecological networks. *The Open Ecology Journal*, 7–24. <https://doi.org/10.2174/1874213000902010007>
- Drinkwater, R., Schnell, I. B., Bohmann, K., Bernard, H., Veron, G., Clare, E., ... Rossiter, S. J. (2018). Using metabarcoding to compare the suitability of two blood-feeding leech species for sampling mammalian diversity in North Borneo. *Molecular Ecology Resources*, 19(1), 105–117. <https://doi.org/10.1111/1755-0998.12943>
- Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., & Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology*, 30(12), 1904–1916. <https://doi.org/10.1111/1365-2435.12659>
- Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., ... Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One*, 3(7), e2802. <https://doi.org/10.1371/journal.pone.0002802>
- Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, 11(4), 839–850. <https://doi.org/10.1046/j.1365-294X.2002.01485.x>
- Flynn, J. M., Brown, E. A., Chain, F. J. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, 5(11), 2252–2266. <https://doi.org/10.1002/ece3.1497>
- Ford, C. S., Ayres, K. L., Toomey, N., Haider, N., Van Alphen Stahl, J., Kelly, L. J., ... Wilkinson, M. J. (2009). Selection of candidate coding DNA barcoding regions for use on land plants. *Botanical Journal of the Linnean Society*, 159(1), 1–11. <https://doi.org/10.1111/j.1095-8339.2008.00938.x>
- Fritz, G. N. (1983). Biology and ecology of bat flies (Diptera: Streblidae) on bats in the genus *Carollia*. *Journal of Medical Entomology*, 20, 1–10.
- Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS One*, 10(10), e0138432. <https://doi.org/10.1371/journal.pone.0138432>
- González-Varo, J. P., Arroyo, J. M., & Jordano, P. (2014). Who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies. *Methods in Ecology and Evolution*, 5(8), 806–814. <https://doi.org/10.1111/2041-210X.12212>
- Guimarães, P. R. Jr, Jordano, P., & Thompson, J. N. (2011). Evolution and coevolution in mutualistic networks. *Ecology Letters*, 14(9), 877–885. <https://doi.org/10.1111/j.1461-0248.2011.01649.x>
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hebert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, 270(Suppl. 1), S96–S99. <https://doi.org/10.1098/rsbl.2003.0025>
- Hebert, P. D. N., Dewaard, J. R., Zakharov, E. V., Prosser, S. W. J., Sones, J. E., McKeown, J. T., ... La Salle, J. (2013). A DNA “barcode blitz”: Rapid digitization and sequencing of a natural history collection. *PLoS One*, 8(7), e68535. <https://doi.org/10.1371/journal.pone.0068535>
- Hemprich-Bennett, D. R., Oliveira, H. F. M., Le Comber, S. C., Rossiter, S. J., & Clare, E. L. (2018). Assessing the impact of taxon resolution on network structure, with implication for comparative ecology. *BioRxiv*. Retrieved from <http://biorxiv.org/content/early/2018/06/28/357376.abstract>
- Hudson, L. N., Emerson, R., Jenkins, G. B., Layer, K., Ledger, M. E., Pichler, D. E., ... Reuman, D. C. (2013). Cheddar: Analysis and visualisation of ecological communities in R. *Methods in Ecology and Evolution*, 4(1), 99–104. <https://doi.org/10.1111/2041-210X.12005>
- Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N., & Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Research*, 21(9), 1552–1560. <https://doi.org/10.1101/gr.120618.111>
- Ivanova, N. V., Clare, E. L., & Borisenko, A. V. (2012). DNA barcoding in mammals. In W. Kress & D. Erickson (Eds.), *DNA barcodes. Methods in Molecular Biology (Methods and Protocols, Vol. 858)*. Totowa, NJ: Humana Press.
- Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., ... Woodward, G. (2009). Review: Ecological networks – beyond food webs. *Journal of Animal Ecology*, 78(1), 253–269. <https://doi.org/10.1111/j.1365-2656.2008.01460.x>
- Kaiser-Bunbury, C. N., Muff, S., Memmott, J., Müller, C. B., & Caflisch, A. (2010). The robustness of pollination networks to the loss of species and interactions: A quantitative approach incorporating pollinator behaviour. *Ecology Letters*, 13(4), 442–452. <https://doi.org/10.1111/j.1461-0248.2009.01437.x>
- King, R. A., Read, D. S., Traugott, M., & Symondson, W. O. C. (2008). Molecular analysis of predation: A review of best practice for DNA-based approaches. *Molecular Ecology*, 17(4), 947–963. <https://doi.org/10.1111/j.1365-294X.2007.03613.x>
- Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), 8369–8374. <https://doi.org/10.1073/pnas.0503123102>
- Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., ... Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8), 2923–2928. <https://doi.org/10.1073/pnas.0709936105>
- Lim, V.-C., Clare, E. L., Littlefair, J. E., Ramli, R., Bhasu, S., & Wilson, J.-J. (2017). Impact of urbanisation and agriculture on the diet of fruit bats. *Urban Ecosystems*, 21(1), 61–70. <https://doi.org/10.1007/s11252-017-0700-3>
- Martín González, A. M., Dalsgaard, B., & Olesen, J. M. (2010). Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, 7(1), 36–43. <https://doi.org/10.1016/j.ecocom.2009.03.008>
- Martinez, N. D. (1991). Artifacts or attributes? Effects of resolution on the little rock lake food web. *Ecological Monographs*, 61(4), 367–392. <https://doi.org/10.2307/2937047>

- McCann, K. (2000). The diversity-stability debate. *Nature*, 405(6783), 228–233. <https://doi.org/10.1038/35012234>
- McCann, K. (2007). Protecting biostructure. *Nature*, 446(7131), 29. <https://doi.org/10.1038/446029a>.
- McDonald-Madden, E., Sabbadin, R., Game, E. T., Baxter, P. W. J., Chadès, I., & Possingham, H. P. (2016). Using food-web theory to conserve ecosystems. *Nature Communications*, 7, 10245. <https://doi.org/10.1038/ncomms10245>
- Memmott, J., Waser, N., & Price, M. (2004). Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society of London B-Biological Sciences*, 271(1557), 2605–2611.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network Motifs: Simple building blocks of complex networks. *Science*, 298(5594), 824–827.
- Nielsen, J. M., Clare, E. L., Hayden, B., Brett, M. T., & Kratina, P. (2018). Diet tracing in ecology: Method comparison and selection. *Methods in Ecology and Evolution*, 9(2), 278–291. <https://doi.org/10.1111/2041-210X.12869>
- Nuismer, S. L., Jordano, P., & Bascompte, J. (2013). Coevolution and the architecture of mutualistic networks: Coevolving networks. *Evolution*, 67(2), 338–354. <https://doi.org/10.1111/j.1558-5646.2012.01801.x>
- Piñol, J., San Andrés, V., Clare, E. L., Mir, G., & Symondson, W. O. C. (2014). A pragmatic approach to the analysis of diets of generalist predators: The use of next-generation sequencing with no blocking probes. *Molecular Ecology Resources*, 14(1), 18–26. <https://doi.org/10.1111/1755-0998.12156>
- Pocock, M. M. J. O., Evans, D. M. D., & Memmott, J. (2012). The robustness and restoration of a network of ecological networks. *Science*, 335(6071), 973–977.
- Poisot, T., Stouffer, D. B., & Kéfi, S. (2016). Describe, understand and predict: Why do we need networks in ecology? *Functional Ecology*, 30, 1878–1882. <https://doi.org/10.1111/1365-2435.12799>
- Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, 21(8), 1931–1950. <https://doi.org/10.1111/j.1365-294X.2011.05403.x>
- Porco, D., Rougerie, R., Deharveng, L., & Hebert, P. D. N. (2010). Coupling non-destructive DNA extraction and voucher retrieval for small soft-bodied Arthropods in a high-throughput context: The example of Collembola. *Molecular Ecology Resources*, 10(6), 942–945. <https://doi.org/10.1111/j.1755-0998.2010.2839.x>
- Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, 7(3), 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>
- Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. *PLoS One*, 8(7), e66213. <https://doi.org/10.1371/journal.pone.0066213>
- Reid, F. A. (2009). *A field guide to the mammals of central america and southeast mexico*. Oxford, UK: Oxford University Press.
- Roslin, T., & Majaneva, S. (2016). The use of DNA barcodes in food web construction — terrestrial and aquatic ecologists unite!. *Genome*, 59(9), 603–628. <https://doi.org/10.1139/gen-2015-0229>
- Salinas-Ramos, V. B., Herrera Montalvo, L. G. G., León-Regagnon, V., Arrizabalaga-Escudero, A., & Clare, E. L. (2015). Dietary overlap and seasonality in three species of mormoopid bats from a tropical dry forest. *Molecular Ecology*, 24(20), 5296–5307. <https://doi.org/10.1111/mec.13386>
- Schloss, P., Westcott, S. L., Ryabin, T., Harrmann, M., Hollister, E. B., Lesniewski, R. A., ... Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75(23), 7537–7541. <https://doi.org/10.1128/AEM.01541-09>
- Sedlock, J. L., Krüger, F., & Clare, E. L. (2014). Island bat diets: Does it matter more who you are or where you live? *Molecular Ecology*, 23(15), 3684–3694.
- Simmons, N. B. (2005). Order Chiroptera. In D. E. Wilson, & D. M. Reeder (Eds.), *Mammal Species of the world: a taxonomic and geographic reference*, (Vol. 1, 3rd ed., pp. 312–529). Baltimore, MA: Johns Hopkins University Press.
- Smith, M. A., Woodley, N. E., Janzen, D. H., Hallwachs, W., & Hebert, P. D. N. (2006). DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3657–3662. <https://doi.org/10.1073/pnas.0511318103>
- Strona, G., & Lafferty, K. D. (2016). Environmental change makes robust ecological networks fragile. *Nature Communications*, 7, 12462. <https://doi.org/10.1038/ncomms12462>
- Symondson, W. O. C. (2002). Molecular identification of prey in predator diets. *Molecular Ecology*, 11(4), 627–641. <https://doi.org/10.1046/j.1365-294X.2002.01471.x>
- Thébault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329(5993), 853–856. <https://doi.org/10.1126/science.1188321>
- Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2016). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources*, 16(3), 714–726. <https://doi.org/10.1111/1755-0998.12490>
- Toju, H. (2015). High-throughput DNA barcoding for ecological network studies. *Population Ecology*, 57(1), 37–51. <https://doi.org/10.1007/s10144-014-0472-z>
- Vázquez, D. P., Morris, W. F., & Jordano, P. (2005). Interaction frequency as a surrogate for the total effect of animal mutualists on plants. *Ecology Letters*, 8(10), 1088–1094. <https://doi.org/10.1111/j.1461-0248.2005.00810.x>
- Wirta, H. K., Hebert, P. D. N., Kaartinen, R., Prosser, S. W., Várkonyi, G., & Roslin, T. (2014). Complementary molecular information changes our perception of food web structure. *Proceedings of the National Academy of Sciences of the United States of America*, 111(5), 1885–1890. <https://doi.org/10.1073/pnas.1316990111>
- Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., ... Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, 314(5800), 787–790. <https://doi.org/10.1126/science.1132294>
- Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, 11(2), 236–244. <https://doi.org/10.1111/j.1755-0998.2010.02920.x>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Clare EL, Fazekas AJ, Ivanova NV, et al. Approaches to integrating genetic data into ecological networks. *Mol Ecol*. 2019;28:503–519. <https://doi.org/10.1111/mec.14941>