

# Influences of sampling effort on detected patterns and structuring processes of a Neotropical plant–hummingbird network

Jeferson Vizontin-Bugoni<sup>1,2\*</sup>, Pietro K. Maruyama<sup>1</sup>, Vanderlei J. Debastiani<sup>3</sup>, L. da S. Duarte<sup>3</sup>, Bo Dalsgaard<sup>2</sup> and Marlies Sazima<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Ecologia, Universidade Estadual de Campinas (Unicamp), Cx. Postal 6109, CEP 13083-862, Campinas, SP, Brazil; <sup>2</sup>Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100, Copenhagen Ø, Denmark; and <sup>3</sup>Programa de Pós-Graduação em Ecologia, Universidade Federal do Rio Grande do Sul – UFRGS, Porto Alegre, RS, Brazil

## Summary

1. Virtually all empirical ecological interaction networks to some extent suffer from under-sampling. However, how limitations imposed by sampling incompleteness affect our understanding of ecological networks is still poorly explored, which may hinder further advances in the field.

2. Here, we use a plant–hummingbird network with unprecedented sampling effort (2716 h of focal observations) from the Atlantic Rainforest in Brazil, to investigate how sampling effort affects the description of network structure (i.e. widely used network metrics) and the relative importance of distinct processes (i.e. species abundances vs. traits) in determining the frequency of pairwise interactions.

3. By dividing the network into time slices representing a gradient of sampling effort, we show that quantitative metrics, such as interaction evenness, specialization ( $H_2'$ ), weighted nestedness (wNODF) and modularity ( $Q$ ; QuanBiMo algorithm) were less biased by sampling incompleteness than binary metrics. Furthermore, the significance of some network metrics changed along the sampling effort gradient. Nevertheless, the higher importance of traits in structuring the network was apparent even with small sampling effort.

4. Our results (i) warn against using very poorly sampled networks as this may bias our understanding of networks, both their patterns and structuring processes, (ii) encourage the use of quantitative metrics little influenced by sampling when performing spatio-temporal comparisons and (iii) indicate that in networks strongly constrained by species traits, such as plant–hummingbird networks, even small sampling is sufficient to detect their relative importance for the frequencies of interactions. Finally, we argue that similar effects of sampling are expected for other highly specialized subnetworks.

**Key-words:** connectance, forbidden links, modularity, nestedness, network metrics, neutrality, NODF, plant–pollinator networks, QuanBiMo, specialization

## Introduction

In the last decades, the understanding of mutualistic plant–animal interactions at the community scale has greatly advanced due to an increasing use of network approaches (Jordano 1987; Bascompte 2009; Dormann *et al.* 2009; Vázquez *et al.* 2012, 2015; Schleuning *et al.* 2014; Kissling & Schleuning 2015). This has revealed sev-

eral consistent patterns in the structure of bipartite plant–animal networks. Notably often only a small proportion of possible links are actually realized, resulting in low connectance (Jordano 1987); networks are often nested and modular (Bascompte *et al.* 2003; Olesen *et al.* 2007); the degree distribution is skewed with most species having few links and few species having many links (Jordano, Bascompte & Olesen 2003); and there is high asymmetric dependence between partners in a given network (Jordano 1987; Bascompte, Jordano & Olesen 2006). These struc-

\*Correspondence author: E-mail: jbugoni@yahoo.com.br

tural properties are expected to be associated with community stability and maintenance (Bascompte, Jordano & Olesen 2006; Thébaud & Fontaine 2010), ecosystem functioning (Schleuning, Fründ & Garcia 2015) and to have implications for conservation (Tylianakis *et al.* 2010).

An array of ecological, historical and evolutionary processes may influence network structure (Vázquez, Chacoff & Cagnolo 2009b; Vázquez *et al.* 2009a; Dalsgaard *et al.* 2013; Schleuning *et al.* 2014; Vizentin-Bugoni, Maruyama & Sazima 2014; Martín González *et al.* 2015; Renoult *et al.* 2015). Additionally, chance meeting governed by species abundances, i.e. neutrality, may determine the structure of networks (Stang, Klinkhamer & van der Meijden 2007; Vázquez, Chacoff & Cagnolo 2009b). In this regard, several recent studies have investigated the relative importance of distinct processes in structuring mutualistic networks. Most of these studies have found species abundances as a major factor determining interactions in the networks, with a complementary role of species traits (e.g. Stang, Klinkhamer & van der Meijden 2007; Vázquez, Chacoff & Cagnolo 2009b; González-Castro *et al.* 2015; Olito & Fox 2015). Nevertheless, in some more specialized systems, such as Neotropical plant–hummingbird networks, contrasting results were found with species traits predicting interactions better than abundances (e.g. Maruyama *et al.* 2014; Vizentin-Bugoni, Maruyama & Sazima 2014). Importantly, sampling may also influence the detected network structure, making it critical to consider in order to disentangle biological processes from methodological shortcomings (Blüthgen *et al.* 2008; Ulrich 2009; Vázquez *et al.* 2009a; Chagnon 2015). Thus, limitations related to sampling may reduce our understanding of the structure of interaction networks (Vázquez, Chacoff & Cagnolo 2009b) and make spatio-temporal and taxonomic comparisons problematic (Blüthgen *et al.* 2007; Chacoff *et al.* 2012; Fründ, McCann & Williams 2015). Despite the fact that sampling incompleteness may influence network patterns, most studies provide no estimate of sampling effort, but assume that interactions in the given community were sufficiently sampled to describe the associated network (Ollerton & Cranmer 2002; Vázquez *et al.* 2009a; Gibson *et al.* 2011; Chacoff *et al.* 2012; Rivera-Hutinel *et al.* 2012; Fründ, McCann & Williams 2015). Considering that even estimates of species making up the community suffer from difficulties of sampling (Gotelli & Colwell 2001), limitations of sampling should be analogous for interaction networks which represent distinct combinations of species.

To date, only a handful of studies have explicitly evaluated the effects of sampling incompleteness on the description of network structure. These studies found high variation in the number of detected species and total number of links and suggest that some metrics, that is aggregated statistics describing network patterns, are prone to sampling bias (e.g. Goldwasser & Roughgarden 1997; Banašek-Richter, Cattin & Bersier 2004; Nielsen & Bascompte 2007; Blüthgen *et al.* 2008; Dormann *et al.*

2009; Martinez *et al.* 1999; Rivera-Hutinel *et al.* 2012; Fründ, McCann & Williams 2015). Nevertheless, some important gaps remain to be addressed regarding the importance of sampling completeness. Notably, little is known about the influence of sampling effort on quantitative network structure, that is metrics calculated from networks that take into account the strength of interactions (but see Blüthgen, Menzel & Blüthgen 2006 and Blüthgen *et al.* 2008 for some indices and Fründ, McCann & Williams 2015 for numerous metrics using simulated data) and we know of only one attempt at investigating the effects of sampling incompleteness on the understanding of the processes structuring ecological networks (Olito & Fox 2015). Moreover, all information about the influence of sampling on detected network structures is based on either simulated networks (e.g. Fründ, McCann & Williams 2015) or is from less specialized and potentially incompletely sampled networks from temperate regions (e.g. Nielsen & Bascompte 2007; Hegland *et al.* 2010; Chacoff *et al.* 2012; Rivera-Hutinel *et al.* 2012). Although a reasonable description of plant–pollinator network structure may be possible with a sampling focused during the peak-season in temperate areas (Hegland *et al.* 2010), such a short sampling span is unlikely to be enough for networks from tropical areas, where many pollinators' life and/or activity span longer periods and their food plants typically present staggered flowering all year-round.

By employing an unprecedented sampling effort across 2 years and focusing on a specialized and relatively easily sampled subnetwork of plants and hummingbirds, we built an interaction network with unusually high sampling completeness. This characteristic of our study system offers an unique opportunity to evaluate the susceptibility of network patterns and structuring processes under little variation in the network species richness, as well as greater confidence that unobserved interactions actually do not happen. Specifically, we ask (i) to what extent is network structure, as measured by both binary and quantitative metrics, affected by sampling effort? and (ii) does sampling effort affect the relative importance of distinct processes (i.e. species abundance vs. traits) in determining pairwise frequencies of interaction in the network? As a model system, we use a plant–hummingbird subnetwork embedded in a larger diverse network from a species-rich area in the Brazilian Atlantic Rainforest, where forbidden links are known to play a major role in structuring the interactions (Vizentin-Bugoni, Maruyama & Sazima 2014).

## Materials and methods

### STUDY AREA AND DATA COLLECTION

The data used here represent an update of our previous study (Vizentin-Bugoni, Maruyama & Sazima 2014) by covering additional plant species and one more year of sampling. Data were collected along 12 000 m of trails in the Atlantic Rainforest

from Santa Virgínia Field Station at Serra do Mar State Park, SE Brazil (23°17'S–23°24'S and 45°03'W–45°11'W), from September 2011 to August 2013, over 4–10 days per month. We sampled interactions by observing at least three individuals of ornithophilous or potentially hummingbird-pollinated plants, away 100 m (or more) from each other along the trails. We included in the network all native species blooming at least once during our 2 years of sampling and which were possible to observe for 50 h (Table S1, Supporting information). Five species with slightly lower sampling were also included: *Edmundoa lindeni* (Regel) Leme (38 h), *Macrocarpaea rubra* Malme, *Sinningia cooperi* (Paxton) Wiehler, *Wittrockia superba* Lindm. (44 h each) and *Nidularium rutilans* E. Morren (46 h). In total, we carried out 2716 h of focal observation in which we identified the visiting hummingbirds that touched anthers and stigmas (Table S2) as well as the precise moment the visits occurred across the 50 h of sampling for each plant species. Individual-based rarefaction indicates that most links in the community were recorded (Fig. S1) and links richness estimation suggests that we observed *c.* 82% of all interactions (123 observed links from  $150.3 \pm 16.9$  expected by Chao 1 estimator; Chao 1984; Chacoff *et al.* 2012).

We monitored flower and hummingbird phenologies by assigning their presence or absence monthly (Tables S3 and S4). We did not consider a finer time scale for phenology as it was not possible to sample all coflowering plant species within the same hour or day. For each of the hummingbird-pollinated species, we measured corolla depth from *c.* 10 flowers collected in the field (Table S5) and hummingbird bill lengths were measured from museum specimens (Table S6). In order to better estimate the hummingbirds' ability to access nectar in deep flowers, we added 80% as a correction for tongue extension to bill length estimates. This correction is based on measures from *Selasphorus rufus* tongue extension (Grant and Temeles 1992) and keeps the proportionality in which longer-billed species tend also to have longer tongues (Paton & Collins 1989) (Table S6). Using an alternative 33% threshold to correct for tongue extension (as in Vizentin-Bugoni, Maruyama & Sazima 2014) did not alter our results (Table S10, see model "M1"). Plant abundance was quantified as the total number of open flowers counted per species monthly along all trails (Table S1). In order to obtain data for all species, hummingbird abundances were measured as the proportion of days a species was recorded in our 130 days in the field (Table S2). We used this measure because some rarer species were not recorded during our counting of hummingbirds. The reliability of this estimation of hummingbird abundance is supported by its positive and strong correlation with the number of aural and visual contacts of species across ten transects sampled monthly (100 m; Fig. S2; Table S7). All other sampling details on phenological, morphological and abundances data followed Vizentin-Bugoni, Maruyama & Sazima (2014).

## DATA ANALYSIS

### Interaction networks along a sampling effort gradient

Plant–hummingbird interaction data were assembled into a quantitative bipartite network, with pairwise interaction frequencies representing the number of legitimate visits between a given hummingbird and plant species. Thus, each cell ( $a_{ij}$ ) represents the number of interactions between a pollinator ( $i$ ) and plant ( $j$ ) species. To construct a sampling gradient, we pooled all interactions

observed into time slices of 1 h of sampling effort; for instance, the first-hour time slice was composed by all interactions recorded by plants and pollinators in the first hour of observation to each plant species. If a plant species was not visited in this interval, it was not included in this specific time slice. Then, by summing time slices sequentially, we created a gradient of networks with increasing sampling effort. Therefore, the gradient was composed of 50 networks, from one to 50 h of observation for each plant species in the community.

### Network patterns

In order to evaluate how sampling effort influences network patterns, we computed network metrics for the 50 networks with accumulated sampling effort. We first calculated descriptors of networks including the number of plants, richness of pollinators, number of links (i.e. pairwise combinations) and number of interactions recorded. Moreover, we also calculated binary and quantitative network metrics which are widely used in the literature and cover distinct network properties: *Connectance* is defined as the proportion of possible links actually observed in the network; *Interaction evenness* is a measure of variation in frequencies of interactions, that is visits, among distinct links (e.g. Bersier, Banašek-Richter & Cattin 2002). *Specialization* was quantified by the  $H_2'$  index, which is an application of information theory to quantitative networks and can be interpreted as a measure of how species partition their interactions in the network (Blüthgen, Menzel & Blüthgen 2006); *Nestedness* was measured by NODF index, both for binary and quantitative networks, the latter denoted as wNODF (Almeida-Neto & Ulrich 2011). *Modularity* ( $Q$ ) was also estimated for both binary and quantitative matrices using the QuanBiMo optimization algorithm (Dormann & Strauss 2014). As the QuanBiMo algorithm is stochastic, the values found can be slightly different between runs; we followed Maruyama *et al.* (2014) and Schleuning *et al.* (2014) and account for this by choosing the higher values from 10 independent runs set to  $10^7$  swaps to each network. The significance of metrics was assessed by comparing the observed values to those obtained by 1000 null model randomizations, with the exception of modularity, which was tested against 100 runs due to large computational time required by the algorithm. Calculation of NODF significance was performed using the software ANINHADO and tested with the null model in which interactions are distributed proportionally to the marginal totals of the network (Guimarães & Guimarães 2006). For binary  $Q$ , we used the *shuffle.web* function that relocates entries in the matrix keeping the original dimensions and, for all quantitative indices, we used the *vaznull* null model which keeps the marginal totals and the connectance in the network, both ran in R package *bipartite* (Dormann, Gruber & Fründ 2008). Metric values were considered significant if they did not overlap 95% of confidence intervals of the randomized values.

### Rarefaction-like approaches for interaction networks

One can argue that the order of the focal observations in the field used to assemble the network could affect metric values, especially across the sampling gradient. For instance, if the first plant sampled was an individual located in a less visible spot, then it could be less visited not by barriers related to the plant trait, but by some stochasticity. Therefore, this would affect the values of

metrics calculated for each time slice. In order to account for this potential bias, we used an analysis inspired in sample-based rarefaction method used for species richness (Gotelli & Colwell 2001). In this case, our samples were time slices of 1 h of observation to each plant species, with the respective links (and number of visits) observed. As we observed 50 h, each plant species also had 50 time slices. We then randomly assembled networks with accumulating sampling effort from 1 to 50 h of accumulated observation. To each class of sampling effort (1–50 h), we generated 1000 randomized networks and calculated all network metrics to contrast to the observed variation in the metric values with increasing sampling.

We also checked the robustness of our findings by simulating an individual-based rarefaction-like gradient of sampling effort, also inspired on Gotelli & Colwell (2001). We successively removed 10% of the interactions recorded from the complete matrix (50 h), from 10% to 90% of interactions removal and all metrics were recalculated for each removal level. We then performed 1000 repetitions for each removal level and calculated all above mentioned metrics. Results are presented in Fig. S3 which can be directly compared with Fig. 1. As the results were roughly similar regardless of the methods, we kept the ‘sample-based rarefaction-like method’, which better reflect sampling procedures in the field conducted through timed observations (see Gibson *et al.* 2011).

### Processes entangling interactions

We investigated the relative importance of phenological overlap, morphological matching and species abundance (or combinations of them) in predicting pairwise frequencies of interactions. We used the conceptual and analytical framework proposed by Vázquez, Chacoff & Cagnolo (2009b) and subsequent adaptations by Vizontin-Bugoni, Maruyama & Sazima (2014) which is based on probability matrices (null models) and likelihood analysis. We produced eight probability matrices based on Phenological overlap (**P**) – the probability of interaction between two species is given by the number of months a hummingbird and a flowering plant species overlapped in their occurrence ( $a_{ij} = 0$ –24 months); Morphological matching (**M**) – an interaction is allowed ( $a_{ij} = 1$ ) only if the hummingbird bill + tongue length estimates exceed the floral corolla depth, otherwise the probability of interaction is zero ( $a_{ij} = 0$ ); Abundances (**A**) – the probability of an interaction is proportional to the multiplication of the relative abundances of plant and hummingbird species in the community. In this sense, two abundant species have higher probability to interact than rarer species. Note that in **A**, all species combinations are allowed ( $a_{ij} > 0$ ), while in **P** and **M**, some interactions are considered ‘forbidden links’ ( $a_{ij} = 0$ ) when species do not overlap in time or have no morphological matching, respectively. In order to investigate potential effects of combined mechanisms on the network, we also multiplied these matrices (**A**, **P** and **M**) in all possible combinations by element-wise multiplication (Hadamard product), that is **PM**, **PA**, **AM** and **APM**. Finally, all matrices were normalized by dividing each cell by the matrix total. As a benchmark, we produced an additional matrix in which all species had the same probability to interact (**Null**).

We used likelihood analysis and model selection to compare the capacity of each probabilistic matrix (**A**, **P**, **M**, **AP**, **AM**, **PM**, **APM** and **Null**) in predicting frequencies of pairwise

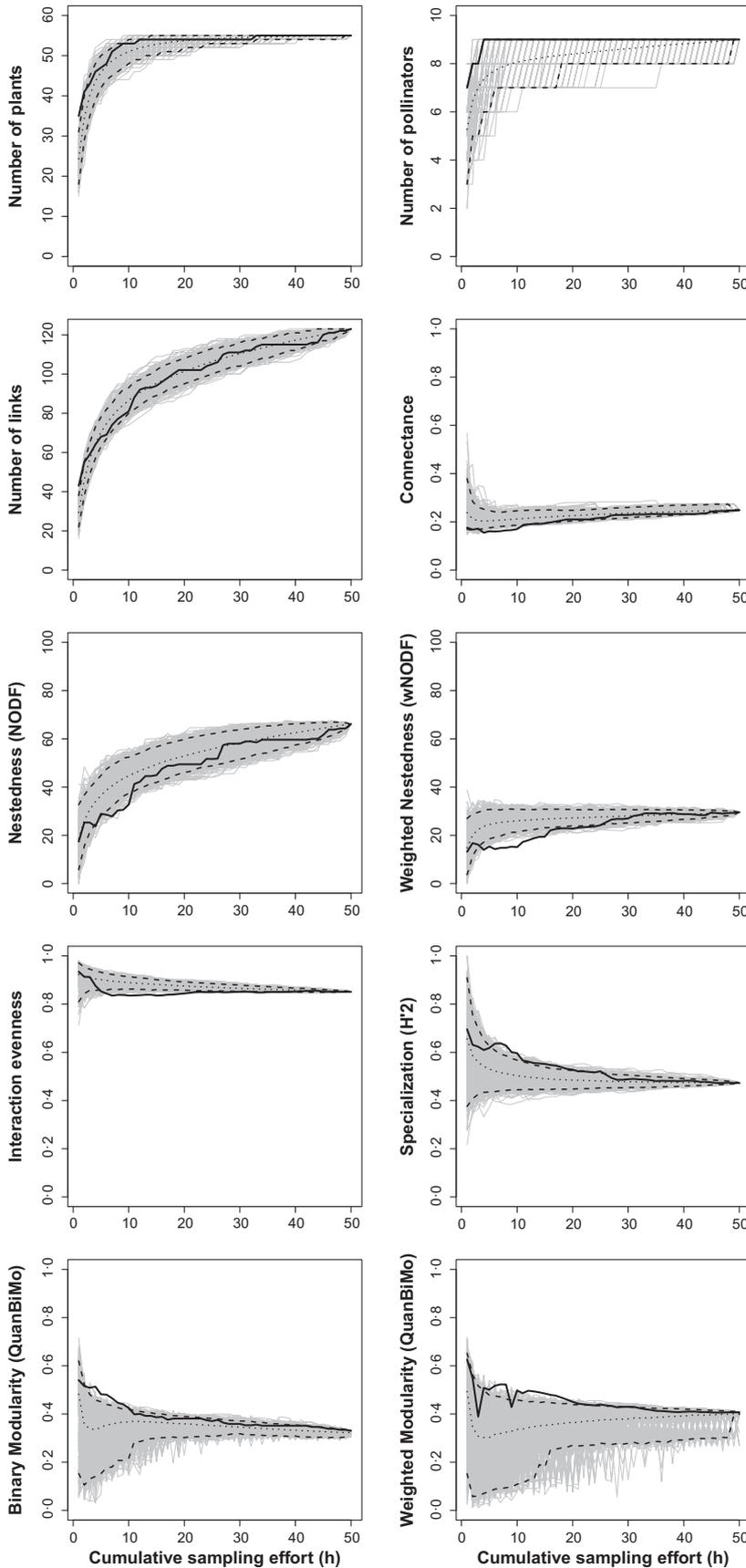
interactions in the observed matrix. Model matrices were considered equal when  $\Delta AIC \leq 2$  (Burnham & Anderson 2002). To penalize models with distinct complexities, we used as parameters the number of species contained in the matrix (plant + hummingbird species) multiplied by the number of probabilistic matrices considered in the model (i.e. **A**, **P** and **M** = 1; **AP**, **AM** and **PM** = 2; and **APM** = 3; see Table S8). In order to identify whether the best predictor changes across the sampling gradient, we repeated this analysis in networks built with cumulative sampling effort (1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50 h). Analyses on probability matrices were conducted using *bipartite* package in R, assuming that the probability of interactions follows a multinomial distribution (Vázquez, Chacoff & Cagnolo 2009b). Finally, in order to estimate the fit of the best model to the data, we performed Spearman’s rank-order correlation between the observed frequency of pairwise interactions and its corresponding probability of interactions from the best model (i.e. the PM model, see Results). We repeated the correlation for each time slice of sampling effort and presented the coefficient of correlation in Fig. 2.

## Results

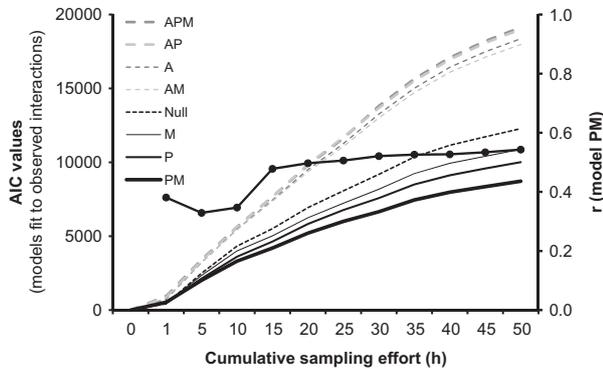
### NETWORK STRUCTURE AND METRICS VARIATION ACROSS THE SAMPLING EFFORT GRADIENT

We recorded 55 hummingbird-pollinated plants and nine hummingbird species, which performed 2793 visits distributed among 123 distinct pairwise interactions. The plant–hummingbird network presented moderate connectance (0.25), high interaction evenness (0.85; 95% CI: 0.66–0.68), moderately high specialization ( $H_2' = 0.47$ ; 95% CI: 0.11–0.18) and high nestedness when considering binary information (NODF = 66.15; 95% CI: 37.7–38.2), but a non-nested pattern when considering quantitative data (wNODF = 29.50; 95% CI: 42.5–55.2). Conversely, the network was non-modular in the binary version ( $Q = 0.33$ ; 95% CI: 0.31–0.34), but significantly modular when considering quantitative information ( $Q = 0.41$ ; 95% CI: 0.03–0.13), with four distinct modules (Fig. 1; Tables S9 and S11).

All nine hummingbird species were recorded with 3 h of observation for all plants. Although some plant species only received visits after 33 h of sampling and were then included in the network, 96% of the species received at least one visit before completing 10 h of observations (Fig 1; Table S9). In spite of the inclusion of most plants and hummingbirds early in the sampling gradient, new links were still recorded close to 50 h of sampling. Connectance, interaction evenness and  $H_2'$  achieved reasonably stable values around 10–15 h, and  $H_2'$  became significantly higher than expected by the null model after 5 h. On the other hand, values of NODF increased consistently along the sampling gradient, although already significant also after 5 h. Weighted nestedness, wNODF, achieved an asymptotic value around 30 h, but was only significant under very small sampling (1 h of observation). Binary modularity,  $Q$ ,



**Fig. 1.** Network patterns described by metrics calculated on the observed network (black line) or expected (grey overlapping lines) by randomizations (1000 iterations) of the data across increasing sampling effort of a plant–hummingbird network in SE Brazil. Dashed black lines indicate mean and 95% CI. Note that quantitative metrics were, in general, more robust to low sampling effort. Note also the observed sequence of sampling (black lines), in general, falls out the networks assembled randomly (1000 grey lines) from a pool of observed samples (1 h time slices), which suggest that random changes in the order of plant individuals observed had some effect on the value of network metrics, especially under low sampling.



**Fig. 2.** Ability of eight models to predict observed frequency of interaction between pairwise species over increasing sampling effort in a hummingbird–plant network. Models are probability matrices based on species abundance (**A**), phenological overlap (**P**), morphological matching (**M**) and all possible combinations among them. The **Null** model is a benchmark model that assumes all interactions have the same probability to occur. Note that after 5 h of sampling effort the model **PM**, which includes both phenological overlap and bill-corolla (morphological) matching, had the best ability to predict pairwise interaction while those models including **A** had worst fits, even worse than the **Null** model. This means that the higher importance of forbidden links in detriment of abundances was identified even under low sampling. We also show the change in coefficient of correlation between the observed frequency of pairwise interactions and the frequency predicted by the best model **PM** (right axis, black line with dots in the plot). Note that the correlation coefficient tends to increase with sampling effort (up to *c.* 15 h). All correlations were  $P < 0.001$ .

progressively decreased with sampling completeness but was always non-significant; quantitative  $Q$  showed oscillation under small sampling but stabilized after 20 h of observation. Although quantitative  $Q$  was significant after just 1 h of observation, the number of identified modules stabilized only after 15 h. The order of sample slices that we used to assemble the networks across the sampling gradient had no or only small influence on most metrics. Only connectance, wNODF and binary modularity presented slightly different values from those obtained in randomly reassembled networks under small sampling effort (<15 h; compare black lines with grey trends in Fig. 1, see also Fig. S3).

#### PROCESSES STRUCTURING THE NETWORK ACROSS THE SAMPLING EFFORT GRADIENT

With very small sampling effort ( $\leq 2$  h), the **Null** model had the best predictive ability of the observed network (Fig. 2, Table S10). However, after 3 h of observation the model **PM**, which includes both phenological overlap and bill-corolla matching, had consistently the best ability to predict the frequency of pairwise interactions (Fig. 2). Furthermore, after 15 h of sampling effort, the coefficient of correlation remained consistent with a slight increase up to  $r = 0.543$  in the most complete network (50 h).

## Discussion

We provide a thorough evaluation on the effects of sampling effort on ecological networks by analysing widely used binary and quantitative metrics as well as on the processes defining interactions in networks. Our findings show that low sampling effort to some extent influences our understanding of interaction patterns, but has minor influence on the identification of processes structuring a tropical plant–hummingbird subnetwork. Specifically, we found that network metrics are not all equally affected by sampling: quantitative metrics tend to be more robust than binary ones, and the significance of network metrics changes along the sampling gradient. These findings were consistent regardless of the methods applied to create the gradient of sampling effort, that is sample-based or individual-based rarefaction-like methods. Furthermore, inferences on the relative importance of processes determining frequencies of interaction were biased only under very small sampling effort, and we consistently identified the importance of traits as the main drivers of interactions as sampling increased.

#### ARE DETECTED NETWORK PATTERNS BIASED AT LOW SAMPLING EFFORT?

The Santa Virgínia network was structurally similar to other tropical plant–hummingbird networks, presenting more plants than hummingbird species, moderate connectance, intermediate specialization, and high modularity (Maruyama *et al.* 2014, 2015; Vizentin-Bugoni, Maruyama & Sazima 2014; Martín González *et al.* 2015). The order of the focal observations, that is time slices, to assemble the network had some effect on connectance, wNODF and binary modularity, at least under smaller sampling effort (*c.* 15 h; Fig. 1). This further reinforces the idea that poorly sampled networks might lead to wrong inferences on network patterns. In addition to the methods chosen to assemble the networks (Gibson *et al.* 2011), our findings thus reveal that the order and choice of which individual plant will be observed (and how many of them) can also influence some of the observed metrics. Since species richness in our system stabilizes after *c.* 10 h, detected biases on the network metrics beyond 10 h of sampling reflect the effect of addition of links and distribution of the interactions among links, more than the dimension of the network, as suggested previously (see below).

Effects of low sampling effort on network structure depend on the metric considered, and two different biases emerged: in the metric values *per se* and their significances. To understand these potential biases is crucial when comparing network metric values between webs, either to investigate geographical (e.g. Ollerton & Cranmer 2002; Olesen and Jordano 2002; Banašek-Richter, Cattin & Bersier 2004; Dalsgaard *et al.* 2013; Schleuning

et al. 2014; Martín González et al. 2015) or temporal patterns (e.g. Petanidou et al. 2008).

Connectance is known to be strongly biased by the addition of new species in the network (e.g. Jordano 1987; Banašek-Richter, Cattin & Bersier 2004; Nielsen & Bascompte 2007; Rivera-Hutinel et al. 2012; Fründ, McCann & Williams 2015; but see Martínez et al. 1999). Previous studies with plant–insect pollinator networks have found that connectance tends to decrease along the sampling gradient (Nielsen & Bascompte 2007; Rivera-Hutinel et al. 2012), which may be caused by the discovery of new species happening faster than of new links. This is not the case in our system, since network size reached a constant relatively quickly, that is most hummingbird species were recorded visiting a particular plant species early in the sampling gradient. This may partly be because hummingbirds are very active and need to frequently visit flowers to cope with their high metabolism (Suarez 1992); thus, they are readily recorded and incorporated in the interaction network. In contrast to plant–hummingbird networks, one should expect stronger effects of sampling incompleteness on connectance in other sub-networks such as orchid–pollinator systems in which many species produce no or scarce nectar and interactions are rarely recorded (e.g. Ackerman, Rodriguez-Robles & Melendez 1994) or larger ‘full’ networks, that is including several animal taxa (e.g. Donatti et al. 2011; Danieli-Silva et al. 2012).

Another widely used binary metric, NODF, consistently increased with species and links inclusion in the network (Fig. 1). The same was found by Rivera-Hutinel et al. (2012), who argued that the increase in NODF could be related to the increase in network size. However, this argument is not supported here since network size did not increase much after *c.* 10 h of sampling, nevertheless NODF value kept increasing. This suggests that NODF is dependent on the detection of links for highly connected species, which are major contributor for high nestedness in networks (Almeida-Neto et al. 2008). Thus, the use of NODF to compare nestedness estimates across time, space or systems should be done carefully, and it is safer to compare this metric preferably for networks that had reached both a stable size and number of links. In this sense, binary nestedness seems to be more influenced by sampling incompleteness than previously thought (Nielsen & Bascompte 2007). In contrast, quantitative nestedness (wNODF) was more robust, but still tended to increase with sampling intensity, as also found by a simulation study (Fründ, McCann & Williams 2015). Even with stable network size, this metric was progressively affected by the detection of new links (10–35 h) and first stabilized after *c.* 35 h of sampling. This suggests that high sampling effort in order to accumulate most of the links may be needed before its use for spatio-temporal comparisons. The same may be expected for binary modularity ( $Q$ ), though less so, since its value decreased consistently up to *c.* 15 h of sampling effort. Interestingly, the same trend of

decreasing binary modularity along a sampling gradient was found for a plant–pollinator network from Chilean deciduous forest (Rivera-Hutinel et al. 2012). This suggests that the addition of new links had the effect of blurring the boundaries of modules.

Interaction evenness, specialization index ( $H_2'$ ) and quantitative modularity ( $Q$ ) were less prone to biases since all these weighted metrics achieved stable values already after *c.* 10 h of sampling, which argue for their use in comparative studies. The stabilization of all these metrics coincided with the stabilization of network size, which suggest that they are depended on network dimensions but little affected by addition of rarer links. In relation to  $H_2'$  specifically, our results agree with Blüthgen, Menzel & Blüthgen (2006) that found this metric to be less affected by sampling incompleteness; however, contrary to what they suggested, it is possible that this index is also affected by network size. The influence of network size on  $H_2'$  could explain the slight overestimation found by simulations of this index value on poor sampled networks (Fründ, McCann & Williams 2015). In regard to modularity, however, despite the stability of its value, modules identification under small sampling were sensitive to addition of species and new links and interactions (Table S9), suggesting that modules conformation cannot be safely assessed in poorly sampled networks. Also, the high variation of the detected  $Q$  in the randomizations (grey lines in Fig. 1) under small sampling effort suggests that this metric is influenced by the identity of both species and links that are added. In sum, our empirical findings suggest that quantitative metrics are more robust to sampling effort than binary ones, as suggested previously for some food webs metrics (Bersier, Banašek-Richter & Cattin 2002; Banašek-Richter, Cattin & Bersier 2004) and simulated bipartite networks (Fründ, McCann & Williams 2015). Additionally, our results point out that network dimensions have some effect on most of the estimates metric values, since higher variation was prevalent under small sampling effort when network size was variable.

When considering metrics significance, small sampling leads to wrong conclusions depending on the metrics considered, as NODF and  $H_2'$  became significant and wNODF became non-significant after *c.* 5 h of sampling. In short, both findings regarding the variation in metrics value and significance suggest that description of network patterns is susceptible to bias by missing species, links and visits and an accumulated sampling effort of *c.* 10–15 h for each plant species seems necessary for our system consisting of plant–hummingbird interactions.

#### ARE PROCESSES STRUCTURING NETWORKS BIASED AT LOW SAMPLING EFFORT?

As many other tropical plant–pollinator communities, in Santa Virgínia, there was a strong variation in hummingbird bill and corolla length; plants presented short flowering periods and species bloomed sequentially

(Tables S3–S6). All these factors together seem to be the primary mechanisms determining interactions between plants and hummingbirds, no matter how abundant species are (Wolf, Stiles & Hainsworth 1976; Maruyama *et al.* 2014; Vizentin-Bugoni, Maruyama & Sazima 2014). This conclusion is already apparent with only  $\geq 3$  h of sampling, when morphological matching and temporal overlap appeared as the best predictors of the frequency of interactions. Interestingly, this happened before the stabilization of network dimensions (i.e. species richness), suggesting that when forbidden links are the most important predictors of frequencies of interactions, their importance is revealed even in incompletely sampled networks (here, with only 6% of the total sampling effort). The only previous study that also evaluated the influence of sampling on the inferences of processes determining interaction frequencies used a temperate plant–pollinator network (Olito & Fox 2015). They consistently found the phenological overlap as the best predictor of pairwise interactions. Despite of the more than twofold sampling completeness of our tropical subnetwork (i.e. 82% here vs. 37% of the estimated links identified in Olito & Fox 2015), both studies agree that the identification of the major processes structuring ecological networks may be possible even under small sampling effort.

Although we present an analysis for a single ecological subnetwork within a larger network, our findings may also apply to other types of mutualistic systems in the tropics, such as plant–frugivorous bird and plant–hawkmoth subnetworks, in which there is high morphological variation among species in traits important for determining interactions (e.g. sizes of mouth apparatus, fruits and flowers) and fruiting or flowering periods tend to be sequentially organized among species (Moermond & Denslow 1985; Cocucci, Moré & Sérsic 2009; Amorim, Wyatt & Sazima 2014; González-Castro *et al.* 2015). It may also apply to antagonistic networks and entire food webs when traits are dominant structuring factors (Eklöf *et al.* 2013). On the other hand, larger networks such as those including multiple groups of pollinators, for example bees, flies, moths and birds (Danieli-Silva *et al.* 2012), or multiple taxonomic groups of seed dispersers, for example birds, large and small mammals and fishes (Donatti *et al.* 2011), could each be influenced differently by sampling effort. This could be generated, for instance, by distinct rates of species and interactions accumulation among groups within the same network due differences in physiology, morphology and behaviour, for example. Thus, different taxonomic groups may demand distinct sampling efforts in order to satisfactorily describe the network structure. Geographically, we may also expect differences in the influence of sampling effort as species-rich communities in the tropics may be more prone to bias than temperate counterparts.

In sum, detected network patterns may be biased depending on the sampling effort employed to build the network and the metrics considered; evidence suggests

that binary metrics are more influenced by sampling than quantitative metrics. However, for more specialized networks in which traits have a strong role in determining interactions, processes structuring the networks may be identified even under small sampling effort. The latter may even be the case for more generalized systems, such as temperate plant–insect pollination networks (Olito & Fox 2015). We argue that ecologists should better investigate the extent to which sampling artefacts bias network patterns, especially for larger networks in diverse tropical ecosystems. To circumvent sampling bias, we suggest using metrics little influenced by sampling (i.e. certain quantitative metrics) and gather more intensively sampled networks. This may include the use of complementary sampling methods, for example focal observation, spot censuses and pollen loads analysis, which may be valuable ways to improve link detection.

## Acknowledgements

We thank Instituto Florestal and Santa Virgínia Field Station staff for permission and facilities to collect the data. We also thank Carsten Dormann for suggestions on modularity calculations; Pedro Jordano and an anonymous reviewer for valuable suggestions; also Pedro J. Bergamo, Marcela Firens, Rafael Gomes, Cintia Kameyama, Raquel F. Monteiro and Marina Wolowski for assistance on plant identification; and Vinicius A. G. Bastazini for assistance with Fig. 1. Financial support was provided by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) through a PDSE scholarship (processo: 99999.008012/2014-08) and a Ph.D scholarship to JVB, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) through a Ph.D scholarship to PKM and a grant to MS; finally, JVB and BD thank the Danish National Research Foundation for its support of the Center for Macroecology, Evolution and Climate.

## Data accessibility

Uploaded as online supporting information.

## References

- Ackerman, J.D., Rodriguez-Robles, J.A. & Melendez, E.J. (1994) A meager nectar offering by an epiphytic orchid is better than nothing. *Biotropica*, **26**, 44–49.
- Almeida-Neto, M. & Ulrich, W. (2011) A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling and Software*, **26**, 173–178.
- Almeida-Neto, M., Guimarães, P., Guimarães, P.R., Loyola, R.D. & Ulrich, W. (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239.
- Amorim, F.W., Wyatt, G.E. & Sazima, M. (2014) Low abundance of long-tongued pollinators leads to pollen limitation in four specialized hawkmoth-pollinated plants in the Atlantic Rain forest, Brazil. *Naturwissenschaften*, **101**, 893–905.
- Banašek-Richter, C., Cattin, M.-F. & Bersier, L.-F. (2004) Sampling effects and the robustness of quantitative and qualitative descriptors of food webs. *Journal of Theoretical Biology*, **226**, 23–32.
- Bascompte, J. (2009) Mutualistic networks. *Frontiers in Ecology and Environment*, **7**, 429–436.
- Bascompte, J., Jordano, P. & Olesen, J.M. (2006) Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, **312**, 431–433.
- Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003) The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9383–9387.

- Bersier, L.-F., Banašek-Richter, C. & Cattin, M.F. (2002) Quantitative descriptors of food web matrices. *Ecology*, **83**, 2394–2407.
- Blüthgen, N., Menzel, F. & Blüthgen, N. (2006) Measuring specialization in species interaction networks. *BMC Ecology*, **6**, 9.
- Blüthgen, N., Menzel, F., Hovestadt, T., Viala, B. & Blüthgen, N. (2007) Specialization, constrains and conflicting interests in mutualistic networks. *Current Biology*, **17**, 341–346.
- Blüthgen, N., Fründ, J., Vázquez, D.P. & Menzel, F. (2008) What do interaction network metrics tell us about specialization and biological traits? *Ecology*, **89**, 3387–3399.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference*, 2nd edn. Springer-Verlag, New York, NY, USA.
- Chacoff, N.P., Vázquez, D.P., Lomáscolo, S.B., Stevani, E.L., Dorado, J. & Padrón, B. (2012) Evaluating sampling completeness in a desert plant–pollinator network. *Journal of Animal Ecology*, **81**, 190–200.
- Chagnon, P.-L. (2015) Characterizing topology of ecological networks along gradients: the limits of metrics' standardization. *Ecological Complexity*, **22**, 36–39.
- Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, **11**, 265–270.
- Cocucci, A.A., Moré, M. & Sérsic, A.N. (2009) Restricciones mecánicas en las interacciones planta–polinizador: estudio de casos en plantas polinizadas por esfíngidos. *Interacciones planta–animal y la conservación de la biodiversidad* (eds R. Medel, R. Zamora, M. Aizen & R. Dirzo), pp. 43–59. CYTED, Madrid, Spain.
- Dalsgaard, B., Trøjelsgaard, K., Martín González, A.M., Nogués-Bravo, D., Ollerton, J., Petanidou, T. et al. (2013) Historical climate-change influences modularity and nestedness of pollination networks. *Ecography*, **36**, 1331–1340.
- Danieli-Silva, A., de Souza, J.M.T., Donatti, A.J., Campos, R.P., Vicente-Silva, J., Freitas, L. et al. (2012) Do pollination syndromes cause modularity and predict interactions in a pollination network in a tropical high-altitude grassland? *Oikos*, **121**, 35–43.
- Donatti, C.I., Guimarães, P.R., Galetti, M., Pizo, M.A., Marquitti, F.M.D. & Dirzo, R. (2011) Analysis of a hyper-diverse seed dispersal network: modularity and underlying mechanisms. *Ecology Letters*, **14**, 773–781.
- Dormann, C.F., Gruber, B. & Fründ, J. (2008) Introducing the bipartite package: analysing ecological networks. *R News*, **8**, 8–11.
- Dormann, C.F. & Strauss, T. (2014) A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, **5**, 90–98.
- Dormann, C.F., Fründ, J., Blüthgen, N. & Gruber, B. (2009) Indices, graphs and null models: analyzing bipartite ecological networks. *Open Ecology Journal*, **2**, 7–24.
- Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P. et al. (2013) The dimensionality of ecological networks. *Ecology Letters*, **16**, 577–583.
- Fründ, J., McCann, K.S. & Williams, N.M. (2015) Sampling bias is a challenge for quantifying specialization and network structure: lessons from a quantitative niche model. *Oikos*, doi:10.1111/oik.02256.
- Gibson, R.H., Knott, B., Eberlein, T. & Memmott, J. (2011) Sampling method influences the structure of plant–pollinator networks. *Oikos*, **120**, 822–831.
- Goldwasser, L. & Roughgarden, J. (1997) Sampling effects and the estimation of food-web properties. *Ecology*, **78**, 41–54.
- González-Castro, A., Yang, S., Nogales, M. & Carlo, T.A. (2015) Relative importance of phenotypic trait matching and species' abundances in determining plant–avian seed dispersal interactions in a small insular community. *AoB Plants*, in press, doi:10.1093/aobpla/plv017.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.
- Grant, V. & Temeles, E.J. (1992) Foraging ability of rufous hummingbirds on hummingbird flowers and hawkmoth flowers. *Proceedings of the National Academy of Sciences*, **89**, 9400–9404.
- Guimarães, P.R. Jr & Guimarães, P. (2006) Improving the analyses of nestedness for large sets of matrices. *Environmental Modelling and Software*, **21**, 1512–1513.
- Hegland, S.J., Dunne, J., Nielsen, A. & Memmott, J. (2010) How to monitor ecological communities cost-efficiently: the example of plant–pollinator networks. *Biological Conservation*, **143**, 2092–2101.
- Jordano, P. (1987) Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. *American Naturalist*, **129**, 657–677.
- Jordano, P., Bascompte, J. & Olesen, J.M. (2003) Invariant properties in coevolutionary networks of plant–animal interactions. *Ecology Letters*, **6**, 69–81.
- Kissling, W.D. & Schleuning, M. (2015) Multispecies interactions across trophic levels at macroscales: retrospective and future directions. *Ecography*, **38**, 346–357.
- Martín González, A.M., Dalsgaard, B., Nogués-Bravo, D., Graham, C.H., Schleuning, M., Maruyama, P.K. et al. (2015) The macroecology of phylogenetically structured hummingbird–plant networks. *Global Ecology and Biogeography*, **24**, 1212–1224.
- Martinez, N.D., Hawkins, B.A., Dawah, H.A. & Feifarek, B.P. (1999) Effects of sampling effort on characterization of food-web structure. *Ecology*, **80**, 1044–1055.
- Maruyama, P.K., Vizentin-Bugoni, J., Oliveira, G.M., Oliveira, P.E. & Dalsgaard, B. (2014) Morphological and spatio-temporal mismatches shape a Neotropical savanna plant–hummingbird network. *Biotropica*, **46**, 740–747.
- Maruyama, P.K., Vizentin-Bugoni, J., Dalsgaard, B., Sazima, I. & Sazima, M. (2015) Nectar robbery by a hermit hummingbird: association to floral phenotype and its influence on flowers and network structure. *Oecologia*, **178**, 783–793.
- Moermond, T.C. & Denslow, J.S. (1985) Neotropical avian frugivores: patterns of behavior, morphology, and nutrition, with consequences for fruit selection. *Ornithological Monographs*, **36**, 865–897.
- Nielsen, A. & Bascompte, J. (2007) Ecological networks, nestedness and sampling effort. *Journal of Ecology*, **95**, 1134–1141.
- Olesen, J.M., Bascompte, J., Dupont, Y.L. & Jordano, P. (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19891–19896.
- Olesen, J.M. & Jordano, P. (2002) Geographical patterns in plant–pollinator mutualistic networks. *Ecology*, **83**, 2416–2424.
- Olito, C. & Fox, J.W. (2015) Species traits and abundances predict metrics of plant–pollinator network structure, but not pairwise interactions. *Oikos*, **124**, 428–436.
- Ollerton, J. & Cranmer, L. (2002) Latitudinal trends in plant–pollinator interactions: are tropical plants more specialised? *Oikos*, **98**, 340–350.
- Paton, D.C. & Collins, B.G. (1989) Bills and tongues of nectar-feeding birds: a review of morphology, function and performance, with intercontinental comparisons. *Australian Journal of Ecology*, **14**, 473–506.
- Petanidou, T., Kallimanis, A.S., Tzanopoulos, J., Sgardelis, S.P. & Pantis, J. (2008) Long-term observation of a pollination network: fluctuation in species and interactions, relative invariance of network structure and implications for estimates of specialization. *Ecology Letters*, **11**, 564–575.
- Renoult, J.P., Blüthgen, N., Binkenstein, J., Weiner, C.N., Werner, M. & Schaefer, H.M. (2015) The relative importance of color signaling for plant generalization in pollination networks. *Oikos*, **124**, 347–354.
- Rivera-Hutinel, A., Bustamante, R.O., Marín, V.H. & Medel, R. (2012) Effects of sampling completeness on the structure of plant–pollinator networks. *Ecology*, **93**, 1593–1603.
- Schleuning, M., Fründ, J. & Garcia, D. (2015) Predicting ecosystem functions from biodiversity and mutualistic networks: an extension of trait-based concepts to plant–animal interactions. *Ecography*, **38**, 380–392.
- Schleuning, M., Ingmann, L., Strauß, R., Fritz, S.A., Dalsgaard, B., Dehling, D.M. et al. (2014) Ecological, historical and evolutionary determinants of modularity in weighted seed-dispersal networks. *Ecology Letters*, **17**, 454–463.
- Stang, M., Klinkhamer, P.G.L. & van der Meijden, E. (2007) Asymmetric specialization and extinction risk in plant–flower visitor webs: a matter of morphology or abundance? *Oecologia*, **151**, 442–453.
- Suarez, R.K. (1992) Hummingbird flight: sustaining the highest mass-specific metabolic rates among vertebrates. *Experientia*, **48**, 565–570.
- Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, **329**, 853–856.
- Tylianakis, J.M., Laliberté, E., Nielsen, A. & Bascompte, J. (2010) Conservation of species interaction networks. *Biological Conservation*, **143**, 2270–2279.
- Ulrich, W. (2009) Ecological interaction networks: prospects and pitfalls. *Ecological Questions*, **11**, 17–25.
- Vázquez, D.P., Chacoff, N.P. & Cagnolo, L. (2009b) Evaluating multiple determinants of the structure of mutualistic networks. *Ecology*, **90**, 2039–2046.
- Vázquez, D.P., Blüthgen, N., Cagnolo, L. & Chacoff, N.P. (2009a) Uniting pattern and process in plant–animal mutualistic networks: a review. *Annals of Botany*, **103**, 1445–1457.

- Vázquez, D.P., Lomáscolo, S.B., Maldonado, M.B., Chacoff, N.P., Dorado, J., Stevani, E.L. *et al.* (2012) The strength of plant–pollinator interactions. *Ecology*, **93**, 719–725.
- Vázquez, D.P., Ramos-Jiliberto, R., Urbani, P. & Valdovinos, F.S. (2015) A conceptual framework for studying the strength of plant–animal mutualistic interactions. *Ecology Letters*, **18**, 385–400.
- Vizentin-Bugoni, J., Maruyama, P.K. & Sazima, M. (2014) Processes entangling interactions in communities: forbidden links are more important than abundance in a hummingbird–plant network. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20132397.
- Wolf, L.L., Stiles, F.G. & Hainsworth, F.R. (1976) Ecological organization of a tropical, highland hummingbird community. *Journal of Animal Ecology*, **45**, 349–379.

Received 2 July 2015; accepted 8 October 2015  
Handling Editor: Daniel Stouffer

## Supporting Information

Additional Supporting Information may be found in the online version of this article.

**Appendix S2.** Figures showing the asymptotic trend of links in a rarefaction, a correlation between hummingbird species abundances and their frequencies of occurrences and an individual-based rarefaction for network metrics.

**Figure S1.** Individual-based rarefaction curve for links with 95% confidence intervals (grey lines). Note the asymptotic tendency from accumulated number of links with the accumulation of visits observed in the network. We ran this analysis in EstimateS 9.1.0 (Colwell 2013) and calculated confidence intervals using unconditional variances as suggested by Colwell *et al.* (2012, doi:10.1093/jpe/rtr044)

**Figure S2.** Spearman correlation between frequency of occurrence during 130 days of field work and number of contacts during counts in transects. Observations were conducted over a 2 year period and include nine hummingbird species from Santa Virgínia Field Station, southeastern Brazil.

**Figure S3.** Network metrics calculated under a simulated sampling effort gradient created by interaction removals (individual based rarefaction-like approach) of a plant–hummingbird network (individual-based rarefaction framework) in SE Brazil. We simulated rarefaction-like sampling reduction by removing successively 10% of interactions creating class of removal from 90% to 10% removals. Network metrics were recalculated (1000 iterations) and their values (gray overlapped lines) and mean (black line) were plotted. Dashed black lines indicate mean and 95% CI. Note that the results do not differ importantly from Figure 1 but metrics were less variable across the gradient of sampling effort using this approach.

**Table S1.** Abundances of 55 hummingbird-pollinated species quantified along 12 000 m of trails in the Atlantic Rainforest at Santa Virgínia Field Station, southeastern Brazil. Number of flowers (or inflorescences, for Asteraceae species) was counted monthly from September 2011 to August 2013 and relative abundances indicate the relative proportion of flowers accounted by each species calculated on the total number of flowers.

**Table S2.** Abundances of nine hummingbird species occurring in the Atlantic Rainforest from September 2011 to August 2013 at Santa Virgínia Field Station, southeastern Brazil. Contacts in transects

indicate the total number of aural and visual contacts with individuals counted monthly across 10 transects (100 m each); frequency of occurrence is the proportion of days in which a species was recorded across 12 000 m of trails percurred and over 130 days of fieldwork; and the relative frequency is the relative proportion accounted by each species calculated on the frequency of occurrences.

**Table S3.** Plant phenology quantified by the monthly presence/absence of flowers of 55 hummingbird-pollinated species from September 2011 to August 2013 in the 12 000 m of trails in the Atlantic Rainforests at Santa Virgínia Field Station, southeastern Brazil. Species acronym according Table A1. In total, 130 days of sampling were spread along the years.

**Table S4.** Hummingbird phenology indicated by the monthly presence/absence of flowers from September 2011 to August 2013 in the 12 000 m of trails in the Atlantic Rainforests at Santa Virgínia Field Station, southeastern Brazil. Species acronym according Table A2.

**Table S5.** Minimum corolla depth in 55 hummingbird-pollinated species in Atlantic Rainforests from September 2011 to August 2013 in the 12 000 m of trails in the Atlantic Rainforest at Santa Virgínia Field Station, southeastern Brazil. Corolla depth was measured as the internal distance from the base of nectar chamber to the distal portion of the flower (*i.e.* effective corolla length, *sensu* Wolf, Stiles & Hainsworth 1976), which represents the minimum mouth apparatus length needed to a hummingbird access the nectar legitimately. Species acronym according Table A1.

**Table S6.** Bill length (exposed culmen), estimated tongue extension and bill+tongue estimation of nine hummingbird species in the Atlantic Rainforest at Santa Virgínia Field Station, southeastern Brazil. \*Tongue extension estimated based on *Selasphorus rufus* measures, which is around 80% of the bill length (Grant & Temeles 1992). Species acronym according to Table S2.

**Table S7.** Geographical coordinates of starting and ending points from the ten transects (100 m long) where we counted hummingbirds monthly at Santa Virgínia Field Station, southeastern Brazil.

**Table S8.** Number of parameters used to penalize model complexity in each of eight models described in Table S10. These numbers of parameters were defined according to number of plant and animal species in the matrix and to number of variables included in the models. Smaller networks tend to be easier to predict, so the number of species in the matrices was included in the model's penalization to account for the increasing network size along the sampling effort gradient.

**Table S9.** Network metrics over increasing sampling completeness in the hummingbird-plant network in Santa Virgínia. Between parenthesis are shown values expected by the null model (95% confidence interval) and bold indicates significant differences between observed and expected by the null model.

**Table S10.** AIC values indicating the ability of eight models to predict observed frequency of interaction between pairwise species over increasing sampling effort in the Santa Virgínia hummingbird-

plant network. Models are probability matrices based on species abundance (**A**), phenological overlap (**P**) and morphological matching (**M**) and all possible combination among them. **Null** model is a benchmark model that assumes all interactions have the same probability to occur. Note that **Null** is the best predictive model under very small sampling ( $\geq 3$  h), but after 3 h of cumulative sampling the **PM** model, which includes both phenological overlap and bill-corolla (morphological) matching, had the best ability to predict pairwise interaction. Also note that all models including **A** had the worst fits, even worse than the **Null** model. Because data on tongue extension in hummingbirds is scarce in the literature, we also recreate a morphological model (**MI**) considering the tongue extension as 1/3 of the bill length. However, these models presented minor influence on the results

because they performed similarly to the model **M**; thus we discussed just model **M** in the text.

**Table S11.** Plant–hummingbird network from Atlantic Rainforest at Santa Virgínia Field Station, southeastern Brazil assembled with 50 h of observation to each plant species (see detail in 'Materials and methods' for details). Plant and hummingbirds species names follow Tables S1 and S2, respectively.