

# Geographical distances and the similarity among parasite communities of conspecific host populations

R. POULIN<sup>1\*</sup> and S. MORAND<sup>2</sup>

<sup>1</sup>Department of Zoology, University of Otago, P.O. Box 56, Dunedin, New Zealand

<sup>2</sup>Centre de Biologie et d'Ecologie tropicale et méditerranéenne, Laboratoire de Biologie Animale (UMR 5555 CNRS), Université de Perpignan, 66860 Perpignan, France

(Received 16 February 1999; revised 14 April 1999; accepted 14 April 1999)

## SUMMARY

The geographical distance between conspecific host populations is no doubt a key determinant of the likelihood that exchanges of parasite species occur between these populations. This variable must therefore be taken into account in studies that compare parasite species richness or similarity among host populations. This paper presents a multivariate approach, based on the permutation of matrices, that allows all pairwise geographical distances between host populations to be included as independent variables. The method is illustrated with 3 separate data sets on parasite communities of conspecific fish from different lakes. In 2 of 3 cases, geographical distances among lakes had a significant influence on the similarity of their parasite communities. The effect of geographical distance on species richness of parasite communities also proved important in 2 of the 3 case studies. These examples demonstrate the pervasive influence of distances among host populations on their parasite communities, and the need to properly control for them in statistical analyses.

Key words: distance matrices, geography, lake characteristics, permutations, species richness.

## INTRODUCTION

The processes determining the evolution of animal communities have received much attention from ecologists (Strong *et al.* 1984; Ricklefs & Schluter, 1993). Parasite communities are good models for such studies because they can often be investigated at several scales. Studies at the level of the parasite component community, i.e. all the parasites exploiting a host population, have been numerous in recent years (Esch, Bush & Aho, 1990; Poulin, 1997). In fish hosts, it is common practice to sample the parasite component communities of fish of the same species but from different water bodies, and then attempt to relate the species richness of component communities or their similarities to the characteristics of the water body in which they occur (Kennedy, 1978; Marcogliese & Cone, 1991; Hartvigsen & Halvorsen, 1994). For instance, Kennedy (1978) has reported that the species richness of metazoan parasites of brown trout, *Salmo trutta*, correlates positively with lake surface area and negatively with altitude of the lake above sea level among British lakes. Hartvigsen & Halvorsen (1993) found the same patterns in Norwegian lakes. This approach has merit because it can serve to identify key factors facilitating or preventing the recruitment of new parasite species into the component community.

\* Corresponding author: Department of Zoology, University of Otago, P.O. Box 56, Dunedin, New Zealand. Tel: +64 3 4797983. Fax: +64 3 4797584. E-mail: robert.poulin@stonebow.otago.ac.nz

The geographical distance between component communities, however, might be an important source of similarity between them both in terms of species composition and species richness, as recently emphasized by Hartvigsen & Halvorsen (1994). When studying parasite communities at a higher scale, i.e. when comparing the parasite faunas of different host species in order to elucidate which host properties are associated with high parasite species richness, it is now widely accepted that one must control for phylogenetic relatedness between host species (Poulin, 1995; Gregory, Keymer & Harvey, 1996). Surely, if short phylogenetic distances between host species can result in similarities in their parasite faunas, so could short geographical distances between conspecific host populations affect the similarities between their parasite component communities. Exchanges of parasite species over evolutionary time are facilitated among phylogenetically adjacent host species, as they are over ecological time among geographically adjacent host populations of the same species (Poulin, 1998). It is therefore essential to make sure that the effect of distance is taken into account.

One way in which investigators of parasite component communities of fish have attempted to control for the confounding effect of geographical distances has been to include as a property of each lake the distance between the lake and the nearest other lake (e.g. Kennedy, 1978; Hartvigsen & Kennedy, 1993). This was a necessary oversimplification, however, imposed by the statistical methods available at the time. Pairs of lakes do not occur in isolation in the

landscape; they are part of a mosaic of lakes and rivers, with exchanges of parasite species occurring among all lakes with respective probabilities determined by the distance between pairs of lakes. One lake could be 10 km from the nearest other lake, with no other lakes occurring within a 50 km radius; another lake could also be 10 km from its closest neighbour, but with a further 10 lakes occurring within the 50 km radius. The distance measure used by earlier investigators (Kennedy, 1978; Hartvigsen & Kennedy, 1993) would fail to distinguish between these very different scenarios, and would treat the two lakes as though they had identical probabilities of being colonized by new parasite species.

In this paper, we present a statistical method based on the permutation of distance matrices to control simultaneously for the influence of geographical distances between all pairs of lakes in studies of parasite component communities of fish. The method has been used in parasite community ecology before, to control for phylogenetic distances in a comparative analysis of species richness across rodent species (Feliu *et al.* 1997). We use 3 published case studies to illustrate how it can also be used to control for geographical distances among populations of the same host species, and to show how the role of geographical distances is pervasive and sometimes more important than any other characteristic of the host populations.

#### MATERIALS AND METHODS

We chose 3 case studies in which sufficient information was presented on both the parasite communities of fish and on the characteristics of the lakes surveyed. The 3 studies examined here are investigations of (1) component communities of metazoan parasites in brown trout, *Salmo trutta*, from 10 British lakes (Kennedy, 1978); (2) component communities of helminth parasites of brown trout, *S. trutta*, from 10 reservoirs in southwest England (Kennedy, Hartvigsen & Halvorsen, 1991; Hartvigsen & Kennedy, 1993); and (3) component communities of intestinal helminths in carp, *Cyprinus carpio*, from 6 Chinese lakes (Nie *et al.* 1999).

Each study was treated separately. In each analysis, we looked at 2 dependent variables, the species richness of component communities and the similarity in species composition between component communities. Similarity between pairs of lakes was estimated using Jaccard's index, computed as  $c/(a+b-c)$ , where  $a$  and  $b$  are the species richness in the 2 communities being compared, and  $c$  is the number of parasite species they have in common. We determined whether the following independent variables had any influence on the above 2 dependent variables: geographical distance among lakes, lake

surface area, lake depth (not available for British lakes or reservoirs), lake altitude (not available for Chinese lakes), and lake age (in the case of English reservoirs only). Geographical distances among all pairs of lakes were obtained from maps; the other 3 lake variables were obtained from the original studies.

We used a method based on the permutations of distance matrices to perform a multivariate analysis of the determinants of component community richness and composition. The method was described by Legendre, Lapointe & Casgrain (1994). All computations were performed using the program PERMUTE 3.4 (by P. Casgrain, available on the internet at <http://alize.ere.umontreal.ca/~casgrain/>); other programs capable of permutations on matrices would be just as suitable and would give the same results. Two of our variables, the Jaccard indices and geographical distances among lakes, were initially in the form of matrices representing all possible pairwise comparisons between lakes. The other variables (species richness, lake surface area, lake depth, lake altitude, reservoir age) were lake traits that were transformed into distance matrices by computing the Euclidean distance (i.e. the absolute value of the difference) between values for all possible pairs of lakes. Multiple regressions were performed on the values in the matrices, and repeated after each of 999 random permutations of the dependent variable matrix. Partial regression coefficients were obtained for each independent variable and their significance tested; the significance was calculated as the probability of obtaining, in the 999 random rearrangements of matrices, a regression coefficient greater than or equal to the observed one. Because of the number of permutations we performed, 0.001 was the lowest possible probability value. Since we want to enter in the regression only the independent variables that contribute significantly to the explanation of the dependent variable, we used a stepwise regression approach with a backward elimination procedure (see Legendre *et al.* 1994). In other words, all independent variables were initially included, and at each step we dropped the variable whose partial regression coefficient had the highest  $P$  value (provided it was equal to or greater than 0.100) and repeated the multiple regression.

#### RESULTS

##### *Parasites of trout in British lakes*

In the analysis of species richness, geographical distance among lakes was dropped as an explanatory variable, and only lake surface area and lake altitude proved good predictors of species richness (Table 1). This result agrees with the conclusions of Kennedy (1978), which were based on single-factor analyses.

Table 1. Effect of geographical distance and other independent variables on parasite species richness and similarity among component communities of brown trout from 10 British lakes

(An optimal subset of explanatory variables was selected using a stepwise regression. Numbers are partial regression coefficients as well as the coefficient of determination,  $r^2$ ; the significance (in parentheses) is derived from 999 random permutations.)

Dependent variable		
Independent variable	Step 1	Step 2
Species richness		
Geographical distance	-0.057 (0.398)	—
Lake surface area	0.462 (0.001)	0.461 (0.002)
Lake altitude	0.339 (0.029)	0.343 (0.026)
$r^2$	0.329 (0.018)	0.325 (0.011)
Species similarity (Jaccard Index)		
Geographical distance	-0.307 (0.030)	-0.306 (0.028)
Lake surface area	0.075 (0.423)	—
Lake altitude	-0.449 (0.009)	-0.450 (0.005)
$r^2$	0.284 (0.017)	0.278 (0.006)

In our multivariate analysis of similarity in parasite species composition, however, geographical distance was retained as a predictor, with the value of the Jaccard Index being inversely proportional to the distance between lakes (Table 1).

*Parasites of trout in English reservoirs*

Our analysis of the predictors of species richness retained only geographical distance among reservoirs as a significant predictor (Table 2). Geographical distance correlated negatively with species richness, but since richness values in our analysis are in fact contrasts in richness between pairs of reservoirs, the

result means that similarity in species richness increases with increasing distance among reservoirs. This counter-intuitive finding corresponds exactly to what Hartvigsen & Kennedy (1993) reported based on single-factor analyses. In our multivariate analysis of similarity among reservoirs in parasite species composition, geographical distance was dropped as a predictor, and only reservoir surface area correlated strongly with the Jaccard Index, with reservoirs of similar surface area sharing more parasite species than reservoirs with different areas (Table 2).

*Parasites of Chinese carp*

The authors of the original study performed no formal analysis of the role of lake features on the parasite component communities (Nie *et al.* 1999). In our analyses, we found that geographical distance contributed to the variability in both species richness and species composition among lakes (Table 3). Shorter geographical distances between lakes were associated with a greater similarity in parasite species composition between them (i.e. a high Jaccard Index), and nearby lakes harboured numbers of parasite species that were more similar than those of distant lakes. The latter effect is indicated by the positive relationship between geographical distance and species richness since, as mentioned above, richness values in the matrix are in fact differences in richness between pairs of lakes. The influence of geographical distance was statistically stronger on species similarity than on species richness (Table 3). Of the 2 lake characteristics considered, only lake surface area influenced the parasite community, correlating positively with species richness only (Table 3).

Table 2. Effect of geographical distance and other independent variables on parasite species richness and similarity among component communities of brown trout from 10 English reservoirs

(An optimal subset of explanatory variables was selected using a stepwise regression. Numbers are partial regression coefficients as well as the coefficient of determination,  $r^2$ ; the significance (in parentheses) is derived from 999 random permutations.)

Dependent variable				
Independent variable	Step 1	Step 2	Step 3	Step 4
Species richness				
Geographical distance	-0.224 (0.056)	-0.200 (0.058)	-0.203 (0.055)	-0.218 (0.050)
Reservoir surface area	0.038 (0.348)	0.028 (0.383)	—	—
Reservoir altitude	0.064 (0.349)	—	—	—
Reservoir age	0.177 (0.154)	0.166 (0.144)	0.166 (0.172)	—
$r^2$	0.079 (0.578)	0.075 (0.395)	0.075 (0.200)	0.047 (0.120)
Species similarity (Jaccard Index)				
Geographical distance	-0.121 (0.234)	-0.161 (0.152)	—	—
Reservoir surface area	-0.342 (0.001)	-0.325 (0.001)	-0.305 (0.004)	—
Reservoir altitude	-0.107 (0.253)	—	—	—
Reservoir age	0.204 (0.111)	0.221 (0.102)	0.235 (0.068)	—
$r^2$	0.185 (0.078)	0.176 (0.037)	0.151 (0.022)	—

Table 3. Effect of geographical distance and other independent variables on parasite species richness and similarity among component communities of carp from 6 Chinese lakes

(An optimal subset of explanatory variables was selected using a stepwise regression. Numbers are partial regression coefficients as well as the coefficient of determination,  $r^2$ ; the significance (in parentheses) is derived from 999 random permutations.)

Dependent variable	Step 1	Step 2	Step 3
Independent variable			
Species richness			
Geographical distance	0.674 (0.028)	0.583 (0.053)	—
Lake surface area	0.794 (0.033)	0.674 (0.049)	—
Lake depth	0.344 (0.120)	—	—
$r^2$	0.614 (0.040)	0.511 (0.047)	—
Species similarity (Jaccard Index)			
Geographical distance	-0.820 (0.033)	-0.803 (0.011)	-0.773 (0.022)
Lake surface area	-0.105 (0.304)	-0.082 (0.263)	—
Lake depth	-0.066 (0.307)	—	—
$r^2$	0.607 (0.063)	0.603 (0.026)	0.597 (0.024)

## DISCUSSION

Our results show that the geographical distance between component communities of fish parasites is often the best, most general predictor of the similarity in species composition. The effect of geographical distance on species richness of component communities also proved important, though it may be less predictable. These conclusions are based on studies of parasite communities of trout and carp, and these may not be typical fish host species. Both tend to harbour host-specific and widespread parasite species; it would be important to validate the role of geographical distance using other host species that harbour different types of parasites. The main contribution of our study, however, is to introduce a multivariate method that allows the influence of geographical distance, if any, to be controlled in investigations of the determinants of species richness and composition of parasite communities. The approach has limitations similar to those of other multivariate methods. For instance, the number of lakes (or host populations) studied must be greater than the number of variables included in the analysis, otherwise there are not enough degrees of freedom to test for relationships (P. Casgrain, personal communication). Thus data from studies based on very few lakes (e.g. Hartvigsen & Halvorsen, 1993) cannot be re-analysed using permutation of distance matrices. However, for most large-scale studies, the method we describe should be perfectly adequate.

Despite their fish hosts being physically confined to one lake, exchanges of parasites among different lakes can take several routes (Esch *et al.* 1988). Parasites with complex life-cycles use hosts other than fish at some stages in their development; for

instance, helminths using birds as definitive hosts will have a greater dispersal potential than other helminths. Human activities that lead to the introduction of new parasites in certain lakes can also contribute to exchanges of parasite species among habitats. Whether parasite exchanges are mediated by bird movements, human activities or other processes, geographical distances among lakes will influence the probability of these exchanges. Birds from one lake are more likely to visit a nearby lake than a geographically remote one, and fishermen are more likely to release fish from one lake into another local lake than into a very distant one. This is not always true, of course: birds can be very selective about the lakes they visit, whether they are near or far, and fishermen can decide to introduce, legally or not, fish to distant areas. Generally, though, geographical distances among lakes are likely to mask the effects of other variables and, over time, become the best predictor of parasite species similarity among fish populations.

It would be more accurate to say that it is the true isolation of a lake in a landscape that matters, and that geographical distances among lakes are only a surrogate measure of this isolation. For instance, a fish species in one lake may be a long distance away from another lake containing conspecifics, even though its home lake may be surrounded by several nearby lakes all containing other species. Also, distant lakes are not truly isolated when we consider anadromous fish species: all lakes are linked with one another by the sea. Thus geographical distances may not always mirror true isolation (Esch *et al.* 1988). There are also many factors that promote dissimilarity between lakes in the composition of parasite communities, and these may sometimes be stronger than those, such as geographical closeness,

that promote similarity (Kennedy *et al.* 1991; Hartvigsen & Kennedy, 1993). The action of these factors could mean that distant lakes usually have dissimilar parasite communities, but that near lakes do not necessarily have similar parasite communities. Nevertheless, because geographical distances are often likely to be important and because they are a good index of true isolation, their effect should be measured in statistical analyses; if they have none then no harm is done by including them.

We used studies of fish parasites to illustrate the statistical approach and the effect of geographical distances on parasite component communities. The same approach, however, should be used in all comparative studies of parasite component communities from several host populations, whatever type of hosts are investigated. The method would apply equally well for conspecific marine fish from different reefs, conspecific birds from different islands, or conspecific mammals from different regions of the same continent (see examples in several chapters of Esch *et al.* 1990). It may well be that the influence of geographical distance is weaker among component communities that are not physically isolated from one another, such as the lake communities we analysed. Exchanges of parasite species should be more common when physical barriers are absent, thus rapidly negating any effect of large distances between communities (see Poulin, 1997). Still, since the effect of geographical distance can be estimated and/or removed from statistical analyses, there is no reason not to do so.

Permutation methods are now widely used in biology and offer a wide range of advantages over traditional statistical tests (Manly, 1991). There are several computer programs other than PERMUTE, all easily available, which perform the sort of matrix permutations described in this paper (such as RT 2.1, available from B. Manly, Centre for Applications of Statistics and Mathematics, University of Otago, Dunedin, New Zealand). It should therefore become a requirement for future studies of the determinants of species composition and richness in component communities to control for the confounding and often important effect of geographical distances. Given the traditional view that host populations represent islands for parasite colonization (Kuris, Blaustein & Alió, 1980) and the main predictions of island biogeography theory regarding the distance between islands and the sources of colonizing species (MacArthur & Wilson, 1967), this seems a very reasonable recommendation.

We are grateful to C. R. Kennedy for sending us unpublished information on the geographical location of the British lakes he studied. R. P. thanks the Université de Perpignan for financial support during his visit.

## REFERENCES

- ESCH, G. W., BUSH, A. O. & AHO, J. M. (1990). *Parasite Communities: Patterns and Processes*. Chapman & Hall, London.
- ESCH, G. W., KENNEDY, C. R., BUSH, A. O. & AHO, J. M. (1988). Patterns in helminth communities in freshwater fish in Great Britain: alternative strategies for colonization. *Parasitology* **96**, 519–532.
- FELIU, C., RENAUD, F., CATZEFLIS, F., HUGOT, J.-P., DURAND, P. & MORAND, S. (1997). A comparative analysis of parasite species richness of Iberian rodents. *Parasitology* **115**, 453–466.
- GREGORY, R. D., KEYMER, A. E. & HARVEY, P. H. (1996). Helminth parasite richness among vertebrates. *Biodiversity and Conservation* **5**, 985–997.
- HARTVIGSEN, R. & HALVORSEN, O. (1993). Common and rare trout parasites in a small landscape system. *Parasitology* **106**, 101–105.
- HARTVIGSEN, R. & HALVORSEN, O. (1994). Spatial patterns in the abundance and distribution of parasites of freshwater fish. *Parasitology Today* **10**, 28–31.
- HARTVIGSEN, R. & KENNEDY, C. R. (1993). Patterns in the composition and richness of helminth communities in brown trout, *Salmo trutta*, in a group of reservoirs. *Journal of Fish Biology* **43**, 603–615.
- KENNEDY, C. R. (1978). An analysis of the metazoan parasitocoenoses of brown trout *Salmo trutta* from British lakes. *Journal of Fish Biology* **13**, 255–263.
- KENNEDY, C. R., HARTVIGSEN, R. & HALVORSEN, O. (1991). The importance of fish stocking in the dissemination of parasites throughout a group of reservoirs. *Journal of Fish Biology* **38**, 541–552.
- KURIS, A. M., BLAUSTEIN, A. R. & ALIÓ, J. J. (1980). Hosts as islands. *American Naturalist* **116**, 570–586.
- LEGENDRE, P., LAPOINTE, F. -J. & CASGRAIN, P. (1994). Modelling brain evolution from behavior: a permutational regression approach. *Evolution* **48**, 1487–1499.
- MACARTHUR, R. H. & WILSON, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton.
- MANLY, B. F. J. (1991). *Randomization and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- MARCOGLIESE, D. J. & CONE, D. K. (1991). Importance of lake characteristics in structuring parasite communities of salmonids from insular Newfoundland. *Canadian Journal of Zoology* **69**, 2962–2967.
- NIE, P., YAO, W. J., GAO, Q., WANG, G. T. & ZHANG, Y. A. (1999). Diversity of intestinal helminth communities of carp from six lakes in the flood plain of the Yangtze River, China. *Journal of Fish Biology* **54**, 171–180.
- POULIN, R. (1995). Phylogeny, ecology, and the richness of parasite communities in vertebrates. *Ecological Monographs* **65**, 283–302.
- POULIN, R. (1997). Species richness of parasite assemblages: evolution and patterns. *Annual Review of Ecology and Systematics* **28**, 341–358.
- POULIN, R. (1998). *Evolutionary Ecology of Parasites: From Individuals to Communities*. Chapman & Hall, London.

RICKLEFS, R. E. & SCHLUTER, D. (1993). *Species Diversity in Ecological Communities : Historical and Geographical Perspectives*. University of Chicago Press, Chicago, USA.

STRONG, D. R., SIMBERLOFF, D., ABELE, L. G. & THISTLE, A. B. (1984). *Ecological Communities : Conceptual Issues and the Evidence*. Princeton University Press, Princeton, USA.