Trait matching and sampling effort shape the structure of the frugivory network in Afrotropical forests

Clémentine Durand-Bessart1,2*, Norbert J. Cordeiro3,4, Colin A. Chapman5,6,7,8*, Katharine Abernethy9,10, Pierre-Michel Forget11, Colin Fontaine2*, and François Bretagnolle1*

1Biogeosciences, UMR 6282, Université Bourgogne Franche-Comté-CNRS, 21000 Dijon, France; 2Centre d’Écologie et des Sciences de la Conservation, CESCO, UMR 7204, MNHN-CNRS-SU, 75005 Paris, France; 3Department of Biology (mc WB 816), Roosevelt University, 430 S. Michigan Avenue, Chicago, IL 60605, USA; 4Science & Education, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605, USA; 5Wilson Center, 1300 Pennsylvania Avenue NW, Washington, DC 20004, USA; 6Department of Anthropology, Center for the Advanced Study of Human Paleobiology, The George Washington University, Washington, DC 20037, USA; 7School of Life Sciences, University of KwaZulu-Natal, Scottsville, 3201 Pietermaritzburg, South Africa; 8Shaanxi Key Laboratory for Animal Conservation, Northwest University, 710069 Xi’an, China; 9African Forest Ecology Group, School of Natural Sciences, University of Stirling, Stirling, FK9 4LA UK; 10Institut de Recherches en Écologie Tropicale, CENAREST, Gros Bouquet, 21444 Libreville, Gabon; 11Muséum National d’Histoire Naturelle, UMR 7179 MÉCADEV CNRS-MNHN, 1 Avenue du Petit Château, 91800 Brunoy, France

Summary

- Frugivory in tropical forests is a major ecological process as most tree species rely on frugivores to disperse their seeds. However, the underlying mechanisms driving frugivore–plant networks remain understudied. Here, we evaluate the data available on the Afrotropical frugivory network to identify structural properties, as well as assess knowledge gaps.
- We assembled a database of frugivory interactions from the literature with >10,000 links, between 807 tree and 285 frugivore species. We analysed the network structure using a block model that groups species with similar interaction patterns and estimates interaction probabilities among them. We investigated the species traits related to this grouping structure.
- This frugivory network was simplified into 14 tree and 14 frugivore blocks. The block structure depended on the sampling effort among species: Large mammals were better-studied, while smaller frugivores were the least studied. Species traits related to frugivory were strong predictors of the species composition of blocks and interactions among them. Fruits from larger trees were consumed by most frugivores, and large frugivores had higher probabilities to consume larger fruits.
- To conclude, this large-scale frugivory network was mainly structured by species traits involved in frugivory, and as expected by the distribution areas of species, while still being limited by sampling incompleteness.

Introduction

Biodiversity depends on multiple ecological interactions, and, in the last few decades, one of the chief goals of community ecology was to characterize the complex patterns of interactions among species and the factors shaping them (Olesen et al., 2007; Coelho & Rangel, 2018). Such interaction network approaches are particularly relevant to understand the responses of communities to disturbances of anthropogenic or natural origins (e.g. Galetti et al., 2006; Tylianakis et al., 2007).

Bipartite mutualistic interaction networks, involving beneficial interactions between two categories of species such as plants and seed dispersers or pollinators, share consistent patterns in their topological structures (Bascompte & Jordano, 2007). These mutualistic networks have been mainly studied at local or community scale, in particular regarding the heterogeneity of species generalism, with few species being highly connected and most species being poorly connected (Jordano et al., 2003). From this observation, nestedness in plant–frugivore networks was described as the overlap in interaction among species of increasing generalism (Bascompte et al., 2003; Almeida-Neto et al., 2008). By contrast, modularity arose as a term that identified groups of interacting species (modules) that are linked more densely than with other species of the network (Olesen et al., 2007). Both nestedness and modularity are predicted to relate to community stability (Thebault & Fontaine, 2010), and several processes have highlighted both ecological patterns. These nonexclusive processes include filters that constrain the possibility of links, such as the spatial (distribution matching or distribution overlap), temporal species matching (phenological matching), trait-based processes (trait matching) and phylogenetic relationships (Vázquez et al., 2009a,b; González-Varo & Travesset, 2016;
Most plant–frugivore relationships concerned local community or subset of organisms that restrict the interpretation of the structure of the network. With the increasing availability of data on species interactions, mutualistic networks have also been studied at regional or biome scales (e.g. de Almeida & de Mikich, 2018; Redhead et al., 2018; Windsor et al., 2023). Such large-scale networks, based on the aggregation of observed interactions at different sites, often named metaweb (Maiorano et al., 2020), provide more robust estimates of the topological descriptors of the networks (Quintero et al., 2022). Additionally, the integration of large-scale data allows exploring ecological and biogeographical patterns (Redhead et al., 2018; Windsor et al., 2023) and the eco-evolutionary processes shaping them (Windsor et al., 2023). Finally, large-scale networks are great tools for proposing conservation priorities or for restoring ecosystems (de Almeida & de Mikich, 2018; Windsor et al., 2023).

The recent development of latent block models offers a new perspective to understand the structure of ecological networks (Leget et al., 2015; de Manincor et al., 2020; Bar-Hen et al., 2022). Latent block models are parametric clustering methods that group species into blocks, according to their pattern of interaction, with distinct blocks of the two set of nodes of the bipartite network. Then, species in the same block of animals probabilistically share the same interactions with species from other blocks of plants and vice versa. Contrary to nestedness and modularity approaches that identify a single predefined architectural pattern, latent block models look for the structure that best explains the data. Species within blocks can therefore be considered as functionally redundant as they have a similar interaction pattern. Latent block models can not only reveal a modular or a nested structure but also highlight structures that are more complex, by allowing different numbers of blocks for each side of the network and estimating probabilities of interaction among each block couple (Fig. S1). The probabilities of interactions linking blocks identify whether a block, for example of trees, maintains a strong and/or specific relationship with a particular block of frugivores rather than with many blocks of frugivores. Therefore, the identification of the functional and taxonomic characteristics of the species within blocks appears promising to investigate (1) the relative role of functional convergence with potential associated syndromes (Ronce & Clobert, 2012; Valenta & Nevo, 2020) and (2) phylogenetic inertia with species sharing inherited traits (Olesen et al., 2007).

Very few studies investigate the structure of tropical plant–frugivore networks (de Almeida & de Mikich, 2018), and this is particularly lacking in Afrotropical forests. These forests are particularly affected by the global decline in large wildlife – downsizing crisis – and plants due to the unsustainable human extraction (Abernethy et al., 2013; Malhi et al., 2016). Afrotropical forests are some of the last remaining areas where megafauna and megafauna persist at densities that maintain their ecological functions (Terborgh et al., 2016; Berzaghi et al., 2018), and their decline raises major concerns for the future functioning and resilience of tropical ecosystems (Markl et al., 2012; Abernethy et al., 2013; Beaune et al., 2013; Effiom et al., 2013). In these forests, most trees produce fleshy fruits that depend on frugivores to disperse their seeds (Abernethy et al., 2013; Effiom et al., 2013), while being important food resources for the forest fauna, particularly the megafauna (Gautier-Hion et al., 1985; Beaune et al., 2013; Bush et al., 2020). At the continental scale, many site-scale studies on frugivory have documented diets of African frugivores, mostly on charismatic species (e.g. primates and elephants), but at a more regional level, interactions remain poorly sampled. Thus, the overall structure of tree–frugivore networks in Afrotropical forests is largely unexplored (Dugger et al., 2019).

We know from other parts of the world that frugivory interactions are mostly governed by trait matching, with the size of fruits and seeds ingested constrained by the size of the fruit eaters (Kitamura et al., 2002; Forget et al., 2007; Donatti et al., 2011; Dehling et al., 2016). We also know that frugivore body mass is related to specialization, with the biggest species tending to be more generalist than the smaller ones, as they are able to ingest both small and large fruits and seeds (Trolliet et al., 2019; Godínez-Alvarez et al., 2020). These large-bodied species increase network cohesion and thereby network stability (Vidal et al., 2014). Investigating such patterns related to body, fruit and seed size in Afrotropical forest frugivory networks is thus a priority to identify, understand and mitigate the consequences of the downsizing crisis in these forests.

Recently, de Almeida & de Mikich (2018) proposed an approach of concatenating local information to assemble a global network so that structural properties could be ascertained. Their approach, which advanced the understanding of ecological processes of network structure in Neotropical communities, inspired our study on Afrotropical frugivory networks. Here, we assemble and analyse a database aggregating the current knowledge on tree–frugivore interactions in Afrotropical forests. We quantify the database’s sampling completeness and then analyse the structure of the corresponding interaction network using a latent block model. We investigated the relationships between the block structure of the network and species taxonomy, geographic distribution, trait and conservation status as well as sampling effort. More specifically, we answered two main questions: (1) What are the determinants of the species composition of blocks? In other words, do species from the same block share particular characteristics? (2) What are the determinants of the probability of interactions of tree and frugivore blocks? In other words, are block couples with high probability of interaction made of species with matching characteristics? Our analysis first indicates that the sampling of frugivory interactions in Afrotropical forests is still far from complete and that this affects the structure that our analysis highlights. Second, we found that despite this sampling effect, the blocks we identified species grouped together with similar traits rather than similar taxonomy and distribution. Third, we found that the probability of interactions among blocks was related to expected trait and distribution matching among species. In addition, to provide an overview of current knowledge on frugivory interaction in Afrotropical forests, our study brings evidence of the relevance of block models to relate network structure to ecological processes at regional scales. Unlike most studies on...
mutualistic networks, which tend to have local scale focus (e.g. Donatti et al., 2011; Schleuning et al., 2011; Carreira et al., 2020), the results from our study provide important conceptual ideas on network structure at large spatial scales, here on the broader Afrotropical forests.

Materials and Methods

Study sites and database

We assembled a tree–frugivore interaction dataset from literature by searching both the Web of Science and Google Scholar. The search terms used were frugivor*, seed dispers*, tree-frugivore interaction, and the genus or guild name of tree and vertebrate species that inhabit Afrotropical forests. We selected literature sources presenting data on endozoochory and synzoochory with trees, palms or shrubs taller than 3 m (hereafter tree).

A total of 256 literature sources were selected (Appendix A), listing 10,547 interactions – one interaction being the consumption of the fruit of a given tree species by a given vertebrate species. This involved 807 tree and 285 frugivore species and included forests across Africa (Fig. 1). We focused our study on the 6022 unique interactions, that is an interaction between one species of tree and one frugivore species. We removed duplicates to avoid the redundancy and bias towards certain interactions.

In addition to species’ identity, we included their order, family, geographic distribution and conservation status. Distributional data were obtained from the African Plant Database of the Geneva Botanical Garden for trees, and for frugivores, these data were derived from IUCN (https://www.iucnredlist.org). Following Droissart et al. (2018), we considered three biogeographical regions as categorical variables: West Africa (W), Central Africa (C) and Albertine Rift montane (East Africa, E). Each species was assigned to one category: W, C, E, WC, WE, CE or A (whole geographical area). For conservation status, we used the species status given in the IUCN Red List of Threatened Species (2020) considering species with a critically endangered, endangered, vulnerable or near-threatened status as ‘threatened’ and species with a least concern status as ‘not threatened’.

We derived species traits of trees and frugivores from the different literature sources (Appendix A). For frugivores, we recorded body mass (g). For trees, we recorded fruit and seed dimensions (length and width in cm), number of seeds per fruit, average height (m) and wood density (g cm$^{-3}$). Tree height and wood density reveal aspects of life strategy of trees, which impact their relationships with frugivores. For example, pioneer species have often low wood densities. The wood density values for each tree species were obtained in R (R Core Team, 2021) with the function getWoodDensity from the BIOMASS package (Rjou-Méchain et al., 2017) using data from Chave et al. (2009).

Sampling completeness of the tree–frugivore interactions

To assess the sampling completeness of the dataset, we used accumulation curves for both species and interactions, examining the difference between the estimated richness calculated by the non-parametric estimator Chao 2, that is the asymptote of the accumulation curve, and the observed richness (Costa et al., 2016).

Fig. 1 (a) Location and contribution of the different study sites included in the database. Circles represent study sites, with their size being proportional to the number of species studied and their colour related to the number of studies per site. The green areas show forest cover > 30%. (b) Proportion of unique interactions involving the different frugivore groups included in the database, in clockwise order starting with ruminants, birds (large and small), carnivores, pigs, bats, primates (apes and monkeys), elephant and rodents. (c) Proportion of unique interactions involving the different tree orders included in the database (only orders involved in > 0.5% of unique interactions are presented).
We estimated the expected richness of tree and frugivore species using the full dataset of 10,547 interactions on the overall Afrotropical region and the three biogeographical regions separately. We considered one interaction record between a plant species and an animal species registered per site within each study. Similarly, we estimated the expected number of interactions per frugivore taxonomic group as in Fig. 1b, comparing it to the number of interactions included in our database. The expected Chao 2 estimator values were obtained with the `specpool` function from the `vegan` package in R (R Core Team, 2021; Chao et al., 2009).

Network structure: block model analysis

We used a latent block model to analyse the structure of the presence–absence interaction matrix between tree and frugivore species. This parametric method, based on regression models and latent variables, uses statistical inference to assign species with similar interaction patterns into groups called blocks and estimates probabilities of interaction among blocks. The number of blocks, the probabilities for species to belong to each block and the probabilities of species interaction between block couples are estimated by maximizing the integrated complete-data likelihood (Newman, 2016). This allows one to highlight network structure in a very flexible way, without predefining a structure beforehand, except the existence of blocks. We used the `blockmodels` R package (Leger, 2016) adapted for bipartite networks, with the Bernoulli family to match our binary adjacency matrix.

Relative contribution of sampling effort, species distribution area, traits and taxonomy to the block structure of the frugivory network

Several nonexclusive factors could originate the block structure or, in other words, could make pattern of interaction similar within groups of species. Among those, species could have a similar pattern of interaction because they share (1) the same distribution area and thereby the same pool of potential partners, or (2) traits involved in the choice of similar interacting partners due to a common evolutionary history. In addition, the sampling effort for each species might also affect the block structure of the network, for example species with low sampling effort having very few interactions being grouped in the same block, while species with high sampling effort having many interactions being grouped in other blocks.

To evaluate the relative contribution of these factors to the block structure of the network, despite the correlations among them, we performed a random forest model as implemented in the R package `randomForest` (Liaw & Wiener, 2002). This classification method relies on building decision trees from bootstrap samples (Fox et al., 2017) that allow one to incorporate categorical (taxonomy and distribution area) as well as continuous variables (species traits), and support possible interactions and collinearity among variables, as expected in our case among species traits and taxonomy (De & Fabricius, 2000; Loh, 2014). We quantified the variable importance by using the mean decrease in Gini index that measures how each variable contributes to the homogeneity of the classification resulting from a random forest model. For each variable, the mean decrease in Gini index is the average aggregation values of Gini index over all decision trees (Fox et al., 2017). Higher the mean decrease in Gini index values, higher the importance of the variable.

We fitted one random forest classification model for frugivore blocks and another one for tree blocks. For frugivores, we included the effects of the body mass (log-transformed), the taxonomic group and the distribution area of species, as well as the number of publications in the database to account for sampling effort. For trees, we had pairs of highly correlated trait variables, such as fruit length and width, seed length and width. For these cases, we selected the trait with the highest variable importance in determining the blocks, that is fruit length and seed length. We further added tree height, wood density and seed number as well as the taxonomic order, its distribution area and the number of publications. For tree and frugivore models, we used the seven geographic categories described previously as descriptors of distribution area.

In a second step, to visualize potential differences in tree traits among blocks, we ran a principal component analysis (PCA) (FACTOMINER; Husson et al., 2018). The variables used in the PCA were the seed length, seed width, fruit length, fruit width, tree size, number of seeds per fruit and wood density. The first dimension explained 38.26% of the variation and had a positive loading with the fruit size – width and length – and seed size – width and length – components. Dimension 2 explained 24.67% of the variance and had a positive loading with the number of seeds and a negative loading with wood density; it differentiated species with numerous seeds and low wood density from species with few seeds and high wood density. Dimension 3 explained 12.89% of the variation and had a positive loading with tree size.

Determinants of the summed-interaction-probability of blocks and average species traits

We used the summed-interaction-probability of blocks as the sum of the probabilities of interaction with other blocks estimated by the latent block model. A high summed-interaction-probability indicates strong interactions with many blocks it interacts with.

To investigate the relationships between block summed-interaction-probability and the average trait values of the species within blocks, we used a generalized linear model with a negative binomial family distribution. We included the summed-interaction-probability of blocks as the dependent variable and the average trait values of the species within blocks, as well as the number of publications that include species from the blocks, as covariates. The average trait values were average body mass (log-transformed) for frugivore blocks, and the average species coordinate from the PCA as described previously using dimensions 1–3 (Fig. S2). We selected the best models (Table S1), using the `MuMin` package (Barton, 2019).
Determinants of the probability of interaction between block couples

The probability of interaction between species from one frugivore block and one tree block might be related to the overlap in distribution areas between the species belonging to this block couple (distribution matching). Similarly, this probability might also be affected by the match in the mean trait value of the species in each block (trait matching). We quantified the proportion of 'species couples' having some overlap in their distribution area as the number of tree–animal species couples with overlapping distribution areas over the total number of possible species pairs between the two blocks. We performed a linear mixed-model regression, with the probability of interaction between tree and frugivore blocks (logit transformed) as the dependent variable and with (a) the average body mass (log-transformed) per frugivore block, (b) the mean coordinate of the first dimension of the PCA, (c) the mean coordinate of the second dimension of the PCA, (d) the mean coordinate of the third dimension of the PCA, (e) the proportion of species with overlapping distribution areas, as well as (f) the mean number of studies by tree block and (g) the mean number of studies by frugivore blocks. We implemented interactions between (a) and (b), (a) and (c), (a) and (d). We included tree and frugivore block identity as random factors, to account for pseudo-replication. We selected the best model (Table S2), using the muMIn package (Barton, 2019).

Results

Sampling completeness of the dataset

The sampling completeness for species richness reached 72% for frugivore species and 73% for tree species. Concerning the richness of interactions, our dataset only reached a sampling completeness of 42% (Fig. S3). The sampling completeness for interaction richness did not differ when considering the biogeographical region separately: 43% for Western Africa, 42% for Central Africa and 42% for Eastern Africa.

The sampling completeness of interactions by frugivore groups ranged from 3% to 66%, with elephants having the highest sampling completeness, followed by primates, bats, ruminants and birds with intermediate values, and pigs, carnivores and rodents having the lowest values (Table 1).

Table 1 Sampling completeness of interaction richness for frugivore groups.

<table>
<thead>
<tr>
<th>Number of species</th>
<th>Number of observed interactions</th>
<th>Estimated total number of interactions ± SE</th>
<th>Sampling completeness (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodents</td>
<td>21</td>
<td>199</td>
<td>6578.2 ± 3794.9</td>
</tr>
<tr>
<td>Carnivores</td>
<td>4</td>
<td>27</td>
<td>182.7 ± 127.1</td>
</tr>
<tr>
<td>Pigs</td>
<td>2</td>
<td>9</td>
<td>44.9 ± 25.4</td>
</tr>
<tr>
<td>Birds</td>
<td>141</td>
<td>2151</td>
<td>6354.5 ± 318.2</td>
</tr>
<tr>
<td>Ruminants</td>
<td>13</td>
<td>421</td>
<td>931.5 ± 81.2</td>
</tr>
<tr>
<td>Bats</td>
<td>20</td>
<td>341</td>
<td>734.0 ± 72.3</td>
</tr>
<tr>
<td>Primates</td>
<td>48</td>
<td>2767</td>
<td>5604.6 ± 192.1</td>
</tr>
<tr>
<td>Elephant</td>
<td>1</td>
<td>209</td>
<td>315.5 ± 28.9</td>
</tr>
</tbody>
</table>

The number of observed interactions corresponds to the number of unique interactions in our database. The estimated total number of interactions corresponds to the asymptotic values calculated by the Chao 2 estimator, with standard error (±SE). The sampling completeness represents the proportion of observed interactions over the estimation; the completeness range is presented in brackets.

Structure of the tree–frugivore network: blocks and interaction probability

The latent block model found the best block combination with 14 tree blocks and 14 frugivore blocks (Fig. 2). Most species were attributed with high certainty to their respective blocks, with 95% of frugivores and tree species having a probability higher than 0.8 of belonging to their block (Fig. S4).

The number of species within each block was highly variable. In particular, one block of frugivores and one block of trees, respectively named F14 and T14, encompassed a much higher number of species than the others: respectively, 105 and 427 species or 37% and 55% of the total number of species. The remaining 13 frugivore and tree blocks contained four to 33 species (mean = 14.6) and 12–84 species (mean = 40) respectively.

The probabilities of interactions between species of frugivore and tree blocks estimated by the latent block model were typically low (Fig. S5). Only 11% of the interactions had a probability > 0.5, while 63% had a probability lower than 0.1 (Fig. 2). The richest blocks F14 and T14 had probabilities of interactions < 0.1 with any other block, except for the interactions between T14 and F2 that reached 0.15 (Fig. 2). The average number of bibliographic sources per species and per block was 39.36 for tree blocks and 109.57 for frugivore blocks; this number was minimal for blocks T14 and F14, with on average 2.70 and 3.07 bibliographic sources per species respectively. The low probabilities of blocks T14 and F14 interacting with the other blocks were likely due to a lack of information regarding the interactions of their respective species.

Determinants of the species composition of tree and frugivore blocks

For both tree and frugivore blocks, the random forest model that best predicted the assignation of tree and frugivore species to their block included all traits, taxonomy, distribution area and sampling effort variables. These models had a 58% and 56% accuracy respectively for tree and frugivore blocks, indicating that nearly 60% of species were correctly assigned to their block with this set of predictors. To investigate the contribution of the predictor, we looked at the variable importance with the mean decrease in Gini index that quantifies how much the performance of the classification relies on the different variables included in the model (Fig. 3). For frugivores, the predictors with the highest importance were the body mass of the species, followed by the sampling effort, and subsequently by the distribution area and the frugivore taxonomic groups (Fig. 3a). For tree blocks, the pattern was different, with the sampling effort on tree species

New Phytologist (2023) 237: 1446–1462
www.newphytologist.com © 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.
having the highest variable importance by far, followed by a set of four tree traits including fruit and seed sizes and by taxonomic order, distribution area and seed number (Fig. 3b).

We further illustrated the relatively high variable importance of fruit and seed size using a PCA (Figs 4, S6). Six blocks tended to have distinct trait values: blocks T2, T3, T4, T7, T10 and T11. Blocks T2 and T3 had, on average, larger seeds and fruits, and also larger trees than most of the other blocks. Blocks T4, T7, T10 and T11 had, on average, smaller fruits and seeds, but while T7 was composed of smaller trees, T4 and T11 had larger trees (Figs 4, S6). Blocks T1, T4, T5 and T6 had, on average, larger trees. The second dimension of the PCA, driven by the number of seeds and wood density, did not strongly discriminate the various blocks (Fig. S6).

Despite a rather low variable importance of taxonomy for both tree and frugivore blocks, there was some taxonomic clustering within blocks, with some blocks presenting a high proportion of one or two taxonomic groups of trees or frugivores (Fig. 2). For frugivore blocks, blocks F1, F2, F5 and F6 had a high proportion of species from Central Africa. Block F3 had a high proportion of species from East Africa, with blocks F11 and F12 that also had species widespread in all three regions. Blocks F9 and F10 were mostly composed of species from West Africa. Blocks F7, F8 and F14 had a high proportion of species widespread across all three regions (Fig. 2).

Determinants of the summed-interaction-probability of blocks

The summed-interaction-probability of blocks, measured as the sum of the probabilities that species from a block have to interact...
with species from all other blocks, was for frugivores only related to the mean number of publications per species ($P = 0.002$; Table S3; Fig. 5a), and not to the mean body mass of frugivores (Table S3). For tree blocks, we found that block summed-interaction-probability was positively related to the third dimension of the PCA, related to tree size ($P = 0.002$; Table S3; Fig. 5b), but the other dimensions, as well as the number of publications by species, had no significant effect (Table S3).

Determinants of the probability of interaction between block couples

We found that the probability of interaction between tree and frugivore block couples was related to the statistical interaction between the mean value of dimension 1 of the PCA – representing the size of fruits and seeds of trees – over the species of the tree block and the mean body mass of the species of the frugivore block (Table S4). This indicated that while the probability to interact with small fruits and seeds was not affected by the mean body mass of the frugivore blocks, the probability to interact with larger fruit and seed increased with the mean body mass of the frugivore blocks (Figs 6a, S7). We also found a negative relationship between the probability of interactions between block couples and the statistical interaction between the mean value of dimension 3 of the PCA over the species of the tree block and the mean body mass of the species of frugivore blocks. This indicated that small trees had a higher probability of interactions with larger frugivores, while large trees had similar probability of interactions with small and large frugivores (Fig. 6b). As expected, the higher the mean distribution overlap was among the species between block couples, the higher the probability for them to interact was (Table S4). The mean number of studies per species and per block of trees and frugivores was also positively related to the probability of interactions between block couples, suggesting that blocks with frequently studied species had a higher probability of interactions with each other. We also found a significant positive relationship with the mean body mass of frugivore blocks, suggesting that blocks containing large frugivores had higher probabilities of interactions with tree blocks (Fig. S8). We found a negative relationship with dimensions 1 and 2 of the PCA on tree traits and a positive relationship with dimension 3 (Fig. S8). This indicates that blocks containing trees with smaller seeds and fruits – dimension 1 – had higher probability of interactions with frugivore blocks, as well as blocks with

Fig. 3 Results of random forest models illustrating the relative importance of different model variables (measured by the mean decrease in Gini index) in predicting the assignment of species to their blocks. The greater the mean decrease in Gini index, the more the associated variable contributes to predicting the assignment of species in their block. (a) Frugivores species and (b) tree species. Nb studies, number of studies; Geo. area, biogeographical regions.

Fig. 4 Representation of the tree species through the first and third axes of the principal component analysis on tree traits, regarding their blocks. The first axis (dimension 1) is positively related to the size of fruits and seeds, with large fruits and seeds having positive values. The third axis (dimension 3) differentiates small and large trees, with large trees having positive values (see Fig. S2). The ellipses are confidence ellipses representing the mean value of the block around the barycentre. Blocks are represented by different colours.
trees with high wood density and low number of seeds – dimension 2 – and blocks with taller tree species – dimension 3.

Overall, these results suggest that the structure of the network depends not only on the spatial distribution of species but also on species traits such as fruit and seed size and frugivore body mass making the imprint of trait matching visible in the block structure of the network.

Discussion

Here, we provide an assessment of current knowledge of frugivory interactions between trees and frugivores in a continent-wide scale of Afrotropical forests. Our analysis first emphasizes that we still have important gaps of information about frugivory interactions in this part of the world. Second, the latent block model approach allowed us to simplify the full network into a limited number of frugivore and tree blocks, that is groups of species with similar patterns of interaction. Third, we showed that this block structure is mainly related to both sampling effects and species traits linked to frugivory interactions. In what follows, we discuss these results in terms of the important drivers of structure in this network, and we also address how this analysis elucidates concerns about the ecological functioning of threatened Afrotropical forests.

Fig. 5  Relationships between summed-interaction-probability for frugivore (a) and tree blocks (b) with (a) mean number of studies by frugivore species inside frugivore blocks and (b) mean coordinates of tree species on dimension 3 of the principal component analysis (PCA) related to tree size; where each dot represent a block. The grey shading represent the confidence intervals.

Fig. 6  Relationships between the probability of interactions between tree and blocks (logit), and mean biological traits of tree and frugivore species. (a) How the mean seed size and fruit size of trees (represented by dimension 1 of the principal component analysis (PCA)) and the mean body mass of frugivores affect the probability of interaction among blocks. (b) How the size of trees (represented by dimension 3 of the PCA) and the mean body mass of frugivores affect the probability of interaction among blocks. Each point represents one of the interactions between one frugivore and one tree block. The colour of the point indicates the mean body mass of the frugivores, from dark blue for large frugivores to red for small frugivores. Blue and red lines represent fitted values for a mean frugivore body mass of 11.71 and 3.61 respectively.
The random forest analysis revealed that an important part of the block structure identified by the latent block model was related to variation in sampling effort among species. This was particularly evident for one block of frugivores and one block of trees that grouped together species that were poorly represented in the publications evaluated in this study. These two blocks are therefore more related to the absence of data rather than to the actual structure of the network. Importantly, these two blocks were the ones including the highest number of species, highlighting that we still miss data on the interactions of many species to fully uncover the structure of Afrotropical frugivory network (González-Varo & Traveshet, 2016). The sampling completeness of our dataset was indeed moderate and varied substantially among vertebrate groups, with large frugivores reaching higher sampling completeness. However, even for primates, a very well-studied group, only half of the interactions were known. There are comparatively little data on small birds, bats, rodents and carnivores, although they can contribute significantly to frugivory and seed dispersal (Seltzer et al., 2013; Carreira et al., 2020; Godínez-Alvarez et al., 2020). To gain a deeper understanding of these networks, the bias towards large vertebrates should be compensated by additional sampling on lesser-known species and by tree-centred studies with systematic day and night observations of fruit eaters.

Species traits, in particular those related to frugivory interactions, such as frugivore body mass or seed and fruit length for trees also had high importance in the composition of blocks that was expected from long-term studies on frugivory interactions mostly outside Afrotropical forests (Gautier-Hion et al., 1985; Donatti et al., 2011; Bender et al., 2018; Ong et al., 2022). The structure highlighted by the latent block model further stresses the functional aspects of frugivory, making this approach very relevant to study networks of ecological interactions (Bar-Hen et al., 2022). Interestingly, the variable importance of such traits was higher than the ones of taxonomic groups or distribution areas. This suggests first that, despite known phylogenetic signal in frugivory interactions (Gautier-Hion et al., 1985; Donatti et al., 2011), these traits might not be fully conserved and that some convergences might be present. It is also possible that these traits contain more information on frugivory interactions than the taxonomic levels we used in our analysis. Finally, our analyses of frugivory interactions indicate that, despite the large area covered by our data, the species composition of the blocks we identified was more related to species traits than an imprint of the spatial distribution of species.

For frugivores, the mean body mass differed among blocks, and while it seems positively related to block summed-interaction-probability, the literature bias towards large frugivore species involved in this pattern did not allow us to confirm what was shown in previous studies (Donatti et al., 2011). For trees, their height as well as fruit and seed length differed among blocks, but tree height only was positively related to block summed-interaction-probability. The absence of relationships between block summed-interaction-probability and fruit or seed length was unexpected as a negative relationship was found in Asian forests (Kitamura et al., 2002). The increase in block summed-interaction-probability with tree height indicated that tall trees were consumed by most frugivore species, which also fit with previous findings showing that some frugivores, like birds and bats, prefer taller trees (Duncan & Chapman, 1999), a trait often associated with late-successional canopy species that primarily interact with generalist species (Schleuning et al., 2011).

In evaluating the probability of interactions among blocks, we found that they were related to the overlap in the distribution area of species between block couples. This was expected as species with nonoverlapping distributions cannot interact and our network data covered an area greater than the distribution area of most species. We also found that the match in the functional traits of species between block couples had great importance, with large-bodied frugivores interacting with a large spectrum of fruits and seed size and yet being the main disperser of large-seeded fruits. Such relationships between network structure and trait compatibility between frugivores and fruits and seeds of trees had already been highlighted in the forests of Gabon (Gautier-Hion et al., 1985), for African primates and hornsbills (Poulson et al., 2001, 2002) and in other tropical forests (Donatti et al., 2011; Hawes & Peres, 2014); all these examples are compatible with a nested structure (Balcomb & Chapman, 2003; Vidal et al., 2013).

Functional redundancy for frugivory interactions, an important concept in plant–animal interactions, could further be interpreted by the blocks identified by the latent block model. As species within blocks share similar patterns of interactions, blocks grouping many species may be considered as having a high redundancy. Trees inside blocks share similar frugivores, while frugivores of the same blocks feed on similar tree species. Therefore, the loss of one tree or frugivore species inside a block may have its ecological role substituted by another species in the block. However, our analysis does not account for variation in dispersal effectiveness among frugivores, which would be necessary to understand whether redundancy of species within blocks is actually realistic (Calvino-Cancela & Martín-Herrero, 2009; Gonzalez-Castro et al., 2015, 2022). By contrast, blocks with few species and therefore low redundancy, or blocks grouping many globally threatened species might require particular conservation attention. This may be the case for four frugivore blocks (F2, F5, F9 and F10) with a high proportion of large species: apes, ruminants or monkeys that are threatened (Osuri et al., 2020). The loss of entire blocks could lead to the weakening of the network, and this could happen even before extinctions occur (Valiente-Banuet et al., 2015). In Afrotropical forests where large species are rapidly declining, due to hunting and logging, negative consequences on seed dispersal and thereby tree recruitment have been predicted and already shown by several studies (e.g. Vanthomme et al., 2010; Abernethy et al., 2013; Beaune et al., 2013; Effiom et al., 2013). Implications of their loss for Afrotropical forests are profound, even with an uneven sampling of frugivory interactions.

Our study exemplifies the relevance of latent block model to study the structure of ecological networks, showing the imprint of ecological processes on the revealed block structure despite the incompleteness of the data. The processes we highlighted...
were mainly related to the species traits involved in frugivory, as well as to overlap in distribution areas that could be expected from the spatial scale of our study. However, the full understanding of the structure of the network and associated processes is still limited by sampling incompleteness. Information on the role of the different frugivores play in seed dispersal, particularly of poorly studied groups such as small birds, bats, rodents and carnivores, is also needed to improve our understanding of changes that will result from defaunation and deforestation and, at the same time, improve future conservation strategies.

Acknowledgements

The authors thank Vincent Porcher for his help in data representation, and Etienne Akomo-Okoou and Gislain W. Ebang Ella, researchers from the Institut de Recherches en Ecologie Tropicale in Libreville, for helpful discussion. The research project was supported by the Agence Universitaire de la Francophonie. Colin A. Chapman was supported by the Wilson Foundation while working on this project. We would also like to thank the editor and anonymous reviewers who helped us improve this manuscript.

Author contributions

FB conceived the study. CF and CD-B took part in designing the study. CD-B and CF performed data analysis. FB assembled the dataset. CD-B wrote the first draft of the manuscript. NC, CC, KA and P-MF contributed substantially to the ideas and writing of the manuscript. CF and FB contributed equally to this work.

ORCID

Katharine Abernethy https://orcid.org/0000-0002-0393-9342
François Bretagnolle https://orcid.org/0000-0003-0164-840X
Colin A. Chapman https://orcid.org/0000-0002-8827-8140
Norbert J. Cordeiro https://orcid.org/0000-0003-2807-0636
Clémentine Durand-Bessart https://orcid.org/0000-0001-9896-5078
Colin Fontaine https://orcid.org/0000-0001-5367-5675
Pierre-Michel Forget https://orcid.org/0000-0002-9252-974X

Data availability

The data that support the findings of this study are available in the following doi: 10.5281/zenodo.7313425.

References


Appendix A

Literature sources used in the study


Cooke CA. 2012. The feeding, ranging, and positional behaviors of Cercocebus torquatus, the red-capped mangabey. Sette Camara Gabon: a phylogenetic perspective. Doctoral dissertation, Ohio State University, Columbus, OH, USA.

Cordeiro NJ, Munisi BJ. 2008. The feeding, ranging, and positional behaviors of Cercocebus torquatus, the red-capped mangabey. Sette Camara Gabon: a phylogenetic perspective. Doctoral dissertation, Ohio State University, Columbus, OH, USA.


Djigo-Djossou S, Kone I, Fandohan AB, Djigo JG, Huyenen MC, Sinbin S. 2015. Habitat use by white-thighed colobus in the Kikélé Sacred Forest:


Fimbel C. 1994. The relative use of abandoned farm clearings and old forest habitats by primates and a forest antelope at Tiwai, Sierra Leone, West Africa. *Biological Conservation* 70: 277–286.


Supporting Information

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

Fig. S1 Example of the treatment of plant–animal interactions networks by modularity and latent block model.

Fig. S2 Results showing the correlation among tree traits from the principal component analysis.

Fig. S3 Accumulation curves showing the sampling completeness of the frugivory network.

Fig. S4 Graphs of the probability of each tree and frugivore species belonging to its block.

Fig. S5 Graph of the probability of interaction between tree blocks and frugivore blocks.

Fig. S6 Clustering of species by blocks regarding their trait values on the principal component analysis.

Fig. S7 Matrix showing the probability of tree and frugivore blocks interacting following species traits.

Fig. S8 Graphs showing the relationships between the probability of interactions between blocks and species traits.

Table S1 Result from the model selection of models including the sum of probabilities by blocks and species traits and number of studies by blocks.

Table S2 Results of the selection of model including the probabilities of interaction between tree and frugivore blocks.

Table S3 Results from the generalized linear model showing the relationships with the sums of probability between tree and frugivore blocks.

Table S4 Results from the linear mixed regression showing the relationships between the probability of interactions between blocks and species traits.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.