Molecular and behavioural evidence for gene flow between host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae)

Michele Drès & James Mallet*

Galton Laboratory, Department of Biology, University College London, 4 Stephenson Way, London, NW1 2HE, UK

*Corresponding author: j.mallet@ucl.ac.uk

Summary

Larch and pine associated populations of *Zeiraphera diniana* (Lepidoptera: Tortricidae) differ in a number of heritable traits, but pheromone-mediated cross-attraction occurs between them in the wild. Using a quartet mate choice design (one male and one female of each type per cage) we estimate that, following cross-attraction by pheromones, the subsequent probability of hybridization is approximately 28%. We also examined molecular data, and were unable to distinguish between the races on the basis of 695bp of mitochondrial COI, tRNA-leucine, and COII gene sequence. Both results support earlier field studies suggesting that larch- and pine-feeding populations are host races that hybridize at an appreciable level in the wild. The shared mitochondrial haplotypes we observed are also consistent with ongoing and successful gene flow between the two host races.

Keywords

biotypes, gene flow, host races, hybridization, mitochondrial DNA, introgression, mate choice

Introduction

The larch budmoth *Zeiraphera diniana* Guenée (Lepidoptera: Tortricidae) inhabits coniferous forests throughout the Palaearctic, where it feeds on larch (*Larix* spp.), pine (*Pinus cembra*) and spruce (*Picea* spp.) (Berryman, 1986). Larch-feeding populations in the French and Swiss Alps undergo large fluctuations in density (up to 10^5 fold over a ~9 year period, Baltensweiler 1993) and are a model system for the study of population biology (e.g. Bjørnstad *et al*. 2002). The species is also a serious pest which causes wide-scale defoliation of larch at peak densities. Pine-feeding *Z. diniana* larvae are often found in sympatry with cycling larch populations, but field observations suggest that density fluctuations on this host are much smaller. (Baltensweiler W., Dormont, L. pers. comm.)

The larch- and pine-feeding populations appear to be host races (Emelianov *et al*. 1995, Drès and Mallet 2002); they are genetically distinct and associate with different hosts, but there is evidence of appreciable hybridization where they overlap. Adults tend to roost on their native host species in the wild (Emelianov *et al*. 2003) and in the laboratory they prefer to alight (Emelianov *et al*. 2003) and lay eggs (Bovey and Maksymov 1959) on the host species from which they were collected as larvae. In experimental conditions both types of larvae survive well on larch, but larch-race larvae rarely survive on pine (Baltensweiler 1977). The time of hatching of each race is synchronised with the bud burst of its host, approximately two weeks later in Alpine pine than in Alpine larch (Baltensweiler W, pers. comm.).

Fifth instar larch-race larvae are black while pine-race larvae are predominantly yellow-green (Baltensweiler 1993), and components of female pheromones are released in different ratios by each race. The differences in host preference, hatch time, larval colour and

pheromone composition are heritable, and there is also divergence in allozyme frequency and at AFLP loci (Emelianov *et al*. 1995, 2003, 2004).

Nonetheless, adults of pine and larch races overlap for several weeks (Baltensweiler W., pers. comm.) in the field, and long range, pheromone-mediated cross-attraction occurs at low frequency (Emelianov *et al*. 2001). In the laboratory, hybrid and backcross progeny can be bred with a normal sex ratio and no evidence for hybrid inviability, sex ratio distortion, or sterility (Baltensweiler 1977; Emelianov *et al*. 1995; Drès 2000).

Genetic differentiation between the populations has been studied using allozymes and AFLP markers. Large frequency differences were found at three of 13 polymorphic allozyme loci employed (*Mdh*-s, *Pgm*, and sex-linked *Idh*-s), but significant differences were absent at the remainder (Emelianov *et al*. 1995). AFLP analysis has revealed a heterogeneous pattern of genomic divergence between races similar to that seen in the allozyme study; AFLP markers that differed between races were clustered in just a few genomic regions. Analysis of AFLP genotypes also led to identification of three putatively hybrid individuals in wild-caught collections (Emelianov *et al*. 2004). These results are compatible with a hypothesis of host races; gene flow may maintain homogeneity across most genomic regions, while host-specific selection acting on regions near *Mdh*-s, *Pgm*, *Idh*-s, and on chromosomes with many divergent AFLP loci prevents exchange at these sites (Emelianov *et al*. 1995, 2004).

Here, we investigate genetic divergence between the races further using mitochondrial DNA (mtDNA) sequence information. The mitochondrial genome is maternally inherited and physically unlinked to nuclear loci that are likely to be subject to host-associated selection. MtDNA also tends to diverge more rapidly between populations than nuclear

DNA (Graur & Li 1999; Tomita *et al*. 1999; Simon *et al*. 1994), especially in Lepidoptera (Sperling 1990). The *Z. diniana* W-chromosome (equivalent to the *Drosophila* Y chromosome; female Lepidoptera are heterogametic ZW) is little differentiated at AFLP loci compared to some highly divergent chromosomes (Emelianov *et al*. 2004), suggesting that gene flow via female genomes may be high enough to homogenise variation between host races (in the absence of selection, but see James & Ballard 2003). Thus, an absence of mtDNA divergence between the populations would be consistent with the hypothesis of ongoing maternally derived gene flow. Conversely, mtDNA differentiation between races would indicate that current gene flow – at least as mediated by females – is restricted.

We examined an mtDNA fragment encoding part of external loop 6, transmembrane helix 12 (M12), and the carboxy terminus of cytochrome oxidase I (COI), transfer RNA-leucine (tRNA-leu), and the amino terminus, M1, interval, M2, linking and partial copper binding regions of cytochrome oxidase II (COII). With the exception of the region encoding tRNA-leu, this is a region of low mtDNA sequence conservation, and contains peaks of mtDNA nucleotide diversity (reviewed in Simon 1994, Caterino and Sperling 1999).

Because it can still be difficult to distinguish between ongoing and recently ceased gene flow on the basis of genetic data alone, we also conducted behavioural experiments to obtain direct measurements of mating behaviour. Gene flow between larch- and pine-races involves a number of steps, beginning with pheromone-mediated cross attraction of males to females of the opposite race (Emelianov *et al*. 2001). Here, we estimate the frequency with which subsequent hybrid matings take place under competitive conditions, and test the ability of hybrids to obtain matings with non-hybrids.

Materials and Methods

Collections

Fourth and fifth instar *Z. diniana* larvae were collected in 1998-9 from six sites in the Western (French) and Eastern (Swiss) Alps (for a map, see Emelianov et al., 1995): (1) Bois les Ayes, France (LA, 44°50′ N, 6°39′E), (larch and pine races); (2) Bois les Combes, France (LC, 44°54′ N, 6°34′E), (larch race); (3) Montgénèvres, France (MG, 44°56′, 6°43′ E), (larch race); (4) Tueda, France (T, 45°23′N, 6°43′E), (pine race); (5) Pontresina, Switzerland (P, 46°29′ N, 9°54′ E), (larch and pine races); (6) Sils, Switzerland (S, 46°42′N, 9°27′ E), (larch race); and (7) Val Bever, Switzerland (VB, 46°26′N, 9°50′E), (larch and pine races). Individuals were labelled according to collection site and host species (e.g. LA(L) = Les Ayes (larch race)).

Rearing

Larvae were individually reared to pupation on newly flushing larch (*Larix* spp.) using methods described in Emelianov *et al.* (1995, 2003) and Drès (2000). Rearing took place in a constant temperature (CT) room at 17-21°C with a reversed day night cycle (day: 21:00 – 13:00, twilight: 13:00 – 14:00). Larch pupae and 5th instar larvae were removed to an incubator (day: 21:00-13:00) and chilled as necessary at 12°-14°C to synchronize their development with the more slowly developing pine-race insects. Adults were allowed to mature for two to eight days before use in mate choice experiments.

DNA isolation, PCR and sequencing

All samples used in mtDNA analysis were collected from sites LA, LC, MG, P and VB. DNA was extracted as described by Bernardi *et al*. (2004) from thoraces and stored at –80°C until use. MtDNA was amplified with the primers 'Eva' (Bogdanowicz *et al*. 1993) and 'George III' (A. Brower, www.ent.orst.edu/brower/primers.htm) in a Perkin-Elmer Gene Amp 2400 PCR machine using a thermal cycling regime of 94°C for 5 min; 35 cycles of 94°C for 40 sec, 50°C for 40 sec, 72°C for 90 sec; and 72°C for 10 min. PCR products were purified using Qiagen QIAquick or QIAEX II Gel Extraction kits according to the manufacturer's instructions, and sequenced using one or more of the primers: (i) George III, (ii) ZDF-1 (GCTTCTCCTTTAATAGAAC) (iii) ZDF-2 (CTTCCACCAGCTGAACAT), and (iv) ZDR-1 (GTTCTATTAAAGGAGAGGCTCTATTTTG). With the exception of George III all sequencing primers were designed from *Z. diniana* sequence. DNA sequences were aligned using Clustal W Version 1.7 (Thompson *et al*. 1994). Putative amino acid differences between sequences were determined with reference to the *Drosophila* mitochondrial genetic code (Clary and Wolstenhome 1985).

Quartet mate choice experiments

The majority of 'quartet' experiments conducted were of larch race *vs.* pine race moths (n=96). We also performed a limited number of tests using hybrids: F1 hybrid vs. larch race (n=5), and F2 larch x hybrid backcross vs. larch race (n=1). To assess mate choice and distinguish propensity to mate from host choice, each experiment involved four moths - one virgin male and one virgin female of each type being tested, *i.e.* 408 individuals in total (see

also Davies *et al*. 1997). Each quartet of moths was enclosed in a 1 litre cylindrical cage before 13:00 (twilight) and observed every 30 minutes between 13:00 and 19:30 until a single mating took place, whereupon the test was terminated and the remaining individuals discarded. In a pilot study of 22 experiments, all matings took between 30 minutes and several hours to complete, and commenced before 19:30. Cages contained fresh sprigs of larch and pine to promote normal behaviour. Individual adults are not identifiable to host race, therefore, prior to the experiment, one male and one female from each quartet was marked on one wing with a felt tipped pen to enable subsequent identification. The unmated male and female were removed from each cage as soon as a mating was observed, and were not used in subsequent experiments. Wherever possible, all members of a quartet eclosed within four days of one another. In 16 larch *vs.* pine experiments this was not the case, although in 11 of these it was possible to ensure that both members of the same sex had paired eclosion dates.

Likelihood ratio (G) tests were used to determine whether tendency to mate was independent of biotype of the opposite sex. Marginal totals of the contigency table were used to calculate relative mating propensities of males, and females, of each race (c.f. methods used in Davies *et al*. 1997; Jiggins *et al*. 2001). Cross-products of mating propensities provided expected frequencies of each type of cross under a null hypothesis of random mating (Casares *et al*. 1998).

For each race and sex, G tests or (where sample sizes were small) Fisher exact tests were also used to examine whether tendency to hybridize and mating propensity were independent of identification marking, year of experiment, and age.

Results

mtDNA sequencing

The 695 bp mtDNA consensus sequence amplified (GenBank accession no. xxxxx) corresponds to bp 2776 – 3485 of the *Drosophila yakuba* reference sequence (Clary and Wolstenhome 1985) and has an A-T composition of 78.3%. Thirty-four sequences were obtained, 19 from larch-race samples and 15 from pine-race samples. Twenty-six sequences were 695bp long, and eight sequences, LA(P)1, LA(P)2, LC(L)1, S(L)1, VB(L)1, P(P)1, P(P)2and VB(P)1 were 468bp long (Table 1) - these extended less far into COII, terminating in the M2 structural region, and were obtained with the George III sequencing primer only. Nucleotide information for all 15 variable sites in the 695bp region is summarised in Table 1 and Fig. 1. In total, six haplotypes were detected, the consensus haplotype shared by most larch - and pine-race samples, and five variant haplotypes hereafter referred to as L1, L2, L3, L4, and P1, where L (larch) and P (pine) indicate host race origin. All substitutions are described relative to the consensus haplotype. Haplotypes L1 and L3 were scored in 468bp samples, and therefore may contain additional undetected substitutions in COII .

Estimates of sequence variability presented here refer to the 695bp mtDNA fragments, and are conservative because additional sites of variation may have been present in 468bp samples. In the larch-race, seven third-codon and one first-codon sites were polymorphic, giving rise to a total nucleotide variability of 1.2%. In the pine-race nucleotide variability across the same region was 1.4%, and consisted of one first-codon, one second-codon, and eight third-codon variable sites. One first-codon substitution shared by haplotypes L3 and P1 caused a putative valine to isoleucine replacement, and one second-codon substitution in

the pine race haplotype L4 resulted in the putative replacement of methionine with threonine. All other nucleotide substitutions between haplotypes were synonymous.

Mate choice experiments

Mating was non-random with respect to host race (See Table 2, G = 13.56, ν = 1, P=0.0002). In total, 28% of all matings were hybrid (25% pine female x larch males, 3% larch female x pine male). Pine females had greater mating propensities than larch females: 84.4 (75.5 - 91.0)% *vs.* 15.6 (9.0-24.5)%, and pine males greater mating propensities than larch males: 62.5 (52.0-72.2)% *vs.* 37.5 (27.8 – 48.0)%.

The overall frequency of hybridization was independent of marking in males and females of both races (Table 3). Mating propensity was independent of marking in larch and pine males, and in pine females. There was a nominally significant lack of independence between marking and mating propensity in larch females (Fisher Exact Test, P=0.033, Table 4); this result did not remain significant after Bonferroni correction for multiple tests. Excluding data from experiments involving marked larch females (n=46) did not significantly affect the overall rate of hybridization, nor did excluding the small 1998 dataset (n=8 tests), or data from experiments in which moths differed in age by more than four days (n= 16). Therefore all data were retained in the final analysis.

Three of the five larch *vs.* hybrid mate choice tests resulted in hybrid female x larch male matings. The remaining matings were larch female x larch male (n=1) and hybrid female x hybrid male (n=1). No larch female x hybrid male matings were observed, but this is perhaps not surprising given the small sample size (in addition, we did no tests involving

pine race and hybrids). The single larch *vs.* backcross experiment resulted in a backcross female x larch male mating.

Discussion

mtDNA differentiation between the races.

Larch- and pine-associated populations of *Z. diniana* could not be distinguished on the basis of the 695 bp of mtDNA sequenced here; the majority of haplotypes were identical. In 26 haplotypes of 695bp, only VB(L)-2 differed from the consensus sequence solely in the region absent from shorter fragments, therefore few apparently consensus 468bp haplotypes are likely to have differed from this common sequence.

It is possible that some non-consensus haplotypes occur only in one race, but if so they are not abundant. Haplotypes L1, L2 and L4, for example, were scored only in larch-race samples. Haplotypes L2 and L4 were unique, and L1 was found in a pair of larch-race samples collected ~290 km apart. However, due to the rarity of these haplotypes, their absence in the pine host race could be due to sampling effects. The most strongly divergent haplotype, P1 from the pine race, appears to be shared with the larch race judging from the 468 bp haplotype sequence L3. These data as a whole give convincing evidence for minimal divergence in mitochondrial DNA sequence between host races.

Mate choice experiments

There was a strong interaction between the host race of a mated female and the host race of her mate. Therefore, assortative mating was evident, independently of differences in mating propensity. The low mating propensity of both males and females from the larch race in these experiments applied both to assortative and cross-matings, suggesting that some rearing effect or real biological differences between the host races is a cause of low mating propensity. However, it is hard to separate the effects of mating propensity from mate choice purely on the basis of mating count data such as those obtained here (Davies *et al*. 1997).

All insects were reared to pupation on larch, because pine resin flowing from cut pine shoots prevents successful rearing of larvae on pine in the lab. Host plant chemistry potentially affects cuticular hydrocarbons (Stennet and Etges 1997) and these compounds are known to play a key role in the mate choice behaviour of many insect species (Coyne *et al*. 1994; Ferveur 1997; Singer 1998; Tregenza and Wedell 1997), though their importance in *Z. diniana* is unknown. However it seems unlikely that any effect of host foodplant on the chemistry of pine-race insects influenced results presented here. The majority of pine-race insects used in mate choice experiments were collected from the wild as late fifth instar larvae and consumed little or no larch in captivity, yet they mated successfully and strongly assortatively with other individuals from the pine race.

Gene flow between the host races

The lack of mitochondrial DNA differentiation between larch- and pine- races of *Z. diniana* is consistent with our earlier estimate of hybridization of approximately 3% per generation (Emelianov *et al*. 2003), and with the exchange of genetic information between the races via

backcrossing (Emelianov *et al*. 2004). At least some of this gene flow must be female-mediated, as is also suggested by results of the mate choice experiments. Female-mediated gene flow is possible provided female hybrids and their female offspring are viable and fertile, as here (see Emelianov *et al*. 2003).

It is possible that we failed to detect existing differentiation. The fragment of mitochondrial DNA sequenced here contains several peaks of COI and COII nucleotide diversity in Lepidoptera and other insects (Caterino and Sperling, 1999; Lunt et al., 1996; Simon et al., 1994), but other areas of mtDNA such as the hypervariable region may be even more informative (Harrison 1989; Lunt *et al*. 1996; Simon *et al*. 1994).

It is also possible that absence of molecular differentiation is explained by a very recent cessation of gene flow, rather than by ongoing gene flow. Several studies have detected only very low or nil levels of mtDNA divergence between populations for which there is scant evidence of hybridization (unlike this study), suggesting mtDNA differentiation between some reproductively isolated species is either absent or remains undetected. For example, among three species of *Yponomeuta* ermine moths (Yponomeutidae) less than 1% divergence was detected across 2.3kb of COI and COII (Sperling *et al*. 1995). No divergence was found between the two Eurasian species of pierid butterfly, *Colias croceus* and C. *erate poliographus* in 333bp of COI (Pollock *et al*. 1998), although the authors suggest that sequencing a longer fragment might reveal differences. Similarly, variation across a 410bp COI sequence of 39 *Lycaeides idas* and *L. melissa* (Lycaenidae) failed to distinguish between the species (Nice and Shapiro 1999), and host related species of Astraptes skippers differed by as little as 0.2-0.3% at COI (Hebert *et al*. 2004). Harrison (1989) reviews other examples.

Because of the potential ambiguity in the molecular data, we also addressed the question of gene flow using behavioural experiments. The quartet mate-choice tests demonstrate that close-range mate choice in *Z. diniana* is only moderately assortative with respect to host race, and that hybrid females, at least, can successfully compete with non-hybrids to mate with larch-race males. Thus, we provide direct evidence that hybridization occurs readily between the forms, and can lead to introgression. Additional evidence that this gene flow occurs in the wild comes from field studies of pheromone-mediated cross attraction, and from the collection of larvae with F1 and backcross AFLP genotypes (Emelianov *et al*. 2001, 2004).

The fairly high rate of hybridization in these experiments also contrasts strongly with the rarity of cross-attraction in long-range pheromone-based field tests (Emelianov *et al*. 2001). The relative strengths of the various stages to cross-attraction can be estimated as follows. On average, larch females and pine females tend to settle as adults on the wrong host tree in mixed forest about 14% of the time (Emelianov *et al*. 2003). Females on the wrong hosts tend to attract more males of the wrong host race (22.0%) than females alighting on the correct hosts (6.2%; Emelianov *et al*. 2001). Thus, pheromonal cross attraction can be estimated, on average, as 0.86 x 0.062 + 0.14 x 0.220, or 8.4%. Finally, assuming females do not compete in attracting males and mate according to the experiments described in Table 2, the average cross-mating due to close-range interactions is [3/(12+3)+24/(57+24)]/2 = (20.0%+29.6%)/2 = 24.8%. Cross-mating frequency is almost symmetrical with 20% of larch females mating with pine males, and 29.6% of pine females mating with larch males (G_1 = 1.17, P≈0.5, based on comparing these best estimates of female cross-mating tendencies with their average, 24.8%). Thus, the most strongly assortative phase of the mating sequence consists of the long-range pheromone attraction (91.6% assortative). Overall, the varying routes by which

cross-mating can occur lead to approximately 2.2-3.8% hybridization overall per generation (see calculations in Emelianov *et al*. 2003), a value similar to the finding of a few percent of genetically identified hybrids in the field (Emelianov *et al*. 2004). Host races can be defined as "genetically differentiated, sympatric populations of parasites that use different hosts and between which there is appreciable gene flow" (Drès and Mallet 2002). The data gathered here support the view that the divergent forms of *Zeiraphera diniana* inhabiting larch and pine are host races under this definition, rather than "good species".

Acknowledgments

We thank Werner Baltensweiler for hospitality in Switzerland and advice on all aspects of *Z. diniana* biology; Laurent Dormont, Igor Emelianov and Yvonne Graneau for assistance with collections and for discussions in the lab; and Angela Kingman, Richard Marguerie and Hazel Smith for help with molecular techniques.

References

Baltensweiler, W. 1977 Colour-polymorphism and dynamics of larch budmoth populations (*Zeiraphera diniana* Gn., Lep. Tortricidae). *Mitteil. Schweiz. Entomol. Gesellsch.* **50**, 15-23.

Baltensweiler, W. 1993 A contribution to the explanation of the larch bud moth cycle, the polymorphic fitness hypothesis. *Oecologia (Berl.)* **93**, 251-255.

Baltensweiler, W., Priesner, E., Arn, H. & Delucchi, V. 1978 Unterschiedliche Sexuallockstoffe bie Lärchen- und Arvenform des Grauen Lärchenwicklers (*Zeiraphera diniana* Gn., Lep. Tortricidae). *Mitteil. Schweiz. Entomol. Gesellsch.* **51**, 133-142.

Bernardi, H.J., Bucciarelli, G., Costagliola, D., Robertson, D.R. & Heiser, J.B. 2004 Evolution of coral reef fish *Thalassoma* spp. *Marine Biol.* **144**, 369-375.

Berryman, A.A. 1986 On the dynamics of blackheaded budworm populations. *Canad. Entomol.* **118**, 775-779.

Bjørnstad, O.N., Peltonen, M., Liebhold, A.M. & Baltensweiler, W. 2002 Waves of larch budmoth outbreaks in the European Alps. *Science* **298**, 1020-1023.

Bogdanowicz, S.M., Wallner, W.E., Bell, J., Odell, T.M. & Harrison, R.G. 1993 Asian gypsy moths (Lepidoptera: Lymantriidae) in North America: evidence from molecular data. *Ann. Entomol. Soc. Amer.* **86**, 710-715.

Bovey, P. & Maksymov, J.K. 1959 Le problème des races biologiques chez la tordeuse grise du mélèze *Zeiraphera griseana* (Hb.). Note préliminaire. *Vierteljahrsschrift Naturforsch. Ges. Zürich* **104**, 264-274.

Casares, P., Carracedo, M.C., del Rio, B., Pineiro, R., Garcia Florez, L. & Barros, A.R. 1998 Disentangling the effects of mating propensity and mate choice in *Drosophila*. *Evolution* **52**, 126-133.

Clary, D.O. & Wolstenhome, D.R. 1985 The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J. Molec. Evol.* **22**, 252-271.

Coyne, J.A., Crittenden, A.P. & Mah, K. 1994 Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science* **265**, 1461-1464.

Davies, N., Aiello, A., Mallet, J., Pomiankowski, A. & Silberglied, R.E. 1997 Speciation in two neotropical butterflies: extending Haldane's rule. *Proc. Roy. Soc. Lond. B* **264**, 845-851.

Drès, M.A. 2000. Gene flow between host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae). Ph.D. Diss, University of London.

Drès, M. & Mallet, J. 2002 Host races in plant-feeding insects and their importance in sympatric speciation. *Phil. Trans. Roy. Soc. Lond. B* **357**, 471-492.

Emelianov, I., 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 the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae): polymorphism, host races or sibling species? *Heredity* **75**, 416-424.

Emelianov, I., Marec, F. & Mallet, J. 2004 Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proc. Roy. Soc. Lond. B* **271**, 97-105.

Emelianov, I., Simpson, F., Narang, P. & Mallet, J. 2003 Host choice promotes reproductive isolation between host races of the larch budmoth *Zeiraphera diniana*. *J. Evol. Biol.* **16**, 208-218.

Engels, W.R. 1993 Contributing software to the internet: the Amplifying program. *Trends Biochem. Sci.* **18**, 448-450.

Ferveur, J.F. 1997 The pheromonal role of cuticular hydrocarbons in *Drosophila melanogaster*. *BioEssays* **19**, 353-358.

Graur, D. & Li, W.-H. 1999 *Fundamentals of Molecular Evolution*. Sunderland, Mass., USA: Sinauer.

Harrison, R.G. 1989 Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* **4**, 6-11.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H. & Hallwachs, W. 2004 Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci., USA* **101**, 14812-14817.

James, A. & Ballard, W.O. 2003 Mitochondrial genotype affects fitness in *Drosophila simulans*. *Genetics* **164**, 187-194.

Jiggins, C.D., Naisbit, R.E., Coe, R.L. & Mallet, J. 2001 Reproductive isolation caused by colour pattern mimicry. *Nature* **411**, 302-305.

Lunt, D.H., Zhang, D.-X., Szymura, J.M. & Hewitt, G.M. 1996 The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molec. Biol.* **5**, 153-165.

Nice, C.C. & Shapiro, A.M. 1999 Molecular and morphological divergence in the butterfly genus *Lycaeides (Lepidoptera: Lycaenidae) in North America.* J. Evol. Biol. **26**, *473-484.*

Pollock, D.W., Watt, W.B., Rashbrook, V.K. & Iyengar, E.V. 1998 A molecular phylogeny for *Colias* butterflies and their relatives (Lepidoptera, Pieridae). *Ann. Entomol. Soc. Amer.* **91**, 524-531.

Priesner, E. & Baltensweiler, W. 1987 Studien zum Pheromon-Polymorphismus von *Zeiraphera diniana* Gn. (Lep., Tortricidae). 1. Pheromon-Reaktionstypen männlicher Falter in europäischen Wildpopulationen, 1978-85. *J. Appl. Entomol.* **104**, 234-256.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. 1994 Evolution, weighting, and phylogenetic usefulness of mitochondrial genes with a compilation of conserved PCR primers. *Ann. Entomol. Soc. Amer.* **87**, 651-701.

Singer, T.L. 1998 Roles of hydrocarbons in the recognition systems of insects. *Amer. Zoologist* **38**, 394-405.

Sperling, F.A.H. 1990 Natural hybrids of *Papilio* (Insecta: Lepidoptera): poor taxonomy or interesting evolutionary problem? *Canad. J. Zool.* **68**, 1790-1799.

Sperling, F.A.H., Landry, J.-F. & Hickey, D.A. 1995 DNA-based identification of introduced ermine moth species in North America (Lepidoptera: Yponomeutidae). *Ann. Entomol. Soc. Amer.* **88**, 155-162.

Stennett, M.D. & Etges, W.J. 1997 Premating isolation is determined by larval rearing in cactophilic *Drosophila mojavensis*. III. Epicuticular hydrocarbon variation is determined by use of different host plants in *Drosophila mojavensis* and *Drosophila arizonae*. *J. Chem. Ecol.* **23**, 2803-2824.

Thompson, J.D. 1994 Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.

Tomita, K., Ueda, T., Ishiwa, S., Crain, P.F., McCloskey, J.A. & Watanabe, K. 1999 Codon reading patterns in *Drosophila melanogaster*: a unique wobble role in animal mitochondrial DNA. *Nucl. Acids Res.* **27**, 4291-4297.

Tregenza, T. & Wedell, N. 1997 Definitive evidence for cuticular pheromones in a cricket. *Anim. Behav.* **54**, 979-984.

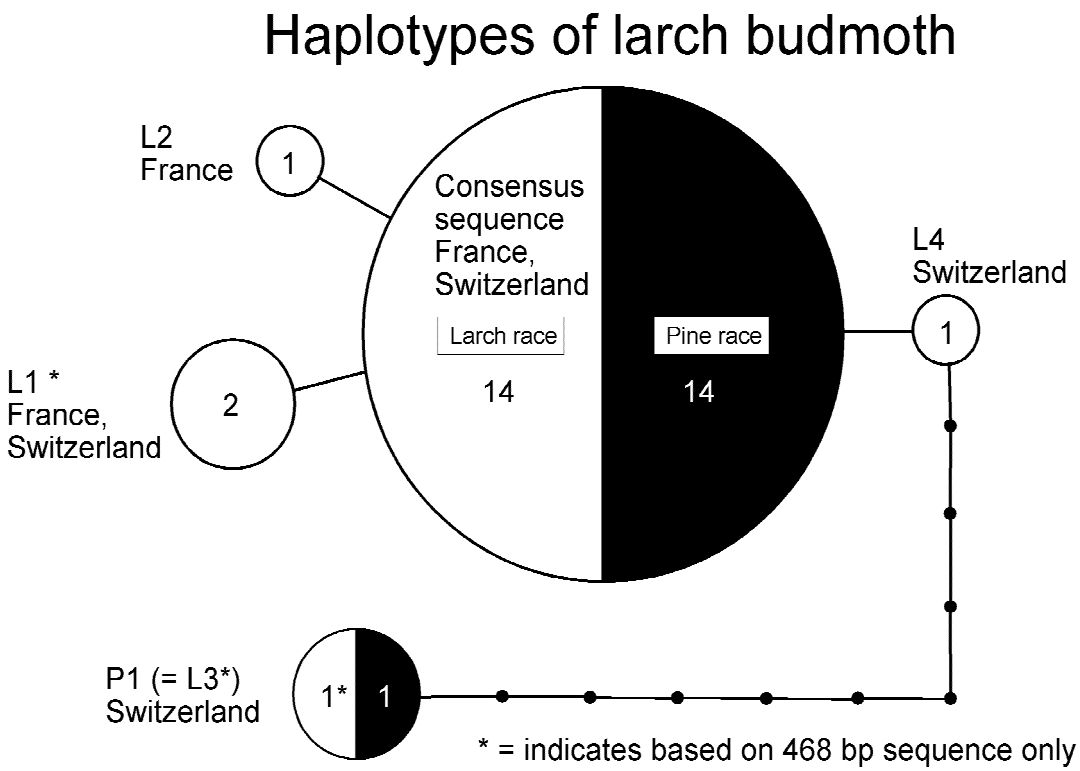

Figure 1. Minimum spanning network of larch budmoth mitochondrial haplotypes
Each haplotype is represented by a circle (numbers of haplotypes are shown inside). Haplotypes not found, but inferred to have been present, are shown as dots. Most haplotypes were based on 695 bp of sequence; eight haplotypes for which only 468 bp were scored have been included, based on the assumption that their sequence is identical to the most closely related sequences in the unsampled region.

Tables

Table 1. Variable sites in *Zeiraphera diniana* mtDNA sequences

Locality (host race) - sequence number	haplotype	position relative to the *Drosophila yakuba* reference sequence (Clary and Wolstenhome 1985)												
		2814	2832	2844	2895	3167	3196	3232	3346	3350	3352	3358	3451	3471
consensus sequence	consensus	**A**	**T**	**C**	**A**	**G**	**A**	**T**	**T**	**C**	**G**	**T**	**T**	**T**
Larch race														
LA(L)-1	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(L)-2	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(L)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)-1*	L1	.	.	.	.	**A**‡	.	.	–	–	–	–	–	–
LC(L)-2	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)-4	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)-1	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)-2	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)-3	L2	.	**C**	.	.	.	.	.	.	.	.	.	.	.
P(L)-1	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)-2	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)-4	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(L)-1*	L3 (=P1)	**C**	.	**T**	**G**	.	**G**	**C**	–	–	–	–	–	–
VB(L)-2	L4	.	.	.	.	.	.	.	.	.	**A**	.	.	.
VB(L)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(L)-4	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
S(L)-1*	L1	.	.	.	.	**A**‡	.	.	–	–	–	–	–	–
Pine race														
LA(P)-1*	consensus	.	.	.	.	.	.	.	–	–	–	–	–	–
LA(P)-2*	consensus	.	.	.	.	.	.	.	–	–	–	–	–	–
LA(P)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)-4	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)-5	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-1*	consensus	.	.	.	.	.	.	.	–	–	–	–	–	–
P(P)-2*	consensus	.	.	.	.	.	.	.	–	–	–	–	–	–
P(P)-3	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-4	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-5	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-6	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-7	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)-8	consensus	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(P)-1*	consensus	.	.	.	.	.	.	.	–	–	–	–	–	–
VB(P)-2	P1(=L3)	**C**	.	**T**	**G**	.	**G**	**C**	**C**	**T**	**A**	**C**	**C**	**C**‡‡

All substitutions are synonomous unless otherwise indicated
* based on 468 bp sequences only
_ base unknown
‡ putative valine to isoleucine substitution
‡‡ putative methionine to threonine substitution

Table 2. Larch vs. pine host race mating experiments

	larch male	pine male
larch female	12	3
pine female	24	57

(n=96, G=13.56, v = 1, P=0.0002)

Table 3. Results of G-tests and Fisher Exact tests of independence between marking and tendency to hybridize.

host race	sex	G	P
Larch	Female	-	0.85 (P exact)
Larch	Male	0.06	0.81
Pine	Female	0.06	0.81
Pine	Male	-	0.14 (P exact)

(For each test v = 1)

Table 4. Results of G-tests and Fisher Exact tests of independence between marking and mating propensity.

host race	sex	G	P
larch	female	-	0.03* (P exact).
larch	male	0.12	0.73
pine	female	2.70	0.10
pine	male	0.30	0.58

(For each test v = 1. The symbol " * " indicates nominally significant result (did not remain significant after Bonferroni correction for multiple tests))