npg

Spatial structure of shell polychromatism in populations of *Cepaea nemoralis*: new techniques for an old debate

A Bellido[1], L Madec[2], J-F Arnaud[2,3] and A Guiller[3]

[1]UMR CNRS 6553, Equipe 'Evolution des populations et des aspèces' Service d'Ecologie du sol et de Biologie des Populations, Station Biologique de Paimpont, 35380 Plélan-le-Grand, France; [2]UMR CNRS 6553, Equipe 'Evolution des populations et des aspèces' Service de Zoologie et Ecophysiologie, Campus de Beaulieu, Avenue du Gal Leclerc, 35042 Rennes cedex, France; [3]UMR CNRS 6553, Equipe 'Evolution des populations et des aspèces' Service de Parasitologie Pharmaceutique, Avenue du Pr. Léon Bernard, 35043 Rennes cedex, France

A conspicuous shell polychromatism is observed in colonies of *Cepaea nemoralis* from western France (Brittany). The present study is intended to search for a spatial structure of shell features at this scale and to infer evolutionary processes from the observed patterns. We used a database of morph frequencies (six composite phenotypes were retained) measured on 213 samples regularly distributed on the whole studied area. Data analysis was based on two distinct multivariate methods leading to the following steps: (i) to search for a structure without reference to environmental conditions with a method (global principal components analysis: GPCA) which takes into account the spatial information by means of a neighbouring relationship between sampling points (Delaunay triangulation); (ii) to test the structuring power of environmental conditions by means of two explanatory factors (distance from the sea, altitude) involved in a redundancy analysis (RDA); (iii) to search for a spatial structure using residuals of the previous analysis, ie, after removing effects of environmental conditions. Global covariance accounted for 26.4% of the total variance, leading to a highly significant autocorrelation for each phenotype (step 1). Geographical mapping of factorial scores resulting from global analysis showed a well structured littoral zone and a strong southern-northern inland differentiation. Sixteen percent of the total variance was expressed in RDA but all morphs were not equally concerned. After removing environmental effects, a significant spatial structure still remains but was essentially caused by random processes. We argue for the importance of these last phenomena.
Heredity (2002) **88**, 75–82. DOI: 10.1038/sj.hdy/6800012

Introduction

In the recent past, patterns of geographical variation in shell polychromatism of *Cepaea nemoralis* have been the subject of various evolutionary interpretations (see for a review of ideas, Jones *et al*, 1977; Clarke *et al*, 1978; Lamotte, 1988; Gould and Woodruff, 1990; Cook, 1991, 1998). With the exception of local situations where consequences of selective pressures and/or random fluctuations on phenotype distributions are clear, it often remains difficult to disentangle the effects of these processes, leading sometimes to conflicting points of view. Thus, during the last 50 years, Lamotte and co-workers have accumulated data suggesting the minor part of selection, especially at a microgeographical scale, to explain the distribution of genetic variation and the maintenance of polymorphism (see however the linear relations revealed by canonical analyses between phenotypic constitution of colonies and environmental factors

for different French regions in Gerdeaux (1978), Khemici (1987), and Ratel (1987)), whereas the majority of papers based on composite phenotypes highlight the influence of ecological conditions or of a frequency-dependent predation (eg, Clarke *et al*, 1978; Cain, 1983). However, the detection of a stable lack of association between morph frequencies and habitat background, which has led to the concept of 'area effects' (Cain and Currey, 1963), needs to consider historical events and/or gene interactions experienced by the populations (for different points of view, see Clarke, 1966; Cameron and Pannett, 1985; Goodhart, 1987). Now, such discrepancies are primarily viewed in *Cepaea* as a consequence of habitat instability related to human activities or, more precisely for area effects, the result of invasions from refugia in which genetic differentiation has occurred (Davison and Clarke, 2000).

Here, after a brief outline of multivariate statistics using neighbouring relationships between sample units, we will look at a significant structuring of phenotypic frequencies on a regional scale, in Brittany, north-western France (Figure 1). In this region, colonies are polymorphic for yellow and pink at the shell colour locus (brown shells are not considered because of a mean regional fre-
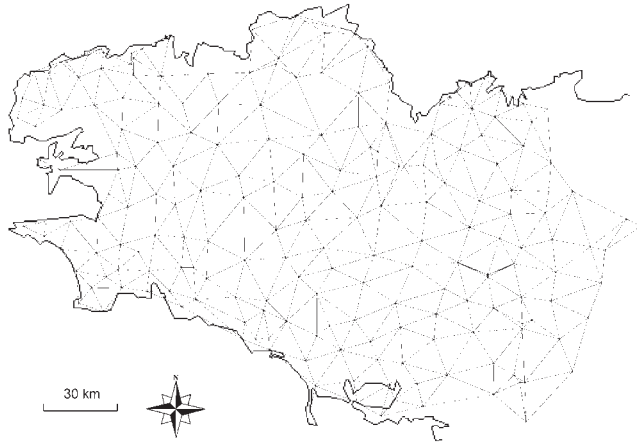
Figure 1 Delaunay neighbourhood graph performed with 213 sampling sites. The average distance between connected points is 12.3 km, the maximum distance is 27.1 km. This graph was modified according to main geographical obstacles.

quency <1%), and for zero-, mid- and five-banded. Three decisions which have led this work have to be mentioned and justified: (i) the bias towards consideration of regional evolutionary pressures only is accounted for, insofar as they are established at such a large scale for a sufficiently long time to detect their influence on the spatial distribution of gene frequencies; (ii) the choice of the region, Brittany, where the influence of ecological conditions, already studied by Guerrucci-Henrion (1966), is adequately summarized here by two measurements (distance from littoral, altitude), and where effects related to genetic drift are strongly suspected because of a population structure often based on well isolated colonies of few individuals; and (iii) the use of statistics which avoid partial interpretations deduced from previous univariate analyses (examples in references cited above) and which split the genetic variation according to selective and random patterns.

The study is carried out following three steps: (i) search for a spatial structure without reference to the spatial variation of ecological conditions; (ii) test of the structuring power of environmental conditions by means of the two explanatory factors (redundancy analysis); and (iii) search for a spatial structure using residuals of the previous analysis, ie, after removing effects of environmental conditions.

Materials and methods

Sampling
We used in part a database made up more than 30 years ago (Guerrucci-Henrion, 1966), recently supplemented by a sampling effort concentrated in zones under-represented in the initial work. Thus, more than 350 localities have been sampled in Brittany, but an analysis based on such various sample units could be biased by several sources of frequency or habitat changes, with potential consequences on the observed spatial patterns. First, the oldest samples were collected more than 30 years ago and it is well established that substantial frequency changes over time can occur (Murray and Clarke, 1978; Cain *et al*, 1990; Cook and Pettitt, 1998; Cowie and Jones, 1998; Cook

et al, 1999). However, temporal variation alone seems insufficient to affect the regional trends of polychromatism and then, to distort a spatial interpretation of the data at this scale for the following reasons. Such significant changes have only been detected at a local scale and they need a longer time in *Cepaea* to become appreciable, especially if they result from a modification in selective pressures (Cook, 1998). Moreover, tests carried out on six Breton populations surveyed at least twice, several years apart, show only one significant temporal fluctuation of the morph frequencies (unpublished results). Secondly, as our goal was to collect samples which should express the phenotypical composition of sites rather than the composition of a particular deme, microgeographical heterogeneity of morph frequencies, often observed in linear colonies, was ignored. Consequently, some collecting areas were expanded as much as necessary to reach the minimum sample size of 20 individuals, without however exceeding 1000 m^2 that was less than the panmictic area of the species (Lamotte, 1951).

Only 213 samples were retained to carry out analyses for the following reasons: (i) a minimum sample size of 20 individuals was the prerequisite to reduce random fluctuations of morph frequencies; (ii) only one sample per multi-sampled locality was retained from the initial work in 1966, in order to reach a uniform weighting of geographical zones over the whole sampled area, especially in over-sampled southern zones (Quimper, Vannes). Such an elimination avoids the statistical bias related to unbalanced sampling for methods involving neighbourhood structures. In addition, the rarest shell patterns in the whole sampling (frequency <1%) were removed from the analyses so the six morphs retained include more than 97% of the total number: these 18 659 individuals were distributed according to the following percentages: (i) P00000 (P0B: pink shell without band) = 18.2%; (ii) P00300 (P1B: pink shell with band 3 present) = 18.3%; (iii) P12345 (P5B: pink shell with all the bands present) = 13.1%; (iv) Y00000 (Y0B: yellow shell without band) = 18.3%; (v) Y00300 (Y1B: yellow shell with band 3 present) = 20.5%; (vi) Y12345 (Y5B: yellow shell with all the bands present) = 11.6%.

Spatial structure analysis
In the first and third steps of the study (see above), we used a method which takes into account the spatial information (geographical coordinates) in multivariate analyses by means of a neighbouring relationship between sampling points (Thioulouse *et al*, 1995). This method, which does not involve a regular sample scheme, needs firstly to define a neighbourhood structure of the sites by means of a connection network. It is deduced in the present study from the Delaunay triangulation (Green and Sibson, 1977; see also Heywood, 1991 for a discussion on algorithms used to define networks of paired neighbours) (Figure 1). The resulting neighbouring matrix $M = [w_{ij}]$ is a symmetric matrix of between-site connections ($w_{ij} = 1$ if sites i and j are connected, $w_{ij} = 0$ otherwise).

For the analysis of spatial structure of a single variable x, total variance VAR(x) is partitioned between global variability GV(x) and local variance LV(x) according to Thioulouse *et al* (1995):

$$VAR(x) = GV(x) + LV(x), \text{ with}$$

$$VAR(x) = \sum_i p_i.(x_i - \bar{x}_D)^2$$

$$GV(x) = \sum_j \sum_i p_{ij} (x_i - \bar{x}_D)(x_j - \bar{x}_D)$$

$$LV(x) = \sum_j \sum_i p_{ij} (x_i - x_j)^2$$

and for localities *i* and *j*,

$$p_{ij} = \frac{w_{ij}}{W}; \ W = \sum_{ij} w_{ij}; \ p_i = \frac{1}{W} \sum_j w_{ij}; \ \bar{x}_D = \sum p_i x_i$$

Whereas local analysis focuses on the variability between neighbours, leading to the search of 'boundaries' (sharp transitions between sites, eg, Durand *et al* (1999)), global analysis defines groups of sites on the basis of similarity between neighbours (maximal spatial autocorrelation). For a given variable (morph frequency), one calculates a relative global autocovariance according to $RGV(x) = GV(x)/VAR(x)$ which is merely a p_i-*weighted* Moran index (*I*).

In the multivariate case, this method also allows the introduction of Geary and Moran indices (Cliff and Ord, 1973) into multivariate analyses. In our study, the maximization of total, global and local variability is extended to *p* variables through principal component analysis (PCA). Thus, three kinds of PCA may be carried out according to the weight matrix in the statistical triplet: (i) total PCA (Xd, Ip, D); (ii) local PCA (LPCA) with (Xd, Ip, D-P); and (iii) global PCA (GPCA) with (Xd, Ip, P), with Ip = (n × n) identity matrix:

$$D = diag(p_1,...p_i...p_n), \text{ and } P = \frac{1}{W} M \text{ and } M = [w_{ij}]_{nn}$$

The global analysis of the spatial structure is in fact a multivariate Moran index analysis. The row (sites) factorial scores are the linear combinations of the six morph frequencies which maximize global variability. Similarly to the univariate case, spatial correlation associated with each factorial axis is a Moran index of p_i-weighted site scores.

As the present paper focuses on analyses of global variability, autocorrelation will be detailed by computing an autocorrelation index (Moran's *I*) for each variable: (i) on the whole data set from the neighbourhood graph; and (ii) for each distance class (see results), leading to the construction of a correlogram which represents the (dis)similarity between pairs of sampled sites as a function of their geographical distance (Sokal and Oden, 1978).

Redundancy analysis
In the second step of the study, we used redundancy analysis (RDA or PCAIV) to measure the effect of two environmental factors on fluctuations of the six selected morph frequencies. A precise description of the method is given in van den Wollenberg (1977), Sabatier *et al* (1989) and Lebreton *et al* (1991). These frequencies (*f*) are used as quantitative variables of interest after standardisation of the *asin*($\sqrt{f}$) value. The two environmental factors, taken as instrumental variables in the table crossing sites and morphs, are 'distance from littoral' (DL) and 'altitude'. The choice of these factors is based on the assumption that they represent an integrating information of environmental pressures which directly influence the

biology of *Cepaea* in Britanny. DL was already suggested for its influence on the distribution of several morphs in this region (Guerrucci-Henrion, 1966). Preliminary analyses led to use of $\log^2(DL)$ transformation to yield a linear relation with phenotype frequencies. For each site, the available altitude measurement is the average altitude of the corresponding district.

In the third step, global analysis was performed using residual variance of morph frequencies, which is independent from environmental pressures included in the two structuring factors DL and altitude (orthogonal space to 'DL * Altitude').

Software availability
All calculations and graphs were made with ADE-4 (Thioulouse *et al*, 1997). The package can be obtained freely by anonymous FTP to pbil.univ-lyon1.fr, in the /pub/mac/ADE/ADE4 directory. A WWW documentation and downloading page is available at: http://pbil.univ-lyon1.fr/ADE-4.html, which also provides access to updates and user support through the ADEList mailing list.

Results

In the spatial analysis of morph frequencies (step 1), global covariance accounted for 26.4% of the total variance, leading to a highly significant autocorrelation for each phenotype (Table 1). On the basis of opposing contributions of P5B-P1B and Y0B, factorial scores of sites on the first axis of GPCA projected onto a geographical map allow the visualization of two principal patterns (Figure 2): (i) a well-structured littoral zone with a combination of morph frequencies contrasting with the inland one; (ii) a strong southern–northern inland differentiation (zone 1a–1b *vs* zone 2). The opposition between Y1B (+ other banded morphs) and P0B (+ Y0B) on axis 2 leads to a new structure based on a marked NE-SW gradient in the NE quadrant (zone 2 *vs* zone 3) contrasting with an apparent homogeneity of inland and littoral in zone 4. However, despite the nine points considered in the

Table 1 Unweighted Moran index performed on Delaunay network before and after extraction by RDA of the effect of environmental factors

	before RDA		*after RDA*		*SS*
	I Moran	*z-test (P)*	*I Moran*	*z-test (P)*	
Univariate analysis					
Y0B	0.231	$3 \ 10^{-8}$	0.166	$6.1 \ 10^{-5}$	0.738
Y1B	0.349	$<10^{-8}$	0.173	$2.8 \ 10^{-5}$	0.937
Y5B	0.155	$1.7 \ 10^{-4}$	0.071	0.0728	0.027
P0B	0.339	$<10^{-8}$	0.156	$1.6 \ 10^{-4}$	0.448
P1B	0.178	$1.9 \ 10^{-5}$	0.113	$5.6 \ 10^{-3}$	0.178
P5B	0.317	$<10^{-8}$	0.170	$4.0 \ 10^{-5}$	0.808
Global analysis					
F1	0.394	$<10^{-8}$	0.206	$3.9 \ 10^{-7}$	
F2	0.314	$<10^{-8}$	0.198	$8.7 \ 10^{-7}$	

(*P*): probability related to the z-test. *SS*: sum of squares of differences between Moran index values for each distance class; this measure of discrepancy between univariate correlograms is used for a purely descriptive purpose.
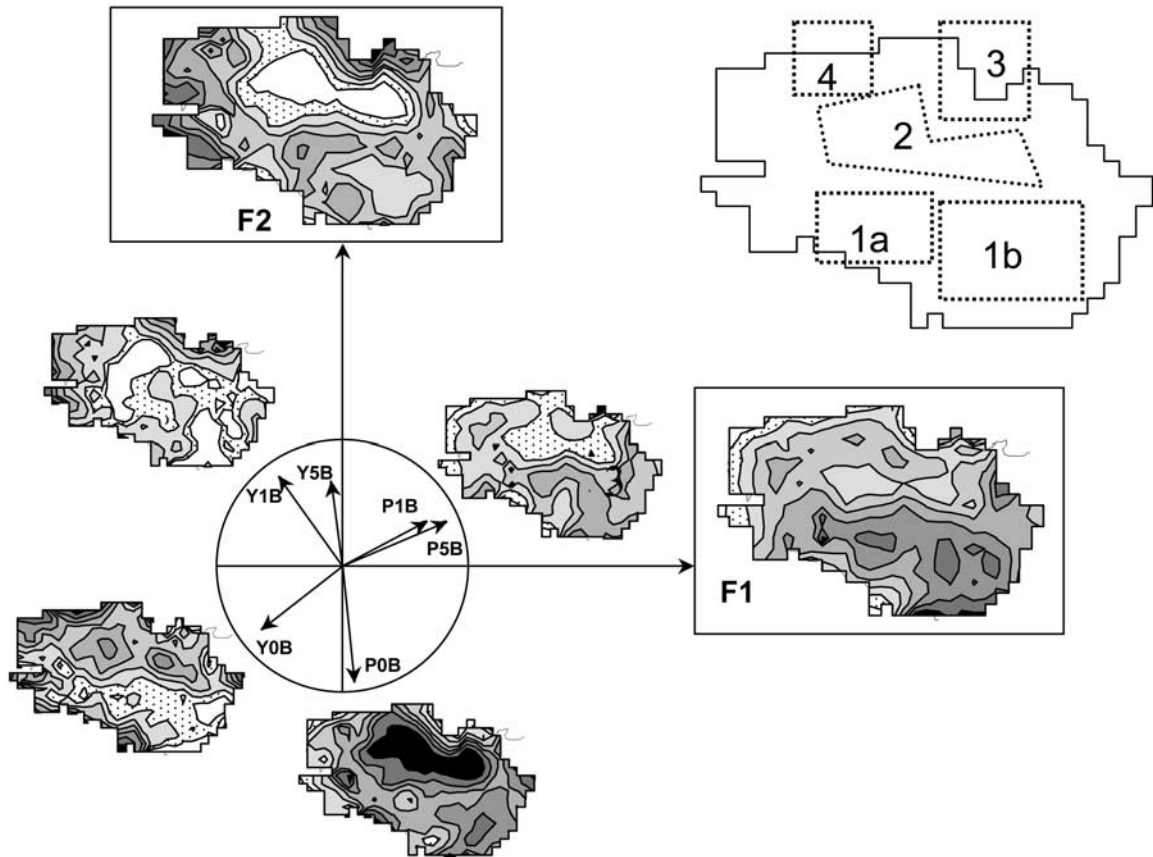
Figure 2 Graphical representation of GPCA results. (i) Projection of factorial site scores (F1 and F2) from global analysis onto the geographical map. (ii) Projection of standardized morph frequencies. Drawing of contour curves (eight grey levels) is based on the estimate of a variable at each node of a grid superimposed onto the geographical map (interpolation procedure). We use a two-dimensional locally weighted regression (Cleveland, 1979) in which nine neighbours were taken into account. Level scale (not showed) for each map is performed in order to obtain equal numbers in each class. (iii) Vectorial representation of correlations of morph frequencies with factorial scores. (iv) Localization of characteristic zones (1a, 1b, 2, 3, 4) described in the text.

locally weighted regression, this observation has possibly to be related to the fact that there is no sample site near the sea in zone 4 (Figure 1).

In step 2, influence of environmental factors on these spatial patterns was expressed by the percentage of inertia extracted by instrumental variables in RDA. Thus, percentages associated with DL and altitude, which were significantly correlated ($r = 0.424$, $P < 0.001$), were respectively 10% and 6% (Figure 3b). However, the different morphs were not concerned with the same intensity by environmental pressures: results reported in Table 2 (R^2) revealed their highly significant influence on the variation in opposite directions of Y1B and P0B or, more generally, the opposition between yellow and pink shells (Figure 3).

In the third step, spatial analysis was based on partitioning of the variance not explained in RDA, which led to a global covariance of 0.697 and a local variance of 4.31 (Table 2). Factorial scores of sites on the first two axes of GPCA are mapped as in step 1 (Figure 4). On the basis of opposition between unbanded morphs and banded ones, axis 1 (53% of the total inertia) shows again a 'island-like' pattern located in the centre of the region and which has the same features as in most of the littoral sites. Axis 2 (28% of the total inertia) mirrored an inland-littoral gradient which is complicated in the N-E quadrant and the most western part by an inversion of the general trend. This axis was essentially explained by factorial coordinates of Y1B in sites from zone 4.

Moran indices based on Delaunay network (Table 1) or distance classes (Figure 5) were computed before and after RDA. Firstly, an important decrease of the values was observed for the most spatially structured frequencies (Y1B, P0B, P5B) after removing the variance related to environmental factors (Table 1). Moran index correlograms performed on the first factor of global analysis before RDA schematically showed three different parts: (i) from 0 to 110–120 km: a progressive decline of I from strong positive values to negative ones; (ii) from 120 to 170–180 km: fluctuations of I around 0 (*X*-axis); and (iii) a strong decrease of I for the most distant pairs of sites (NW-SE direction, see Figure 5). After RDA, Moran indices are lower and lead to a correlogram in only two parts: (i) a positive autocorrelation ($I_{[0-10]} = 0.25$; $P = 0.001$) at very short distances and a decrease of I with a value equal to 0 for the [20–30] or [30–40] distance class according to the morph (correlograms per morph not shown); and (ii) I close to 0 with very slight fluctuations over remaining distance classes, except the last one for which a significant positive autocorrelation was detected. Then, the main difference between these two correlograms was a lack of spatial structure in residuals of morph frequency
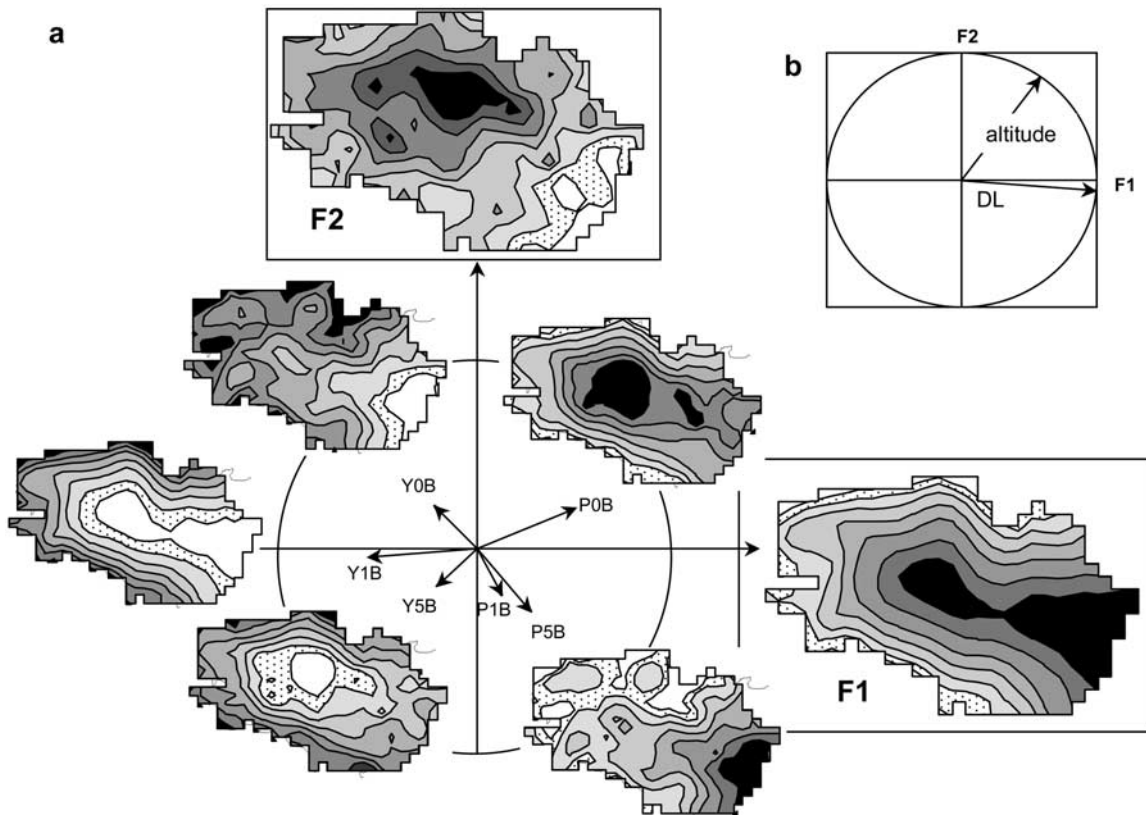
Figure 3 (**a**) Graphical representation of RDA performed with DL and altitude. Small sized maps show the projected table of the dependent variables (morph frequencies) onto the RDA space. (**b**) Vectorial representation of the correlations between environmental variables and RDA factors. See Figure 2 for explanations.

Table 2 Decomposition of variance in RDA and in orthogonal RDA. Spatial analysis is performed on variance from orthogonal RDA

	RDA		Orthogonal RDA	Global covariance		Local Variance	Total Variance	Univariate Moran Index	
	R^2	col. %	(1)	GV (2)	% (2)/0.697	LV	VAR (3)	unweighted (2)/(1)	p_i-weighted (2)/(3)
Y0B	0.087	9.0	0.913	0.152	21.76	0.70	0.86	0.166	0.177
Y1B	0.286	29.7	0.714	0.122	17.57	0.56	0.69	0.173	0.178
Y5B	0.074	7.7	0.926	0.066	9.46	0.85	0.92	0.071	0.072
P0B	0.277	28.8	0.723	0.111	15.92	0.64	0.75	0.156	0.147
P1B	0.069	7.2	0.931	0.105	15.08	0.87	0.97	0.113	0.108
P5B	0.171	17.7	0.829	0.141	20.20	0.68	0.82	0.170	0.172
Total	0.964	100.0	5.036	0.697	100.00	4.31	5.01		
Row %	16.07		83.93	13.92		86.08	100		

Note: difference between variance of the orthogonal RDA (5.04) and total variance of spatial analysis (5.01) is due to row-weighting based on the number of connections taken into account. The usual Moran index (under hypothesis of equal weighting for each point) is the ratio of global variance to unweighted initial variance. Similarly, the weighted Moran index is obtained by dividing (2) with (3).

from a distance of 30 km, which expressed a significant influence of environmental factors in the first one.

Information for each phenotype was obtained following the procedure described in Table 1. Four of the six phenotypes were characterized by a spatial structure similar to the one illustrated in Figure 5, that is a high SS value mainly related to the discrepancy between the two correlograms from the [170–180] distance class. In the same way, Y5B and P1B are poorly spatially structured,

ie, an initial positive value of *I* which decreases up to 0 at 30–50 km, followed by random fluctuations of *I* around 0 in subsequent classes.

Discussion

Knowing that a conspicuous shell polymorphism was observed in Breton colonies of *Cepaea nemoralis*, the present study is primarily intended to search for a spatial
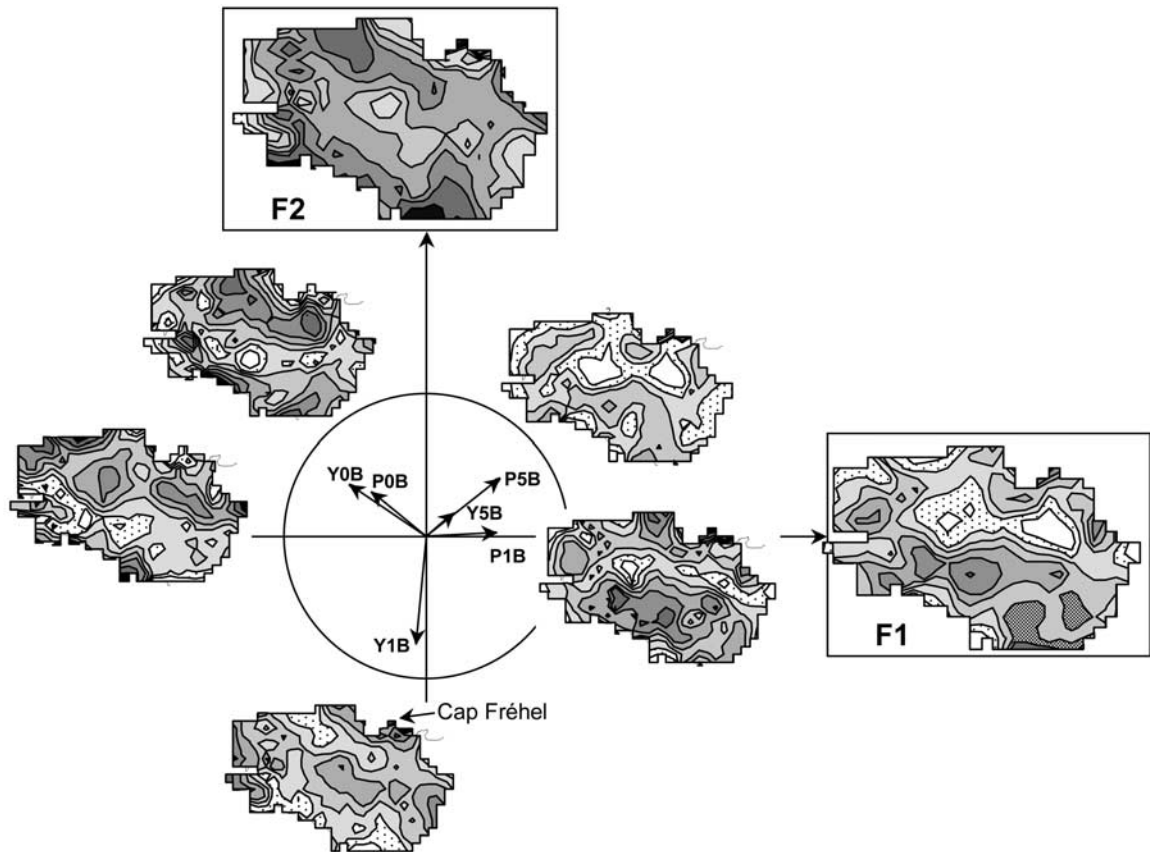
Figure 4 Graphical representation of GPCA after extraction of variance explained by DL and altitude. See Figure 2 for explanations.
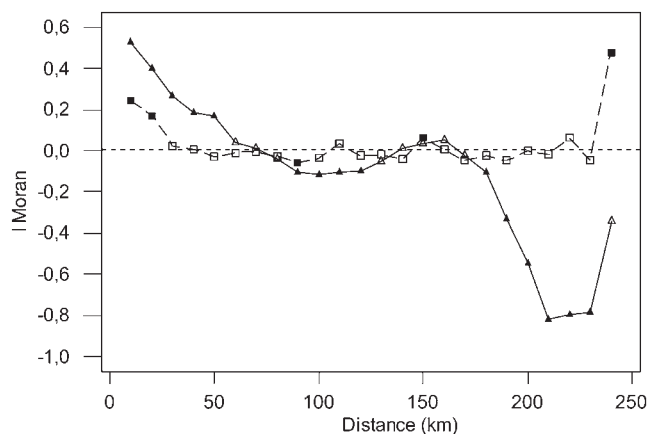

Figure 5 Moran Index correlogram (distance class: 10 km) performed on the first factor of global analysis before (triangles) and after (squares) extraction of the effect of environmental variables. Filled symbols represent significant ($P < 0.05$) indices for a given distance class.

structure of shell features at this regional scale and to infer evolutionary processes from the observed patterns by means of multivariate methods. Of the eight forces described by Jones *et al* (1977), we have only retained the climate as a selective one by means of two measurements, 'DL' and 'altitude' which essentially mirrored hygrometric conditions, with consideration neither for the genetics of polychromatism, ie, without accounting for each banding and colour type, nor for other environmental

pressures already listed as operating on shell polymorphism at a local scale. This decision, partially based on the extensive work of Guerrucci-Henrion (1966) in two zones of southern Brittany, could appear to be an excessive simplification of the reality but is easily justified in two ways. Firstly, as regards regional trends of polymorphism, few environmental pressures acting at fine scale are also liable to spatial structuring at a larger scale, ie, on the basis of a large sample of human-modified habitats with approximately the same vegetation background and microtopography. Moreover, local ecological differences between sites in Brittany are often related to the proximity of the sea so, its potential multi-way influence is largely integrated in the factor 'DL'. Secondly, whatever the region studied, its seems illusory to attempt a factor by factor decomposition of a variation created and maintained by the combination of a great number of phenomena, some of them being perhaps unknown until now, some others not clearly expressed in phenotype frequencies because field populations are seldom at equilibrium (Cook, 1998). Our goal is only to provide a synthetic and accurate picture of geographical structure which will potentially lead to further insight into the main classes of evolutionary pressures. In this way, among the numerous techniques available to analyze data variability (see Guiller *et al*, 1998; Sokal *et al*, 1997, 1998; Durand *et al*, 1999 for critical reviews), combining GPCA and RDA is especially appropriate here because of a partitioning of the total variance of morph frequencies, leading to geographical maps from which noisy variation in the initial data set is smoothed and the respective parts of each cate-

gory of pressures clearly delineated. Thus, the level curve representation of site factorial scores related to the first GPCA suggests that the decrease of Moran's *I* in the first part of the correlogram on Figure 5 should be partially due to clinal gradients from littoral to inland including the influence of altitude, with similar morph frequencies in neighbouring populations (highly significant positive values of *I*), ie, located at the same distance from the coast and/or at the same altitude. This possibly adaptive phenomenon finds its expression mainly in an increase of pink frequency at the expense of yellow with the distance from littoral (axis 1) but also, in quite limited directions, an increase of unbanded shells according to an increase of altitude (axis 2). The first result was observed by Guerrucci-Henrion (1966), Harvey (1971) and Mazon *et al* (1987) and related to an increase of rainfall or atmospheric humidity from littoral to inland. The second one was not reported, essentially because of an inadequate sampling of the whole region by Guerrucci-Henrion. The second part of the correlogram [110–170 km] is characterized by random fluctuations of *I* around 0 because distance classes include pairs of contrasted sites (littoral *vs* inland) together with northern *vs* southern littoral populations which have homogeneous phenotype frequencies. Furthermore, spatial structure mapped in Figure 1 shows a strong opposition between sites located at the ends of an axis NW-SE, which explains highly significant negative values of *I* in the third part of the correlogram ([190–230 km]). However, the final increase of *I*, which concerns the last distance class, was quite surprising but should be carefully considered because this class includes only 17 paired-populations. This *a priori* conflicting pattern would be enhanced if the sites from south-east of Brittany were considered, because of a sharp increase of yellow morph frequencies in this zone, without a corresponding change in environmental features.

After removing environmental effects (DL, altitude), differences between maps before and after RDA for each morph are all the more pronounced since the variance of the involved phenotype was environmentally induced. However, the spatial structure remains significant and is based on almost equivalent contributions of each morph, excluding Y5B which is never really structured at this scale. Figure 5 shows that this significant pattern results for the most part from the persistence of positive correlations at small distances, even if all *I* values strongly decrease after RDA, especially for Y1B and P0B (Table 1). More precisely, assuming that environmental pressures were removed, the remaining spatial structure highlights an increasing phenotypic differentiation with physical isolation of colonies within a range of about 30 km, but our experiment is obviously not appropriate to test for an isolation by distance at this scale. As a meaningful example, a sharp contrast (high local variance) was identified in the zone of 'Cap Fréhel' (Figure 4) which is related to a strong increase of one-banded shells, particularly of Y1B, in colonies located in an area of approximately 20 km width. However, the processes which led to such a pattern remain unclear.

After 30 km, residual frequencies seem to randomly fluctuate from one colony to the other (Figure 5), which is not really confirmed by examination of factorial scores surfaces which show many homogeneous zones but they are undetectable with such a correlogram (Figure 4). More generally, we can point out that usual one-dimen-

sional correlograms are unable to distinguish data similarities arising from a fixed structure (eg, area effect or selection gradient – see Sokal *et al*, 1997) from those resulting in a real isolation by distance and, obviously, to detect an eventual anisotropic effect of the structuring factor. Mapping factorial coordinates of colonies according to a maximization of global correlation (GPCA/zones of homogeneous morph frequencies) or local variance (LPCA/zones of abrupt transition – results not shown) after RDA, leads chiefly to the detection of phenotypic area effects, following the general definition by Jones *et al* (1980), Woodruff and Gould (1980) and Gould and Woodruff (1990). There is historical evidence for *Cepaea* in Brittany of strongly isolated groups of populations, often of small size, which are established in open habitats with recurrent anthropogenic disturbances. Landscape perturbations by man may operate at different time and space scales: ie, (i) ancient destructions of woodlands or, more recently, of hedgerows which concern all the region; and (ii) present microgeographical events like seasonal burnings or chemical treatments, but their main consequence is always a strong reduction of population size, sometimes an extinction. Then, at different scales, founder events followed by a slow neighbourhood diffusion leading today to secondary contacts of phenotypically differentiated zones seem to be a suitable explanation for the patterns observed in Brittany, as in other regions (see Cameron *et al*, 1980; Cameron and Dillon, 1984; Cameron and Pannett, 1985). Whatever their historical origin (bottleneck, colonization of a new area), founder events may give rise for the most part to the strong local variance observed in all our analyses. However, we have to retain that all the environmental sources of phenotypic variation have not been identified.

The results which we have discussed here in a preliminary way highlight the great part of random processes in spatially modeling shell polymorphism in Breton colonies of *Cepaea nemoralis*. The data analysis steps that were used allow an integrative approach of geographical variability of polychromatism, because of the interpretation of spatial trends derived from 'multivariate phenotypes' and consideration of only the main dichotomy of evolutionary pressures into environmental *vs* random. For the same region, random patterns must now be deeply investigated in the light of LPCA results and confronted with those related to the sibling species *Cepaea hortensis*. Multivariate analyses will also allow a comparative study of data sets concerning regions with contrasted environments (selective pressures, landscape components) and/or different past history.

References

Cain AJ (1983). Ecology and ecogenetics of terrestrial molluscan populations. In: Russel-Hunter WD (ed) *The Mollusca*, Academic Press: New York, pp 597–647.

Cain AJ, Currey JD (1963). Area effects in *Cepaea. Philos Trans R Soc London [B]* **246**: 1–81.

82

Cain AJ, Cook LM, Currey JD (1990). Population size and morph frequency in a long-term study of *Cepaea nemoralis*. *Proc Roy Soc Lond* **240**: 231–250.

Cameron RAD, Dillon PJ (1984). Habitat stability, population histories and patterns of variation in *Cepaea*. *Malacologia* **25**: 271–291.

Cameron RAD, Pannett DJ (1985). Interaction between area effects and variation with habitat in *Cepaea*. *Biol J Linn Soc* **24**: 365–379.

Cameron RAD, Carter MA, Palles-Clark MA (1980). *Cepaea* on Salisbury Plain: patterns of variation, landscape history and habitat stability. *Biol J Linn Soc* **14**: 335–358.

Clarke B (1966). The evolution of the morph-ratio clines. *Am Nat* **100**: 389–402.

Clarke B, Arthur W, Horsley DT, Parkin DT (1978). Genetic variation and natural selection in pulmonate snails. In: Fretter J, Peake J (eds) *The Pulmonates*, Academic Press: New York, pp 220–270.

Cleveland WS (1979). Robust locally weighted regression and smoothing scatterplots. *J Am Statist Assoc* **74**: 829–836.

Cliff AD, Ord JK (1973). *Spatial Autocorrelation*. Pion: London.

Cook LM (1991). *Genetic and Ecological Diversity: The Sport of Nature*. Chapman & Hall, London.

Cook LM (1998). A two-stage model for *Cepaea* polymorphism. *Philos Trans R Soc London [B]*, **353**: 1577–1593.

Cook LM, Pettitt WA (1998). Morph frequencies in the snail *Cepaea nemoralis*: changes with time and their interpretation. *Biol J Linn Soc* **64**: 137–150.

Cook LM, Cowie RH, Jones JS (1999). Change in morph frequency in the snail *Cepaea nemoralis* on the Marlborough Downs. *Heredity* **82**: 336–342.

Cowie RH, Jones JS (1998). Gene frequency changes in *Cepaea* snails on the Marlborough Downs onver twenty-five years. *Biol J Linn Soc* **65**: 233–255.

Davison A, Clarke B (2000). History or current selection? A molecular analysis of 'area effects' in the land snail *Cepaea nemoralis*. *Proc Roy Soc Lond* **267**: 1399–1405.

Durand JD, Guinand B, Bouvet Y (1999). Local and global multivariate analysis of geographical mitochondrial DNA variations in *Leuciscus cephalus* L. 1758 (Pisces: Cyprinidae) in the Balkan Peninsula. *Biol J Linn Soc* **67**: 19–42.

Gerdeaux D (1978). Le polymorphisme de *Cepaea nemoralis* en Provence : ses relations avec le milieu. *Arch Zool Exp Génér* **119**: 565–584.

Goodhart CB (1987). Why are some snails visibly polymorphic, and others not ?. *Biol J Linn Soc* **31**: 35–58.

Gould SJ, Woodruff DS (1990). History as a cause of area effects: an illustration from *Cerion* on great Inagua, Bahamas. *Biol J Linn Soc* **40**: 67–98.

Green PJ, Sibson R (1977). Computing Dirichlet tesselations in the plane. *The Computer J* **21**: 168–173.

Guerrucci-Henrion MA (1966). Recherches sur les populations naturelles de *Cepaea nemoralis* en Bretagne. *Arch Zool Exp Génér* **107**: 369–546.

Guiller A, Bellido A, Madec L (1998). Ordination and genetic distances : the model of the landsnail *Helix aspersa* in Maghreb. *Syst Biol* **47**: 208–227.

Harvey PH (1971). Populations of *Cepaea nemoralis* from southwestern France and northern Spain. *Heredity* **27**: 353–363.

Heywood JS (1991). Spatial analysis of genetic variation in plant populations. *Annu Rev Ecol Syst* **22**: 335–355.

Jones JS, Leith BH, Rawlings P (1977). Polymorphism in *Cepaea*: a problem with too many solutions? *Annu Rev Ecol Syst* **8**: 109–143.

Jones JS, Selander RK, Schnell GD (1980). Patterns of morphological and molecular polymorphism in the land snail *Cepaea nemoralis*. *Biol J Linn Soc* **14**: 359–387.

Khemici E (1987). *Recherches sur la variabilité phénotypique des populations de Cepaea nemoralis* (L.) et *Cepaea hortensis* (M.) (Gastéropodes Pulmonés) de Normandie. Thèse 3e cycle, Paris VI.

Lamotte M (1951). Recherche sur la structure génétique des populations naturelles de *Cepaea nemoralis* L. *Bull Biol Fr* **35** (Suppl): 1–239.

Lamotte M (1988). Facteurs influençant la diversité du polymorphisme de la coquille dans les populations naturelles de *Cepaea nemoralis* Moll. *Haliotis* **18**: 131–157.

Lebreton JD, Sabatier R, Banco G, Bacou AM (1991). Principal component and correspondence analysis with respect to instrumental variables: an overview of their role in studies of structure – activity and species – environment relationships. In: Devillers J, Karcher W (eds) *Applied Multivariate Analysis in SAR and Environmental Studies*, ECSC, EEC, EAEC: Netherlands, pp 84–114.

Mazon LI, de Pancorbo MAM, Vicario A, Aguirre AI, Estomba A, Lostao CM (1987). Distribution of *Cepaea nemoralis* according to climatic regions in Spain. *Heredity* **58**: 145–154.

Murray J, Clarke B (1978). Changes of gene frequency in *Cepaea nemoralis* over 50 years. *Malacologia* **17**: 317–330.

Ratel MO (1987). *Recherches sur la variabilité phénotypique des populations de Cepaea nemoralis* (L.) et *Cepaea hortensis* (M.) (Gastéropodes Pulmonés) de la région parisienne. Thèse 3e cycle, Paris VI.

Sabatier R, Lebreton JD, Chessel D (1989). Principal component analysis with instrumental variables as a tool for modelling composition data. In: Coppi R, Bolasco S (eds) *Multiway Data Analysis*, Elsevier Sciences Publ: North Holland, pp 341–352.

Sokal RR, Oden NL (1978). Spatial autocorrelation in biology. 1. Methodology. *Biol J Linn Soc* **10**: 199–228.

Sokal RR, Oden NL, Thomson BA (1997). A simulation study of microevolutionary inferences by spatial autocorrelation analysis. *Biol J Linn Soc* **60**: 73–93.

Sokal RR, Oden NL, Thomson BA (1998). Local spatial autocorrelation in biological variables. *Biol J Linn Soc* **65**: 41–62.

Thioulouse J, Chessel D, Champely S (1995). Multivariate analysis of spatial patterns: a unified approach to local and global structures. *Environ Ecol Stat* **2**: 1–14.

Thioulouse J, Chessel D, Dolédec S, Olivier JM (1997). ADE-4: a multivariate analysis and graphical display software. *Statist Comput* **7**: 75–83.

Wollenberg AL, van den (1977). Redundancy analysis. An alternative for canonical correlation analysis. *Psychometrika* **42**: 207–219.

Woodruff DA, Gould SJ (1980). Geographic differentiation and speciation in *Cerion* a preliminary discussion of patterns and processes. *Biol J Linn Soc* **14**: 389–416.