

LRH: TIMOTHY H. VINES ET AL.
RRH: STABILITY OF A MOSAIC HYBRID ZONE

THE MAINTENANCE OF REPRODUCTIVE ISOLATION IN A
MOSAIC HYBRID ZONE BETWEEN THE FIRE-BELLIED
TOADS *BOMBINA BOMBINA* AND *B. VARIEGATA*.

T.H.Vines^{1,2,5}, S.C. Köhler², M. Thiel², I. Ghira³, T.R. Sands¹,
C.J. MacCallum⁴, N.H. Barton¹ and B. Nürnberger²

¹ *Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland*

² *Department Biologie II, Ludwig-Maximilians-Universität, Karlstr. 23, D-80333 München, Germany*

³ *Department of Zoology, Babes-Bolyai University, Str. Kogalniceanu, Nr. 1, 3400 Cluj-Napoca, Romania*

⁴ *Trends in Ecology and Evolution, Elsevier Science London, 84 Theobald's Road, London, WC1X 8RR, England*

⁵ Author for correspondence (tvines@holyrood.ed.ac.uk)

ABSTRACT

Mosaic hybrid zones arise when ecologically differentiated taxa hybridize across a network of habitat patches. Frequent interbreeding across a small scale patchwork can erode species differences that might have been preserved in a clinal hybrid zone. In particular, the rapid breakdown of neutral divergence sets an upper limit to the time for which differences at marker loci can persist. We present here a case study of a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata* (Anura: Discoglossidae) near Apahida in Romania. In our 20 x 20 km study area, we detected no evidence of a clinal transition but found a strong association between aquatic habitat and mean allele frequencies at four molecular markers. In particular, pure populations of *B. bombina* in ponds appear to cause massive introgression into the surrounding *B. variegata* gene pool found in temporary aquatic sites. Nevertheless, the genetic structure of these hybrid populations was remarkably similar to those of a previously studied transect near Pescenica (Croatia), which had both clinal and mosaic features: estimates of heterozygote deficit and linkage disequilibrium in each country are similar. In Apahida, the observed strong linkage disequilibria should stem from an imperfect habitat preference that guides most (but not all) adults into the habitats to which they are adapted. In the absence of a clinal structure, the inferred migration rate between habitats implies that associations between selected loci and neutral markers should break down rapidly. While plausible selection strengths can maintain differentiation at those loci adapting the toads to either permanent or temporary breeding sites, the divergence at neutral markers must be transient. The hybrid zone may be approaching a state in which the gene pools are homogenized at all but the selected loci, not dissimilar from an early stage of sympatric divergence.

Key words.— linkage disequilibrium, gene flow, migration, selection, habitat preference, speciation

When ecologically differentiated taxa interbreed, the ultimate outcome of hybridization may critically depend on the distribution of habitat in the zone of contact. While a smooth clinal transition should form across an abrupt ecotone, a more interspersed mixture of habitat types may result in a mosaic of distinct genotypes (Harrison and Rand 1989). The associated increase in the rate of hybridization should accelerate introgression at neutral loci. Taxon differences that do not confer adaptation to one or the other habitat may disappear, and the situation may ultimately resemble a case of early sympatric divergence driven by local adaptation. Moreover, the conditions for the maintenance of adaptive differences become more stringent as the mosaic takes up an increasingly large area between the two pure gene pools. Explaining such combinations of broad geographic gradients with fine-scaled divergence may require a hybrid between existing theories, which apply to either smooth clines or to metapopulations of discrete demes (Endler 1977; Barton and Whitlock 1997; Pannell and Charlesworth 2000). Such theory could show how stability and the strength of the barrier to gene flow will vary within a hybrid zone as a function of the distribution of habitat. Here, we describe a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata* in Romania. Our analysis of these data suggests that neutral differences between the two taxa may be collapsing rapidly.

Adaptive differences across an ecotone may be maintained indefinitely even when they are only weakly selected (Haldane 1948). In a habitat mosaic, their maintenance is assured provided that the strength of selection s exceeds the rate of migration m . In a continuously distributed population, the habitat patches must be larger than the characteristic distance $l = \sigma/\sqrt{s}$, where σ is the dispersal range (Slatkin 1973). However, when the organisms do not disperse at random, as is assumed in these models, but show a preference for the habitat to which they are adapted, then a mosaic pattern of genotypes may exist even on smaller spatial scales. Habitat preference thus plays a similarly critical role in fine-scaled mosaic hybrid zones as it does in sympatric divergence (Via 2001). The additional immigration of pure types from the periphery helps to sustain the mosaic pattern.

As the mosaic widens, the influence of pure types must weaken at the centre. There may be a critical threshold distance at which the centre reverts to a simple patch system that may or may not maintain differentiation depending on the local balance of dispersal among patches, selection, and the strength of habitat preference. Examples of mosaics apparently maintained by associations with patchily distributed habitat include *Gryllus* ground crickets (Rand and Harrison 1989; Harrison and Bogdanowicz 1997), *Sceloporus* lizards (Sites et al. 1995) and *Allonemobius* crickets (Howard 1986; Britch et al. 2001), although the degree to which these associations are due to selection or an active habitat preference is currently unknown. In the case of *Iris*, individual plants are associated with different habitats over very small scales. This is probably

due to a combination of strong ecological selection and limited seed dispersal (Arnold and Bennett 1993).

Mosaic patterns may also arise from random drift if the number of migrants between demes (Nm) or the neighborhood size is small (Wright 1943). A particularly important form of drift is due to long range dispersal into the space between two expanding populations (Ibrahim et al. 1996), or into unoccupied patches in the mosaic from the parental populations (Nichols and Hewitt 1994). Examples of patchwork distributions apparently generated by dispersal and drift (as opposed to habitat associations) include *Solenopsis* ants (Shoemaker et al. 1996), *Mus* (Hauffe and Searle 1993), *Limnoporus* water striders (Klingenberg et al. 2000) and *Palaemonetes* shrimps (Garcia and Davis 1994). An interesting example where both factors may be involved is the hybrid zone between the crickets *Chorthippus brunneus* and *C. jacobsi*, where habitat associations alone cannot satisfactorily explain the deviations in allele frequency away from a smooth cline. The excess variation here could be explained by long distance dispersal from pure populations into the area of contact, either when the zone was initially forming or after the recurrent extinction of hybrid populations (Bridle et al. 2001; Bridle and Butlin 2002).

More generally, hybrid zones represent barriers to gene flow that delay but cannot ultimately prevent introgression of neutral traits into the opposite gene pool (Barton and Bengtsson 1986). The initial association between neutral and selected loci must break down as long as some recombination occurs between them. However, the time scale of this process is typically very slow, such that original differences at marker loci are still diagnostic today in most recent (i.e. post-glacial) hybrid zones. The higher rate of hybridization in mosaic hybrid zones should hasten the decay of differences in neutral allele frequencies. This rate of decay depends on the total selection strength and on the way this selection is distributed across the genetic map. The barrier to gene flow becomes more effective as the total selection increases and as it is distributed over an increasing number of loci (Barton and Bengtsson 1986). While most hybrid zones may be far from equilibrium with regard to neutral loci, the same need not be true for those loci that are under selection. Despite extensive introgression at marker loci, the organisms may continue to be very well adapted to the habitat patches in which they are found. Mosaic hybrid zones are thus promising places in which to investigate the erosion rates of both selected and neutral traits under hybridization, a issue which has received much recent attention (e.g. Reiseberg et al. 1999; Wu 2001; Lenormand 2002).

In this paper, we present data from a mosaic hybrid zone between the fire-bellied toads *B. bombina* and *B. variegata* in Romania. These two taxa have been diverging for around 3-5 million years, and have adapted to different breeding habitats. While *B. bombina* reproduces in semi-permanent ponds, *B. variegata* is a typical puddle breeder. Nevertheless, they produce abundant fertile hybrids wherever their ranges meet (Szymura 1993). Several of its features

should adapt *B. variegata* to reproduction in and dispersal between ephemeral habitats: a robust skeleton, thicker skin (to reduce water loss), larger eggs and faster developing tadpoles. *B. bombina* is smaller, and produces a larger number of slow growing and consistently less active tadpoles. The latter trait is thought to reduce their visibility to pond based predators (Lawler 1989, Kruuk and Gilchrist 1997; Vorndran et al. 2002). *B. bombina* is found in lowlands and flood plains throughout Central and Eastern Europe, whereas *B. variegata* is found at higher altitudes in Western and Central Europe (Fig. 1a).

[Figure 1]

Bombina hybrid zones in Poland, Croatia and the Ukraine consist of narrow clines 5-10 km wide, with extensive pure populations on either side (Szymura and Barton 1986, 1991; MacCallum et al. 1998; A. Yanchukov, pers. comm. 2001). These clines are concordant for a range of morphological traits and allozyme markers, which appears to be a consequence of strong linkage disequilibria across the genome (Sanderson et al. 1992, Nürnberger et al. 1995). There is also evidence for reduced hybrid viability in Poland (Koteja 1984) and Croatia (Kruuk et al. 1999a). In the Pescenica transect in Croatia, an active preference of adults for either ponds or puddles contributes to the association between habitat and marker genotype in the cline centre (MacCallum et al. 1998). This mechanism may also be operating in the Ukrainian transect where the same habitat association is observed (A. Yanchukov, pers. comm. 2001). Although no systematic survey of habitats was made in Poland, habitat associations cannot be strong, because the hybrid zones fit closely to a set of smooth clines (Szymura and Barton 1991). As we will show, these clinal patterns contrast sharply with Romania, where there is no steep cline in genotype frequencies. Instead, the distribution of toads reflects the distribution of habitats in the area, with *B. bombina* alleles being associated with pond-like habitats and *B. variegata* with puddles. In this paper, we attempt to quantify how the hybrid zone in Romania differs from the clinal pattern seen in Pescenica, Croatia, and then explore the consequences of this mosaic structure on the fate of selected and neutral differences between *B. bombina* and *B. variegata*.

METHODS

In 1999, we sampled waterbodies extensively around Cluj county, in the NW of the Transylvanian Plain of Romania, and collected 336 toads from 37 sites (Fig. 1a and b). Adult toads were caught by hand or with a net and immediately anesthetized in 0.2% MS222 (3-amino benzoic acid ethyl ester, Sigma). A toe was taken as a tissue sample and stored in 99.9% ethanol. They were subsequently scored for 3 diagnostic marker loci: two SSCPs (*Bb7.4* and *Bv24.11*) and one microsatellite (*Bv12.19*), all of which segregate independently (B. Nürnberger, pers. comm. 2000). These samples provide a wider regional context for the Apahida area (see Fig. 1b), and they were not included in the detailed genetic analyses (see below).

Toad Collections

In 2000 and 2001 we focused our attention in an area 20 by 20 km, our main study area near the village of Apahida on the Transylvanian plain, Romania (Fig. 1b and c). The south west corner of this area was located at 46°50'N, 23°47'E. Here, the landscape consists of rolling hills dissected by small river valleys flowing to the nearby Somes river to the north west. The soil is mostly sandy loam, becoming more clay rich in the valleys. The altitude ranges from 300 to 550 m above sea level, and the vegetation is either small arable strip fields or pasture. There are some small woodlands and few patches of forest in the area, principally beech, oak and hornbeam. Toads were collected from 93 sites: 745 individuals from 70 sites in 2000, and a further 189 toads from 23 new sites in 2001. Only three of the sites were large ponds, the majority being either drainage ditches or tractor wheel ruts. One group of sites consisted of 10 excavated holes around 3-4 m deep and 3-5 m in diameter. Several smaller, isolated watering holes were also found in other parts of the study area. Adult toads were caught and anesthetized as before. After a photograph was taken of the toads' individual belly pattern, a toe was taken as a tissue sample from either the right (2000) or the left (2001) foot and stored in 99.9% ethanol. We also measured snout-vent and tibiofibula lengths, and recorded the presence of nuptial pads (found only on breeding males), dorsal warts and dorsal spots. Recaptures were simply re-photographed and released.

Quantifying the difference in habitat use

The major habitat difference between *B. bombina* and *B. variegata* is the persistence of the waterbody through the season, which correlates well with many physical and ecological variables such as size of the water body and abundance of aquatic vegetation. Sites also vary in age and thus in successional stage, which is exemplified by the 10 water holes mentioned above: the first was dug in the 1950's and the last in the mid 1990's (Ilie Tudorescu, pers. comm. 2000). Clear ecological differences exist between them. To quantify the permanence of a site, the

following variables were measured (following MacCallum et al. 1998): width, depth, % emerged vegetation, % submerged vegetation and % bank vegetation (subdivided into three height classes: % <15 cm, % 15-50 cm, % > 50 cm), along with the location on our 20 by 20 km grid (i.e. the x, y coordinates).

To enable direct comparison between the Romanian habitats and those sampled in Croatia, we calculated a discriminant function axis on both data sets jointly based on a sample of typical ponds and puddles. All variables were transformed to improve their normality (log for continuous variables, arcsine for percentages: Sokal and Rohlf 1981) and the discriminant function was found for 152 sites (25 ponds and 36 puddles in Romania, 23 and 68 in Croatia) using the stepwise routine in SPSS version 5.5 (SPSS Inc., Chicago). The results are given in Table 1. The four retained variables were width, % emerged vegetation, depth and % submerged vegetation. With the exception of substituting % bank cover <15 cm for % submerged vegetation, the variables are the same as those in the axis calculated by MacCallum et al. (1998). The distinction between the two habitat types was highly significant ($F_{4,146} = 91.3, P < 10^{-6}$). The function was then calculated for all sites, and the discriminant score was rescaled to run from 0 (ponds) to 1 (puddles), using the most pond-like and the most puddle-like sites in the dataset as a whole as endpoints. Both these were in Croatia, the minimum and maximum values in Romania were 0.15 and 0.91 respectively. We denote this axis by H .

[Table 1]

Genotyping methods

The 2000 and 2001 animals were scored for *Bb7.4*, *Bv12.19* and *Bv24.11*, along with an additional unlinked microsatellite locus *Bv24.12*. The GenBank accession numbers for these loci are AF472441, AF472423, AF472425 and AF472426, respectively. Two competing ways of scoring locus *Bv12.19* were found for 479 of the 748 animals genotyped in 2000, depending on whether or not weak peaks in the curve output of our automatic sequencer (AlfExpress, Amersham Biosciences) were scored as alleles. Including these faint alleles gave only a weak heterozygote deficit $F_{IS} = 0.09$, whereas leaving them out gave $F_{IS} = 0.54$. The former is similar to the deficit at *Bv12.19* for the other 269 samples from 2000 ($F_{IS} = 0.08$) and the 206 toads genotyped in 2001 ($F_{IS} = 0.02$). Since there was no reason to believe that *Bv12.19* was subject to greatly different evolutionary forces between the samples, the faint alleles were included.

Based on the analysis of pure populations from Romania, there were two *B. variegata* alleles and one *B. bombina* allele at *Bb7.4*, *Bv12.19* and *Bv24.11*, whereas *Bv24.12* had five characteristic of *B. variegata*, one characteristic of *B. bombina*, and one of exactly intermediate length. Since the latter was found on only four occasions, and then only in hybrid sites, it was left unscored (this has no detectable effect on the results). The *B. variegata* alleles were combined

to make all the loci biallelic, alleles are henceforth labelled either 0 (*B. bombina*) or 1 (*B. variegata*). Mis-scoring rates for these loci have been found to be around 0.5% in a parallel study (B. Nürnberger, unpublished data). We summarize the state of a toad's genome by the total number of '1' alleles carried, forming a hybrid index (HI) which runs from 0 for pure *B. bombina* to 8 for pure *B. variegata*. It is important to bear in mind that this measure is an estimate based on four markers, and individuals with a 'pure' HI of 0 or 8 may well be introgressed at other loci (Boecklen and Howard 1997).

Statistical approaches

We use maximum likelihood techniques to estimate heterozygote deficits and linkage disequilibria (Edwards 1972; Hill 1974; see Szymura and Barton (1991) and MacCallum et al. (1998) for details). All calculations were performed on the Macintosh software package Analyse (by S.J.E. Baird), available at <http://helios.bto.ed.ac.uk/evolgen>. Briefly, the goodness-of-fit of an estimate for either heterozygote deficit (F_{IS}) or standardized linkage disequilibrium (R) is given by the natural logarithm of its likelihood (LogL), the difference between two LogL estimates is distributed as $\chi^2/2$ with 1 degree of freedom in large samples, provided that the estimate is not at the boundary. Heterogeneity between sites in F_{IS} can be assessed by comparing log likelihoods when F_{IS} is held constant or allowed to vary between them. Following MacCallum (1994), we account for heterozygote deficit when estimating linkage disequilibrium by assuming for a pair of loci that a proportion F_{IS} of the offspring is produced by union of identical gametes, and is thus homozygous. The remainder of gametes, $1 - F_{IS}$, unite at random. We standardize pairwise linkage disequilibrium (D) by the allele frequencies to give $R_{ij} = D_{ij} / \sqrt{p_i q_i p_j q_j}$. This facilitates comparisons between sites with different allele frequencies. However, this standardization does not fully remove the effect of allele frequencies: the full range of R values (-1 to 1) is only possible when $p = q = 0.5$ (Lewontin 1988).

RESULTS

The spatial pattern of hybridization between *Bombina bombina* and *B. variegata* in the area around Apahida (Romania) differs strikingly from the narrow transition zones that have been described in Poland, Croatia and in the Ukraine (Szymura and Barton 1986, 1991; MacCallum et al. 1998; A. Yanchukov, pers. comm. 2001). There, allele frequency clines were found in bands 6 – 9 km wide, located at ecotones between forested hills and open plains and flanked on either side by extensive areas containing only pure animals. In contrast, the hybrid populations to the east of Apahida (Romania) are about 20 km from pure *B. variegata* populations to the NW in the foothills of the Apuseni mountains and 100 km from the main expanse of pure *B. bombina* in the Hungarian plains (Fig. 1b). Nevertheless, populations of the latter are found sporadically in large ponds across the entire Transylvanian Plains, often in close proximity to the puddle-dwelling *B. variegata* (Stugren 1959; Stugren and Vancea 1968).

Our collections across an area of about 40 x 40 km around the city of Cluj confirm this extended mosaic (Fig. 1b). Isolated *B. bombina* sites in areas dominated by *B. variegata* have been documented before on the Slovakian karst plateau (Gollmann et al. 1988) and in the Mátra mountains of Hungary (Gollmann 1987). Furthermore, in Kostajnica, Bosnia, large allele frequency differences were observed between pond and puddle sites even when they were in close proximity (Szymura 1988, 1993). In the following analyses, we compare the spatial and genetic structure of the Apahida (Romania) study area to the previously described hybrid zone in Pesčenica, Croatia (MacCallum et al. 1998; our Fig. 1d). The latter has a strong clinal structure across an altitudinal gradient, yet also has a mosaic pattern at its centre. Around Apahida, in contrast, there is no evidence for a cline over the same spatial scale (Fig. 1c), and neither is there is a comparable ecological gradient (see below). It should be noted that the genetic data for Croatia were from four allozyme loci, whereas here four neutral DNA markers are used. Since the allozymes *Ldh-B* and *Mdh-1* and SSCP loci *Bb7.4* and *Bv24.11* are closely concordant in a Ukrainian transect (A. Yanchukov, pers. comm. 2001), and since all loci are concordant with each other (Szymura and Barton 1991), we consider results based on either allozymes or DNA markers to be directly comparable.

Spatial distribution of allele frequencies and habitats

In the majority of sites, the mean frequency of *B. variegata* alleles (p) among sampled adults ranged from 0.5 to 0.8. Only four sites had $p < 0.4$, and in these, only 13 out of 33 animals contained only *B. bombina* alleles (hybrid index: HI = 0). Overall, the adult sample was dominated by hybrid individuals (Fig. 2). To explore the relative importance of a clinal transition and of the spatial distribution of habitat, we regressed the x and y -coordinates and the discriminant habitat score H for each site onto the allele frequency p . The fitted regression was $p = 0.53 -$

$0.009x - 0.002y + 0.44H$. The overall fit of the model is significant: $F_{3,89} = 15.3$, $P < 0.0001$. Neither the x or the y spatial location axes were significant ($P = 0.14$ and $P = 0.66$ respectively), and the habitat variable was of overwhelming importance ($F_{1,89} = 27.6$, $P < 0.0001$; see Figure 3). When the same analysis was applied to MacCallum's (1994) data on the Pescenica transect the best fit was $p = 0.93 - 0.05x - 0.01y + 0.38H$, which was also significant: $F_{3,86} = 63.7$, $P < 0.0001$. Spatial location was significant in this case (x : $F_{1,86} = 82.7$; y : $F_{1,86} = 9.9$; $P < 0.005$ for both), as expected from the layout of the transect along the SW-NE axis (Fig. 1d). The habitat score H was also highly significant ($F_{1,86} = 11.2$, $P = 0.001$). A partial Mantel test (Smouse et al. 1986; Manly 1991) comparing the relative contribution of physical distance and difference in habitat score to the difference in p between pairs of sites gives a qualitatively similar result for both countries.

[Figure 2] [Figure 3]

We can describe the spatial distribution of habitat by a least-squares regression of the habitat axis H against the x and y coordinates. There is a significant relationship between east-west position and habitat type in Croatia ($F_{1,90} = 19.7$, $P < 0.0001$), but not north-south ($P > 0.1$). There is no relationship in Romania ($P > 0.1$ for both x and y). This suggests that different structures of the hybrid zones in Croatia and Romania might be due to a different spatial distribution of habitat. Note that these data only address occupied sites rather than all waterbodies. Collecting data on the latter is hampered by the difficulty of defining a 'usable' waterbody from the perspective of a toad. We are confident that the majority of occupied sites in the study area in 2000 and 2001 were sampled.

Association between habitat and allele frequency

In Romania, we can gauge the overall allele frequency difference between ponds ($H = 0$) and puddles ($H = 1$) from the slope of a regression of p onto the habitat axis (Fig. 3). This gives $\Delta p = 0.39$. Getting a comparable figure for Croatia is difficult: the steep cline confounds any quantification of habitat associations over large spatial scales. We can, however, examine how the difference in p corresponds to the difference in habitat score between nearby sites. Taking only those pairs that are less than 1 km apart (an estimate of the lifetime dispersal distance derived from the Polish and Croatian transects: Szymura and Barton 1991; MacCallum et al. 1998), the relationships in Croatia and Romania are broadly similar: $\Delta p = 0.16 \Delta H$ and $\Delta p = 0.30 \Delta H$ respectively (Fig. 4). It is difficult to test whether there is any significant difference between these relationships, because the regressions are based on pairs of sites rather than on independent data, and hence significance will be overestimated. (This does not bias the regression estimate, however).

[Figure 4]

These slopes provide a second estimate of the maximum difference in allele frequency between ponds and puddles (i.e. when $\Delta H = 1$, $\Delta p = 0.16$ and 0.30 in Croatia and Romania respectively); we consider them a more reliable reflection of habitat association because they are estimated over similar scales, of the same order as the dispersal range (see Discussion). We note that estimates of habitat association were also made in two different ways in a previous analysis of the Croatian data. MacCallum et al. (1998) calculated the difference in allele frequency between adjacent ponds and puddles ($\leq 300\text{m}$ apart) as 0.25 ; this figure is probably less accurate, however, as it is based on only 7 pairs of sites. MacCallum (1994) fitted a model of a sigmoid cline to the whole Croatian dataset, including the habitat as an additional explanatory variable. This gave $\Delta p = 0.15$, which is close to our simpler estimate of 0.16 .

Genotype frequencies: heterozygote deficit and linkage disequilibria

We divided the sites into seven groups by their mean allele frequency and obtained F_{IS} estimates across all loci within each group. When F_{IS} is estimated as a common value across all loci, the peak is on the *B. bombina* side (max. $F_{IS} = 0.22$ at $p = 0.21$; Fig. 5a). Since there are few sites in the three left hand groups ($n = 1, 2$ and 3 respectively), this asymmetry does not result in a better fit for a cubic model ($F_{3,6} = 1.34$, $P = 0.33$) over a simple linear one ($F_{1,6} = 52.05$, $P = 0.005$) when the means are weighted by their support limits. However, the means for each bin show a similar asymmetric pattern to that seen in Croatia (MacCallum et al. 1998; our Fig. 5b), where the maximum F_{IS} is 0.23 at $p = 0.32$ when a weighted regression is used. (Note that the equivalent regressions in MacCallum et al. 1998 are unweighted; the cubic fit is the best for their data with either method). We show the cubic curve in Figure 5a for comparison with the Croatian data.

The standardized linkage disequilibria between loci, $R_{ij} = D_{ij} / \sqrt{p_i q_i p_j q_j}$ were estimated across all sites. In these computations, the non-zero estimates of F_{IS} were taken into account in order to remove any undue inflation of disequilibrium through correlations of genes within loci (see Materials and Methods). Table 2 gives values of R estimated across all sites for each pair of loci. Across all sites and loci, $R = 0.090$ (support limits: $0.083, 0.097$). This indicates that combinations of genes found in the parental taxa are in excess, despite their constant breakup by recombination. To examine how linkage disequilibrium changes with allele frequency, R was estimated for all loci within each of the seven groups of sites used above (Fig. 5c). Again, the maximum R estimate ($R = 0.38$ at $p = 0.28$; Fig. 5c) was shifted towards *B. bombina*-like hybrid populations, similar to Croatia (there $R = 0.39$ at $p = 0.39$; Fig. 5d). However, as for heterozygote deficit, F_{IS} , the cubic regression ($F_{3,6} = 4.98$, $P = 0.11$) did not give a signifi-

cantly better fit than a linear ($F_{1,6} = 18.59$, $P = 0.02$). Again, we show the cubic curve fit for comparison with the Croatian plot.

[Table 2] [Figure 5]

DISCUSSION

Our survey of genotype frequencies across the *Bombina* hybrid zone around Apahida is in striking contrast with the narrow clines seen in Poland and Croatia. No steep gradient in allele frequencies was apparent; instead, there was a fine-scaled mosaic with strong divergence in marker frequency between different habitats. The frequency distribution of pure and hybrid populations was asymmetric. Pure *B. bombina* populations were only found in sporadically occurring large ponds, whereas *B. variegata*-like hybrid populations inhabited the much more abundant temporary sites in the surrounding landscape (Figs. 2 & 3). This asymmetry is even more remarkable considering that the nearest extensive pool of entirely pure *B. bombina* lies 100 km away in the Hungarian plains: local strongholds of *B. bombina* appear to cause the observed massive introgression of *B. bombina* alleles into the surrounding *B. variegata*-like gene pool, in which 83% of adults were recombinant at our four marker loci.

Despite the very different spatial pattern of the hybrid zones near Apahida and near Pescenica, Croatia, their local adult populations had broadly similar genetic structures (see also MacCallum et al. 1998). When sites were pooled by their allele frequency, the maximum heterozygote deficit was $F_{IS} = 0.23$ in Croatia and 0.21 in Romania (Fig. 5a and b). The analogous maxima for the standardized linkage disequilibrium, R , were 0.40 and 0.38, respectively (Fig. 5c and d). Moreover, both parameters showed an asymmetric pattern in both regions, being stronger in sites with more *B. bombina*-like toads. Associations between habitat and marker frequencies were of similar strength ($\Delta p = 0.16$ in Croatia vs. 0.30 in Romania). The Polish transects (Szymura and Barton 1991) showed somewhat weaker linkage disequilibrium ($R = 0.22$), but no heterozygote deficit, and little evidence of habitat associations. However, their clinal pattern was similar to that in Croatia. Surveys in Slovakia (Gollmann 1988) and Kostajnica, Bosnia (Szymura 1988) are also consistent with a mosaic distribution created by an active habitat preference. The mosaic pattern in Romania raises several issues, which we address in turn: why do the Romanian and Croatian hybrid zones differ? How is strong linkage disequilibrium maintained within a broadly sympatric distribution? And, under what circumstances will either selected or neutral divergence persist in the face of migration and hybridization?

Why do the zones in Apahida and Pescenica differ?

The absence of a cline in allele frequencies sets the hybrid zone in Apahida apart from that in Pescenica. The explanations for this difference in structure include a) inherent differences in the hybridizing taxa in each area; b) differences in history of the two zones; and c) from differences in the spatial distribution of habitat.

Regional differences have been documented within each taxon. *B. variegata* is subdivided into a western and a Carpathian form: Nei's genetic distance D_N based on 9 allozyme loci is 0.16 compared to $D_N = 0.65$ between *B. bombina* and *B. variegata*. There is less differentiation between the northern and southern *B. bombina* ranges ($D_N = 0.09$). Near Apahida, the southern form of *B. bombina* hybridizes with the Carpathian form of *B. variegata*, a combination that has not previously been studied. The hybrids here may have a relatively stronger habitat preference or disperse over longer distances, both of which would promote the observed sympatric distribution. At present it is difficult to assess these hypotheses because we lack comparable estimates for the parameters involved. The association between habitat and genotype does appear to be stronger in Apahida than Pescenica (cf. regressions of Δp on ΔH , Fig. 4), but this could also result from the weaker clinal influence in the Romanian hybrid zone.

Differences in the history of the hybrid zones may account for their different structures. Two populations may meet in a straight front or in a patchy mosaic (Ibrahim et al. 1996). Random drift and local extinction may generate different levels of patchiness (e.g. Bridle et al. 2001), although neither of these stochastic effects can explain the strong habitat associations in Apahida. Alternatively, we may have sampled the two zones at different stages on their way to the same final state.

There is stronger evidence that the spatial distribution of habitat contributes to the difference in hybrid zone structure. The habitat score shows a clinal pattern in Pescenica that parallels the allele frequency cline, while no relationship exists between either spatial location and either habitat score or allele frequency in Apahida. However, these data are based only on occupied sites. One could argue a similar habitat mosaic exists in both locations, but the influx of pure types in Pescenica maintains a clinal structure. This implies the existence of ponds on the *B. variegata* side and puddles on the *B. bombina* side that are suitable but unoccupied. We find this unlikely for two reasons. First, the hybrid zone itself would have to prevent the movement of toads from one side to the other, even though both habitat types were available. Second, the hilly wooded terrain and the open flood plain are generally unfavorable for the establishment of ponds and puddles respectively. We therefore find it more plausible that differences in the layout of habitat contribute strongly to the structure of the hybrid zones in the two countries. Clearly, we cannot be certain that the other factors outlined above do not play a role.

How is strong linkage disequilibrium maintained within a broadly sympatric distribution?

In hybrid zones, strong linkage disequilibrium (D) are most likely caused by the admixture of populations, especially between unlinked neutral markers (Barton and Gale 1993, Kruuk et al. 1999b). In a strictly clinal hybrid zone, the influx of pure genotypes from the periphery should therefore largely account for any observed D . Habitat preference within a mosaic hybrid zone

has the same effect. It reduces the migration between adjacent sites of the opposite type, but those adults that do move introduce allele combinations typical of their source population. As the strength of habitat preference increases, the migrants become rarer, yet each contributes proportionally more to the observed D (Kruuk 1997). The Pescenica transect is an example of a hybrid zone which linkage disequilibria can be explained by a combination of clinal influence and habitat preference.

Given the much larger distances to the nearest source of pure individuals of either taxon in Apahida, we have argued above for a much weaker clinal influence. This implies an important role for habitat preference in generating D . Direct evidence should come from observations of non-random movement with regard to habitat. In Pescenica, there was a significant association between genotype and habitat over distances that were much shorter than the observed within-season dispersal range of adults (MacCallum et al. 1998).

In Apahida, it is more difficult to demonstrate non-random movement, because suitable sites are further apart, which makes the study of dispersal by mark-recapture less efficient. However, two observations indicate the existence of an active preference. First, the presence of pure *B. bombina* populations would otherwise imply extremely strong selection against *B. variegata* adults immigrating from the surrounding temporary sites. In addition, our sample included 13 adjacent man-made sites in varying successional stages (distance between adjacent sites ~ 5 m). Even on this scale, there was a significant correlation between genotype and habitat score ($r = 0.64$, $P = 0.006$). In the following, we assume that the local habitat associations in Romania are mainly due to active preference, as in Croatia.

Estimating the rate of introgression of pure B. bombina

Since linkage disequilibria in hybrid zones is largely a product of admixture between genetically differentiated populations, it can provide an estimate of dispersal. In strictly clinal hybrid zones, the mean dispersal distance can be determined in the context of a diffusion model (Barton and Gale 1993). In Apahida, however, there is no spatial cline. In order to assess the strength of evolutionary forces that operate in this mosaic, we instead consider a model of discrete habitat patches in a short term migration-selection balance.

Strong linkage disequilibria are seen in *B. bombina*-like hybrid populations and are mainly due to the introgression of pure *B. bombina* genotypes from ponds. This is suggested by the observation of 'pure' ($HI=0$) *B. bombina* adults in temporary sites with p as high as 0.59. These toads are unlikely to have been generated locally (Fig. 6), and so can be assumed to be newly arrived immigrants. Their presence is determined more by habitat than by the proximity to ponds: the sites in question have a significantly more pond-like habitat score than the remainder of the temporary sites ($t=4.87$, $P < 0.001$), yet the mean distance to the large pond (site 293) is 2.78

km. In the following, we concentrate on intermediate sites with $0.2 < p < 0.6$ and determine for each one of them the immigration rate m of pure *B. bombina* that would generate the locally observed D . We then use the mean of these estimates to ask whether plausible selection strengths could hold this zone in migration-selection balance.

Since these parameter estimates are based on equilibrium relationships, we need to consider the demography of the toad populations. Aside from the large pond (site 293), all sites are temporary in that they dry up before the end of summer in most years. Nonetheless, many of them will attract toads at the beginning of the season for several to many years in a row. As pairwise D equilibrate rapidly ($D_{t+1} = 0.5 D_t$ for unlinked loci), the estimation of migration rates from the current level of D appears justified. From mark-recapture studies in Croatia and Romania (MacCallum et al. 1998; Nürnberger et al., submitted; T. Sands, unpublished data) we know that the total pool of adults that visit a site per year is several times larger than the number present on any one sampling date. We therefore assume that the composition of this adult assemblage is approximately stable on the scale of, say, ten generations and that its size is large enough that deterministic theory can be applied.

[Figure 6]

In general, migration at a rate m between two populations that differ in allele frequency by Δp at each of two loci generates pairwise linkage disequilibrium proportional to $(m \Delta p^2)/r$, where r is the recombination rate between the two loci measured after dispersal and before reproduction. Here, we consider the immigration from pond populations with $p \sim 0$, so that Δp_i for a given intermediate site i is roughly equal to p_i . We assume that, after random mating, strong selection on the offspring maintains a constant Δp_i . Then, m_i can be estimated using $D_i = (m_i \Delta p_i^2)/r$. For the 23 populations in the range $0.2 < p < 0.6$, this yields an average immigration rate $m = 0.19$ (s.d. 0.19). Details of the computations are given in Appendix. The mean F_{IS} in these sites is 0.17; in our model $F_{IS} = m p / q (1-m)$, which gives 0.16 for $m = 0.19$ and $p = 0.4$. The migration rate inferred from linkage disequilibrium can thus also account for the observed deficit of heterozygotes.

The computation of a single migration rate between two subgroups of populations in Apahida is clearly an oversimplification. For example, adults with $HI = 0$ make up 11% of the samples in intermediate sites, which is about half of the proportion needed to account for the observed linkage disequilibria. The remainder may stem from the dispersal of recombinants: disequilibrium in the nine sites with $0.55 < p < 0.6$ can be accounted for by immigration of $m = 0.26$ from sites with $p = 0.2$. This observation suggests that the offspring of *B. bombina* immigrants are contributing to linkage disequilibria in the more *B. variegata*-like populations, and therefore introgression into the wider *B. variegata* pool must be occurring at an appreciable rate. Nonethe-

less, we use our estimate of $m = 0.19$ as a yardstick in the following consideration of selection strengths.

Will divergence be maintained in the traits adapting toads to ponds and puddles?

If, for simplicity, we divide the genome into selected and neutral loci, the question about the stability of this hybrid zone has two parts: a) will differentiation be maintained with respect to those traits that mediate adaptation (including the habitat preference), and b) for how long will frequency differences at neutral markers persist? Considering the first question initially for single locus, a locally adaptive allele will be maintained if the selection favoring it is stronger than the rate of migration ($s > m$: Haldane 1932 p. 210). As the number of selected loci increases, linkage disequilibria build up, and both the per locus selection (s) and the total selection across n loci ($S = ns$) determine the dynamics of the hybrid zone (Barton 1983). Moreover, the question of stability depends on precisely how selection acts. Since many of the differences between *B. bombina* and *B. variegata* are adaptations to different habitats, we concentrate on selection against alleles in the wrong habitat.

As a first approximation, we may assume that the influx of *B. bombina* alleles into populations with intermediate allele frequency ($m = 0.2$ into sites with $p > 0$, see above) is balanced by selection in these sites. We have made some numerical calculations for a symmetrical model with $n = 5$ to 20 unlinked loci, following Barton and Shpak (2000). If selection favours genotypes that are adapted to the local environment, selection is able to counter higher rates of immigration, simply because the immigrant genotypes have low fitness. With fitness $\text{Exp}[-2Sq]$, where q is the proportion of *bombina* alleles, any amount of migration can be balanced by strong selection. For example, $m = 0.2$ can be countered by selection on 20 loci of $S \sim 1.7$, which implies a fitness of immigrant pure genotypes ($q = 1$) of $\text{Exp}[-2S] = 3.3\%$, and of F1's of 18%. These values may not be inconsistent with the mild hybrid inviability observed in the lab (Koteja 1984, Kruuk et al. 1999a), as immigrant fitness may be much lower in the field.

This model assumes immigration from only one parental taxon (*B. bombina*). Since there must also be migration into the intermediate sites from more *B. variegata*-like populations (Fig. 2, see above), the swamping effect will be reduced, both because allele frequencies are under weaker pressure, and because linkage disequilibria will be higher. Moreover, the high inferred migration rate suggests that the *B. variegata*-like animals may constitute a single large population that uses a range of temporary habitat for reproduction. In this much larger gene pool, the *B. bombina* immigrants make up a smaller fraction ($m \sim 0.05$), so that less selection will suffice to maintain allele frequency differences. The fitnesses above are therefore an upper limit to the strength of selection that could indefinitely maintain local adaptation in the *B. variegata* pool.

Since this limit is not implausible, we have no reason to doubt the stability of this hybrid zone as far as selected loci are concerned.

What strength of genetic barrier is required to explain the current differences at neutral loci?

If indeed the traits that confer adaptation to temporary sites are largely intact in the *B. variegata*-like hybrid populations, then our data on marker alleles demonstrate the ongoing breakdown of neutral divergence: there are hardly any pure *B. variegata* populations in the area. On their own, neutral allele frequency differences should dissipate very rapidly, over an approximate timescale of $c. 1/m$ (here: five) generations. The fact that marker loci are informative in hybrid zones over much longer times scales is due to selection on linked loci, which creates a genetic barrier to gene flow (Barton 1983). In the present context, its effect is to reduce the actual migration rate to a lower effective rate m_e and hence to slow the inevitable decay of neutral divergence.

For the scenario above, selection in intermediate populations against incoming *B. bombina* alleles of $S \sim 1.7$ on 20 unlinked loci would reduce the migration rate at interspersed neutral loci to $m_e = 0.09 m$ (Barton and Shpak, 2000). Thus, a selection scheme that would stably maintain the adaptation of the toads to ponds and puddles, respectively, could sustain the divergence at marker loci for only 55 generations or, assuming a generation time of 5 years, for about 280 years. Clearly, selection schemes can be devised that generate larger barrier strengths. For example, Barton and Bengtsson (1986, Fig.4a) calculated the reduction in gene flow for a model of weak additive selection on n evenly spaced loci, with the neutral marker embedded at the centre. With 40 selected loci and $S = 1.7$ on one such chromosome, gene flow would be reduced 100 fold. This model implies the existence of one locus every 4 cM with $s = 0.04$ and could have sustained neutral divergence for 2500 years. Without empirical data on the total strength of selection and on its distribution across the genome, we cannot currently determine how long the allele frequency differences have persisted. Nonetheless, it is useful to note that neutral divergence in this and other mosaic hybrid zones is necessarily transient and thus poses constraints on feasible combinations of hybrid zone age and selection.

To summarize, the fire-bellied toads *B. bombina* and *B. variegata* form an extended mosaic hybrid zone near Apahida, Romania that differs remarkably in its spatial structure from previously analyzed clines (Poland: Szymura and Barton 1991; Croatia: MacCallum et al. 1998). Pure *B. bombina* populations occur sporadically in large ponds in open hilly terrain. they are the source of massive introgression into *B. variegata*-like populations in the abundant temporary aquatic sites throughout the area. Based on linkage disequilibria in intermediate sites, we derived an estimate of this asymmetric migration rate which also accounts neatly for the observed heterozygote deficit. Since the conditions for the maintenance of divergence are more stringent in mosaic as opposed to clinal hybrid zones, we investigated the minimal selection

strengths that are required to stabilize this hybrid zone. While plausible selection coefficients could maintain adaptive divergence, the high rate of introgression at neutral markers implies that non-critical taxon differences are eroding. Ultimately, this hybrid zone may come to resemble an early stage of ecological divergence in sympatry.

ACKNOWLEDGMENTS

We thank Gyongyver Mara and Trabantul Galbena for enthusiastic field assistance, Alex Hofmann and Robert Sieglstetter for access to their unpublished data, Bruni Förg-Brey and Gabi Praetzel for help in the lab. Helpful comments on a previous version of the manuscript were provided by R. Ennos, J. Szymura, F. Balloux, J. Bridle, L. Kruuk, F. Bonhomme, M. Arnold and two anonymous reviewers. This work was supported by Natural Environment Research Council studentships to THV and TRS and Deutsche Forschungsgemeinschaft grant Nu 51/2-1 to BN.

LITERATURE CITED

- Arnold, M. L., and B. D. Bennett. 1993. Natural hybridization in Louisiana irises: genetic variation and ecological determinants. Pp. 115-139 *in* R. G. Harrison, ed. *Hybrid zones and the evolutionary process*. Oxford University Press, Oxford.
- Asmussen, M. A., and M. M. Orive. 2000. The effects of pollen and seed migration on nuclear-cytoplasmic systems. I. Nonrandom associations and equilibrium structure with both maternal and paternal cytoplasmic inheritance. *Genetics* 155:813-831.
- Barton, N. H. 1983. Multilocus clines. *Evolution* 37:454-471.
- . 2000. Estimating multilocus linkage disequilibria. *Heredity* 84:373-389.
- Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. *Heredity* 57:357-376.
- 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.
- Barton, N. H., and M. Shpak. 2000. The effects of epistasis on the structure of hybrid zones. *Genet. Res.* 75:179-198.
- Barton, N. H., and M. C. Whitlock. 1997. The evolution of metapopulations. Pp 183-210 *in* I. Hanski and M. E. Gilpin, eds. *Metapopulation dynamics: ecology, genetics, and evolution*. Academic Press, San Diego, CA.

- Boecklen, W. J., and D. J. Howard. 1997. Genetic analysis of hybrid zones: number of markers and power of resolution. *Ecology* 78:2611-2616.
- Bridle, J. R., S. J. E. Baird, and R. K. Butlin. 2001. Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* 55:1832-1843.
- Bridle, J. R., and R. K. Butlin. 2002. Mating signal variation and bimodality in a mosaic hybrid zone between *Chorthippus* grasshopper species. *Evolution* 56:1184-1198.
- Britch, S. C., M. L. Cain, and D. J. Howard. 2001. Spatio-temporal dynamics of the *Allonemobius fasciatus* - *A. socius* hybrid zone: a 14-year perspective. *Mol. Ecol.* 10:627-638.
- Edwards, A. W. F. 1972. *Likelihood*. Cambridge Univ. Press, Cambridge.
- Endler, J. A. 1977. *Geographic variation, speciation and clines*. Princeton Univ. Press, Princeton.
- Garcia, D. K., and S. K. Davis. 1994. Evidence for a mosaic hybrid zone in the grass shrimp *Palaemonetes kadiakensis* (Palaemonidae) as revealed by multiple genetic markers. *Evolution* 48:376-391.
- . 1987. *Bombina bombina* and *B. variegata* in the Mátra mountains (Hungary): New data on distribution and hybridization (Amphibia, Anura, Discoglossidae). *Amphibia-Reptilia* 8:213-224.
- Gollmann, G., P. Roth, and W. Hödl. 1988. Hybridization between the fire-bellied toads *Bombina bombina* and *Bombina variegata* in the karst regions of Slovakia and Hungary. *J. Evol. Biol.* 1:3-14.
- Haldane, J. B. S. 1932. *The causes of evolution*. Longmans, New York.
- . 1948. The theory of a cline. *J. Genet.* 48:277-284.
- Harrison, R. G., and D. M. Rand. 1989. Mosaic hybrid zones and the nature of species boundaries. Pp 110-133 in D. Otte and J. Endler, eds. *Speciation and its consequences*. Sunderland, MA, Sinauer Associates.
- Harrison, R. G., and S. M. Bogdanowicz. 1997. Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution* 51:493-505.

- Hauffe, H. C., and J. B. Searle. 1993. Extreme karyotypic variation in a *Mus musculus/domesticus* hybrid zone: the tobacco mouse story revisited. *Evolution* 47:1374-1395.
- Hill, W. G. 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33:229-239.
- Howard, D. J. 1986. A zone of overlap and hybridization between two ground cricket species. *Evolution* 40:34-43.
- Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansions. *Heredity* 77:282-291.
- Klingenberg, C. P., J. R. Spence, and C. K. Mirth. 2000. Introgressive hybridization between two species of waterstriders (Hemiptera: Gerridae: *Limnopus*): geographical structure and temporal change of a hybrid zone. *J. Evol. Biol.* 13:756-765.
- Koteja, E. 1984. Mortality of the embryos of the fire-bellied toad (*Bombina bombina*), yellow-bellied toad (*Bombina variegata*) and their hybrids. M. Sc diss., Inst. Zoologii, Uniw. Jagiellon-ski, Krakow.
- Kruuk, L. E. B. 1997. Barriers to gene flow: a *Bombina* (fire-bellied toad) hybrid zone and multilocus cline theory. Ph. D. diss., University of Edinburgh.
- Kruuk, L. E. B., and J. Gilchrist. 1997. Mechanisms maintaining species differentiation: predator mediated selection in a *Bombina* hybrid zone. *Proc. Roy. Soc. Lond. B* 264:105-110.
- Kruuk, L. E. B., J. S. Gilchrist, and N. H. Barton. 1999a. Hybrid dysfunction in fire-bellied toads (*Bombina*). *Evolution* 53:1611-1616.
- Kruuk, L. E. B., S. J. E. Baird, K. S. Gale, and N. H. Barton. 1999b. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics* 153:1959-1971.
- Lawler, S. P. 1989. Behavioral response to predators and predation risk in four species of larval anurans. *Anim. Behav.* 38:1039-1047.
- Lawler, S. P. 1989. Behavioral response to predators and predation risk in four species of larval anurans. *Anim Behav.* 38:1039-1047.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183-189.

Lewontin, R. C. 1988. On measures of gametic disequilibrium. *Genetics* 120:849-852.

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

MacCallum, C. J., B. Nürnberger, N. H. Barton, and J. M. Szymura. 1998. Habitat preference in a *Bombina* hybrid zone in Croatia. *Evolution* 52:227-239.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

Nichols, R. A., and G. M. Hewitt. 1994. The genetic consequences of long distance dispersal during colonization. *Heredity* 72:312-317.

Nürnberger, B., N. H. Barton, C. J. MacCallum, J. Gilchrist, and M. Appleby. 1995. Natural selection on quantitative traits in a *Bombina* hybrid zone. *Evolution* 49:1224-1238.

Pannell, J. R., and B. Charlesworth. 2000. Effects of metapopulation processes on measures of genetic diversity. *Phil. Trans. R. Soc. Lond. B* 355:1851-1864.

Pounds, N. J. G. 1979. A historical geography of Europe 1500-1840. Cambridge Univ. Press, Cambridge.

Rand, D., and R. G. Harrison. 1989. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution* 43:432-449.

Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 152:713-727.

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.

Shoemaker, D. D., K. G. Ross, and M. L. Arnold. 1996. Genetic structure and evolution of a fire ant hybrid zone. *Evolution* 50:1958-1976.

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

- Sites, J. W., N. H. Barton, and K. M. Reed. 1995. The genetic structure of a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae) in Central Mexico. *Evolution* 49:9-36.
- Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35:627-632.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*. W.H. Freeman and Co., New York, NY.
- Stugren, B. 1959. Eidonomische Untersuchungen an *Bombina* Oken (Amph., Discoglossidae) aus dem Gurghiu-Tale. *Zool. Jb. Syst.* 86:382-394.
- Stugren, B., and S. Vancea. 1968. Geographic variation of the Yellow Bellied Toad (*Bombina variegata*) (L.) from the Carpathian mountains of Romania and the USSR. *J. Herpetol.* 2:97-105.
- Szymura, J. M. 1988. Regional differentiation and hybrid zones in the fire-bellied toads *Bombina bombina* and *B. variegata* in central Europe (in Polish). *Rozprawy Habilitacyjne* No. 147, Uniwersytet Jagiellonski, Krakow.
- . 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 *B. variegata*, near Cracow in Southern Poland. *Evolution* 40:1141-1159.
- Szymura, J. M., and N. H. Barton. 1991. The genetic structure of a hybrid zone between the fire bellied toads *Bombina bombina* and *B. variegata*: comparisons between loci and between transects. *Evolution* 45:237-261.
- Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* 16:381-390.
- Vorndran, I. C., E. Reichwald, and B. Nürnberger. 2002. Does differential susceptibility to predation on tadpoles stabilize the *Bombina* hybrid zone? *Ecology* 86: 1648 - 1659.
- Wright, S. 1943. Isolation by distance. *Genetics* 28:114-138.
- Wu, C. I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* 14:851-865.

TABLES

Table 1: Discriminant function coefficients (standardized) and their Wilks λ , which measures the effect of a given term on the function. Variables were entered together, and are ordered by their contribution to the function. The second and third columns give the mean coefficient for ponds and puddles respectively. All Wilks λ gave $P < 0.0001$. The final column gives the correlation of each variable with the discriminant function.

	Overall	Ponds	Puddles	Wilks λ	Correlation
Width	0.50	2.48	0.79	0.41	0.75
Emergent Veg	0.79	-7.21	-0.22	0.34	0.58
Depth	2.19	-2.24	4.93	0.29	0.57
Submerged Veg	0.66	0.67	-1.58	0.28	0.34
Constant	0.49	-5.60	-5.11		
Group Mean		2.33	-1.05		

Table 2: Pairwise standardized linkage disequilibria R across all sites. The maximum likelihood value of R for each pair of loci is given in the main body of the table. The column on the right hand side lists the averages per locus and an overall estimate, given that R is constrained to be constant across all sites and loci (bottom figure). All pairs of loci gave a significant excess of parental genotypes (i.e. $\Delta\log L > 2$).

	<i>Bb7.4</i>	<i>Bv12.19</i>	<i>Bv24.11</i>	Average
<i>Bb7.4</i>				0.145
<i>Bv12.19</i>	0.085			0.078
<i>Bv24.11</i>	0.203	0.105		0.153
<i>Bv24.12</i>	0.148	0.046	0.153	0.115
				0.090

FIGURE Legends

Figure 1. Locations of the study sites (A) and population mean frequencies of *B. variegata* alleles (B-D). In panel A, the location of the Romanian (panel B) and the Croatian (panel D: Pescenica) hybrid zones are shown on the map of Europe. Panel B shows the regional samples across Cluj county and shows the location of the intensively studied Apahida area (panel C).

Figure 2. The distribution of genotypes from the Apahida study area. The histogram shows all 954 individuals classified according to the number of *variegata* alleles that they carry at our four diagnostic marker loci (ind HI axis). Sites were grouped by their mean allele frequency into bins of width 0.05 (p mean axis). The 140 individuals not scored at four loci were classified in the range 0-8 by interpolation. See Figure 2 of MacCallum et al. (1998) for the analogous chart for the Pescenica transect.

Figure 3. Mean frequency of *B. variegata* alleles per population (p) as a function of the habitat score H in Apahida. The regression is $p = 0.46 + 0.39 H$ ($F_{1,92} = 30.2$, $P < 10^{-6}$).

Figure 4. The difference in *B. variegata* allele frequency (Δp) as a function of the difference in habitat score (ΔH) for pairs of sites under 1 km apart for a) Pescenica and b) Apahida. To ensure that the intercepts of the regressions (solid lines) are zero, the data include both possible orderings of each pair of sites.

Figure 5. Maximum likelihood estimates of F_{IS} and R for binned sites across all loci in Apahida and in Pescenica. The vertical lines show the 2-unit support limits. Also shown are cubic regressions on the overall data, fitted by least squares, see text for details. Sites are binned by allele frequency (p). Panels B and D were redrawn after MacCallum et al. (1998).

Figure 6. The percentage of toads with pure *B. bombina* marker genotypes as a function of p in each site. The line is $(1 - p)^8$, the expected frequency of animals with HI = 0 under Hardy-Weinberg proportions.

Figure 2.

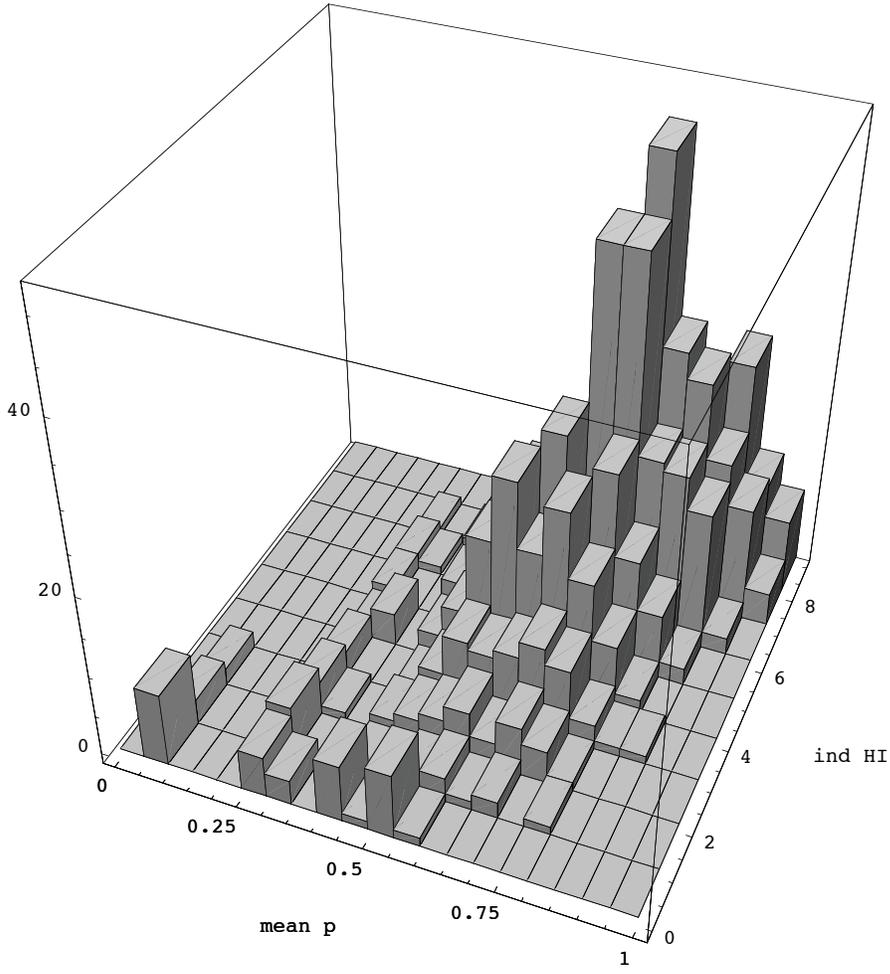


Figure 3.

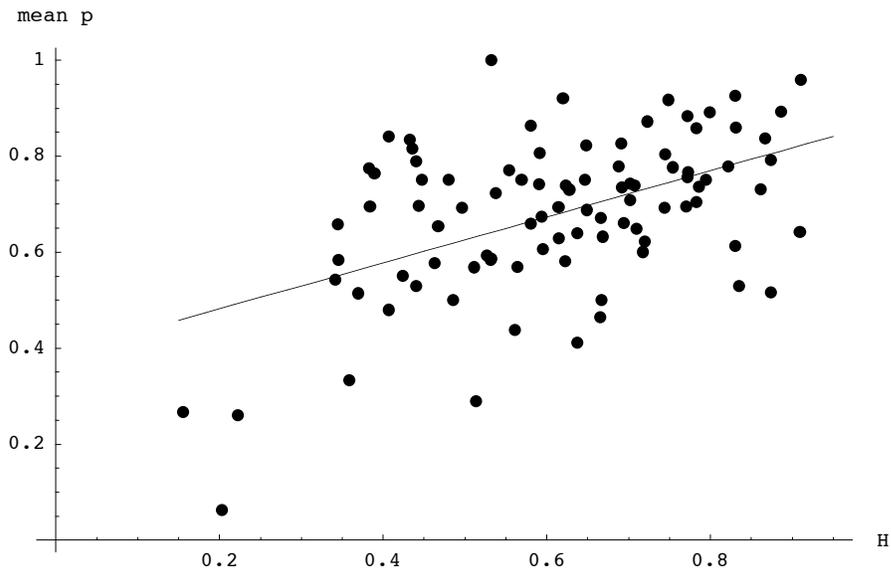


Figure 4

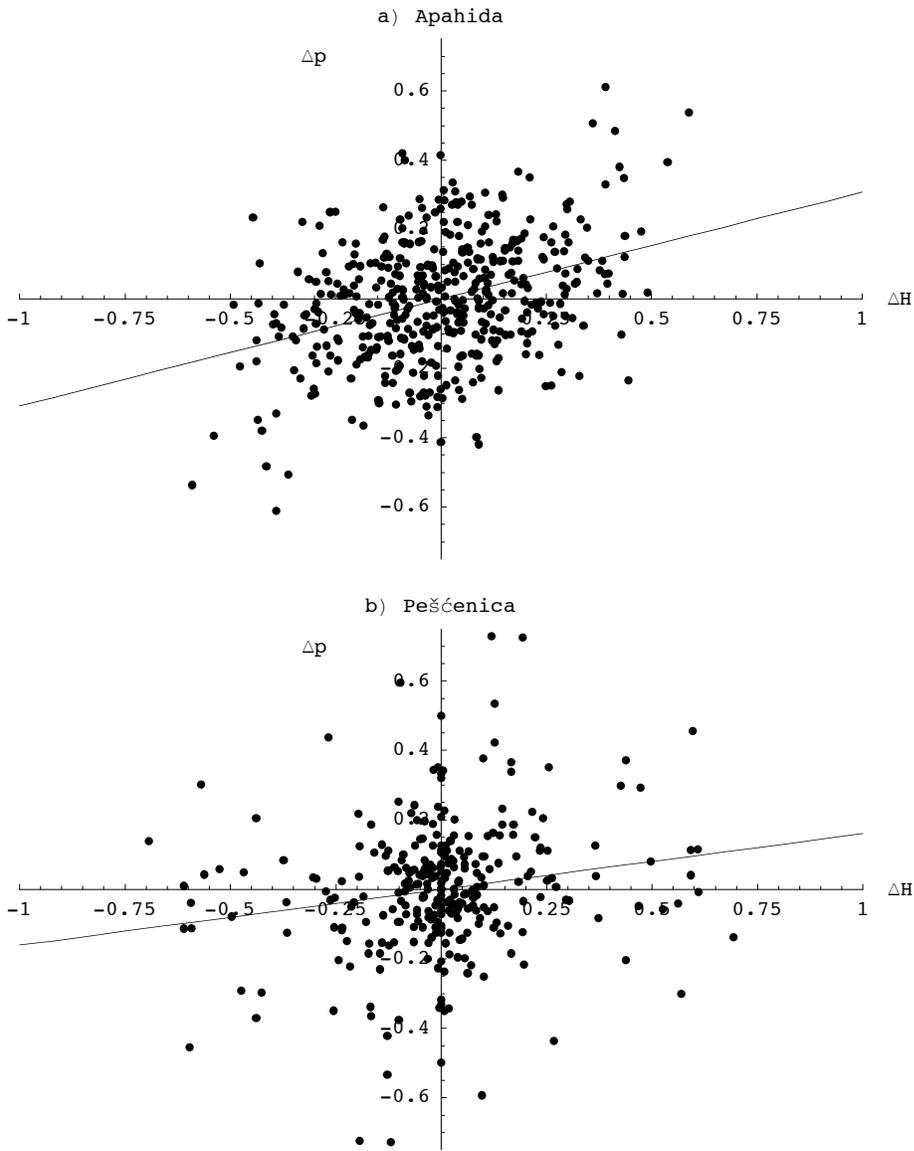
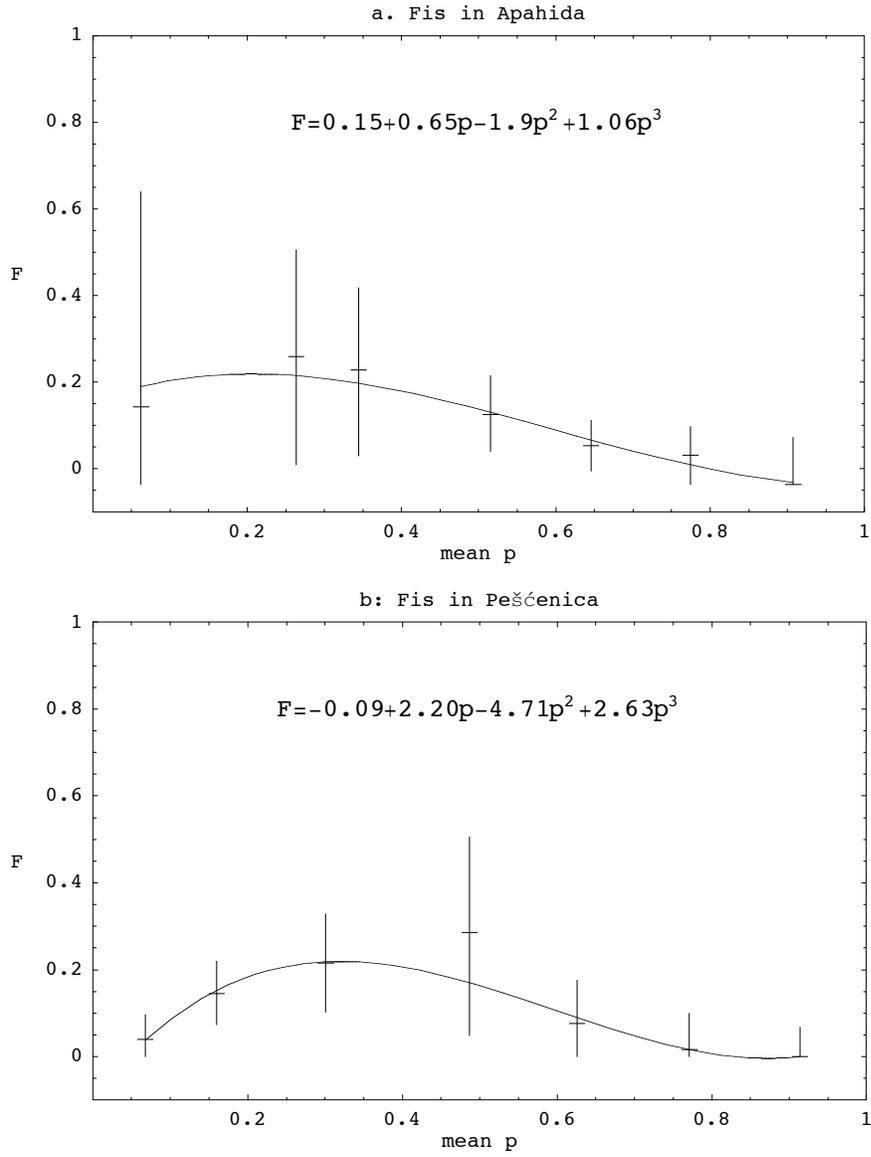


Figure 5



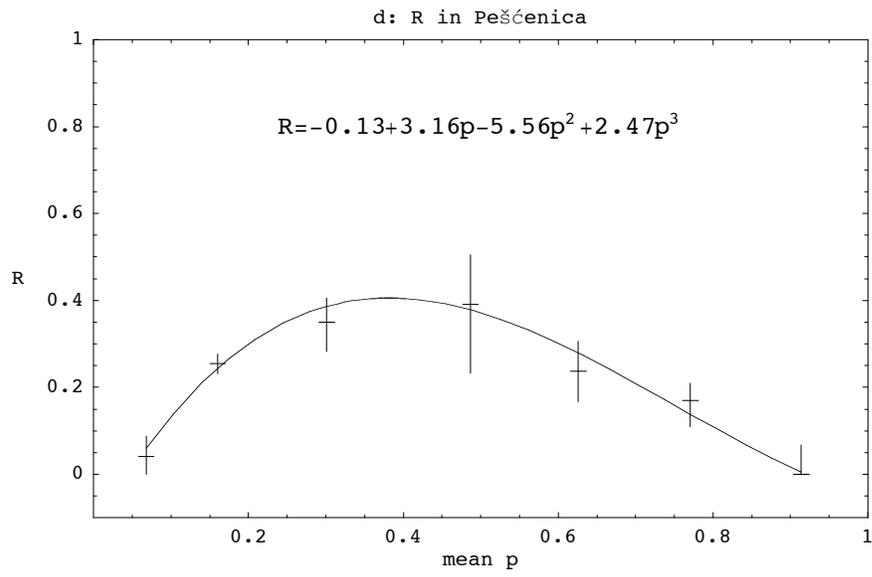
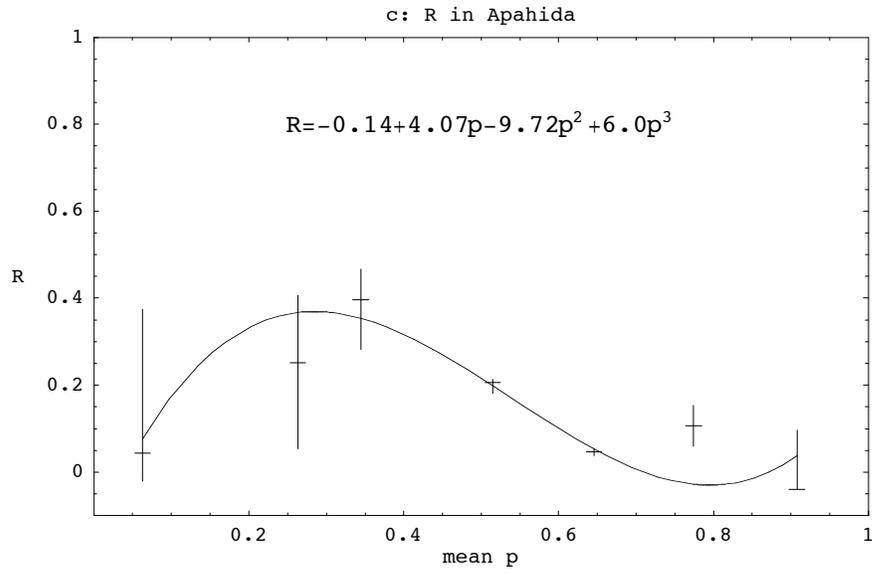
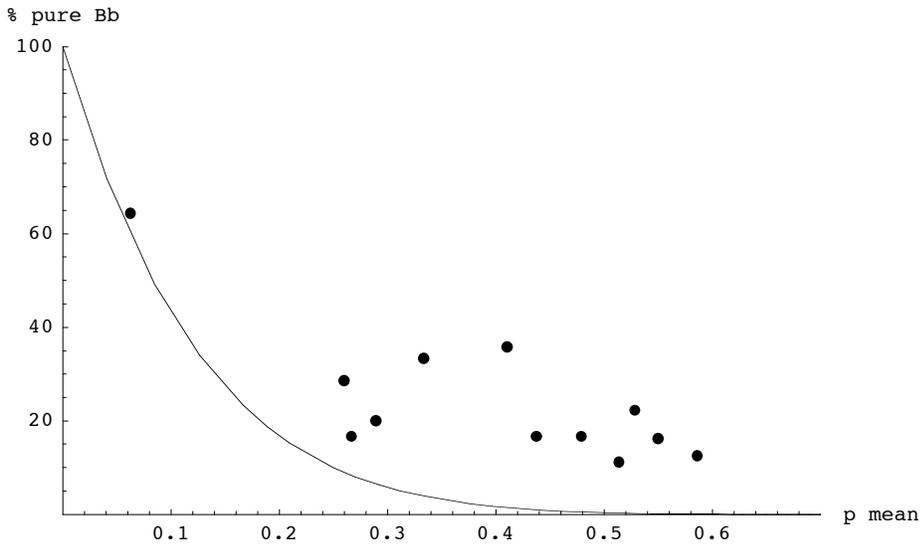


Figure 6



APPENDIX

Linkage disequilibrium generated by admixture

Immigration at a rate m between populations which differ in allele frequency by Δp at each of two loci generates pairwise linkage disequilibrium proportional to $m \Delta p^2 / r$. However, the precise relationship depends on detailed assumptions about the life cycle, the stage at which D is measured, and the forces acting on allele frequency. The formulae given by Asmussen and Orive (2000) and Barton (2000) assume that there is immigration from two source demes, such that allele frequencies remain at equilibrium. Barton and Gale (1993) consider immigration from a single deme, such that allele frequency differences and linkage disequilibrium decay at a steady rate. Here, we suppose that strong selection acts to impede gene flow, so that marker frequencies change slowly despite high rates of immigration. We must therefore find the linkage disequilibrium when immigration is countered by selection at linked loci.

Assume that individuals mate at random within demes, producing zygotes in Hardy-Weinberg proportions. Juveniles are subject to multiplicative selection, which again maintains Hardy-Weinberg proportions. Note that we need not assume that selection acts directly on the two loci in question: the change in genotype frequencies depends on the marginal fitnesses, which may be determined by selection at linked loci. Adults migrate between demes, and then mate at random to begin the next generation. Migration may represent movement from place to place, or mixing between habitats due to imperfect habitat preference. Genotype frequencies are measured in breeding adults after migration. We assume that immigrants come from a source deme with allele frequencies $p = 0$ at both loci, and that there is an equilibrium between selection and immigration; this determines the selection coefficients as a function of immigration rate.

The heterozygote deficit in this model can readily be shown to be $mp / q(1 - m)$. There is no simple explicit solution for the linkage disequilibrium as a function of immigration rate and allele frequency. However, the relation can be calculated numerically. For low migration rates, $D \sim 3 m \Delta p^2$, as expected from formulae which assume a quasi-equilibrium between immigration, and with recombination $r = 1/2$ for unlinked loci (Barton 2000). For higher migration rates, disequilibrium remains close to this expectation when allele frequency is low, but falls below it for higher allele frequencies. In the extreme case, $D \sim 0.5 m \Delta p^2$ for $m = 0.5, p = 0.5$. Since we will be concentrating on migration rates less than $m \sim 0.3$, and allele frequencies in the range 0.2 to 0.4, we take the relation as $D \sim 2 m \Delta p^2$.