Evolution, 49(6), 1995, pp. 1224–1238

NATURAL SELECTION ON QUANTITATIVE TRAITS IN THE *BOMBINA* HYBRID ZONE

BEATE NÜRNBERGER,[1] NICK BARTON, CATRIONA MACCALLUM, JASON GILCHRIST, AND MICHAEL APPLEBY
Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom
[1]*E-mail: bnur@castle.ed.ac.uk*

Abstract.—Observations on the means, variances, and covariances of quantitative traits across hybrid zones can give information similar to that from Mendelian markers. In addition, they can identify particular traits through which the cline is maintained. We describe a survey of six traits across the hybrid zone between *Bombina bombina* and *Bombina variegata* (Amphibia: Discoglossidae) near Pešćenica in Croatia. We obtained laboratory measuments of the belly pattern, skin thickness, mating call, skeletal form, egg size, and the developmental time of tadpoles. Although offspring from hybrid populations showed no evidence of reduced viability, a third of the F_1 families failed completely, irrespective of the direction of the cross. All traits differed significantly between the taxa. Clines in belly pattern, skin thickness, mating call, and skeletal form were closely concordant with clines in four diagnostic enzyme loci. However, the cline in developmental time was displaced towards *bombina,* and the cline in egg size was displaced towards *variegata.* This discordance could be because the traits are not inherited additively or because they are subject to different selection pressures. We favor the latter explanation in the case of developmental time. We show that moderate selection acting directly on a trait suffices to shift its position; rather stronger selection is needed to change its width appreciably. Within hybrid populations, there are significant associations among quantitative traits, and between traits and enzymes. Phenotypic variances also increase in hybrid populations. These observations can be explained by linkage disequilibria among the underlying loci. However, the average magnitude of the covariance between traits is about half that expected from the linkage disequilibria between enzyme loci. The discrepancy is not readily explained by nonadditive gene action. This puzzle is now unresolved and calls for further investigation.

Key words.—Cline, egg size, hybrid zone, linkage disequilibrium, quantitative genetics, tadpole development.

Received March 10, 1994. Accepted September 6, 1994.

Stable hybrid zones give rise to distinct clines in genetic or phenotypic traits. They owe their existence to natural selection, which counteracts the otherwise inevitable erosion of species differences through interbreeding. The analysis of selection in hybrid zones cannot only explain why two taxa remain distinct but may also provide insights into the genetic basis of the traits that maintain the distinction. Hybridization produces recombinant genotypes and phenotypes, thus allowing us to study the hitherto elusive relationship between these two organismal levels. Any trait that forms a stable cline in which two taxa meet and hybridize is a candidate for such an analysis. In this sense, hybrid zones truly are evolutionary laboratories (Hewitt 1988).

In principle, selection can act in two different ways. First, hybrids may be less fit because they harbor an incompatible mix of genomes. Such "tension zones" (Key 1968) are habitat independent and thus in principle not tied to any particular location. Alternatively, different types may be favored on either side of an environmental gradient, which therefore determines the position of the hybrid zone. These two scenarios are not mutually exclusive. For example, range expansion after differential adaptation in allopatry may give rise to a hybrid zone in which the two different habitats adjoin. Randomly accumulated genetic incompatibilities may then contribute to the maintenance of this hybrid zone. Both types of selection are likely to play a role in the hybrid zone between the fire-bellied toad, *Bombina bombina,* and the yellow-bellied toad, *Bombina variegata.* In this paper, we identify particular traits that affect fitness in this system and study the effect of selection on the pattern of phenotypes and genotypes.

These two European toads differ profoundly in many characters. At the molecular level, the two taxa differ in allozymes (Szymura 1983, 1988, 1993), mitochondrial DNA (mtDNA) (Szymura et al. 1985), and albumin (Maxson and Szymura 1984). Nei's genetic distance between *B. bombina* and the four different subgroups of *B. variegata* ranges from 0.37 to 0.59 (Szymura 1988, 1993). The mtDNA sequence has diverged by 5.6% to 7.0% (Szymura et al. 1985, unpubl. data). Assuming a molecular clock, these data place the split between the two taxa into the pre-Pleistocene, 2–7 mya (Szymura 1988, 1993). One would expect a concomitant buildup of reproductive isolation even under a uniform selection regime. In fact, hybrid offspring suffered increased embryonic mortality and showed more frequent developmental abnormalities in a transect near Cracow in Poland (Koteja 1984; Szymura and Barton 1986).

A long list of ecological differences separates the two taxa (Szymura 1993). *Bombina variegata* lives in upland terrain where it reproduces in temporary puddles. In contrast, *B. bombina* inhabits the plains and breeds in more permanent water. Consequently, the hybrid zone tends to follow the edge of mountain ranges (Arntzen 1978; Szymura 1988). The different life-styles are reflected by several specific adaptations. In comparison to *B. bombina, B. variegata* lays a smaller number of larger eggs, which give rise to relatively earlier metamorphosing offspring (Rafińska 1991). In addition, *B. variegata* has a thicker skin (Czopkowa and Czopek 1955), which presumably lowers the risk of desiccation in keeping with its more terrestrial habit. *Bombina variegata* also possesses relatively longer leg bones (Michałowksi 1961) and a generally stronger skeleton, which should facilitate dispersal over land between temporary breeding sites. The molecular and phenotypic data taken together suggest that both selection

against hybrids and selection by habitat play a role in the maintenance of this hybrid zone.

This study concentrates on four of these presumptive adaptations: egg size, developmental time to metamorphosis, skin thickness, and skeletal proportions. We also include in our analysis the very different male mating calls (Lörcher 1969; Sanderson et al. 1992) and the striking divergence in belly coloration. The approach has been indirect: rather than studying natural selection in the field, we infer it from the comparison of the widths and relative locations of the clines formed by the six quantitative traits. Traits that deviate from others in these parameters must be determined by a different balance of evolutionary forces in the hybrid zone (Butlin et al. 1991). More generally, we use this data set to probe our understanding of genotypic and phenotypic clines based on existing theory and identify questions that call for further theoretical development. In the following, we outline the arguments that underlie the analysis.

Many hybrid zones, including that in *Bombina,* can be explained by a model in which cline width reflects the balance between dispersal and selection (Haldane 1948; Bazykin 1969; Slatkin 1973). Stronger selection narrows the cline, whereas increased dispersal blurs the otherwise sharp transition from one type to the other. The robust relationship among cline width, dispersal, and selection suggests that strong selection must be acting to produce the observed clines in diagnostic enzyme markers in *Bombina* ($s = 0.22$ in two Polish transects; Szymura and Barton 1991); yet, it is difficult to believe that these markers are more than weakly selected. The apparent contradiction is resolved by the linkage disequilibria, which are generated by the influx of pure genomes into the hybrid zone. They represent statistical associations and can thus be present among loci that are not physically linked. Consequently, some of the selection experienced by a given locus is due to other loci with which it is associated.

The strength of this "effective selection" on a given locus depends on the relative rates of selection and recombination. The dynamics are best understood in the case of selection acting purely against heterozygotes (Barton 1983). If selection dominates, then hybrid genotypes are removed from the gene pool before they can be broken up by recombination. Consequenctly, linkage disequilibria are reinforced, and the effective selection per locus remains high. A population in the center of the zone would be expected to consist mostly of parentals and F_1s, with only a few F_2 and backcross individuals. In contrast, if recombination is the relatively faster process, it randomizes the genetic background of a given locus and thus weakens the effective selection acting on it. The analysis shows, however, that even in this case of "weak coupling" a given locus is still affected by selection on other loci throughout the genome. Furthermore, simulations (Baird 1995) have shown that it may take thousands of generations before the analytic equilibrium is reached. Relatively stronger disequilibrium is expected in younger hybrid zones. The *Bombina* hybrid zone fits the weak coupling case: there is a wide range of recombinants, and yet significant linkage disequilibria among allozyme loci are found in the center of the zone ($D_{max} = 0.139$, MacCallum 1994).

Interactions among selected loci due to linkage disequilibrium are reflected in the individual clines (Barton 1983). They should become steeper and also more concordant, that is, equal in both width and position. Moreover, sharp steps in allele frequency may arise in the center (Barton 1979). Similarly, neutral loci, that are interspersed with selected ones in the genome, are expected to form concordant clines due to their shared selective background (Barton 1986). These predictions extend to quantitative traits. Assume two traits, each of which is determined by a separate set of additive loci. If these loci are no more than weakly selected, then the mean phenotypes should form concordant clines as well.

Thus, linkage disequilibria complicate the search for key traits in the hybrid zone, because they impose concordance on characters that are subject to different degrees of direct selection. We focus our attention therefore on those traits that deviate from a common pattern. Deviations arise in several ways. Relatively strong selection on a particular trait should cause its cline to become steeper relative to others. In a habitat-dependent hybrid zone, the selection coefficients may reverse in different places for different traits. A sufficiently large spatial offset should shift a given cline away from the rest. Nonadditive inheritance may also cause shifted clines. In this case, the underlying changes in allele frequency are concordant, but the phenotype is a nonlinear function of these frequencies and therefore deviates. The latter two cases will in general not be easy to distinguish, but a very large shift is more likely due to an environmental offset than to nonadditivity.

Linkage disequilibria among loci that affect different traits also give rise to phenotypic covariances in hybrid populations. Although pleiotropy has the same effect, it is unlikely to produce consistent covariances among a diverse collection of traits. Phenotypic covariances in the center of the zone should thus yield a measure of average linkage disequilibrium, which not only provides insight into the degree of coupling within hybrid genomes but can also yield independent estimates of the average dispersal distance. We present a formal derivation of this relationship below.

This paper has two aims. First, we set out to understand the selection pressures that maintain the *Bombina* hybrid zone. The analysis of cline width and position separates a set of four concordant clines from the remaining two which form unique patterns. Habitat dependent selection is likely to cause one of these. This approach is complemented by a study of offspring survival in pure and laboratory crosses to assess the importance of genetic incompatibilities. Second, we study the covariances among characters in hybrid versus pure populations and discuss how linkage disequilibria can be estimated from them. We aim to show how quantitative traits can be analyzed in a similar way to Mendelian markers, thus developing methods that can be applied to other hybrid zones.

Materials and Methods

During the summers of 1991 and 1992, we collected 443 adults from 20 different sites across a transect located 15 km southeast of Zagreb near the village of Pešćenica in Croatia (fig. 1, table 1). After their belly patterns had been photographed for individual identification, the toads were kept at ambient temperature in Edinburgh in two 2× 2 m enclosures and were fed twice a week on *Calliphora* maggots and oc-

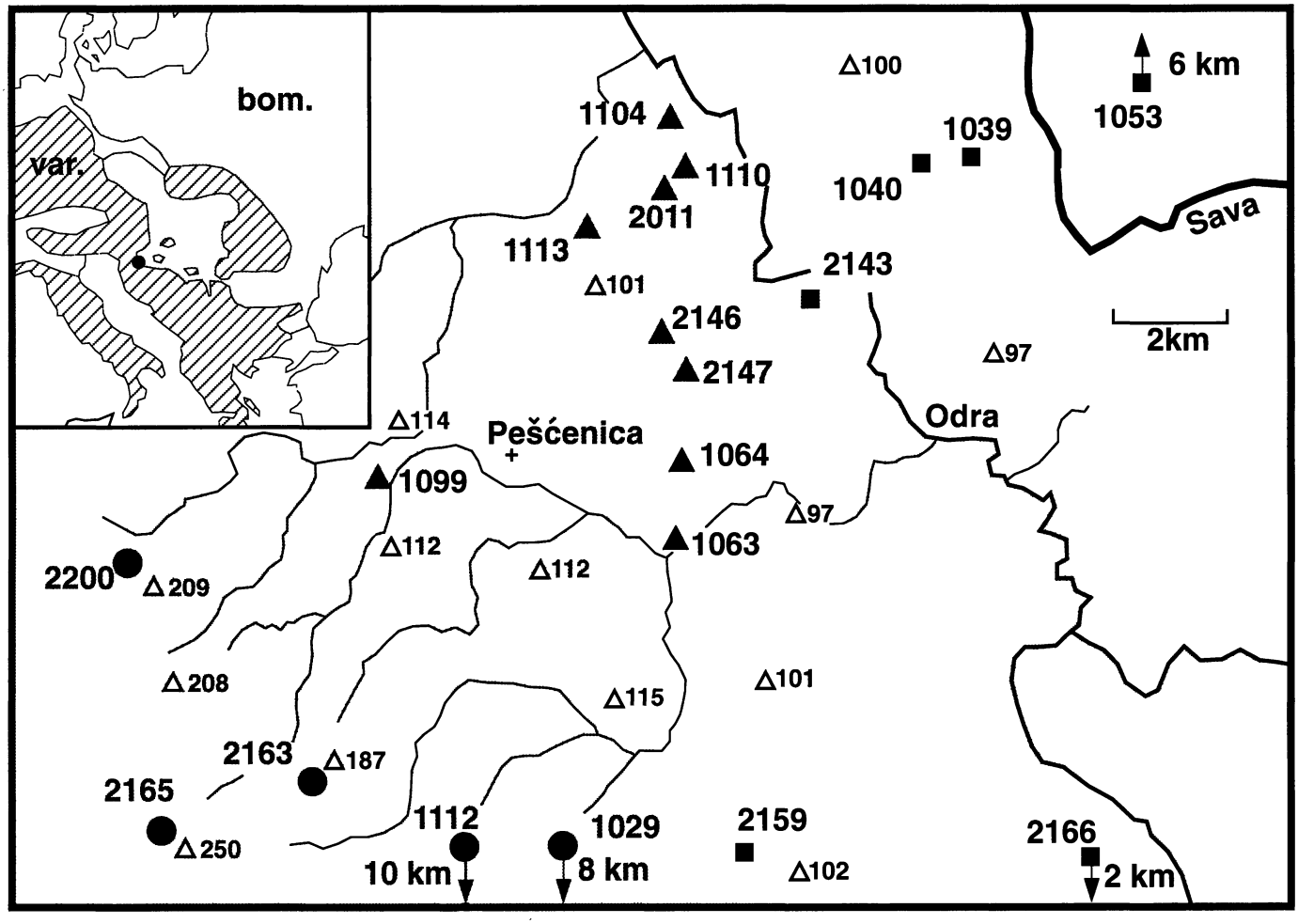

FIG. 1. Map of the Pešćenica transect with sampling locations. Different symbols are used for the three categories: ● *variegata,* ▲ hybrid, and ■ *bombina.* The mean frequency of *variegata* alleles and estimates of F_{IS} and D are given in table 1. The numbers next to small open triangles give elevations in meters. Shaded areas represent forest. The inset (redrawn after Szymura and Barton 1986) shows the distribution of the two taxa in central and Eastern Europe; Pešćenica is indicated by a filled circle.

casionally on crickets (*Achaeta domestica*). They have since been returned to the exact collection sites.

All animals were scored for allozyme variation at five loci: *Ak, Mdh-1, Ldh-1, Idh-1,* and *Gpi.* Single clipped toes were homogenized, and enzyme variants were separated in horizontal starch gels. Detailed protocols have been published by Szymura (1976a,b, 1983) and MacCallum (1994). Staining was performed using agar overlays (Shaw and Prasad 1970). The genotypes of a much larger number of toads from more than 80 locations have been studied by MacCallum (1994); the results will be published elsewhere.

The sampling locations were classified based on the average frequency of *variegata* alleles. A value of 0.2 or less defines a "pure" *bombina* population, whereas "pure" *variegata* sites have means of at least 0.8. All sites in between are referred to as hybrid (table 1). Figure 1 shows that all *variegata* sites are located on an upland ridge at the southwestern corner of the transect. They are all nonpermanent and sparsely vegetated bodies of water (mostly water-filled wheel ruts). As expected, *bombina* populations were found in shallow ponds (one exception being an overgrown wheel rut) in the lowland of the Odra and Sava rivers. The hybrid sites also lie at low elevations and south of the Odra, but are more similar to the *variegata* sites in habitat (for details, see MacCallum 1994).

The belly pattern of *Bombina bombina* consists of relatively small and unconnected red spots, whereas *Bombina variegata* has large and contiguous areas of yellow color. This variation was quantified by counting the number of connections between particular areas on the toads' ventral surface that are consistently colored in both taxa. This scheme due to J. M. Szymura (Szymura and Barton 1991) gives a score from 0 to 10. *B. bombina* is found at the lower end and *B. variegata* at the upper end of this scale.

The thickness of the epidermis ("skin thickness") was measured from 20μm thick hematoxylin-stained cross sections of clipped toes. A camera lucida was used to superimpose the image of a digitizing pad onto the field of view of a compound microscope (Leitz LM-LUX, 40× magnification). Ten transects across the epidermis were measured and averaged per animal.

To obtain skeletal measurements, toads were anaesthetized

TABLE 1. Enzyme genotypes across the transect. The number of individuals per sample (N) is followed by the frequencies of *variegata* alleles at the four diagnostic loci, and their average ($\bar{p}$). F_{is} is the most likely estimate of the heterozygote deficit (assumed the same across loci). D (mle) gives the most likely estimate of pairwise linkage disequilibrium, taking into account the observed heterozygote deficit, F_{is}. This is made on the assumption that the standardized disequilibrium ($R = D_{ij}/\sqrt{p_i q_i p_j q_j}$) is constant across pairs; D (mle) = $R\bar{p}\bar{q}$. Likelihood-ratio tests show no significant heterogeneity in F_{is} or R across loci. D (cov) is calculated in the same way as for quantitative traits, by substituting the average covariance between loci into equation (A3). Averages for *bombina,* hybrid and *variegata* populations are weighted by sample size (N).

Site	N	Allele frequencies				$\bar{p}$	F_{is}	D (mle)	D (covariance)
		AK	MDH	LDH	IDH				
1053	17	0.059	0.000	0.059	0.030	0.037	0.000	0.005	0.003
1039	60	0.067	0.008	0.125	0.043	0.061	0.127	0.000	0.058
1040	40	0.075	0.012	0.125	0.057	0.067	0.222	0.004	0.066
2166	42	0.083	0.012	0.107	0.073	0.069	0.289	0.004	0.068
2143	36	0.181	0.069	0.154	0.125	0.132	0.192	0.046	0.106
2159	14	0.143	0.107	0.179	0.107	0.134	0.158	0.034	0.134
bombina	209						0.181	0.012	0.012
1104	13	0.231	0.154	0.269	0.154	0.202	0.578	0.100	0.219
1064	33	0.303	0.167	0.242	0.297	0.252	0.264	0.061	0.105
1063	37	0.392	0.176	0.324	0.311	0.301	0.182	0.055	0.066
2011	21	0.286	0.500	0.381	0.357	0.381	0.325	0.134	0.216
2147	10	0.300	0.450	0.600	0.444	0.449	0.511	0.121	0.222
1110	13	0.462	0.462	0.731	0.654	0.577	0.250	0.109	0.157
1113	6	0.833	0.500	0.583	0.500	0.609	0.000	0.099	0.069
1099	33	0.742	0.697	0.833	0.894	0.792	0.122	0.029	0.047
2146	6	0.667	0.917	0.917	0.667	0.792	0.130	−0.012	−0.025
hybrids	172						0.287	0.071	0.113
1112	6	0.833	0.583	1.000	1.000	0.854	0.322	0.064	0.039
2163	7	0.643	0.929	1.000	1.000	0.893	0.000	−0.019	−0.099
2165	18	0.833	0.944	0.944	1.000	0.931	0.000	−0.003	−0.008
2200	10	0.900	0.950	1.000	0.889	0.935	0.779	−0.005	−0.013
1029	37	0.960	0.973	1.000	1.000	0.983	0.376	−0.000	−0.001
variegata	78						0.178	0.002	−0.002

by submersion in 2% MS222 (3-aminobenzoic acid ethyl ester, Sigma), laid flat on their back, and X-rayed. Figure 2 shows the 12 measurements that were digitized from the X-ray photographs. Small bones (e.g., those in the arm) could not be measured reliably because their images were not sufficiently focused to determine the endpoints. Separate right and left measurements were taken in bilaterally symmetric bones for an analysis of fluctuating asymmetries. No allometry was found, as the plots of ln(bone length) against ln(body size) were linear with a slope of one. To remove correlations with body size, the measurements were divided by the sphenethmoid-ischium distance (points 13 and 24 in fig. 2). All measures were then log transformed. The overall difference in skeletal proportions was expressed through a discriminant function that was computed on the two pure types. This linear combination of the individual measurements maximizes the ratio of between-group to within-group variance (Dillon and Goldstein 1984).

Egg size and developmental time were measured in a large-scale breeding experiment, which also provided data on the survival of hybrid versus pure offspring. Table 2 lists the sample sizes in terms of the number of families per trait and type of cross. Two wk prior to the onset of the experiment (August 1992), the adults were brought into the laboratory (22.5°C, 14:10 h light/dark cycle), sorted into pairs, and fed every other day on crickets. Animals were mated randomly within populations. The crosses were categorized as *bombina,* hybrid, or *variegata* based on the population from which the parents originated. An exception to this rule was 15 pairs for which males and females were taken from opposite ends of the hybrid zone. They will be referred to as "F_1" crosses below. Note that males and females of some hybrid crosses differed greatly in their enzyme score and may thus have produced essentially F_1 offspring. Four genetic markers, however, can provide only a rough idea of the F or backcross generation to which a toad belongs. We therefore do not differentiate among different hybrid crosses in this analysis. To induce reproduction, the toads were injected with 250 I.U. of human chorionic gonadotropin (Sigma). Pairs were set up in shallow water in 30× 30 cm plastic containers with a clay slate and some nylon string to which females could attach their eggs.

Freshly laid egg clutches were placed in 500 ml of 10% standard amphibian saline. The diameter of five eggs per pair was measured to the nearest 0.02 mm through a dissecting microscope fitted with a graticule. These values were averaged and converted into an estimate of egg volume under the assumption that the eggs are spherical. No measurements were taken past the blastula stage. Close monitoring of 11 families revealed an increase in mean egg volume from one-cell stage to blastula of 0.12 mm^3, which corresponds to 2.5% of the range in family means. Because the total amount and the rate of change varied among families, no correction was made for this error.

In the majority of crosses, 13 hatchlings were reared individually in polystyrene tumblers (maximum $N = 25$ in some F_1 crosses). The members of a family were distributed over 13 blocks to account for temperature variation within the laboratory. The average temperature was 22.5°C and varied between 24° and 21°C from the top shelf to the bottom shelf. The bottom of each tumbler had been replaced with a taut

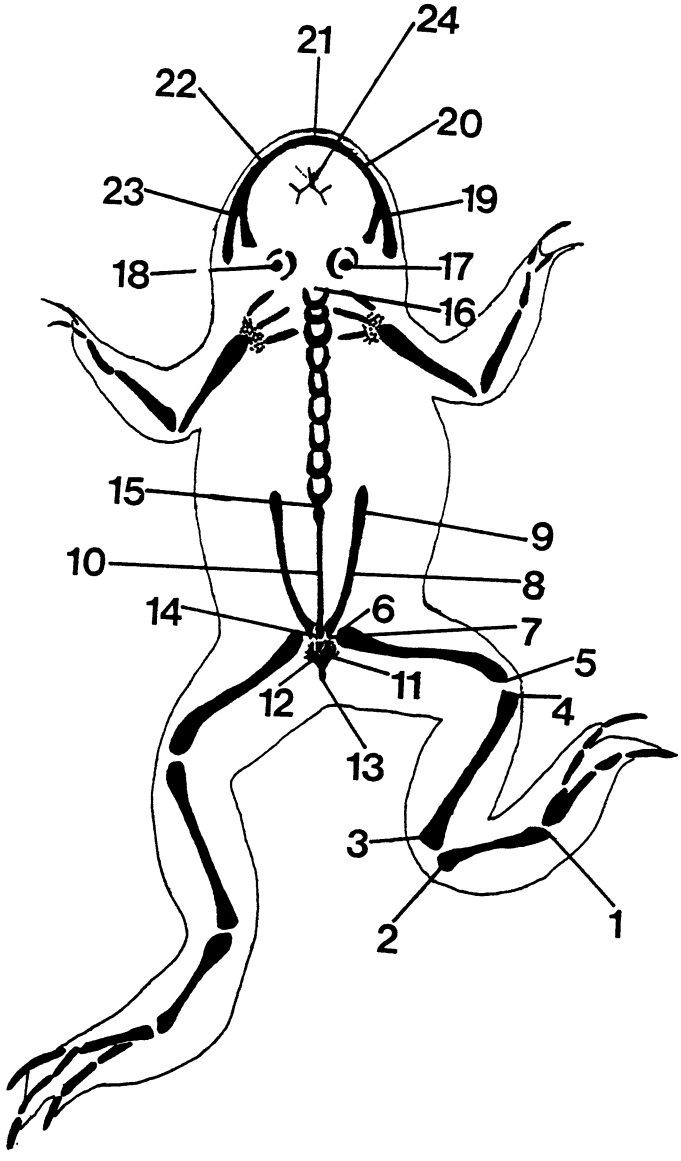

FIG. 2. Tracing of a toad X-ray with key to the skeletal measurements. Bone lengths and characteristic distances were measured between marked points as follows: 1–2, calcaneus; 3–4, tibiofibula; 5–6, femur; 11–12, pubis; 13–24, ischium-sphenethmoid; 14–15, urostyle; 15–16, vertebral column (unfused); 17–18, otic distance; and 21–24, sphenethmoid-snout. A seventh-order polynomial was fitted to a series of points along the ilia (7, 8, and 9 represent a subset). The ilium angle was measured as the tangent to this curve at 8. The ilium distance is the distance between 8 and the urostyle. The maxilla curvature is based on a fifth-order polynomial (19 to 23) and defined as the radius at 21.

nylon mesh. Tumblers were placed in groups of nine inside plastic boxes, and the boxes were filled with water such that each tumbler contained 250 ml (box capacity: 6 liters). With this setup, water could be removed each week with a hose connected to a vacuum pump and replaced from a reservoir of dechlorinated tap water without any handling of the tadpoles. A pilot study had shown that relatively more tadpoles died on the days following complete water changes. Therefore, only 80% of the water was replaced each week in this study. This procedure did not affect tadpole survival. This design excludes competition by direct interaction. Because no more than one member of a family was placed in a given box, it also controls for possible kinship effects on tadpole development (Jasieński 1988). We note, however, that growth inhibitors (e.g., Steinwascher 1978), if present, could have diffused among tumblers.

The tadpoles were fed *ad libitum* on boiled nettle leaves (*Urtica dioica*). Uneaten food was removed daily to maintain a uniform water quality. We measured developmental time as the number of days from oviposition to metamorphosis (Gosner stage 42: both arms extended, Gosner 1960). Weight, snout-vent length, and tail length at metamorphosis were recorded. These size measurements will be published elsewhere. The toadlets were reared individually in their tumblers for the following three months. The boxes were screened in, the water level was lowered, and a piece of foam, wedged into the tumbler, served as an island. Daily rations consisted of 5 to 10 anaesthetized *Drosophila* flies dusted with a vitamin supplement (Vionate™).

Mating calls were recorded after males had been injected with 250–375 I.U. of gonadotropin. As is typical for anurans (Duellman and Trueb 1986), call characteristics in *Bombina* change as a function of temperature (Lörcher 1969; Sanderson et al. 1992). We therefore controlled the water temperature in the recording boxes with aquarium heaters to an average of 23.5°C (range: 23°C–24°C, one recording at 25°C). This range is contained within the upper end of the temperature distribution in shallow ponds and puddles around Pešćenica in May and June (MacCallum 1994). Since *Bombina* males alter their call characteristics in response to other males in a chorus (Gilchrist unpubl. data), the data presented here are based on solo recordings only. Recordings were made with a Sony Walkman WM-D6C Professional™ and a Shure Prologue16L-LC™ microphone. The calls were analyzed with Signalyze™ (Keller 1992) on an Apple Macintosh computer.

Bombina calls consist of relatively short pulses that are separated by longer silent intervals. This pattern can be repeated for long periods (several minutes to almost an hour, Lörcher 1969). The cycle length is defined as the time elapsed from the beginning of one pulse to the onset of the next (Sanderson et al. 1992). The variance in cycle length within a recording is often very high, especially in *B. bombina.* No clear distinction exists between a very long call interval and a temporary interruption of the call sequence. This may at least in part be a laboratory artifact. We therefore quantified cycle length as the average of the four shortest cycles within a recording of standard length (typically 20 s, 30 s for calls with very long cycle length). Pulse duration and fundamental frequency were measured on the corresponding four pulses. Regressions of the call parameters on temperature, male weight and ischium-sphenethmoid distance within *B. bombina* and within *B. variegata* were found to be nonsignificant. Cycle length measures were logtransformed to correct for nonnormality of the data.

RESULTS

Offspring Survival

We divide the survival data into three stages: hatching success, tadpole survival, and the survival of newly meta-

TABLE 2. Samples sizes of the breeding experiment. The family sizes per cross type, the number of families from which data was collected per trait and cross type, and the total number of metamorphs are listed.

	bombina	Hybrid	*variegata*	F_1	Total
No. of tadpoles reared per family	13	13	13	13–25	
No. of families					
Egg size	42	43	30	14	129
Hatching success	40	41	30	14	126
Tadpole survival	37	41	16	13	107
Developmental time*/survival of recent metamorphs	30	33	16	9	88
Total no. of metamorphs	296	369	159	117	941

* Mean developmental time per family was estimated from families with more than two metamorphs.

morphosed toadlets. The average proportion of tadpoles hatching per family is very similar in the four types of crosses: 0.45 in F_1, 0.46 in *bombina,* 0.49 in hybrids, and 0.52 in *variegata* crosses. There are no significant differences among these. This comparatively low hatching success is in part due to the hormone injection, which prompts females to lay immature as well as mature eggs (J. M. Szymura pers. comm. 1994). Although we do not find evidence for a habitat-independent disadvantage of hybrid embryos, such an effect would have to be strong to be detected with these data.

The average proportion of hatchlings that survive to metamorphosis shows the same rank order as before: 0.54 in F_1, 0.66 in *bombina,* 0.74 in hybrids, and 0.87 in *variegata* crosses. The distributions are quite dissimilar (fig. 3). All *variegata* families are clustered around the mean, but in contrast, both the *bombina* and the hybrid histograms show considerable spread, suggesting heterogeneity among families in tadpole survival. Dividing (arbitrarily) the families into three classes (<20% survival, 20%–60%, >60%), the F_1 distribution differs significantly from those of *bombina* and *variegata,* whereas the comparison with hybrids is insignificant (G_2 = 17.73, 9.34, 1.62, respectively). A distinctly bimodal pattern is found in the F_1 crosses: in 4 of 13 families, no tadpoles reached metamorphosis, whereas more than 60% transformed in 8 others. It is hard to test the significance of this pattern. However, no families fell in the interval 0.385–0.643; this would occur with a probability of only 4.5% if the probability density increased linearly from low to high survival, with the observed mean of 0.54. Note that the survival per family is uncorrelated with the direction of the cross (fig. 3D).

During the first 3 mo past metamorphosis, 57% of all deaths occurred in the first 6 d past metamorphic climax; that is, while the tail is resorbed and the animal assumes its toadlike body proportions. The average proportion of deaths per family during this critical time interval does not differ among types. The mean proportion surviving is 0.82 in *variegata,* 0.86 in hybrids, 0.90 in *bombina,* and 0.98 in F_1 crosses. There was one *variegata* family in which all 12 metamorphs died within the first 6 d and one hybrid family in which 6 out of 12 toadlets died. For all other families, the proportion surviving was greater than 0.5. Except for the two outliers, there were no differences among the distributions.

Differences in Quantitative Traits between the Taxa

Table 3 lists the means and standard deviations for the eight measured variables. They are computed separately for *bombina,* hybrid, and *variegata* populations. As expected, animals from *variegata* populations have the most connected belly pattern, which is reflected in their high spot score. Compared with *bombina,* they also have a significantly thicker epidermis. The estimates, however, are not as disparate as those reported by Czopkowa and Czopek (1955, 22.8 vs. 65.2μm). Their use of dorsal skin rather than tissue from toes probably explains at least part of this discrepancy.

Egg volume differs more than two fold between *bombina* and *variegata* populations. The time from egg to metamorphosis is on average 1 wk shorter in *variegata* than in *bombina.* However, in contrast to all other measures for which we found significant differences between the two taxa, developmental time in hybrids does not take on an intermediate value. Instead, it essentially coincides with that of *variegata.* The analysis below resolves this pattern on a finer scale and suggests a possible cause.

The two taxa show pronounced differences in skeletal proportions. Highly significant differences are found in 9 of the 12 individual bone measures (table 4). The discriminant function based on all 12 measures distinguishes very successfully between *bombina* and *variegata* (eigenvalue = 4.31). The pooled within-group correlations of the individual variables with the discriminant score are the most reliable measure of each variable's contribution to the function (Dillon and Goldstein 1984). As shown in table 4, the tibiofibula and femur lengths are the two dominant measures (r = 0.61 and 0.56, respectively), whereas the third major leg bone, the calcaneus, is less important (r = 0.25). Recall that all bone lengths were standardized for body size. Thus, the main difference in skeletal proportion is due to the *relatively* longer leg in *variegata.* This confirms earlier results by Michałowski (1961). All other measures make only minor contributions to the function.

There was no evidence for an increase in fluctuating asymmetry in hybrid populations. The (left-right) variances in *bombina,* hybrids, and *variegata* were homogeneous for femur, tibiofibula, and calcaneus length (Scheffé-Box test: P > 0.25, 0.12, and 0.11, respectively). The (left-right) variances for the ilium-urostyle distance were heterogeneous (P < 0.01), but the maximum was observed in pure *bombina.*

Among the three call parameters, only the cycle length differs significantly between the two taxa: the mean for *bombina* is 2.14 s whereas the mean for *variegata* is 0.64 s. These values are comparable to those reported by Lörcher (1969) and Sanderson et al. (1992). In contrast with the two earlier studies, there is no cline in either pulse duration or fundamental frequency across the Pešćenica transect. In all three

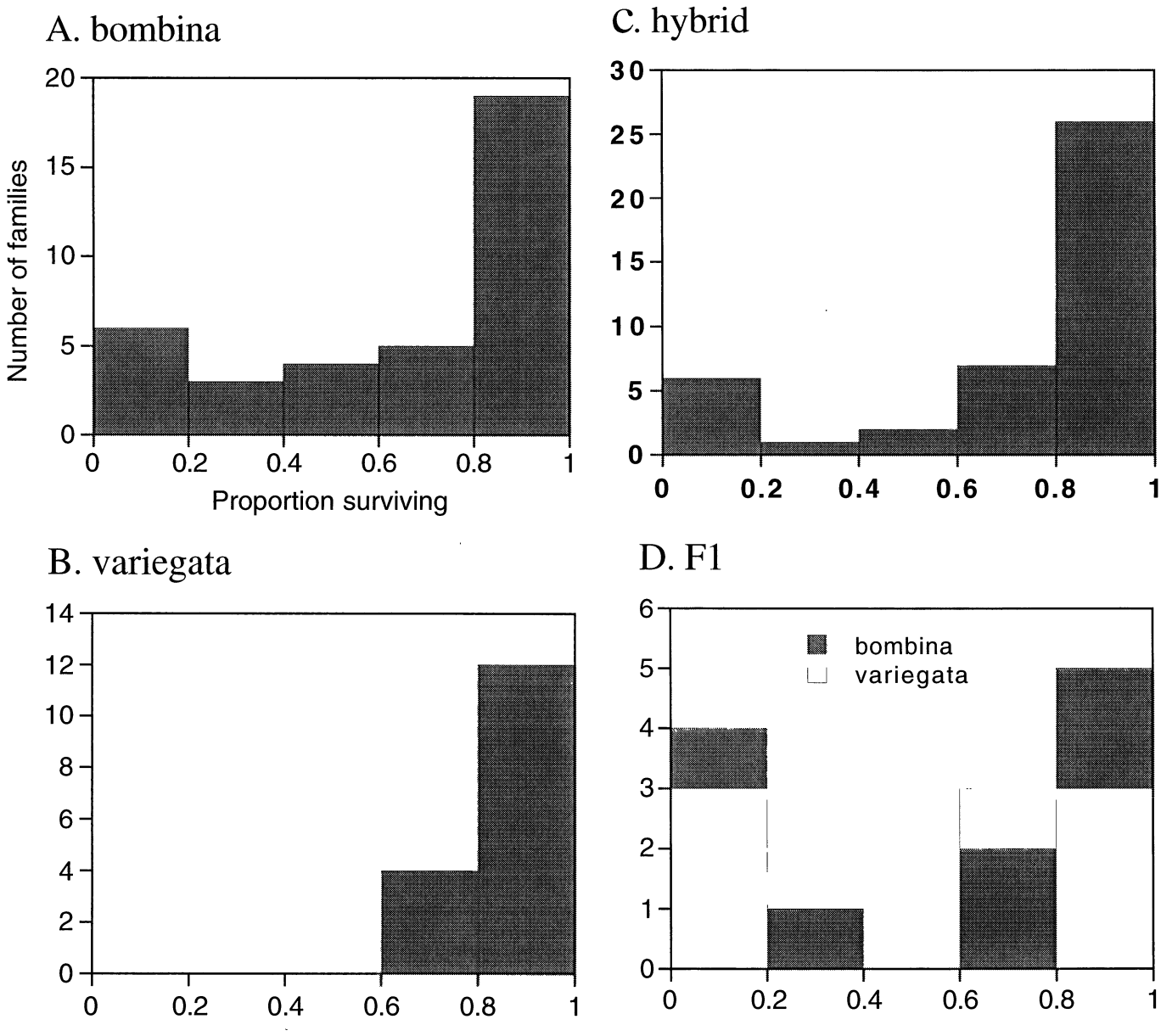

FIG. 3. Distributions of tadpole survival per family in the four cross types. In panel D, the height of each column represents the total number of F_1 families in each survival class. Different shading indicates the direction of the crosses. The inset legend identifies the maternal parent.

TABLE 3. Means, standard deviations, and sample sizes for the eight quantitative traits. Data are presented separately for *bombina*, hybrid, and *variegata* populations. The last column gives the results of *t*-tests comparing the *bombina* and *variegata* means. Nonnormally distributed traits, marked by an asterisk, were tested with the Mann-Whitney *U*-test.

	bombina			Hybrid			*variegata*			*b* vs. *v*
	Mean	SD	*N*	Mean	SD	*N*	Mean	SD	*N*	*P*
Spot score	1.76	1.62	170	3.87	2.80	160	6.98	1.42	113	0.0001*
Skin thickness (μm)	40.22	6.63	130	45.46	7.75	114	48.47	6.83	72	0.0001
Egg volume (mm^3)	1.96	0.48	53	2.44	0.96	57	4.54	0.77	33	0.0001
Developmental time† (d)	53.02	4.81	30	46.14	3.76	33	46.40	3.35	16	0.0001*
Skeletal proportions (discriminant score)	−1.70	1.02	122	−0.20	1.77	115	2.51	1.06	84	0.0001
ln(cycle length [ms])	7.43	0.18	40	6.93	0.40	37	6.45	0.11	27	0.0001
Fundamental frequency (Hz)	595.5	34.7	40	579.0	39.7	37	585.0	60.3	27	0.5090
Pulse duration (ms)	196.61	40.33	40	197.36	40.82	37	214.77	39.18	27	0.1052*

† *N* equals the number of families (cf. table 2).

TABLE 4. Pooled within-group correlations of the individual skeletal measurements with the discriminant function. Also listed are the *F* and *P*-values from the univariate ANOVAs (sample size for *bombina*: 122, and for *variegata*: 84).

Measurement	Correlation coefficient	Univariate *F*	Significance
Tibiofibula	0.6147	327.5	<0.0001
Femur	0.5620	273.7	<0.0001
Ilium distance	0.3342	96.8	<0.0001
Sphenethmoid-ischium	0.2588	58.0	<0.0001
Calcaneus	0.2472	53.0	<0.0001
Ilium angle	−0.2044	36.2	<0.0001
Sphenethmoid-snout	0.1913	31.7	<0.0001
Maxilla curvature	0.1577	21.55	<0.0001
Pubis	0.1470	18.72	<0.0001
Vertebral column	0.0616	3.3	0.0712
Urostyle	−0.0445	1.7	0.1918
Otic	0.0017	0.002	0.9602

data sets, however, cycle length stands out as the most divergent call parameter. We exclude pulse duration and fundamental frequency from the following analysis.

Analysis of Population Means

To compare the clines in quantitative traits, we use the enzyme data as a point of reference. Szymura and Barton (1991) inferred from the degree of introgression in the Polish transects that each of these loci is no more than weakly selected (s = 0.38%). They are also unlinked (Szymura and Farana 1978). Their relatively steep and concordant clines (MacCallum 1994) should therefore be a direct consequence of the linkage disequilibria in hybrid populations. We expect concordance between the mean frequency of *variegata* alleles and the population means of a given quantitative trait except in the following two cases: (1) the trait in question in subject to direct and strong selection, and/or (2) the trait is nonadditively inherited. Concordance results in a linear regression of phenotypic populations means on the mean frequency of *variegata* enzyme alleles.

Figure 4 shows the regressions for all six traits. The spot score and the discriminant score show an almost perfect linear correlation with the hybrid index ($r \approx 0.97$ in both cases). Skin thickness increases nearly linearly at the *bombina* end of the cline. On the *variegata* side, however, the gradient becomes noisier and shallower. Aside from two outliers, the regression for cycle length appears linear. Note, however, that cycle length, spot score, and discriminant score do not change much past the marker frequency of 0.8. More interesting patterns are shown by egg volume and developmental time. Egg volume shows little change up to a mean marker allele frequency of 0.4 and then rises steeply towards the *variegata* side. The picture for developmental time comes as a surprise: the minimum values are found not in pure *variegata* populations but in two hybrid populations on the *variegata* side of the gradient. The populations with the highest frequency of *variegata* alleles have means comparable to populations in the center of the zone.

To test whether the deviations from linearity are significant, quadratic polynomials were fitted to each plot, and *F*-ratios were computed for the quadratic mean squares (Snedecor and Cochran 1980, p. 400). Table 5 shows that only egg size and developmental time deviate significantly from the null hypothesis. The developmental time cline is stable to logarithmic transformation.

Not only do these two traits deviate from the enzymes, they are also offset relative to each other. Developmental time declines significantly at the *bombina* end of the cline, where egg volume is almost constant. Conversely, the dip in developmental time on the *variegata* side is not reflected in egg volume. This implies that faster development in *variegata* is not strictly a consequence of its larger egg size. The F_1 crosses corroborate this conclusion. The mean developmental time in the four F_1 crosses with a *variegata* mother is 47.32, compared with the value for pure *variegata* crosses of 46.40. Similarly, the average across the five F_1 families with a *bombina* mother is 48.66 d, compared with 53.20 d in pure *bombina* crosses. Both differences are not significant, though the *bombina* comparison is suggestive (Mann-Whitney *U*-test, $P = 0.0811$).

Covariances and Estimates of Linkage Disequilibrium

Linkage disequilibria in hybrid populations should cause an increase in the covariance among quantitative traits in the center of the zone. We computed the covariances among all quantitative traits except for developmental time, but including the hybrid index, that is, the number of *variegata* alleles at the four marker loci (table 6, above the diagonal). Developmental time was ignored because it could not be attributed to any particular adult. Pooled within-population estimates were obtained for the three groups (*bombina,* hybrid, and *variegata*) after subtraction of the respective population means. A similar correction for sex differences made no appreciable difference and was therefore not applied to the data in table 6.

Of the 14 trait pairs, 12 show higher covariances in hybrid than in pure populations. Only two exceptions were found, in which *variegata* estimates exceed those for hybrids (skin thickness vs. cycle length and vs. skeletal discriminant function). Traits characteristic of the parental taxa are associated with each other: for example, the relatively shorter cycle length and higher spot score in *variegata* translates into a negative covariance between these traits in hybrid populations. In principle, covariances between quantitative traits could be due to correlated effects of nongenetic, "environmental" fluctuations, or to pleiotropy. Neither seem likely for such disparate traits, and neither would give the consistent difference in covariance between hybrid and "pure" populations. Note that none of the covariances are significant within *bombina* or *variegata* populations.

If our reasoning is correct, one should be able to estimate the strength of linkage disequilibrium from the covariance in hybrid populations. As shown in the Appendix (see also Sanderson et al. 1992), if two traits, z and z', are influenced by the additive effects of separate sets of genes, each in Hardy-Weinberg proportions, then the covariance between them is proportional to the average linkage disequilibrium between diagnostic alleles, D^*, and to the differences in means across the hybrid zone: $\mathrm{cov}(z, z') = \Delta z \Delta z' D^*/2$.

The lower triangle of table 6 shows estimates of D^* com-

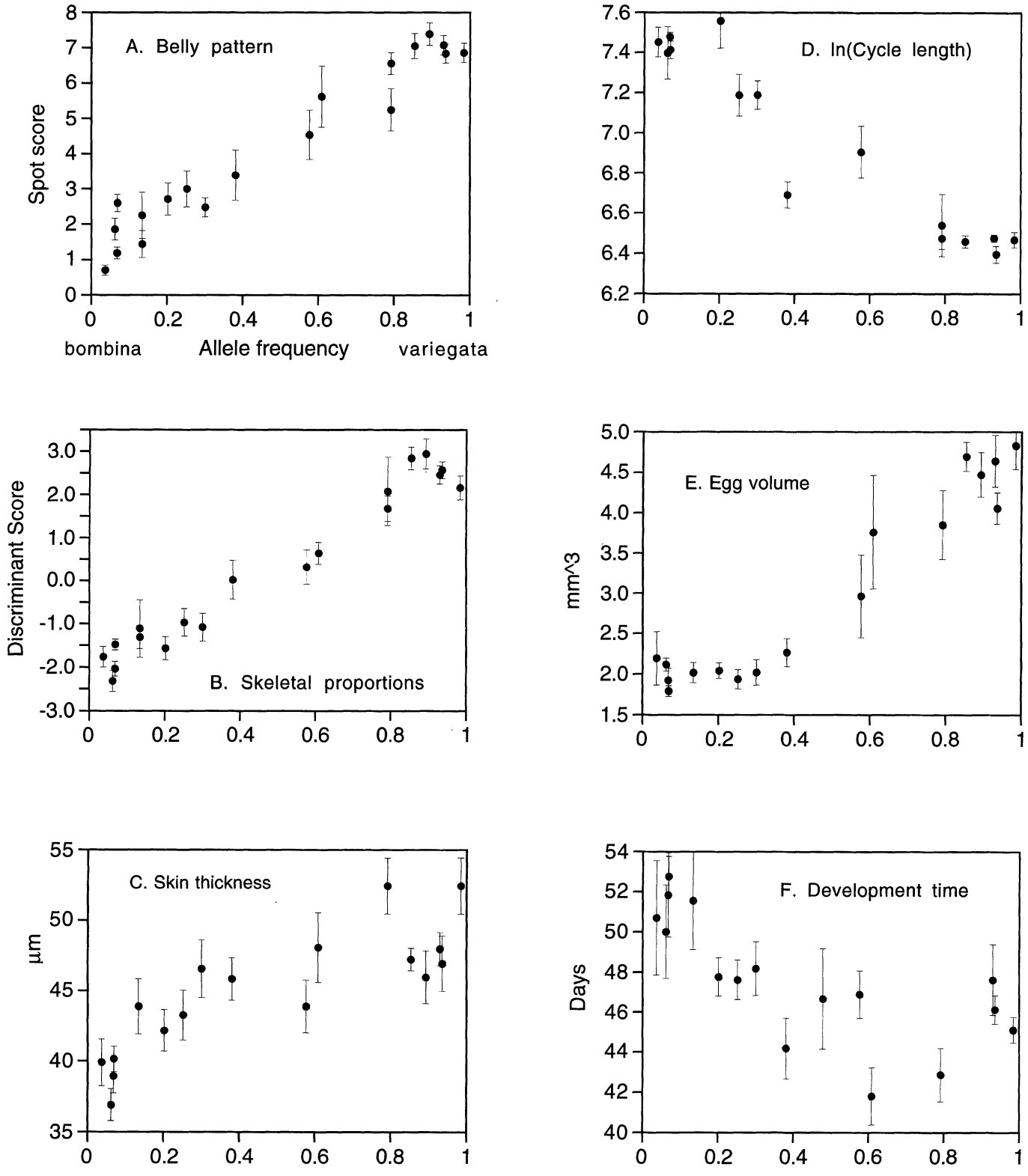

FIG. 4. Plots of population means in the six quantitative traits as a function of the mean frequency of *variegata* alleles at four diagnostic marker loci. Error bars represent one standard error of the mean.

TABLE 5. Test for deviations from linearity. A quadratic polynomial was fitted to each plot. Listed are the results of *F*-tests comparing the mean squares of the quadratic term to the residual variance. NS, not significant.

Trait	df	*F*	*P*
Spot score	1,16	0.0371	NS
Skeletal proportions	1,16	0.1637	NS
Cycle length	1,12	0.9606	NS
Skin thickness	1,14	2.7612	NS
Egg size	1,14	6.6689	<0.025
Developmental time	1,12	12.9269	<0.005

puted with this formula. The Δz values were obtained by extrapolating from the regressions of population means on mean *variegata* allele frequency (see fig. 4) to the extreme allele frequencies (0, 1). The mean D^*-values in the three categories are 0.001 (*bombina*), 0.059 (hybrid), and 0.010 (*variegata*). The maximum-likelihood estimates for the within-population linkage disequilibrium between diagnostic enzyme loci were calculated for the same set of individuals, and gave 0.012 for *bombina,* 0.071 for hybrid, and 0.002 for *variegata* (table 1). Although there is rough agreement between these figures, they are not strictly comparable: the maximum-likelihood estimates have been adjusted downward for the observed significant deficit of heterozygotes (eq. A3). To make a robust comparison, we estimate the linkage disequilibria D^* between allozymes in exactly the same way as for the quantitative traits, from the average pairwise covariances between the four diagnostic loci. This gives $D = 0.113$ for hybrid populations—substantially more than the value of 0.059 estimated by the same method from the quantitative traits. All but 1 of the 14 covariances involving quantitative traits are lower than expected from the covariance between enzymes (table 6).

This discrepancy could exist because alleles at the quantitative trait loci are not strictly additive. In the Appendix, the relation between covariance and linkage disequilibrium is calculated for a polygenic model that is additive across loci, but which allows for dominance. All the allele frequency clines are assumed to be concordant. Provided that the direction of dominance varies across loci, such that the phenotpic clines are concordant with the change in allele frequencies (as is observed for most trait pairs; fig. 4), then equation (A3) is always correct. However, if there is consistent dominance such that the allele frequency and trait clines are discordant, then covariance is greatest where the clines are steepest. This is in fact a more fundamental and robust relation: with *arbitrary* dominance, the covariance is exactly proportional to the product of gradients in the characters ($\mathrm{cov}[z,z'] = [\sigma^2/2r][\partial z/\partial x][\partial z'/\partial x]$, where $[\partial z/\partial x]$ is the gradient in trait mean, and r is the harmonic mean recombination rate [eq. A9]). This is approximately true with arbitrary epistasis (see Appendix).

We can test this prediction with the discordant egg-size cline. Separate estimates of D^* were computed for *bombina*-like hybrid populations (frequency range: 0.2–0.5), where the egg size cline is flat, and for *variegata*-like hybrid populations (range: 0.5–0.8), which show a steep gradient in this trait (table 7, cf. fig. 4E). Although the average estimate of D* is remarkably consistent across both categories for all other traits, egg volume stands out with a very low value in the *bombina*-like sites and a very high value at the *variegata* end of the range. In fact, these two estimates represent the extremes in the table. This result supports the argument that the covariances between traits are due to linkage disequilibria generated by mixing across the cline. However, because most trait clines are concordant with the allozyme clines, the covariances should still be given by equation (A3), regardless

TABLE 6. Estimates of covariances and D^*. The individual measurements were adjusted for the respective population mean, and the residuals were lumped within each of the three categories. For each trait combination, the estimates are given in the order: *bombina* (top), hybrids (middle), and *variegata* (bottom). Covariances are listed above the diagonal; variances are on the diagonal (italics). Sample sizes are in parentheses. The values in bold below the diagonal are estimates of D^*. Covariances that differ significantly (Bonferroni-adjusted) from zero are marked by an asterisk. Δz-values indicate the maximum trait difference across the transect (see text).

Δz	Enzyme 8.00	Spot score 6.31	Egg volume 3.09	ln(cycle length) −1.18	Skin thickness 11.10	Discrim. score 5.19
Enzyme	*0.8677* (122)	0.3290 (122)	−0.0398 (46)	−0.0468 (27)	−0.7986 (108)	0.1263 (96)
	4.2369 (109)	2.6226 (109)*	0.8061 (46)*	−0.2022 (26)*	0.7379 (96)	1.6990 (86)*
	0.5081 (70)	−0.0708 (70)	−0.1063 (30)	0.0079 (14)	−0.0559 (57)	−0.0034 (57)
Spot score	**0.0130**	*2.1583* (170)	0.0267 (52)	−0.0037 (37)	0.0868 (130)	0.1403 (122)
	0.1039	*5.5643* (160)	0.5901 (56)*	−0.1791 (33)	2.2279 (114)	1.7414 (115)*
	−0.0028	*2.0029* (113)	−0.0630 (33)	−0.0810 (25)	1.2834 (72)	0.0156 (84)
Egg volume	**−0.0032**	**0.0027**	*0.2015* (52)	—	0.1282 (46)	−0.0083 (47)
	0.0653	**0.0605**	*0.4670* (56)	—	1.9960 (41)	0.2696 (44)
	−0.0086	**−0.0065**	*0.5168* (33)	—	0.4693 (27)	−0.0991 (28)
ln(cycle length)	**0.0099**	**0.0010**	—	*0.0310* (37)	0.0999 (33)	−0.0103 (26)
	0.0429	**0.0481**	—	*0.0432* (33)	−0.0945 (28)	−0.0890 (25)
	−0.0017	**0.0218**	—	*0.0120* (25)	−0.1420 (19)	−0.0187 (22)
Skin thickness	**−0.0180**	**0.0025**	**0.0075**	**−0.0153**	*36.8830* (130)	−0.0360 (105)
	0.0166	**0.0635**	**0.1165**	**0.0145**	*50.3659* (114)	1.3330 (92)
	−0.0013	**0.0366**	**0.0274**	**0.0217**	*41.3663* (72)	1.8113 (62)
Discrim. score	**0.0061**	**0.0086**	**−0.0010**	**0.0034**	**−0.0013**	*0.8580* (122)
	0.0819	**0.1063**	**0.0337**	**0.0295**	**0.0463**	*1.7435* (115)
	−0.0002	**0.0010**	**−0.0124**	**0.0061**	**0.0629**	*1.0538* (84)

TABLE 7. Average D^* estimates per trait and hybrid category. The per-site mean *variegata* allele frequencies in category I lie in the interval [0.2–0.5], whereas category II covers the range [0.5–0.8]. The linkage disequilibria among enzymes are estimated both by maximum likelihood and from the covariance between them, as in table 6. The average heterozygote deficit (F_{is}) is also shown.

Trait	Category I	Category II
Heterozygote deficit (F_{is})	0.306	0.139
Enzyme score (mle)	0.0822	0.0499
Enzyme score (cov)	0.1361	0.0665
Spot score	0.0880	0.0835
Egg volume	0.0323	0.1617
Cycle length	0.0493	0.0485
Skin thickness	0.0500	0.0542
Skeletal proportions	0.0688	0.0528

of the mode of inheritance. The downward bias in our estimates of linkage disequilibrium from phenotypic covariances cannot therefore be explained by nonadditive inheritance.

For five of six traits, the phenotypic variances are increased in hybrids (table 6, diagonal entries). Since this pattern would be expected only in one-third of the cases by chance, the probability of finding it for at least five of six traits is only $(1/3)^6 + 6(1/3)^5(2/3) = 1.78\%$. This increase could be due to greater developmental instability (cf. section on fluctuating asymmetry), increased heterozygosity at individual loci, if few genes are involved (Lande 1981; Sanderson et al. 1992), or it could be caused by linkage disequilibria. From our estimates of D, we can compute the expected increase in variance ($\Delta\text{var}[z] = [\Delta z]^2[1 + F_{is}]D^*/2$; cf. eq. A3). For $D = 0.113$ based on allozymes, the average ratio of predicted to observed values is 1.2. If one uses $D^* = 0.059$ based on the covariance among quantitative traits, this ratio drops to 0.66; that is, the observed variance exceeds expectation by one-third.

DISCUSSION

Our survey of six quantitative traits across the *Bombina* hybrid zone at Pešćenica uses a parallel analysis to those of Mendelian markers (Szymura and Barton 1986, 1991; MacCallum 1994). The comparison of cline width and position reveals a group of four traits whose clines parallel allozyme frequencies. However, the cline in developmental time appears to be displaced towards *bombina,* and the cline in egg size is displaced towards *variegata.* In hybrid populations, phenotypic variance is greater for most traits, and there are significant associations between traits. Both phenomena are likely to be due to linkage disequilibria generated by the mixing of divergent populations. However, the associations are only about half as strong as expected from the linkage disequilibria between enzyme alleles.

Our laboratory breeding experiment showed no evidence that offspring from wild-caught hybrids have reduced viability. This contrasts with Koteja's (1984) findings of a five fold increase in embryonic mortality in hybrid populations near Cracow in Poland. Our data on hatching success rule out such a strong effect in the Pešćenica transect, despite their limited resolution. It appears that different degrees of genetic compatibility exist in different parts of the hybrid zone. This conclusion is supported by the fact that the two transects involve different subgroups of *Bombina variegata* (Nei's genetic identity = 0.85; Szymura 1993).

F_1 crosses, however, show a striking dichotomy in tadpole survival. Some of these families failed completely, whereas others exhibited above average survival. The outcome appears not to depend on the direction of the cross. But the severe inviability of some F_1's might have little effect on the actual maintenance of the hybrid zone, if F_1 matings (or their equivalent) are rare in nature. Incompatible combinations of alleles would be strongly selected against in the first generation and thus be transmitted to subsequent hybrid generations at a much reduced frequency. There are examples of such amelioration of F_1 incompatibility involving chromosome rearrangements in shrews and mice (Searle 1986, 1991) and in the grasshopper *Chorthippus parallelus* (Virdee and Hewitt 1994).

The concordance among the clines of four quantitative traits could in principle be generated by identical selection on each trait. However, it is difficult to imagine how, for example, the belly coloration and the mating call could play equally important roles in the hybrid zone. Linkage disequilibria present a much more plausible explanation. Hence, on the one hand, we cannot resolve the varying degrees of selection on spot score, skeletal proportions, skin thickness, and mating call; on the other hand, the similarity of their clines strengthens our conclusion that egg size and developmental time are indeed subject to a different balance of evolutionary forces. Although genetic drift could produce similarly deviating clines, it is not a likely explanation in this system: the average dispersal distance of the toads is roughly 1 km per generation (MacCallum 1994), which is large relative to the cline width in Pešćenica (4.7 km). Given the observed toad densities, the neighborhood size must be large as well. Sampling events associated with the establishment of the zone should have long disappeared, and current drift can be considered negligible. Indeed, the residual variation in allele frequency between loci within sites and around the overall cline is small ($F_{st} = 0.0068$ and 0.025, respectively; MacCallum 1994).

On the basis of additional biological information, we argue that all of the four concordant traits are likely to be under some direct selection. The belly pattern in *Bombina* is clearly aposematic: the toads exude noxious chemicals from their skin and, when prodded by a predator, arch their back such that further prodding will roll them over and expose the brightly colored underside (the so-called "Unken reflex"). The larger amount of color on the *variegata* belly may have a more startling effect on the predators encountered by this more terrestrial toad.

Our results on the mating call agree with the work of Sanderson et al. (1992) in the Cracow transect, which showed that the cline for cycle length is of the same width and position as that for enzyme markers. Behavioral observations (Lörcher 1969; Seidel 1988; Barandun 1990) suggest that the main function of the call may be to mediate male-male competition rather than form the basis for female choice. As such, it could be important in the maintenance of the hybrid zone: the advance of *variegata* males into *bombina* territory may be hindered given that *variegata* has the much quieter call.

Finally, the thicker skin and relatively longer leg should aid in *variegata*'s more frequent movements over land between temporary breeding sites.

The nonlinearity of the egg size cline disappears from both data sets if one plots egg volume on a logarithmic scale, which is consistent with multiplicative gene action. An offset environmental gradient is the obvious alternative explanation. At present, we cannot distinguish between these two hypotheses of differential selection and nonadditive gene action.

The most interesting pattern is found in the cline of developmental time. Contrary to our expectation, the minimum time required to reach metamorphosis is not found among offspring from pure *variegata* populations but in hybrids on the *variegata* side of the gradient. It is difficult to imagine an underlying genetic system that would map a linear change in allele frequencies into the observed phenotypic gradient. A difference in the shape of the underlying allele frequency cline is the only alternative explanation.

Our field observations suggest a possible reason why the minimum is found among hybrids. All pure *variegata* populations are found in temporary sites ("puddles") on the hillside, whereas most of our *bombina* samples come from permanent lowland ponds. Hybrids are found in an intermediate habitat type, that is, lowland puddles. During two field seasons that were affected by drought, the lowland puddles had dried by the end of May, whereas breeding aggregations of *variegata* in the hills persisted through most of June. This implies that selection for rapid development is strongest in lowland puddles. The genes for fast development in the hybrid zone are most likely derived from *variegata* populations, which would explain why the minimum is found on that side of the gradient. If this reasoning is correct, then only a subset of both *bombina* and *variegata* genes can become established in hybrid sites. We intend to test this hypothesis in the field with genetically marked tadpole cohorts.

That the clines in developmental time and egg size are offset relative to each other comes as a surprise. Rafińska (1991) reared pure *bombina* and *variegata* families in the laboratory and showed that larger eggs in *variegata* gave rise to larger hatchlings that metamorphose earlier. Her results suggested a causal relationship between these two variables. However, because only pure animals were used in Rafińska's study, direct genetic effects on developmental time that are independent of egg size could not be detected. Both the discordance between egg size and developmental time and the similar developmental time of reciprocal F_1 crosses indicate that such direct effects exist. This finding parallels those of Brown (1967) and Petranka et al. (1987) at the interpopulational level.

The hybrid zone described here raises two quantitative questions. What range of selection pressures would be consistent with the similarity in widths between traits that are expected to be under very different strengths of selection? What strength of selection is required to displace the clines in egg size and developmental rate away from the rest of the hybrid zone? A full answer to these questions would require simulation of clines at many loci. In the following, we attempt an approximate answer. On dimensional grounds, the width of clines in quantitative traits should be proportional to $\sigma/(s'V_g)^{1/2}$, where s' is the strength of stabilizing selection towards a spatially varying optimum, z_{opt} $\{W \approx \exp[-s'(z - z_{opt})]^2/2\}$, V_g is the additive genetic variance, and σ is the average dispersal distance (Slatkin 1978).

We incorporate linkage disequilibrium by supposing that overall selection generates a barrier to gene flow, equivalent to a physical barrier to dispersal. At a single locus, the allele frequency u is perturbed by selection on other loci through the addition of a term $(\sigma^2/2r)(\partial u/\partial x)(\partial \log(\bar{W})/\partial x)$ to the diffusion equation, where $\bar{W}$ is the mean fitness, x denotes the location along a linear transect, and r is the harmonic mean recombination rate between the selected loci and the observed marker (Barton 1986). Exactly the same argument applies to an additive polygenic trait, with mean z. If we suppose that stabilizing selection, s', acts to favor an optimum of $z_{opt} = -1/2$ for $x < 0$, and $z_{opt} = +1/2$ for $x > 0$, then the cline in the mean is given by:

$$\frac{\partial z}{\partial t} = 0 = \frac{\sigma^2}{2}\frac{\partial^2 z}{\partial x^2} + \frac{\sigma^2}{2r}\frac{\partial z}{\partial x}\frac{\partial \log(\bar{W})}{\partial x} - s'V_g(z - z_{opt}). \quad (1)$$

The first term represents gene flow by diffusion; the second, the effect of linkage disequilibria with other selected loci (the barrier effect); and the third, direct stabilizing selection on the trait. It is exact in the limit of weak selection, provided that selection is not explicitly dependent on frequency or location (note that fitness in the model is independent of location *except* at $x = 0$, which is not covered by eq. 1). The analogous equation for allele frequencies is accurate even when selection is quite strong (Barton and Gale 1993). To solve equation (1), we must consider V_g to be constant across the zone. This assumption is restrictive, because the genetic variance is expected to increase as a result of linkage disequilibrium and is observed to do so. However, a variable V_g should not alter the outcome qualitatively, and the actual solution should lie between the solution derived by assuming V_g constant at the maximum and at the minimum.

Figure 5 shows solutions to this equation, using for illustration parameter values estimated from the Polish transects (Szymura and Barton 1991). The solid curve in figure 5A shows numerical solutions of equation (1) (see figure legend for details). Strong selection ($s'V_g > 0.02$, say) maintains a cline that is much narrower than the barrier and that is thus not much affected by it. For reference, the dotted line to the right shows the analytical solution in the limiting case of strong selection without a barrier ($w = \sigma[2/s'V_g]^{1/2}$; Slatkin 1978). At the other extreme, when selection is weak, the width is still proportional to $\sigma(2/s'V_g)^{1/2}$, but the slope is reduced by a factor $\bar{W}^{1/r}$, which represents the barrier effect. The left dotted line shows this analytical result in the limiting case of weak selection. Between these two extremes, there is a region in which changes in selection make relatively little difference to the width. For example, the width changes by a factor of 1.64 between $s'V_g = 0.01$ and 0.1, instead of the expected factor of $(0.1/0.01)^{1/2} = 3.16$. Given the inaccuracies in detecting differences in width, direct selection pressures might vary by more than an order of magnitude without producing obvious differences in cline shape.

Clines are more sensitive to displacement than to variation in the strength of selection. Figure 5B shows results from

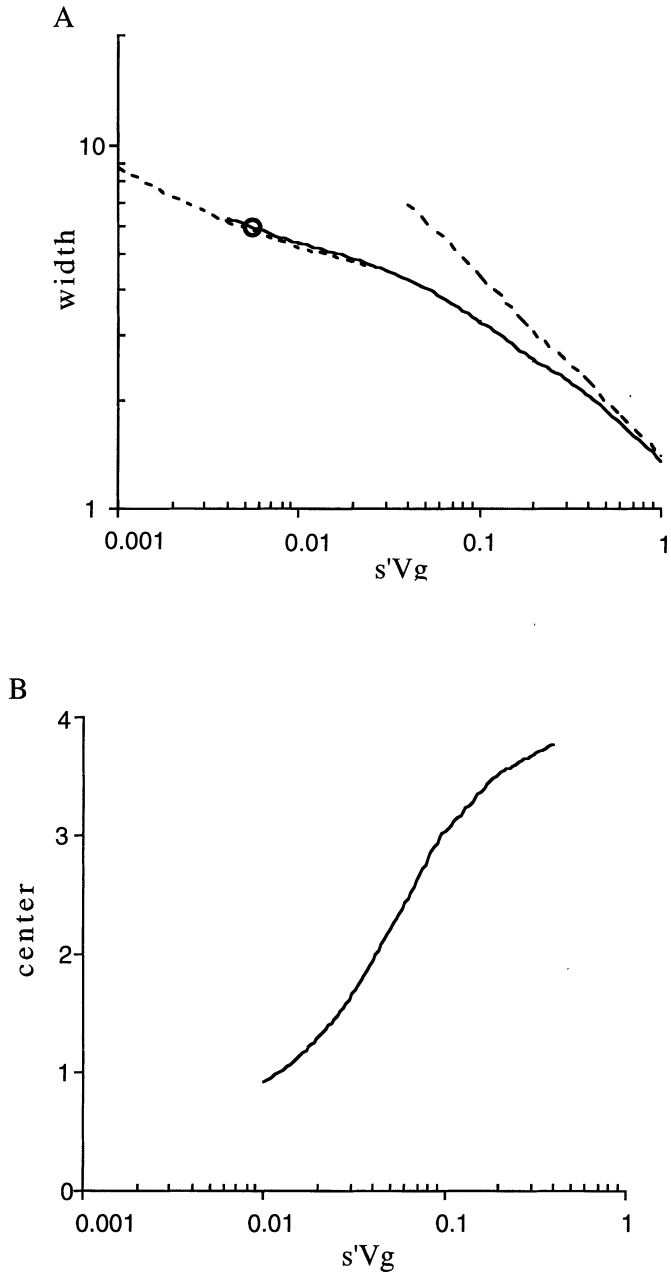

FIG. 5. The effect of a barrier to gene flow on concordance. (A) shows the relation between the width of a cline in a quantitative trait and the strength of the direct selection on it, s'. Selection is scaled relative to the genetic variance (V_g). The solid curve shows numerical solutions to equation (1) where the optimum changes abruptly at the center of the hybrid zone. The dotted curve on the right shows the limiting solution for strong selection [$\sigma\sqrt{2/(s'V_g)}$]. The left dotted curve shows the limit for weak selection, where the hybrid zone appears as a barrier of strength B [$w = \bar{W}_0^{1/r}(B + \sigma\sqrt{2/s'V_g})$, derived by solving equation (1) in the limit of small $s'V_g$)]. The open circle indicates the observed width for allozymes. Panel B shows the relation between the position of the cline in the quantitative trait, and $s'V_g$. Solutions were derived using the Runge Kutta routine NDSolve of Mathematica (Wolfram 1991). Mean fitness decreases to a minimum of 0.58 at the center, and was assumed for simplicity to take the form $\log(\bar{W}) = \log(0.58)/\mathrm{Cosh}(2x/w')^2$, where the underlying clines change over a width $w' = 6$ km. The dispersal distance was $\sigma = 1$ km, and the harmonic mean recombination rate $r = 0.25$. These parameters give a barrier of strength $B = 32.1$ km. This is somewhat less than that estimated in Szymura and Barton 1991, because there, mean fitness was assumed to follow a stepped pattern (Barton 1983).

equation (1) for the case in which direct selection changes sign 4 km away from a hybrid zone of width 6 km. As selection becomes weaker, the cline is pulled towards the center; throughout the range shown in the graph, its width is close to the value $\sigma(2/s'v_g)^{1/2}$ expected with no barrier. A shift of 2 km could be produced by selection of strength $s'V_g = 0.04$; because genetic variance reduces mean fitness under stabilizing selection by a factor $\exp(-s'V_g/2)$, this represents a load of only 2% on the population. Although the arguments here are more accurate than that used in the discussion of Sanderson et al. (1992), they are approximate, and should be checked against simulations in which the genetic variance as well as the mean are allowed to change. Nevertheless, they suggest that the similar widths of the clines in all these quantitative traits are consistent with moderately strong direct selection on the traits; the discordant position of the clines in egg size and developmental time could also be produced by moderately strong direct selection ($s'V_g \approx 0.04$, say).

In hybrid populations, quantitative traits showed consistent associations with each other and with diagnostic allozymes. These are best explained by linkage disequilibria, generated by the mixing of divergent populations. The theory set out in the Appendix shows that the covariance between quantitative traits should be given by $(\sigma^2/2r)(\partial z/\partial x)(\partial z'/\partial x)$, regardless of the mode of inheritance. This gives a robust method for estimating the dispersal rate, σ^2, solely from phenotypic data. However, the data from the Pešćenica transect yield covariances between traits which are about half that expected from the linkage disequilibria between allozymes, even when the clines are concordant. It seems unlikely that there are weaker linkage disequilibria between loci responsible for the quantitative traits than those between allozymes, since the former should, if anything, be under stronger selection and therefore have narrower clines and stronger associations. Sanderson et al.'s (1992) estimates of D both from allozyme loci and from the covariance between cycle length and allozymes agreed closely. Their study, however, was based on a smaller number of individuals from one central population only. We are currently pursuing this issue through computer simulation and by developing molecular markers to identify the loci (or at least, chromosomes) responsible for quantitative variation.

The analysis presented here shows how clines in quantitative traits can give similar information to clines in Mendelian markers. Variation in the position and width of clines can suggest traits that are subject to selection. This approach has been used successfully in *Chorthippus* (Butlin et al. 1991; Virdee and Hewitt 1994); however, it is complicated in *Bombina* by strong linkage disequilibria, which make it harder to disentangle different traits. Associations among traits are themselves valuable, because they give a method for estimating linkage disequilibria from purely phenotypic data. We hope that our work will encourage others to move from analyses that concentrate on Mendelian markers towards an un-

derstanding of geographic variation in the traits responsible for adaptation and reproductive isolation.

Acknowledgments

The project would not have been possible without F. Perović's extensive knowledge of the natural history of the Pešćenica area, and his assistance in the field. Particular thanks are due to the Perović family for their generous hospitality. The Croatian Museum of Natural History and the Croatian Ministry of the Environment were helpful in granting all the necessary permits. J. Szymura assisted with allozyme techniques and in sharing unpublished data from his original survey of the area. M. Davidson and K. Grant prepared the histological specimens, and G. Patterson volunteered time and expertise in X-raying our toads. All members of L. Partridge's lab generously provided us with toad food on a daily basis, in the form of uncountably many spare *Drosophila.* G. Malarky and M. Oli stoically coped with much tedious toad care. We thank W. G. Hill, L. Kruuk, D. Rand, J. Szymura, and an anonymous reviewer for helpful comments on the manuscript. This research was supported by a grant from the Natural Environment Research Council (GR3/8002) to N.B.

Literature cited

Arntzen, J. W. 1978. Some hypotheses on the postglacial migrations of the fire-bellied toad, *Bombina bombina* (L.) and the yellow-bellied toad, *Bombina variegata.* Journal of Biogeography 5: 339–345.

Baird, S. J. E. 1995. The mixing of genotypes in hybrid zones: a simulation study of multilocus clines. Evolution 49(6):1038–1045.

Barandun, J. 1990. Reproduction of yellow-bellied toads *Bombina variegata* in a man-made habitat. Amphibia-Reptilia 11:277–284.

Barton, N. H. 1979. Gene flow past a cline. Heredity 42:333–339.

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

———. 1986. The effects of linkage and density-dependent regulation on gene flow. Heredity 57:415–426.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, Oxford.

Bazykin, A. D. 1969. Hypothetical mechanism of speciation. Evolution 23:685–687.

Brown, H. A. 1967. Embryonic temperature adaptations and genetic compatibility in two allopatric populations of the Spadefoot toad, *Scaphiopus hammondi.* Evolution 21:742–761.

Butlin, R. K., M. G. Ritchie, and G. M. Hewitt. 1991. Comparisons among morphological characters and between localities in the *Chorthippus parallelus* hybrid zone (Orthoptera: Acrididae). Philosophical Transactions of the Royal Society of London, B 334:297–308.

Czopkowa, G., and J. Czopek. 1955. The vascularization of the respiratory surfaces in *Bombina variegata.* Bulletin de l'Académie Polonaise de Sciences Cl. II, 3:313–318.

Dillon, W. R., and M. Goldstein. 1984. Multivariate analysis—methods and applications. Wiley, New York.

Duellmann, W. E., and L. Trueb. 1986. Biology of amphibians. McGraw-Hill, New York.

Gosner, K. L. 1960. A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16:183–190.

Haldane, J. B. S. 1948. The theory of a cline. Journal of Genetics 48:277–284.

Hewitt, G. M. 1988. Hybrid zones—natural laboratories for evolutionary studies. Trends in Ecology and Evolution 3:158–167.

Jasieński, M. 1988. Kinship ecology of competition: Size hierarchies in kin and nonkin laboratory cohorts of tadpoles. Oecologia 77:407–413.

Keller, E. 1992. Signalyze user's manual–Version 2.0. InfoSignal™, Seattle.

Key, K. H. L. 1968. The concept of stasipatric speciation. Systematic Zoology 17:14–22.

Koteja, E. 1984. Embryonic mortality of the fire-bellied toad (*Bombina bombina*), the yellow-bellied toad (*Bombina variegata*), and their hybrids. [In Polish.] Master's Thesis. Jagiellonian University, Cracow, Poland.

Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99:541–553.

Lörcher, K. 1969. Vergleichende bio-akustische Untersuchungen an der Rotund Gelbbauchunke, *Bombina bombina* (L.) and *B.v.variegata* (L.). Oecologia 3:84–124.

MacCallum, C. 1994. Adaptation and habitat preference in a hybrid zone between *Bombina bombina* and *Bombina variegata* in Croatia. Ph.D. diss. University of Edinburgh.

Maxson, L. R., and J. M. Szymura. 1984. Relationships among discoglossid frogs: an albumin perspective. Amphibia-Reptilia 5:245–252.

Michałowksi, J. 1961. Studies on species characteristics in *Bombina variegata* (L.) and *Bombina bombina* (L.). Applying the L: T indicator to the classifying purposes. Acta Zoologica Cracow 6:51–59.

Petranka, J. W., A. Sih, L. B. Kats, and J. R. Holomuzki. 1987. Stream drift, size-specific predation, and the evolution of ovum size in an amphibian. Oecologia 71:624–630.

Rafińska, A. 1991. Reproductive biology of the fire-bellied toads, *Bombina bombina* and *B. variegata* (Anura: Discoglossidae): egg size, clutch size, and larval period length differences. Biological Journal of the Linnean Society 43:197–210.

Sanderson, N., J. M. Szymura, and N. H. Barton. 1992. Variation in mating call across the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata.* Evolution 46:595–607.

Searle, J. B. 1986 Factors responsible for a karyotypic polymorphism in the common shrew, *Sorex araneus.* Proceedings of the Royal Society of London B Biological Sciences 229:227–298.

———. 1991. A hybrid zone comprising staggered chromosomal clines in the house mouse (*Mus musculus domesticus*). Proceedings of the Royal Society of London B Biological Sciences 246:121–132.

Seidel, B. 1988. Die Struktur, Dynamik und Fortpflanzungsbiologie einer Gelbbauchunkenpopulation (*Bombina variegata variegata,* L. 1758, Discoglossidae, Anura, Amphibia) in einem Habitat mit temporären Kleingewässern im Waldviertel (Niederösterreich). Ph.D. diss. University of Vienna.

Shaw, C. R., and R. Prasad. 1970. Starch gel electrophoresis of enzymes—a compilation of recipes. Biochemical Genetics 4: 297–320.

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

———. 1978. Spatial patterns in the distribution of polygenic characters. Journal of Theoretical Biology 70:213–228.

Snedecor, G. W., and W. G. Cochran. 1980. Statistical methods, 7th ed. Iowa State University Press, Ames.

Steinwascher, K. 1978. Interference and exploitation competition among tadpoles of *Rana utricularia.* Ecology 59:1039–1046

Szymura, J. M. 1976a. New data on the hybrid zone between *Bombina bombina* and *Bombina variegata* (Anura, Discoglossidae). Bulletin Academia Polonica Scientatis Class. II 24:355–363.

———. 1976b. Hybridisation between discoglossid toads *Bombina bombina* and *B. variegata* in Southern Poland as revealed by the electrophoretic technique. Zeitschrift for Zoologische Systematik und Evolutionsforschung 14:227–236.

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

———. 1988. Regional differentiation and hybrid zones between the fire-bellied toads *Bombina bombina* (L.) and *B. variegata* (L.) in Europe. [In Polish.] Rozprawy Habilitacyjne No. 147. Cracow: Uniwersytet Jagiellonski.

———. 1993. Analysis of hybrid zones with *Bombina*. Pp. 261–289 *in* R.G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, Oxford.

Szymura, J. M., and N. H. Barton. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata,* near Cracow in Southern Poland. Evolution 40:1141–1159.

———. 1991. The genetics structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata:* Comparison between transects and between loci. Evolution 45:237–261.

Szymura, J. M., and I. Farana. 1978. Inheritance and linkage analysis of five enzyme loci in interspecific hybrids of toadlets, genus *Bombina.* Biochemical Genetics 16:307–319.

Szymura, J. M., C. Spolski, 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.

Virdee, S. R., and G. M. Hewitt. 1994. Clines for hybrid dysfunction in a grasshopper hybrid zone. Evolution 48:392–407.

Wolfram, S. 1991. Mathematica—a system for doing mathematics by computer, 2d ed. Addison-Wesley, Redwood City, Calif.

Corresponding Editor: J. Curtsinger

Appendix: Covariances between Nonadditive Traits

First, consider two quantitative traits, z and z', influenced by the additive effects of separate sets of genes. The numbers of loci per trait are n and n', respectively. The phenotypes, z and z', are given by

$$z = \sum_{i=1}^{n} \alpha_i x_i + E, \tag{A1A}$$

$$z' = \sum_{i=1}^{n} \alpha'_j x'_j + E', \tag{A1B}$$

where α_i is the average effect of, say, a *variegata* allele at the ith locus, x_i indicates the number of *variegata* alleles at that locus (0, 1 or 2), and E and E' are the environmental deviations. Given our assumptions, the covariance between the two traits must be genetic, and in the absence of pleiotropy, due entirely to linkage disequilibrium, D_{ij}:

$$\begin{aligned}\operatorname{cov}(z, z') &= \sum_{i=1}^{n}\sum_{j=1}^{n'} \alpha_i\alpha'_j \operatorname{cov}(x_i, x'_j) \\ &= 2\sum_{i=1}^{n}\sum_{j=1}^{n'} \alpha_i\alpha'_j(1 + F_{is})D_{ij}.\end{aligned} \tag{A2}$$

The factor $(1 + F_{is})$ arises because if the state of the two homologous genes at each locus are correlated (F_{is}), then the covariance between the contributions x_i, x'_j includes a component due to the associations across chromosomes as well as within chromosomes: $\operatorname{cov}(x_i, x'_j) = 2(1 + F_{is})D_{ij}$. We wish to replace the α's with measurable quantities. Consider the difference in mean of a particular trait between populations fixed for *bombina* alleles on the one hand and those fixed for *variegata* alleles on the other hand, Δz. This difference equals the sum of effects of all *variegata* alleles, $\Sigma_{2n}\,\alpha_i$. Thus we can write

$$\operatorname{cov}(z, z') = \frac{1}{2}\Delta z\Delta z'(1 + F_{is})D^*. \tag{A3}$$

Here, D^* is the average linkage disequilibrium between diagnostic loci. If the alleles underlying a given distinguishing trait are not fixed on either side of the zone, then the effect of alleles at the ith locus on Δz is devalued by Δp_i, the difference in allele frequency across the cline. Similarly, the amount of linkage disequilibrium generated by dispersal is reduced from D^* to $D^*\Delta p_i\Delta p_j$. The value D^* is thus an estimate of the linkage disequilibrium *between loci fixed across the hybrid zone* (see Sanderson et al. 1992 for details).

Now, consider a polygenic model that is additive across loci, but which allows for dominance: alleles have effects $(-\alpha, \delta, \alpha)$ at each locus. All the clines in allele frequency are assumed to be concordant. The means $\bar{z}$, $\bar{z}'$ are then

$$\bar{z} = \sum_{i=1}^{n} \alpha_i(p - q) + 2\delta_i pq, \tag{A4A}$$

$$\bar{z}' = \sum_{j=1}^{n'} \alpha'_i(p - q) + 2\delta'_i pq. \tag{A4B}$$

Assuming Hardy-Weinberg, the covariance is

$$\begin{aligned}\operatorname{cov}(z, z') = 2\sum_{i=1,j=1}^{n,n'} D^*_{ij}\{&[\alpha_i - \delta_i(p - q)][\alpha'_j - \delta'_j(p - q)] \\ &+ 2\delta_i\delta_j D^*_{ij}\},\end{aligned} \tag{A5}$$

where D^* is the linkage disequilibrium between each pair of loci, and the sum is across all pairs of loci. The variance of a single trait is given by the obvious extension. Consider for simplicity diagnostic loci, and label alleles such that p changes from 0 to 1 at each locus. One then expects that if disequilibrium is built up by dispersal, and if all clines have the same width, D^*_{ij} should be the same for all pairs of loci. Moreover, the changes in mean across the hybrid zone are not affected by dominance and are still given by $\Delta z = \Sigma_i\,\alpha_i$, $\Delta z' = \Sigma_j\,\alpha'_j$. Equation (A3) is then always fairly accurate. It is exactly correct if $p = ½$, because then equation (A5) reduces to equation (A2). If both traits have a mix of dominant and recessive alleles such that the mean changes linearly ($\Sigma\,\delta_i = 0$), it is also exactly correct, for all p.

Equation (A5) assumes Hardy-Weinberg proportions. If F_{is} is included, the expression is more complicated, but still reduces to equation (A3) if the mean changes linearly ($\Sigma\,\delta_i = 0$). This analytic result was confirmed by constructing two concordant quantitative traits from the data on the four diagnostic enzymes, using sites 2147 and 1011 (table 1). The first trait was determined by two genes (one dominant, one recessive: $\delta = +\alpha, -\alpha$), and the second in the same way, but by the other two genes. Averaging over all possible permutations that satisfy this scheme, the covariance between the traits was exactly the same as between two additive traits.

We can generalize to discordant clines by considering the case in which linkage disequilibria are maintained by a balance between dispersal and recombination:

$$D^*_{ij} = \frac{\sigma^2}{r_{ij}}\frac{\partial p_i}{\partial x}\frac{\partial p_j}{\partial x}. \tag{A6}$$

This approximation is derived assuming weak selection, but is accurate even for selection as strong as is typical of the *Bombina* hybrid zone (Barton and Gale 1993). Allele frequencies change from 0 to 1 across the hybrid zone. Differentiating equation (A4) gives the gradients in means:

$$\frac{\partial z}{\partial x} = \sum_{i=1}^{n} 2\frac{\partial p_i}{\partial x}[\alpha_i - \delta_i(p_i - q_i)], \tag{A7A}$$

$$\frac{\partial z'}{\partial x} = \sum_{i=1}^{n'} 2\frac{\partial p_j}{\partial x}[\alpha'_j - \delta'_j(p_i - q_i)]. \tag{A7B}$$

Substituting equation (A6) into equation (A5), and approximating the r_{ij} by their harmonic mean gives

$$\begin{aligned}\operatorname{cov}(z, z') = 2\sum_{i=1,j=1}^{n,n'} &\frac{\sigma^2}{r}\frac{\partial p_i}{\partial x}\frac{\partial p_j}{\partial x} \\ &\cdot[\alpha_i - \delta_i(p - q_i)][\alpha'_j - \delta'_j(p_j - q_j)].\end{aligned} \tag{A8}$$

Because we assume no pleiotropy, the two sets of loci are independent, and this sum can be separated into a product, given the exact relation:

$$\operatorname{cov}(z, z') = \frac{\sigma^2}{2r}\frac{\partial z}{\partial x}\frac{\partial z'}{\partial x}. \tag{A9}$$

A much more general (though approximate) argument can be made, which applies to epistasis as well as dominance. Consider a trait that is an arbitrary function of a set of additive traits, each scaled to run from 0 to 1: $z = \Sigma_i\, f_i(y_i)$, such that $\partial z/\partial x = \Sigma_i\,(\partial f_i/\partial y_i)(\partial y_i/\partial x)$, and approximately (for small fluctuations in the y's), $\operatorname{cov}(z, z') = \Sigma_i\,(\partial f_i/\partial y_i)(\partial f_i/\partial y'_j)\operatorname{cov}(y_i, y'_j)$. Because for additive traits, $\operatorname{cov}(y_i, y'_j) = (\sigma^2/2r)(\partial y_i/\partial x)(\partial y'_j\,\partial x)$, equation (A9) holds.