

## TECHNICAL REPORT

# Measuring the number of co-dominants in ecological communities

ZAAL KIKVIDZE<sup>1\*</sup> AND MASAHIKO OHSAWA<sup>2</sup>

<sup>1</sup>*Laboratory of Ecology, Faculty of Science, Chiba University, Yayoi-cho 1-33, Inage-ku, Chiba 263 8522, Japan and* <sup>2</sup>*Laboratory of Biosphere Functions, Institute of Environmental Studies, University of Tokyo, Hongo 7-3-1, Tokyo 113 0033, Japan*

We suggest a concept that allows the objective determination of the number of co-dominants in a community. We define co-dominants as a subset of species that are more abundant and more uniformly distributed than other species in a given sample. We compare the sample with a model community and use Simpson's diversity index to estimate the apparent number of co-dominants. Dominant species determined in this way are responsible for 70–90% of the total measure of abundance in the sample. The statistical significance of the apparent number of co-dominants may be assessed by a randomization test.

**Key words:** abundance distribution; community structure; diversity; dominance; Simpson's index.

## INTRODUCTION

Biological diversity has two components: the number of species (richness) and their relative abundance (evenness). Ecologists have developed many indices that combine these two components (for a review see Magurran 1988); however, recent studies point to Simpson's index for its certain mathematical advantages. These include its ability to produce unbiased estimations from a sample of reasonable size, its predictable dependence on sample size/sampling effort that permits accurate extrapolations, and its ability to measure similarity between communities (Lande 1996; Smith & Wilson 1996; Kikvidze 2000; Lande *et al.* 2000). In this article, we attempt to show that Simpson's index may be used: (i) for an objective estimation of the number of dominant species in a commu-

nity; and (ii) as a statistic for a Monte Carlo technique to test whether the estimated number of dominants fits a hypothetical model of abundance distribution.

## NUMBER OF CO-DOMINANTS

Distinguishing dominant and subordinate species within a community is an important initial step of any ecological study (Grime *et al.* 1988). The problem resembles the American game 'Catch 22'; to know the true dominants we shall understand the principal mechanisms governing the given community, but to understand the mechanisms we need to know the dominants beforehand and to concentrate the research on these dominant species. Initially, we can only base our judgment on abundance data and estimate the number of apparent dominants as candidates for being true dominants; then we can amend this list step-by-step with our improved knowledge of the mechanisms. However, analytical procedures to directly address the question of which species may be estimated

---

\*Author to whom correspondence should be addressed. Email: zaal@nature.s.chiba-u.ac.jp

Received 11 June 2001.

Accepted 5 December 2001.

dominants are surprisingly scant. Dominance is considered a difficult concept to discuss in quantitative terms (Dajoz 1977; Greig-Smith 1983), although it may be characterized by a diversity measure such as Simpson's index (Magurran 1988). In general, dominants are designated subjectively (e.g. Murray *et al.* 1999).

When we attempt to describe diversity by richness only, we create a 'flat' image of community structure: we ignore differences in relative abundances and implicitly give equal share  $1/S$  to each species ( $S$  is the number of species). Now let us try a 'black-and-white' conversion of such an image: (i) let us ignore the share of subordinate species (we still recognize their presence, but give them zero share; by this the richness  $S$  will not change); and (ii) we ignore differences in the relative abundance of dominants by giving equal share  $1/A$  to each dominant species ( $A$  is the number of co-dominants). Such a view also has a mechanistic implication. In particular, we consider dominants a consistent part of the community, and we expect them to occur within a given site because they are best fitted to the local environmental conditions and are relatively strong competitors. Conversely, subordinate species are a chaotic element of a community that despite lower fitness and/or weaker competitive ability still occur for various reasons, such as mass effect and disturbance. Accordingly, we postulate that the abundance distribution of co-dominants should be more uniform than that of the whole assemblage. Therefore, we can compare the actual dominance ranking in a community with the 'black-and-white' model in which a certain subset of species, or co-dominants, have equal share of the community and the remaining species are zero. We can then identify the model most similar to the sample by various criteria; we postulate that the number of co-dominants in this model estimates the apparent number of dominants in the sample (see also Ohsawa 1984).

We suggest a dominance measure as the principal criterion of the similarity between the sample and its 'black-and-white' model: the dominance in the sample must be equal to the dominance in the model. This can be measured using Simpson's index:

$$\lambda = \sum_{i=1}^S x_i^2,$$

where  $x_i$  is the relative abundance of the  $i$ th species and  $S$  is richness. Simpson's index  $\lambda$  measures the probability that two randomly chosen individuals from a given community are the same species (Lande 1996; Lande *et al.* 2000). Consequently, this probability will be equal for both the sample and its model at  $\lambda_{\text{sample}} = \lambda_{\text{model}}$ . Note that in the model,  $x_i = 1/A$  when  $i \leq A$ , and  $x_i = 0$  when  $i > A$  (species ranked descending in respect to their relative abundance); this gives  $\lambda_{\text{model}} = 1/A$ . Consequently, when  $\lambda_{\text{sample}} = \lambda_{\text{model}}$ , then  $\lambda = 1/A$ , or  $A = 1/\lambda$ . By this, we attribute the apparent number of co-dominants to the reciprocal of Simpson's index.

There is also a technical consequence of the equality  $\lambda_{\text{sample}} = \lambda_{\text{model}}$ . In particular,  $\lambda$  is related to the variance of relative abundance in a sample:

$$\begin{aligned} \text{Var} * (S-1) &= \sum_{i=1}^S (x_i - 1/S)^2 \\ &= \sum_{i=1}^S x_i^2 - 1/S = \lambda - 1/S. \end{aligned}$$

As we keep the richness  $S$  equal for the sample and its model, at equal dominance we will have  $\text{Var}_{\text{sample}} = \text{Var}_{\text{model}}$ . Therefore, the similarity of the sample and its 'black-and-white' model may also be seen in the equality of their variances.

We analyzed data from various sources (79 samples from plant communities and 25 samples from animal communities), and estimated the apparent number of co-dominants and their share in the samples using the procedure outlined above, namely as  $A = 1/\lambda$ ; to calculate the share of co-dominants we rounded up the value of  $1/A$  and picked the corresponding number of most abundant species from the community and then summed their relative abundance (Table 1). The share of co-dominants is  $\sim 0.8$  and this value shows remarkably low variation when compared to other indices (e.g. it is almost 13-fold less variable than the apparent number of co-dominants as shown by the corresponding values of the coefficient of variation). Thus, in different communities we may find various numbers of co-dominants, but their share will be large and robustly constant.

As we estimate the apparent number of co-dominants based on Simpson's index, we may expect that the former will retain the mathematical advantages of the latter. In particular, we examined the relationship between co-dominants and

**Table 1** The mean values of the apparent number of co-dominants ( $A$ ) and their share in 104 different communities

Index	Mean $\pm$ SD	CV
Richness ( $S$ )	24.25 $\pm$ 13.43	0.554
Apparent number of co-dominants ( $A = 1/\lambda$ )	8.19 $\pm$ 6.39	0.78
Share of co-dominants	0.815 $\pm$ 0.05	0.061

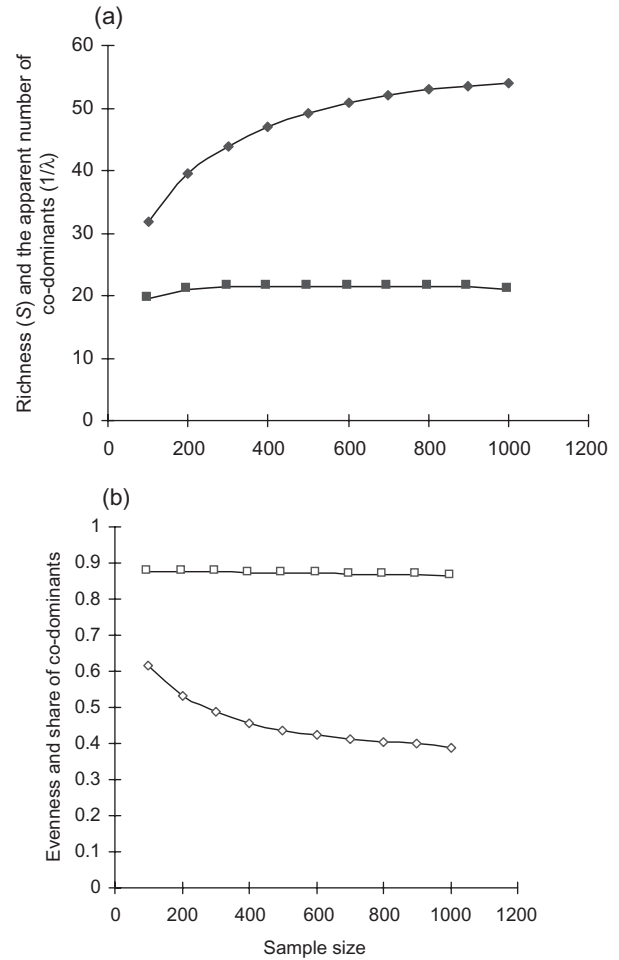
CV is the coefficient of variation ( $CV = SD/\text{mean}$ ).

sample size. In all tested cases, the share was as constant as the value of the apparent number of co-dominants  $A = 1/\lambda$  over a wide range of sample sizes, whereas richness and evenness were not. Figure 1 shows an example of one of the tests conducted on a hay meadow in the central Caucasus Mountains (altitude 2100 m). Data on species presence/absence were collected with a 10 cm  $\times$  10 cm square placed randomly 250 times within a 5 m  $\times$  5 m plot. We calculated average curves based on 100 randomized orders of the counting squares. Sample size was measured as the number of hits. Evenness was calculated as  $(1/\lambda)/S$ .

**ALTERNATIVE APPROACHES AND EXAMPLES**

A method of dominance analysis suggested by Ohsawa (1984) may be considered an alternative approach for the determination of co-dominants. This method was devised as a modification of the multi-indicator method of Yeates (1968). In a community dominated by a single species, its relative dominance may be stated as 1. If, however, two species share dominance the relative abundance of each should ideally be 0.5, or if there are three co-dominants it should be 0.33(3), and so on. The number of dominant species  $T$  is defined as the number of species that shows the least distance between the actual dominance values and the expected share of the corresponding co-dominant-number model. The distance ( $E$ ) is calculated by the following equation:

$$E^2 = \sum_{i=1}^T (x_i - 1/T)^2 + \sum_{i=T+1}^S (x_i - 0)^2 \quad (1)$$



**Fig. 1.** Dependence of community structure characteristics on sample size. (a) Richness ( $\blacklozenge$ ) versus the reciprocal of Simpson's index ( $\blacksquare$ ). (b) evenness ( $\blacklozenge$ ) versus the share of co-dominants ( $\square$ ).

where  $x_i$  is the relative frequency of the  $i$ th species (ranked descending),  $E$  is the Euclidean distance between the model and the given sample,  $S$  is richness (the number of species in a sample), and  $T$  is the number of co-dominants. Algebraic conversions of equation 1 produce:

$$E^2 = \sum_{i=1}^S x_i^2 - \left( 2 * \sum_{i=1}^T x_i - 1 \right) / T.$$

The first member of the left part of this equation is nothing else than Simpson's concentration  $\lambda$  (Simpson 1949). Therefore, this equation may be re-written as:

$$E^2 = \lambda - \left( 2 * \sum_{i=1}^T x_i - 1 \right) / T.$$

**Table 2** Determining the number of co-dominants by the minimum Euclidean distance between the model ( $T$ ) and the samples

$T$	Proportion	Cumulative proportion	$(2*\sum x_i - 1)/T$	$E^2$
Plot 1				
1	0.917	0.917	0.834	0.012
2	0.069	0.986	0.486	0.36
3	0.014	1	0.333	0.513
Plot 2				
1	0.498	0.498	-0.004	0.429
2	0.415	0.913	0.413	0.012
3	0.062	0.975	0.317	0.108
4	0.025	1	0.25	0.175

Obviously,  $\min E$  depends on the cumulative abundance function  $\sum x_i$ ; in particular,  $E^2$  reach its minimum at maximum values of  $(2\sum x_i - 1)/T^*$ , as demonstrated by the examples shown in Table 2. Reciprocals of Simpson's index for these examples equal 1.18 and 2.35 for plots 1 and 2, respectively. The rounded values of  $1/\lambda$  coincide with the estimates of the dominance analysis. In general, estimates produced by  $A$  and  $T^*$  are tightly correlated ( $r^2 = 0.96$ ,  $P < 0.0001$ ,  $n = 79$ ). However, Simpson's index is easier to calculate and, therefore, suits randomization tests.

Another approach is to judge the number of dominants subjectively. We deliberately produced simple imaginary examples to demonstrate how the methods work and we believe that an experienced ecologist would produce the same estimates of co-dominants for these examples. However, the next empirical example shows how ambiguous subjective estimations can be.

Figure 2 displays the frequency distribution of 54 species found in a hay meadow plot (data are the same as Fig. 1). Neither the distribution nor its dominance–diversity curve produces a clear indication of the dominants and subordinates within this community. At some points the frequency declines sharply, as if there were interruptions, and we have no idea of which of these 'interruptions' divides the co-dominant species from the rare species. The number produced by dominance analysis using Euclidean distance is  $T^* = 19$ . Simpson's index produces the apparent

number of co-dominants  $A = 21$ . If we examine the curve at these points, we can see an 'interruption' from species number 20 to species number 21. It would be fair to state that for the first approach there are 20 co-dominants in this sample. Naturally, the life history and behavior of particular species can correct this estimate, but this needs to be examined in a further study.

## MONTE CARLO ASSESSMENT

The next step is an assessment of the obtained number of co-dominants. An abundance model can provide a reference for such an assessment when used as a null hypothesis for a randomization test (Wilson 1993). Randomization methods do not require any assumptions on data distribution, but rather generate an empirical distribution from the observed sample. This advantage is especially important with the small sample sizes that are common in ecological studies, and randomization tests are increasingly replacing standard statistical tests in ecological research (Slade 1999; Fortin & Jacquez 2000).

We used the random fraction model as a null hypothesis for our randomization procedure. This model is a stochastic analog of a log normal distribution (Keating 1998). Although the original model considers only resource partitioning, we have included other processes. In particular, we believe that the relative abundance of different species is affected by, for example, competition for resources (Tokeshi 1993), metabolic requirements that imply plant size and allometric scaling relationships (Brown 1999), mass-effect from adjacent areas and migrations from remote areas combined with irregular disturbances. Accordingly, we attribute the total abundance to a 'stick' that is fractionated stochastically by many ecological processes; we set the 'stick' equal to one, then randomly break it into two pieces (splitting ratio was allowed within the range of 0.05–0.95). Then we randomly selected one of the pieces and broke it in two, yielding a total of three pieces. We repeated this process until there were  $S$  pieces, equal to the sample richness. The length of the  $i$ th piece represents the relative abundance  $x_i$  of the  $i$ th species. We used  $A = 1/\lambda$  as a test statistic; we repeated the randomization 1000 times and

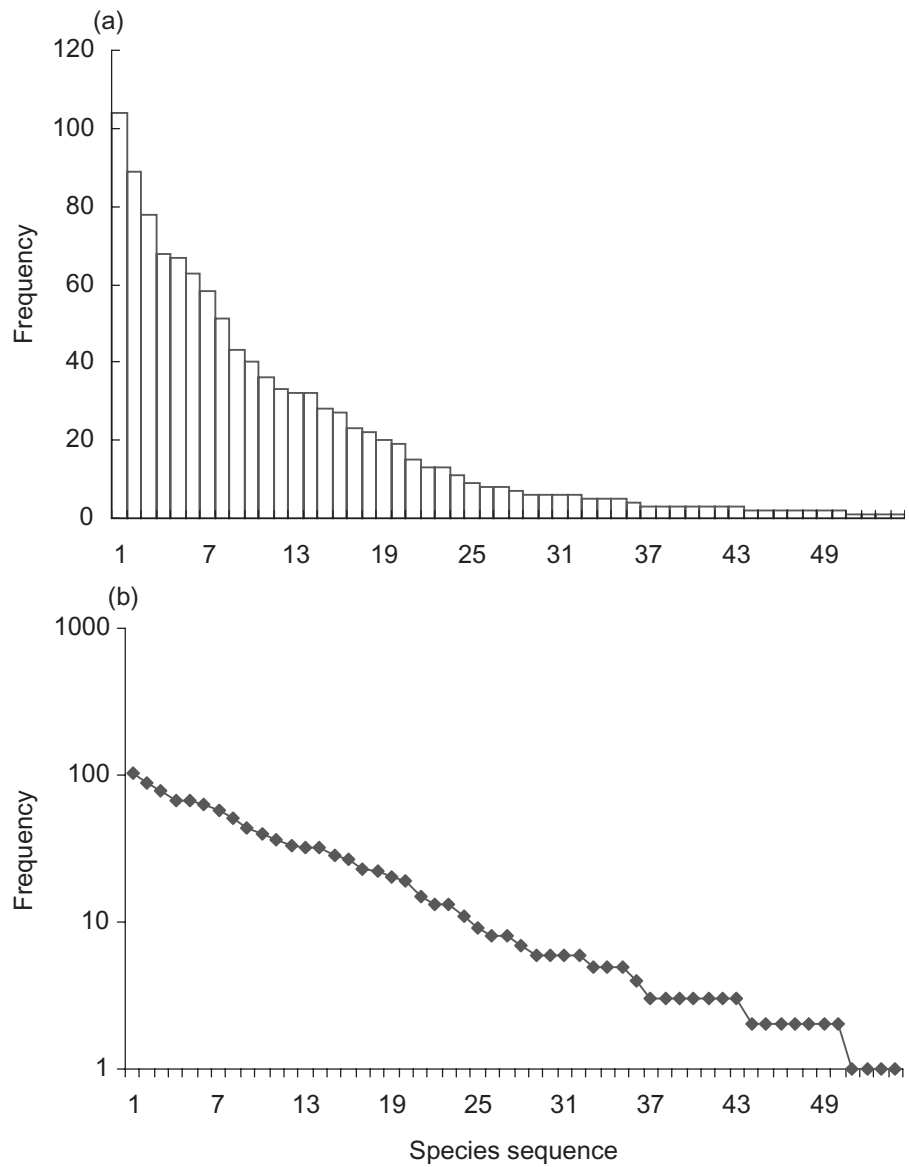


Fig. 2. (a) An example of frequency distribution and (b) corresponding dominance-diversity curve from a hay meadow community in central Caucasus sampled with small squares placed randomly.

Table 3 Examples of randomization test

Location	Community type	Richness	Index		Significance of difference ( <i>P</i> )	Data source
			$A = 1 \lambda^{-1}$ Observed	Expected		
W Europe remote forests	Birds	35	20.98	6.87	<0.001	Magurran (1988, Table 4)
	Birds	35	8.37	6.87	0.306	Magurran (1988, Table 4)
	Birds	26	1.96	5.99	0.041	Magurran (1988, Table 4)
SW Japan	Forest	35	7.42	6.87	0.363	Hara <i>et al.</i> (1996)
NE Japan	Forest	37	1.64	7.14	0.001	Hara (1991)
S Taiwan	Forest	67	20.27	9.24	0.016	Hara <i>et al.</i> (1997)
The Caucasus grasslands	Grazed	68	18.06	9.21	0.056	Kikvidze (2000)
	Mown	76	28.07	9.68	0.001	Kikvidze unpublished

calculated the randomized values of this statistic, whereas the significance level was determined from the number of randomization values equal to, or more extreme than, that observed.

Such a null hypothesis assumes a prevalence of stochastic processes in the formation of the abundance distribution. For any given richness this model produces a 'central' mean value of  $A = 1/\lambda$ , which we took as the expected value. If the observed and expected values do not differ significantly, the observed pattern is primarily determined by the stochastic random fraction mechanism. However, if any well-developed ecological process reduces stochasticity, then we will find a significant difference between the observed and expected values of  $A = 1/\lambda$ . Table 3 shows some examples of the randomization test applied to data from bird, forest and grassland communities. Examples are given that: (i) conform well to the random fraction model; (ii) show significantly more co-dominants than expected; and (iii) show significantly less co-dominants than expected. It is important to note that we could not find a sample with significantly less  $A = 1/\lambda$  in grasslands; this is consistent with a common view regarding the high diversity of alpine and subalpine grasslands.

Different ecological processes may reduce stochasticity in abundance distributions. For instance, positive plant-plant interactions at a neighbor scale and environmental patchiness at a community scale can support more co-dominants than expected. Conversely, competition with a highly tolerant species can trim down the number of co-dominants to a level significantly less than expected by the random fraction null model. An assessment of the apparent number of co-dominants by a randomization technique can generate hypotheses about the factors responsible for the observed pattern in community structure and, hence, can point where to direct further studies examining the processes and mechanisms.

## CONCLUSIONS

The reciprocal of Simpson's index is a recommended diversity measure (Hill 1973; Magurran 1988; Smith & Wilson 1996), yet it lacks a straightforward interpretation. Our attempt shows that  $1/\lambda$  may be viewed as the apparent number of

co-dominants in a given community. We consider such an interpretation useful because it is simple and conforms well to our intuitive understanding of dominants as species that have the major share of a community. This new view combines the mathematical advantages of Simpson's index, which include: (i) an objectiveness of measurement; (ii) a dependence on sample size; and (iii) the testability for statistical significance (e.g. by a randomization test). By estimating the dominants, this approach can distinguish abundant and rare species in a community; this property may be used for conservation purposes.

## ACKNOWLEDGEMENTS

We thank Michael Palmer, Bastow Wilson for their valuable comments.

## REFERENCES

- BROWN J. H. (1999) Macroecology: progress and prospect. *Oikos* **87**: 3–14.
- DAJOZ R. (1977) *Introduction to Ecology*. Hodder & Stoughton, London.
- FORTIN M.-J. & JACQUEZ G. M. (2000) Randomization tests and spatially autocorrelated data. *ESA Bulletin* **81**: 201–205.
- GREIG-SMITH P. (1983) *Quantitative Plant Ecology*. Blackwell Scientific Publications, Oxford.
- GRIME J. P., HODGESON J. G. & HUNT R. (1988) *Comparative Plant Ecology*. Unwin-Hyman, Boston.
- HARA M. (1991) Structure of a Japanese beech forest at Mt Kurikoma, north-eastern Japan. *Saito-Ho-on Kai Museum Research Bulletin* **59**: 43–55.
- HARA M., HIRATA K., FUJIHARA M., OONO K. & HSIEH CH-F. (1997) Floristic composition and stand structure of three evergreen broad-leaved forests in Taiwan, with special reference to the relationship between micro-landform and vegetation pattern. *Natural History Research* **4**: 81–112.
- HARA M., HIRATA K. & OONO K. (1996) Relation between macro-landform and vegetation structure in an evergreen broad-leaved forest on Okinawa Island, S-W Japan. *Natural History Research* **3**: 27–35.
- HILL M. O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**: 427–432.
- KEATING K. A. (1998) Estimating species richness:

- the Michaelis-Menten model revisited. *Oikos* 81: 411–416.
- KIKVIDZE Z. (2000) Modelling species richness and diversity in grassland communities of the Central Caucasus. *Oikos* 89: 123–127.
- LANDE R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.
- LANDE R., DE VRIES P. & WALLA T. (2000) When species accumulation curves intersect: implications for ranking diversity using small samples. *Oikos* 89: 601–605.
- MAGURRAN A. E. (1988) *Ecological Diversity and its Measurement*. Croom-Helm, London.
- MURRAY R. B., RICE B. L., KEITH D. A., MYERSCOUGH P. J., HOWELL J., FLOYD A. G., MILLS K. & WESTOBY M. (1999) Species in the tail of rank-abundance curves. *Ecology* 80: 1806–1816.
- OHSAWA M. (1984) Differentiation of vegetation zones and species strategies in the sub-alpine region of Mt Fuji. *Vegetatio* 57: 15–52.
- SIMPSON E. H. (1949) Measurement of diversity. *Nature* 163: 688.
- SLADE N. (1999) Randomization tests using standard statistical software. *ESA Bulletin* 80: 179–180.
- SMITH B. & WILSON J. B. (1996) A consumers's guide to evenness indices. *Oikos* 76: 70–82.
- TOKESHI M. (1993) Species abundance patterns and community structure. *Advances in Ecological Research* 24: 111–186.
- WILSON J. B. (1993) Would we recognise a Broken-Stick community if we found one? *Oikos* 67: 181–183.
- YEATES M. H. (1968) *An Introduction to Quantitative Analysis in Economic Geography*. McGraw-Hill, New York.