Evolution, 40(6), 1986, pp. 1141–1159

GENETIC ANALYSIS OF A HYBRID ZONE BETWEEN THE FIRE-BELLIED TOADS, *BOMBINA BOMBINA* AND *B. VARIEGATA*, NEAR CRACOW IN SOUTHERN POLAND

JACEK M. SZYMURA
Department of Comparative Anatomy, Jagiellonian University, ul. Karasia 6, Krakow 30-060, Poland

AND

NICHOLAS H. BARTON
Department of Genetics and Biometry, University College, 4 Stephenson Way, London NW1 2HE, U.K.

Abstract.—The fire-bellied toads *Bombina bombina* and *B. variegata* differ extensively in biochemistry, morphology, and behavior. We use a survey of five diagnostic enzyme loci across the hybrid zone near Cracow in Southern Poland to estimate the dispersal rate, selection pressures, and numbers of loci which maintain this zone. The enzyme clines coincide closely with each other and with morphological and mitochondrial DNA clines. Although the zone lies on a broad transition between environments suitable for *bombina* and *variegata*, the close concordance of diverse characters, together with increased aberrations and mortality in hybrids, suggest that the zone is maintained largely by selection against hybrids. There are strong "linkage disequilibria" between each pair of (unlinked) enzyme loci ($\bar{R} = 0.129$ [2-unit support limits: 0.119–0.139]). These are probably caused by gene flow into the zone, and they give an estimate of dispersal ($\sigma = 890$ [790–940] m gen$^{-1/2}$). The clines are sharply stepped, with most of the change occurring within 6.15 (5.45–6.45) km, but with long tails of introgression on either side. This implies that the effective selection pressure on each enzyme marker (due largely to disequilibrium with other loci) is $s^* = 0.17$ (0.159–0.181) at the center but that the selection acting directly on the enzyme loci is weak or zero ($s_e < 0.0038$). The stepped pattern implies a barrier to gene flow of 220 (48–415) km. This would substantially delay neutral introgression but would have little effect on advantageous alleles; the two taxa need not evolve independently. Strong selection is needed to maintain such a barrier: hybrid populations must have their mean fitness reduced by a factor of 0.65 (0.60–0.77). This selection must be spread over a large number of loci to account for the concordant patterns and the observed cline widths ($N = 300$ [80–2,000]).

Received August 12, 1985. Accepted June 9, 1986

Hybrid zones are found in a wide variety of organisms and may involve a wide range of characters (Barton and Hewitt, 1985). Several mechanisms have been proposed to account for them: simple mixing of neutral alleles following secondary contact (Hafner et al., 1983), selection favoring hybrids within narrow ecotones (Moore, 1977), selection against hybrids (Zaslavskii, 1967; Bazykin, 1969, 1973; Key, 1968), or selection favoring different alleles in different environments (Fisher, 1930; Haldane, 1948). Diverse views are also held on their importance in evolution. It has been suggested that hybrid zones may eventually be strengthened into barriers which separate fully isolated biological species, through the reinforcement of premating isolation (Fisher, 1930; Dobzhansky, 1940), the accumulation of modifiers (Fisher, 1930; Clarke, 1966; Endler, 1977), the clumping of zones at different genetic loci in the same spatial position (Key, 1981), or through their effect as barriers to gene flow at loci not themselves involved in maintaining the zone (White, 1968, 1978).

In this paper, we will show that the pattern of genotype frequencies across a hybrid zone in *Bombina* can give robust and quantitative information about the strength and nature of the selective forces maintaining the zone and about its effect as a barrier to gene flow. In particular, we show that although strong selection acts to maintain genetic differentiation at several hundred loci, gene flow at other loci is not likely to be impeded sufficiently to prevent advantageous alleles crossing between the two taxa.

The hybrid zone between the European fire-bellied toads *Bombina bombina* and

TABLE 1. Differences between *Bombina bombina* and *B. variegata*.

	B. bombina	*B. variegata*	References
Distribution	Lowlands of Eastern and Central Europe	Mountainous and hilly regions of Western and Southern Europe, and Carpathians	Andren et al. (1984), Arntzen (1978)
Habitat	Largely aquatic	Largely terrestrial	Madej (1973)
Breeding sites	Large, permanent waters	Temporary pools, small ponds	
Skin thickness (epidermis/dermis)	134.5 μ (22.8 μ/111.7 μ)	296.6 μ (65.2 μ/231.4 μ)	Czopkowa and Czopek (1955)
Mean fecundity (largest clutches observed)	363 eggs (509, 547, 689)	116 eggs (204, 233, 294, 297)	A. Rafińska (pers. comm.)
Mean number of eggs per clump (Range)	32 (9–76)	17 (4–58)	A. Rafińska (pers. comm.)
Egg size	1.4 mm	1.9 mm	A. Rafińska (pers. comm.)
Development time at 20°C (egg to toadlet)	73–75 days	61–63 days	A. Rafińska (pers. comm.)
Breeding behavior	Prolonged breeder, territorial	Explosive breeder, nonterritorial	Szymura (unpubl.) (see also Lörcher [1969])
Vocal sacs	Present	Absent	Boulenger (1886)
Rate of mating calls	22 min^{-1}	95 min^{-1}	Lörcher (1969)
Duration of mating calls	210 msec	160 msec	Lörcher (1969)
Fundamental frequency of mating calls	530 Hz	580 Hz	Lörcher (1969)
DNA content per nucleus	18.8 pg	21.1 pg	Olmo et al. (1982)
Chromosome number (similar karyotypes)	24	24	Morescalchi (1965), Wickbom (1949)
Nei's genetic distance	0.49 ± 0.13; 39 loci		Szymura (1983)
Immunological albumin distance	2–4 IDU		Maxson and Syzmura (1984)
Mitochondrial DNA divergence	$F = 0.21$; $P = 0.094 \pm 0.011$ substitutions per nucleotide (16 restriction enzymes)		Szymura et al. (1985)

Bombina variegata is almost ideal for our analysis. These two taxa are quite distinct in morphology and behavior (Table 1). Many of their differences can be regarded as adaptations to the different environments in which *bombina* and *variegata* live (lowland and mountain, respectively). There are also extensive biochemical differences, including immunological (Maxson and Szymura, 1984), electrophoretic (Gollmann, 1984; Szymura, 1976, 1983) (Nei's $D = 0.49 \pm 0.13$), DNA content (Olmo et al., 1982), and mitochondrial DNA (Szymura et al., 1985) divergence. Five electrophoretic loci were chosen for this study for the following reasons: they are fixed for alternative alleles on either side of the zone; they can be scored from amputated toes, so that large samples can be taken without damaging the population; and Mendelian inheritance and independent assortment have been demonstrated in laboratory crosses (Szymura and Farana, 1978).

The ranges of *B. bombina* and *B. variegata* meet in Central Europe along a wide front, extending from Austria along the southern edge of the Danube Valley to the Black Sea and completely surrounding the Carpathian Mountains along their foothills (Fig. 1). The contact in Germany is not well

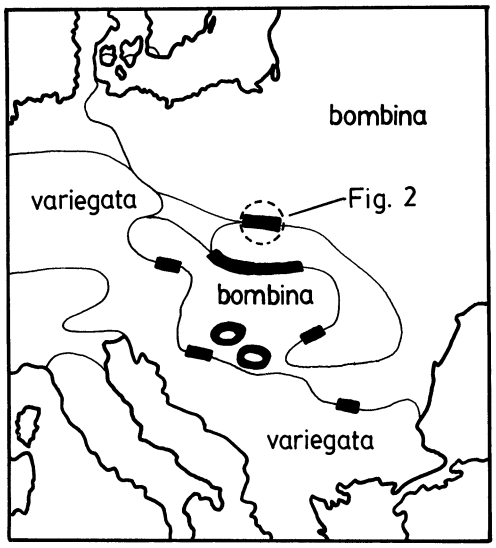

FIG. 1. The distribution of *Bombina bombina* and *B. variegata* across Central and Eastern Europe. Thick lines indicate regions of known hybridization; however, hybridization is likely wherever the two taxa meet.

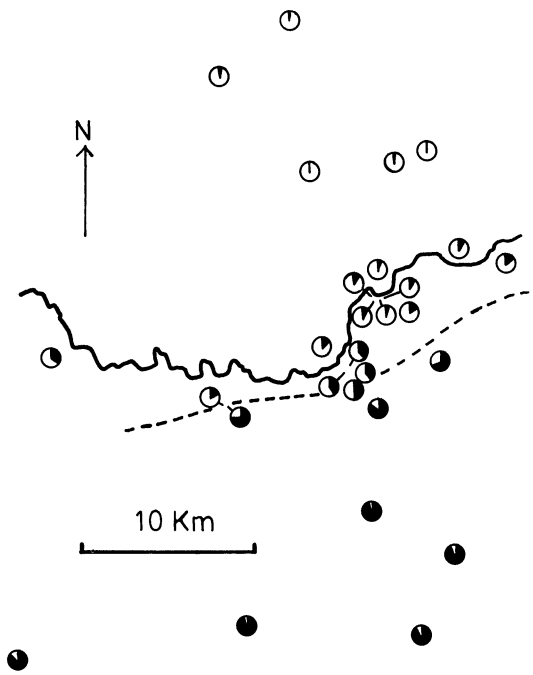

FIG. 2. The average frequency of *variegata* alleles across the transect near Cracow. The morphological hybrid zone (Michałowski, 1958) is indicated by the dotted line, and the Vistula by the thicker line.

understood. In Poland, *B. bombina* and *B. variegata* meet along the foothills of the Carpathians (Horbulewicz, 1927, 1933; Michałowski, 1958; Madej, 1964*a*, 1964*b*, 1973); here, we describe a detailed survey across a transect just west of Cracow, involving a total of 1,988 individuals from 34 samples, taken from 29 sites (Fig. 2, Table 2). We aim not only to make inferences about population structure and selective mechanisms in *Bombina* from the pattern of genotypes found across the hybrid zone, but also to illustrate some general methods which may be useful in analyzing similar data from other hybrid zones.

MATERIALS AND METHODS

Electrophoresis

Toads were sampled during the breeding season (May–July) from 1974 to 1981 across the hybrid zone west of Cracow in southern Poland. Each collection site was usually a single pond, or in the case of *B. variegata*-like hybrid populations, a series of pools, puddles or ditches which lay within a radius not exceeding 100 meters. Direct and indirect measures of movement in *Bombina*, described below, give us confidence that we sampled panmictic areas.

Animals were brought into the laboratory, where electrophoresis was performed according to well established procedures (Szymura, 1976; Szymura and Farana, 1978). Five enzyme loci (lactate dehydrogenase [*Ldh-1*], malate dehydrogenase [*Mdh-1*], creatine kinase [*Ck*], adenylate kinase [*Ak*], and phosphoglucose isomerase [*Gpi*]) were scored for each individual. These loci are diagnostic for *B. bombina* and *B. variegata* in Poland (Szymura, 1976, 1983), and the formal genetics of their electrophoretic variation have been determined (Szymura and Farana, 1978). There is no evidence of linkage among any of the five loci. Since 1976, after finding that genotypes can be determined from amputated toes, all toads have been released at their place of capture. This reduces the disturbance caused by sampling large numbers of animals.

Statistics

The electrophoretic survey gives us the number of each genotype found in each site. We will condense these observations by es-

timating the parameters of various genetic models. These estimates will be made using the method of likelihood (Hacking, 1965; Edwards, 1972); that is, by judging the plausibility of a model by the likelihood with which it would give the observed data. When two discrete hypotheses are compared, the significance of a difference in $\log_e$ likelihood (ΔL) depends on the difference in the numbers of degrees of freedom associated with the two hypotheses (ν): a hypothesis in which many parameters have been estimated from the data necessarily gives a greater likelihood than one which has not been fitted to the data. Since, in large samples, the difference in $\log_e$ likelihood (ΔL_ν) is distributed approximately as $(\frac{1}{2})\chi^2_{[\nu]}$, we will use the conventional 95% limits of the χ^2 distribution as our criterion of significance. (In large samples, these two-unit support limits correspond to 95% confidence limits [Edwards, 1972].) Estimates will be presented in two ways: either by plotting a graph of likelihood against the parameter values, or, more usually, by giving the most likely value, together with the range of values which is within two units of $\log_e$ likelihood of this maximum.

The two main statistical problems will be to estimate linkage disequilibria and cline shape; selection pressures, dispersal rates, etc. will be inferred from these estimates. The programs used to estimate linkage disequilibrium and cline shape are available from N. H. Barton.

In estimating linkage disequilibrium, we will assume that diploid genotypes are in Hardy-Weinberg proportions, thus reducing the problem to the estimation of haploid gamete frequencies (see below). We further simplify the problem by assuming that third- and higher-order disequilibria are zero; this allows us to make separate estimates of pairwise disequilibrium for each pair of loci. The adequacy of this assumption can be checked by examining the distribution of the "hybrid index," defined as the number of *variegata* alleles in each toad. In the samples from Kopanka and Skawina, at least, this distribution is not significantly different from that expected if higher disequilibria are neglected. (We intend to return to the complex problem of estimating higher-order disequilibria in a later paper.) Having made these approximations, we fix our estimates of the allele frequencies in the population to equal the frequencies observed in the sample. We then calculate the expected genotype frequencies and their corresponding likelihoods for a range of pairwise disequilibria; this is the method used by Hill (1974).

The second major problem is to estimate cline shape from the average allele frequency at each site. In principle, this is straightforward: we calculate the likelihood of a range of parameters describing cline shape (Fig. 5), and choose the most likely combination as our estimate. However, two difficulties arise. First, we show below that allele frequencies fluctuate randomly from site to site, most likely as a result of drift. So, in calculating the likelihood of a given cline, we should allow for two sources of deviation: simple sampling error and random variations in the true frequency at each site. We can estimate the standardized variance of the underlying allele frequencies ($F_{ST} = [\mathrm{var}(p)]/pq$) by comparing the shapes of clines at different loci (see below). The variance in average frequency of *variegata* alleles around the expected cline is $V = (1/N) + (F_{ST}/5)$, where N is the number of alleles sampled, summed over the five loci. This allows us to define an "effective sample size," $N_e = 1/V$. Without this procedure, we would give undue weight to a few very large samples and would lose much information from the smaller samples. The $\log_e$ likelihood of a model which predicts an allele frequency of P is the sum over samples of $N_e\{\bar{p}[\log(P)] + \bar{q}[\log(Q)]\}$, where $\bar{p}$ and $\bar{q}$ are the average frequencies of *variegata* and *bombina* alleles.

The second difficulty in estimating cline shape is that we wish to estimate six parameters describing the cline (its width, position, and the size and rate of decay of the tails of introgression on either side; Figs. 5 and 6). Finding the maximum of a function of six parameters is not trivial, and finding the range of values within two units of the maximum is still harder. In our problem, the likelihood surface cannot be approximated by a bell-shaped surface (Fig. 6a, b). Hence, Fisher's (1925) method of information, which is based on a Taylor expansion around the maximum, cannot be used.

We have therefore chosen a new method, based on the Metropolis algorithm (Metropolis et al., 1953; Binder, 1979). This generates a random walk through parameter space, with probability density equal to the likelihood. One begins with an arbitrary combination of parameters. Then, one of the parameters is changed by a random amount. The distribution of these changes is irrelevant, provided that their expectation is zero and provided that the distribution does not depend on the parameter values. For efficiency, it should be chosen so that roughly equal numbers of changes are accepted and rejected. If this change increases the likelihood, it is accepted; if it decreases the likelihood (from ℓ to ℓ', say), then it is only accepted with probability ℓ'/ℓ. This process is repeated, running through each parameter in turn. It generates a random walk, which tends to increase the likelihood if the parameters are far from the maximum, but which eventually settles into an equilibrium distribution proportional to the likelihood. The results can be presented either by plotting a scatter diagram of two parameters (a two-dimensional cross-section through the six-dimensional parameter space; Fig. 6), or by plotting a histogram of each parameter and reading off the most likely estimate and its support limits. The accuracy of the estimates was checked by replication; the estimates given here are based on a series of runs, each beginning with a warm-up of 3,000 iterations, followed by 12,000 iterations during which the likelihood is accumulated. Independent runs gave parameter estimates which varied much less than the range spanned by the support limits.

Results

Position and Stability of the Enzyme Clines

As well as coinciding with each other (see below), the enzyme clines in *Bombina* coincide with the morphological hybrid zone mapped out by Michałowski (1958; Fig. 2). The same is true of the contacts near Zagreb in Yugoslavia (Karaman, 1922; Szymura, unpubl.), and in Austria (Gollmann, 1984). The enzyme clines at Cracow also coincide with the cline in mitochondrial DNA (Szymura et al., 1985). This coincidence, together with direct observations of morphology over the past 30–80 years suggests that the zone is not changing rapidly (op. cit.; Mehely, 1905 [Mecsek Mts., Hungary]; Horbulewicz, 1927, 1933 [Western Ukraine]). However, it is harder to judge the stability of the zone over longer periods. Biochemical (Szymura, 1983) and paleontological (Sanchiz and Młynarski, 1979) evidence suggests that *B. bombina* and *B. variegata* diverged in the Pliocene, 3–4 Myr B.P. Their ranges have subsequently fluctuated, withdrawing to southern refuges during the Pleistocene glaciations and expanding to the north in warmer periods.

At present, the distribution of *bombina* and *variegata* follows the distribution of lowland and mountain habitats (Fig. 1), and the hybrid zone near Cracow lies in a transition region where the Carpathian foothills rise gently from the flat Vistula valley. In general, precipitation increases to the south, though there is considerable microclimatic variation due to varied relief. On the uplands north of the Vistula river, the predominant soil is loess on limestone. The flood plain of the Vistula is covered with sand, alluvial soil, or loess, while the Carpathian foothills have clayey mountain soils. North of the Vistula flood plain, there is little precipitation, and the limestone is well drained; as a result, ephermeral ponds do not last long enough for *B. variegata* to develop through metamorphosis. Temporary pools are more common on the Carpathian foothills and in the valleys draining them. Permanent artificial ponds are present throughout the transect but seem to be more important breeding sites towards the north of the transect. Often, permanent ponds south of the hybrid zone contain no *Bombina* at all.

Despite the broad correlation between the distribution of the two taxa and the environments to which they are best adapted, the center of the hybrid zone does not seem to follow closely any obvious environmental contour, and it is certainly not located at any narrow ecotone. The faunistic, floristic, and edaphic transition between the lowlands and the Carpathians occurs over a region much broader than the hybrid zone and involves a complex pattern of interdig-

itation and blending. In contrast, most of the transition between *bombina* and *variegata* occurs over a region 6 km wide, located within the flood plain of the Vistula (Fig. 2).

The widespread hybridization of *B. bombina* and *B. variegata* along their contact, situated between lowland and mountain, has probably been a regular feature of their history, rather than the result of recent human intervention. There have been permanent human settlements near Cracow (including settlements at Tyniec and Skawina) as early as 6,000 B.P., and human activity has increased steadily since then (Żaki, 1974). The time of greatest change would have been more than 10 centuries ago, with the spread of agriculture and deforestation. *B. bombina* were found in Sweden in the recent past and are found on the Danish Isles at present (Andren et al., 1984); these areas were accessible to the toads via a land connection 7,000 years ago. This suggests that *bombina* colonized the lowlands of Poland, and *variegata* the Carpathians, somewhat before this time, soon after the last glaciation, 8,000–9,000 B.P. Taking the generation time as 3 years, this implies that the zone near Cracow is 2,300–3,000 generations old.

Deviations from Hardy-Weinberg Proportions

For each sample and each locus, we can compare the likelihood that the corresponding population is in Hardy-Weinberg proportions, with the likelihood that there is some deficit (F_{IS}) of heterozygotes (see Table 4 for a summary of our notation and estimates). Only two populations out of 34 showed significant deviations. One of these (Jeziorzany) indicated a deficit of heterozygotes at *Mdh-1* ($\Delta L_1 = 4.20$). One would expect such deviations to occur by chance when many tests are made. However, the other population (Świnna Poręba) is clearly not in Hardy-Weinberg proportions. Combining all five loci, $\Delta L_5 = 30.05$. There are no significant differences in F_{IS} between loci; overall, $F_{IS} = 0.65$ (0.49–0.81). This large deviation is caused by three *bombina* genotypes that are embedded in a predominantly *variegata* sample (Table 2).

For the remaining 33 samples, the total difference in $\log_e$ likelihood associated with deviations from Hardy-Weinberg proportions is $\Delta L_{33\times5} = 74.79$; this is not statistically significant. However, we can make a more sensitive test for deviations by partitioning this statistic into a component due to a consistent heterozygote deficit, equal across all loci and all samples, and a component due to variation around this average. There is an indication of a slight average deficit ($F_{IS} = 0.017$ (0–0.034); $\Delta L_1 = 1.94$), but no evidence of any variation between loci or between samples ($\Delta L_{164} = 72.82$).

For simplicity, we have so far neglected the linkage disequilibria between loci. These will increase the chance that several loci in the same sample will show apparent deviations and so will lead to overestimates of the extent of such deviations. However, we show below that disequilibria are not extremely large and so are not likely to cause much error. Moreover, we have tested each pair of loci for deviations from the two-locus genotype frequencies expected under random mating: none of these tests were significant.

Coincidence of Clines at Different Loci

Since there is little, if any, deviation from either single-locus or pairwise Hardy-Weinberg proportions, we need only consider the frequencies of the $2^5 = 32$ haploid genotypes. We will consider first the five allele frequencies. (Throughout, p denotes the frequency of the *variegata* allele, q the frequency of the *bombina* allele).

There is a clear transition at all loci from *bombina* alleles in the north to *variegata* alleles in the south (Table 2, Fig. 2), a result expected from previous morphological studies in this area (Michałowski, 1958). Perhaps the most striking feature of this transition is the concordance of patterns at different loci: to a first approximation at least, all loci seem to change at the same point, and over the same distance. Despite this broad concordance, the large sample sizes allows us to detect slight variations between loci. In most samples, allele frequencies differ significantly: overall, $\Delta L_{34\times4} = 272.72$, which is highly significant. This heterogeneity can be illustrated by a graph plotting the allele frequencies at each of the five

TABLE 2. Allele frequencies and distribution of the hybrid index. X is the distance from the center of the hybrid zone (Fig. 2); positive values indicate the *variegata* side of the zone (to the south), and dashes correspond to sites far from the main transect. N is the number of alleles sampled, $\bar{P}$ is the average frequency of *variegata* alleles, and $\bar{R}$ is the standardized linkage disequilibrium, averaged over all pairs of loci. B, H, F_1, and V are the frequencies of apparently pure *bombina*, hybrids, apparent F_1's and apparently pure *variegata* genotypes, respectively. The F_1's are counted as hybrids, so that $B + H + V = 1$. Tabled values for the hybrid index are the numbers of individuals with 0, 1, . . . , 10 *variegata* alleles. The number of individuals heterozygous at all loci is included within class "5," but is also given in brackets. The apparent F_1 individual from Świnna Poręba was homozygous for a sixth marker enzyme and so must be the product of more than one generation of hybridization.

Site	Date	X (km)	N	$\bar{P}$	$\bar{R}$	B	H	(F_1)	V	Hybrid index 0	1	2	3	4	5 (F_1)	6	7	8	9	10
Bedkowska Valley	1975	−19.3	420	0.019	—	0.833	0.167	0	0	35	7									
Pisary (near Rudawa)	1976	−18.0	460	0.022	—	0.783	0.217	0	0	36	10									
Rząska	1974	−9.2	370	0.038	—	0.703	0.297	0	0	26	9	1	1							
Mydlniki	1979	−9.6	650	0.034	0.000	0.667	0.333	0	0	40	18	2								
Aleksandrowice	1975	−11.5	200	0.025	—	0.800	0.200	0	0	16	2	2								
Kraków-Przegorzaly	1975	—	350	0.044	0.013	0.610	0.390	0	0	25	14	2								
Kraków 114	1976	—	350	0.143	−0.031	0.229	0.771	0	0	8	11	10	5	1						
Kraków Bielany	1981	−4.7	230	0.126	0.254	0.696	0.304	0	0	16	6	1								
Tyniec 2	1975	−3.6	350	0.126	0.254	0.343	0.657	0	0	12	13	5	2	2		1				
Tyniec 2	1977	−3.6	980	0.101	0.050	0.378	0.622	0	0	37	31	21	6	3						
Tyniec 4	1977	−3.4	538	0.093	0.049	0.471	0.529	0	0	16	11	6	1							
Tyniec 3	1977	−3.4	2,314	0.103	0.039	0.376	0.624	0	0	88	81	42	16	5	2 (0)					
Tyniec 5	1977	−3.35	600	0.100	0.104	0.400	0.600	0	0	24	22	8	2	4						
Tyniec 1	1974	−1.8	240	0.204	0.409	0.409	0.591	0	0	9	5	2	4	2				1	1	
Jeziorzany	1975	−2.4	250	0.140	0.072	0.280	0.720	0	0	7	6	7	5							
Sidzina	1974	1.4	330	0.745	0.155	0	0.970	0	0.030			1	1	1		7	4	5	13	1
Kopanka 1	1975	−0.8	384	0.323	0.166	0.053	0.947	0.053	0	2	7	3	8	11	4 (2)	1	1	1		
Kopanka 1	1976	−0.8	670	0.382	0.166	0.039	0.961	0.013	0	3	9	11	5	11	11 (1)	9	8			
Kopanka 2	1979	−0.7	660	0.458	0.067	0.015	0.985	0.045	0	1	2	4	11	15	9 (3)	17	4	3		
Kopanka 2	1981	−0.7	3,564	0.485	0.152	0.003	0.997	0.031	0	1	17	28	45	57	69 (11)	57	47	19	12	
Samborek (near Skawina)	1978	−0.25	274	0.391	−0.046	0	1.000	0	0		1	4	5	5	7 (0)	3	1			
Podbory Skawińskie	1979	0.0	600	0.523	0.160	0.017	0.967	0.033	0.017	1		4	7	7	12 (2)	14	9	5		1
Podbory Skawińskie	1981	0.0	1,300	0.521	0.111	0	1.000	0.008	0		6	4	15	17	30 (1)	25	17	13	3	
Podbory Skawińskie East	1981	0.0	130	0.469	0.128	0	1.000	0	0		1	2	2	1	2 (0)	2	2	1		
Skawina	1975	1.85	120	0.833	0.188	0	0.667	0	0.333						2 (0)		1	2	3	4
Wielkie Drogi 1	1976	—	480	0.744	0.006	0	0.958	0.021	0.042						6 (1)	8	8	14	10	2
Wielkie Drogi 1	1977	—	492	0.761	0.034	0	0.940	0.040	0.060						3 (2)	9	11	10	13	3
Wielkie Drogi 2	1978	—	450	0.153	0.379	0.400	0.600	0.026	0	18	13	5	2	1	3 (1)		3			
Półwieś	1975	—	190	0.279	0.466	0.263	0.737	0	0	5	5	1	1	2	1 (0)		1	1		
Radziszów	1976	7.0	380	0.939	—	0	0.474	0	0.526								1	3	14	20
Głogoczów	1976	11.0	510	0.875	0.260	0.020	0.647	0	0.333	1						2	2	10	19	17
Lanckorona	1976	12.2	310	0.971	—	0	0.320	0	0.680									1	7	17
Rudnik	1976	14.7	460	0.948	0.013	0	0.457	0	0.543									4	17	25
Świnna Poręba	1974	—	400	0.888	0.822	0.075	0.200	0.025	0.725	3					1 (1)		1	1	5	29

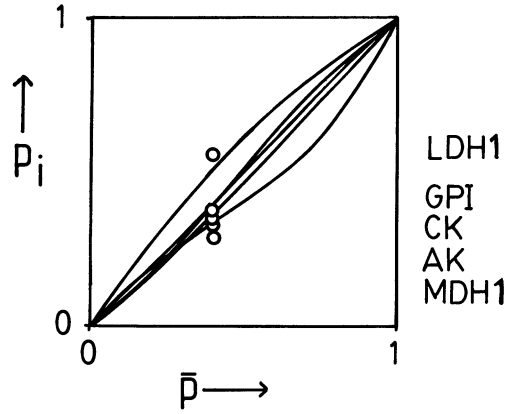

FIG. 3. Comparison of clines at different loci. The frequency at each locus (p_i) has been plotted against the average frequency ($\bar{p}$) in each sample. Space only permits one example to be plotted (Kopanka 1); the solid lines show the curves fitted by polynomial regression (see text and Table 3).

loci (p_i) against the average frequency ($\bar{p} = \Sigma p_i/5$) in the sample (Fig. 3). If the clines coincided exactly, all the points would lie on the diagonal. In fact, they are scattered around the diagonal. This scatter has two causes. First, the clines vary in position and width. Such variations in cline shape can be described by fitting a polynomial which relates the allele frequency at a particular locus (p_i) to the average ($\bar{p}$):

$$p_i = \bar{p} + 2\bar{p}\bar{q}(\alpha + (\bar{p} - \bar{q})\beta).$$

The first component, α, describes an increase in frequency of *variegata* alleles above the average, or in other words, a shift of the cline towards *bombina*'s territory. The second component, β, describes an increase of *variegata* alleles on the *variegata* side and a decrease on the *bombina* side: in other words, a narrowing of the cline. In units of average cline width (defined as 1/maximum slope [Endler, 1977]), α is twice the shift in position, and β is the decrease in cline width below the average. The estimated positions and widths are shown in Figure 3 and Table 3.

Even after removing heterogeneity in position and width, significant random variation in allele frequency remains ($\Delta L_{128} = 131.46$). The amount of variation at different loci seems similar and gives an estimate of the standardized variance F_{ST} (= Var[p]/$\bar{p}\bar{q}$) of 0.0067 (0.0034–0.0100). (This estimate has been multiplied by a factor $\frac{5}{4}$, to allow for the fact that $\bar{p}$ was estimated from the observed allele frequencies.) There are significant differences between sites in the average variability (F_{ST}) ($\Delta L_{33} = 37.81$).

In all subsequent analyses, we ignore these minor variations in frequency between loci and consider only the average frequency of *variegata* alleles, $\bar{p}$.

Cline Shape

The two-dimensional pattern of allele frequencies ($\bar{p}$; Fig. 2) has been reduced to a more tractable one-dimensional transect (Fig. 4a) by measuring the distance of each sample, X, from the center of the morphological hybrid zone, which is indicated on Figure 2. Some samples are too far from the main transect for their distances from the center to be found reliably; these have been omitted from this part of the analysis. We have also pooled successive samples from the same site, since these do not differ significantly in average allele frequency. This reduces the number of distinct points to 23 (Table 1).

The average of the five clines (Fig. 4a) is

TABLE 3. Comparisons between clines at different loci. α and β are coefficients giving the position and width of each cline, respectively, ±1 standard error. $\bar{R}$ is the average standardized linkage disequilibrium between each pair of loci, and F_{ST} is the standardized variance of fluctuations about the fitted cline.

			$\bar{R}$				
Locus	α	β	*Ldh-1*	*Mdh-1*	*Ck*	*Ak*	F_{ST}
Ldh-1	0.24 ± 0.04	−0.02 ± 0.09	—				0.0072 (0.0005–0.0045)
Mdh-1	−0.05 ± 0.04	0.07 ± 0.09	0.140	—			0.0030 (0.0000–0.0088)
Ck	−0.19 ± 0.04	−0.20 ± 0.08	0.100	0.075	—		0.0092 (0.0016–0.0187)
Ak	−0.01 ± 0.04	0.01 ± 0.09	0.092	0.167	0.144	—	0.0055 (0.0000–0.0128)
Gpi	−0.02 ± 0.07	0.09 ± 0.10	0.167	0.171	0.050	0.181	0.0087 (0.0013–0.0179)
							0.0067 (0.0036–0.0101)

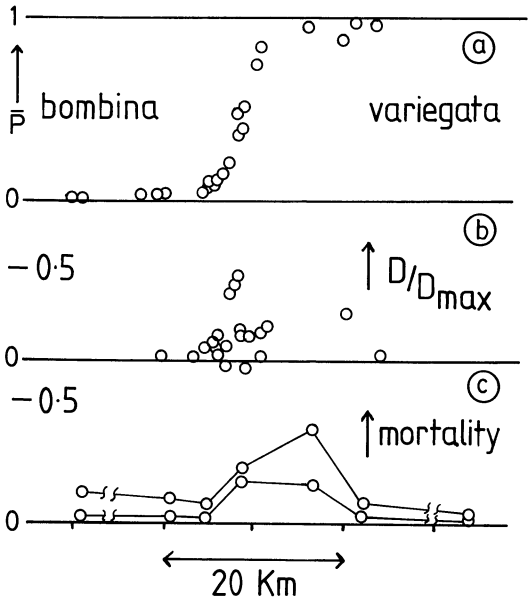

FIG. 4. a) The average allele frequency, plotted against distance from the morphological zone (X in Table 2). b) The standardized linkage disequilibrium (R), averaged over all pairs of loci. c) Mortality from cleavage to gastrulation (lower curve), and from gastrulation to feeding (upper curve), measured from clutches produced by 10 pairs of *Bombina*, from each of seven populations spanning the hybrid zone (Koteja, 1984).

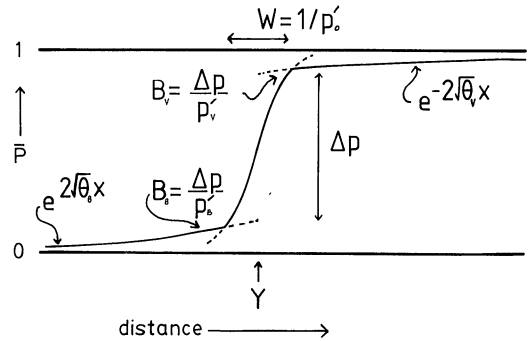

FIG. 5. The model of cline shape (see text for explanation).

reasonably smooth. Most of the change in allele frequency is concentrated in a narrow region, approximately 6 km wide. If selection were acting independently at each locus we would expect a smooth, sigmoidal cline in which the allele frequencies fall away exponentially over a distance of the same order as the width of the central step (Endler, 1977). For example, with selection against heterozygotes, the allele frequency in the tails should fall away as $\exp(-4x/w)$, where w is the cline width (defined as 1/maximum slope) (Bazykin, 1969). Other forms of selection will give very similar patterns (Karlin and Richter-Dyn, 1976). However, introgressing alleles are found in the most distant samples, up to 20 km from the center of the zone (Figs. 2 and 4a), at very much higher frequency than would be expected from simple models of clines.

Possible causes of this "stepped" pattern are discussed below. The problem here is to describe and estimate the shape of the cline. Ideally, we would derive the cline shape from some genetical model (e.g., that in Barton, 1983). However, the calculations involved in finding the shape of a multilocus cline are too hard for this to be practicable. Instead, we use a model which gives the same shape as the theoretical prediction at the center and in the tails, but which differs in between, just at the edge of the hybrid zone. Since few samples are found in this region, our approximation should cause little error. The model is described in Figure 5; the notation is summarized in Table 4. Six parameters are involved: y, w, θ_b, θ_v, B_b, and B_v. In the middle of the cline, $p = (1 + \tanh[2(x - y)/w])/2$; the center of the cline is at a position y, and has width w. This sigmoid curve is spliced onto two exponential tails. The shape of each tail is characterized by two parameters. θ describes the rate of decay ($p \propto \exp(-4x\sqrt{\theta}/w)$; θ is a dimensionless parameter equal to the square of the ratio between the scale over which the tail decays and the cline width. Since the characteristic distance over which allele frequencies change is inversely proportional to the square root of the selection pressure causing the change (Slatkin, 1973), this parameter can be interpreted as the ratio between the selection acting directly against introgression of each enzyme allele (s_e), and the effective selection on each allele within the hybrid zone (s^*). The two parameters θ_b and θ_v measure the rates of decay of the tails on the *bombina* and the *variegata* sides of the hybrid zone. The size of the tails (that is, the total proportion of introgressing alleles) is described by B, the ratio between the difference in allele frequency across the

TABLE 4. Notation and parameter estimates.

ΔL_ν	The difference in $\log_e$ likelihood between hypotheses which differ by ν degrees of freedom.
N	Number of alleles sampled from each site, summed over the five enzyme loci.
$F_{IS} = 0.017$ (0–0.034)	Deficit of heterozygotes, averaged over all sites and all loci.
$F_{ST} = 0.0067$ (0.0034–0.0100)	Standardized variance of allele frequency.
$R = 0.129$ (0.119–0.139)	Standardized pairwise linkage disequilibrium (defined by $R = D/\sqrt{pquv}$, where p, q, u, and v are the allele frequencies at the two loci.
$\sigma = 890$ (790–940) m gen$^{-1/2}$	Dispersal rate (defined as the standard deviation in distance between parent and offspring).
p'	Gradient in frequency of the *variegata* allele, dp/dx.
$w = 6.15$ (5.45–6.45) km	Cline width (defined as 1/maximum gradient = $1/p'_{max}$)
$\theta_b = 0.011$ (0–0.025)	Measures of the rate of decay of the tails of introgression into *bombina*
$\theta_v = 0$ (0–0.011)	and *variegata*, respectively; these tails have the form $\exp(-x\sqrt{\theta}/w)$ (Fig. 5).
$B_b = 160$ (48–430) km	Measures of the barrier to the flow of genes into *bombina* and *variegata*,
$B_v = 280$ (48–400) km	respectively. $B \equiv \Delta p/p'$, where Δp is the size of the step in allele frequency caused by the hybrid zone.
$n = 300$ (80–2,000)	The number of loci at which selection acts against heterozygotes.
$r = 0.123$	The harmonic mean recombination rate between an enzyme marker and the n selected loci.
$(\bar{W}_h/\bar{W}_p) = 0.65$ (0.60–0.77)	The mean fitness of a population at the center of the hybrid zone, relative to the mean fitness of a pure population outside the zone (*bombina* and *variegata* are treated as equivalent here).
$S = 0.86$ (0.52–1.02)	The total selection pressure against hybrids (defined as $-2\log_e[\bar{W}_h/\bar{W}_p]$).
$s^* = 0.170$ (0.159–0.181)	The effective selection pressure on an enzyme marker (defined by $w = \sqrt{\sigma^2/s^*}$).
$s_e = 0.0016$ (0–0.0038)	The selection acting directly on each enzyme allele, outside the hybrid zone (estimated from $\theta = s_e/s^*$).
$s = 0.0027$ (0.0005–0.0065)	The selection acting against heterozygotes at loci responsible for reproductive isolation (in this model, $S = ns$).

central step (Δp), and the gradient of allele frequency near the step (p'):$B = \Delta p/p'$. This has the dimensions of a distance. If the central step is caused by some barrier to gene exchange, B is a measure of the strength of that barrier (Nagylaki, 1976). The two parameters B_b and B_v measure the sizes of the tails on the *bombina* and *variegata* sides of the zone.

The likelihood of various values of these parameters has been calculated numerically using the Metropolis algorithm described above; the estimates are summarized in Table 4. The density of points in Figure 6 is proportional to the likelihood of the pair of parameters plotted on the two axes. The position and width of the enzyme clines are estimated quite accurately (Fig. 6c). However, the shapes of the tails of introgression are harder to determine (Fig. 6a,b). In particular, it is hard to separate the two parameters, B and θ; the low frequency of introgressing alleles could be due to a very strong barrier to gene flow, combined with a rather slow rate of decay beyond the barrier (B large, θ small), or conversely, by a relatively weak barrier, but a rapid rate of decay (B small, θ large). Nevertheless, there is clearly a significant step at the center, equivalent to a distance B of at least 48 km, and most likely $\approx$220 km (taking the average of the estimates for either side of the zone). The rate of decay of the tails is also significantly slower than would be expected from the sharp central step (i.e., $\theta \ll 1$). Indeed, these data do not rule out the possibility that there is no decay at all ($\theta = 0$). However, the absence of foreign alleles in more distant samples (Szymura, 1976, unpubl.) make this unlikely.

The shapes of the tails on either side differ significantly: the likelihood surfaces for the *bombina* and *variegata* tails (Fig. 6a,b) hardly overlap. However, this asymmetry should be treated cautiously: there are rather few samples on the *variegata* side of the zone, and the asymmetry is due to one of these (Głogoczów), which shows an unusu-

ally high frequency of *bombina* alleles, and unusually high disequilibrium. (This pattern is weakened but still persists, even when the single pure *bombina* genotype in the Głogoczów sample is excluded.) In the subsequent analysis, we use estimates of B and θ averaged over both sides of the zone.

Linkage Disequilibrium

The value of the coefficient of linkage disequilibrium (D) is constrained by the allele frequencies and must necessarily be smaller when the population is less polymorphic ($-pq < D < pq$ if all p's are equal). We therefore give our estimates in terms of the correlation coefficient, $R_{ij} = D_{ij}/\sqrt{pq_i pq_j}$, which lies between -1 and $+1$. (Note that disequilibria have only been estimated for sites where both the relevant loci are polymorphic.) There are significant positive disequilibria between all pairs of loci, though the estimates vary from pair to pair (Table 3). There is also variation among sites. Świnna Poręba shows much stronger disequilibrium between all pairs of loci ($\bar{R} = 0.822$) than any other site (see above). There is significant variation among the remaining sites for most pairs of loci ($\Delta L_{252} = 205.72$; Fig. 4b). Averaging over all pairs and all sites except Świnna Poręba, $R = 0.129$ (0.119–0.139). The variations between different pairs of loci can be partly explained by differences in the shapes of the corresponding clines. Wider clines show weaker disequilibria (Fig. 7a), and there is an indication that disequilibria are also weaker for clines that are staggered (Fig.7b).

Dispersal

The distribution of distances travelled by individually marked toads at Kopanka is shown in Figure 8; the standard deviation of distance travelled over one year was 250 m. These *Bombina* mature in their third year and can live more than ten years. Taking the generation time (conservatively) as three years, we can make a rough estimate of the standard deviation of distance between parent and offspring: $\sigma = 430$ m gen$^{-1/2}$; it is this parameter that determines the rate of gene flow (Nagylaki, 1974). This estimate is consistent with studies of *B. variegata* by Beshkov and Jameson (1980) and Flis (1984); the latter found that the mean distance travelled by adults increases with time: 104, 197 and 243 m in consecutive years. Little is known about migrations of juveniles (Płytycz and Bigaj, 1984).

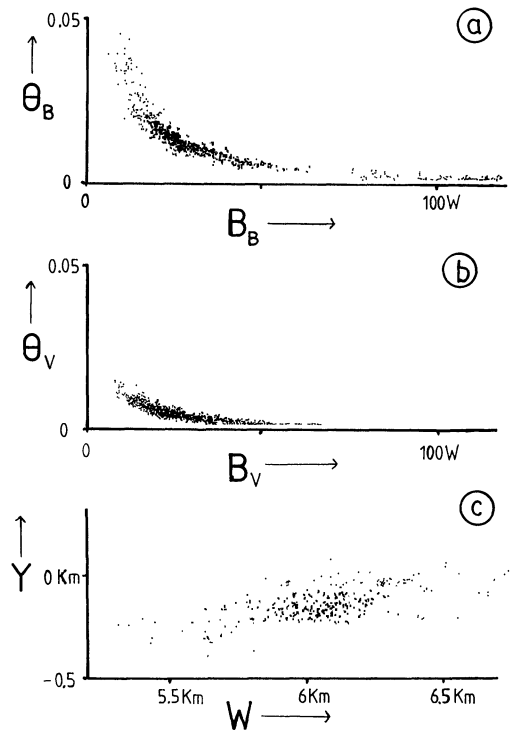

FIG. 6. The likelihood of different cline shapes. Each point represents one step in a random walk, with probability density equal to the likelihood. a and b) Rate of decay of the tails (θ), and barrier strength (B) on the *bombina* and *variegata* sides of the zone, respectively. c) Position (y) and width (w) of the zone.

Hybrid Viability and Morphological Aberrations

Two lines of evidence suggest that there is selection against hybrids: embryonic mortality and morphological aberrations. The mortality of embryos, especially in the first part of embryonic development, seems to be higher than that of the parental species, though this is based on small numbers of crosses. Mean mortality ranged from 8.8 to 37.2% in hybrid populations, as compared to 3.8 to 12.5% outside the hybrid zone. The combined mortality of 49 hybrid clutches in the first developmental period was 6.0% (SD = 12.9%), and 15.6% (SD = 17.5%) in the second period. Mortality for 19 clutches from parental populations was 1.1% (SD =

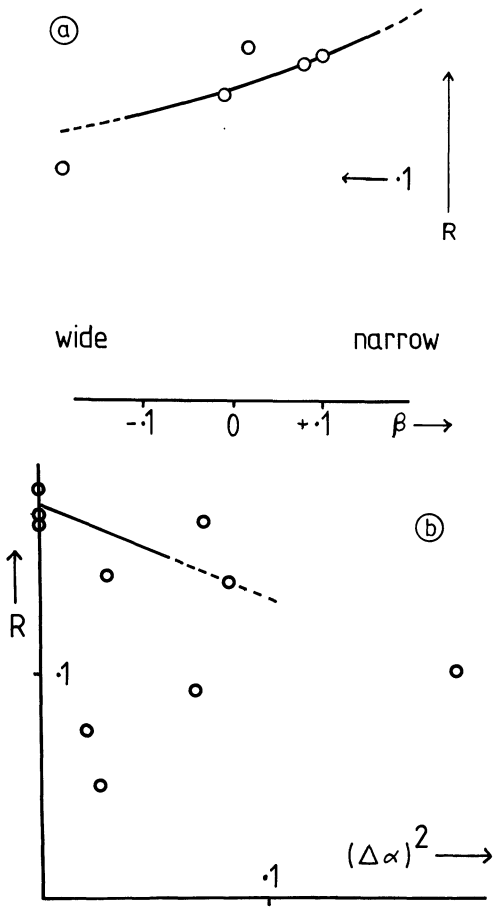

FIG. 7. a) The five points show, for each locus, the standardized linkage disequilibrium (R), averaged over the other four loci and over all sites, as a function of the width of the cline (β). b) The ten points show, for each pair of loci, the standardized linkage disequilibrium (R), averaged over all sites, as a function of the distance between the clines ($\Delta\alpha$). The solid curves in each diagram show the relations expected if linkage disequilibrium is generated by dispersal. (See Table 3 and text for details.)

1.5%) for *B. bombina* and 8.1% (SD = 11.1%) for *B. variegata* (Koteja, 1984).

Individuals from the hybrid zone show various developmental anomalies, which are less frequent in either pure *B. bombina* or *B. variegata*. Although one would expect greater morphological variation in genetically enriched hybrid populations, it seems more likely that the higher level of abnormal phenotypes is caused by a breakdown of developmental homeostasis. Hybrid tadpoles, collected in the field, demonstrate various deformations of tooth row patterns (Czaja, 1980), which may impair feeding. Adult individuals may show a number of vertebral column anomalies (Madej, 1965). The frequency of these anomalies is about twice as high in the hybrid zone as in parental populations (44.5% vs. 23.2%; Szymura, unpubl.); similar effects of hybridization were also observed by Madej (1965). Finally, the characteristic spotting patterns on the dorsal and ventral sides of the firebellied toads tend to be less regular and more asymmetric in the hybrid zone (Kukuła, 1979).

Summary of Results

All populations, apart from the aberrant sample from Świnna Poręba, are very close to Hardy-Weinberg proportions ($F_{IS} = 0.017$ (0–0.0034)). Clines at different loci coincide closely, though there is some variation in shape and some random fluctuation ($F_{ST} = 0.0067$ (0.0034–0.01)). The clines also coincide with morphological and mitochondrial DNA clines, and they lie on a broad environmental transition between habitats suitable for *bombina* and *variegata*. However, there is no exact relation between local environment and cline position. The clines are sharply stepped, with most of the change occurring over 6 km, but with long tails of introgression on either side. There are strong positive "linkage disequilibria" between each of these pairs of unlinked loci ($R = 0.129$). Individuals move considerable distances ($\sigma \approx 430$ m gen$^{-1/2}$). Natural hybrids show a higher frequency of morphological aberrations, and their offspring show signs of increased mortality in the laboratory.

Inferences

We will now show that the pattern of genotype frequencies across the hybrid zone in *Bombina* can be explained by theoretical models of multilocus clines (Barton, 1983; Barton and Bengtsson, 1986). This allows us to use observations of linkage disequilibrium (R) and cline shape (w, B, θ) to estimate dispersal rate, selection strength, and numbers of loci under selection. These indirect estimates can then be compared with direct measurements to test the plausibility of our explanation.

Linkage Disequilibrium.—The substan-

tial positive linkage disequilibrium found in the *Bombina* hybrid zone must be maintained by some force which is strong enough to counter the independent segregation of these unlinked genes. In any set of coincident clines, disequilibrium will be produced by the dispersal of parental combinations of alleles into the center (Li and Nei, 1974; Slatkin, 1975). Positive disequilibrium may also be generated if parental combinations of alleles confer greater fitness. However, when many genes are involved ($\log(n) \gg 1$), which we will show to be likely in *Bombina,* dispersal will generate more disequilibrium than will epistasis (Barton, 1983).

If dispersal is the main process involved, then the strength of disequilibrium, relative to its maximum possible value, R, will be greatest in the center, and will decline towards the edges. This is because genomes which have introgressed across the zone will have been exposed to more generations of recombination and segregation than will those in the center. Putting this quantitatively, the strength of disequilibrium is approximately $D_{ij} = \sigma^2 p'_i p'_j / r_{ij}$ for $D \ll pq$; σ^2 is the dispersal rate (variance in parent-offspring distance), and $p'_i p'_j$ are the gradients of the clines (see Barton, 1982). Hence, R is proportional to $p'_i p'_j / \sqrt{pq_i pq_j}$, which declines exponentially outside the hybrid zone. This expectation fits with the pattern seen in Figure 4b, apart from the sample from Głogoczow, which we discuss below. We can also examine the variation in disequilibrium among different pairs of clines. The expected relations between average disequilibrium and cline shape are shown by the solid curves in Figure 7; these were calculated by integrating the product of gradients across the hybrid zone, assuming that the two clines have the observed position and width (α and β) and have the shape predicted by equation 23 in Barton (1983). The average linkage disequilibrium should be lower for wider clines (Fig. 7a), and for clines that are staggered (Fig. 7b). However, although the observed relations are consistent with the generation of disequilibrium by dispersal, other processes might be expected to give similar patterns: for example, epistasis ϵ would give $R = \epsilon pq_i pq_j / r_{ij} \sqrt{pq_i pq_j}$.

At the center of the zone (that is, at the point where the cline is steepest), dispersal should give $R = \sigma^2 p'^2 / pqr = 4\sigma^2 / w^2 r$ (where w is the cline width and $r = \frac{1}{2}$ for unlinked loci). The average R across the transect is 0.129; the value of R at the center, and for two clines which coincide precisely, will be rather higher. Extrapolating from Figures 4b and 7, we reach a figure of $R \approx 0.17$. Hence, the dispersal rate needed to produce this disequilibrium is $\sigma = w\sqrt{Rr/4} = 890$ (790–940) m gen$^{-\frac{1}{2}}$. (We do not give support limits for R, since the uncertainties here arise from between-site and between-locus variation, rather than from sampling errors. The limits for σ only include the main contribution to the error, which arises from estimating cline width.)

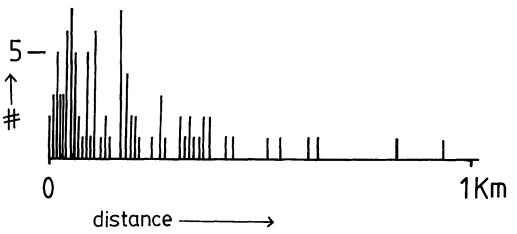

FIG. 8. Distances moved by 93 adult *Bombina* at Kopanka, over one year. Mean square distance: σ = 0.25 km year^{-1}.

This is rather higher than the direct measurement of ≈430 m gen$^{-\frac{1}{2}}$. However, the latter is likely to be an underestimate. Firstly, such direct measures inevitably miss those individuals who move very long distances; if the distribution of distances moved is leptokurtic (as in Fig. 8), this may cause a large error. Secondly, much gene flow may occur through the extinction and recolonization of local populations. The drying up of ponds inhabited by hundreds of toads has been observed by one of us (JMS) in Tyniec on several occasions; marked *Bombina* were recovered in less than a month's time in nearby ponds, 200–400 meters away. Dispersal rate may therefore vary greatly over time, and the data on movement at Kopanka could represent a stagnant period in the population's life. Finally, the generation time of *Bombina* may be rather more than the three years assumed in extrapolating from movements over a single year. More generally, indirect genetic measures of gene flow in a variety of organisms are usually substantially higher than direct measures (Slat-

kin, 1985). The estimate of $\sigma = 890$ m gen$^{-1/2}$ therefore seems reasonable.

We can now use this estimate of dispersal to find the strength of selection needed to maintain the enzyme clines (s^*): the narrower a cline is relative to dispersal, the stronger the selection required to maintain it. The width of any cline maintained by a balance between dispersal and selection is approximately equal to the characteristic scale of selection (Slatkin, 1973), which we define here as $\ell = \sqrt{8\sigma^2/s^*}$, where s^* is a measure of the selection pressure maintaining the cline. For a cline maintained by heterozygote disadvantage s^*, the width $w = \ell$; for other models, $w \simeq \ell$. Our estimates of σ and w give $s^* = 0.170$ (0.159–0.181). It is important to point out that this is not the total selection pressure acting on all the differences between *bombina* and *variegata*; neither is it the selection pressure acting on each individual enzyme locus. It is the "effective selection pressure" at a single locus that would be required in order to maintain a cline of the observed width. In the hybrid zone, we know that there are substantial linkage disequilibria, even between unlinked loci. The shape of the cline at a particular locus depends on the selection pressure on all loci in linkage disequilibrium with the observed locus, an effect analogous to "hitch-hiking" in a single population. The effective selection pressure is the selection acting on the individual enzyme locus and on all those loci in disequilibrium with it: $s^*_i = \Sigma s_j D_{ij}$.

Cline Shape.—The most striking feature of the enzyme clines is that they are sharply stepped, with most change occurring within 6 km, but with alleles introgressing tens of kilometers to either side. The most obvious explanation for this pattern is suggested by the strong disequilibria at the center of the zone. At the center, the "effective selection pressure," s^*, maintaining the cline at each locus is made up of contributions from many other correlated loci. The clines will therefore be steep. Outside the hybrid zone, we expect linkage disequilibria to be much weaker: by the time a chromosome has passed through the hybrid zone, the disequilibria will have decayed, the selection should act on each individual locus. The clines should therefore tail away rather slowly. It is clear that the enzyme alleles are selected against very weakly, if at all, in a foreign environment or a foreign genetic background. The estimates of the rate of decay of the tails of introgression (θ) give the selection on these loci as $s_e = 0.16\%$ (0–0.38%).

The step induced by the interactions between genes in the center of a set of coincident clines can be thought of as a barrier to gene flow. Neutral or weakly selected alleles in *bombina* can only flow into *variegata* by recombining into the genetic background of *variegata*; this recombination must occur within the narrow hybrid zone and must take place before selection eliminates the foreign alleles with which the neutral allele we are following is associated. The strength of such a barrier is given by B, which is the ratio between the size of the step in allele frequency and the gradient on either side. In this case, $B = 220$ (48–415) km (averaging over gene flow in the two directions). This can be thought of as the length of unimpeded habitat that would present an equivalent obstacle to neutral intergradation (Barton, 1979).

The strength of the barrier depends primarily on the ratio between the rate of selection and the rate of recombination. The relationship between the barrier strength and this ratio is remarkably simple and depends little on how selection acts. In particular, selection might act to favor different genotypes in different environments, or it might act against hybrids. Provided that fitnesses are not frequency-dependent, and provided that disequilibria are not extremely strong, $B = w(\bar{W}_h/\bar{W}_p)^{-1/r}$ (Barton and Bengtsson, 1986). Here, $\bar{W}_h/\bar{W}_p$ is the mean fitness of a hybrid population relative to a parental population, and r is the harmonic mean recombination rate between the locus we are considering and all the loci involved in maintaining the genetic differences between the taxa. We have already estimated B and w, and we can calculate r from the number of chromosomes and chiasma frequency in *Bombina*. There are 12 pairs of metacentrics and about one chiasma per arm in both taxa (Morescalchi, 1965). Since the chromosomes vary in length, we can take the effective number to be (say) eight pairs. If we assume that n loci are under selection and that these are spread evenly over eight

chromosomes, each of map length 1 Morgan, then each will be $\approx 8/n$ Morgans from its nearest neighbors. If we assume that the enzyme loci are about the same distance from their neighbors, and average over all possible positions in the genome, we find that $r^{-1} \approx 2(\log(n) - 1.627)$. The harmonic mean recombination rate decreases somewhat as the number of loci increases; taking $n = 300$, we find that $r \approx 0.123$. Combining this with estimates of B and w, we find that $\bar{W}_h/\bar{W}_p = 0.65$ (0.60–0.77). We can thus make the strong prediction that hybrid populations must have substantially reduced mean fitness in order to account for the shape of the enzyme clines. The high frequency of morphological aberrations in the zone and the suggestion of increased mortality of hybrid embryos in the laboratory make this plausible and suggest that selection acts primarily against hybrids, rather than favoring different genotypes in different environments.

It is possible that the stepped pattern is partly caused by a few long-distance migrants moving across the zone within a single generation. Such movements do occur. At Świnna Poręba, three adults with apparently pure *B. bombina* genotypes probably represent toads invading *B. variegata* territory along the narrow Skawa River Valley from abundant *B. bombina* populations occupying the Vistula Valley near Zator, about 15 km north of Świnna Poręba (see Michałowski [1958] for details of the *Bombina* distribution in that region). However, there seems to be very little long-distance migration across the main transect. If we consider those populations more than 3 km from the center of the zone, or (where their position relative to the center is uncertain) with the rarest allele at less than 15%, then there are no putative F_1 individuals, and only one putative parental genotype (at Głogoczów) out of 948 individuals (Table 2). An influx of migrants at this rate of $\approx 10^{-3}$ per generation would introduce foreign alleles at about the same rate as they diffuse across the zone; the influx by recombination through the zone is proportional to the gradient at the edge of the cline: $\sigma^2 p'/2 \simeq \sigma^2/2B \approx 2 \times 10^{-3}$. However, long-distance migrants will be subject to strong selection, since they carry an entirely foreign genome, whereas those alleles diffusing across the zone will be subject to very much weaker selection: roughly s_e (<0.0038) as compared with S (≈ 0.8). Leptokurtic dispersal therefore seems likely to have a negligible effect on cline shape, in this case.

Number of Loci under Selection.—We now have three distance scales: the barrier strength, B (which is a measure of the size of the tails of introgression), the width of the enzyme clines, w, and the dispersal distance, σ (estimated from the linkage disequilibria). What can these measures tell us about the number of genes which are responsible for the reproductive isolation between *bombina* and *variegata*? We have used the ratio between the barrier strength and the width of the enzyme clines (B/w) to give an estimate of the total selection pressure maintaining the hybrid zone, and the ratio between the width of the enzyme clines and the dispersal rate gives the effective selection pressure on each enzyme marker. The ratio between these two selection pressures is smaller than the actual number of loci under selection, since the "effective selection pressure" is the net effect of all loci in linkage disequilibrium with the marker locus. However, if we assume some specific model of selection, this ratio can be calculated as a function of the actual number of loci and the ratio between selection and recombination (Barton, 1979, 1983; Barton and Hewitt, 1982). The precise relation depends to some extent on the way selection acts; we will suppose that selection against heterozygotes, s, acts multiplicatively at n loci, giving a total selection pressure of $S = ns$ and a mean fitness at the center of $\exp(-S/2)$. This is the simplest model and leads to results typical of more realistic schemes (Barton, 1983; Barton and Bengtsson, 1986). It gives $s = 0.27\%$ (0.05%–0.65%), and $n = 300$ (80–2,000) loci.

This method is similar to that used by Barton and Hewitt (1981*b*), who compared the width of the cline in hybrid inviability in the grasshopper *Podisma pedestris* with the dispersal rate. Both these distances were measured directly; here, we have inferred them from electrophoretic data. It is important to note that we do not assume that the enzyme markers are typical of the se-

lected genes or, indeed, that they are under any selection at all.

Quite apart from these calculations, it is clear that the electrophoretic pattern in *Bombina* cannot be accounted for by selection at a few (<10, say) genes. First, the strength of the barrier to gene flow caused by a hybrid zone increases as the number of loci involved increases; unless the total selection pressure is extremely strong, a strong barrier can only be generated when many loci are involved. This is primarily because as selection is spread over more loci, the selection on each locus becomes weaker, and the clines become correspondingly broader (Barton and Hewitt, 1982). Second, the similarity in patterns at these five unlinked loci argues that selection cannot be concentrated on a few chromosomes. Finally, the extensive biochemical divergence and, more to the point, the many adaptive morphological differences (Table 1) suggest that our estimate that several hundred loci are under selection is reasonable.

Discussion

The coincident and sharply stepped clines at each of five enzyme loci, together with the strong linkage disequilibria between these loci, imply that there is a substantial barrier to gene flow between *bombina* and *variegata,* which is caused by the cumulative effects of weak selection at several hundred loci. The loose correlation between the position of the hybrid zone and local habitat, and the evidence of inviability and morphological aberrations in hybrids, suggest that selection acts primarily against hybrids, though different alleles may well be favored in different places. However, our inferences about gene flow, selection strength, and numbers of loci do not depend on whether the clines are maintained by selection against hybrids or by environmental heterogeneity.

The sharp step in each enzyme cline implies a barrier to gene flow, averaged over both directions, of $\approx$220 km. This would delay the introgression of neutral alleles for $(B/\sigma)^2 \approx 60{,}000$ generations, about the same delay as would be caused by an unimpeded stretch of habitat 220 km across (Barton and Hewitt, 1985). However, even a slightly advantageous allele will penetrate the hybrid zone much more rapidly. This is because, whereas a large proportion of neutral genes must cross in order to reduce differentiation across the zone appreciably, only a small proportion of advantageous alleles need cross before increasing exponentially in frequency. For example, the advance of an allele with an advantage of only $S = 1\%$ would be delayed (assuming no dominance) for only $(1/S)\log((B/\sigma)\sqrt{S}) \simeq 320$ generations (Barton, 1979). This is much less than the time it would take for the wave of advance of such an allele to move over 220 km of free habitat ($\approx$1,750 generations; Fisher, 1937). Thus, although selection maintains considerable divergence for many characters (Table 1) and greatly impedes neutral introgression, *bombina* and *variegata* need not evolve in complete independence and cannot be considered, in the strict sense, as separate biological species (sensu Mayr, 1942). Nevertheless, the extensive divergence between these two taxa would lead the majority of systematists (including one of us: JMS) to regard them as good species; this example suggests that the pattern of gene flow may not be an appropriate basis for taxonomy.

The long, shallow tails of introgression of the electrophoretic alleles imply that these alleles are only weakly selected and may be neutral ($s_e < 0.38\%$). If they are neutral, then the calculation above shows that selection at other loci could maintain differentiation for long periods ($\approx$60,000 generations, $\approx$180,000 years) and certainly for much longer than the age of the present hybrid zone. However, molecular clock arguments suggest that *bombina* and *variegata* have been diverging for several million years (Szymura, 1983; Maxson and Szymura, 1984; Szymura et al., 1985); this will be an underestimate if gene flow has impeded molecular divergence. Over such long periods, the hybrid zone should not present a barrier strong enough to prevent introgression or to allow the accumulation of neutral differentiation in the first place. If we consider the two taxa to constitute populations that are effectively panmictic over these long time scales, then the exchange of as few as one individual per generation should impede divergence by drift (Kimura and Maruyama,

1971). Of course, the taxa are in fact distributed over wide geographic areas, and their ranges have presumably often been disrupted by climatic changes. These theoretical arguments may not, therefore, be directly applicable. However, they do suggest that the electrophoretic alleles may have diverged as a result of weak selection, rather than by drift alone. In that case, the basis of the molecular clock in *Bombina,* and in general, becomes unclear, and the estimates of divergence time must be treated cautiously.

Despite the strong selection pressures, which should favor the reinforcement of mating isolation in hybrid zones, and the extensive theoretical discussion of this possibility, there is remarkably little evidence that such reinforcement has actually evolved (Barton and Hewitt, 1985). The apparently random mating in the *Bombina* hybrid zone is especially surprising, since the pure forms show strong ecological preferences, have different mating calls, and, most directly, show strong mating preferences in the laboratory: mating a *bombina* male to a *variegata* female requires stimulation of the male with a gonadotrophin injection to induce amplexus. (The reverse combination is easily accomplished; Michałowski and Madej, 1969.) In some areas, mating preferences do allow the two types to retain their integrity in sympatry: in Yugoslavia, for example, there is a bimodal distribution of the electrophoretic hybrid index (Szymura, unpubl.). Such isolation may explain the occasional presence of pure *bombina* at Głogoczów and Świnna Poręba, on the *variegata* side of the zone. However, it is clear that across the zone at Cracow, and also in Austria (Gollmann, 1984), mating preferences have broken down, rather than being reinforced. This breakdown may be because the two parental types are replaced with a wide variety of recombinant genotypes. Once hybrid populations are formed, any allele which causes assortment for some character will only gain an advantage from reducing the proportion of alleles in the wrong physical or genetic background at loci with which the character is in linkage disequilibrium. This advantage will therefore be, at most, approximately $s^* = 17\%$ and may be swamped by gene flow from outside the zone, where the modified mating system is presumably at a disadvantage (Barton and Hewitt, 1981*a*). However, further work on mating behavior in the zone is needed before this can be more than speculation.

Finally, what does our knowledge of the present nature and effects of the differences between *bombina* and *variegata* tell us about the origin of these differences? The genetic basis of reproductive isolation may provide one of the few clues to the mechanism of speciation (Templeton, 1981). The large number of loci considered here and the variety of characters which have diverged strongly suggest that reproductive isolation has evolved in many small steps, rather than in one or a few drastic "genetic revolutions." Natural selection has clearly been important in producing many of the adaptive morphological differences (Table 1). However, it is hard to judge the relative importance of (for example) random drift and natural selection from present information (Barton and Charlesworth, 1984). The continual process of evolution may have erased the evidence of its own mechanism, leaving us with a record as imperfect as that from fossils. If so, we must be content with an understanding of contemporary and observable evolutionary processes.

Acknowledgments

We are grateful to J. Mitton and W. P. Hall for their suggestions and help with earlier versions of the statistical analysis. The manuscript was much improved by the helpful comments of Dorothy Currie, Gunther Gollmann, Godfrey Hewitt, Julian MacLean, and Jim Mallet. Thanks are also due to Tina Tsang for her careful typing. This work was supported by the Exchange Agreement between the Polish Academy of Sciences and the Royal Society, and by grants from the Polish Academy of Sciences (project MR-II/6), the Royal Society, the Nuffield Foundation, and the Science and Engineering Research Council.

Literature Cited

Andren, C., G. Nilson, and R. Podloucky. 1984. The fire-bellied toad *Bombina bombina* (L.). SEH Conservation Committee, Council of Europe, Goteborg, Sweden.

Arntzen, J. W. 1978. Some hypotheses on postgla-

cial migrations of the fire bellied toad, *Bombina bombina* (L.) and the yellow bellied toad, *Bombina variegata*. J. Biogeog. 5:339–345.

BARTON, N. H. 1979. Gene flow past a cline. Heredity 43:333–339.

———. 1982. The structure of the hybrid zone in *Uroderma bilobatum*. Evolution 36:863–866.

———. 1983. Multilocus clines. Evolution 37:454–471.

BARTON, N. H., AND B. O. BENGTSSON. 1986. The barrier to genetic exchange between hybridising populations. Heredity. *In press.*

BARTON, N. H., AND B. CHARLESWORTH. 1984. Genetic revolutions, founder effects, and speciation. Ann. Rev. Ecol. Syst. 15:133–164.

BARTON, N. H., AND G. M. HEWITT. 1981*a*. Hybrid zones and speciation, pp. 109–145. *In* W. R. Atchley and D. S. Woodruff (eds.), Evolution and Speciation: Essays in Honour of M. J. D. White. Cambridge Univ. Press, Cambridge, U.K.

———. 1981*b*. The genetic basis of hybrid inviability in the grasshopper *Podisma pedestris*. Heredity 47: 367–383.

———. 1982. Hybrid zones as barriers to gene flow, pp. 341–359. *In* G. S. Oxford and D. Rollinson (eds.), Protein Polymorphism: Adaptive and Taxonomic Significance. Blackwell, Oxford, U.K.

———. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst. 16:113–148.

BAZYKIN, A. D. 1969. Hypothetical mechanisms of speciation. Evolution 23:685–687.

———. 1973. Population genetics of disrupting and stabilizing selection. II. Systems of adjacent populations and populations within a continuous area. (In Russian.) Genetika 9:156–166.

BESHKOV, V. A., AND D. L. JAMESON. 1980. Movements and abundance of the yellow-bellied toad *Bombina variegata*. Herpetologica 36:365–370.

BINDER, K. 1979. Monte Carlo Methods. Springer-Verlag, Berlin, W. Ger.

BOULENGER, G. A. 1886. On two European species of Bombinator. Proc. Zool. Soc. Lond. 1886:499–501.

CLARKE, B. C. 1966. The evolution of morph ratio clines. Amer. Natur. 100:389–400.

CZAJA, A. 1980. Studies on the variability of the oral apparatus of tadpoles of *Bombina bombina, Bombina variegata,* and their natural hybrid populations. (In Polish.) M.S. Thesis. Jagiellonian Univ., Cracow, Poland.

CZOPKOWA, G., AND J. CZOPEK. 1955. The vascularisation of the respiratory surfaces in *Bombina variegata*. Bull. Acad. Polon. Sci., Cl. II 3:313–318.

DOBZHANSKY, TH. 1940. Speciation as a stage in evolutionary divergence. Amer. Natur. 74:312–321.

EDWARDS, A. W. F. 1972. Likelihood. Cambridge Univ. Press, Cambridge, U.K.

ENDLER, J. A. 1977. Geographic Variation, Speciation, and Clines. Princeton Univ. Press, Princeton, N.J.

FISHER, R. A. 1925. Statistical Methods for Research Workers. Oliver and Boyd, London, U.K.

———. 1930. The Genetical Theory of Natural Selection. Oxford Univ. Press, Oxford, U.K.

———. 1937. The wave of advance of advantageous genes. Ann. Eugen. 7:355–369.

FLIS, E. 1984. Studies on migrations in a population of the yellow-bellied toad *Bombina variegata* in the Beskid Wyspowy Mts. (In Polish.) M.S. Thesis. Jagiellonian University, Cracow, Poland.

GOLLMANN, G. 1984. Allozymic and morphological variation in the hybrid zone between *Bombina bombina* and *Bombina variegata* (Anura, Discoglossidae) in northeastern Austria. Z. Zool. Syst. Evolut.-forsch. 22:51–64.

HACKING, I. 1965. Logic of Statistical Inference. Cambridge Univ. Press, Cambridge, U.K.

HAFNER, J. C., D. J. HAFNER, J. L. PATTON, AND M. F. SMITH. 1983. Contact zones and the genetics of differentiation in the pocket gopher *Thomomys bottae* (Rodentia, Geomyidae). Syst. Zool. 32:1–20.

HALDANE, J. B. S. 1948. The theory of a cline. J. Genet. 48:277–284.

HILL, W. G. 1974. Estimation of linkage disequilibrium in random mating populations. Heredity 33: 229–239.

HORBULEWICZ, L. 1927. Die geographische Verbreittung der *Bombinator* und *Triton*-Arten in Bereiche der Bezirke Sambor, Drohobycz, Stryj (Kleinpolen). (In Polish, with German summary). Bull. Int. Acad. Polon. Sci. Let., Cracovie, Ser. B 1–2:87–111.

———. 1933. Die geographische Verbreittung der *Bombinator*-Arten in Bereiche der Bezirke Dobromil-Przemysl-Jarosław. (In Polish, with German summary.) Kosmos (Lwow) A 58:209–223.

KARAMAN, S. 1922. Uber unsere Bombinatorarten. (In German.) Glasnik. Kroat. Naturwiss. Ges. Zagreb 34:63–70.

KARLIN, S., AND N. RICHTER-DYN. 1976. Some theoretical analyses of migration-selection interactions in a cline: A generalised two-range environment, pp. 659–706. *In* S. Karlin and E. Nevo (eds.), Population Genetics and Ecology. Academic Press, N.Y.

KEY, K. H. L. 1968. The concept of stasipatric speciation. Syst. Zool. 17:14–22.

———. 1981. Species, parapatry, and the morabine grasshoppers. Syst. Zool. 30:425–458.

KIMURA, M., AND T. MARUYAMA. 1971. Pattern of neutral polymorphism in a geographically structured population. Genet. Res. 18:125–131.

KOTEJA, E. 1984. Embryonic mortality of the fire bellied toad (*Bombina bombina*), the yellow bellied toad (*Bombina variegata*), and their hybrids. (In Polish.) M.S. Thesis. Jagiellonian University, Cracow, Poland.

KUKUŁA, K. 1979. Comparison of morphological and biochemical variation in populations from the hybrid zone between *Bombina bombina* and *Bombina variegata*. (In Polish.) M.S. Thesis. Jagiellonian University, Cracow, Poland.

LI, W. H., AND M. NEI. 1974. Stable linkage disequilibrium without epistasis in subdivided populations. Theoret. Popul. Biol. 6:173–183.

LÖRCHER, K. 1969. Vergleichende bio-akustiche Untersuchungen an den Rot- und Gelbbauchunke, *Bombina bombina* (L.) und *B. v. variegata* (L.). (In German, with English summary.) Oecologia 3:84–124.

MADEJ, Z. 1964*a*. Studies on the fire bellied toad (*Bombina bombina,* Linnaeus, 1761) and yellow bellied toad (*Bombina variegata,* Linnaeus, 1758)

of Upper Silesia and Moravian Gate. Acta Zool. Cracov. 9:291–336.

———. 1964*b*. Bellied toads (*Bombina* Oken, 1816) of the Sadecki Beskid Mts. Acta Zool. Cracov. 9: 761–774.

———. 1965. Variations in the sacral region of the spine in *Bombina bombina* (Linnaeus, 1761) and *Bombina variegata* (Linnaeus, 1758) (Salientia, Discoglossidae). Acta Biol. Cracov., Ser. Zool. 8: 185–197.

———. 1973. Ecology of European bellied toads. (*Bombina* Oken, 1816). (In Polish, with English summary.) Przezl. Zool. (Wrocław) 17:200–204.

MAXSON, L., AND J. M. SZYMURA. 1984. Relationships among discoglossid frogs: An albumin perspective. Amphibia-Reptilia 5:245–252.

MAYR, E. 1942. Systematics and the Origin of Species. Columbia Univ. Press, N.Y.

MEHELY, L. 1905. Die herpetologischen Verhaltmise des Mecsek-Gebirges und der Kapela. (In German.) Ann. Hist. Mus. Nat. Hungary (Budapest) 3:256–316.

METROPOLIS, N., A. ROSENBLUTH, M. ROSENBLUTH, A. TELLER, AND E. TELLER. 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21:1087–1092.

MICHAŁOWSKI, J. 1958. Geographische Verbreittung der Unken (*Bombina* Oken) zwischen den Flussen Weichsel, Skawa und Raba (Woiwodschaft Krakow). (In Polish, with Russian and German summaries.) Acta Zool. Cracov. 3:247–283.

MICHAŁOWSKI, J., AND Z. MADEJ. 1969. Studies on the relationship of *Bombina bombina* (Linnaeus) and *Bombina variegata* (Linnaeus). III. Taxonomic characters in both species from laboratory and interspecific hybrids. Acta Zool. Cracov. 14:173–200.

MOORE, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. Q. Rev. Biol. 52:263–278.

MORESCALCHI, A. 1965. Osservazioni sulla cariologia di *Bombina*. (In Italian, with English summary.) Boll. Zool. 32:207–218.

NAGYLAKI, T. 1974. Conditions for the existence of clines. Genetics 80:595–615.

———. 1976. Clines with variable migration. Genetics 83:867–886.

OLMO, E., A. MORESCALCHI, V. STINGO, AND G. ODIERNA. 1982. Genome characteristics and the systematics of the Discoglossidae (Amphibia: Salientia). Monitore Zool. Ital. (N.S.) 16:283–299.

PŁYTYCZ, B., AND J. BIGAJ. 1984. Preliminary studies on the growth and movements of the yellow-bellied toad, *Bombina variegata* (Anura: Discoglossidae). Amphibia-Reptilia 5:81–86.

SANCHIZ, F. B., AND M. MŁYNARSKI. 1979. Remarks on the fossil anurans from the Polish Neogene. Acta Zool. Cracov. 24:153–174.

SLATKIN, M. 1973. Gene flow and selection in a cline. Genetics 75:733–756.

———. 1975. Gene flow and selection in a two-locus system. Genetics 81:209–222.

———. 1985. Gene flow in natural populations. Ann. Rev. Ecol. Syst. 16:393–430.

SZYMURA, J. M. 1976. Hybridisation between discoglossid toads *Bombina bombina* and *B. variegata* in southern Poland as revealed by the electrophoretic technique. Z. Zool. Syst. Zool.-forsch. 14:227–236.

———. 1983. Genetic differentiation between hybridizing species *Bombina bombina* and *Bombina variegata* (Salientia, Discoglossidae) in Poland. Amphibia-Reptilia 4:137–145.

SZYMURA, J. M., AND I. FARANA. 1978. Inheritance and linkage analysis of five enzyme loci in interspecific hybrids of toadlets, genus *Bombina*. Biochem. Genet. 16:307–319.

SZYMURA, J. M., C. SPOLSKY, AND T. UZZELL. 1985. Concordant change in mitochondrial and nuclear genes in a hybrid zone between two frog species (genus *Bombina*). Experientia 41:1469–1470.

TEMPLETON, A. R. 1981. Mechanisms of speciation—A population genetic approach. Ann. Rev. Ecol. Syst. 12:23–48.

WHITE, M. J. D. 1968. Models of speciation. Science 158:1065–1070.

———. 1978. Modes of Speciation. Freeman, San Francisco, CA.

WICKBOM, T. 1949. Further cytological studies on Anura and Urodela. Hereditas (Lund) 35:33–48.

ŻAKI, Z. 1974. Die Archaeologie Klienpolens in Fruhen Mittelalten. (In Polish, with German summary.) Prace Komisji Archeologicznej P.A.N. Wyd. P.A.N., o/Krakow, Cracow, Poland.

ZASLAVSKII, V. A. 1967. Reproductive mutual elimination as an ecological factor. (Ecological consequences of the genetic interaction of populations). (In Russian; English translation: British Library, RTS 5428.) J. Gen. Biol. 28:3–11.

Corresponding Editor: M. Slatkin