Full title: Assessing the impact of taxon resolution on network structure

Short running title: Resolution in bipartite networks

David R. Hemprich-Bennett, (hemprich.bennett@gmail.com)^{1,2},

Hernani F. M. Oliveira, (oliveiradebioh@gmail.com)¹,

Steven C. Le Comber, (s.c.lecomber@qmul.ac.uk)^{1,3},

Stephen J. Rossiter, (s.j.rossiter@qmul.ac.uk)¹,

Elizabeth L. Clare, (e.clare@qmul.ac.uk)¹

Addresses:

¹School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Rd,
London, United Kingdom E1 4NS

²Department of Zoology, University of Oxford, Zoology Research and Administration Building,
11a Mansfield Road, Oxford, United Kingdom OX1 3SZ

³Deceased

Correspondence: David Hemprich-Bennett

Abstract

Constructing ecological networks has become an indispensable approach in understanding how different taxa interact. However, the methods used to generate data in network research varies widely among studies, potentially limiting our ability to compare results meaningfully. In particular, methods of classifying nodes vary in their precision, likely altering the architecture of the network studied. For example, rather than being classified as Linnaean species, taxa are regularly assigned to morphospecies in observational studies, or to Molecular Operational Taxonomic Units (MOTUs) in molecular studies, with the latter defined based on an arbitrary threshold of sequence similarity. Although the use of MOTUs in ecological networks holds great potential, especially for allowing rapid construction of large datasets of interactions, it is unclear how the choice of clustering threshold can influence the conclusions obtained. To test the impact of taxonomic precision on network architecture, we obtained and analyzed 16 datasets of ecological interactions, inferred from metabarcoding and observations. Our comparisons of networks constructed under a range of sequence thresholds for assigning taxa demonstrate that even small changes in node resolution can cause wide variation in almost all key metric values. Moreover, relative values of commonly used metrics such as robustness were seen to fluctuate continuously with node resolution, thereby potentially causing error in conclusions drawn when comparing multiple networks. In observational networks, we found that changing node resolution could, in some cases, lead to substantial changes to measurements of network topology. Overall, our findings highlight the importance of classifying nodes to the greatest precision possible, and demonstrate the need for caution when comparing networks that differ with respect to node resolution, even where taxonomic groups and interaction types are similar. In such cases, we

recommend that comparisons of networks should focus on relative differences rather than absolute values between the networks studied.

Key Words: food webs, metabarcoding, MOTU, network ecology, node resolution, species interactions

Introduction

The construction of ecological networks has become an indispensable approach in understanding how different taxa interact, as well as how such interactions are affected by biotic and abiotic factors (Baldock et al. 2015, Orford et al. 2016). It has become routine to generate networks to study diverse relationships, from mutualism (Jordano et al. 2003) to parasitism (Lafferty et al. 2006), carnivory (Wirta et al. 2015) and indirect interactions (Melián and Bascompte 2002). Researchers then typically assess these networks using a suite of metrics that quantify the diversity or distribution of interactions (Memmott et al. 2004, Kaiser-Bunbury and Blüthgen 2015).

Despite their increasing use, ecological networks frequently include unresolved nodes, where species identities are not known (Bascompte et al. 2003, Montoya et al. 2006, Pocock et al. 2012). Yet while the impacts of unresolved nodes and thus mixed resolution have been cited as a fundamental problem in network ecology (Ings et al. 2009), their consequences for the analysis and interpretation of ecological data have been largely overlooked. Work to date has mostly concentrated on unipartite networks, generating conflicting findings on the robustness of the network metrics to taxonomic resolution (Martinez 1993, Thompson and Townsend 2000, Woodward 2010). Bennett et al (2019) stated that in bipartite networks, various characteristics such as modularity and nestedness may be incorrectly measured if taxonomic resolution fails to capture the interactions accurately.

61 The potential problems surrounding imperfect node resolution are a longstanding issue for
62 traditional networks that, typically, rely on morphology. In such cases it is frequently impossible
63 to distinguish among cryptic taxa, especially where expert taxonomic identification is unavailable,
64 or where morphologically-diverse species may be misidentified as multiple species. As a result,
65 the ‘true’ nodes existing in nature may be erroneously clumped together or split in the network
66 dataset analyzed, and networks may contain a mixture where some nodes are classified to species,
67 and some to a higher taxonomic level or morphospecies. An increasing number of studies have
68 used molecular methods to identify species interactions as an alternative to morphology generating
69 greater resolution. For example, DNA barcoding has been shown to reveal more nodes in host-
70 parasitoid networks than could be seen from rearing data alone, with measurable changes in
71 network structure (Wirta et al. 2014). Despite this, DNA sequences might not always contain
72 sufficient phylogenetic information for node delimitation, potentially leading to the same mixed
73 resolution in networks cited as problematic in traditional analyses.

74 The development of high throughput sequencing (HTS) provides new opportunities in
75 ecology. In particular, network ecologists are now able to screen mixed samples for multiple taxa
76 and thereby obtain data from often numerous interactions at the same time (Pompanon et al. 2012).
77 These ‘metabarcoding’ techniques overcome the difficulty of observing ecological interactions
78 (Clare et al. 2009), and/or of inferring interactions where samples, such as stomach contents,
79 contain no identifiable remains (Piñol et al. 2013). A major challenge in current metabarcoding
80 research is interpreting the millions of sequences generated, which are frequently not possible to
81 fully identify due to the lack of reference sequences from known taxa. A common solution is to
82 augment identifications with sequences classified into Molecular Operational Taxonomic Units
83 (MOTUs), which are used as taxonomic proxies, including as nodes in interaction networks (Floyd

et al. 2002, Clare et al. 2016). MOTUs are best thought of as equivalent pools of genetic diversity partitioned by a uniformly-applied threshold of genetic divergence, but which may not be equivalent to accepted taxonomic levels (see original definition in Floyd et al. 2002). Previous results have shown that the generation of MOTUs can be sensitive to the choice of thresholds as well as to the algorithms used and other parameters; consequently, MOTU counts can vary by orders of magnitude (Flynn et al. 2015, Clare et al. 2016), with substantial differences in associated diversity estimates (Bachy et al. 2013, Egge et al. 2013). The MOTU sequence divergence threshold used can vary dramatically; bacterial studies originally used a 3% threshold as standard, (Yang et al. 2013) and this threshold has also been adopted by many within the eukaryotic metabarcoding community (Brown et al. 2015). However there is no special biological meaning behind this, and it was established for the 16S gene in bacteria rather than commonly used Eukaryotic loci. More relaxed thresholds have been applied (Salinas-Ramos et al. 2015) to limit MOTU inflation, and Amplicon Sequence Variants (ASVs), which rely on sequencing error profiles have become common in bacterial work, replacing a strict threshold approach (Callahan et al. 2016). Studies may similarly vary in other aspects that will inform the choice of MOTU threshold, including type of genetic marker (Wang et al. 2010), genomic region (Huber et al. 2009, Engelbrektson et al. 2010), target taxa (Pentinsaari et al. 2016), and expected level of sequencing error (Clare et al. 2016).

The impact of altering MOTU threshold (and thus number of nodes) on the results of metabarcoding studies has rarely been investigated. In a study of dietary overlap, Clare et al. (2016) found that altering clustering parameters significantly altered MOTU number but had minimal effect on measures of niche overlap when all samples were treated exactly the same way (thus any MOTU inflation or clumping was equivalent in all cases). In contrast, networks are likely to be

more sensitive to such changes, given that topology is critically dependent on the level of connectance among nodes (Poisot and Gravel 2014), and that the ability of an ecological network to withstand perturbations (stability) is thought to arise from the buffering effect of weak interactions (McCann 2000, Kéfi et al. 2019). The unknown effects of node resolution are also likely to apply to some traditional (observation based) networks, in which nodes may be resolved to different taxonomic levels within a single network (Ings et al. 2009), for example, in the presence of cryptic taxa (e.g. (Carvalho et al. 2008, Heleno et al. 2010, Pocock et al. 2012)).

To establish the impact of node delimitation on network architecture and its consequence for interpreting differences among networks, we collated multiple datasets of ecological interactions including both traditional observation-based and metabarcoding based data. For each dataset we then built networks using varying node resolutions and compared them using some of the most commonly-used network level metrics (Dormann et al. 2009). We made the following two predictions. First, we expected that across networks and data types, altering the resolution at which nodes are delimited would result in similar changes in measured network metrics. For example, when artificially reducing the number of nodes, the measured connectance of each network will increase, and no networks would be found to have a decreased connectance value. Second, we predicted that when comparing multiple ecological networks for a given metric, the rank order of the measured values will remain the same, such that metric x will always be greater in network a than in network b , regardless of the resolution of the network. Thus, our interpretation of how these networks differ from each other would not be affected by the manner of their generation. Our findings, however, revealed unexpected and inconsistent responses to changing levels of node delimitation within the molecular and observational datasets, highlighting potentially serious caveats in comparative studies of network dynamics.

Methods

To assess the impact of resolution on network measurements, we compared molecular and traditional networks. In order to be included, these datasets needed to meet stringent criteria for both data analysis procedure and resolution. Metabarcoding data is known to be influenced by factors such as the choice of sequencing locus, PCR primers, DNA extraction method and sequencer used (Zinger et al. 2019) and even the chemistry of the chosen sequencing protocol. In addition, many of the bioinformatic choices can influence how taxa are resolved (Deagle et al. 2019). In order to overcome these challenges and provide the most statistically sound analysis, we restricted the datasets to those generated using identical protocols at a single facility and then reanalyzed these ourselves using a single analytical pipeline to ensure consistency. This limits the number of uncontrolled variables introduced during data production. For example, we limited our choice to insect consuming species where more data was available but did not expand this to other feeding guilds, such as frugivory, which would require a different locus, primers, bioinformatic approach to construct a molecular network (e.g. MOTU clustering is not easily applied to plants). In observational networks similar stringent limitations on data were imposed. To reduce variability, we restricted the datasets to those of a single interaction type (in this case frugivory do to the general high resolution of nodes in these datasets), and then only included networks where every node was assigned to species level with full taxonomy available from the R package “Taxize” (Chamberlain and Szocs 2013) or a literature search, to avoid taxonomic inconsistencies at interim taxonomic levels.

Metabarcoding-based networks

Seven molecular datasets met the criteria for inclusion. These included guano samples collected from bats surveyed as part of studies conducted at sites in the USA (Gordon et al. 2018), Jamaica, Costa Rica and Malaysia (authors' unpublished data). All bats were captured under permit in either mist-nets or harp traps. For details of sites and trapping methods see Appendix S1: Table S1. To generate predator-prey datasets, we undertook metabarcoding of guano from individual insectivorous bats. Molecular procedures have been published elsewhere and PCR details are described in Appendix S1: Section S1. In brief, DNA was extracted using the QIAamp Stool Mini Kit (Qiagen, UK) with protocol modifications from Zeale et al. (2011) and Clare et al. (2014). Amplification, gel electrophoresis, amplicon size selection, clean up and sequencing were conducted at the Biodiversity Institute of Ontario, University of Guelph (Canada) using COI primers ZBJ-ArtF1c and ZBJ-ArtR2c (Zeale et al. 2011) modified with the dual adaptor system (Clare et al. 2014). Sequencing was performed on the Ion Torrent (Life Technologies) sequencing platform following Clare et al., (2014) with 192 samples (2 x 96 well plates) in a run using a 316 chip and following the manufacturer's guidelines but with a 2x dilution.

Sequences were de-multiplexed according to forward and reverse MIDs (allowing two mismatches and two indels). MIDs, primers and adapters were then removed (http://hannonlab.cshl.edu/fastx_toolkit). Amplicons of 147-167 bp were retained (target amplicon length = 157bp) and collapsed into unique haplotypes (http://hannonlab.cshl.edu/fastx_toolkit). All of these steps were performed in Galaxy (Afgan et al. 2016). We then removed singletons using a custom-written script.

For each dataset, we generated MOTUs using the Uclust algorithm (Edgar 2010) in QIIME (Caporaso et al. 2010) at 35 clustering similarity thresholds, from 0.91 to 0.98 with increments of 0.002. Sequence files were then converted into binary interaction matrices where nodes at the top

level were bat species and nodes at the bottom level were MOTU, where a value of 1 for a_{ij} denotes a positive interaction, of predator i consuming prey item j . To generate networks, the resulting interaction matrices were simplified by combining columns containing bats of the same species (e.g. if two individuals of species i consumed prey item j , $a_{ij} = 2$). As metabarcoding data is known to be subject to biases caused by choices such as the reagents, locus, PCR primers, DNA extraction method, and sequencer used (Zinger et al. 2019), we restricted the datasets in this analysis to ones generated identically at a single facility, and only using guano samples from insectivorous bats. This limits the number of uncontrolled variables involved in analysis (e.g. the inclusion of frugivory would require a different locus and different bioinformatics processing, and as clustering is not optimal for plant DNA, data produced this way could not be included here).

For each of the 245 networks generated (35 clustering thresholds x seven datasets) we calculated each of the metrics under the function `networklevel` in the ‘Bipartite’ R package (Dormann et al. 2008) using a custom wrapper script that is available as the package ‘LOTUS’ (<https://github.com/hemprichbennett/LOTUS>, DOI: 10.5281/zenodo.1297081), compiled for R (R Core Team 2019). All metrics were either classified as qualitative or quantitative, based on whether they are binary or incorporate information on interaction strength (see Appendix S1: Table S2).

Using these data, two sets of comparisons were made (see Appendix S1: Table S1). In the most severe scenario, seven datasets from diverse groups of bats in multiple continents, climatic conditions and habitat types are compared, referred to as the ‘Global molecular dataset’. We also compared a subset of two of these bat-arthropod datasets, which were sampled at the same location in Guanacaste, Costa Rica, in two consecutive seasons (wet and dry), referred to as the ‘Guanacaste molecular dataset (subset of two networks)’. This subset was collected for use in a separate ecological comparison study (currently in prep), and here serves as an example of how

comparisons of different treatments in a single ecological study can vary when network resolution is altered.

To assess the effect sizes of the clustering threshold, individual dataset, and the interaction between these terms, we used a linear mixed effects model in the R package ‘lme4’ (Bates et al. 2015) with the measured metric value fitted as the response variable, and dataset (e.g. dataset *A* from the USA, or dataset *B* from Malaysia), clustering threshold, and the interaction between dataset and clustering threshold as fixed effects, and allowed a random intercept for each dataset. In this, the effect strength of dataset or clustering on the response variable are of little interest: we expect datasets to have different structures and that using different clustering thresholds will affect the values of the metrics. That is, we expect the measured (absolute) values to change, but that the relative values will be similar (dataset *a* is always measured as providing a far greater value than dataset *b*). Of interest here is the interaction term, since a significant dataset*clustering threshold interaction suggests that the slopes of the datasets (judged by the metric in question) vary as a consequence of changing clustering threshold. While we expect that for a given metric the slopes will go in the same direction (i.e. all slopes will either be positive or negative), if the angle of the slopes vary substantially between datasets it shows that conclusions drawn when comparing networks can vary depending on the clustering threshold used when generating them. Thus, the *F* values of the interaction – the amount of variance in the model attributable to the interaction – is used as a measure of the extent to which the datasets respond differently to changes in threshold (strictly, whether the slopes of the relationship between clustering threshold and metric vary between networks). Of special interest was the ability to compare these interaction terms between multiple metrics, to see which metrics are most strongly affected by altered clustering thresholds. To create a measurement for this for each mixed effects model, the *F* values of the effect size of

the network’s identity were then divided by the F values for effect size of the interaction between network identity and clustering level.

From this same analysis, we also looked at the ranges over which the rank order of the different datasets was unchanged: i.e. where changing a threshold of dataset generation does not stop dataset *a* returning a greater value of a metric than dataset *b*. This is of interest as the clustering thresholds which most accurately represent the underlying taxonomic diversity are unknown, however if there are ranges of clustering thresholds within which network topology are relatively unchanged, they may represent a range which is relatively reliable for network metabarcoding. We did not include compartment diversity or number of species on the higher network level in this analysis, as these values were unchanged. All molecular analyses are available in the GitHub repository https://github.com/hemprichbennett/network_otus.

Observational Networks

A total of nine published datasets were compiled of interactions between plants and vertebrate frugivores from the Galapagos and the Canary Islands (Nogales et al. 2016), and interactions between plants and frugivorous birds in Brazil (Galetti and Pizo 1996), Japan (Noma 1996), Malaysia (Lambert 1989), Mexico (Kantak 1979) and Spain (Jordano 1985, Rezende et al. 2007). These datasets were unique in that all nodes of both network levels were resolved at species-level. We then retrieved the corresponding order, family and genus level data from online databases using the package ‘Taxize’ (Chamberlain and Szocs 2013). The resulting dataset is referred to as the ‘Global observational dataset’.

To determine the impact of incomplete node resolution on network architecture for each of these datasets, we reanalyzed the interactions by relabelling a given proportion of randomly selected nodes so as to reduce the taxonomic resolution; for example, species names were replaced

with the corresponding genus. If two nodes then had the same identity, they were collapsed together to become a single node with the sum of its parent nodes' interactions. Thus if *Solanum lycopersicum* and *S. vespertilio* were both simplified to become *Solanum*, there would now be a single *Solanum* node containing the sum of their interactions. For a given proportion of randomly selected nodes, re-labelling was then performed for increasing proportions at increments of 0.1, until all nodes were relabelled (i.e. 0.1 to 1.0). Because of the random nature of the re-labelling, this was repeated 100 times for each network level, taxonomic level (Genus, Family and Order), and proportion of nodes. As the datasets from Nogales et al (2016) were binary (noting only the presence or absence of an interaction), the subsequent networks made by relabelling these two networks were also constrained to be binary and they were not included in the analyses of any quantitative metrics.

We then measured each available network-level metric using the 'Bipartite' package (Dormann et al. 2008) to summarize the structure of each of the simplified networks. To determine the impact of imperfect node resolution on network structure, we ran mixed effects models using the R package 'lme4' (Bates et al. 2015). In our models the metric value was the response variable, the proportion of nodes relabelled, the network level (plant or seed consumer) being relabelled, and the taxonomic level being relabelled to were fitted as fixed effects, and the dataset being relabelled was fitted as a random effect with a random intercept. To visualize the changes occurring to the network metric measurements with node relabelling, we made pairwise comparisons between each of the simplified networks. To aid interpretation, these comparisons were restricted to iterations where the same network level was being relabelled and to the same taxonomic level (e.g. only comparing iterations where the frugivores were relabelled, and only being relabelled to Order level). We then plotted the percentage of these pairwise comparisons in which at least one

combination of iterations gave an erroneous rank order: e.g. if for each of the 36 pairs of networks, there was a combination of the relabelling events where relabelling 10% of nodes from network *a* and 20% of nodes from network *b* gave a rank order differing from that of the original networks, this location would receive a value of 100%. All observational analyses are available in the GitHub repository https://github.com/hemprichbennett/network_clustering_observations.

Focal metrics

For analyses of both metabarcoding and observation-based networks, we focus in particular on the metrics functional complementarity (Blüthgen and Klein 2011), H2' (Blüthgen et al. 2007), modularity (Newman and Girvan 2004, Dormann and Strauss 2014), nestedness temperature (Atmar and Patterson 1993), NODF (Almeida-Neto et al. 2008), and robustness (Memmott et al. 2004). Functional complementarity (Blüthgen and Klein 2011, Devoto et al. 2012, Peralta et al. 2014), calculates the level to which nodes in a network level have complementary non-overlapping interactions, through measuring the branch lengths of a functional dendrogram of their interaction dissimilarity with values between 0 (no complementarity), and 1 (perfect complementarity). H2' is a measure of the specialization of both levels of a bipartite network, (Blüthgen et al. 2006), designed for comparing the specialization between multiple networks. Modularity measures the level to which species interactions form discrete clusters of dense interactions, with values between 0 (no modularity) and 1 (perfectly modular structure) (Newman and Girvan 2004, Dormann and Strauss 2014). Nested temperature (Atmar and Patterson 1993) and NODF (Almeida-Neto et al. 2008) are two metrics describing the 'nestedness' of a network: the level to which the interactions of the specialists in a network are a subset of the interactions of the generalists. 'Robustness' is a measure of how tolerant a system is to extinction cascades, measuring the area underneath the curve of a plotted secondary extinction simulation (Memmott

et al. 2004). These metrics were chosen as they are among the most commonly used metrics in network ecology, and are relatively independent of sampling effects (Fründ et al. 2016). As such they are felt to be of especial interest to the ecological community, and able to give more reliable conclusions for this study than other metrics. Results and plots from the analyses of all other network-level metrics can be found in Appendix S1.

Results

Metabarcoding-based networks

Our analyses of the Global molecular dataset (seven networks) revealed that the absolute values of most metrics were sensitive to the MOTU clustering threshold applied (Figure 1 and 2), reflecting changes in underlying network structure. Trends in summary metrics with MOTU threshold were seen to differ in both the magnitude and/or the direction. For example, the metric ‘functional complementarity’ for the higher network level (i.e. bats) showed an increase with threshold for some datasets, but a decrease for others, with a high F value associated with the interaction term (Figure 2). In contrast, the metric ‘NODF’ showed relatively consistent directional responses to threshold, as seen by an intermediate F value (Figure 2).

Due to this variation in the behavior of metrics with changes in threshold, the resulting final rank order to the datasets was also seen to vary depending on the metric used for a given MOTU threshold. For example, while we observed no change in the rank order of the datasets based on ‘nestedness’, the rank order based on ‘robustness’ switched almost continuously throughout all thresholds used (Figures 1 and 3). Thus, we found that in our largest comparisons

between all molecular networks the outcome was critically dependent on the precise choice of threshold.

Our more restricted comparison of the Guanacaste molecular dataset (subset of two networks) was generated from data collected during separate seasons (wet and dry). The two networks in this dataset were thus predicted to be relatively similar, and yielded considerably more consistent conclusions than the Global molecular dataset (seven networks) (Figures 3 and 4). Although absolute values of metrics typically varied in response to threshold, the rank order of metrics derived for the two datasets was more stable than that recorded in the case of the Global molecular dataset. For example, the metric ‘robustness’ was always higher for the wet season than the dry season, thereby preserving the rank order of this pair of datasets (Figure 4), compared to the former comparison of seven datasets in which the rank order of this metric varied considerably.

Observational networks

Our analyses of the Global observational dataset showed that, for the majority of metrics, conclusions based on the rank order were sensitive to the proportion of nodes being relabelled. In pairwise comparisons of each dataset, at various combinations of proportions of nodes relabelled (Figure 5), it was possible to create an erroneous rank order for each of the focal metrics when coarsely identifying nodes. We found that the focal metrics appeared to differ in their sensitivity to node relabelling, showing a pattern of erroneous findings being more likely when identifying the higher network level (seed consumers) coarsely than for the lower network level (plants). The changes in conclusions were most marked when identifying nodes to Order level, with increasing reliability when identifying to Family and Genus. Such changes were possible when coarsely identifying a low proportion of nodes (0.1-0.25). For every focal metric, it was possible to alter

the rank order by coarsely identifying some nodes, but the identity of the dataset (used as a random effect in the model) was still the most important in predicting the metric value (Table 1).

Discussion

Our analyses of observational and molecular datasets reveal that node resolution critically impacts the structure of ecological networks, and that this can lead to wide variation in the magnitude and behavior of commonly reported metric values. We further show that inherent variation in measured values can lead to erroneous conclusions in comparisons of networks, although these problems appear less evident in comparisons of ecologically-matched datasets. These findings therefore have important implications for the issue of node resolution, a long-standing challenge in network ecology that has become a topic of increasing interest in light of the proliferation of sequence data.

Resolution and ecological network analysis

Newly available DNA metabarcoding approaches are expected to be transformative in ecological network research by allowing large volumes of data to be generated rapidly (Kaartinen et al. 2010, Wirta et al. 2014, Evans et al. 2016). Unlike traditional approaches to network construction, in which interacting taxa are commonly identified based on observations, these methods rely on the concept of MOTUs. Despite these differences in methodology, our comparison revealed that both types of method are prone to related issues.

A key result was that in both observation-based and metabarcoding-based networks, altering taxonomic resolution led to often dramatic changes in the numbers of nodes, which in the latter case varied by several orders of magnitude. This is worrying because the number of nodes, and their consequence for connectance, are widely considered strong determinants of many

354 frequently-measured characteristics of network structure (Poisot and Gravel 2014, Chagnon 2015).
355 For example, higher numbers of nodes will increase the proportion of weak links in networks,
356 whereas reducing the number of nodes may cause networks to appear more generalized. Such
357 trends also have broad implications for theoretical interpretations, with the distribution of link
358 strength seen to play a pivotal role in the stability of ecosystems (McCann 2000, Solé and Montoya
359 2001).

360 Other key network metrics that showed strong responses to node resolution included those
361 based on interaction distribution. In some cases, such as robustness, this led to widespread
362 variation in the rank order of networks. Robustness for the higher network level showed a rapid
363 increase with an increasing number of nodes on the lower network level, showing that an inflated
364 estimation of the dietary richness available to a consumer reduces the perceived likelihood of
365 extinction of higher node species. Robustness is commonly used in forecasting ecosystem
366 resilience to species loss, and has been linked to ecological restoration (Pocock et al. 2012). When
367 numbers of nodes increased, this was associated with increases in metrics of specialization, such
368 as $H2'$. However, as the magnitude of this effect varied between datasets it frequently caused
369 switches in their rank order (Figures 2 and 3). Interestingly, altering node resolution had a large
370 effect on network specialization as measured by $H2'$ for molecular datasets (Figures 1-3), but not
371 for the observational datasets (Figures 5 and 6). The molecular datasets may be more susceptible
372 to node resolution inducing changes in measures of specialization, as few bat nodes are interacting
373 with a high number of MOTU nodes. Molecular networks typically have more nodes than networks
374 built upon observations (Wirta et al. 2014, Macgregor et al. 2019) and so although bats have
375 especially diverse diets, it is reasonable to expect such datasets to be typified by high number of
376 nodes and a skewed degree distribution. We also found that descriptors of ecological interactions

among taxa at the same network level were highly labile. For example, some metrics related to niche-use, such as niche overlap (Rudolf and Lafferty 2011) and C-score (Stone and Roberts 1990), varied widely, possibly due to inflated resource partitioning arising from the over-splitting of MOTUs (Clare 2014). On the other hand, we found that functional complementarity – an alternative measure of niche differentiation based on distance matrices (Devoto et al. 2012, Peralta et al. 2014) – was less sensitive to threshold used, giving fewer alterations in rank order in the molecular dataset. Nestedness describes the extent to which interactions involving specialists comprise subsets of those involving generalists, and is a pattern seen across diverse networks in nature (Nielsen and Bascompte 2007). Our analyses show that with an increasing clustering threshold, both nestedness metrics used here decreased with node resolution (Almeida-Neto et al. 2008, Atmar and Patterson 1993). The network level being clustered in the observational analysis frequently had an effect on any pairwise comparisons being made between networks (Figure 5), although for each focal metric the effect size of the network level being clustered was smaller than that of the dataset being clustered (Table 1). Pairwise comparisons were generally most-affected by clustering the seed consumers, possibly because there were typically a lower number of taxonomic Families and Orders of seed consumers than plants in each network.

Our findings on the impact of node resolution complement previous assertions that network dimension and sampling intensity may affect multiple network metrics (Dormann et al. 2009). Fründ et al. (2016) demonstrated that qualitative metrics summarizing ecological specialization (e.g. generality) are especially sensitive to sample size, but argued that where such biases were predictable, these metrics still hold value provided that interpretations are restricted to relative values. On the other hand, quantitative analogues that take account of interaction strength were reported to be more robust to sample sizes (Fründ et al. 2016), a result also supported by our

own observations from node resolution (Figure 2). Weighted networks and metrics contain far more information than their binary counterparts, and so even when multiple nodes are collapsed into a single node, the information loss is minimal compared to that retained by the rest of the network.

There is significant debate about the extent to which metabarcoding-based research can be quantitative (Lamb et al. 2019). Arguments have been made for incorporating the number of sequences obtained per MOTU in a sample as a proxy for the biomass consumed (Deagle et al. 2019), or instead simply using the presence or absence of a MOTU in a sample (Clare 2014, Andriollo et al. 2019). In the metabarcoding-based component of this study we used the presence or absence of a prey item in an individual bat, and then summed the interactions of all bats of a given species. Whilst this should not be interpreted as providing information on the biomass in a given ecological interaction, it is generally felt to be the most reliable approach for the study of generalist consumers (Andriollo et al. 2019).

Our results show that resolution is a problem common to networks based on both DNA barcoding and observations. In observation-based networks we found that low levels of relabelling, representing a coarse identification of only a low proportion of nodes, was frequently enough to change the rank order of network comparisons. Given that the difficulty of identifying all nodes to species level means many published networks include a mix of species-level and more coarsely identified nodes, this challenges the reliability of such studies' findings (e.g. Rezende et al. 2009, Pocock et al. 2012, Baldock et al. 2015, Kantsa et al. 2018). The effects of relabelling on network measurements appears to have been more severe when used on the observational dataset (Figures 5 and 6), which is problematic as in such networks it is typically more difficult to identify nodes by visual than molecular means. Such issues will continue to be present in studies that either

incorporate novel technologies such as cameras (Gray et al. 2018, Sritongchuay et al. 2019a, 2019b), or avoid classifying nodes to species level and instead opt to classify them to non-taxonomic levels, such as functional groups. These results highlight the need for network ecologists to identify all nodes to uniform resolution with the greatest level of precision possible and importantly, to use identical methods and resolution for the comparisons of any networks. Moreover, when taxa are common to multiple networks in a study, it is crucial for them to be identified to the same taxonomic resolution within each treatment. When nodes in networks are unresolved, they can either represent cases where species are erroneously clumped together into single nodes, or single species may be erroneously used as multiple nodes. These two phenomena are represented here by the observational (erroneous clumping) and MOTU (presumed erroneous splitting) analyses.

In the context of metabarcoding, which looks set to become an important tool in network ecology, the assigning of sequences to species is highly challenging, especially where sequences are short and contain limited information. Steps towards achieving a solution might involve combining data from multiple loci, or, where samples contain sufficiently intact DNA, generating longer sequences. Regardless, it is important to recognize that one or few loci will rarely resolve all species in a complex data set, and network ecologists will thus continue to rely on MOTUs at least in part. While most programs to date classify MOTUs by splitting genetic diversity according to a single threshold, it is well known that interspecific divergence will vary widely across both loci and taxonomic groups (Johns and Avise 1998, Pentinsaari et al. 2016). Emerging approaches offer the means to balance over-splitting of MOTUs against retaining sequencing errors (Frøslev et al. 2017), however, ultimately an adaptive approach- in which specific thresholds can be fitted to different taxonomic groups- might further aid taxonomic precision. Molecular analyses are also

generally unable to detect cannibalism in taxonomic groups where it is likely to occur (Berry et al. 2015) and in some scenarios can be prone to false positives due to secondary ingestion; where an item ingested by prey is detected as secondary predation (Toju and Baba 2018). As in traditional networks, it is vital that the exact same molecular and bioinformatics procedures be used in the comparison of any two networks, and to aid future comparisons all code should be made openly available for transparency. It is encouraging that in the subset of our metabarcoding data collected using a paired protocol and designed specifically for comparative study, findings were much less vulnerable to showing changes in conclusions based on network resolution (Figures 3 and 4). To reduce potential uncontrolled variation caused by comparing networks with multiple interaction types, the analyses reported here are restricted to those of bat-insect networks and observational networks of frugivory. For the metabarcoding networks it was necessary to restrict our analysis to data generated using a single laboratory protocol and sequencing technology (Ion-Torrent), to control for multiple factors known to influence metabarcoding studies (Zinger et al. 2019). Analysis protocols are often specific to the primer used, the error profile of the sequencing platform, and clustering protocols such as the MOTU approach used here are not appropriate for some taxonomic approaches. For just one example, different primers will generate amplicons of different lengths. Sequence data of different lengths clustered at the same percentage cut off point would generate different estimates of node resolution (1% of 100bp is different than 1% of 200bp) introducing an uncontrolled factor in statistical analysis. As a consequence, we have limited our inclusion of data to a strict criteria to control these variables. Our analysis of observational networks of frugivory required us to only use datasets in which each node was identified to species level, however this also restricted the data available to us in a similar way. While we expect our

findings to be broadly generalizable, we note that extrapolations from this relatively small number of networks should be made with caution.

Finally, we conclude that our ability to make meaningful interpretations regarding ecological networks critically depends on the nature of the underlying data and its processing. We further show that precise metric values can be arbitrary, and while relative values in comparative studies may be more reliable, effect sizes are likely to be the most important criteria when deciding if these values are biologically meaningful. Regardless of the technique used to generate network data, issues are likely to persist in the delimitation of nodes and therefore any conclusions drawn from study of the network as a whole. As such, we recommend that researchers identify network nodes to the greatest precision possible, and acknowledge the limitations of their datasets. Overall, we suggest that caution must be taken when comparing values from multiple networks, especially where node resolution differs.

Acknowledgements

Dedicated to the memory of Steven Le Comber (1966-2019). We thank Rob Knell for discussions about our analysis. For help with data collection in Malaysia we thank Victoria Kemp, Jamiluddin Jami, Arnold James, Mohd. Mustamin, Ampat Siliwong, Sabidee Mohd. Rizan and Najmuddin Jamal. We also thank Henry Bernard, Eleanor Slade, Owen Lewis, Matthew Struebig and other members of the LOMBOK consortium for facilitating research in Sabah, and we are especially grateful to the Sabah Biodiversity Council (permit JKM/MBS.1000-2/2 (374)), Sabah Forest Department, Yayasan Sabah, Sime Darby and Benta Wawasan for granting access. For data collection in Costa Rica, we thank Bernal Rodriguez Herrera, Daniel Janzen, Winnie Hallawachs, Roger Blanco, Maria Marta Chavarria and Alejandro Masis for all the support conducting this research in Sector Santa Rosa (of ACG); Edgar Jimenes for the help during fieldwork. Research

was performed under permit R-07-2015-OT-CONAGEBIO and R-08-2015-OT-CONAGEBIO, from the Ministry of Environment and Telecommunications (MINAET) and Comisión Nacional para la Gestión de la Biodiversidad (CONAGEBIO). For data collection in Jamaica, we thank Susan König and Matthew Emrich for access to unpublished network data. For the Big Bend dataset, we thank Loren Ammerman, Sally Ivens, Rowena Gordon, Joanne Littlefair, John Ratcliffe and M. Brock Fenton for contributing their network data. Big Bend data were collected in 2011 under permit 2011 BIBE-2009-SCI-0013, 2014 under BIBE-2013-SCI-0023, 2015 under SPR-0994-703 and 2016 under BIBE-2015-SCI-0022. This research utilized Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT (<http://doi.org/10.5281/zenodo.438045>) and the University of Oxford's Advanced Research Computing (ARC) facility (<http://dx.doi.org/10.5281/zenodo.22558>). This work was funded by studentships awarded to DRHB (Queen Mary University of London) and HFM (Science Without Borders, Brazil), a Natural Environment Research Council grant (NE/K016407/1) to SJR, and a Royal Society grant (RG130793) to ELC.

Statement of authorship: DRHB, SJR and ELC conceived of the project, DRHB facilitated fieldwork in Malaysia, DRHB, HFMO and ELC undertook field collections, DRHB, SCLC and HFMO analyzed the data, and DRHB wrote the manuscript with input from all authors.

Literature cited

- Afgan, E., D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Grüning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, and J. Goecks. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research* 44:W3–W10.
- Almeida-Neto, M., P. Guimarães, P. R. Guimarães, R. D. Loyola, and W. Ulrich. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* 117:1227–1239.
- Andriollo, T., F. Gillet, J. R. Michaux, and M. Ruedi. 2019. The menu varies with metabarcoding practices: A case study with the bat *Plecotus auritus*. *PLOS ONE* 14:e0219135.
- Atmar, W., and B. D. Patterson. 1993. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* 96:373–382.
- Bachy, C., J. R. Dolan, P. López-García, P. Deschamps, and D. Moreira. 2013. Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *The ISME Journal* 7:244–255.
- Baldock, K. C. R., M. A. Goddard, D. M. Hicks, W. E. Kunin, N. Mitschunas, L. M. Osgathorpe, S. G. Potts, K. M. Robertson, A. V. Scott, G. N. Stone, I. P. Vaughan, and J. Memmott. 2015. Where is the UK’s pollinator biodiversity? The importance of urban areas for flower-visiting insects. *Proceedings of the Royal Society B: Biological Sciences* 282:20142849.

- 531 Bascompte, J., P. Jordano, C. J. Melián, and J. M. Olesen. 2003. The nested assembly of plant–
532 animal mutualistic networks. *Proceedings of the National Academy of Sciences of the*
533 *United States of America* 100:9383–9387.
- 534 Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models
535 Using lme4. *Journal of Statistical Software* 67:1–48.
- 536 Bennett, A. E., D. M. Evans, and J. R. Powell. 2019. Potentials and pitfalls in the analysis of
537 bipartite networks to understand plant–microbe interactions in changing environments.
538 *Functional Ecology* 33:107–117.
- 539 Berry, O., C. Bulman, M. Bunce, M. Coghlan, D. C. Murray, and R. D. Ward. 2015. Comparison
540 of morphological and DNA metabarcoding analyses of diets in exploited marine fishes.
541 *Marine Ecology Progress Series* 540:167–181.
- 542 Blüthgen, N., and A.-M. Klein. 2011. Functional complementarity and specialisation: The role of
543 biodiversity in plant–pollinator interactions. *Basic and Applied Ecology* 12:282–291.
- 544 Blüthgen, N., F. Menzel, and N. Blüthgen. 2006. Measuring specialization in species interaction
545 networks. *BMC Ecology* 6.
- 546 Blüthgen, N., F. Menzel, T. Hovestadt, B. Fiala, and N. Blüthgen. 2007. Specialization,
547 Constraints, and Conflicting Interests in Mutualistic Networks. *Current Biology* 17:341–
548 346.
- 549 Brown, E. A., F. J. J. Chain, T. J. Crease, H. J. MacIsaac, and M. E. Cristescu. 2015. Divergence
550 thresholds and divergent biodiversity estimates: can metabarcoding reliably describe
551 zooplankton communities? *Ecology and Evolution* 5:2234–2251.

- 552 Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes.
553 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature*
554 *Methods* 13:581–583.
- 555 Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer,
556 A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E.
557 Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder,
558 J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld,
559 and R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing
560 data. *Nature Methods* 7:335–336.
- 561 Carvalho, L. G., Y. M. Buckley, R. Ventim, S. V. Fowler, and J. Memmott. 2008. Apparent
562 competition can compromise the safety of highly specific biocontrol agents. *Ecology*
563 *Letters* 11:690–700.
- 564 Chagnon, P.-L. 2015. Characterizing topology of ecological networks along gradients: The limits
565 of metrics' standardization. *Ecological Complexity* 22:36–39.
- 566 Chamberlain, S., and E. Szocs. 2013. taxize - taxonomic search and retrieval in R. *F1000Research*.
- 567 Clare, E. L. 2014. Molecular detection of trophic interactions: emerging trends, distinct advantages,
568 significant considerations and conservation applications. *Evolutionary Applications*
569 7:1144–1157.
- 570 Clare, E. L., F. J. J. Chain, J. E. Littlefair, and M. E. Cristescu. 2016. The effects of parameter
571 choice on defining molecular operational taxonomic units and resulting ecological analyses
572 of metabarcoding data. *Genome* 59:981–990.

- 573 Clare, E. L., E. E. Fraser, H. E. Braid, M. B. Fenton, and P. D. N. Hebert. 2009. Species on the
574 menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular
575 approach to detect arthropod prey. *Molecular Ecology* 18:2532–2542.
- 576 Clare, E. L., W. O. C. Symondson, H. Broders, F. Fabianek, E. E. Fraser, A. MacKenzie, A.
577 Boughen, R. Hamilton, C. K. R. Willis, F. Martinez-Nuñez, A. K. Menzies, K. J. O.
578 Norquay, M. Brigham, J. Poissant, J. Rintoul, R. M. R. Barclay, and J. P. Reimer. 2014.
579 The diet of *Myotis lucifugus* across Canada: assessing foraging quality and diet variability.
580 *Molecular Ecology* 23:3618–3632.
- 581 Deagle, B. E., A. C. Thomas, J. C. McInnes, L. J. Clarke, E. J. Vesterinen, E. L. Clare, T. R.
582 Kartzinel, and J. P. Eveson. 2019. Counting with DNA in metabarcoding studies: How
583 should we convert sequence reads to dietary data? *Molecular Ecology* 28:391–406.
- 584 Devoto, M., S. Bailey, P. Craze, and J. Memmott. 2012. Understanding and planning ecological
585 restoration of plant–pollinator networks. *Ecology Letters* 15:319–328.
- 586 Dormann, C. F., J. Frueund, N. Bluethgen, and B. Gruber. 2009. Indices, graphs and null models:
587 analyzing bipartite ecological networks. *The Open Ecology Journal* 2:7–24.
- 588 Dormann, C. F., B. Gruber, and J. Fruend. 2008. Introducing the bipartite package: Analysing
589 ecological networks. *R News* 8:8–11.
- 590 Dormann, C. F., and R. Strauss. 2014. A method for detecting modules in quantitative bipartite
591 networks. *Methods in Ecology and Evolution* 5:90–98.
- 592 Egge, E., L. Bittner, T. Andersen, S. Audic, C. de Vargas, and B. Edvardsen. 2013. 454
593 Pyrosequencing to Describe Microbial Eukaryotic Community Composition, Diversity and
594 Relative Abundance: A Test for Marine Haptophytes. *PLOS ONE* 8:e74371.

- Engelbrektson, A., V. Kunin, K. C. Wrighton, N. Zvenigorodsky, F. Chen, H. Ochman, and P. Hugenholtz. 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *The ISME journal* 4:642–647.
- Evans, D. M., J. J. N. Kitson, D. H. Lunt, N. A. Straw, and M. J. O. Pocock. 2016. Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology* 30:1904–1916.
- Floyd, R., E. Abebe, A. Papert, and M. Blaxter. 2002. Molecular barcodes for soil nematode identification. *Molecular Ecology* 11:839–850.
- Flynn, J. M., E. A. Brown, F. J. J. Chain, H. J. MacIsaac, and M. E. Cristescu. 2015. Toward accurate molecular identification of species in complex environmental samples: testing the performance of sequence filtering and clustering methods. *Ecology and Evolution* 5:2252–2266.
- Frøslev, T. G., R. Kjøller, H. H. Bruun, R. Ejrnæs, A. K. Brunbjerg, C. Pietroni, and A. J. Hansen. 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications* 8:1–11.
- Fründ, J., K. S. McCann, and N. M. Williams. 2016. Sampling bias is a challenge for quantifying specialization and network structure: lessons from a quantitative niche model. *Oikos* 125:502–513.
- Galetti, M., and M. A. Pizo. 1996. Fruit eating by birds in a forest fragment in southeastern Brazil. *Revista Brasileira de Ornitologia - Brazilian Journal of Ornithology* 4:71–79.
- Gordon, R., S. Ivens, L. K. Ammerman, M. B. Fenton, J. E. Littlefair, J. M. Ratcliffe, and E. L. Clare. 2018. Molecular diet analysis finds an insectivorous desert bat community

- 617 dominated by resource sharing despite diverse echolocation and foraging strategies.
618 Ecology and Evolution.
- 619 Gray, R. E. J., R. M. Ewers, M. J. W. Boyle, A. Y. C. Chung, and R. J. Gill. 2018. Effect of tropical
620 forest disturbance on the competitive interactions within a diverse ant community.
621 Scientific Reports 8:1–12.
- 622 Heleno, R., I. Lacerda, J. A. Ramos, and J. Memmott. 2010. Evaluation of restoration effectiveness:
623 community response to the removal of alien plants. Ecological Applications 20:1191–1203.
- 624 Huber, J. A., H. G. Morrison, S. M. Huse, P. R. Neal, M. L. Sogin, and D. B. Mark Welch. 2009.
625 Effect of PCR amplicon size on assessments of clone library microbial diversity and
626 community structure. Environmental Microbiology 11:1292–1302.
- 627 Ings, T. C., J. M. Montoya, J. Bascompte, N. Blüthgen, L. Brown, C. F. Dormann, F. Edwards, D.
628 Figueroa, U. Jacob, J. I. Jones, R. B. Lauridsen, M. E. Ledger, H. M. Lewis, J. M. Olesen,
629 F. J. F. V. Veen, P. H. Warren, and G. Woodward. 2009. Review: Ecological networks –
630 beyond food webs. Journal of Animal Ecology 78:253–269.
- 631 Johns, G. C., and J. C. Avise. 1998. A comparative summary of genetic distances in the vertebrates
632 from the mitochondrial cytochrome b gene. Molecular Biology and Evolution 15:1481–
633 1490.
- 634 Jordano, P. 1985. El ciclo anual de los Paseriformes frugívoros en el matorral mediterráneo del sur
635 de España: importancia de su invernada y variaciones interanuales. Ardeola 32:69–94.
- 636 Jordano, P., J. Bascompte, and J. M. Olesen. 2003. Invariant properties in coevolutionary networks
637 of plant–animal interactions. Ecology Letters 6:69–81.

- 638 Kaartinen, R., G. N. Stone, J. Hearn, K. Lohse, and T. Roslin. 2010. Revealing secret liaisons:
639 DNA barcoding changes our understanding of food webs. *Ecological Entomology* 35:623–
640 638.
- 641 Kaiser-Bunbury, C. N., and N. Blüthgen. 2015. Integrating network ecology with applied
642 conservation: a synthesis and guide to implementation. *AoB PLANTS* 7:plv076.
- 643 Kantak, G. E. 1979. Observations on Some Fruit-Eating Birds in Mexico. *The Auk* 96:183–186.
- 644 Kantsa, A., R. A. Raguso, A. G. Dyer, J. M. Olesen, T. Tscheulin, and T. Petanidou. 2018.
645 Disentangling the role of floral sensory stimuli in pollination networks. *Nature*
646 *Communications* 9:1041.
- 647 Kéfi, S., V. Domínguez-García, I. Donohue, C. Fontaine, E. Thébault, and V. Dakos. 2019.
648 Advancing our understanding of ecological stability. *Ecology Letters* 22:1349–1356.
- 649 Lafferty, K. D., A. P. Dobson, and A. M. Kuris. 2006. Parasites dominate food web links.
650 *Proceedings of the National Academy of Sciences* 103:11211–11216.
- 651 Lamb, P. D., E. Hunter, J. K. Pinnegar, S. Creer, R. G. Davies, and M. I. Taylor. 2019. How
652 quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* 28:420–
653 430.
- 654 Lambert, F. 1989. Fig-eating by birds in a Malaysian lowland rain forest. *Journal of Tropical*
655 *Ecology* 5:401–412.
- 656 Macgregor, C. J., J. J. N. Kitson, R. Fox, C. Hahn, D. H. Lunt, M. J. O. Pocock, and D. M. Evans.
657 2019. Construction, validation, and application of nocturnal pollen transport networks in
658 an agro-ecosystem: a comparison using light microscopy and DNA metabarcoding.
659 *Ecological Entomology* 44:17–29.
- 660 Martinez, N. D. 1993. Effects of Resolution on Food Web Structure. *Oikos* 66:403–412.

- 661 McCann, K. S. 2000. The diversity–stability debate. *Nature* 405:228–233.
- 662 Melián, C. J., and J. Bascompte. 2002. Food web structure and habitat loss. *Ecology Letters* 5:37–
663 46.
- 664 Memmott, J., N. M. Waser, and M. V. Price. 2004. Tolerance of pollination networks to species
665 extinctions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*
666 271:2605–2611.
- 667 Montoya, J. M., S. L. Pimm, and R. V. Solé. 2006. Ecological networks and their fragility. *Nature*
668 442:259–264.
- 669 Newman, M. E. J., and M. Girvan. 2004. Finding and evaluating community structure in networks.
670 *Physical Review E* 69:026113.
- 671 Nielsen, A., and J. Bascompte. 2007. Ecological networks, nestedness and sampling effort. *Journal*
672 *of Ecology* 95:1134–1141.
- 673 Nogales, M., R. Heleno, B. Rumeu, A. González-Castro, A. Traveset, P. Vargas, and J. M. Olesen.
674 2016. Seed-dispersal networks on the Canaries and the Galápagos archipelagos: interaction
675 modules as biogeographical entities. *Global Ecology and Biogeography* 25:912–922.
- 676 Noma, N. 1996. Annual Fluctuations of Sapfruits Production and Synchronization within and inter
677 Species in a Warm Temperate Forest on Yakushima Island. *Tropics* 6:441–449.
- 678 Orford, K. A., P. J. Murray, I. P. Vaughan, and J. Memmott. 2016. Modest enhancements to
679 conventional grassland diversity improve the provision of pollination services. *Journal of*
680 *Applied Ecology* 53:906–915.
- 681 Pentinsaari, M., H. Salmela, M. Mutanen, and T. Roslin. 2016. Molecular evolution of a widely-
682 adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports* 6:35275.

- 683 Peralta, G., C. M. Frost, T. A. Rand, R. K. Didham, and J. M. Tylianakis. 2014. Complementarity
684 and redundancy of interactions enhance attack rates and spatial stability in host–parasitoid
685 food webs. *Ecology* 95:1888–1896.
- 686 Piñol, J., V. San Andrés, E. L. Clare, G. Mir, and W. O. C. Symondson. 2013. A pragmatic
687 approach to the analysis of diets of generalist predators: the use of next-generation
688 sequencing with no blocking probes. *Molecular Ecology Resources* 14:18–26.
- 689 Pocock, M. J. O., D. M. Evans, and J. Memmott. 2012. The Robustness and Restoration of a
690 Network of Ecological Networks. *Science* 335:973–977.
- 691 Poisot, T., and D. Gravel. 2014. When is an ecological network complex? Connectance drives
692 degree distribution and emerging network properties. *PeerJ* 2:e251.
- 693 Pompanon, F., B. E. Deagle, W. O. C. Symondson, D. S. Brown, S. N. Jarman, and P. Taberlet.
694 2012. Who is eating what: diet assessment using next generation sequencing. *Molecular*
695 *Ecology* 21:1931–1950.
- 696 R Core Team. 2019. R: A language and environment for statistical computing. Vienna, Austria.
- 697 Rezende, E. L., E. M. Albert, M. A. Fortuna, and J. Bascompte. 2009. Compartments in a marine
698 food web associated with phylogeny, body mass, and habitat structure. *Ecology Letters*
699 12:779–788.
- 700 Rezende, E. L., J. E. Lavabre, P. R. Guimarães, P. Jordano, and J. Bascompte. 2007. Non-random
701 coextinctions in phylogenetically structured mutualistic networks. *Nature* 448:925–928.
- 702 Rudolf, V. H. W., and K. D. Lafferty. 2011. Stage structure alters how complexity affects stability
703 of ecological networks. *Ecology Letters* 14:75–79.

- 704 Salinas-Ramos, V. B., L. G. H. Montalvo, V. León-Regagnon, A. Arrizabalaga-Escudero, and E.
705 L. Clare. 2015. Dietary overlap and seasonality in three species of mormoopid bats from a
706 tropical dry forest. *Molecular Ecology* 24:5296–5307.
- 707 Solé, R. V., and M. Montoya. 2001. Complexity and fragility in ecological networks. *Proceedings*
708 *of the Royal Society of London. Series B: Biological Sciences* 268:2039–2045.
- 709 Sritongchuay, T., A. C. Hughes, and S. Bumrungsri. 2019a. The role of bats in pollination
710 networks is influenced by landscape structure. *Global Ecology and Conservation*
711 20:e00702.
- 712 Sritongchuay, T., A. C. Hughes, J. Memmott, and S. Bumrungsri. 2019b. Forest proximity and
713 lowland mosaic increase robustness of tropical pollination networks in mixed fruit orchards.
714 *Landscape and Urban Planning* 192:103646.
- 715 Stone, L., and A. Roberts. 1990. The checkerboard score and species distributions. *Oecologia*
716 85:74–79.
- 717 Thompson, R. M., and C. R. Townsend. 2000. Is resolution the solution?: the effect of taxonomic
718 resolution on the calculated properties of three stream food webs. *Freshwater Biology*
719 44:413–422.
- 720 Toju, H., and Y. G. Baba. 2018. DNA metabarcoding of spiders, insects, and springtails for
721 exploring potential linkage between above- and below-ground food webs. *Zoological*
722 *Letters* 4:4.
- 723 Wang, W., Y. Wu, Y. Yan, M. Ermakova, R. Kerstetter, and J. Messing. 2010. DNA barcoding of
724 the Lemnaceae, a family of aquatic monocots. *BMC Plant Biology* 10:205.

- 725 Wirta, H. K., P. D. N. Hebert, R. Kaartinen, S. W. Prosser, G. Várkonyi, and T. Roslin. 2014.
726 Complementary molecular information changes our perception of food web structure.
727 Proceedings of the National Academy of Sciences 111:1885–1890.
- 728 Wirta, H. K., E. J. Vesterinen, P. A. Hambäck, E. Weingartner, C. Rasmussen, J. Reneerkens, N.
729 M. Schmidt, O. Gilg, and T. Roslin. 2015. Exposing the structure of an Arctic food web.
730 Ecology and Evolution 5:3842–3856.
- 731 Woodward, G. 2010. Integrative Ecology: From Molecules to Ecosystems. Elsevier.
- 732 Yang, C., Y. Ji, X. Wang, C. Yang, and D. W. Yu. 2013. Testing three pipelines for 18S rDNA-
733 based metabarcoding of soil faunal diversity. Science China. Life Sciences 56:73–81.
- 734 Zeale, M. R. K., R. K. Butlin, G. L. A. Barker, D. C. Lees, and G. Jones. 2011. Taxon-specific
735 PCR for DNA barcoding arthropod prey in bat faeces. Molecular Ecology Resources
736 11:236–244.
- 737 Zinger, L., A. Bonin, I. G. Alsos, M. Bálint, H. Bik, F. Boyer, A. A. Chariton, S. Creer, E. Coissac,
738 B. E. Deagle, M. D. Barba, I. A. Dickie, A. J. Dumbrell, G. F. Ficetola, N. Fierer, L.
739 Fumagalli, M. T. P. Gilbert, S. Jarman, A. Jumpponen, H. Kauserud, L. Orlando, J. Pansu,
740 J. Pawlowski, L. Tedersoo, P. F. Thomsen, E. Willerslev, and P. Taberlet. 2019. DNA
741 metabarcoding—Need for robust experimental designs to draw sound ecological
742 conclusions. Molecular Ecology 28:1857–1862.

Tables

Table 1a-f: Summary statistics of the mixed effects models used on the Global observational dataset. Note that the lower network level, and the taxonomic level of ‘species’ are here used as the baseline values.

Functional complementarity, higher level					
Effect	Term	Estimate	Std error	Statistic	P. value
Fixed	Clustering threshold	42.19	7.64	5.52	<0.001
Fixed	Network level (upper)	177.42	4.38	40.51	<0.001
Fixed	Taxonomic level: genus	13.29	5.35	2.48	0.01
Fixed	Taxonomic level: order	45.97	5.37	8.56	<0.001
Random	Random effect: dataset	1625.26			
Random	Residual	507.86			

H2'					
Effect	Term	Estimate	Std. error	Statistic	P. value
Fixed	Clustering threshold	-0.03	<0.001	-41.19	<0.001
Fixed	Network level (upper)	-0.01	<0.001	-28.69	<0.001
Fixed	Taxonomic level: genus	<0.001	<0.001	7.22	<0.001
Fixed	Taxonomic level: order	-0.01	<0.001	-25.67	<0.001
Random	Random effect: dataset	0.16			
Random	Residual	0.05			

749

Modularity					
Effect	Term	Estimate	Std. error	Statistic	P. value
Fixed	Clustering threshold	-0.02	<0.001	-33.76	<0.001
Fixed	Network level (upper)	-0.02	<0.001	-43.25	<0.001
Fixed	Taxonomic level: genus	0.01	<0.001	14.56	<0.001
Fixed	Taxonomic level: order	-0.01	<0.001	-25.3	<0.001
Random	Random effect: dataset	0.09			
Random	Residual	0.04			

750

Nestedness					
Effect	Term	Estimate	Std. error	Statistic	P. value
Fixed	Clustering threshold	-1.21	0.05	-24.63	<0.001
Fixed	Network level (upper)	0.13	0.03	4.66	<0.001
Fixed	Taxonomic level: genus	0.28	0.03	8.08	<0.001
Fixed	Taxonomic level: order	-0.28	0.03	-8.04	<0.001
Random	Random effect: dataset	6.43			
Random	Residual	3.25			

751

752

NODF					
Effect	Term	Estimate	Std. error	Statistic	P. value
Fixed	Clustering threshold	7.99	0.07	107.51	<0.001
Fixed	Network level (upper)	0.17	0.04	3.96	<0.001
Fixed	Taxonomic level: genus	-2.36	0.05	-45.61	<0.001
Fixed	Taxonomic level: order	0.63	0.05	12.21	<0.001
Random	Random effect: dataset	12.84			
Random	Residual	4.9			

753

754

Robustness, higher level					
Effect	Term	Estimate	Std. error	Statistic	P. value
Fixed	Clustering threshold	-0.06	<0.001	-46.22	<0.001
Fixed	Network level (upper)	-0.02	<0.001	-22.12	<0.001
Fixed	Taxonomic level: genus	<0.001	<0.001	1.03	0.3
Fixed	Taxonomic level: order	-0.02	<0.001	-20.61	<0.001
Random	Random effect: dataset	0.09			
Random	Residual	0.08			

Figure captions

Figure 1. Changes in selected metrics in the Global molecular dataset (seven networks).

The X-axis shows the clustering threshold to generate an individual network, the Y-axis shows the changes in individual network calculations. Lines at 93% and 97% denote the most commonly-used range of clustering thresholds.

Figure 2. Metric reliability when analyzing the Global molecular dataset (seven networks),

as described by the F value of the network divided by the F value of the interaction term between network and clustering threshold used. A high value indicates all slopes covarying, whilst a low value indicates greatly varying gradients, showing potentially poor reliability in inter-network comparisons.

Figure 3. Differences in network rankings as a function of clustering threshold for different

metrics for the Global molecular dataset (seven networks) and Guanacaste molecular dataset (subset of two networks). The Global molecular dataset comprises a range of datasets of insectivorous bats and their prey from around the world, whereas the Guanacaste molecular dataset is from bats and their prey studied at a single location in two consecutive seasons. Continuous lines show ranges over which the rank order of the networks analyzed remains unchanged; in each row, each line or dot shows a rank order that differs from those immediately to its left or right. So, for example in the comparison between seven networks, ‘togetherness, lower’ is unchanged across the entire range of thresholds, while Shannon diversity is consistent from 91-93.5% and again (with a different rank order) from 94.5-97.5%. Lines at 93% and 97% denote the most commonly-used range of clustering thresholds.

779

780 Figure 4. Changes in selected metrics in the Guanacaste molecular dataset (subset of two
781 networks). The X-axis shows the clustering threshold to generate an individual network, the Y-
782 axis shows the changes in individual network calculations. Lines at 93% and 97% denote the most
783 commonly-used range of clustering thresholds.

784

785 Figure 5. The percent of pairwise comparisons of networks in the Global observational
786 dataset in which a given combination of node simplification could create an erroneous conclusion
787 (i.e. those that different from the conclusion drawn with no relabelling) for the focal network
788 metrics.

Network

- Big Bend
- Guanacaste normal, 2009
- La Selva wet, 2015
- Windsor Cave
- Guanacaste dry, 2015
- Guanacaste wet, 2015
- SAFE Project

Global molecular dataset (seven networks)

Guanacaste molecular dataset (subset of two networks)

Guanacaste molecular dataset (subset of two networks)

Ecology

Network —●— Guanacaste dry, 2015 —●— Guanacaste wet, 2015

The percent of pairwise comparisons of networks in the Global observational dataset in which a given combination of node simplification could create an erroneous conclusion (i.e. those that differ from the conclusion drawn with no relabelling) for the focal network metrics.

152x76mm (600 x 600 DPI)