

Opinion is intended to facilitate communication between reader and author and reader and reader. Comments, viewpoints or suggestions arising from published papers are welcome. Discussion and debate about important issues in ecology, e.g. theory or terminology, may also be included. Contributions should be as precise as possible and references should be kept to a minimum. A summary is not required.

Community nestedness and the proper way to assess statistical significance by Monte-Carlo tests: some comments on Worthen and Rohde's (1996) paper

Bernard Hugué, ORSTOM, Laboratoire d'Ecologie des Eaux Douces et des Grands Fleuves, Université Lyon I, F-69622 Villeurbanne cedex, France (hugué@biomserv.univ-lyon1.fr) and Jean-François Guégan, ORSTOM/U.M.R. 5554-5 du C.N.R.S., Station Méditerranéenne de l'Environnement Littoral, Université de Montpellier II, 1 quai de la Daurade, F-34200 Sète, France.

In 1994, Guégan and Hugué found that the infra-communities of parasites on the African cyprinid fish *Labeo coubie* exhibited significant nested-subset structure and were not random assemblages (Guégan and Hugué 1994). In a recent analysis of our data, Worthen and Rohde (1996) reached the opposite conclusion. These authors suggested that we assessed the statistical significance of our Monte-Carlo test in an inappropriate way. However, in the light of recent reviews on the use of computer-intensive statistical methods in biology (Manly 1991, Crowley 1992, Potvin and Roff 1993), it seems to us that our approach is

simulations. Since pioneering works in the 1970s, Monte-Carlo methods have become common tools for community ecologists. The use of these methods has been facilitated by the availability of fast micro-computers because they need computer simulations to mimic the stochastic process assumed to have generated the data under a specific null hypothesis. This increasing importance of computation-intensive statistical methods (randomization tests, Monte-Carlo methods, bootstrapping procedures) in biology has largely motivated recent reviews emphasizing the versatility of these methods due to their distribution-free nature (Manly

studies dealing with community nestedness (including that of Worthen and Rohde 1996) incorporates some drawbacks. Our note is thus mainly a discussion about the proper way to assess statistical significance by Monte-Carlo tests, a point we believe to be of concern for many community ecologists.

Community nestedness was first formulated by Patterson and Atmar (1986) to describe a particular non-random pattern of species occurrence along a species richness gradient of insular communities. A completely nested design occurs when species that compose a depauperate island community constitute a proper subset of those inhabiting richer islands. As a result, a set of such island communities arranged by species richness presents a nested series. Patterson and Atmar (1986) devised an index of nestedness (N) to quantify this pattern. This index equals zero when the presence/absence matrix is perfectly nested, and it grows larger as nestedness decreases. The statistical significance of the observed value of N is then tested by Monte-Carlo

subsequent workers, such as Worthen and Rohde (1996), does not take this non-parametric property into account. These authors developed computer programs to generate mean and variance of N by Monte-Carlo simulations under two different null hypotheses of species distribution within localities (RANDOM0 and RANDOM1). Under the RANDOM0 hypothesis, each species has the same probability to be drawn during the construction of a simulated community whereas under RANDOM1 each species has a probability to be drawn weighted by its observed occurrence. When the simulations are completed, the statistical significance of the observed value is evaluated by comparing the z -score (the difference between the observed value and the simulated mean value divided by the simulated standard deviation) with a standard normal distribution. Thus, it is assumed that N is normally distributed under the null hypothesis, which represents a conventional parametric approach. Empirical evidences suggest that, under the two currently used null hypotheses, N some-

(DNK)
Vol 80 n°3



times presents a skewed distribution (Wright and Reeves 1992, Worthen and Rohde 1996) and it is, at best, approximately normally distributed. These conditions lead to a biased estimation of p , the statistical significance, that does not converge toward the true value with increasing number of simulations. Moreover, the calculation of z -score is based on estimated mean and variance and, thus, it is exposed to an unknown amount of error even if the assumption of normality is met. As pointed out by Wright and Reeves (1992), N

out by Worthen and Rohde (1996), their study is one of the largest investigations of nestedness in a colonization-dominated system in which they conducted 38 statistical tests of parasite communities nestedness for fish host species from different habitats and with distinct life-histories. This made their study an ideal candidate for the use of methods allowing a quantitative and simultaneous analysis of several tests of the same hypothesis, such as meta-analysis (Arnqvist and Wooster 1995) or Fisher's combined probabilities (Sokal and

rection should be integrated in the calculation of the z -score.

The procedure used by Worthen and Rohde (1996) following Patterson and Atmar (1986) differs actually from the correct way to assess statistical significance by Monte-Carlo tests or other computer intensive methods

tion of probabilities of individual tests. Consequently, the results must be treated with caution if systematically biased probabilities are included, as is possible if z -scores are used.

Spurred by Worthen and Rohde's (1996) re-analysis of our data set (Guégan and Huguény 1994), we re-

the incidence totals of matrices it generates do not approximate those of the observed matrix (Wright and Reeves 1992, Hugueny pers. obs.). Obviously, improvements of Monte Carlo algorithms are needed in this field but, whatever the algorithm used, our message is that the use of z -scores should be prohibited.

Acknowledgements – We would like to thank S. Dolédec and D. H. Wright for useful comments and D. Simberloff for his continuous encouragement to develop Monte-Carlo approaches in parasite community ecology.

References

- Andrén, H. 1994. Can one use nested subset pattern to reject the random sample hypothesis? Examples from boreal bird communities. – *Oikos* 70: 489–491.
- Arnqvist, G. and Wooster, D. 1995. Meta-analysis: synthesizing data from multiple studies. – *Evolutionary Ecology* 9: 23–47.
- Guégan, J.-F. and Hugueny, B. 1994. A nested parasite species subset pattern in tropical fish: host as major determinant of parasite infracommunity structure. – *Oecologia* 100: 184–189.
- Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. – Chapman and Hall, London.
- Patterson, B. D. and Atmar, W. 1986. Nested subsets and the structure of insular mammalian faunas and archipelagos. – *Biol. J. Linn. Soc.* 28: 65–82.
- Potvin, C. and Roff, D. A. 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics? – *Ecology* 74: 1617–1628.
- Poulin, R. 1996. Richness, nestedness, and randomness in parasite infracommunity structure. – *Oecologia* 105: 545–551.
- Sokal, R. R. and Rohlf, F. J. 1981. *Biometry*, 2nd ed. – Freeman, San Francisco.
- Worthen, W. B. and Rohde, K. 1996. Nested subset analyses of colonization-dominated communities: metazoan ectoparasites of marine fishes. – *Oikos* 75: 471–478.
- , Carswell, M. L. and Kelly, K. A. 1996. Nested subset patterns of colonial invertebrates. *See acknowledgements, p. 47.*