Grazing and the vanishing complexity of plant association networks in grasslands
Alexandre Génin, Thierry Dutoit, Alain Danet, Alice Le Priol, Sonia Kéfi

To cite this version:
Alexandre Génin, Thierry Dutoit, Alain Danet, Alice Le Priol, Sonia Kéfi. Grazing and the vanishing complexity of plant association networks in grasslands. Oikos, 2021, 130 (4), pp.541-552. 10.1111/oik.07850. hal-03123198

HAL Id: hal-03123198
https://hal.science/hal-03123198
Submitted on 9 Mar 2021
Grazing and the vanishing complexity of plant association networks in grasslands

Alexandre Génin
Thierry Dutoit
Alain Danet
Alice le Priol
Sonia Kéfi

1 ISEM, CNRS, Univ. Montpellier, IRD, EPHE, Montpellier, France
2 Avignon Université, Aix Marseille Univ, CNRS, IRD, IMBE, Avignon, France.
3 Centre d’Ecologie et des Sciences de la Conservation, CNRS, MNHN, Sorbonne Université, Paris, France.
4 Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
Abstract

Interactions are key drivers of the functioning and fate of plant communities. A traditional way to measure them is to use pairwise experiments, but such experiments do not scale up to species-rich communities. For those, using association networks based on spatial patterns may provide a more realistic approach. While this method has been successful in abiotically-stressed environments (alpine and arid ecosystems), it is unclear how well it generalizes to other types of environments.

We help fill this knowledge gap by documenting how the structure of plant communities changes in a Mediterranean dry grassland grazed by sheep using plant spatial association networks. We investigated how the structure of these networks changed with grazing intensity to show the effect of biotic disturbance on community structure.

We found that these grazed grassland communities were mostly dominated by negative associations, suggesting a dominance of interference over facilitation regardless of the disturbance level. The topology of the networks revealed that the number of associations were not evenly-distributed across species, but rather that a small subset of species established most negative associations under low grazing conditions. All these aspects of spatial organization vanished under high level of grazing as association networks became more similar to null expectations.

Our study shows that grazed herbaceous plant communities display a highly non-random organization that responds strongly to disturbance and can be measured through association networks. This approach thus appears insightful to test general hypotheses about plant communities, and in particular understand how anthropogenic perturbations affect the organization of ecological communities.
Keywords

Plant–plant interactions, Grazing, Association networks, Ecological Networks, Spatial patterns, Co-occurrence
Acknowledgements

The authors warmly thank the staff from the Réserve Nationale des Coussouls de Crau for the numerous discussions and their help with logistics, with a specific thought to L. Tatin and A. Wolff. We also thank the members of the BioDICée team at ISEM along with J. Azuara and C. Kouchner for their help with field work, and D. Pavon (IMBE) for his help with plant identification. AG would like thank T. Daufresne, S. Boudsocq and J. Delarivière from Eco&Sols (INRA) for their help with soil processing, sampling and survey design. The study has been supported by the TRY initiative on plant traits (http://www.try-db.org), maintained by J. Kattge and G. Bönisch (Max Planck Institute for Biogeochemistry, Jena, Germany), and supported by DIVERSITAS/Future Earth and the German Centre for Integrative Biodiv. Research (iDiv) Halle-Jena-Leipzig.

Author contributions

AG led conceptual framing, field work, data analyses and manuscript writing. AD and AlP helped with conceptual framing and field work. SK and TD provided significant support and advice all along the project. All authors contributed substantially to the manuscript writing and editing.
Grazing and the vanishing complexity of plant association networks in grasslands

Abstract

Interactions are key drivers of the functioning and fate of plant communities. A traditional way to measure them is to use pairwise experiments, but such experiments do not scale up to species-rich communities. For those, using association networks based on spatial patterns may provide a more realistic approach. While this method has been successful in abiotically-stressed environments (alpine and arid ecosystems), it is unclear how well it generalizes to other types of environments.

We help fill this knowledge gap by documenting how the structure of plant communities changes in a Mediterranean dry grassland grazed by sheep using plant spatial association networks. We investigated how the structure of these networks changed with grazing intensity to show the effect of biotic disturbance on community structure.

We found that these grazed grassland communities were mostly dominated by negative associations, suggesting a dominance of interference over facilitation regardless of the disturbance level. The topology of the networks revealed that the number of associations were not evenly-distributed across species, but rather that a small subset of species established most negative associations under low grazing conditions. All these aspects of spatial organization vanished under high level of grazing as association networks became more similar to null expectations.

Our study shows that grazed herbaceous plant communities display a highly non-random organization that responds strongly to disturbance and can be measured through association networks. This approach thus appears insightful to test general hypotheses about plant communities, and in particular understand how anthropogenic perturbations affect the organization of ecological communities.
Introduction

Ecological interactions are fundamental bricks of ecological communities and influence key properties of ecological systems such as productivity, stability or response to perturbations (Tilman 1982, Thebault and Fontaine 2010). This is in particular true of plant-plant interactions, which can determine the fate of a given community following environmental perturbations (Verdú and Valiente-Banuet 2008). In some ecosystems, such as drylands (arid and semi-arid grasslands), certain “nurse” plant species improve local environmental conditions around them, and thereby increase the recruitment, growth and survival of other species below or close to their canopy (Callaway 2007). Such facilitative interactions have been shown to influence how ecosystems respond to changes in temperature or biotic stress (Kéfi et al. 2007). Understanding the effect of global changes in environmental conditions thus requires knowledge on how those changes will affect species interactions.

A traditional approach to measure interactions between plant species is to use pairwise experiments (Engel and Weltzin 2008). However, this can only work for a small set of species (Graff and Aguiar 2011, Holthuijzen and Veblen 2016), because the number of experiments required to measure interactions among all the pairs of a diverse set of species is prohibitively high. In addition, pairwise experiments do not always preserve the environmental setting of plant communities, so interactions measured through experiments may be different from those that occur in situ (Engel and Weltzin 2008). For instance, this can happen because of indirect interactions (when an interaction between a pair of species is altered by the presence of another; Wootton 1994, Levine et al. 2017). To overcome this, recent work has sought to revisit an old approach (Vries 1954) and use the spatial associations of pairs of species in a given community as proxies for their interactions (Saiz and Alados 2011). Put intuitively, this approach is based on a principle of co-occurrence: individuals of species that facilitate
each other are expected to be found next to each other more frequently than expected by chance (this is referred to as a positive association hereafter). Conversely, individuals of species that have a net negative interaction, i.e. exhibit interference with each other, should also exclude one another spatially (hereafter, a negative association). Therefore, the way sessile species are spatially organized emerges at least in part from species interactions. However, approaches attempting at recovering pairwise species interactions from spatial associations have been shown to often fail (Freilich et al. 2018, Rajala et al. 2019, Blanchet et al. 2020). This may be because intra-specific variability in interaction strengths (Guimarães 2020), or confounding factors affect the patterns of spatial associations (e.g. shared habitat preferences, D’Amen et al. 2018), but also because we often lack the statistical power to estimate with precision the high number of pairwise associations in diverse communities (Rajala et al. 2019). Nonetheless, recent research has shown that aggregated community-level properties such as the balance between negative and positive interactions could be predicted with some accuracy using spatial association patterns (Barner et al. 2018). The analysis of associations thus could, if not precisely predict pairwise interactions, be a promising way of coarsely evaluating the general structure of interaction networks in plant communities.

Another interesting aspect of the association approach is the use of networks. The structure of a community as captured by pairwise associations can be represented as a network, where nodes are species, and links represent the significant spatial associations between pairs of species. A sign can be attributed to each link, either positive when two species tend to cluster in space (for positive associations), or negative when they tend to segregate in space (for negative associations). The absence of a link between two nodes in such network materializes the absence of any significant spatial relationship between two species. Such signed networks are often referred to a “spatial association networks” or simply “association networks” (used hereafter), as well as “spatial ecological networks” (e.g. Saiz et al. 2018) or “co-occurrence networks” (Delalandre and Montesinos-Navarro 2018).
Similar to food webs, they provide a unified way to describe the overall structure of a plant community. For example, the importance of negative associations in a community is formally captured by the proportion of negative links in its association network (the proportion of negative associations). The topology of the networks (which species is linked to which) also provides essential information on functioning: in a community, species that have many positive links with others could allow identifying nurse species which facilitate many others; such presence of nurse species is expected to lead to a high heterogeneity in the number of positive associations across species (Saiz et al. 2018). Conversely, in a plant community where interactions are more symmetrical, such heterogeneity in the number of significant associations per species should be lower. Association networks thus provide a rich description of the structure of plant communities. Although they do not replace experimental measurements of interactions, they provide a glimpse into the structure of communities, and an opportunity to improve the current knowledge regarding the drivers of community functioning (Losapio et al. 2019).

For many ecosystems, forecasting the upcoming ecological changes of the next decades requires understanding how species interactions drive ecological dynamics. However, this is rendered difficult by the fact that interactions themselves change with environmental conditions (e.g. precipitation, Tielbörger and Kadmon 2000), abiotic stress levels (Callaway et al. 2002), or disturbance regimes (Saiz and Alados 2012). Anticipating the upcoming changes in ecological systems, and plant communities in particular, thus require the knowledge of how environmental conditions affect interactions, which can be done by focusing on the way communities change along spatial gradients of increasing stress.

A major source of biotic stress in plant communities is grazing, which causes recurrent disturbance that drives not only taxonomic and functional composition (Díaz et al. 2007), but also the structure of species interactions. Grazing effects on association networks have mostly been studied in drylands so
far, probably because of the striking spatial structure of the vegetation there, which is organized in vegetation patches rather than homogeneously spread out in space (Verdú and Valiente-Banuet 2008). This specific spatial structure arises from shrubs providing protection against grazing for other plants, an interaction known as “associative protection”. This mechanism has been extensively confirmed by pairwise experiments outside our study area (Baraza et al. 2006, Graff et al. 2007, Graff and Aguiar 2011), and within it to some extent (Buisson et al. 2015). It has been shown to strongly influence the structure of association networks by producing a unimodal pattern of variation of positive associations along grazing gradients (Saiz and Alados 2012). Associative protection dominates at intermediate levels of grazing, but is absent either when grazing is low and net negative interactions between neighbours are most common (interference), or when grazing is very high, and disturbance is too high for interactions to matter since grazers eat and trample the majority of the plant biomass. However, it is unknown whether such mechanisms would be important enough to drive the structure of spatial association networks in milder climates, and among herbaceous plants.

To help reduce this knowledge gap, we studied how grazing affected the topology of plant association networks in a Mediterranean grassland. Contrasting with previous studies, our work focuses on the variations of associations in fully-herbaceous communities without shrubs, more constrained by grazing, less by climate, and with a majority of annuals, which have often been ignored in association studies (e.g. Soliveres and Maestre 2014). We investigated the effect of sheep pressure on several key structural aspects of association networks that have been shown to vary in previous studies (Saiz and Alados 2012, Saiz et al. 2018): (i) the total number of associations (i.e. whether plant communities are strongly spatially-structured by associations); (ii) the importance of positive and negative associations and (iii) the heterogeneity of association networks (i.e. a measure of the variability of the number of associations across species; Saiz et al. 2018). Based on seminal hypotheses regarding plant interactions, we hypothesized the following regarding the trends in the proportion of association types along the
H1. at low grazing pressure, interference between plant species dominates and the number of negative associations is therefore at its maximum along the gradient (Michalet et al. 2006):

H2. at intermediate grazing pressure, positive associations should be at their maximum due to the possible presence of associative protection (Maestre et al. 2009, Smit et al. 2009)

H3. at high grazing pressures, where biotic disturbance levels are the highest, both negative and positive interactions should have a low impact on spatial structure, and thus there should be fewer significant plant associations, both negative and positive (Graff and Aguiar 2011)

After testing these hypotheses and describing the changes in the structure of the association networks, we discuss how these can be used to improve the assessment of plant community structure.
Methods

Sampling site

We carried out this work in a Mediterranean dry grassland located in the south of France, La Crau (longitude 4.882 E; latitude 43.548 N), where vegetation has been grazed by sheep for at least 2000 years (Badan et al. 1995). Plant communities are diverse and mostly composed of herbaceous plants (on average 30 ± 12 species/m² in this study, mean ± s.d.), with very few perennial higher shrubs, leading to an outstanding steppe-like landscape. Grazing in La Crau is intense relative to other systems and has been centered around numerous sheepfolds for centuries (Saatkamp et al. 2020), around which ruderal nitrophilous species dominate as a response to the high grazing and trampling pressure and soil nutrient levels (Figure (Devaux et al. 1983). Further from sheepfolds, typical dominant species include *Brachypodium retusum* P. Beauv. with interspersed *Thymus vulgaris* L. and *Asphodelus ayardii* Jahand. & Maire (Molinier and Tallon 1950, Buisson and Dutoit 2006). This very-well defined gradient constitutes an ideal “natural experiment” setting, as around sheepfold the effects of grazing pressure strongly dominate over other environmental characteristics (Figure 1 and Supplementary Information 1).

Surveys

We selected six sites, i.e. six sheepfolds, spread out in the study area of 10 500 ha of a steppe-like ecosystem (Buisson and Dutoit 2006). Surveys were carried out during the plant growing season (April 15 – June 10), which is the time of the year where approximately 80% of the species can be detected and identified. This was done during the years 2016 and 2017. The survey period was comprised within the grazing season, which is from February to June.
Computing species spatial associations requires characterizing the spatial structure of plant communities, which we did using pairs of 5-meter long transects. The grazing gradient is known to produce a change in vegetation composition, which has been classified into successive types that form “belts” around a given sheepfold (Molinier and Tallon 1950, Buisson and Dutoit 2006). Around the sheepfold, species that are most nitrophilous or grazing-resistant predominate, such as Urtica sp. and Onopordum illyricum L. (Figure 1a, red and yellow areas). Moving away from the sheepfold, these are replaced by plant assemblages in which Trifolium spp. dominates (Figure 1a, green areas), until the last type of community is reached, with high covers of Brachypodium retusum (Figure 1a, white dashed areas; Supplementary Information 1). This latter type constitutes most of the vegetation in the study area and represents communities where grazing pressure is at its lowest.

We identified each belt in space based on a preliminary quadrat-based survey, and placed the transects in the different vegetation belts to document a wide range of grazing levels (Supplementary Information 1). We laid out two replicate transects in each belt, separated from at least 10m but no more than 75m so that we could consider that the two transects were under similar levels of grazing pressure. The starting point and direction of each transect were chosen randomly by throwing a pen. To make sure that no environmental factor would affect spatial associations (e.g. two species co-occurring because they share a similar microhabitat), we repeated the throw whenever the environmental conditions appeared to vary along a transect (e.g. when vehicle tracks, water puddle, ant nest or any other visible heterogeneity was present). For each transect, we laid out a measuring tape, and recorded the length over which every part of plant individuals overlapped (Figure 2). Individual positions could be determined with an accuracy of 2 mm, as estimated from repeated measurements: this very fine spatial grain was chosen to allow the detection of positive and negative associations alike (Araújo and Rozenfeld 2014) as it was close to the typical length over which plants intersected the transect (2 to 3 mm). At some sites, some intermediate vegetation belts were absent or not identifiable in the field and
were therefore not sampled. As a result, one site had two missing belts (i.e. two pairs of transects), and two others had one belt missing (i.e. one pair of transects). In two sites, we carried out the sampling both in 2016 and 2017 to increase our sample size, yielding a final dataset with 2 to 8 pairs of transects per site.

Measuring species associations

Our statistical analyses comprised two broad steps. First, we computed the association networks (one association network for each pair of transects) and summarized them into community-level metrics (e.g. the total number of negative links). Then, we investigated how these community-level metrics changed along the grazing gradient.

Computing networks and community-level metrics

Computing association networks can be done by defining a metric of association between two species, and assigning a positive (resp. negative) link when the metric for the pair of species is above (resp. below) a reference threshold (Sanderson and Pimm 2015). Here, we used the total overlap between individuals of a pair of species to measure pairwise association (this is equivalent to using the number of co-occurrence when using discrete sampling). In more formal terms, for a given species $i$ intersecting $n$ times in a transect, and a species $j$ intersecting $m$ times, the total overlap $O_{ij}$ within a transect is computed as follows:

$$O_{ij} = \sum_{l=1}^{n} \sum_{k=1}^{m} \int_{x=0}^{x=L} I_{lk}(x)dx$$

where $L$ is the length of the transect and $I_{lk}(x)$ is equal to 1 when both individual $l$ and $k$ are present at position $x$ along the transect, and 0 elsewhere (Figure 2). As we sampled two transects per vegetation belt, which can be considered as replicates, we computed $O_{ij}$ for both transects, and used their total as the metric of association between pairs of species. It is worth noting that $O_{ij} = O_{ji}$ for any $i$ and $j$, so
For each pair of species, the above procedure yields an observed value of total overlap. It remains to be determined whether this value is significantly high or low, i.e. whether species significantly cluster or segregate in space. To do so we used a randomization-based approach that compared the observed patterns to that of a null expectation with the same total cover for each species, but random spatial structure. We took the observed pair of replicate transects and randomized the position of each segment over which plant individuals intersected the transect (one rectangle in Figure 2b). We did so by placing all segments at a new, random position in the pair of replicate transects (a given individual segment could thus occur in one of the two replicate transects, but be in the other one after randomization). New random positions were redrawn if necessary so that no segment in the random transects would extend past the last segment in the observed transects. Put graphically, no segment would be placed past the rightmost rectangle in Figure 2b. We generated 1999 pairs of such random transects, and recomputed the total overlap $O_{ij}$ between each species, thus obtaining a null distribution of total overlap for each species pair. Pairwise associations were then classified into positive, negative or neutral by using a cutoff value $\alpha$ applied to the null distribution. A positive association was considered to occur between a pair of species when $1 - \alpha/2$ % of the null distribution was below the observed value. Conversely, a negative association was retained when $\alpha/2$ % of the null distribution was above the observed value. We used a threshold of 25 % (i.e. $\alpha = 0.25$), so for example a positive association was retained when the observed overlap was above 87.5 % of the null distribution. The value of $\alpha$ represents a tradeoff – lower values produce association networks in which species exhibit strong spatial patterns (spatial aggregation or segregation), but the resulting networks have very few links, so community-level metrics may be unreliable. As the choice of such cutoff value may affect the results, we carried out a sensitivity analysis to make sure that the results were robust to a range of thresholds (Supplementary Information 3).
For each pair of transects, this first step of the analysis yields a set of $A^+$ significant positive and $A^-$ significant negative associations. This set can be thought as a network where nodes constitute species and positive and negative associations are positive or negative links between them (Figure 1, b1). We summarized these networks into six community-level metrics:

- $K^+$, and $K^-$, the fraction of positive and negative links in the network, computed respectively as $A^+/A_{\text{max}}$ and $A^-/A_{\text{max}}$, where $A_{\text{max}}$ is the maximum possible number of links, $S(S - 1)/2$ for an undirected network with $S$ species and no self-interaction

- $L$, the total fraction of links computed as $(K^+ + K^-)$, which measures the density of links in the association network, and thus the general importance of associations on community structure

- $R$, the ratio of positive to negative links, defined as $(K^+ - K^-)/(K^+ + K^-)$ (Díaz-Sierra et al. 2017)

- $H^+$ and $H^-$, measuring the heterogeneity of associations, as defined in Estrada 2010. This index captures how close a network is to a star-graph, i.e. whether only one outlying species establishes all links (maximum index value of one), or whether all species have the same number of links (minimum value of zero). Such index is well-suited to assess heterogeneity in small networks for which distribution-fitting is unreliable (Estrada 2010). This index is computed as follow:

$$H = \frac{\sum_{i,j\in E} (k_i^{0.5} - k_j^{0.5})^2}{N - 2 \sqrt{N - 1}}$$

where the sum is done over the set of all edges in the network $E$. $i$ and $j$ describe the nodes involved in an edge, with their corresponding degrees $k_i$ and $k_j$. $N$ here is the total number of nodes with degree non-zero (i.e. that have at least one link). We computed this metric on networks made up only of positive links (for positive associations) or negative links (for
negative associations), which yielded $H^+$ and $H^-$.  

**Cover-corrected community-level metrics**

Despite being taken into account in the null model to derive pairwise associations, species covers still have an effect on the community-level metrics defined above. To illustrate this, we can consider the example of rare species. By definition, pairs of rare species have low covers, hence placing them randomly in a transect will yield a null distribution of overlap mostly composed of zeroes. Because the overlap cannot be negative, such species pair cannot have an observed overlap below the random expectation, i.e. a negative association. As a result, as the number of rare species changes along a gradient, the proportion of negative associations ($K^-$) will vary, masking the biologically-relevant changes in the spatial behavior of species. A similar dependence on species’ cover distribution affects all community-level metrics ($L$, $K^+$, $K^-$, $R$, $H^+$, and $H^-$), so it is necessary to control for this effect to make meaningful statements about changes in plant associations along a gradient (Tylianakis and Morris 2017, Pellissier et al. 2018, Saiz et al. 2018). This bias is further detailed in the Supplementary Information 4.

We corrected community-level metrics using the following procedure (Figure 1b). For each pair of transects, we created 999 'null' pairs, with randomized positions of individuals (mixing individuals between transects, i.e. carrying the same randomization as in the first step of the analysis). We built association networks using the same procedure, this time starting with the randomized pairs of transects. For each pair of replicate transect, we obtained this way 999 association networks that represent the null expectation of random spatial structure, but similar species total covers. We then computed the community-level metrics on the null networks to obtain, for each pair of replicate transects, 999 null values for each metric. For example, we obtained null values for $K^-$ (the proportion of negative associations) for each pair of replicate transects. We then computed the deviation of the
observed value to its null distribution, using a z-score (or standardized effect size, SES; following example given for $K^-)$:

$$K_{\text{SES}}^- = (K^- - \mu_{\text{null}})/\sigma_{\text{null}}$$

where $\mu_{\text{null}}$ is the mean of the null distribution for the community-level metric, and $\sigma_{\text{null}}$ is its standard deviation Figure 1, b3). 999 null replicates were considered enough to estimate with reasonable precision these two parameters. By focusing on the changes in this z-score instead of the raw value of each community-level metric, we measured the changes in a given aspect of the association network, controlling for the effect of changing species covers along the gradient (Tylianakis and Morris 2017, Pellissier et al. 2018).

**Characterizing the grazing gradient**

In La Crau, the local grazing pressure (the number of sheep*time unit.ha$^{-1}$) is known to decrease with the distance to the sheepfold, but this relationship has not been explicitly measured for most sites. Previous studies (Dureau 1998; Supplementary Information 2) suggest that it decreases as the inverse of distance then reaches a minimum, i.e. that grazing pressure $G$ can be described by the following function:

$$G(x) = G_{\text{inf}} + \frac{G_0 - G_{\text{inf}}}{1 + \tau x} \frac{\tau G_0 + xG_{\text{inf}}}{\tau + x} \quad (\text{equation 1})$$

where $x$ is the distance to the sheepfold, $G_{\text{inf}}$ is the grazing pressure far from the sheepfold, $G_0$ is the grazing pressure at the sheepfold, and $\tau$ characterizes the spatial extent of the gradient.

Soil characteristics are expected to follow closely the grazing pressure. This is particularly true for soil nutrients that are spatially redistributed as grazers feed in a given area and defecate in another (Steinauer and Collins 1995, Selbie et al. 2015). In our area, we thus expected the variations in the amount of nutrients in the soil to closely match those in grazing pressure. We therefore carried out
sampling to measure soil properties as a function of the distance to the sheepfold. We sampled soil wherever a pair of transects was carried out, and complemented these samples with additional ones where the grazing gradient was very extended (see Supplementary Information 1). Soil was sampled by mixing subsamples of the upper first five centimeters (after scraping sheep dung) from a circular area of radius 4 m. Soil sampling took place in October 2017, when sheep flocks were absent.

We used the total Nitrogen content ($N_{tot}$) in the soil (as measured by the Kjeldahl method; Bremner 1960) to characterize the grazing pressure at a given distance from the sheepfold. We assumed that the total $N$ measured from soil samples was linearly-related to the (unmeasured) grazing pressure, i.e. that $N_{tot} = aG(x) + b$ (with $a$ and $b$ being the slope and the intercept, respectively). Based on equation 1, this yields the expected relationship between distance $x$ and nitrogen content $N_{tot}$, which we fitted to the empirical data:

$$N_{tot}(x) = \frac{\tau N_{tot,0} + x N_{tot,inf}}{\tau + x} \quad \text{(equation 2)}$$

Because we had a reduced number of samples per site (2 to 4) and standard non-linear regression was prone to overfitting, we carried out the regression using a Bayesian setting and weakly conservative priors. We used the site as a random effect on the estimates of $\tau$, $N_{tot,0}$ and $N_{tot,inf}$ along with a Gaussian residual error (see Supplementary Information 2 for details on parameter estimation).

We checked that the estimated values for $N_{tot,0}$ and $N_{tot,inf}$ were consistent with previous reported values for the area (Römermann et al. 2005), and then used the estimated $N_{tot}$ in the soil at the given distance to the sheepfold as a proxy for grazing pressure.

**Effect of grazing pressure on associations**
We characterized the effect of grazing on species associations by investigating the trends of community-level metrics along the gradient of grazing pressure (as measured by $N_{tot}$). To test whether the trends were linear or unimodal (hump-shaped), we fitted both a straight-slope and a second-order polynomial model, with the focal metric as response (e.g. $R_{SES}$), the estimated $N_{tot}$ as predictor and the site as a random effect on the trend coefficients. We carried out this regression using a Bayesian setting with uninformative priors, and retained the model that had the lowest error when carrying leave-one-out (loo) cross-validation, following recommendations from Vehtari et al. (2017). We used normal priors with mean zero and standard deviation 50 for regression coefficients, and an exponential prior with rate 0.1 for the residual error parameter.

All analyses were carried out in R v4.0.2 (R Core Team 2020), with regressions carried out using the package loo v2.3.1 (Vehtari et al. 2017) and brms v2.14.0 (Bürkner 2017). Data and code used for the analyses are available at https://datadryad.org/stash/dataset/doi:10.5061/dryad.ns1rn8prd [private during peer-review]. Trait data to compute the percentage of ruderal species was obtained from the TRY database (Kühn et al. 2004, Kattge et al. 2020).
Results

We found that the grazing gradient around sheepfolds had a strong effect on soil conditions, as total nitrogen of the soil was doubled (model estimation of $1.35 \pm 0.36 \text{ g.kg}^{-1}$ far from sheepfolds vs. $2.85 \pm 0.27 \text{ g.kg}^{-1}$ at the entrance; Figure 3a). Such values are consistent with what has been previously reported in the study area (Römermann et al. 2005, Saatkamp et al. 2020). The effects of grazing on soil surface conditions were evidenced by the increase in bare ground from $20 \pm 7\%$ to $42 \pm 14\%$ (Figure 3b) and dung cover from $9 \pm 5\%$ to $92 \pm 12\%$ (Figure 3c). Grazing favored higher covers of ruderal species ($17 \pm 8\%$ to $63 \pm 5\%$, Figure 3d) and reduced species richness ($49 \pm 6$ species far, vs $10 \pm 1$ close to sheepfolds; Figure 3e).

Regarding plant association metrics, model selection favored linear relationships (straight lines) for all relationships between association-based community-level metrics and grazing pressure (Figure 4,5). In particular, we found no support for unimodal variations of the number of positive associations $K_{SES}^+$ along the gradient (Figure 4c).

Grazing was found to increase the number of links in plant communities ($L_{SES}$, Figure 4a), as positive associations decayed quicker than negative associations with grazing (Figure 4b, 4c). Overall, there was an excess of negative associations compared to positive associations ($K_{SES}^-$ was always above zero and $K_{SES}^+$ almost always below zero; Figure 4b, 4c). Both negative and positive associations were found to approach the null expectation (i.e. a value of $K_{SES}^-$ and $K_{SES}^+$ close to zero) as sheep pressure increased (Figure 4b, 4c). Consistent with these patterns, the association ratio $R_{SES}$ was always negative regardless of the grazing level, and close to the null expectation where grazing was the highest (Figure 4d). This suggests that these communities were strongly structured by spatial associations at low grazing pressure, but became increasingly close to the null expectation with increased grazing.
The heterogeneity in negative associations $H_{\text{SES}}$ decreased with grazing pressure (Figure 5a). In other words, plant species of a given community had similar numbers of negative associations with each other under high grazing conditions and more variable ones under low grazing. We found no equivalent effect of grazing on the heterogeneity of positive associations $H_{\text{SES}}^+$ (Figure 5b).

All the above trends were robust to the choice of $\alpha$ (Supplementary Information 3).
Discussion

Using continuous transects documenting the spatial associations of plant individuals in a Mediterranean dry grassland, we found non-random trends in spatial associations along the grazing intensity gradient. Negative associations were found to dominate over positive ones under low grazing, with a highly-variable number of negative association per species. Grazing was found to reduce all these characteristics to their null expectations, overall making the communities less structured by interspecific associations at high grazing levels.

Association trends

Changes in association trends

We found that negative associations dominated where grazing pressure was lower (H1), which is consistent with interference driving the assembly of communities (Graff et al. 2007). As grazing pressure increases, it appears that this imprint of interference on spatial patterns vanishes. This is consistent with the known effects of grazing in grasslands: as biomass is removed from dominant species, ground-level light increases and thus spatial exclusion by taller competitors is reduced (Borer et al. 2014, Odriozola et al. 2017). However, such decrease in negative interactions is usually associated with an increase in species richness that we did not observe here (Figure 3e). This suggests that the trends observed here reflect more than a sole reduction in interference. Given the very high disturbance levels around sheepfolds (more than 1000 sheep being present daily close to sheepfolds), another factor altering spatial patterns is the direct reduction in covers due to the sheep presence (Adler et al. 2001, Graff and Aguiar 2011). Disturbance may be too high to allow plants to grow enough during the grazing season to preempt space and produce significant spatial patterns (Alados et al. 2004). Disentangling this direct effect of disturbance on spatial associations from that of the processes occurring within the plant community would require additional independent evidence, and constitutes a
limit of relying only on association networks.

Positive associations exhibited no unimodal trend along the grazing gradient (H2). Instead, we found that they declined linearly with increasing grazing. This contrasts with previous work reporting unimodal trends along grazing gradients both experimentally (Smit et al. 2009) and using association (Saiz and Alados 2012, Le Bagousse-Pinguet et al. 2012). A possible explanation for this pattern is that our study was carried out in areas where the grazing pressure is always relatively high compared to other work using association networks. Stocking rates in La Crau often reach above 1 ind.ha\(^{-1}\) (Buisson et al. 2015), well above the maximum rate of 0.64 ind.ha\(^{-1}\) reported by Saiz et al. (2012). It is thus likely that we could only document the reduction in the importance of interactions under high disturbance, as suggested by Graff et al. (2013). Moreover, positive associations were found to be less frequent than expected by chance at all grazing levels, suggesting that facilitation is rare in our system.

This general absence of facilitation, despite some existing experimental evidence (Buisson et al. 2015), could emerge from the particular morphology of the plants in the study area. The study area lacks tall woody shrubs which are a major source of positive associations, as they often act as nurse species under which herbaceous plants grow in dry areas (Saiz and Alados 2012). Compared to the study of Saiz and Alados (2012), at our sites, plants are more functionally similar (forbs and grasses of similar heights), which limits the possibility of plants to aggregate in space by growing on top of, or below, each other, and thus the proportion of positive associations. This could explain in particular why the species *Brachypodium retusum*, which exhibits mostly negative associations in our work, was found to engage mostly in positive ones in Saiz and Alados (2012). By reducing the necessary complementarity in height and morphology of the plants for spatial aggregation, i.e. increasing functional similarity, grazing could reduce the prevalence of positive associations in our study area.

**Heterogeneity in associations**
While plant ecology has mostly focused on describing the relative importance of different types of interactions (e.g. the relative importance of facilitation vs. interference; Callaway et al. 2002, Michalet et al. 2006, Soliveres et al. 2015), it has seldom focused on the structural aspects of plant-plant interaction networks. For example, facilitation in abiotically-stressed environments can be asymmetric, with key nurse plants facilitating many others, or more symmetric, where plants with similar lifeforms buffer each other against harsh conditions (Lin et al. 2012). In our study, such asymmetry was found in negative associations. We found that they were more heterogeneous when grazing was low. This suggests that under low grazing conditions, inter-specific interference is asymmetric, with few species establishing a high number of negative links, and does not arise from all plants being equally more competitive. This maps well onto limiting-similarity acting where competition for resources dominates (Chesson 2000), which favors functional divergence in traits (Valiente-Banuet and Verdú 2008) and plant strategies. Some plants may engage in “Competitive confrontation” (sensu Novoplansky 2009) by maximizing interference with their neighbors, while other may display “Competitive avoidance” or tolerance and avoid such behavior (Gruntman et al. 2017). This diversity in plant strategies may be a factor explaining the heterogeneity in negative links. As grazing increases, communities converge towards grazing-tolerant strategies (Carmona et al. 2012 and Figure 3d), resulting in the observed reduction of network heterogeneity. Such hypotheses could be tested with theoretical models exploring the link between plant traits and strategies, and the structure of interaction networks in plant communities. More generally, this line of thought could also be extended to other network properties such as modularity, nestedness (Verdú and Valiente-Banuet 2008) or structural balance (Saiz et al. 2017). Research on food webs has produced several of such theoretical models of network assembly from species traits (e.g. Williams and Martinez 2000): equivalent work for plant association networks is still burgeoning (Lin et al. 2012).
Moving forward with associations

Using associations as proxies to estimate plant-plant interactions seems to have been deemed a reasonable solution for plants in arid drylands, as evidenced by the breadth of papers relying on this method (e.g. Verdú and Valiente-Banuet 2008, Saiz and Alados 2012, Delalandre and Montesinos-Navarro 2018). This may be motivated by the fact that facilitation (e.g. between nurse plants and their protécted species, a type of interaction which is particularly frequent in arid ecosystems) has been shown to be well-captured by association patterns (Freilich et al. 2018). However, our work suggests that plant-plant associations provide valuable information in other ecosystems as well. Strengthening this approach further would require additional evidence to assess whether interactions in a community effectively give rise to the expected association patterns. Such piece of evidence could come from independent trait data (Soliveres et al. 2014). For example, increased functional similarity between two plants could be associated with a higher likelihood of observing a negative association between them (Conti et al. 2017). Another avenue to put the association approach to the test would be to rely on experimental approaches directly measuring interaction strengths (Choler et al. 2001). Given the ballooning number of studies based on plant associations (Losapio et al. 2019), it becomes more and more timely to test such assumptions.

Another hurdle lying ahead of association-based work is methodological in nature. Much work outside of plant ecology shows that many factors can bias the interaction strengths estimated from spatial associations, such as the scale of sampling (Araújo and Rozenfeld 2014, McNickle et al. 2018, Delalandre and Montesinos-Navarro 2018), the species’ habitat preferences (Morueta-Holme et al. 2016) or the method used for inference (Barner et al. 2018). Working at the local scale, where the imprint of ecological interactions is generally thought to be stronger, and with sessile species may alleviate some of those shortcomings, but probably not all of them. For example, shared residual
differences in micro-habitat may drive some of the associations between species. Here, our sampling
avoided any apparent fine-scale variations in soil characteristics, but residual bias may still be present.
For example, in our transects, the presence of pebbles (typical size of 5-20 cm) may make plant
individuals only grow in those remaining areas where pebbles are not present. This can produce
artefactual positive associations that are not due to any biotic interaction, as plant will cluster in the
remaining areas with free bare ground. In our case, this effect is unlikely to affect the trends given the
large predominance of negative associations, but it highlights the limits to the precision of association-
based approaches.

Relying on spatial patterns to make general statements about interactions in plant communities will
also require a standardization of methods. One acute point worth highlighting when considering trends
in association networks along gradients is controlling for changes in species abundances (Pellissier et
al. 2018). If networks are being compared without correction, the trends may not reflect changes in
species’ behaviors but rather changes in cover. For example, rare species cannot produce negative
associations, because even though they may strongly exclude other species in space, their total cover is
not sufficient to produce significant spatial segregation (Supplementary Information 4). Reporting
ecologically-meaningful variations in community-level metrics (e.g. the total number of negative links,
$K^-$) must be done with indices that control for species covers (e.g. $K_{SES}^-\)$. Misinterpreting this property
of associations can lead and has led to spurious interpretation of patterns (e.g. Calatayud et al. 2020).

To adequately anticipate how ecological communities will respond to global changes, it is necessary to
better map empirical interaction networks. This may not always be possible, as pairwise experiments
may be prohibitively expensive, or previous knowledge may be missing for the ecosystem of interest
(e.g. Kéfi et al. 2015). Association-based approaches provide an alternative yet objective basis to map
interactions in situ, preserving the environmental setting in which they occur, and for systems where
previous knowledge is scarce. They could thus greatly complement traditional approaches, but come with their own sets of methodological challenges. Obtaining detailed and accurate empirical interaction networks will thus require leveraging the complementarity of both experimental and observational approaches.

References


Soliveres, S. et al. 2015. Moving forward on facilitation research: response to changing environments and effects on the diversity, functioning and evolution of plant communities: Facilitation,


Figure 1. General variation of vegetation around a sheepfold and procedure used to calculate community-level network properties. Vegetation can be divided into different “belts” (a1), whose spatial extents depend on the grazing pressure (a2) and which are most extended towards the south-east because of the sheepfold buffering the herds against dominant winds (a1). We carried out two transects per vegetation type (i.e. per belt) in the south-east direction, leading to one observed association network and 999 ‘null’ networks per pair of transect (b1). Both observed and null networks were summarized into summary metrics, for example $R$, the relative frequency of positive to negative associations (b2). The deviation of the observed summary metric from the null expectation was then computed using standardized effect size, yielding cover-corrected community-level metrics such as $R_{SES}$ (b3).
Figure 2. Design of the transect surveys. Panel (a) represents a top-down view of plant individuals spread along a line transect (black line). We recorded all intersections of the line with plant parts, along with a class of height (b). The total overlap between species was then computed along the transect. For example, here $O_{12}$, the total overlap between species 1 and 2 is given by $O_{12} = l_1 + l_2$. We used pairs of transects in the field but in this figure, only one transect is represented for simplicity.
Figure 3. Variations in soil nitrogen, bare ground, dung cover, percentage of ruderal species, species richness as a function of distance to the sheepfold. The shape of the points and the separate trend lines indicate the different sites. Red trend lines are fits of a saturating function (the relationship described in equation 2), with different link functions depending on the nature of the response variable (proportional, continuous or discrete). The percentage of ruderal species is the proportion of the total plant cover made up of ruderal species *sensu* Grime’s CSR classification (Grime 1977).
Figure 4. Total fraction of links ($L_{SES}$), fraction of negative ($K_{SES}^-$), positive links ($K_{SES}^+$) and association ratio ($R_{SES}$) as a function of grazing pressure. Trend lines indicate linear regressions along the grazing gradient, as measured through total Nitrogen, with dashed lines representing the 95% credible interval on the predicted mean. Red and blue backgrounds highlight the areas corresponding to positive and negative values, respectively. Point shapes indicate different sites.

Figure 5. Trends of the heterogeneity of negative ($H_{SES}^-$, a) and positive ($H_{SES}^+$, b) links. Red and blue
backgrounds highlight the areas corresponding to positive and negative values, respectively. Trend lines indicate linear regressions along the grazing gradient, as measured through total Nitrogen, with dashed lines representing the 95% credible interval on the predicted mean. The posterior distribution for the $H_{SES}^+$ slope included zero between its 2.5% and 97.5% quantiles, suggesting no effect of grazing, so the trend is represented with a dotted line.
1. Description of the site and sampling protocol

In La Crau, herds of sheep are not within enclosures but lead by shepherds within open, extensive but delimited areas (most of them above 70 ha, with a maximum of 540 ha; Tatin et al., 2013). Most of these areas include a sheepfold built next to a shelter for the shepherd. Sheep are gathered in the sheepfold at night or, because of the large size of the herds (on average 1,600 individuals; Tatin et al., 2013), in temporary enclosures next to it. As a result, sheep spend more time near the sheepfold than elsewhere in the grazing area, resulting in a natural gradient of grazing pressure and other impacts related to the presence of sheep (e.g., nutrient enrichment through dung/urine deposition and trampling). This gradient is most extended towards the south-east, which is the main direction of strong dominant winds.

Vegetation communities respond to this grazing gradient, resulting in a strong species turnover with ruderal, nitrophilous communities near the sheepfold (Figure S1, and typical “Coussouls”-type communities (dominated by *Brachypodium retusum* and *Asphodelus ayardii*) where grazing is at its lowest, away from the sheepfold (Molinier & Tallon, 1950; Gomila, 1987). Previous work often classifies vegetation communities in four types, two being the ones described above, along with two others corresponding to intermediate levels of grazing, one defined by the presence of *Onopordum illyricum*, and another by the dominance of *Trifolium subterraneum* (Molinier et Tallon, 1950). In the absence of accurate measurement of the grazing pressure, we reused these four known vegetation types to guide our surveys and ensure that we documented a large range of grazing levels.

We first carried a quadrat-based survey and recorded the presence/absence of all species in a grid of 5 by 5 20x20 cm subquadrats (resulting in a quadrat size of 1m²). The first quadrat was placed as close as possible to the sheepfold entrance. The spacing between quadrats then followed a geometric series, i.e., the second quadrat was placed at \( d_2 = 2m \) from the first, then subsequent quadrats were placed such that \( d_{n+1} = b \cdot d_n \). \( b \) was between 1.2 and 1.5 depending on each site. Adjustments to this spacing rule were made to avoid local disturbances unrelated to grazing (e.g., dirt roads). We recorded in a 3x6m area around each quadrat the presence of indicator species, which were those used in the literature to define the four aforementioned vegetation types (Molinier & Tallon, 1950; Gomila, 1987). Basing the sampling on these indicator species allowed us to make sure we documented all vegetation types, hence a large range of grazing pressures, even in the absence of its direct measurement. At a given site, we stopped surveying once two quadrats were done in the “Coussouls”-type community (the criterion used was *Brachypodium retusum* at more than 50% cover in a quadrat).

Based on indicator species, we assigned in the field each quadrat to one of the four types of community. We thus obtained ranges of quadrats for each site and type of plant community. For example, for a site, we could have quadrat 1 to 3 in the most nitrophilous community, then 3-5 in the first intermediate community, 5-8 in the second intermediate community, then 8-10 in the “Coussouls”-type community. We then carried out a pair of transect in each community type, and
because this sampling was repeated twice for some sites, it resulted in at most 8 pairs of transects per site (i.e. 16 transects in total).

Figure S1 - Graphical summary of the different types of surveys carried out for a given site. Four plant community types have been identified based on a set of indicator species, as indicated by colors and labels. Labels indicate notable species present in each belt. Quadrats were carried out at different distances from the sheepfold entrance. In each belt, a pair of 5m transects was laid out to measure the spatial distribution of species individuals. Soil was also sampled at least four times per site at the average distance of each pair of transects.

In addition to quadrats and transects, soil samples were taken between Oct. 16 and Oct. 31 2017 (when sheep are absent). One sample was taken at the average distance of each transect pair, sometimes complemented by samples placed at quadrat positions so that soil characteristics could be interpolated at each quadrat and transect position. This resulted in 4 to 6 soil samples per site.
Each soil sample was taken by mixing 16 regularly-spaced sub-samples in a circle of 4 m around a center point. All surface litter was scraped and approximately 125 cL of soil was sampled at each point between 0 and 10 cm. All subsamples were mixed in a bowl and a compound sample of 400 g was taken. This latter sample was dried for at least 72 h at 40°C, sieved to 2 mm and sent for analyses (Laboratoire Teyssier, 5 route des Junchas, 26460 Bourdeaux, France). Soil analyses included pH (in water), Metson total CEC (Cation Exchange Capacity), total CaCO$_3$, total organic matter, total Nitrogen, total Phosphorus (Olsen method), exchangeable K, Mg, Ca and Na.

### 2. Estimating Grazing pressure

A stated earlier, grazing in La Crau is centered around sheepfolds. During the day the shepherd will lead the sheep out and come back at night close to the sheepfold (Figure 2). As a result, disturbance due to the presence of sheep is strong next to the sheepfold, and decreases as one moves away from it. However, there are very few direct measurements of the shape of this relationship.

![Aerial view of a sheepfold (Google maps), in which the nitrophilous plant communities (dark green) are clearly visible. (b) The sheep stays close to the sheepfold at night, then the herd is conducted during the day by the shepherd: as a result, sheep spend more time close to the sheepfold than far from it.](image)

To our knowledge, only one study (Dureau, 1998) attempted to measure such relationship directly. Extracting the reported values and plotting grazing pressure against distance, we obtain the following relationship:
The above plot suggests that grazing pressure is well-described by the following relationship (drawn as a red line above), where $x$ is the distance to the sheepfold:

$$G(x) = G_{inf} + \frac{G_0 - G_{inf}}{1 + \lambda x}$$

where $\lambda$ characterizes how fast the grazing pressure decays as a function of distance to the sheepfold entrance. Writing $\tau = 1/\lambda$ and $\Delta G = G_{inf} - G_0$, this equation can be rewritten as:

$$G(x) = \frac{\tau G_0 + (G_0 + \Delta G)x}{\tau + x}$$

In this equation,

- $\tau$ characterizes the extent of the spatial gradient
- $G_0$ characterizes the grazing pressure near the sheepfold entrance
- $\Delta G$ characterizes the difference in grazing pressure between the areas near the entrance and areas far from the sheepfold (“at infinite distance”).

We found that many soil-related variables showed trends that could be described by such relationship, suggesting that soil parameters could be used as a proxy for the position of a given transect along the grazing gradient (Figure S4).
We focused on the variations of total nitrogen to estimate grazing pressure because it was expected to be the most directly correlated to herbivore defecation around and in enclosures (Selbie, et al. 2015, Steinauer and Collins 1995), and was well-measured by the soil analyses (compared to phosphorus for instance for which small concentrations had similar measured values). Other soil variables were more indirectly linked to sheep presence (e.g. pH), or results reported by the soil analyses were unreliable (e.g. total Ca, were values are only 0 or 1). It is worth noting however that most soil variables showed correlated trends along the grazing gradient.

Assuming that the total N in the soil is linearly related to sheep activity (i.e. that $N \approx a \cdot G + b$ with $a$ and $b$ being constants), we can model its relationship with distance to obtain an estimate of grazing index. In the main text analyses, we removed three outliers from the regression that did not fit model assumptions (points with $N > 0.0035$ in figure S4). These outliers exhibited a much higher content in $N_{tot}$ that was predicted by an inverse-distance relationship, and were probably the result of sample contamination (probably a sampling error involving remaining dung in the soil sample).

For the analyses described in the main text, we fitted a model of the following form:

$$N(x) = \frac{\tau N_0 + (N_0 + \Delta N)x}{\tau + x}$$

(3)
where:

- \( N_0 \) is the estimated soil Nitrogen content at the sheepfold entrance
- \( \Delta N \) is the estimated difference in soil Nitrogen content between areas far from the sheepfold (where grazing pressure is at its minimum) and areas close to it
- \( \tau \) is the spatial extent of the gradient (proportional to the distance at which grazing pressure has decayed by half of its maximum value). It represents a distance (in meters).

We fitted this model using a bayesian setting and the following conservative priors (to prevent overfitting) on population-level effects:

\[
N_0 \sim \mathcal{N}(\mu_N, 3 \times \sigma_N) \\
\Delta N \sim \mathcal{N}(0, 3 \times \sigma_N) \\
\tau \sim \mathcal{N}(200, 200)
\]

where \( \sigma_N \) is the standard deviation of the observed values of total nitrogen, pooled for all sites. We used a site-level random effect on the estimates of \( N_0, \Delta N \) and \( \lambda \), and used the following Student distributions with three degrees of freedom and the following parameters:

\[
N_0 \sim \text{Student}_3(\text{mean} = 0, \text{sd} = 5 \times \sigma_N) \\
N \sim \text{Student}_3(\text{mean} = 0, \text{sd} = 5 \times \sigma_N) \\
\tau \sim \text{Student}_3(\text{mean} = 0, \text{sd} = 200)
\]

We modelled a correlated group-level effect between \( N_0 \) and \( \Delta N \) because it was found to produce better fits (checked visually).

The model was fit using the R package \texttt{brms} (see the included file \texttt{functions.R}, line 109 for the implementation).

For all pairs of transect, we took the average predicted total N value as an index of grazing pressure.

### 3. Sensitivity to association cutoff \( \alpha \)

We retained a positive association between two species when their total overlap was above the 87.5% quantile of the null distribution, i.e. using \( \alpha=0.25 \) cutoff (100% - \( \alpha/2 = 0.875 \) or 87.5%). Similarly, we retained negative associations when the total overlap was below the 12.5% quantile (\( \alpha/2 \)). \( \alpha \) here represents a tradeoff. Larger values of \( \alpha \) yield networks with a higher proportion of links, which may be interesting to compute aggregate statistics such as the proportion of links (\( L \)). However, these associations can be weak (i.e. they only weakly aggregate or segregate in space). Smaller values of \( \alpha \) will yield networks with links between plants that are only strongly positively or negatively associated, but these networks can be very sparse (low number of links) and thus some community-level metrics (e.g. the total number of links) can be badly estimated.

In figure S5, we investigate the influence of the value of \( \alpha \) on the trends we observe along the gradient. We report the estimates on the slopes of the linear trends displayed in Figure 4 and 5 in main text, but with varying values of \( \alpha \).
Figure S5 - Results of sensitivity analyses. Black bars indicate the 2.5% and 97.5% quantiles of the credible interval on the slope, and the black points the mean estimate. Red crosses indicate values of $\alpha$ for which a quadratic model was selected over a linear model to describe the trends of the network metric along the gradient. The dashed vertical line indicates the value chosen for the trends reported in main text (0.25).

For most metrics, it appears that the trends stay stable for a wide variety of $\alpha$ cutoffs, suggesting that our conclusions are robust and not the result of a specific choice of value for $\alpha$. While some values of $\alpha$ yield quadratic models, this behavior is sporadic and does not reflect the general pattern.

4. Species’ abundances and bias in associations

The ability of a species to exhibit negative or positive associations varies with its abundance. Specifically, species under a given total cover cannot exhibit positive associations. We run here a small simulation with synthetic data to highlight this bias in association networks.

We first consider a transect with two large individuals totalling 50% cover each. Figure S6 represents a transect with such individuals, which is of the same nature as the data recorded in this study.
Figure S6 - A transect with two large individuals. The x-axis is the length along the transect, and the y-axis is the height of the plant individuals (see also Figure 2b in main text).

In such situation, the total observed overlap is 0.2 (20%). The null distribution of overlap (given randomized positions of the individuals along the transect) for the species pair is in Figure S7.

Figure S7 - Null distribution of overlap for a transect with a pair of species with high cover (50%). The continuous lines indicate the quantiles above or under which a significant positive or negative association is retained, respectively (for $\alpha = 0.25$). The dashed line indicates the observed value.

We see that the null distribution has values above and below the cutoff used to retain a positive or negative associations (the red and blue areas, respectively). Thus, the two species can exhibit both a positive or a negative association, depending on their observed overlap.

Note that this distribution has no values above 0.5 (50%), which is the maximum possible overlap.
between the two species.

We now consider the case of two rare species with low cover (10%):

Figure S8 - A transect with two small individuals. The x-axis is the length along the transect, and the y-axis is the height of the plant individuals (see also Figure 2b in main text).

We run the same analysis as before to obtain the null distribution of overlap between the two species (Figure S9).

Figure S9 - Null distribution of overlap for a transect with a pair of species with low cover (10%). The continuous lines indicate the quantiles above or under which a significant positive or negative association is retained, respectively (for $\alpha = 0.25$). The dashed line indicates the observed value. Note that the observed value and the lower quantile lines are superposed.
This distribution is heavily skewed, and has a high excess of zeroes (no overlap between the two species). As a result, the lower cutoff is zero (i.e. the 12.5% quantile is zero), which is also the natural bound for the metric of overlap, which cannot be negative. As a result, such two rare species cannot exhibit an overlap lower than the cutoff, and thus cannot exhibit a negative association, regardless of their ability to exclude spatially other species.

We can produce this null distribution of overlaps for different values of cover of the two species and track when the lower quantile of the null distribution is zero. Simulations show that for the case above, the lower quantile of the null distribution is zero up to the cover of 41%. This means that two species with a cover below this number will never exhibit a negative association for statistical reasons.

This asymmetrical detection of associations has a very strong effect on the estimates of the community scale metrics (e.g. the total number of negative links), as natural communities have a few abundant species and a lot of rare ones. To correctly interpret trends along gradients where species abundances change, it is thus necessary to compare, not the raw values of community metrics, but how much they deviate to what is expected given the distribution of abundances (see Methods in the main text).

References


