THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1156

History or current selection? A molecular analysis of 'area effects' in the land snail *Cepaea nemoralis*

Angus Davison* **and Bryan Clarke**

Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

We have used molecular variation in microsatellite and mitochondrial DNA to throw light on the origins of enigmatic geographical patterns, known as 'area effects', in the shell polymorphisms of the land snail *Cepaea nemoralis*. Our aim was to assess the relative importance of recent selection and historical events in the formation of these patterns. On the Marlborough Downs in Wiltshire, England, the 'type locality' for area effects, the frequencies of microsatellites are significantly associated with the frequencies of alleles for shell banding. A less clear association is found between microsatellites and shell colour. Mitochondrial haplotypes show no significant relationships. Although the correlated geographical patterns could be the results of random genetic drift from an initially uniform array of populations, the magnitudes of the patterns, and of the correlations between them, seem too strong to have arisen by drift since the last glaciation. Our results suggest that invasions from refugia have been the most important factors in forming area effects.

Keywords: evolution; isolation-by-distance; microsatellite; phylogeography; polymorphism; pulmonate

1. INTRODUCTION

It is important to understand how geographical patterns of gene frequencies arise and are maintained. In order to do so, we must separate the effects of historical factors, such as drift and selection in refugia, from more immediate explanations, such as current selective regimes.

Studies on land snails have clarified the relative roles of chance and adaptation in evolution (Cain & Sheppard 1950; Cain & Currey 1963; Ochman *et al*. 1983; Gould & Woodruff 1990; Johnson *et al*. 1993; Cook 1998). It was once thought that genetic variation in the colour and banding of snail shells was entirely random and non-adaptive (Dobzhansky 1937; Mayr 1942; Lamotte 1951). Then Cain & Sheppard (1950, 1952, 1954) studied *Cepaea nemoralis* (L.), and reported that in some lowland areas of Britain there are clear correlations between the frequencies of shell phenotypes, the kinds of habitat in which they live, and the effects of natural selection by predators, notably song thrushes.

Later, Cain & Currey (1963) found that on chalk downlands, where song thrushes are relatively rare, particular colour and banding alleles predominate over large areas, apparently regardless of the habitat. Groups of similar populations are separated from other such groups by steep gradients of gene frequencies, over a few hundred metres, and often in places where the habitat seems to be uniform. Cain & Currey attributed these 'area effects' to selection by cryptic features of the microclimate. Since then, other explanations have been suggested. They include genetic drift and co-adaptation (Goodhart 1963), the selective accumulation of modifiers in morph-ratio clines (Clarke 1966), and differentiation in refugia produced by changing uses of land (Cameron *et al*. 1980). Area effects have now been found in several other species of snails (see, for example, Clarke 1968; Gould & Woodruff 1990; Johnson *et al*. 1993).

*Author for correspondence (a.davison@hgmp.mrc.ac.uk).

While the area effects on the downlands of Wiltshire and Berkshire persist 25 years after they were first described (Cowie & Jones 1998; Cook *et al*. 1999), their origins are still matters for debate. A few workers have looked for geographical correlations between the frequencies of allozymes and those of shell genotypes. Johnson (1976) found that an area effect in Berkshire 'reflects substantial genetic differentiation', but Ochman *et al*. (1983) reported that the frequencies of allozymes in Pyrenean *C. nemoralis* correlate neither with shell genotypes nor with climate or vegetation. Similarly, Jones *et al*. (1980) found no correlation between allozymes and area effects on a sand-dune in north Wales. Ochman and his colleagues suggested that the geographical patterns of allozymes in the Pyrenees reflect the isolation of snails in Pleistocene refugia, whereas the patterns of shell genotypes have become locally adjusted to modern environments.

In the present study we examine mitochondrial and microsatellite sequences, using techniques developed by Terrett *et al*. (1996), Thomaz *et al*. (1996) and Davison (1999, 2000*a*), to investigate the origins of area effects in *C. nemoralis*. We compare the geographical patterns in the DNA sequences with those in shell characters, using populations from the 'type locality' for area effects, the Marlborough Downs in Wiltshire, England.

2. MATERIAL AND METHODS

(a) *Sample collection*

The study covered an area of 'high' chalk downland, the Marlborough Downs, about 3 km × 6 km in extent and reaching an altitude of 270 m (above sea level), situated between Swindon and Marlborough in Wiltshire (figure 1). It includes many of the sites studied by Cain & Currey (1963). Samples of *C. nemoralis* were collected in April 1997 and June 1998. Each sample was taken from as small an area as possible, usually less than 10 m × 10 m. Snails were collected at five other sites: Anglesey (Wales), Dovedale (central England), Kilmartin (Scotland), Hepscott (north-eastern England) and Mullaghmore (Ireland).

Received 20 January 2000 *Accepted* 23 March 2000

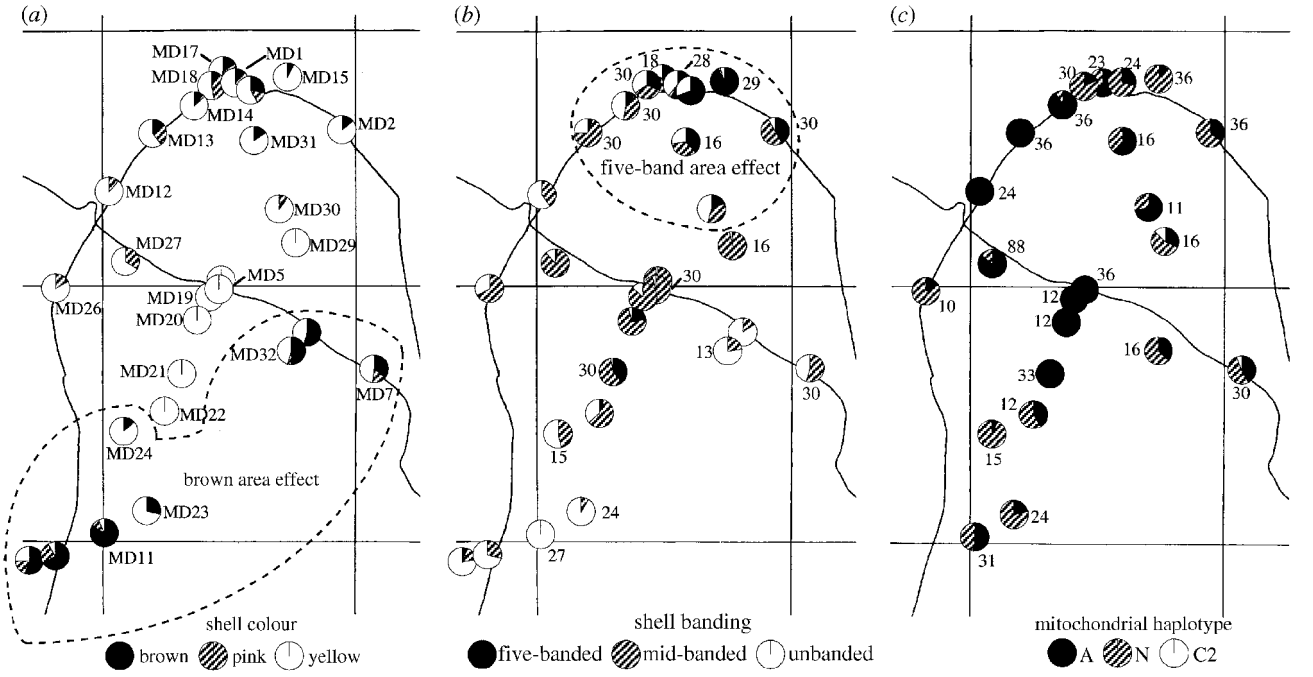

Figure 1. Maps of the Marlborough Downs for shell colour (*a*), shell banding (*b*) and mitochondrial haplotype (*c*). Each grid square corresponds to 3 km × 3 km. The dotted lines show the location of the area effects described by Cain & Currey (1963). The names of sites that were used for DNA analysis are shown in (*a*), the sample size used for the microsatellite analysis in (*b*), and the sample size used for the mitochondrial analysis in (*c*).

The positions of the sites on the Marlborough Downs are shown in figure 1*a*, and the sample sizes used in the analyses of microsatellites and mitochondria are shown in figure 1*b*,*c*, respectively. The grid references of the sites are given in the electronic Appendix A, which can be found at The Royal Society Web site.

(b) *DNA isolation, amplification and electrophoresis*

DNA was extracted from the snails using Nucleon[1] Phytopure kits (Nucleon Biosciences, Scotlab Ltd, Scotland). Methods for the amplification and description of the mitochondrial haplotypes are given by Davison (2000*b*). Five sets of microsatellite primers were amplified (Davison 1999). Allelic lengths were determined by reference to control samples and to a DNA-sequencing ladder.

(c) *Statistical analysis*

The first method used for relating the microsatellite data to area effects was a 'constant row total multiple correspondence analysis' (CRT-MCA) in the software package ADE-4 (Thioulouse *et al.* 1997). CRT-MCA is a way of displaying multivariate gene frequencies. The parameters are related to *F*-statistics, although pooling of populations is not needed. All loci are considered simultaneously, and the contribution of each locus to population differentiation can be estimated. CRT-MCA has been used to study phases of colonization in populations of *Bufo*, and genetic differentiation in a zone of hybridization between taxa of *Sorex* (Guinand 1996; Guinand & Easteal 1996; Lugon-Moulin *et al.* 1999).

A second analysis, linear multidimensional scaling (SYSTAT v. 8, SPSS Science), was used to summarize the microsatellite data on two axes. This method is related in function to principal-component analysis, but has two advantages for our purposes. It can usually fit a suitable model in fewer dimensions, and it does not require a linear relationship between calculated distances and observed dissimilarities. The goodness-of-fit statistic that the procedure sets out to minimize is the stress, a low value (< 0.15) indicating a good fit (Borg & Groenen 1997). In our analysis a Pearson distance (correlation) matrix of allele frequencies was entered. The success in reducing the data to two axes was judged not only by the stress, but also by Shepard diagrams (plots of distances derived from the multidimensional scaling (MDS) against the observed distances).

GENEPOP v. 3.1 (Raymond & Rousset 1995) was used to test for deviations from Hardy–Weinberg expectations at microsatellite loci, and for evidence of linkage disequilibrium.

Methods of estimating differences between pairs of samples at microsatellite loci assume either infinite alleles or stepwise mutations. Weir & Cockerham's (1984) estimator of F_{ST}, θ, assumes infinite numbers of possible alleles, and is most often used for the analysis of allozyme frequencies. R_{ST} (Slatkin 1995; modified by Goodman (1997)) assumes stepwise mutations, and has been considered more suitable for microsatellites. In this study both estimators were calculated because it is not clear which model is appropriate for the microsatellites of *C. nemoralis*. F_{ST} was estimated using GENEPOP v. 3.1 (Raymond & Rousset 1995) and R_{ST} using R_{ST}-calc (averaging variance components) (Goodman 1997). Departures from zero were investigated using the permutation tests in the two packages.

The analysis of regression was used to detect associations between distances based on DNA, those based on shell phenotypes, and those based on geography (Manly 1997). Significance was tested by randomizing particular matrices using the program RT (Manly 1997; Mantel 1967). Geographical distances were transformed to logarithms, and genetic distances were either untransformed or transformed by $D/(1-D)$, where D is the untransformed distance (F_{ST} or R_{ST}) as recommended by Rousset (1997).

All multiple tests of significance were corrected by the 'sequential Bonferroni' method (Rice 1989).

3. RESULTS

(a) *Shell polymorphisms*

Thirty-two populations from the Marlborough Downs were scored for shell colour and banding, using the criteria of Cain & Currey (1963). The frequencies of yellow, pink, brown, unbanded, mid-banded and five-banded snails were mapped (figure 1*a*,*b*) and compared with the area effects described by Cain & Currey. In agreement with both the original survey and a more recent one (Cowie & Jones 1998), we found that the brown morph is common in the south of the region. Elsewhere it is mostly absent, except for a patch in the north where Cain & Currey found frequencies up to 25% and we found frequencies up to 28%. In this region, distinguishing between brown and pink morphs is unusually difficult. Five-banded snails occur almost exclusively in the north, except for a few colonies in the centre of the region (see figure 1*b*) where there are snails with very faint bands, often only visible near the lip. Between the northern 'five-banded' area and the southern 'brown' area there is a third area in which yellow mid-banded shells predominate.

(b) *Microsatellites*

Sixteen populations from the Marlborough Downs and five from the rest of Britain and Ireland were scored at five compound microsatellite loci. There were significant deficiencies of heterozygotes within populations in 16 out of 96 tests (involving four different loci in 14 populations). The heterozygote deficiencies could have arisen if the sampling areas had exceeded those of Wrightian 'neighbourhoods', or because of rare null alleles. There is no consistency between loci in departures from Hardy–Weinberg within any single population, and therefore the second explanation is favoured over the first. It is possible that the null allele problem may have arisen due to occasional poor amplification of the larger allele. Significant associations within populations between loci (whether microsatellites, mitochondria, or shell characters) were found in six out of 450 comparisons. Of these, one was between two microsatellite loci, two were between two shell loci (for colour and banding, which are genetically linked), and three were between the two kinds of loci. None involved mitochondria.

Using the CRT-MCA, we projected the first two axes, together representing 36.6% of the variation in microsatellites, on to a map (figure 2). The first axis, accounting for 20.7% of the variation, corresponds rather closely to the five-banded area effect (figure 2*a*). Neighbouring colonies differ more when they come from separate area effects than they do when they come from the same one. The second axis, representing 15.9% of the variation, separates off the brown area (figure 2*b*), but also includes three populations from the five-banded area (MD1, MD2 and MD31). The sharpest transition is between the brown area and the adjacent populations of MD21, MD5 and MD29. Four of the five microsatellite loci contribute almost equally to the first two axes of the CRT-MCA. The least polymorphic locus (*Cne*15) hardly contributes at all.

With multidimensional scaling, and including all the populations in the British Isles (stress = 0.13), the first axis indicated that the populations on the Marlborough Downs are distinct from all the other populations in Britain and Ireland except that in Dovedale, geographically the nearest. When the analysis was restricted to the Marlborough Downs (stress = 0.10) the two axes together

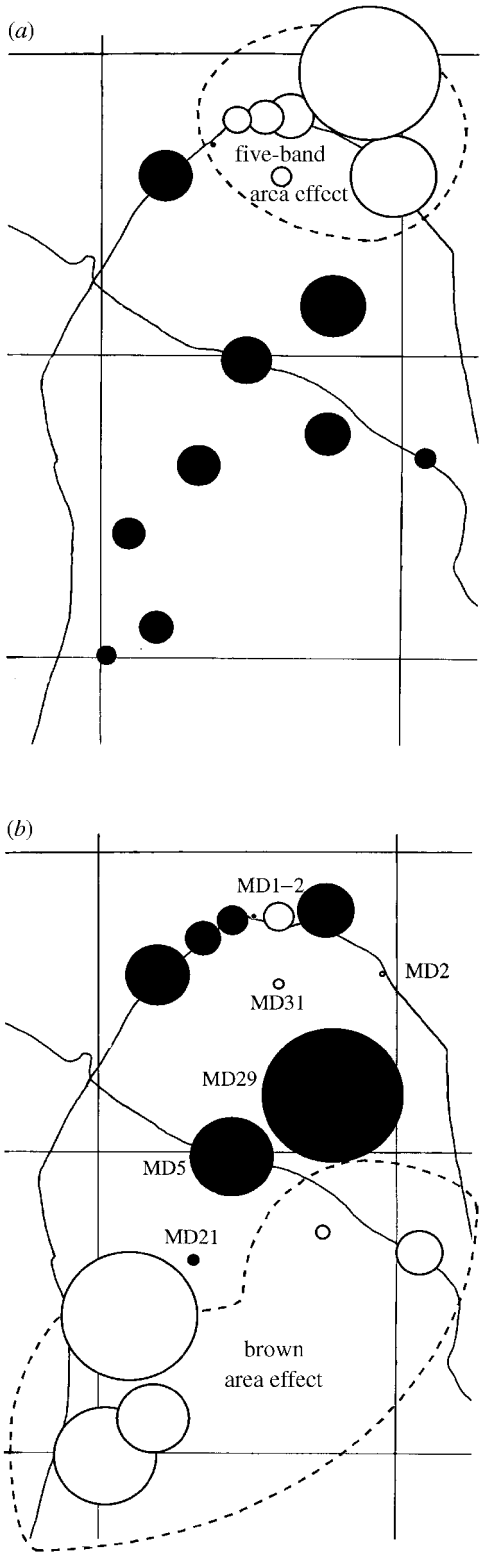

Figure 2. The first two axes from the CRT-MCA projected onto the Marlborough Downs. Open circles show positive values, filled circles show negative values, and the size of the circle is proportional to the magnitude. Axis 1, 20.7% of the variation (*a*); axis 2, 15.9% of the variation (*b*).

Table 1. *Estimates of* F_{ST} (*Weir & Cockerham 1984*) *and* R_{ST} (*Slatkin 1995; Goodman 1997*) *at each locus and across all loci on two spatial scales: within the Marlborough Downs, and between the Marlborough Downs and the other populations from the British Isles. A sample population from the Marlborough Downs* (*MD1*) *was used for the latter comparison*

	within Marlborough Downs		between Marlborough Downs and others	
locus	F_{ST}	R_{ST}	F_{ST}	R_{ST}
*Cne*1	0.141	0.123	0.238	0.660
*Cne*6	0.095	0.169	0.234	0.538
*Cne*10	0.144	0.121	0.486	0.659
*Cne*11	0.085	0.058	0.233	0.530
*Cne*15	0.091	0.112	0.543	0.584
all loci	0.117	0.117	0.351	0.595

separated all the populations in the five-banded area, except MD13, from all the populations in the other areas. Populations in the brown area were not discriminated (data not shown).

Estimates of genetic differentiation between populations (F_{ST} and R_{ST}) are given in table 1. Comparing populations within the Marlborough Downs, the two measures are roughly equal. Comparing populations outside the Marlborough Downs, or between the Marlborough Downs and the rest, R_{ST} values exceed F_{ST} values, particularly in comparisons between distant colonies.

Plots of genetic distance against geographical distance were made for each microsatellite locus, and for all loci combined (figure 3). Whether F_{ST} or R_{ST} was used, there were significant positive associations at *Cne*6, *Cne*10, and all loci combined (table 2). Transformed and untransformed data gave similar results. In all cases, the relationships were nonlinear, geographically distant populations being genetically more similar than expected. This indicates a patchiness in the occurrence of the microsatellite alleles (see Endler 1977). The patches seem to be correlated with the area effects, but not exactly coincident with them.

We looked directly for correlations between genetic distances based on microsatellites and those based on shell phenotypes. There were highly significant associations between distances based on each of four microsatellite loci and distances based on shell banding (divided into 'unbanded', 'mid-banded' and 'five-banded'; see table 3). Only one microsatellite locus (*Cne*15) gave distances that were significantly associated with those for shell colour.

Since shell banding is controlled by at least two unlinked loci, one (*B*) determining the presence or absence of bands, and the other (*M*) converting five-banded to mid-banded (Cain *et al*. 1960), we also tested

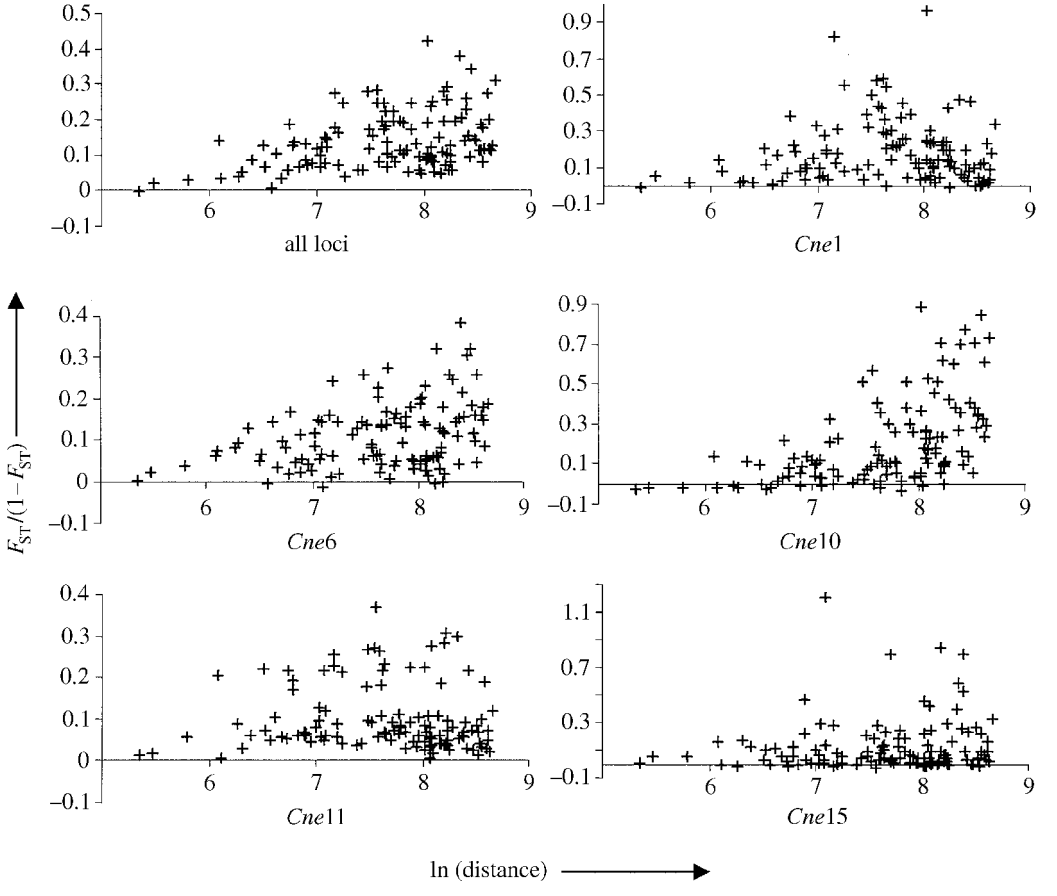

Figure 3. Plots of geographical distance between populations on the Marlborough Downs against F_{ST} for all loci combined, and for individual loci. Genetic distance was transformed by $F_{ST}/(1-F_{ST})$. Similar regressions were obtained using untransformed F_{ST}, or R_{ST}.

Table 2. *Regression coefficients for genetic distance estimators (as dependent variables) against the geographical distance transformed to logarithms. Genetic distance was transformed into $D(1-D)$, where D is F_{ST} or R_{ST}, following the recommendations of Rousset (1997)*

(Table-wide sequential Bonferroni methods (Rice 1989) were used to test for significance.)

marker type	estimator	loci	regression coefficient (b)	probability b (one-tailed)	significance level
mtDNA	F_{ST}	—	0.144	0.3260	—
microsatellites	F_{ST}	all loci	0.042	0.0001	**
		*Cne*1	0.015	0.2431	—
		*Cne*6	0.039	0.0002	**
		*Cne*10	0.156	0.0001	**
		*Cne*11	−0.004	0.3414	—
		*Cne*15	0.029	0.1185	—
	R_{ST}	all loci	0.085	0.0001	**
		*Cne*1	0.001	0.4934	—
		*Cne*6	0.233	0.0003	**
		*Cne*10	0.160	0.0001	**
		*Cne*11	0.010	0.3301	—
		*Cne*15	0.043	0.0311	*

*$p < 0.05$, **$p < 0.01$.

correlations between microsatellite distances and those at each banding locus individually. Highly significant associations were found using all the microsatellite loci combined (table 3). Individually, *Cne*6 and *Cne*10 were associated with the *B* locus, whereas *Cne*l, *Cne*10 and *Cne*11 were associated with the *M* locus. Note that *Cne*10 and *Cne*11 are probably linked genetically to the *M* locus, the estimated frequencies of recombination being 22.6 and 37.5%, respectively (Davison 1999).

(c) *Mitochondria*

Mitochondrial haplotypes were studied in 23 populations from the Marlborough Downs. Both British lineages, *A* and *N* (Terrett *et al*. 1996; Thomaz *et al.* 1996; Davison 2000*a*), occur at most sites. The exceptions are in the centre of the region, where only the *A* type was found (figure 1*c*). This 'area effect' for a mitochondrial haplotype appears to be very roughly coincident with the yellow mid-banded area, but the association is not formally significant.

4. DISCUSSION

Three major area effects were observed: one with an excess of five-banded snails in the north, one with an excess of yellow mid-banded snails in the centre, and one with an excess of brown snails in the south (figure 1*a*,*b*). The patterns are essentially the same as those described by Cain & Currey (1963) and by Cowie & Jones (1998, based on a survey carried out in 1985).

The multiple correspondence analysis of microsatellites divided the Marlborough Downs into its major area effects, the first axis separating the five-banded area (figure 2*a*), and the second the brown area (figure 2*b*). The sharp clines in the frequencies of shell genotypes between area effects were reflected by similar clines in the frequencies of microsatellite alleles. It is clear that particular arrays of microsatellites occur in patches, and that these patches coincide closely, although not exactly, with the area effects. Because the frequencies of microsatellites are probably not altered locally by natural selection, the coincidence immediately suggests that the two forms of variation reflect a common pattern of historical events, for example in the founding of the populations (Ibrahim *et al.* 1996). We must, however, consider other possible explanations.

The correlations between genetic distances and geographical distances might be the results of migration and drift acting on an array of populations that were initially homogeneous (Wright 1946). *C. nemoralis* arrived in Britain in the late Holocene, *ca.* 10 000 years before the present (Kerney *et al*. 1980) so that there is a limited time in which the area effects could have arisen. Endler (1977) made some computer simulations of migration and drift in a two-dimensional 'stepping stone' model, and observed that they produced area effects in fewer than 1000 generations (for *C. nemoralis* roughly 3000 years), although they were neither as extensive nor as different from each other as those on the Marlborough Downs. Endler's model suggests that scattered barriers to gene flow could generate larger and stronger effects.

It is possible that migration and drift have produced two sets of area effects, one in shell genotypes and the other in microsatellites. If so, they could show coincident clines, and significant correlations in gene frequencies. Random processes often produce such correlations when the patches are large in relation to the spaces between samples (i.e. where the values in the samples are associated geographically), but Endler's results suggest that the correlations would be weak unless other factors, such as barriers to gene flow, were involved.

A second possibility is that the frequencies of shell genotypes are the results of cryptic selection, as suggested by Cain & Currey (1963), but that the microsatellites vary through drift and migration. Once more the correlations between them could be accidental. This explanation, however, is unsatisfactory because it multiplies hypotheses, particularly when the loci are unlinked. If one pattern of variation is due to drift and migration, why not the other?

Table 3. *Regression coefficients for microsatellite distance estimators (as dependent variables) against estimates of banding and colour distances based on phenotype and genotype (italics)*

(Table-wide sequential Bonferroni methods (Rice 1989) were used to test for significance.)

		banding			colour (= *C* locus)		
marker type	loci	regression coefficient (*b*)	probability *b* (one-tailed)	significance level	regression coefficient (*b*)	probability *b* (one-tailed)	significance level
mtDNA	—	0.086	0.1404	—	0.010	0.4577	—
microsatellites	all loci	0.134	0.0001	**	0.016	0.2721	—
	*Cne*1	0.161	0.0001	**	− 0.030	0.2563	—
	*Cne*6	0.059	0.0043	**	− 0.003	0.4502	—
	*Cne*10	0.265	0.0001	**	0.012	0.4153	—
	*Cne*11	0.117	0.0001	**	0.002	0.4693	—
	*Cne*15	0.049	0.1175	—	0.139	0.0036	**
		B locus			*M* locus		
mtDNA	—	0.004	0.4748	—	0.005	0.4757	—
microsatellites	all loci	0.048	0.0066	**	0.078	0.0001	**
	*Cne*1	− 0.015	0.3249	—	0.228	0.0001	**
	*Cne*6	0.055	0.0018	**	0.102	0.0822	—
	*Cne*10	0.148	0.0001	**	0.226	0.0005	**
	*Cne*11	0.015	0.2276	—	0.109	0.0001	**
	*Cne*15	0.047	0.1024	—	0.002	0.4425	—

** $p < 0.01$.

In practice it is extremely difficult to test whether a particular concordance is due to selection or drift (or both), because the statistics depend so strongly on assumptions about the history of the populations. For example, we could retrospectively postulate barriers to explain by random events almost any pattern of frequencies. Barriers would promote concordance between loci whatever the causes of differentiation. The only way unequivocally to show natural selection experimentally is to demonstrate a clear mechanistic connection between changes of gene frequency and the action of the selective agent. While this has been done for climatic selection on shell colour and banding in *Cepaea*, the connection between climate and area effects remains insecure, with the possible exception of some areas on the tops of mountains (Jones 1973; Heath 1975; Clarke *et al*. 1978).

Whatever the difficulties, there is a generalization that we believe to be valid; if one set of loci is selectively neutral with respect to the ecological or other differences between localities, then the higher the concordance, the stronger is the argument that the area effects reflect a common history. This history may have been rooted in the shared occupation of a region surrounded by barriers to gene flow (Cameron *et al*. 1980), or in a shared pattern of colonization (Ibrahim *et al*. 1996; Hewitt 1996). The methods used here cannot discriminate between them, but both histories involve differentiation in refugia. The present concordances are strong (figure 2), and so they argue strongly for a historical explanation.

The data offer some hints about history. Two microsatellite loci are geographically associated with the *B* (banding) locus, but only one with the *C* (colour) locus (table 3). This is initially puzzling because *B* and *C* are tightly linked in a supergene. The explanation may be that some populations are mixed descendants of more than one founding event. Indeed the multivariate analysis suggests this. The second axis separates snails from the brown area, but also includes some northern populations containing browns, suggesting that the two groups share some of their ancestry.

The two divergent mitochondrial lineages that occur in Britain, the *A* and *N* types, have different geographical distributions, the *N* type predominating in eastern Britain and the *A* type predominating in the west (Davison 2000*a*). There is a zone down the centre of the country, including the Marlborough Downs, where both haplotypes occur together. The simplest explanation of this pattern is that there were two separate invasions, one spreading through eastern Britain, the other in the west. Their coexistence in the centre would then indicate a meeting and hybridization between the two sets of invaders. The rough geographical correspondence between haplotype *A* and yellow mid-banded snails on the Downs could well be a relic of invaders from the west with a particular combination of mitochondrial and nuclear genes.

We are grateful to Dr Sara Goodacre, Dr Christopher Wade, Dr John Brookfield and two anonymous reviewers for helpful discussions and comments on the manuscript. Chris Jiggins and Nic Flanagan helped collect snails on the Marlborough Downs. We thank the UK Natural Environment Research Council for support.

REFERENCES

Borg, I. & Groenen, P. 1997 *Modern multidimensional scaling: theory and applications*. New York: Springer.

Cain, A. J. & Currey, J. D. 1963 Area effects in *Cepaea*. *Phil. Trans. R. Soc. Lond.* B **246**, 1–81.

Cain, A. J. & Sheppard, P. M. 1950 Selection in the polymorphic land snail *Cepaea nemoralis*. *Heredity* **4**, 275–294.

Cain, A. J. & Sheppard, P. M. 1952 The effects of natural selection on body colour in the land snail *Cepaea nemoralis*. *Heredity* **6**, 217–231.

Cain, A. J. & Sheppard, P. M. 1954 Natural selection in *Cepaea*. *Genetics* **39**, 89–116.

Cain, A. J., King, J. M. B. & Sheppard, P. M. 1960 New data on the genetics of polymorphism in the snail *Cepaea nemoralis* (L.). *Genetics* **45**, 393–411.

Cameron, R. A. D., Carter, M. A. & Palles-Clark, M. A. 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 morph-ratio clines. *Am. Nat.* **100**, 389–402.

Clarke, B. 1968 Balanced polymorphism and regional variation in land snails. In *Evolution and environment* (ed. E. T. Drake), pp. 351–368. Yale University Press.

Clarke, B., Arthur, W., Horsley, D. T. & Parkin, D. T. 1978 Genetic variation and natural selection in pulmonate molluscs. In *The pulmonates* (ed. V. Fretter & J. Peake), pp. 219–270. London: Academic Press.

Cook, L. M. 1998 A two stage model for *Cepaea* polymorphism. *Phil. Trans. R. Soc. Lond.* B **353**, 1577–1593.

Cook, L. M., Cowie, R. H. & Jones, J. S. 1999 Change in morph frequency in the snail *Cepaea nemoralis* on the Marlborough Downs. *Heredity* **82**, 336–342.

Cowie, R. H. & Jones, J. S. 1998 Gene frequency changes in *Cepaea* snails on the Marlborough Downs over twenty-five years. *Biol. J. Linn. Soc.* **65**, 233–255.

Davison, A. 1999 Isolation and characterization of long compound microsatellite repeat loci in the land snail, *Cepaea nemoralis* L. (Mollusca, Gastropoda, Pulmonata). *Mol. Ecol.* **8**, 1760–1761.

Davison, A. 2000*a* An east–west distribution of divergent mitochondrial types in the land snail, *Cepaea nemoralis*. *Biol. J. Linn. Soc.* (In the press.)

Davison, A. 2000*b* The inheritance of divergent mitochondria in the land snail, *Cepaea nemoralis* (Mollusca: Gastropoda). *J. Mollusc. Stud.* **66**, 143–147.

Dobzhansky, T. 1937 *Genetics and the origin of species*. New York: Columbia University Press.

Endler, J. A. 1977 *Geographic variation, speciation, and clines*. Princeton University Press.

Goodhart, C. B. 1963 'Area effects' and non-adaptive variation between populations of *Cepaea* (Mollusca). *Heredity* **18**, 459–465.

Goodman, S. J. 1997 R_{ST}-calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Mol. Ecol.* **6**, 881–887.

Gould, S. J. & Woodruff, D. S. 1990 History as a cause of area effects: an illustration from *Cerion* on Great Inagua, Bahamas. *Biol. J. Linn. Soc.* **40**, 67–98.

Guinand, B. 1996 Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biol. J. Linn. Soc.* **58**, 173–195.

Guinand, B. & Easteal, S. 1996 Multivariate patterns of genetic differentiation support complex colonization schemes in *Bufo marinus* populations. *Evolution* **50**, 944–951.

Heath, D. J. 1975 Colour, sunlight and internal temperatures in the land-snail *Cepaea nemoralis* (L.). *Oecologia* **19**, 29–38.

Hewitt, G. M. 1996 Some genetic consequences of ice ages, their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276.

Ibrahim, K. M., Nichols, R. A. & Hewitt, G. M. 1996 Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**, 282–291.

Johnson, M. S. 1976 Allozymes and area effects in *Cepaea nemoralis* on the western Berkshire Downs. *Heredity* **36**, 105–121.

Johnson, M. S., Murray, J. & Clarke, B. 1993 The ecological genetics and adaptive radiation of *Partula* on Moorea. *Oxf. Surv. Evol. Biol.* **9**, 167–238.

Jones, J. S. 1973 Ecological genetics and natural selection in molluscs. *Science* **182**, 546–552.

Jones, J. S., Selander, R. K. & Schell, G. D. 1980 Patterns of morphological and molecular polymorphism in the land snail *Cepaea nemoralis*. *Biol. J. Linn. Soc.* **14**, 359–387.

Kerney, M. P., Preece, R. C. & Turner, C. 1980 Molluscan and plant biostratigraphy of some late Devensian and Flandrian deposits in Kent. *Phil. Trans. R. Soc. Lond.* B **291**, 1–43.

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

Lugon-Moulin, N., Brünner, H., Wyttenbach, A., Hausser, J. & Goudet, J. 1999 Hierarchical analyses of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora: Soricidae). *Mol. Ecol.* **8**, 419–431.

Manly, B. J. F. 1997 *Randomization, bootstrap and Monte Carlo methods in biology*, 2nd edn. London: Chapman & Hall.

Mantel, N. 1967 The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.

Mayr, E. 1942 *Systematics and the origin of species*. New York: Columbia University Press.

Ochman, H., Jones, J. S. & Selander, R. K. 1983 Molecular area effects in *Cepaea*. *Proc. Natl Acad. Sci. USA* **80**, 4189–4193.

Raymond, M. & Rousset, F. 1995 GENEPOP, version 1.2: population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.

Rice, W. R. 1989 Analyzing tables of statistical tests. *Evolution* **43**, 223–225.

Rousset, F. 1997 Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219–1228.

Slatkin, M. 1995 A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–462.

Terrett, J. A., Miles, S. & Thomas, R. H. 1996 Complete DNA sequence of the mitochondrial genome of *Cepaea nemoralis* (Gastropoda: Pulmonata). *J. Mol. Evol.* **42**, 160–168.

Thioulouse, J., Chessel, D., Doledec, S. & Olivier, J. M. 1997 ADE-4: a multivariate analysis and graphical display software. *Stat. Comp.* **7**, 75–83.

Thomaz, D., Guiller, A. & Clarke, B. 1996 Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proc. R. Soc. Lond.* B **263**, 363–368.

Weir, B. S. & Cockerham, C. C. 1984 Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Wright, S. 1946 Isolation by distance under diverse systems of mating. *Genetics* **31**, 39–59.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

An electronic appendix to this paper can be found at (http://www.pubs.royalsoc.ac.uk).