little attention has been devoted to understanding the under-

aspects of communities through time and across space (Ackerly 2003, Blois et al. 2013a, Dalsgaard et al. 2013). However, interactions among species also have played a major role in structuring community composition and functioning (Jablonski 2008, Blois et al. 2013c, Wisz et al. 2013). Given recent interest in understanding how climate change may lead to new biotic interactions and unexpected ecological dynamics (Zarnetske et al. 2012, Blois et al. 2013c), there is a critical need to disentangle the joint e ects of abiotic and biotic factors on community dynamics.

Previous work on assemblage structure has quantified community pattern as a single index – such as the number of species or the number of checkerboard pairs – that is then subject to null model analysis. Even a moderately-sized assemblage (e.g. a dataset with multiple sites and multiple species at each site) contains many potential species pairs, however, each of which may exhibit positive, negative, or random associations. In many cases, single metrics that summarize an entire assemblage can be deceptive (Ulrich and Gotelli 2012), and it is more instructive to analyze individual pairs of species (Sfenthourakis et al. 2006). Gotelli and Ulrich (2010) use an empirical Bayes approach (Efron 2005) to control for the potentially large number of false positives that can emerge with the analysis of many species pairs. is kind of analysis allows for a determination of the relative frequency of positively, negatively, and randomly associated species pairs. However, non-random species associations are not necessarily caused by species interactions. us, a central dilemma is how to distinguish non-random species associations produced by actual species interactions from those produced by environmental filtering or dispersal limitations. All three processes can operate singly or in concert to generate both positive and negative species associations.

Ecologists working with modern faunas often explicitly or implicitly limit comparisons to a set of environmentally similar and spatially adjacent sites for which dispersal limitation is unlikely to be important (Phillips et al. 2003, Zhang et al. 2011). In such systems, it is reasonable to attribute non-random species associations to species interactions. e e ects of species interactions certainly can be scaled up to larger spatial and temporal domains (Jablonski 2008, Gilman et al. 2010, Baiser et al. 2012, Blois et al. 2013c), and MacArthur (1972) argued explicitly for this scaling in his final book, Geographical ecology. However, the e ects of environment, dispersal, and history become progressively more important at larger spatial and temporal scales, and it is di cult to untangle them from the e ects of species interactions (Ricklefs 2004).

is dilemma is illustrated clearly in fossil records. ese records usually encompass timescales at which environment, dispersal, and biotic interactions are all potentially important controls on species distributions and the associations among species, yet usually their e ects cannot be directly observed. Often we have information only about species occurrences across space and through time, and perhaps information about past environments. Only rarely can we infer actual biotic interactions in fossil systems (Wilf et al. 2001, Kowalewski 2002, Currano et al. 2010, Peñalver et al. 2012, Blois et al. 2013c), making it di cult to confidently attribute the causes of past species associations to the influence of environmental similarity, interactions with other species, or other factors.

Here, we provide a framework for inferring the importance of biotic interactions, dispersal limitation, and abiotic e ects on positive and negative species associations. is framework can be applied to species associations measured at any spatial or temporal scale, but we illustrate it in an analysis of eastern North American plant assemblages based on fossil pollen data from the past 21 000 yr. Previous work on both individual species and communities has demonstrated that changes in fossil pollen assemblages across space and time are tightly linked with climate, particularly in the latest Pleistocene and early Holocene (Grimm et al. 1993, Williams et al. 2002, Shuman et al. 2004, Yu 2007, Blois et al. 2013a). Indeed, the tight linkages between vegetation and climate make fossil pollen data an excellent proxy for reconstructing past climates (Viau et al. 2006, Bartlein et al. 2011). Additionally, the recognition of individualistic species responses to deglaciation, the resulting formation of no-analog communities during the Pleistocene–Holocene transition, and the attribution of these communities to no-analog climates (Williams et al. 2004) suggest that species interactions should not be dominant drivers of community patterns, especially during the height of the no-analog period from 17–11 kyr BP. Other evidence suggests that not all changes in vegetation can be attributed exclusively to climate. For example, at several sites in the Great Lakes region, the loss of megaherbivores and their associated species interactions at the end of the Pleistocene may have contributed to the formation of no-analog plant

Methods

Theoretical framework

Null model analyses of species co-occurrence

A large literature on null model analyses of species co-occurrence has accumulated over the past 80 yr (Harvey et al. 1983, Gotelli and Graves 1996). e initial impetus for these analyses was to ask whether co-occurrence patterns

many statistical tests is quite high. A similar problem arises in the analysis of microarrays, in which the expression levels of thousands of potentially non-independent genes are assayed with parametric or non-parametric statistical tests (Kammenga et al. 2007). For null model analysis of the co-occurrence of individual species pairs (Gotelli and Ulrich 2010), we adapted an empirical Bayes approach originally proposed by Efron (2005) for this problem of screening large numbers of non-independent tests. In brief, *C*-scores for each species pair are rescaled to a [0,1] range and binned into histogram categories. Next, the simulated data are binned in a similar way, and the mean and 95% confidence interval of the *C*-scores of simulated species pairs in each bin is calculated. Finally, the original *C*-score values within each bin are ordered from smallest to largest *C*-scores. For the Pairs analysis, pairs of species are retained whose *C*-scores are above the simulated mean for the bin (Bayesian mean criterion), and which would be statistically significant if the species pair was treated as an independent test. is 'double screen' reduces some of the false positives that would arise by simply retaining all species pairs for which the uncorrected association (aggregated or segregated) yielded $p < 0.05$. For the bins that are near 0.0, these largest *C* scores will represent aggregated species pairs. For the bins that are near 1.0, these largest *C* scores will represent segregated species pairs.

is is less conservative than a cut-point based on the 95% confidence interval for bin deviations, but more conservative than an unadjusted count of significant pairs, and usually more conservative than a sequential Bonferroni correction in which the pairs are ordered by their p values and a cuto is imposed that is determined by both the individual p-value and its rank. Benchmark tests of the Pairs algorithm show that it is e ective (though not perfect) at controlling for false positives while still allowing for detection of a relatively small subset of non-random species pairs from a binary presence–absence matrix (Gotelli and Ulrich 2010). We ran the Pairs analysis for each data matrix to identify the subset of species pairs that exhibited strong aggregation or segregation.

Identifying the causes of non-randomness

Null model analysis has been a successful tool for identifying non-random patterns of species associations. But the analysis cannot, by itself, point to the causes of such segregated or aggregated patterns. Here we consider explicitly two major classes of mechanisms that might lead to non-random associations of species pairs: dispersal limitations and habitat or climate (environmental) filtering of species into groups with similar environmental niches. All significant pairs that did not show signals of significant environmental variation or dispersal limitation may provide evidence of a significant species interaction, though it is also possible that environmental factors not considered in this analysis could contribute to non-random associations. To infer the roles of environmental factors and dispersal limitation, we move beyond the results of the standard null model tests with additional analysis of the characteristics of the sites. As described below, we focus specifically on subsets of sites that di er significantly in either environmental characteristics or spatial location.

Site classifications

For any particular pair of non-random species, each site can be assigned to one of four mutually exclusive classes, based on the presence of one species (1,0) or (0,1), both species (1,1) or neither (0,0) (Fig. 1). e average characteristics of sites assigned to these 4 classes will generate additional patterns that can be used to distinguish among di erent causes of non-randomness. If the site characteristics are measured as continuous variables (such as average annual precipitation), sets of sites can be compared with ANOVA or t-tests. If the site characteristics are measured as discrete variables (such as depositional environment for fossil materials), sets of sites can be compared with a 2-way contingency table analysis. In each case, the analysis will pinpoint whether characteristics of sites vary systematically based on the presence or absence of each species member in the pair.

For segregated species pairs, the critical comparison is between sites that have one species (1,0) and sites that have the other species (0,1) (i.e. allotopic sites). If species interactions are the critical factor in producing segregation, these two classes of allotopic sites should not di er systematically in either their environmental characteristics or their spatial arrangement. Such species interactions might include pairwise competition or predation, but also might reflect indirect e ects of other species. For aggregated species pairs, the critical comparison is between sites that have both species (1,1) (i.e. syntopic sites) and sites that have neither species (0,0). If species interactions are important in producing aggregation, these syntopic and empty sites should not di er systematically in either their environmental characteristics or their spatial arrangement. Such pairwise interactions might include pairwise mutualism or commensalism (or even predation), but also might reflect indirect e ects of other species.

Environment tests

For each site within a data matrix, we have di erent measures of environment, either continuous (e.g. annual precipitation or temperature, as in this study; see Climate and distance data) or categorical (e.g. soil type or depositional environment). In the case of continuous measures, a one-way ANOVA can be used to compare the environment between the allotopic sites of segregated pairs ((1,0) vs (0,1)) and between the syntopic sites and empty sites of aggregated pairs ((1,1) vs (0,0)). e null hypothesis is that site characteristics do not di er systematically between these pairs of site classifications. For the categorical measures, a two-way contingency table can be used to classify the sites. For the segregated species pairs, we counted the frequency of each environmental type for the two kinds of allotopic sites ((1,0) and (0,1)). For the aggregated species pairs, we counted the frequency of each environmental type for the syntopic sites (1,1) and the empty sites (0,0). For both kinds of two-way data tables, we used a chi-square test of association. e null hypothesis was that the frequencies of di erent environmental types did not di er among the site classes. If this null hypothesis is rejected, a parsimonious interpretation is that environmental associations are at least partly responsible for segregated or aggregated patterns of species occurrence (Fig. 1, 2).

Figure 1. Hypothetical patterns of species associations on the landscape under nine scenarios.

Random

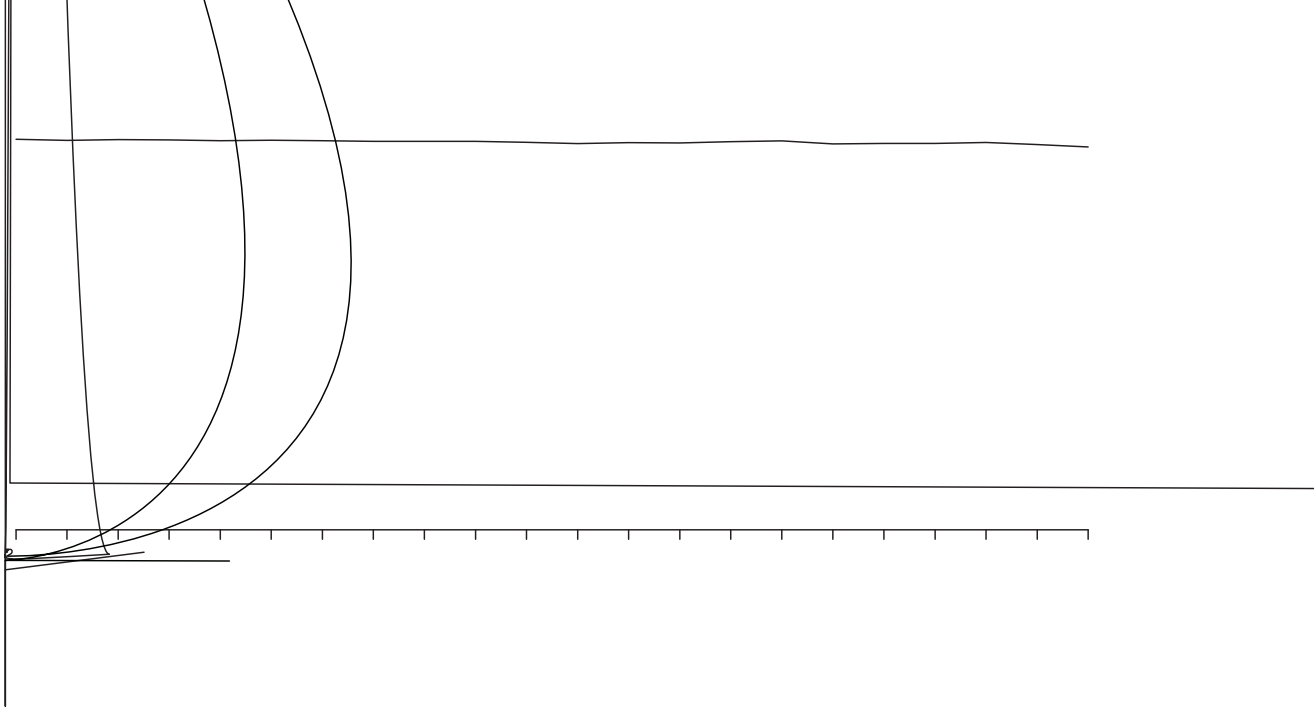

Table 2. The proportion of segregated and aggregated pairs that can be attributed to climate, dispersal limitation, both, or neither process through time. Note that the total number of segregated or aggregated pairs per time slice may be different than the totals in Table 1 because statistical significance could not be assessed for some taxon pairs.

Age (kyr BP)	Proportion of segregated pairs				Proportion of aggregated pairs			
	Distance	Climate	Both	Neither	Distance	Climate	Both	Neither
0	0	0	1	0	0	0.1	0.9	0
1	0	0	1	0	0	0	1	0
2	0	0	1	0	0	0	0.9474	0.0526
3	0	0	1	0	0	0	1	0
4	0	0	1	0	0	0	1	0
5	0	0	1	0	0	0	0.96	0.04
6	0	0	1	0	0	0	1	0
7	0	0	1	0	0	0	0.9375	0.0625
8	0	0	1	0	0	0	0.92	0.08

Table 3. The taxon pairs involved in significant segregations or aggregations through time.

Age (kyr BP)	Segregated pairs	Aggregated pairs
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		

unsurprising. e method of testing individual pairs imposes a strong screen for type I error, but more importantly, most taxa occur in relatively few sites. For good statistical reasons, it is di cult to assert that segregated pairs are non-random when both members of the pair are relatively rare, though we can detect aggregation more easily in this case. But without any additional evidence, the most parsimonious interpretation of the observation that two rare species do not co-occur frequently is that the pattern is due to chance. us, our approach is an inherently conservative method to begin with, but avoids falsely attributing biological processes to patterns that are more parsimoniously accounted for by simple sampling properties of the data.

Of the non-random subset, there were more aggregated than segregated pairs in most time periods (Table 1, Fig. 3). ese results for fossil assemblages form an interesting contrast with a recent meta-analysis of pairwise associations in 272 presence–absence matrices for modern assemblages (Gotelli and Ulrich 2010). In modern assemblages, most species pairs also showed random associations, although these tended to be concentrated in data matrices from a relatively small number of studies. However, the non-random fraction for modern assemblages was dominated by segregated species pairs, with a 4-fold excess of perfectly segregated checkerboard pairs compared to the most

If we had imposed a false discovery screen for both the environmental and spatial overlap tests, we might have found more pairs that would be classified as examples of biotic interactions. However, most taxon pairs showed strong e ects of both spatial and environmental segregation, so there would not be that many pairs classified as non-significant by both tests after screening for false positives.

Taxon aggregations were also infrequent, but more common than segregations. In 12 of 22 time slices there were taxon pairs for which aggregations could not be attributed to climate or large-scale spatial overlap between syntopic versus empty sites. ese were most common between 16 and 11 kyr BP. Such pairs might reflect positive biotic interactions such as direct mutualisms, indirect e ects such as shared pollinators or exclusion from the same sites due to a shared competitor or predator, or unmeasured habitat or climatic associations. e occurrence of potential positive biotic interactions in the latest Pleistocene could also provide sup-

Cardillo, M. and Meijaard, E. 2010. Phylogeny and co-occurrence of mammal species on southeast Asian islands. – Global Ecol. Biogeogr. 19: 465–474.

implications for species distribution modelling. – Biol. Rev. 88: 15–30.
Yu, Z. 2007. Rapid response of fore