RESEARCH PAPER

Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association

Spyros Sfenthourakis,* Evangelos Tzanatos and Sinos Giokas

Section of Animal Biology, Department of Biology, University of Patras, University Campus, GR-26500 Patras, Greece

ABSTRACT

Aim To test whether congeneric species are significantly associated with one another in space, either positively or negatively. Also, to provide a framework for a causal investigation of co-occurrence patterns by a parallel comparison of interactions in geographical and ecological data matrices.

Location For the analysis of congeneric species' co-occurrences we used 30 matrices covering a wide range of taxa and geographical areas, while for the causal investigation we used the distribution of 50 terrestrial isopod species on 20 islands and 264 sampling stations in the central Aegean archipelago, as well as a number of ecological variables for each sampling station.

Methods We developed a software program (cooc) that incorporates the species-by-species approach to co-occurrence analysis using EcoSim's output of prior null model analysis of co-occurrence. We describe this program in detail, and use it to investigate one of the most common assembly rules, namely, the decreased levels of co-occurrence among congeneric species pairs. For the causal analysis, we proceed likewise, cross-checking the results from the geographical and the ecological matrices. There is only one possible combination of results that can support claims for direct competition among species.

Results We do not get any strong evidence for widespread competition among congeneric species, while most communities investigated do not show significant patterns of species associations. The causal analysis suggests that the principal factors behind terrestrial isopod species associations are of historical nature. Some exceptional cases are also discussed.

Main conclusions Presence/absence data for a variety of taxa do not support the assembly rule that congeneric species are under more intense competition compared to less related species. Also, these same data do not suggest strong interactions among species pairs, regardless of taxonomic status. When significant species associations can be seen in such matrices, they mainly reflect the effects of history or of habitat requirements.

Keywords

Assembly rules, competition, community structure, co-occurrence, macroecology, species association.

*Correspondence: S. Sfenthourakis, Section of Animal Biology, Department of Biology, University of Patras, University Campus, GR-26500 Patras, Greece. Tel./fax: +302610 969268; E-mail: sfendo@upatras.gr

INTRODUCTION

Patterns of species co-occurrence and exclusion have interested ecologists for more than fifty years (Elton, 1946; Cole, 1949; Williams, 1964; Terborgh & Weske, 1969; Pianka, 1973, 1994: 264; Schluter, 1984). In the 1970s, these patterns were discussed within the theoretical framework of the 'community assembly rules' (Diamond, 1975; Connor & Simberloff, 1979; Weiher & Keddy, 1999).

Like other patterns emerging from presence-absence matrices, significant co-occurrence is best detected by the use of a Monte Carlo simulation approach, because such matrices can easily

show 'structure' due to stochastic phenomena (Gotelli & Graves, 1996). Of course, the Monte Carlo approach is not immune to the effects of random or hidden structure [e.g., Narcissus effect (Colwell & Winkler, 1984)], so there has been considerable debate about which set of assumptions leads to the construction of the proper 'null space' (Manly, 1995; Sanderson *et al.*, 1998; Gotelli, 2000; Gotelli & Entsminger, 2001, 2003; Manly & Sanderson, 2002). Until recently, species co-occurrence patterns in whole communities were identified using one of the various 'ensemble' indices developed for this purpose, conveniently included in the free software EcoSim (Gotelli & Entsminger, 2004). Some years ago, Sanderson (2000) proposed a simple Monte Carlo approach for the detection of co-occurrence for every species pair in a presence/absence matrix. This approach, using what Sanderson (2004) called 'the natural metric', potentially identifies the species that most probably interact, positively or negatively.

Zaman & Simberloff (2002) criticized species by species approaches, on the grounds that occurrences of species pairs are not independent, making it difficult to assess significance levels. However, if a species *i* tends to co-occur with species *j* and *k*, this does not necessarily imply that species *j* and *k* also co-occur. Species *i* and *j* may co-occur in different sites than those where species *i* and *k* are present. The co-occurrence of species *j* and *k* must be tested independently. Of course, the occurrence of every species may be indirectly constrained by the occurrences of other species when the richness of each site is conditioned by the effect of area (or other 'independent' factors). Nevertheless, this is a general quality for such presence-absence matrices, and it affects all approaches but, at the same time, this is the very reason for the use of randomised matrices.

In previous work (Sfenthourakis *et al.*, 2004), we developed an EcoSim-accessory software routine (cooc) that follows the principles of Sanderson (2000, 2004). cooc is able to identify significantly positive (co-occurrence) or negative (exclusion) associations among all species pairs in a presence-absence matrix. Here, we explore this approach in more detail to investigate certain questions on co-occurrence and assembly rules, and we suggest a framework for the causal investigation of these patterns.

A central question in the 'assembly rules' discussion concerned co-occurrence between congeneric species. In principle, congeneric species should compete more intensely for resources and, therefore, they should exhibit higher levels of mutual exclusion compared to pairs of species belonging to different genera. This is because congeners are recent descendants of a common ancestor; therefore their current niche is the product of niche divergence and should overlap more compared with more distantly related species (depending also on the mode of speciation and the phylogenetic patterns in each genus). Further, congeneric species tend to be more similar (both morphologically and ecologically) to one another than to species belonging to other genera. However, empirical evidence supporting this argument has been weak (*cf.* Gotelli & McCabe, 2002), and precise methods for identifying co-occurring congeneric pairs have been applied on very few community matrices (e.g., Connor & Simberloff, 1979; Sanderson, 2000, 2004). In the present paper, we apply such a method to a

number of matrices including a variety of taxa, in an attempt to find evidence for, or against, this claim.

A second question we address herein concerns the underlying factors that may lead to significant patterns of co-occurrence or exclusion. We know from evolutionary biology and biogeography that species are not randomly distributed in space. This means that besides interactions among species (e.g., competition), the incidences of species are also affected by the distribution of appropriate habitats as well as historical factors. In most cases, the spatial distribution of habitats also shows some structure, and species exploiting similar habitats will reflect it. Thus, the causal basis of co-occurrence might be historical, biological (in the sense of direct interaction, such as symbiosis) and/or ecological. Similarly, 'exclusion' might be due to history, competition, or different habitat requirements (if the respective habitats do not co-occur). In order to test the effect of these three categories of factors, we compiled data both on species presence per sampling station, and on several ecological variables recorded at these same sampling stations. The combination of these data gives the 'ecological matrix': 'presence' of species per ecological factor, a matrix that can also be analysed for exploitation of similar habitats. Competing pairs of species should covary negatively (exclusion) in the geographical matrix, but positively (co-occurrence) in the ecological matrix (i.e., exploiting similar habitats). All other combinations of co-occurrence and/or exclusion among the geographical and the ecological matrices exclude possibilities for true competition: (a) when species covary positively in both matrices, either they exploit similar habitats, they share a common history, and/or their interaction is either neutral or symbiotic, (b) negative interaction in both matrices is inconsistent with the hypothesis of competitive exclusion, since the species do not exploit similar habitats anyway, (c) positive interaction in the geographical matrix and negative in the ecological (species with different habitat requirements but consistent geographical distribution) may be due to a common history. In addition to these combinations, there should also be pairs deviating in one of the two matrices only. In Table 1 we summarize the factors that would lead to each case. Of course, these are simplified interpretations, but such a simplification is sometimes necessary as a first step in causal investigations.

METHODS

Species by species co-occurrence analysis

To determine which species pairs exhibit higher or lower co-occurrence frequency than expected by chance (due to the matrix structure), we developed cooc, a software program written in C++, that works on the output files produced by the Co-occurrence routine of EcoSim (Gotelli & Entsminger, 2004). cooc follows the principles of Sanderson (2000, 2004) for a 'natural co-occurrence' index. First, cooc counts how many times each species in the original matrix occurs together with each other species. Next, it does the same for a user-defined number of simulated 'null' matrices (produced by EcoSim) and provides a frequency distribution of co-occurrences for each species pair. cooc then

Table 1 Plausible assumptions concerning causality of significant co-occurrence patterns, based on the simultaneous analysis of geographical and ecological data matrices at the species pair level. The geographical matrix contains species presence-absence data, and the ecological matrix contains binary data on species per ecological variable. Positive interaction means higher co-occurrence than expected by chance, negative means less co-occurrence than expected by chance, and none signifies the absence of deviating co-occurrence pattern. Here, allopatry is not necessarily meant in the strict sense, but denotes a tendency

| | Ecological matrix | | |
	Interaction **Positive**	**Negative**	**None**
Geographic matrix	**Positive** common ecology and history — neutral/symbiotic	common history	common history
	Negative competition — ecological 'allopatry'	irrelevance — historical 'allopatry'	historical 'allopatry'
	None common ecology and history — neutral	different ecology	random occurrence

compares the observed number of each species pair's co-occurrences in the original matrix to this distribution. If the number falls outside the distribution or within user-defined tails, cooc designates the respective species pair as significantly different from a random association. The software also produces two other output files, one with the observed co-occurrences of each species pair, and the second with the frequency distribution of each species pair co-occurrences in the null matrices, so that the user can work further on these data. For the present analysis, and in accordance to the principles presented by Sanderson *et al.* (1998), we used the constraint of fixed column and row sums for the construction of null matrices. However, in contrast with these authors, we chose the sequential swapping algorithm (see next section), following the rationale of Gotelli & Entsminger (2003). We obtain very similar results using the more time consuming knight's tour algorithm (see Sanderson, 2000; but *cf.* Gotelli & Entsminger, 2001). The desired number of output null matrices can be selected as an option in EcoSim; for the present analysis we used 5000 null matrices in each case. With the current version of cooc, in a common modern PC it is possible to process matrices with up to 200 species and 270 sampling units. The software is available from the authors upon request.

A note on statistical properties

The main procedure of cooc is simply a counting of species by species co-occurrences in various matrices (original and nulls). This means that the outcome of the procedure depends only on the quality of the randomization algorithm and the confidence limits defined by the user.

The properties of the randomization algorithm used in the present paper (sequential swapping) have been widely discussed (see references in the previous section), and it seems to be fairly dependable, even though it may produce a slightly biased null matrix space. In a recent paper, Miklós & Podani (2004) have presented a modified version of this algorithm, the 'trial swap', that is able to produce an unbiased null matrix space, and according to these authors, it performs better than the independent and sequential swapping algorithms of EcoSim. Still another approach is presented by Zaman & Simberloff (2002).

Nevertheless, the aim of our present work is not the evaluation of randomization algorithms. We want to present a method that can take advantage of any possible algorithm and use its output for further analysis. It is self-evident, though, that as soon as a more efficient algorithm becomes widely accepted and available, the user of cooc can choose to use this instead.

The probability that each species pair co-occurrence could be due to chance can be defined *ad hoc* for each species pair (e.g., see Connor & Simberloff, 1979), taking into account the incidence of each species and calculating the probabilities of common presence for the specific matrix. Such a procedure would demand the acceptance of a conventional significance level (usually 5%), and would be unrealistically slow for medium to large matrices. Therefore, the use of a conventional 5% (two-tailed) for the frequency distribution herein, assuming the proper null matrix space, is a reasonable choice for our procedure. The use of artificial matrices with 'known' structure is not instructive in our case: (a) in the extreme case of a perfect checkerboard distribution, the final outcome will depend on the unoccupied sites by the species of each pair, that is, on the possible presence-absence configurations in the randomized matrices, and (b) if we place 1s and 0s randomly, the resulting configuration will necessarily have structure (species pairs will, or will not, co-occur in some sites), and it is this very structure that we want to evaluate by the simulation and the use of cooc. No matrix is truly 'random' *with respect to the occurrence of each species pair*.

We should note, though, that the procedure followed herein is somewhat biased against positive associations (co-occurrences) among species, because the extreme co-occurrences (i.e., those involving co-occurrence in most, or in all of the sites in the matrix) are not recognized as significant. For instance, if two species co-occur in 9 out of a total of 10 sites included in the matrix, cooc will not count them as significantly co-occurring (since the possible alternative presence/absence configurations in the null space are very limited). Of course, for 'omni-present' species, there is only one possible configuration under the restriction of fixed row and column sums, so that such species will not appear in the output file of cooc. Partly for this reason, we focus mainly on negative associations (mutual exclusions), which are more interesting as far as congeneric species are concerned.

Overall deviations from random patterns

To test whether species pairs co-occur more (or less) often than expected by chance, we analysed 30 published presence-absence matrices. Many matrices used in similar studies (e.g., the 96 matrices used by Gotelli & McCabe, 2002), and available on the web (e.g., Atmar & Patterson, 1995), do not contain species names, so are not useful in our analysis herein.

In accordance with Sanderson (2004), we can assume that up to 5% of all possible species pairs in a matrix could deviate (co-occur more or less often than expected) by chance. Since the 5% threshold is arbitrary, we also checked the possibility of 10% chance deviations. In cases where species richness is low enough so that the possible species combinations are restricted, this higher threshold might be more appropriate, as it leads to a more strict evaluation of co-occurrence (only matrices where more than 10% of the possible species pairs deviate would be considered). In any case, the threshold for random expectation chosen does not affect the main tendencies, as we can always check the relative contribution of positive (co-occurrence) and negative (exclusion) species associations in each case (taking into account the afore-mentioned bias against positive associations).

The matrices we used cover a wide variety of taxonomic groups, from birds and mammals to land snails and terrestrial isopods, and a relatively wide geographical range. Obviously, the designation of genera may differ among taxonomic groups, an issue that might complicate the comparative evaluation of results. Nevertheless, in any taxon, congeners should be more similar to one another than with species assigned to different genera, so the expectation of increased intrageneric competition still holds. One matrix contains data for a single genus (the beetle genus *Hegeter* on the Canary islands), so all expected pairs should be congeneric. We used this matrix to check the signs (positive — negative) of the probable deviations. Table 2 lists the data matrices we used, with references and information regarding their size, the taxonomic group involved and their geographical origin, as well as a summary of the co-occurrence analysis results.

For comparative reasons, we also checked the C-score (Stone & Roberts, 1990), one of the 'ensemble' indices, for the same matrices. The C-score is based on the number of 'checkerboard units' among species pairs (cases where the one species is present whereas the other is absent and vice-versa). Checkerboard units (CU) are calculated by the formula:

$$CU = (r_i - S)(r_j - S)$$

where S is the number of shared sites (sites containing both species) and r_i and r_j are the row totals for species i and j. The C-score is the average of all possible checkerboard pairs, calculated for species that occur at least once in the matrix. In a competitively structured community the C-score should be larger than expected by chance. Significance levels are provided via Monte Carlo simulation (see Gotelli & Entsminger, 2004). The results are given in Table 3.

Co-occurrence of congeneric species

In the same 30 datasets we checked for possible unexpected deviations of congeneric species. If the total number of deviating congeneric pairs was significantly lower or higher than the number expected by chance (i.e., according to the actual percentage of congeneric pairs in the original matrix), we considered the congeners as actually deviating. Significant deviations from expected proportions were tested using the χ^2 test. The results are presented in Table 2.

Causal approach to species associations

For the analysis of co-occurrence causality, as presented in Introduction, we used data on terrestrial isopods of Aegean islands (Sfenthourakis, 1994, 1996a,b; Sfenthourakis *et al.*, 2004) for which we have detailed collecting data at the sampling station level. More particularly, for each sampling station we used presence-absence data of a number of ecological parameters related to habitat preferences of terrestrial isopods. The species per ecological variable matrix is produced by the combination of this matrix with that of species per sampling station. The total list of ecological variables used (41) is presented in Table 4. The complete matrix contains a total of 264 sampling stations on 20 islands, and 50 species. Since the number of variables may affect the results, we also used reduced matrices of ecological variables. We produced two different reduced matrices. In the first, biologically similar variables (e.g., similar types of vegetation) were merged, a process that left 31 variables. In order to produce the second reduced matrix, we applied successive cluster analyses (Jaccard's index, UPGMA) on the variables matrix and in each step we merged variables that were grouped at a similarity level ≥ 90%; we then ran the analysis again, until all variables were grouped at a level < 90%. The reduced matrix produced contained only 12 variables. This procedure increases the discrepancies among species, since variables with similar 'occupancies' were merged. Nevertheless, it is biologically meaningful because it reflects more strictly the ecological range exploited by each species (by merging interrelated variables).

RESULTS

Overall deviations of species pairs

Deviating species pairs have been found in 29 out of the 30 matrices (Table 2). In total, negative relations (exclusions) prevail (469 exclusions versus 380 co-occurrences), but the ratio of co-occurrences to exclusions is strongly in favour of exclusions in only 14 of the 30 matrices. Under the assumption of 5% chance deviations, only 8 of the matrices show a significant number of deviations (> 5%), while just 3 matrices pass through the more strict 10% filter. These are the rodents on sand dunes at the Mojave and Great Basin Deserts (38%), terrestrial isopods of Aegean islands when listed by sampling station (22%), and Darwin's finches on 17 of the Galapagos islands (13%). The mammals in North American forests (9%), land snails on small

Table 2 The databases used for the analysis of co-occurrence patterns among congeneric species. The results of the analysis are given as numbers of deviating total and congeneric pairs (n.a. = non applicable). Asterisks mark those cases where the observed number of deviating congeneric pairs differs from expected proportions (χ^2, $P < 0.05$). The expected number of congeneric pairs for the beetles of the genus *Hegeter* coincides with the expected 5%, since all species in the matrix belong to the same genus

Reference	Taxonomic group	Geographic origin	Species × sites	Species pairs	Congeneric pairs	Deviating pairs (no.+, no.−)	Deviating congeneric pairs (expected by chance)	Mode of interaction (no.+, no.−)
Sfenthourakis 1994, 1996b	Terrestrial isopods	Central Aegean islands, Greece (sampling stations)	50 × 264	1225	50	265 (153+, 112−)	11 (11)	2+, 9−
Sfenthourakis 1994, 1996b	Terrestrial isopods	Central Aegean islands, Greece (islands)	50 × 20	1225	50	45 (21+, 24−)	2 (2)	2−
Argano and Manicastri 1995	Terrestrial isopods	Circum-Sardinian islands	51 × 29	1275	61	39 (16+, 23−)	2 (2)	1+, 1−
Rodriguez 1991	Terrestrial isopods	Canary islands	60 × 11	1770	268	27 (0+, 27−)	3 (4)	3−
Taiti and Howarth 1996	Terrestrial isopods	Hawaii	47 × 16	1081	27	8 (4+, 4−)	0 (0)	n.a.
Willemse 1984	Orthoptera	Central Aegean islands, Greece	40 × 22	780	6	22 (14+, 8−)	0 (2)	n.a.
Mühle *et al.* 2000	Coleoptera: Buprestidae	Kyklades islands, Greece	27 × 10	351	47	3 (0+, 3−)	0 (0)	n.a.
Mühle *et al.* 2000	Coleoptera: Buprestidae	N. Sporades islands, Greece	44 × 13	946	93	8 (3+, 5−)	1 (1)	1+
Garcia-Barros 2004 (web-site)	Coleoptera: Tenebrionidae (genus *Hegeter*)	Canary islands	16 × 7	120	120	0	0 (6)*	n.a.
Dennis *et al.* 2001	Lepidoptera	Central Aegean islands, Greece	34 × 7	561	3	1 (0+, 1−)	0 (0)	n.a.
Gathorne-Hardy *et al.*, 2000	Termites	Krakatau	19 × 3	171	12	1 (1+, 0−)	0 (0)	n.a.
Zarfdjian and Economidis 1989	Rotifera	North Greece, lakes	41 × 12	820	30	8 (2+, 6−)	0 (0)	n.a.
Zarfdjian and Economidis 1989	Cladocera	North Greece, lakes	47 × 12	1081	51	9 (5+, 4−)	0 (0)	n.a.
Botsaris 1996	Land snails	Islets of Saronikos Gulf, Greece	58 × 77	1653	27	140 (67+, 73−)	3 (2)	1+, 2−
Riedel 1992	Land snails (Zonitidae)	Central Aegean islands, Greece	13 × 27	78	11	2 (1+, 1−)	0 (0)	n.a.
Wiktor 2001	Slugs	Central Aegean islands, Greece	17 × 19	136	59	5 (3+, 2−)	1 (2)	1+
Cameron *et al.* 1996	Land snails	Porto Santo islands, Madeira	56 × 6	1540	124	8 (0+, 8−)	1 (1)	1−
McLain and Pratt 1999	Fish	Caribbean islands (fringing reefs)	41 × 18	820	57	12 (4+, 8−)	2 (1)	2+
Chondropoulos 1986, 1989 Gruber 1986 Foufopoulos and Ives 1999	Reptiles	Central Aegean islands, Greece	28 × 34	378	10	21 (12+, 9−)	2 (1)	1+, 1−
Clark 1989	Reptiles	Islands of Saronikos Gulf, Greece	21 × 19	210	6	8 (1+, 7−)	0 (0)	n.a.
Sanderson 2000	Birds (Darwin's finches)	Galapagos islands	13 × 17	78	19	10 (2+, 8−)	7 (2)*	2+, 5−
Magioris 1994	Birds (nesting)	Central Aegean islands, Greece	54 × 41	1431	31	109 (32+, 77−)	0 (2)	n.a.
Fernandez-Juricic 2002	Birds	Spain (parks)	29 × 26	406	9	8 (1+, 7−)	0 (0)	n.a.
Kelt and Brown 1999	Mammals	Mojave and Great Basin deserts (sand dunes)	12 × 13	66	9	25 (25+, 0−)	6 (3)	6+
Conroy *et al.* 1999	Mammals	Alaska	23 × 24	253	7	16 (2+, 14−)	1 (0)	1−
Heaney 1986	Mammals	Luzon, Mindoro, etc., Philippines	47 × 8	1081	55	4 (0+, 4−)	1 (0)	1−
Heaney 1986	Mammals	Mindanao, Philippines	35 × 9	595	22	4 (0+, 4−)	0 (0)	n.a.
Lomolino and Davis 1997	Mammals	N. American forests	23 × 23	253	18	23 (8+, 15−)	2 (2)	1+, 1−
Millien-Parra and Jaeger 1999	Mammals	Japan	55 × 12	1485	40	40 (16+, 24−)	1 (1)	1−
Crowell 1986	Mammals	Islands of Maine	38 × 9	703	7	17 (3+, 14−)	0 (0)	n.a.

Table 3 Matrices with significant C-score values, and their respective *P* value. All these matrices are competitively structured (i.e., observed value is significantly larger than expected by chance)

	obs	sim	P (obs ≥ sim)
Terrestrial isopods, central Aegean islands, Greece (sampling stations)	586.69	577.58	< 0.001
Terrestrial isopods, central Aegean islands, Greece (islands)	9.67	9.39	< 0.001
Terrestrial isopods, Canary islands	1.53	1.39	0.0002
Terrestrial isopods, circum-Sardinian islands	11.9	11.6	0.005
Buprestidae, N. Sporades islands, Greece	0.98	0.86	0.02
Land snails, Saronikos Gulf, Greece	40.78	39.55	0.006
Slugs, central Aegean islands, Greece	7.15	6.82	0.005
Reptiles, central Aegean islands, Greece	15.75	15.05	< 0.001
Reptiles, Saronikos Gulf, Greece	2.75	2.52	0.006
Birds, Galapagos islands	4.27	3.16	< 0.001
Nesting birds, Aegean islands, Greece	14.73	13.98	< 0.001
Mammals, Alaska	5.66	5.20	0.008
Mammals, Luzon, Mindoro *et al.*, Philippines	0.68	0.59	0.03
Mammals, N. American forests	6.79	6.33	0.02
Mammals, Japan	2.1	1.9	< 0.001
Mammals, Island of Maine	0.60	0.55	0.05

obs: observed value of C-score.
sim: mean value of C-score in 5000 simulated matrices.

islets of Saronikos Gulf, Greece (8%), nesting birds on Aegean islands (8%), Alaskan mammals (6%), and reptiles on Aegean islands (5.5%) complete the list to the 5% threshold. It is noteworthy that the mammals (rodents) of the sand dunes at the Mojave and Great Basin Deserts show exclusively co-occurrences. Also, Aegean terrestrial isopods show increased levels of deviations when listed by sampling station, whereas the species per island matrix does not give more deviations than expected by chance. A more detailed discussion of the results concerning this matrix can be found in Sfenthourakis *et al.* (2004).

All significant C-scores (16 matrices) suggest competitively structured matrices (the observed C-score value is higher than the mean value of the simulated matrices). The deviating matrices include 7 of the matrices mentioned above, while the matrix with the larger number of deviating pairs (rodents of Mojave and Great Basin Deserts) is not included. This must be related to the fact that this matrix is dominated by co-occurring species, while the C-score index checks for 'checkerboard' distributions (mutually excluded species).

There is no apparent correlation between negative species associations identified by our procedure and C-score results in the remaining 10 deviating matrices. For instance, a matrix with 5 negative associations in a total of 946 species pairs (Buprestidae in northern Sporades islands) gives a significant C-score, while one with 8 negative associations out of 780 pairs (Orthoptera of central Aegean) does not.

Co-occurrence of congeneric species

Of the 29 datasets with significantly deviating pairs, 16 involve congeneric species. In the vast majority of cases, the numbers of deviating congeneric pairs do not differ from those expected

by chance (χ^2 test, $P < 0.05$, see Table 2). Pronounced differences are found only in 3 out of the 29 cases, namely, Darwin's finches in Galapagos, rodents at sand dunes in the Mojave and Great Basin Deserts, and the beetle genus *Hegeter* on the Canary islands.

In total, we could not find any clear evidence for increased prevalence of negative relationships among congeneric species, since a manifest pattern of excessive exclusions, highly suggestive of competition, can be seen only for Darwin's finches on the Galapagos islands. Competition among these birds has been documented also by long-standing observations regarding character displacement and adaptive radiations, studies of paradigmatic nature in evolutionary biology (Lack, 1947; Grant, 1986). In Table 5 we present in detail the results from this dataset, with information on the occurrence patterns of the 'deviating' species.

Geographic versus ecological co-occurrence

The analysis of the original species per ecological variable matrix gave 59 deviating species pairs out of a total of 1225 possible pairs, a number less than 5%, while the species per sampling station matrix gave 265 such pairs (22%). Only 14 of these pairs were common between the two matrices (see Table 6), a number not different from what we would expect by chance (13). Of these, eleven co-occurred more than expected by chance in both matrices (common ecology and history, neutral/symbiotic interaction), two co-occurred more in the ecological matrix but less in the geographical one (competition), and one co-occurred less in the ecological and more in the geographical matrix (common history).

The analysis of the species per island matrix resulted in 45 deviating pairs (< 5%), 23 of which were also present (and,

Table 4 The list of 41 ecological variables registered in each sampling station as present/absent, and then used for the construction of the ecological (species per variable) matrix

1.   Altitude 0–300 m a.s.l.
2.   Altitude 300–600 m a.s.l.
3.   Altitude 600–900 m a.s.l.
4.   Maquis vegetation (evergreen sclerophyllus scrubs — *Quercus coccifera*, *Q. ilex*, *Pistacia lentiscus*, *Juniperus* spp.)
5.   Broadleaved vegetation
6.   Coniferous vegetation
7.   Vegetation consisting of *Arbutus* spp. or/and *Erica* spp.
8.   Other arboreal vegetation (*Ceratonia siliqua*, *Crataegus* spp. etc.)
9.   Phrygana vegetation (*Sarcoporium spinosum*, *Cistus* spp., *Genista* spp. etc.)
10. *Asphodelus* spp. formations
11. Graminae
12. Halophilous plants
13. Helophilous plants
14. Olive groves
15. Fruit-bearing plantations
16. Cultivated herbaceous plants
17. Riparian vegetation with *Platanus orientalis*
18. Riparian vegetation with *Acer* spp.
19. Other riparian vegetation (e.g., *Arundo donax*, *Nerium oleander*, *Myrtus communis*, *Vitex agnus-castus*, etc.)
20. Dry riverbanks
21. Freshwater
22. Sea-spray effects
23. Rich in shelters (stones, logs, etc.)
24. Sparse shelters
25. Very few shelters
26. Stonewalls
27. Buildings
28. Obvious effects of recent fire
29. Roads
30. Calcareous substrate
31. Mixed calcareous and non-calcareous substrate
32. Acidic substrate
33. Sandy and/or soft soil
34. Soil of intermediate softness
35. Muddy soil
36. Hard soil
37. Rich leaf-litter layer
38. Sparse leaf-litter layer
39. Very humid soil
40. Intermediate soil humidity
41. Hardly humid soil

expectedly, with the same type of interaction) in the results of the species per sampling station matrix analysis. Only 2 of the pairs appeared also in the ecological matrix, exactly as many as we would expect by chance. One of them (*Ligidium ghigii* × *Armadillidium ameglioi*) co-occurred in the geographical matrix, but was negatively related in the ecological (indicating common history), while the other (*Bathytropa granulata* × *Proporcellio quadriseriatus*) gave the opposite combination, indicating

competition. We should note that the latter pair gave the same results also for the species per sampling station matrix.

The first reduced ecological matrix (31 variables) gave 31 deviating pairs (< 5%), 11 negatively and 20 positively associated. Of these, 17 were identical with those of the complete ecological matrix. Among the remaining 14 there were 5 more cases of common positive associations with the species per sampling station matrix. All common pairs were 'double positives' (common ecology and history, neutral/symbiotic interaction). No common pair was found with the species per island matrix.

Finally, the second reduced ecological matrix (12 variables) gave 21 deviating pairs (< 5%), 16 positively and 5 negatively associated. Only 4 of these pairs do not appear in the results of the complete ecological matrix, adding just one case of common co-occurrence with the species per sampling station matrix. In these results the second case of putative competition identified in the comparison of the complete ecological variable matrix with the species per island matrix appeared again (co-occurrence in the former versus competition in the latter, for *Platyarthrus lindbergi* × *Armadillidium vulgare*). Again, no common pair was found with the species per island matrix.

Overall, then, there are 2 cases of possible competition, 17 cases indicating common ecology and history, but neutral/symbiotic interaction, while all other cases of deviation in one of the matrices suggest either historical or ecological effects on patterns of coexistence, according to the combinations in Table 1.

DISCUSSION

The generality of common competitive exclusion among congeneric species, relative to non-congeners, was not supported by the analysis of 30 presence-absence matrices undertaken here. Few congeneric pairs of species showed any pattern at all. The same holds also for all species pairs, even though there is some discrepancy with the results produced by the application of the 'ensemble' index, C-score. According to this index, nearly half of the matrices are competitively structured. The reasons for this discrepancy could be related to the different patterns examined by the two procedures, since our algorithm counts coupled presences, while C-score counts 'checkerboard units', that is, cases of mutual exclusion in 2×2 units, no matter what the total possible configurations of the respective species occurrences might be.

According to our analysis, most communities checked do not exhibit significant species by species interactions. According to the results of the causal analysis involving the cross-check of distribution and ecological datasets, most cases of deviation can be attributed to the effects of common history and/or habitat requirements. Of course, the taxonomically biased approach usually taken in such analyses (since data sets are formed on a taxonomic basis) may obscure species interactions with members of other taxonomic groups that share similar resources. This could be an interesting direction for future research, where the method we propose herein can play some role.

It is useful now to examine more closely the few cases where deviations from randomness were found.

Table 5 The results of the Darwin's finches dataset (from Sanderson, 2000) analysis. Ten deviating species pairs have been found (out of 78 total species pairs), 2 of which co-occur more often (co-occurrence) and 8 less often (competition) than expected by chance. For each species pair, total number of co-occurrences, as well as occurrences of each member species, are also given (in a total of 17 islands). It should be noted that in this dataset all islands host at least one congeneric pair and that most species (9 out of 13) occur in 10 or more islands. One species, *Camarhynchus pauper*, occurs on one island only, two more (*Geospiza conirostris* and *Cactospiza heliobates*) on two islands, and one species (*Cactospiza pallida*) on 6 islands. Of these species, only *G. conirostris* participates in deviating pairs, and this happens with species that occur on many islands but never on the same island with it

		occurrence of species in each pair		
species pair	association	co-occurrences	1st	2nd
Geospiza fortis × *Geospiza fuliginosa*	co-occurrence	13	13	14
Geospiza fortis × *Geospiza difficilis*	competition	7	13	10
Geospiza fortis × *Geospiza scandens*	co-occurrence	12	13	12
Geospiza fortis × *Geospiza conirostris*	competition	0	13	2
Geospiza fuliginosa × *Geospiza difficilis*	competition	7	14	10
Geospiza difficilis × *Geospiza scandens*	competition	6	10	12
Geospiza scandens × *Geospiza conirostris*	competition	0	12	2
Geospiza conirostris × *Camarhynchus psittacula*	competition	0	2	10
Geospiza conirostris × *Camarhynchus parvulus*	competition	0	2	10
Geospiza conirostris × *Platyspiza crassirostris*	competition	0	2	11

Table 6 Terrestrial isopod species pairs whose co-occurrence frequency deviates from that expected by chance in both matrices (species per sampling station and species per ecological variable), and the presumed causal explanation of the observed combination of interactions. The two pairs that are suspect for direct competition are given in boldface

	Type of interaction in each matrix		
Species pairs that deviate from random in both matrices	geographic	ecological	Causal explanation
Ligidium ghigii × *Paraschizidium album*	positive	negative	common history
Cretoniscellus strinatii × *Trichodillidium malickyi*	positive	positive	common ecology/history
Bathytropa granulata × *Rodoniscus anophthalmus*	positive	positive	common ecology/history
Bathytropa granulata × *Leptotrichus kosswigi*	positive	positive	common ecology/history
Bathytropa granulata × *Porcellio laevis*	positive	positive	common ecology/history
Bathytropa granulata* × *Proporcellio quadriseriatus	**negative**	**positive**	**competition**
Bathytropa granulata × *Protracheoniscus babori*	positive	positive	common ecology/history
Bathytropa granulata × *Schizidium hybridum*	positive	positive	common ecology/history
Chaetophiloscia elongata × *Armadillidium marmoratum*	positive	positive	common ecology/history
Chaetophiloscia elongata × *Armadillidium vulgare*	positive	positive	common ecology/history
Platyarthrus lindbergi* × *Armadillidium vulgare	**negative**	**positive**	**competition**
Porcellio werneri × *Armadillidium ameglioi*	positive	positive	common ecology/history
Trachelipus aegaeus × *Armadillidium vulgare*	positive	positive	common ecology/history
Paraschizidium aegaeum × *Schizidium tinum*	positive	positive	common ecology/history

The Darwin's finches matrix gave 10 deviating pairs (the sequential swapping and the knight's tour algorithm both gave identical results), in contrast to the 14 pairs mentioned by Sanderson (2000) (13 are mentioned in the text, but 14 are clearly shown in his Fig. 5, p. 336). Our algorithm 'failed' to identify 4 cases of non-congeneric exclusion (*Geospiza difficilis* × *Camarhynchus psittacula*, *G. difficilis* × *Platyspiza crassirostris*, *G. difficilis* × *Certhidea olivacea*, and *G. conirostris* × *Cactospiza pallida*) found by Sanderson. Since this author does not explicitly describe the software he used, we cannot evaluate further these differences. Nevertheless, we should note that one of the negatively related pairs recognized by Sanderson, but not by us, involves a species (*C. olivacea*) that is present on all of the islands, something that raises further questions about what 'mutual exclusion' may mean in this context. We would expect that, given the fixed column and row sums assumption, any species would appear together with an 'omni-present' one exactly the same number of times in all simulated matrices, therefore it should not appear as deviating.

The rodents at sand dunes of Mojave and Great Basin Deserts show increased co-occurrence, even among congeneric species. This might be a result of the highly nested structure of the

respective matrix. In fact, increased nestedness means an increased frequency of co-occurring species pairs, but this does not always lead to significant results. The details of the interplay between nestedness and (significant) co-occurrence is another interesting issue for future investigation. As Kelt & Brown (1999) suggest, factors other than simple taxonomic relatedness, such as body size or functional relationships, might be important in shaping co-occurrence patterns. Again, the use of COOC can help towards pattern identification under any similar assumption, as long as the respective data are available for each species. At sites with extremely uniform ecological conditions, such as desert sand dunes, the Narcissus effect may be responsible for the shaping of observed patterns, since only those species able to live in the respective habitats are present in each community. The causal approach taken in the present work, as exemplified by Aegean terrestrial isopods, may offer an insight to the exact nature of species associations.

The case of the beetle genus *Hegeter* on Canary islands is strikingly different, since there are no significantly associated species pairs, while 6 were to be expected by chance. It is plausible to assume that the distribution of these species is controlled by other factors (e.g., special habitat requirements) in a way that pure geographical information is inadequate to identify patterns at the species by species level. It would be interesting to check a broader dataset including species of other related genera plus ecological factors, but such information was not available to us.

The fact that Aegean terrestrial isopods exhibit important patterns of non-congeneric species associations when analysed at the sampling station level, calls for an explanation involving ecological factors. Nevertheless, for the vast majority of pairs, history appears to be the principal causal factor because most of theses cases fall into the last column and first two rows of Table 1 ('none' ecological — 'positive' or 'negative' geographical association). Only two cases of apparent competition are supported, a result not surprising for animals like terrestrial isopods, whose biology does not offer many opportunities for direct competition.

Finally, we should note that the causal approach presented in this paper is not intended to definitively resolve the problem of identifying competition, but represents a plausible second step of analysis after the documentation of a significant pattern concerning species associations. The detection of real patterns is a prerequisite for any further attempt towards causal interpretation, but to make good use of its results, it is necessary to have some plausible set of assumptions concerning the causality behind these patterns. We believe that the framework summarized in Table 1 can provide a useful tool, complementary to the 'favoured states' approach followed by Kelt & Brown (1999), for the analysis of biotic communities assembly.

ACKNOWLEDGEMENTS

REFERENCES

Argano, R. & Manicastri, C. (1995) Gli isopodi terrestri delle piccole isole circumsarde (Crustacea, Oniscidea). *Biogeographia*, **18**, 283–298.

Atmar, J.W. & Patterson, B.D. (1995) *The nestedness temperature calculator: a visual basic program, including 294 presence-absence matrices.* AICS Research, Inc., University Park, NM, and The Field Museum, Chicago.

Botsaris, I. (1996) *A study on the biogeography of land snails of Saronikos Gulf.* PhD Thesis, University of Athens, Greece (in Greek, English abstract).

Cameron, R.A.D., Cook, L.M. & Hallows, J.D. (1996) Land snails on Porto Santo — adaptive and non-adaptive radiation. *Proceedings of the Royal Society of London Series B*, **351**, 309–327.

Chondropoulos, B.P. (1986) A checklist of Greek reptiles. I. The lizards. *Amphibia-Reptilia*, **7**, 217–235.

Chondropoulos, B.P. (1989) A checklist of the Greek reptiles. II. The snakes. *Herpetozoa*, **2**, 3–36.

Clark, R. (1989) A check list of the herpetofauna of the Argo-Saronic Gulf district, Greece. *British Herpetological Society Bulletin*, **28**, 8–24.

Cole, L.C. (1949) The measurement of interspecific association. *Ecology*, **30**, 411–424.

Colwell, R.K. & Winkler, D.W. (1984) A null model for null models in biogeography. *Ecological communities: conceptual issues and the evidence* (ed. by D.R. Strong Jr., D. Simberloff, L.G. Abele and A.B. Thistle), pp. 344–359. Princeton University Press, Princeton.

Connor, E.F. & Simberloff, D. (1979) The assembly of species communities: chance or competition? *Ecology*, **60**, 1132–1140.

Conroy, C.J., Demboski, J.R. & Cook, J.A. (1999) Mammalian biogeography of the Alexander Archipelago of Alaska: a north temperate nested fauna. *Journal of Biogeography*, **26**, 343–352.

Crowell, K.L. (1986) A comparison of relict versus equilibrium models for insular mammals of the Gulf of Maine. *Island Biogeography of Mammals* (ed. by L.R. Heaney and B.D. Patterson). *Biological Journal of the Linnean Society*, **28**, 37–64.

Dennis, R.L.H., Shreeve, T.G., Olivier, A. & Coutsis, J.G. (2001) Contemporary geography dominates butterfly diversity gradients within Aegean archipelago (Lepidoptera: Papilionoidea, Hesperioidea). *Journal of Biogeography*, **27**, 1365–1383.

Diamond, J.M. (1975) Assembly of species communities. *Ecology and evolution of communities* (ed. by M.L. Cody and J.M. Diamond), pp. 342–444. Harvard University Press, Cambridge.

Elton, C. (1946) Competition and the structure of ecological communities. *Journal of Animal Ecology*, **15**, 54–68.

Fernandez-Juricic, E. (2002) Can human disturbance promote nestedness? A case study with breeding birds in urban habitat fragments. *Oecologia*, **131**, 269–278.

Foufopoulos, J. & Ives, A.R. (1999) Reptile extinctions on land-bridge islands: Life-history attributes and vulnerability to extinction. *American Naturalist*, **153**, 1–25.

Garcia-Barros, E. (2004) *Biogeography and evolution. Exercises on the Canary islands.* http://www.uam.es/personal_pdi/ciencias/egb/canarias/canaries.htm.

Gathorne-Hardy, F.J., Jones, D.T. & Mawdsley, N.A. (2000) The recolonization of the Krakatau islands by termites (Isoptera), and their biogeographical origins. *Biological Journal of the Linnean Society*, **71**, 251–267.

Gotelli, N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, **81**, 2606–2621.

Gotelli, N.J. & Entsminger, G.L. (2001) Swap and fill algorithms in null model analysis: rethinking the knight's tour. *Oecologia*, **129**, 281–291.

Gotelli, N.J. & Entsminger, G.L. (2003) Swap algorithms in null model analysis. *Ecology*, **84** (2), 532–535.

Gotelli, N.J. & Entsminger, G.L. (2004) *EcoSim: Null models software for ecology, Version 7.* Acquired Intelligence Inc. and Kesey-Bear. Jericho, VT 05465. http://garyentsminger.com/ecosim.htm.

Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology.* Smithsonian Institution Press, Washington, D.C.

Gotelli, N.J. & McCabe, D.J. (2002) Species co-occurrence: a meta-analysis of J.M. Diamond's assembly rules model. *Ecology*, **83**, 2091–2096.

Grant, P.R. (1986) *Ecology and evolution of Darwin's finches.* Princeton University Press, Princeton.

Gruber, U. (1986) *Podarcis erhardii* (Bedriaga, 1876) — Ägäische Mauereidechse. *Handbuch der Reptilien und Amphibien Europas*, *Vol. 2/II, Echsen 3* (ed. by W. Böhme), pp. 25–49. Aula, Wiesbaden.

Heaney, L.R. (1986) Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. *Island biogeography of mammals* (ed. by L.R. Heaney and B.D. Patterson). *Biological Journal of the Linnean Society*, **28**, 127–165.

Kelt, D.A. & Brown, J.H. (1999) Community structure and assembly rules: confronting conceptual and statistical issues with data on desert rodents. *Ecological assembly rules: perspectives, advances, retreat* (ed. by E. Weiher and P. Keddy), pp. 75–107. Cambridge University Press, Cambridge.

Lack, D. (1947) *Darwin's finches.* Cambridge University Press, Cambridge.

Lomolino, M.V. & Davis, R. (1997) Biogeographic scale and biodiversity of mountain forest mammals of western North America. *Global Ecology and Biogeography Letters*, **6**, 57–76.

Magioris, S. (1994) The avifauna of the Cyclades (Aegean Sea). *Hellenic Zoological Archive*, **2**, 1–16.

Manly, B.F.J. (1995) A note on the analysis of species co-occurrences. *Ecology*, **76**, 1109–1115.

Manly, B.F.J. & Sanderson, J.G. (2002) A note on null models: justifying the methodology. *Ecology*, **83**, 580–582.

McLain, D.K. & Pratt, A.E. (1999) Nestedness of coral reef fish across a set of fringing reefs. *Oikos*, **85**, 53–67.

Miklós, I. & Podani, J. (2004) Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, **85** (1), 86–92.

Millien-Parra, V. & Jaeger, J.-J. (1999) Island biogeography of the Japanese terrestrial mammal assemblages: an example of a relict fauna. *Journal of Biogeography*, **26**, 959–972.

Mühle, H., Brandl, P. & Niehuis, M. (2000) *Catalogus Faunae Graeciae: Coleoptera, Buprestidae.* Ed. Röble, Augsburg.

Pianka, E.R. (1973) The structure of lizard communities. *Annual Review of Ecology and Systematics*, **4**, 53–74.

Pianka, E.R. (1994) *Evolutionary ecology,* 5th edition. Harper-Collins College Publishers, N.Y.

Riedel, A. (1992) *The Zonitidae (sensu lato) (Gastropoda, Pulmonata) of Greece.* Fauna Graeciae Vol. V, Hellenic Zoological Society, Athens, Greece.

Rodriguez, R.S. (1991) *Estudio taxonómico y faunístico de los isópodos terrestres del Archipélago Canario.* PhD thesis, Univ. Autonoma de Barcelona, Spain (in Spanish, English summary).

Sanderson, J.G. (2000) Testing ecological patterns. *American Scientist*, **88**, 332–339.

Sanderson, J.G. (2004) Null model analysis of communities on gradients. *Journal of Biogeography*, **31**, 879–883.

Sanderson, J.G., Moulton, M.P. & Selfridge, R.G. (1998) Null matrices and the analysis of species co-occurences. *Oecologia*, **116**, 275–283.

Schluter, D. (1984) A variance test for detecting species associations with some example applications. *Ecology*, **65**, 998–1005.

Sfenthourakis, S. (1994) *Biogeography, systematics and ecological aspects of terrestrial isopods in central Aegean islands.* Ph.D. thesis, University of Athens, Greece (in Greek, English abstract).

Sfenthourakis, S. (1996a) The species-area relationship of terrestrial isopods (Isopoda: Oniscidea) from the Aegean archipelago (Greece): a comparative study. *Global Ecology and Biogeography Letters*, **5**, 149–157.

Sfenthourakis, S. (1996b) A biogeographic analysis of terrestrial isopods (Isopoda, Oniscidea) from central Aegean islands (Greece). *Journal of Biogeography*, **23**, 687–698.

Sfenthourakis, S., Giokas, S. & Tzanatos, E. (2004) From sampling stations to archipelagos: investigating aspects of the assemblage of insular biota. *Global Ecology and Biogeography*, **13**, 23–35.

Stone, L. & Roberts, A. (1990) The checkerboard score and species distributions. *Oecologia*, **85**, 74–79.

Taiti, S. & Howarth, F.G. (1996) Terrestrial isopods from the Hawaiian Islands (Isopoda: Oniscidea). *Bishop Museum Occasional Papers*, No. 45, 59–71.

Terborgh, J. & Weske, J.S. (1969) Colonization of secondary habitats by Peruvian birds. *Ecology*, **50**, 765–782.

Weiher, E. & Keddy, P. (1999) Assembly rules as general constraints on community composition. *Ecological assembly rules: perspectives, advances, retreat* (ed. by E. Weiher and P. Keddy), pp. 251–271. Cambridge University Press, Cambridge.

Wiktor, A. (2001) *The slugs of Greece (Arionidae, Milacidae, Limacidae, Agriolimacidae — Gastropoda, Stylomatophora).* Fauna Graeciae Vol. VIII, Natural History Museum of Crete — Hellenic Zoological Society, Irakleio, Greece.

Willemse, F. (1984) *Catalogue of the Orthoptera of Greece.* Fauna Graeciae Vol. I, Hellenic Zoological Society, Athens, Greece.

Williams, C.B. (1964) *Patterns in the balance of nature.* Academic Press, N.Y.

Zaman, A. & Simberloff, D. (2002) Random binary matrices in biogeographical ecology — Instituting a good neighbor policy. *Environmental and Ecological Statistics*, **9**, 405–421.

Zarfdjian, M.H. & Economidis, P.S. (1989) Listes provisoires des rotifères, cladocéres et copépodes des eaux continentales grecques. *Biologia gallo-hellenica*, **15**, 129–146.

BIOSKETCHES

Dr Spyros Sfenthourakis is a Lecturer of Ecology at the Department of Biology of the University of Patras (Greece). His research interests include island biogeography, elevational gradients, the terrestrial invertebrate biodiversity of Greece, and the systematics, distribution and ecology of terrestrial isopods.

Evangelos Tzanatos has a MSc in Environmental Biology and works for his PhD thesis at the Department of Biology (University of Patras, Greece). His research focuses on the evaluation of the methodology of fish stock assessment and of integrated fisheries management. He constructs the necessary models in C++.

Dr Sinos Giokas is a Lecturer of Animal Biodiversity at the Department of Biology of the University of Patras (Greece). His research interests include the study of speciation, ecology, phylogeny and population biology of land snails, the assemblage of insular biotas, and the terrestrial invertebrate biodiversity of Greece.