





LETTER

Joint species distributions reveal the combined effects of host plants, abiotic factors and species competition as drivers of species abundances in fruit flies

Benoit Facon^{1*}  | Abir Hafsi^{2*} | Maud Charlery de la Masselière² | Stéphane Robin³ | François Massol⁴  | Maxime Dubart⁵ | Julien Chiquet³ | Enric Frago⁶  | Frédéric Chiroleu² | Pierre-François Duyck^{2,7}  | Virginie Ravigné²

¹INRAE, UMR PVBMT, Saint Pierre, France

²CIRAD, UMR PVBMT, Saint Pierre, France

³Laboratoire MMIP – UMR INRA 518/AgroParisTech, Paris, France

⁴Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR 9017 – CHIL – Center for Infection and Immunity of Lille, Lille, France

⁵Univ. Lille, CNRS, UMR 8198 –Evo-Eco-Paleo, Lille, France

⁶CIRAD, UMR CBGP, Montferrier sur Lez, France

⁷IAC, Equipe ARBOREAL, Nouméa, Nouvelle-Calédonie

Correspondence

Benoit Facon, Pôle de protection des plantes – UMR PVBMT – 7, chemin de l'IRAT – 97410 St Pierre, France.
Email: benoit.facon@inrae.fr

Funding information

Agence Nationale de la Recherche, Grant/Award Number: ANR-10-EQPX-20 and ANR-17-CE32-011; European Regional Development Fund, Grant/Award Number: GURDT I2016-1731-0006632

Editor: Tim Coulson

Abstract

The relative importance of ecological factors and species interactions for shaping species distributions is still debated. The realised niches of eight sympatric tephritid fruit flies were inferred from field abundance data using joint species distribution modelling and network inference, on the whole community and separately on three host plant groups. These estimates were then confronted the fundamental niches of seven fly species estimated through laboratory-measured fitnesses on host plants. Species abundances depended on host plants, followed by climatic factors, with a dose of competition between species sharing host plants. The relative importance of these factors mildly changed among the three host plant groups. Despite overlapping fundamental niches, specialists and generalists had almost distinct realised niches, with possible competitive exclusion of generalists by specialists on Cucurbitaceae. They had different assembly rules: Specialists were mainly influenced by their adaptation to host plants, while generalist abundances varied regardless of their fundamental host use.

KEYWORDS

community ecology, niche modelling, performance, phytophagous insects, preference, specialisation

INTRODUCTION

The search for fundamental processes underlying species distributions is among the oldest challenges in ecology (Diamond, 1975; Gotelli & Graves, 1996). Understanding assembly processes could also be crucial for coping with

global changes and habitat loss currently affecting both abiotic conditions and species distributions (Adler & Hille Ris Lambers 2008). Species distributions can be determined by several factors such as environmental filtering, interspecific interactions, regional species pool and dispersal (D'Amen et al., 2018; Jabot et al., 2020; Müller

*Co-first authors

This article has been reviewed and recommended by Peer Community in Ecology: Joaquín Hortal (2021) Understanding the interplay between host-specificity, environmental conditions and competition through the sound application of Joint Species Distribution Models. Peer Community in Ecology, 100080. <https://doi.org/10.24072/pci.ecology.100080>

et al., 2011; Nakadai et al., 2018). Despite decades of research, estimating the relative importance of these processes on species distributions has proven particularly complex (Pollock et al., 2014; Zurell et al., 2018). Most of these processes imprint species distributions in a scale-dependent manner (Meynard et al., 2013). For instance, abiotic factors are generally thought to determine large-scale species ranges, whereas interspecific interactions would influence species distributions at smaller spatial scales (Heikkinen et al., 2007; Thuiller et al., 2015; but see Gotelli et al., 2010 and Araújo & Rozenfeld, 2014).

Phytophagous insects are among the most diverse and abundant groups of terrestrial animals and a major component of ecosystems due to their tight interaction with primary producers and their sometimes important economic impacts and invasive potential (Roy et al., 2015). Knowledge of the main determinants of insect occurrence on particular plant species and their potential to colonise and persist in a given area is however still limited. In particular, the importance of interspecific competition in structuring phytophagous insect communities has been a controversial issue (Kaplan & Denno, 2007). Many experimental studies conclude that interspecific competition plays a primary role (Denno et al., 1995; Kaplan & Denno, 2007), but the consistent absence of negative co-occurrence patterns in natural phytophagous insect communities suggests otherwise (Brazeau & Schamp, 2019; Tack et al., 2009). This apparent discrepancy could result from the regulation of phytophagous insect populations below competitive levels through shared predators or parasites (Hairston et al., 1960). Phytophagous insects would also rarely overexploit their hosts, leaving sufficient plant material for competition with other species to be mild (Kaplan & Denno, 2007). Ecological differences between species could lower the intensity of competition (Stewart et al., 2015). Lastly, some phytophagous arthropods could benefit from previous attack of the host plant by other species (Godinho et al., 2016). The importance of competition relative to ecological conditions in shaping phytophagous insect distributions thus remains an open question and demands appropriate testing (Augustyn et al., 2016; Nakadai et al., 2018).

Whether biotic interactions affect species distributions should be uncovered from proper analysis of patterns of co-occurrence. Species interactions are expected to affect species occurrence, for example, competition should cause checkerboard patterns of occurrence. However, species occurrences also result from common or diverging species dependence on confounding environmental factors. Species that share the same abiotic niche will frequently co-occur without necessarily interacting (Blanchet et al., 2020; Wisz et al., 2013). Conversely, negative co-occurrence patterns may simply result from diverging ecological requirements. As a consequence, estimating the effect of species interactions on species distributions first requires properly

characterising species abundance's responses to environmental variables (Pollock et al., 2014), which is the object of species distribution modeling (SDM) approaches (Elith & Leathwick, 2009). A particular class of SDM approaches, joint species distribution models (JSDM), attempts to infer the relationships between species abundances and environmental variables, explicitly accounting for the interdependence of species distributions using multivariate regression methods (Pollock et al., 2014). In addition to estimates of the effect of habitat filtering on species distributions, these approaches provide residual covariances between species abundances, that is, covariances not explained by environmental factors. Residual covariances result from species interactions and a diversity of other factors, as, for example, missing covariates, so that there is no simple relationship between species interactions and residual covariances (Zurell et al., 2018).

To further track species interactions, a growing body of literature pleads for using independent knowledge of species traits in species distribution modelling (Kraft et al., 2008; Lavorel et al., 1997; Poisot et al., 2015), which is still seldom done. Explicitly comparing estimates of fundamental niche, that is, measures of fitness in controlled conditions and in absence of species interactions, to estimates of realised niche, that is, inferred species abundances' responses to environmental factors, could shed new light on the gap between fundamental and realised niches and the importance of species interactions in shaping species distributions.

In the case of phytophagous insects, an obvious feature of the environment to account for is host plant identity. Host plants can be treated as any environmental cofactor, and their effects on species abundances can be inferred directly from adequate abundance data (Ferrier & Guisan, 2006). Host plants impose a specific challenge because modelling a phytophagous community as a whole relies on the assumption that the interdependence of species abundance does not depend on host plants. But as intraguild interactions mostly occur in/on plant organs, they may be modulated by plant species identity, with possible consequences for species occurrence patterns (Ulrich et al., 2017). Analysing competition patterns on different host plants could therefore allow detecting the role of host plants in shaping species co-occurrences.

Here, we aimed at disentangling the roles of host plant species, abiotic factors and interspecific interactions on the distributions of eight fruit fly species (Tephritidae) occurring in sympatry on a diversity of host plants and in highly variable abiotic conditions. The study system, which comprises four generalist species, three specialists of Cucurbitaceae and one specialist of Solanaceae, presents key advantages to tackle community assembly questions. First, these species occupy a small island in south-western Indian Ocean (Réunion, 2512 km²) where they are considered the main actors in the guild of fruit-eating phytophagous

arthropods (Quilici & Jeuffraut, 2001). Second, the local environment is characterised by important variability in elevation (from 0 to 3000 m), climatic conditions, land use and plant distributions (Duyck et al., 2006a). Observational and experimental studies have suggested that climatic factors could influence local tephritid distributions (Duyck et al., 2004). Climatic factors were even found more influential than host plant diversity in allowing coexistence in an analysis of the distributions of the four generalist species on four host plants (Duyck et al., 2008). Lastly, competition between the eight species has repeatedly been advocated to shape this community. First, host use strategies largely overlap, opening possibilities of competition (Duyck et al., 2008; Quilici & Jeuffraut, 2001). Second, the arrival of one generalist species on the island has constrained the host ranges of some resident species, without complete exclusion (Charlery de la Masselière et al., 2017a). Third, larval competition experiments involving a subset of plant species and abiotic conditions have evidenced hierarchical competition interactions among the generalist species (Duyck et al., 2006b).

Here, we confronted a long-term field dataset describing abundances of the eight fly species on 21 host plants with laboratory measures of fundamental host use obtained for seven of the fly species on the same plants. We first modelled joint species distributions using Poisson-lognormal (PLN) modelling (Chiquet et al. 2019) and conducted model selection among various combinations of host plant species and ecological covariates (representing temperature, rainfall, elevation, land use and date). Residual correlations estimated under the selected model were further dissected to identify significant traces of unexplained co-occurrences. Second, we assessed whether knowledge on fundamental host use was sufficient to explain field species abundances by accounting for host plant species either directly or through estimates of female preference and larval performance in laboratory conditions. Finally, we tested for a potential dependence of community structuring factors on host plants by replicating the analyses on three subsets of plants: Cucurbitaceae, Solanaceae and the other plants.

MATERIAL AND METHODS

Species abundance table

Field campaigns were conducted over a period of 18 years (1991–2009) to identify potential host plants for Tephritidae on the whole island including orchards, gardens and natural areas. These surveys were assembled in a previous study (Charlery de la Masselière et al., 2017a) and used here as species abundance table. Each observation corresponds to the numbers of individual

flies of each species recorded from a set of fruits sampled in one location at a specific date. For each sample, the collected fruits were counted and weighted, before being stored until adult fly emergence. To avoid keeping samples that could have suffered from a transportation or storage issue, only samples with at least one individual fly were kept. Among these, we further selected samples with GPS coordinates belonging to one of the 21 host plants characterised in the laboratory (see below). Of the 12,872 initial samples, we therefore kept 4918 samples and a total of 97,351 individual flies. Samples covered 104 field sessions all year-round over the study period (Tables S1 and S2) and originated from 380 sites well distributed over the island (Figure 1a). Additional details on sample collection can be found in Appendix S1.

Ecological covariables

Sample GPS coordinates allowed retrieving ecological and climatic characteristics from GIS information available on the CIRAD Agricultural Web Atlas for Research (AWARE, <https://aware.cirad.fr>). Each sample was associated with a month, year, land use category, elevation, three pluviometry descriptors (minimal rainfall in the 20% most humid years, minimal rainfall in the 20% driest years, median annual rainfall between 1986 and 2016) and three temperature variables (minimal, mean, maximal annual temperature between 1987 and 2017) (Figure 1a, Appendix S1). To account for correlations between variables, a factorial analysis of mixed data (FAMD) was conducted on all 10 variables using *FactoMineR* (Lê et al., 2008), producing 10 uncorrelated dimensions subsequently used as ecological covariates (Figure S2 for FAMD details).

Species traits

For all species but *Dacus ciliatus*, fundamental host use, that is, fly fitness in optimal abiotic conditions and without antagonists, was characterised using four traits describing larval performances and female preferences on 21 plant species. Female preferences were the numbers of eggs laid during 24 h on each fruit species in the ‘no-choice’ experiment of de la Masselière et al. (2017b). Larval performances (survival probability s , development time T and pupal weight w) were obtained from Hafsi et al., (2016) for 17 plant species, and in the current study for *Coffea arabica*, *Solanum mauritanium*, *Syzygium jambos* and *Syzygium samarangense*, using the same methods. Larval performance traits were combined into a single performance trait using the formula:

$$Perf = s \times w / T$$

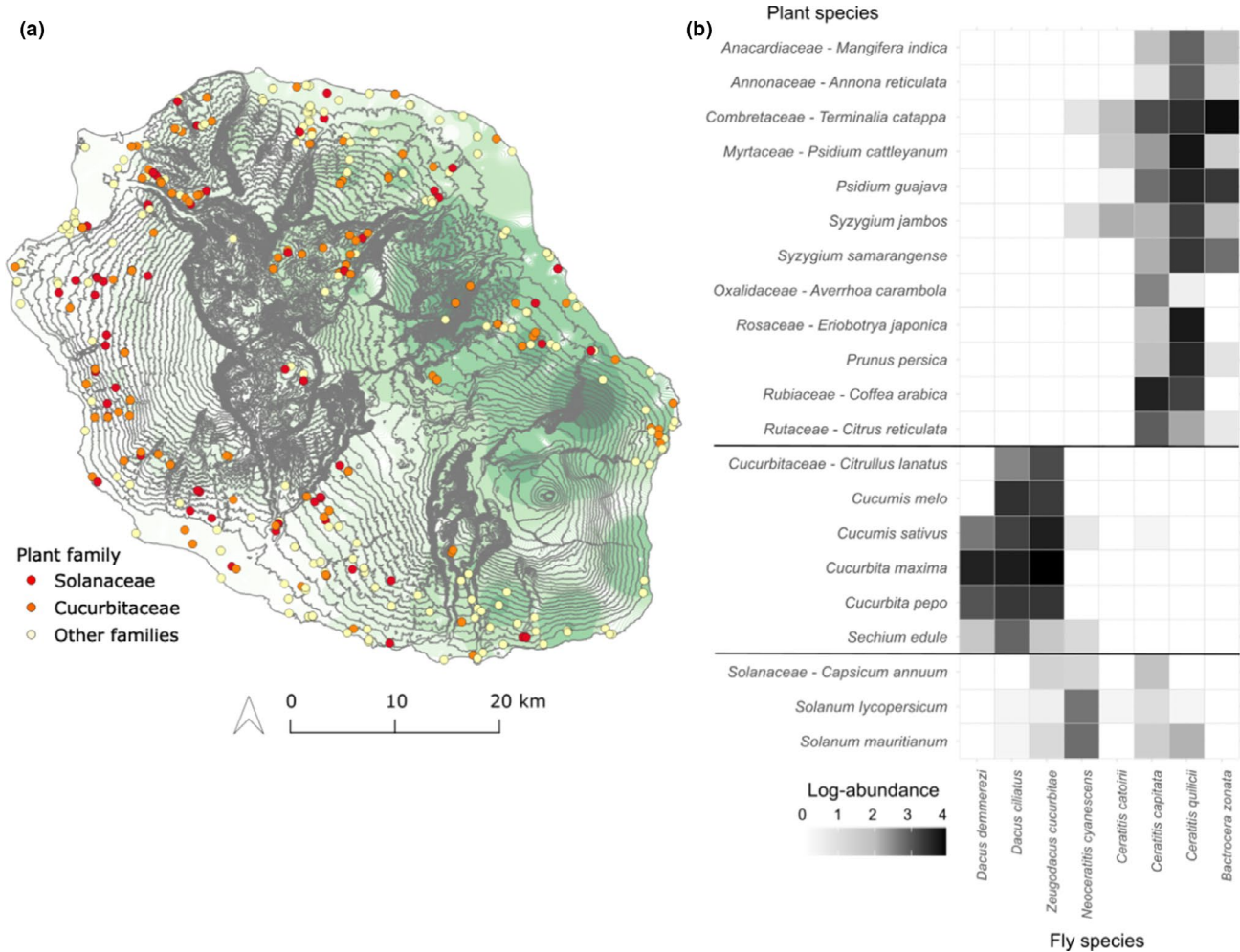


FIGURE 1 Characteristics of the species abundance dataset. (a) Sampling sites in Réunion. Dot colors refer to the family of the sampled plant (Solanaceae $n = 259$, Cucurbitaceae $n = 2347$, other families $n = 2285$ samples). Elevation, represented by grey isoclines, ranges from 0 along the coast to >3000 m, in the centre of the island, and strongly correlates with annual temperature. Pluviometry (here median rainfall over 1986–2016) is represented by green isohyets. The island is separated into two contrasted rainfall regimes: very humid all year-round in the east and drier, especially during winter, in the west. (b) Fly species abundances on the 21 studied plant species

Preference and performance traits were log-transformed before being included as covariates in statistical models.

Statistical analysis

Datasets

Because determinants of species distributions may depend on host plant identity, we replicated all analyses on the 21-plant dataset and on three sub-datasets: (1) Cucurbitaceae only (3 fly \times 6 plant species, 2347 samples), (2) Solanaceae only (3 fly \times 3 plant species, 259 samples) and (3) other plant families (4 fly \times 12 plant species, 2285 samples).

Statistical modelling

Joint variations in fly species abundances were modelled using Poisson-lognormal (PLN) models with the

$PLN()$ function in the *PLNmodels* R package (Chiquet et al., 2018, 2019). A PLN model is a multivariate mixed generalised linear model, where each species count is assumed to arise from a Poisson distribution with a parameter resulting from fixed effects of covariates and a random lognormal effect. Random effects associated with all species observed in a sample are jointly sampled from a multivariate lognormal distribution. The residual variance–covariance of the multivariate distribution reveals species abundance covariations still unexplained after controlling for confounding environmental covariates and differences in sampling efforts. Samples differed in fruit number and weight, inducing uncontrolled variation in sampling effort among samples, potentially leading to spurious associations between species abundances. Consequently, the (log-transformed) fruit weight of each sample was added as an offset to every tested model (fruit number was also tested, yielding identical conclusions, not shown). Model diagnostics were conducted using the

R package *DHARMA* (Hartig, 2020). Covariations between species abundances unexplained by covariates were further investigated using the *PLNnetwork()* function, which adjusts the considered model under a sparsity constraint on the inverse of the variance-covariance matrix, that is, constraining the number of edges in the resulting estimated network. The stability of the resulting species associations was estimated as their selection frequency in bootstrap subsamples of the StARS model selection procedure (range 0–1; Liu et al., 2010; Appendix S3).

Model selection design

The importance of plant species identity, ecological covariates and species interactions to explain species abundances was approached by model selection using the extended BIC criteria (Chen & Chen, 2008). First, we compared models (listed in Table 1) including either no covariate (Model 1–0), plant species as a cofactor (Model 1–1), ecological variables (all 10 FAMD dimensions

described above, Model 1–3) or both plant species and ecological variables (Model 1–5). Second, we evaluated the importance of residual species abundance covariations. *PLNmodels* enables fitting models where the residual variance–covariance matrix is constrained to be diagonal. Such models assume no possible interaction between species. We therefore compared all models with their diagonal counterpart (Models 1–2, 1–4 and 1–6). Third, we estimated how well knowledge of the fundamental niche (preference and performance) explained field abundances. *PLNmodels* does not yet allow accounting for species traits (covariates describe samples but not species). To cope with this limitation, we considered that assuming that species distribute according to their fundamental niche implies negligible species interactions, with species distributing independently from one another. Using this assumption, we fitted models separately for each fly species, obtained their likelihoods and numbers of parameters and computed the BIC of the seven-species dataset as

$BIC = k \ln(n) - 2 \ln(L)$, where k and L are the sums of the numbers of parameters and likelihoods over the seven

TABLE 1 Model selection on the 21-plant dataset ($n = 4918$)

Models	Covariates	Residual matrix	K	L	BIC	Δ_{BIC}
A) Model Set 1 (21 plants \times 8 flies)						
Model 1–5	Plant + Eco	Full	284	–27,664.3	57,742.8	0.0
Model 1–6	Plant + Eco	Diagonal	256	–27,997.8	58,171.7	428.9
Model 1–2	Plant	Diagonal	176	–28,608.9	58,713.9	971.1
Model 1–1	Plant	Full	204	–28,784.0	59,302.1	1559.2
Model 1–3	Eco	Full	124	–35,888.8	72,831.6	15,088.8
Model 1–4	Eco	Diagonal	96	–36,598.4	74,012.9	16,270.1
Model 1–0	None	Full	44	–37,228.1	74,830.3	17,087.4
B) Model Set 2 (21 plants \times 7 flies)						
Model 2–6	Plant + Eco	Diagonal	224	–20,753.9	43,374.1	0.0
Model 2–5	Plant + Eco	Full	245	–20,723.1	43,487.5	113.4
Model 2–2	Plant	Diagonal	154	–21,563.1	44,409.4	1035.3
Model 2–1	Plant	Full	175	–21,504.7	44,467.5	1093.4
Model 2–12	Preference + Performance + Eco	Diagonal	98	–23,924.4	48,665.3	5291.2
Model 2–10	Performance + Eco	Diagonal	91	–24,387.4	49,533.0	6158.9
Model 2–11	Preference + Performance	Diagonal	28	–25,689.9	51,613.0	8238.9
Model 2–8	Preference + Eco	Diagonal	91	–25,516.9	51,792.1	8418.0
Model 2–9	Performance	Diagonal	21	–26,443.3	53,061.6	9687.5
Model 2–4	Eco	Diagonal	84	–27,082.1	54,864.0	11,489.9
Model 2–7	Preference	Diagonal	21	–27,628.9	55,432.8	12,058.7
Model 2–3	Eco	Full	105	–27,321.5	55,517.8	12,143.7
Model 2–0	None	Full	35	–31,060.1	62,411.8	19,037.7

Models are ranked by increasing BIC (from best to worst). K is the number of parameters. L is the log-likelihood. Δ_{BIC} is the BIC difference between any focal model and the best one.

single-species datasets and n is the number of samples. Following this principle, we built models where the host plant cofactor was replaced by either preference (Models 2–7 and 2–8), or performance (Models 2–9 and 2–10), or both preference and performance (Models 2–11 and 2–12). In addition, all previous models were re-evaluated on the dataset excluding the species lacking fundamental host use estimates (Models 2–0 to 2–6).

RESULTS

All host plants

The variance–covariance structure of the complete dataset was first inferred by fitting a PLN model without any covariate (Model 1–0). The obtained residual variance-covariance matrix (Figure 2a) revealed a sharp distinction between (1) the four generalists (*Bactrocera zonata*, *Ceratitis quilicii*, *Ceratitis capitata* and *Ceratitis catoirii*), (2) the three specialists of Cucurbitaceae (*D. ciliatus*, *Dacus demmerezi* and *Zeugodacus cucurbitae*) and the specialist of Solanaceae (*Neoceratitis cyanescens*). *N. cyanescens* abundances showed very low covariances with other species. The two other groups showed positive within-group covariances and negative among-group covariances. This variance–covariance structure suggested strong separation of the realised niches of the three groups, likely mediated by host plants.

Considering model selection between models without species traits, that is, models including combinations of plant species and ecological covariates (Table 1), model ranking was equivalent on the eight-species (Models 1–0 to 1–6) and seven-species (Models 2–0 to 2–6) datasets, except that diagonal models tended to be slightly better than their full-matrix counterparts on the latter. This result was further shown robust to sample bootstrapping (Appendix S3).

For both datasets, the selected model included both ecological variables and host plant species as covariates (Models 1–5 and 2–6), which strongly improved the BIC ($\Delta\text{BIC} = 17,087.4$ and $18,935.7$ respectively) as compared with the basic models 1–0 and 2–0. Models with host plant species ranked close to the selected model with a moderately inflated BIC relative to the selected model (Model 1–2: $\Delta\text{BIC} = 971.1$, Model 2–2: $\Delta\text{BIC} = 933.2$) and an important reduction in residual covariances relative to the basic model (compare Figure 2a and c). Including ecological variables alone greatly increased BIC (Model 1–3: $\Delta\text{BIC} = 15,088.8$, Model 2–4: $\Delta\text{BIC} = 11,387.9$) and caused a mild reduction of residual covariances (compare Figure 2a and b).

The selected model showed a good fit, with a tendency to overestimate low abundance values (Figure S3). As expected, model diagnostics revealed an excess of zeros and some over dispersion (Appendix S3) which PLN modeling is robust to (Chiquet et al. 2019). No temporal or spatial autocorrelation was detected in model residuals. Regression coefficients relative to host plants, that is, the response of species abundances to host plants, had

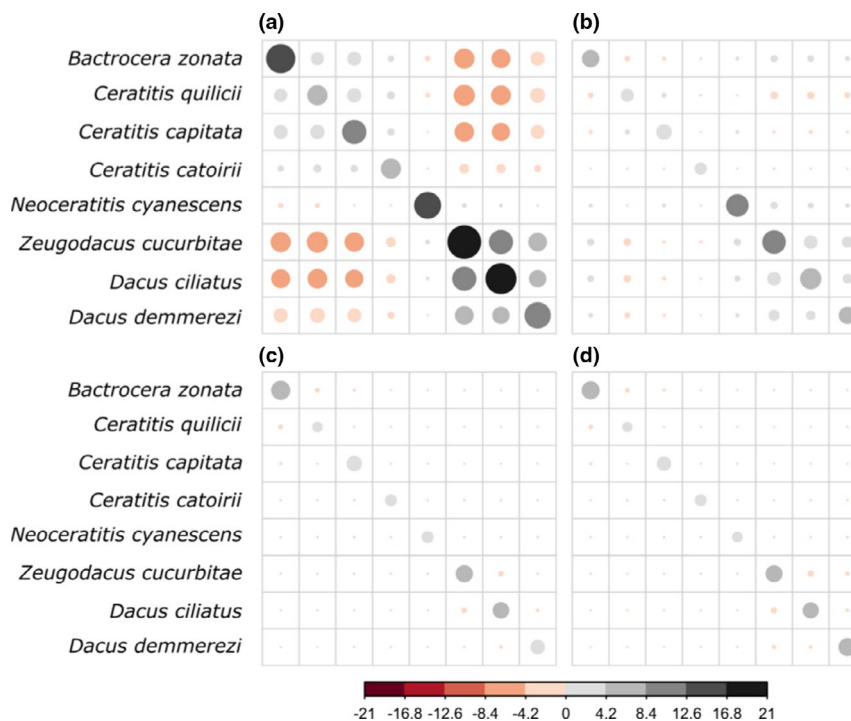


FIGURE 2 Residual variance–covariance matrices obtained after PLN model fitting on species abundances. (a) Without any covariate (Model 1–0). (b) With ecological covariates (Model 1–3). (c) With plant species as a cofactor (Model 1–1). (d) With both plant species and ecological covariates (Model 1–5)

a strongly bimodal distribution (Figure 3a, right panel). For each fly, some plant species had very low coefficients (<-80), meaning a negligible effect of the corresponding host plant on fly abundances (approx. $\exp(-80) \approx 1E-35$). These plants were considered as non-hosts. The other plants had much stronger coefficients (>-14) and were interpreted as hosts. Overall, this realised host range was narrower than the laboratory-measured fundamental host range (Figure 3a, left panel), particularly for generalists. Only 10 fly–plant associations (out of 147) showed a reverse pattern (zero laboratory-measured fitness and yet strong regression coefficient), suggesting marginal difficulties with measuring fitness in laboratory conditions. Among plants inferred as hosts, coefficients correlated positively with fly laboratory-measured fitness for specialists but not for generalists (Figure 3b).

The coefficients relative to the first two axes of the FAMD on ecological covariates could be interpreted as responses of fly abundances to rainfall, temperature and elevation (Figure 4). Rainfall had a low effect on

the abundances of *B.zonata*, *C.capitata*, *C.quilicii* and *D.demmerezi*. *C.catoirii*, *N.cyanescens* and *Z.cucurbitae* showed a propensity towards warm high-rainfall areas, while *D.ciliatus* seemed to prefer colder drier climates. *C.quilicii* and *C.catoirii* were not much affected by temperature. *B.zonata*, *C.capitata*, *D.ciliatus* and *Z.cucurbitae* should thrive in low-elevation warm climates. *D.demmerezi* and to a lesser extent *N.cyanescens* seemed to prefer colder higher-elevation climates.

The selected model had full residual variance–covariance matrix on the eight-species dataset and diagonal residual matrix on the seven-species dataset. On the eight-species dataset, its diagonal version ranked second ($\Delta BIC = 428.9$). Congruently, the residual matrix inferred under the selected model (Figure 2d) had very low and always negative covariance values for all species pairs. This suggested possible, though weak, antagonist interactions. The largest residual correlations were observed between the three specialists *D.demmerezi*, *D.ciliatus* and *Z.cucurbitae* (residual correlations ranging from

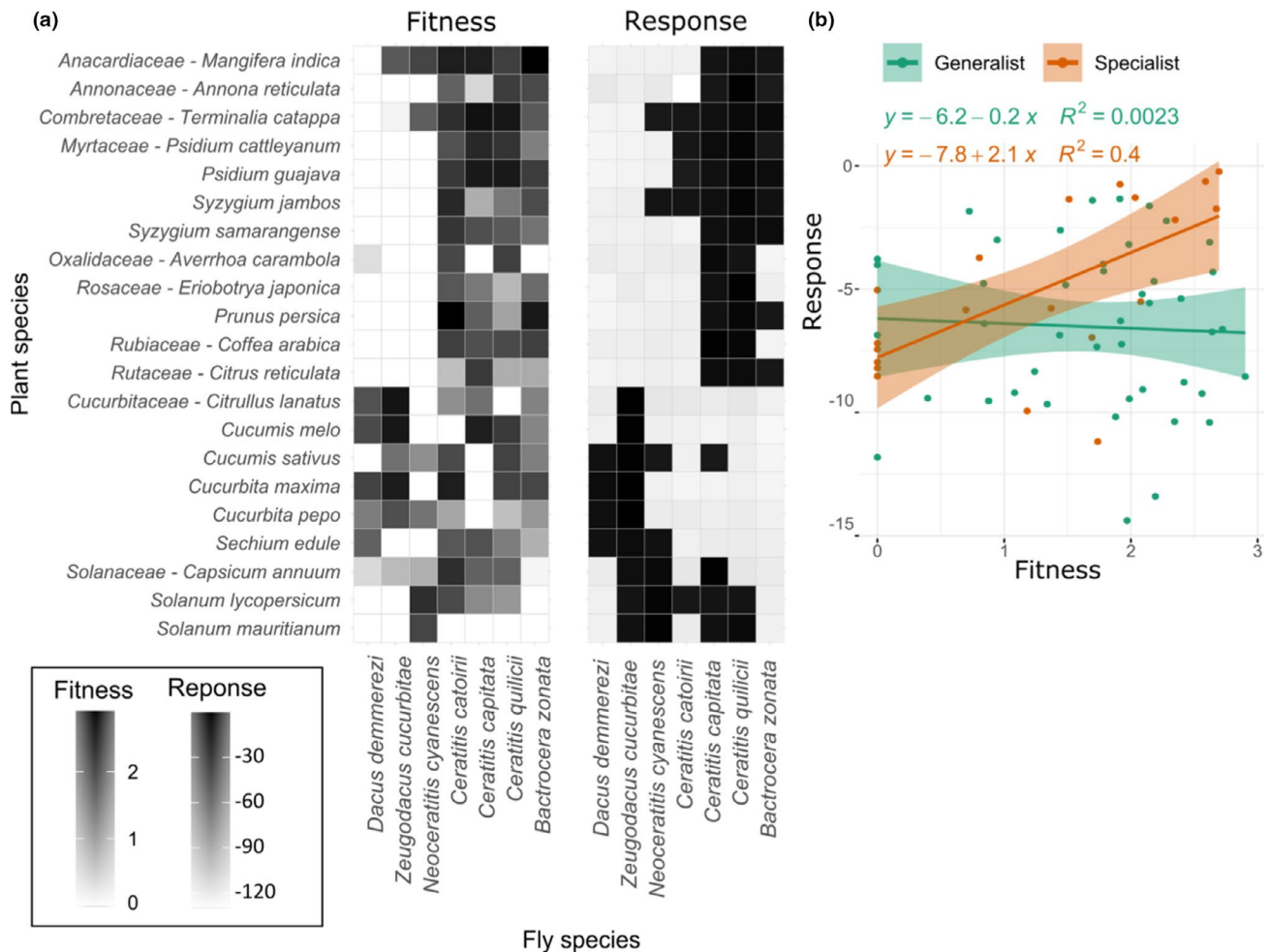


FIGURE 3 Comparison of fly species' fundamental and realised host use. (a) Fundamental host use as measured in the laboratory (left panel—fitness is the logarithm of the product preference and performance) and realised host use as inferred from regression coefficients relative to host plants in the fullest model (right panel—response is obtained from Model 2–5 on the seven-species dataset). (b) Relationship between inferred responses to host plants and laboratory-measured fitness for specialists in orange and generalists in green on hosts detected as such in the field. Lines represent linear regressions with slope 95% confidence intervals as shadowed areas

−0.031 to −0.112) and between the generalists *B. zonata*, *C. quilicii* and *C. capitata* (residual correlations from −0.025 to −0.090). Network inference, applied to the selected model, converged to one significant interaction between *D. ciliatus* and *Z. cucurbitae*, with a stability of 0.99 (i.e. detected in 99 out of 100 network inferences on bootstrapped data). These results suggested that at the scale of the whole community, competitive interactions between fly species only weakly affected their joint distributions. Fly species abundances were mainly explained by host plants and ecological covariates.

Considering models accounting for species fundamental niche (Table 1B), the best model with species traits included female preference, larval performance and ecological variables (Model 2–12, $\Delta\text{BIC} = 5189.2$). It ranked intermediate between the model with ecological variables alone and the best model. Such good performance of models with species traits suggests that the fundamental host range of fly species is an important determinant of fly species joint distributions. Whether in combination with ecological covariates or not, models with larval performance were slightly better than models with female preference.

Specific groups of plants

Detailed model comparisons are provided in Appendix S2.

Cucurbitaceae hosted *D. ciliatus*, *D. demmerezi* and *Z. cucurbitae* (Table S5). On this subcommunity,

the selected model was again the one including host plant species, ecological covariates and a full residual variance–covariance matrix. It performed better than its diagonal version. Network inference yielded two significant negative interactions between *D. ciliatus* and *Z. cucurbitae* (stability = 1.0) and between *D. demmerezi* and *Z. cucurbitae* (stability = 0.56). Models including plant species alone did not perform much better than those with ecological covariates alone, congruently with the idea that host plants affect fly distributions more similarly within the Cucurbitaceae family than at the 21-plant dataset scale. The model with both species traits and ecological covariates ranked third just after the diagonal version of the selected model.

Solanaceae hosted *N. cyanescens*, *C. capitata* and *C. quilicii* (Table S6). For these, the selected model included host plant species, ecological covariates and a diagonal residual matrix, and congruently, network inference yielded no significant residual covariance. Models with species traits were almost as good as their equivalent with host plant as a cofactor. Female preference and larval performance performed equally well and almost as good as both traits together, suggesting important correlation between the traits.

The last dataset considered all families other than Cucurbitaceae or Solanaceae with *B. zonata*, *C. capitata*, *C. catoirii* and *C. quilicii* (Table S7). The selected model included host plant species, ecological covariates and a full residual variance–covariance matrix. It performed slightly better than its diagonal version, and network inference yielded one significant interaction between *B. zonata* and

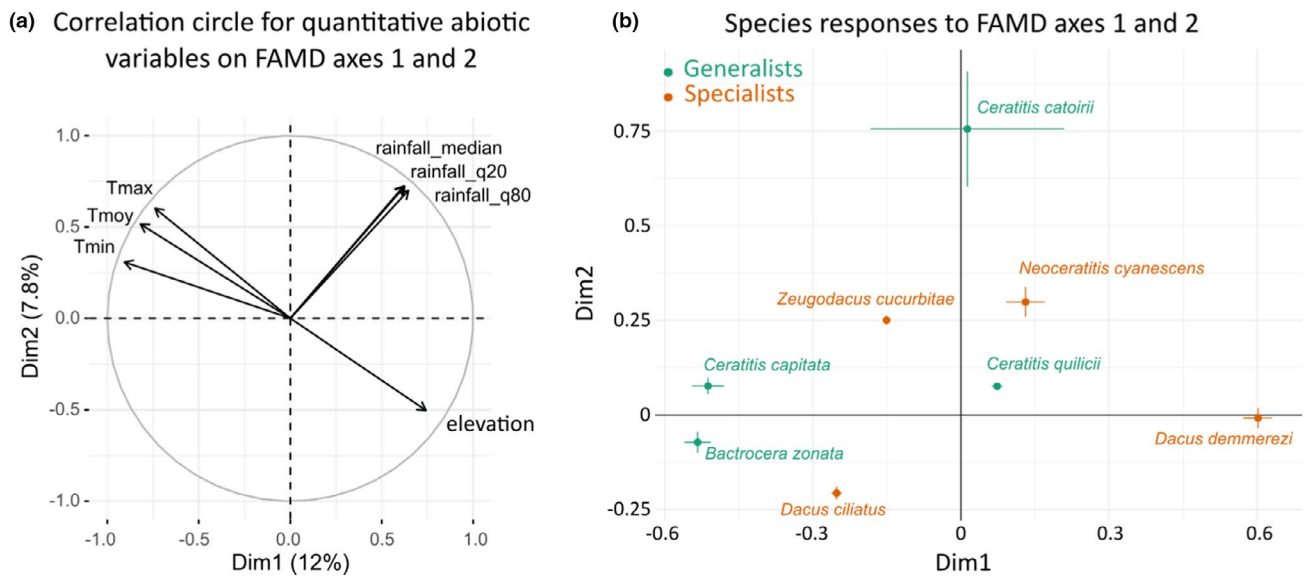


FIGURE 4 Species abundance responses to ecological variables. (a) Correlation circle on the first and second axes of the FAMD on ecological variables (Dim1 and Dim2, respectively). The first axis contrasts warm low-altitude sites and colder high-altitude sites. The second axis is a gradient of rainfall and maximal temperature (see Figure S3 for details on axes contributions). (b) Regression slopes relative to Dim1 and Dim2, inferred under the fullest model on the 21 plant × 8 fly species dataset (Model 1–5). Error bars represent approximate confidence intervals ($1.96 \times$ standard errors). For the first axis (Dim1), negative slopes (e.g. *Bactrocera zonata* and *Ceratitis capitata*) can be interpreted as a positive effect of temperatures and a negative effect of elevation on species abundances. For the second axis (FAMD 2), positive slopes (e.g. *Ceratitis catoirii*) can be interpreted as a positive effect of rainfall on species abundances

C. quilicii (stability =1.0). Contrary to the other subcommunities, the model with only ecological covariates performed well, almost as good as the model with host plant alone, suggesting redundancy between ecological information and plant identity and a low importance of plant identity. Species traits did not clearly improve model fit.

DISCUSSION

The determinants underlying the structure of a community of eight tephritid fly species were deciphered. Modelling joint species abundances without accounting for any covariate (only intercepts and an offset) confirmed a major role of host use strategy on fly species abundances. Species abundances co-varied positively among generalists and among specialists and negatively between species of each of these groups. Common responses to environmental factors may cause positive residual correlations in species abundances, while divergent responses will imply negative correlations, potentially leading to incorrect interpretations of species interactions (Dormann et al., 2018; Ovaskainen et al., 2016). Accounting for environmental covariates strongly improved model fit and made all residual covariances almost completely vanish, particularly among groups, suggesting that no important environmental factor structuring the community has been missed. Obtaining long-term abundance data on all a priori relevant species of a given community is rarely possible. Here, over the study period, one other fruit fly species (*Carpomya vesuviana*) was mentioned locally. It was considered very rare and only present on Indian jujube (*Ziziphus mauritiana*) in dry low-elevation areas (Quilici & Jeuffrault, 2001; Franck *et al.* 2017). It was not detected in the 204 jujube samples of the full dataset. In addition, not all 108 plant species considered as potential fruit fly hosts were included. To evaluate the effect of omitted plants, we conducted the analysis on all plants with at least 10 samples with emerging flies (6434 samples, 36 plant species including the 21 studied in the laboratory) and obtained strikingly similar results (model ranking and estimates of regression coefficients; Appendix S3), comforting the idea that most relevant factors have been accounted for.

Detection of competitive interactions

Some residual covariances remained non-negligible after accounting for fly species' response to host plants between some generalists and between some specialists of Cucurbitaceae. They were all negative, suggesting a possible minor role of antagonistic interactions within specialists and within generalists. Only one of these residual covariations resisted the network inference process on the whole dataset (*D. ciliatus* and *Z. cucurbitae*, the two most abundant specialists of Cucurbitaceae).

Two more significant covariations were detected when focusing on Cucurbitaceae (*Z. cucurbitae* and *D. demmerezi*) and on other plant families (*B. zonata* and *C. quilicii*), suggesting possible dependence of species interactions on host plants. On Cucurbitaceae, although qualitatively congruent with other independent empirical measures of host range (Vayssières et al., 2008) and climatic niche (Vayssières & Carel, 1999), host plant species and abiotic factors only moderately improved model fit. All three specialist flies found on Cucurbitaceae are able to thrive on any plant of this family (Charlery de la Masselière et al., 2017b), and competitive interactions between these fly species are highly plausible (Vayssières et al., 2008). On other plant families, both host plants and abiotic factors clearly improved model fit, congruently with former interpretations of the system (Duyck et al., 2008; Duyck et al., 2006a). Responses of generalist species to abiotic factors were strikingly congruent with former independent laboratory experiments (Duyck et al., 2006a; Duyck & Quilici, 2002; Duyck et al., 2004). There was redundancy between host plants and abiotic factors. Many of these plants are exploited but not planted (e.g. Myrtaceae). Their distributions are therefore more dependent on ecological factors than those of Solanaceae and Cucurbitaceae, which are mainly cultivated throughout the year in Réunion. On Solanaceae, no residual covariation was detected. Plant identity was the main determinant of species abundances, congruently with the idea that Solanaceae impose adaptive challenges on their fruit consumers through a variety of toxic compounds (Brévault et al., 2008), rendering host adaptation the main factor driving species abundances.

A current competition ghost

Among-group covariation between specialists and generalists was mainly attributable to fly species' adaptation to host use with a minor contribution of abiotic factors. Previous studies have highlighted differences in host adaptation between these fly species (Hafsi et al., 2016). Contrary to the specialists, which are mainly able to use their preferred hosts, the four generalists can thrive in numerous plant species and have weak female preferences (Charlery de la Masselière et al., 2017b). Accordingly, specialists were seldom found in plants other than Cucurbitaceae or Solanaceae. However, these results do not explain why generalists were so rarely found on Cucurbitaceae and Solanaceae. Competitive exclusion with specialists would be a natural hypothesis to explain this absence (Nakadai et al., 2018). Here, no competition among groups was detected. It is possible, however, that competition has already operated and that competitive exclusion has been so strong that generalists cannot be found on Cucurbitaceae. In PLN modelling, such absence could be interpreted as negative response of generalist abundances to Cucurbitaceae and be encapsulated in

plant cofactor slopes rather than in residual covariances. It is precisely when competition is intense enough to cause niche partitioning that it can no longer be detected. This result evokes a well-known paradox in ecology termed ‘the ghost of competition past’ (Connell, 1980) according to which the observed differentiation in niches is the result of past interspecific competition.

To escape the paradox, knowledge about the fundamental niches of species through eco-evolutionary approaches could help settle whether species interactions are an important driver of species assemblages (Augustyn et al., 2016; Dormann et al., 2018). Laboratory measurements of larval performance and female preference on host plants were used in replacement of plant identity. Congruently with the community being essentially driven by host use, preference and performance clearly improved model fit. Interestingly, performance was more informative than preference, which was expected from previous knowledge that generalists’ preferences are uncorrelated to their performances (Charlery de la Masselière et al., 2017b). If competition truly shapes species abundances, and has not been detected, it is to be found in the difference between models with plant identity as a cofactor and models with species traits instead. Here, we found a difference suggesting that competitive exclusion is at work. In terms of importance, from model rankings, host use patterns were the most important factor shaping species abundances, followed by abiotic factors and possibly a dose of competition.

This predominance of host plants as a structuring factor of phytophagous insect communities has been much debated, but congruent studies exist. In analyses of insect communities along road verges, Schaffers et al., (2008) found that the composition of plant communities was a much better predictor of insect and spider assemblages than environmental variables. Similarly Nakadai et al., (2018) found that sharing of host species was predominant among butterflies of the Japanese archipelago, suggesting that interspecific resource competition may not effectively determine community assembly patterns at regional scales. In an earlier review on the importance of competition in insects, Denno et al., (1995) pointed that only weak to moderate effect of competition should certainly be expected in phytophagous insects such as tephritids due to their high mobility and weak aggregation behaviours. Experimental manipulations of competitive interactions in the field could offer a promising way to test the validity of the present inferences. These experiments would also be useful to unveil the role of other biotic interactions (e.g. natural enemies), as forces capable of modulating interspecific competition between fruit flies.

Generalists versus specialists

Overall specialists and generalists almost had very distinct realised host uses with different assembly rules.

That specialists and generalists form separate interaction networks has already been highlighted, for example, among soil microbial species (Barberan et al., 2012). It is well known that the predictability of assemblages differs between generalist and specialist phytophagous insects (Müller et al., 2011). This has led to the hypothesis that specialists would assemble according to the species-sorting paradigm of metacommunity ecology (Leibold et al., 2004), whereby species occurrences are mainly driven by habitat heterogeneity and local adaptation. Generalists’ assembly rules, however, would rather follow a mass-effect paradigm (Shmida & Wilson, 1985), according to which sink populations, where the species is maladapted, can persist through a migration influx from source populations (Müller et al., 2011). Our results confirm this hypothesis. For specialists, a good agreement between the inferred host plant effects on species abundances and the laboratory measures of host adaptation suggested that specialists were mainly filtered by host plant characteristics. In contrast, generalists displayed no relationship between inferred and laboratory-measured host plant effects, suggesting that generalists were found on some hosts where their fitness is low and at low density on good hosts. Besides, generalists use fruits whose availability is highly variable over time. Contrary to specialists, most of their hosts are not available all year long (Figure S4). This temporally variable habitat may trigger a dynamics of local extinction and recolonisation, in which the roles of migration and stochasticity become more important than that of host adaptation and in which coexistence is possible despite fundamental niche overlap (Chesson, 2000; Connell, 1980).

ACKNOWLEDGEMENTS

This work is dedicated to Serge Quilici who has been leading prolific research on the tephritids of Réunion his whole career. Cirad technicians Jim Payet and Serge Glénac made this study possible through their invaluable expertise with fly rearing and ecology. We also thank them as well as Antoine Franck, Christophe Simiand and Patrick Turpin for collecting field data over the years. Thomas Brequigny contributed to measuring larval traits during his internship. Thicenters work used images acquired within the framework of the CNES Kalideos device (Réunion site), which benefited from the ‘Programme Investissements d’Avenir’ EQUIPEX of the French ‘Agence Nationale de la Recherche’ on project GEOSUD bearing the reference ANR-10-EQPX-20. The images also required financial support the French Ministry of Agriculture and field data transmitted by the ‘Syndicat du Sucre de la Réunion’ and the ‘SAFER de la Réunion’. BF, FC, FM, JC, MD, SR and VR received the financial support of the French ‘Agence Nationale de la Recherche’ project NGB (ANR-17-CE32-011). EF, FC, PFD and VR were funded by the European Union (European Regional

Development Fund, ERDF contract GURDT I2016-1731-0006632), the Conseil Régional de La Réunion and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD). AH was funded by the ‘Ministère de l’Enseignement supérieur et de la Recherche Scientifique de la Tunisie’. This study used the facilities provided by the Plant Protection Platform (3P, IBISA), Saint-Pierre, Réunion, France.

CONFLICT OF INTEREST

The authors declare that they have no financial conflict of interest with the content of this article.

AUTHORSHIPS

BF, VR and PFD conceived the study. AH and MCM performed the laboratory experiments. BF and VR processed the data with assistance from FM, JC, SR, MD, EF and FC for the statistical analyses. BF and VR wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ele.13825>.

DATA AVAILABILITY STATEMENT

Data and script are available online: <http://doi.org/10.5281/zenodo.4569532>

ORCID

Benoit Facon  <https://orcid.org/0000-0001-8970-6840>
 François Massol  <https://orcid.org/0000-0002-4098-955X>
 Enric Frago  <https://orcid.org/0000-0001-8817-1303>
 Pierre-François Duyck  <https://orcid.org/0000-0001-5484-1970>

REFERENCES

- Adler, P.B., Lambers, H.R.J. (2008) The influence of climate and species composition on the population dynamics of ten prairie forbs. *Ecology*, 89, 3049–3050.
- Araújo, M.B. & Rozenfeld, A. (2014) The geographic scaling of biotic interactions. *Ecography*, 37, 406–415.
- Augustyn, W.J., Anderson, B. & Ellis, A.G. (2016) Experimental evidence for fundamental, and not realized, niche partitioning in a plant–herbivore community interaction network. *Journal of Animal Ecology*, 85, 994–1003.
- Barberan, A., Bates, S.T., Casamayor, E.O. & Fierer, N. (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal*, 6, 343–351.
- Blanchet, F.G., Cazelles, K. & Gravel, D. (2020) Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23, 1050–1063.
- Brazeau, H.A. & Schamp, B.S. (2019) Examining the link between competition and negative co-occurrence patterns. *Oikos*, 128, 1358–1366.
- Brévault, T., Duyck, P.F. & Quilici, S. (2008) Life-history strategy in an oligophagous tephritid: the tomato fruit fly, *Neoceratitis cyanescens*. *Ecological Entomology*, 33, 529–536.
- Charlery de la Masselière, M., Facon, B., Hafsi, A. & Duyck, P.F. (2017b) Diet breadth modulates preference-performance relationships in a phytophagous insect community. *Scientific Reports*, 7, 16934.
- Charlery de la Masselière, M., Ravigné, V., Facon, B., Lefeuvre, P., Massol, F., Quilici, S. et al. (2017a) Changes in phytophagous insect host ranges following the invasion of their community: Long-term data for fruit flies. *Ecology and Evolution*, 7, 5181–5190.
- Chen, J. & Chen, Z. (2008) Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 95, 759–771.
- Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31, 343–366.
- Chiquet, J., Mariadassou, M. & Robin, S. (2018) Variational inference for probabilistic poisson PCA. *The Annals of Applied Statistics*, 12, 2674–2698.
- Chiquet, J., Mariadassou, M. & Robin, S. (2019). Variational Inference for Sparse Network Reconstruction from Count Data. In *Proceedings of the 36th International Conference on Machine Learning*, edited by Kamalika Chaudhuri and Ruslan Salakhutdinov, 97, 1162–71. Proceedings of Machine Learning Research. Long Beach, California, USA: PMLR.
- Connell, J.H. (1980) Diversity and the coevolution of competitors, or the ghost of competition past. *Oikos*, 35, 131–138.
- D’Amen, M., Mod, H.K., Gotelli, N.J. & Guisan, A. (2018) Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. *Ecography*, 41, 1233–1244.
- Denno, R.F., McClure, M.S. & Ott, J.R. (1995) Interspecific interactions in phytophagous insects: competition reexamined and resurrected. *Annual Review of Entomology*, 40, 297–331.
- Diamond, J.M. (1975) Assembly of species communities. In: Cody, M.L. & Diamond, J.M. (Eds.) *Ecology and evolution of communities*. Harvard Univ. Press, pp. 342–344.
- Dormann, C.F., Bobrowski, M., Dehling, D.M., Harris, D.J., Hartig, F., Lischke, H. et al. (2018) Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, 27, 1004–1016.
- Duyck, P.F., David, P., Junod, G., Brunel, C., Dupont, R. & Quilici, S. (2006b) Importance of competition mechanisms in successive invasions by polyphagous Tephritids in La Réunion. *Ecology*, 87, 1770–1780.
- Duyck, P.F., David, P., Pavoine, S. & Quilici, S. (2008) Can host-range allow niche differentiation of invasive polyphagous fruit flies (Diptera: Tephritidae) in La Réunion? *Ecological Entomology*, 33, 439–452.
- Duyck, P.F., David, P. & Quilici, S. (2006a) Climatic niches partitioning following successive invasions by fruit flies in La Réunion. *Journal of Animal Ecology*, 75, 518–526.
- Duyck, P.F. & Quilici, S. (2002) Survival and development of different life stages of three *Ceratitis* spp. (Diptera: Tephritidae) reared at five constant temperatures. *Bulletin of Entomological Research*, 92, 461–469.
- Duyck, P.F., Sterlin, J.F. & Quilici, S. (2004) Survival and development of different life stages of *Bactrocera zonata* (Diptera: Tephritidae) reared at five constant temperatures compared to other fruit fly species. *Bulletin of Entomological Research*, 94, 89–93.
- Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697.
- Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43, 393–404.
- Godinho, D.P., Janssen, A., Dias, T., Cruz, C. & Magalhães, S. (2016) Down-regulation of plant defence in a resident spider mite species and its effect upon con- and heterospecifics. *Oecologia*, 180, 161–167.

- Gotelli, N.J. & Graves, G.R. (1996) *Null models in Ecology*. Washington DC: Smithsonian Institution Press.
- Gotelli, N.J., Graves, G.R. & Rahbek, C. (2010) Macroecological signals of species interactions in the Danish avifauna. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 5030–5035.
- Hafsi, A., Facon, B., Ravigné, V., Chiroleu, F., Quilici, S., Chermiti, B. et al. (2016) Host plant range of a fruit fly community (Diptera: Tephritidae): Does fruit composition influence larval performance? *BMC Ecology*, 16, 40.
- Hairston, N.G., Smith, F.E. & Slobodkin, L.W. (1960) Community Structure, Population Control, and Competition. *American Naturalist*, 94, 421–425.
- Hartig, F. (2020). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.3.3.0. <https://CRAN.R-project.org/package=DHARMA>
- Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Körber, J.H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, 16, 754–763.
- Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J. (2020) Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*, 23, 1330–1339.
- Kaplan, I. & Denno, R.F. (2007) Interspecific interactions in phytophagous insects revisited: a quantitative assessment of competition theory. *Ecology Letters*, 10, 977–994.
- Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, 322, 580–582.
- Lavorel, S., McIntyre, S., Landsberg, J. & Forbes, T.D.A. (1997) Plant functional classifications: from general groups to specific groups based on response to disturbance. *Trends in Ecology & Evolution*, 12, 474–478.
- Lê, S., Josse, J. & Husson, F. (2008) FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25, 1–18.
- Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. et al. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Eco. Lett.*, 7, 601–613.
- Liu, H., Roeder, K. & Wasserman, L. (2010) Stability Approach to Regularization Selection (Stars) for High Dimensional Graphical Models. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 2*, 1432–40. USA.
- Meynard, C.N., Lavergne, S., Boulangeat, I., Garraud, L., Van Es, J., Mouquet, N. et al. (2013) Disentangling the drivers of metacommunity structure across spatial scales. *Journal of Biogeography*, 40, 1560–1571.
- Müller, J., Stadler, J., Jarzabek-Müller, A., Hacker, H., ter Braak, C. & Brandl, R. (2011) The predictability of phytophagous insect communities: host specialists as habitat specialists. *PLoS One*, 6, e25986
- Nakadai, R., Hashimoto, K., Iwasaki, T. & Sato, Y. (2018) Geographical co-occurrence of butterfly species: the importance of niche filtering by host plant species. *Oecologia*, 186, 995–1005.
- Ovaskainen, O., Roy, D.B., Fox, R. & Anderson, B.J. (2016) Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods in Ecology and Evolution*, 7, 428–436.
- Poisot, T., Stouffer, D.B. & Gravel, D. (2015) Beyond species: why ecological interaction networks vary through space and time. *Oikos*, 124, 243–251.
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M. et al. (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5, 397–406.
- Quilici, S. & Jeuffraut, E. (2001) *Plantes-hôtes des mouches des fruits: Maurice, Réunion, Seychelles*: PRMF/COI, Imp. Graphica, St André, La Réunion.
- Roy, H.E., Rorke, S.L., Beckmann, B., Booy, O., Botham, M.S., Brown, P.M.J. et al. (2015) The contribution of volunteer recorders to our understanding of biological invasions. *Biological Journal of the Linnean Society*, 115, 678–689.
- Schaffers, A.P., Raemakers, I.P., Sykora, K.V. & ter Braak, C.J.F. (2008) Arthropod assemblages are best predicted by plant species composition. *Ecology*, 89, 782–794.
- Shmida, A. & Wilson, M.V. (1985) Biological determinants of species diversity. *Journal of Biogeography*, 12, 1–20.
- Stewart, A.J.A., Bantock, T.M., Beckmann, B.C., Botham, M.S., Hubble, D. & Roy, D.B. (2015) The role of ecological interactions in determining species ranges and range changes. *Biological Journal of the Linnean Society*, 115, 647–663.
- Tack, A.J.M., Ovaskainen, O., Harrison, P.J. & Roslin, T. (2009) Competition as a structuring force in leaf miner communities. *Oikos*, 118, 809–818.
- Thuiller, W., Pollock, L.J., Gueguen, M. & Münkemüller, T. (2015) From species distributions to meta-communities. *Ecology Letters*, 18, 1321–1328.
- Ulrich, W., Jabot, F. & Gotelli, N.J. (2017) Competitive interactions change the pattern of species co-occurrences under neutral dispersal. *Oikos*, 126, 91–100.
- Vayssières, J.F. & Carel, Y. (1999) Les Dacini (Diptera: Tephritidae) inféodés aux Cucurbitaceae à la Réunion: gamme de plantes-hôtes et stades phénologiques préférentiels des fruits au moment de la piqûre pour des espèces cultivées. *Annales de la Société entomologique de France*, 35, 197–202.
- Vayssières, J.F., Carel, Y., Coubes, M. & Duyck, P.F. (2008) Development of immature stages and comparative demography of two cucurbit-attacking fruit flies in Reunion island: *Bactrocera cucurbitae* and *Dacus ciliatus* (Diptera Tephritidae). *Environmental Entomology*, 37, 307–314.
- Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F. et al. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 88, 15–30.
- Zurell, D., Pollock, L.J. & Thuiller, W. (2018) Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography*, 41, 1812–1819.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Facon, B., Hafsi, A., Charlery de la Masseliere, M., Robin, S., Massol, F., Dubart, M., Chiquet, J. et al. (2021) Joint species distributions reveal the combined effects of host plants, abiotic factors and species competition as drivers of species abundances in fruit flies. *Ecology Letters*, 00, 1–12. <https://doi.org/10.1111/ele.13825>