



Do ecological specialization and functional traits explain the abundance–frequency relationship? Arable weeds as a case study

Guillaume Fried¹ | Laura Armengot² | Jonathan Storkey³ |
Bérenger Bourgeois^{4,5} | Sabrina Gaba^{6,7} | Cyrille Violle⁸ | François Munoz⁹

¹Anses, Laboratoire de la Santé des Végétaux, Unité Entomologie et Plantes Invasives, Montferrier-sur-Lez, France

²FiBL, Research Institute of Organic Agriculture, Frick, Switzerland

³Sustainable Agriculture Sciences, Rothamsted Research, Harpenden, Hertfordshire, 2JQ, United Kingdom

⁴Centre de Synthèse et d'Analyse sur la Biodiversité – Fondation pour la Recherche sur la Biodiversité, Montpellier, 34000, France

⁵Département de Phytologie, Université Laval, Québec, QC G1V 0A6, Canada

⁶USC Agripop, Centre d'Etudes Biologiques de Chizé, INRAE, Villiers-en-Bois, 79360, France

⁷Centre d'Etudes Biologiques de Chizé, UMR 7372, CNRS & Univ. La Rochelle, Villiers-en-Bois, F-79360, France

⁸CEFE Univ Montpellier – CNRS – EPHE – IRD – Univ Paul Valéry, Montpellier, 34293, France

⁹Laboratoire Interdisciplinaire de Physique, Université Grenoble-Alpes, Grenoble, 34081, France

Correspondence

Anses, Laboratoire de la Santé des Végétaux, Unité Entomologie et Plantes Invasives, 34988 Montferrier-sur-Lez, France.
Email: guillaume.fried@anses.fr

Funding information

Fondation pour la Recherche sur la Biodiversité, Grant/Award Number: Cesab DISCO-WEED Working Group

Handling Editor: Hanno Schaefer

Abstract

Aim: The abundance–frequency relationship (AFR) is among the most-investigated patterns in biogeography, yet the relative contributions of niche-based processes related to ecological strategies, and of neutral processes related to spatial colonization–extinction dynamics, remains uncertain. Here, we tested the influences of ecological specialization and functional traits on local abundance and regional frequency, to determine the contribution of niche-based processes.

Location: France and the UK.

Taxon: Vascular plants.

Methods: We used two arable weed surveys covering 1,544 fields in Western Europe (France, UK), along with functional traits related to resource acquisition, resource requirements, flowering phenology and dispersal. We quantified specialization both to arable habitat and to individual crop types, and performed phylogenetic path analyses to test competing models accounting for direct and indirect relationships between traits, specialization, abundance and frequency. We performed the analyses for all species in each country, as well as for a subset of the most abundant species.

Results: Local abundance of weeds increased with their regional frequency, but the relationship became negative or null when considering only the most abundant weeds. Specialization to arable habitat and to individual crop type either had a similar or opposite effect on regional frequency and local abundance explaining these positive and negative relationships, respectively. Regional frequency was not directly explained by any trait but indirectly by resource requirement traits conferring specialization to the arable habitat. Conversely, high local abundance was directly related to low seed mass, high SLA, early and short flowering.

Main conclusions: Direct/indirect effects of functional traits on local abundance/regional frequency, respectively, supports a significant role of niche-based processes in AFR. Neutral spillover dynamics could further explain a direct linkage of abundance and frequency. Similar causal paths and consistent influences of traits on specialization and abundance in the two studied regions suggest genericity of these findings.

KEYWORDS

generalist–specialist, neutral processes, niche-breadth, occupancy, path analysis, weed biogeography

1 | INTRODUCTION

Understanding why locally abundant species are often regionally widespread, whereas locally rare species are narrowly distributed, is a fundamental question in biogeography (Köckemann, Buschmann, & Leuschner, 2009; Soininen & Heino, 2005). A positive relationship between local abundance and regional frequency (abundance-frequency relationship, hereafter AFR) has been found in 80% of published studies, for many taxonomical groups: microbes, algae, bryophytes, vascular plants, arthropods, mammals or birds (Gaston et al., 2000). Contrastingly, only 5% of studies examined by Gaston

(1996) reported a negative relationship questioning the processes at play in the AFR. The nature and role of local and regional processes driving the pattern thereby remain uncertain and the study of the AFR should now shift from pattern documentation to a more process-based perspective.

The way local abundance relates to regional frequency can reflect the influence of several mechanisms such as local assembly processes (Keddy, 1992), spatial dynamics across heterogeneous environments (Leibold et al., 2004) and biogeographic history (Jetz, Rahbek, & Colwell, 2004). In recent years, studies on AFR focused on neutral, niche-independent spatial colonization–extinction

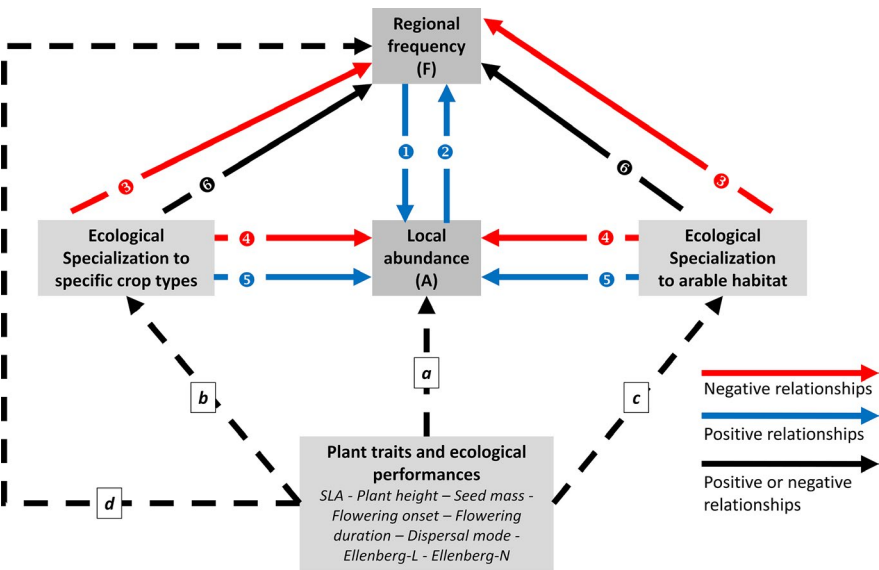


FIGURE 1 Relationships between regional frequency (F), local abundance (A), ecological specialization and functional traits. Arrows represent the directionality of relationships based on theoretical predictions, and numbers refers to the hypotheses proposed to explain the AFR in the Introduction and in the legend opposite. Dashed arrows indicate elementary relationships between traits, specialization, local abundance and regional frequency, which are analysed to assess the contributions of processes underlying the AFR

Legend		
Hypothesis	Arrows	Notes
<i>Neutral-based process</i>		
<i>Spillover</i>	1, 2	Widespread species are more likely to colonise empty patches; while locally abundant species are more likely to disperse to surrounding sites: positive AFR expected
<i>Niche-based processes</i>		
<i>Jack of all trades, master of all</i>	3, 4	Habitat generalists are both more widespread and locally abundant : positive AFR expected
<i>Jack of all trades, master of none</i>	5	Habitat specialists are more locally abundant in their preferred habitat: positive AFR if arrow 6 is positive, negative AFR if arrow 6 is negative.
<i>Resource availability hypothesis</i>	6	Direction of the relationship between habitat specialization and regional frequency will be determined by the frequency of the preferred habitat in the landscape or the dataset of sample points

dynamics or conversely on the role of niche differences among species (Gregory & Gaston, 2000; Heino & Grönroos, 2014; Verberk, Van Der Velde, & Esselink, 2010). From the neutral perspective, regionally widespread species can more easily colonize empty patches and reinforce local populations, thereby increasing local abundance (Hubbell, 2001), a process sometimes called *spillover* (e.g. Rand, Tylaniakis, and Tschardtke (2006), Figure 1, arrow 1). Furthermore, locally abundant species can send more dispersers to surrounding sites and spread more efficiently at regional scale (arrow 2), thus increasing regional frequency (Shmida & Wilson, 1985).

From the niche-based perspective, ecological specialization and habitat filtering can also shape the AFR. Variation in ecological specialization (habitat breadth) should influence the AFR in several ways. Species occurring in diverse environmental conditions (hereafter *habitat generalists*) can occupy more sites and be regionally more frequent than species with narrower ecological niche (hereafter *habitat specialists*) (Brown, 1984). The *jack-of-all-trade-master-of-all* hypothesis states that habitat generalists are not only more widespread but also locally more abundant than habitat specialists, although this latter idea has received little support (Krasnov, Poulin, Shenbrot, Mouillot, & Khokhlova, 2004; Lawton, 1993). According to this hypothesis, ecological specialization would negatively relate with regional frequency (arrow 3) but also with local abundance (arrow 4), hence leading to a positive AFR. Alternatively, the *jack-of-all-trade-master-of-none* hypothesis states that habitat specialists can have higher performance in their preferred habitat than habitat generalists (Devictor et al., 2010; Futuyma & Moreno, 1988). According to this hypothesis, ecological specialization should be positively correlated to local abundance (arrow 5). The regional distribution of the preferred habitat of these specialist species will then determine their regional frequency according to the *resource availability* hypothesis (Hanski,

Kouki, & Halkka, 1993), arrow 6. If the species are specialists of a narrowly distributed habitat, a negative relationship with the regional frequency would be expected since they will be regionally rarer than habitat generalists (Gaston, 1996). Conversely, if specialists use a widespread habitat or if regional frequency is only measured in their preferred habitat, a positive relationship would be expected (Gaston (1996). The combination of the *jack-of-all-trade-master-of-none* (arrow 5) and *resource availability* (arrow 6) hypotheses should thus yield either a positive or negative AFR.

Still from a niche-based perspective, species functional traits are expected to be constrained towards some 'optimum' values, for which fitness is maximized under the local environmental conditions (Violle et al., 2007). If species functional traits reflect niche-dependent variations in local performance, they should help to decipher the contribution of niche-based processes in the AFR (McGill, Enquist, Weiher, & Westoby, 2006; Violle et al., 2007). We expect that traits related to resource acquisition, resource requirements and competition can explain variation in local abundance (arrow *a*) and variation in specialization (arrows *b* and *c*) while traits related to dispersal and recruitment (wind dispersal, seed mass) can entail variation in regional frequency (Lloret, Medail, Brundu, & Hulme, 2004) (arrow *d*). An open question is to what extent functional traits influence regional frequency directly (arrow *d*), or rather indirectly by affecting local abundance (arrow *a*) and/or ecological specialization (arrow *b* or *c*).

Moreover, locally rare and abundant species are likely to differ in the AFR in relation to distinct neutral- and niche-based processes affecting them. One hypothesis is that abundant species should be mainly affected by environmental factors, while rare species should be mostly influenced by dispersal limitation and stochastic dynamics (Siqueira et al., 2012). Conversely, rare species can be more specialized and more influenced by the environment and local processes, while abundant species could be more generalists and therefore

TABLE 1 Summary statistics of traits used and sources of the data

Traits	France (BVG Dataset) N = 122 species			UK (FSE Dataset) N = 102 species			Sources
Quantitative traits	Min-Max	Median	Mean	Min-Max	Median	Mean	
Plant height (m)	0.2–5	0.8	0.97	0.1–5	1	0.99	France: Tison and De Foucault (2014); UK: Clapham, Tutin, and Moore (1989)
SLA (mm ² /mg)	6.3–53.7	26.6	27.5	10.9–53.7	26.6	27.3	Kleyer et al. (2008)
Seed mass (g)	0.015–203.4	1.1	4.8	0.02–30.3	0.7	2.4	Kew (2018)
Flowering onset (month)	1–8	5.0	5.1	1–8	6.0	5.5	France: Tison and De Foucault (2014); UK: Clapham et al. (1989)
Flowering duration (month)	1–12	4.0	5.0	1–12	4.0	4.2	
Ellenberg-N	2–9	7.0	6.6	2–9	7.0	6.8	France: Julve (1998), UK: Hill, Mountford, Roy, and Bunce (1999)
Ellenberg-L	4–9	7.0	7.1	4–9	7.0	7.0	
Qualitative traits							
Dispersal mode							Julve (1998)
Wind	n = 29			n = 31			
Other means	n = 93			n = 71			



more influenced by regional and dispersal processes (Pandit, Kolasa, & Cottenie, 2009). Testing the AFR either for all species, or separately for the most abundant ones, would help to understand how rare and abundant species affect the AFR pattern and therefore to determine the underlying processes.

We tested the contributions of neutral- and niche-based processes to the AFR using arable weeds as a model. Weed flora in arable fields represents a relevant model to address biogeographical and macroecological theories (Mahaut et al., 2020). First, the crop environment is relatively homogeneous across regions making an analysis of the drivers of the AFR at large scale more tractable by avoiding spurious influences of uncontrolled environmental factors (Ikegami, Wandrag, Duncan, & Hulme, 2019). Second, arable fields are subject to highly frequent disturbances so that weed populations undergo broad demographic fluctuations; thus, neutral dispersal dynamics can greatly influence local abundances (Perronne, Le Corre, Bretagnolle, & Gaba, 2015). Conversely, strong constraints associated with farming practices (regular tillage and weeding, including herbicides) and biotic factors (competition with crop plants) enhance the influence of niche-based processes in weed community assembly (Fried, Norton, & Reboud, 2008). Third, crop fields delineate spatially distinct weed assemblages, thus providing a relevant case to address meta-community dynamics. In this context, dispersal plays a key role both at the landscape level (Bourgeois, Gaba, Plumejeaud, & Bretagnolle, 2020), through natural or human-assisted dispersal (e.g. through combine harvester), and at regional or even inter-continental scale, due to contamination of crop seeds. Fourth, plants in arable fields cover a wide spectrum of abundance and specialization (Munoz et al., 2020), from specialist weeds adapted to farming practices since millennia, to casual weeds occurring more often in other open habitats (grasslands, fringes), the presence of which is expected to rely more on dispersal from adjacent environments (Metcalf, Hassall, Boinot, & Storkey, 2019). Moreover, among specialist weeds, some are specific to certain crop types while others are able to grow in all crop types (Fried, Petit, & Reboud, 2010). For these reasons, weeds represent an excellent model for investigating the interplay of neutral- and niche-based processes, from local to regional scales, and thus addressing fundamental questions of biogeography and macroecology.

We analysed field plot composition in different crop types in a broad study area in France and the UK, to quantify local abundances, ecological specialization to arable habitat and to specific crop types, and regional frequency of arable weed species. We selected eight traits representing competitive, regenerative and dispersal abilities of weed species (Table 1). We devised a structural equation model to test theoretical predictions of Figure 1 and to answer the following questions: (i) is there a positive AFR for arable weeds, (ii) what is the contribution of rare versus abundant weeds to the overall pattern? (iii) how does ecological specialization to arable habitat or to specific crop types influence AFR? and (iv) do functional traits directly influence regional abundances, or indirectly through local performance and specialization to the arable habitat?

2 | MATERIALS AND METHODS

2.1 | Study areas and weed sampling

We analysed two nationwide surveys of arable weed communities conducted in France: *Biovigilance Flore*, BVG hereafter (Fried et al., 2008), and in the UK: *Farm scale evaluation*, FSE hereafter (Heard et al., 2003). The detailed sampling protocols are described in Fried et al. (2008) and Heard et al. (2003), respectively. Appendix S1 summarizes the protocols and presents how we selected samples to allow comparison between the two regions. The regional frequency of a species was defined as the proportion of sampled fields where the species was present. Local mean abundance was calculated as the average density in sites where the species was present. Local mean abundance and regional frequency were log-transformed prior to analyses (Gaston et al., 2000). We first analysed local abundance and regional frequency in the 'arable field' habitat, in UK and France separately, whatever the crop type. Second, we analysed relationships between local abundance and regional frequency separately for the four main crop types found in both French and UK surveys (maize, wheat, oilseed rape and sugarbeet).

2.2 | Functional traits and ecological specialization

The eight selected traits are summarized in Table 1 with their unit, the distributions of trait values in the species pools from France and UK weed surveys, and the sources of trait data. Canopy height at maturity represents the ability to compete for light with neighbouring plants, especially with crop species (Westoby, Falster, Moles, Vesk, & Wright, 2002). Specific leaf area (SLA) represents the ability to acquire and use resources during the favourable period, and is positively correlated to the relative growth rate of weed seedlings (Storkey, 2004). A specific flowering period (early or late) can lead to enhanced fitness of weeds depending on crop phenology: early flowering in winter-sown crops, such as winter wheat and winter oilseed rape, versus late flowering in spring-sown crops such as sugarbeet and maize (Perronne et al., 2015). Generally, species emerging early during the crop growing season, with early flowering onset and short life cycle, should better escape competition and herbicide treatments, and produce seeds before crop canopy closure (Gaba et al., 2017). Because of high disturbance and spatiotemporal variability in management practices, functional traits affecting weed dispersal ability and persistence in seed bank (dispersal mode and seed mass) should also determine the ability of weeds to colonize and persist in arable fields. In the context of intensive agriculture, high seed production (often associated to low seed mass) can relate to greater seed bank and counterbalance high mortality rate due to herbicide applications (Storkey, Moss, & Cussans, 2010). Species were also assigned to wind dispersal mode based on the presence of adaptations for long-distance dispersal by wind, which is expected to explain higher regional

frequency. Finally, two resource requirement traits were considered, namely, Ellenberg indicator values for light (Ellenberg-L) and nitrogen (Ellenberg-N). High rate of fertilization in arable fields should favour more nitrogen-demanding species. In the following analyses, plant height and seed mass were log-transformed to ensure normality.

We defined ecological specialization at two levels: between arable and non-arable habitats (here, *ecological specialization to arable habitat*), and across crop types within the arable habitat (here, *ecological specialization to crop types*). Ecological specialization to arable habitat differentiated weeds more often occurring in arable fields (habitat specialists) from those also often growing in other surrounding herbaceous habitats (habitat generalists). We used a fidelity measure to represent ecological specialization to the arable habitat, based on the relative frequency of a species in a reference habitat compared to other habitats (Chytrý, Tichý, Holt, & Botta-Dukát, 2002). In France, we assessed weed frequency in 5,382 vegetation samples

in arable fields (the complete BVG dataset), and in 96,438 samples covering diverse grassland habitats in the DIVGRASS database (Violle et al., 2015). In the UK, we assessed weed frequency in 268 samples in arable fields (the complete FSE dataset), and in 15,756 samples from other habitats excluding arable and horticulture habitats (Carey et al., 2008). We quantified the *phi* coefficient of association:

$$\phi = \frac{N \cdot n_p - n \cdot N_p}{\sqrt{n \cdot N_p \cdot (N - n) \cdot (N - N_p)}}$$

with *N* the total number of samples used (101,820 for France, 16,024 for UK), *N_p*, the number of samples in arable habitats (5,382 for France and 268 for UK, respectively), *n*, the number of occurrences of the species in the dataset and *n_p*, the number of occurrences of the species in arable habitats.

Second, we quantified specialization to crop types within each regional survey, as the coefficient of variation (*SD/mean*) of weed

TABLE 2 List of path models tested. F stands for regional frequency, A for local abundance, Phi for ecological specialization to arable habitats, ESC for ecological specialization to specific crop type, Traits for the plant traits and ecological performances

Model group	Model subgroup	#Model	Formulas
Common to all models (all path models link traits to abundance and specialization)		1–22	Phi ~ Traits, A ~ Traits, ESC ~ flowering duration + flowering period
No relationships between local abundance and regional frequency		1	F ~ Traits
		2	F ~ Phi+ESC
		3	F ~ Phi+ESC, A ~ Phi+ESC
		4	F ~ Phi+ESC, F ~ Traits
		5	F ~ Phi+ESC, A ~ Phi+ESC, F ~ Traits
Frequency is only a causal <i>parent</i> (the path models do not link traits to frequency)		6	A ~ F
		7	A ~ F+Phi + ESC
Frequency is determined <i>indirectly</i> by traits (the path models follow the arrows a, b, c in Figure 1 to link traits to frequency)	Regional frequency is determined by local abundance (arrow 2 on Figure 1)	8	F ~ A
		9	F ~ Phi+ESC + A, A ~ Phi+ESC
		10	F ~ Phi+ESC + A
		11	F ~ A, A ~ Phi+ESC
	Regional frequency determined local abundance (arrow 1 on Figure 1)	12	F ~ Phi+ESC, A ~ F
		13	F ~ Phi+ESC, A ~ Phi+ESC + F
Frequency is determined <i>directly</i> by traits (the path models follow the arrow d in Figure 1 to link traits to frequency)		14	A ~ F, F ~ Traits
		15	A ~ Phi+ESC, F ~ Traits
		16	A ~ F+Phi + ESC, F ~ Traits
Frequency is determined both <i>directly</i> and <i>indirectly</i> by traits (the path models follow the arrows a, b, c, d in Figure 1 to link traits to frequency)	Regional frequency is determined by local abundance (arrow 2 on Figure 1)	17	F ~ A, F ~ Traits
		18	F ~ Phi+ESC + A, F ~ Traits
		19	F ~ A, A ~ Phi+ESC, F ~ Traits
		20	F ~ Phi+ESC + A, A ~ Phi+ESC, F ~ Traits
	Regional frequency determined local abundance (arrow 1 on Figure 1)	21	F ~ Phi+ESC, A ~ F, F ~ Traits
		22	F ~ Phi+ESC, A ~ Phi+ESC + F, F ~ Traits

Note: All models including Phi and ESC in their formulas are testing different niche-based hypotheses (represented by arrows 3, 4, 5 and/or 6 on Figure 1). In addition, all models, except #1 to 5, include a test of neutral-based hypotheses (represented by arrows 1 and/or 2 on Figure 1). Best models (with $\Delta\text{CICc} < 2$) are in bold characters.

abundance (Devictor et al. 2010) across the four crop types, wheat, maize, oilseed rape and sugarbeet.

2.3 | Structural Equation Models

2.3.1 | Design of causal models

We designed path analyses to examine the relationships displayed in Figure 1 between regional frequency, local abundance and ecological specialization, and to assess how functional trait values directly or indirectly influenced these three variables. Path analysis is a form of structural equation modelling (Shipley, 2000) assessing putative causal relationships between two or more variables. The first step is to build a path diagram representing causal relationships between variables. We formulated the initial path model based on the relationships hypothesized in Figure 1. We assumed that functional traits should allow grasping niche-based processes underlying AFR so that some relationships were included and analysed in all models: (i) local abundance of weed species depending on trait values (Shipley, Vile, & Garnier, 2006), (ii) ecological specialization of species to a particular habitat, here arable fields, depending on their trait values (Futuyma & Moreno, 1988), (iii) specialization to specific crop types mainly depending on phenological traits, here flowering onset and duration (Perronne et al., 2015).

We designed the alternative path models to assess direct and indirect relationships (Table 2). We considered direct influences of

traits on regional frequency (model paths including arrow *d*, Figure 1). Alternatively, traits could indirectly influence regional frequency through their influences on local abundance and/or ecological specialization (model paths including arrows *a*, *b* and *c*). Thus, we built alternative models with direct effects (models #14, #15, #16, see Table 2), indirect effects (models #8, #9, #10, #11, #12, #13) and both direct and indirect effects (models #17, #18, #19, #20, #21, #22). Furthermore, regional frequency could be determined by local abundance (arrow 2 on Figure 1), or local abundance could be determined by regional frequency (arrow 1 on Figure 1), yielding distinct subgroups of path models. We also built models in which regional frequency exclusively influenced local abundance and was not determined by traits and ecological specialization (models #6, #7). Finally, we designed models with no relationship between local abundance and regional frequency (models #1, #2, #3, #4, #5). Table 2 summarizes the different groups of path models. We indicate how the models relate to the hypotheses represented by the arrows in Figure 1, and specify the direct or indirect influences of functional traits on regional frequency.

2.3.2 | Model comparison

We compared alternative causal models with the d-separation test developed by Shipley (2000). d-separation predicts a minimal set of conditional probabilistic independencies that must all be true if the causal model is correct (Verma & Pearl, 1988). We selected the

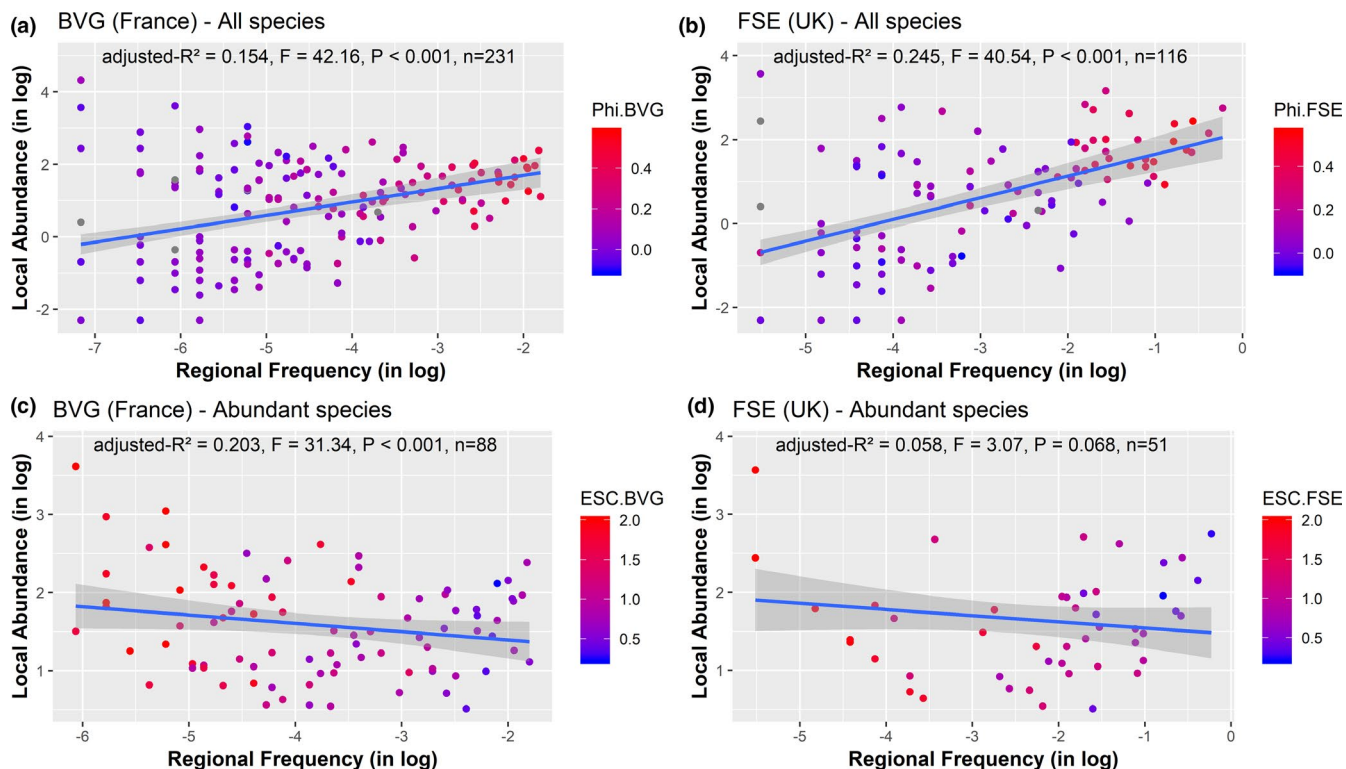


FIGURE 2 The relationship between regional frequency and mean local abundance at occupied sites in arable weed flora for different subsets of species. Model statistics based on Phylogenetic Generalized Least Squares regressions. Colours represent the degree of ecological specialization to arable habitats (Phi) and the degree of ecological specialization to crop types (ESC) from low (blue) to high (red) values

best causal models based on the C statistic Information Criterion (CICc), a modified version of Akaike Information Criterion (AIC) considering all possible links between explanatory variables

(Hardenberg & Gonzalez-Voyer, 2013). Following the rule of thumb of Burnham and Anderson (2003), we selected the best models with $\Delta\text{CICc} < 2$.

TABLE 3 Best model(s) selected and related path coefficients (=standardized regression coefficients) between the variables for the BVG dataset (France) and for different subsets of species of increasing abundance (based on A which is the log of species abundance (individuals/m²))

n (number of species)	122	Subsets of weed species					
		119	115	113	101	91	88
Abundance threshold	All	A > -2	A > -1.5	A > -1	A > -0.5	A > 0	A > 0.5
Best model(s) selected							
#1	-	-	-	-	-	3	2
#9	1	1	1	1	-	2	-
#10	2	2	2	-	1	-	-
#12	-	-	-	-	-	1	1
#13	-	-	-	-	-	-	-
Regional frequency							
Local abundance	0.27	0.20	0.19	0.18	0.09	-	-
Spe. arable habitat	0.47	0.52	0.52	0.52	0.48	0.49	0.49
Spe. crop type	-0.45	-0.44	-0.44	-0.46	-0.52	-0.50	-0.50
Local abundance							
Regional frequency	-	-	-	-	-	-0.23	-0.25
Spe. arable habitat	0.33	0.33	0.33	0.34	-	-	-
Spe. crop type	0.18	0.28	0.27	0.35	-	-	-
Ellenberg-N	-0.08	-0.03	-0.05	-0.11	-0.11	0.03	0.04
Ellenberg-L	0.12	0.11	0.11	0.04	0.18	0.15	0.09
Flowering duration	-0.30	-0.30	-0.26	-0.28	-0.11	-0.23	-0.23
Flowering onset	-0.33	-0.36	-0.32	-0.43	-0.23	-0.36	-0.37
Wind dispersal	-0.37	-0.22	-0.26	-0.12	-0.31	-0.51	-0.47
Seed mass	-0.11	-0.08	-0.08	-0.08	0.01	-0.01	-0.04
Plant height	-0.01	-0.04	-0.09	0.07	-0.12	-0.11	-0.11
SLA	0.24	0.24	0.24	0.16	0.15	0.18	0.22
Specialization to arable habitat							
EllenbergN	0.24	0.26	0.26	0.29	0.31	0.37	0.37
Ellenberg-L	0.16	0.15	0.15	0.19	0.19	0.15	0.14
Flowering duration	0.04	0.04	0.05	0.06	0.19	0.15	0.15
Flowering onset	-0.12	-0.11	-0.11	-0.11	0.01	-0.01	-0.02
Wind dispersal	0.13	0.19	0.19	0.36	0.41	0.29	0.22
Seed mass	0.01	0.03	0.03	0.03	0.06	0.03	0.03
Plant height	0.04	0.04	0.04	0.02	-0.01	0.01	0.00
SLA	0.24	0.22	0.22	0.20	0.14	0.14	0.14
Specialization to crop type							
Flowering duration	-0.07	-0.08	-0.08	-0.09	-0.12	-0.13	-0.13
Flowering onset	0.08	0.07	0.07	0.06	0.15	0.13	0.14

Note: Best models (with $\Delta\text{CICc} < 2$) are ordered: '1' = best model, '2' = second best model, '3' = third best model; '-' indicates that the model or the variable has not been selected. Values in bold character are significant ($p < 0.05$).

2.3.3 | Acknowledging phylogenetic relatedness

In classical path analyses, coefficients between variables are estimated by generalized least squares (GLS). In our case, because of phylogenetic relatedness, species do not represent independent data points for such analyses. Some of the relationships identified between regional frequency and local abundance, and between indices of commonness and species attributes, could thus result from inherited similarity among evolutionarily related taxa (Harvey & Pagel, 1991). Therefore, we used phylogenetic path analysis

TABLE 4 Best model(s) selected and related path coefficients (=standardized regression coefficients) between the variables for the FSE dataset (UK) and for different subsets of species of increasing abundance (based on A which is the log of species abundance (individuals/m²))

n (number of species)	102	Subsets of weed species					
		94	93	85	69	61	51
Abundance threshold	All	A>-2	A>-1.5	A>-1	A>-0.5	A > 0	A > 0.5
Best model(s) selected							
#1	-	-	-	-	1	1	1
#9	2	1	1	1	-	-	-
#10	-	-	-	-	-	-	-
#12	-	-	-	-	-	-	-
#13	1	-	-	-	-	-	-
Regional frequency							
Local abundance	-	0.19	0.20	0.16	-	-	-
Spe. arable habitat	0.30	0.19	0.19	0.17	0.20	0.18	0.17
Spe. crop type	-0.68	-0.71	-0.71	-0.73	-0.76	-0.77	-0.76
Local abundance							
Regional Frequency	0.71	-	-	-	-	-	-
Spe. arable habitat	0.39	0.69	0.70	0.73	-	-	-
Spe. crop type	0.49	0.20	0.22	0.24	-	-	-
Ellenberg-N	-0.02	0.07	0.07	0.10	0.04	-0.07	0.10
Ellenberg-L	0.09	-0.01	0.00	-0.04	0.17	0.15	-0.18
Flowering duration	-0.30	-0.24	-0.21	-0.14	0.34	0.12	0.08
Flowering onset	-0.22	-0.26	-0.23	-0.13	0.21	0.02	0.03
Wind dispersal	0.14	0.04	0.01	0.06	-0.54	-0.50	-0.56
Seed mass	-0.23	-0.12	-0.10	-0.12	-0.12	-0.16	-0.36
Plant height	0.17	0.02	0.01	0.06	0.33	0.37	0.31
SLA	0.04	0.01	0.01	0.02	0.05	0.12	-0.06
Specialization to arable habitat							
Ellenberg-N	0.20	0.22	0.21	0.25	0.18	0.18	0.21
Ellenberg-L	0.14	0.09	0.10	0.11	0.17	0.15	0.02
Flowering duration	0.53	0.67	0.67	0.79	1.00	1.03	1.15
Flowering onset	0.25	0.38	0.38	0.50	0.69	0.71	0.95
Wind dispersal	-0.09	-0.11	-0.12	0.2	-0.10	-0.18	-0.04
Seed mass	0.05	0.09	0.09	0.09	0.12	0.13	0.17
Plant height	-0.18	-0.21	-0.22	-0.15	-0.09	-0.13	-0.33
SLA	-0.03	-0.06	-0.05	-0.05	0.00	-0.01	0.03
Specialization to crop type							
Flowering duration	-0.45	-0.56	-0.57	-0.73	-0.85	-0.91	-0.93
Flowering onset	-0.16	-0.22	-0.23	-0.38	-0.53	-0.59	-0.56

Note: Best models (with ΔCICc < 2) are ordered: ‘1’= best model, ‘2’= second best model, ‘3’= third best model; ‘-’ indicates that the model or the variable has not been selected. Values in bold character are significant (p < 0.05).

(Gonzalez-Voyer & Von Hardenberg, 2014), hereafter PPA, in which path coefficients were estimated with phylogenetic generalized least squares (PGLS). PGLS is a generalized least squares model where phylogenetic relationships among species influence the covariance of residuals. PPA assesses relationships between variables while controlling for non-independence due to phylogenetic relatedness. The phylogeny was derived from a dated and comprehensive megaphylogeny of spermatophytes, using the function *S.PhyloMaker* provided by Qian and Jin (2016) with scenario 'S3' (See Appendix S2 in Supporting Information for details). We performed all statistical analyses with R version 3.4.3 using library 'phylopath' (van der Bijl, 2018).

3 | RESULTS

The BVG dataset reported 231 species in 1,293 fields across France, and the FSE dataset reported 116 species in 251 fields across the UK. Table S3 in Appendix S3 provides the full species list, with corresponding regional frequency, mean local abundance and ecological specialization for each region. We found a positive relationship between regional frequency and local abundance in both datasets (Figure 2a, 2b). For the 82 species present in both the BVG and FSE datasets, ecological specialization (Pearson's $r = 0.70$, $p < 0.001$) and regional frequency (Pearson's $r = 0.71$, $p < 0.001$) were highly correlated across the two datasets, while local mean abundances

were less but still well correlated (Pearson's $r = 0.32$, $p = 0.002$). Subsequent phylogenetic path analyses (PPA) were restricted to 122 and 102 species that occurred in more than 1% of at least one of the four crop species surveyed in BVG and FSE, respectively.

The best PPA models (with $\Delta\text{CICc} < 2$) were models #9 and #10 for the BVG dataset (Tables 2 & 3, Table S4 in Appendix S4) and models #13 and #9 for the FSE dataset (Tables 2 & 4, Table S5 in Appendix S5). In these three models, functional traits indirectly influenced regional frequency. In model #9 (selected for both datasets), local abundance was determined by both traits and ecological specialization, while regional frequency was determined indirectly by traits through ecological specialization and local abundance (Figure 3). For the FSE dataset, the best model was model#13, which was very similar to model#9, except that local abundance was determined by regional frequency rather than the reverse (Figure 4). Model #10 (second best model for BVG) is similar to model #9 except that local abundance is not determined by ecological specialization (see Figure S6 in Appendix S6). These three models were also most often selected when analysing each crop type separately (model#13:5 times, model#9:3 times, model#10:1 time, see Tables S7 in Appendix S7). In addition, model#12, very close to model#13 (see Table 2), was selected 5 times.

For both datasets, ecological specialization to arable habitats was positively related to local abundance and regional frequency while ecological specialization to crop type was positively related to local abundance but negatively related to regional frequency

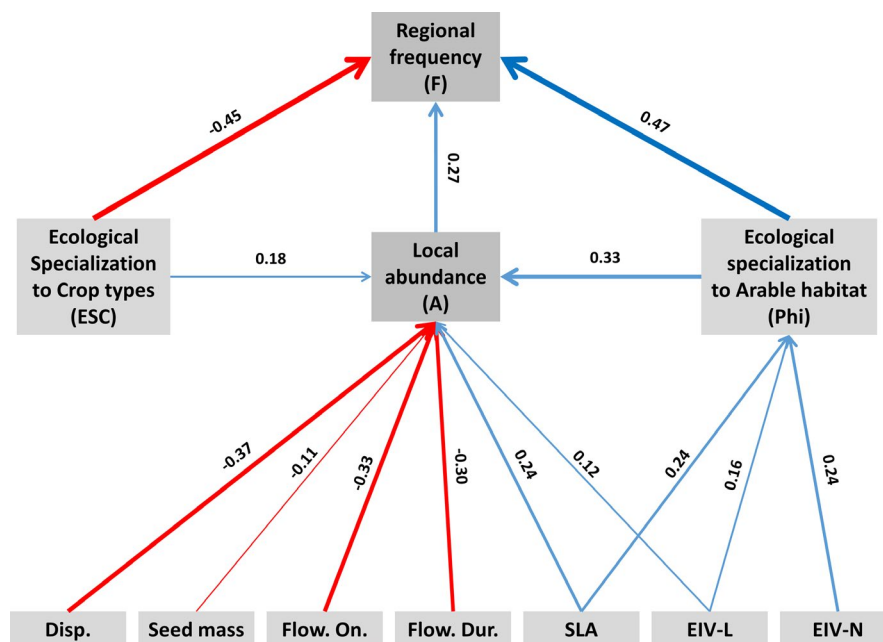


FIGURE 3 Phylogenetic path analysis (PPA) with arrows representing direct and indirect effects of explanatory variables on ecological specialization, local mean abundance and regional frequency of arable weed flora in BVG dataset (France). Blue arrows represent positive effects, and red arrows represent negative effects significant at p values < 0.05 . Line thickness is proportional to the standardized regression coefficients. This figure represents the best-fit model (model #9, see Table 2) according to the CICc-based approach (i.e. a modified version of Akaike Information Criterion) that considers all possible links between explanatory variables. Disp.: wind dispersal, Flow. Dur.: flowering duration, Flow. On.: flowering onset, SLA: specific leaf area, EIV-L: Ellenberg Indicator Value for light, EIV-N: Ellenberg Indicator Value for nitrogen

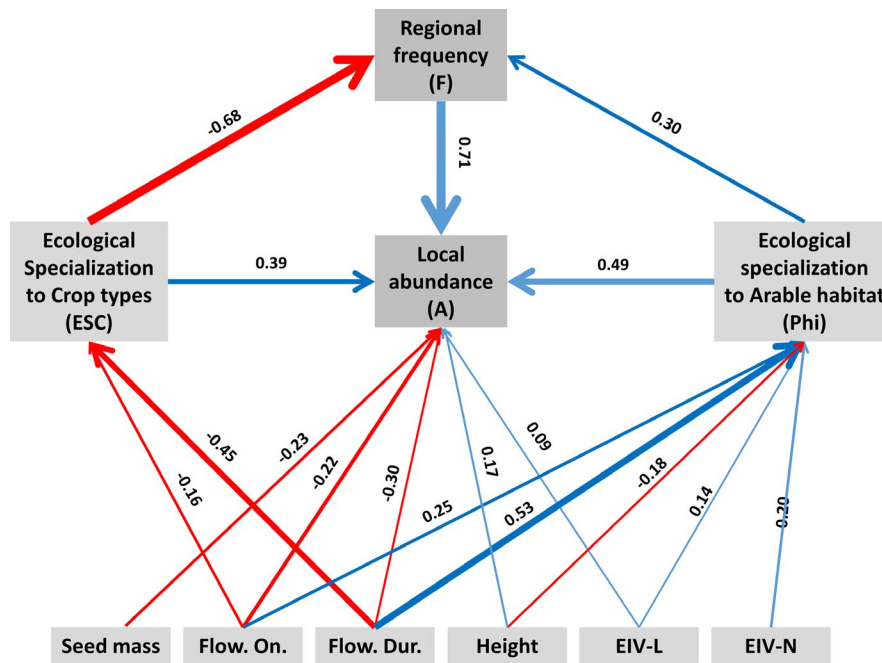


FIGURE 4 Phylogenetic path analysis (PPA) with arrows representing direct and indirect effects of explanatory variables on ecological specialization, on local abundance and regional frequency of arable weed flora in FSE dataset (UK). Blue arrows represent positive effects, and red arrows represent negative effects significant at p values < 0.05 . Line thickness is proportional to the standardized regression coefficients. This figure represents the best-fit model (model #13, see Table 2) according to the CICC-based approach (i.e. a modified version of Akaike Information Criterion) that considers all possible links between explanatory variables. Flow. Dur.: flowering duration, Flow. On.: flowering onset, EIV-L: Ellenberg Indicator Value for light, EIV-N: Ellenberg Indicator Value for nitrogen

(Tables 3 & 4). We performed PPA to subsets of increasingly abundant weeds in the BVG and FSE datasets. We found that models #10 and #9 remained best in BVG dataset, up to a certain threshold of minimal local abundance, although the strength of the AFR continuously decreased (Table 3). Above the threshold, model #12 became the best model and regional frequency became negatively related to local abundance (Table 3, Figure 2c). Likewise, the best models remained #13 or #9 in FSE dataset up to a threshold of minimal local abundance, with similar decrease in the strength of the AFR than in the BVG dataset. Above the threshold, model #1 became the best model meaning that there were no relationship between regional frequency and local abundance of these abundant species (Table 4, Figure 2d).

In the best PPA models for both regions (i.e. #9, #10, #13), a high degree of ecological specialization to arable fields was determined by high values of Ellenberg-N and Ellenberg-L, for both datasets (Figures 3 and 4). In addition, SLA for BVG, and flowering onset and duration for FSE, were positively related to ecological specialization, while plant height was negatively associated in the FSE dataset (Figures 3 and 4). A short flowering period and early flowering were associated with specialization to crop type for the FSE dataset. Higher local abundance was related to high values of Ellenberg-L, early flowering onset, short flowering duration and to low seed mass in both datasets. For BVG, absence of wind dispersal and high SLA were also related to higher local abundance. For FSE, high stature was also positively related to local abundance.

Tables 3 and 4 provide the path coefficients for BVG and FSE analyses, respectively. The trait values explaining weed local abundance slightly differed according to crop type (see Tables S7 in Appendix S7), with, for example, changing influence of flowering onset depending on the sowing date of the crop, that is, positive for spring-sown crops and negative for autumn-sown crops.

4 | DISCUSSION

Our study shows how niche-based processes, through key functional traits, contribute to the interrelationships between local abundance, regional frequency and ecological specialization. We performed original phylogenetic path analyses (Hardenberg & Gonzalez-Voyer, 2013) to assess the role of the ecological attributes while controlling the phylogenetic dependence across species. First, we reported a consistent and positive relationship between regional frequency and local abundance for arable weeds in the two studied regions (France and the UK), and across several crop types (winter wheat, maize, oilseed rape, sugarbeet), in line with previous results in other organisms (Brown, 1984; Gaston et al., 2000). The selected models incorporated both paths where regional frequency influenced local abundance and models where local abundance influenced regional frequency, indicating that there is no preferential directionality in the relationship. Second, we found that the relationship shifted to negative or null when restricted to the locally most abundant

species. The degree of ecological specialization to arable habitat and crop types explained the positive and the negative relationships, respectively, highlighting that ecological specialization is a key factor driving the AFR, although a few studies have explicitly considered its role (see Verberk et al. (2010); Heino and Grönroos (2014)). Third, local abundance was directly determined by functional traits related to resource acquisition, competition and phenology, while regional frequency was not determined by dispersal syndrome but indirectly determined by traits related to specialization to arable habitats such as resource requirement traits for light and nitrogen.

4.1 | Contributions of niche-based processes: specialization and species traits

Our results are consistent with niche-based hypotheses predicting influences of niche breadth (Gaston et al., 2000) and functional traits (Violle et al., 2007). First, our results support the 'jack-of-all-trades, master of none' hypothesis (Figure 1), with higher local abundance for specialist species. Specialization to arable habitat was measured based on the relative frequency of weeds within and outside arable habitat (using independent datasets), while the regional frequency was based on occurrences in arable habitat only. Therefore, specialization and regional frequency represented independent and complementary facets of the ability of weeds to spread in diverse ecological contexts. Habitat generalists were those able to occupy other open habitats, such as grasslands and hedges, which often surround cultivated fields, but our study showed that these generalists were not more frequent in cultivated fields. This result contradicts the 'jack-of-all-trades-master-of-all' hypothesis, which predicts that species able to occupy a wide range of contexts should be both locally abundant and widespread (Brown, 1984). Our results conversely support the idea that habitat specialists (of arable habitat) could be better adapted and more frequent and abundant in their preferred habitat than habitat generalists. As expected, when the regional frequency was measured only in the habitat for which ecological specialization is measured, the combination of the 'jack-of-all-trades, master of none' and the 'resource availability' hypotheses yielded a positive AFR (Figure 1).

We also examined the role of ecological specialization within arable habitat, at the level of crop types, that is, the ability to occur in a single or diverse crop types. The strongest and most consistent pattern was a negative relationship between specialization to crop type and regional frequency, meaning that crop-generalist weeds were regionally more widespread in arable fields confirming the *resource availability* hypothesis (Hanski et al., 1993). This supports the idea that crop-generalist weeds, by being able to thrive in different types of crops, can be regionally more frequent across crop fields (Fried et al., 2010). Conversely, crop-specialist weeds achieved higher mean local abundance than crop-generalist weeds as expected by the *jack-of-all-trade-master-of-none* hypothesis. In this case, this hypothesis explains the albeit weaker negative relationship between regional frequency and local abundance for the most abundant species in the

datasets. This happens when the crop type used by the specialists is limited in area over which regional frequency is measured, limiting its potential distribution range (Gaston, 1996).

All the selected path models indicated that species traits influenced directly local abundance and specialization, but indirectly regional frequency. The effects of individual traits matched most of our expectations: locally abundant species in arable fields had high SLA, small seeds and were early flowering, which is consistent with a ruderal strategy (short life cycle, high fecundity) adapted to frequent disturbances in productive habitats (Grime, 1974). Species with short flowering period and no wind-dispersal mechanisms also had higher local abundance. Non wind-dispersed weeds can invest more into local offspring, enhancing local persistence in habitat specialists, while greater dispersal would be less costly for habitat generalists able to cope with a variety of ecological conditions (Olivieri, Michalakakis, & Gouyon, 1995). Short flowering period is associated with specific timing of flowering and to specialization to particular crop types (at least in FSE), as expected for species that mimic the life cycle of a particular crop (Fried et al., 2010).

No direct relationship was found between traits and regional frequency (Murray, Thrall, Gill, & Nicotra, 2002) contradicting the expectation that regional frequency would be driven by traits related to dispersal syndrome and colonization ability (wind dispersal, seed mass). Rather, owing to a strong relationship between regional frequency and specialization to arable habitat, we found that traits related to specialization to arable habitat indirectly influenced regional frequency. This result establishes new perspectives on the linkage between traits and regional frequency, supporting and expanding the conclusions of Lososová, Chytrý, and Kuhn (2008) that showed that regional frequency was determined by flowering periods, and requirements for temperature and nutrients. The traits determining ecological specialization were only partly the same as those determining local abundance in BVG (e.g. Ellenberg-L and SLA), and in FSE, some traits influenced local abundance and specialization to arable habitat in opposing directions (plant height, flowering onset and flowering duration). In this regard, flowering duration is a key trait with opposed influences on abundance and specialization to a specific crop on the one hand, and on specialization to arable habitat and regional frequency on the other hand. Weeds with a long flowering period can complete their cycle at different times of the year, making them more adapted to diverse crop types (Storkey et al., 2010) and allowing higher regional frequency. Conversely, a short flowering period can be associated with greater performance only in specific crops (Perronne et al., 2015), which could explain the positive association of this trait value to local abundance.

The overall positive AFR can be explained by the contrast between locally rare and narrowly distributed habitat generalists (non-specialists of arable habitat) versus locally abundant and widespread habitat specialists of arable fields (Figure 2a & 2b). The fact that the relationship disappeared or became negative for the most abundant species further underlined that the rare non-specialist species of arable habitat greatly contributed to the overall positive relationship. This is related to a distinctly 'upper triangular' pattern, in which all

combinations of regional frequency and local abundance seem possible except high frequency/low abundance, as already observed for other plant communities (Thompson, Hodgson, & Gaston, 1998). While abundant weed species have trait values adapted to local conditions in arable fields (niche-based processes), non-specialist weed species occurring in arable fields with poorly adapted traits should rely more on spillover from neighbouring habitats (meta-community dynamics) as suggested by the distinction between core persistent species and occasional species (Magurran & Henderson, 2003).

4.2 | Contribution of neutral processes in meta-community dynamics

In the best path models, regional frequency was either a causal factor of local abundance (models #12 and #13), or conversely, regional frequency was determined by local abundance (models #9 and #10). With competing models pointing in opposite directions, it is difficult to conclude on the main directionality of the AFR and both directions can make sense under neutral dynamics (Verberk et al., 2010). First, the influence of regional frequency on local abundance suggests that dispersal dynamics across weed communities can maintain weed pressure and spillover. Weed dispersal plays a key role both in space (exchanges between fields, Bourgeois et al. (2020)) and in time (through the seed bank, Mahaut, Fried, and Gaba (2018)) by reducing the probability of local extinction (rescue effect) and increasing local abundance (Hubbell, 1997). In addition, since specialization to arable habitat was positively related to local abundance, the dispersal dynamic of abundant species from regional to local level is assumed to occur predominantly for weeds that are specialist of arable habitat and, therefore, mainly between arable fields. Second, models where regional frequency depends on local abundance (#9 and #10) suggests that greater local abundance and persistence could also increase the ability to send dispersers and allow a source-sink dynamic. For example, the higher the abundance of a weed in a field, the higher the chance that it will be harvested with the crop and dispersed by the harvester or with contaminated seeds for sowing at larger scales.

4.3 | Conclusions

Our study suggests that the understanding of abundance-frequency relationships in biogeography and macroecology should be improved (i) by integrating the roles of functional traits and ecological specialization in a comprehensive model and (ii) by testing alternative causal models using path analyses. Examining the directions of the abundance-frequency relationship, and the direct or indirect influences of species traits and specialization on abundance and frequency allows better deciphering the contributions of niche-based and neutral processes. We showed that specific trait values influenced directly local abundance of weed species and indirectly regional frequency through trait values associated with specialization to arable habitat. These traits relationships give an insight into niche-based

mechanisms that likely drive variations in species local abundance and regional frequency, and should partly explain the low strength of the relationship (due to distinct trait values and mechanisms affecting abundance and frequency). Finally, testing the abundance-frequency relationship for different subsets of species, with or without rarest species, helps to identify how the latter species contribute to the overall pattern. In arable fields, neutral processes can explain the presence of rare generalist species while niche-based processes explain the success of abundant specialist weeds. Together, these distinct processes generate the positive abundance-frequency relationship observed in arable weeds.








ACKNOWLEDGEMENTS

The authors thank Lisa Norton for the extraction of Country Side surveys data, the DIVGRASS consortium for species distribution data, Christine Meynard, Helen Metcalfe, Lucie Mahaut, Joël Chadoeuf, Vincent Bretagnolle as well as three anonymous reviewers for constructive comments on a previous version of the manuscript.

DATA AVAILABILITY STATEMENT

Regional frequency (Fr, in log), local abundance (A, in log) and fidelity index (Phi) to arable habitat for 122 species in the BVG dataset and 102 species in the FSE dataset are available in Supporting Information. Data on traits and phylogeny can be retrieved from Dryad: <https://doi.org/10.5061/dryad.zw3r2286h>

ORCID

Guillaume Fried  <https://orcid.org/0000-0002-3653-195X>
 Laura Armengot  <https://orcid.org/0000-0002-9820-9667>
 Jonathan Storkey  <https://orcid.org/0000-0003-1094-8914>
 Bérenger Bourgeois  <https://orcid.org/0000-0002-2885-2518>
 Sabrina Gaba  <https://orcid.org/0000-0002-7145-6713>
 Cyrille Violle  <https://orcid.org/0000-0002-2471-9226>
 François Munoz  <https://orcid.org/0000-0001-8776-4705>

REFERENCES

- Bourgeois, B., Gaba, S., Plumejeaud, C., & Bretagnolle, V. (2020). Weed diversity is driven by complex interplay between multi-scale dispersal and local filtering. *Proceedings of the Royal Society B: Biological Sciences*, 287, 20201118. <https://doi.org/10.1098/rspb.2020.1118>
- Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, 124, 255–279. <https://doi.org/10.1086/284267>
- Burnham, K. P., & Anderson, D. R. (2003). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer Science & Business Media.
- Carey, P., Wallis, S., Chamberlain, P., Cooper, A., Emmett, B., Maskell, L., ... Reynolds, B. (2008). Countryside survey: UK results from 2007.
- Chytrý, M., Tichý, L., Holt, J., & Botta-Dukát, Z. (2002). Determination of diagnostic species with statistical fidelity measures. *Journal of Vegetation Science*, 13, 79–90. <https://doi.org/10.1111/j.1654-1103.2002.tb02025.x>
- Clapham, A. R., Tutin, T. G., & Moore, D. M. (1989). *Flora of the British Isles*. UK: Cambridge University Press.
- Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., ... Mouquet, N. (2010). Defining and measuring ecological

- specialization. *Journal of Applied Ecology*, 47, 15–25. <https://doi.org/10.1111/j.1365-2664.2009.01744.x>
- Fried, G., Norton, L. R., & Reboud, X. (2008). Environmental and management factors determining weed species composition and diversity in France. *Agriculture Ecosystems & Environment*, 128, 68–76. <https://doi.org/10.1016/j.agee.2008.05.003>
- Fried, G., Petit, S., & Reboud, X. (2010). A specialist-generalist classification of the arable flora and its response to changes in agricultural practices. *BMC Ecology*, 10, 20. <https://doi.org/10.1186/1472-6785-10-20>
- Futuyma, D. J., & Moreno, G. (1988). The evolution of ecological specialization. *Annual Review of Ecology and Systematics*, 19, 207–233. <https://doi.org/10.1146/annurev.es.19.110188.001231>
- Gaba, S., Perronne, R., Fried, G., Gardarin, A., Bretagnolle, F., Biju-Duval, L., ... Reboud, X. (2017). Response and effect traits of arable weeds in agro-ecosystems: A review of current knowledge. *Weed Research*, 57, 123–147. <https://doi.org/10.1111/wre.12245>
- Gaston, K. J. (1996). The multiple forms of the interspecific abundance-distribution relationship. *Oikos*, 211–220. <https://doi.org/10.2307/3546192>
- Gaston, K. J., Blackburn, T. M., Greenwood, J. J., Gregory, R. D., Quinn, R. M., & Lawton, J. H. (2000). Abundance–occupancy relationships. *Journal of Applied Ecology*, 37, 39–59. <https://doi.org/10.1046/j.1365-2664.2000.00485.x>
- Gonzalez-Voyer, A., & Von Hardenberg, A. (2014). *An introduction to phylogenetic path analysis*. Modern phylogenetic comparative methods and their application in evolutionary biology (pp. 201–229). Springer.
- Gregory, R. D., & Gaston, K. J. (2000). Explanations of commonness and rarity in British breeding birds: Separating resource use and resource availability. *Oikos*, 88, 515–526. <https://doi.org/10.1034/j.1600-0706.2000.880307.x>
- Grime, J. P. (1974). Vegetation classification by reference to strategies. *Nature*, 250, 26–31. <https://doi.org/10.1038/250026a0>
- Hanski, I., Kouki, J., & Halkka, A. (1993). Three explanations of the positive relationship between distribution and abundance of species. In R. E. S. Ricklefs (Ed.), *Species diversity in ecological communities: Historical and geographical perspectives* (pp. 108–116). Chicago: University of Chicago Press.
- Hardenberg, A. V., & Gonzalez-Voyer, A. (2013). Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution: International Journal of Organic Evolution*, 67, 378–387. <https://doi.org/10.1111/j.1558-5646.2012.01790.x>
- Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford University Press.
- Heard, M. S., Hawes, C., Champion, G. T., Clark, S. J., Firbank, L. G., Haughton, A. J., ... Hill, M. O. (2003). Weeds in fields with contrasting conventional and genetically modified herbicide-tolerant crops. I. Effects on abundance and diversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358, 1819–1832.
- Heino, J., & Grönroos, M. (2014). Untangling the relationships among regional occupancy, species traits, and niche characteristics in stream invertebrates. *Ecology and Evolution*, 4, 1931–1942. <https://doi.org/10.1002/ece3.1076>
- Hill, M. O., Mountford, J., Roy, D., & Bunce, R. G. H. (1999) Ellenberg's indicator values for British plants. ECOFACT Volume 2 Technical Annex. Institute of Terrestrial Ecology.
- Hubbell, S. P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, 16, 9–21. <https://doi.org/10.1007/s003380050237>
- Hubbell, S. P. (2001). The unified neutral theory of species abundance and diversity. Princeton University Press, Princeton, NJ.
- Hubbell, S. P. (2004) *Quarterly Review of Biology*, 79, 96–97.
- Ikegami, M., Wandrag, E., Duncan, R. P., & Hulme, P. E. (2019). Length of cultivation determines native and non-native weed richness in crop fields worldwide. *Biological Invasions*, 21, 363–375. <https://doi.org/10.1007/s10530-018-1828-7>
- Jetz, W., Rahbek, C., & Colwell, R. K. (2004). The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, 7, 1180–1191. <https://doi.org/10.1111/j.1461-0248.2004.00678.x>
- Julve, P. (1998). Baseflor. Index botanique, écologique et chorologique de la flore de France. In.
- Keddy, P. A. (1992). Assembly and response rules: Two goals for predictive community ecology. *Journal of Vegetation Science*, 3, 157–164. <https://doi.org/10.2307/3235676>
- Kew, R. B. G. (2018). Seed Information Database (SID). Version 7.1. Retrieved from: <http://data.kew.org/sid/> (May 2018).
- Kleyer, M., Bekker, R. M., Knevel, I. C., Bakker, J. P., Thompson, K., Sonnenschein, M., ... Peco, B. (2008). The LEDA Traitbase: A database of life-history traits of Northwest European flora. *Journal of Ecology*, 96, 1266–1274.
- Köckemann, B., Buschmann, H., & Leuschner, C. (2009). The relationships between abundance, range size and niche breadth in Central European tree species. *Journal of Biogeography*, 36, 854–864. <https://doi.org/10.1111/j.1365-2699.2008.02022.x>
- Krasnov, B. R., Poulin, R., Shenbrot, G. I., Mouillot, D., & Khokhlova, I. S. (2004). Ectoparasitic “jacks-of-all-trades”: Relationship between abundance and host specificity in fleas (Siphonaptera) parasitic on small mammals. *The American Naturalist*, 164, 506–516. <https://doi.org/10.1086/423716>
- Lawton, J. H. (1993). Range, population abundance and conservation. *Trends in Ecology & Evolution*, 8, 409–413. [https://doi.org/10.1016/0169-5347\(93\)90043-O](https://doi.org/10.1016/0169-5347(93)90043-O)
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., ... Gonzalez, A. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, 7, 601–613. <https://doi.org/10.1111/j.1461-0248.2004.00608.x>
- Lloret, F., Medail, F., Brundu, G., & Hulme, P. E. (2004). Local and regional abundance of exotic plant species on Mediterranean islands: Are species traits important? *Global Ecology and Biogeography*, 13, 37–45. <https://doi.org/10.1111/j.1466-882X.2004.00064.x>
- Lososová, Z., Chytrý, M., & Kuhn, I. (2008). Plant attributes determining the regional abundance of weeds on central European arable land. *Journal of Biogeography*, 35, 177–187.
- Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422, 714–716. <https://doi.org/10.1038/nature01547>
- Mahaut, L., Cheptou, P.-O., Fried, G., Munoz, F., Storkey, J., Vasseur, F., ... Bretagnolle, F. (2020). Weeds: Against the Rules? *Trends in Plant Science*. <https://doi.org/10.1016/j.tplants.2020.05.013>
- Mahaut, L., Fried, G., & Gaba, S. (2018). Patch dynamics and temporal dispersal partly shape annual plant communities in ephemeral habitat patches. *Oikos*, 127, 147–159. <https://doi.org/10.1111/oik.04415>
- McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21, 178–185. <https://doi.org/10.1016/j.tree.2006.02.002>
- Metcalfe, H., Hassall, K. L., Boinot, S., & Storkey, J. (2019). The contribution of spatial mass effects to plant diversity in arable fields. *Journal of Applied Ecology*, 56, 1560–1574. <https://doi.org/10.1111/1365-2664.13414>
- Munoz, F., Fried, G., Armengot, L., Bourgeois, B., Bretagnolle, V., Chadoeuf, J., ... Gaba, S. (2020). Ecological Specialization and rarity of arable weeds: Insights from a Comprehensive Survey in France. *Plants*, 9, 824. <https://doi.org/10.3390/plants9070824>

- Murray, B. R., Thrall, P. H., Gill, A. M., & Nicotra, A. B. (2002). How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. *Austral Ecology*, 27, 291–310. <https://doi.org/10.1046/j.1442-9993.2002.01181.x>
- Olivieri, I., Michalakakis, Y., & Gouyon, P.-H. (1995). Metapopulation genetics and the evolution of dispersal. *The American Naturalist*, 146, 202–228. <https://doi.org/10.1086/285795>
- Pandit, S. N., Kolasa, J., & Cottenie, K. (2009). Contrasts between habitat generalists and specialists: An empirical extension to the basic metacommunity framework. *Ecology*, 90, 2253–2262. <https://doi.org/10.1890/08-0851.1>
- Perronne, R., Le Corre, V., Bretagnolle, V., & Gaba, S. (2015). Stochastic processes and crop types shape weed community assembly in arable fields. *Journal of Vegetation Science*, 26, 348–359. <https://doi.org/10.1111/jvs.12238>
- Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, 9, 233–239. <https://doi.org/10.1093/jpe/rtv047>
- Rand, T. A., Tylianakis, J. M., & Tscharnkte, T. (2006). Spillover edge effects: The dispersal of agriculturally subsidized insect natural enemies into adjacent natural habitats. *Ecology Letters*, 9, 603–614. <https://doi.org/10.1111/j.1461-0248.2006.00911.x>
- Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling*, 7, 206–218. https://doi.org/10.1207/S15328007SEM0702_4
- Shipley, B., Vile, D., & Garnier, É. (2006). From plant traits to plant communities: A statistical mechanistic approach to biodiversity. *Science*, 314, 812–814. <https://doi.org/10.1126/science.1131344>
- Shmida, A., & Wilson, M. V. (1985). Biological determinants of species diversity. *Journal of Biogeography*, 12, 1–20. <https://doi.org/10.2307/2845026>
- Siqueira, T., Bini, L. M., Roque, F. O., Marques Couceiro, S. R., Trivinho-Strixino, S., & Cottenie, K. (2012). Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. *Ecography*, 35, 183–192. <https://doi.org/10.1111/j.1600-0587.2011.06875.x>
- Soininen, J., & Heino, J. (2005). Relationships between local population persistence, local abundance and regional occupancy of species: Distribution patterns of diatoms in boreal streams. *Journal of Biogeography*, 32, 1971–1978. <https://doi.org/10.1111/j.1365-2699.2005.01342.x>
- Storkey, J. (2004). Modelling seedling growth rates of 18 temperate arable weed species as a function of the environment and plant traits. *Annals of Botany*, 93, 681–689. <https://doi.org/10.1093/aob/mch095>
- Storkey, J., Moss, S. R., & Cussans, J. W. (2010). Using assembly theory to explain changes in a weed flora in response to agricultural intensification. *Weed Science*, 58, 39–46. <https://doi.org/10.1614/WS-09-096.1>
- Thompson, K., Hodgson, J. G., & Gaston, K. J. (1998). Abundance–range size relationships in the herbaceous flora of central England. *Journal of Ecology*, 86, 439–448. <https://doi.org/10.1046/j.1365-2745.1998.00264.x>
- Tison, J.-M., & De Foucault, B. (2014). *Flora Gallica: Flore de France*. Mèze (France): Biotope Éditions.
- van der Bijl, W. (2018). Phylopath: Easy phylogenetic path analysis in R. *PeerJ*, 6, e4718.
- Verberk, W. C. E. P., Van Der Velde, G., & Esselink, H. (2010). Explaining abundance–occupancy relationships in specialists and generalists: A case study on aquatic macroinvertebrates in standing waters. *Journal of Animal Ecology*, 79, 589–601. <https://doi.org/10.1111/j.1365-2656.2010.01660.x>
- Verma, T., & Pearl, J. (1988). *Influence diagrams and d-separation*. Computer Science Department: UCLA.
- Violle, C., Choler, P., Borghy, B., Garnier, E., Amiaud, B., Debarros, G., ... Viovy, N. (2015). Vegetation ecology meets ecosystem science: Permanent grasslands as a functional biogeography case study. *Science of the Total Environment*, 534, 43–51. <https://doi.org/10.1016/j.scitotenv.2015.03.141>
- Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, 116, 882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>
- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A., & Wright, I. J. (2002). Plant ecological strategies: Some leading dimensions of variation between species. *Annual Review of Ecology and Systematics*, 33, 125–159. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150452>

BIOSKETCH

Guillaume Fried is a plant ecologist interested in trait-based approaches and large-scale analyses of spatial and temporal changes in the structure and composition of arable weed communities. He and the other authors collaborate on the CESAB-FRB DISCO-WEED project which aims at disentangling the role of anthropic disturbances and ecological processes on weed community assembly (see DISCO-WEED project, <https://www.fondationbiodiversite.fr/la-frb-en-action/programmes-et-projets/le-cesab/discoweed/>).

Authors' contributions: GF, FM, JS and CV conceived the ideas; GF, JS, LA and FM worked on the species distribution and traits datasets to extract useful information; GF and LA analysed the data; and GF led the writing with assistance from LA, JS, BB, SG, CV and FM.

SUPPORTING INFORMATION

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

How to cite this article: Fried G, Armengot L, Storkey J, et al. Do ecological specialization and functional traits explain the abundance–frequency relationship? Arable weeds as a case study. *J. Biogeogr.* 2020;00:1–14. <https://doi.org/10.1111/jbi.13980>