A TRAIT-BASED TEST FOR HABITAT FILTERING: CONVEX HULL VOLUME

WILLIAM K. CORNWELL,1,4 DYLAN W. SCHWILK,2 AND DAVID D. ACKERLY3

1Department of Biological Sciences, Stanford University, Stanford, California 94305 USA
2U.S. Geological Survey, Western Ecological Research Center, Sequoia-Kings Canyon Field Station, 47050 Generals Highway, Three Rivers, California 93271 USA
3Department of Integrative Biology, University of California, Berkeley California 94720 USA

Abstract. Community assembly theory suggests that two processes affect the distribution of trait values within communities: competition and habitat filtering. Within a local community, competition leads to ecological differentiation of coexisting species, while habitat filtering reduces the spread of trait values, reflecting shared ecological tolerances. Many statistical tests for the effects of competition exist in the literature, but measures of habitat filtering are less well-developed. Here, we present convex hull volume, a construct from computational geometry, which provides an \( n \)-dimensional measure of the volume of trait space occupied by species in a community. Combined with ecological null models, this measure offers a useful test for habitat filtering. We use convex hull volume and a null model to analyze California woody-plant trait and community data. Our results show that observed plant communities occupy less trait space than expected from random assembly, a result consistent with habitat filtering.

Key words: chaparral; community assembly; convex hull; habitat filtering; plant; seed mass; specific leaf area; trait; wood density.

INTRODUCTION

Ecological theory has proposed a view of community assembly as a series of nested processes, described as "rules" or "filters," that leads to a certain suite of species coexisting at a given site (Diamond 1975, Weiher and Keddy 1999). These processes, which integrate interactions among species as well as between species and the abiotic environment, can be coarsely divided into two categories; competition and habitat filtering. While both may act simultaneously, these two assembly processes have different effects on the ecological strategies, and thus the functional traits, of coexisting species.

Classical competition theory predicts a limit to the ecological similarity of coexisting species (MacArthur and Levins 1967). Many authors have argued that a species’ morphological traits can be used to infer its ecological strategy (see Ricklefs 1990). Therefore an assembly process that results in a limit to the similarity of coexisting species will lead to an even spacing of species across strategy axes, which can be tested empirically as even-spacing in trait space (Ricklefs and Travis 1980, Stubbs and Wilson 2004). The presence of competition can then be detected as deviations from a null model of assembly (Gotelli and Graves 1996). The multivariate nature of this question was recognized early, and multivariate statistical tests for limiting similarity are well-developed in the literature (Stubbs and Wilson 2004, and citations therein).

Habitat filtering can be thought of as a reduction in the range of successful strategies among coexisting species (van der Valk 1981, Keddy 1992, Weiher et al. 1998, Weiher and Keddy 1999). For example, in alpine environments species without traits that convey freezing tolerance are excluded. Species with those traits are successful, and, as a result, the community of alpine species is homogenous with respect to these traits when compared to a regional pool. Analogous to limiting similarity, habitat filtering is a multivariate question, which may be reflected in the range, area, or volume of traits, depending on the dimensionality of trait space considered. For example, Reich et al. (1999) suggest that ecological processes prevent species with high nitrogen but low photosynthetic rates from being successful, and McClain et al. (2004) argued that in some habitats ecological processes restrict the volume of morphospace occupied by benthic gastropods. However, statistical tests for habitat filtering as an ecological process are not well developed (see Weiher and Keddy 1999).

In this report, we present the convex hull volume, a multivariate measure derived from computational geometry (Preparata and Shamos 1985), as the basis for a statistical test for habitat filtering. Convex hull volume quantifies the volume of trait space occupied by species in a community regardless of the shape of the distribution and represents the multivariate equivalent of range. Convex hull volume is easily extensible from one to \( n \) dimensions, and, in combination with null
models, it can easily take into account the multivariate trait correlation structure in the species pool.

As a test case, we use this new measure to characterize the volume of trait space for California woody-plant communities. To test for nonrandom assembly processes in our empirical data, we compare the observed convex hull volumes of communities to results from a null model of random assembly from a species pool. Specifically, we examine the following question: Is the trait volume of California woody-plant communities significantly less than expected by chance? In other words, is there evidence for the effects of multivariate habitat filtering in this natural system?

**METHODS**

**Quantifying trait volume: convex hull volume**

The multivariate range of ecological strategies displayed by a set of species in an $n$-dimensional trait space is the volume of the space occupied by the species. Although many shapes could be used to enclose the points (e.g., the smallest cube or sphere), the convex hull, defined as the smallest convex set enclosing the points (Preparata and Shamos 1985, Barber et al. 1996), is the most reasonable as it reduces the amount of empty space compared to cubes or spheres (see Fig. 1).

Calculation of convex hulls is a well-studied problem in computational geometry and the method has diverse application in computer graphics. In addition, two-dimensional convex polygons have been used in ecology to estimate species ranges from point observations (e.g., Kerley et al. 2002). In a multivariate trait space, the following criterion defines which portions of the space should be found within the convex hull: if in an $n$-dimensional trait space we find two individuals $a$ and $b$ in our community, one with trait values $(a_1, a_2, \ldots, a_n)$ and the other with trait values $(b_1, b_2, \ldots, b_n)$, then for any value $t$ with $0 \leq t \leq 1$, an individual with trait values 

$$[ta_1 + (1-t)b_1, ta_2 + (1-t)b_2, \ldots, ta_n + (1-t)b_n]$$

will also be in our convex hull volume. For example, when $t = 0.5$, the above statement translates to: for any two individuals $a$ and $b$ in our trait space, the individual whose trait values are the means of $a$ and $b$ will also be in our estimate of the ecologically “allowed” part of trait space.

Finding $n$-dimensional convex hulls is a complex computational problem. We use the Qhull program of Barber and Huhdanpaa (“Quickhull” algorithm described in Barber et al. [1996]) to calculate the convex hull of sets of multidimensional points and to calculate the volume (or hyper-volume) of that hull. We wrote a script interface to Qhull that allows the user to input trait and plot data easily; our program, TraitHull, is available online.$^5$ The Quickhull algorithm is also implemented in Matlab version 6.x or higher (see function: convhull).

**Study system**

The field sampling for this project occurred at Jasper Ridge Biological Preserve. The 481-ha preserve, located in the eastern foothills of the Santa Cruz Mountains, California, USA, contains a diverse array of woody vegetation types including broadleaf evergreen forest and chaparral. The climate is mediterranean-type, with mean annual rainfall of 605 mm, ~80% of which falls between November and March.

In any study of the ecological role of functional traits, trait selection is crucial. Four traits (Appendix A) were selected from the largely orthogonal suites of traits identified in previous work on plant trait variation (Westoby et al. 2002, Diaz et al. 2004). Specific leaf area (SLA) is part of a suite of traits associated with leaf life span and the “leaf economics” spectrum of fast-to-slow resource capture (Reich et al. 1999, Wright et al. 2004). Wood density has been shown to correlate with resistance to embolism (Hacke et al. 2001) and to be part of a suite of traits related to water relations and minimum seasonal water potential (Ackerly 2004). Seed
mass is a crucial feature of dispersal and regeneration ecology (Coomes and Grubb 2003). Leaf area is related to energy balance of leaves (Givnish 1987) and to a suite of morphological traits known as Corner’s rules (Ackerly and Donoghue 1998). Here we hypothesize that specific combinations of these four functional traits that will not be viable in each habitat, leading to a reduction in the multivariate range at a given site.

**Trait and plot sampling**

Using ArcGIS (Environmental Systems Research Institute, Redlands, California, USA), we randomly located 44 20 × 20 m plots across the parts of the preserve dominated by woody vegetation. To ensure adequate sampling of all of the common vegetation types, we required a minimum of five plots to be located in each of five woody vegetation types as determined by a previous vegetation map drawn from aerial photographs. Fifty-four native species occurred in the survey with a median plot species richness of 12 and a maximum richness of 17 species. One plot contained only three species, which precludes four-dimensional analysis. It was excluded from the analyses described below.

We measured specific leaf area (fresh area/dry mass) and individual leaf area for two individuals of each species in each of the 44 plots (a total of >1500 measurements). To characterize species’ wood density, which is more laborious, we sampled 3-yr-old shoots from five representative individuals spread out across the species distribution. We removed the pith, phloem, and bark, measured fresh volume by water displacement and determined dry mass after drying for >2 d at 70°C. Reported values are oven dry mass/fresh volume. The seed mass data set was assembled from several sources including field collections, previous studies at Jasper Ridge (Ackerly et al. 2002, Ackerly 2004), and literature sources (Baker 1972, Keeley 1991, Young and Young 1992, USDA 2004, USFS 2004). In cases in which multiple measurements from different sources were available, we used the mean of all observations as a species mean.

**Null behavior of convex hull volume**

We explored the behavior of convex hull volume through two sets of randomization trials. First, we tested the effect of trait dimensionality by constructing pools of species with 1–5 orthogonal traits. The traits all had mean = 0 and standard deviation = 1 and were drawn from a normal distribution. We then drew species without replacement from these pools. At each species richness and for each dimensionality the draw was repeated 500 times. We could then compare the relationship between species richness and trait volume across different numbers of dimensions.

Second, we tested the effect of correlations among traits on convex hull volume. To do this, we generated species pools with four traits with differing degrees of colinearity (pairwise $r^2$ values ranging from 0 to 0.5). We then randomly sampled from these vectors and calculated convex hull volume as described above (see **Quantifying trait volume: convex hull volume**). In addition, we compared our empirical trait by species data ($z$ scores in four dimensions; range in $r^2 = 0.00–0.19$) to four orthogonal traits (mean = 0, $sd = 1$).

**Alternative models of trait scaling**

Functional traits are measured in different units and show different variances across species, and finding the most useful way to combine traits into one multivariate space is not straightforward (see Foote 1997). Moreover, the relative importance of a trait with respect to community assembly is difficult to determine, especially in natural systems, and is not necessarily directly linked with its observed variance. It is not appropriate to calculate convex hull on untransformed trait data, as a simple shift in measurement scale (e.g., grams to milligrams) will arbitrarily change the relative weighting given to a trait. For this study we explore three approaches to trait weighting: equal weighting, global weighting, and logarithmic weighting.

Traits may be weighted equally by standardizing each trait with respect to the mean and variance in the local species pool (using a $z$ transform to mean = 0, $sd = 1$; see Appendix A). The drawback to this approach is that the measure of trait volume is specific to the pool of species included in that study, and comparisons across studies are not possible. A second approach is global weighting, in which each trait is scaled relative to the global mean and variance for that trait. Recently compiled global data sets for plant traits now allow us to quantify the variation found across thousands of species (see Appendix A; Brown 1997, Wright et al. 2004, Moles et al. 2005). We use these data sets to scale each axis to the global mean and standard deviation in each trait and then proceed with the convex hull analysis. The advantage to this approach is that a range of 1 $sd$ of seed size would then be the same whether measured in California or Panama. A technical problem is that global data sets will continue to grow, and a single standard is required to maintain comparability across studies.

The third approach weights the traits by their observed variance on a logarithmic scale. Log-scaling makes measures of spread (variance, range, etc.) independent of the units in which they were measured, while preserving the different variances in the samples (see Appendix A). For example seed mass varies in this data set by almost six orders of magnitude while wood density varies by less than three-fold. Similarly to the global scaling, trait volume measures using log-scaled traits are independent of pool and can be compared across studies.

In the example below, the traits we chose were relatively orthogonal ($r^2 = 0.00–0.19$). If traits are more correlated, or in cases where the number of traits is large
relative to the number of species per plot, principal components could be calculated as a first step before finding the convex hull volume using a smaller number of principal component axes.

Species means and intraspecific variation

It is worth noting that most previous studies of community assembly and functional traits have relied on species means as the estimate of a species trait value at every site in each study. One advantage of the convex hull method is that it allows us to incorporate intra-specific variation (due to plasticity and genotypic differentiation) into an analysis. We conducted a preliminary test, using leaf area and specific leaf area, the only two traits in this study with data for each species at each site. We tested whether trait volume, calculated with species means, differed from trait volume calculated with the actual measurements from each plot. In this example, we found no significant effect of incorporating intraspecific variation (results not shown). However, this may not be the case for all data sets.

Null model and significance testing

Our null model for community assembly was a random or lottery assembly process drawing from the overall pool of 54 native species found in the sampled plots at Jasper Ridge. The sample size and geographic range of the species pool has been shown to be an important consideration for studies of community assembly (Law and Morton 1996). In this case the small (481-ha) size and mosaic nature of the vegetation types at Jasper Ridge suggest that over ecological time scales all species would be capable of dispersal to any given site.

Convex hull volume and other statistical measures of the spread or range of a variable are highly correlated with sample size (= species richness) in both random draws and observed data. In order to incorporate this correlation into our null model, we conducted 500 random draws at each species richness. Observed communities with a given number of species were then compared to the null draws at the same species richness. This approach is known as the “RA3” algorithm and was first described by Lawlor (1980; see also Winemiller and Pianka 1990, Gotelli and Graves 1996).

We used the null model to generate expectations for convex hull volume and the range and variance of individual traits. Significance testing was done on a collective basis for the 43 plots and separately for the two most common habitats, chaparral and broadleaf evergreen forest. We used a paired nonparametric test (two-tailed Wilcoxon signed-ranks test) to test our empirical data against the expectation from the null model (Sokal and Rohlf 1995).

Results

Null patterns

Under the null model of random assembly of species, for fully orthogonal traits, the convex hull volume increases monotonically with species richness. The shape of this relationship depends on the dimensionality of the trait space considered, with convex hull volume in higher dimensional spaces showing an increasingly linear relationship with species richness (see Appendix C). With fewer dimensions, there are fewer ways that the addition of a new species can increase the range of observed traits, so the “rate of return” with addition of new species diminishes at higher species richness.

The effect of correlation between traits is to decrease the total volume of trait space available to species as well as to decrease the slope of the species richness-convex hull volume relationship and to increase the curvature as the convex hull saturates more quickly (see Appendix C). These changes become more pronounced as the correlation between traits increases, reflecting the effective decrease in dimensionality of correlated data.

Empirical vs. null patterns

Two-dimensional convex hulls for two representative plots are presented (see Appendix D). Cumulatively, in our samples of woody plant communities, trait volume is significantly lower than would be expected through random assembly (Fig. 2). Species in 40 out of 43 plots occupied less trait space than would be expected by chance. Using a two-tailed nonparametric test the distribution of plot data was highly significantly different from the expectation ($P < 0.001$). The same result and significance level was found for all three models of trait scaling. Two univariate ranges, SLA ($P < 0.001$) and leaf size ($P < 0.001$), and three univariate variances, SLA ($P < 0.001$), leaf size ($P < 0.001$), and wood density ($P = 0.014$), were also significantly less than the expectation.

For the analysis by vegetation class, we found significant habitat filtering in both the chaparral and broadleaf evergreen forest ($P < 0.002$ for both habitats) and no significant difference between these two habitats in terms of the magnitude of the effect ($t$ test, $P = 0.610$).

Using global scaling we found a mean volume of 2.09 (standard deviates to the fourth power) for the 20 $\times$ 20 m plots sampled in this study. Using log-scaling we calculated a mean volume of 0.082 log units raised to the fourth power. These results can be directly compared with parallel studies in other systems using the same set of traits. The results of pairwise and three-way combinations of traits are reported in Appendix B.

Discussion

In our study, we found that convex hull volume provides a quantititative, multivariate measure of the amount of trait volume occupied by species in a community. Combined with a null model of community assembly, convex hull volume provides a powerful method to test for the effect of habitat filtering. In the empirical example we found that woody plant species co-occurring in local communities occupy less multivariate functional trait space than would be expected by
chance (Fig. 2). This result is consistent with a habitat-level filter that limits the multivariate range of trait space that species can occupy at a given site (van der Valk 1981, Keddy 1992). Consistent with the multivariate result, the variances of three out of four univariate traits were less than expected by chance, and the range of two out of four was less than expected by chance. There was a significant reduction in convex hull volume in both chaparral and broadleaf evergreen forest and no significant difference in the magnitude of the effect between habitats, which suggests that for these traits the strength of habitat filtering did not change with habitat.

The individual traits that showed decreased range and variance, specific leaf area, leaf size, and wood density, are linked to species' functional strategies with respect to metabolism, energy balance, and water relations. Low specific leaf area is strongly associated with long leaf life span (Reich et al. 1999) and with an allocation strategy that increases photosynthetic capacity per unit leaf area (Wright and Westoby 2002). Leaf size influences the conductance of the leaf boundary layer. The reduced boundary layer effect for smaller leaves may lower leaf temperatures on exceptionally hot days (Givnish 1987). High wood density has been shown to correlate with resistance to drought-induced embolism (Hacke et al. 2001). The topographic gradients at Jasper Ridge lead to marked variation in moisture levels, thermal load, and light intensity (Ackerly et al. 2002). Thus, we interpret the results for these three traits related to physiological strategies, and their multivariate combination, as evidence for restrictions on the viable range of trait values in each sampled habitat and a shift in this range across habitats. The functional significance of trends in the mean trait values across abiotic gradients within Jasper Ridge are explored elsewhere (W. K. Cornwell and D. D. Ackerly, unpublished manuscript).

Trait scaling

Alternate models of trait scaling lead to highly correlated results across plots ($r^2$ between 0.993 and 1.000) and similar results relative to the null model. This result suggests that trait scaling may not have a large effect on the interpretation of results within a study. Across studies, however, scaling options that allow for quantitative comparison and meta-analysis of results will be critical to broader studies of the multivariate range of ecological strategies in natural systems. For example, scaling traits in a comparable way allows the question of whether the range of ecological strategies differs across global gradients in latitude or precipitation.

The log-scaling option is attractive for its simplicity and the lack of a required standard and is the preferred option in cases when a global data set of trait values is not available. However, it implicitly gives more weight to traits that vary over more orders of magnitude, and there is no a priori reason to believe that differences in the absolute range of variation correspond to intrinsic functional significance of the respective traits. The global-scaling option places equal weight to each trait (relative to the global variance), which also may not correspond directly to functional significance, but we believe it is less arbitrary than the range of log values. If global scaling is to be adopted by the plant ecology research community, an effort is required to establish global reference distributions for each trait, complementing recent efforts to standardize measurement protocols (see Cornelissen et al. 2003).
Related measures of trait dispersion

A previous measure of morphological disparity (Ricklefs and Travis 1980) that is based on the product of standard deviates of community samples on principal components analysis (PCA) axes. For our data set, this measure is moderately correlated with convex hull volume \( r^2 = 0.580 \); we interpret the lack of a tighter correlation as based in a conceptual difference between the two constructs. While convex hull volume increases monotonically with species richness in both randomly assembled and empirical communities, the Ricklefs and Travis method does not, and the so of PCA loadings can decrease at higher diversity if the added species are near the multivariate mean. Nonetheless, the Ricklefs and Travis measure is also significantly lower in our data relative to random assemblies \( P < 0.001 \) using the RA3 method.

The Ricklefs and Travis method and a related measure, the sum of the community ranges of PCA axes (Foote 1992), both essentially use a rectangular construct to quantify trait volume. We interpret convex hull volume as an advance on rectangular methods in that the convex hull more accurately quantifies trait volume by excluding the “missing corners” of triangular distributions (e.g., Fig. 1a). Triangular distributions for trait values have been documented in plant systems and may be common (Cornelissen 1999).

Convex hull volume as a measure of trait volume is also related to measures of evolutionary morphological diversification (Foote 1997) and to measures of functional diversity as it relates to ecosystem function (see Petchey et al. 2004). Unlike these measures of multivariate trait spread, convex hull volume focuses on the multivariate range of trait values, which is consistent with a model of habitat filtering as a community assembly process.

Conclusion

Recent work on plant functional traits has established important axes that differentiate plant strategies (Westoby et al. 2002), and further work has shown that species are arrayed along functional continua rather than in discrete groups (Diaz et al. 2004, Wright et al. 2004). Multivariate methods that utilize this continuous nature of functional traits in the context of community assembly will provide a more nuanced understanding of the processes that generate and maintain a spread in functional strategies. Furthermore, the convex hull measure can be combined with tests of limiting similarity to address the complementary questions of species interactions and limiting similarity (Ricklefs and Travis 1980, Stubbs and Wilson 2004).

Acknowledgments

Jasper Ridge Biological Preserve staff, R. Freund, E. Giffenbaum, and T. Carlson helped immensely during the fieldwork. I. Wright and A. Moles generously provided data from global data sets on leaf and seed traits. Thanks to M. Mayfield, N. Kraft, M. Boni, K. Preston, and two anonymous reviewers for comments that improved this manuscript. Funding was provided by an NSF graduate fellowship to W. K. Cornwall and by NSF 0078301 to D. D. Ackerly.

Literature Cited


APPENDIX A

Summary statistics for the distribution of trait values for both global data sets and Jasper Ridge (54 woody species) (Ecological Archives E087-085-A1).

APPENDIX B

Convex hull volumes calculated using trait subsets and log and global scaling (Ecological Archives E087-085-A2).

APPENDIX C

Behavior of convex hull volume in a null model with respect to species richness and the dimensionality of trait space considered (Ecological Archives E087-085-A3).

APPENDIX D

Two-dimensional convex hulls containing two representative communities shown relative to the Jasper Ridge pool of 54 woody species (Ecological Archives E087-085-A4).