Linking multidimensional functional diversity to quantitative methods: a graphical hypothesis-evaluation framework

Kate S. Boersma[1,6], Laura E. Dee[2], Steve J. Miller[2], Michael T. Bogan[3], David A. Lytle,[4] and
Alix I. Gitelman[5]

[1]*Department of Biology, University of San Diego, San Diego, California 92110, USA*
[2]*Bren School of Environmental Science & Management, University of California at Santa Barbara, Santa Barbara, California 93117, USA*
[3]*Department of Environmental Science, Policy & Management, University of California at Berkeley, Berkeley, California 94720, USA*
[4]*Department of Integrative Biology, Oregon State University, Corvallis, Oregon 97331, USA*
[5]*Department of Statistics, Oregon State University, Corvallis, Oregon 97331, USA*

Abstract. Functional trait analysis is an appealing approach to study differences among biological communities because traits determine species' responses to the environment and their impacts on ecosystem functioning. Despite a rapidly expanding quantitative literature, it remains challenging to conceptualize concurrent changes in multiple trait dimensions ("trait space") and select quantitative functional diversity methods to test hypotheses prior to analysis. To address this need, we present a widely applicable framework for visualizing ecological phenomena in trait space to guide the selection, application, and interpretation of quantitative functional diversity methods. We describe five hypotheses that represent general patterns of responses to disturbance in functional community ecology and then apply a formal decision process to determine appropriate quantitative methods to test ecological hypotheses. As a part of this process, we devise a new statistical approach to test for functional turnover among communities. Our combination of hypotheses and metrics can be applied broadly to address ecological questions across a range of systems and study designs. We illustrate the framework with a case study of disturbance in freshwater communities. This hypothesis-driven approach will increase the rigor and transparency of applied functional trait studies.

Key words: community assembly; disturbance; functional diversity; multidimensional trait space; multivariate analysis; ordination; trait-based ecology.

Introduction

For over a century, ecologists have sought to explain observed differences among biological communities. As a result, many approaches have been developed to simplify species assemblages into quantifiable units using indices such as species richness, diversity, and evenness (e.g., MacArthur 1965, Whittaker 1972, Hill 1973). Recently, trait-based methods have emerged as an appealing way to quantify and explain community differences (Weiher and Keddy 1995, Petchey and Gaston 2002, McGill et al. 2006, Suding et al. 2008). Functional traits are a useful currency with which to understand community dynamics because they can determine how species

respond to the environment ("response traits") and affect ecosystem functioning ("effect traits"; Naeem and Wright 2003). Further, communities that have no species in common will share functional traits (e.g., body size), and trait values can be compared among individuals within and among communities. Thus trait analyses may highlight patterns across ecosystems that are not apparent in taxonomic analyses (e.g., Lavorel et al. 2013, Parravicini et al. 2014). The sheer quantity of recent functional diversity (FD) studies indicates that a functional approach to ecology is an attractive means to detect similarities or differences in community assembly patterns.

Many approaches have been developed to quantify the composition of functional trait states within and among communities ("multivariate trait composition"). Some of these methods, such as multivariate ordination, account for trait non-independence (Petchey and Gaston

2006) and the potential for suites of traits to respond in concert to the environment (Mouchet et al. 2010). Currently, there are over 20 metrics to measure various facets of multivariate trait composition, and a large body of reviews, frameworks, and primers attempting to guide ecologists through this growing analytical complexity (Petchey et al. 2004, Petchey and Gaston 2006, Mouchet et al. 2010, Cadotte et al. 2011, Schleuter et al. 2011, Kleyer et al. 2012, de Bello et al. 2013, Mason et al. 2013, Mouillot et al. 2013b, Spasojevic et al. 2014). This diversity of approaches is inspiring; however, it requires informed metric selection that is appropriate to address specific research questions.

Researchers have begun to apply these quantitative FD approaches to rigorously test ecological hypotheses. For instance, Mason and colleagues (2013) applied a functional approach to examine coexistence mechanisms in woody plant communities through convergence and divergence in resource-use traits relative to soil fertility. Similarly, Laliberté and colleagues (2013) examined convergence in plant functional traits to test the influence of environmental filtering and niche overlap on plant community composition along gradients of primary productivity and grazing disturbance. These two studies are exemplary because they explicitly outlined a hypothesis of functional difference (trait convergence/divergence) and selected a quantitative FD metric from the literature (functional richness; Mason et al. 2005) to test the hypothesis and interpret results. However, convergence and divergence are not the only ways that FD can differ among communities; other patterns may emerge and it is not always clear which aspects of the structure of trait space are most relevant for testing alternate ecological hypotheses of functional difference.

The complexity of combining information on multiple traits for multiple taxa makes it challenging to formulate ecologically informed, multivariate hypotheses and then use these hypotheses to select appropriate quantitative methods in a rigorous and transparent manner. An intuitive framework is needed to guide hypothesis-driven selection of quantitative FD methods. Then, the hypotheses of interest can determine appropriate quantitative analyses to test them, rather than selecting metrics post-hoc or adopting methods designed for other applications without thorough assessment. A hypothesis-driven approach can also determine when metrics do not yet exist for detecting possible FD responses of interest, and therefore motivate development of new quantitative approaches. This way, researchers can ensure that ecological questions drive methodological development instead of vice versa.

We present an intuitive framework that outlines a broad set of ecological hypotheses about functional differences among communities and provides direct ways to test these hypotheses. This framework uses graphical hypotheses of functional responses to disturbance to guide the selection, application, and interpretation of FD methods. We highlight the need for ecologically motivated metric development and selection by describing a novel quantitative metric that quantifies functional turnover between communities. We demonstrate the framework with a case study from an aquatic ecosystem.

A Graphical Hypothesis Framework

It is widely appreciated that multidimensional trait space (Fig. 1) can be a useful foundation on which
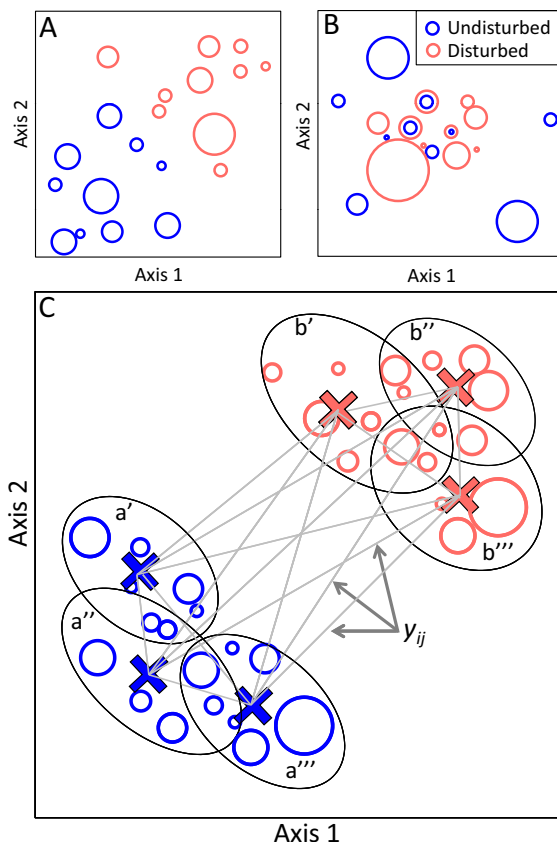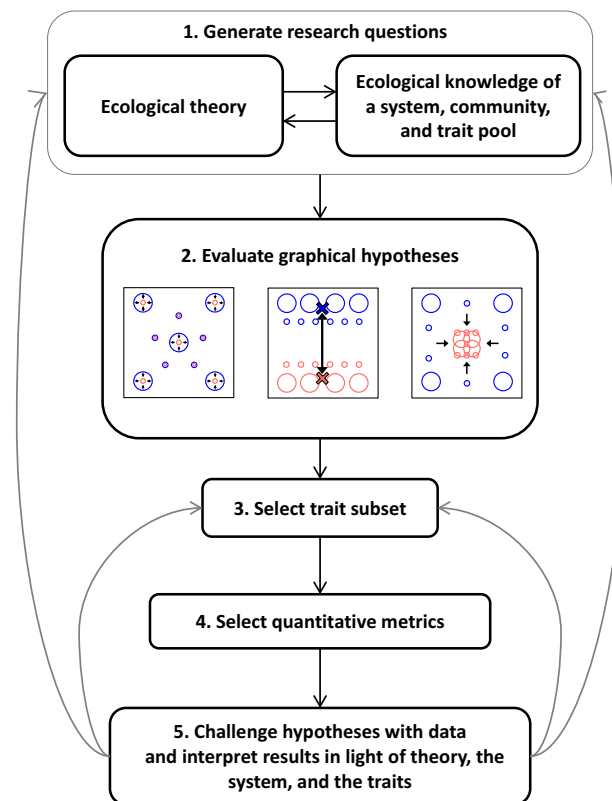

Fig. 1. Multidimensional trait space. Each circle represents a combination of functional trait states ("trait combination"), and circle size reflects the abundance of individuals with that trait combination. Blue circles represent trait combinations in an undisturbed community and red circles in a disturbed community (as per Mouillot et al. 2013b). (A) Two communities with non-overlapping trait composition (e.g., complete functional turnover between communities). Disturbance causes a shift in the location of each community's central tendency or *functional centroid*. (B) Species in the disturbed community occupy a smaller area of trait space than the undisturbed community and thus contain a narrower *range* of trait combinations (e.g., functional convergence). (C) Multiple communities of each disturbance type to demonstrate the nested levels of biological information that can be represented in trait space. Ellipses a′, a″, and a‴ surround the trait combinations of species in three replicate samples of undisturbed communities, and ellipses b′, b″, and b‴ encircle three replicate samples of disturbed communities. The Xs indicate the location of the multivariate functional centroids of each community, and the gray lines represent the distances between functional centroids (y_{ij}). If all species points were removed and only the functional centroids remained, panel C would become an ordination of six communities in trait space. Individual-scale information could also be included in this representation, nested within each point.

to build research questions and generate hypotheses (Rosenfeld 2002). The functional trait composition of a community can be visualized with species' combinations of trait states as points in a Euclidean functional "trait space" (Fig. 1). This trait space can be constructed by two or more traits to generate an *n*-dimensional space. Ordination techniques such as nonmetric multidimensional scaling or principal components analysis can be applied to create a reduced-dimensionality approximation of the distances between objects in the original *n*-dimensional trait space. The position of each point on the ordination is determined by a specific combination of trait values (hereafter "trait combination," as per de Bello et al. 2005), and therefore points that are closer together represent species with similar trait composition. If species abundance information is available, the number of individuals possessing a particular trait combination can be indicated by point size (Rosenfeld 2002, Mouillot et al. 2013*b*). This way, trait information on many species can be combined visually to characterize the multivariate trait composition of entire communities (e.g., Mouchet et al. 2010, de Bello et al. 2013, Mouillot et al. 2013*b*).

While it is relatively easy to create plots of multivariate trait space, we still lack a framework for consistently using these techniques to select quantitative approaches that address ecological questions. We believe that visualizations of ecological phenomena in multivariate trait space can fill this gap. Our approach is straightforward, intuitive, and quantitative: (1) design research questions based on ecological theory and system-specific ecology; (2) evaluate graphical hypotheses of expected differences in trait space; (3) select a subset of traits to address the research questions; (4) select or develop metrics that quantify the characteristics of trait space necessary to test the hypotheses; and (5) challenge the hypotheses with data and interpret differences in trait space in light of the results (Fig. 2). Steps 1, 3, and 5 are reasonably well established, and we illustrate their execution in a case study. Our primary contribution is to facilitate the selection and development of appropriate FD metrics for hypothesis testing (Fig. 2, step 4) by visualizing ecological hypotheses in trait space (Fig. 2, step 2). This process can indicate which combinations of metrics are required to detect relevant changes in trait space. The selection of appropriate metrics for hypothesis testing has been a conceptual hurdle in the current literature; explicitly linking ecological hypotheses to functional metrics will help researchers rigorously address a broad array of ecological questions.

Step 1. Research questions are designed using system-specific ecological knowledge and ecological theory. How will a trait-based approach address outstanding questions regarding ecological processes?

Step 2. The arrangement of trait combinations in multidimensional trait space (Fig. 1) is used to visualize and evaluate hypotheses regarding changes in functional composition of communities. We provide five hypotheses of community assembly in response to disturbance (Fig. 3), although this approach can be applied to other ecological contexts as well. Hypotheses are described in the main text.

Step 3. Outcomes of FD analyses are highly sensitive to trait selection (Petchey and Gaston 2006), and the characteristics of the traits will determine which FD metrics are mathematically feasible and biologically informative (Podani 2009, McCune 2014).

Step 4. An extensive literature guides researchers to select metrics or combinations of metrics to address each hypothesis given the trait subset (Fig. 4). New metrics may be required to capture relevant changes in trait space, as we highlight with the development functional distance to quantify functional turnover.

Step 5. This step follows the scientific method and is outlined with the case study. Theory and system knowledge play a key role in interpreting outcomes of tests using FD metrics.

Fig. 2. A graphical hypothesis-evaluation framework. We use graphical hypotheses representing ecological patterns to inform the selection of quantitative functional diversity metrics and guide the interpretation of results in the context of system-specific ecology.

Graphical hypotheses

We use visualizations of ecological hypotheses in trait space to motivate informed metric selection (or development) and quantitative hypothesis testing (Fig. 3). Our visualization of *hypotheses* complements the visualization of *metrics* in Mouillot et al. (2013*b*); however, in our approach trait space also forms the basis for linking questions to quantitative methods, not just as a way to conceptualize the metrics themselves. These ecological hypotheses are not mutually exclusive, as we illustrate in the context of disturbance. We consider the role of disturbance as a categorical environmental driver (disturbed/undisturbed) to highlight the complementary nature of our work and that of previous authors (Mouillot et al. 2013*b*). However, the approach can be readily extended to general questions of spatial or temporal differences in community structure, beyond the binary treatments described here. Further, it is not limited to tests of disturbance ecology; our intent is to provide a visualization tool to guide hypothesis testing and the selection of analytical methods that may be useful across a variety of systems and ecological questions.

We present five graphical hypotheses of how disturbed and undisturbed communities may differ in multivariate trait composition (Fig. 3). We use the term "graphical hypotheses" to refer to visual representations of ecological phenomena in multivariate trait space, not a probabilistic graphical modeling approach (Koller et al. 2007). The hypotheses represent fundamental ecological phenomena that are commonly observed in published studies. They can be modified and applied broadly to describe patterns of community distinction in many systems. The hypotheses draw from the fields of community assembly theory (Samuels and Drake 1997, Leibold et al. 2004, Leibold and McPeek 2006), restoration ecology (Matthews and Spyreas 2010, Ruhí et al. 2013, Laughlin 2014), disturbance ecology (Houseman et al. 2008, Mouillot et al. 2013*b*), and functional ecology (Mouchet et al. 2010, Mason and de Bello 2013, Mason et al. 2013, Mouillot et al. 2013*b*), among others. Our goal is not to pioneer novel hypotheses of community assembly but instead to demonstrate how to use FD analyses to test this broad set of hypotheses in a rigorous and ecologically informed manner.

We illustrate our graphical hypotheses with two example communities, representing disturbed and undisturbed conditions (Fig. 3). We represent differences in functional trait composition between the communities in a multidimensional trait space as described in Figure 1. In most applications of trait space in the published literature, each unique trait combination represents a species and concentric points represent functionally redundant species that share trait combinations (Rosenfeld 2002). However,

Null Hypothesis: Undisturbed and disturbed communities have identical distributions of trait combinations. Neither abundance nor composition of trait combinations differs.

Equal Impact Hypothesis: Trait composition is identical between undisturbed and disturbed communities. The total abundance of trait combinations differs between communities and these effects are identical across all species.

Trait Abundance Shift Hypothesis: Trait composition is identical in undisturbed and disturbed communities, yet the relative abundance of individuals with each trait combination differs.

Functional Turnover Hypothesis: Undisturbed and disturbed communities occupy different regions of trait space, reflecting different locations of their functional centroids.

Convergence/Divergence Hypothesis: Undisturbed and disturbed communities share no trait combinations in common and occupy distinct quantities of trait space, although the location of the functional centroid does not not differ between communities.
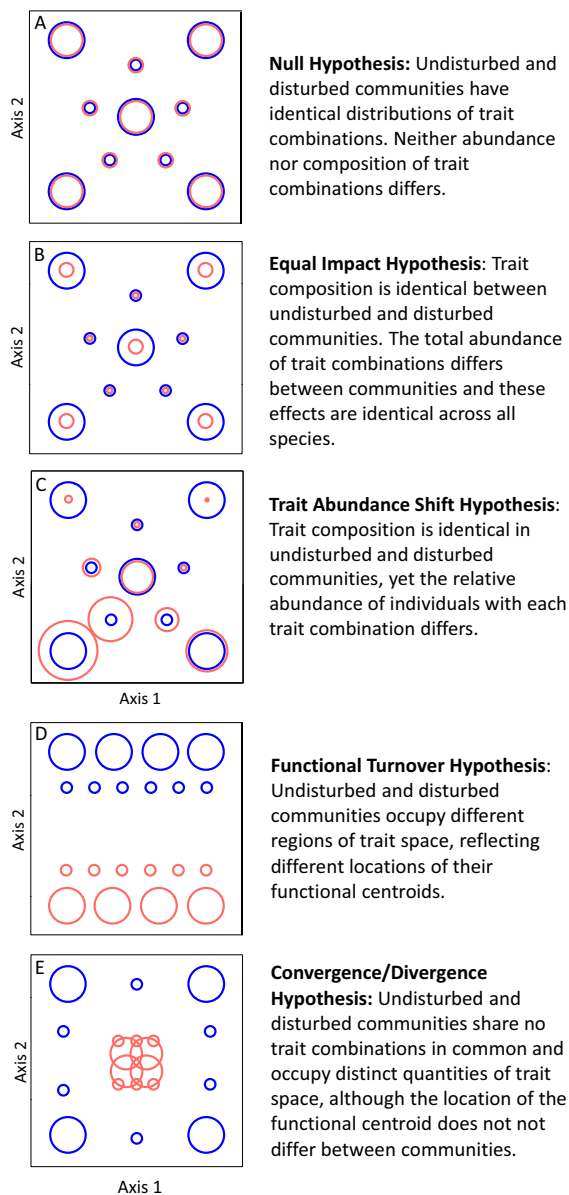
Fig. 3. Five graphical hypotheses of how disturbance may alter the distribution of trait combinations in multidimensional trait space. Each hypothesis is a multivariate ordination, where each point represents a trait combination and circle size is determined by the abundance of individuals with that particular combination. Blue circles represent trait combinations of species in an undisturbed community, and red circles represent those of a disturbed community (as per Mouillot et al. 2013*b*). For illustration, each community contains 10 trait combinations (or species), and we assume that both communities exist within the same regional species and trait pools.

it is also possible that species' trait combinations could vary within species via genetic variation, phenotypic plasticity, or other processes (Miner et al. 2005, Bolnick et al. 2011). Therefore, we refer to points in trait space as "trait combinations." The hypotheses were designed to incorporate species

abundances but may be subset and modified for datasets containing only presence/absence information. While each hypothesis in Fig. 3 represents one disturbed community sample and one undisturbed community sample, the basic framework and quantitative analyses are intended to be performed on replicate communities, such as replicates within treatments, sites, or time points (Fig. 1C), as we demonstrate with the case study.

We present the following hypotheses:

H1: Null Hypothesis (Fig. 3A)

In this scenario, there is no difference in trait composition or the total number of individuals ("total abundance") between disturbed and undisturbed communities. In other words, there are no differences in the abundance, range, or central tendency of trait states. The two communities may be exposed to similar community/trait structuring processes (Weiher and Keddy 1995) or may contain distinct species with the same trait combinations, reflecting complete functional overlap (Mouillot et al. 2013a). Note that insufficient evidence to reject the null hypothesis may occur because the selected traits are inadequate to detect a real difference between communities, emphasizing the importance of careful selection of ecologically relevant traits (Fig. 2, Step 3).

H2: Equal Impact Hypothesis (Fig. 3B)

The range and central tendency of trait combinations do not differ between disturbed and undisturbed communities, yet the total number of individuals differs and these impacts are equal across all trait combinations. In other words, *total abundance* differs but the *relative abundance* of each trait combination in relation to other trait combinations within each community remains unchanged. Equal impacts might arise from largely abiotic processes that have similar per capita effects across species. For example, this could occur if a disturbed community contains a smaller quantity of a shared resource or has poorer overall habitat quality than an undisturbed community, resulting in overall lower species abundances. Alternatively, mild nutrient enrichment may generate the same pattern but opposite in effect, in which all trait combinations increase in abundance by the same amount following disturbance.

H3: Trait Abundance Shift Hypothesis (Fig. 3C)

All trait combinations present in the undisturbed community are also present in the disturbed community and vice versa; however, the number of individuals with each trait combination differs between communities. In other words, traits that are favorable in one community may be unfavorable in the other, producing differences in the distribution of abundances in trait space but no differences in the range or identity of trait combinations. This pattern could occur if

conditions under each disturbance type are not extreme enough to cause the local extirpation of species (taxonomic extinction) or trait combinations (functional extinction) as in the Functional Turnover or Convergence/Divergence Hypotheses. Changes in the abundance of some trait combinations relative to others, without the complete loss or addition of any one combination, may be important warning signs of impending functional extinctions and functional turnover (Säterberg et al. 2013).

H4: Functional Turnover Hypothesis (Fig. 3D)

If disturbance introduces or eliminates some trait combinations altogether, disturbed and undisturbed communities may share few species or trait combinations and occupy distinct regions of trait space. In this case, the location of the central tendency of each community differs, while the range and total abundance of trait combinations may remain unchanged. This outcome could occur if disturbed and undisturbed communities are exposed to different environmental conditions that reduce the persistence of species with particular trait combinations from the shared regional species pool, representing species sorting and environmental turnover of species/traits (Poff 1997, Leibold et al. 2004, Webb et al. 2010). Functional turnover may also be expected in classic Clementsian succession in which species at a single site are replaced through time (Clements 1916), which could result in a change in a community's central tendency in trait space. In *Selecting metrics to test the hypotheses*, we develop and describe a new metric for detecting functional turnover that can be applied to test this hypothesis.

H5: Convergence/Divergence Hypothesis (Fig. 3E)

Of all of the hypotheses, convergence/divergence is the most prominent in the current FD literature, likely because of its intuitive interpretation in the context of ecological theory. Graphically, convergence occurs when species in the disturbed community occupy a smaller area of trait space than those in the undisturbed community (Mason et al. 2005), reflecting a smaller range of trait values. This could occur if disturbance acts as an environmental filter (Poff 1997, Leibold et al. 2004, Grime 2006) that reduces the trait combinations in the disturbed community to a subset of those existing in the undisturbed community (Leibold et al. 2004, Webb et al. 2010) or if functionally homogenous species replace a more functionally diverse community. Alternately, disturbance may create functional niches for colonization of the disturbed community by species that were excluded from the undisturbed community, causing functional divergence (e.g., Cadotte 2007). Divergence can be represented as an expansion of the area species occupy in trait space. Expansion may occur if disturbance causes the local extinction of a competitively dominant species or keystone predator (Paine 1966), or facilitates the
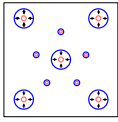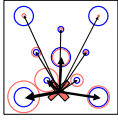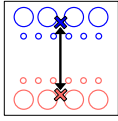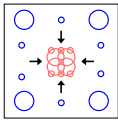
Hypothesis	Metric requirements		Metric	Hypothesis tests
Null	-		TAb, FDisp, FDist, FRic	None of the metrics differ between treatments.
Equal impact	Detect differences in abundance between communities that are evenly distributed across trait combinations.		Total abundance (TAb)	Total abundance of trait combinations differs between A and B (Welch's *t* test). FDisp, FDist, and FRic do not differ.
Trait abundance shift	Detect differences in the relative abundance of traits in each community.		Functional dispersion (FDisp; Laliberté and Legendre 2010)	Functional dispersion differs between A and B (Welch's *t* test). FDist and FRic do not differ. TAb may or may not differ.
Functional turnover	Detect differences in the location of communities' functional centroids in trait space.		Functional distance (FDist; this study)	Functional distance between A and B is greater than null expectation (Mixed effects modeling). FRic, FDisp, and TAb may or may not differ.
Convergence/Divergence	Detect differences in the range of trait space occupied by each community.		Functional richness (FRic; Villegér et al. 2008)	Functional richness differs between A and B (Welch's *t* test). FDisp will also differ. FDist and TAb may or may not differ.

Fig. 4. Using graphical hypotheses to select quantitative metrics. This figure represents our decision process to select metrics for testing hypotheses about responses to disturbance. "Metric requirements" are factors considered when selecting each metric, as described in "*Selecting metrics to test the hypotheses*.". "Hypothesis test" indicates which tests are needed to support each hypothesis. Hypotheses are not mutually exclusive and support may be found for multiple hypotheses at once. Note: the graphical hypotheses depicted in Fig. 3 are simplified diagrams that compare a single disturbed community replicate with a single undisturbed community replicate, but hypothesis-testing requires that there are multiple replicates of communities in Treatments A and B (undisturbed and disturbed). TAb, total abundance; FDisp, functional dispersion; FDist, functional distance; FRic, functional richness.

establishment of invasive species without supplanting native taxa (e.g., Hejda and de Bello 2013).

Selecting metrics to test the hypotheses

A key challenge we seek to address is the selection of quantitative methods that capture the changes in trait space necessary to support or reject the hypotheses. Fig. 4 illustrates our decision process to select quantitative metrics prior to analyses that test the five graphical hypotheses (Fig. 2, Step 4). First, we used the graphical representations to determine which aspects of the functional space must be measured to test the hypotheses of ecological change (Fig. 4, Metric requirements). Then we identified metrics or combinations of metrics with mathematical properties that quantify these aspects and can be used to compare them between disturbed and undisturbed communities (Fig. 4, Metric). Our recommendations in Fig. 4 come from consulting the extensive quantitative FD literature (Petchey et al. 2004, Mason et al. 2005, 2013, Petchey and Gaston 2006, Mouchet et al. 2010, Cadotte et al. 2011, Schleuter et al. 2011, Kleyer et al. 2012, de Bello et al. 2013, Mouillot et al. 2013*b*, Spasojevic et al. 2014). We did not identify an intuitive existing

approach we believed was suitable to test differences in the location of the central tendency of trait combinations as in the Functional Turnover Hypothesis, so we designed a new metric, *functional distance* (described *described in Functional distance: a novel metric*). The final step of metric and method selection is to determine appropriate statistical hypothesis tests (Fig. 4, Hypothesis tests).

We identified two published FD metrics that are appropriate to test the graphical hypotheses: functional richness (Villéger et al. 2008) and functional dispersion (Laliberté and Legendre 2010), and created a third, called functional distance. When used in combination with the abundance of individuals with each trait combination, these metrics enable quantitative tests of how functional trait differences between disturbed and undisturbed communities support or do not support the five hypotheses (Fig. 4). The three multivariate metrics are derived from pairwise Gower dissimilarity (Gower 1971) in trait composition between replicate communities. Next we describe how and why our evaluation of the graphical hypotheses (Fig. 3) led us to select each of the FD metrics (Fig. 4). We also discuss metric limitations and cases where combinations of metrics may be required.

Functional dispersion.—The Trait Abundance Shift Hypothesis (Fig. 3C) involves differences in the proportion of individuals with particular trait combinations that do not result in the addition or loss of any combinations entirely, as occurs with the Functional Turnover and Convergence/Divergence Hypotheses. To demonstrate differences in the relative abundance of trait combinations that would indicate a trait abundance shift, one must quantify the distribution of abundances in trait space. Functional dispersion measures the mean distance of all points to the abundance-weighted functional centroid of a community (Laliberté and Legendre 2010), where the functional centroid is a composite mean trait value of all trait combinations in a community that is weighted by the abundance of individuals possessing each trait combination (the Xs in Figs. 1C and 4). Conceptually, functional dispersion quantifies the mean difference between points by comparing each trait combination to a hypothetical representative trait combination (the functional centroid). Differences in dispersion between disturbed and undisturbed communities suggest that disturbance favors some trait combinations over others, and this can be reflected in differences in the relative abundance of individuals with different trait combinations.

Functional richness.—Determining support for the Convergence/Divergence Hypothesis (Fig. 3E) requires a metric that is not dependent upon the relative abundance of trait combinations. Functional richness measures the overall area of trait space (convex hull) occupied by species in a community (Mason et al. 2005, Cornwell et al. 2006, Villéger et al. 2008, Podani 2009). Significant differences in this metric between disturbed and undisturbed samples indicate disturbance-induced convergence or divergence in functional composition that may not be associated with a shift in the location of the functional centroid (Villéger et al. 2008). This metric is the one most commonly used in existing FD studies that test hypotheses about responses to disturbance (e.g., Laliberté et al. 2013, Mason et al. 2013), likely because differences in this metric are the most straightforward to interpret in the context of a system's ecology.

Functional distance: a novel metric.—As described in "*Functional dispersion*," above, functional dispersion detects changes in the *relative abundance* of trait combinations and functional richness detects changes in the *range* of trait combinations (Fig. 4). However, to test whether disturbance causes the addition or loss of trait combinations in an entire region of trait space as in the Functional Turnover Hypothesis (Fig. 3D), it is also necessary to determine whether disturbance affects the location of a community's functional centroid, irrespective of species abundances (Fig. 1C). The distance between the functional centroids of two communities (y_{ij}) measures differences in trait space that may result from low functional overlap between

the communities. While many metrics exist, we did not find an intuitive, published metric that could account for the non-independence of pairwise comparisons between replicate communities, so we created a novel metric that we refer to as *functional distance*. We define functional distance as the distance between the non abundance-weighted functional centroids of two communities in trait space (gray lines in Fig. 1C). While it is straightforward analytically to calculate the functional distance between two communities in multidimensional space, calculating pairwise distances between multiple communities that have been grouped (by sampling occasion, site, or treatment), as in the case of replicate disturbed and undisturbed communities, is more complex. Multiple pairwise distances associated with a given site or community (values within a row or column in a community matrix) are non-independent, and this non-independence creates a correlated error structure that biases calculation of meaningful confidence intervals or standard errors (Clarke et al. 2002). We address this problem using mixed effects models with disturbed/undisturbed as a fixed effect and community as a random effect. A similar method has been applied with genetic distances (Van Strien et al. 2012) and landscape distances (Bellamy et al. 2003), but to our knowledge this is its first application to measure functional distance.

To quantify functional distance, we converted the community matrix to a presence/absence matrix of 1s and 0s and then calculated the functional centroid for each community from the Gower distance trait matrix. This is easily obtained using function fdisp() in the R package FD (Laliberté and Legendre 2010, Laliberté and Shipley 2011). This function applies principal coordinates analysis to the distance matrix to generate centroid coordinates for each community and corrects for negative eigenvalues as described by Anderson (2006). Then, we calculated pairwise Euclidean distances between the centroids of all disturbed and undisturbed communities (functional distances) and examined the effect of disturbance on these distances using a mixed-effects model of the following form

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where Y_{ij} is the functional distance between the centroids of communities i and j, μ is the fixed effect of disturbed/undisturbed (0/1), α_i and β_j are random effects to account for correlations between pairwise distances that have a community in common, and ε_{ij} is an independent error term. Following the suggestion of Van Strien et al. (2012), who addressed this issue with genetic distances, we used Markov chain Monte Carlo simulation to repeatedly resample distances from the distance matrix and compare models with and without the disturbance term. Finally, we determined model performance by examining posterior means and the 95% credible interval. When the superior model included the disturbance term and the credible interval for the disturbance term did not overlap zero, we considered

communities i and j to have demonstrated a disturbance-induced shift in multivariate trait space.

Combinations of Metrics

As illustrated in Fig. 4, a combination of metrics is necessary to determine support for the hypotheses. For example, support for the Equal Impact Hypothesis (Fig. 3B) requires evidence of a difference in total abundance, the *lack* of a difference in functional dispersion and functional richness, and functional distance that is not significantly different from zero. The Trait Abundance Shift Hypothesis (Fig. 3C), Functional Turnover Hypothesis (Fig. 3D), and Convergence/Divergence Hypothesis (Fig. 3E) can all generate differences in functional dispersion. Therefore, determining which of these hypotheses may be supported in a particular case study also requires the inclusion of functional distance and functional richness. For example, disturbance may cause the relative abundance of trait combinations to shift without functional turnover or convergence/divergence, but convergence/divergence and functional turnover are always accompanied by shifts in trait abundances. These nuances are ecologically informative and understanding them is essential to making appropriate inferences from FD studies.

Under some circumstances further scrutiny may be necessary to differentiate among hypotheses. For example, if disturbance causes species to be replaced by other taxa with trait combinations that are the same distance from the functional centroid, this functional turnover would not be detected by any of metrics combinations we propose. Situations like this one highlight the importance of visualizing and interpreting differences in trait space in light of system-specific ecological knowledge. Such cases can usually be avoided by incorporating ecological observations or by combining taxonomic and functional analyses as we demonstrate with our case study, below. Integrating species and trait information will produce a fuller understanding of how communities function and allow researchers to isolate and interpret any irregularities.

A Case Study

We demonstrate the framework with a published dataset on stream invertebrate communities before and after catastrophic stream drying (Bogan and Lytle 2011). In this study, researchers sampled the aquatic invertebrate community in a small, isolated stream (French Joe Canyon) in southeast Arizona for eight years (2002–2009), before and after a severe drought and resultant stream drying event. Samples taken before drought were classified as "undisturbed" and those taken after as "disturbed."

Step 1: Generate research question.—We asked: how does stream drying affect the functional composition of aquatic invertebrate communities?

Step 2: Evaluate graphical hypotheses.—The five hypotheses are all ecologically feasible in this case study, and we imagine this will be true in many other applications as well. Taxonomic analyses documented species turnover when drought-sensitive taxa were eliminated from aquatic communities and replaced by taxa that were resilient to drought (Bogan and Lytle 2011). Consequently we expected this case study to support the Functional Turnover Hypothesis. However, given the limited information on how aquatic communities respond to unprecedented drought and the potential for multiple hypotheses to be at play concurrently, we set out to test all five hypotheses.

Step 3: Select the trait subset.—Outcomes of FD analyses are highly sensitive to trait identity, and the inclusion of more traits does not necessarily result in more accurate or informative outcomes (Petchey and Gaston 2006). We used background knowledge of aquatic invertebrate communities to choose seven categorical traits for this analysis that are associated with biological responses to drought in arid-land streams: body size, functional feeding group, dispersal ability, locomotion, voltinism, respiration, and diapause (Boersma et al. 2014, Schriever et al. 2015).

Step 4: Select the quantitative metrics.—We used the metric combinations and statistical tests described in Fig. 4 to evaluate the five hypotheses. Because of our expectation of functional turnover, it was important to include the novel metric, *functional distance*, since neither functional richness nor functional dispersion alone could test the Functional Turnover Hypothesis (Fig. 3D). After calculating the metrics and examining their empirical distributions, we compared total abundance, richness and dispersion between disturbed and undisturbed groups using Welch's t tests. We calculated functional distance using the mixed effects modeling approach described earlier. All analyses were conducted using R version 2.14.1 (R Development Core Team 2011) and packages MCMCglmm, lme4, FD, and vegan (Hadfield 2010, Bates et al. 2011, Laliberté and Shipley 2011, Oksanen et al. 2012).

Step 5: Challenge hypotheses with data and interpret results in light of system-specific ecology.—Counter to our expectations, we did not detect a difference in the location of the functional centroid, i.e., in the *functional distance* between disturbed and undisturbed communities (MCMC, lower credible interval = −0.002, upper credible interval = 0.020). Despite a lack of support for the Functional Turnover Hypothesis, there was evidence that functional richness was lower in the disturbed post-drying communities than in the undisturbed pre-drying communities (Welch's t test, $t = 2.541$, df = 14.447, $P = 0.023$) and moderate evidence of the same pattern in functional dispersion (Welch's t test, $t = 1.920$, df = 17.959, $P = 0.071$). Finally, total abundance was higher in the disturbed community than in the undisturbed community (Welch's t test, $t = -3.311$, df = 16.133, $P = 0.004$). In sum, the combined evidence

of no functional turnover with differences in functional richness, functional dispersion, and total abundance indicates support for the Convergence/Divergence Hypothesis (Fig. 3E) and a lack of support for the other four hypotheses (Fig. 4).

System-specific ecology is as important when interpreting the results of FD analyses as it is when generating research questions. Disturbed post-drought communities experienced an influx of colonists when flow returned after the stream drying event (Bogan and Lytle 2011), and these new colonists were functionally similar to some (but not all) of the species present in the community before drying. The functionally homogeneous colonists replaced taxa that had diverse trait combinations at the edges of multivariate trait space, which produced the overall reduction in functional richness we observed, without functional turnover.

CONCLUSIONS

Our hypothesis-evaluation framework was motivated by the need for an intuitive and ecologically informed way to select quantitative FD methods prior to data analysis. The growing number of quantitative FD approaches makes it essential that researchers justify metric selection in light of trait-based hypotheses, rather than simply adopting published techniques without assessment. Our framework facilitates this process by using graphical hypotheses in multivariate trait space to guide the selection, application, and interpretation of quantitative methods. We propose a set of metrics to test common ecological hypotheses of functional community assembly, including a new metric to measure functional distance among communities. This approach will encourage more consistent application of multivariate FD methods in a rigorous manner, while honoring the importance of system-specific ecological knowledge and observed community assembly patterns.

Our five graphical hypotheses represent functional trait responses to environmental drivers in two categories (in our example: disturbed and undisturbed), and we envision many applications for which two categories are sufficient (e.g., treatment/control, before/after). However, our flexible approach could easily be adapted to accommodate more than two groups, such as multiple time points or samples along a gradient or to consider effect traits and test hypotheses concerning ecosystem functioning (Naeem and Wright 2003, Suding et al. 2008). The a priori selection of quantitative tools to address hypotheses also helps researchers identify potential needs to design novel metrics if appropriate ones do not exist, as we have done here with our functional distance metric to measure functional turnover among replicate communities.

The recent surge in FD studies suggests that a functional approach to ecology is an appealing means to detect community assembly patterns that transcend specific ecosystems. However, it is challenging to conceptualize ecological patterns in multivariate trait space and select quantitative methods to test them. Our framework will address this challenge and thereby increase the applicability of functional approaches and enable functional ecology to continue to grow into a practical and rigorous discipline.

LITERATURE CITED

Anderson, M. J. 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62:245–253.

Bates, D., M. Maechler and B. M. Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42. https://cran.r-project.org/web/packages/lme4/

Bellamy, P. E., P. Rothery, and S. A. Hinsley. 2003. Synchrony of woodland bird populations: the effect of landscape structure. Ecography 26:338–348.

Bogan, M. T., and D. A. Lytle. 2011. Severe drought drives novel community trajectories in desert stream pools. Freshwater Biology 56:2070–2081.

Bolnick, D. I., P. Amarasekare, M. S. Araujo, R. Burger, J. M. Levine, M. Novak, V. H. Rudolf, S. J. Schreiber, M. C. Urban, and D. A. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. Trends in Ecology & Evolution 26:183–192.

Boersma, K. S., M. T. Bogan, B. A. Henrichs and D. A. Lytle. 2014. Invertebrate assemblages of pools in arid-land streams have high functional redundancy and are resistant to severe drying. Freshwater Biology 59:491–501.

Cadotte, M. W. 2007. Concurrent niche and neutral processes in the competition-colonization model of species coexistence. Proceedings of the Royal Society B 274:2739–2744.

Cadotte, M. W., K. Carscadden, and N. Mirotchnick. 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecology 48:1079–1087.

Clarke, R., P. Rothery, and A. Raybould. 2002. Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. Journal of Agricultural, Biological, and Environmental Statistics 7:361–372.

Clements, F. E. 1916. Plant succession: an analysis of the development of vegetation. Carnegie Institute of Washington.

Cornwell, W. K., D. W. Schwilk, and D. D. Ackerly. 2006. A trait-based test for habitat filtering: convex hull volume. Ecology 87:1465–1471.

de Bello, F., J. Lepš, and M.-T. Sebastià. 2005. Predictive value of plant traits to grazing along a climatic gradient in

the Mediterranean. Journal of Applied Ecology 42:824–833.

de Bello, F., C. P. Carmona, N. W. H. Mason, M.-T. Sebastià, and J. Lepš. 2013. Which trait dissimilarity for functional diversity: trait means or trait overlap? Journal of Vegetation Science 24:807–819.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:857–874.

Grime, J. P. 2006. Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. Journal of Vegetation Science 17:255–260.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software 33:1–22.

Hejda, M., and F. de Bello. 2013. Impact of plant invasions on functional diversity in the vegetation of Central Europe. Journal of Vegetation Science 24:890–897.

Hill, M. 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54:427–432.

Houseman, G. R., G. G. Mittelbach, H. L. Reynolds, and K. L. Gross. 2008. Perturbations alter community convergence, divergence, and formation of multiple community states. Ecology 89:2172–2180.

Kleyer, M., S. Dray, F. Bello, J. Lepš, R. J. Pakeman, B. Strauss, W. Thuiller, and S. Lavorel. 2012. Assessing species and community functional responses to environmental gradients: which multivariate methods? Journal of Vegetation Science 23:805–821.

Koller, D., N. Friedman, L. Getoor, and B. Taskar. 2007. Graphical models in a nutshell in L. Geetor and B. Taskar, editors. Introduction to Statistical Relational Learning. MIT Press.

Laliberté, E., and P. Legendre. 2010. A distance-based framework for measuring functional diversity from multiple traits. Ecology 91:299–305.

Laliberté, E. and B. Shipley. 2011. FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-11.https://cran.r-project.org/web/packages/FD/

Laliberté, A. S., D. A. Norton, and D. Scott. 2013. Contrasting effects of productivity and disturbance on plant functional diversity at local and metacommunity scales. Journal of Vegetation Science 24:834–842.

Laughlin, D. C. 2014. Applying trait-based models to achieve functional targets for theory-driven ecological restoration. Ecology Letters 17:771–784.

Lavorel, S., et al. 2013. A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. Journal of Vegetation Science 24:942–948.

Leibold, M. A., and M. A. McPeek. 2006. Coexistence of the niche and neutral perspectives in community ecology. Ecology 87:1399–1410.

Leibold, M. A., et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

MacArthur, R. 1965. Patterns of species diversity. Biological Review 40:510–533.

Mason, N. W. H., and F. de Bello. 2013. Functional diversity: a tool for answering challenging ecological questions. Journal of Vegetation Science 24:777–780.

Mason, N. W. H., D. Mouillot, W. G. Lee, and J. B. Wilson. 2005. Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos 111:112–118.

Mason, N. W. H., F. de Bello, D. Mouillot, S. Pavoine, and S. Dray. 2013. A guide for using functional diversity indices

to reveal changes in assembly processes along ecological gradients. Journal of Vegetation Science 24:794–806.

Matthews, J. W., and G. Spyreas. 2010. Convergence and divergence in plant community trajectories as a framework for monitoring wetland restoration progress. Journal of Applied Ecology 47:1128–1136.

McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends in Ecology and Evolution 21:178–185.

Miner, B. G., S. E. Sultan, S. G. Morgan, D. K. Padilla, and R. A. Relyea. 2005. Ecological consequences of phenotypic plasticity. Trends in Ecology & Evolution 20:685–692.

Mouchet, M. A., S. Villéger, N. W. H. Mason, and D. Mouillot. 2010. Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Functional Ecology 24:867–876.

Mouillot, D., et al 2013a. Rare species support vulnerable functions in high-diversity ecosystems. PLoS Biology 11:e1001569.

Mouillot, D., N. A. J. Graham, S. Villéger, N. W. H. Mason, and D. R. Bellwood. 2013b. A functional approach reveals community responses to disturbances. Trends in Ecology & Evolution 28:167–177.

Naeem, S., and J. P. Wright. 2003. Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem. Ecology Letters 6:567–579.

Oksanen, J., G. F. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens and H. Wanger. 2012. vegan: Community Ecology Package. R package version 2.0-3.https://cran.r-project.org/web/packages/vegan/

Paine, R. T. 1966. Food web complexity and species diversity. American Naturalist 100:65–75.

Parravicini, V., et al. 2014. Global mismatch between species richness and vulnerability of reef fish assemblages. Ecology Letters 17:1101–1110.

Petchey, O. L., and K. J. Gaston. 2002. Functional diversity (FD), species richness and community composition. Ecology Letters 5:402–411.

Petchey, O. L., and K. J. Gaston. 2006. Functional diversity: back to basics and looking forward. Ecology Letters 9:741–758.

Petchey, O. L., A. Hector, and K. J. Gaston. 2004. How do different measures of functional diversity perform? Ecology 85:847–857.

Podani, J. 2009. Convex hulls, habitat filtering, and functional diversity: mathematical elegance versus ecological interpretability. Community Ecology 10:244–250.

Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391–409.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: www.r-project.org

Rosenfeld, J. S. 2002. Functional redundancy in ecology and conservation. Oikos 98:156–162.

Ruhí, A., D. Boix, S. Gascón, J. Sala, and X. Quintana. 2013. Nestedness and successional trajectories of macroinvertebrate assemblages in man-made wetlands. Oecologia 171:545–556.

Samuels, C. L., and J. A. Drake. 1997. Divergent perspectives on community convergence. Trends in Ecology & Evolution 12:427–432.

Säterberg, T., S. Sellman, and B. Ebenman. 2013. High frequency of functional extinctions in ecological networks. Nature 499:468–470.

Schleuter, D., M. Daufresne, F. Massol, and C. Argillier. 2011. A user's guide to functional diversity indices. Ecological Monographs 80:469–484.

Schriever, T. A., M. T. Bogan, K. S. Boersma, M. Cañedo-Argüelles. K. L.Jaeger J. D.Olden, and D. A. Lytle. 2015.

Hydrology shapes taxonomic and functional structure of desert stream invertebrate communities. Freshwater Science 34:399–409.

Spasojevic, M. J., S. Copeland, and K. N. Suding. 2014. Using functional diversity patterns to explore metacommunity dynamics: a framework for understanding local and regional influences on community structure. Ecography 37:239–249.

Suding, K. N., S. Lavorel, F. S. Chapin, J. H. C. Cornelissen, S. Díaz, E. Garnier, D. Goldberg, D. U. Hooper, S. T. Jackson, and M.-L. Navas. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. Global Change Biology 14:1125–1140.

Van Strien, M. J., D. Keller, and R. Holderegger. 2012. A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. Molecular Ecology 21:4010–4023.

Villéger, S., N. W. H. Mason, and D. Mouillot. 2008. New multidimensional functional diversity indices for a multi-faceted framework in functional ecology. Ecology 89:2290–2301.

Webb, C. T., J. A. Hoeting, G. M. Ames, M. I. Pyne, and N. LeRoy Poff. 2010. A structured and dynamic framework to advance traits-based theory and prediction in ecology. Ecology Letters 13:267–283.

Weiher, E., and P. A. Keddy. 1995. Assembly rules, null models, and trait dispersion: new questions from old patterns. Oikos 74:159–164.

Whittaker, R. J. 1972. Evolution and measurement of species diversity. Taxon 21:213–251.

CONCEPTS & SYNTHESIS