Evolution, 59(10), 2005, pp. 2170–2184

GENETIC DRIFT WITHIN A PROTECTED POLYMORPHISM: ENIGMATIC VARIATION IN COLOR-MORPH FREQUENCIES IN THE CANDY-STRIPE SPIDER, *ENOPLOGNATHA OVATA*

G. S. OXFORD
Department of Biology (Area 18), University of York, P.O. Box 373, York YO10 5YW, United Kingdom
E-mail: gso1@york.ac.uk

Abstract.—The candy-stripe spider, *Enoplognatha ovata*, exhibits a striking color polymorphism comprising three morphs. A number of lines of evidence strongly suggest that this polymorphism is maintained by natural selection: its presence in a sister species, *E. latimana*; the physical nature of the variation; the virtual lack of monomorphic populations; the highly consistent rank-order of morphs within populations; and the presence of large-scale clines associated with climatic variables. However, the absence of selection is equally strongly suggested by very local surveys of morph frequencies over space and time, perturbation experiments, and a variance in morph frequency between populations that is virtually independent of spatial scale. In addition, local spatial patterns in one study site (Nidderdale, Yorkshire, England) have been explained in terms of intermittent drift over half a century ago, a hypothesis supported here by the distributions of four other genetic markers (two allozyme and two visible polymorphisms). A heuristic model is suggested that reconciles these apparently contradictory messages regarding the importance of drift and selection in this system. It is proposed that when allele frequencies of the color morph *redimita* lie between approximately 0.05 and 0.3, the Δq on q plot is very shallow, so that within this region, where the majority of populations lie, selection is weak and drift is the major force determining local morph frequencies. However, outside this range of frequencies, powerful selection acts to protect the polymorphism. This model may apply to polymorphisms in other species and explain why evidence of selection in natural populations is often elusive.

Key words.—Allozyme variation, *Enoplognatha latimana*, intermittent drift, intermittent selection, scale-independent variation.

Received January 20, 2005. Accepted July 25, 2005.

Between the 1930s and 1960s, color polymorphisms provided the principal material that allowed the study of genetic variation in natural populations, and enabled the emerging discipline of ecological genetics to establish a firm foundation in Britain (Ford 1964; Cain and Provine 1991). Since the 1960s other genetic markers have become widely available with the introduction of allozyme electrophoresis and various DNA technologies. Despite its demise in the context of a genetic marker, color polymorphism remains a prominent feature of many plant and animal groups, particularly within the arthropods, although the underlying significance of this variation is still obscure in the majority of cases.

The evolutionary significance of an organism's external phenotype can be assessed in a number of ways. For example, the mapping of morphological features onto molecular phylogenies has demonstrated that convergence is a widespread phenomenon and with it the implication that the phenotypes are of adaptive value. Thus, with respect to coloration, Hoekstra and Price (2004) discussed the convergence of plumage characteristics in orioles (*Icterus* spp.) and Gillespie (2004) showed how, among Hawaiian *Tetragnatha* spiders, color morphs associated with specific habitats have evolved independently on different islands. In some cases the loci responsible for shifts in color have been identified allowing an assessment of the degree of genetic change underlying convergence between species (Mundy et al. 2004) and within them (Nachman et al. 2003; Hoekstra et al. 2004). Although such convergent changes are almost certainly driven by natural selection, observations in the field are still required to determine the nature of the selective forces involved in each case.

A more traditional approach to elucidating the evolutionary mechanisms responsible for the presence of visible variation within species is to examine spatial and temporal patterns of morph frequencies in the light of those expected under different hypotheses (e.g., Endler 1986; Jones et al. 1977; Halkka and Halkka 1990). This line of approach has had some spectacular successes, for example the evolution of melanism in moths (Cook 2003), although there are still very few cases where we have real insights into the relative importance of evolutionary processes such as selection and drift (O'Hara 2005).

One reason for the lack of information on evolutionary forces acting on visible polymorphisms may lie in the fact that temporal morph-frequency changes, which could be correlated with shifts in selective forces (Endler 1986), may be slow and require observations over extended periods of time. Another may be a lack of, or inconsistent, associations between morph frequencies and monitored environmental variables. This might suggest, among other things, that inappropriate variables are being measured, that the characters concerned are truly neutral, or that selection is operating intermittently but for much of the time is too weak to overcome the stochastic forces associated with small populations. In species with relatively long generation times there may also be an appreciable lag between a change in selection and consequential shifts in morph frequencies (Cook 1998). Determining which of these or other possibilities is the case is often extremely difficult.

Here I present an analysis of published and new data relating to a color polymorphism in the candy-stripe spider *Enoplognatha ovata* (Clerck) (Theridiidae), and use the patterns that emerge to assess the relative roles played by various evolutionary processes in the maintenance of the visible var-

iation and the determination of morph frequencies across a number of geographical scales. The patterns are shown to be enigmatic: some strongly support a major role for genetic drift, while others are explicable only in terms of natural selection. Possible evolutionary scenarios that could reconcile these contradictory conclusions are explored and a testable heuristic model is developed.

Ecology and Genetics

Enoplognatha ovata is widely distributed across Europe and eastern Asia. It also occurs on both the eastern and western coasts of central North America, almost certainly as a result of human introduction (Oxford and Reillo 1994). The species is a strict annual and frequents open habitats, such as road verges, waste ground, and domestic gardens, which contain low-growing, broad-leaved vegetation. Because a gravid female requires a broad-leaf retreat in which to produce and guard its usually single egg sac, the population structure mirrors the patchiness of appropriate vegetation (Oxford 1993a). Within suitable vegetation, spider densities are often very high compared to those in less appropriate intervening areas (Oxford and Shaw 1986). This concentration of spiders within relatively stable vegetation patches, and their annual life cycle, means that genetic data can readily be collected from discrete, identifiable populations (and generations) over both space and time. During the usual surveying period, females are enclosed within rolled leaves so that sampling is unbiased with respect to color morphs.

Throughout its range, the species exhibits a striking color polymorphism comprising three morphs (Oxford 1983; Wise and Reillo 1985). The opisthosoma (abdomen) can be plain yellow (var. *lineata* Clerck), yellow with two dorsolateral carmine stripes (var. *redimita* Clerck), or yellow with the dorsal surface of the opisthosoma entirely carmine (var. *ovata* Clerck). The morphs are determined by three alleles, C_l, C_r and C_o, respectively, at an autosomal locus (*C*) (Oxford 1983; Reillo and Wise 1988a). It seems highly likely that the yellow background color of the opisthosoma of all morphs is determined by a locus other than *C*, and that C_l is, in effect, a null allele that leaves the basic coloration unmodified (analogous to the ABO blood group system in humans where the *H*-locus determines the antigenic properties of red blood cells, which are then modified by alleles I_A and/or I_B, but not by the null-allele I_O; see, for example, Hartl and Jones 1998). C_r and C_o, in contrast, are alleles that determine the presence and distribution of red pigment. The phenotypic expression of the *C* alleles is complicated by epistasis with a putative polymorphic regulatory locus (*R*) that is postulated to be very tightly linked to, and *cis*-acting on, the *C* locus (Oxford 1983, 1985a). The two *R* alleles determine the stage of development at which the color alleles C_r and C_o are expressed. Color alleles located adjacent to the R_e (early) regulatory allele are expressed in both sexes from the third or fourth instars onward. Color alleles adjacent to the R_l (late) regulatory allele are activated in females only after the final molt and are not expressed at all in males, where they effectively become null alleles. In mature females the color morphs exhibit a simple dominance hierarchy, irrespective of the regulatory-locus variation, with *ovata* dominant to *redimita*, and both dominant to *lineata* (Oxford 1983; Reillo and Wise 1988a). Fortunately this is the sex and life stage normally scored. In addition to this variation, the opisthosoma can bear two dorsolateral rows of black spots with up to six (occasionally seven) in each row. The presence or absence of spots seems to be controlled by a major locus but the number of spots may be polygenically determined (Oxford 1989).

Materials and Methods

Color-Morph Frequencies

Color morphs of mature females were routinely scored by sampling rolled leaves (see above) during July and August. For all temporal and some snapshot surveys, females were scored nondestructively in situ, otherwise individuals were mass preserved in 70% ethanol (Oxford and Reillo 1993). Spatial surveys of color-morph frequencies have been conducted across a number of geographical scales: Nidderdale, Yorkshire, England (Oxford 1976; Oxford and Shaw 1986); Pembrokeshire, Wales (Oxford 1991); Sweden (G. S. Oxford and B. Gunnarsson, unpubl. data); mainland Britain (Oxford 1985b); mainland Europe (Oxford and Reillo 1993). Additional information from central Europe was published in Hippa and Oksala (1979). Data for eastern North America are available in Reillo and Wise (1988b,c), Wise and Reillo (1985) and Reillo (1989). For collecting sites not considered in the other papers, figure 1 in Reillo (1989) was enlarged and morph frequencies estimated from the histograms.

Temporal surveys have been undertaken in Nidderdale (Oxford 1976; Oxford and Shaw 1986), for some populations over a period of 30 years (generations). Temporal information is also available for some populations in eastern North America (Reillo and Wise 1988c) and central Europe (Hippa and Oksala 1979).

Regulatory Polymorphism

The estimation of allele frequencies at the regulatory locus requires either rearing spiders to maturity in the laboratory or sampling a population in the field twice within a season. Early patterns establish in both sexes in the third or fourth instars, whereas the color-morph frequencies in mature females will represent the combined early- and late-developing patterns. Populations were scored in mid-June, when males are in their penultimate or final instars and females in at least their fourth instar (early sample), and again in late July to mid-August when all females are mature (late sample); males by this time are dead (Oxford 1985a). Studies from which data on the regulatory polymorphism can be extracted were made in Nidderdale (Oxford 1983, 1985a, unpubl. data), eastern North America (Reillo and Wise 1988c), and Finland (Hippa and Oksala 1979). Hippa and Oksala (1979) and Gerhardt (1921) provided some additional information from central European countries. The proportion of the early regulatory allele is expressed as the frequency of, for example, *redimita* in the early sample divided by its frequency in the late sample (Oxford 1985a). These calculations assume that the proportion of early-developing *redimita* in the population remains constant between the two sampling dates.

Allozymes

The pattern of variation at enzyme-coding loci was sought in the Nidderdale populations using cellulose acetate electrophoresis (Richardson et al. 1986). Females were sampled toward the end of the reproductive season when their removal could have little effect on their offspring's fitness. Individual animals were homogenized in two volumes of 50 mM Tris buffer (pH 8.0), centrifuged, and the supernatant applied to Titan III cellulose acetate sheets (Helena Laboratories, Beaumont, TX). Appropriate running buffers were taken from Richardson et al. (1986) and stain recipes and procedures adapted from those of Hebert and Beaton (1993). Twenty-five enzyme systems were initially tested but only five proved to stain reliably and/or show interpretable variation—amylase (Amy), phosphoglucose isomerase (Pgi), lactate dehydrogenase (Ldh), malate dehydrogenase (Mdh) and phosphoglucomutase (Pgm). Of these, the *Amy* and *Pgi* loci were highly polymorphic in all populations, whereas the *Ldh, Mdh*, and *Pgm* loci were monomorphic in most populations, and very weakly polymorphic in the rest. Initial samples for electrophoresis were collected from Nidderdale sites D, B, E, Q, K, AD, Jb, AC, AB, and N (see Oxford and Shaw 1986) in 1996 ($n = 40$ to 44 individuals per site); these were analyzed for variation at the *Pgi* locus alone. In 1998, further samples were taken from the same selection of sites except that G was added and AC was not resampled. Of the 40 individuals collected per site, typically 24 were scored for the five polymorphic loci mentioned above. For site AD only 20 spiders could be found and all were scored. Spiders were stored at −80°C between collection and processing for electrophoresis.

Statistical Analyses

Statistical analyses were performed in Minitab (ver. 13; Minitab Inc., State College, PA) or using macros written for Excel spreadsheets (Excel 2000, Microsoft Corp., Redmond, WA). Comparisons of variances in morph frequencies among surveys used a jackknife technique (Layard 1973) because of the sensitivity of other methods to the nonnormality of underlying distributions (Box 1953; Miller 1974). The pseudovalues thus generated were subjected to a one-way ANOVA and interpreted using appropriate post-hoc comparisons. Ninety-five percent confidence intervals for binomial data were determined using the adjusted Wald procedure as the exact method was shown by Agresti and Coull (1998) to be overly conservative, particularly at low sample sizes. Implementation followed Samuels and Witmer (2003, p. 209).

To compare the impact of the regulatory polymorphism on color-morph frequencies, an index of polymorphic diversity (H) was calculated using $H = -\Sigma_{i=1}^{s} p_i (\ln p_i)$ where p_i is the frequency of the ith phenotype and s is the number of phenotypes in the population. H is a function of s. To remove this effect, evenness (J) was calculated as H/H_{max} (equivalent to $H/\ln s$), which ranges from 0 to 1 (Magurran 2004). In some cases, multiple statistical tests are performed but adjustments to the significance level below which the null hypothesis is rejected to allow for this were not made. Perneger (1998) and Moran (2003) discuss the mathematical, logical, and practical problems associated with standard adjustments for multiple statistical testing; for example, Bonferroni and sequential Bonferroni.

Results

Patterns in the extensive data on color-morph frequencies in *E. ovata* were sought on two levels: (1) variation in Nidderdale and (2) ''snapshot'' surveys encompassing different geographic scales.

Variation in the Nidderdale Populations

As well as considering the color polymorphism itself, variation at other loci is also relevant. The regulatory locus affects the expression of color, and this, together with the spotting and allozyme loci, provides genetic markers that illuminate the possible causes of population differentiation at the color locus.

Color-morph frequencies

Oxford and Shaw (1986) provided details of color-morph frequencies in the 44 sites sampled in Nidderdale between 1970 and 1984. Annual surveys of most of the sites continued until summer 1989, with additional sampling in 1995 and 1999. A few sites were scored for specific purposes in other years too. Since 1996 sites S, T, and AN have been used for experiments in which color-morph frequencies have been deliberately altered (G. S. Oxford, unpubl. data) by removal of individuals. Color-morph data for these sites and years are therefore omitted from what follows but the populations are used for the estimation of allele frequencies at the regulatory locus, which should not be affected by the perturbations.

Figure 1 charts the frequency of the *redimita* morph in populations adjacent to the principal road running through the study site (populations AG through N; see Oxford and Shaw 1986) over four time periods. Although sampling in Nidderdale began in 1970, information was available for very few sites prior to 1974. The final period is over 11 rather than five years because of more sporadic sampling after 1989. The fine-scale geographic variation in morph frequency is apparent during the more complete survey years of 1979–1983 and 1984–1988, with dramatic changes occurring on a scale of tens of meters. The distribution of *redimita* frequencies is retained during this 10-year period and is also clearly detectable in the less complete surveys of 1974–1978 and 1989–1999, when cumulative sample sizes per site were generally smaller. The broad-scale pattern of population differentiation has persisted for at least 30 generations.

Regulatory-locus variation

A consequence of the regulatory polymorphism is that males and females within a population can have very different color-morph frequencies and this could influence how selection operates on this system. In Nidderdale the *ovata* morph always develops early (Oxford 1983, 1985a) suggesting just one combination of alleles at the linked loci, C_oR_e. However, the *redimita* morph can develop early or late indicating the presence of two allelic combinations, C_rR_e and C_rR_l, respectively.

Between populations there is marked variation in the fre-

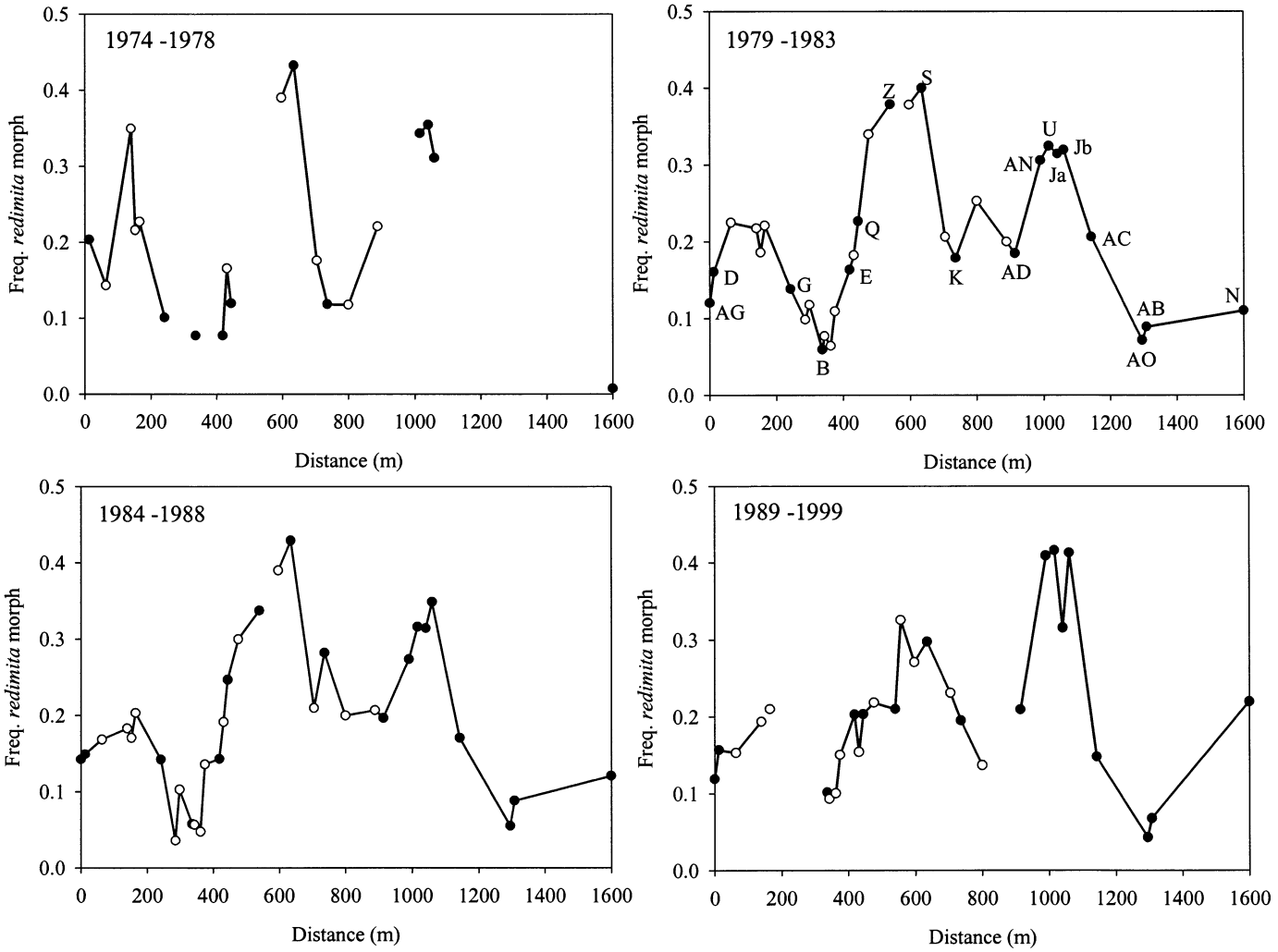

FIG. 1. Frequency of the *redimita* morph in sites ordered along the principal road running through the Nidderdale study area. The four plots show sequential time frames, as indicated. Sites referred to with respect to other genetic markers are shown as solid symbols and identified in the 1979–1983 panel.

quency of the early regulatory allele associated with C_r, ranging from zero to one over a matter of a few hundred meters (Oxford 1985a, unpubl. data). Figure 2 shows the distribution of this allele in populations along the road through the Nidderdale study site. Three principal areas are defined. One, to the southwest, has a high frequency of the early allele (populations AG to K) although within it three populations, D, Q, and S, have a frequency of the late regulatory allele significantly different from zero. This then gives way to a low-frequency area (populations AN to AB) via a transitional site, AD. Within this region, some sites provide no evidence at all for the early regulatory allele. Finally, the most northerly site, N, appears to be fixed for the early allele. Further sampling of sites to the southeast of D and G (sites I, O, A, L, AF) confirm that the early-allele frequency dramatically decreases with distance from G in a direction orthogonal to the main road, a pattern suggested in Oxford (1985a; Table 1A). Estimates of the frequencies of the early regulatory allele in different years are consistent within sites.

The impact of the late regulatory allele is to reduce the frequency of *redimita* in males and immature females, and thus affect the polymorphic diversity within and between populations through the season. Table 1A shows the estimated frequencies of early-developing *redimita* within the *redimita* class (Oxford 1985a, unpubl. data), the evenness (J) values for the early and late samples, and the effect of the late regulatory allele in changing evenness. In all populations, the value of J increases from early to late samples, including those for which there is no statistical evidence for the late regulatory allele. Detecting the late regulatory allele from field samples is at best crude (Oxford 1985a) and low frequencies may well be present in all Nidderdale populations. Because the frequency of the *redimita* morph in males and immature females is generally low, the effect of the late regulatory allele is to increase *redimita* and reduce *lineata* frequencies in mature females and thus increase J, which would be maximized if all morphs were at equal frequencies. There is no significant correlation between the frequency of the early regulatory allele within *redimita* and the frequency of the *redimita* morph (frequencies arcsine transformed) among

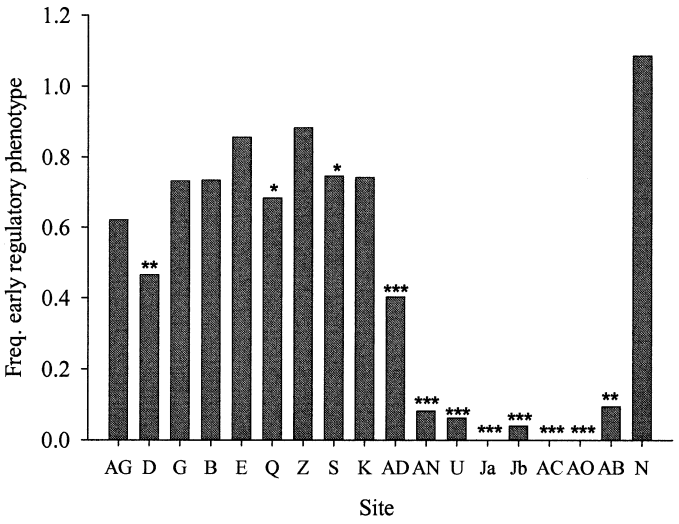

FIG. 2. Plot of estimated frequencies of the early-regulatory phenotype associated with the *redimita* morph in Nidderdale populations ordered in sequence along the principal road through the study site. The frequency of the early-regulatory phenotype in site N is greater than one because of the vagaries of sampling. The results of 2×2 contingency chi-squared tests assessing the null hypothesis that the frequency of the late-developing regulatory phenotype is zero are indicated: * $0.05 > P > 0.01$; ** $0.01 > P > 0.001$; *** $P < 0.001$. The actual distances between sites are shown in Figure 1. Data from (Oxford 1985a, unpubl. data).

populations. Similarly, no association is found when *redimita* and *ovata* are combined as a red-pigmented class.

Black spotting

The frequency of black spotting also varies between the Nidderdale populations as shown in Table 1A (Oxford 1989, unpubl. data). In three sites, spotting frequencies in males and females differ significantly (Oxford 1989), but for the purpose of investigating broad geographical patterns these data have been pooled within sites. Figure 3 shows the spatial pattern in populations arranged in order along the principal road. Overall, frequency differences between sites are highly significantly heterogeneous ($G_{13} = 270.2$, $P \ll 0.001$). A post-hoc simultaneous test procedure (Sokal and Rohlf 1995, pp. 723–724) identified a number of overlapping homogeneous sets. Three of them, together comprising all of the sites, did not overlap in terms of their statistical groupings, and these are indicated in Figure 3. It is clear that the heterogeneity between populations describes a pattern of peaks and troughs with high frequencies in sites B to Z; and, in N, low frequencies in sites U to AB and intermediate frequencies in sites D, G, K, and AD. Sites I, A, and L, not on the principal road but lying to the southeast of D and G, also show intermediate frequencies of spotting (Table 1A; Oxford 1989).

Allozymes

Allele frequencies at the two polymorphic loci assayed are shown in Table 2. Each locus possessed two major alleles with sporadic minor alleles. For both loci there is significant heterogeneity in allele frequencies among populations (*Pgi-2* vs. other alleles combined, $G_{10} = 94.94$, $P \ll 0.001$; *Amy-3* vs. other alleles combined, $G_9 = 25.39$, $P = 0.0026$). A post-hoc simultaneous test procedure identified two overlapping groups of sites for both *Pgi* and *Amy* (Fig. 4). *Pgi* presents a clear spatial pattern of peaks and troughs with just one site, K, linking the two groups. The picture for *Amy*, for which sample sizes are lower, is less clear with sites D and G forming a high *Amy-3* frequency group and sites E and Q a low frequency group. All other sites are common to these two groups. Evidence for linkage disequilibrium was sought between *Pgi* and *Amy*, and between these loci and the color phenotypes, using the procedures in POPGENE (http://www.ualberta.ca/~fyeh/info.htm). None was detected.

Snapshot Surveys

Color-morph frequencies

Color-morph frequencies in mature females have been scored in a number of geographical surveys. Figure 5 displays the frequencies of the three morphs for samples ($n \geq 40$) from all surveys except that conducted by Hippa and Oksala (1979) in which *redimita* and *ovata* were not separated. The most striking feature to emerge from Figure 5 is the extreme bunching of samples within phenotypic space. There is also no obvious relationship between the geographical scale of a survey and the degree of morph-frequency variation revealed (see below).

Ignoring for the moment the considerable heterogeneity between populations within surveys, weighted average morph frequencies for all samples, irrespective of sample size, are given in Table 3. In all cases but one there is concordance in the overall rank order of morphs with *lineata* the most, and *ovata* the least, common. The glaring exception to this pattern is in Finland where, overall, the *lineata* morph is in the minority and this is also the case for most individual populations (Hippa and Oksala 1979).

For comparisons of morph frequencies between surveys it is convenient to concentrate on the bottom recessive *lineata* morph, as this is present in all populations. Analyses were performed on arcsine-transformed *lineata* frequencies for samples of 40 or more individuals ($n = 1118$). In no case was the data distribution within each survey significantly different from normal (one-sample Kolmogorov-Smirnov test). Nidderdale was entered in the analysis twice, both as individual samples collected over space and time with sample sizes of 40 or more ($n = 439$), and also as aggregated samples within sites over time ($n = 44$). A one-way ANOVA showed that there are highly significant differences between surveys in the mean proportion of *lineata* ($F_{8,1162} = 35.47$, $P < 0.001$) with Finland showing the lowest frequency of *lineata* and eastern North America the highest. Post-hoc tests (Fisher's pairwise comparisons with individual error rate set to 0.05) revealed the groupings indicated in Table 4. Eastern North America and Finland are significantly different from all other surveys. The two Nidderdale entries were homogeneous and also differed significantly from all others. The remaining surveys combined in three overlapping groupings, as shown. The mean *lineata* frequencies in Tables 3 and 4 differ slightly because the latter have been derived from means calculated after arcsine transformation and are based on sample sizes of 40 or more individuals.

TABLE 1. Estimated frequencies of the early regulatory allele within the red-pigmented class in Nidderdale (Oxford 1985a, unpubl. data) and in Finland and other countries (Hippa and Oksala 1979). For each site, polymorphic evenness values (*J*) are given for the early sample (males and immature females) and the late sample (mature females) within a season, and the percentage change in *J* between the samples. In (A) all morphs present in the site contribute to the *J*-values whereas in (B) *redimita* and *ovata* are combined for this calculation. *fE*/*fT*, the frequency of red-pigmented morphs in the early sample divided by the frequency in the late sample, estimates the frequency of the early regulatory allele. Percent change, [*J*(late) − *J*(early)]/*J*(late) × 100; *P*, probability associated with 2 × 2 contingency chi-squared or Fisher's exact tests on, in (A), numbers of *lineata* and *redimita* and in (B), numbers of *lineata* and (*redimita* + *ovata*), in early and late samples. This tests the null hypothesis that the proportion of late-developing morphs is zero. ns, not significant. *fS*, frequency of the black spotting phenotype. In (B), sites 2–15 are in Finland; 21–23, Denmark; 24, Germany; 27–28, Switzerland.

(A) Nidderdale; *redimita* and *ovata* separated							(B) Finland and other countries; *redimita* and *ovata* combined					
Site	*fE*/*fT*	Early sample *J*	Late sample *J*	Percent change	*P*	*fS*	Site	*fE*/*fT*	Early sample *J*	Late sample *J*	Percent change	*P*
A	0.273	0.175	0.384	54.4	<0.001	0.638	2	0.492	0.908	0.928	2.12	0.001
B	0.734	0.456	0.480	4.9	ns	0.864	5d	0.453	0.906	0.871	−3.92	<0.001
D	0.465	0.218	0.356	38.8	0.003	0.683	6	0.588	0.811	0.984	17.55	0.033
E	0.856	0.546	0.568	3.9	ns	0.974	7	0.405	0.872	0.851	−2.51	<0.001
G	0.732	0.369	0.469	21.4	ns	0.627	8	0.493	0.946	0.830	−13.91	<0.001
I	0.751	0.723	0.745	3.0	ns	0.622	9b	0.422	0.822	0.966	14.87	<0.001
Ja	0	0	0.924	100.0	<0.001	0.318	11	0.326	0.750	0.928	19.22	<0.001
Jb	0.040	0.084	0.617	86.5	<0.001	0.443	12	0.780	0.791	0.887	10.82	ns
K	0.742	0.505	0.524	3.6	ns	0.659	14	0.605	0.999	0.738	−35.39	<0.001
L	0.027	0.040	0.635	93.7	<0.001	0.519	15	0.621	0.999	0.754	−32.63	<0.001
N	1.084	0.445	0.516	13.7	ns	0.758	21	0.964	0.746	0.760	1.93	ns
O	0.200	0.221	0.558	60.3	<0.001	—	22	0.422	0.478	0.801	40.30	0.019
Q	0.683	0.543	0.619	12.3	0.014	0.908	23	0.376	0.544	0.915	40.58	ns
S	0.746	0.801	0.911	12.1	0.023	—	24	0	0	0.913	100.00	0.013
U	0.062	0.210	0.698	69.9	<0.001	0.434	27	0	0	0.552	100.00	ns
Z	0.882	0.888	0.931	4.6	ns	0.907	28	0	0	0.497	100.00	ns
AB	0.094	0.072	0.440	83.8	0.003	0.270						
AC	0	0.078	0.590	86.8	<0.001	—						
AD	0.403	0.247	0.465	46.8	<0.001	0.616						
AF	0.038	0.067	0.740	90.9	<0.001	—						
AG	0.621	0.439	0.601	26.8	ns	—						
AN	0.082	0.220	0.660	66.7	<0.001	—						
AO	0	0	0.398	100.0	<0.001	0.370						

The overall rank order of morphs mentioned above also applies to the majority of individual collections. In samples of 40 or more spiders across all surveys in which *redimita* and *ovata* were scored separately, 92.3% had *lineata* > *redimita* > *ovata* and 97.9% had *lineata* as the most frequent morph. The Nidderdale samples were entered as separate scores for each site in each year and these of course will not be independent. Excluding Nidderdale makes virtually no difference to the figures (now 93.6% and 96.7%, respectively). The pattern of samples with specific rank orders of morphs varies little as sample size increases (Table 5). However, if samples of 10 to 30 individuals (median 19.5, $n = 35$) and those of 40 or more (median 51, $n = 672$) are compared for numbers in rank orders *lineata* > *redimita* > *ovata* versus other orders, the difference is highly significant (Fisher's exact test, $P = 0.00063$). The frequency of *lineata* > *redimita* > *ovata* is lower in the smaller samples, but is still the predominant rank order (75% versus 93.6%). Monomorphic *lineata* samples will be scored within the rank-order category *lineata* > *redimita* = *ovata*. In the small sample-size group, four of five collections in this rank-order category were monomorphic (80%) compared with 12 of 14 (86%) in the large sample-size group. Some of the apparently monomorphic populations may actually be polymorphic but with a second morph not detected because of the vagaries of sampling. Upper 95% confidence limits for a second morph can be calculated for these samples using the adjusted Wald method (Samuels and Witmer 2003). For the smallest monomorphic sample, $n = 13$, the upper limit is 0.271, whereas for the largest, $n = 117$, the upper limit is 0.039.

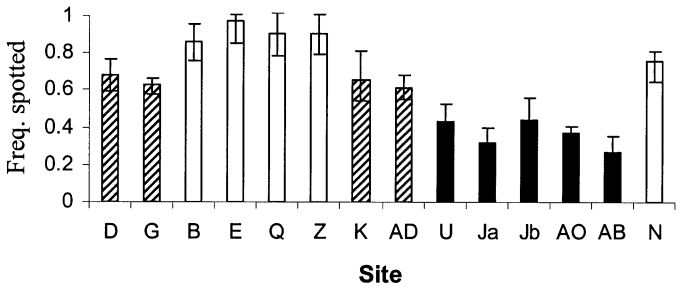

FIG. 3. Plot of estimated frequencies of the black-spotted morph in Nidderdale populations ordered in sequence along the principal road through the study site. The different shadings refer to three homogeneous groupings revealed by a post-hoc simultaneous test procedure (see text). Error bars are the 95% confidence intervals based on the adjusted Wald procedure (see text). The actual distances between sites are shown in Figure 1. Data from Oxford (1989, unpubl. data).

Variances in the frequency of *lineata* among surveys were compared using the jackknife method. A one-way ANOVA revealed a highly significant difference ($F_{8,1155} = 7.98$, $P < 0.001$) and post-hoc pairwise comparisons indicated the groupings shown in Table 6. The variances are unrelated to the area the surveys covered (Table 3) (Spearman's rank cor-

TABLE 2. Genetic variation at two enzyme-coding loci in Nidderdale populations. Numbers of each allele are shown together with the frequencies of alleles *Pgi-2* and *Amy-3*, which are plotted in Figure 4.

Site	No. of alleles				Total alleles	Frequency of *Pgi-2*	Site	No. of alleles				Total alleles	Frequency of *Amy-3*
	Pgi-1	*Pgi-2*	*Pgi-3*	*Pgi-4*				*Amy-1*	*Amy-2*	*Amy-3*	*Amy-4*		
D	0	26	103	1	130	0.208	D	0	23	25	0	48	0.521
G	1	7	40	0	48	0.167	G	1	24	23	0	48	0.479
B	0	43	91	0	134	0.321	B	0	33	15	0	48	0.313
E	0	53	79	0	132	0.402	E	0	36	10	2	48	0.208
Q	0	56	80	0	136	0.412	Q	0	40	8	0	48	0.167
K	0	30	100	0	130	0.231	K	2	25	20	1	48	0.417
AD	0	12	112	0	124	0.097	AD	0	25	15	0	40	0.375
J	0	16	114	0	130	0.123	J	0	40	8	0	48	0.167
AC	0	10	72	0	82	0.122	AB	0	33	15	0	48	0.313
AB	1	39	88	0	128	0.313	N	0	32	16	0	48	0.333
N	0	16	120	0	136	0.118							

relation $r_s = -0.323$, df = 6, $P = 0.435$, Nidderdale means omitted) reinforcing the impression from Figure 5 that variation in color-morph frequencies is scale independent.

In two studies, attempts have been made to seek associations between morph frequencies and a small number of physical (geographic, altitudinal, and climatic) factors. In Britain, the frequency of the *lineata* morph was most highly correlated with wetness indices (positive association with annual rainfall or number of wet days) in most of the analyses, although the amount of variation explained by this relationship in the total dataset was only 15% (Oxford 1985b). In continental Europe, by contrast, January temperature was the most highly significant explanatory variable (negative association) for populations on a Netherlands to Italy transect, and July temperatures (negative association) for populations along the west coast of France (Oxford and Reillo 1993). The different environmental factors highlighted as the most important explanatory variable in these surveys may, in part, be a result of the use of area surveys in Britain and approximately linear transects in mainland Europe. For example, if a transect were taken up the west coast of Britain, where the climate is wetter than in the east, wetness indices may no longer be the most important factors influencing morph frequencies. To test this proposition, data from the most westerly 100 km Ordnance Survey map squares sampled by Oxford (1985b; squares NC, NG, NM, NN, NS, NX, SD, SH, SN, SS, SW, and SX) were subjected to logistic multiple-regression analysis, the technique employed by Oxford and Reillo (1993). A total of 117 samples was included, all consisting of 40 or more individuals, and a binomial probability distribution assumed. The explanatory environmental variables are described in Oxford (1985b). In the final regression model (data not shown) by far the most significant relationship ($P \ll 0.001$, $r^2 = 0.14$) was a negative one between the frequency of the *lineata* morph and January temperature, which parallels the results from the Netherlands to Italy transect in continental Europe. Thus, the lack of concordance between surveys in the environmental factor identified as the most important correlate of *lineata* morph frequencies may depend to some extent on the sampling protocol employed and need not detract from the fact that in all surveys very powerful associations were identified using a small number (seven in the European survey; nine in the British) of key physical explanatory variables. In all cases only a small proportion of the total variation in color-morph frequencies was explained by the relationships.

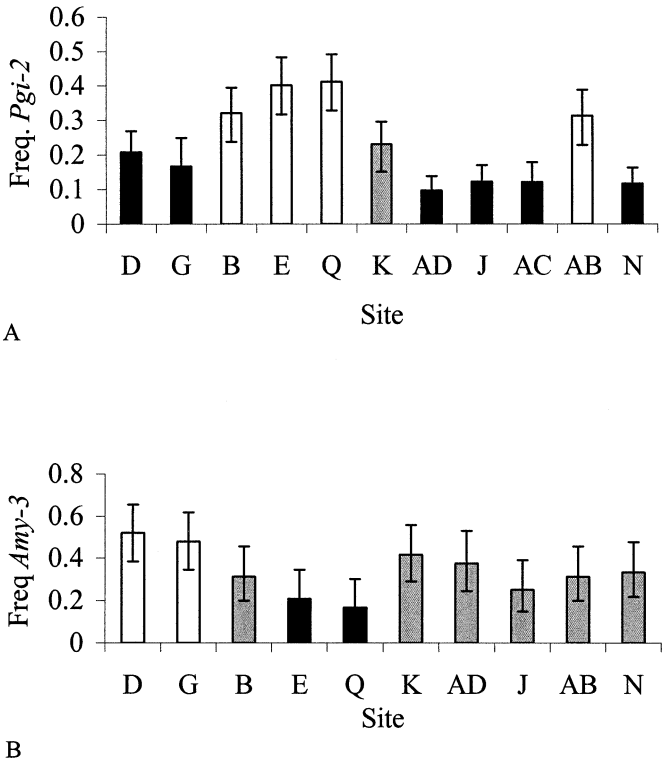

FIG. 4. Plots of the frequencies of allozyme alleles *Pgi-2* (A) and *Amy-3* (B) in Nidderdale populations ordered in sequence along the principal road through the study site. Sets of sites indicated in black and white are significantly different from each other using a post-hoc simultaneous test procedure. Those shown in gray are not significantly different from either the black or the white sets. Error bars are the 95% confidence intervals based on the adjusted Wald procedure (see text). The actual distances between sites are shown in Figure 1.

Regulatory-locus variation

As mentioned earlier, the regulatory-locus variation is important in that it determines when, and in which sex, the color alleles are expressed, and is one factor affecting the visible diversity within populations. Other than Nidderdale,

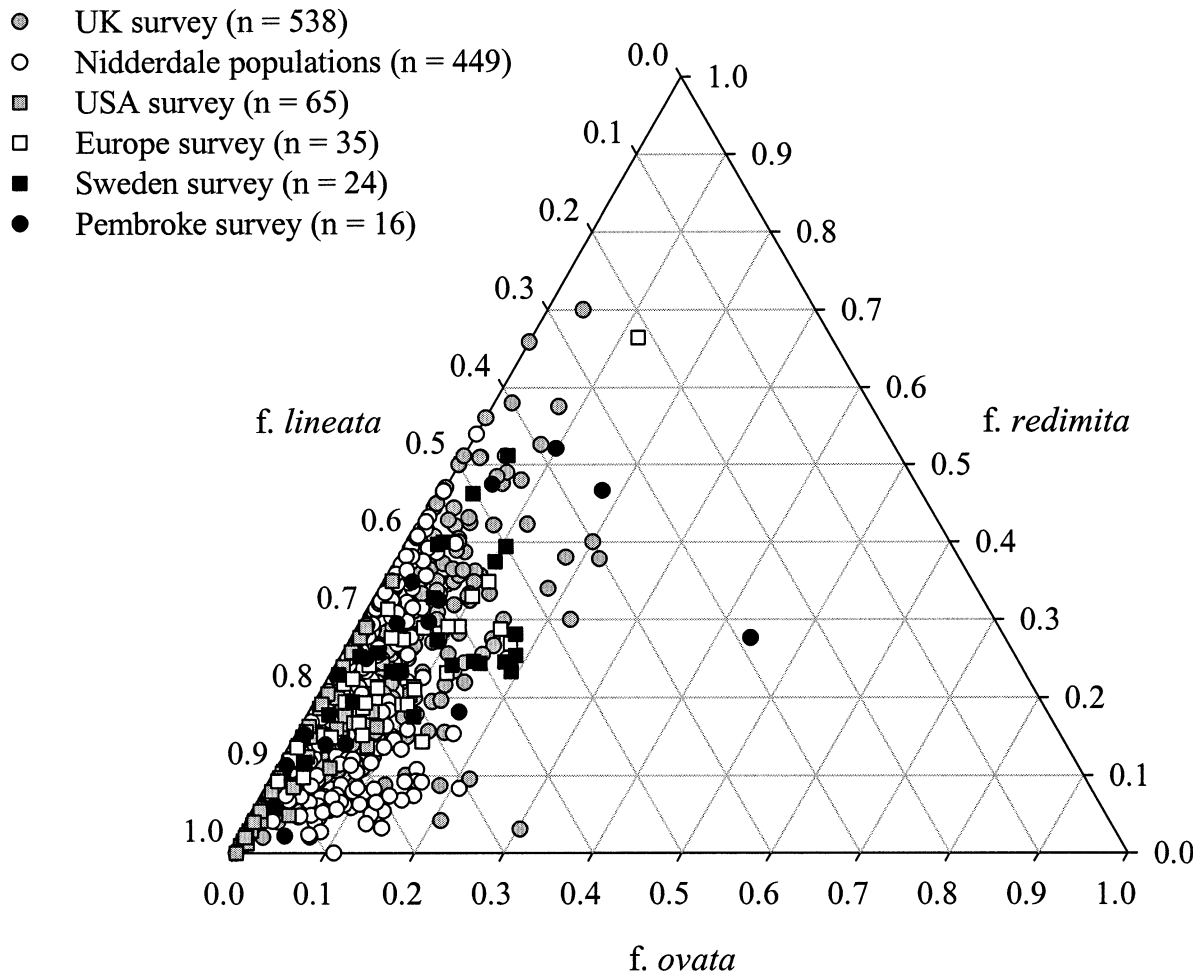

FIG. 5. Plot of *lineata*, *redimita*, and *ovata* morph frequencies for all samples ($n \geq 40$) across six geographical surveys. The Nidderdale data are included within the United Kingdom survey as a single entry averaging all sites over all years.

information on the regulatory polymorphism is available principally for two areas, eastern North America and Finland, although there is both rearing and statistical evidence for the presence of the late regulatory allele elsewhere in Britain (Oxford 1983, 1992, unpubl. data) and central Europe (Hippa and Oksala 1979).

Of the coastal Maine populations of *E. ovata* investigated by Reillo and Wise (1988c), 18 are analyzable for the presence of the late regulatory allele (their table 2). None of the samples was heterogeneous when mature males and immature females were compared, and only one was significant when these were pooled and compared with mature females. Many expected values fell below five including the one significant comparison (population G, $\chi^2_2 = 8.871$, $P = 0.012$) where four of the six expected values were in this category. Low expected values will lead to an inflated rejection rate. Given this, and the number of tests performed, there seems little indication that the late regulatory allele is present in any of the populations sampled, although low frequencies will not be detected by this method. If the late regulatory allele is present in some or all populations at low levels, there should be a trend such that red-pigmented morphs are generally more frequent in mature females than in mature males plus immature females. If a + is assigned to a sample where the frequency of a red-pigmented morph is higher in mature females, and a − for deviations in the other direction, for *redimita* the ratio of +:− is 7:11, and for *ovata* 10:7. The *redimita* ratio is in the opposite direction to that expected if late regulatory alleles are present, and the *ovata* ratio is not significant different from 1:1 on a one-tailed binomial test. A one-tailed test is appropriate here because only a deviation in the direction of higher frequencies of red-pigmented morphs in mature females is of interest. The lack of evidence for the late regulatory allele is further supported by the results of laboratory rearing studies in which all *redimita* and *ovata* morphs derived from two coastal Maine populations developed their patterns early in development (Reillo and Wise 1988a).

In a later study, Reillo (1989) sampled from a much broader geographical area in eastern North America and for many locations included scores for both males and females. Unfortunately he combined penultimate and mature stages within the sexes on the assumption that the early-developing allele is fixed in North American populations. In males, no differences in morph frequencies between these two developmental stages are expected but in females combining will blur the

TABLE 3. Mean color-morph frequencies in mature female *Enoplognatha ovata* in different regional surveys, ordered according to increasing frequency of *lineata*. Weighted frequencies are those derived by totaling numbers across all populations (irrespective of sample size) so that large samples contribute proportionately more than small samples to the overall means. Data sources: Nidderdale, U.K. (Oxford 1976; Oxford and Shaw 1986; G. S. Oxford, unpubl. data); Pembrokeshire, U.K. (Oxford 1991); Finland (Hippa and Oksala 1979); Sweden (G. S. Oxford and B. Gunnarsson, unpubl. data); U.K. 1980 and 1981 (Oxford 1985b); North America (Reillo 1989); Europe (Oxford and Reillo 1993). Areas (km^2) are very approximate and include intervening land where sampling was along transects (Europe, North America) and across two island groups separated by mainland (Sweden). Note that for Finland, the distinction between *redimita* and *ovata* morphs was not recognized. *n*, the total number of spiders scored during the surveys.

	Weighted frequencies				
	lineata	*redimita*	*ovata*	*n*	Approx. area (km^2)
Finland	0.3187	0.6813		1393	2800
Sweden	0.6157	0.2963	0.0880	3193	25,000
Pembrokeshire	0.6826	0.2342	0.0833	1405	140
Europe	0.7370	0.2070	0.0560	4768	735,000
U.K. 80	0.7345	0.2358	0.0297	12,867	229,000
U.K. 81	0.7464	0.2194	0.0342	22,810	229,000
Nidderdale	0.7887	0.1881	0.0233	86,679	1.4
North America	0.8507	0.1264	0.0228	13,755	437,000

distinction between early- and late-developing morphs. Nevertheless, a comparison of morph frequencies in males and females could suggest the existence of the late regulatory allele. Data were analyzed from Reillo's sites 1 through 35 but excluding populations 2 to 5, considered in more detail above. Only two of the samples showed significant heterogeneity between morph frequencies in males and females (31, $\chi^2_1 = 5.367$, $P = 0.021$; 32, Fisher's exact test $P = 0.020$), and both were in the direction of a higher proportion of *redimita* in females. If the direction of morph-frequency differences between males and females is scored as before, then for *redimita* the +:− ratio is 19:9 (null hypothesis 1:1, one-tailed binomial test, $P = 0.044$), and for *ovata* 2:3. Thus, *redimita* tends to occur at a higher frequency in females than in males, and this is just formally significant. Seligy (1971) reared material from New York State and Ontario but made no mention of color patterns appearing for the first time in mature females or being absent in males. Thus, although there is no evidence from reared spiders for the presence of the late regulatory allele in eastern North American populations, the trends in Reillo's (1989) data are in a direction consistent with a low-frequency occurrence. Of course, other factors could also lead to a reduced frequency of *redimita* in males.

On the basis of material scored in a number of sites in Finland, Hippa and Oksala (1979) concluded that *redimita* and *ovata* morphs are extremes of a continuum of red pigmentation. Their spiders were therefore categorized as white (*lineata*) or red (*redimita* + *ovata*). Some of their collections allow the estimation of the early and late regulatory alleles in the composite ''red'' morph (Table 1B). In no case was there significant heterogeneity in morph frequencies between males and immature females, thus, these were pooled for comparison with mature females. Statistical evidence for the presence of the late regulatory allele was found in nine of the 10 Finnish collections (sites 2 through to 15, Table 1B), but the effect on the polymorphic diversity index in early and late samples varies. In some populations, morph frequencies were more equal in mature females compared with males and immature females (*J* increases) but in others the reverse was true (*J* decreases). For comparison, the overall effect on *J* of combining *redimita* and *ovata* morphs in Nidderdale was very slight (data not shown) because the *ovata* morph is rare or absent in most populations. The average frequency of the early regulatory allele (within the combined red class) is similar in the Nidderdale and Finnish samples (Nidderdale 0.426, Finland 0.520, $F_{1,27} = 0.43$, nonsignificant, frequencies arcsine transformed). However, the average frequency of the red morph in mature females is very different (Nidderdale 0.212, Finland 0.644, $F_{1,27} = 76.78$, $P < 0.001$, frequencies arcsine transformed). It is a combination of high frequencies of the red morph and of the late regulatory allele that leads to the reduction of polymorphic evenness in some Finnish populations over the reproductive season.

Hippa and Oksala (1979) also reported data from European countries other than Finland. Some of these collections were subsequently found also to contain a sister species, *E. latimana*, which shares the color polymorphism (Hippa and Oksala 1982). Only the pure *E. ovata* samples are considered in Table 1B (sites 21 to 28). Sample sizes were generally small, but for two sites (22, Rödby, Denmark; 24, Hamburg, Germany) there is statistical evidence for the late regulatory allele, and all show deviations in a direction consistent with the presence of this allele. All populations show an increase in the polymorphic evenness index between males and immature females, and adult females. Finally, Gerhardt (1921) sampled *E. ovata* in the vicinity of Breslau (now Wroclaw), Poland. Here, females exhibited all three color morphs, with *lineata* and *redimita* at comparable frequencies, but males were mostly *lineata*, seldom *redimita*, and never *ovata*. This suggests that in this population the *ovata* and some of the *redimita* color alleles were *cis* with the late regulatory allele.

TABLE 4. Mean *lineata* frequencies (back-transformed from radians) from the eight surveys and the homogeneous groupings indicated by post-hoc pairwise tests. Surveys joined by a line are not significantly different from one another. Nidderdale was entered twice, once as all individual samples collected over both space and time with sample sizes of 40 or more ($n = 439$, Nidd), and once as aggregated samples within sites over time ($n = 44$, Nidd M). Note that back-transformed mean morph frequencies differ slightly from the means calculated from raw frequencies in all samples in Table 3.

Finland	Sweden	Pembrokeshire	Europe	U.K. 80	U.K. 81	Nidd M	Nidd	N. America
0.3525	0.6201	0.6829	0.7328	0.7457	0.7518	0.7913	0.7976	0.8946

TABLE 5. Rank orders of the three color morphs *lineata* (l), *redimita* (r), and *ovata* (o) and their frequencies in samples of differing sample sizes (all surveys except Nidderdale and Finland).

	Sample size					
Rank	10–19	20–29	30–39	40–49	50–99	>100
l > r > o	0.6667	0.8235	0.9500	0.9342	0.9261	0.9519
l > r = o	0.1667	0.1176	0.0500	0.0175	0.0311	0.0107
l = r > o	0.1111	0	0	0.0307	0.0350	0.0107
l > o > r	0	0.0588	0	0.0044	0.0078	0
r > l > o	0	0	0	0	0	0.0053
o > l > r	0.0556	0	0	0.0132	0	0.0214
Total	18	17	20	228	257	187

DISCUSSION

The data from both the temporal and spatial surveys presented above, together with other already published information, can be used to explore the nature of the evolutionary forces that impinge on the color polymorphism in *E. ovata*. Before discussing this it is necessary to consider the nature of the color variation. Most workers have reported no problems distinguishing three distinct color morphs except for the rare occurrence of individuals that seem to be intermediate between *redimita* and *ovata*. These are found more frequently in laboratory-reared material and may represent *ovata* in which the red pigment has, for environmental reasons, not developed fully (Oxford 1983). On the basis of spiders collected in Finland, however, Hippa and Oksala (1979) questioned the validity of the classical morphs and claimed there was continuous variation for red pigmentation between *redimita* and *ovata*. Critically, they also sampled in Sweden, Denmark, Germany, Switzerland, and Italy without revising their opinion regarding the nature of the red-pigment variation. Others who have surveyed *E. ovata* in these countries have had no difficulties scoring the traditional morphs (Nielsen 1932; Oxford and Reillo 1993; G. S. Oxford and B. Gunnarsson, unpubl. data). It seems likely that Hippa and Oksala overinterpreted the intramorph variation inevitably present between individuals.

The enigmatic nature of the forces affecting color variation in *E. ovata* becomes obvious when attempts are made to reconcile conclusions from local- and large-scale surveys, temporal trends, perturbation experiments, and the nature of species-level attributes. Some aspects show clear indications that stochastic forces are paramount in determining morph frequencies, whereas others strongly suggest the action of powerful natural selection. The strands of evidence for stochastic, and for selective, processes can now be drawn together.

Evidence for Stochastic Forces

The most intensively studied populations are those in Nidderdale where the degree of spatial genetic differentiation among sites situated along the verges of the principal road through the study area is exceptional. All populations occur within a linear distance of only 1.6 km yet show marked peaks and troughs on a scale of 10 to 100 meters, the broad patterns of which have remained stable for at least 30 generations (Fig. 1). Four lines of evidence suggest that powerful local selection is absent. First, experiments both in Nidderdale (Oxford and Shaw 1986) and in North America (Reillo and Wise 1988b) in which populations have been artificially perturbed away from their current morph frequencies have failed to show the rebound expected if local selection were operating (Endler 1986). Second, Oxford and Shaw (1986) modeled the Nidderdale populations and tested whether the observed annual variation in the estimated frequencies of the *lineata* allele was compatible with genetic drift. They found it was and that incorporating directional selection into the model did not improve the fit for the majority of populations. However, the degree of variation over time was less than would have been expected on the basis of drift alone. This could be a result of low-level migration although some form of balancing selection could not be eliminated. Third, the lack of scale-dependent variability between populations in the various surveys (Table 6, Fig. 5) is difficult to reconcile with the selective determination of morph frequencies. The fact that the small study area in Nidderdale houses populations of comparable variability to those sampled over vastly greater scales in North America and Europe, and that the variance among local populations in Pembrokeshire, Wales, is the highest of any survey, challenges explanation in terms of selection by climatic factors or predation. Finally, in Nidderdale (and elsewhere) there are no associations between morph frequencies within a site and obvious environmental

TABLE 6. Comparison of the variances in *lineata* morph frequencies across surveys, arranged in order of increasing value. Nidderdale is entered twice, once based on numbers for a site totaled across all sampling years (Nidd M) and once for individual samples from sites in all years (Nidd). Variances are shown although the homogeneous groupings indicated were generated by post-hoc pairwise comparisons after one-way ANOVA on jackknifed pseudovalues. Solid lines link surveys for which the variances were not significantly different. The dotted line links two nonadjacent surveys that were not significantly different.

Nidd M	Nidd	Europe	N. America	U.K. 81	U.K. 80	Sweden	Finland	Pembrokeshire
0.0064	0.0082	0.0086	0.0092	0.0143	0.0155	0.0169	0.0241	0.0426

features such as altitude, aspect, or vegetation type, all of which will affect the microclimate and other attributes such as prey (and perhaps predator) abundance. If selection is determining morph frequencies within populations, one has to postulate local selective forces that change spatially over short distances but are constant over decades. Obvious candidates when dealing with a color polymorphism are sight-hunting predators such as birds, but it is inconceivable that selection from such a source would be so constant on the spatial and temporal scales involved here.

If selection is not the reason for the microspatial morph-frequency patterns in Nidderdale, what is? Oxford and Shaw (1986) argued that intermittent drift could be responsible for establishing the current distributions and highlighted a major disturbance in the area during the late 1940s when the principal road was widened and metaled. They suggested that, as a result of this event, only small pockets of spiders survived and that within them stochastic forces led to disparate allele frequencies. These pockets then acted as sources for recolonization as the broad-leaved vegetation on the verges reestablished. If this is the correct explanation, evidence for this perturbation should be discernable at other loci within the same populations. The datasets for the distributions of black spotting, regulatory alleles, and allozymes presented here confirm a highly significant degree of population differentiation, with a distinct spatial pattern of peaks and troughs along the linear array of sites (Figs. 2–4). Thus, five genetic markers show patterns of spatial differentiation of roughly similar wavelengths, an observation consistent with the intermittent drift hypotheses. Patterns of variation at the marker loci are not all independent of one another. A correlation matrix across all Nidderdale sites of the arcsine-transformed frequencies of the early regulatory phenotype, spotting, *Pgi-2*, *Amy-3*, and the three color morphs revealed three significant associations, one of which, the inevitable negative correlation between the frequencies of *lineata* and *redimita*, is of no interest. The other two were between the frequency of the early regulatory allele and spotting ($r = 0.805$, df = 15, $P = 0.0001$), and between the frequency of the *ovata* morph and spotting ($r = 0.616$, df = 15, $P = 0.008$). Although such associations could be generated by linkage disequilibrium and/or correlated selection, some parallel distributions might also be expected on the intermittent drift hypothesis, especially if the number of populations surviving the perturbation was small.

Evidence for Natural Selection

The lines of evidence for selection acting on the color polymorphism are all indirect, but cumulatively they are very persuasive. First, the color polymorphism in *E. ovata* has been maintained for a very long period of time (probably in excess of 100,000 years), as attested by its presence in a sister species, *E. latimana* (Snazell 1983; Oxford 1985b, 1991, 1992; Oxford and Reillo 1993). All three color morphs are common to both species, although, to date, only one individual of *E. latimana* has been reported with the *ovata* pattern (G. S. Oxford in Harvey et al. 2002). Both species also share the regulatory and black-spotting systems (Oxford 1992), so similar conclusions must apply to these loci too. The presence of polymorphisms that transcend species boundaries strongly suggests that they have been retained from a common ancestor and are therefore highly unlikely to be selectively neutral (Golding 1992; Richman 2000).

Second, the polymorphism is present in virtually all populations. The surveys of color-morph frequencies involved 1171 samples containing 40 or more spiders, and of these just 12 (1.0%) were monomorphic for color, in all cases *lineata*. Seven of these were from North America, three from the British (1981) survey, and two from Nidderdale. It is possible that these monomorphic samples were taken from polymorphic populations with low frequencies of the red-pigmented morphs. As shown above, even the population from which the largest monomorphic sample ($n = 117$) was taken could contain a second morph at a frequency of about 4%. Given the preponderance of the bottom recessive *lineata* morph in most populations, a high proportion of new colonies founded by single gravid females would be expected to be monomorphic. Indeed, the probability that a random, single-mated female from a population in which the *lineata* morph frequency in both sexes is 0.7 would found a new, monomorphic *lineata* colony is almost 50%. This percentage would be greater in eastern North America where *lineata* frequencies are much higher, although with red-pigmented morph frequencies so low the stochasticity of sampling would be expected to yield higher numbers of apparently monomorphic populations. Of course, it is not known how often colonies are founded by lone mated females. Many *E. ovata* populations are relatively small and isolated, conditions that would be expected to lead to a high population turnover and repeated founder events. This population structure is also conducive to continuous genetic drift, which will eventually lead to monomorphism. Although survey sample size and effective population size will be very poorly correlated except in Nidderdale, it is the case that samples of less than about 30 (in the British surveys, which involved data gathering by the general public, a minimum sample size of 40 was set) tend to come from small, low-density populations. Yet the vast majority of even small samples are polymorphic for color. So, despite the spatially patchy nature of the habitat favored by *E. ovata* and the opportunities this should create for stochastic fixation of morph frequencies, the occurrence of polymorphism within populations of this species is remarkably high and suggests that the color variation is actively maintained by selection. In small populations, this selection would have to be strong to overcome stochastic forces and this might explain why, if founder populations are monomorphic for *lineata*, any *redimita* alleles introduced via occasional migration successfully establish (for example site N in Nidderdale discussed by Oxford and Shaw 1986, and below).

Third, there is great stability in the relative frequencies of the three color morphs. Although mean frequencies of the *lineata* morph do vary significantly among surveys, the rank order of *lineata* > *redimita* > *ovata* is preserved on average within surveys and in the vast majority of individual samples (Table 5, Fig. 5). Over 90% of adequately sampled populations ($n \geq 40$), and even 75% of small samples ($10 \leq n \leq 30$), show this pattern. This consistency, across large geographical scales, is not compatible with what one might expect if the color phenotypes are selectively neutral. This con-

clusion is reinforced by the fact that the rank order of *lineata* and *redimita* in *E. latimana* is, for the majority of samples, the same as in *E. ovata* (Oxford and Reillo 1993). Not included in the consideration above are the data from Finland (Hippa and Oksala 1979) where, overall, *lineata* is in the minority (0.319).

A fourth line of evidence for the selective maintenance of the color polymorphism comes from the blatant nature of phenotypic differences, suggesting that visible appearance rather than some pleiotropic effect of the color alleles (Goodhart 1987) is a crucial feature. The distinctiveness of the variation is greatly enhanced by the deposition of a continuous layer of matt-white guanine below the hypodermis, the tissue containing the yellow and red pigments. Occasionally individual spiders are found with a much reduced guanine layer and there the background against which the colors are contrasted is the brown of the digestive mass. In these individuals the red and yellow hypodermal pigments are extremely difficult to discern. Therefore, if pleiotropic effects of the color alleles are the reason for the maintenance of the polymorphism, the visible aspects, if selected against, could be virtually eliminated. It has been argued elsewhere (Oxford 1998) that storage of guanine is metabolically costly and would not be expected to evolve in the absence of overriding selection for its colorant effect. Both of these considerations suggest that color per se is a key element of the polymorphism.

Finally, the large-scale associations between color-morph frequencies and climatic variables also suggest selection. In the United Kingdom, the possibility that morph frequencies and the major explanatory climatic variables, wetness indices, were both contingent on geography was examined and rejected (Oxford 1985b). The climatic associations for the two European transects were inconsistent inter se and with the U.K. results, but as shown by subsampling the British data, the survey design can affect conclusions. The important point is that despite using a small number of climatic variables (to reduce Type I errors) powerful associations were found in all three surveys. The mechanistic process(es) connecting climatic factors with morph frequencies are unknown.

The morph frequencies discussed above are those of mature females. The late regulatory allele, when present, has a major effect on the polymorphic diversity of color within populations during the active season. Certainly in British and central and north European populations this allele seems to be widespread. In some Nidderdale sites with just the *lineata* and *redimita* morphs and with the late regulatory allele fixed, the color polymorphism is only manifest in mature females and is openly exhibited for only one or two weeks before females establish themselves within rolled leaves (Oxford 1985a). This has clear implications for some possible modes of selection, for example visual predation. However, there are no significant associations between the arcsine-transformed frequencies of the early regulatory allele and of the adult female *lineata* morph in either data set (Nidderdale, $r = -0.055$, df $= 21$, $P = 0.804$; Finland/Europe, $r = -0.557$, df $= 9$, $P = 0.075$), and the overall effect of the regulatory variation on polymorphic diversity is inconsistent; increasing it in Nidderdale and generally decreasing it in Finland.

Synthesis

Thus, there are two sets of arguments regarding the evolutionary significance of the color polymorphism in *E. ovata* that seem mutually exclusive. One set suggests strongly not only that selection is maintaining the polymorphism, and has done so for a long period of time, but also that the overall rank order of morph frequencies is tightly controlled, even in small populations. The other set of arguments, based largely on perturbation experiments and analyses of local spatial and temporal variation, provides no evidence for such selection. However, any selection must act at the local scale.

How can these opposing arguments be reconciled? There seem three possible models. First, the color variation considered here might reflect the ghost of selection past (e.g., Janzen and Martin 1982)—it may once have been selectively maintained but, because of environmental changes, selection is no longer acting. The polymorphism and current morph frequencies distributions may retain the signal of previous selection because of inertia inherent in large, highly structured populations. Second, epistatic and genotype-environment interactions may determine optimum color-morph frequencies within local populations; that is, there is a rugged fitness landscape (Gavrilets 2004). Finally, strong selection protecting the polymorphism may operate only at extreme allele frequencies; at intermediate frequencies selection is so weak that stochastic forces predominate in local populations.

On the first hypothesis, selection, at some time in the past, became weaker or disappeared entirely leaving the color variation as a neutral character. Although there are no data with which to address this hypothesis directly, the persistence of the almost ubiquitous morph-frequency patterns described above requires large, structured populations. However, many populations are small and occupy isolated broad-leaved vegetation islands in landscapes that are otherwise unsuitable, for example, pasture, sand dunes, and moorland; yet these too are polymorphic for color and have the same rank order of morphs. Even populations in extensive vegetation patches undergo fluctuations in size of at least an order of magnitude, which in low years can yield female population estimates well below 100 (Oxford 1993b). There is ample opportunity for both continuous and intermittent drift to generate populations fixed for one allele and yet they are extremely rare. The second hypothesis, where color-morph frequencies reflect local fitness peaks, is at odds with the lack of movement following perturbations. It is also inconsistent with the observation that four other marker loci show very similar but not coincidental patterns of spatial differentiation in the Nidderdale populations. If epistatic or genotype-environment interactions are invoked to explain the peaks and troughs for color then, for consistency, similar interactions must be responsible for differentiation at the other loci too. The chance of this, on such a small spatial scale, is unlikely.

The third hypothesis, of intermittent selection, leads to an empirical model of how, and when, selection may operate. This is encapsulated in the hypothetical Δq on q plot in Figure 6, expressed in terms of the frequency of the *redimita* allele. The polymorphism certainly seems to be protected in some way so the plot has a negative slope where it crosses the $\Delta q = 0$ axis. However, the gradient at this point is extremely

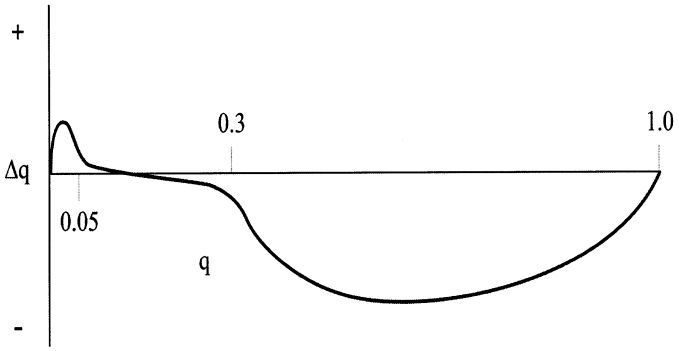

FIG. 6. Hypothetical plot of Δq on q, where q = the frequency of the *redimita* allele. The contribution of *ovata* to the polymorphism is ignored. The allele-frequency values of 0.05 and 0.3 are those that yield the observed range of morph frequencies of *redimita* in British and mainland European populations.

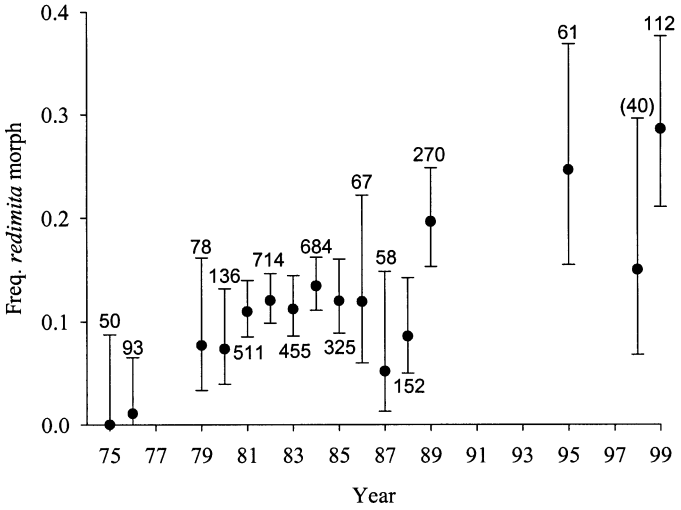

FIG. 7. Plot of *redimita* frequency against year in site N, Nidderdale. Sample sizes are shown for each point and will approximately equal the mature female population size (Oxford 1993b). The exception is for 1998 when a predetermined number of spiders was sampled for electrophoresis. Error bars are the 95% confidence intervals based on the adjusted Wald method (see text).

shallow and remains so for q-values between approximately 0.05 and 0.3. These allele frequencies yield *redimita* phenotype frequencies of between 0.1 and 0.5, the range observed in British and western and central European populations (the more sporadic and rarer *ovata* morph is ignored here). Outside this 0.05 to 0.3 allele-frequency region, selection must increase markedly to protect the polymorphism, even in small populations. The exact shape of the plot is unimportant as long as it has these necessary properties. Thus at normally observed morph frequencies, that is, within the area of the Δq plot with the shallow gradient, selection is weak and, given frequently small population sizes, may effectively be absent. The color morphs, under these circumstances, act as quasi-neutral characters as drift overrides selection. Once stochastic processes move allele frequencies outside this area, powerful selection operates to return them. The polymorphism is therefore, at the same time, protected (explaining the evidence for selection discussed above) and also quasi-neutral (explaining the lack of evidence for selection in local populations). The perturbation experiments conducted to date (Oxford and Shaw 1986; Reillo and Wise 1988b), merely altered frequencies within the quasi-neutral zone and, as a consequence, failed to identify a return to former frequencies. The major prediction of the model is that morph frequencies perturbed outside the quasi-neutral region should quickly be returned to it—experiments of this sort are in progress. If this heuristic model is correct it suggests that the quasi-neutral zone may vary geographically, for example in North American compared with Finnish populations.

Observations of natural populations with extreme morph frequencies should also provide evidence for this model. Selection should drive allele frequencies into the quasi-neutral zone, but once there, as selection is relaxed, stochastic processes should lead to greater fluctuations in allele frequencies as a result of continuous drift. One Nidderdale population provided an opportunity to examine this prediction. In 1975, when first examined, the sample from site N was monomorphic for *lineata* (n = 50). In 1976 one *redimita* was observed (n = 93) and subsequently the frequency of the *redimita* morph increased to between 0.15 and 0.3 (Fig. 7). The regression of arcsine-transformed *redimita* frequency, weighted by sample size, against time is highly significant (b = 0.014, df = 14, P = 0.001). Inspection of Figure 7 suggests that there may indeed be greater variation in *redimita* frequency in the later years but interpretation is complicated by the fact that sample sizes are also generally smaller during this period. If selection is responsible for the increase in the frequency of *redimita*, its mean strength can be estimated using the quotient of the frequencies of the dominant to the recessive phenotype (Cain et al. 1990; Cook et al. 1999). The slope of the plot of log(frequency *redimita*/frequency *lineata*) against time estimates log(w), where w is the fitness of the *redimita* morph. The regression is significant (b = 0.0387, SE = 0.0104, df = 13, P = 0.0026) and yields a fitness estimate for *redimita* of 1.093 (data from 1975 are ignored because the frequency of *redimita* was zero). Rescaling this to one, the average relative fitness of the *lineata* morph during this period was 0.915, representing an 8–9% disadvantage. Migration cannot be responsible for this increase in *redimita* frequency because adjacent populations are generally smaller, have lower frequencies of *redimita* (over all years, mean frequency in site AB = 0.086, and in site AO = 0.059) and are very different at the regulatory and spotting loci.

The conclusions from these studies are that natural selection of some sort is responsible for maintaining the color polymorphism in *E. ovata* (and, by extension, in *E. latimana*) and for determining the broad bounds of the morph-frequency distribution in any one geographical area. However, within these distributions, selection is probably very weak and stochastic processes primarily influence local morph frequencies. Thus, in the majority of populations at any one time the polymorphism is effectively neutral. The model proposed might apply more widely and explain why evidence for selection in natural populations may be elusive. In some respects, this proposal is similar to the two-stage model suggested by Cook (1998) to explain the shell color and banding polymorphisms in the snail *Cepaea*. The *Enoplognatha* model differs principally in that, for the majority of populations,

drift is the normal engine of morph-frequency change and selection is only intermittently brought to bear when morph frequencies are perturbed beyond certain limits. In neither of these examples is the nature of the selective regime protecting the polymorphism fully understood.

Acknowledgments

I am very grateful for the assistance of B. R. Johnson, R. H. Oxford, and R. J. Pitts in scoring the Nidderdale populations in the early years, and for the enormous efforts expended by over 100 volunteers during the British surveys. A. Bennett and S. Smout contributed the allozyme data, and P. J. Mayhew kindly ran the logistic regression analysis in GLIM. B. Gunnarsson (University of Göteborg) collaborated with the spider survey in Sweden during 1997, for which we acknowledge the financial support of the Swedish Institute and the Swedish Science Research Council. I also thank L. M. Cook, T. J. Crawford, J. D. Currey, M. H. Williamson, and two anonymous referees for their helpful comments on drafts of the paper.

Literature Cited

Agresti, A., and B. A. Coull. 1998. Approximation is better than ''exact'' for interval estimation of binomial proportions. Am. Stat. 52:119–126.

Box, G. E. P. 1953. Non-normality and tests on variances. Biometrika 40:318–335.

Cain, A. J., and W. B. Provine. 1991. Genes and ecology in history. Pp. 3–28 *in* Berry R. J., T. J. Crawford, and G. M. Hewitt, eds. Genes in ecology. Blackwell, Oxford, U.K.

Cain, A. J., L. M. Cook, and J. D. Currey. 1990. Population size and morph frequencies in a long-term study of *Cepaea nemoralis*. Proc. R. Soc. Lond. B 240:231–250.

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

———. 2003. The rise and fall of the carbonaria form of the peppered moth. Q. Rev. Biol. 78:399–417.

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

Endler, J. A. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

Ford, E. B. 1964. Ecological genetics. Methuen, London.

Gavrilets, S. 2004. Fitness, landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.

Gerhardt, U. 1921. Vergleichende Studien über die Morphologie de Mänlichen Tasters und die Biologie de Kopulation de Spinnen. Arch. Naturg. Abt. A 87:78–247.

Gillespie, R. G. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science 303:356–359.

Golding, B. 1992. The prospects for polymorphisms shared between species. Heredity 68:263–276.

Goodhart, C. B. 1987. Why are some snails visibly polymorphic, and others not? Biol. J. Linn. Soc. 31:35–58.

Halkka, O., and L. Halkka. 1990. Population genetics of the polymorphic meadow spittlebug, *Philaenus spumarius* (L.). Evol. Biol. 24:149–191.

Hartl, D. L., and E. W. Jones. 1998. Genetics: principles and analysis. 4th ed. Jones and Bartlett, Sudbury, U.K.

Harvey, P. R., D. R. Nellist, and M. G. Telfer. 2002. Provisional atlas of British spiders (Arachnida, Araneae). Vol. 1. Biological Records Centre, Huntingdon, U.K.

Hebert, P. D. N., and M. J. Beaton. 1993. Methodologies for allozyme analysis using cellulose acetate electrophoresis. Helena Laboratories, Beaumont, TX.

Hippa, H., and I. Oksala. 1979. Color polymorphism of *Enoplognatha ovata* (Clerck) (Araneae, Theridiidae) in western Europe. Hereditas 90:203–212.

———. 1982. Definition and revision of the *Enoplognatha ovata* (Clerck) group (Araneae, Theridiidae). Entomol. Scand. 13: 213–222.

Hoekstra, H. E., and T. Price. 2004. Parallel evolution is in the genes. Science 303:1779–1780.

Hoekstra, H. E., K. E. Drumm, and M. W. Nachman. 2004. Ecological genetics of adaptive color polymorphism in pocket mice: geographical variation in selected and neutral genes. Evolution 58:1329–1341.

Janzen, D. H., and P. S. Martin. 1982. Neotropical anachronisms: the fruits the gomphotheres ate. Science 215:19–27.

Jones, J. S., B. H. Leith, and P. Rawlings. 1977. Polymorphism in *Cepaea*: a problem with too many solutions? Annu. Rev. Ecol. Syst. 8:109–143.

Layard, M. J. W. 1973. Robust large-sample tests for homogeneity of variances. J. Am. Stat. Assoc. 68:195–198.

Magurran, A. E. 2004. Measuring biological diversity. Blackwell, Oxford, U.K.

Miller, R. G. 1974. The jackknife: a review. Biometrika 61:1–15.

Moran, M. D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100:403–405.

Mundy, N. I., N. S. Badcock, T. Hart, K. Scribner, K. Janssen, and N. J. Nadeau. 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. Science 303:1870–1873.

Nachman, M. W., H. E. Hoekstra, and S. L. D'Agostino. 2003. The genetic basis of adaptive melanism in pocket mice. Proc. Natl. Acad. Sci. USA 100:5268–5273.

Nielsen, E. 1932. The biology of spiders. Vol. II. Levin and Munksgaard, Copenhagen.

O'Hara, R. B. 2005. Comparing the effects of genetic drift and fluctuating selection on genotype frequency changes in the scarlet tiger moth. Proc. R. Soc. Lond. B 272:211–217.

Oxford, G. S. 1976. The color polymorphism in *Enoplognatha ovatum* (Clerck) (Araneae: Theridiidae)—temporal stability and spatial variability. Heredity 36:369–381.

———. 1983. Genetics of color and its regulation during development in the spider *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae). Heredity 51:621–634.

———. 1985a. Geographical distribution of phenotypes regulating pigmentation in the spider *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae). Heredity 55:37–45.

———. 1985b. A countrywide survey of color morph frequencies in the spider *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae): evidence for natural selection. Biol. J. Linn. Soc. 24: 103–142.

———. 1989. Genetics and distribution of black spotting in *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae), and the role of intermittent drift in population differentiation. Biol. J. Linn. Soc. 36:111–128.

———. 1991. Visible morph-frequency variation in allopatric and sympatric populations of two species of *Enoplognatha* (Araneae: Theridiidae). Heredity 67:317–324.

———. 1992. *Enoplognatha ovata* and *E. latimana*: a comparison of their phenologies and genetics in Norfolk populations. Bull. Br. Arachnol. Soc. 9:13–18.

———. 1993a. Genetic drift in a visibly polymorphic spider, *Enoplognatha ovata*: antidote to the peppered moth. J. Biol. Ed. 27:115–120.

———. 1993b. Components of variation in population size in the spider *Enoplognatha ovata sensu stricto* (Clerck) (Araneae: Theridiidae). Bull. Br. Arachnol. Soc. 9:193–202.

———. 1998. Guanine as a colorant in spiders: development, genetics, phylogenetics and ecology. Pp.121–131 *in* P. A. Selden, ed. Proceedings of the 17th European Colloquium of Arachnology, Edinburgh 1997. British Arachnological Society, Burnham Beeches, Bucks, U.K.

Oxford, G. S., and P. R. Reillo. 1993. Trans-continental, visible morph-frequency variation at homologous loci in two species of spider, *Enoplognatha ovata* and *E. latimana*. Biol. J. Linn. Soc. 50:235–253.

———. 1994. The world distributions of species within the *Eno-*

plognatha ovata group (Araneae: Theridiidae): implications for their evolution and for previous research. Bull. Br. Arachnol. Soc. 9:226–232.
Oxford, G. S., and M. W. Shaw. 1986. Long-term variation in colormorph frequencies in the spider *Enoplognatha ovata* (Clerck) (Araneae: Theridiidae): natural selection, migration and intermittent drift. Biol. J. Linn. Soc. 27:225–249.
Perneger, T. V. 1998. What's wrong with Bonferroni adjustments? Brit. Med. J. 316:1236–1238.
Reillo, P. R. 1989. Color polymorphism in the spider *Enoplognatha ovata* (Araneae: Theridiidae): broad-scale morph-frequency variation in northeastern North America. Am. Midl. Nat. 122: 199–203.
Reillo, P. R., and D. H. Wise. 1988a. Genetics of color expression in the spider *Enoplognatha ovata* (Araneae: Theridiidae) from coastal Maine. Am. Midl. Nat. 119:318–326.
———. 1988b. An experimental evaluation of selection on color morphs of the polymorphic spider *Enoplognatha ovata* (Araneae: Theridiidae). Evolution 42:1172–1189.
———. 1988c. Temporal and spatial patterns of morph-frequency variation among coastal Maine populations of the polymorphic spider *Enoplognatha ovata* (Araneae: Theridiidae). Am. Midl. Nat. 120:337–354.
Richardson, B. J., P. R. Baverstock, and M. Adams. 1986. Allozyme electrophoresis: a handbook for animal systematics and population studies. Academic Press, New York.
Richman, A. 2000. Evolution of balanced genetic polymorphism. Mol. Ecol. 9:1953–1963.
Samuels, M. L., and J. A. Witmer. 2003. Statistics for the life sciences. 3rd. ed. Pearson Education, Upper Saddle River, NJ.
Seligy, V. L. 1971. Postembryonic development of the spider *Enoplognatha ovata* (Clerck)(Araneae, Theridiidae). Zool. J. Linn. Soc. 50:21–31.
Snazell, R. 1983. On two spiders recently recorded from Britain. Bull. Br. Arachnol. Soc. 6:93–98.
Sokal, R. R., and F. J. Rohlf. 1995. Biometry. 3rd ed. W. H. Freeman and Co., New York.
Wise, D. H., and P. R. Reillo. 1985. Frequencies of color morphs in four populations of *Enoplognatha ovata* (Araneae: Theridiidae) in eastern North America. Psyche 92:135–144.

Corresponding Editor: K. Hughes