

Host choice promotes reproductive isolation between host races of the larch budmoth *Zeiraphera diniana*

I. EMELIANOV, F. SIMPSON, P. NARANG* & J. MALLET

Galton Laboratory, Department of Biology, University College London, London, UK

Keywords:

adult host choice;
hybridization;
speciation;
Zeiraphera.

Abstract

The chances for sympatric speciation are improved if ecological divergence leads to assortative mating as a by-product. This effect is known in parasites that find mates using host cues, but studies of larch- and pine-feeding races of the larch budmoth (*Zeiraphera diniana*, Lepidoptera: Tortricidae) suggest it may also occur when mate attraction is via sex pheromones that are independent of habitat. We have previously shown that females releasing pheromones on or near their own host attract more males of their own race than if placed on the alternative host. This host effect would enhance assortative mating provided adults preferentially alight on their native hosts. Here we investigate alighting preferences in natural mixed forest using a novel likelihood analysis of genotypic clusters based on three semidiagnostic allozyme loci. Both larch and pine females show a realized alighting preference for their own host of 86%. The equivalent preferences of males were 79% for the larch race and 85% for the pine race. These preferences are also detectable in small-scale laboratory experiments, where alighting preferences of larch and pine races towards their own hosts were, respectively, 67 and 66% in females and 69 and 63% in males. Pure larch race moths reared in the laboratory had alighting choice similar to moths from natural populations, while hybrids were intermediate, showing that alighting preferences were heritable and approximately additive. The field estimates of alighting preference, coupled with earlier work on mate choice, yield an estimated rate of natural hybridization between sympatric host races of 2.2–3.8% per generation. Divergent alighting choice enhances pheromone-mediated assortative mating today, and is likely to have been an important cause of assortative mating during initial divergence in host use. Because resources are normally ‘coarse-grained’ in space and time, assortative mating due to ecological divergence may be a more important catalyst of sympatric speciation than generally realized.

Introduction

Sympatric speciation has been regarded as unlikely because it is difficult to understand how genes for divergent ecology become correlated with genes for mate

Correspondence: James Mallet, Galton Laboratory, Department of Biology, University College London, 4 Stephenson Way, London NW1 2HE, UK. Tel.: +44 207 679 7412; fax: +44 207 383 2048; home page: <http://abacus.gene.ucl.ac.uk/jim/>

*Present address: Institute of Psychiatry, King's College London, London SE5 8AF, UK.

choice in the face of gene flow and recombination (Felsenstein, 1981; Futuyma, 1998). The chances of speciation improve sharply when an individual's choice of habitat enhances the probability of mating with other individuals making the same choice (Colwell, 1986; Butlin, 1990; Rice & Hostert, 1993; Bush, 1994; Dieckmann & Doebeli, 1999). Host races of plant parasites, such as the apple maggot *Rhagoletis pomonella*, provide some of the best examples. Strong differences in adult host choice between apple and hawthorn host races of *Rhagoletis* lead inevitably to assortative mating because

the host plant provides the only long-range mate-finding cues (Feder *et al.*, 1994; Filchak *et al.*, 2000).

However, the generality of this kind of adaptive speciation remains questionable (Futuyma, 1998; Tregenza & Butlin, 1999). Mate-finding in many animals uses habitat-independent long-range cues, such as song (birds, grasshoppers, frogs), or volatile pheromones (moths and many other insects), and speciation seems correspondingly less likely to arise via ecological divergence. The work described here investigates a moth with a habitat-independent pheromone-based mating system. Based on this study, we argue that correlations between habitat choice and mate choice may be more universal than generally appreciated.

The larch budmoth *Zeiraphera diniana* has sympatric host races that feed on European larch (*Larix decidua*) and Cembran pine (*Pinus cembra*) in mountainous areas of Europe. The larch race is renowned for outbreaks in larch forests near the treeline in the Alps. Its population density may fluctuate as much as 10⁵-fold between outbreak and crash on an 8–9 years cycle, and outbreaks are accompanied by long-range mass migrations; dispersal during these events may connect populations hundreds of kilometres apart (Baltensweiler *et al.*, 1977; Baltensweiler & Rubli, 1999). Larch and pine races interbreed freely in captivity and display only weak assortative mating in the laboratory (Drès, 2000). The near-absence of short-range assortativity, combined with high dispersal ability of the larch budmoth is likely to lead to hybridization in the wild. In spite of the potential for hybridization, there are many differences between host races, including female sex-pheromones and male pheromone response (Baltensweiler *et al.*, 1978), population dynamics (Baltensweiler *et al.*, 1977), genetic markers (Emelianov *et al.*, 1995) and a number of host use traits, including larval colour, timing of egg hatch (Day, 1984), and host choice during oviposition (Bovey & Maksymov, 1959).

Divergent sex pheromones cause assortative mating of *Z. diniana* host races (Baltensweiler *et al.*, 1978), but we have recently shown that the degree of assortativeness also depends on the host plant. The probability of cross-attraction can be greatly increased if females ‘call’ (release sex pheromones) from the other race’s host, or from within neighbourhoods with a preponderance of the other race’s host (Emelianov *et al.*, 2001). If this host effect is to enhance assortative mating, there must be a tendency for females and males of each host race to congregate on or around their own host species. In laboratory tests, larch and pine race females tend to lay about 70–90% of their eggs on their own hosts (Bovey & Maksymov, 1959), but those results were obtained in small-scale experiments under unnatural conditions. Furthermore, host choice for oviposition does not necessarily correlate with a tendency of females to alight and call for males on a particular host, and no data at all exist for alighting preference of males.

In this paper, we use a novel likelihood-based analysis of genetic population structure to calculate the fraction of adults of *Z. diniana* on their own and alien hosts in natural mixed forests, and from this we estimate the alighting probabilities of both sexes in the wild. We also study alighting choice, as well as its inheritance, in the laboratory. Finally, we estimate the overall hybridization rate between host races in nature by combining the probability that females alight on an alien host, the previously estimated probability of pheromone cross-attraction (Emelianov *et al.*, 2001), and mating preferences at close range (Drès, 2000).

Methods

Study sites

Field collections were made in 1998 and 1999 near Bever (9°53'E, 46°33'N, alt. 1708 m) and Pontresina (9°54'E, 46°30'N, alt. 1805 m), which are situated about 5 km apart in side valleys of the Upper Engadine valley in eastern Switzerland. Both sites are forested with an approximately equal mix of European larch and Cembran pine, and are inhabited by sympatric pine- and larch-feeding populations of the larch budmoth.

Field collections

Larvae

Fourth and fifth instar larvae of *Z. diniana* were collected in 1998 and 1999 from both larch and pine. Small samples collected in 1998 were used to determine local frequencies of allozymes within the larvae of each host race, in order to check the results from likelihood analysis of the adults from the same year. Larvae collected in 1999 were reared to adulthood, and were then used in the laboratory study of host alighting preference.

Adults

Adults were collected from both sites in 1998 in parts of the forest where the ratio of larch to pine trees was close to 1:1. In 1998, both host races were common, based on our larval survey earlier the same year. We shook lower branches of randomised nonadjacent trees (10 of each host species in Pontresina, 30 of each in Bever) in the early morning when air temperature was below 8 °C, the activity threshold of *Z. diniana* moths. Moths that fell from the branches were collected on cotton sheets placed underneath the trees, frozen in liquid N₂ and shipped to the laboratory for allozyme analysis.

Allozyme electrophoresis

The host races can be distinguished at semidiagnostic allozyme markers. These are: sex-linked isocitrate dehydrogenase (*Idh*), and two unlinked autosomal loci,

malate dehydrogenase (*Mdh*) and phosphoglucomutase (*Pgm*; see Emelianov *et al.*, 1995). In 1998, adults reared from larvae and adults collected directly from host trees were scored for these three allozymes using cellulose acetate electrophoresis. In our host race analysis, we were interested only in alleles that were diagnostic or that differed significantly in frequency, so we lumped rarer alleles with the common ones of similar mobility, leaving two major alleles at each locus: for *Idh*, *f* (corresponding to allele 2.40 of our earlier study) and *s* (1.00), for *Mdh*, *f* (5.50) and *s* (1.00), and for *Pgm*, *f* (1.12) and *s* (1.00); see Emelianov *et al.* (1995) for details of electrophoretic conditions and alleles.

Analysis of host alighting preference in the field

The host races differ strongly in the frequency of allozymes, although none are fixed (Table 1, see also Emelianov *et al.*, 1995). Using allozymes, individual females can be identified to a particular host race with an estimated 4% (larch race) to 10% (pine race) error. In males, however, the error is more than an order of magnitude lower (Emelianov *et al.*, 2001) because of the extra allelic information from *Idh* (females are hemizygous for the sex-linked *Idh* locus in *Zeiraphera* because, as in all Lepidoptera, they are the heterogametic sex). Since we are interested in population patterns, rather than in identifying individuals *per se*, and in behaviour of both males and females, rather than just of males as in the earlier paper (Emelianov *et al.*, 2001), we here employ a more powerful likelihood analysis of multilocus allozyme genotypes to estimate fractions of host races among the adults on each host, and to estimate host alighting preference of each host race.

Multilocus genotypes within any sample may fall into one or more 'genotypic clusters' (Mallet, 1995; Feder, 1998), and it is possible to test the null hypothesis that only a single genotypic cluster is present (consisting of a single set of allele frequency parameters, with each allele independent of all others). Rejection of the null hypothesis implies that more than one genotypic cluster, each with a separate set of allele frequency parameters, explains disequilibria and Hardy-Weinberg deviations

in the data better than a single genotypic cluster. In the present case of two host races, there are only two sets of allele frequencies, and the statistical analysis carried out allows estimation of relative abundances and allele frequencies of each host race, and their support limits. Feder (1998) defines species on the basis that 95% of individuals can be identified to a particular genotypic cluster. However, in genotypic data of this kind the probability that each individual belongs to a particular genotypic cluster will depend strongly on the amount of genotypic data obtained: for instance, the analysis of three-locus genotypic data obtained here leads to the rather strange result under Feder's criterion that larch and pine males of *Zeiraphera diniana* fall into separate host-associated species, whereas hemizygous females from larch and pine are merely host races. Instead, the question we ask in our analysis is similar to that of the 'partitioning' analysis of Pritchard *et al.* (2000): we ask whether more than one genotypic cluster (i.e. separate populations differing in allele frequency), as well as what allele frequencies in each cluster and what relative frequency of the clusters, can best explain genotypic patterns in the data.

To estimate the fraction of individuals that are larch race beaten from larch (λ_L) and from pine (λ_P) trees, we make the simplifying assumption that there are no hybrids between the two forms. Because hybridization is likely to occur in nature (and is indeed estimated from the combined behavioural data below), this assumption is strictly incorrect. Nonetheless, for the purposes of our study the assumption is adequate because the hybridization rate is at most only a few per cent (see Results).

The estimation procedure is as follows. Assuming the alleles are independent within each host race (i.e. Hardy-Weinberg and linkage equilibrium within each host race – as found: see Emelianov *et al.*, 1995), the probabilities of observing each genotype of a given host race can be calculated given the allele frequencies for *f* alleles at *Idh*, *Mdh* and *Pgm* (p_L , q_L , r_L , respectively, for the larch race, and p_P , q_P , r_P for the pine race). For example, the expected genotypic fraction g_{jhP} among pine females of the *Idh*-*Mdh*-*Pgm* genotype $j = f-fs-ss$ on host h is $p_P 2q_P (1 - q_P) (1 - r_P)^2$. Similarly, the expected fraction g_{khL} of

Table 1 Frequencies of 'fast' alleles in larch and pine host races collected in 1998.

Locus-allele	Frequencies [support limits]		Frequencies estimated via likelihood	
	Larch race larvae (no. of genomes)	Pine race larvae (no. of genomes)	Larch race adults (no. of genomes)	Pine race adults
<i>Idh-f</i>	(29) 0.552 [0.369, 0.726]	(25) 0.960 [0.832, 0.998]	(645) 0.530	0.893
<i>Mdh-f</i>	(62) 0.226 [0.132, 0.342]	(42) 0.976 [0.897, 0.999]	(954) 0.152	0.990
<i>Pgm-f</i>	(58) 0.172 [0.089, 0.285]	(42) 0.310 [0.182, 0.461]	(954) 0.222	0.218

the genotype $k = ss-ss-sf$ among larch males is $(1 - p_L)^2 (1 - q_L)^2 2r_L(1 - r_L)$. For a particular sex, the expected fractions of genotype i for larch (g_{ihL}) and pine (g_{ihP}) races can now be combined to give the overall expected fraction for each genotype on host h :

$$g_{ih} = \lambda_h g_{ihL} + (1 - \lambda_h) g_{ihP},$$

where λ_h represents the mixture fraction in that sex that is larch race on tree species h , and $(1 - \lambda_h)$ is the fraction that is pine race. The likelihood of this genotype is then given by $g_{ih}^{n_{ih}}$ where n_{ih} is the number of individuals with genotype i actually found on host h . The overall log likelihood for λ_h is then $\sum n_{ih} \log_e(g_{ih})$.

So far, we have developed a method to estimate the fractions of individuals on a given tree species that are larch race (λ_h) from the field-collected adults, but we are here more interested in the fractions of individuals of host race r that alight on a particular host tree (μ_r), as a measure of the realized host alighting preference of that host race (see also Feder *et al.*, 1994; who experienced a similar estimation problem in a similar context). We should remember that μ_r will be a 'realised' rather than actual preference, because the actual preference of each individual interacts with the spatial structure of host trees through encounter rate. Assuming that local population densities of larch and pine race are equal, the preference of larch moths for alighting on larch is $\mu_L = \lambda_L N_L / (\lambda_L N_L + \lambda_P N_P)$, while the preference of pine moths for alighting on larch is $\mu_P = (1 - \lambda_L) N_L / [(1 - \lambda_L) N_L + (1 - \lambda_P) N_P]$, where N_L is the total number of individuals found on larch, and N_P is the total number found on pine. Conversely, the fraction of larch race choosing pine is therefore $1 - \mu_L$, and the fraction of pine race choosing pine is similarly $1 - \mu_P$. In order to use the likelihood analysis above to estimate the realized preferences μ_r directly, we rearrange to find explicit solutions for λ_h in terms of μ_r :

$$\lambda_L = \mu_L [N_L(1 - \mu_P) - N_P \mu_P] / N_L(\mu_L - \mu_P),$$

and

$$\lambda_P = (1 - \mu_L) [N_L(1 - \mu_P) - N_P \mu_P] / N_P(\mu_L - \mu_P).$$

The overall likelihoods for μ_L , μ_P , p_L , q_L , r_L , p_P , q_P , r_P were then maximized using the 'Solver' algorithm in a spreadsheet program over both pine and larch tree samples for each site and sex. For the purposes of this analysis, the allele frequencies p_L , q_L , r_L , p_P , q_P , and r_P were assumed the same in males and females for each host race within any site; these allele frequencies for each host race are estimated as 'nuisance parameters' during the procedure, and can be compared with the allele frequencies found in larvae from the same year.

We used the differences in log likelihood ($\Delta \log_e L$) as a measure of inference in distinguishing realized host alighting preference between samples. In large samples $G = 2\Delta \log_e L$ is distributed as a χ^2 (Edwards, 1972; Rohlf & Sokal, 1981). To assess the error of estimation, we use

support limits, defined as parameter values with log_e likelihood values 2 units lower than the maximum; these are asymptotically equivalent to 95% confidence limits (Edwards, 1972), and were obtained by manually varying the parameter of interest, and re-maximizing the likelihood while leaving all other parameters free to vary, until the $\log_e L$ was 2.00 units below the maximum. This procedure also includes a test of the null hypothesis for a single genotypic cluster: there is evidence for presence of both host races within any sample when support limits obtained in this way include neither $\lambda_h = 0$ nor $\lambda_h = 1$ (or equivalently, neither $\mu_r = 0$ nor $\mu_r = 1$).

Measurement of host alighting preference in the laboratory

Virgin females and males were placed individually in 1-L clear plastic cages containing a pair of similar 10 cm cuttings of European larch and Cembran pine planted in clean wet sand 5 cm apart. The experiment was performed in a controlled-climate room ($17 \pm 1^\circ\text{C}$ with 16 h light/8 h dark cycle with 30 min dawn and 60 min twilight, consisting of gradual transition from complete darkness to full brightness, and vice-versa). The settling position of each individual was recorded every 30 min for 3 h before dark and 2 h after dark, the period of maximal pheromone activity in the larch budmoth (Emelianov *et al.*, 2001) for at least two nights. For the first 3–5 min of each 30 min observation cycle, the position of each moth was recorded: on larch, on pine, or on the inner surface of the cage (= 'null'). The cage was then tapped gently until the moths flew up and alighted on the inner surface of the cage. During the remaining 25–27 min moths were free to initiate a flight and choose a new alighting position, which was in turn recorded at the beginning of the next observation cycle.

After completing these experiments, these same adults were crossed in all four possible directions. Fecundity of pine females was low, and few progenies from pine race mothers were obtained. Therefore, only larch \times pine and larch \times larch crosses could be used for estimation of alighting in laboratory-reared moths. All individuals were reared on artificial diet from third instar to pupation to control as far as possible for the effect of larval environment. Early instar larvae do not survive on artificial diet, and were therefore reared from first to second instar on larch shoots, whatever their genotype. Host alighting preferences of the resultant adult males (12 larch \times larch families and 21 larch \times pine families) and females (12 larch \times larch, and 19 larch \times pine families) were tested in the same way as for their parents. There was some evidence for heterogeneity between broods within each cross type, suggesting correlations among individuals within broods. Thus the individual was not an appropriate level of analysis for tests involving these broods. Therefore, for all analyses

involving broods (i.e. larch \times larch or larch \times pine), the response data of the entire brood, consisting of the data for all brood members, took the place of the data for the 'individual' in the statistical analyses described below. This method most efficiently uses the data from each individual to compare between brood types, while still recognizing that some of the variation between brood types is due to uninteresting brood-to-brood 'error' variation.

Beta-binomial tests of laboratory alighting preference

Estimates of alighting differences between strains (i.e. host races or F_1 hybrids) were obtained using a likelihood method, as follows. Choices made by each moth during individual observation cycles can be assumed to follow a binomial distribution, but average choice may differ between individuals within each strain. It is therefore important to allow for variation between individuals within strains when testing for differences between strains. Assuming that the former follows a beta distribution, we obtain a beta-binomial distribution for host alighting choice overall. The beta distribution is used here because it can be parameterized to fit a variety of data distributed between 0 and 1, whether bimodal, skewed, or simple binomial-like. Thus data that has a variance greater than a simple binomial is easily accommodated by the betabinomial distribution. Likelihood analysis of strain-specific alighting preference was carried out using the program BETABINO written by Ziheng Yang (see appendix in Jiggins *et al.*, 2001). Using this method, we examined (1) choice between alighting on larch vs. alighting on pine shoots (host alighting preference), and (2) choice between alighting on a shoot vs. alighting on the wall of the container (null-choice

probability). BETABINO also calculates standard errors for each parameter from the curvature of the likelihood surface near the maximum likelihood value; if there is heterogeneity between individuals within a strain, this standard error will be larger than the standard error expected under a simple binomial model.

Results

Host alighting preference in the field

Overall allele frequencies estimated from larvae and adults collected from larch and pine trees in 1998 are shown in Table 1. Allele frequencies for the host races estimated from larval populations and those estimated by likelihood analysis of adults resting on each tree species are not significantly different (Table 1). No significant differences in the alighting preferences reported below occur whether we use the larval allele frequencies to estimate adult alighting parameters, or, as here, estimate allele frequencies and alighting preferences within a single analysis of adult genotypes. Genotypes of adult males and females collected from different hosts in Bever and Pontresina are shown in Appendices 1 and 2, respectively. The fractions of larch race among moths found on either larch or pine (λ_h) and realized preferences for larch (μ_L) of each host race and their support limits are shown in Table 2. Realized alighting preferences of moths differed strongly and highly significantly between races for both sexes at both sites (Fig. 1, Table 2); however, allele frequencies and realized preferences (μ_r) showed only slight differences between sites (Fig. 1, Tables 1 and 2) which were marginally significant in males ($G_8 = 15.76$, $P < 0.05$), but not in females ($G_8 = 9.78$, ns). Within the larch race, females tended to

Table 2 Alighting preference estimated from adults collected in the field.

Site	Sex	Fraction of larch race moths	Parameter value (support limits)	Realized preference for larch	Parameter value (support limits)
Bever	Male	On larch		Of larch race	
	Female	λ_L	0.76 (0.60, 0.88)	μ_L	0.77 (0.64, 0.87)
Pontresina	Male	λ_L	0.91 (0.85, 0.95)	μ_L	0.84 (0.79, 0.90)
	Female	λ_L	0.60 (0.34, 0.85)	μ_L	0.83 (0.61, 0.97)
Overall	Male	λ_L	0.84 (0.63, 0.96)	μ_L	1.00 (0.89, 1.00)
	Female	λ_L	0.71 (0.57, 0.83)	μ_L	0.79 (0.68, 0.88)
Bever	Male	On pine		Of pine race	
	Female	λ_P	0.90 (0.84, 0.94)	μ_P	0.86 (0.81, 0.91)
Pontresina	Male	λ_P	0.11 (0.06, 0.19)	μ_P	0.12 (0.06, 0.18)
	Female	λ_P	0.24 (0.14, 0.32)	μ_P	0.14 (0.07, 0.21)
Overall	Male	λ_P	0.08 (0.01, 0.24)	μ_P	0.22 (0.10, 0.33)
	Female	λ_P	0.00 (0.00, 0.12)	μ_P	0.16 (0.05, 0.29)

This table shows larch-based maximum likelihood estimates: the fraction of larch moths among total individuals collected either from larch (λ_L) or from pine (λ_P), and realized preference of larch and pine race individuals (μ_L , μ_P) for larch. The fraction of pine race moths on each host h is given by $(1 - \lambda_h)$, and the realized preference of host race r for pine is given by $(1 - \mu_r)$.

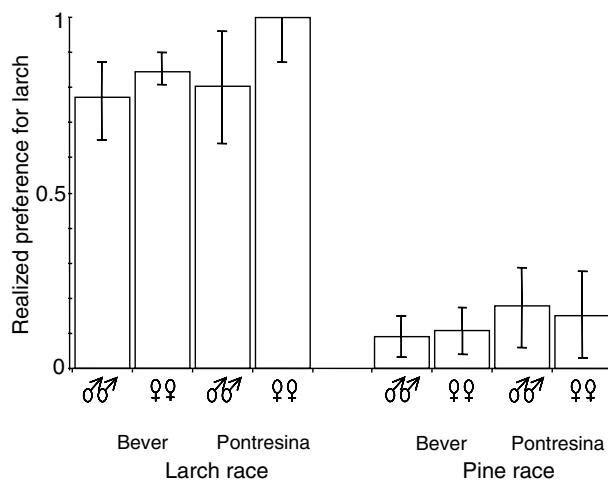


Fig. 1 Realized alighting preference for larch in the field (μ_L for larch race, μ_P for pine race). Error bars show two-unit support limits (asymptotically equivalent to 95% confidence limits) around the maximum likelihood estimate.

show a stronger realized alighting preference for larch than males at both Bever and Pontresina, and overall (see support limits in Table 2), but in the pine race, females did not differ significantly from males. In spite of the strong preferences, however, in almost all cases there was highly significant evidence for both host races being present in samples, and that preference for the native host was nonabsolute. The only exception is that no larch females at Pontresina were found on pine ($\lambda_P = 0$), and that therefore larch females were estimated to have an absolute preference for larch ($\mu_L = 1$) at that site.

Alighting behaviour of pure races and hybrids in the laboratory

Despite considerable overlap (Fig. 2a,b), wild females of larch and pine races also showed a significant preference for alighting on their own host in the laboratory ($G_1 = 34.29, P < 0.001$) (Fig. 3). The same was also true for males ($G_1 = 19.97, P < 0.001$) (Fig. 3). Our analysis and experimental design controlled for potential family effects and behavioural effects of sex pheromones, thus the host-specific alighting behaviour of males is due to an attraction to the host itself. After one generation of laboratory rearing, larch females and, to a lesser extent, larch males retained the larch preference, whereas hybrids showed no preference; they behaved in a manner intermediate between pure races (Figs 2c,d and 3). Laboratory-reared larch females did not differ from wild-collected females ($G_1 = 0.028, \text{ns}$) but were significantly different from pine females ($G_1 = 24.6, P < 0.001$) and from hybrid females ($G_1 = 5.00, P < 0.05$). Similarly, laboratory males did not differ from

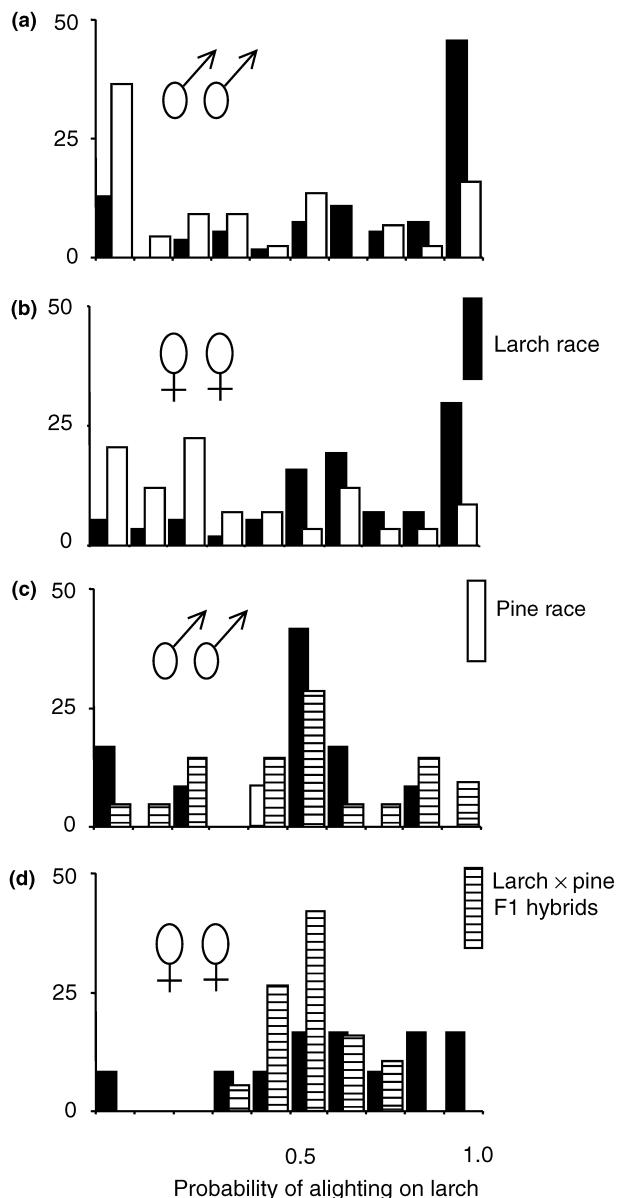


Fig. 2 Host alighting behaviour in the laboratory. Distribution of individual host alighting choice of adults collected as larvae – (a) (males), (b) (females). Distribution of family mean host alighting choice of laboratory-reared larch \times larch progeny and F_1 hybrid broods – (c) (males), (d) (females).

wild larch males ($G_1 = 2.12, \text{ns}$), although they were also not significantly different from hybrid males ($G_1 = 1.92, \text{ns}$). ‘Null’ choices occurred at a high rate during our experiments, consistently above 50% of observations. Larch females did not differ significantly in terms of null-choice probability from pine females, and larch males did not differ in this respect from pine males. However, pure-race males made significantly more null choices (83%) than females (62%) (Fig. 4; $G_1 = 97.26, P < 0.001$).

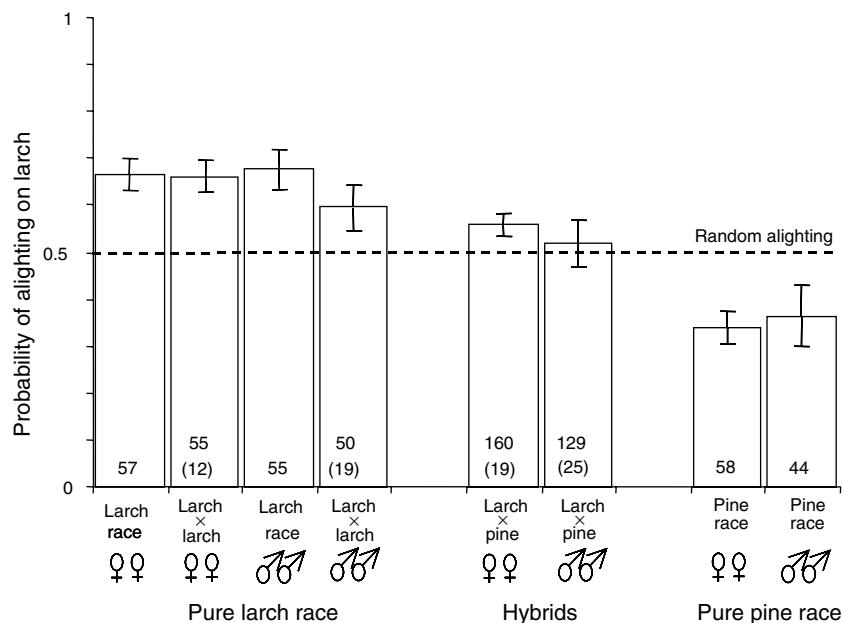


Fig. 3 Host alighting behaviour in the laboratory: beta-binomial mean probabilities of alighting on larch. Error bars are the standard errors estimated from the curvature of the likelihood surface. Numbers at the base of each column are numbers of individuals (numbers in parentheses are nos. of broods, as used in statistical analyses and estimation of standard errors).

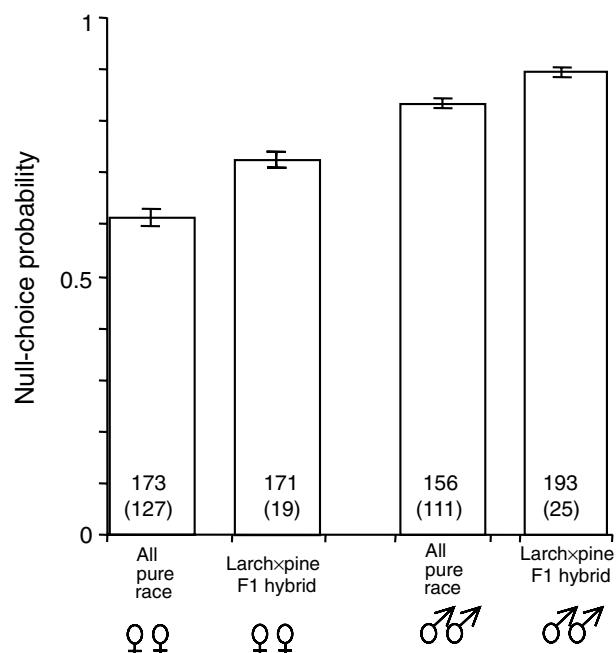


Fig. 4 Beta-binomial null-choice probabilities in pure race and hybrid males and females. Error bars are standard errors estimated from the curvature of the likelihood surface. Numbers at the base of each column are numbers of individuals (numbers in parentheses are nos. of independent broods and individuals, as used in statistical analyses and estimation of standard errors).

Hybrids made significantly more null choices than pure-race moths (Fig. 4), 72% for females ($G_1 = 19.97$, $P < 0.001$), and 89% for males ($G_1 = 18.52$, $P < 0.001$).

Discussion

Host choice, alighting behaviour, assortative mating, and speciation via pleiotropy

From this study it is clear that (i) adult females of larch and pine host races of *Zeiraphera diniana* prefer to alight on their native host in both field and the laboratory; (ii) the same is true, perhaps to a somewhat lesser extent, for pure-race males; (iii) laboratory hybrids of both sexes have no significant preference and choose larch or pine alighting substrates approximately at random. Our results show that host alighting preference is most strongly expressed with real host trees in the field, but is also demonstrable even in small-scale laboratory tests with small twigs of host material. The laboratory experiments indicate that the strong interracial differences in host alighting observed in the field are heritable rather than nongenetic, for instance inability of moths to disperse from hosts on which they developed as larvae.

It has been long recognized that a sufficiently strong tendency of adults to alight on their larval hosts can lead to adaptive speciation via assortative mating of host-specific populations when the host is used as a mating rendezvous (Tregenza & Butlin, 1999). Empirical studies have demonstrated this kind of assortative mating in tephritid (Craig *et al.*, 1993; Feder *et al.*, 1994) and agromyzid (Tavormina, 1982) flies, chrysomelid beetles (Kreslavsky *et al.*, 1981), and aphids (Via, 1999). In all of these cases, males are thought to rely on host cues in order to find mates. *Zeiraphera* differs in that the host is not used as a direct cue for mate finding. In this species, as well as in the majority of Lepidoptera, mate-finding is

via long-range sex-pheromones whose chemical composition does not depend on host (although some effects of host plants on mating have been found; see Landolt & Phillips, 1997).

The long-range attraction of *Zeiraphera* males to calling females is usually highly assortative. This is mainly caused by pheromone differences, but we have shown that the degree of assortative attraction depends strongly on the host tree from which females call, and on the ratio of host and nonhost trees in the neighbourhood of the caller: cross-attraction is reduced when females call from a tree of their own host species, and from neighbourhoods where their own host is abundant (Emelianov *et al.*, 2001). However, in order for this host effect to result in assortative attraction in the field, adults of both sexes must preferentially settle on or near their own hosts. Divergent host preference is of course expected in females, whose choices during egg-laying are likely to have direct fitness effects on their offspring. Host-specific alighting preferences in males, on the other hand, seem somewhat less likely. Nonetheless, the interaction between cross-attraction and the host from which females call is best explained by a propensity of males to congregate around their native hosts (Emelianov *et al.*, 2001). In this paper, genetic analyses of field populations, together with laboratory choice experiments, show that males do indeed express host-specific alighting behaviour. The host-associated alighting preferences we have uncovered in both males and females, together with the earlier results on host-dependent assortative attraction, thus show how sympatric ecological divergence can enhance assortative mating. We have little evidence on the particular question of whether host races of *Z. diniana* initially diverged in sympatry. Instead our results in *Zeiraphera* are more important in showing that this sympatric route to assortative mating is likely in general.

Inheritance of host alighting behaviour in the laboratory

Inheritance of host alighting behaviour of both males and females of phytophagous insects has not to our knowledge been investigated previously. Post-alighting oviposition behaviour has been studied, and may show either additive (Pennacchio *et al.*, 1994; Messina & Slade, 1997; Sezer & Butlin, 1998) or dominant (Lu & Logan, 1995) patterns of inheritance. We here show that differences in alighting preference between host races are heritable, and that the preferences of hybrids is intermediate between parental races, suggesting approximately additive inheritance. These results pave the way for further research addressing many important questions. For example: How many genes or regions of the genome are involved in control of the host alighting behaviour? Is the host alighting behaviour of males and females controlled by the same loci? Are the genes involved in

oviposition preference the same as those involved in alighting preference (as seems likely)?

Probability of hybridization between host races of *Z. diniana*

If the host alighting behaviour of female *Z. diniana* were random, then, given the probabilities of cross-atraction, that is 3.3 and 37.7% for larch females calling, respectively, from larch and pine, 9.1 and 6.3% for pine females calling from pine and larch (Emelianov *et al.*, 2001), the average hybridization probability for larch and pine females would be, respectively, 20.5 and 7.7%. Interestingly, this high probability of hybridization would exist despite strongly differentiated pheromone-signalling systems. The actual probability of alighting on an alien host in the field, however, is here estimated as 14% for females of both larch and pine races (Table 2). Assuming these alighting preferences are typical, we can combine our data on long-range cross-atraction measured at Bever (Emelianov *et al.*, 2001) and estimate the overall probability of hybridization between host races. The high fidelity of the females' host alighting choice ensures that larch females ($\mu_L = 0.86$) would be approached by males 91.9% of whom are larch race; i.e. $0.86 \times (1 - 0.033) + 0.14 \times (1 - 0.377)$. Thus, assuming random mating at close range, the probability that larch females hybridize is 8.1%. Similarly, for a pine female ($\mu_P = 0.86$), the probability that she mates with a pine male is 91.3% (i.e. $0.86 \times 0.909 + 0.14 \times 0.937$) and her probability of hybridization is 8.7%, again assuming no close-range assortative mating. (Note: the very slight differences between estimates given in Emelianov *et al.* (2001) and here are due to a minor change in assumptions adopted in this paper. We here relax our former assumption that allele frequencies are the same as those found in larvae).

However, mating behaviour at close range is not completely random. Laboratory mate choice experiments show that in 1:1 mixed populations, the chances that larch females mate with larch and pine males are 0.800 and 0.200, respectively, while the equivalent probabilities for pine females are 0.296 and 0.704 (Drès, 2000). Using the figures above, the overall probability bias towards assortative mating for larch females will be $(0.919 \times 0.800) : (0.081 \times 0.200) = 0.7352 : 0.0162$. Thus, the probability that larch females mate with pine males should be $0.0162 / (0.0162 + 0.7352) \approx 2.2\%$. Similarly, the overall assortative mating bias for pine females will be $(0.913 \times 0.704) : (0.087 \times 0.296) = 0.6427 : 0.0258$, and the probability of hybridization between pine females and larch males should be $0.0258 / (0.6427 + 0.0258) \approx 3.8\%$. On average, hybridization rates are expected to be about 3%.

These estimates are strongly dependent on a series of modelling assumptions, as well as on the particular details of where and when we carried out our field studies. For this reason, we do not attempt to estimate

the support limits on our estimates – factors other than statistical errors are probably more important overall. For instance, we assume that: (1) mating takes place in populations with equal densities of females of each host race; (2) relative densities of larch and pine males are as found in Bever in 1997 (Emelianov *et al.*, 2001); (3) larch and pine host trees are evenly scattered and at equal density; (4) alighting behaviour of virgin females and males is similar to that of adults of unknown mating status collected in the field during this study; (5) the probability that a female mates with an attracted male is as predicted by close-range mate choice experiments (Drès, 2000) – if instead females mate with the first male to arrive, for example, the hybridization rate might be higher, and dependent only on relative rates of long-range attraction; (6) females mate once; or, if more than once, further matings are independent of and as assortative as the initial mating. Clearly, at least some of these assumptions, particularly those about relative population density and dispersion will be violated some of the time, or perhaps even most of the time. In particular our estimate of hybridization will be approximately correct only for the demographic situation – approximately 1:1 ratio of each host race – present at the study sites during our fieldwork. In most areas, however, one or other race will predominate. This can be caused by variability in forest composition from almost pure larch to almost pure pine, or due to strong periodic oscillations of population density in the larch race (Baltensweiler *et al.*, 1977). Such density differences will alter the interaction between host choice and mate choice by affecting the local ratio of larch and pine race individuals. Nonetheless, our purpose here is to gain an idea of the maximal rate of hybridization under natural conditions of sympatry and equal densities, rather than to obtain a globally accurate probability of hybridization over the whole range of the two forms.

Comparison with estimated hybridization rates in similar studies

Our hybridization estimates are similar to those made for other insects that use host plants rather than long-range mating cues. For instance, in the chrysomelid beetle *Lochmaea caprea* the probability of hybridization between birch and willow races in sympatry is 1–3% (Kreslavsky *et al.*, 1981); for hawthorn and apple races of the tephritid apple maggot *Rhagoletis pomonella* the probability is about 6% (Feder *et al.*, 1994); and for alfalfa and clover races of pea aphids *Acyrtosiphon pisum* it is 9–11% (Via, 1999). In another case where neither habitat cues nor long-range pheromones are involved in mate attraction, hybridization also occurs at a similar rate: in parapatry between two closely related species of *Heliconius* butterflies the rate is about 5% (Mallet *et al.*, 1998).

The fact that none of these estimates exceed 10% is probably not accidental. Continued gene flow between the host races should quickly obliterate any differenti-

ation unless divergent selection prevents fusion. If selection is only moderate, which is likely, low rates of gene flow will be a requirement for the maintenance of a bimodal genotypic distribution, or, in other words, for the coexistence of two recognizable, genetically distinct genotypic clusters. Possible mechanisms preventing genetic homogenization are selective mortality of hybrids, due to a sensitivity to secondary chemistry or physical defences of the host, or to an asynchrony of hybrid egg hatch with the timing of bud-burst. Selection for synchronization between egg hatch and bud-burst is known in a close relative, the spruce-feeding *Zeiraphera canadensis*, where juvenile mortality is strongly affected by the timing of bud-burst (Ostaff & Quiring, 2000). Another possibility is that hybrids are less fit for other ecological reasons: for example, the increase of null choice alighting probability seen in our experiments with laboratory hybrids may have some fitness-related effects in the field. Although we do not yet understand the causes of selection, it is clear that some such process must occur, or the differences between the host races would have broken down.

Host preference, imperfect sympatry, and reproductive isolation

We have shown that the attraction of *Zeiraphera* to their own hosts contributes strongly to assortative mating. Provided that an evolutionary shift in female host preference is also expressed in males, as here, evolution of divergent host choice in a hypothetical ancestor of *Zeiraphera* host races would cause some degree of reproductive isolation, even in the absence of pheromone differentiation. Hybridization between host-specific populations will be maximal when the environment is fine-grained in space and time (hosts are not clumped). A uniform fine-grained environment or 'ideal sympatry' is typically assumed in models of sympatric speciation, e.g. Dieckmann & Doebeli (1999), but hybridization rates decrease when this 'ideal sympathy' is violated. Because rapid host choice evolution under intense directional or disruptive selection seems not improbable, and because hosts and their host-specific parasites are almost always distributed in a somewhat coarse-grained, nonideal sympathy, the possibilities for sympatric speciation seem much more substantial than normally imagined in simplified models. Of course, to some extent, this 'imperfect sympathy' becomes increasingly similar to an allopatric or parapatric distribution as hosts and their parasites become more clumped spatially and temporally. To avoid getting caught up in a sterile terminological debate about the meaning of sympathy, an alternative way of phrasing our general conclusion is as follows: – a consideration of population structure of hosts and behaviour of parasites suggests that host shifts may often be sufficient to initiate speciation in parasites, even in species that do not use the host directly as a cue in mate choice, and even in the absence of geographical isolation.

Acknowledgments

We thank Prof. W. Baltensweiler for advice and encouragement, D.A. Knapp, C. Morgan, D. Ostaff, I. Phillips and Z. Yang for helpful support during the experimental and statistical analysis phases of this work. This research was supported by the Biotechnology and Biological Sciences Research Council.

References

- Baltensweiler, W., Benz, G., Bovey, P. & Delucci, P. 1977. Dynamics of larch budmoth populations. *Annu. Rev. Entomol.* **22**: 79–100.
- Baltensweiler, W., Priesner, E., Arn, H. & Delucci, V. 1978. Unterschiedliche Sexuallockstoffe bei Larchen- und Arvenform des Grauen Larchenwicklers (*Zeiraphera diniana* Gn., Lep. Tortricidae). *Mitt. Schweiz. Entomol. Ges.* **51**: 133–142.
- Baltensweiler, W. & Rubli, D. 1999. Dispersal: an important driving force of the cyclic population dynamics of the larch budmoth, *Zeiraphera diniana* Gn. *For Snow Landscape Res.* **74**: 3–153.
- Bovey, P. & Maksymov, J.K. 1959. Le problème des races biologiques chez la tordeuse grise du mélèze, *Zeiraphera griseana* (Hb.). *Vierteljahrsschr. Naturforsch. Ges. Zur.* **104**: 264–274.
- Bush, G.L. 1994. Sympatric speciation in animals – new wine in old bottles. *Trends Ecol. Evol.* **9**: 285–288.
- Butlin, R.K. 1990. Divergence in emergence time of host races due to differential gene flow. *Heredity* **65**: 47–50.
- Colwell, R.K. 1986. Population structure and sexual selection for host fidelity in the speciation of hummingbird flower mites. In: *Evolutionary Process and Theory* (S. Karlin & E. Nevo, eds), pp. 475–495. Academic Press, London.
- Craig, T.P., Itami, J.K., Abrahamson, W.G. & Horner, J.D. 1993. Behavioral evidence for host race formation in *Eurosta solidaginis*. *Evolution* **47**: 1696–1710.
- Day, K. 1984. Phenology, polymorphism and insect-plant relationships of the larch budmoth, *Zeiraphera diniana* (Guenee) (Lepidoptera, Tortricidae), on alternative conifer hosts in britain. *Bull. Ent. Res.* **74**: 47–64.
- Dieckmann, U. & Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature* **400**: 354–357.
- Drès, M. 2000. *Gene Flow Between Host Races of the Larch Budmoth Zeiraphera Diniana (Lep. Tortricidae)*. PhD Thesis, University College London, London.
- Edwards, A.W.F. 1972. *Likelihood*. Cambridge University Press, Cambridge.
- Emelianov, I.M., Drès, M., Baltensweiler, W. & Mallet, J. 2001. Host-induced assortative mating in host races of the larch budmoth. *Evolution* **55**: 2002–2010.
- Emelianov, I., Mallet, J. & Baltensweiler, W. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the larch budmoth) – polymorphism, host races or sibling species? *Heredity* **75**: 416–424.
- Feder, J.L. 1998. The apple maggot fly *Rhagoletis pomonella*. Flies in the face of conventional wisdom about speciation?. In: *Endless Forms: Species and Speciation* (D. J. Howard & S. H. Berlocher, eds), pp. 291–308. Oxford University Press, New York.
- Feder, J.L., Opp, S.B., Wlazlo, B., Reynolds, K., Go, W. & Spisak, S. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proc. Natl. Acad. Sci. USA* **91**: 7990–7994.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals. *Evolution* **35**: 124–138.
- Filchak, K.E., Roethle, J.B. & Feder, J.L. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* **407**: 739–742.
- Futuyma, D.J. 1998. *Evolutionary Biology*, 3rd edn. Sinauer, Sunderland.
- Jiggins, C.D., Linares, M., Mallet, J., Naisbit, R.E., Salazar, C. & Yang, Z. 2001. Sex-linked hybrid sterility in a butterfly. *Evolution* **55**: 1631–1638.
- Kreslavsky, A.G., Mikheev, A.V., Solomatin, V.M. & Gritzenko, V.V. 1981. Genetic exchange and isolating mechanisms in sympatric races of *Lochmaea caprea* (Coleoptera, Chrysomelidae). *Zool. Zhurnal* **60**: 62–68.
- Landolt, P.J. & Phillips, T.W. 1997. Host plant influences on sex pheromone behavior of phytophagous insects. *Annu. Rev. Entomol.* **42**: 371–391.
- Lu, W.H. & Logan, P. 1995. Inheritance of host-related feeding and ovipositional behaviors in *Leptinotarsa decemlineata* (Coleoptera, Chrysomelidae). *Environ. Entomol.* **24**: 278–287.
- Mallet, J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**: 294–299.
- Mallet, J., McMillan, W.O. & Jiggins, C.D. 1998. Estimating the mating behavior of a pair of hybridizing *Heliconius* species in the wild. *Evolution* **52**: 503–510.
- Messina, F.J. & Slade, A.F. 1997. Inheritance of host-plant choice in the seed beetle *Callosobruchus maculatus* (Coleoptera: Bruchidae). *Ann. Entomol. Soc. Am.* **90**: 848–855.
- Ostaff, D.P. & Quiring, D.T. 2000. Role of the host plant in the decline of populations of a specialist herbivore, the spruce budmoth. *J. Anim. Ecol.* **69**: 263–273.
- Pennacchio, F., Digilio, M.C., Tremblay, E. & Tranfaglia, A. 1994. Host recognition and acceptance behavior in two aphid parasitoid species – *Aphidius ervi* and *Aphidius microlophii* (Hymenoptera, Braconidae). *Bull. Ent. Res.* **84**: 57–64.
- Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Rice, W.R. & Hostert, E.E. 1993. Laboratory experiments on speciation – what have we learned in 40 years? *Evolution* **47**: 1637–1653.
- Rohlf, F.J. & Sokal, R.R. 1981. *Biometry*, 2nd edn. W.H. Freeman and company, San Francisco.
- Sezer, M. & Butlin, R.K. 1998. The genetic basis of oviposition preference differences between sympatric host races of the brown planthopper (*Nilaparvata lugens*). *Proc. R. Soc. Lond. B Biol. Sci.* **265**: 2399–2405.
- Tavormina, S.J. 1982. Sympatric genetic-divergence in the leaf-mining insect *Liriomyza brassicae* (Diptera, Agromyzidae). *Evolution* **36**: 523–534.
- Tregenza, T. & Butlin, R.K. 1999. Speciation without isolation. *Nature* **400**: 311–312.
- Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution* **53**: 1446–1457.

Received 25 July 2002; revised 18 September 2002; accepted: 26 November 2002

Appendix 1 Genotypes of adult males collected in 1998 from larch and pine.

Idh	Mdh	Pgm	On larch			On pine		
			Bever	Pontresina	Total	Bever	Pontresina	Total
ss	ss	ss	6	1	7	0	0	0
ss	ss	sf	4	1	5	0	0	0
ss	ss	ff	0	0	0	0	0	0
ss	fs	ss	0	0	0	0	0	0
ss	fs	sf	1	0	1	0	0	0
ss	fs	ff	0	0	0	0	0	0
ss	ff	ss	0	0	0	0	1	1
ss	ff	sf	0	0	0	0	0	0
ss	ff	ff	0	0	0	0	0	0
fs	ss	ss	4	3	7	1	1	2
fs	ss	sf	3	0	3	3	1	4
fs	ss	ff	2	0	2	0	0	0
fs	fs	ss	1	2	3	1	0	1
fs	fs	sf	1	2	3	1	0	1
fs	fs	ff	0	0	0	0	0	0
fs	ff	ss	3	1	4	9	2	11
fs	ff	sf	2	3	5	4	3	7
fs	ff	ff	1	0	1	2	1	3
ff	ss	ss	2	1	3	2	0	2
ff	ss	sf	1	0	1	0	0	0
ff	ss	ff	1	0	1	1	0	1
ff	fs	ss	3	1	4	0	0	0
ff	fs	sf	0	0	0	0	1	1
ff	fs	ff	0	0	0	0	0	0
ff	ff	ss	5	2	7	45	5	50
ff	ff	sf	0	1	1	11	10	21
ff	ff	ff	0	0	0	3	2	5
Totals			40	18	58	83	27	110

Appendix 2 Genotypes of adult females collected in 1998 on larch and pine.

Idh	Mdh	Pgm	On larch			On pine		
			Bever	Pontresina	Total	Bever	Pontresina	Total
s	ss	ss	29	6	35	4	0	4
s	ss	sf	19	0	19	4	0	4
s	ss	ff	0	0	0	0	0	0
s	fs	ss	11	1	12	0	0	0
s	fs	sf	5	1	6	0	0	0
s	fs	ff	1	0	1	1	0	1
s	ff	ss	1	0	1	4	1	5
s	ff	sf	0	0	0	1	0	1
s	ff	ff	0	0	0	0	0	0
f	ss	ss	33	4	37	4	0	4
f	ss	sf	15	4	19	7	0	7
f	ss	ff	1	1	2	0	0	0
f	fs	ss	12	3	15	2	1	3
f	fs	sf	9	0	9	4	1	5
f	fs	ff	1	0	1	0	0	0
f	ff	ss	12	1	13	51	8	59
f	ff	sf	3	3	6	26	9	35
f	ff	ff	0	0	0	4	1	5
Totals			152	24	176	112	21	133