



The impact of non-reproductive plant species on assessments of community structure and species co-occurrence patterns

Brandon S. Schamp, Lonnie W. Aarssen, Gillian S.J. Piggott & Sneha K. Dante

Keywords

Checkerboard patterns; Clonal extension; Community assembly; Competition; Dispersal; Null model; Species association; Species co-distribution; Species segregation

Nomenclature

Gleason & Cronquist (1991)

Received 14 September 2015

Accepted 29 February 2016

Co-ordinating Editor: Miquel De Cáceres

Schamp, B.S. (Corresponding author, brandon.schamp@algomau.ca)¹,

Aarssen, L.W. (aarssenl@queensu.ca)²,

Piggott, G.S.J. (gillian.piggott14@gmail.com)²,

Dante, S.K. (skdante8@gmail.com)²

¹Department of Biology, Algoma University, Sault Ste. Marie, ON P6A 2G4, Canada;

²Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada

Abstract

Aims: Studies of community structure and co-occurrence patterns rely on the premise that community data reflect where species successfully grow and which species they grow with. However, plant censuses generally do not distinguish between species with reproductive individuals and those only represented by non-reproductive individuals. We tested whether inclusion of non-reproductive species, which may not reflect success in that location, significantly impacts evaluations of community structure and co-occurrence.

Location: Queen's University Biological Station, Ontario, Canada, old-field plant communities.

Methods: We quantified the impact of non-reproductive species in two plant communities by comparing community structure and co-occurrence patterns when non-reproductive species were included or excluded.

Results: Including non-reproductive species significantly increased plot-level species richness in both communities (54% and 13% increases), altered species evenness in both communities, significantly impacted beta-diversity among plots in one site, and disproportionately impacted assessments of diversity in species-rich plots. Excluding non-reproductive species resulted in reduced negative co-occurrence patterns in both communities, with a substantially larger impact in one community. In that community, the impact of non-reproductive species was even more pronounced when abundance data were used in analysis, and when pair-wise co-occurrence patterns were assessed. Additionally, including non-reproductive species drastically decreased the number of species pairs with perfect negative co-occurrence across sites, indicating that these species can add 'noise' to co-occurrence patterns. We examined possible explanations for the presence of non-reproductive species. In one community, non-reproductive species were 22 times less abundant (per plot) than reproductive species within plots, although they were not rare overall. Differences in the number of non-reproductive species per plot across our focal communities were not clearly driven by differences in clonality, stress from extreme weather or low N. While these patterns are consistent with the interpretation that non-reproductive species are present due to mass effects, this possibility requires further research.

Conclusions: Including non-reproductive plant species in censuses can significantly impact assessments of community structure and species co-occurrence. The divergent impact of their inclusion on our two communities highlights the possibility that excluding non-reproductive species from surveys may remove noise from community data and clarify theories of plant species co-existence.

Introduction

A principal goal in ecology is to understand the mechanisms responsible for existing variation in patterns of diversity and relative abundance of species across different communities (e.g. Diamond 1975; Gotelli & McCabe 2002; Kikvidze et al. 2005). While research has demonstrated that mechanisms such as competition and dispersal play important roles in the assembly of natural communities (e.g. Schoener 1983; Aarssen & Epp 1990; Weiher & Keddy 1995; Pärtel et al. 1996; Flinn et al. 2010), understanding the relative contribution of multiple mechanisms in natural systems remains a challenge. Progress may benefit from reconsideration of some of the basic assumptions ecologists hold about the value of simple census methods in estimating community structure and investigating assembly rules (e.g. Gotelli & Colwell 2001). Here, we examine the impact of two different census methods on estimates of plant community structure and negative co-occurrence patterns.

Plant census methods for angiosperms can either tally all species found in a given sample plot (Full Census), or only those achieving reproduction in the census year in that plot (Flower Census). Censuses that include species that do not reproduce in a given plot are arguably more common because they require a single visit, as opposed to a relatively labour-intensive collection of visits across the growing season that is required to determine which species have had at least one representative achieve reproduction in each plot. In our experience, many of these non-reproductive species are represented by very small plants that are unlikely to survive and reproduce in that location. There is an implicit assumption, albeit a debatable one, that data including these non-reproductive species represent reliable information about the co-existence of the species in sample plots. Including non-reproductive species in censuses gives a liberal measure of plot diversity, while excluding them yields a conservative one. The extent to which non-reproductive species influence assessments of community structure and co-occurrence patterns in different communities is currently unknown.

The use of census data to interpret natural patterns of species co-existence has been universal. For example, diversity–productivity relationships, perhaps the most studied plant community pattern (Grace 1999), rely on the premise that plot-level measures of diversity reflect co-existence. The fact that co-occurrence tests are not called co-existence tests is an acknowledgement of the potential for other processes, including dispersal, to introduce noise into these relationships (e.g. Schamp et al. 2015). While it is clear that several well-used negative co-occurrence models have acceptable statistical power (Gotelli 2000), evidence suggests that these tests produce evidence of significant negative co-occurrence less than half of the

time (41% of analyses; Götzenberger et al. 2012). It is possible that noise from non-reproductive species contributes to this low level of support for negative co-occurrence patterns (for other potential influences, see Götzenberger et al. 2012). If non-reproductive species are in the process of being excluded, they are falsely labelled as positively co-occurring with other species in that sample plot, and potentially adding, inappropriately, negative co-occurrence patterns with species in other plots. As such, when non-reproductive species are present in plots where they are destined to be out-competed, their impact on negative co-occurrence patterns does not reflect competitive sorting, but may instead reflect the influence of dispersal. The corollary is that when non-reproductive species are excluded from community censuses (and other influences like abiotic variation are minimized), the assessment of negative co-occurrence patterns may more closely reflect the outcome of species interactions.

Early explorations of co-occurrence were conducted with a goal of understanding how competitive interactions, and other processes, contribute to patterns of species distribution (Diamond 1975). Competition, under this scenario, prevents the co-occurrence of species that are unequal competitors for a shared set of resources. Positive co-occurrence patterns, wherein pairs of species are frequently found growing together, are also possible, and may be indicative of facilitation, equivalent competitive abilities or niche differences not involving horizontal space. Considerable effort has gone into investigating the efficacy and appropriateness of different null model algorithms and co-occurrence indices and much progress has been made (Connor & Simberloff 1979; Stone & Roberts 1990; Manly 1995; Sanderson et al. 1998; Gotelli 2000; Gotelli & Entsminger 2001; Fayle & Manica 2010; Ulrich & Gotelli 2010; Gotelli & Ulrich 2012). Measures of co-occurrence for specific species pairs have also been developed (Sfenthourakis et al. 2004; Gotelli & Ulrich 2010; Veech 2013).

Dispersal can influence patterns of community structure and negative co-occurrence. Seed addition experiments, for example, have demonstrated the potential for dispersal to augment plant diversity (Turnbull et al. 2000). Furthermore, it is clear from both theoretical and experimental work that dispersal can impact co-occurrence patterns (e.g. Ulrich 2004; Bell 2005; Schamp et al. 2015). Dispersal can impact co-occurrence patterns in multiple ways. If a species is incapable of dispersing to all sites under consideration, this can produce negative co-occurrence patterns that are not generated by competitive interactions (Pielou & Pielou 1968; Zalewski & Ulrich 2006; Heino 2013). Also, dispersal can reduce competitive exclusion, influencing co-occurrence patterns (Heino & Grönroos 2013; Schamp et al. 2015). We examine the possibility that non-reproductive plant species may contribute information about

dispersal's contribution to community structure and negative co-occurrence patterns.

Abiotic variation can also contribute to community structure and negative co-occurrence patterns if some plots under consideration differ significantly in abiotic conditions, such that different subsets of species within the focal system matrix are incapable of successfully growing in all plots (Dullinger et al. 2007; Reitalu et al. 2008; Heino 2013). This issue should be less problematic when analyses focus on relatively small communities that are sufficiently homogeneous that all plots are within the fundamental niches of all focal species. Focusing on a small community also reduces the likelihood that the absence of a species from some plots is due to dispersal limitation. In such communities, patterns of richness, evenness and co-occurrence are more likely to reflect the influence of competition. However, even at small scales dispersal can influence both community structure and co-occurrence patterns (Heino & Grönroos 2013; Schamp et al. 2015).

In this study, we examine how the inclusion of non-reproductive species (i.e. those whose resident plants are 'vegetative' only) in plot censuses impacts patterns of community structure and negative species co-occurrence. We test two predictions: (1) the inclusion of non-reproductive species in census data significantly impacts the evaluation of community structure; (2) including non-reproductive species significantly impacts the assessment of negative co-occurrence patterns. Additionally, we conduct a number of analyses directed at understanding the nature of non-reproductive species within our focal communities and examining the possibility that they are present due to dispersal.

Methods

Study sites

The two focal communities were located at the Queen's University Biological Station, Chaffey's Locks, Ontario, Canada (44°34' N, 76°20' W). Both were old-field communities on abandoned agricultural fields that had not been ploughed or tilled for more than 70 yrs. Community plot censuses were performed by setting up grids of 1 m × 1 m plots within an 80 m × 80 m area in Site 1 and a 50 m × 50 m area in Site 2. Each study field was bordered on three sides by dense forest and on one side by a road, and in each case the study grid was centrally located in the field such that there were no edge effects; in both cases, plot grids were over 100 m from the road. The study communities were chosen to be relatively homogeneous topographically. The two communities were approximately 750 m apart, separated by a road as well as a matrix of mixed deciduous forest. Site 1 contains 61 vascular plant species, the majority of which are perennial

(e.g. *Poa pratensis*, *Elymus repens*, *Asclepias syriaca*) including some forage species typical of the region (e.g. *Phleum pratense*, *Vicia cracca*, *Trifolium pratense*). Site 2 is also dominated by herbaceous perennials, and contains 37 plant species, the most common of which are *P. pratensis*, *P. pratense*, *Cerastium arvense*, *Potentilla recta*, *V. cracca* and *Rumex acetosella*.

Data collection

Data collection took place between May and August for both sites; Site 1 data were collected in 2006 and Site 2 in 2009. Fifty focal plots in each site were chosen randomly from the available grid of plots; one plot from Site 1 was excluded from analysis because vegetation was repeatedly disturbed (i.e. flattened) by a deer that chose to sleep there. In both communities, two methods of plot-level census were conducted on a bi-weekly basis. The first plot census method included all species found in each plot, regardless of whether they flowered (Full census), and the second method included only species that flowered that season (Flower census). The Full and Flower censuses represent liberal and conservative estimates of species co-existence respectively. The Flower census should reduce the impact of mass effects on plot-level species composition by excluding all individuals for which evidence of successful reproduction in that plot does not exist.

In Site 1, we also collected plot-level abundance data. Over the course of the field season, we monitored plots bi-weekly and counted and tagged all plants that were reproductive (using yarn at the base). The abundance of each species, reproductive and non-reproductive, in each plot was determined when plots were harvested in the last 2 wks of Aug 2006 and cross-referenced with data from bi-weekly monitoring. Abundance was quantified as the number of rooted units per species, with a rooted unit representing a transition point between above-ground and below-ground tissues or a collection of shoots that are attached at or slightly below ground level (e.g. a single tiller for rhizomatous or tuft-forming grasses). Thus, a rhizomatous plant attached below the soil was likely counted as multiple rooted units. Rooted units were enumerated as an estimate of the number of functional individuals, which allowed us to assess the proportion of individuals reproducing for each species, an approach we believe would have proven error-rich had we assessed abundance using cover data. Disturbance to plots during the intensive bi-weekly monitoring was reduced though the use of benches ($L \times W \times H = 2.2 \text{ m} \times 0.75 \text{ m} \times 1.0 \text{ m}$) on which researchers knelt while tagging and counting reproductive plants. Benches were large enough that the legs of the bench could be located 1 m from the plot.

Seed addition in Site 1

Site 1 data were collected as part of a multi-purpose study. As such, in the autumn of 2005, prior to the 2006 censuses, 250 seeds of each of 15 non-local species were seeded into all 49 study plots. These species (Appendix S2 Table S1) are common to old-field plant communities in the area, but were not present in Site 1. This was done to examine the impact of variation in community structure on invasion success and also to understand how the influx of non-local dispersers, which are unlikely to achieve flowering in the census year, affected co-occurrence patterns. This paper concerns only community structure and co-occurrence patterns using the community census from 2006; we report analyses both with and without invaders. The impact of invaders, as we report, was negligible, with only seven of these species appearing in the 2006 census, and to only a limited extent (three plants reaching reproduction). This is likely due to the limited colonization, and the fact that unlike mass effects, the invasion treatment was applied to all plots equally.

Analysis

Non-reproductive species and community structure

To understand how the inclusion or exclusion of non-reproductive species impacts community structure in general, we tested whether plot-level species richness differed significantly between Full and Flower censuses using Wilcoxon signed rank tests. We also compared rank abundance curves for the two census methods to understand how each method alters species occupancy patterns among plots. We assessed the evenness of these abundance patterns using E_{var} , a measure of evenness that is not correlated with species richness (Smith & Wilson 1996). Next, we compared β -diversity for the two census methods in each study site. Beta-diversity among plots was calculated with presence-absence data using the Jaccard index of compositional dissimilarity. We used the `vegdist`, `betadisper` and `anova` functions, part of the `vegan` package in R, to test for differences in β -diversity between census methods in each site. Using this method, β -diversity is calculated as the distance from each plot to the group centroid in principal coordinate space, as set out in Anderson (2006). These analyses were carried out in R (R Foundation for Statistical Computing, Vienna, AT).

We used reduced major axis regression to determine the slope of the relationship between the number of reproductive species and the total number of species in each plot (both reproductive and non-reproductive) for both sites. We then used a t -test on slopes to determine whether these slopes differ significantly across the two communities. If the slopes were equal, this suggests that the Full census

can be accepted as having been predictive of the Flower census in previous studies.

Non-reproductive species and co-occurrence patterns

To compare co-occurrence patterns across the two census methods, we used three approaches. First, we assessed co-occurrence at the community level using the C -score metric in combination with the fixed-fixed independent swap algorithm with 30 000 swaps to generate null distributions. To do so, we used matrices in which rows represent species and columns represent sites. The C -score metric is the mean number of checkerboards observed across all pairs of species and is calculated as $\sum(S_i - Q)(S_j - Q)/[(R)(R - 1)/2]$, where S_i is the sum of row i , S_j is the sum of row j , Q is the number of sites in which both species in a given pair are present, and R is the number of rows in the matrix (Stone & Roberts 1990). We compared the two census methods within each community using the standardized effect size (SES; e.g. Sanders et al. 2003), which is calculated as the number of SDs that the observed C -score (C_{obs}) for the study community is away from the mean C -score in the null distribution (C_{sim}), defined here by 5000 randomized C -scores. This was calculated as $(C_{\text{obs}} - C_{\text{sim}})/S_{\text{sim}}$ with S_{sim} equal to the SD of C -scores in the null distribution.

Because co-occurrence tests produce a single value for each census method (P -value or SES), we could not test whether observed differences between censuses were statistically significant. However, we used a jackknifing procedure in which we iteratively conducted co-occurrence tests on all possible matrices with $n - 1$ plots, where n is the total number of plots. This does not produce independent measures of co-occurrence, but does give a measure of how stable our co-occurrence estimate is. Thus, we conducted null model tests for each combination of 49 plots in Site 1 (49 combinations) and each combination of 50 plots in Site 2 (50 combinations; Schamp et al. 2015). Similar jackknifing methods have been used to assess the stability of phylogenetic hypotheses (Lanyon 1985).

Second, we tested whether a community-level null model that incorporated abundance data would also produce different results across the two census methods. We used the IT randomization method (Ulrich & Gotelli 2010) in combination with the CA_{ST} metric to compare co-occurrence patterns across census methods in Site 1, for which we had abundance data. The IT randomization algorithm fills null matrices cell-by-cell, adding individuals randomly, but weighted by the number of individuals of that species found in the community, and by the number of individuals in that particular plot. This iterative process continues until row and column sums match that of the original community matrix (Ulrich & Gotelli 2010).

Because this model does not conserve the species richness distribution during randomization, the question it addresses differs subtly from that addressed by co-occurrence tests using the fixed–fixed swap algorithm, because it acknowledges that interactions may shape both co-occurrence and richness patterns. The CA_{ST} metric is similar to the C -score but using abundance data. CA_{ST} is the mean number of abundance checkerboards per pair of species in the matrix, calculated as $CA_{ST} = (4CA)/[m(m - 1)n(n - 1)]$, where CA is the number of abundance checkerboards, m the number of sites, and n the number of species in the focal matrix. Abundance checkerboards are assessed for all 2×2 submatrices.

Third, we assessed whether census method significantly impacted the number of significantly negatively co-occurring species pairs identified using the empirical Bayes method (<http://www.keib.umk.pl/oprogramowanie/?lang=en>; Gotelli & Ulrich 2010). This analysis was also done for Site 1, and used the C -score metric in combination with the independent swap algorithm with 1000 matrix randomizations and a false error rate correction (Benjamini & Yekutieli 2001).

Exploring the source of non-reproductive species

We conducted several analyses to better understand the mechanism(s) responsible for non-reproductive species. Non-reproductive species may be present for a variety of reasons. Local mass effects may disperse species to plots where they are capable of growing, but incapable of competing given the local composition. It is also possible that non-reproductive species are simply in the pre-reproductive stages of growth, are present but stunted due to stress from low resources, or stressful weather conditions, or that clonality is responsible. We investigated each of these possibilities, although we envision this work as an early exploration of the nature of non-reproductive species that will inform more direct experiments in the future.

First, we tested whether non-flowering species made co-occurrence patterns noisier. This would be expected if these species are present in some plots due to mass effects and disrupted clear patterns of competitive dominance among species pairs (when counted as present), but this noisiness would not be expected for the other explanations above. We assessed the ‘noisiness’ of co-occurrence patterns by determining the degree to which pairs of species in our two focal communities formed more perfect checkerboard patterns across pairs of species when non-reproductive species were excluded (i.e. whether non-reproductive species disrupt/reduce perfect checkerboards). The Checker metric enumerates the number of perfect checkerboards among all possible $n(n - 1)/2$ pairs of species (Diamond 1975; Gotelli 2000). For two hypothetical species, a perfect

checkerboard is observed when they are not found growing together in any plot; the Checker metric is therefore more conservative, requiring a higher standard of evidence of negative co-occurrence. We compared the total number of perfect checkerboards (Checker score) among all pairs of species for Full and Flower censuses in both sites to test whether the inclusion of non-reproductive species reduced the number of perfect checkerboards among species pairs. We used the same jackknifing procedure for these tests that we used with tests using the C -score metric.

Second, we used Site 1 data to test whether the abundance of species in plots that had at least one reproductive representative was significantly higher than the abundance of species that had no reproductive representatives (Mann–Whitney U -test). If non-reproductive species were simply young plants, present via clonality, or the result of resource stress, they should not be any less abundant than species with some reproductive individuals. In contrast, if these species were present because of local dispersal, we expected that they would be significantly less abundant in plots than species with some reproductive representatives.

Third, to explore the possibility that non-reproductive species in plots were explained by clonal extension, we determined the number of species in each community and each plot that are capable of such extension via rhizomes or stolons (Gleason & Cronquist 1991). Using these data, we tested whether the large difference among our two sites in the number of non-reproductive species corresponded with a difference in the proportion of non-reproductive species per plot that were rhizomatous or stoloniferous.

Finally, we tested whether non-reproductive species might be present by virtue of stress from low resources (Site 1 – for which we had soil nutrient data) or stress from unusual weather in the two collection years. For Site 1, we used soil nitrate and ammonia data from soil cores collected from each plot in Aug 2006 to test whether plots with low available N contained more non-reproductive species (i.e. were negatively correlated). Because the two sites were sampled in different years, we were able to use weather station data from the Queen’s University Biological Station spanning 1997–2010 to determine whether the relatively high number of non-reproductive species observed in Site 1 could be attributed to an unusual weather year that restricted growth of some plants.

Results

Community structure and non-reproductive species

The inclusion of non-reproductive species (i.e. Full census) resulted in pronounced differences in community structure. Plot-level species richness was significantly higher for the Full census than for the Flower census in both focal communities (Fig. 1a, b; Wilcoxon signed rank tests,

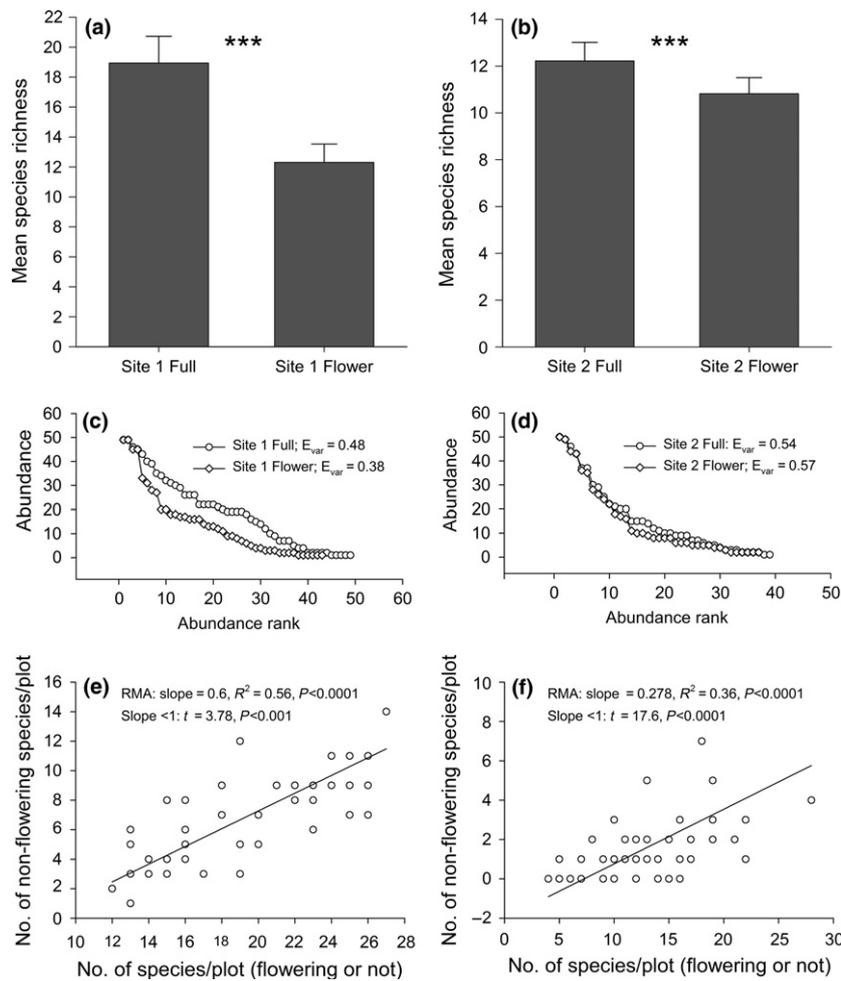


Fig. 1. Plot-level community structure comparing the two census methods, illustrating how the inclusion of non-reproductive plant species (Full census) significantly affects plot-level species richness (panels **a**, **b**; Wilcoxon signed rank tests, *** indicates $P < 0.001$) as well as how rank abundance patterns change when censuses only focus on species that achieve flowering in the study year (panels **c**, **d**). Evenness, measured in terms of the number of plots occupied across species (E_{var} , Smith & Wilson 1996) was lower for the Flower census than for the Full census in Site 1 but higher in Site 2. Whiskers for species richness represent 95% confidence intervals for the means, illustrating the breadth of variation in plot-level species richness. Panels **e** and **f** illustrate reduced major axis (RMA) regression of the number of non-reproductive vascular plant species per plot on the total number of vascular plant species per plot, including those with or without reproductive representatives for Site 1 (**e**) and Site 2 (**f**). RMA regression was used because the regression was undertaken not to establish a causative or predictive relationship, but to establish whether the number of non-reproductive species was proportionate to the overall plot diversity. A two-tailed student's t -test established that the RMA slope was significantly less than 1 for both sites.

$P < 0.001$). For Site 1, mean plot-level species richness was 54% higher for the Full census than for the Flower census (Full: $\bar{x} = 18.9$; Flower: $\bar{x} = 12.3$; Fig. 1a). The inclusion of non-reproductive species also markedly altered the shape of rank abundance curves and evenness of occupancy patterns (E_{var} ; Smith & Wilson 1996; Fig. 1c, d). Beta-diversity among plots was significantly lower for the Full censuses in Site 1 (ANOVA, $df = 1$, $P = 0.019$; Full $\bar{x} = 0.40$; Flower $\bar{x} = 0.44$), but did not differ across censuses in Site 2 (ANOVA, $df = 1$, $P = 0.764$; Full $\bar{x} = 0.42$; Flower $\bar{x} = 0.42$).

The relationship between species richness per plot and the number of non-reproductive species per plot for the

two census methods differed in slope for the two communities (Site 1: slope = 0.69; Site 2: slope = 0.28; $t = 4.16$, $P < 0.0001$; Fig. 1e, f), indicating that the Full census is not a constant predictor of the Flower census across communities.

Co-occurrence patterns and non-reproductive species

Our community-level assessment of co-occurrence patterns using presence-absence data uncovered differences across the two census methods. In both sites, there was significant negative-co-occurrence, regardless of the census method; however, in both sites, negative co-occurrence

was stronger for the Full census compared to the Flower census (Fig. 2). In Site 1, the inclusion of non-reproductive species (Full census) increased the effect size of the co-occurrence test by 58% compared to when non-reproductive species were excluded (Flower census: 20% higher for Site 2; Fig. 2a, b). A similar pattern was observed for Site 1

when invaders were included; however, SES values were uniformly lower across the two census methods (Appendix S1 Fig. S1).

Analysing negative co-occurrence patterns for Site 1 using abundance data and a null model designed to use with these data produced similar results to those produced

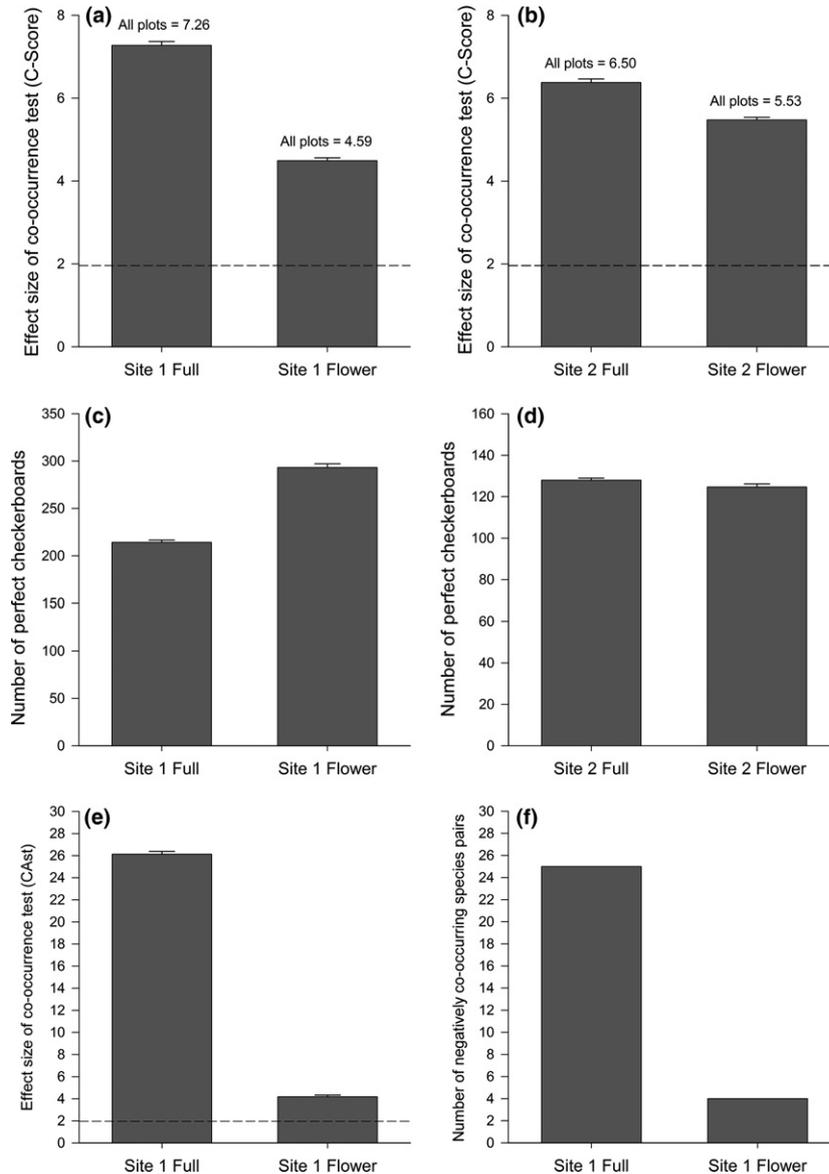


Fig. 2. A comparison of co-occurrence patterns for the two census methods (Full vs Flower) for Site 1 (panel a, c, e) and Site 2 (panel b, d, f). Panels a and b compare the standardized effect size (SES) of co-occurrence tests within communities conducted using the C-score metric and the independent swap algorithm. Panels c and d illustrate how the inclusion of non-reproductive species impacts the number of perfect checkerboards among pairs of species. The SES of full sample tests are given above the bars in panels (a and b). Panel e presents a comparison of the SES of co-occurrence patterns for the two census methods for Site 1 using abundance data in conjunction with the CA_{ST} co-occurrence metric and IT randomization algorithm. Panel f illustrates the number of significantly negatively co-occurring species pairs found using presence data for Site 1 with the empirical Bayes method, using the C-score with 1000 matrix randomizations and a false error rate correction (Benjamini & Yekutieli 2001). Dashed lines represent the z-score of 1.96, above which negative co-occurrence patterns are generally accepted as being higher than expected by chance. Whiskers (panels a–e) illustrate 95% confidence intervals resulting from a jackknifing procedure in which co-occurrence tests were performed for the community for all possible combinations of plots with a single plot removed.

using presence–absence data; negative co-occurrence patterns were much more pronounced when non-flowering species were included (Full census). This pattern was, however, more pronounced than when presence–absence data were used in analysis (524% higher; Fig. 2e). Additionally for Site 1, the number of significantly negatively co-occurring pairs of species identified using the empirical Bayes method (www.uni.torun.pl/~ulrichw; Gotelli & Ulrich 2010) was much higher for the Full census compared to the Flower census (525% higher; Fig. 2f).

Exploring the source of non-reproductive species

For Site 1, there were more perfect checkerboards among species pairs (i.e. a higher Checker score) for the Flower census (Fig. 2c); thus, the inclusion of non-reproductive species reduced the number of species pairs with perfect checkerboards, adding noise. For Site 2, the number of perfect checkerboards observed was similar for the two census methods (Fig. 2d).

In Site 1, species in plots that were represented by at least one reproductive individual were significantly more abundant than those that did not have a single individual reach flowering in the census year (Fig. 3; Mann–Whitney U , $P < 0.001$). The difference was large, with the mean plot-level abundance for species with reproductive representatives at 235.5 compared to a mean abundance of species without reproductive representatives of 10.5.

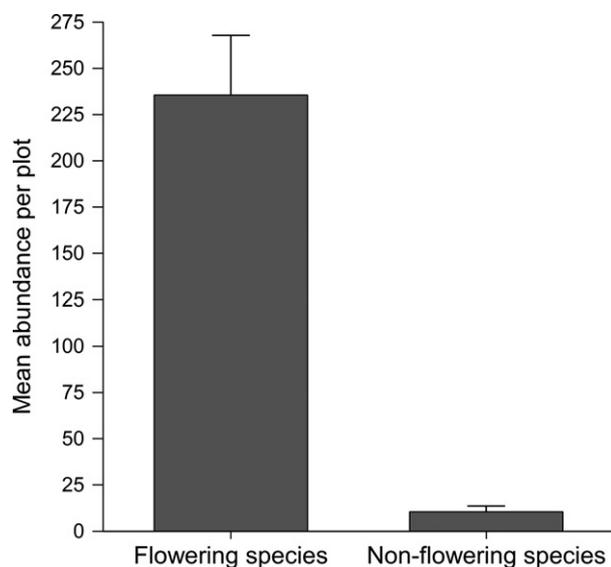


Fig. 3. Mean abundance of species in plots that have at least one individual flowering, compared to the abundance of species in plots that have no individuals flowering. Whiskers are the 95% confidence limits. Species with reproductive representatives were significantly more abundant than those with no reproductive representative (Mann–Whitney U , $P < 0.001$).

Overall, Site 1 contained a higher proportion of species with rhizomes or stolons than did Site 2 (Site 1: 30 of 51 species; Site 2: 18 of 37 species); however, the two sites did not differ significantly in the proportion of non-reproductive species with rhizomes or stolons per plot (Mann–Whitney U , $P = 0.51$). In fact, the proportion of non-reproductive rhizomatous or stoloniferous species per plot was slightly higher in site 2 (Site 1: 0.54; Site 2: 0.57). Consequently, differences in the proportion of non-reproductive species per plot among the two sites did not correspond to differences in the proportion of clonal species in those plots.

In Site 1, for which we have soil N data, ammonia plus nitrate in plots was positively correlated with the number of non-reproductive species in plots ($r = 0.33$, $P = 0.024$); this was in opposition to the negative correlation expected if nutrient stress was responsible for non-reproductive species. With respect to mean temperature and total rainfall from May to Aug, neither 2006 or 2009 were extreme weather years compared to the years spanning 1997–2010 (Appendix S1 Fig. S2). A summary of the analyses carried out here is presented in Appendix S2 Table S2.

Discussion

Community structure and non-reproductive species

We found evidence that the inclusion of non-reproductive species in plot censuses can have a significant impact on assessments of species richness, evenness and β -diversity. Differences were inconsistent among the two sites, highlighting the potential for census method to play an unpredictable but powerful role in understanding multiple aspects of community structure in different communities (Fig. 1).

The two census methods do not uniformly impact plots with different levels of diversity (Fig. 1e, f). While plot-level diversity is strongly positively related for the two census methods, the slopes of these relationships were significantly different across the two study sites. This result clearly indicates that the Full census is not a reliable estimate of the Flower census, although it is possible that one of the two sites studied here is anomalous.

Assessments of plot-level diversity are used to assess species co-existence. For example, diversity–productivity studies are generally pursued with the goal of understanding how diversity relates to niche differentiation. Given that the Flower census requires multiple censuses across the growing season, it is likely that most diversity–productivity studies have included non-reproductive species. It is apparent from our results that the Flower census, which enumerates only reproductive species in plots, may be a more reasonable estimate of species co-existence in sample plots. For example, some of the

considerable variation in diversity–biomass relationships observed in natural systems (Grace et al. 2007; Adler et al. 2011) may result from site-to-site variation in how similar community measures are for the two census methods. The variable contribution of non-reproductive species across communities, along with the potential for these to contribute inconsistent and potentially unreliable evidence of plot-level co-existence, suggests that comparisons of diversity across communities would benefit from the exclusion of non-reproductive species from censuses (i.e. using the Flower census).

Co-occurrence patterns and non-reproductive species

Including non-reproductive species in plot censuses (Full census) resulted in more pronounced patterns of negative co-occurrence in both sites (Fig. 2a, b, e, f). This pattern was consistent regardless of whether the co-occurrence null model examined presence–absence or abundance data, or when a pair-wise co-occurrence test was conducted. These large differences indicate how consequential the decision to include or exclude non-reproductive species in plot censuses is when examining patterns of co-occurrence, regardless of what co-occurrence analysis is used. The impact of non-flowering species on co-occurrence was more pronounced when abundance data were used. One possible explanation is that because IT randomization scheme does not fix species richness per plot, rare species may play a more powerful role in this null model than in the fixed–fixed null model.

An important question is just how common and how pronounced significant negative co-occurrence patterns are (e.g. Gotelli & McCabe 2002; Götzenberger et al. 2012). Analyses of multiple plant communities, such as those by Gotelli & McCabe (2002) for example, could greatly change depending on whether non-reproductive species were included in plot-level censuses. The degree to which non-flowering species may affect our understanding of co-occurrence patterns clearly depends on whether communities more closely resemble Site 1, where non-reproductive species are very common, or Site 2, where they are less so.

The Flower census method is a conservative estimate of co-occurrence patterns, requiring a higher standard of evidence for co-occurrence in a given site/plot. In a sense, this is similar to the use of the Checker metric, which has a more stringent requirement for co-occurrence, but also has undesirable type II error rates compared with the *C*-score metric (Gotelli 2000). As such, for plant communities, using the *C*-score metric with the Flower census method retains the more desirable error rates associated with the *C*-score, while maintaining the higher evidentiary standard for co-occurrence that is characteristic of the

Checker metric. It is clear the difference can be substantial (Fig. 2a).

Exploring the source of non-reproductive species

For Site 1, non-reproductive species strongly reduced the degree to which species pairs formed perfect checkerboards (Fig. 2c), consistent with what we would expect if non-reproductive species contribute noise to a system in which interspecific interactions are contributing to negative co-occurrence patterns. Additionally, species with reproductive representatives in Site 1 are more than 20 times more abundant in plots than species which did not have any reproductive representatives, suggesting strongly that these non-reproductive species may not be capable of success in competition with other species in these plots (Fig. 3). It is worth clarifying here that this is a per-plot analysis; as such, low abundance within plots is not related to rarity in the community (i.e. low plot occupancy; Fig. 1c, d). For example, *Trifolium aureum* occurred in 82% of plots in Site 1, but counting only species with reproductive representatives led to its removal from 29 plots. In fact, there is a positive correlation between the number of plots a species is found in, and the number of plots that species occurs in only as a non-reproductive plant ($r = 0.4$, $P = 0.002$). Thus, non-reproductive species are not simply subordinate species throughout the community, but are subordinate within particular plots. As such, the apparent challenge posed by non-reproductive species cannot simply be addressed by focusing on common species. Furthermore, while these species are rare in plots, they are not the only rare species in plots. Plots also contain rare (i.e. low-abundance) species that are reproducing. Removal of these species would inappropriately reduce the signal of negative co-occurrence in the community. These patterns of checkerboard disruption and abundance of reproductive and non-reproductive species are consistent with the interpretation that non-reproductive species are present as a result of local mass effects, at least in Site 1, and inconsistent with other explanations, including that non-reproductive species are young, and accurately reflect species co-existence/co-occurrence patterns. We could not calculate similar plot-level abundance comparisons for Site 2, and non-reproductive species, of which there were significantly fewer per plot in Site 2 (Fig. 1b), did not reduce the number of perfect checkerboards among species pairs for this site (Fig. 2d). This clearly reflects that variable contribution of non-reproductive species in different communities.

We found no support for several explanations for non-reproductive species in plots. Differences in non-reproductive species in plots across our sites did not correspond to

differences in the number of species that produce rhizomes or stolons. At this time, clonality does not appear to be responsible for non-reproductive species; however, more research is necessary to confirm this. Additionally, in Site 1, available N was positively correlated with the number of non-reproductive species in plots, which is not supportive of the notion that abiotic stress was responsible for non-reproductive plants. In fact, this pattern is more consistent with competitive suppression under higher resources. Finally, neither site was censused during an unusually stressful weather year, as indicated by how 2006 and 2009 fit into distributions of yearly temperature and precipitation data for the region (Appendix S1, Fig. S2).

There may be other explanations for non-reproductive species in plots; however, at this time, the combination of low abundance of non-reproductive species relative to others, as well as a lack of support for alternatives, suggest a role for mass effects in delivering non-reproductive species to plots, particularly in Site 1. Furthermore, variation in the effect of non-reproductive species on community structure and co-occurrence patterns among our two sites is consistent with the expectation that the role of mass effects will differ across communities. Based on our results, we hypothesize that local dispersal (i.e. mass effects) is a candidate to explain many non-reproductive species. We anticipate this work will stimulate further research exploring how variable the contribution of non-reproductive species is across a larger selection of communities, and whether mass effects are primarily responsible for non-reproductive species in plant communities.

Conclusions

This work considers the contribution of non-reproductive species to plant community structure, recognizing a previously unstudied component of plant diversity. We show conclusively that the inclusion of non-reproductive species in plot censuses can strongly impact measures of plot-level diversity, relative abundance and negative co-occurrence patterns, and that this impact can vary across communities. For Site 1, it is clear that non-reproductive species are significantly less abundant in plots than species that have reproductive representatives and that the inclusion of non-reproductive species greatly reduces the number of perfect checkerboards among species pairs for Site 1, a pattern expected if these plants contribute noise rather than a true signal of co-occurrence patterns. These patterns, in conjunction with a number of unsupported alternative explanations, suggest a role for local mass effects that requires further investigation. Our results, and particularly their variability across two nearby communities, also suggest that the inclusion of non-reproductive species in plant

community censuses may impair our ability to clearly compare community structure across abiotic gradients or communities.

Acknowledgements

We thank M. Neytcheva, J. Chambers, J. Waugh, E. Thompson, D. Kleiman, A. Tracey and S. Dombroskie for help with the collection of field data, and M. Lavender for helpful comments. This research was supported by the Natural Sciences and Engineering Research Council of Canada grants to LW Aarssen and BS Schamp, and by an Ontario Graduate Scholarship to BSS.

References

- Aarssen, L.W. & Epp, G.A. 1990. Neighbor manipulations in natural vegetation: a review. *Journal of Vegetation Science* 1: 13–30.
- Adler, P.B., Seabloom, E.W., Borer, E.T., Hillebrand, H., Hautier, Y., Hector, A., Harpole, W.S., O'Halloran, L.R., Grace, J.B. (...) & Yang, L.H. 2011. Productivity is a poor predictor of plant species richness. *Science* 333: 1750–1753.
- Anderson, M.J. 2006. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62: 245–253.
- Bell, G. 2005. The co-distribution of species in relation to the neutral theory of community ecology. *Ecology* 86: 1757–1770.
- Benjamini, Y. & Yekutieli, D. 2001. The control of false discovery rate in multiple testing under dependency. *Annals of Statistics* 29: 1165–1188.
- Connor, E.F. & Simberloff, D. 1979. The assembly of species communities: chance or competition? *Ecology* 60: 1132–1140.
- Diamond, J.M. 1975. Assembly of species communities. In: Cody, M.L. & Diamond, J.M. (eds.) *Ecology and evolution of communities*, pp. 342–444. Harvard University Press, Cambridge, MA, US.
- Dullinger, S., Kleinbauer, I., Pauli, H., Gottfried, M., Brooker, R., Nagy, L., Theurillat, J.P., Holten, J.L., Abdaladze, O. (...) & Garbherr, G. 2007. Weak and variable relationships between environmental severity and small-scale co-occurrence in alpine plant communities. *Journal of Ecology* 95: 1284–1295.
- Fayle, T.M. & Manica, A. 2010. Reducing over-reporting of deterministic co-occurrence patterns in biotic communities. *Ecological Modelling* 221: 2237–2242.
- Flinn, K.M., Gouhier, T.C., Lechowicz, M.J. & Waterway, M.J. 2010. The role of dispersal in shaping plant community composition of wetlands within an old-growth forest. *Journal of Ecology* 98: 1292–1299.
- Gleason, H.A. & Cronquist, A. 1991. *Manual of the vascular plants of northeastern United States and adjacent Canada*. New York Botanical Garden Press, New York, 986 pp.
- Gotelli, N.J. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81: 2606–2621.

- Gotelli, N.J. & Colwell, R.K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4: 379–391.
- Gotelli, N.J. & Entsminger, G.L. 2001. Swap and fill algorithms in null model analysis: rethinking the knight's tour. *Oecologia* 129: 281–291.
- Gotelli, N.J. & McCabe, D.J. 2002. Species co-occurrence: a meta-analysis of J.M. Diamond's Assembly Rules Model. *Ecology* 83: 2091–2096.
- Gotelli, N.J. & Ulrich, W. 2010. The empirical Bayes approach as a tool to identify non-random species associations. *Oecologia* 162: 463–477.
- Gotelli, N.J. & Ulrich, W. 2012. Statistical challenges in null model analysis. *Oikos* 121: 171–180.
- Götzenberger, L., de Bello, F., Brathen, K.A., Davison, J., Dubuis, A., Guisan, A., Leps, J., Lindborg, R., Moora, M. (...) & Zobel, M. 2012. Ecological assembly rules in plant communities – approaches, patterns and prospects. *Biological Reviews* 87: 111–127.
- Grace, J.B. 1999. The factors controlling species density in herbaceous plant communities: an assessment. *Perspectives in Plant Ecology, Evolution and Systematics* 2: 1–28.
- Grace, J.B., Anderson, T.M., Smith, M.D., Seabloom, E. & Andelman, S.J. 2007. Does species diversity limit productivity in natural grassland communities? *Ecology Letters* 10: 680–689.
- Heino, J. 2013. Environmental heterogeneity, dispersal mode, and co-occurrence in stream macroinvertebrates. *Ecology and Evolution* 3: 344–355.
- Heino, J. & Grönroos, M. 2013. Does environmental heterogeneity affect species co-occurrence in ecological guilds across stream macroinvertebrate metacommunities? *Ecography* 36: 926–936.
- Kikvidze, Z., Pugnaire, F.I., Brooker, R.W., Choler, P., Lortie, C.J., Michalet, R. & Callaway, R.M. 2005. Linking patterns and processes in alpine plant communities: a global study. *Ecology* 86: 1395–1400.
- Lanyon, S.M. 1985. Detecting internal inconsistencies in distance data. *Systematic Zoology* 34: 397–403.
- Manly, B.F.J. 1995. A note on the analysis of species co-occurrences. *Ecology* 76: 1109–1115.
- Pärtel, M., Zobel, M., Zobel, K. & van der Maarel, E. 1996. The species pool and its relation to species richness: evidence from Estonian plant communities. *Oikos* 75: 111–117.
- Pielou, D.P. & Pielou, E.C. 1968. Association among species of infrequent occurrence: the insect and spider fauna of *Polyporus betulinus* (Bulliard) Fries. *Journal of Theoretical Biology* 21: 202–216.
- Reitalu, T., Prentice, H.C., Sykes, M.T., Lönn, M., Johansson, L.J. & Hall, K. 2008. Plant species segregation on different spatial scales in semi-natural grasslands. *Journal of Vegetation Science* 19: 407–416.
- Sanders, N.J., Gotelli, N.J., Heller, N.E. & Gordon, D.M. 2003. Community disassembly by an invasive species. *Proceedings of the National Academy of Sciences of the United States of America* 100: 2474–2477.
- Sanderson, J.G., Moulton, M.P. & Selfridge, R.G. 1998. Null matrices and the analysis of species co-occurrences. *Oecologia* 116: 275–283.
- Schamp, B.S., Arnott, S.E. & Joslin, K.L. 2015. Dispersal strength influences zooplankton co-occurrence patterns in experimental mesocosms. *Ecology* 96: 1074–1083.
- Schoener, T.W. 1983. Field experiments on interspecific competition. *The American Naturalist* 122: 240–285.
- Sfenthourakis, S., Tzanatos, E. & Giokas, S. 2004. From sampling stations to archipelagos: investigating aspects of the assemblage of insular biota. *Global Ecology and Biogeography* 13: 23–35.
- Smith, B. & Wilson, J.B. 1996. A consumer's guide to evenness indices. *Oikos* 76: 70–82.
- Stone, L. & Roberts, A. 1990. The checkerboard score and species distributions. *Oecologia* 85: 74–79.
- Turnbull, L.A., Crawley, M.J. & Rees, M. 2000. Are plant populations seed-limited? A review of seed sowing experiments. *Oikos* 88: 225–238.
- Ulrich, W. 2004. Species co-occurrences and neutral models: reassessing J. M. Diamond's Assembly Rules. *Oikos* 107: 603–609.
- Ulrich, W. & Gotelli, N.J. 2010. Null model analysis of species associations using abundance data. *Ecology* 91: 3384–3397.
- Veech, J.A. 2013. A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* 22: 252–260.
- Weiherr, E. & Keddy, P.A. 1995. Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos* 74: 159–164.
- Zalewski, M. & Ulrich, M. 2006. Dispersal as a key element of community structure: the case of ground beetles on lake islands. *Diversity and Distributions* 12: 767–775.

Supporting Information

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

Appendix S1 (Figures S1–S2):

Comparison of co-occurrence patterns when species with seed additions are included for Site 1.

Histograms illustrating how temperature and rainfall data in the study years fit into longer-term patterns for that region.

Appendix S2 (Tables S1–S2):

A list of non-local species seeded into Site 1 prior to the census year.

A summary of analyses performed.