

## Instructions for calculating indices of Functional Diversity with the file “FunctDiv.xls”

Last modified: May 2008

Macro excel file: Jan (Šuspa) Lepš ([suspa@prf.jcu.cz](mailto:suspa@prf.jcu.cz))

Instruction manual: Francesco de Bello ([fradebello@ctfc.es](mailto:fradebello@ctfc.es)) & Jan Lepš

If you use the excel file please cite it as:

*Lepš J., de Bello F., Lavorel S. & Berman S. (2006). Quantifying and interpreting functional diversity of natural communities: practical considerations matter. Preslia 78: 481–501*

The macro is freely available at: <http://botanika.bf.jcu.cz/suspa/FunctDiv.php>

## In short

### *What does the macro calculate?*

- the RAO index for species and functional trait diversity. The indices are calculated for each sample (e.g., alpha diversity; but the macro can be used to calculate also beta and gamma diversity)
- different types of functional trait diversity indices based on Mason et al. (2005)
- the trait community weighted means (trait mean per sample weighted by species relative abundance)
- the relative abundance of species in a sample from a matrix of species x sample with the 'absolute' abundance of species.

### *Which kind of data is needed?*

- a species x sample composition data with the occurrence or, better, the abundance of each species x each sample (e.g. based on cover, biomass, n° of individuals etc..)
- a trait x species matrix (details how to prepare this matrix are given in section 3).

### *What kind of traits types can be handled?*

- quantitative traits (e.g. height, body size, leaf area) for which a mean and a standard deviation per each species is available (section 3.1)
- categorical and binary traits (e.g. nitrogen fixer vs non-nitrogen fixer, color; see section 3.2 and 3.3). Binary traits are also used as semi-quantitative variables for traits that have not been measured in the field (as taken by data bases, literature etc.). Dummy and fuzzy coding are allowed for these trait types.
- phenological and life cycle data (section 3.4)

## Biggest warnings

- **DO NOT CHANGE THE FILE NAME!!!!** otherwise the macro will not work any more (sorry!)
- **Diversity indices are sensible to the sampling method.** It is strongly recommended to compare diversity indices in data sets where samples have been sampled with the same method (i.e. in term of sample area, species abundance measurement, for example cover or n° of individuals etc..)
- **For matrices with many species the calculation of trait dissimilarity and Rao index might take a while** (a good excuse to boil some water for a tea!)

## 1. What is the macro about? What is functional diversity?

The excel file “FunctDiv.exl” contains macros for the calculation of some of the available indices of Functional Diversity (FD) in the literature (Mason et al. 2005; Botta-Dukat 2005; Petchey & Gaston 2006; Lepš et al. 2006; de Bello et al. 2006, 2007, 2008). The indices that can be calculated take into account two types of data:

- (a) the species’ traits (i.e. ‘matrix species x traits’)
- (b) the species relative abundance in different samples/plots (i.e. ‘matrix species x sample’)

This way, the indices calculated are thought to include information on the relative abundance of the species and do not calculate directly indices based only on species presence/absence (note: some tips on how to use this kind of data are also given in the text).

The macro has been originally prepared as an *ad hoc* tool for our personal needs (based on data mostly on plant and insect traits). As there is not such material available yet, other colleagues asked us to use the file and we finally decided to make it available for everybody who might need it. We are conscious that the macro ‘not very sophisticated’ and ‘amateur’ from the esthetic point of view and because it might not present a fancy user interface and the programming approach is very simple. (Warning: Jan insisted that he learned some programming in late seventies using simple Basic coded using punched tape, and his programming skills might have not improved much from this time; he said that inspecting the code itself could cause hard attack to any person familiar with programming.)

Nevertheless the macro works! We tested the functionality of the file with several data sets and with some (courageous) colleagues working on plants, insects, birds and benthic invertebrates (thanks Marco, Marie, Christian, Sandra, Benoit, Karl etc). After some short introductions (as resumed in this text) we were able to obtain reasonable biological results in a short lap of time. Also the file runs on Excel, so most of the people can use it (even if it is not a free software, sorry).

Suggestions and constructive criticism are always welcome (see emails above)!!!

### 1.1. Functional Diversity: a biological issue

The concept of functional diversity (FD) remains rather complex because there are questions about how to define and measure it (Petchey and Gaston 2006). The Functional Diversity, for example can have different definitions. An approach that has gained currency considers the FD as “the extent of functional trait variation (or differences) among the species in a community” (Tilman 2001, Petchey and Gaston 2002) or, in other words, “the overall difference among species in a community in terms of their traits”. For example, two assemblages with a similar amount of species may be more or less functionally diverse depending on how similar/dissimilar the species’ traits are among the species in the communities. Some other researchers (e.g. Diaz & Cabido 2001; Diaz et al. 2007) define functional diversity more broadly as the “kind, range and relative abundance of traits” (which include both indices of trait dissimilarity and community trait mean; see below). Here we use the first definition while also giving the tools to calculate all components mentioned by Diaz and colleagues.

The FD is an important biological issue because is linked to the way space share the niche space available in a community (Mason et al. 2005) and have important consequences on the functioning of ecosystems (Diaz et al. 2007). Trait composition effects on ecosystem functioning can be assessed by calculating two main indices (a) trait community weighted mean (Garnier et al. 2004; reflecting the mass ratio hypothesis) and (b) functional diversity. Both such indices are calculated by the macro.

### 1.2. Functional Diversity indices: basic mathematics

Two main indices of FD can be calculated by the file “FunctDiv.xls”:

- (a) the Rao index of diversity adapted for functional diversity (i.e. that uses species traits to calculate dissimilarity among species; Botta-Dukat 2005; Lepš et al. 2006; de Bello et al. 2008; Lavorel et al. 2008)
- (b) the index proposed by Mason and colleagues (2005), and modified by Lepš et al. 2006, which reflects the overall variance of traits values in a community.

The two indices have rather different formulas but, preliminary results indicate that their variation along environmental gradients is comparable similar (de Bello et al. 2008). Here we give some details on the Rao index, mostly based on Lepš et al. 2006. Please see Mason et al. (2005) for details on the other index (however instructions to calculate the index are discussed here in section 4).

The **Rao** coefficient presents several desirable properties for describing the FD of a community (see Ricotta 2005, Botta-Dukat 2005). In fact, it is a generalized form of the Simpson index of diversity. If the proportion of  $i$ -th species in a community is  $p_i$  (section 4.1 for details) and dissimilarity of species  $i$  and  $j$  is  $d_{ij}$  (section 3 for details), the Rao coefficient has the form:

$$FD = \sum_{i=1}^s \sum_{j=1}^s d_{ij} p_i p_j$$

where  $s$  is the number of species in the community and  $d_{ij}$  varies from 0 (two species have exactly the same traits) to 1 (the two species have completely different traits). If  $d_{ij}=1$  for any pair of species (so each pair of species is completely different), then  $FD$  is the Simpson index of diversity expressed as 1 minus Simpson index of dominance  $D$ , i.e.  $1 - \sum_{i=1}^s p_i^2$  (see e.g., Botta-Dukat 2005 for details).

The Rao index generally reflects the probability that, picking randomly two individuals in a community (i.e. a sample), they are different. For species diversity (i.e. the Simpson index of species diversity) this represents the probability that the individuals are from different taxa. For trait diversity, the Rao index represents the probability that they are functionally different (e.g. for single traits, either they have different trait values or different trait categories). The Rao coefficient is very flexible, and can be used with various dissimilarity measures (for example, Shimatani 2001 used it with taxonomic dissimilarity; asymmetrical measures can be also used, etc.). The main methodological decisions are mainly **how to measure the species dissimilarity**, and how to characterize the proportion of a species in the community. The same decisions, however, have to be made even if we decide for other indices of functional diversity (details in Lepš et al. 2006).

The species dissimilarity  $d$ , or also called ‘distance’, in the formulas is computed in different ways depending on the trait type. For quantitative traits, the **Overlap** in the trait space among pairs of species is calculated (figure next page). The **dissimilarity** is expressed as **1 minus Overlap** and thus it is scaled against between zero (no dissimilarity) and one (maximum dissimilarity). The user can use different types of data to calculate the dissimilarity between pair of species and should be aware on the basis by which this is calculated in the different type of data (**see section 3** for details). There basically are four types of data that can be used to calculate with the macro in the FuncDiv file: (i) *quantitative* data (section 3.1), (ii) *binary* traits or *continuous* traits that range *from zero to 1* (section 3.2), (iii) categorical data with more than two levels coded as dummy variables (section 3.3) and (iv) phonological and life cycle data (section 3.4).

The case of *quantitative values* for species’ traits is the most ideal case, but not often possible the case (because it is more time consuming for field

measurements and because some traits are only categorical, e.g. legume vs non-legume). For quantitative traits, the overlap is calculated as showed in the following figure (from Lepš et al. 2006: to calculate each curve the user needs the average for a given trait in every species and a measure of the Standard Deviation of this average).

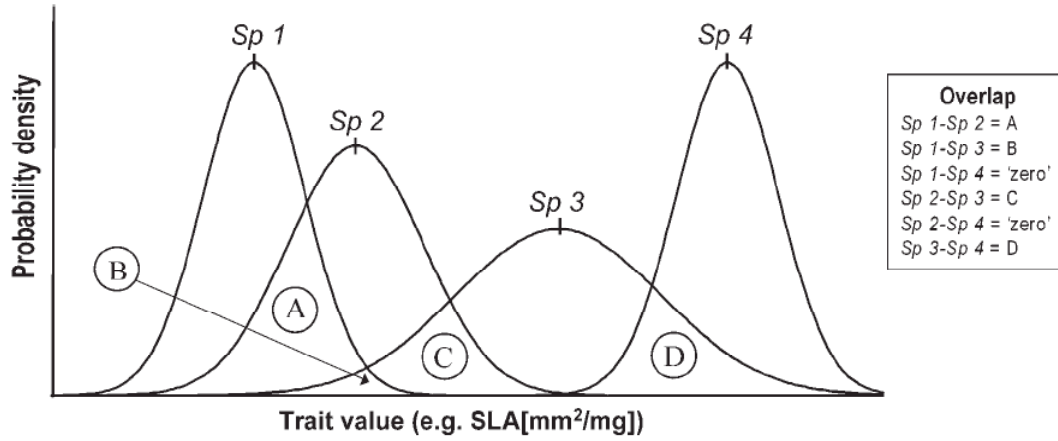


Fig. 3. – Schematic representation of the meaning of species overlap (*O*) based on their probability density (this can be estimated by knowing for example the mean trait value and its standard deviation). Note that the area below the curve is always unity by definition, so that  $1-O$  corresponds for each species to the part not shared with the compared species. Using the normal approximation, the overlap is always positive, but for very different species, the values are so small that can be practically considered zero. Also note that with roughly constant variability, the more the mean differ, the smaller overlap.

Please, note, that the standard deviation is a measure of variability between individuals within a species, not the variability of the estimate (so, use standard deviation, not standard error of mean). In the macro, the probability density of each species is approximated by a normal distribution, and then, the integral of the minimum value of the two compared probability density functions is calculated numerically (as a sum of very narrow rectangles!!).

Finally, the FunctDiv.xls file calculates various versions of the **Mason et al (2005)** index. Because the Mason et al. index is in fact an algebraic function of weighted standard deviation of log transformed trait values, the macro enables calculation of the SD log transformed trait values (weighted by the relative representation of the species) – macro “sdlog”. Because the log transformation is appropriate for some traits only (see Lepš et al. 2006), other macro (“sdnonlog”), calculates weighted standard deviation of original trait values. Also, to calculate these indices, the **macro calculates a community trait weighted mean** (Garnier et al. 2004), which is a very important index of functional trait composition that reflects components of the Mass Ratio Hypothesis (Garnier et al. 2004). The index, with  $x_i$  reflecting the trait mean for a given  $i$ -th species is as follows:

$$\bar{x} = \sum_{i=1}^S p_i x_i$$

## 2. Steps in calculating the FD

There are two basic steps in the calculation of the FD indices. **First** (section 3) the user will calculate the trait dissimilarity (i.e.  $d_{ij}$ ) among all pair of species (this step is not needed for the Mason index, so just jump it if you are using this approach!). To calculate  $d_{ij}$  a different macro will be used depending on the type of data traits available (called, depending on the trait type used 'disbin', 'discateg' and 'disim' and 'phenol'). **Then**, (section 4), when the dissimilarity is calculated, the user will be able to calculate the desired FD with another set of macros (depending on the FD indeed needed).

As commented above the user needs always two kinds of matrices to calculate the FD:

- (a) species x trait matrix (details in section 3)
- (b) species x sample matrix (details in section 4)

## 3. Calculate the dissimilarity among species

To compute the dissimilarity between every pair of species, the user has first to prepare an appropriate *species x traits matrix*. This matrix will include (**see Example 1**) the species name in the first column (the maximum species that can be considered to calculate species dissimilarity is at the moment **limited to 361** for the Rao index, not for Mason's indices). The species name can be either the real Latin-taxonomical name or any other type of label (no limits in characters is applied). In the other columns the user will introduce the traits info. As commented above the traits can be of different types, which correspond to using different sheets, different data inputs and different macros in the file (it is not so difficult as it sounds!).

### 3.1. Quantitative traits

Ideally the user have, for every species, a *quantitative value* for a given trait (as leaf area or body size) plus a standard deviation (SD) for this average (i.e. based on the measurements on the trait over several individuals; Cornellissen et al. 2003). For these trait type he user will use the sheet "**Data**" and will introduce the data this way (see **Example 1**): two columns for every trait, being the first (i.e. left one) the trait value (e.g. in m<sup>2</sup> grams...whatever, the measure unit is not important) and, in the second (e.g. the right one), the SD for that trait for every species. In the case of several quantitative traits, the different couple of columns for every trait will follow one to the other. Here below an example of how the traits should be introduced.

### Example 1 (introduction of *quantitative traits* into the relative 'Data' sheet)

Trait 1 (e.g. SLA here) and relative SD

Trait 2 (e.g. height here) and relative SD (sorry for the typo!!)

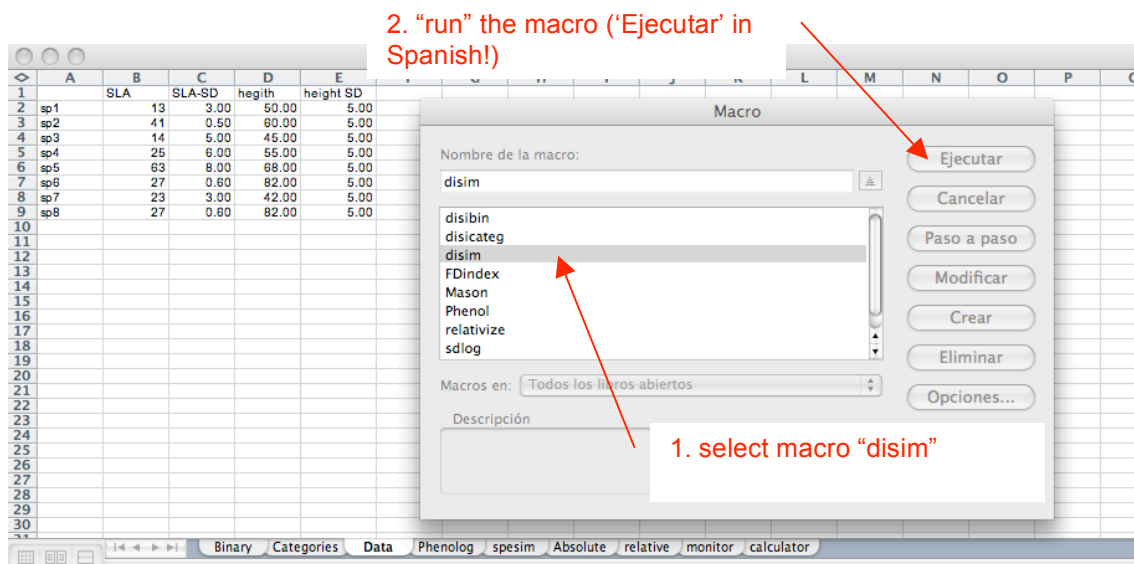
Use "Data" sheet!

	A	B	C	D	E	F	G	H	I
1		SLA	SLA-SD	height	height SD				
2	sp1	13	3.00	50.00	5.00				
3	sp2	41	0.50	60.00	5.00				
4	sp3	14	5.00	45.00	5.00				
5	sp4	25	6.00	55.00	5.00				
6	sp5	63	8.00	68.00	5.00				
7	sp6	27	0.60	82.00	5.00				
8	sp7	23	3.00	42.00	5.00				
9	sp8	27	0.60	82.00	5.00				

Binary Categories Data Phenolog spesim Absolute relative monitor calculator

Then the user is ready to calculate the trait dissimilarity between all pair of specie (the dissimilarity is calculated as shown in the figure in the section 1.2). The user should go to the excel command "Tools-Macro" (if English language is applied) and open the macro window as shown in the **Example 2**, chose the **macro called 'disim'** and then run the macro (click on "run"! ). The macro will ask you information on the number of species (8 in our example), and the number of **traits (called "variables"; 2 in our example)**.

### Example 2 (running the macro 'disim' with quantitative traits)



After the macro is run, the **results** are displayed in the **sheet "spesim"** as shown in the **Example 3**. The values range from 0 (no dissimilarity between a given pair of species) 1 (maximum dissimilarity). For example, the dissimilarity between sp1 and sp2 for SLA is 1 (i.e. no overlap, complete functional



difference) while the dissimilarity between sp4 and sp7 (line 22 in this case) is 0.35 (see circles in the **Example 3**; 0.35 means that, in the field, these two species are functionally identically for SLA in 35% of the cases). The user might decide to copy all these results in a separate file for further controls or for other analysis he/she might be interested to do with dissimilarity (see Petchey & Gaston 2006). Note, in fact, that each time trait dissimilarity is calculated, the new results are overwritten on the old ones. The macro also calculates the average dissimilarity between pairs of species in terms of all traits together (see Example 3). Note that the program does not report the original species names introduced in the “data” sheet.

**Example 3** (results of dissimilarity between all pair of species in the ‘spesim’ sheet)

	A	B	C	D	E	F
	species1	species2	SLA	height	average	
1	1	2	0.682685	0.841343		
2	1	3	0.255977	0.382923	0.31945	
3	1	4	0.830574	0.382923	0.606748	
4	1	5	0.999995	0.928136	0.964066	
5	1	6	0.999922	0.998625	0.999274	
6	1	7	0.904416	0.576285	0.740351	
7	1	8	0.999922	0.998625	0.999274	
8	2	3	0.999999	0.866382	0.933191	
9	2	4	0.992491	0.382923	0.687707	
10	2	5	0.995423	0.576285	0.785854	
11	2	6	1	0.972191	0.986096	
12	2	7	1	0.928136	0.964068	
13	2	8	1	0.972191	0.986096	
14	3	4	0.684678	0.682685	0.683682	
15	3	5	0.999841	0.97855	0.989195	
16	3	6	0.987177	0.999784	0.993481	
17	3	7	0.750732	0.235825	0.493278	
18	3	8	0.987177	0.999784	0.993481	
19	4	5	0.993434	0.806395	0.899914	
20	4	6	0.807458	0.993065	0.900262	
21	4	7	0.35369	0.806395	0.580042	
22	4	8	0.807458	0.993065	0.900262	
23	5	6	0.999983	0.838483	0.919233	
24	5	7	0.999753	0.990677	0.995215	
25	5	8	0.999983	0.838483	0.919233	
26	6	7	0.829209	0.999937	0.914573	
27	6	8	6.19E-05	6.29E-05	6.24E-05	
28	7	8	0.829209	0.999937	0.914573	
29						
30						

### 3.2. Binary and continuous traits scaled between 0 and 1

The most common case for which trait information is available is when the user has traits taken from flora, fauna books and/or databases. In this case, to use the FunctDiv file, we propose to transform all trait into a value by scaling it between 0 and 1. This will apply to (a) *binary traits* or (b) *continuous trait*. Let's see some examples to be clearer.

The case of *binary* traits is, for example, when an animal species “has” or “does not have” wings” or a plant “can” or “cannot” fix nitrogen (as ‘legume’ in **Example 4**). In this case the user has either zero or 1. For this trait type, the dissimilarity between pair of species is 1 when the two species are from different groups (one is coded “0” and the other one “1”) and 0 when they are not different (see results in **Example 6**). Note that for some apparently strange

(‘crazy’ might be the best word here; but actually just for being pragmatic, see below) each trait needs to be characterized by **two headers**. Therefore every trait (columns) will be characterized by a two lines (headers) repeating the same information (changing this will alter the results, see below).

**Example 4** (introduction of *binary* and *continuous* traits in the Binary sheet)

Two headers for every traits (reasons explained in Examples 7 –8)

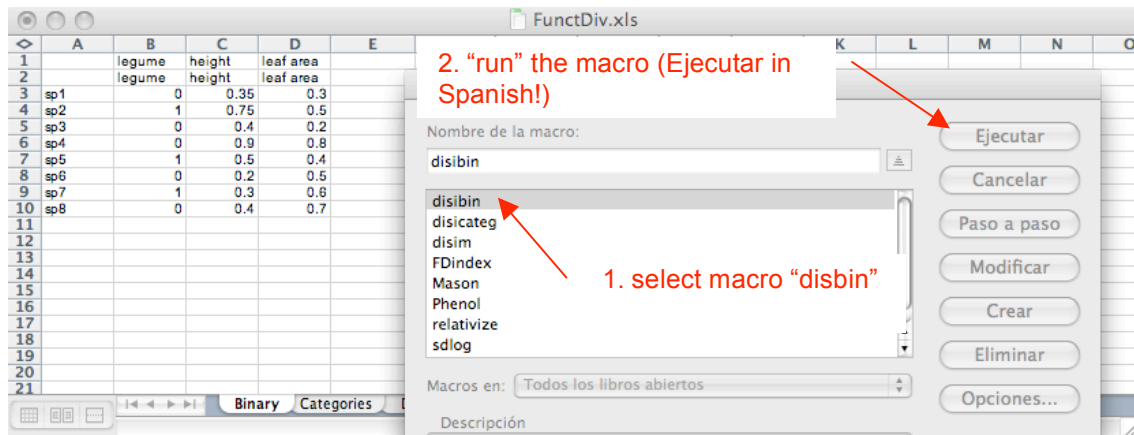
Use the “Binary” sheet

	A	B	C	D	E	F	K	L	M
1		legume	height	leaf area					
2		legume	height	leaf area					
3	sp1	0	0.35	0.3					
4	sp2	1	0.75	0.5					
5	sp3	0	0.4	0.2					
6	sp4	0	0.9	0.8					
7	sp5	1	0.5	0.4					
8	sp6	0	0.2	0.5					
9	sp7	1	0.3	0.6					
10	sp8	0	0.4	0.7					
11									
12									
13									
14									
15									

In the same ‘binary’ sheet the user can introduce semi-quantitative traits that need to be scaled between zero and 1 (as ‘height’ or ‘leaf area’ in **Example 4**). This is a way to deal with quantitative traits for which we have not measured the standard errors (this is the case, for example, of the size or plant height or body size taken from flora/fauna books). To resolve this, the traits should be scaled from zero (minimum value in the data set) and 1 (maximum value). For example, if the minimum plant height in the data set is 10 cm and the maximum is 110 cm, then any other height (e.g. 60 cm) could be scaled somehow in this range (e.g. it can give 0.5). Logarithmic scales can be used to calculate such proportions. The dissimilarity in this case is calculated as the difference between two species. For example, if the trait is height and we have two species having 0.35 (species 1, Example 4) and 0.75 (species 2, Example 4), the dissimilarity will be 0.4 (Example 6).

When the data is well introduced into the “binary” sheet, the user is ready to calculate the trait dissimilarity between all pair of specie. The user should go, as shown in the section 3.1 and example 2) to the excel command “Tools-Macro” and open the macro window (as shown in the **Example 5**), chose the macro called “disibin” and then run the macro (click on “run”!). The macro will ask you again the information on the number of species (always 8 in our example), but this time it will understand automatically how many traits are available (indeed an intelligent machine!). As a matter of fact the macro will automatically count the number of different traits, by looking at the number of columns with different header in the first row. That’s one of the reasons why the two headers are needed.

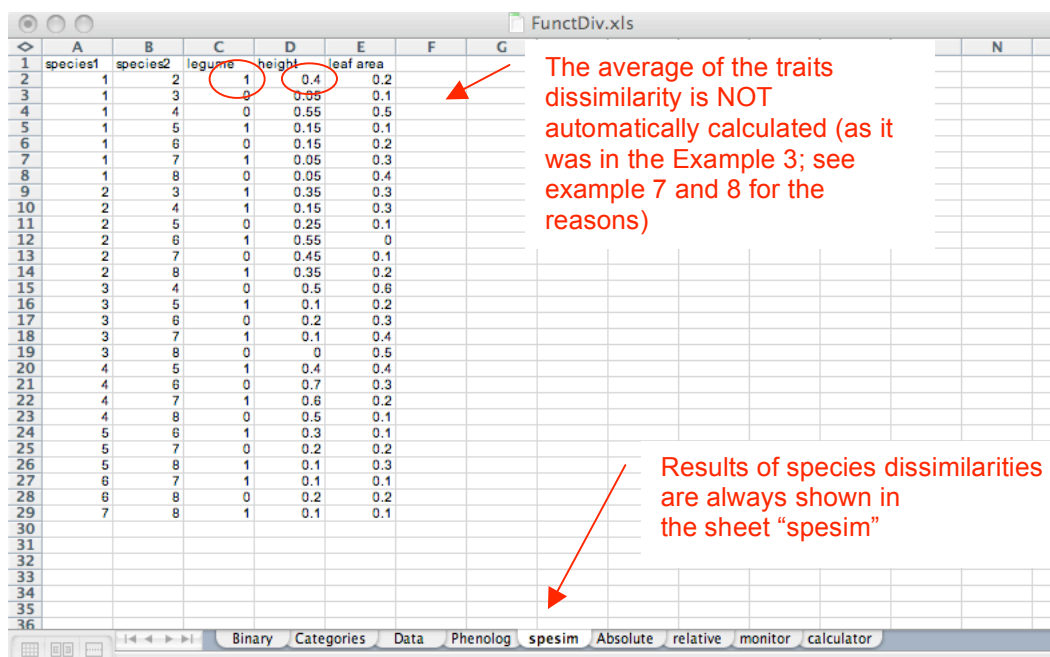
**Example 5** (calculating the dissimilarity for *binary* and *continuous traits* using the ‘disbin’ macro)



After the macro was run, the results will be displayed in the sheet “spesim” (Example 6). **Note that, each time, a different macro (either disbin, discateg or disim) is used to calculate the dissimilarity between species, the data are overwritten in the specsim sheet.** So it is better to keep this clean.

In the results (**Example 6**). The values range from 0 (no dissimilarity between a given pair of species) 1 (maximum dissimilarity). In this case, the macro DOES NOT calculate the average dissimilarity between all pairs of species in terms of all traits together (as in Example 3).

**Example 6** (results of dissimilarity)



The reason of the presence of two headers for every column (and for the absence of automatically calculated averages) is because the user, by changing the first header can change the way the dissimilarity is calculated. If, following the Example 4-6, he/she puts the same header to “height” and “leaf area” in the first row (for example calling it “plantsize”) the user will be able to calculate directly the dissimilarity for the two traits together (**Examples 7-8**).

**Example 7** (changes in the headers change the calculations)

	A	B	C	D	E	F	G	H	I	J	K
1		legume	plantsize	plantsize							
2		legume	height	leaf area							
3	sp1	0	0.35	0.3							
4	sp2	1	0.75	0.5							
5	sp3	0	0.4	0.2							
6	sp4	0	0.9	0.8							
7	sp5	1	0.5	0.4							

**Example 8** (changes in the headers change the calculations and the way the results are shown)

	A	B	C	D	E	F	G	H	I	J	K
1	species1	species2	legume	plantsize							
2	1	2	1	0.3							
3	1	3	0	0.075							
4	1	4	0	0.525							
5	1	5	1	0.125							

### 3.3. Categorical traits coded as dummy variables

The user might also introduce that data for categorical (=nominal) traits with more than two categories, coded as several *dummy variables* (**Example 9**). For example the color of a flower can be yellow-black-red-blue. This gives four categories. So the user should build the traits as four columns where, for each species (row), the sum must be 1 (for example if the flower is blue, then he/she will put one to the column for blue flower and zero for other colors). A special case of dummy variables is the use of *fuzzy coding*. If we have, for example, a flower that is half red and half blue we can code 0.5 for the “red” column and 0.5 for the “blue” one, and zero for the other categories. If we have another trait as life form for plants (GF in Example 9), a plant that can behave either as hemicriptophyte and chamaephyte will have 0.5 for both categories. The user will use the excel **sheet “categories”** for these kind of data. In the matrix a different header for every trait should be given in the first row, while on the second there are the names of the categories in which the traits are divided (see **Example 9**).

### Example 9 (introduction of *dummy* traits)

The first header is the trait type (e.g. flower color)

The second refers to the categories of the trait

The total per species for a given trait is 1

Use the "Categories" sheet

	FLWcolor	FLWcolor	FLWcolor	FLWcolor	GF	GF	GF	GF	GF	GF
	yellow	black	red	blue	Therop.	Hemicri.	Chamaep.	Phanaer.	Geophy.	
sp1	0	0	0	0	1	0	0.5	0.5	0	0
sp2	0	0	0	0	1	0	0	0	0	1
sp3	0	0	0.5	0.5	0.25	0.5	0.25	0	0	0
sp4	0	1	0	0	0	1	0	0	0	0
sp5	1	0	0	0	1	0	0	0	0	0
sp6	0	0.5	0.5	0	0	0	0	0	1	0
sp7	0	1	0	0	0	0	0.5	0.5	0	0
sp8	1	0	0	0	1	0	0	0	0	0

When the data is introduced in the "Categories" sheet, the user is ready to calculate the dissimilarity between all pair of specie. The user should go to the excel command "Tools-Macro" and open the macro window as shown in the **Example 2 and 5** (for the other macros), chose the macro called "**disicateg**" and then run the macro (click on "run"! ). The macro will ask you information on the number of species (8 in our example). The macro will automatically count the number of traits, by the number of columns with different header in the **first** row. The basic difference between this macro (i.e."disicateg" using data in the "Categories" spreadsheet) from the "Binary" data (macro "disibin", section 3.2) is that, with 'disicateg', the values in columns belonging to a single trait **MUST** sum up to one for a species. Consequently, the dissimilarity for species one and two for a trait is calculated as  $(\sum_i (x_{1,i} - x_{2,i}))^2 / 2$  ( $i$  is the category index) . The resulting value is one, if the two species do not share any trait. For crisp classification, the values are either 0 or 1, for fuzzy classification, the values range between 0 and 1 (see results in **Example 10**)

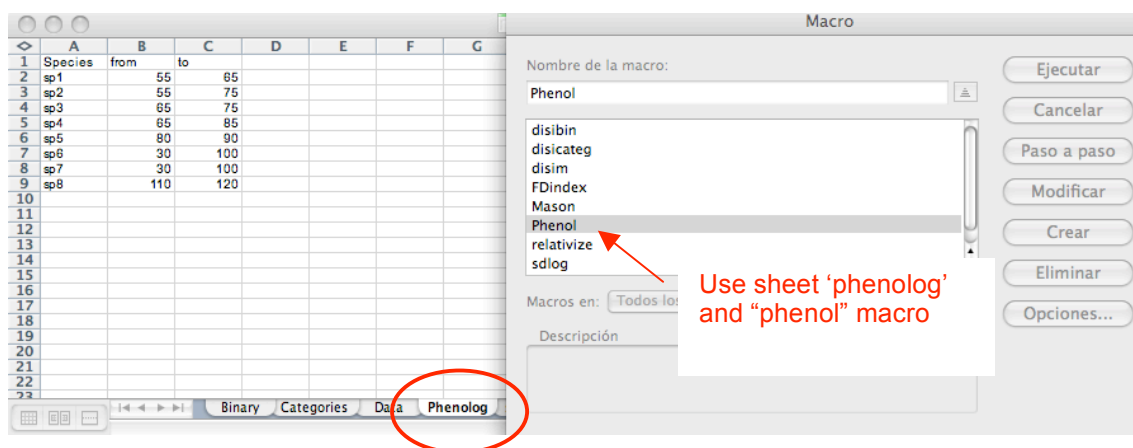
### Example 10 (some of the results of dissimilarity for dummy traits in the spesim sheet)

species1	species2	FLWcolor	GF
1	2	0	1
1	3	0.5	0.25
1	4	1	0.5
1	5	1	1
1	6	1	1
1	7	1	0.5
1	8	1	1
2	3	0.5	1
2	4	1	1
2	5	1	1
2	6	1	1
2	7	1	1
2	8	1	1
3	4	1	0.5
3	5	1	0.75

### 3.4. Phenological and life cycle data

Sometimes, the user have data which can be defined as ‘circular’, i.e. the starting flowering date in a plant (or a trait defined by some inclination or angle). How to calculate the dissimilarity between two species flowering one in December and the other one in January? If we count the number of months (12<sup>th</sup> and 1<sup>st</sup>) in the year the dissimilarity will wrongly be estimated (11). We can thus introduce in another sheet (“phenolog”) data similar on plant flowering period (or corners, life cycle). This sheet calculates the dissimilarity between pair of species as **1 - (number of days when both the species flower together/number of days during which at least one of them is flowering)**. This value might be site specific. To introduce the data the user should thus have data of the Julian day when each species is starting and ending the flowering, and introduce the data as shown in the **Example 11**. Most detailed reason why such kind of data should not be calculated as simple quantitative values is discussed in Leps et al. 2006. The macro runs similarly to macro presented in section 3.1, 3.2 and 3.3 and the result of species dissimilarity will be also displayed in the working sheet “spesim”.

**Example 12** (Introducing data on plant flowering period (‘from’ = starting julian day; ‘to’ ending day) in the ‘phenolog’ sheet).



**Example 13** (results from the macro ‘phenol’ shown in the spesim sheet)

FunctDiv.xls												
	A	B	C	D	E	F	G	H	I	J	K	L
1	species1	species2	Phenolog	DisSimil								
2	1	2	0.47819									
3	1	3	0.952381									
4	1	4	0.967742									
5	1	5	1									
6	1	6	0.84507									
7	1	7	0.84507									
8	1	8	1									
9	2	3	0.47819									
10	2	4	0.845181									
11	2	5	1									
12	2	6	0.704225									
13	2	7	0.704225									
14	2	8	1									
15	3	4	0.47819									
16	3	5	1									
17	3	6	0.84507									
18	3	7	0.84507									
19	3	8	1									
20	4	5	0.769231									
21	4	6	0.704225									
22	4	7	0.704225									
23	4	8	1									



## 4. Calculate the Functional Diversity

Once the user has calculated the dissimilarity among species' traits (needed only for the Rao index), most of the work has already been done (congratulation for surviving this difficult step, if we meet someday remind us to invite you for a beer or the preferred non-alcoholic drink)!! Before doing the final step (calculating the FD), the user should now introduce a *species x sample matrix* and be conscious of the different possibilities he/she has here and the various implications of the choice.

### 4.1. Species x plot matrix

This is the matrix where the user says in which samples (or sometimes also called/defined as “plots” or “communities”) he found the species and in which abundance they were in each sample. If we take the 8 species of the previous examples we can introduce their abundance in the appropriate sheets. Ideally the macros for Rao and Mason indexes need the information of  $p_i$  (see for example the Rao FD formula at page 4 and the Simpson index of species diversity at page 5), which is the relative abundance of species in a plot (this means that the total for every sample should sum 1; see **Example 14**). To calculate the relative frequency the user should consider the following. In general, the users have a “ $n_i$ ” e.g. the number of individuals (or number of contacts, biomass, or what ever measure of species abundance) for every  $i$ -th species in the sample. Then sum of the “ $n$ ” of all the species gives “ $N$ ”, i.e. the sum the total number of individuals for a given sample. The relative abundance for a given  $i$ -th species is  $n_i/N$ , dividing the number of individuals for a given species for  $N$  (for example, if the species1 has 25 individuals, or grams, in sample1 and  $N$  is 100, then the relative abundance is 0.25; see **Example 14 and 15**). The user that wants to use the formula for FD for **presence/absence data** (i.e. no measure of species abundance are available or the user does not want to consider them), can solve this by introduce the same value for the species abundance for all species present in a sample (with the total always summing 1). This is the case for example of column B (Sample1 with 4 species  $1/4=0.25$ ) in the Example 14:

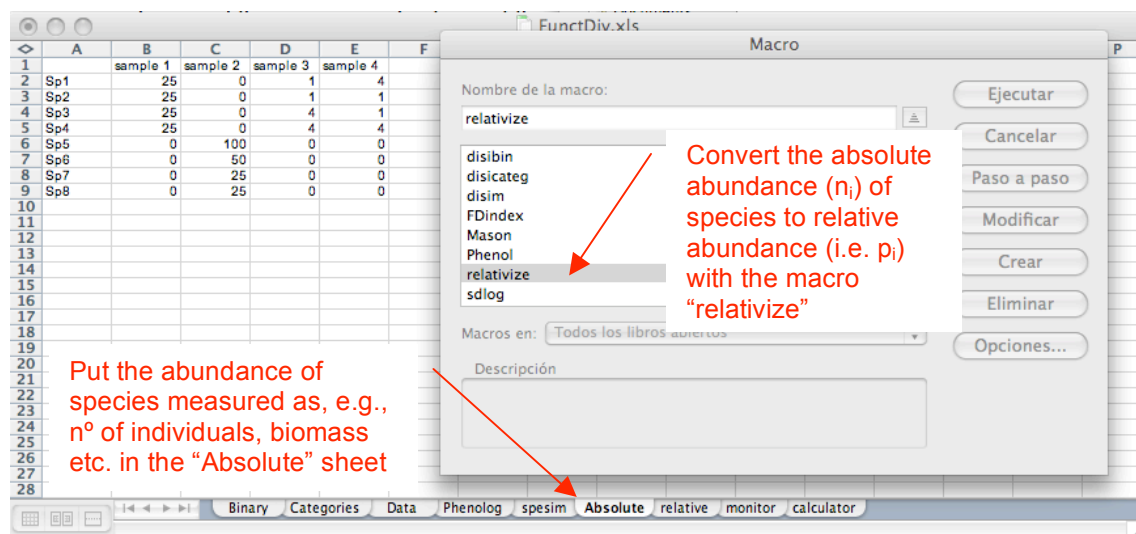
**Example 14** (example of the species x sample matrix having, for each species and plot a given relative abundance. The total for each column need to be = 1)

	A	B	C	D	E	F	G	H	I	J	K	L
1		sample 1	sample 2	sample 3	sample 4							
2	Sp1	0.25	0	0.1	0.4							
3	Sp2	0.25	0	0.1	0.1							
4	Sp3	0.25	0	0.4	0.1							
5	Sp4	0.25	0	0.4	0.4							
6	Sp5	0	0.5	0	0							
7	Sp6	0	0.25	0	0							
8	Sp7	0	0.125	0	0							
9	Sp8	0	0.125	0	0							
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												

Binary Categories Data Phenolog spesim Absolute relative monitor calculator

Very often the users don't have ready-made data of species relative abundance but rather “absolute” values (i.e. with “ $n_i$ ”). This is the case of having rough data of, e.g., the number of individuals, or biomass, or cover for each species. To make life easier a macro (i.e. ‘relativize’) was added to the file to calculate directly the relative abundance of species from the ‘absolute’ abundance of species. For this, the “absolute” data need to be introduced in the “Absolute” sheet and the user need to run the macro called “relativize” (see **Example 15**).

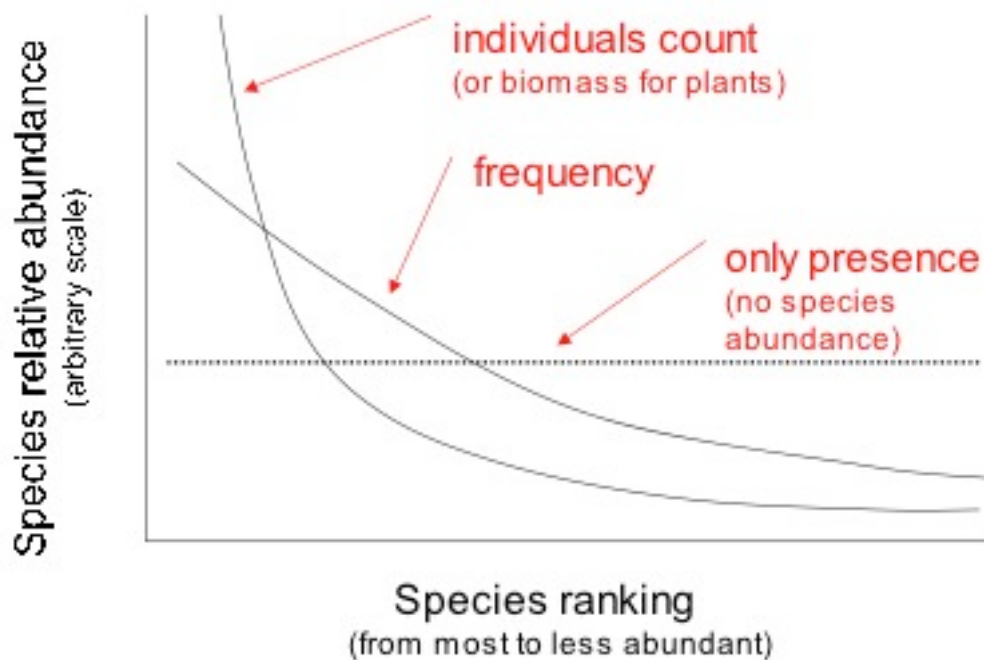
**Example 15** (transforming the absolute abundance of species to relative abundance, i.e.  $p_i$  always needed for FD Rao, Simpson, Mason indices and the community weighted trait mean. Put the data into the “Absolute” sheet and use macro “relativize”)



The user should be aware that different measures of species abundance give a different weight to the dominant species in a community (see details in Lepš et al. 2006; de Bello et al. 2007). There are three main common ways to quantify the relative abundance of species in communities: counts of individuals (population density), frequency or cover estimates, and biomass (then each measure can be transformed, e.g. log-transformed, also affecting dominance of species; de Bello et al. 2007). It is well known to field ecologists that these quantities are not equivalent. An important difference between biomass- and frequency-based calculations of relative abundances is that evenness in the former is usually lower than in the latter (usually much lower when the frequency is based on large basic sampling units). It is hence not uncommon for 5-10 species, if not fewer, to make up the majority of the biomass (e.g. 80%), while a larger number of species (10-20) may be needed to achieve the same threshold on a frequency basis (particularly when using relatively large basic sampling units). Similarly, the biomass values (and also the number of individuals in animal studies) vary over several orders of magnitude among species, whereas the cover varies less and frequency even less. The way of quantifying species' relative abundances has thus important consequences for calculations of compound diversity (including functional diversity) indices because “artificially” the evenness of a community.



**Example 16:** (Effect of different measures of species abundance on the relative abundance of the species. The example shows the species' ranking for an hypothetical community where the user uses the count of individuals or the biomass for plants (maximum importance to most abundant species), the frequency or cover and the case where all the species have the same abundance (only presence/absence data considered), which mimics the highest evenness).



As the FD indices in the macro use the relative abundance of the species ( $p_i$ ), it is clear that a different weight to the traits of the dominant species (more when the species-rank curves are less flat) and subordinate species. Different measure of the species relative frequency gives rather different values for a given community (de Bello et al. 2007). From preliminary observation we observed that the more dominance we have in a sample the more is the chance to have species (i.e. Simpson) and functional diversity calculated with the Rao index are correlated.

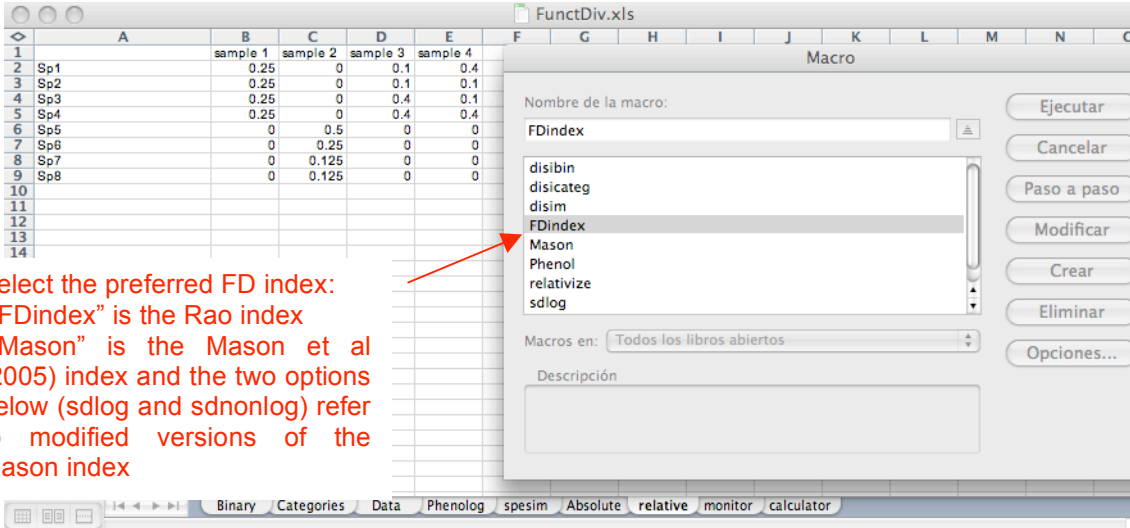
#### 4.2. Calculate the FD indices (finally!!!!)

To finally calculate the Rao FD for each sample, the user will use the results of the dissimilarity among species (those that are in the “spesim” sheet, e.g. the ones for the **Example 3**, for quantitative traits), and the data on the page of the relative abundance of the species (those that are in the “relative” sheet **Example 14**; **Note that for the RAO index, the data in the other sheets are not used; for the Mason index the sheet used are ‘data’ and ‘relative’**).

##### THE RAO index

The user should go to the excel command “Tools-Macro” and open the macro window as shown in the **Example 17**. Then he/she should select the desired FD index. By choosing the macro called “FDindex”, and then run the macro (click on “run”!), the user will calculate the Rao index of diversity. The macro will ask you information on the number of species (8 in our example), the number of samples (also 4 in our example). **The user will be asked also about some crazy information, i.e. the number of the column where the dissimilarity starts in the sheet “spesim” and the ending column (Example 18)**. This means, from our Example 3, with two traits, that the dissimilarity for the first traits was on column C (the third column) so that the “dissimilarity starting column” is equal to 3 (the first two columns refer to species identity; please note the starting column will be always the third in the ‘spesim’ sheet, but the user might be interested to use only some traits starting from another column) and the ending column is where is the last trait (column E, i.e. 5<sup>th</sup> column; i.e. dissimilarity ending column in equal to 5 in the example; note that in the Example 3 the column E is the average of the columns C-D, so E is the dissimilarity for two traits together).

##### Example 17 (ready to calculate the FD!)



Select the preferred FD index:  
-“FDindex” is the Rao index  
-“Mason” is the Mason et al (2005) index and the two options below (sdlog and sdnonlog) refer to modified versions of the Mason index

Finally the user has some results (**Example 19**)!!!!!! **The results are ALWAYS DISPLAYED in the ‘RELATIVE’ SHEET**. For every sample the user gets a value of the FD calculated in terms of the traits considered (in our Example 3, SLA and height). Beside this, the first result (first row) is given by the calculation

of the Simpson index of species diversity (see section 1.3). To calculate a compound index of functional diversity in terms of various (multiple) traits the user can make a simple average of the FD calculated by single traits (Lepš et al. 2006).

**Example 18** (select the dissimilarity starting and ending columns; note that you can start the macro irrespectively of which of the excel sheet are active at the moment)

In the sheet 'spesim' the column C (i.e. the third column from the left) hosts the first information on trait dissimilarity and column E (i.e. n° 5 from the left) the last one.

Put 3 if you want to use SLA as a first trait

Microsoft Excel dialog box: dissimilarity starting column (3), Aceptar, Cancelar.

	A	B	C	D	E
	species1	species2	SLA	hegith	ave
1					
2	1	2	0.255977	0.382923	0.8
3	1	3	0.830574	0.382923	0.606748
4	1	4	0.999995	0.928136	0.964066
5	1	5	0.999922	0.998625	0.999274
6	1	6	0.904416	0.576285	0.740351
7	1	7	0.999999	0.866382	0.933191
8	2	3	0.992491	0.382923	0.687707
9	2	4	0.995423	0.576285	0.785854
10	2	5			
11	2	6			
12	2	7			
13	3	4	0.684678	0.682685	0.683682
14	3	5	0.999841	0.97855	0.989195
15	3	6	0.987177	0.999784	0.993481
16	3	7	0.750732	0.235825	0.493278
17	4				
18	4				
19	4				
20	5				
21	5				
22	6				
23	6				

**Example 19** (results for the calculation of the Rao index of diversity, i.e. Simpson species diversity, first line, and functional diversity one line for each trait, with the same names as introduced in the 'spesim' sheet; **The results are ALWAYS DISPLAYED in the 'RELATIVE' SHEET**).

Results for Simpson species diversity and FD for two traits (+ average for the two traits)

	A	B	C	D	E	F	G	H	I
		sample 1	sample 2	sample 3	sample 4				
1									
2	Sp1	0.25	0	0.1	0.4				
3	Sp2	0.25	0	0.1	0.1				
4	Sp3	0.25	0	0.4	0.1				
5	Sp4	0.25	0	0.4	0.4				
6	Sp5	0	0.5	0	0				
7	Sp6	0	0.25	0	0				
8	Sp7	0	0.125	0	0				
9	Sp8	0	0.125	0	0				
10									
11									
12									
13									
14	Simpson	0.75	0.65625	0.66	0.66				
15	SLA	0.6327052	0.2884147	0.5818537	0.5447196				
16	hegith	0.4259405	0.4728544	0.471589	0.341528				
17	average	0.5293228	0.3805346	0.5267214	0.4431238				
18									
19									
20									
21									
22									
23									

The results indicate for example that, with all plots having the same number of species, sample 1 has the highest Simpson species diversity (i.e. the lowest dominance and highest evenness) but, for example, not the highest FD for height (sample 2 and 3 have higher).

## The MASON and COMMUNITY TRAIT MEAN indices

The user, by calculating a modified version of the Mason index (e.g., mostly using the macro called 'sdnonlog'), will be also allowed to **calculate the community weighted trait mean** (see section 1.3 for the formula). This is therefore a quite useful and time saving macro, even if the index is quite easy to calculate.

Note that the **Mason index**, and its modifications (macro 'sdlog' and 'sdnonlog') use always **the sheets spesim and relative only!!!** Note also that the data on trait standard deviations are not used in the Mason index, so these data is not required. So and so, to avoid changing all the time the data format, the macro will read the trait value every second column (so that columns, 3-5-7 etc....should be left empty if the user does not have data on standard deviation data). This can be important also if we want to calculate the FD Mason also for other type of traits. For example, we can copy in the 'data' sheet, the information of each species being or not a legume (which was the first trait on the 'binary' sheet). See this shown in the **Example 20**.

**Example 20** (we can calculate the modified Mason index and the community weighted trait mean with macro 'sdnonlog', i.e. the last one! The macro uses the data into the sheet 'data'. The information on trait standard deviation (SD), see Example 2, is not needed for this index, so it can be even deleted. If no information on SD is available, the user should nevertheless introduce every single trait in each second column, leaving the other empty. The information of other type of traits, e.g. binary as legume vs non legume, can be copied into the 'data' sheet from other data sheet.

The SD is not needed (but the space-column between traits is needed). Other, non quantitative traits can be used (here copied from the binary sheet).

The sdnonlog macro (the last) calculates the modified Mason index and the community trait weighted mean

The macro use information from the 'data sheet, and not 'spesim'

	A	B	C	D	E	F	G
1		SLA		height		legume	
2	sp1	13		50.00		0	
3	sp2	41		60.00		1	
4	sp3	14		45.00		0	
5	sp4	25		55.00		0	
6	sp5	63		68.00		1	
7	sp6	27		82.00		0	
8	sp7	23		42.00		1	
9	sp8	27		82.00		0	
10							
11							
12							
13							
14							
15							

Macro window details:

- Nombre de la macro: sdnonlog
- Macros en: Todos los libros abiertos
- Descripción: (empty)

To start calculations, the user should go to the excel command "Tools-Macro" and open the macro window as shown in the **Example 20**. Then he/she should select the macro 'sdnonlog' and then run the macro (click on "run"! ). The macro will ask you information on the number of species (8 in our example), the number of samples (also 4 in our example). **The user will be asked also about some how many variables (i.e. traits) are available.** In the Example 20, there are 3 traits available (SLA, height and legume), so the answer is 3!

The results of the macro 'sdnonlog' are shown, as all results, in the 'relative' sheet. In the **Example 21** it is possible to observe two type of results: the community weighted trait mean (above, called 'average') and the modified Mason index which calculate the standard deviation for a given trait in each sample (below, called 'sd'; see Leps et al. 2006 for details). Note that the weighted mean reflect, for binary traits, the relative abundance of different groups of species, e.g. legumes in the example, or in other terms the possibility to find a legume species in a given sample.

**Example 21** (the results of the 'sdnonlog' macro for the community mean, above, and functional diversity below, called here 'sd' because of its formulation)

Community weighted trait mean for each sample. For legume in the example this reflects the relative abundance of legume species in a sample

The modified Mason FD index, called here SD, because it is equal to the standard deviation for a given trait in a sample.

	sample 1	sample 2	sample 3	sample 4
Sp1	0.25	0	0.1	0.4
Sp2	0.25	0	0.1	0.1
Sp3	0.25	0	0.4	0.1
Sp4	0.25	0	0.4	0.4
Sp5	0	0.5	0	0
Sp6	0	0.25	0	0
Sp7	0	0.125	0	0
Sp8	0	0.125	0	0
SLA average	23.25	44.5	21	20.7
height average	52.5	70	51	52.5
legume average	0.25	0.625	0.1	0.1
SLA sd	11.277744	18.540496	8.508819	8.7641314
height sd	5.5901699	12.409674	5.3851648	4.0311289
legume sd	0.4330127	0.4841229	0.3	0.3

In the example, sample 2 has the highest community weighted trait mean for SLA and height, so it means that dominant species in this plot are both taller and with higher SLA leaves (or in other words that the average individual in the community, if species abundance is used, is tall and with high SLA). In the same sample 2 there is also a higher coexistence of different trait types (the FD is higher for SLA, height and legume).

**NOW LET'S CALCULATE FD INDICES WITH YOUR REAL DATA!!!!**

## Acknowledgements

The macro was developed with the support, deep suggestions and enthusiasm of Sandra Lavorel and Karl Grigulis (Laboratoire d'Ecologie Alpine. CNRS, Université Joseph Fourier). Sandra Berman, from the same group studied the sensitivities of the FD indices to different measures of species abundance. Marco Moretti (WLS, Switzerland) tested with great patience and precision the functionality of several versions of the macros with his and Simon Potts (University of Reading) data. Specifically, the work was supported by the EU VISTA project (EVK-2002-00356), the Czech Ministry of Education projects MS 6007665801 and LC 06073, and the project DIVHERBE from the French ACI-ECOGER programme. FdB would also personally supported by the Laboratory of Plant Ecology and Forest Botany in Forestry and Technology Centre of Catalonia (CTFC, Spain). We also thank Marie Vandewalle (University of Lund) and Christian Feld (University of Essen) for being courageous enough to test the macros with their data.

## References (cited and on the topic)

- Botta-Dukat Z. (2005): Rao's quadratic entropy as a measure of functional diversity based on multiple traits. – *J. Veg. Sci.* 16: 533-540.
- Cornelissen J. H. C., Lavorel S., Garnier E., Diaz S., Buchmann N., Gurvich D. E., Reich P. B., ter Steege H., Morgan H. D., van der Heijden M. G. A., Pausas J. G. & Poorter H. (2003): A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. – *Austral. J. Bot.* 51: 335-380.
- de Bello F., Lepš J. & Sebastià M.-T. (2005): Predictive value of plant traits to grazing along a climatic gradient in the Mediterranean. – *J. Appl. Ecol.* 42: 824-833.
- de Bello F., Lepš J. & Sebastià M.-T. (2006): Variations in species and functional plant diversity along climatic and grazing gradients. – *Ecography* 29:801-810
- de Bello, F., Lepš, J., Lavorel, S., & Moretti, M. 2007b. Importance of species abundance for assessment of trait composition: an example based on pollinator communities. *Community Ecology* 8:163-170.
- de Bello, F., Thuiller, W., Lepš, J., Choler, P., Clement, J. C., Macek, P., Sebastia, M. T., Lavorel, S. Partitioning of functional diversity reveals the extent of trait convergence and divergence. *Journal of Vegetation Science*, in press.
- Diaz S. & Cabido, M. (2001): Vive la différence: plant functional diversity matters to ecosystem processes. – *Trends Ecol. Evol.* 16: 646-655.
- Diaz S., Cabido M. & Casanoves F. (1998): Functional implications of trait-environment linkages in plant communities. – In: Weiher E. & Keddy P. (eds), *Ecological Assembly Rules: Perspectives, Advances, Retreats*, p 338-362. Cambridge University Press, Cambridge.
- Díaz, S., Lavorel, S., de Bello, F., Quétier, F., Grigulis, K., & Robson, T.M. 2007. Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America* 104: 20684-20689
- Fukami, T. et al. 2005. Species divergence and trait convergence in experimental plant community assembly. – *Ecology Letters* 8: 1283-1290.
- Garnier, E., Cortez, J., Billes, G., Navas, M.L., Roumet, C., Debussche, M., Laurent, G., Blanchard, A., Aubry, D., Bellmann, A., Neill, C., & Toussaint, J.P. 2004. Plant functional markers capture ecosystem properties during secondary succession. *Ecology* 85: 2630-2637.
- Grime J. P. (2001): Plant strategies, vegetation processes and ecosystem properties. – John Wiley & Sons.
- Hooper, D. U. et al. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. – *Ecological Monographs* 75: 3-35.
- Huston, M. A. 1994. Biological diversity: the coexistence of species in changing landscape. – Cambridge University Press.
- Lanta V. & Lepš J. (2006) Effect of functional group richness and species richness in manipulated productivity-diversity studies: a glasshouse pot experiment. *Acta Oecologica-International Journal of Ecology*, 29, 85-96
- Lavorel S. & Garnier E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology*, 16, 545-556
- Lavorel S., Diaz S., Pausas J., McIntyre S., Cornelissen J. H. C., Garnier E., Harrison S. P., McIntyre S., Pausas J., Pérez-Harguindeguy N., Roumet C., & Urcelay C. (2006): Plant functional types: are we getting any closer to the Holy Grail? – In *Terrestrial Ecosystems in a Changing World* (eds J. Canadell, L. F. Pitelka & D. Pataki), in press. Springer-Verlag.
- Lavorel, S., Grigulis, K., McIntyre, S., Garden, D., Williams, N., Dorrough, J., Berman, S., Quétier, F., Thébault, A., & Bonis, A. 2008. Assessing functional diversity in the field – methodology matters! *Functional Ecology* 22:134-147.
- Loreau M., Mouquet N. & Gonzalez A. (2003): Biodiversity as spatial insurance in heterogeneous landscapes. – *Proc. Natl. Acad. Sci.* 22: 12765-12770.
- MacArthur R. H. & Levins R. (1967): The limiting similarity, convergence and divergence of coexisting species. *Amer. Natur.* 101: 377-385.
- Mason N. W. H., MacGillivray K., Steel J. B. & Wilson J. B. (2003): An index of functional diversity. – *J. Veg. Sci.* 14: 571-578.
- Mason N. W. H., Mouillot D., Lee W. G. & Wilson, J. B. (2005): Functional richness, functional evenness and functional divergence: the primary components of functional diversity. – *Oikos* 111: 112-118.

- Mayfield M.M., Boni M.E., Daily G.C. & Ackerly D. (2005) Species and functional diversity of native and human-dominated plant communities. *Ecology*, 86, 2365-2372
- McGill B. J., Enquist B. J., Weiher E. and Westoby M. (2006): Rebuilding community ecology from functional traits. - *Trends Ecol. Evol.* 21: 178-185.
- McIntyre S. and Lavorel S. (2001): Livestock grazing in subtropical pastures: steps in the analysis of attribute response and plant functional types. – *J. Ecol.* 89: 209-226.
- Petchey O. L. & Gaston K. J. (2002): Functional diversity (FD), species richness and community composition. – *Ecol. Lett.* 5: 402-411.
- Petchey O. L. & Gaston, K. J. (2006): Functional diversity: back to basics and looking forward. – *Ecol. Lett.* 9: 741-758.
- Petchey O. L., Hector A. & Gaston K. J. (2004): How do different measures of functional diversity perform? - *Ecology* 85: 847-857.
- Petrů M., Tielborger K., Belkin R., Sternberg M. & Jeltsch F. (2006): Life history variation in an annual plant under two opposing environmental constraints along an aridity gradient. - *Ecography* 29: 66-74.
- Rao, C. R. 1982. Diversity and Dissimilarity Coefficients - a Unified Approach. - *Theoretical Population Biology* 21: 24-43.
- Ricotta C. (2005): A note on functional diversity measures. - *Basic and Applied Ecology* 6: 479-486.
- Shimatani K. (2001): On the measurement of species diversity incorporating species differences. - *Oikos* 93: 135-147.
- Magurran A. E. (2004): *Measuring biological diversity*. – Blackwell, Oxford.
- Stevens, R. D. et al. 2003. Patterns of functional diversity across an extensive environmental gradient: vertebrate consumers, hidden treatments and latitudinal trends. - *Ecology Letters* 6: 1099-1108.
- Stubbs, W. J. and Wilson, J. B. 2004. Evidence for limiting similarity in a sand dune community. - *Journal of Ecology* 92: 557-567.
- Tilman D. (2001): Functional diversity. - In: Levin, S. A. (ed.), *Encyclopedia of biodiversity*, p. 109-120, Academic Press.
- Warwick, R. M. and Clarke, K. R. 1998. Taxonomic distinctness and environmental assessment. - *Journal of Applied Ecology* 35: 532-543.
- Weiher E., von der Werf A., Thompson K., Roderick M., Garnier E. & Eriksson O. (1999): Challenging Theophrastus: a common core list of plant traits for functional ecology. – *J. Veg. Sci.* 10: 609-620.
- Westoby M, Falster D.S., Moles A.T., Vesk P.A, Wright I.J. (2002): Plant ecological strategies: Some leading dimensions of variation between species. *Ann. Rev. Ecol. Syst.* 33: 125-159.
- Westoby M. (1998): A leaf-height-seed (LHS) plant ecology strategy scheme. - *Plant and Soil* 199: 213-227.