Application of ecological genetics techniques to test for selection by habitat on allozymes in *Cepaea nemoralis* (L.)

IAIN F. WILSON*

Department of Genetics, Queens Medical Centre, University of Nottingham, Nottingham NG7 2UH, U.K.

Techniques of ecological genetics were applied to the distribution of gene frequencies at seven polymorphic enzyme loci among populations of *Cepaea nemoralis* in lowland England in an attempt to detect evidence of natural selection operating at the molecular level. In Warwickshire, the frequencies of *MDH-1* alleles differed significantly, and those at *PEP-D* nearly so, between ancient woodlands and old open habitats. However, these differences were not repeated elsewhere. Overall, the pattern of gene frequencies at the molecular level was uncoupled from that at the morphological level, where there was evidence of natural selection on shell colour and banding patterns. The brown phenotype did not appear to be selected in relation to the visible background but, in Gloucestershire, was found to be at a higher frequency on the upper slopes than at lower altitudes along the Cotswold escarpment.

Keywords: allozymes, *Cepaea nemoralis*, ecological genetics, natural selection, neutral theory, shell phenotypes.

Introduction

The existence of enormous amounts of molecular genetic variation in populations of organisms presents the question of whether the variation is maintained by natural selection or by stochastic forces. Kimura (1968) originally developed his neutral theory to explain the diversity detected by starch-gel electrophoresis (Harris, 1966; Hubby & Lewontin, 1966; Lewontin & Hubby, 1966). The theory maintains that advantageous mutations are the exception and that molecular evolution occurs by random genetic drift. Nevertheless, given the differences between habitats in environmental parameters and the apparently deterministic nature of metabolic pathways, it would not be surprising if alternate alleles at a locus performed their metabolic functions at different rates or in different ways, especially in poikilothermic organisms. Such an hypothesis could explain why molecular genetic diversity persists in natural populations. The present paper describes the distribution of allozyme frequencies in the gastropod *Cepaea nemoralis* (L.) in four counties in lowland England and asks whether

habitat type is a factor in determining patterns of molecular genetic variation.

Natural selection through visual predation by the song thrush, *Turdus ericetorum*, can operate differentially between habitats on shell characters of *C. nemoralis* (Cain & Sheppard, 1954) and selection may maintain the polymorphism (Sheppard, 1951; Clarke, 1962). Although there are many examples in the literature of differences between habitats in phenotype frequencies in other species (e.g. Ford, 1975), there are only a few examples of differences at the molecular level (e.g. Koehn *et al.*, 1980; Watt *et al.*, 1983; Powers *et al.*, 1991), and the neutral theory has been generally accepted by population geneticists. However, two recent papers (McDonald, 1991; Tatarenkov & Johannesson, 1994) have reported further evidence of selective differences between allozymes at some loci, where frequencies are associated with components of the environment.

Enzyme variation in *C. nemoralis* has been investigated since the mid-1960s and the inheritance of variants at more than 20 loci has been confirmed (e.g. Brussard & McCracken, 1974; Johnson, 1979; Johnson *et al.*, 1984). However, none of this genetic variation has yet been found to be associated with habitat (e.g. Johnson, 1976; Ochman *et al.*, 1983). Instead there are often discontinuous distributions

*Present address: Fermentation Laboratory, Microbiology Department, University College, Galway, Ireland.

of allozyme frequencies analogous to area effects for visual characters (Cain & Currey, 1963a).

Area effects have variously been explained by local adaptation (Cain & Currey, 1963a,b), founding effects and coadaptation (Goodhart, 1963), genetic drift and local extinctions (Wright, 1978) and selection modifying morph-ratio clines (Clarke, 1966). Cameron and his colleagues, however, give a more comprehensive explanation based upon landscape history (Cameron *et al.*, 1980; Cameron & Dillon, 1984; Cameron & Pannett, 1985). They show that in regions with stable vegetation the snails show variation with habitat, whereas in less stable regions they show area effects.

In the present paper the effect of habitat on the distribution of molecular variation in *C. nemoralis* is compared with its effect on visible shell characters. As selective differences between habitats on allozymes may be slight, only habitats known to be long-established are compared. Pairs of samples from ancient woodlands and old open habitats were collected and analysed for consistent differences in their gene frequencies at seven allozyme loci and four shell colour and banding loci.

Materials and methods

Study areas and collecting of *C. nemoralis*

Ancient woodlands are rare in Britain and within them collecting adequate sample sizes of live individuals was difficult. Woodland samples were taken, where possible, deep within each wood to reduce any effect of migration from surrounding open habitats, and as far away as possible from grassy rides, or areas showing evidence of recent disturbance. If these ideal conditions could not be met then a collection was made in as typical a wooded area as could be found. Each open habitat sample was collected within 2 km of a woodland sample. Collecting was restricted to an area of 400 m² to ensure that snails came from within the panmictic unit of *C. nemoralis* (Lamotte, 1951). Only samples containing at least 20 live individuals were retained for analysis. An effort was made to collect only adult snails but this was not always possible. Juvenile *C. hortensis* were excluded on the basis of the distribution of dark pigment in the foot.

Samples were first collected in the summer of 1987 from populations in Lincolnshire, Warwickshire and Oxfordshire; one pair of samples was collected from Lincolnshire, eight from Warwickshire and eight from Oxfordshire. In the summer of 1989, after finding a significant association of allele

frequencies with habitat at one enzyme locus, and a near-significant association at another, pairs of samples were collected from populations in Gloucestershire.

Individuals were scored by the methods of Jones *et al.* (1977) for shell colour and banding phenotypes at four loci: *C* (^B, brown; ^P, pink; ^Y, yellow), *B* (^O, bands absent; ^B, bands present), *U* (³, midbanded; ⁻, unmodified) and *T* (⁰⁰³⁴⁵, upper two bands absent; ⁻, unmodified). In addition the phenotypes 'dark' and 'effectively unbanded' (EU) were scored (Cameron & Pannett, 1985). Because of dominance, heterozygotes at *C* could not be scored consistently and those at *B* were unscorable.

Habitat classification

The species of plants present at the sample sites were recorded along with the OS grid reference and altitude. Each wood was surveyed to obtain an overall impression of its diversity and to locate any areas of recent planting, past habitation or other major disturbance. The absence of dog's mercury (*Mercurialis perennis*) was used as an indicator of changes in land use or of locally acidic soils, which were highlighted by the presence of bracken (*Pteridium aquilinum*), bluebells (*Endymion non-scriptus*) and birch (*Betula pendula*).

Woodlands were classified by the methods of Rodwell (1991). The two community types that were commonly encountered are: (1) W8 [Ash-maple-dog's mercury (*Fraxinus excelsior*; *Acer campestre*; *M. perennis*). Subcommunities: a, *Primula vulgaris*-*Glechoma hederacea*; b, *Anemone nemorosa*-*Ranunculus ficaria*; c, *Deschampsia cespitosa*; d, *Hedera helix*; e, *Allium ursinum*; f, *Acer pseudoplatanus*-*Ulmus glabra*; g, *Cornus sanguinea*-*Campanula latifolia*] and (2) W12 [Beech-dog's mercury (*Fagus sylvatica*; *M. perennis*). Subcommunities: a, *Sanicula europaea*; b, *F. excelsior*-*M. perennis*; c, *Taxus baccata*]. One sample was collected from a locally distributed community-type within a W8 woodland: W5 [Alder-sedge (*Alnus glutinosa*; *Carex paniculata*). Subcommunities: a, *Phragmites australis*; b, *Lycopus europaeus*; c, *Chrysosplenium oppositifolium*]. A woodland type usually associated with acidic soils but which was sometimes locally distributed among W8 and W12 woodlands was also encountered: W10 [Oak-bracken-brambles (*Quercus robur*; *P. aquilinum*; *Rubus fruticosus*). Subcommunities: a, *Castanea sativa* coppice; b, Typical; c, *Corylus avellana*-*H. helix*; d, *Holcus lanatus*; e, *A. pseudoplatanus*-*Oxalis acetosella*].

Most samples from open habitats were collected

from long grass and mixed vegetation alongside hedgerows, usually next to roads.

The soil types reported in the soil maps of England and Wales (Avery *et al.*, 1975; Soil Survey Staff, 1983) were recorded.

Assessing habitat ages

The methodology of Rackham (1980) and Peterken (1981), which involves interpreting archaeological remains, evidence of past management, woodland-indicator and nonindicator species, abundance and diversity of vegetation, fauna, map features, written documents and names of woods, was used to assess habitat ages. Sample sites were located on the first OS map of each region (Ordnance Survey, 1970; reprints of maps first published in the 1830s) and the contemporary habitat was noted. The Domesday for each county was consulted for any historical information. Professor Robert Cameron kindly gave information on ancient woodlands in Warwickshire, and Jonathan Spencer of the former Nature Conservancy Council gave information on woodlands in Lincolnshire, Warwickshire (NCC, 1986a) and Oxfordshire (NCC, 1986b; Gibson, 1988). Information on those in Gloucestershire was obtained from the local nature conservation trust (Regini *et al.*, 1987; Hawkins *et al.*, 1988).

Estimating the age of a woodland is complicated by the fact that those aged 400 years or so and contiguous with older woodlands can develop floral characteristics of primary woodland. Assessing the ages of open habitats is also problematic. The numbers and types of tree and shrub species along a 30 pace length of hedgerow were recorded along with species found in the field layer, and ages of the hedgerows were estimated from Hooper's predictive equation (in Pollard *et al.*, 1974). The estimates, however, have large variances and must be used cautiously. Nevertheless, in some areas, notably Lincolnshire and Warwickshire, land enclosures are well-documented and it appears that hedgerows in these counties are likely to be about 200 years old. The presence of woodland indicator species in a hedgerow indicates that it is either a woodland relic or is, or was for a long time, in contact with ancient woodland. The nomenclature adopted here for referring to a particular sample is -w for woodland and -o for open habitat.

Starch gel electrophoresis

Snails were killed by freezing at -20°C and hepatopancreas samples were gently homogenized for 30 s

and spun at 1700 g for about 20 min at 4°C . Samples were then either placed on ice for immediate use or frozen at -80°C .

Full descriptions of electrophoresis methods are given by Harris & Hopkinson (1976). Buffer systems were those used by Selander *et al.* (1971) and Johnson *et al.* (1984). Six enzymes encoded by 11 loci were investigated in hepatopancreas tissue: malate dehydrogenase (EC 1.1.1.37; *MDH-1* and -2), diaphorase (EC 1.8.1.4; *DIA-1*), superoxide dismutase (EC 1.15.1.1; *SOD-1*, -2 and -3), phosphoglucosyltransferase (EC 5.4.2.2; *PGM-2*), phenylalanyl-proline dipeptidase (EC 3.4.13.9; *PEP-D*) and leucyl-leucyl-leucine tripeptidase (EC 3.4.-.-; *PEP-F*, *PEP-E* and *PEP-S*). The mobilities of bands (f , fast; m , intermediate; s , slow) relative to the commonest allele at each locus were recorded. Variation was detected at seven loci: *MDH-1*, *DIA-1*, *SOD-2*, *PGM-2*, *PEP-D*, *PEP-E* and *PEP-F*. The inheritance of allelic variants at *DIA-1* was confirmed by breeding studies (Wilson, 1993).

Statistical analysis

As heterozygotes cannot be scored reliably and as there is epistasis between B^O , U and T and between U^3 and T , χ^2 and Fisher's exact tests were carried out on the numbers of midbanded in banded and on those of 00345 among shells neither unbanded nor midbanded. As many tests were made, the Bonferroni procedure (in Weir, 1990) was used to reset the level of significance.

Allele frequencies at the seven polymorphic loci were calculated for each population. The calculation of allele frequencies at *PEP-F* was complicated by the presence of a null allele (*PEP-Fⁿ*) in some samples, so the method of Bernstein (in Hedrick, 1985) was used to estimate allele frequencies, which were adjusted to sum to unity.

Analysis of gene frequencies was performed using BIOSYS-1 (Swofford & Selander, 1989). The exact test of Haldane (1954) was used to determine whether observed allozyme frequencies were in Hardy-Weinberg equilibrium. Generalized Linear Interactive Modelling (GLIM; Aitken *et al.*, 1989) was used to compute the ANOVA to test for differences in gene frequencies between habitats. Estimated allele numbers at the four shell loci and the seven enzyme loci were fitted to a binomial model linked to the logit function. The model partitioned the transformed variances between habitats, counties and pairs, which were compared with the residual variance. Other models were applied to test for significant effects of longitude, altitude, soil type and

ages of open habitats upon allele frequency distributions.

Cluster analysis was performed using Nei's (1978) unbiased genetic distance (also used here to calculate phenotypic distances from estimated recessive allele frequencies) to construct dendrograms with Unweighted Pair-Group Method with Arithmetic Averaging (UPGMA). Geographical distances between the samples in each pair were calculated from the OS grid references and rounded to the nearest 10 metres.

The FORTRAN program of Black & Krafur (1985) was used to calculate two-locus composite linkage disequilibrium coefficients from the allozyme frequencies and from those estimated for the recessive

phenotypes at the shell colour and banding loci. Ohta's methods (1982a,b) were applied, in the hope of distinguishing between deterministic and stochastic effects that can produce linkage disequilibrium.

Results

The pair number, county, OS grid reference, estimated minimum age, habitat type, altitude and soil type of the sample sites are reported in Table 1. The woodlands in Lincolnshire, Warwickshire and Oxfordshire were mostly type W8. Those in Gloucestershire were mostly type W12. Although all woodlands will have been managed, most appear to be either ancient or ancient modified by planting.

Table 1 Pair number (Pr), county (C: L, Lincolnshire; W, Warwickshire; O, Oxfordshire; G, Gloucestershire), OS grid reference, estimated minimum age (r, recent; a, ancient; m, ancient woodland modified by planting; numbers refer to centuries before present), habitat type (Hb), altitude (Alt) in metres above sea level and soil type of *Cepaea nemoralis* sample sites

Pr	Site	C	Woodland					Open habitat				
			OS	Age	Hb*	Alt	Soil*	OS	Age	Hb†	Alt	Soil*
1	Holywell	L	SK 997 167	r	w8f	70	343c	SK 997 169	2	g/h	60	343c
2	Ufton	W	SP 383 625	a	w8a	100	712b	SP 377 616	2	g/h/t	110	411a
3	Chesterton	W	SP 342 575	a/3	w8a	90	572f	SP 347 579	2	g/h/t	90	411b
4	Bath Copse	W	SP 324 557	1/a	w8a	90	411b	SP 321 558	2	g/h	70	411b
5	Lighthorn Rough	W	SP 323 546	a/3	w8c	110	712g	SP 326 546	5	g/h	110	712g
6	Itchington Holt	W	SP 368 558	3/a	w8c	130	711m	SP 365 544	2	g/h	120	711t
7	Walton	W	SP 287 507	m	w8c	70	712b	SP 293 513	3	g/h	90	411a
8	Bath Hill	W	SP 291 532	m	w8	80	712b	SP 289 532	2	g	60	411a
9	Bowshot	W	SP 303 529	2/a	w8c	110	711m	SP 300 523	5	g/h/w	110	711m/411b
10	Marley - 1	O	SP 477 079	a	w8	80	343a	SP 479 083	2	g/w	70	712b
11	Marley - 2	O	SP 475 075	a	w8	130	343a	SP 478 072	2	g/t	130	343a
12	Marley East	O	SP 481 077	2	w8c	80	712b	SP 482 070	2	g/t	110	343a
13	Oaken Holt	O	SP 463 068	a	w8a	100	712b	SP 459 066	2	g/h	70	712b
14	Stroud Copse	O	SP 445 076	a	w8	80	712b	SP 444 078	5	g/h	60	712b
15	Wytham - 1	O	SP 456 089	a	w8	110	343a	SP 450 088	2	g/h	70	712b
16	Wytham - 2	O	SP 462 092	a	w8	110	343a	SP 474 089	2	g/h	80	511h
17	Tubbs Copse	O	SP 469 063	a	w5a/w8	90	712b	SP 467 061	2	g/h	70	712b
18	Thrupp	G	SO 868 032	a	w12b	170	343a	SO 874 025	7	g/h	130	411a
19	Dark	G	SO 828 024	a	w12b	130	343a	SO 839 029	6	g/h/w	80	572i
20	Pen	G	SO 822 025	a	w12b	130	572i	SO 816 033	5	g/h	70	572i
21	Stanley	G	SO 815 022	a	w12b	160	572i	SO 809 032	6	g/h	60	572i
22	Coaley	G	SO 789 001	a	w12b	200	343a	SO 775 011	2	g/h	60	572i
23	Woodchester	G	SO 832 009	a	w12b	120	343a	ST 845 999	0.5-2	g	110	411a
24	Bowldown	G	ST 846 915	a	w12c	140	343a	ST 838 929	6	g/h/t	170	411a
25	Oldbury	G	ST 816 891	m	w8	150	411a	ST 824 893	3	g/h	150	411a/343a
26	Alderley	G	ST 778 908	a	w12b	130	343a	ST 794 912	2	g	170	343a
27	Claypits	G	SO 865 037	a	w12b	140	343a	SO 858 043	8	g/h	50	343b

*See Materials and methods.

†g, grass; h, hedgerow; t, trees; w, woodland relic.

Twenty-one open habitat samples were collected from mixed vegetation with grass adjacent to hedge-rows, two of which contained woodland relic indicator species. The remainder were collected from open grassland with mixed vegetation, one of which contained woodland relic indicator species. Most of the open habitats in Warwickshire and Oxfordshire appear to be aged about 200 years while most of those from Gloucestershire appear to be older than 500 years.

Shell colour and banding polymorphisms

The percentages of the various shell phenotypes are reported in Table 2, which gives the sample sizes, the proportion of juveniles in each sample and the distance between the members of each pair of samples. Because of epistasis the percentage of 00300 among banded shells, and that of 00345 among banded shells that are not midbanded, is reported. Two samples from open habitats contained

Table 2 Sample size (*N*), percentage of juveniles (*J*) and of different shell phenotypes (*Y*, yellow; *P*, pink; *B*, brown; *EU*, 'effectively unbanded'; *D*, 'dark'; *U*, unbanded; *M*, midbanded among banded; 345, upper two bands missing among banded not midbanded) in woodland and open habitat samples of *Cepaea nemoralis*, and distance (*Dist*) in metres between the members of each pair (*Pr*) of samples. The overall mean percentages and standard deviations (*SD*), and those for each county (abbreviations as in Table 1) are reported

Pr	Woodland										Open habitat										Dist (m)
	<i>N</i>	<i>J</i>	<i>Y</i>	<i>P</i>	<i>B</i>	<i>EU</i>	<i>D</i>	<i>U</i>	<i>M</i>	345	<i>N</i>	<i>J</i>	<i>Y</i>	<i>P</i>	<i>B</i>	<i>EU</i>	<i>D</i>	<i>U</i>	<i>M</i>	345	
1	21	38	48	52	0	14	29	0	0	14	20	5	85	15	0	10	0	0	10	0	200
2	21	33	5	91	5	76	81	5	40	58	28	25	64	36	0	32	29	0	7	27	1080
3	25	20	16	84	0	56	52	36	19	8	22	23	59	41	0	27	0	27	0	0	640
4	20	15	30	70	0	80	55	30	36	56	30	37	63	37	0	70	27	37	42	18	320
5	23	74	0	65	35	74	74	17	58	25	26	50	58	42	0	38	31	0	19	24	300
6	20	20	15	85	0	55	60	25	33	10	45	20	64	36	0	49	20	20	19	21	1430
7	30	27	37	63	0	57	33	33	15	24	40	33	40	47	13	55	30	25	7	36	850
8	36	28	39	58	3	61	53	33	29	18	29	17	62	31	7	48	31	31	20	6	200
9	21	76	48	43	9	81	43	38	46	29	22	27	0	100	0	45	64	0	23	29	670
10	21	38	19	81	0	14	14	10	0	0	21	67	86	14	0	43	0	10	11	24	450
11	28	25	11	75	14	39	43	29	15	0	25	20	72	28	0	28	0	0	0	16	420
12	24	17	17	75	8	58	63	46	23	0	24	21	46	42	12	67	50	29	47	11	710
13	20	10	15	75	10	30	30	10	22	0	29	31	72	28	0	24	3	0	24	0	450
14	20	25	10	65	25	60	50	40	8	27	51	45	72	14	14	55	14	39	10	14	220
15	27	4	11	85	4	70	67	19	18	50	24	21	33	67	0	25	8	8	0	9	610
16	23	17	9	83	9	83	74	35	73	0	29	7	93	4	3	48	7	24	9	25	1240
17	24	8	42	46	12	63	63	54	18	0	20	30	75	25	0	40	25	5	26	14	280
18	20	15	15	85	0	85	80	30	57	17	21	0	52	19	29	86	43	5	65	14	920
19	30	27	3	90	7	100	97	20	83	75	25	12	64	12	24	84	32	44	29	40	1210
20	21	14	0	100	0	95	100	67	86	0	22	27	41	50	9	59	41	5	48	0	1000
21	21	14	19	76	5	95	81	52	80	50	31	0	23	55	23	26	29	23	4	0	1170
22	20	35	5	85	10	100	95	50	90	100	22	27	27	41	32	73	68	32	60	0	1720
23	20	60	10	65	25	100	90	15	94	0	20	10	35	35	30	75	55	40	50	17	1640
24	22	23	41	41	18	77	50	32	47	0	31	10	71	29	0	52	19	29	32	0	1610
25	20	10	0	75	25	70	75	40	33	0	20	5	65	30	5	40	25	25	20	0	820
26	20	10	5	75	20	55	55	35	23	0	23	4	78	9	13	87	22	39	71	25	1650
27	20	20	5	90	5	85	75	30	57	17	20	5	20	65	15	80	70	50	20	25	920
Mean	23	26	18	73	9	68	62	31	41	21	27	21	56	35	8	50	27	20	25	15	848
SD	4	18	15	15	10	24	22	16	28	27	8	16	22	21	11	21	21	16	21	12	511
Mean _W	25	37	24	70	6	67	56	27	35	28	30	29	51	46	3	46	29	17	17	20	735
SD	6	24	17	16	12	11	15	11	14	19	8	11	22	22	5	13	17	15	13	12	528
Mean _O	24	18	17	73	10	52	50	30	22	10	28	30	69	28	4	38	13	14	16	14	512
SD	3	11	11	13	7	24	20	17	22	19	10	18	20	20	6	18	17	15	16	8	251
Mean _G	21	23	10	78	12	86	80	37	65	26	24	10	48	34	18	66	40	29	40	12	1271
SD	3	15	12	16	10	15	17	16	25	37	4	10	21	19	11	21	19	15	22	14	363

Fig. 1 Scatter diagram of percentage yellow against percentage 'effectively unbanded' shells of *Cepaea nemoralis*. Open habitat samples, open circles; woodland samples, filled triangles. Pair numbers refer to those reported in Table 1.

no juveniles whereas one woodland sample contained 76 per cent juveniles. Nevertheless, allele frequencies of juveniles and adults were pooled because none of the samples showed significant differences between them when the Bonferroni procedure was applied.

Figure 1 shows a plot of percentage yellow against percentage EU shells with the members of each pair of samples linked. The data fall neatly into two groups with most open habitats in the upper left quadrat and most woodlands in the lower right quadrat. Twenty-two pairs of samples demonstrate clearly the differentiation between habitats consistent with the model of visual selection proposed by Cain & Sheppard (1954). There were, however, a few anomalies.

The most unusual sample was 9-o which was fixed for pink. However, this site may be strongly influenced by a nearby, probably ancient, woodland. Sample 10-w had a lower percentage of EU than expected, and samples 12-o, 18-o and 26-o had percentages of EU that were higher than their respective woodland samples. The unusual position of 26-o is accentuated by the percentage of EU in the woodland sample which was the lowest among Gloucestershire woodlands. The Lincolnshire woodland sample had the highest percentage of yellow shells among all woodland samples and also had an unusually low percentage of EU shells. This woodland appears to be recent, although it may have

Table 3 ANOVA of shell phenotypes among samples of *Cepaea nemoralis* from Warwickshire, Oxfordshire and Gloucestershire

Shell phenotype	F-ratio		
	Habitat	County	Pair number
'Dark'	55.69***	10.05***	1.75
Yellow	45.39***	2.29	0.75
Pink	44.19***	0.94	0.76
EU	13.98**	8.58**	1.56
00300	10.81**	9.59***	1.54
00000	5.42*	2.75	1.16
00345	0.29	3.70*	1.46
Brown	0.08	4.53*	0.88

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Phenotypes are listed in decreasing order of F -ratio attributed to habitat type.

been until recently contiguous with a small ancient relic woodland, and the pair of samples was excluded from further analysis except for comparison in UPGMA.

ANOVA (Table 3) shows that there was a highly significant effect of habitat upon the distribution of yellow, pink and midbanded shells and upon that of the phenotypes 'dark' and EU, and a significant effect upon the distribution of unbanded shells. There appeared to be no effect of habitat upon the

distribution of brown or 00345 shells. Brown shells showed an interesting distribution along the Cotswold escarpment. Excluding the pairs of samples from the plateau, samples from a higher altitude each had a lower frequency of brown shells than their pair regardless of habitat type (Tables 1 and 2).

There was highly significant regional variation (Table 3) upon the distribution of midbanded, 'dark' and EU shells. There was also significant regional

UPGMA on shell colour and banding loci (Fig. 2) showed that the samples mostly cluster into two groups, one comprising woodlands and the other open habitats.

Enzyme polymorphisms

Allele percentages at the seven variable enzyme loci are reported in Table 4. None of the samples showed significant differences in allele frequencies between the life stages when the Bonferroni procedure was applied and they were therefore pooled. Samples conformed to Hardy-Weinberg equilibrium.

Overall mean allozyme frequencies at each locus were similar in the two habitats. However, among the samples from Warwickshire and Oxfordshire *MDH-1'* was at a slightly higher frequency in wood-

Fig. 2 UPGMA dendrogram of 54 samples of *Cepaea nemoralis* based on estimated allele frequencies, adjusted for epistasis by C^B , at four shell colour and banding loci. Pair numbers refer to those reported in Table 1.

lands than in open habitats (Table 5; $P < 0.05$). This difference was mostly caused by the nearly consistent difference between habitats in Warwickshire. The second largest effect of habitat was at *PEP-D* ($P > 0.10$) whereas alleles at the other loci showed very little effect. These findings prompted the extension of the survey to include samples from Gloucestershire and to analyse them for allele frequencies at *MDH-1* and *PEP-D*.

Among Gloucestershire samples there was a higher frequency of *MDH-1^f* in four of the 10 woodlands than in their paired open habitats. Taking the Gloucestershire data into account, the effect of habitat over the three counties was not significant (Table 5; $P > 0.10$), nor was it upon the frequency of alleles at *PEP-D*.

In contrast to the cluster analysis of shell colour and banding loci, UPGMA on the seven enzyme loci (Fig. 3) showed no evidence of clustering of samples by habitat. ANOVA showed that the regional differentiation apparent in Fig. 3 was highly significant at each enzyme locus (Table 5). In addition, there was highly significant microgeographical variation at *MDH-1* and *PEP-D*.

Linkage disequilibrium

The colour and banding loci showed linkage disequilibria in more than 50 per cent of the samples (Wilson, 1993). The association was mostly between

brown and unbanded shells, which can be explained by epistasis between *C^B* and *B*. Ten samples had a significant excess of pink unbanded shells and 19 samples had a significant deficiency. Five samples had a significant excess of yellow unbanded shells and 22 samples had a deficiency. There was, however, no significant systematic disequilibrium between any of the loci investigated.

Discussion

As expected, the frequencies of phenotypes at the shell colour and banding loci in *C. nemoralis* differ between habitats in Warwickshire (Currey *et al.*, 1964; Cameron & Pannett, 1985) and Oxfordshire (Cain & Sheppard, 1954), consistent with the model of visual predation proposed by Cain & Sheppard (1954) but unlike the distribution of phenotypes found in areas with unstable histories (Cameron *et al.*, 1980; Cameron & Dillon, 1984).

There is, of course, highly significant regional variation, and this is more marked for allozyme loci than for the shell loci. Although *MDH-1* and *PEP-D* both show significant microgeographical variation there is little evidence (although this could be a result of the small number of samples collected in each county) of the marked allozyme differentiation within areas less than 10 km × 10 km found by Johnson (1976) on the Berkshire Downs, where there was concordance between area effects for midbanded and spread-banded shells and alleles at two enzyme loci. It is apparent from Fig. 3 that levels of genetic differentiation are greater in Warwickshire than in Oxfordshire. This probably reflects the fact that geographical distances between samples are greater in Warwickshire.

On a larger scale Jones *et al.* (1980) and Caugant *et al.* (1982) found that patterns of differentiation of alleles at shell colour and banding loci and at molecular loci were independent of each other among populations of *C. nemoralis* from North Wales and the Pyrenees. However, in a much larger survey Ochman *et al.* (1983) found that there were regional patterns of concordance at allozyme loci with little differentiation at shell loci. In Warwickshire, Cameron & Pannett (1985) found microgeographical variation at shell colour and banding loci resembling area effects that was not detected here. Further analysis showed that this was probably a result of the smaller number of samples collected in this study (Wilson, 1993). In addition, in that county, they found no effect of habitat upon the distribution of midbandeds, and consequently a smaller effect upon that of effectively unbandeds.

Table 5 ANOVA of the common alleles of *Cepaea nemoralis* at (a) two enzyme loci among 52 samples and (b) seven enzyme loci among 32 samples

Allozyme	F-ratio		
	Habitat	County	Pair number
(a) Among 52 samples			
<i>MDH-1^f</i>	2.36	5.62**	3.00**
<i>PEP-D^f</i>	0.76	13.14***	4.67***
(b) Among 32 samples			
<i>MDH-1^f</i>	8.39*	10.81**	3.62**
<i>PEP-D^f</i>	2.68	26.31***	4.52**
<i>DIA-1^m</i>	0.87	12.76**	1.14
<i>PEP-F^s</i>	0.48	173.70***	1.54
<i>SOD-2^s</i>	0.17	217.40***	1.77
<i>PEPE^s</i>	0.08	35.10***	2.09
<i>PGM-2^s</i>	0.00	17.64***	1.62

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Alleles are listed in decreasing order of *F*-ratio attributed to habitat type.

Fig. 3 UPGMA dendrogram of 34 samples of *Cepaea nemoralis* based on allele frequencies at seven enzyme loci. Pair numbers refer to those reported in Table 1.

It is possible therefore that patterns of regional differentiation at enzyme loci reflect founder effects, modified by other historical stochastic events, rather than natural selection, whereas selective differences between shell colour and banding phenotypes in different environments explain their contemporary distribution, but that natural selection on a regional scale cannot be ruled out. In lowland England, where selection has shaped the distribution of gene frequencies at shell colour and banding loci, there is evidence of microgeographical variation at these loci (Cameron & Pannett, 1985) and at enzyme loci (the present study) indicative of changes in landscape history, but which have had a less marked effect

than those upon downland populations of *Cepaea* (Cameron & Dillon, 1984).

Although the discontinuity in the distribution of brown shells along the Cotswold escarpment may be indicative of a sharp cline between different area effects it is not concordant with the pattern of allozyme distributions and requires further investigation.

The results reported here, comparing ancient woodlands and old open habitats, show that whereas there has been adequate time for selection to act at shell colour and banding loci there is little evidence of selection at enzyme loci. Although there seem to be consistent differences between habitats in the distribution of alleles at *MDH-1* among samples

from Warwickshire, this is not repeated elsewhere. The mode of differentiation at the molecular level in populations of *C. nemoralis* in lowland England generally appears to be one of local and regional differentiation with either little effect of habitat type, contrasting markedly with the pattern at the morphological level, or with an effect that is too small to be detected by the techniques employed here.

Acknowledgements

I thank Professor Bryan Clarke for his supervision during this research. Professor Robert Cameron and Jonathan Spencer gave information on habitats. Drs Mark Beaumont, Jane Farquharson, Nigel Harding and John Stephen helped with fieldwork and in discussion, while Joyce Ellis and Dave Walters gave technical help. The research was supported by a NERC PhD studentship.

References

- AITKEN, M., ANDERSON, D., FRANCIS, B. AND HINDE, J. 1989. *Statistical Modelling in GLIM*. Clarendon Press, Oxford.
- AVERY, B. W., FINDLAY, D. C. AND MACKNEY, D. 1975. *Soil Map of England and Wales, 1:1,000,000*. Soil Survey of England, Harpenden.
- BLACK, W. C. AND KRAFSUR, E. S. 1985. A FORTRAN program for the calculation and analysis of two-locus linkage disequilibrium coefficients. *Theor. Appl. Genet.*, **70**, 491–496.
- BRUSSARD, P. F. AND MCCrackEN, G. F. 1974. Allozymic variation in a North American colony of *Cepaea nemoralis*. *Heredity*, **33**, 98–101.
- CAIN, A. J. AND CURREY, J. D. 1963a. Area effects in *Cepaea*. *Phil. Trans. Roy. Soc. B*, **246**, 1–81.
- CAIN, A. J. AND CURREY, J. D. 1963b. The causes of area effects. *Heredity*, **18**, 467–471.
- CAIN, A. J. AND SHEPPARD, P. M. 1954. Natural selection in *Cepaea*. *Genetics*, **39**, 89–116.
- CAMERON, R. A. D. AND DILLON, P. J. 1984. Habitat stability, population histories and patterns of variation in *Cepaea*. *Malacologia*, **25**, 271–290.
- CAMERON, R. A. D. AND PANNETT, D. J. 1985. Interaction between area effects and variation with habitat in *Cepaea*. *Biol. J. Linn. Soc.*, **24**, 365–379.
- CAMERON, R. A. D., CARTER, M. A. AND PALLES-CLARK, M. A. 1980. *Cepaea* on Salisbury Plain: patterns of variation, landscape history and habitat stability. *Biol. J. Linn. Soc.*, **14**, 335–358.
- CAUGANT, D., SELANDER, R. K. AND JONES, J. S. 1982. Geographic structuring of molecular polymorphisms in Pyrenean populations of the snail *Cepaea nemoralis*. *Genetica*, **57**, 177–191.
- CLARKE, B. C. 1962. Balanced polymorphism and the diversity of sympatric species. In: Nichols, D. (ed) *Taxonomy and Geography*, pp. 47–70. Systematics Association Publication no. 4, Oxford.
- CLARKE, B. C. 1966. The evolution of morph-ratio clines. *Am. Nat.*, **100**, 389–402.
- CURREY, J. D., ARNOLD, R. W. AND CARTER, M. A. 1964. Further examples of variation of populations of *Cepaea nemoralis* with habitat. *Evolution*, **18**, 111–117.
- FORD, E. B. 1975. *Ecological Genetics*, 4th edn. Chapman and Hall, London.
- GIBSON, C. W. D. 1988. The distribution of 'Ancient woodland' plant species among areas of different history in Wytham Woods, Oxfordshire. In: Kirby, K. J. and Wright, F. J. (eds) *Woodland Conservation and Research in the Clay Vale of Oxfordshire and Buckinghamshire*, pp. 32–40. Nature Conservancy Council, Peterborough.
- GOODHART, C. B. 1963. "Area effects" and non-adaptive variation between populations of *Cepaea* (Mollusca). *Heredity*, **18**, 459–465.
- HALDANE, J. B. S. 1954. An exact test for randomness of mating. *J. Genet.*, **52**, 631–635.
- HARRIS, H. 1966. Enzyme polymorphisms in man. *Proc. R. Soc. B*, **164**, 298–310.
- HARRIS, H. AND HOPKINSON, D. A. 1976. *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland, Amsterdam.
- HAWKINS, W., FLEMING, J., GRIFFITHS, W., RALPHS, I. AND NIEBRZYDOWSKI, S. 1988. *A Revised Inventory of Gloucestershire's Ancient Woodlands: the Stroud and Dursley Valleys*. The Gloucestershire Trust for Nature Conservation, Stonehouse, Gloucestershire.
- HEDRICK, P. W. 1985. *Genetics of Populations*. Jones and Bartlett, Portola Valley, CA.
- HUBBY, J. L. AND LEWONTIN, R. C. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*, **54**, 577–594.
- JOHNSON, M. S. 1976. Allozymes and area effects in *Cepaea nemoralis* on the western Berkshire downs. *Heredity*, **36**, 105–121.
- JOHNSON, M. S. 1979. Inheritance and geographic variation of allozymes in *Cepaea nemoralis*. *Heredity*, **43**, 137–141.
- JOHNSON, M. S., STINE, O. C. AND MURRAY, J. 1984. Reproductive compatibility despite large-scale divergence in *Cepaea nemoralis*. *Heredity*, **53**, 655–665.
- JONES, J. S., LEITH, B. H. AND RAWLINGS, P. 1977. Polymorphism in *Cepaea*: a problem with too many solutions? *Ann. Rev. Ecol. Syst.*, **8**, 109–143.
- JONES, J. S., SELANDER, R. K. AND SCHNELL, G. D. 1980. Patterns of morphological and molecular polymorphism in the land snail *Cepaea nemoralis*. *Biol. J. Linn. Soc.*, **14**, 359–388.
- KIMURA, M. 1968. Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.
- KOEHN, R. K., NEWALL, R. I. E. AND IMMERMAN, F. 1980. Maintenance of an aminopeptidase allele frequency cline by natural selection. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 5385–5389.
- LAMOTTE, M. 1951. Recherches sur la structure génétique des populations naturelles de *Cepaea nemoralis* (L.).

- Bull. Biol. Fr. Belg. Suppl.*, **35**, 1–239.
- LEWONTIN, R. C. AND HUBBY, J. L. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, **54**, 595–609.
- MCDONALD, J. H. 1991. Contrasting amounts of geographical variation as evidence for direct selection: the *Mpi* and *Pgm* loci in eight crustacean species. *Heredity*, **67**, 215–219.
- NCC. 1986a. Notification of Long Itchington and Ufton Woods as an SSSI. Nature Conservancy Council, Peterborough.
- NCC. 1986b. Notification of Wytham Woods as an SSSI. Nature Conservancy Council, Peterborough.
- NEI, M. 1978. Estimation of heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- OGHMAN, H., JONES, J. S. AND SELANDER, R. K. 1983. Molecular area effects in *Cepaea*. *Proc. Natl. Acad. Sci., U.S.A.*, **80**, 4189–4193.
- OHTA, T. 1982a. Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 1940–1944.
- OHTA, T. 1982b. Linkage disequilibrium with the island model. *Genetics*, **101**, 139–155.
- ORDNANCE SURVEY. 1970. *Reprints of the first edition of the one-inch Ordnance Survey of England and Wales*. Sheets 44, 52, 61 and 68. David and Charles, Newton Abbot, Devon.
- PETERKEN, G. F. 1981. *Woodland Conservation and Management*. Chapman and Hall, London.
- POLLARD, E., HOOPER, M. D. AND MOORE, N. W. 1974. *Hedges*. Collins, London.
- POWERS, D. A., LAUERMAN, T., CRAWFORD, D., SMITH, M., GONZALEZ-VILLASENOR, I. AND DIMICHELE, L. 1991. The evolutionary significance of genetic variation at enzyme synthesizing loci in the teleost *Fundulus heteroclitus*. *J. Fish Biol.*, **39** (suppl. A), 169–184.
- RACKHAM, O. 1980. *Ancient Woodland: its History, Vegetation and Uses in England*. Edward Arnold, London.
- REGINI, K., GOUDGE, C., HAWKINS, W., HOLT, I. AND DAVIES, J. 1987. *A Revised Inventory of Gloucestershire's Ancient Woodlands: the Cotswold Plateau*. The Gloucestershire Trust for Nature Conservation, Stonehouse, Gloucestershire.
- RODWELL, J. S. (ed). 1991. *British Plant Communities: Woodland and Scrub*, vol. 1. Cambridge University Press, Cambridge.
- SELANDER, R. K., SMITH, M. H., YANG, S. Y., JOHNSON, W. E. AND GENTRY, J. B. 1971. Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*). Studies in genetics VI. *University of Texas Publications*, **7103**, 49–90.
- SHEPPARD, P. M. 1951. Fluctuations in the selective value of certain phenotypes in the polymorphic land snail *Cepaea nemoralis* (L.). *Heredity*, **5**, 125–134.
- SOIL SURVEY STAFF. 1983. *Soils of England and Wales, 1: 250,000*. Soil Survey of England, Harpenden, Herts.
- SWOFFORD, D. L. AND SELANDER, R. B. 1989. BIOSYS-1. *A Computer Program for the Analysis of Allelic Variation in Population Genetics and Biochemical Systematics*. Release 1.7. University of Illinois, Urbana, IL.
- TATARENKOV, A. AND JOHANNESSON, K. 1994. Habitat related allozyme variation on a microgeographic scale in the marine snail *Littorina mariae* (Prosobranchia: Littorinacea). *Biol. J. Linn. Soc.*, **53**, 105–125.
- WATT, W. B., CASSIN, R. C. AND SWAN, M. S. 1983. Adaptation at specific loci. III. Field behavior and survivorship differences among *Colias* PGI genotypes are predictable from *in vitro* biochemistry. *Genetics*, **103**, 725–739.
- WEIR, B. S. 1990. *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.
- WILSON, I. F. 1993. *Factors Influencing Gene Frequency Distributions in Cepaea nemoralis* (L.). Ph.D. Thesis, University of Nottingham.
- WRIGHT, S. 1978. *Evolution and the Genetics of Populations*, vol. 4, *Variability Within and Among Natural Populations*. University of Chicago Press, Chicago.