

pymfinder: a tool for the motif analysis of binary and quantitative complex networks

Bernat Bramon Mora^{1,*}, Alyssa R. Cirtwill² and Daniel B. Stouffer¹

¹Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

²Department of Physics, Chemistry, and Biology, Linköping University, Linköping, Sweden.

*bernat.bramon@gmail.com

1 Abstract

2 We developed *pymfinder*, a new software to analyze multiple aspects of the so-called network
3 motifs—distinct n -node patterns of interaction—for any directed, undirected, unipartite or
4 bipartite network. Unlike existing software for the study of network motifs, *pymfinder*
5 allows the computation of node- and link-specific motif profiles as well as the analysis of
6 weighted motifs. Beyond the overall characterization of networks, the tools presented in this
7 work therefore allow for the comparison of the “roles” of either nodes or links of a network.
8 Examples include the study of the roles of different species and/or their trophic/mutualistic
9 interactions in ecological networks or the roles of specific proteins and/or their activa-
10 tion/inhibition relationships in protein-protein interaction networks. Here, we show how
11 to apply the main tools from *pymfinder* using a predator-prey interaction network from a
12 marine food web. *pymfinder* is open source software that can be freely and anonymously
13 downloaded from <https://github.com/stoufferlab/pymfinder>, distributed under the
14 MIT License (2018).

15 Introduction

16 The use of network theory has proven insightful in multiple fields, from the study of the
17 spread of disease epidemics (1) to the characterization of neuronal networks (2). In ecology,
18 this approach has been crucial to understanding the ways different species interact with
19 each other, and the network perspective has justly become a central topic in community
20 ecology (3). Over recent years, multiple methods for studying the topology of ecological
21 networks have been successfully developed. Examples include models to generate realistic
22 ecological communities (4) or tools for studying different network metrics such as com-
23partmentalization (5), nestedness (6) or intervality (7). Following these advances, one of
24 the most versatile ways to understand the structure of complex ecological networks is via
25 the so-called network motifs—i.e. the analysis of small subgraphs representing the distinct
26 patterns of interaction involving any set of n species. These subgraphs have been referred
27 to as the ‘building blocks’ of complex networks (8).

28 The study of network motifs has been applied to multiple ecological systems over the
29 recent years, including those composed of trophic (9) and mutualistic interactions (10).
30 Non-ecological examples include in protein-protein interaction networks (11) and transcrip-
31 tional regulation networks (12). There are typically two main approaches that are taken
32 involving network motifs. First, counting the number of appearances of any given n -node
33 pattern of interactions provides an overall perspective of the structure of a network. This
34 has been done in different ecological studies, including the characterization of food webs
35 (13), plant-pollination (14) and host-parasitoid networks (15). Second, other ecological
36 studies have focused on the role of different species (16) and interactions (17), defining
37 their position within the network based on which network motifs they form a part of.
38 Following this work on network motifs, multiple tools for the counting of network mo-
39 tifs have been developed over the last decades (18, 19, 20). Most of the methodological
40 work has focused on providing tools to efficiently quantify the overall structure of directed
41 and undirected unipartite networks—i.e. graphs consisting of one set of interacting nodes.
42 Unfortunately, to our knowledge, we are still lacking general-purpose software to also an-
43alyze bipartite networks—i.e. graphs consisting of two interacting sets of non-overlapping
44 nodes—as well as to quantify the node- and link-specific motif profiles in both unipartite

45 and bipartite networks. In addition, there is no tool to date that allows the user to include
46 information regarding the interaction strengths of a network within the analysis of motifs.
47 In response, we present *pymfinder*, software for motif analysis of network structure plus of
48 the nodes and links of any type of network—i.e. directed/undirected, bipartite/unipartite,
49 and weighted/binary networks.

50 *pymfinder* is an open-source and versatile tool for the study of network motifs and the
51 result of long-standing research involving the study of ecological networks. For example,
52 *pymfinder* was used to shed light on the ecological mechanisms underlying food-web struc-
53 ture (9), which, together with Bascompte 2005 (21) and Camacho et. al. (22), was one
54 of the first studies to put network motifs into a purely ecological context. Building on
55 these foundational studies, network motifs and *pymfinder* were shown to provide a useful
56 way to characterize species' roles, showing them to be evolutionary conserved across com-
57 munities (16). Similarly, the roles of links involving parasite species were characterized
58 through the study of network motifs, generating an understanding of how different types
59 of feeding links are distributed within a food web (17). The same software has also been
60 used to study bipartite networks. For instance, a study on host-parasitoid networks showed
61 how species' roles seem to be conserved over spatial scales as well as consistent over time
62 (15). Perhaps more importantly, the software presented here has also been a central piece
63 of very recent research. For example, the tools in *pymfinder* were used to relate species'
64 roles to multiple ecological traits in five marine food webs, showing that feeding environ-
65 ment is particularly strongly related to such roles (23). Likewise, the variability of species'
66 roles in plant-pollinator communities in the Arctic has recently shown to be related to the
67 variability in community composition (24). Finally, the description of species' roles has
68 also been key to comparing entire networks by means of aligning species to each other,
69 resulting in the identification of common backbones shared across food webs from different
70 ecosystems (25). Overall, the tools included in *pymfinder* are and have been instrumental
71 to the development of a diverse set of projects over the years, and we believe that they
72 have the potential to be valuable for many others. This article describes the main structure
73 of *pymfinder* and showcases some of its principal applications using a detailed ecological
74 dataset as the backdrop.

75 **Design and implementation**

76 **General description**

77 *pymfinder* is a Python library that combines Python methods for network-motif analysis.
78 Some of the engine underneath is a modified version of *mfinder*—a software tool for network-
79 motif detection developed by Kashtan et. al. (8, 18). Originally, *mfinder* was written in C
80 and made available solely as an executable, and we use it within *pymfinder* for its underlying
81 efficiency. The *mfinder* code has been both included and modified here with the explicit
82 consent of Nadav Kashtan, the author of *mfinder* 1.2.

83 As input, *pymfinder* accepts any type of network. That is, the analyses can be performed
84 for both unipartite and bipartite networks. The format in which the networks are passed
85 to the different functions of the package is either as text files, Python arrays or *pymfinder*-
86 objects. Text files must describe the set of links comprising the networks, where each link
87 appears as a separate line in the files. For example, a given line “ $A \ B \ w$ ” would describe
88 a single link $A \rightarrow B$ between nodes A and B , where w represents the strength or weight
89 associated to such link (see Appendix). Similarly, Python arrays need to represent the list
90 of interactions forming the networks. Notice that the direction of the links is important.
91 Therefore, in bipartite networks, nodes of each group need to consistently be placed on
92 the same side of the interactions—e.g. in a plant-pollinator networks the direction of the
93 interactions in the input must all go from a plant to a pollinator (or vice-versa). Importantly,
94 undirected networks can also be analyzed by *pymfinder*; however, any links between two
95 nodes A and B in such networks need to be characterized by the two parallel links $A \ B$ and
96 $B \ A$. The output of *pymfinder*, is a high-level data type (‘class’) that contains different
97 descriptors of the motif composition of the network under study (see Appendix).

98 **Structure of the package**

99 At their core, all of the analyses performed by *pymfinder* are based around the identification
100 of all the different n -node patterns of interaction found within a given network. To do this,
101 *pymfinder* will always start by enumerating the unique motifs/subgraphs that make up the

102 overall structure of the network under study. This analysis can be performed for multiple
103 motif sizes. This is especially important for bipartite networks, where three-node motifs are
104 minimally informative and one needs to explore bigger motifs (15). Notice, however, that
105 increasing the number of nodes can be computationally challenging for unipartite networks
106 since the number of unique motifs quickly increases with their size—i.e. there are 13 unique
107 three-node motifs, 199 unique four-node motifs and 9364 unique five-node motifs.

108 For the sake of simplicity, we will focus most of the description of the methods presented
109 here on the analysis of three-node network motifs. For any given network, this analysis is
110 a three-step process. First, *pymfinder* loops through all the rows i of the adjacency matrix
111 A associated with the network. For each non-zero element a_{ij} found in row A_i , it then
112 searches for any connected element $a_{jk} = 1$, $a_{kj} = 1$, $a_{ik} = 1$, and/or $a_{ki} = 1$, revealing the
113 existence of any motif comprised of the nodes i , j , and k . If i , j and k define a motif and
114 this motif has not already been identified, the corresponding motif and the position of each
115 node within the motif is recorded.

116 Based on this initial motif enumeration, *pymfinder* can perform three subsequent analyses:
117 (i) the analysis of the overall network structure, (ii) the nodes and links' participation in
118 the different motifs, and (iii) the nodes and links' role in each of the motifs.

119 *Motif structure*

120 The most basic application of *pymfinder* is the analysis of the overall motif structure of a
121 given network. In particular, such analysis generates a description of the distribution of
122 distinct n -node patterns of interaction found within the network (up to 8-node motifs). The
123 application also includes the possibility of estimating the null motif composition expected
124 for such network (see Appendix). To generate this null composition, *pymfinder* uses an
125 MCMC algorithm to perform a randomization of the network while preserving the in-
126 and out-degree of the nodes and each node's number of single and double links (26, 27).
127 Comparing the observed motif frequency to the random expectation, the application can
128 be used to determine which interaction patterns are over- or under-represented relative to
129 this null model (9). To do so, *pymfinder* calculates the mean and standard deviation of the
130 null expectation as well as the z-scores for its comparison with the actual observations.

131 An additional feature of *pymfinder* is the possibility of incorporating information regard-
132 ing the link strength into the analysis of the motif structure. This is notable in particular
133 since there is no software available to explore the way the interaction strengths are dis-
134 tributed within networks across motifs. To do so, *pymfinder* will account for each motif
135 within a given weighted network as a function of the strength of the links forming them (Fig.
136 **1**). Note that the algorithm allows the user to choose how the weight of a motif is defined.
137 Specifically, given a motif formed by the set of links with strengths $\{l\} = \{l_1, l_2, \dots, l_L\}$,
138 *pymfinder* will calculate the weight of such motif as $f(\{l\})$, where f is the function defined
139 by the user. By default, *pymfinder* uses the arithmetic mean as the function f . Similar
140 to unweighted networks, analysis of the motif structure of a weighted network returns the
141 average and standard deviation of the weight of each motif, as well as the median and the
142 first and third quartiles.

143 *Motif participation*

144 The study of network motifs can also be used as a way to classify nodes based on which
145 patterns of interactions they are part of. For any given network, this application deter-
146 mines the frequency of appearance of every node across each of the different motifs (Fig.
147 **1b**), defining their participation across these distinct patterns of interactions. This a useful
148 perspective for motif analysis because it provides a node-based description of the networks
149 that can be used to understand the nature of specific nodes (e.g. different species in eco-
150 logical networks or different proteins in protein-protein interaction networks) as well as
151 decomposes the overall structure of the network at a finer resolution (**21**). Similarly, the
152 same analysis can also be performed for the links forming the network. That is, *pymfinder*
153 can quantify the frequency with which every link forms part of each distinct motif. As for
154 the analysis of the overall structure of the networks, the motif participation of both nodes
155 and links can also be calculated for any given motif size up to 8 nodes for weighted and
156 unweighted networks. Again, *pymfinder* will account for each motif within a given weighted
157 network as a function of the strength of the links forming it (Fig. **1**), and the algorithm
158 allows the user to choose this function just as described above for motif structure.

159 *Motif-role profiles*

160 Within any given motif, nodes can play multiple roles. For example, in the two-node motif
161 $A \rightarrow B$, there are two distinct positions A and B , which define two different roles—e.g. a
162 predator and a prey in a food web. In contrast, for the two-node motif $A \leftrightarrow B$, A and B
163 occupy indistinguishable positions; therefore, there is a single distinct role. The same idea
164 can be extended to all n -node motifs. For example, there are 30 distinct node positions and
165 24 distinct link positions across the 13 unique three-nodes motifs. These distinct positions
166 within the different motifs are important because the number of times that a node appears
167 in each of them can be used as a way to define its structural role in a community (16).
168 That is, we can characterize a node’s structural role based on the number of times that it
169 occupies each distinct position of the n -node motifs. *pymfinder* provides a way to determine
170 such n -node motif-role profiles for both the nodes (Fig. 1c) and the links (Fig. 1d) of a
171 given network. Notice, however, that this function can only be run for two- and three-nodes
172 motifs in unipartite networks, and two- to six-nodes motifs in bipartite networks.

173 The analysis of node and link motif-role profiles can also incorporate information regard-
174 ing the strengths of interactions between nodes. As before, consider a motif m formed by
175 the set of nodes $\{i\}$ and the set of links with strengths $\{l\}$. For any node j in $\{i\}$, *pymfinder*
176 calculates the contribution c_{jm} of motif m to any of the positions of j ’s motif-role profile
177 as:

$$c_{jm} = \frac{f(\{l_j\})}{\sum_i f(\{l_i\})} f(\{l\}) \quad (1)$$

178 where $\{l_i\}$ is the set of strengths of all links in m involving node i , and f is a function defined
179 by the user. By default, *pymfinder* again uses the arithmetic mean as f for weighted motif-
180 role profiles. Notice that the contribution $c_{im} = 1$ when ignoring the weights, or f is the
181 arithmetic mean and all weights are equal to the motif size. When analyzing the motif-role
182 profile of a link k forming such motif, the contribution c_{km} is assumed to be exactly equal
183 to its link strength l_k .

184 **Basic tests**

185 To ensure the reliable functioning of *pymfinder*, we included a set of basic tests in the
186 package. All these basic tests are based around the idea of analyzing the structure of
187 artificial networks containing only a single motif of each type for a given motif size—up
188 to five-node motifs for bipartite networks and three-node motifs for unipartite networks.
189 In addition, those networks are also set up so that any given node or link is only involved
190 in a single motif and role. Using these single-motif networks, we tested the functions of
191 *pymfinder* by ensuring that the analysis of such artificial networks does not result in the
192 misrepresentation of any motif, node, link or role.

193 **Results/Discussion**

194 The tools provided by *pymfinder* can be used in a large variety of systems and do not
195 depend on the nature or providence of the networks. To illustrate the capabilities and
196 potential of the software, we outline the study of a food web from a marine ecosystem
197 as a representative study system (28). This specific network describes the predator-prey
198 interactions between approximately 250 of the species found across an extensive area of the
199 Caribbean Sea.

200 We first analyzed the overall three-species motif structure of the network and compared
201 it to the random expectation (Fig. 2b). For this example, we used the z-score values to
202 draw this comparison, which assume normality of the motif distribution. Notice, however,
203 that *pymfinder* also returns the mean number of motif counts in the randomized networks,
204 which allows for other types of statistical analyses. We found that the observed motif
205 distribution is generally significantly different from the random expectation, showing either
206 over- or under-representation relative to the results of the null model used here. This is
207 evidence of a non-random organization of ecological communities (21, 29), which speaks to
208 the eco-evolutionary mechanisms shaping the ways in which different species interact with
209 each other. We then studied the distribution of link weights across motifs to test whether or
210 not different motifs are generally made of different interaction strengths. For this particular
211 example, we log-transformed the link weights to be approximately normally distributed as

212 well as scaled them so that the weakest and strongest links had a weight of zero and one,
213 respectively. In general, we found that interaction strengths are distributed in a similar
214 manner across the different motifs of the network under study (Fig. 2c). Notice that these
215 results are subject to the logarithmic transformation applied to the weight data, which is
216 generally very skewed (28).

217 Following the analysis of the overall motif structure, we examined the motif participation
218 of the different nodes and links that make up this food web. We found that some nodes
219 (e.g., sea cucumbers and algae) share almost identical motif-participation profiles while
220 others (e.g., filefish and sea cucumbers) have very distinct profiles (Fig. 3a). This shows
221 how motifs can be a valuable and insightful way to classify and compare the species across
222 communities. Perhaps more importantly, we observed how the information regarding the
223 interaction strengths forming the motifs changed those motif-participation profiles (Fig.
224 3b). Therefore, adding interaction strengths allowed us to distinguish between the roles
225 of species with similar unweighted profiles. This is important because it suggests that,
226 from a node-specific perspective, interaction strengths are not equally distributed across
227 motifs. The uneven distribution of interaction strengths has important implications for the
228 relationship between network structure and species–interaction strengths and the stability
229 of food webs(30, 31). We also looked at the motif-participation profiles of the links (Fig.
230 3c). We found that those profiles could also be an indicator of the observed differences on
231 the way interaction strengths are distributed across motifs, as suggested by previous work
232 (17).

233 Finally, we studied the motif-role profiles of the species of the marine network. This
234 analysis is similar to the motif participation analysis of nodes and links; however, it provides
235 a finer resolution to the role that different species or links might play in the community.
236 Using the proportion of times that the different species are in each of the 30 unique positions
237 of the three-species network motifs, we performed an analysis of multivariate homogeneity of
238 group dispersions to compare the roles of the species in the network (32). To do this, we first
239 calculated the euclidean distance between the roles of every pair of species in the network,
240 generating a dissimilarity matrix of all species. We then performed a basic clustering
241 analysis of the species-role dissimilarity matrix to find the most distinct groups of roles (Fig.

242 4). Finally, we used the function *betadisper* from the R package *vegan* (33) to perform the
243 Principal Coordinates Analysis (PCoA) of the data. We found four characteristic groups of
244 species presenting very distinct motif-role profiles. Notice that the same analysis can also be
245 done for the motif-role profile of every link in the network. This is useful because it shows
246 the diversity of structural roles in this community and underlines how those profiles could
247 be used to compare species, links or networks within and across ecosystems, environments
248 and biomes (34, 25).

249 Availability and Future Directions

250 *pymfinder* is open source software that can be freely and anonymously downloaded from
251 <https://github.com/stoufferlab/pymfinder>. The documentation of the package is at-
252 tached as supplementary material and the data used to test the software can be found
253 within the github repository. *pymfinder* has been tested to run on any platform that
254 supports Python. *pymfinder* will require you to have the Python modules Numpy and Se-
255 tupuptools installed in your machine. Data used to present the software has been previously
256 published by Bascompte et. al. (28). We are currently working on additional software that
257 uses the weighted motif-role profiles of nodes to efficiently align bipartite networks multiple
258 times.

259 References

- 260 [1] Mark EJ Newman. Spread of epidemic disease on networks. *Physical review E*,
261 66(1):016128, 2002.
- 262 [2] Olaf Sporns. Network analysis, complexity, and brain function. *Complexity*, 8(1):56–60,
263 2002.
- 264 [3] Jordi Bascompte and Pedro Jordano. *Mutualistic networks*. Princeton University Press,
265 2013.
- 266 [4] Richard J Williams and Neo D Martinez. Simple rules yield complex food webs. *Nature*,
267 404(6774):180–183, 2000.

- 268 [5] Roger Guimera, Marta Sales-Pardo, and Luís A Nunes Amaral. Module identification
269 in bipartite and directed networks. *Physical Review E*, 76(3):036102, 2007.
- 270 [6] Jordi Bascompte, Pedro Jordano, Carlos J Melián, and Jens M Olesen. The nested
271 assembly of plant–animal mutualistic networks. *Proceedings of the National Academy
272 of Sciences*, 100(16):9383–9387, 2003.
- 273 [7] Daniel B Stouffer, Juan Camacho, and Luís A Nunes Amaral. A robust measure of
274 food web intervality. *Proceedings of the National Academy of Sciences*, 103(50):19015–
275 19020, 2006.
- 276 [8] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and
277 Uri Alon. Network motifs: simple building blocks of complex networks. *Science*,
278 298(5594):824–827, 2002.
- 279 [9] Daniel B Stouffer, Juan Camacho, Wenxin Jiang, and Luís A Nunes Amaral. Evidence
280 for the existence of a robust pattern of prey selection in food webs. *Proc R Soc Biol
281 Sci*, 274(1621):1931–1940, 2007.
- 282 [10] Carsten F Dormann, Jochen Fründ, Nico Blüthgen, and Bernd Gruber. Indices, graphs
283 and null models: analyzing bipartite ecological networks. *The Open Ecology Journal*,
284 2(1), 2009.
- 285 [11] Esti Yeger-Lotem, Shmuel Sattath, Nadav Kashtan, Shalev Itzkovitz, Ron Milo, Ron Y
286 Pinter, Uri Alon, and Hanah Margalit. Network motifs in integrated cellular networks
287 of transcription–regulation and protein–protein interaction. *Proceedings of the National
288 Academy of Sciences of the United States of America*, 101(16):5934–5939, 2004.
- 289 [12] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the
290 transcriptional regulation network of escherichia coli. *Nature genetics*, 31(1):64, 2002.
- 291 [13] Janis Klaise and Samuel Johnson. The origin of motif families in food webs. *Scientific
292 Reports*, 7(1):16197, 2017.
- 293 [14] María C Rodríguez-Rodríguez, Pedro Jordano, and Alfredo Valido. Functional conse-
294 quences of plant-animal interactions along the mutualism-antagonism gradient. *Ecol-
295 ogy*, 98(5):1266–1276, 2017.

- 296 [15] Nick J Baker, Riikka Kaartinen, Tomas Roslin, and Daniel B Stouffer. Species' roles in
297 food webs show fidelity across a highly variable oak forest. *Ecography*, 38(2):130–139,
298 2015.
- 299 [16] Daniel B Stouffer, Marta Sales-Pardo, M Irmak Sirer, and Jordi Bascompte. Evo-
300 lutionary conservation of species' roles in food webs. *Science*, 335(6075):1489–1492,
301 2012.
- 302 [17] Alyssa R Cirtwill and Daniel B Stouffer. Concomitant predation on parasites is highly
303 variable but constrains the ways in which parasites contribute to food web structure.
304 *Journal of Animal Ecology*, 84(3):734–744, 2015.
- 305 [18] Nadav Kashtan, Shalev Itzkovitz, Ron Milo, and Uri Alon. Efficient sampling algorithm
306 for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*,
307 20(11):1746–1758, 2004.
- 308 [19] Sebastian Wernicke and Florian Rasche. Fanmod: a tool for fast network motif detec-
309 tion. *Bioinformatics*, 22(9):1152–1153, 2006.
- 310 [20] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network
311 research. *InterJournal, Complex Systems*, 1695(5):1–9, 2006.
- 312 [21] Jordi Bascompte and Carlos J Melián. Simple trophic modules for complex food webs.
313 *Ecology*, 86(11):2868–2873, 2005.
- 314 [22] Juan Camacho, Daniel B Stouffer, and Luís A Nunes Amaral. Quantitative analysis of
315 the local structure of food webs. *Journal of Theoretical Biology*, 246(2):260–268, 2007.
- 316 [23] Alyssa R Cirtwill and Anna Eklöf. Feeding environment and other traits shape species'
317 roles in marine food webs. *Ecology letters*, 21(6):875–884, 2018.
- 318 [24] Alyssa R Cirtwill, Tomas Roslin, Claus Rasmussen, Jens Mogens Olesen, and Daniel B
319 Stouffer. Between-year changes in community composition shape species' roles in an
320 arctic plant–pollinator network. *Oikos*, 2018.
- 321 [25] Bernat Bramon Mora, Dominique Gravel, Gilarranz J. Luis, Poisot Thimothée, and
322 Daniel B. Stouffer. Identifying a common backbone of interactions underlying food
323 webs from different ecosystems. *Nature Communications*, 2018. (accepted).

- 324 [26] Ron Milo, Nadav Kashtan, Shalev Itzkovitz, Mark EJ Newman, and Uri Alon. On the
325 uniform generation of random graphs with prescribed degree sequences. *arXiv preprint*
326 *cond-mat/0312028*, 2003.
- 327 [27] James G Sanderson, Michael P Moulton, and Ralph G Selfridge. Null matrices and
328 the analysis of species co-occurrences. *Oecologia*, 116(1-2):275–283, 1998.
- 329 [28] Jordi Bascompte, Carlos J Melián, and Enric Sala. Interaction strength combinations
330 and the overfishing of a marine food web. *Proceedings of the National Academy of*
331 *Sciences of the United States of America*, 102(15):5443–5447, 2005.
- 332 [29] Jonathan J Borrelli. Selection against instability: stable subgraphs are most frequent
333 in empirical food webs. *Oikos*, 124(12):1583–1588, 2015.
- 334 [30] Anje-Margriet Neutel, Johan AP Heesterbeek, and Peter C de Ruiter. Stability in real
335 food webs: weak links in long loops. *Science*, 296(5570):1120–1123, 2002.
- 336 [31] Mark Emmerson and Jon M Yearsley. Weak interactions, omnivory and emergent
337 food-web properties. *P Roy Soc Lond B Bio*, 271(1537):397–405, 2004.
- 338 [32] Marti J Anderson. Distance-based tests for homogeneity of multivariate dispersions.
339 *Biometrics*, 62(1):245–253, 2006.
- 340 [33] Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legen-
341 dre, Dan McGlinn, Peter R. Minchin, R. B. O’Hara, Gavin L. Simpson, Peter Solymos,
342 M. Henry H. Stevens, Eduard Szoecs, and Helene Wagner. *vegan: Community Ecology*
343 *Package*, 2017. R package version 2.4-4.
- 344 [34] Timothée Poisot, Elsa Canard, David Mouillot, Nicolas Mouquet, and Dominique
345 Gravel. The dissimilarity of species interaction networks. *Ecol Lett*, 15(12):1353–1361,
346 2012.

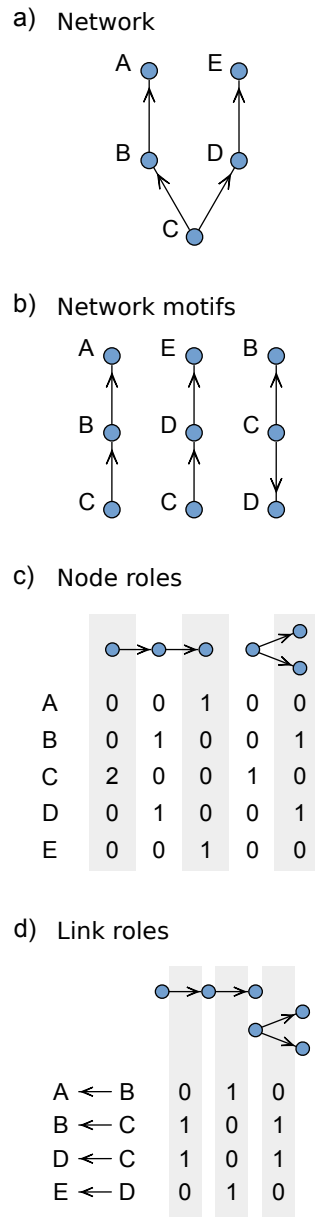


Figure 1: Main components of network-motif analysis. (a) A simple network that could represent a simple ecological community—where nodes would characterize species and the arrows would indicate the interactions between them—or a protein-protein interaction network—where nodes would represent different proteins and the arrows indicate either activation or inhibition. (b) All three-node motifs found in the network from (a); from this classification, we can compute the overall network structure and the number appearances of every node in each motif. (c) The characterization of every node’s motif-role profile. This characterization is based on the number of appearances of every node in each of the unique motif node-positions. (d) The characterization of every link’s motif-role profile, which is based on the number of appearances of every link in each of the unique motif link-positions. Notice that we excluded any motif or role that was not represented in the network.

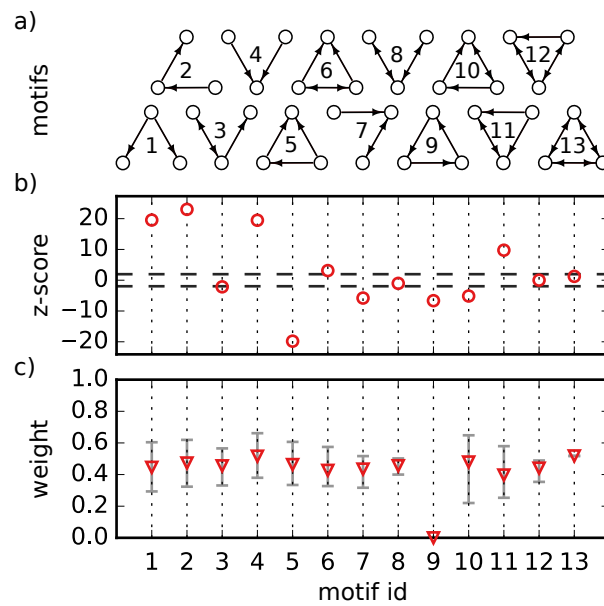


Figure 2: Analysis of the overall motif structure of the marine food web under study. The first panel (a) shows all the possible three-species motif structures. In this case, any arrow indicates the direction of energy flow from a prey to its predator. The second panel (b) presents the z-scores obtained from comparison between the empirical motif frequency and the random expectation. The dotted lines indicate the thresholds for significant over- and under-representation ($z = 1.97$ and $z = -1.97$, respectively). The third panel (c) shows the median weight found for each motif. The error bars represent the first and third quartiles. Note that the motif id given on the x-axis corresponds to the indexing in (a), and that the interaction strengths have been transformed to approximately be normally distributed and strictly positive.

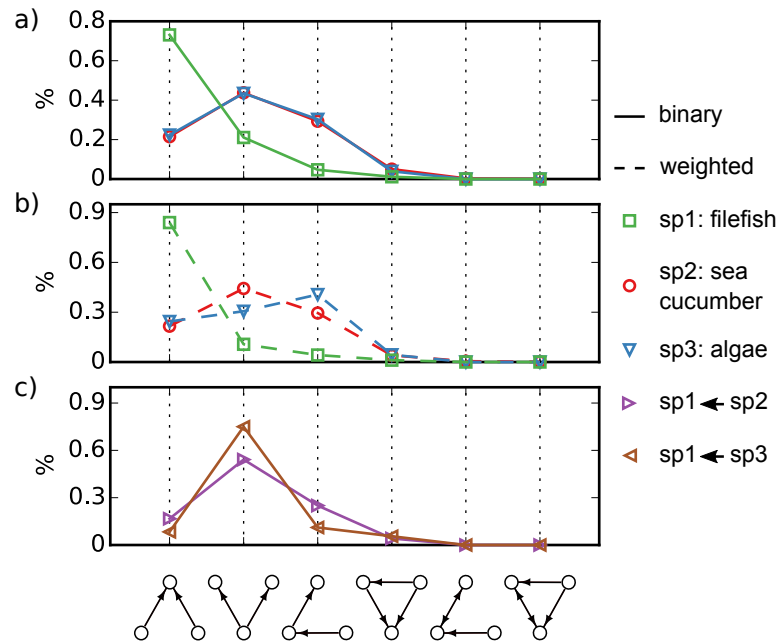


Figure 3: Analysis of the species' motif participation in the marine food web under study. The first panel (a) shows the motif-participation profiles of three representative species from the web; here, every point describes the proportion of times that these species are found in any of the possible motifs. For simplicity, we excluded the seven motifs in which these species never appear. The second panel (b) presents the motif-role profiles for the same three species when adding information regarding the interaction strengths. In this case, every point represents the relative weight associated with the motifs in which each species participates. The third panel (c) shows the motif-participation profiles for the links involving the same three species.

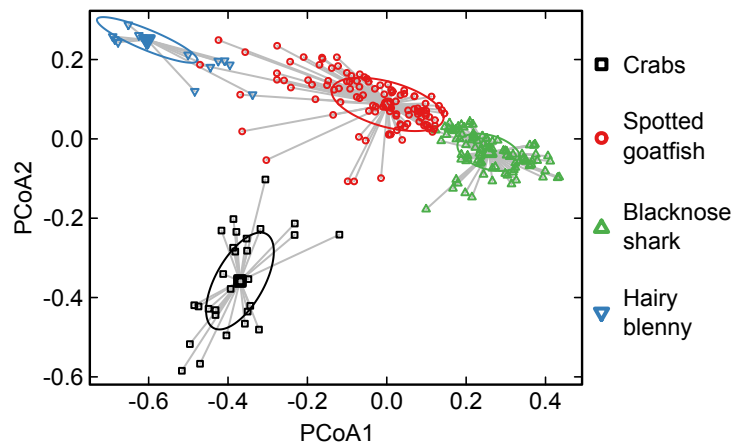


Figure 4: Principal coordinate analysis of the dissimilarity matrix containing the pairwise distances between all nodes' motif-role profiles in the marine food web under study. Every point represents a different species and each color corresponds to a group characterizing a distinct role. The species in the legend are those corresponding to the medoids of each group. The ellipses are the one-standard-deviation ellipses about the group medians.