Alpha Diversity – Tree-Based

SESYNC

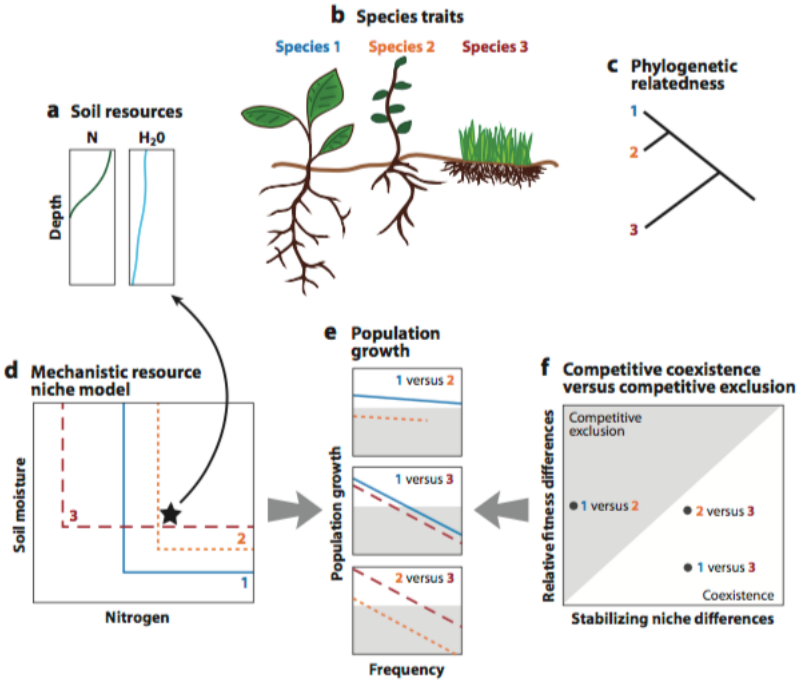
1 February 2017

Community Assembly and Species Co-Existence

- Central Questions:
  - Why do species occur in the community you observe?
  - Why do some pairs of species co-exist while others do not?
- Predictions:
  - Ecologically similar species should co-exist if the abiotic environment is most important
  - Ecologically dissimilar species should co-exist if biotic interactions (e.g. competition) are most important

Community Assembly and Species Co-Existence

- Predictions
  - Ecologically similar species should co-exist if the abiotic environment is most important
  - Ecologically dissimilar species should co-exist if biotic interactions (e.g. competition) are most important
- What do we need to test predictions?
  - Information regarding the ecological similarity of co-existing species.

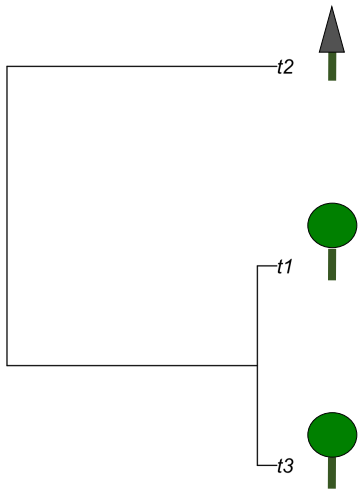


Community Assembly and Species Co-Existence

- How can we measure ecological similarity of co-existing species?
  - Measure species niches or traits
    - Difficult
  - Estimate ecological similarity by using relatedness as a substitute.

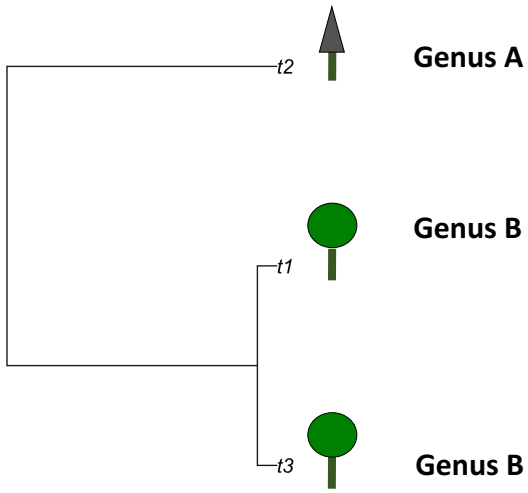
Community Assembly and Species Co-Existence

- Relatedness as a substitute for ecological similarity
  - Charles Darwin was the first to recognize that closely related species should be more ecologically similar due to *common descent*
  - For example two species from one genus should be more ecologically similar to one another than they are to another species from a different genus
    - I.E. 2 *Acer* species are likely more ecologically similar to one another when compared to 1 *Pinus* species



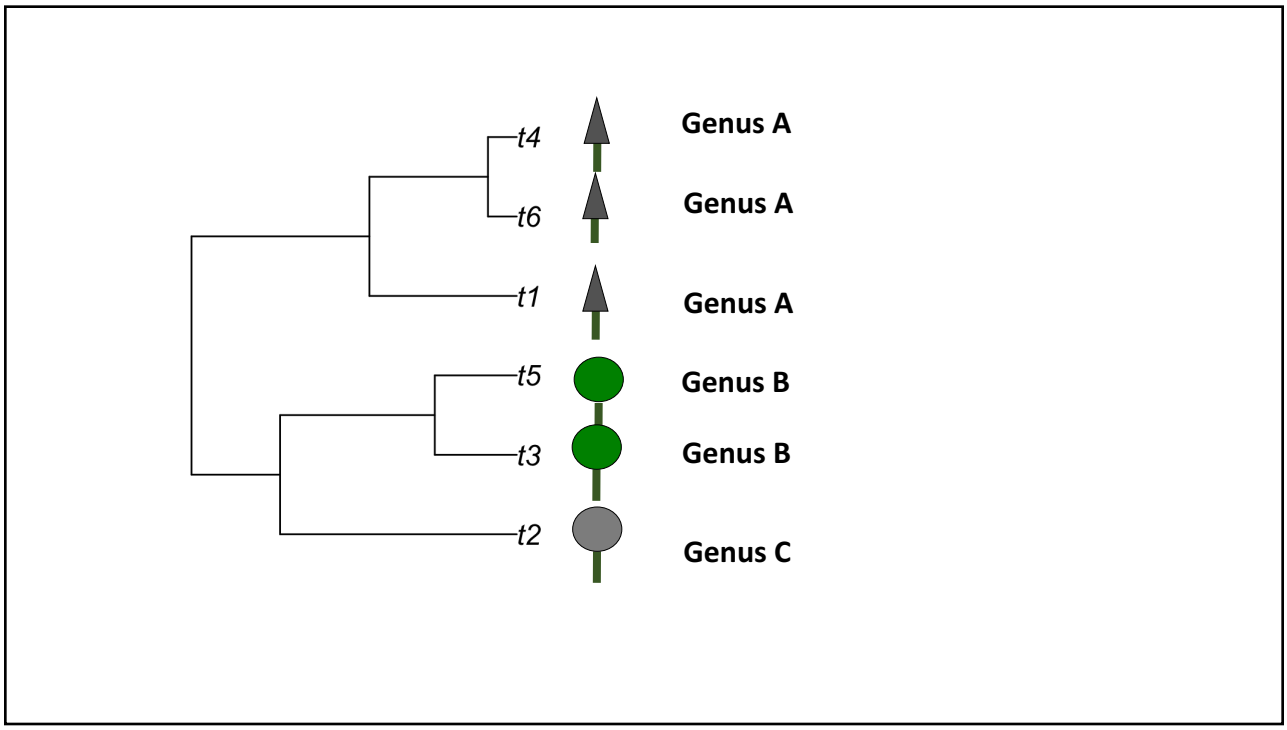
Community Assembly and Species Co-Existence

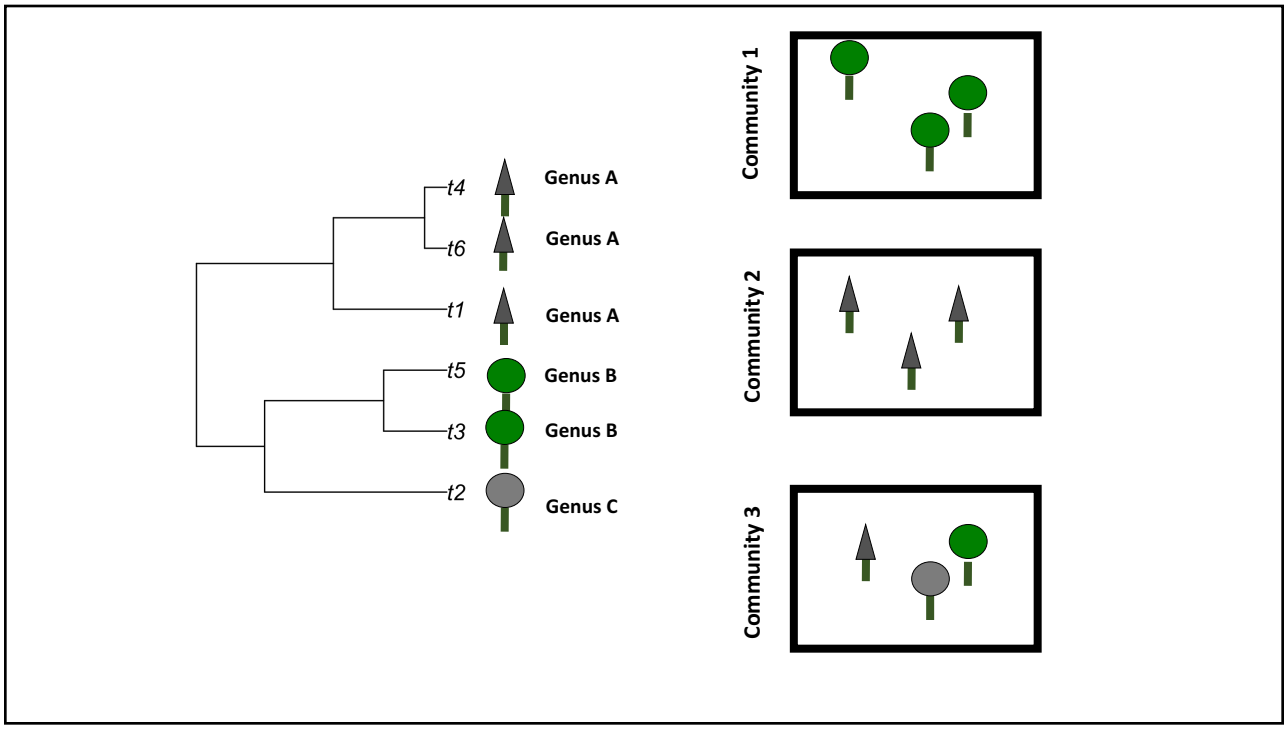
- Relatedness as a substitute for ecological similarity
  - Early plant community ecologists around 1910 (e.g. Jaccard) used Darwin's logic to help test predictions regarding whether the abiotic environment or competition are the major factors controlling species co-existence.
  - Jaccard and others design analyses that calculated the ratio of the *# of genera : # of species* found in a community

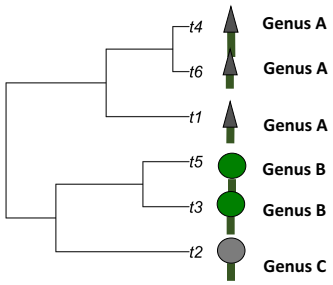


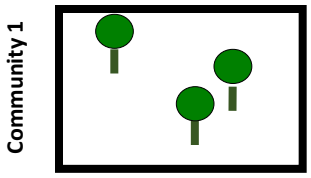
Community Assembly and Species Co-Existence

- Relatedness as a substitute for ecological similarity
  - Jaccard and others design analyses that calculated the ratio of the *# of genera : # of species* found in a community
  - If there is a high number of genera and a low number of species in a community, then species are distantly related in the community.
  - If there is a low number of genera and a high number of species in a community, then species are closely related in the community.


t4
t6
t1
t5
t3
t2
Genus A
Genus A
Genus A
Genus B
Genus B
Genus C

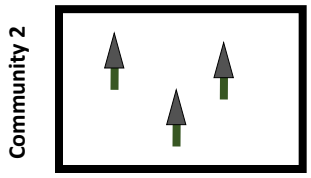

t4
t6
t1
t5
t3
t2
Genus A
Genus A
Genus A
Genus B
Genus B
Genus C
Community 1
Community 2
Community 3





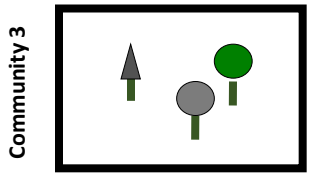
1 genus
3 species

1:3 ratio



1 genus
3 species

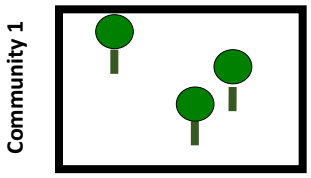
1:3 ratio



3 genus
3 species

3:3 ratio

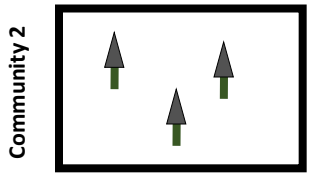
Abiotic filtering
Importance of Environment



1 genus
3 species

1:3 ratio

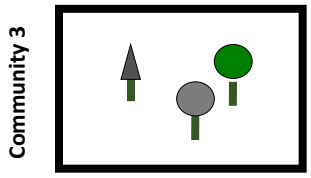
Abiotic filtering
Importance of Environment



1 genus
3 species

1:3 ratio

Importance of Biotic Interaction



3 genus
3 species

3:3 ratio

Community Assembly and Species Co-Existence

- Relatedness as a substitute for ecological similarity
  - Because closely related species are more ecologically similar
    - A high genus : species ratio indicates distantly related and ecologically dissimilar species coexist. This may indicate the importance of species competition
    - A low genus : species ratio indicates closely related and ecologically similar species coexist. This may indicate the importance of abiotic conditions determining species coexistence in a plant community.

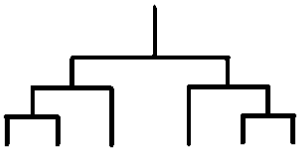
Community Assembly and Species Co-Existence

- The genus : species ratio type of study in plant community ecology started ~1910 and was popular until 1990's.
- A large criticism of genus : species ratio analyses is that it does not account for the different ages of genera and species.

Community Assembly and Species Co-Existence

- A large criticism of genus : species ratio analyses and species diversity analyses is that they do not account for the different ages of genera and species or relatedness in general.
  - Solution = Use phylogenetic trees to estimate the relatedness of coexisting species

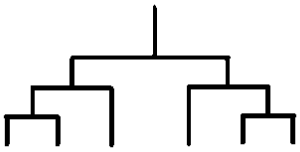
Phylogeny



Phylodiversity

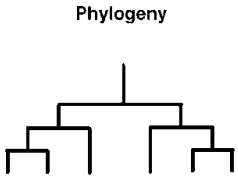
- In the 1990's conservation biologists recognized the biodiversity is not only species diversity.
  - Biodiversity has several axes or dimensions including genetic, functional and phylogenetic diversity

Phylogeny



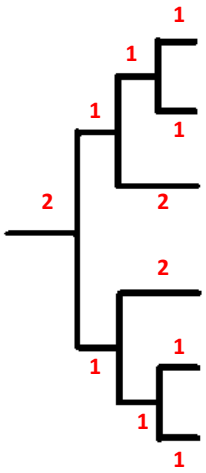
Phylodiversity

- Phylogenetic diversity was first truly formalized by Dan Faith in 1992
  - He proposed a metric called PD that is also commonly referred to as Faith's Index
  - Many additional metrics have now been generated but this metric is still widely used



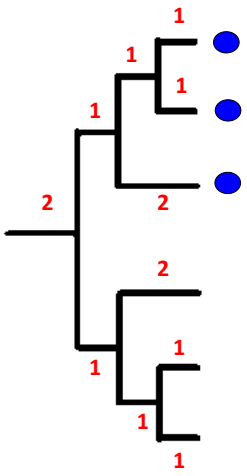
Calculating Faith's Index (PD)

- Begin with a large phylogeny and measure its total branch length (ie the 'tree length') **14**



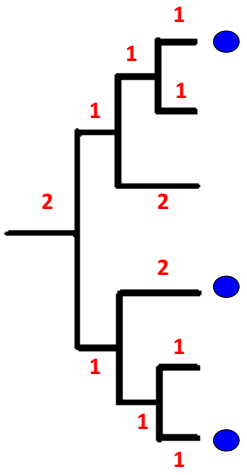
Calculating Faith's Index (PD)

- Begin with a large phylogeny and measure its total branch length (ie the 'tree length') **14**
- Next prune the large phylogeny so that it only contains the species in your community.
- Measure the tree length of the pruned phylogeny **8**
- PD = community tree length / overall tree length = **0.57**



Calculating Faith's Index (PD)

- Begin with a large phylogeny and measure its total branch length (ie the 'tree length') **14**
- Next prune the large phylogeny so that it only contains the species in your community.
- Measure the tree length of the pruned phylogeny **10**
- PD = community tree length / overall tree length = **0.71**



Main Phylogenetic Diversity Indices

$$Faith = \sum_{i}^{n} l_i$$

Main Phylogenetic Diversity Indices

terribly important for analyses of community structure and diversity. This has lead to the development of a version of Faith's index that is weighted by abundance [62] that I will call the Weighted Faith's Index.

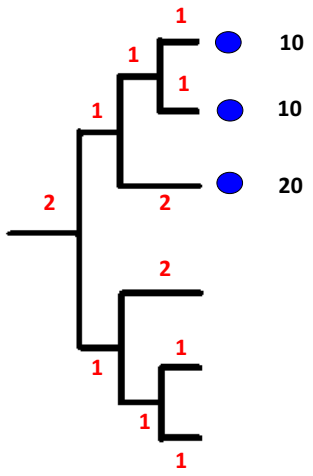
$$Weighted.Faith = n \times \frac{\sum_{i}^{n} l_i \bar{A}_i}{\sum_{i}^{n} \bar{A}_i}$$

Where n is the number of branches in the phylogenetic tree, l_i is the length of the i^{th} branch, and $\bar{A}_i$ is the average abundance of all species subtended by that branch. As you can see, calculating this weighted metric is more complex than simply summing the branch lengths in a phylogeny containing community members. It requires calculating a

Calculating Abundance Weighted Faith's Index (PD)

Branch Length	Average Abundance	Length x abundance
2	40/6 = 6.66	13.33
1	40/3 = 13.33	13.33
1	20/2 = 10	10
1	10/1 = 10	10
1	10/1 = 10	10
2	20/1 = 20	40
1	0/3 = 0	0
2	0/1 = 0	0
1	0/2 = 0	0
1	0/1 = 0	0
1	0/1 = 0	0
	70	96.66

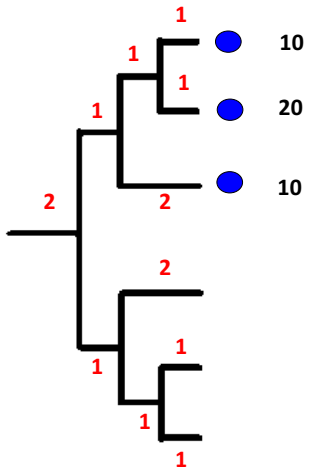
Weighted.pd = 11 * (70/96.66) = 7.966



Calculating Abundance Weighted Faith's Index (PD)

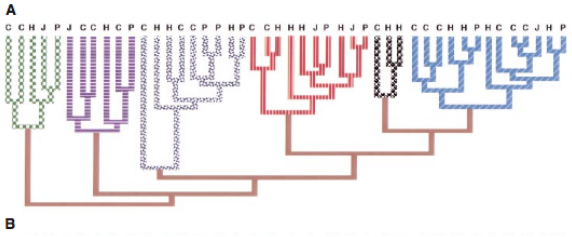
Branch Length	Average Abundance	Length x abundance
2	40/6 = 6.66	13.33
1	40/3 = 13.33	13.33
1	30/2 = 15	15
1	10/1 = 10	10
1	20/1 = 20	20
2	10/1 = 10	20
1	0/3 = 0	0
2	0/1 = 0	0
1	0/2 = 0	0
1	0/1 = 0	0
1	0/1 = 0	0
	75	91.66

Weighted.pd = 11 * (75/91.66) = 9.00



Functional Diversity (FD) via Petchey & Gaston

- FD was popularized by Owen Petchey and Kevin Gaston in 2002 in an Ecology Letters paper
- Takes the exact same calculation as Faith's Index and uses it on a trait dendrogram (essentially a morphological phylogeny)
- Downside is that it MUST use a dendrogram and cannot be used on raw trait distance matrices



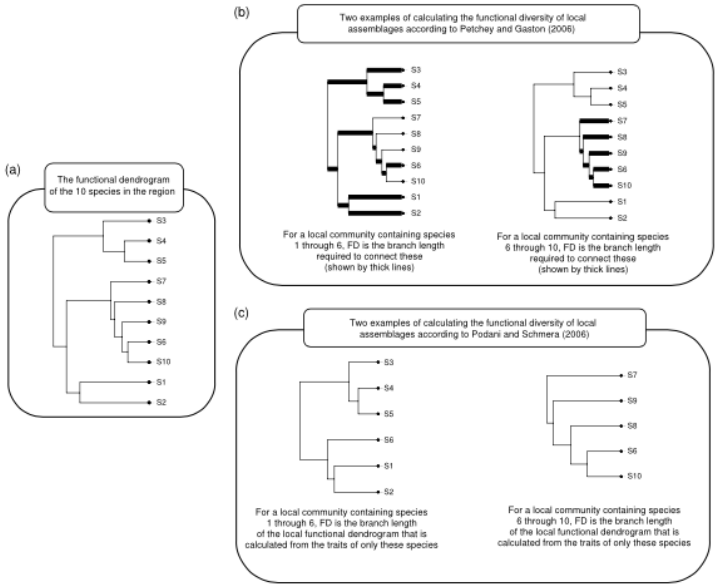


Fig. 2. A regional functional dendrogram can be constructed from functional trait information about the species in the region (left). One method to calculate the FD of the species in a local assemblage is to subset the regional dendrogram, so that the FD is the summed branch length (indicated by thick lines) across the regional dendrogram required to join the species present in the assemblage (Petchey and Gaston 2002b, 2006). Another method is to recalculate a local dendrogram from the trait information of the species present in the local assemblage, and use the summed branch length of this local functional dendrogram (Podani and Schmera 2006).

Functional Diversity (FD) via Petchey & Gaston

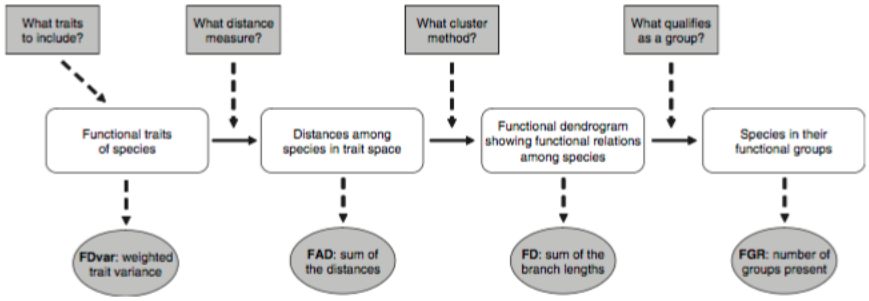


Figure 1 The process of producing a functional classification (unshaded objects) and estimating different measures of functional diversity (shaded ellipses). FDvar (Mason *et al.* 2003); FAD, plant attribute diversity (Walker *et al.* 1999); FD (Petchey & Gaston 2002b); and FGR, functional group richness. Less quantitative approaches implicitly contain all the same steps and decisions. The shaded rectangular boxes represent decisions in the process of making a classification, so that the number of decisions required for each measure increases from left to right.

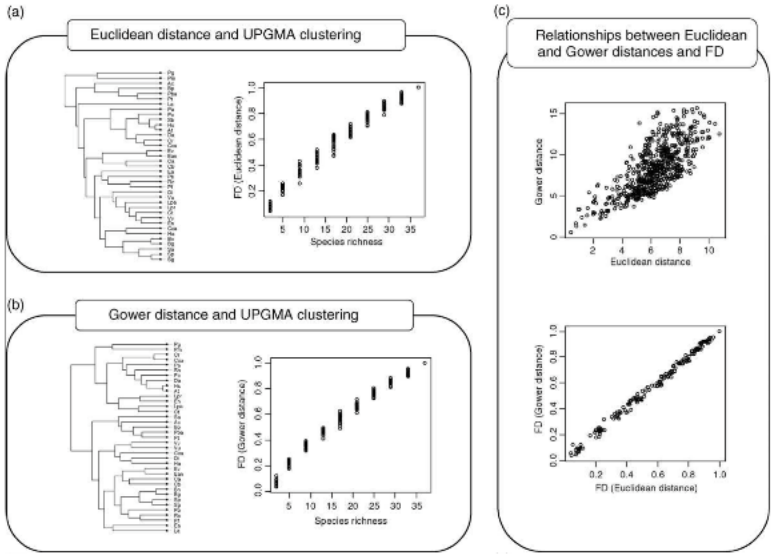


Fig. 1. The effects of Euclidean and Gower distances on analyses of FD for plant species assemblages, calculated from trait information in Chapin et al. (1996). (a) functional dendrogram when the distance measure was Euclidean, and the associated relationship between FD and species richness of 20 random assemblages at each of 10 levels of diversity. (b) functional dendrogram when the distance measure was Gower, and the associated relationship between FD and species richness of the same random assemblages. (c) relationships between Euclidean and Gower distances, and between FD values calculated using those distances.

Functional Diversity (FD) via Petchey & Gaston

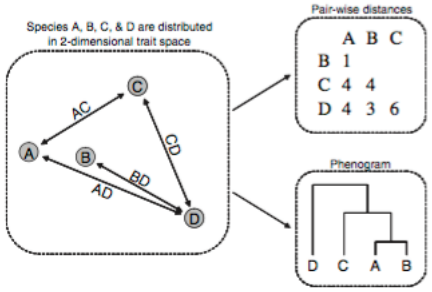
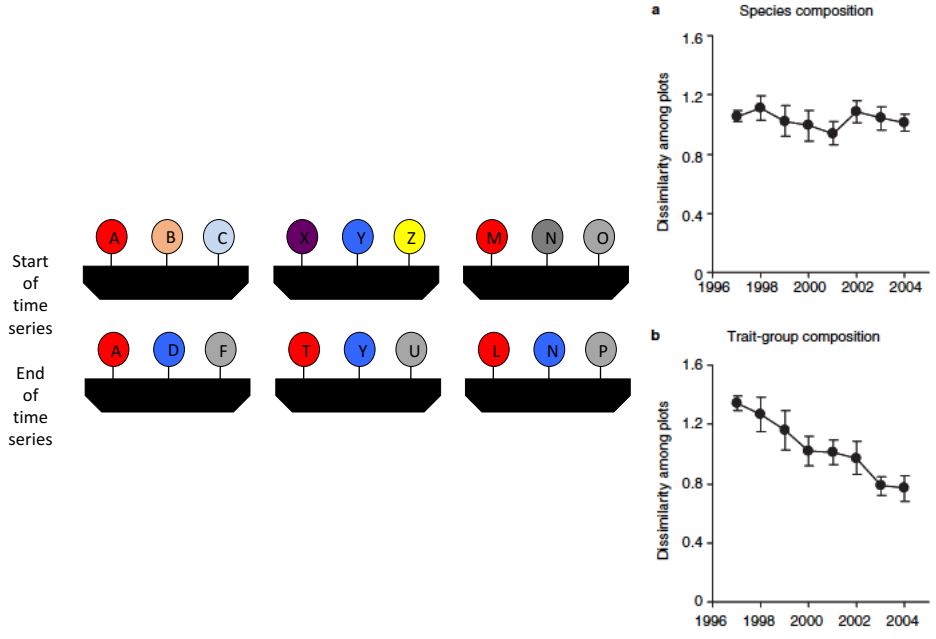


Figure 2 Measuring functional diversity is a problem of how to measure the amount of variation represented by a set of points in multivariate space, for example, species A, B, C and D in the left-most dashed box. The arrows between the species (shaded circles) represent four of the six pair-wise distances. All six pair-wise distances are given in the matrix in the upper-right dashed box. Pair-wise distances are used directly by some measures of functional diversity (Table 1). The phenogram in the lower-right dashed box is a hierarchical description of the distances between species. Some measures of functional diversity work directly on this phenogram (Table 1). In these examples, the distance metric (i.e. Euclidean, Manhattan and Jaccard) is arbitrary, as is the clustering method (e.g. average linkage and minimum linkage) that produced the phenogram.



Fukami et al. 2005 Ecology Letters

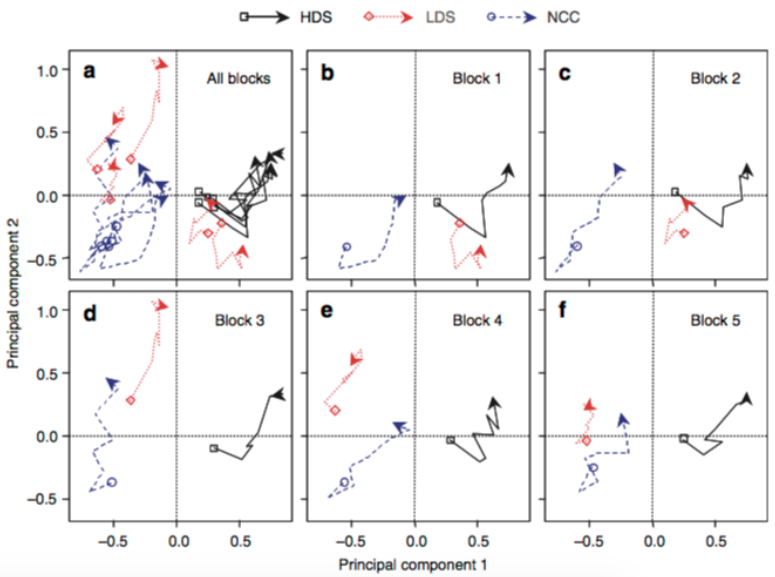


Figure 1 Changes in plant species composition from 1997 (square, diamond or circle) to 2004 (end of arrows). A high-diversity seed mixture (HDS), a low-diversity mixture (LDS), or no species (natural colonization control) (NCC) was sown in 1996 (Table 1). All graphs are based on a single common principal component analysis (see Materials and methods).

We constructed trait groups using the literature information on as many ecologically important species traits as possible (Tutin *et al.* 1964–1980; Grime *et al.* 1988; Thompson *et al.* 1997). These traits were related to life history, growth, dispersal, phenology, mycorrhizal association and other characteristics (see Table S1). Although certain plant traits are plastic, all of the traits we used, possibly except seed weight, are static and do not vary substantially with environmental conditions, making the literature information adequate for our purpose (see also Hérault *et al.* 2005). We ran these trait data through hierarchical clustering using Ward's method (Lepš & Šmilauer 2003) and used 14 clusters as our trait groups (Table 2). The cut-off for the number of clusters was partly determined by the limited species pool (there would be little point in clustering if too many groups were occupied by a single species), and partly by looking at the results of successive iterations. After 14 clusters, the successive subdivisions were difficult to describe in biological terms.

significant response to the sowing treatments, and should thus have negligible effects on the results. We acknowledge that species may not always fall into distinct functional groups and that it is often not straightforward to identify the ecologically most relevant way to construct trait groups (e.g. Lavorel & Garnier 2002; Petchey & Gaston 2002). Nonetheless, we found that non-hierarchical clustering using *K*-means (Legendre & Legendre 1998) produced qualitatively the same pattern as did Ward's method, indicating the robustness of our results to trait construction methods.