Supplementary Material for

Rapid monitoring of ecological persistence

in *Proceedings of the National Academy of Sciences*

Chuliang Song[1,2,3*], Benno I. Simmons[3,*], Marie-Josée Fortin[2],
Andrew Gonzalez[1], Christopher N. Kaiser-Bunbury[3], Serguei Saavedra[4]

[1] Department of Biology, Quebec Centre for Biodiversity Science,
McGill University, Montreal, Canada

[2] Department of Ecology and Evolutionary Biology
University of Toronto, Toronto, Canada

[2] Department of Ecology and Evolutionary Biology,
Princeton University, Princeton, NJ, USA

[3]Centre for Ecology and Conservation, University of Exeter,
Cornwall Campus, Penryn, United Kingdom

[4]Department of Civil and Environmental Engineering,
MIT, Cambridge, MA, USA

* These authors contributed equally to this work and should be considered as joint first authors.

Contents

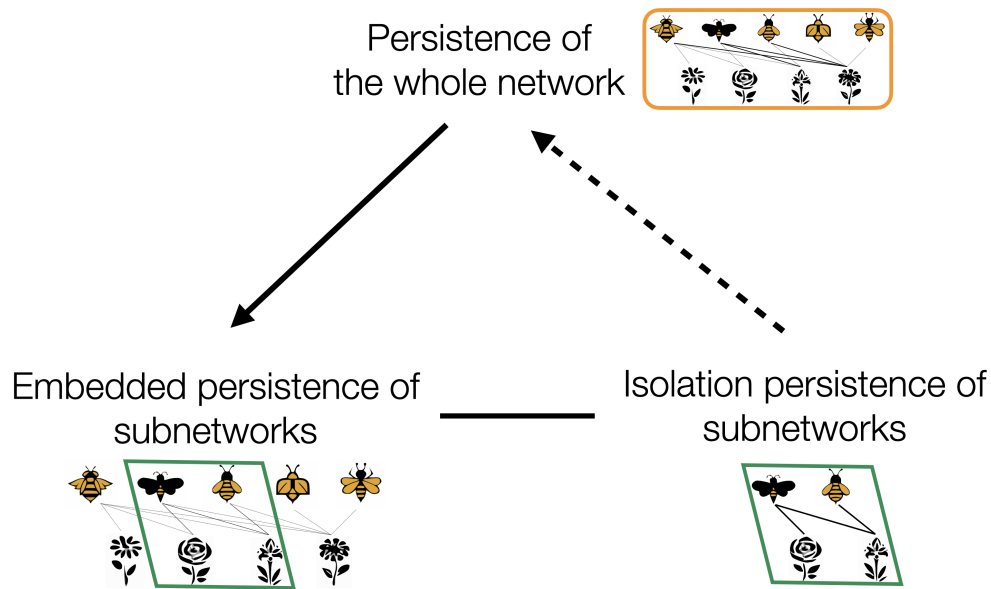# A   Illustration of embedded persistence and persistence in isolation

Figure S1: **From small-scale subnetworks to larger networks**. For illustrative purposes, we consider a hypothetical network structure of a mutualistic network consisting of five pollinator species and four plant species (orange box). Persistence is a network-level property, which makes the network scale the most relevant scale for measuring persistence. However, time and cost constraints and sampling biases limit our knowledge of the network structure. Thus, we most often observe small-scale subnetworks called motifs (green box) that are embedded into the larger network. 'Embedded persistence' refers to persistence of these motifs that are part of larger networks. However, given that the whole network is unobservable, we can instead study whether the subnetworks can persist in isolation, removed from the wider network context, which we call the 'persistence in isolation' of a subnetwork. While it is clear that the persistence of the network determines the embedded persistence of subnetworks, here we show that the persistence in isolation of subnetworks is linked to the persistence of the network as a whole.

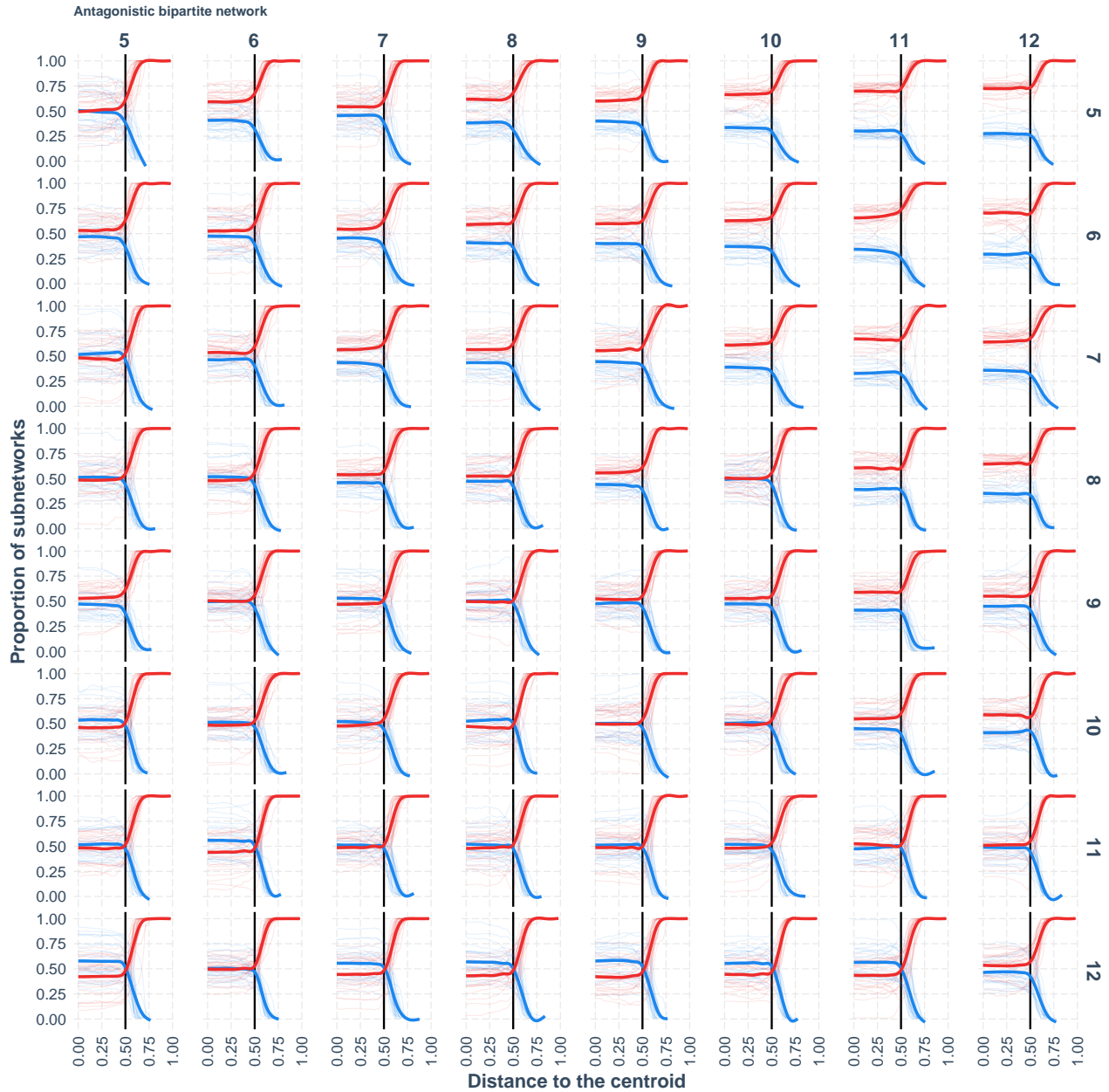# B Robustness of phase transition

B.1 network size

Figure S2: Phase transitions in antagonistic systems. The horizontal panels correspond to the numbers of columns in the bipartite networks, while the vertical panels correspond to the numbers of rows. The x axis denotes the distance to the centroid of the coexistence domain (visualization in Figure 3A). The y axis denotes the proportion of subnetworks that are either persistent in isolation (blue) or non-persistent (red). We find that our main phase transition result is robust to changes in network size, with similar patterns observed in a range of network dimensions.
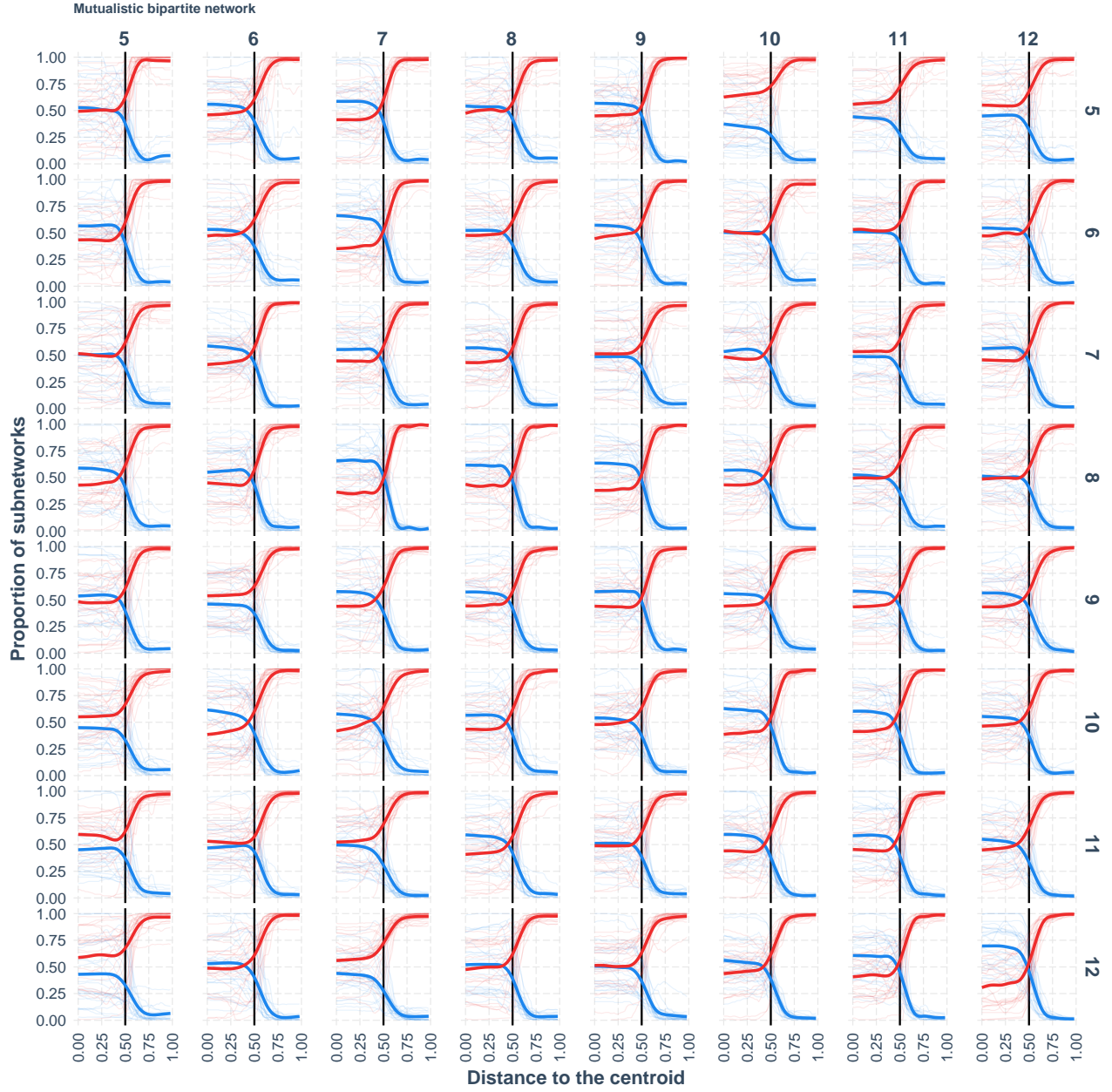
Figure S3: Same as Figure S2 except we focus on mutualistic systems. We find that the results are robust across interaction types, with qualitatively identical findings in antagonistic networks.
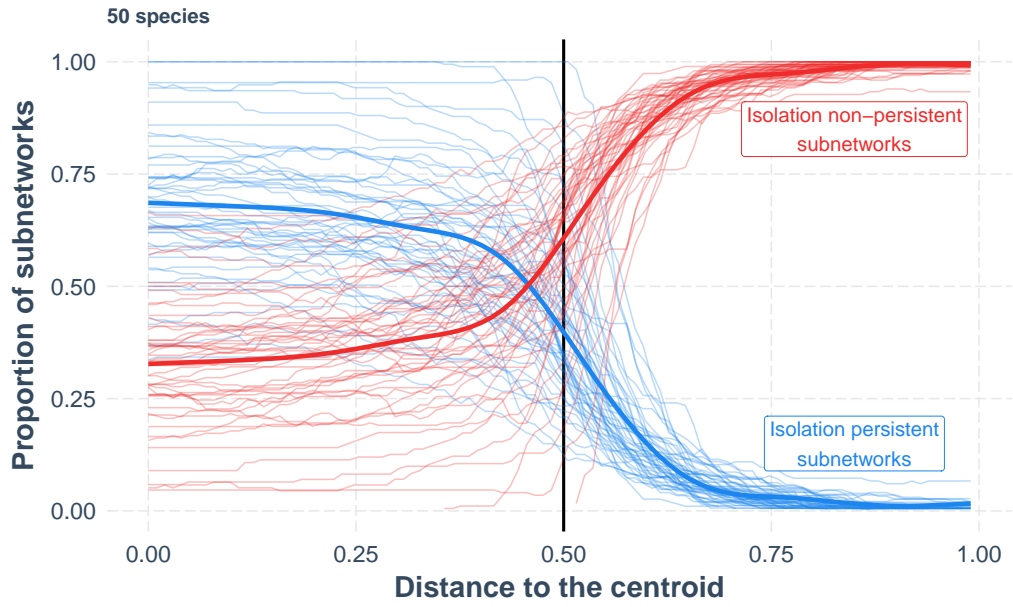
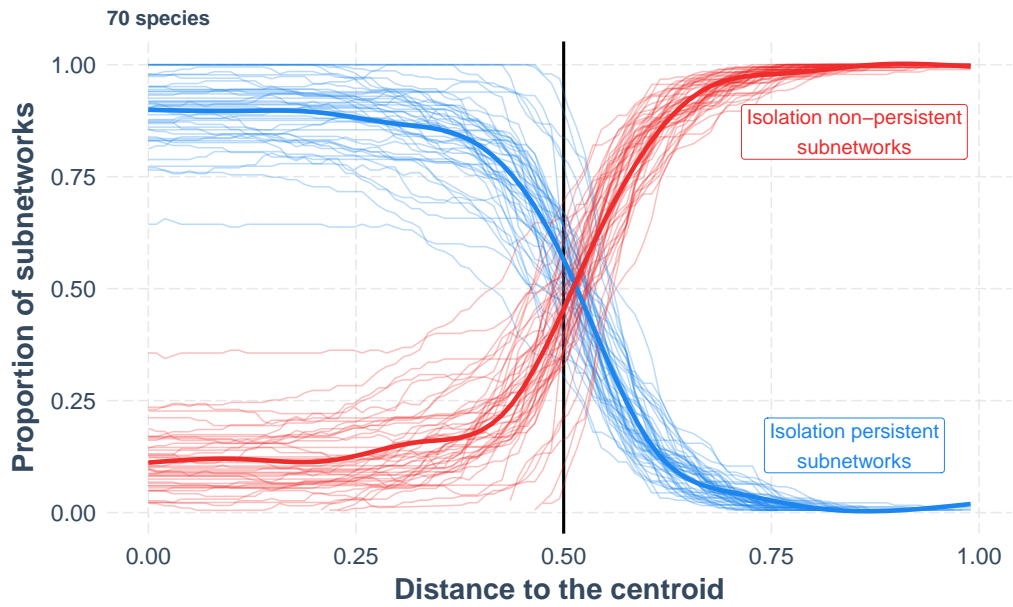Figure S4: The same as Figure 3B except that we consider 50 species.

Figure S5: The same as Figure 3B except that we consider 70 species.
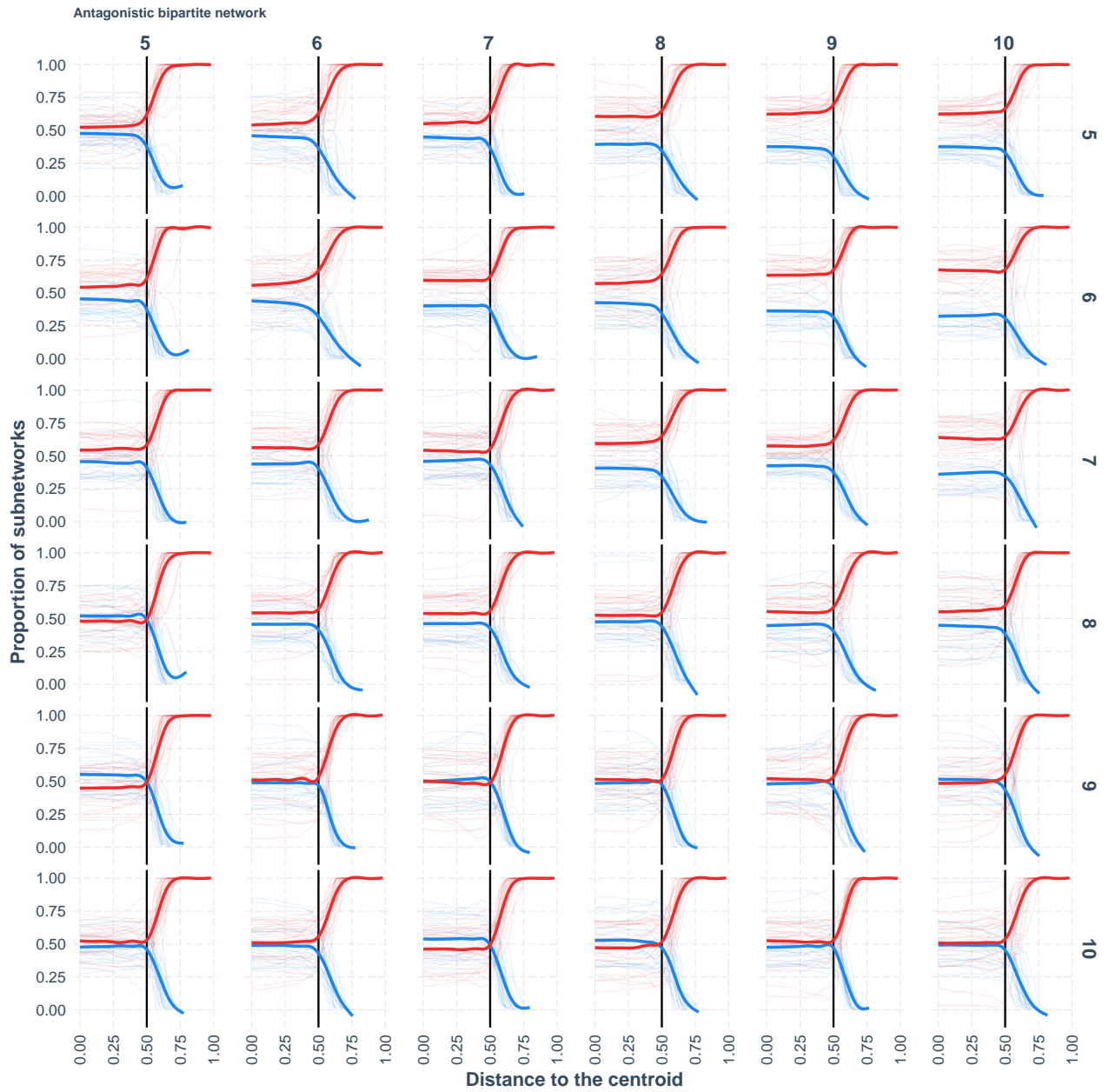
B.2   network structure

Figure S6: Same as Figure S2 except we focus on fully nested structures. We find that the results are robust across network structures, with qualitatively identical findings in Erdős–Rényi networks.
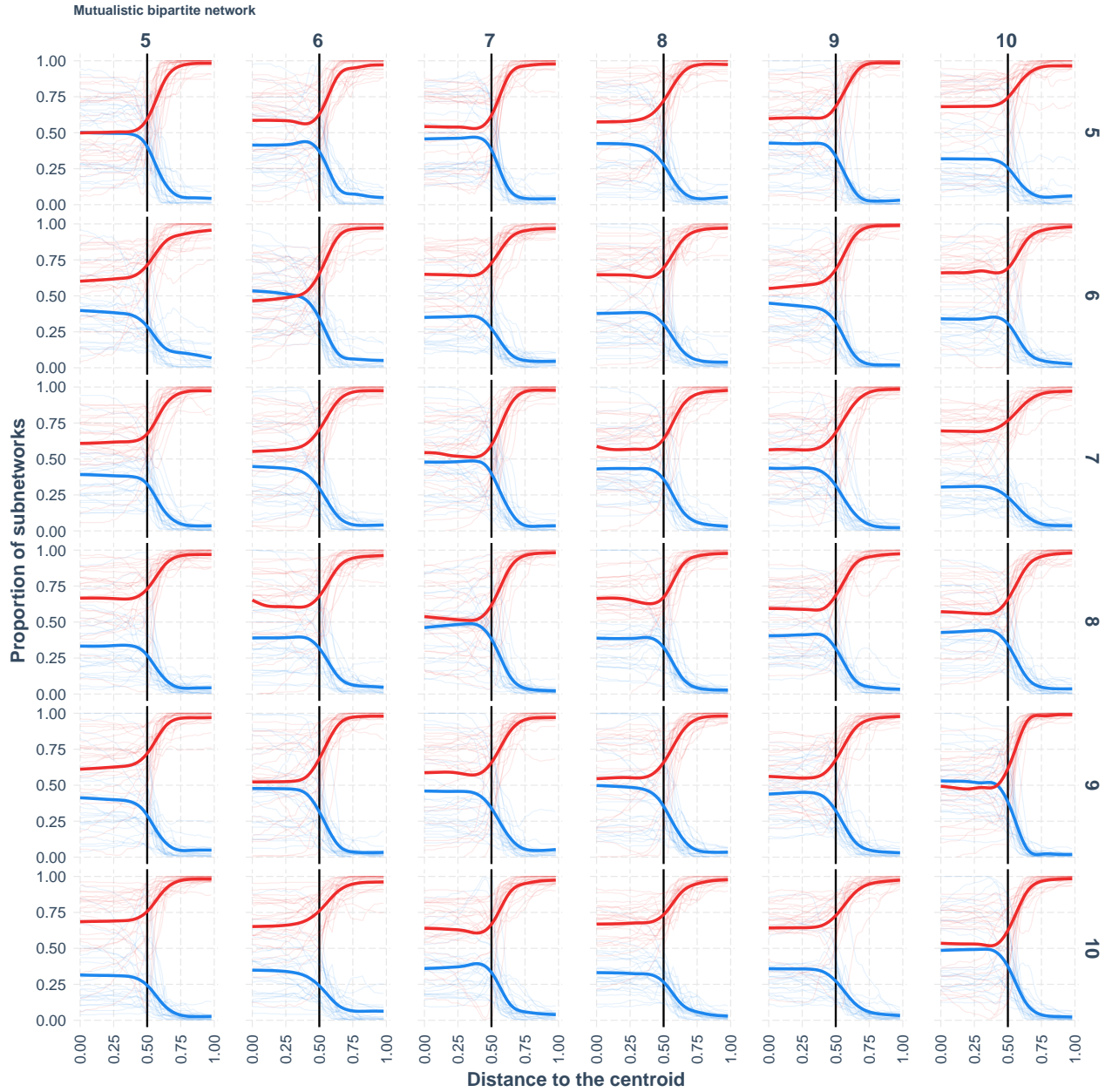
Figure S7: Same as Figure S6 except we focus on mutualistic systems. We find that the results are robust across interaction types, with qualitatively identical findings in antagonistic networks.
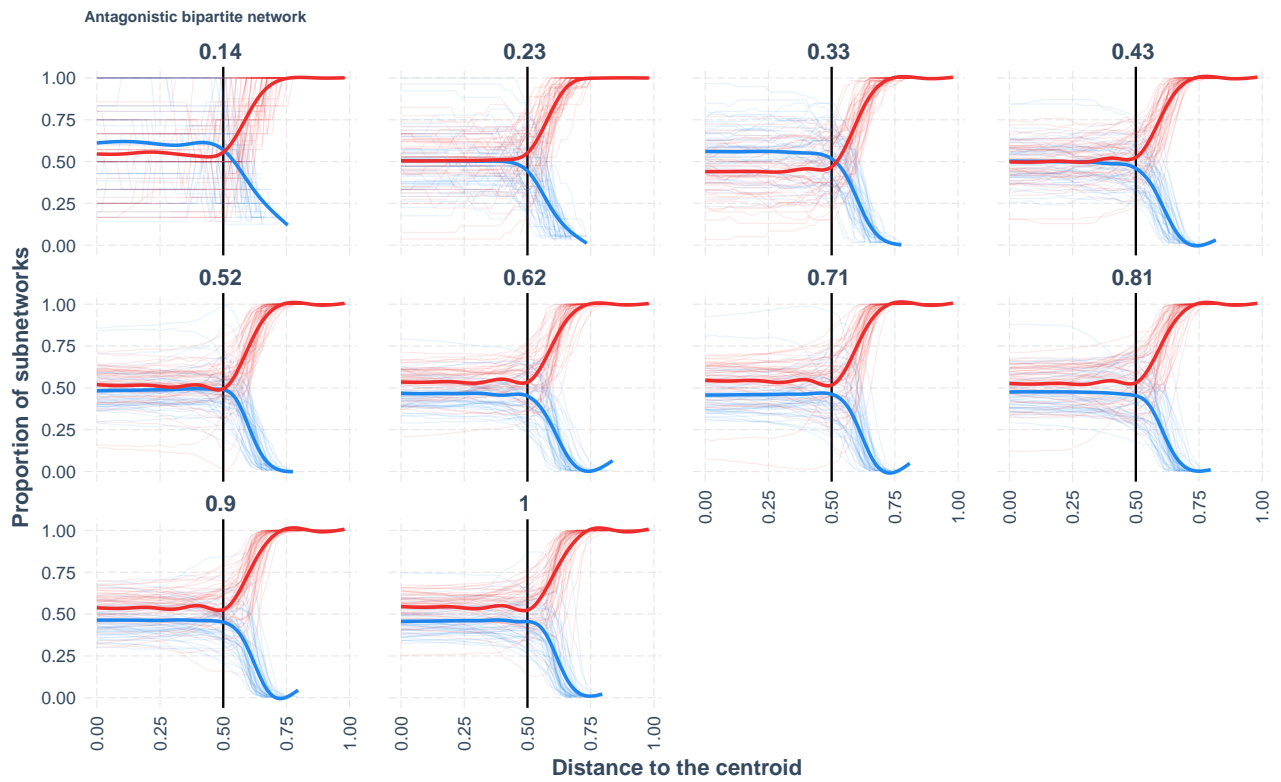
B.3   Connectance

Figure S8: Same as Figure S2 except we change the connectance in the system. We find that the results are robust across connectance.
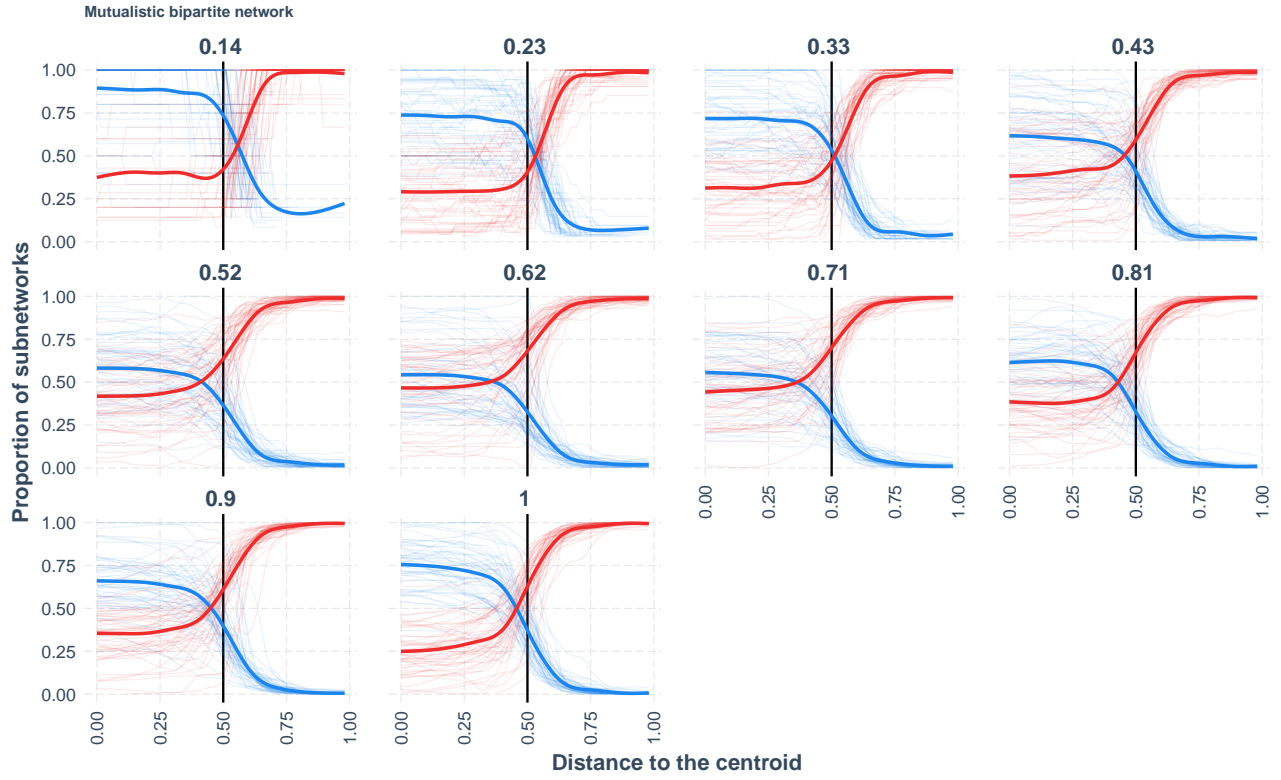
Figure S9: Same as Figure S8 except we focus on mutualistic systems. We find that the results are robust across interaction types, with qualitatively identical findings in antagonistic networks.
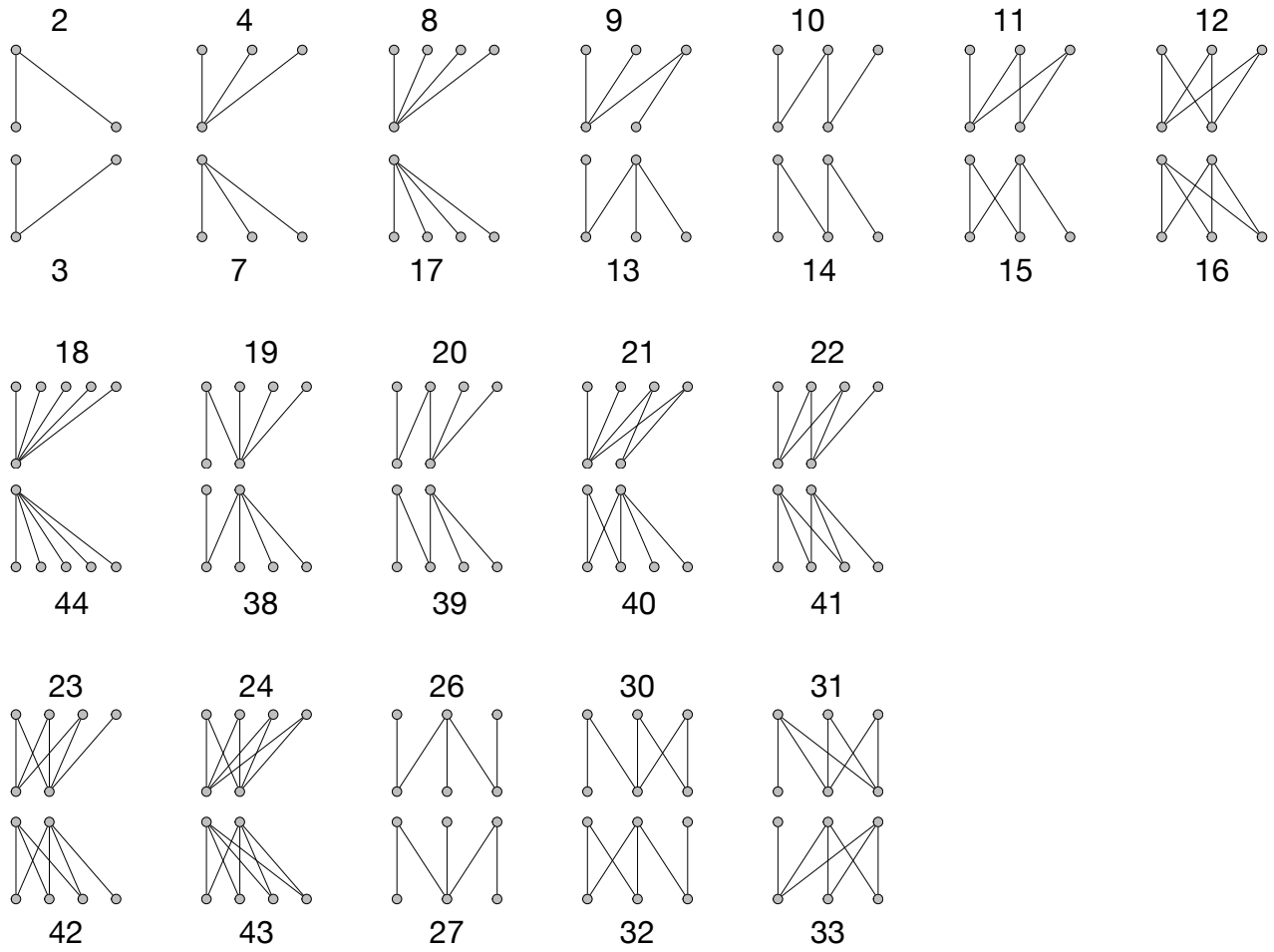
C   Motif pairs

Figure S10: The motif pairs used in the analyses. Note each motif is a transpose of the other motif in the pair. Numbers above motifs refer to motif IDs following (1).
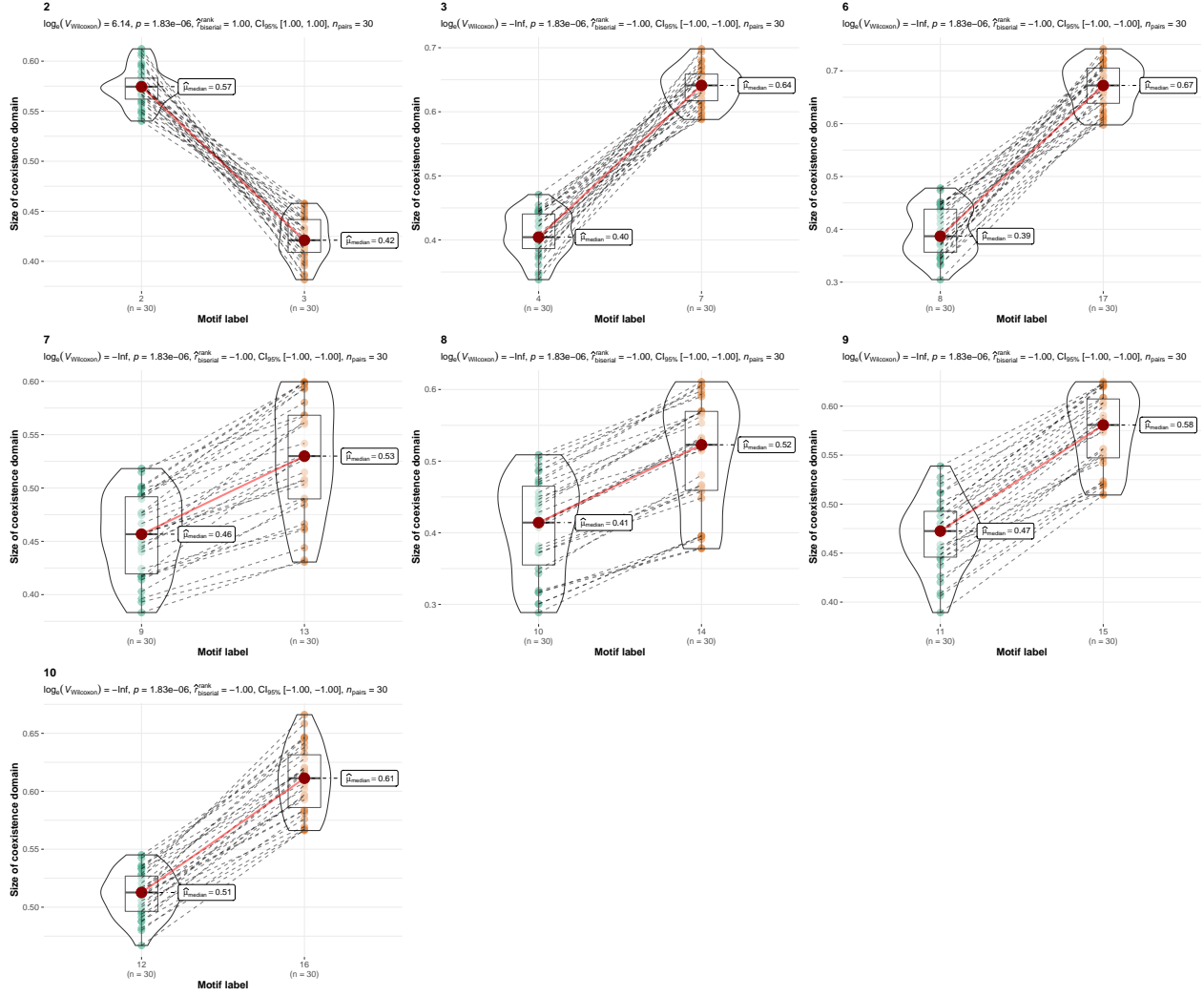
Figure S11: The sizes of coexistence domains between transpose motif pairs differ robustly. Each panel shows a pair of motifs that are transpose to each other. We show here all motifs pairs with three to five species. The x axis shows the motif label, while the y axis shows the size of coexistence domain. Each point denotes a different parametrization of the mutualistic benefit δ and the inter-specific competition strength ρ. Each pair of linked points has *identical* parametrization, thus, the difference between the sizes of coexistence domains solely comes from the trophic constraints. We find robust differences in the sizes of coexistence domains in all motif pairs.

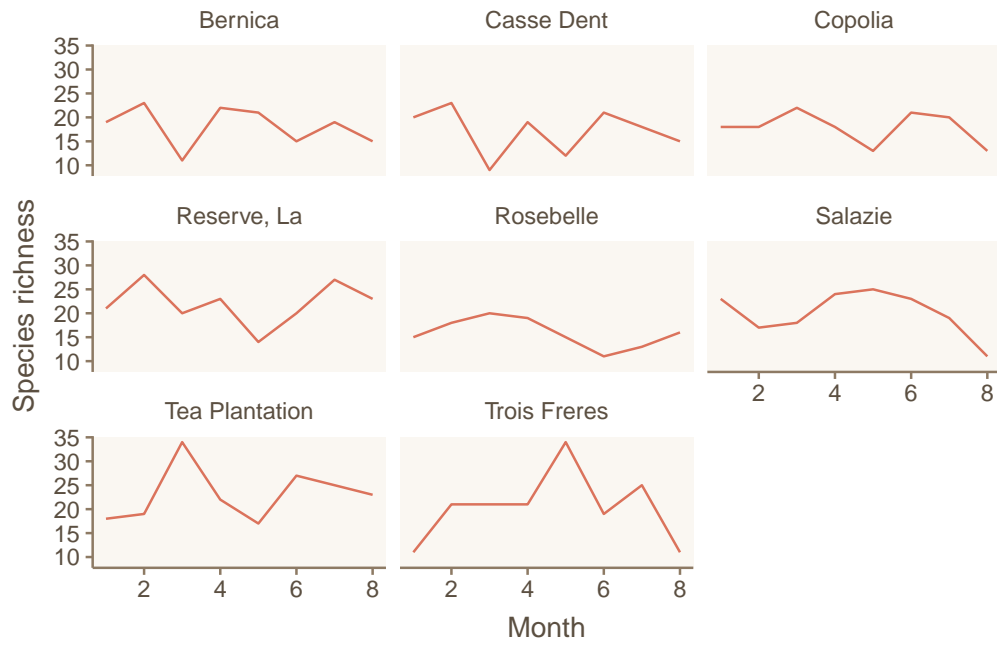# D   Analysis on a dataset of temporal networks
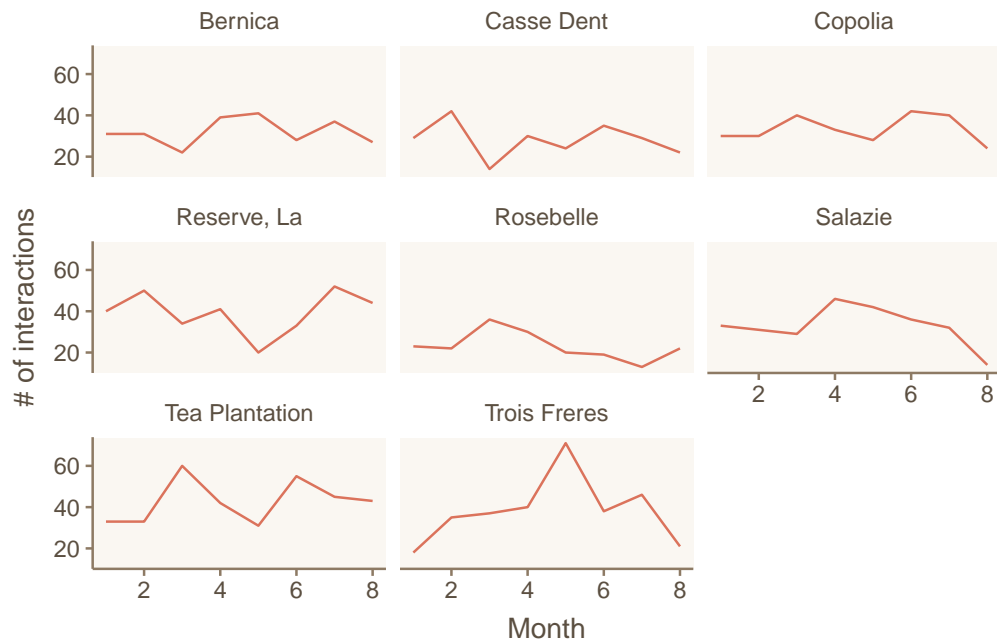

Figure S12: Changes in species richness.

Figure S13: Changes in number of interactions.

Figure S14: **Restored mutualistic networks contain more subnetworks with higher likelihood of persistence in isolation than unrestored disturbed networks.** We use a temporal dataset with 64 networks (8 networks sampled over 8 months) located on the granitic island of Mahé, Seychelles (2). The x axis denotes the eight consecutive months between September 2012 and April 2013. The y denotes the z-scores of a given subnetwork (over- or under-representation of empirical motif frequency compared to motif frequency in randomized networks). The two black horizontal lines correspond to the threshold above or below which a subnetwork (motif) occurs significantly more or less, respectively, than random (z-score = 2 or -2). Each translucent light line corresponds to a different network. The thick lines correspond to the average across 4 different networks. The green lines correspond to the restored networks, while the purple lines correspond to the disturbed networks. Each panel focuses on subnetworks of a given subnetwork type; for visual clarity we show the more persistent in isolation subnetwork of each subnetwork pair (see Figure S15 for results with both motifs in each pair). For 6 out of 7 pairs, the more persistent in isolation subnetworks are significantly over-represented in the networks, and the over-representation is stronger in restored networks than in disturbed networks.
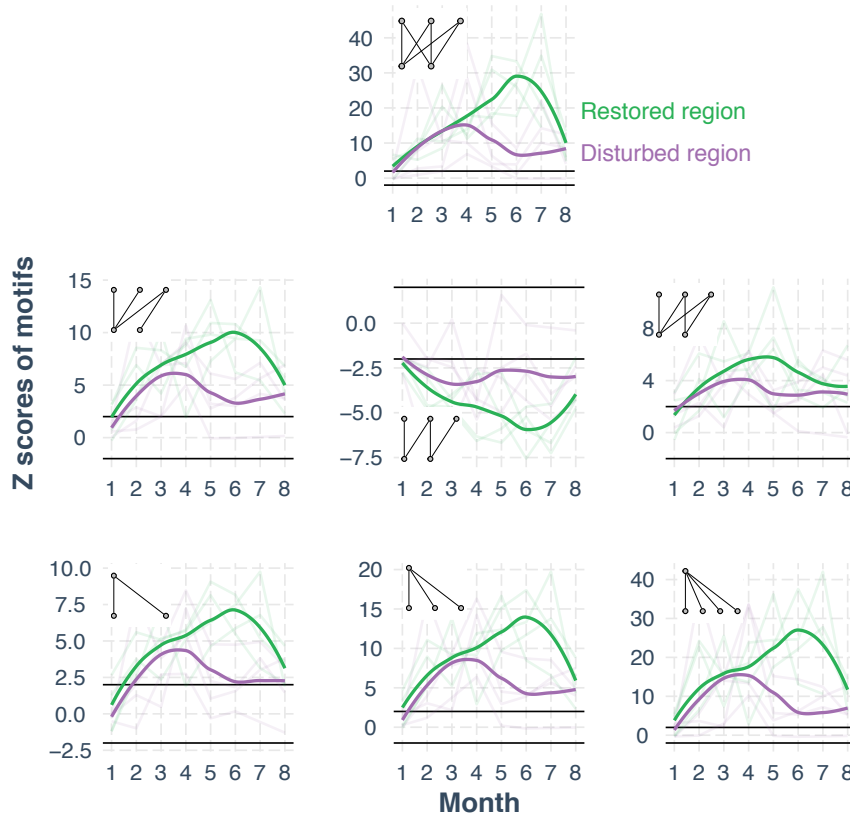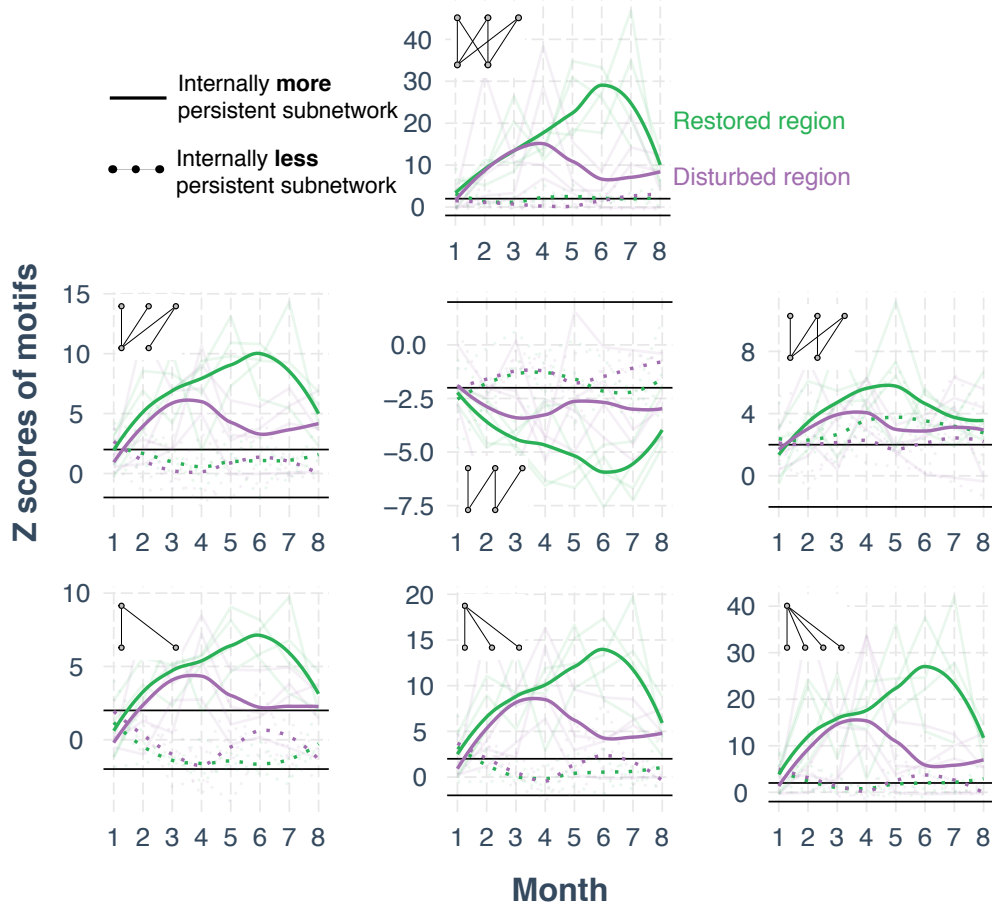
Figure S15: **Restored mutualistic networks contain more subnetworks with higher likelihood of persistence in isolation than unrestored disturbed networks.** We use a temporal dataset with 64 networks (8 networks sampled over 8 months) located on the granitic island of Mahé, Seychelles (2). The x axis denotes the eight consecutive months between September 2012 and April 2013. The y denotes the z-scores of a given motif (over- or under-representation of empirical motif frequency compared to motif frequency in randomized networks). The two black horizontal lines correspond to the threshold above or below which a motif occurs significantly more or less, respectively, than random (z-score = 2 or -2). Each translucent, light line corresponds to a different network. The thick lines correspond to the average across 4 different networks. The green lines correspond to the restored networks, while the purple lines correspond to the disturbed networks. Each panel focuses on subnetworks of a given motif type and its transpose. The persistence in isolation of a motif type and its transpose can be fairly compared without the danger of different parameerizations altering conclusions. The opaque, solid line denotes the more persistent in isolation motif type (the more persistent motif out of the two in each transpose pair), while the dotted line denotes the less persistent in isolation motif type. We show all the bipartite network motif pairs with 3 to 5 species. For all the pairs, the less persistent in isolation subnetworks, regardless whether the networks are restored or not, are not statistically significant present in the empirical networks. For 6 out of 7 pairs, the more persistent in isolation subnetworks are significantly over-represented in the networks, and the over-representation is stronger in restored networks than in disturbed networks.

Figure S16: Same as Figure 5 except we conduct the analysis in each network. The labelling of motifs follow (1, 3).
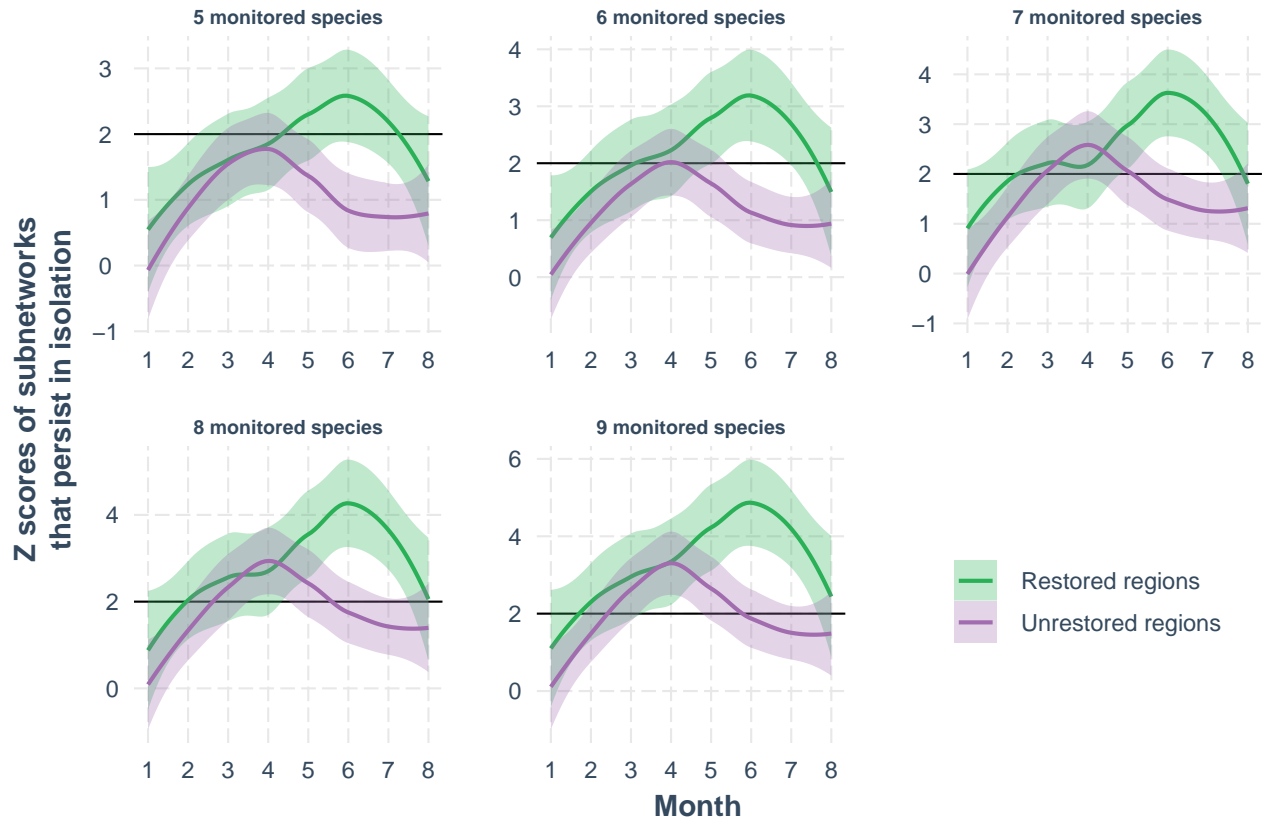
Figure S17: Same as Figure 5 expect we consider only biomonitoring a subset of all species. Each panel represents different numbers (from 5 to 9) of monitored species. The qualitative results remain the same.
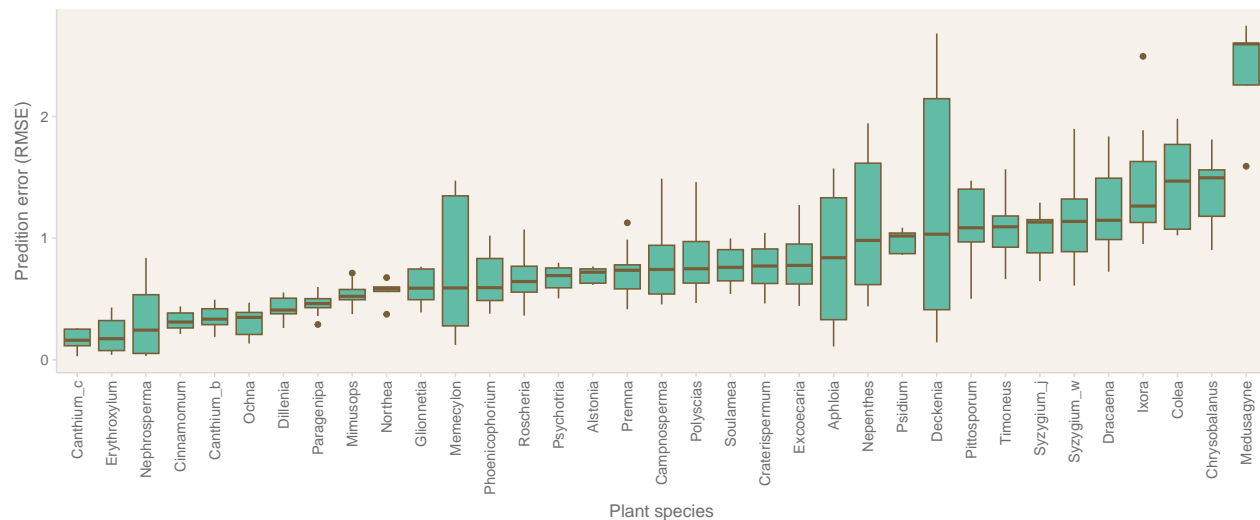
Figure S18: The x axis shows different plant species from the empirical Seychelles dataset. The y axis shows the prediction error (RMSE) when the particular plant species was monitored. The box plot average across different random monitoring cases.
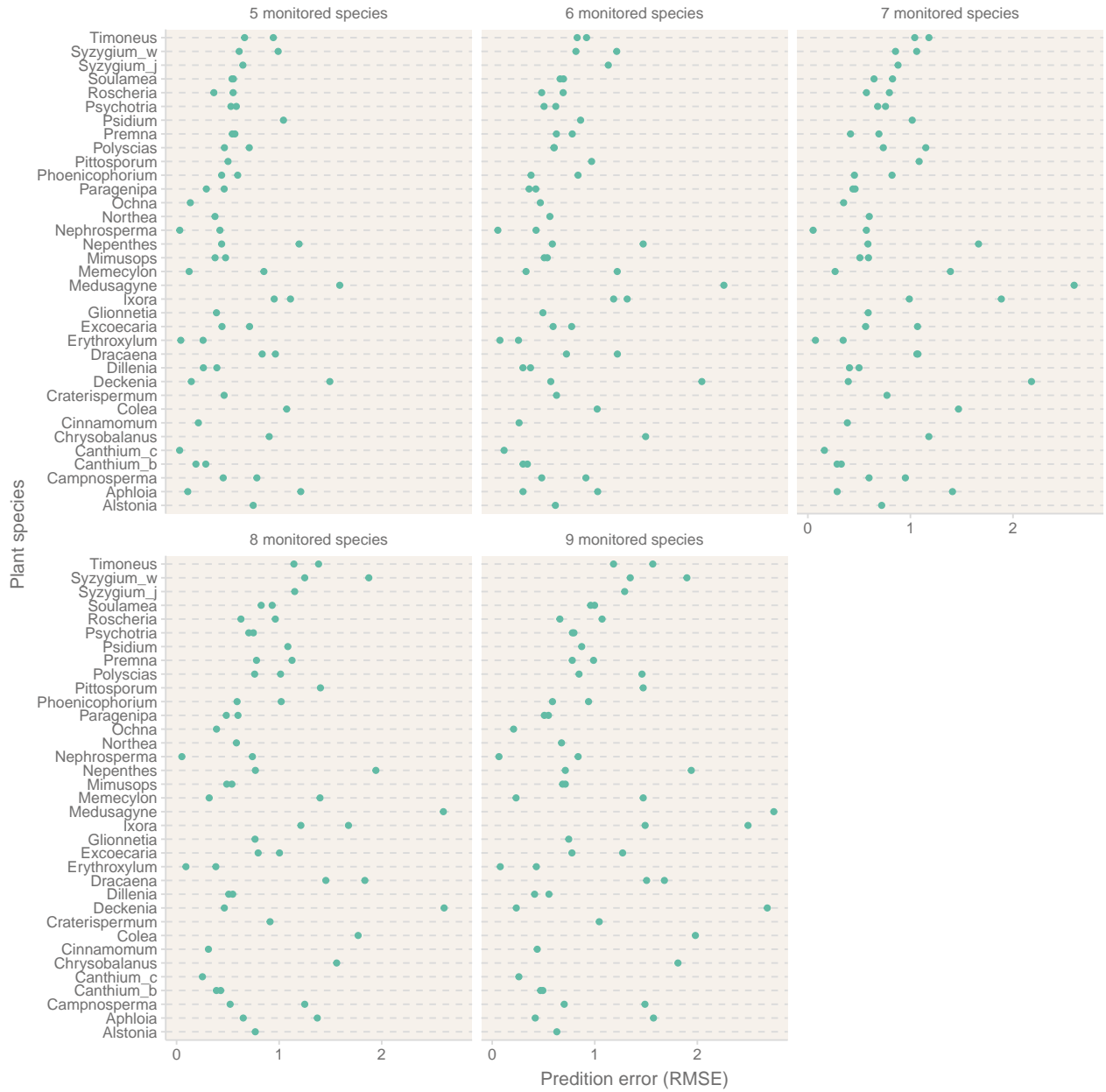
Figure S19: Same as Figure S18 except we separate it by different numbers of monitored plant species.

E   Estimates of sampling costs in empirical networks

Collecting ecological network data is costly. In (4) there were two main costs. First, the cost of observing all woody flowering species. The numbers of hours invested in each network is proportional to the number of species flowering in each network. Each plant species took 3.03 ± 0.62 h of observation time per network. Each hour of observation had another 30 mins 'non-observational' time added (which includes searching and insect processing time). Thus, approximately 4.5h of time is invested per plant species per network. With average US \$20 per hour for skilled researchers (`https://www.salary.com/tools/salary-calculator/general-laborer-sr-hourly` [accessed Nov 17 2022]) and 464 flowering species across the whole sampling period, the approximate cost of observation was $464 \times 4.5 \times 20 = \$41,760$. The analysis in Figure S17 shows that the community could be effectively monitored for persistence with as few as five plant species. We therefore calculated the reduction in observation costs that would result from only having to observe five species per network.

The second major cost came from insect identification. Kaiser-Bunbury et al. (4) collected a total of 490 insect specimens that were not identifiable in the field. According to Fly Evidence insect identification services (`https://flyevidence.co.uk/ecological` [accessed Nov 17 2022]), each specimen cost approximately US \$12 to process, plus US \$60 for each new species encountered. The total identification costs for the study were \$14613. A reduction in the number of plant species monitored reduces the number of insect species that require identification. We therefore also calculated the average reduction in insect identification costs that would result from monitoring 5 random plant species in each community.

Our approach could greatly reduce the sampling cost. We found that the cost of plant observation would be reduced by 31% from US \$41,760 to \$28,800, a saving of \$12,960. The cost of insect identification would be reduced by 74%, from \$14613 to \$3771, a saving of \$10,842. Combined, this represents a cost reduction of 42%, from \$56,373 to \$32,571, a saving of \$23,802. It is important to note that although monitoring a subset of species reduces costs, monitoring costs can be strongly influenced by the species involved, the types of interaction and the target metric of interest. All these features must be considered by decision makers when designing monitoring strategies.

References

[1] Simmons BI, et al. (2019) Motifs in bipartite ecological networks: uncovering indirect interactions. *Oikos* 128:154–170.

[2] Kaiser-Bunbury CN, et al. (2017) Ecosystem restoration strengthens pollination network resilience and function. *Nature* 542:223–227.

[3] Simmons BI, et al. (2019) bmotif: A package for motif analyses of bipartite networks. *Methods in Ecology and Evolution* 10:695–701.

[4] Kaiser-Bunbury CN, Blüthgen N (2015) Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants* 7.