Gene flow between host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae)

by

Michele Anne Drès

A thesis submitted for the degree of
Doctor of Philosophy
of the University of London

September 2000

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

Abstract

This project investigates the host race status of larch and pine associated *Zeiraphera diniana* populations via studies of actual gene flow, hybrid fitness, and mitochondrial DNA. The formation of host races - genetically distinct, host associated biotypes connected by gene flow - is an essential step in sympatric speciation via host shift, and their existence provides strong support for this theory. However, the identification of these forms has been complicated by the use of multiple host race definitions that are often ambigous in an empirical context. Therefore a new, empirically based definition is developed in and employed throughout this thesis. A novel requirement of this definition is that at least 1% actual gene flow (movement of genes via hybridisation) occurs between the races each generation, and leads to backcrossing. Other requirements include sympatry, genetic differentiation, and host fidelity. Previous studies have shown that larch and pine associated *Z. diniana* meet most host race criteria, but little was known about the level of gene flow between the races. The main findings presented here are that the level of actual gene flow (the movement of genes via hybridisation) between the biotypes is approximately 2.4% per generation, that backcrossing is likely, and therefore that they are indeed host races. The extent of actual gene flow was estimated from the combined probabilities of long range, pheromone-mediated cross attraction (measured in the field), and hybridisation in competitive situations (measured in the laboratory). Evidence of backcrossing was obtained from several experiments. Proportions of non-hybrid, hybrid, and backcross broods hatching, and surviving to final larval instars in the laboratory did not differ. Hybrid females were able to obtain matings with larch males in competitive conditions, and hybrids of both sexes are fertile. In addition, there was no evidence of mitochondrial DNA differentiation between the forms across approximately 800bp of mitochondrial cytochrome oxidase I (partial), tRNA-leucine, and cytochrome oxidase II (partial).

Acknowledgements

I would like to thank my supervisor, James Mallet, for all his help throughout this project, and Igor Emelianov, for initiating the *Zeiraphera* host race project. Many thanks also to Werner Baltensweiler, for his generosity with his expertise and his time, and to Yvonne Graneau, for her help and good company in the rearing lab and the field. To Laurent Dormont for help with collections, and Hazel Smith and Richard Marguerie de Rotrou for assistance with PCR. To Mary and John Drès for, among other things, collecting larch branches, and to Peter Drès. To Andrew Hingle, Angela Gouge and Margaret Greenwood for all their help and support. This project was funded by a NERC studentship.

Table of Contents

List of Tables

List of Figures

Preface

Biologists have long suggested a sympatric origin, via host shift, of the myriad closely related, sympatric insect species specialized on different host plants (Bush, 1975; Bush and Smith, 1998; Price, 1980). This thesis addresses a central aspect of host shift speciation theory, the ability of host associated forms to retain genetic differences while undergoing gene flow, by focusing on sympatric biotypes of the moth *Zeiraphera diniana* found on larch and pine trees in the French and Swiss Alps.

Chapter 1 provides a general introduction to host shift speciation theory. A source of confusion in the field is the use of multiple, ambigous definitions of 'host races', intermediate forms whose existence lends strong support to the theory. Consequently, a new, experimentally verifiable host race definition is developed. The new definition is applied to putative host race systems for which appropriate empirical data exists, and the results are discussed with reference to the likelihood of sympatric speciation in plant feeding insects.

Chapter 2 and the Appendix are concerned with evidence for actual gene flow between larch and pine associated *Z. diniana* populations. The term 'actual gene flow' refers simply to the movement of genes from one population to another via hybridisation; no assumption is made about whether the genes that do flow become incorporated, or are lost (Mallet, in press). The Appendix describes a study of long range, pheromone mediated cross attraction carried out in the field. Single virgin females of both host races were enclosed on either their natal or the alternate host type, and the relative proportions of *Z. diniana* males of the opposite host race they attracted were determined. In Chapter 2, laboratory mate choice tests are presented. Experimental moths – one male and one female each of larch and pine races, or larch race and F1 hybrid – were confined in close proximity, and the direction of the first mating to take place was recorded.

Chapters 3 and 4 address the question of whether actual gene flow between the two races leads to gene flow in the more traditional sense, the incorporation of genes from one population into another. The intrinsic fitness of larch, pine, F1 and backcross broods are compared in Chapter 3 using laboratory measures of percentage egg hatch, survival

to final (5^{th}) larval instar, and sex ratio at 5^{th} instar. Finally, in Chapter 4 I consider molecular evidence for gene flow between the races. An 800bp region of mitochondrial DNA encompassing a region from 3' cytochrome oxidase one to 3' cytochrome oxidase two is examined for evidence of differentiation along host race lines.

Format of Thesis Chapters

All chapters have been written in the style of manuscripts for submission to scientific journals, but their formatting has been standardized for presentation in this thesis. None of the chapters has yet been submitted, but it is expected that all will be submitted with my name as first author. Figures and tables are presented at the end of the first chapter to which they are relevant. Chapter 1 provides a general introduction to host shift speciation theory, but the study system is not introduced in detail until Chapter 2. Chapter 2 describes mate choice experiments, that, in combination with the results of the Appendix, allow an estimate of the level of actual gene flow between the host races. The implications of the results of Chapters 3 and 4 with regard to the probability of gene flow between the forms are also briefly discussed in Chapter 2. All chapters and analyses therin are entirely my own work. The Appendix (Emelianov et al., in prep) has been submitted to Evolution, and is senior authored by Igor Emelianov. It has however been included here because I made a substantial contribution to the work.The rearing of larvae in 1997 and 1998 was a collaborative effort between myself and Igor Emelianov, with the help, in 1997, of Yvonne Graneau. Collections from 1997 onwards were made in concert with Werner Baltensweiler, and, variously, Laurent Dormont, Igor Emelianov, Svetlana Emelianova, Yvonne Graneau, and James Mallet. Collections made in 1996, referred to in Chapters 3 and 4, were made prior to the start of this project, by Werner Baltensweiler, Laurant Dormont, and Chris Hartley, and all rearing in this year was carried out by Igor Emelianov.

References

Bush, G. 1975. Modes of animal speciation. *Annual Review of Ecology and Systematics*, 6:339-364.

Bush, G., and J. Smith. 1998. The genetics and ecology of sympatric speciation: A case study. *Researches on Population Ecology*, 40:175-187.

Emelianov, I.,W. Baltensweiler, M. Drès, and J. Mallet. in prep. Host fidelity of larch and pine race *Zeiraphera diniana* (Lepidoptera: Tortricidae).

Mallet, J. in press. Gene Flow. In I. P. Woiwood, D. R. Reynolds and C. D. Thomas (eds.), *Insect Movement: Mechanisms and Consequences*. CAB International, Wallingford.

Price, P. W. 1980. *Evolutionary Biology of Parasites*. Princeton University Press, Princeton.

Chapter 1

Does host race formation in plant feeding insects lead to sympatric speciation?

Abstract

The existence of a continuous array of sympatric, host associated forms, from polymorphisms, through host races with increasing degrees of reproductive isolation, to good species, would be strong evidence of a stable route for sympatric speciation via host shift. A similar pattern has been shown to exist in parapatric biotypes meeting at hybrid zones. However, assessing the pattern in sympatric, plant-feeding insects (a group in which speciation via host shift seems particularly likely) is difficult, in part because multiple, ambigous host race definitions are currently in use, but also because estimates of current levels of gene flow have been achieved in only a handful of likely cases. Here, we outline a new, empirical definition of host races, and apply it to 17 putative host race systems. Of these, only two could be unambiguously identified as such, but a further seven are likely to be host races. We conclude that sympatric speciation seems likely, but caution that current empirical data does not rule out the possibility of a discontinuity in the theoretically stable route from host associated polymorphism to host associated species.

Introduction

The formation of 'host races' - genetically distinct, host associated biotypes connected by gene flow - is an important step in models of sympatric speciation via host shift, which describe how disruptive host associated selection, combined with differences in host use, can drive sympatric populations to speciate. The identification and study of host races has a controversial history, largely due to a close association with sympatric speciation theory (Tauber and Tauber, 1989), and the literature has been reviewed several times, both by authors who believe sympatric speciation is common in parasites (Diehl and Bush, 1984; Strong et al., 1984; Tauber and Tauber, 1989; see also Mopper and Strauss, 1998), and those that do not (Jaenike, 1981; Claridge 1988; see also Claridge et al., 1997). A range of new empirical studies have since been published, as have models that approximate the process of sympatric speciation more closely than their predecessors.

A particular advantage of using host race systems in empirical studies is the ease with which many can be hybridised and backcrossed in the laboratory. This makes it possible to investigate the nature and location of genes responsible for the traits at which they differ, including those which contribute to reproductive isolation in the wild. Although most studies address their formation in phytophagous insects and their role in sympatric speciation, host races have been reported in a variety of systems, and their study is relevant to many aspects of ecological and evolutionary biology.

Female races ('gentes') of the common cuckoo *Cuculus canorus*, for example, are adapted to different avian hosts but undergo male mediated gene flow (Gibbs et al., 1996; Marchetti et al., 1998). Host races have also been reported in the mistletoe genus *Phoradendron* (Clay et al., 1985; Glazner et al., 1988), and their formation appears likely in several other non phytophage systems, including mammalian fleas and ticks (Diehl and Bush, 1984), symbiotic marine shrimp (Duffy, 1996) and crabs (Stevens, 1990), and nematode worms (Jaenike and Selander, 1980).

The existence and formation of host races can be an important consideration in both the control and the conservation of biotypes. For example, Gould (1998), discusses the possible formation of tobacco budworm (*Heliothis virescens*) strains adapted to transgenic insecticidal cultivars. Host race identification, which can be complicated because most

are cryptic, is also important in more traditional forms of pest control. Clarke and Walter (1995), for instance, noted that the introduction of multiple populations of what were presumed to be the same biological control agent have often involved different host associated 'strains' or cryptic species, and that this was likely to have contributed to the failure of several pest control programmes. Conversely, the recognition of host races raises questions of how populations are chosen for conservation. Most conservation policies recognise only species in the traditional 'biological' sense - groups which do not hybridise. Host races, however, contribute to diversity despite their appreciable gene flow, and may be incipient species.

For the remainder of this review, we will however confine ourselves to the role of host races in the sympatric speciation of phytophagous insects. The enormous number of sympatric, closely related insect species specialized on different host plants (e.g. Farrell, 1998), has long led biologists to suggest the likelihood of sympatric speciation in this group (Bush, 1975; Bush and Smith, 1998; Price, 1980). The existence of host races, particularly in taxa containing sympatric host associated species, is strong evidence for speciation via host shift, because they form an important intermediate stage between host use polymorphism and 'full' species.

In this review we aim to

I. outline developments in the theoretical study of host shift speciation
II. clearly define the term 'host race' by a set of empirically testable criteria
III. identify those biotypes so far investigated which meet our criteria, and those which do not
IV. summarize current empirical evidence for the operation of host shift speciation

I Host shift speciation theory

Sympatric versus allopatric speciation

Well established allopatric models of speciation predict that physically separated populations, freed from the cohesive effects of gene flow, and subject to disruptive

selection and/or genetic drift, will accumulate many small differences whose pleiotropic effects eventually combine to form an intrinsic barrier to gene flow composed of pre- and/or post-mating isolating mechanisms (Felsenstein, 1981; Mayr, 1963). Sympatric speciation, however, differs from the theoretically less complex allopatric form in the absence of any physical restriction to gene flow between diverging forms, and by the minimal involvement of genetic drift.

While it has been argued that physical separation of populations is an essential precursor to their genetic divergence (Mayr, 1963), a number of empirical and theoretical studies have shown otherwise. Well documented examples of sympatric speciation via the appearance of impenetrable intrinsic barriers to gene flow are provided by cases of 'instantaneous' speciation by polyploidy in many varieties of domestic and wild plants (White, 1978). More generally applicable theoretical models, supported by laboratory experiments, reveal that that sympatric speciation is possible even in the presence of some gene flow, provided that the relative strength of disruptive selection is high (reviewed in Rice and Hostert, 1993).

Indirect selection for assortative mating

Important early models of sympatric speciation introduced the idea of habitat selection (Maynard Smith, 1966), describing populations in a selectively 'patchy' but geographically contiguous range, whose members, polymorphic at both habitat performance and habitat preference loci, occupied (and mated within) the patches to which they were best adapted. As mating consequently took place between like forms, habitat selection reduced gene flow between populations experiencing divergent selection pressures, shifting the balance between gene flow and disruptive selection in favour of the latter. However, this, and related models were criticised because they assume tight genetic linkage between the locus responsible for habitat and/or mate choice, and that for the trait under differential selection (Rice and Hostert, 1993). In the absence of tight linkage, recombination would destroy any emerging association between habitat preference and performance alleles (Rice and Hostert, 1993), and sympatric speciation seemed unlikely. However, recent models incorporating multiple loci suggest that sympatric speciation involving separate mate choice and diverging loci are not as unlikely as previously believed (Tregenza and Butlin, 1999). Kondrashov and Kondrashov (1999), for example, show that where multiple loci govern a trait under strong disruptive selection, gene flow between opposite phenotypic extremes will be low

even when mating is random, due to the very low fitness of their hybrid offspring. If opposite phenotypes have even weak linkage disequilibrium with (physically unlinked) alleles at a locus governing mate choice, pre-mating isolation often evolves. Dieckmann and Doebeli (1999) reach a similar conclusion from a different model, in which reproductive isolation develops despite the need for linkage disequilibrium between 'ecological' and 'marker' loci. Again, unlike the earlier models, both mating preference and the trait subject to selection are determined by many loci.

Other models particularly relevant to host shift speciation eliminate the need for linkage between mate preference and diverging loci altogether, assuming an alternative scenario in which the trait that first diverges pleoitropically reduces gene flow between the incipient species (Rice, 1984b; Rice and Hostert, 1993). According to these 'pleiotropy' models, the initial reduction in gene flow allows divergence in traits subject to less intense disruptive selection, which in turn depresses gene flow further, and so on, until reproductive isolation is complete (Rice, 1984b; Rice and Hostert, 1993). After divergence in the first trait the process is similar to the gradual accumulation of pre- and post-mating isolating mechanisms believed to take place during allopatric speciation (Rice and Hostert, 1993).

Evolution of host preference

Phytophagous insects, whose life cycle and mating behaviour is often strongly influenced by their host, are particularly likely candidates for a reduction in gene flow via pleiotropy. The trait believed to be involved is host preference, and the partially reproductively isolated forms are host races (Bush, 1975; Bush, 1994).

There is some evidence that a host shift can result from mutations in only a few or even single genes affecting host choice, as phytophagous insects of several genera can complete their life cycle using host plants on which they are never found in the wild. Female pierid and *Heliconius* butterflies, for example, may avoid ovipositing on host plants suitable for larval development (Courtney, 1981; Smiley, 1978). Similarly, the pine-feeding host race of *Zeiraphera diniana* survives as well on larch as the larch form (Baltensweiler, 1977). Ponderosa pine-feeding biotypes of the mountain pine beetle *Dendroctonus ponderosae* even have an increased fitness on jack pine, a host they rarely infest in the wild (Cerezke, 1995). Of course, the ability to survive on an alternative host in the lab does not always translate into the ability to survive in the field, because

ecological factors other than nutrition, such as levels of parasitism and predation, may come into play. Nonetheless, these results suggest that physiological changes necessary for a population to be founded on a host may often be minimal; all that is required is a genetic change in the oviposition preference of females.

There are several avenues by which a host shift could reduce gene flow between phytophagous insect populations. The most straightforward is if both males and females respond to host cues and mate on the host. However, a host shift might also reduce gene flow in less direct ways. For example, host plant chemistry can affect cuticular hydrocarbons (Stennett and Etges, 1997), and these chemicals often play an important role in mate choice (Coyne et al., 1994; Ferveur, 1997; Singer, 1998; Tregenza and Wedell, 1997). Host plant phenology may also influence the developmental timing of insects (Langor, 1989; Wood and Guttman, 1982), and seasonal isolation can be a powerful inhibitor of gene flow (Wood and Guttman, 1982).

Direct selection for assortative mating (reinforcement)

Direct selection for positive assortative mating caused by hybrid disadvantage (reinforcement) is a potentially important evolutionary force, as it increases the likelihood that partially reproductively isolated, genetically divergent populations will speciate (Noor, 1999). However, although recent models suggest that reinforcement is possible, direct experimental evidence for its operation is scant, and it remains controversial (see Noor, 1999 for critical review).

Many studies have focused on narrow hybrid zones, but reinforcement may be more likely in diverging forms that overlap very broadly (Guldemond and Dixon, 1994). Selection for positive assortative mating in populations meeting at hybrid zones occurs only in a narrow area, and theory suggests that any emerging assortative mating alleles are likely to be swamped by gene flow from the rest of the population (Barton and Hewitt, 1981; Butlin, 1990; Sanderson, 1989). One recent model has cast doubt on this idea, predicting that reinforcement is similarly likely whether population range overlap is partial or complete (Liou and Price, 1994), but another predicted that it was more likely to operate in populations inhabiting a mosaic hybrid zone than in those meeting only along a narrow tension zone (Cain et al., 1999).

II. Defining host races.

Difficulties with current host race definitions

Disagreement about the likelihood of sympatric speciation in general is not the only source of controversy in the host race literature. Another important factor is ongoing disagreement about what a host race is; even recent studies describe 'host race' systems which, in the view of other authors, are in fact sibling species, or polymorphic populations.

The need for a consistent definition of host races was first highlighted by Diehl and Bush (1984). The authors discussed several alternatives before proposing what is now perhaps the most widely quoted:

> ''a population of a species that is partially reproductively isolated from other conspecific populations as a direct consequence of adaptation to a specific host''

This definition has the advantage of conciseness, but it is not tailored to the purposes of empirical study because it is unclear exactly what traits biotypes in the field must be shown to possess in order to meet it. For example, it is not obvious what would constitute adaptation to a specific host, given that, in order for the initial host shift to occur, the only adaptation necessary is preference for a new host. Likewise, 'partial reproductive isolation', which potentially includes any trait under disruptive selection or affecting mate choice, is open to interpretation.

Another concise definition discussed by Diehl and Bush was proposed by Mayr (1970):

> ''non-interbreeding sympatric populations, which differ in biology but not, or scarcely, in morphology [and which are] prevented from interbreeding by preferences for different food plants or other hosts.''

As is the case with the previous definition, the above is not ideal for use in empirical studies. It does not, for example, suggest how biotypes whose assortative mating is

caused by differences in plant preferences are detected. In addition, as pointed out by Diehl and Bush (1984), the requirement that the biotypes be non interbreeding in fact describes host associated sibling species under the widely used biological species concept.

Jaenike (1981) was the first to propose a definition of host races consisting of a set of experimentally verifiable criteria by which they can be identified:

1. ‘‘[The populations] are sympatric, so that individuals in breeding condition in one population are within normal cruising range of those in another... ’’
2. ‘‘There must be a statistically significant genetic difference between the populations, suggesting, though not proving, that gene flow between them is not extensive’’
3. ‘‘The genetic difference (2) under consideration cannot be one that is directly related to host selection [unless] both males and females manifest genetic differentiation in host preference, and ... mating takes place on or near the host plant’’
4. ‘‘It must be shown that the genetic difference (2) is not solely the result of natural selection acting on the current generation of individuals’’
5. ‘‘Finally, if the above conditions are met it should be shown, if experimentally feasible, that the genetic difference between the two populations disappears over a period of generations when they are confined to breed on a single food type ... if the genetic differences between the two groups do not disappear ..., or if they do so initially only to reappear in subsequent generations ... then reproductive isolation between them in the field cannot be ascribed to differences in host preference. In this case the two groups represent distinct species, not host races.’’

Criteria (1), (2), and (4) of Jaenike’s definition have been incorporated into the new definition we propose below. However, in our definition the primary difference between host races and host-associated species is the occurrence of appreciable gene flow between the host races. This difference will affect the host race collapsibility of Jaenike’s criterion (5), but is potentially more easily verifiable in studies using genetic markers. While Jaenike’s definition does not exclude the possibility of gene flow between host

races, neither does it explicitly require it, but rather suggests that it is not extensive (Criterion (2)). In addition, in light of current theory Criterion (5) seems unnecessarily stringent and difficult to verify. Recent models of host shift speciation predict that correlated effects of host preference may initially be the sole reason for assortative mating between host races, but also that assortative mating between host races at later stages of divergence will be strengthened by the pleiotropic effects of a variety of behavioural and phenotypic traits, as well as by reinforcement.

A more recent set of criteria, incorporating some of Jaenike's ideas, is proposed by Bush (1992):

1. ''Individuals of different host associated populations in breeding condition must be sympatric''
2. ''Statistically significant genetic differences exist between these sympatric populations that are not directly related to host selection or solely the result of natural selection acting on a single generation''
3. ''Males and females exhibit genetic variation in host preference that results in assortative mating, i.e., mating occurs on the preferred host plant and host preference is under genetic control''
4. ''Males and females show host-associated tradeoffs in fitness''
5. ''There is no evidence of post-mating reproductive incompatibility. Hybrid incompatibility between host associated populations indicates they are sibling species, not host races''

As with Jaenike's definition, there is no direct requirement for gene flow between the races in the wild. In addition, Criterion (5) above seems unnecessary for the same reason as Jaenike's Criterion (5); current theory does not predict that host races, especially those in later stages of divergence, must lack post-mating reproductive incompatibility. Only if post-mating hybrid incompatibility depressed the level of gene flow between biotypes to a very low level would we consider such biotypes separate species. Similarly, the second part of Criterion (3), that mating must occur on the host plant, excludes the possibility of differences in host usage reducing gene flow between populations in more indirect ways. We propose the following set of criteria by which host races may be identified.

Host races:

1.a. *Use different hosts in the wild*

Without host use differentiation, host races cannot exist.

1.b. *Exhibit 'host fidelity', i.e. a preference for their natal host over alternative hosts*

Host fidelity suggests that biotype differentiation is genetic and not merely the result of host associated selection within a single generation.

2. *Coexist in sympatry in at least part of their range*

The coexistence of reproductively mature adults in study populations excludes the possibility that external barriers contribute to a reduction in gene flow.

3.a. *Display a correlation between host choice and mate choice*

Host races can only maintain their genetic differences across generations if mating is not completely random.

3.b. *Undergo actual gene flow with an appreciable frequency ($ 1% per generation), and backcross*

It is this feature alone that distinguishes host races from host associated species. Evidence for this gene flow may be obtained directly, via mark-recapture studies and observation of mating behaviour, or indirectly, via detection of linkage disequilibria between host-associated marker loci in populations on a single host (Barton et al., 1988).

4.a. *Are genetically differentiated at more than one locus*

Unlike members of a polymorphic population, host races consistently maintain correlated allelic differences at multiple loci. There are thus strong linkage disequilibria when host associated populations are grouped. In contrast, the disequilibria used as evidence for reduced gene flow in (3.b.) must be within host associated populations. The loci involved need not be marker loci; if it can be shown that multiple correlated genetic differences exist for morphology, mate choice (see also 3.a), and host survival, such that few intermediates exist, a host race designation may be hypothesised.

4.b. *Are spatially and temporally replicable, i.e. are more genetically differentiated from populations on another host in sympatry (and at the same time) than at least some geographically distant populations on the same host.*

When populations have been shown to meet criteria (1) and (4.a), a 'snapshot' of correlated genetic differences and of differences in host use has been obtained. In populations also shown to meet criterion (3.a.), these correlations are very likely to be stable over time (i.e. in the face of gene flow). Further, indirect evidence for the stability of these correlations should, whenever possible, be obtained by showing that genetic differentiation between host races in sympatry is greater than in at least some allopatric populations on the same host (4.b). In addition, because genetic drift could lead to fleeting differentiation between forms which is not being maintained across generations, it is important to have evidence that such differentiation is stable over time.

Finally, host races are likely to (5.a.) *have a higher fitness on their natal than their alternate host* and (5.b.) *produce hybrids that are less fit than parental forms.* However, host races in the earliest stages of divergence might not possess these characteristics. Thus, these two criteria are suggestive rather than diagnostic of host races, and we do not include them in the definition.

Host races in the continuum from polymorphism to species

Host races are an intermediate stage between polymorphic populations and full species, and are difficult to define because a definition of species has yet to be agreed upon. The most widely quoted species concept in evolutionary biology is the 'biological species concept', which describes species as 'reproductively isolated' populations between which there is minimal gene flow. However, the meaning of this concept has become somewhat unclear, because many sibling taxa, normally considered good species, undergo gene flow and hybridise at measurable rates (Grant and Grant, 1992; Wang et al., 1997). A number of alternative species definitions, including the ecological, mate recognition, cohesion, and phylogenetic concepts have also been proposed (Cracraft, 1989; Paterson, 1985; Templeton, 1989; Van Valen, 1976) but none of these works perfectly in all situations (see Mallet, 1995). However, a common thread linking most species concepts is that they can be viewed as mechanisms by which separate clusters of genotypes originate or are maintained. Ecological concepts, for example, highlight the role of disruptive selection, while the biological concept emphasises the role of pre- and post-mating isolation. Phylogenetic concepts are more concerned with the history of origination. Here, we employ the 'genotypic cluster' criterion of species as groups

separated by multiple correlated genetic differences in sympatry (Mallet, 1995). These correlations should be sufficient to cause a bimodal genotypic distribution, i.e. two groups of genotypes separated by intermediates that are rarer than the genotypes of the main groups (Jiggins and Mallet, 2000).

Correlations between alleles at different polymorphic loci (linkage disequilibria) can only be maintained between populations when disruptive selection is strong relative to inter-population gene flow and recombination of population specific alleles. However, provided selection is strong enough, the movement of genes from one population into another via hybridisation and backcrossing need not be zero, and so, unlike the biological concept, the genotypic cluster definition allows for some gene flow between species.

Like species, host races are clusters of genotypes separated by gaps. The difference between members of the two groups lies in the extent of 'actual gene flow' (the exchange of migrants, *sensu* Mallet 2000) and hybridisation they undergo. Between host races actual gene flow is appreciable, while between species it occurs rarely or not at all. While any dividing level of gene flow is somewhat arbitrary, given that species and host races are part of a continuum, we believe a reasonable figure for practical purposes is 1% per generation, an order of magnitude higher than the rate of hybridisation between good species that are known to hybridise (Grant and Grant, 1992; Mallet et al., 1998).

Host races differ from polymorphic populations in the pattern of differences between forms. Members of polymorphic populations may differ at more than one locus, but these differences are uncorrelated, so that individuals are placed in different groups depending upon the single phenotype or locus examined. In contrast, when members of host races are placed in groups according to multiple criteria, a bimodal distribution is observed (Mallet, 1995, see also Jiggins and Mallet 2000).

Thus, although host races are clearly distinct from polymorphic populations, they are simply a subset of genotypic cluster species, rather than a discrete category. However, the distinction is useful in studies of speciation, as the forms are an essential step in the pathway from polymorphism to species.

III. Case studies: steps in the continuum

The possibility of host race status has been suggested in the literature on a number of insect biotypes. Here, we attempt to identify those which are host races according to our criteria. Our results are summarized in Table 1. The list of cases discussed is not exhaustive, and is biased towards those systems which have been tested for several of the criteria we include in our list. Host race status has also been proposed for a number of biotypes not mentioned here, particularly within the Aphididae (Tauber and Tauber, 1989; Thieme, 1987)

Cases of single polymorphic populations

The mountain pine beetle *Dendroctonus ponderosae* (Coleoptera: Scolytidae) on lodgepole pine *Pinus contorta* and limber pine *P. ponderosae*

Despite some evidence that mountain pine beetles differ along host plant lines, current data is insufficient to support a host race designation. Male-female pairs collected from different hosts are less likely to lay fertile eggs on lodgepole pine than pairs collected from the same host, and, on limber pine only, their progeny have a lower average dry weight and fat content than the progeny of insects collected from the same host (Langor et al., 1990). However, the offspring of all cross types are fertile, differences in egg laying are not observed on limber pine, and neither development time nor mortality differs between any brood type on either host. Allozyme differentiation was initially reported (Sturgeon and Mitton, 1986) but later shown to result from selection acting within single generations (Langor and Spence, 1991). Field collected adults of the two populations do differ morphologically, but possible effects of within generation selection and/or phenotypic plasticity on morphology have not been examined (Langor and Spence, 1991). Insects on lodgepole pine begin emerging approximately one week earlier than those on limber pine, but peak and late emergence on both hosts overlaps for about two months of the year (Langor 1989).

Red or black headed biotypes of the fall webworm *Hyphantria cunea* (Lepidoptera: Tortricidae) on various hosts

The fall webworm consists of polyphagous 'red headed' and 'black headed' larval forms with overlapping host use patterns that may be sibling species (Jaenike and Selander, 1980; McIntee and Nordin, 1983; McLellan et al., 1991). Differentiation along host plant lines within (or between) colour pattern forms has not been intensively studied. However, no significant differentiation in allozyme frequency was found between populations of the red headed form on black walnut and black cherry (Jaenike and Selander, 1980).

The small ermine moth *Yponomeuta padellus* (Lepidoptera: Yponomeutidae) on hawthorn and prune

As is the case with the two examples above, there is not enough evidence to place biotypes of *Yponomeuta padellus* in the host race category. Sympatric, host associated larval populations collected from a single site differed in allozyme frequency in 1978, 1978, and 1990 (Menken, 1981; Menken, 1982; Raijmann and Menken, 2000), but supporting evidence that this differentiation is maintained in the face of gene flow, and is not simply due to within-generation selection is lacking. There is no evidence of pheromone mediated mate choice (Brookes and Butlin, 1994b), and larvae from different hosts do not differ in their preference for or fitness on various hosts in the laboratory (Kooi et al., 1991). Other factors directly affecting the extent of actual gene flow, such as adult host plant choice and hybridisation rates, have not been investigated.

Cases of possible host races

The spiraea aphid *Aphis citricola* (Homoptera: Aphididae) on satsuma and thunberg spiraea

The average emergence time of populations of the spiraea aphid on satsuma and thunberg spiraea differs by approximately one month, and is under genetic control (Komazaki, 1986; Komazaki, 1990). The two forms occur sympatrically, and laboratory-bred hybrids survive well on one of the hosts, thunberg spiraea, under field-cage conditions (Komazaki, 1986). The potential for hybridisation between these forms in the field has not however been investigated in detail, although Komazaki (1986) stresses that

there is considerable (but incomplete) allochronic isolation of adult forms (Komazaki, 1986).

The aphid *Cryptomyzus galeopsidis* (Homoptera: Aphididae) on redcurrant and blackcurrant primary hosts

Populations of the aphid *Cryptomyzus galeopsidis* on redcurrant and blackcurrant primary hosts (the hosts where sexual forms reproduce) are genetically differentiated, but will hybridise and backcross when housed together (Guldemond, 1990b; Guldemond et al., 1994; Guldemond and Dixon, 1994), although the single hybrid clone tested produced fewer mature sexual females than its parents (Guldemond, 1990b). There is some pre-mating isolation between blackcurrant biotype males and redcurrant biotype females, but hybridisation in the opposite direction occurs freely even when males have a choice of females of both biotypes (Guldemond et al., 1994). Males do not appear to differentiate between pheromones produced by females of the two biotypes (Guldemond and Dixon, 1994). Populations on both primary hosts share a secondary host, hemp nettle (where asexual forms produced later in the season feed). Migratory forms of both biotypes tend to prefer their native host, although, in the case of the redcurrant biotype, this preference was not expressed by all clones (Guldemond, 1990a).

Rhagoletis pomonella (Diptera: Tephritidae) on hawthorn and the 'flowering dogwood fly' on flowering dogwood

Although the two forms have recently been recommended for species status under a 'nonstrict' version of the biological species concept (Berlocher, 1999), we use our criteria to tentatively place them in the host race category. The two biotypes are often sympatric, and exhibit only frequency differences at up to seven of 17 polymorphic allozyme loci. However, only some components of gene flow have been directly measured, and the results are inconclusive; the populations experience partial allochronic isolation, slight post mating isolation, and have differences in host preference (Berlocher, 1999).

The pea aphid *Acyrthosiphon pisum* (Homoptera: Aphididae) on alfalfa and clover

Populations of the pea aphid *Acyrthosiphon pisum* differ in the frequency some of allozyme alleles (Via, 1999), but migration does occur between alfalfa and clover populations. A comparison of semi-diagnostic allele frequency within new host plant

fields to that within fields containing established populations suggested that approximately eleven percent of new migrants to clover fields are from alfalfa, and nine percent of migrants to alfalfa come from clover fields (Via, 1999). However, the likelihood of mating between the two forms should they meet remains unknown, and the survival of aphids migrating to the alternative host is much lower than on the natal host (Via, 1991a; Via, 1991b; Via, 1999).

The sawfly *Platycampus luridiventris* (Hymenoptera: Tenthredinidae) on two species of alder

Platycampus luridiventris associated with the alder species *Alnus glutinosa* and *A. incana* may be host races. Sympatric populations differ in larval morphology, female oviposition preference (Heitland and Pschorn-Walcher, 1992), and allozyme allele frequency (Herbst and Heitland, 1994). Larvae of both types develop faster on their natal host (Heitland and Pschorn-Walcher, 1992). However, their potential for hybridisation has not been studied in either the field or laboratory.

Cases of probable host races

The goldenrod gall ballmaker *Eurosta solidaginis* (Diptera: Tephritidae) on two goldenrod species

Biotypes of *Eurosta solidaginis* on different goldenrod hosts are significantly genetically differentiated (Waring et al., 1990), and have a higher survival on their natal than their alternative host (Craig et al., 1997; Craig et al., 1993). Mating is strongly assortative when host plants are present, but much less so when they are absent (Craig et al., 1993; Itami et al., 1998). Differences in emergence times also contribute to reproductive isolation (Craig et al., 1993). A direct estimate of the extent of gene flow between biotypes of *Eurosta* has been difficult to obtain, although in some conditions, such as high spring temperature, which reduces allochronic isolation (Itami et al., 1998) it appears particularly likely. The observation that approximately 3% of insects collected from the wild have the 'intermediate' oviposition preference also seen in laboratory hybrids suggests gene flow, as does the high fitness of some hybrids and backcrosses on particular host plant genotypes (Itami et al., 1998). Average hybrid fitness is however lower than that of pure forms (Craig et al., 1997), so backcrossing may be rare.

The apple maggot fly *Rhagoletis pomonella* (Diptera: Tephritidae) on apple and hawthorn

Apple and hawthorn infesting forms of *Rhagoletis pomonella* are an extremely well characterised pair of host races. Differences in allele frequency at six allozyme loci have been maintained within a single sympatric site throughout eleven years of study (Feder et al., 1998b, see also Feder et al., 1990; Feder et al., 1993) and there is substantial direct and indirect evidence for gene flow between the biotypes.

Field studies of several components of actual gene flow, including host preference and temporal co-occurrence of mature adults, have revealed that the rate of exchange of migrants between the two populations is approximately 6% per generation (Feder et al., 1994). The occurrence of linkage disequilibria between host associated loci within each form also suggests that gene flow is occurring (Barton et al., 1988; Feder and Bush, 1991). However, some caution is required in interpreting the disequilibria as due to gene flow because, in some locations, disequilibria existed between loci whose frequencies were the same in both biotypes (Feder and Bush, 1991). The two forms are likely to mate randomly when they encounter each other on the same host plant (Feder et al., 1994), and there is no evidence of an intrinsic reduction in hybrid fitness (Reissig and Smith, 1978).

The larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae) on larch and pine

Populations of the larch budmoth *Zeiraphera diniana* on larch and pine co-exist in mixed larch and pine forests of the French and Swiss Alps. The two forms display heritable differentiation in a number of phenotypic traits, including female pheromone blend and male pheromone response (Baltensweiler, 1977; Baltensweiler and Priesner, 1988; Emelianov et al., 1995). They also differ in allozyme frequency at three loci, and this differentiation has been stable since at least 1994 (Emelianov et al., 1995, Emelianov pers. com). There is no between-form differentiation in allele frequency at 10 of the loci examined, but very strong differentiation at the remaining three. This highly heterogeneous pattern is consistent with gene-flow-with-selection models of divergence, which predict that some regions of the genome, where selection is strong relative to gene flow, will become strongly differentiated while others, at which the reverse is true, will not. Thus, the distribution of allozyme frequency differences provides indirect evidence

of gene flow (Emelianov et al., 1995). Direct evidence of gene flow between larch and pine associated *Z. diniana* biotypes is dealt with in detail in Chapters 2 and 3 of this thesis.

Cases of sibling species

The Colorado potato beetle *Leptinotarsa decemlineata* (Coleoptera: Chysomelidae) on potato and horse nettle

Colorado potato beetles on potato and horse nettle are completely reproductively isolated in the laboratory (Boiteau, 1998).

The apple maggot fly *Rhagoletis pomonella* (Diptera: Tephritidae) on hawthorn and the blueberry maggot fly *R. mendax* on blueberry

Populations of *R. pomonella* and *R. mendax* contain unique allozyme alleles at eleven loci (Feder and Bush, 1989a). A survey of a total of 426 individuals of both types failed to reveal any putative hybrid genotypes, despite the co-occurrence of sexually mature adults on intertwined hawthorn and blueberry bushes (Feder and Bush, 1989a). This suggests that the rate of actual gene flow between the forms is considerably less that one percent per generation, and so they are considered sibling species.

The brown planthopper *Nilaparvata lugens* (Homoptera: Delphacidae) on weed grass and cultivated rice

Brown planthoppers found on the weed grass *Leerisa hexandra* and on cultivated rice display heritable differentiation in a number of traits, including mating call pulse repetition frequency. Although viable and fertile F_1 hybrids can be produced (Claridge et al., 1985; Heinrichs and Medrano, 1984), laboratory tests of mate choice suggest that the two populations do not hybridise in the wild; only a single putative hybrid was found in a crowded population cage containing males and females of both forms. Furthermore, when insects were played mating calls of members of their own population and those on the alternate host, both males and females responded only rarely, and with reduced vigour, to calls from members of the other population (Claridge et al., 1985).

Treehoppers of the *Enchenopa binotata* (Homoptera: Membracidae) complex on various (see Table 1) hosts

Members of this treehopper complex display varying levels of allele frequency differentiation, and mate assortatively in the lab (Guttman et al., 1981; Wood and Guttman, 1982). Low levels of hybridisation occurred in laboratory conditions in the absence of host plants, but there is no evidence that such hybridisation also occurs in the wild, and considerable evidence for complete allochronic isolation in field conditions, due to the timing of egg hatch with host bud burst (Wood and Guttman, 1982). Although the possibility that some of the forms may be host races has not been completely discounted, the biotypes so far discovered are most likely to hybridise at a rate of less than one percent per generation, and therefore to be sibling species under our definition.

The *Muellerianella* complex (Homoptera: Delphacidae). *Muellerianella* brevipennis, *M. fairmairei*, and *M. extrusa* on several (Table one) grass species.

Members of this complex are, somewhat tentatively, placed in the species category. In the laboratory all three species mate assortatively with conspecifics (Booij, 1982c). Hybrid broods from all interspecific crosses are much smaller than non- hybrid broods, female biased, and contain predominantly infertile males (Booij, 1982c), although backcross broods were bred from some hybrid females (Booij, 1982c). The calls by which *M. brevipennis*, *M. extrusa*, and *M. fairmari* males communicate with potential mates differ (Booij, 1982a) and although the extent to which this affects their long range cross attraction has not been directly investigated, acoustic behaviour strongly influences mate choice in closely related taxa (Booij, 1982a, and references within). Despite this, there is evidence that in some areas the forms may 'accidentally' come into contact with each other, and hybridise. Hosts of different species are sometimes found in close proximity (Booij, 1982b), and two putative *M. farimairei* x *M. brevipennis* hybrid females have been collected from one such site (Booij, 1982c). The balance of current evidence does however point towards a level of hybridisation of less than 1% per generation, and the very low chance of backcrossing also supports their species status.

The fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae) on corn and rice

Rice and corn associated *Spodoptera frugiperda* can produce viable hybrids in the laboratory (Pashley and Martin, 1987; Whitford et al., 1988), but there is evidence of hybrid sterility in at least one cross direction (Pashley and Martin, 1987). Both long range cross attraction and hybridisation at close range are highly (but incompletely) assortative (Pashley et al., 1992), the latter at least in part because females call at different times of the night. The two forms display frequency differentiation at several allozyme loci (Pashley, 1989a), diagnostic differentiation in an RFLP tandem repeat marker (Lu et al., 1992), and, on balance, appear to be cryptic sibling species rather than host races. However, the rates of hybridization in nature have not been estimated.

IV. Empirical evidence for host shift speciation

Do host races exist ?

Yes. By far the most parsimonious interpretation of the *Rhagoletis pomonella* data is that apple and hawthorn associated populations are host races. Replicable genetic differentiation between the two has been exhaustively demonstrated, and there is very strong direct evidence that the level of actual gene flow between the two populations is approximately 6% in areas of overlap.

Similar evidence, discussed throughout the remainder of this thesis, suggests that *Zeiraphera diniana* on larch and pine are also host races, although in this case the direct estimate of gene flow is lower, approximately 2.4% per generation (Chapter 2). *Eurosta solidaginis* on *Solidago altissima* and *S. gigantea*, are also very likely to be host races. Again, genetic differentiation along host plant lines has been clearly demonstrated, and some gene flow between the forms is likely. However, the level of this gene flow has not been directly estimated, and there is evidence in that events after migration (i.e. reduced survival of *E. solidaginis* hybrids) are likely to inhibit backcrossing.

How common are host races?

Phytophagous host races been confirmed in two of the 17 studies described here, approximately 12% of the sympatric, host associated populations in which the possibility has been investigated. This conservative estimate is likely to rise by up to 40%, because current data suggests, or strongly suggests, the presence of seven more, within aphids *Aphis citricola*, *Acythrosiphon pisum*, and *Cryptomyzus galeopsidis*, the sawfly *Platycampus luridiventris*, the goldenrod gall ballmaker *Eurosta solidaginis*, and the apple maggot fly R. *pomonella* vs. the flowering dogwood fly.

The studies carried out to date have concentrated on systems in which the presence of host races seems particularly likely, in most cases because there is evidence of host-associated differentiation between populations of a presumed sympatric species. How common is cryptic, host-associated differentiation in phytophagous insects? Recent studies, benefiting from sensitive protein and DNA markers, suggest that this type of cryptic differentiation could be quite frequent in phytophagous insects. Taxonomic revisions of several taxa, by taking account of new behavioural and DNA or protein based characters, have detected sympatric, monophagous biotypes within presumed polyphagous species – recent examples include bark beetles *Dendroctonus brevicomis* (Kelley et al., 1999) and fruit flies of the genus *Blepharoneura* (Condon and Steck, 1997). Intensive studies of agricultural pests have also begun to uncover population substructuring along host plant lines (e.g. Shufran et al., 2000), the starting point of many studies described here. This differentiation does not seem limited to particular phytophage taxa; a wide variety of phytophage genera are represented.

Consequently, although the number of phytophage host races discovered so far is small, the number of insect systems that conceals them is potentially huge, and, if the pattern observed in the data presented here is representative of these systems, host races could indeed be a common phenomenon.

Because the shortage of detailed studies of host races noted earlier (Tauber and Tauber, 1989) continues, it is not possible to draw a more definite conclusion from the current empirical data. Of the 17 cases described here, there is insufficient data to resolve the status of about 10. The lack of data is most pronounced in relation to criteria (3.a) and (3.b), those dealing with the possibility of gene flow between the forms. Members of most likely systems show differentiation along host plant lines, will interbreed in the

laboratory, and have levels of genetic differentiation compatible with continuing gene flow, but while several components of gene flow have often been studied, direct or indirect estimates of actual gene flow have rarely been obtained.

Is the formation of host races likely to lead to sympatric speciation?

Did host races in sympatry today undergo their divergence during a period of isolation in allopatry? In the case of *Rhagoletis pomonella* on apple and hawthorn, at least, they did not. Historical records show that the apple host was introduced within the range of hawthorn, and it is extremely unlikely that the two were ever isolated in allopatry (Bush, 1969a; Bush, 1994).

A second line of evidence suggesting that host races are likely to form, and to diverge to the point of speciation, is their geographical distribution. A vast number of host specialised insect species have evolved – many millions of beetle species alone (Farrell, 1998). Many of these phytophagous insect species are identifiable primarily by their host, and are capable of producing viable hybrids in the laboratory. Given their numbers, and the fact that sympatric speciation in these groups is theoretically possible, a sympatric origin for these species seems much more parsimonious than the alternative, that all were separated and diverged in allopatry, obtaining their present sympatric distribution only as a result of secondary contact (Tauber and Tauber, 1989).

Further support for host shift speciation has been provided by comparisons of several host associated biotypes in *Rhagoletis*. The observation that *Rhagoletis pomonella* and *Rhagoletis mendax*, the apple and blueberry maggot flies, are good sympatric species, while host races of *Rhagoletis pomonella* exist on apple and haw, provides a continuity argument that the formation of host races in this genus is likely to lead to speciation (Feder et al., 1998b; Payne and Berlocher, 1995). In addition, comparison between the *Rhagoletis suavis* and *pomonella* species groups shows that members of the first have an allopatric or parapatric distribution, and are not specialized on different hosts, whereas members of the latter, which are all sympatric, are also all restricted to different hosts (Bush and Smith, 1998). Again, this argues for an important role of host shift in the formation of sympatric *Rhagoletis* species.

Finally, the main difficulty for the theory of sympatric speciation has always been to explain how selection can cause multilocus differentiation that is correlated with habitat

use and mate choice, in the presence of gene flow. By providing a continuum of examples in which host-associated differentiation is maintained in spite of actual or probable gene flow, the studies discussed here provide empirical evidence for a route to sympatric speciation.

Conclusions

Convincing evidence that host races exist, and are intermediate steps on the route to sympatric speciation has been provided by intensive studies of Tephritid flies of the genus *Rhagoletis*. However, although present empirical studies suggest the possibility that host races are common, it remains difficult to use direct evidence to draw conclusions about the frequency of sympatric speciation by host shift. Although current examples suggest a stable route for sympatric speciation via host shift, the possibility of a break in this route remains. In the majority of studies of potential host races discussed here, firm conclusions about the level of gene flow have not been reached. This measurement is however vital in unambiguously testing the host races status of likely systems according to the criteria suggested in this review, and is the best way of investigating the possibility of 'gene flow gaps' in the current array of sympatric host-associated forms. In addition, a number of insect biotypes which have been presented as host races clearly represent polymorphisms or sibling species. Part of this confusion is likely to have arisen due to ambigous host race criteria which have been difficult to test and open to interpretation.

References

Baltensweiler, W. 1977. Colour polymorphism and dynamics of larch bud moth populations (*Zeiraphera diniana* Gn. Lep. Tortricidae). *Mitteilungen - Schweizerische Entomologische Gesellschaft*, 50:15-23.

Baltensweiler, W. 1993. A contribution to the explaination of the larch budmoth cycle, the polymorphic fitness hypothesis. *Oecologia*, 93:251-255.

Baltensweiler, W., and E. Priesner. 1988. A study of pheromone polymorphism in *Zeiraphera diniana* Gn (Lep, Tortricidae).3. Specificity of attraction to synthetic pheromone sources by different male-response types from two host races. *Journal of Applied Entomology*, 106:217-231.

Barton, N. H., and G. M. Hewitt. 1981. Hybrid zones and speciation. In W. R. Atchley and D. S. Woodruff (eds.), *Evolution and Speciation, Essays in Honour of M.J.D White*, Vol. 109-145. Cambridge University Press, Cambridge.

Barton, N. H., J. S. Jones, and J. Mallet. 1988. No barriers to speciation. *Nature*, 336:13-14.

Berlocher, S. H. 1995. Population structure of *Rhagoletis mendax*, the blueberry maggot. *Heredity*, 74:542-555.

Berlocher, S. H. 1999. Host races or species? Allozyme characterization of the 'flowering dogwood fly', a member of the *Rhagoletis pomonella* complex. *Heredity*, 83:652-662.

Berlocher, S. H., and B. A. McPheron. 1996. Population structure of *Rhagoletis pomonella*, the apple maggot fly. *Heredity*, 77:83-99.

Boiteau, G. 1998. Reproductive barriers between the partially sympatric colorado and false potato beetles. *Entomologia Experimentalis et applicata*, 89:147-153.

Booij, C. J. H. 1982a. Biosystematics of the *Muellerianella* complex (Homoptera, Delphacidae), interspecific and geographic variation in acoustic behaviour. *Zeitschrift Für Tierpsychologie*, 58:31-52.

Booij, C. J. H. 1982b. Biosystematics of the *Muellerianella* Complex (Homoptera, Delphacidae) - Host-Plants, Habitats and Phenology. *Ecological Entomology*, 7:9-18.

Booij, C. J. H. 1982c. Biosystematics of the *Muellerianella* Complex (Homoptera, Delphacidae), Hybridization Studies. *Genetica*, 57:161-170.

Brookes, M. I., and R. K. Butlin. 1994b. Population structure in the small ermine moth *Yponomeuta padellus* - an estimate of male dispersal. *Ecological Entomology*, 19:97-107.

Brown, J. M., W. G. Abrahamson, and P. A. Way. 1996. Mitochondrial DNA phylogeography of host races of the goldenrod ball gallmaker, *Eurosta solidaginis* (Diptera: Tephritidae). *Evolution*, 50:777-786.

Bush, G. 1975. Modes of animal speciation. *Annual Review of Ecology and Systematics*, 6:339-364.

Bush, G., and J. Smith. 1998. The genetics and ecology of sympatric speciation: A case study. *Researches on Population Ecology*, 40:175-187.

Bush, G. L. 1969a. Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletiis* (Diptera:Tephritidae). *Evolution*, 23:237-251.

Bush, G. L. 1992. Host race formation and sympatric speciation in *Rhagoletis* fruit flies (Diptera:Tephritidae). *Psyche*, 99:335-357.

Bush, G. L. 1994. Sympatric speciation in animals - new wine in old bottles. *Trends in Ecology & Evolution*, 9:285-288.

Butlin, R. K. 1990. Divergence in emergence time of host races due to differential gene flow. *Heredity*, 65:47-50.

Cain, M. L., V. Andreasen, and D. J. Howard. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. *Evolution*, 53:1343-1353.

Cerezke, H. F. 1995. Egg gallery, brood production, and adult characteristics of mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae), in three pine hosts. *Canadian Entomologist*, 127:955-965.

Claridge, M. F., 1988. Species concepts and speciation in parasites. In D L Hawksworth (ed). *Prospects in Systematics*, 92-111. Oxford University Press, Oxford

Claridge, M. F., H. A. Dawah, and M. R. Wilson. 1997. Species in insect herbivores and parasitoids - sibling species, host races and biotypes. In M. F. Claridge, H. A. Dawah, and M. R. Wilson (eds). *Species: The units of biodiversity*. Chapman and Hall. London

Claridge, M. F., J. den Hollander, and J. C. Morgan. 1985. The status of weed-associated populations of the brown planthopper, *Nilaparvata lugens* (Stal) - host race or biological species. *Zoological Journal of the Linnean Society*, 84:77-90.

Clarke, A. R., and G. H. Walter. 1995. 'Strains' and the classical biological control of insect pests. *Canadian Journal of Zoology*, 73:1777-1790.

Clay, K., D. Dement, and M. Rejmanek. 1985. Experimental evidence for host races in mistletoe (*Phoradendron tomentosum*). *American Journal of Botany*, 72:1225-1231.

Condon, M. A., and G. J. Steck. 1997. Evolution of host use in fruit flies of the genus *Blepharoneura* (Diptera: Tephrititdae): Cryptic species on sexually dimorphic host plants. *Biological Journal of the Linnean Society*, 60:443-466.

Courtney, S. P. 1981. Coevolution of pierid butterflies and their cruciferous foodplants. 3. *Anthocharis-cardamines* (L) survival, development and oviposition on different hostplants. *Oecologia*, 51:91-96.

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

Cracraft, J. 1989. . In D. Otte and J. A. Endler (eds.), *Speciation and its consequences*, pp. 28-59. Sinauer.

Craig, T. P., J. D. Horner, and J. K. Itami. 1997. Hybridization studies on the host races of *Eurosta solidaginis*: Implications for sympatric speciation. *Evolution*, 51:1552-1560.

Craig, T. P., J. K. Itami, W. G. Abrahamson, and J. D. Horner. 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. *Bulletin of Entomological Research*, 74:47-64.

Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature*, 400:354-357.

Diehl, S. R., and G. L. Bush. 1984. An evolutionary and applied perspective of insect biotypes. *Annual Review of Entomology*, 29:471-504.

Duffy, J. E. 1996. Species boundaries, specialization, and the radiation of sponge-dwelling alpheid shrimp. *Biological Journal of the Linnean Society*, 58:307-324.

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the larch budmoth) - polymorphism, host races or sibling species. *Heredity*, 75:416-424.

Farrell, B. 1998. 'Inordinate fondness' explained: Why are there so many beetles? *Science*, 281:555-559.

Feder, J., S. Berlocher, and S. Opp. 1998a. Sympatric host race formation and speciation in *Rhagoletis* (Diptera:Tephritidae) : A Tale of Two Species for Charles D. In S. Mopper and S. Strauss (eds.), *Genetic Structure and Local Adaptation in Natural Insect Populations:Effects of Ecology, Life History, and Behaviour*, pp. 408-411. Chapman and Hall, London.

Feder, J. L., and G. L. Bush. 1989a. A field test of differential host-plant usage between two sibling species of *Rhagoletis pomonella* fruit flies (Diptera: Tephritidae) and its consequences for sympatric models of speciation. *Evolution*, 43:1813-1819.

Feder, J. L., and G. L. Bush. 1989b. Gene frequency clines for host races of *Rhagoletis pomonella* in the midwestern United States. *Heredity*, 63:245-266.

Feder, J. L., and G. L. Bush. 1991. Genetic variation among apple and hawthorn host races of *Rhagoletis pomonella* across an ecological transition zone in the midwestern United-States. *Entomologia Experimentalis et Applicata*, 59:249-265.

Feder, J. L., C. A. Chilcote, and G. L. Bush. 1989. Are the apple maggot, *Rhagoletis pomonella*, and blueberry maggot, *Rhagoletis mendax*, distinct species - implications for sympatric speciation. *Entomologia Experimentata et Applicata*, 51:113-123.

Feder, J. L., C. A. Chilcote, and G. L. Bush. 1990. Regional, local and microgeographic allele frequency variation between apple and hawthorn populations of *Rhagoletis pomonella* in Western Michigan. *Evolution*, 44:595-608.

Feder, J. L., T. A. Hunt, and L. Bush. 1993. The effects of climate, host-plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella*. *Entomologia Experimentalis et Applicata*, 69:117-135.

Feder, J. L., S. B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America*, 91:7990-7994.

Feder, J. L., J. B. Roethele, B. Wlazlo, and S. H. Berlocher. 1997a. Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America*, 94:11417-11421.

Feder, J. L., U. Stolz, K. M. Lewis, W. Perry, J. B. Roethele, and A. Rogers. 1997b. The effects of winter length on the genetics of apple and hawthorn races of *Rhagoletis pomonella* (Diptera : Tephritidae). *Evolution*, 51:1862-1876.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals. *Evolution*, 35:124-138.

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

Filchak, K. E., J. L. Feder, J. B. Roethele, and U. Stolz. 1999. A field test for host-plant dependent selection on larvae of the apple maggot fly, *Rhagoletis pomonella*. *Evolution*, 53:187-200.

Gibbs, H. L., M. L. Brooke, and N. B. Davies. 1996. Analysis of genetic differentiation of host races of the common cuckoo *Cuculus canorus* using mitochondrial and microsatellite DNA variation. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 263:89-96.

Glazner, T. G., B. Devlin, and N. C. Ellstrand. 1988. Biochemical and morphological evidence for host race evolution in desert mistletoe, *Phoradendron californicum* (Viscaceae). *Plant Systematics and Evolution*, 161:13-21.

Gould, F. 1998. Sustainability of transgenic insecticidal cultivars: Integrating pest genetics and ecology. *Annual Review of Entomology*, 43:701-726.

Grant, P. R., and B. R. Grant. 1992. Hybridization of bird species. *Science*, 256:193-197.

Guldemond, J. A. 1990a. Choice of host plant as a factor in reproductive isolation of the aphid genus *Cryptomyzus* (Homoptera, Aphididae). *Ecological Entomology*, 15:43-51.

Guldemond, J. A. 1990b. Evolutionary genetics of the aphid *Cryptomyzus*, with a preliminary analysis of the inheritance of host plant preference, reproductive performance and host alteration. *Entomologia Experimentalis et Applicata*, 57:65-76.

Guldemond, J. A., and A. F. G. Dixon. 1994. Specificity and daily cycle of release of sex-pheromones in aphids - a case of reinforcement. *Biological Journal of the Linnean Society*, 52:287-303.

Guldemond, J. A., A. F. G. Dixon, and W. T. Tigges. 1994. Mate recognition in *Cryptomyzus* aphids - copulation and insemination. *Entomologia Experimentalis et Applicata*, 73:67-75.

Guttman, S. I., T. K. Wood, and A. A. Karlin. 1981. Genetic differentiation along host plant lines in the sympatric *Enchenopa binotata* Say complex (Homoptera, Membracidae). *Evolution*, 35:205-217.

Heinrichs, E., and F. Medrano. 1984. *Leersia hexandra*, a weed host of the rice brown planthopper, *Nilaparvata lugens* (Stal). *Crop Protection*, 3:77-85.

Heitland, W., and H. Pschorn-Walcher. 1992. Biological differences between populations of *Platycampus luridiventris* feeding on different species of alder (Hymenoptera: Tenthredinidae). *Entomologia Generalis.*

Herbst, J., and W. Heitland. 1994. Genetic differentiation among populations of the sawfly species *Platycampus-luridiventris*, associated with different alder species (Hymenoptera, Tenthredinidae). *Entomologia Generalis*, 19:39-48.

Itami, J., T. Craig, and J. Horner. 1998. Factors affecting gene flow between the host races of *Eurosta solidaginis*. In S. Mopper and S. Strauss (eds.), *Genetic structure and local adaptation in natural insect populations*, pp. 375-404. Chapman and Hall, New York.

Jaenike, J. 1981. Criteria for ascertaining the existence of host races. *American Naturalist*, 117:830-834.

Jaenike, J., and R. K. Selander. 1980. On the question of host races in the fall webworm, *Hyphantria cunea*. *Entomologia Experimentata et Applicata*, 27:31-37

Jiggins, C. D., and J. Mallet. 2000. Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution*, 15:250-255.

Kelley, S. T., J. B. Mitton, and T. D. Paine. 1999. Strong differentiation in mitochondrial DNA of *Dendroctonus brevicomis* (Coleoptera: Scolytidae) on different subspecies of ponderosa pine. *Annals of the Entomological Society of America*, 92:193-197.

Komazaki, S. 1986. The inheritance of egg hatching time of the spiraea aphid, *Aphis citricola* Van der Goot (Homoptera, Aphididae) on the two winter hosts. *Kontyû*, 54:48-53.

Komazaki, S. 1990. Variation in the hatch timing of the overwintering egg among populations of *Aphis spiraecola* Patch (Homoptera: Aphididae) collected from different host plants and localities in Japan. *Applied Entomology and Zoology*, 25:27-34.

Kondrashov, A. S., and F. A. Kondrashov. 1999. Interactions among quantitative traits in the course of sympatric speciation. *Nature*, 400:351-354.

Kooi, R. E., T. P. M. Vandewater, and W. M. Herrebout. 1991. Host plant selection and larval food acceptance by *Yponomeuta padellus*. *Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen Series C-Biological Sciences*, 94:221-232.

Langor, D. W. 1989. Host effects on the phenology, development, and mortality of field populations of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera, Scolytidae). *Canadian Entomologist*, 121:149-157.

Langor, D. W., and J. R. Spence. 1991. Host effects on allozyme and morphological variation of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera, Scolytidae). *Canadian Entomologist*, 123:395-410.

Langor, D. W., J. R. Spence, and G. R. Pohl. 1990. Host effects on fertility and reproductive success of *Dendroctonus ponderosae* Hopkins (Coleoptera, Scolytidae). *Evolution*, 44:609-618.

Liou, L. W., and T. D. Price. 1994. Speciation by reinforcement of premating isolation. *Evolution*, 48:1451-1459.

Lu, Y.-J., M. J. Adang, D. J. Isenhour, and G. D. Kochert. 1992. RFLP analysis of genetic variation in North American populations of th fall armyworm moth *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Molecular Ecology*, 1:199-208.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution*, 10:294-299.

Mallet, J., W. O. McMillan, and C. D. Jiggins. 1998. Estimating the mating behaviour of a pair of hybridising *Heliconius* species in the wild. *Evolution*, 52:503-510.

Marchetti, K., H. Nakamura, and H. Gibbs. 1998. Host race formation in the common cuckoo. *Science*, 282:471-472.

Maynard Smith, J. 1966. Sympatric speciation. *American Naturalist*:637-650.

Mayr, E. 1963. *Animal species and evolution*. Belknap press of Harvard University Press, Cambridge, MA.

Mayr, E. 1970. *Populations, species, and evolution*. Harvard University Press, Cambridge, MA.

McIntee, E. E., and G. L. Nordin. 1983. Electrophoretic comparison of two types of fall webworm, *Hyphantria-cunea* (Drury) (Lepidoptera, Arctiidae), occurring in central Kentucky. *Journal of the Kansas Entomological Society*, 56:190-198.

McLellan, K. A. M., G. L. Nordin, and K. F. Haynes. 1991. Chemical communication and reproductive isolation in 2 Types of the fall webworm (Lepidoptera, Arctiidae). *Annals of the Entomological Society of America*, 84:118-123.

Menken, S. B. J. 1981. Host races and sympatric speciation in small ermine moths, Yponomeutidae. *Entomologia Experimentalis et Applicata*, 30:280-292.

Menken, S. B. J. 1982. Biochemical genetics and systematics of small ermine moths (Lepidoptera, Yponomeutidae). *Zeitschrift fur Zoologische Systematik und Evolutionsforschung*, 20:131-143.

Mopper, S., and S. Strauss. 1998. *Genetic structure and local adaptation in natural insect populations*. Chapman and Hall, New York.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. *Heredity*, 83:503-508.

Pashley, D. P. 1989a. Host associated differentiation in armyworms (Lepidoptera: Noctuidae): An allozymic and mitochondrial DNA perspective. In H. D. Loxdale and J. d. Hollander (eds.), *Systematics Association Special Volume No. 39*, pp. 103-114. Clarendon Press, Oxford.

Pashley, D. P. 1989b. Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): a sibling species complex? *Annals of the Entomological Society of America*, 79:898-904.

Pashley, D. P., A. M. Hammond, and T. N. Hardy. 1992. Reproductive isolating mechansims in fall armyworm host strains (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America*, 85:400-405.

Pashley, D. P., and J. A. Martin. 1987. Reproductive incompatibility between host strains of the fall armyworm (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America*, 80:731-733.

Paterson, H. E. H. 1985. The recognition concept of species. In E. S. Vrba (ed.), *Species and speciation*, pp. 21-29. Transvaal Museum, Pretoria.

Payne, J. A., and S. H. Berlocher. 1995. Phenological and electrophoretic evidence for a new blueberry-infesting species in the *Rhagoletis pomonella* sibling species complex. *Entomologia Experimentata et Applicata*, 75:183-187.

Price, P. W. 1980. *Evolutionary Biology of Parasites*. Princeton University Press, Princeton.

Priesner, E., and W. Baltensweiler. 1987a. A study of pheromone polymorphism in *Zeiraphera diniana* Gn (Lep, Tortricidae). 1. Male pheromonal response types in european wild populations, 1978-85. *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie*, 104:234-256.

Priesner, E., and W. Baltensweiler. 1987b. A study of pheromone polymorphism in *Zeiraphera diniana* Gn (Lep, Tortricidae).2. Pheromonal response types in F1 hybrids between three host races. *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie*, 104:433-448.

Raijmann, L. E., and S. B. J. Menken. 2000. Temporal variation in the genetic structure of host-associated populations of the small ermine moth *Yponomeuta padellus* (Lepidoptera: Yponomeutidae). *Biological Journal of the Linnean Society*, 70:555-570.

Reissig, W. H., and D. C. Smith. 1978. Bionomics of *Rhagoletis pomonella* on *Crataegus*. *Annals of the Entomological Society of America*, 71:155-159.

Rice, W. R. 1984b. Disruptive selection on habitat preference and the evolution of reproductive isolation - a simulation study. *Evolution*, 38:1251-1260.

Rice, W. R., and E. E. Hostert. 1993. Laboratory experiments on speciation - what have we learned in 40 years. *Evolution*, 47:1637-1653.

Sanderson, N. 1989. Can gene flow prevent reinforcement. *Evolution*, 43:1223-1235.

Shufran, K. A., J. D. Burd, J. A. Anstead, and G. Lushai. 2000. Mitochondrial DNA sequence divergence among greenbug (Homoptera: Aphididae) biotypes: evidence for host-adapted races. *Insect Molecular Biology*, 9:179-184.

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

Smiley, J. 1978. Plant chemistry and the evolution of host specificity: New evidence from *Heliconius* and *Passiflora*. *Science*, 201:745-747.

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

Stevens, P. M. 1990. Specificity of host recognition of individuals from different host races of symbiotic pea crabs (Decapoda: Pinnotheridae). *Journal of Experimental Marine Biology and Ecology*, 143:193-207.

Strong, D. R., J. H. Lawton, and T. R. E. Southwood. 1984. *Insects on plants*. Harvard and University Press, Cambridge, MA.

Sturgeon, K. B., and J. B. Mitton. 1986. Allozyme and morphological differentiation of mountain pine beetles *Dendroctonus ponderosae* Hopkins (Coleoptera, Scolytidae) associated with host tree. *Evolution*, 40:290-302.

Tauber, C. A., and M. J. Tauber. 1989. Sympatric speciation in insects: perception and perspective. In D. Otte and J. A. Endler (eds.), *Speciation and its Consequences*, pp. 307-344. Sinauer Associates, Massachusetts.

Templeton, A. R. 1989. In O. D and J. A. Endler (eds.), *Speciation and its consequences*, pp. 3-27. Sinauer.

Thieme. 1987. In J. Holman (ed.), *Population structure, genetics and taxonomy of aphids and Thysanoptera: Proceedings of International Symposia held at Smolenice, Czechoslovakia*. SPB Academic Publishing, The Hague.

Tregenza, T., and R. K. Butlin. 1999. Speciation without isolation. *Nature*, 400:513-513.

Tregenza, T., and N. Wedell. 1997. Definitive evidence for cuticular pheromones in a cricket. *Animal Behaviour*, 54:979-984.

Van Valen, L. M. 1976. Ecological species, multispecies, and oaks. *Taxon*, 25:233-239.

Via, S. 1991a. The genetic structure of host plant adaptation in a spatial patchwork - demographic variability among reciprocally transplanted pea aphid clones. *Evolution*, 45:827-852.

Via, S. 1991b. Specialized host plant performance of pea aphid clones is not altered by experience. *Ecology*, 72:1420-1427.

Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, 53:1446-1457.

Wang, R. L., J. Wakely, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close realatives. *Genetics*, 147:1091-1106.

Waring, G. L., W. G. Abrahamson, and D. J. Howard. 1990. Genetic differentiation among host associated populations of the gallmaker *Eurosta solidaginis* (Diptera, Tephritidae). *Evolution*, 44:1648-1655.

White, M. J. D. 1978. *Modes of Speciation*. Freeman, San Francisco.

Whitford, F., S. S. Quisenberry, T. J. Riley, and J. W. Lee. 1988. Oviposition preference, mating compatibility, and development of two fall armyworm strains. *Florida Entomologist*, 71:234-243.

Wood, T. K., and S. I. Guttman. 1982. Ecological and behavioral basis for reproductive isolation in the sympatric *Enchenopa binotata* complex (Homoptera, Membracidae). *Evolution*, 36:233-242.

Wood, T. K., and S. I. Guttman. 1983. *Enchenopa binoatata* complex: sympatric speciation? *Science,* 220:310-312.

Table 1. Case studies of host associated populations.

'+' : there is data to support the criterion, '-' : there is evidence against the criterion, '+/-' data is conflicting or inconclusive, '?': published data relevent to the criterion was not found

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3 .a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Mountain pine beetle *Dendroctonus ponderosae* Hopkins (Coleoptera : Scolytidae)	Lodgepole pine *Pinus contorta* Ponderosa pine *Pinus ponderosae*	?	+	?	?	+/-	?	-	+/-	No Polymorphism	(Cerezke, 1995; Langor, 1989; Langor and Spence, 1991; Langor et al., 1990) but see also (Sturgeon and Mitton, 1986)
Goldenrod gall ballmaker *Eurosta solidaginis* (Diptera : Tephritidae)	Goldenrod sp. *Solidago altissima* *S. gigantea*	+	+	+	+/-	+	+	+	+/-	?Yes	(Brown et al., 1996; Craig et al., 1997; Craig et al., 1993; Itami et al., 1998; Waring et al., 1990)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Apple maggot fly *Rhagoletis pomonella* (Diptera : Tephritidae)	Apple *Malus pumila* Hawthorn *Crategus mollis*	+	+	+	+	+	+	+	-	Yes	(Berlocher and McPheron, 1996; Bush, 1975; Bush and Smith, 1998; Bush, 1969a; Bush, 1994; Feder et al., 1998a; Feder and Bush, 1989b; Feder and Bush, 1991; Feder et al., 1990; Feder et al., 1993; Feder et al., 1994; Feder et al., 1997a; Feder et al., 1997b; Filchak et al., 1999; Reissig and Smith, 1978)
Apple maggot fly *Rhagoletis pomonella* Blueberry maggot fly *Rhagoletis mendax* (Diptera: Tephritidae)	Hawthorn *Crategus* spp. Blueberry *Vaccinium corymbosum*	+	+	+	-	+	+	?	?	No Sibling species	(Berlocher, 1995; Feder and Bush, 1989a; Feder et al., 1989)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Apple maggot fly *Rhagoletis pomonella* 'Flowering dogwood fly' (Diptera: Tephritidae)	Hawthorn *Crategus mollis* Flowering dogwood *Cornus florida*	+	+	+	+ / -	+	+	?	+	? Yes	(Berlocher, 1999)
Spiraea aphid *Aphis citricola* Van der Goot (Homoptera: Aphididae)	Thunberg spiraea *Spiraea thunbergii* Satsuma *Citrus unshiu*	?	+	+	?	+	+	+	+	?	(Komazaki, 1986; Komazaki, 1990)
Aphids *Cryptomyzus galeopsidis* (Homoptera : Aphididae)	Redcurrant *Ribes rubrum* Blackcurrant *Ribes nigrum*	+	+	+ / -	+ / -	?	?	+	+	?Yes	(Guldemond, 1990a; Guldemond, 1990b; Guldemond and Dixon, 1994; Guldemond et al., 1994)
Pea aphids *Acrthosiphon pisum* Harris (Homoptera : Aphididae)	Alfalfa *Medicago sativa* Red clover *Trifolium pratense*	+	+	?	?	+	+	+	?	?Yes	(Via, 1991a; Via, 1991b; Via, 1999)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
The *Muellerianella* complex: (Homoptera: Delphacidae) *M. brevipennis* *M. fairmairei* *M. extrusa*	Velvet grass *Holcus lanatus* Soft grass *H. mollis* Tufted hair grass *Deschampsia cespitiosa* Purple moor grass *Molina caerulea*	+	+	+	+ / -	+	?	?	+	No Sibling species	(Booij, 1982a; Booij, 1982b; Booij, 1982c)
Brown planthopper *Nilaparvata lugens* (Homoptera : Delphacidae)	Weed grass *Leersia hexandra* Cultivated rice *Oryza sativa*	+	+	+	-	+	+	+	-	No Sibling species	(Claridge et al., 1985; Heinrichs and Medrano, 1984)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Treehoppers *Enchenopa binotata* (Homoptera : Membracidae)	Hoptree *Ptela trifoliata* Bittersweet *Celastrus scandens* Black locust *Robinia pseudacacia* Redbud *Cercis canadensis* Black walnut *Juglans nigra* Butternut *Juglans cinerea* *Viburnum* sp.	+	+	+	-	+	?	?	?	?No Sibling species	(Guttman et al., 1981; Wood and Guttman, 1982; Wood and Guttman, 1983)
Sawfly *Platycampus luridiventris* (Hymenoptera: Tenthredinidae)	Alder spp. *Alnus glutinosa* *Alnus incana*	+	+	?	?	+	?	?	?	?Yes	(Heitland and Pschorn-Walcher, 1992; Herbst and Heitland, 1994)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Fall webworm *Hyphantria cunea* (Lepidoptera : Tortricidae)	Biotypes are not host associated	-	+	n / a	n / a	n / a	n / a	-	?	No Sibling species	(Jaenike and Selander, 1980; McIntee and Nordin, 1983; McLellan et al., 1991)
Larch budmoth *Zeiraphera diniana* (Lepidoptera : Tortricidae)	Larch *Larix* spp. Pine *Pinus cembra*	+	+	+	+	+	+	+ / -	?	Yes	(Baltensweiler, 1977; Baltensweiler, 1993; Baltensweiler and Priesner, 1988; Day, 1984; Emelianov et al., 1995; Priesner and Baltensweiler, 1987a; Priesner and Baltensweiler, 1987b)

Criteria for host race status											
Case study	1.a. Different hosts	1.b. Exhibit host fidelity	2. Coexist in sympatry	3.a. Mate assortatively	3.b. Hybridise and backcross in the wild	4.a. Genetically differentiated	4.b. Differentiation is replicable	(5.a Adapted to natal host)	(5.b. Hybrids less fit than parentals)	Host race status	References
Small ermine moth *Yponomeuta padellus* (Lepidoptera : Yponomeutidae)	Hawthorn *Crataegus monogyna* Blackthorn *Prunus spinosa*	-	+	?	?	+	+	-	?	? No Polymorphism	(Brookes and Butlin, 1994b; Kooi et al., 1991; Menken, 1981; Menken, 1982; Raijmann and Menken, 2000)
Colorado potato beetle *Leptinotarsa decemlineata* Say False potato beetle *Leptinotarsa juncta* Germar	Horse nettle *Solanum carolinense* Nightshade *Solanum dulcmara*	?	+	+	-	?	?	?	n / a	No Sibling species	(Boiteau, 1998)
Fall armyworm *Spodoptera frugiperda* Smith	Corn *Zea mays* L. Rice *Oryza sativa* L.	?	+	+	+ / -	+	+	+	?	? No Sibling species	(Lu et al., 1992; Pashley, 1989a; Pashley, 1989b; Pashley et al., 1992; Pashley and Martin, 1987; Whitford et al., 1988)

Chapter 2

Hybridisation between host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae)

Abstract

A defining trait of host races is the occurrence of appreciable ($1% per generation) 'actual gene flow', or exchange of genes via hybridisation. Using a quartet mate choice design (one male and one female of each of two biotypes per cage), we estimate that the overall degree of hybridisation between larch and pine biotypes of *Z. diniana* is approximately 28%. In addition, the combined probabilities of hybridisation at close range obtained here, and those of long range cross attraction estimated in a previous study suggest that actual gene flow between the biotypes is approximately 2.4% per generation in the wild. Thus, sympatric larch and pine associated populations of *Z. diniana* are host races.

Introduction

In many parasitic organisms mate choice is directly correlated with host choice, often because the host is used as a mating site (but see Chapter 1 and Appendix). If sympatric populations of these species differ in host preference, gene flow between them will be restricted, and, if the two populations also experience disruptive host associated selection, they may begin to diverge at loci where selection is strong relative to gene flow (Maynard Smith, 1966; Rice and Hostert, 1993). When genetic divergence does occur, and the populations remain connected by gene flow, they are referred to as host races (Chapter 1, Diehl and Bush, 1984).

According to the theory of sympatric speciation via host shift, this initial divergence between host races may have the pleiotropic effect of increasing their pre- or post-mating isolation. Host-associated selection on timing of development, for example, a trait that differs between several closely related phytophages on hosts that fruit or flower at different times e.g. (Feder et al., 1993; Groman and Pellmyr, 2000; Horner et al., 1999; Komazaki, 1998; Komazaki, 1990), is known to reduce gene flow between host races of the apple maggot fly *Rhagoletis pomonella* (Feder et al., 1994). However, divergence in myriad other traits could also increase reproductive isolation, a prediction also made by allopatric models. A stronger barrier to gene flow then permits divergence in more weakly selected traits, which in turn tightens restrictions on gene flow further, and so on, so that genetic differentiation and isolation of the host races increases by positive feedback until complete reproductive isolation is achieved (Feder et al., 1993; Rice, 1984b; Rice and Hostert, 1993).

Is speciation via host shift an important process in generating new parasite species? Studies of the Tephritid fruit fly *Rhagoletis pomonella* have shown that host races can form and be maintained entirely in sympatry (Bush, 1992; Feder et al., 1998), and the detection of likely host races in a variety of other taxa, including pea aphids *Acrthosiphon pisum* Harris (Via, 1989; Via, 1991; Via, 1999), the goldenrod gall ballmaker *Eurosta solidaginis* (Craig et al., 1997; Craig et al., 1993; Itami et al., 1998) and the cuckoo *Cuculus canoru*s (Gibbs et al., 1996; Marchetti et al., 1998) suggests that host race formation is common. Observations that host plant usage overlaps between several allopatric *Rhagoletis* species,

but no sympatric species (Bush and Smith, 1998), and that host fidelity appears to be the primary mechanism of reproductive isolation between sympatric *Rhagoletis pomonella* and *Rhagoletis mendax* (Feder and Bush, 1989) provides strong circumstantial evidence that, in *Rhagoletis* at least, host races speciate in sympatry. However, despite these clear examples of the beginning and possible end points of the process, direct evidence that it proceeds from one to the other in nature has only recently begun to emerge, from measurements of the degree of pre- and post-mating isolation, and genetic differentiation between a number of host race systems (Chapter 1). The existence of an array of host races, displaying levels of gene flow and differentiation consistent with successive stages of speciation via host shift, would provide strong evidence that the gap from polymorphism to host associated species can be bridged in sympatry (see Jiggins and Mallet 2000). Current data suggests that such an array does exist (Chapter 1), but comparisons are hampered by a lack of data on current levels of gene flow, which has been directly estimated in only a handful of cases (Chapter 1).

Here, the degree of hybridisation between races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae) specialized on larch (*Larix* spp.) and pine (*Pinus cembra*), and the ability of hybrids to compete with non-hybrids for mates, is examined using laboratory mate choice experiments. This study forms part of a series addressing several factors affecting gene flow between the races, which are designed to permit a direct estimate of their level of gene flow in the wild. The results presented in this chapter complement those of an earlier investigation of pheromone-mediated, long-range cross-attraction in the field (Appendix) to allow an estimate of the minimum level of gene flow between the forms, and an assessment of their host race status. A defining feature of host races is appreciable (≥1% per generation) actual gene flow, combined with backcrossing (Chapter 1). The term ‘actual gene flow’ refers to the movement of genes from one population to another, and makes no assumption about their establishment (Mallet, in press).

The larch budmoth *Zeriaphera diniana* Guenée is found in coniferous forests throughout the Palaearctic, where it feeds on larch (*Larix* spp.), pine (*Pinus cembra*)and spruce (*Picea* spp.) (Berryman, 1986). Populations on larch are a serious economic pest in the French and Swiss Alps, where, at peak densities, approximately 10^5 fold higher than at the troughs of regular (eight to 12 year) population cycles, they frequently cause widespread defoliation of their host trees (Baltensweiler, 1993). The population dynamics of *Z. diniana* on pine are less well understood, but density on this host appears to be more

stable. Sympatric larch and pine associated populations inhabiting mixed stand forests of the French and Swiss Alps differ in a number of heritable traits likely to reduce their extent of gene flow, including (i) female pheromone blend (Guerin et al., 1984), (ii) strength of male response to alternate pheromone blends (Lofstedt, unpublished data) and (iii) timing of peak adult eclosion (Baltensweiler, 1977). The two races also exhibit spatially and temporally replicable differentiation in fifth instar larval colour (larch larvae are black, and pine are greenish yellow), and in allele frequency at the *Pgm*, *Mdh-s* and *Idh-s* allozyme loci (Emelianov et al., 1995). Other differences between the races include later emergence and slower development of pine larvae relative to larch, and a non-diagnostic tendency towards a larger size of larch race adults (Day 1984, W. Baltensweiler, unpublished data).

However, despite the reduction in gene flow suggested by the number and nature of their differences, there is evidence that appreciable gene flow still takes place between the two biotypes. In the laboratory, larch and pine race insects readily produce hybrid and backcross progeny in normal male to female ratios (Baltensweiler, 1977, Emelianov et al., 1995, Chapter 3), there is no evidence of an intrinsic reduction of hybrid fitness in the laboratory (Chapter 3), and field tests reveal cross-attraction of males to females of the opposite type (Appendix 1). Despite their developmental differences, adults of both forms co-exist in the field for several weeks (Baltensweiler, W., pers com). Furthermore, the allozyme differentiation observed between the two forms consists of three highly significant loci (*Pgm*, *Mdh*-s, and sex-linked *Idh*-s), and 10 non significant loci. This pattern is compatible with continuing gene flow, if, for example, the majority of alleles freely flow between populations via backcrossing, but host specific selection on *Pgm*, *Mdh*-s, and *Idh*-s or linked loci restricts the establishment of biotype specific alleles in the alternate form (Emelianov et al., 1995).

Actual gene flow between larch and pine race *Z. diniana* involves two steps, (i) long range, pheromone mediated cross attraction of a male to a calling female of the opposite biotype (Appendix 1), and (ii) hybrid mating. In this chapter I address the likelihood of the latter step, which, unlike pheromone mediated cross attraction, largely dependent on the behaviour of males, may involve both male and female choice, as well as biotype-specific differences in mating propensities.

Materials and Methods

Collections

Collection sites are shown in Figure 1, Chapter 1. Collections of fourth and fifth instar Z. *diniana* were made in June 1996-1999 at (i) Tueda, France (45°23'N, 6°35'E): pine race, (ii) Mongénèvres, France (44°56'N, 6°43'E): larch race, (iii) Bois les Ayes, France (44°50'N, 6°39'E): larch and pine races, (iv) Bois les Combes, France (44°54'N, 6°34'E): larch race, (v) Pontresina, Switzerland (46°29'N, 9°54'E): larch and pine races, (vi) Sils, Switzerland (46°42'N, 9°27'E): larch race, and (vii) Val Bever, Switzerland (46°26'N, 9°50'E): larch and pine races. Biotypes of larvae were determined according to the host tree, larch (*Larix* spp.) or pine (*Pinus* spp.) from which they were collected, and their colour. All insects were shipped back to the laboratory within two weeks of collection.

Breeding and rearing

All laboratory-bred insects were individually reared to pupation on newly flushing larch (*Larix* spp.) foliage in Sterilin plastic vials (15 x 50mm) with cotton wool stoppers. Newly emerged larvae were provided with fresh larch buds not more than 50mm long, replenished every third day, and later instars with increasingly longer needles, until fifth (final) instars were fed on several of the largest buds available, replaced weekly. Pine race insects were reared to adulthood in a constant temperature (CT) room with reversed day / night cycle (day - 21:00hrs-13:00hrs, twilight - 13:00hrs - 14:00 hrs) kept at 17-21°C. Larch pupae and fifth instar larvae were removed as necessary to an incubator with reversed day night cycle for chilling at 12-14°C, in order to synchronize their development with that of the younger, more slowly developing pine insects. Humidity was maintained at 70-80% throughout laboratory rearing. Rearing tubes containing pupae were cleared of all food and debris prior to adult emergence. Insects were sexed either as larvae (pine and F1 hybrids only) according to the presence or absence of testes visible through the cuticle of fifth instars, or (all insects) according to the number of unfused abdominal segments in the pupae (six in the female, and seven in the male, Bradley et al., 1979). Fifth instar larch larvae are difficult to sex as they are opaque under laboratory lighting.

Upon eclosion, adults were allowed to mature (i) for 5 days at 5°C (1997) or (ii) for 2-8 days in a CT room (17-21°C) (all subsequent years) before being used for breeding and/or mate choice experiments. Adults collected in 1996 were used immediately. All adult eclosion dates were recorded.

To breed insects in the laboratory, females not mated in experiments were enclosed with a single male in plastic petri dishes (10 cm diameter x 1.5 cm high) until mating took place. All mated females, including those that mated during experiments, were removed to 250ml clear plastic cylinders containing fresh larch and pine foliage. An egg-laying substrate formed from a rectangle of green textured paper, folded back on itself along both long edges to a final size of 10cm x 4cm, and taped to one wall folded side down was also present in each cylinder. Females had daily access to water, and were fed a 5% sucrose solution thrice weekly. Eggs were stored in the dark at room temperature (1997) or at 17-21°C in the CT room (all other years) for approximately six weeks prior to chilling for five to six months at 2°C in a darkened incubator. Diapause was broken by removing eggs to room temperature (1997) or to the CT room (subsequent years).

Insects from the 1998 and 1999 collections were reared to eclosion in the same manner as laboratory bred insects. The rearing method of insects collected in 1996 differed from that in subsequent years in (i) use of the laboratory bench top in place of CT room, (ii) natural day/night cycle, and (iii) use of smaller (10 x 50 mm.) rearing tubes.

Mate choice experiments

'Quartet' mate choice experiments, consisting of four moths, one virgin male and one virgin female of each type being tested, were designed to allow detection of relative male and female mating propensities, as well as levels of assortative mating (Davies et al., 1996). Three types of experiments were carried out, (a) larch vs. pine, (b) F_1 hybrid vs. larch, and (c) backcross (F_1 hybridxlarch) vs. larch (one experiment). All experiments were conducted in the CT room at the same temperature (17-21°C) and light cycle (day - 21:00hrs-13:00hrs, twilight - 13:00hrs - 14:00 hrs) used during rearing. Each quartet of moths was enclosed in a 1L cylindrical cage before 13:00 (twilight) and observed half hourly between 13:00 and 19:30 until a single mating took place. All mating chambers contained fresh sprigs of larch and pine, useful for promoting normal mating behaviour (Emelianov, pers com). During a pilot study of 22 quartet experiments, all matings took between 30 minutes and several (>six) hours to complete, and only one mating (at 19:30)

took place after 18:00. Prior to their release into the cage, one male and one female randomly chosen from each experiment was marked on one wing with a felt tipped pen, to enable the subsequent identification of all moths. The unmated male and female were removed from each cage as soon as a mating took place, and the mated pair removed the following morning. Thereafter, individuals were not used in further mate choice experiments.

Four sets of experiments were conducted, in (i) May 1998 and (ii) May 1999 using laboratory bred insects (larch vs. pine, and hybrid vs. larch experiments), and (iii) July 1998 and (iv) July 1999, using insects collected from the wild (larch vs. pine experiments only). Laboratory reared adults used in May 1998 were second generation larch, pine, and hybrid descendents of the larvae collected in June 1996 from sites (i) –(vi). Those used in May 1999 were first generation descendents of insects collected in June 1998 from sites (iii), (v) and (vi) and, in the case of a single individual, of a mated larch female collected in July 1998 from site (vii) (Figure 1). All fifth instar offspring of this female had the recessive black coloration typical of the larch race. Adults used in July 1998 and 1999 were collected as larvae in June of the same year (section 2a, Fig. 1). Wherever possible, all adults used in a mate choice experiment were between two and seven days old, and had eclosed within four days of one another.

Statistical procedures

Statistical analyses of larch vs. pine mate choice results were carried out using Exhaustive Chi Square Analysis software (Uitenbroek,) and an exact binomial confidence interval calculator (Pezzullo, 2000).

The null hypothesis that an individual's tendency to mate was independent of the biotype of the opposite sex was analysed using a likelihood ratio (G) test. Marginal totals of a 2x2 contingency table were used calculate relative (larch vs. pine race) male and female mating propensities. The cross-products of these propensities gave the expected frequencies of each of the four possible types of cross (LxL, LxP, PxL, and PxP) under the null hypothesis. (Casares et al., 1998; Sokal and Rohlf, 1995). The distribution of G can be approximated by the chi-squared ($?^2$) distribution when samples are large (expected frequencies =3) (Conahan, 1970; Sokal and Rohlf, 1995). Fisher exact tests, which give exact binomial probabilities, are preferable to likelihood ratio tests where samples are small and the chi-squared distribution may not be a good approximation of the G

distribution (Sokal and Rolf, 1995). Exact binomial confidence intervals for each of the four types of cross were also calculated.

To test for differences in the proportions of the four cross types throughout the evening, binomial confidence intervals for each were calculated within each of two time periods, (i) from 13:00 (twilight) up to and including 15:00 (+1 hour after dark), and (ii) from 15:00 up to and including 18:00 (+4 hours after dark). Two crosses were omitted from this analysis as their time of mating was not recorded. Approximately equal numbers of matings (57 and 56 respectively) occurred across each time period. Differences in the proportions of each type of cross between years, and in the mating propensity of marked vs. unmarked moths were tested in the same manner. Finally, Fisher exact tests (where expected frequencies of one or more categories were <3) and likelihood ratio tests were used to test for an association between marking and tendency to hybridise.

Estimating the level of hybridisation between larch and pine biotypes

Long range cross attraction between larch and pine race Z. *diniana* has been studied in the wild (Appendix), and, at a conservative estimate, assuming females always call from their own host, is at least 9.1 (5.7-12.4)% in the case of larch males attracted to pine females, and 3.3 (0.76-5.8)% in the case of pine males attracted to larch females (Appendix 1).

Therefore, the proportion of actual gene flow from the larch race to the pine can be conservatively estimated as

$$\frac{1}{[d(l) \times m(l)] + 1}$$

where d(l) is the ratio (or ‘density bias’) of pine race to larch race males attracted by pine race females, and
m(l) is the ratio (or ‘mating bias’) of pine female x pine male to pine female x larch male matings

Similarly, the proportion of actual gene flow from the pine biotype to the larch can be conservatively estimated as

$$\frac{1}{[d(p) \times m(p)] + 1}$$

where d(p) is the ratio (density bias) of larch race males to pine race males attracted by larch females, and
m(p) is the ratio (mating bias) of larch female x larch male to larch female x pine male matings

These estimates are conservative because if females settle on and call from the non-natal host, they are likely to attract higher numbers of males of the opposite host race (Appendix). However, the extent of female host fidelity has not yet been estimated.

Results

Larch vs. pine experiments.

The results for larch vs. pine experiments (one larch female, one larch male, one pine female, and one pine male per quartet) are summarized in Table 1. Mating between larch and pine race moths was not random (G=13.562, *?* = 1, p=0.0002), but 28.13 (19.42-38.22)% of all matings were hybrid, 3.13 (0.07-8.86)% between larch females and pine males, and 25.00 (16.72-34.88)% between pine females and larch males (Table 1). The proportions of all four types of cross remained constant across two time periods, (i) from 13:00 (twilight) up to and including 15:00 (+1 hour after dark, and (ii) from 15:00 up to and including 18:00 (+4 hours after dark) (Figure 3). There was no significant effect of identification marks on the hybridisation rate of pine females (p = 0.6866), larch females (p = 0.5769), pine males (p = 0.2456), or larch males (p=0.9135) (Tables 2-5).

Eighty eight of the larch vs. pine experiments were conducted in 1999, and the remaining eight in 1998. Excluding the 1998 data set had no significant effect on the observed

proportion of any cross type (Figure 4), and therefore results from both years were combined in the final analysis. In 15 experiments, eclosion dates of some or all members of a quartet were unknown and / or separated by more than four days, although in seven of these the eclosion dates of both members of the same sex were paired. Excluding this 'unmatched' data set had no significant effect on the observed proportion of any cross type (Figure 5), and these results were also retained in the final analysis.

As shown in Figure 2, pine females had significantly higher mating propensities than larch females (84.4% vs. 15.6%), and pine males significantly higher mating propensities than larch males (61.5% vs. 37.5%). There was no significant effect of marking on the mating propensities of pine females, larch females, pine males, or larch males (Figure 6).

Larch vs. *F_1 hybrid experiments*

The number of hybrids available for mate choice experiments was limited, due to their use in a variety of other studies, including mapping of colour pattern loci and analysis of hybrid pheromone production. Three of the five larch vs. hybrid mate choice tests that were conducted resulted in F_1 female x larch male matings, but no F_1 male x larch female matings were observed (Table 6). A single larch vs. F2 backcross (F1xlarch) experiment was also conducted, and resulted in a backcross female x larch male mating. Due to small sample sizes, no statistical analyses were carried out on experiments involving hybrid forms.

Direct estimate of gene flow

In the case of pine females on pine trees, the overall bias in the direction of matings (pine female x pine male vs. pine female x larch male) is

[ratio of pine race to larch race males attracted by pine race females = 0.9091 / 0.0909 (from Appendix)] x [ratio of pxp to pxl matings = 0.7037 / 0.2962 (from Table 1)]

= 23.75 : 1

and therefore the percentage of actual gene flow from the pine to the larch race estimated as

$$[1/(23.73+1)] \times 100\%$$
$$= 4.0\%$$

similarly, in the case of larch females on larch trees, the overall bias in the direction of matings (larch female x larch male vs. larch female x pine male) is

[ratio of larch race males to pine race males attracted by larch race females = 0.967 / 0.033 (from Appendix)] x [ratio of lxl to lxp matings = 0.8 / 0.2 (from Table 1)]

$$= 117.21:1$$

and therefore the percentage of actual gene flow from the larch to the pine race estimated as

$$[1/(117.21+1)] \times 100\%$$
$$= 0.8\%$$

giving an average estimate of 2.4% actual gene flow between the races.

Discussion

Mating propensity vs. *mate choice.*

The quartet mate choice design employed here allowed three parameters, relative female mating propensity, relative male mating propensity, and mating discrimination according to biotype to be estimated (Casares et al., 1998; Davies et al., 1997). The mating propensity of pine moths of both sexes was significantly higher than that of larch moths, and this effect was particularly pronounced in the case of females (Figure 2). However, the significant G-test of independence between tendency to mate and biotype of the opposite sex ($p=0.0002$) reveals that differences in the observed proportions of the four crosses were not solely due to differences in the mating propensities of the four moth

types. Although additional behavioural data might allow individual contributions of male vs. female choice, and mate choice vs. mating propensity to the overall rate of hybridisation to be assessed (Davies et al., 1997), it is not possible to separate the effects of these phenomena using data obtained in this series of experiments. The causes of the differences in mating propensities are also unknown, and could include differences in overall 'choosiness', overall vigour, or both.

In addition, it should be noted that, because no insects had access to food or water prior to mating, there is a possibility that mate choice and/or propensity was affected by lack of nutrients. Hingle (2000), for example, found that nutritional stress causes female stalk eyed flies *Cyrtodiopsis dalmanni* to lose their preference for wide-eyespan males, although Gray (1999) found no effect of female nutritional condition on mate choice in house crickets *Acheta domesticus*. However, the insects used here appeared vigorous at the time of experiments: males began courtship almost immediately the experiments were set up, and females flew to avoid them.

Finally, the necessity of rearing of all laboratory bred larvae on larch (pine cuttings secrete copious amounts of resin) could also have affected mating. behaviour. Host plant chemistry is known to affect cuticular hydrocarbons (Stennett and Etges, 1997), and these compounds can play an important role in insect mate choice (Coyne et al., 1994; Ferveur, 1997; Singer, 1998; Tregenza and Wedell, 1997). The importance of cuticular hydrocarbons in *Z. diniana* mating behaviour is unknown. However, the majority of the pine insects used in these experiments were collected from the wild as late fifth instar larvae, and consumed little or no larch while in captivity (Drès, unpub. data).

Lack of temporal isolation

The absence of any significant difference in the relative proportions of the four crosses throughout the evening suggests that sympatric larch and pine *Z. diniana* do not experience any temporal isolation due to differences in mating activity throughout the evening. This result is consistent with previous observations from the field that the timing of pheromone mediated cross attraction does not differ between the races (Appendix 1).

Estimate of actual gene flow

The estimate of actual gene flow obtained in this study is based on the assumption that females always call from their own host. Because the percentage cross attraction of pine males to larch females is much higher (37.7% vs. 3.3%) when larch females call from pine trees (Emelianov et al. 2000), the real level of actual gene flow may be higher. The extent of host fidelity of the biotypes is currently being assessed via a combination of field observations and laboratory experiments (Emelianov et al, in prep.).

It should also be noted that the estimates presented here are only 'snapshots' of gene flow levels; Itami et al. (1998) stress the importance of variation in factors affecting hybridisation across space and time. Relative abundances of larch and pine moths, for example, which are likely to differ dramatically according to larch biotype population density, might affect hybridisation rates, as may the relative abundance and position of larch and pine trees in mixed stand forest. Emelianov et al. (Appendix) found that individual larch or pine trees (and the females on them) were more likely to be close to 'alien' males when surrounded by host trees of the opposite type.

Competitive mating ability of hybrids

Although the number of hybrid vs. nonhybrid mating tests carried out here was small, the large fraction (3/5) of hybrid female x larch male matings observed here suggests that backcrossing at close quarters, at least in this direction, is unlikely to be rare. Furthermore, laboratory studies have not revealed any evidence of intrinsic fitness differences between hybrid and non hybrid broods (Chapter 3). However, the ability of hybrid females to attract non hybrid males, and of hybrid males to locate non hybrid females in the field has yet to be tested.

Host race status

Other studies have demonstrated that sympatric larch- and pine-associated *Z. diniana* exhibit host fidelity (Bovey and Maksymov, 1959; Emelianov et al., in prep), are genetically differentiated (reviewed in Emelianov et al., 1995), and are significantly more likely to attract mates of same biotype than of the alternate biotype (Appendix 1). Therefore, the two forms meet six of the seven host race criteria proposed in Chapter 1. Here, it has been shown that (a) at a conservative estimate, the level of actual gene flow

between larch and pine race *Z. diniana* is 4.0% in the larch to pine direction, and 0.8% in the opposite direction, and (b) that hybrid females are able to successfully compete against nonhybrids to obtain matings with larch males. Thus, there is now also considerable evidence that they fulfil the remaining criterion, i.e. they undergo actual gene flow with an appreciable frequency (š1% per generation), and backcross (Chapter 1).

Conclusions

When males and females of both *Z. diniana* races are in close proximity, mating is moderately assortative with respect to host race; approximately 28% of matings are hybridisations. Differences in relative male and female mating propensities do exist, but either these are not the sole cause of assortative mating, or interactions between males and females are also responsible for some of the differences in mating rates.

Despite this assortativeness, the overall level of actual gene flow (gene movement via hybridisation) between the races is approximately 2.4% per generation. The direction of movement seems to be mainly from the larch race to the pine, because both hybridisation, and long range cross attraction (to females calling from their native host, Appendix 1), take place more readily between larch males and pine females. However, an estimate of female host fidelity will be necessary to verify the directional bias and upper limit of actual gene flow between these biotypes, since cross attraction of pine males to larch females is significantly higher when the females call from larch trees.

We have also shown that F_1 hybrid females can successfully compete with larch females to gain matings with larch males, suggesting that hybridisation between the forms leads to gene flow in the more usual sense, i.e. incorporation of genes from one population into the other. The level of actual gene flow observed, and the likelihood of backcrossing strongly suggests that larch and pine races of *Z. diniana* are host races, and therefore represent an intermediate step between sympatric host associated polymorphism and sympatric species.

References

Baltensweiler, W. 1977. Colour polymorphism and dynamics of larch bud moth populations (*Zeiraphera diniana* Gn. Lep. Tortricidae). *Mitteilungen - Schweizerische Entomologische Gesellschaft*, 50:15-23.

Baltensweiler, W. 1993. A contribution to the explaination of the larch bud moth cycle, the polymorphic fitness hypothesis. *Oecologia*. 93:251-255

Bovey, P., and J. K. Maksymov. 1959. Le problème des races biologiques chèz la tordeuse grise du mélèze *Zeiraphera griseana* (Hb.). *Vierteljahresschrift der Naturforsch. Ges. Zurich*, 104:264-274.

Bradley, J. D., W. G. Tremewan, and A. Smith. 1979. *British Tortricid Moths*. The Ray Society, London.

Bush, G., and J. Smith. 1998. The genetics and ecology of sympatric speciation: A case study. *Researches on Population Ecology*, 40:175-187.

Bush, G. L. 1992. Host race formation and sympatric speciation in *Rhagoletis* fruit flies (Diptera:Tephritidae). *Psyche*, 99:335-357.

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

Conahan, M. A. 1970. The comparative accuracy of the likelihood ratio and chi-squared as approximations to the exact multinomial test. Ed.D. thesis, Leigh University.

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

Craig, T. P., J. D. Horner, and J. K. Itami. 1997. Hybridization studies on the host races of *Eurosta solidaginis*: Implications for sympatric speciation. *Evolution*, 51:1552-1560.

Craig, T. P., J. K. Itami, W. G. Abrahamson, and J. D. Horner. 1993. Behavioral evidence for host-race formation in *Eurosta solidaginis*. *Evolution*, 47:1696-1710.

Davies, N., A. Aiello, J. Mallet, A. Pomiankowski, and R. E. Silberglied. 1997. Speciation in two neotropical butterflies: Extending Haldane's rule. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 264:845-851.

Diehl, S. R., and G. L. Bush. 1984. An evolutionary and applied perspective of insect biotypes. *Annual Review of Entomology*, 29:471-504.

Emelianov, I., W. Baltensweiler, M. Drès, and J. Mallet. in prep. Host fidelity of larch and pine race *Zeiraphera diniana* (Lepidoptera: Tortricidae). .

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the Larch Budmoth) - polymorphism, host races or sibling species. *Heredity*, 75:416-424.

Feder, J. L., S. H. Berlocher, and S. B. Opp. 1998. Sympatric host race formation and speciation in *Rhagoletis* (Diptera:Tephritidae): A tale of two species for Charles D. In S. B. Mopper and S. Y. Strauss (eds.), *Genetic Structure and Local Adaptation in Natural Insect Populations*, pp. 408-441. Chapman and Hall, New York.

Feder, J. L., and G. L. Bush. 1989. A field test of differential host-plant usage between two sibling species of *Rhagoletis pomonella* fruit flies (Diptera: Tephritidae) and its consequences for sympatric models of speciation. *Evolution*, 43:1813-1819.

Feder, J. L., T. A. Hunt, and L. Bush. 1993. The effects of climate, host-plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella*. *Entomologia Experimentalis et Applicata*, 69:117-135.

Feder, J. L., S. B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America*, 91:7990-7994.

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

Gibbs, H. L., M. L. Brooke, and N. B. Davies. 1996. Analysis of genetic differentiation of host races of the common cuckoo *Cuculus canorus* using mitochondrial and microsatellite DNA variation. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 263:89-96.

Gray, D. A. 1999. Intrinsic factors affecting female choice in house crickets: Time cost, female age, nutritional condition, body size, and size-relative reproductive investment. *Journal of Insect Behaviour*, 12:691-700.

Groman, J. D., and O. Pellmyr. 2000. Rapid evolution and specialization following host colonization in a yucca moth. *Journal of Evolutionary Biology*, 13:223-236.

Guerin, P., W. Baltensweiler, H. Arn, and H. Buser. 1984. Host race pheromone polymorphism in the larch budmoth. *Experientia*, 40:892-894.

Hingle, A. P. 2000. Condition dependent female mate-preference in the stalk eyed fly *Cyrtodiopsis dalmanni*, Ph.D. Thesis. University College London.

Horner, J. D., T. P. Craig, and J. K. Itami. 1999. The influence of oviposition phenology on survival in host races of *Eurosta solidaginis*. *Entomologia Experimentalis et Applicata*, 93:121-129.

Itami, J., T. Craig, and J. Horner. 1998. Factors affecting gene flow between the host races of *Eurosta solidaginis*. In S. Mopper and S. Strauss (eds.), *Genetic structure and local adaptation in natural insect populations*, pp. 375-404. Chapman and Hall, New York.

Komazaki, S. 1990. Variation in the hatch timing of the overwintering egg among populations of *Aphis spiraecola* Patch (Homoptera: Aphididae) collected from different host plants and localities in Japan. *Applied Entomology and Zoology*, 25:27-34.

Komazaki, S. 1998. Difference of egg diapause in two host races of the spiraea aphid, *Aphis spiraecala*. *Entomologia Experimentalis et Applicata*, 89:201-205.

Mallet, J. in press. Gene Flow. In I. P. Woiwood, D. R. Reynolds and C. D. Thomas (eds.), *Insect Movement: Mechanisms and Consequences*. CAB International, Wallingford.

Marchetti, K., H. Nakamura, and H. Gibbs. 1998. Host race formation in the common cuckoo. *Science*, 282:471-472.

Maynard Smith, J. 1966. Sympatric speciation. *American Naturalist*:637-650.

Pezzullo, J. P. 2000. Interactive statistics page. :http://members.aol.com/johnp71/confint.html (8.2000).

Rice, W. R. 1984b. Disruptive selection on habitat preference and the evolution of reproductive isolation - a simulation study. *Evolution*, 38:1251-1260.

Rice, W. R., and E. E. Hostert. 1993. Laboratory experiments on speciation - what have we learned in 40 years. *Evolution*, 47:1637-1653.

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

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*. W H Freeman and Company, New York.

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

Tregenza, T., and N. Wedell. 1997. Definitive evidence for cuticular pheromones in a cricket. *Animal Behaviour*, 54:979-984.

Uitenbroek, D. Simple Interactive Statistical Analysis. :http://home.clara.net/sisa/twoby2.html (6.2000).

Via, S. 1989. Field estimation of variation in host plant use between local populations of *pea aphids* from 2 crops. *Ecological Entomology*, 14:357-364.

Via, S. 1991. Specialized host plant performance of pea aphid clones is not altered by experience. *Ecology*, 72:1420-1427.

Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, 53:1446-1457.

Table 1. Direction of crosses, larch vs. pine biotype experiments
n=96
numbers in brackets are 95% confidence intervals
G=13.562, *v* = 1, p=0.0002

	Larch male	**Pine male**
Larch female	12 *12.5 (6.63-20.82)%*	3 *3.13 (.065-8.86)%*
Pine female	24 *25.00 (16.72-34.88)%*	57 *59.38 (48.87-69.29)%*

Table 2. Direction of mating, marked vs. unmarked mated pine females
n=79
G=0.163, *v*=1, p=0.6866

	Mate	
Pine female	**Larch male**	**Pine male**
Marked	12	32
Unmarked	11	24

Table 3. Direction of mating, marked vs. unmarked mated larch females *n*=13 p_{exact} = 0.5769		
	Mate	
Larch female	**Larch**	**Pine**
Marked	8	2
Unmarked	3	0

Table 4. Direction of mating, marked vs. unmarked mated pine males *n*=58 p_{exact} = 0.605		
	Mate	
Pine male	**Larch**	**Pine**
Marked	0	29
Unmarked	2	27

Table 5. Direction of mating, marked vs. unmarked mated larch males *n*=34 G=0.012, *v*=1, p=0.9135		
	Mate	
Larch male	**Larch**	**Pine**
Marked	6	13
Unmarked	5	10

Table 6. Direction of crosses, larch biotype vs. F_1 hybrid experiments *n=5*		
	Larch male	**F_1 male**
Larch female	1	0
F_1 female	3	1

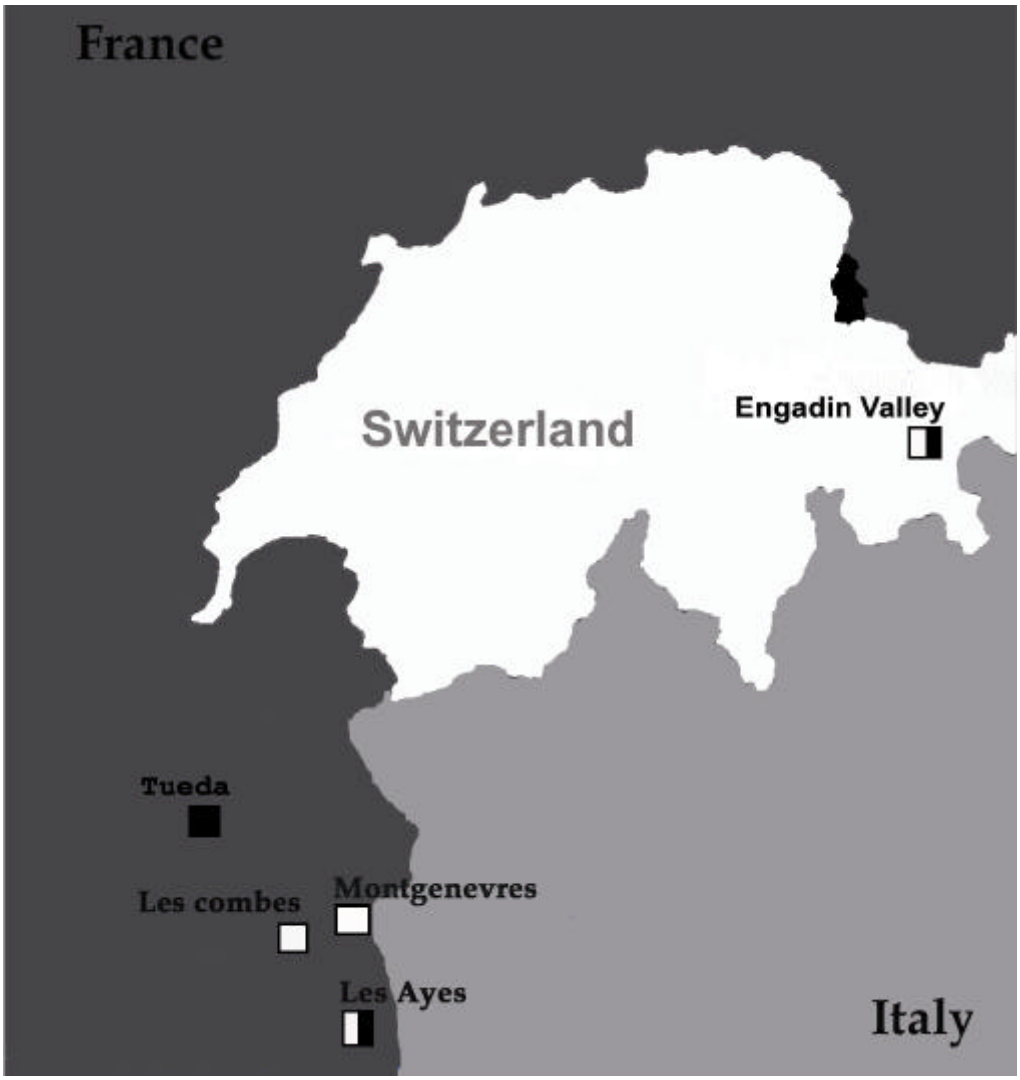

Figure 1. Map of collection sites. Adapted from Emelianov et al., 1995.

Larch and pine race larvae present

Larch race larvae only

Pine race larvae only

Engadin Valley contains Pontresina (larch and pine race larvae), Sils (larch race larvae only), and Val Bever (larch and pine race larvae) sites.

Figure 2. Relative proportions of female matings obtained by pine vs. larch females, and male matings obtained by pine vs larch males. Error bars are exact binomial 95% confidence intervals

Figure 3. Effect of time of night on cross type. Error bars are exact binomial 95% confidence intervals

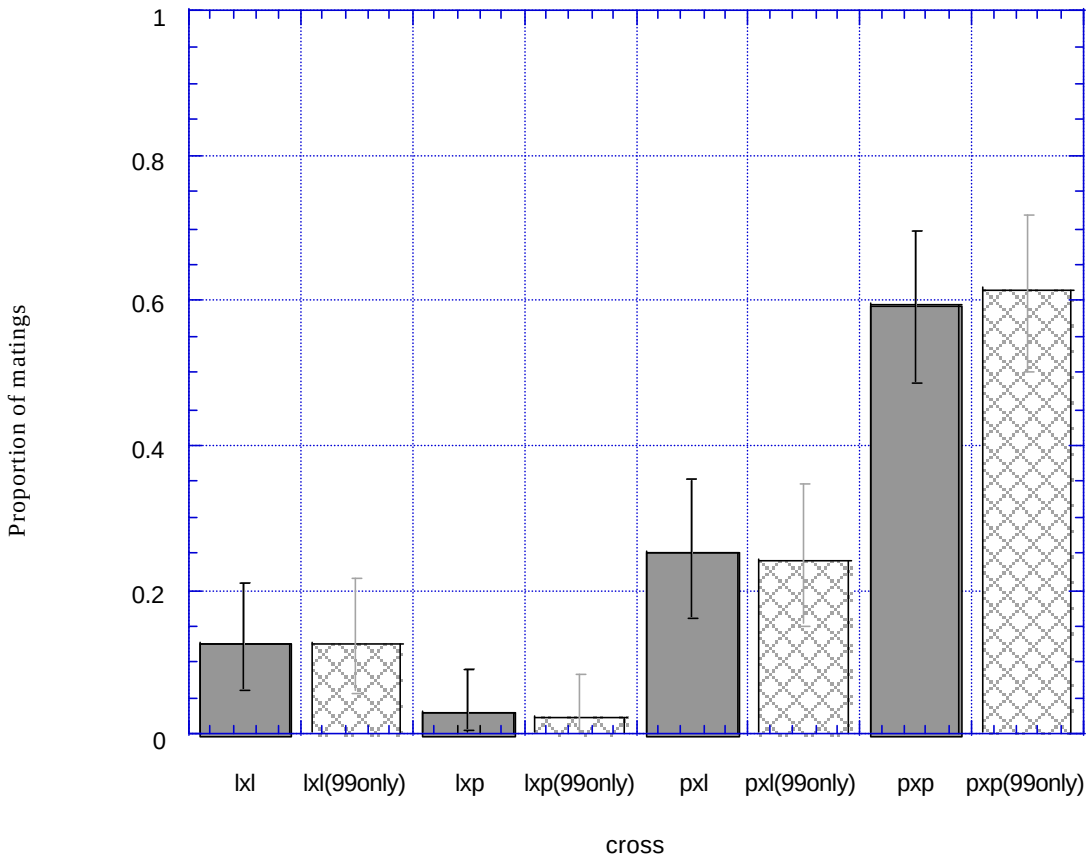

Figure 4. Effect of exluding 1998 data on cross type frequencies. Error bars are exact binomial 95% confidence intervals

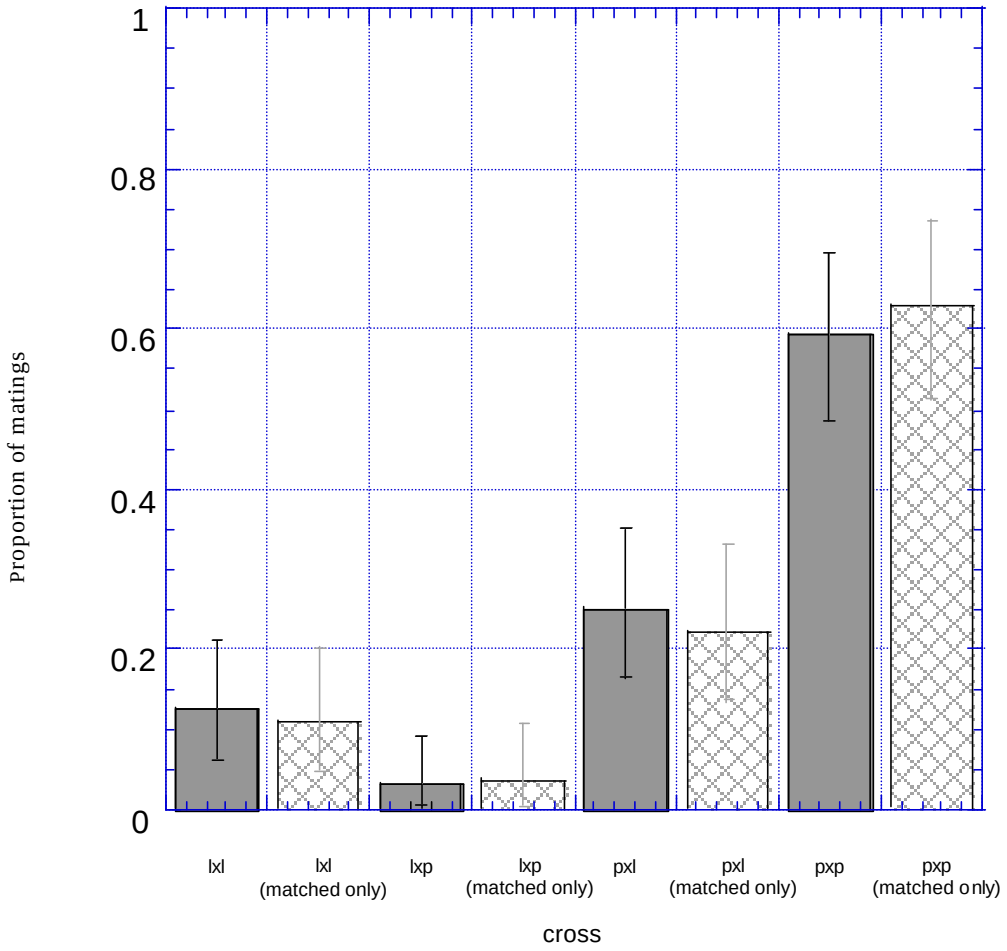

Figure 5. Effect of exluding ‘unmatched’ experiments on cross type frequencies. Error bars are exact binomial 95% confidence intervals

Figure 6. Relative proportions of marked and unmarked mated pine females, larch females, pine males and larch males. Error bars are 95% binomial confidence intervals.

Chapter 3

Relative fitness of larch budmoth *Zeiraphera diniana* host races and their hybrids

Abstract

Larch and pine host races of *Zeiraphera diniana* maintain genetic differences in several traits while sustaining actual gene flow of approximately 2.4% per generation. Given the extent of gene flow, it can be assumed that these differences are maintained via disruptive selection, and that selection is operating against hybrid forms in the wild. However, we found no evidence that larch, pine, hybrid or backcross (hybrid x larch) broods differ in their percentage of hatching eggs, or larvae surviving to final (5^{th}) instar in the laboratory. As with previous studies, we also failed to find any evidence of deviation from 50:50 sex ratios in any brood type. Thus, ecological factors are predicted to play an important role in maintaining differences between these biotypes.

Introduction

The larch budmoth *Zeiraphera diniana* Guenée (Lepidoptera: Tortricidae) is a Palaearctically distributed parasite of larch (*Larix* spp.), pine (*Pinus* spp.) and spruce (*Picea* spp) (Berryman, 1986). The insect is a serious economic pest of coniferous forests throughout the French and Swiss Alps, where larch feeding populations at the peak of eight to ten year population density cycles are responsible for widespread defoliation of their host trees (Baltensweiler, 1993). The population dynamics of Alpine pine feeding *Z. diniana*, however, are less well understood, but appear to be considerably more stable (Baltensweiler pers. com).

Alpine larch and pine associated *Z. diniana* populations are also likely to be host races (Chapter 1). The biotypes exhibit host fidelity in terms of roosting, resting, and oviposition (Bovey and Maksymov, 1959; Emelianov et al., in prep), coexist in sympatry in at least two mixed stand forests (Emelianov et al. 1995), and mate assortatively along host plant lines (Chapter 2, Appendix 1). They are also significantly differentiated at three allozyme loci (Emelianov et al., 1995), and display heritable differentiation in larval colour (fifth instar larch larvae are black, while pine larvae are greenish-yellow) (Baltensweiler, 1993) female pheromone blend (Guerin et al., 1984), and male pheromone response (Baltensweiler and Priesner, 1988, Loftstedt, unpublished data). These differences are maintained in the face of actual gene flow, (exchange of genes via hybridisation, *sensu* Mallet 2000) of at least 2.4% per generation (Chapter 2, Appendix 1). Previous studies have shown that hybrids are fertile (Baltensweiler 1977, Emelianov et al., 1995, Chapter 2), and that, in the laboratory at close range, hybrids and at least some backcross types can successfully compete with non hybrid forms to obtain matings (Chapter 2). However, backcrossing in the wild could be inhibited by a number of factors: an inability of hybrids to attract, locate, or compete for mates (sexual selection), ecologically based natural selection against hybrids, or an intrinsic reduction in hybrid fitness (post-mating isolation). Here, we investigate the latter through laboratory measures of the egg viability, and survival to fifth instar of pine, larch, F_1 hybrid, and F_2 (hybridx larch) broods.

In addition, we test the broods for deviations from a 50:50 sex ratio. Haldane's rule predicts that intrinsic, developmental difficulties experienced by hybrid animals are

more likely to be expressed in the heterogametic sex, often as a deficit of this sex at birth (Haldane, 1922). As is the case with other Lepidoptera, the heterogametic sex in *Z. diniana* is the female. Sex ratio distortion is not expected, however, as previous studies have failed to detect any evidence of sex ratio bias within non-hybrid, F_1 hybrid or backcross broods (Baltensweiler, 1977; Emelianov et al., 1995).

Materials and methods

Collections and rearing

The laboratory bred *Z. diniana* broods used in experiments presented here consist of two cohorts, (i) the 1998 cohort, and (ii) the 1998 cohort, both of which were reared according to the same protocol. Members of the 1998 cohort were the second laboratory generation bred from fourth and fifth instar larvae collected in June 1996 from (1) Bois les Ayes, France (44°50' N, 6°39'E), (2) Bois les Combes, France (44°54' N, 6°34'E), (3) Montgénèvres, France (44°56', 6°43' E), (4) Tueda, France (45°23'N, 6°35'E), (5) Pontresina, Switzerland, (6) Sils, Switzerland (46°42'N, 9°27' E), and (7) Val Bever, Switzerland (46°26'N, 9°50'E). Members of the 1999 cohort were bred from (a) a laboratory population consisting of the 1998 cohort supplemented with the offspring of mated larch females (see below) collected in August 1997, from sites (5)-(7), or, (b) from individuals collected as fourth and fifth instar larvae from sites (5)-(7) in June 1998. A map of collection sites is presented in Figure 1, Chapter 2.

In order to ensure that multiple sibling groups were sampled, the 1996 and 1998 larval collections were made from as many trees as possible at each site. Larvae were classified as larch or pine race according to the host tree, *Larix* spp. or *Pinus cembra*, from which they were collected. The mated females collected in 1997 were either beaten from the branches of host trees at temperatures below the threshold necessary for flight, or collected on sheets placed near ultraviolet lamps (Emelianov et al, in prep.). Because fifth instar larch and pine race larvae differ in colour, and the larch race phenotype is recessive to that of the pine race (Igor Emelianov, unpub. data), it was possible to unambiguously identify pure larch offspring of mated larch females, and only these broods were used in the laboratory breeding programme. Wild collected larvae and

females were transported back to the laboratory within two weeks of collection, where they were reared in the same manner as laboratory bred insects.

In 1996 and 1997 rearing took place on the laboratory bench top, and in 1998 and 1999 in a constant temperature (CT) room kept at 17-21 °C, and a cooled incubator kept at 12-14 °C. Humidity was maintained at 70-80% throughout laboratory rearing.

To produce laboratory bred progeny, females not mated in the 1998 mate choice experiments (Chapter 2) were individually enclosed in clear plastic petri dishes (10cm diameter x 1.5cm high) with single males until mating took place. All mated females, including those which mated during experiments, were removed to 250ml transparent plastic cages containing fresh larch and pine foliage. A rectangle of green textured paper, folded back on itself along both long edges to a final size of 10cm x 4cm, was taped to one wall of each cage, and served as an egg laying substrate. Females had daily access to water, and were fed a 5% sucrose solution thrice weekly. Eggs were kept in the dark at room temperature (1997) or at 17-21°C in the CT room (all subsequent years) for approximately six weeks prior to chilling at 2°C for five to six months, necessary to induce an obligatory diapause. Eggs were kept in the dark throughout diapause, which was broken by their removal to the laboratory bench top (1997) or to the CT room (all subsequent years).

Larvae were individually reared to pupation on newly flushing larch (*Larix* spp.) foliage in 15 x 50mm Sterilin plastic vials with cotton wool stoppers. Newly emerged larvae were provided with fresh larch buds not more than 50mm long, replenished every third day, and final (fifth) instars with several of the largest available larch buds, replenished weekly. Pine form insects were reared to adulthood in a constant temperature (CT) room with reversed day / night cycle (day - 21:00hrs-13:00hrs, twilight - 13:00hrs - 14:00 hrs) maintained at 17-21°C. Early larch and hybrid instars (1^{st} to 4^{th} instar) were also reared exclusively in the CT room, but, because pine race larvae hatch later and develop more slowly than larch race larvae (Day, 1984), 5^{th} instars and pupae were removed as necessary to an incubator with reversed day night cycle for chilling at 12-14°C, in order to synchronize their development with that of the pine insects. Rearing tubes containing pupae were cleared of all food and debris prior to adult emergence. Upon eclosion, adults were allowed to mature (i) for five days at 5C (1997) or (ii) for three-eight days in a CT room (17-23C) (all subsequent years) before being used for breeding.

Measures of fitness

Within the 1998 cohort, the proportions of (i) eggs that hatched, and, where possible (ii) larvae surviving to fourth instar, and (iii) the sex ratio at fourth instar, were recorded for 13 larch x larch (L), 1 pine x pine (P), 7 larch x pine (LP), 3 pine x larch (PL), 10 LxLP, 2 LxPL, 1 PLxL , and 2 LPxL broods. The host race of the female parent of each brood is the first listed. Insects were sexed either as pupae, according to their number of unfused abdominal segments (seven in the male and six in the female, Bradley, 1979), or, in the case of the more translucent pine race and hybrid larvae, according to the presence or absence of testes visible through the cuticle of the fifth instar. No measures of fitness were recorded after fifth instar, due to differences in the rearing temperatures of pine vs. larch (and some hybrid) broods after this developmental stage. In the 1999 cohort, the proportion of hatching eggs only was recorded for 12 L, 3 LxPL, 2 LP, 5 P, 1 PL, and 1 PLxL brood.

Statistical procedures

Statistical analyses were carried out using JMP version 3.1.6 for the Macintosh (Institute, 1996) and an exact binomial confidence interval calculator (Pezzullo, 2000). Due to the relatively small number of broods examined, pine female x larch male (PL) and larch male x pine female (LP) broods were pooled to create a single 'hybrid' category, and all four combinations of hybrid x larch crosses pooled to create a single 'backcross' category.

Differences in egg viability between cross types and cohorts were analysed using a logistic regression with proportion of hatched eggs as the dependent (Y) variable. Maximum likelihoods for models, (a) with cross type as the main factor, (b) with cohort as the main factor, and (c) including an interaction between cross type and cohort, were calculated. The fits of these models to the data set were compared using effect likelihood ratio tests, which computed twice the difference in log likelihoods between a model including all factors (effects), and models from which various factors were absent (Edwards, 1972; Sokal and Rohlf, 1995). Differences in chances of survival to 5th instar between cross types of the 1998 cohort were analysed in a similar manner, with proportion of larvae surviving to fifth instar as the dependent variable. Binomial confidence intervals were calculated for all 1998 cohort sex ratios.

Results

The results of statistical analyses are presented in Tables 1-4 and Figures 1-3. Because several broods (2L, 10L, 17LP, 23PL, 24PL, 29LXLP, 35LXPL, 41PLXL, 54P, 55P, 57P, 60LP, Figure 1) consisted of less than 10 eggs, and had correspondingly large binomial confidence intervals around their observed proportions of egg hatch and survivorship, both of these fitness measures were calculated with and without broods of less than 10 eggs. Excluding small broods from the analysis did not effect the significance of any result (Tables 1-4, and below).

The logistic regression analysis failed to reveal a significant effect of either cross type (all broods: $p=0.34$, broods ≥10 eggs only: $p=0.94$) or cohort (all broods: $p=0.50$, broods ≥10 eggs only: $p=0.22$) on the proportion of hatching eggs, indicating that the eggs of larch, pine, hybrid and backcross broods were equally viable, and that total proportion of viable eggs did not differ between cohorts. Within the 1998 cohort, there was no significant effect of cross type on the proportion of larvae surviving to 5th instar (total broods: $p=0.18$, broods ≥10 eggs only: $p=0.14$). Finally, although sample sizes were small, there was no evidence of sex ratio distortion at fourth instar within any brood.

Discussion

This experiment allowed three measures of intrinsic fitness, percentage egg hatch, survival to late (fifth) instar, and sex ratio, to be compared between larch (L), pine (P) (percentage egg hatch only), hybrid (LP or PL) and hybridxlarch (LxLP, LxPL, LPxL, or PLxL) *Z. diniana* broods reared in the laboratory. No significant difference in either percentage egg hatch or survival to fourth instar was observed between any cross types. Measurements of sex ratios were consist with those of Baltensweiler (1977) and Emelianov et al. (1995) who also failed to detect any sex ratio bias in hybrid and backcross broods reared in the laboratory.

Unfortunately, because of cross type specific differences in rearing protocols necessary for the subsequent use of brood members in mate choice (Chapter 1) and pheromone production (Emelianov et al. in prep) experiments, it was not possible to evaluate survivorship beyond fifth instar. However, several hybrid and backcross insects, including members of a single hybrid x pine brood, were successfully reared to adulthood (Chapter 1 and unpublished data). A more serious shortcoming of this analysis was the extremely small number of pine forms, and the lack of hybrid x pine broods, caused by a failure of the majority of mated pine females to lay eggs in the laboratory. Essentially, this means that one of the controls was lacking, and it remains possible that there are differences in larch and pine race survival, or that hybrid x pine broods suffer developmental difficulties not present in the opposite backcross. However, this does not alter the result that gene flow from the pine to the larch race, at least, is unlikely to be hindered by developmental obstacles.

Finally, because different types of crosses may have different fitnesses (Arnold and Hodges 1995) the necessity of pooling different combinations of backcross and hybrid crosses in the comparison could have masked developmental differences of rare forms. It is possible that, for example, the rare hybrid cross PL had a reduced fitness relative to parental types.

Despite these queries, the balance of evidence so far, from this and other laboratory studies, suggests that there is little or no difference in the intrinsic fitness of hybrid, backcross, and parental *Z. diniana*. This situation contrasts with that observed in another pair of likely host races, biotypes of the goldenrod gall ballmaker *Eurosta solidaginis*, whose F_1 hybrids have an average survivorship much lower than that of parental forms, but is similar to that observed in host races of the apple maggot fly *Rhagoletis pomonella,* for which there is no evidence of any intrinsic reduction in hybrid fitness or fertility (Reissig and Smith, 1978).

The results of this experiment do not suggest that there is no selection against hybrid forms of *Z. diniana*, but rather that selection is unlikely to be in the form of intrinsic fitness differences. In fact, because the genetic differentiation that defines the host races is maintained in the face of gene flow, there must be disruptive, host-associated selection on these traits (Chapter 1) and therefore selection against hybrids or backcrosses. The potential causes of ecologically based hybrid inviability are myriad. One possibility is

that some hybrid larvae may be badly camouflaged, and therefore subject to greater predation than non-hybrids; greenish-yellow pine race larvae bore into young pine shoots, or enclose themselves in webbing next to pine shoots of the same colour, but black 5th instar larch larvae normally enclose themselves in webbing against the dark larch tree bark or in larch cones. It could also be instructive to measure the relative fitness of hybrids and parentals under more competitive conditions; Goday-Herrera et al. (1994), for example, found that, due to their lower average feeding rate, the viability of interspecific hybrids of *Drosophila puvani* and *D. gaucha* was significantly reduced in the presence of the parental species. The ability of hybrid *Z. diniana* to successfully compete with parental larch forms is likely to be particularly important during times of peak larval population density, when hosts rapidly become saturated with larvae. Furthermore, given the differences in emergence times of larch and pine race larvae, which coincide with differences in the timing of bud burst of the two host trees (Day, 1984), it is possible that hybrid larvae may emerge at the wrong time to exploit one or both of the parental hosts. This situation is observed in crosses between host associated biotypes of the spiraea aphid *Aphis citricola* on cherry and thunberg spiraea. When reared outside, hybrids survived well on cherry, but had a fitness of almost zero on the alternate host because they emerged before bud burst (Komazaki, 1986). However, problems in timing might not be related to emergence alone, given that *Zeiraphera larvae* develop at different rates throughout their life cycle.

Conclusions

The percentage of hatching eggs, and larvae surviving to final (5th) instar did not differ between larch race, pine race, hybrid, or backcross (F1 x larch race) *Z. diniana* broods reared in the laboratory, nor did any brood deviate from a 50:50 sex ratio. Thus, there does not appear to be any intrinsic, developmentally based reduction in hybrid or backcross fitness, although differences apparent only at later stages of the life cycle, or under competitive conditions have not been tested for. Nevertheless, selection must be acting against hybrid forms, because the host races remain separated by a number of heritable traits while sustaining actual gene flow of approximately 2.4% per generation. This suggests that ecological factors have played a central role in the divergence of *Z. diniana* host races.

References

Baltensweiler, W. 1977. Colour polymorphism and dynamics of larch bud moth populations (*Zeiraphera diniana* Gn. Lep. Tortricidae). *Mitteilungen - Schweizerische Entomologische Gesellschaft*, 50:15-23.

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

Baltensweiler, W., and E. Priesner. 1988. A study of pheromone polymorphism in *Zeiraphera diniana* Gn (Lep, Tortricidae).3. Specificity of attraction to synthetic pheromone sources by different male-response types from two host races. *Journal of Applied Entomology*, 106:217-231.

Berryman, A. A. 1986. On the dynamics of blackheaded budworm populations. *Canadian Entomologist*, 118:775-779.

Bovey, P., and J. K. Maksymov. 1959. Le problème des races biologiques chèz la tordeuse grise du mélèze *Zeiraphera griseana* (Hb.). *Vierteljahresschrift der Naturforsch. Ges. Zurich*, 104:264-274.

Day, K. 1984. Phenology, polymorphism and insect-plant relationships of the larch budmoth, *Zeiraphera diniana* (Guenee) (Lepidoptera, Tortricidae), on alternative conifer hosts in Britain. *Bulletin of Entomological Research*, 74:47-64.

Edwards, A. W. F. 1972. *Likelihood*. Cambridge University Press, Cambridge.

Emelianov, I., W. Baltensweiler, M. Drès, and J. Mallet. in prep. Host fidelity of larch and pine race *Zeiraphera diniana* (Lepidoptera: Tortricidae).

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the Larch Budmoth) - polymorphism, host races or sibling species. *Heredity*, 75:416-424.

Goday-Herrera, R., B. Burnet, K. Connolly, D. Grey, and I. Weir. 1994. Disruption of the structure of larval foraging behaviour in interspecific hybrids in *Drosophila*. *Heredity*, 72:260-268.

Guerin, P., W. Baltensweiler, H. Arn, and H. Buser. 1984. Host race pheromone polymorphism in the larch budmoth. *Experientia*, 40:892-894.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12:101-109.

SAS Institute. 1996. JMP Statistics for the Macintosh version 3.1.6.

Komazaki, S. 1986. The inheritance of egg hatching time of the spiraea aphid, *Aphis citricola* Van der Goot (Homoptera, Aphididae) on the two winter hosts. *Kontyû*, 54:48-53.

Pezzullo, J. P. 2000. Interactive statistics page. :http:/ / members.aol.com /johnp71 /confint.html (8.2000).

Reissig, W. H., and D. C. Smith. 1978. Bionomics of *Rhagoletis pomonella* on *Crataegus*. *Annals of the Entomological Society of America*, 71:155-159.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*. W H Freeman and Company, New York.

Effect Likelihood ratio tests				
Source of variation	No. parameters	*DF*	*Likelihood ratio Chi Square*	*P-value*
Cross Type	3	3	3.3584058	0.3396
Cohort	1	1	0.4486136	0.5030
Cross Type X Cohort	3	3	3.8348971	0.2799

Table 1. Variation in proportion of hatching eggs with cohort and cross type.

Cross type is (i) larch; (L), (ii) pine; (P), (iii) hybrid; (LP) or (PL),or (iv) hybridxlarch; (L)x(LP), (L)x(PL), (LP)x(L), or (PL)x(L).

Effect Likelihood ratio tests				
Source of variation	No. parameters	*DF*	*Likelihood ratio Chi Square*	*P-value*
Cross Type	3	3	0.4225012	0.9356
Cohort	1	1	1.4843019	0.2231
Cross Type X Cohort	3	3	1.8760841	0.5985

Table 2. Variation in proportion of hatching eggs with cohort and cross type, excluding broods of less than 10 eggs

Cross type is (i) larch; (L), (ii) pine; (P), (iii) hybrid; (LP) or (PL), or (iv) hybridxlarch; (L)x(LP), (L)x(PL), (LP)x(L), or (PL)x(L).

Effect Likelihood ratio tests				
Source of variation	No. parameters	*DF*	*Likelihood ratio Chi Square*	*P-value*
Cross Type	2	2	3.4780357	0.1757

Table 3. Variation in proportion of 1998 cohort larvae surviving to fifth instar with cross type.

Cross type is (i) larch; (L), (ii) hybrid; (LP) or (PL), or (iii) hybridxlarch; (L)x(LP), (L)x(PL), (LP)x(L), or (PL)x(L).

Effect Likelihood ratio tests				
Source of variation	No. parameters	*DF*	*Likelihood ratio Chi Square*	*P-value*
Cross Type	2	2	3.8964721	0.1425

Table 4. Variation in proportion of 1998 cohort larvae surviving to fifth instar with cross type, excluding broods of less than 10 eggs (see Figure 2.)

For description of cross type see Table 3.

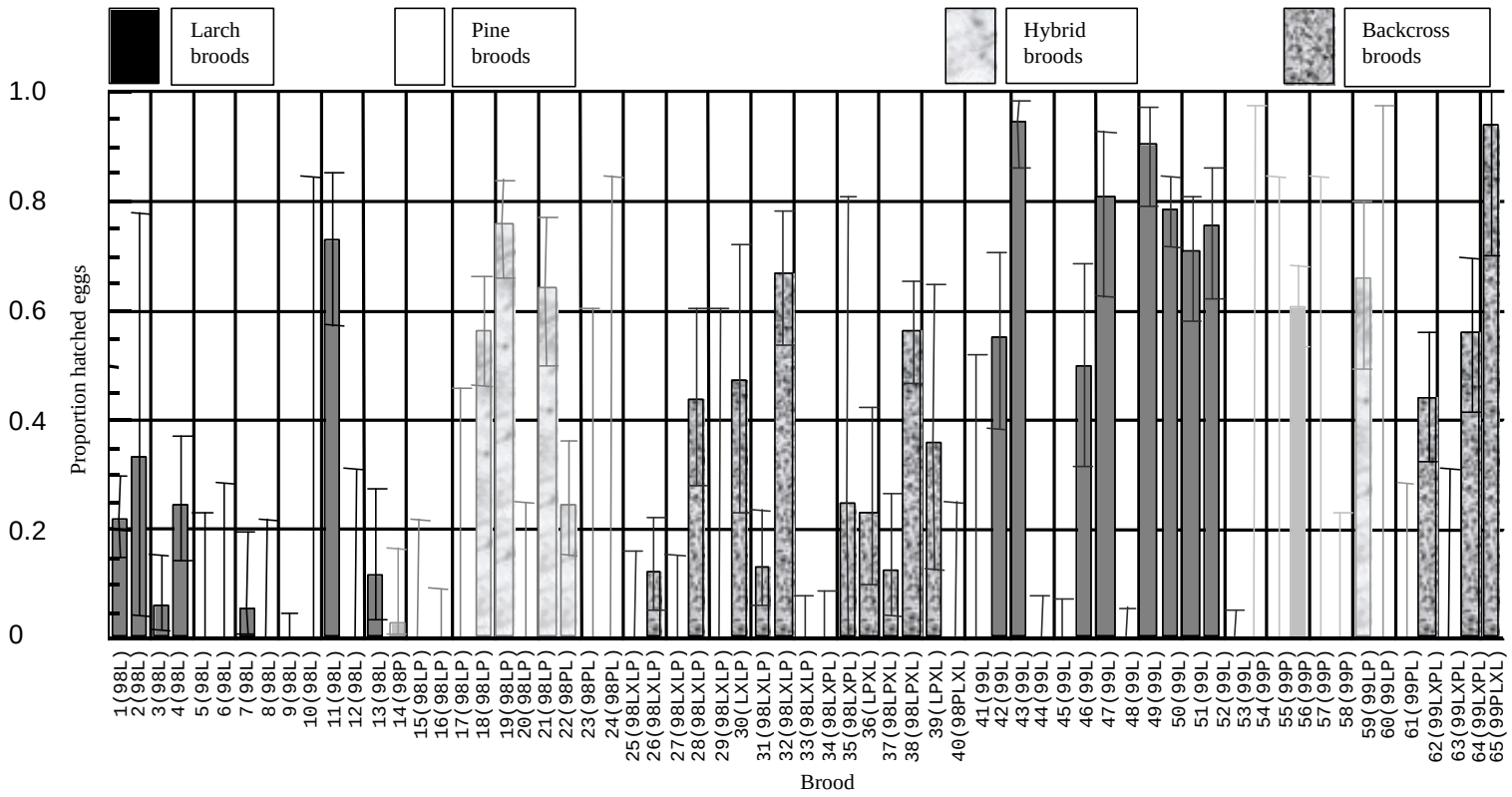

Figure 1. Proportions of hatching eggs, 98 and 99 cohort broods. Year and direction of cross indicated beneath each brood. Error bars are binomial 95% confidence limits.

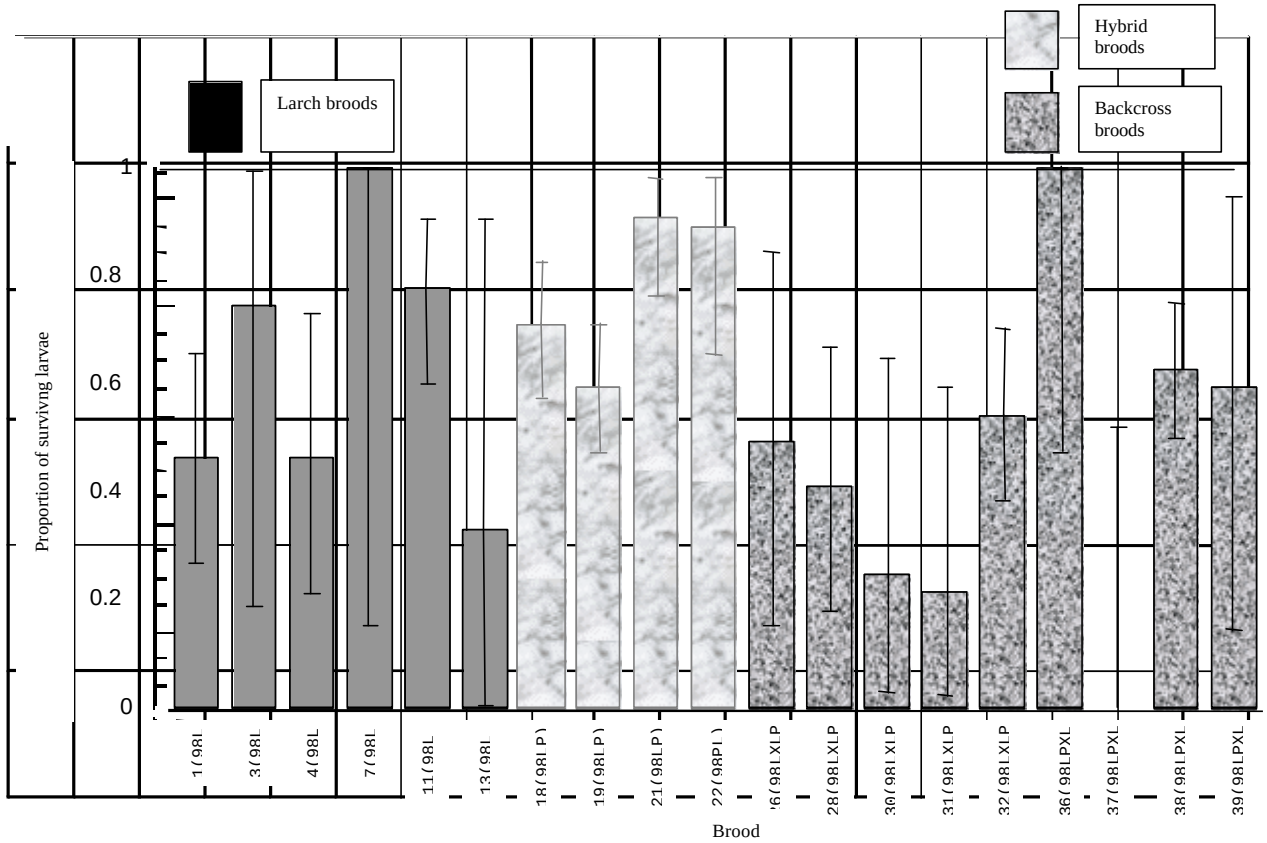

Figure 2. Proportions of larvae surviving to fifth instar, 98 cohort broods. Year and direction of cross indicated beneath each brood. Error bars are binomial 95% confidence limits.

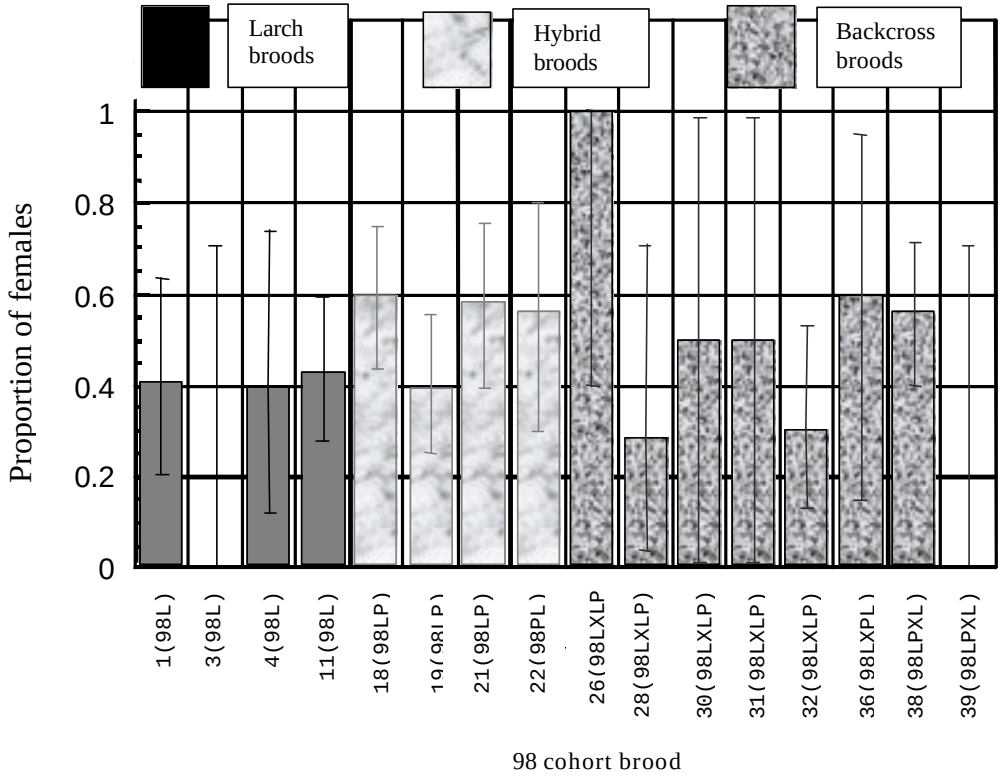

Figure 3. Proportion of females, 98 cohort broods. Brood type and cross is indicated beneath each brood. Error bars are binomial 95% confidence intervals.

Chapter 4

Lack of mitochondrial DNA divergence between host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae)

Abstract

Larch and pine associated races of *Zeiraphera diniana* cannot be distinguished on the basis of nucleotide sequence across an 811 bp region of mitochondrial COI (partial), tRNA-leu, and COII (partial). This lack of differentiation supports the hypothesis that the two biotypes are host races that undergo gene flow, and also suggests that at least some of this gene exchange is female-mediated. However, some caution is necessary whenever the genetic state of a population is used to make inferences about current levels of gene flow, and this data does not rule out a very recent origin of reproductive isolation.

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). *Z. diniana* is a model system for the study of population cycles, as larch feeding populations of the French and Swiss Alps undergo one of the most extreme density fluctuations of any organism, 10^5 fold over an eight to ten year period (Baltensweiler, 1993). The species is a serious pest of this region, and its larvae are responsible for wide scale defoliation of larch trees at the peak of each cycle. Pine feeding *Z. diniana* larvae are often found in sympatry with cycling larch populations, but their density on this host, although less well studied, appears to be more stable (Baltensweiler, pers. com.).

Larch and pine biotypes of *Z. diniana* are also examples of genetically differentiated host races (Chapter 2). Both races exhibit host fidelity in the field, where they tend to roost overnight on their native host (Emelianov et al., 1995), and in laboratory choice chambers, where they favour their natal host as a resting (Emelianov et al., in prep) and egg laying (Bovey and Maksymov, 1959) site.

In addition to host choice, the races exhibit heritable differentiation in a number of other phenotypic traits. Fifth instar larch larvae are normally solid black while pine larvae are yellow-green (Baltensweiler, 1993), and the timing of emergence of each form is tuned to the bud burst of its host, approximately two weeks later for Alpine pine than for Alpine larch trees (Day, 1984). Furthermore, although the difference is not diagnostic, larch adults tend to be larger than pine adults (W. Baltensweiler, unpublished data). The two forms are also pheromone races. Larch race females produce a pheromone consisting primarily of trans -(E)11-tetradecenyl acetate (E11-14:Ac), while the pheromone of pine race females is a 1000:1 blend of trans-(E)9-dodecenyl acetate (E9-12:Ac): (E11-14:Ac) (E11-14:Ac) (Baltensweiler et al., 1978).

Gene flow between larch and pine races of *Z. diniana* is restricted. In experiments involving artificial pheromone lures (Priesner and Baltensweiler, 1987) and confined virgin females (Appendix 1), males responded more strongly to the pheromone of their own host race. Pheromone-mediated long range attraction (Appendix), and mating at

short range (Chapter 2) is positively assortative with respect to natal host, and the peak of adult emergence of the larch and pine races differs by approximately two weeks in our study sites (Baltensweiler, pers. com.). Despite these barriers, larch and pine races continue to undergo appreciable gene flow. Field emergence of the two forms normally overlaps for about two weeks (Baltensweiler, pers. com) and combined estimates of long range cross attraction in the field, and short range hybridisation in the laboratory suggest that they hybridise at a rate of approximately 2.4% per generation (Chapter 2).

Genetic differentiation between larch and pine race *Z. diniana* has previously been studied using allozyme loci (Emelianov et al., 1995). Of the 13 polymorphic loci investigated, significant frequency differences between larch and pine races exist at three loci, *Mdh*-s, *Pgm*, and sex linked *Idh*-s, and are absent at the remaining 10 (Emelianov et al., 1995). Although there are alternative explanations for this pattern (Emelianov et al., 1995) it is possible that continuing gene flow between the races maintains homogeneity across most loci, while host specific selection at *Mdh*-s, *Pgm*, *Idh*-s or closely linked loci reduces exchange at these genes. We would like to investigate the level of genetic divergence between the host races further, using additional genetic markers. Here, we examine mitochondrial DNA sequence differentiation within and between host races of *Z. diniana* across an 811 bp fragment of 3' cytochrome oxidase I (COI), the transfer RNA leucine gene (tRNA-leu), and 5' cytochrome oxidase II (COII).

The mitochondrial cytochrome oxidase I and II genes encode transmembrane proteins of cytochrome oxidase *c*, one of a series of respiratory enzyme complexes involved in pumping electrons across the inner mitochondrial membrane (Saraste, 1999). Cytochrome oxidase I consists of alternating internal loops, transmembrane helices involved in metal binding and electron pumping, and external loops. The majority of the cytochrome oxidase II protein lies outside the membrane, and is involved in copper binding. The COII copper binding region is preceded by two membrane spanning regions, a membrane bound internal loop, and a linking region of unknown function (Saraste, 1990). The region sequenced in this study consisted of external loop six (E6) (partial), transmembrane helix twelve (M12), and the carboxy terminus (COOH) of COI, the entire transfer RNA leucine gene, and the amino terminus (NH_2), transmembrane helix one (M1), interval region (I1), transmembrane helix two (M2), linking (Lnk) and copper binding(Cu) (partial) regions of COII. With the exception of the tRNA leucine gene, which, like other mitochondrial tRNA genes, is subject to more structural and functional constraints than protein coding sequences (Simon et al., 1994), this is a region

of low mtDNA sequence conservation. The M12 and COOH terminus of COI, and the MI and Lnk regions of COII have contained peaks of mtDNA nucleotide diversity in several other studies (reviewed in Simon, 1994), including a recent comparison of several species and subspecies butterflies of the genus *Papilio* (Caterino and Sperling, 1999).

Mitochondrial DNA is a useful tool for the detection of genetic differentiation between closely related taxa (Harrison, 1989; Simon, 1994). Its maternal mode of inheritance renders its effective population size approximately half that of nuclear DNA, and increases the chance that selectively advantageous or neutral mutations will spread, as does its haploid copy number and absence of crossing over at meiosis (Simon et al., 1994). Furthermore, the ability of individual mitochondrial tRNAs to recognise four codons, instead of the usual three, increases the redundancy in the mitochondrial genetic code, and allows more 'silent' third-codon replacements than in nuclear DNA (Li and Graur, 1991; Tomita et al., 1999). In mammals, mitochondrial genomes evolve up to 10 times more rapidly than the nuclear genome (Simon et al., 1994), and, although variability may not be as high in insects, *Drosophila* mtDNA genes also seem to evolve quickly, at similar rates to the most rapidly evolving nuclear genes (Simon et al., 1994). The potential utility of mitochondrial DNA in the study of *Z. diniana* host races, and similar systems, is enhanced by the fact that the entire mitochondrial genome operates as a single, maternally inherited marker, and is physically unlinked to sites on the nuclear genome that may be subject to host associated selection (Harrison 1989). In this case the loci may include divergent allozyme (or closely linked) loci (Emelianov et al., 1995a), and loci controlling larval colour pattern and emergence time (Chapter 3). Thus, any mtDNA differentiation between *Z. diniana* could be a sign of asymmetries between male and female mediated gene flow, and examination of mitochondrial DNA markers could, for example, allow quantification of the frequency with which *Z. diniana* females oviposit on the alternate host in the field.

Materials and Methods

Collections and rearing

Collections of fourth and fifth instar *Z. diniana* larvae were made in June 1996 and 1999 at six sites in the European Alps; (1) Bois les Ayes, France (44°50' N, 6°39'E): larch and pine races, (2) Bois les Combes, France (44°54' N, 6°34'E): larch race, (3) Montgénèvres, France (44°56', 6°43' E): larch race, (4) Pontresina, Switzerland (46°29' N, 9°54' E): larch and pine races, Sils, Switzerland (46°42'N, 9°27' E): larch race, and Val Bever, Switzerland (46°26'N, 9°50'E): larch and pine races (Figure 1., Chapter 2). Individuals were named according to their collection site as follows; LA=Bois les Ayes, France, LC=Bois les Combes, France, MG=Montgénèvres, France, P=Pontresina, Switzerland, S=Sils, Switzerland, VB=Val Bever, Switzerland. Each individual name was followed by a suffix indicating its host race; (L)= larch race, or (P)=pine race.

In order to ensure that multiple sibling cohorts were sampled, collections were made from as many trees as possible at each site. Larvae were classified as larch or pine race according to the host tree, *Larix* spp. or *Pinus cembra*, from which they were collected, and individually reared to pupation on newly flushing larch (*Larix* spp.) in 15x50 mm Sterilin plastic vials with cotton wool stoppers. Insects were transported back to the laboratory within two weeks of collection, and the majority of rearing took place in a constant temperature (CT) room kept at 17-21 °C (1999), and a cooled incubator kept at 12-14 °C (5^{th} instar larch larvae only, 1998 and 1999), or on the laboratory bench top (1997). Humidity was maintained at 70-80% throughout laboratory rearing. Rearing tubes containing pupae were cleared of all food and debris prior to adult emergence. Insects were sexed as pupae according to their number of unfused abdominal segments, seven in the male and six in the female (Bradley et al., 1979), or, in the case of the more translucent pine race, according to the presence or absence of testes in fifth instar larvae. Adults were frozen at - 20 or - 80 °C for molecular analysis.

DNA isolation

All insects were dissected on dry ice. Heads, legs and wings were removed, and thoraces placed in 2.5 ml eppendorf tubes, immersed in either in (i) 250µl TNES extraction buffer

(100mM Tris pH 8.0, 400mM NaCl, 100mM EDTA pH8.0, 0.5% SDS) and 20µl 50µg / ml Proteinase K (10 samples) or (ii) liquid nitrogen (30 samples), and ground to a fine paste or powder using disposable plastic pestles. Samples homogenised in liquid nitrogen were subsequently mixed with 250µl TNES and 2µl 50µg / ml Proteinase K. After incubation overnight at 50°C, samples were vortexed for 5 seconds, mixed with 270µl 2.6M NaCl and 20µl Chloroform, vortexed for a further 5 seconds, and centrifuged at 13,000 rpm for five minutes. The resulting supernatants were removed to fresh tubes, mixed with a double volume of isopropanol, and centrifuged at 13,000 rpm for 30 minutes. The new supernatants were discarded. After washing with 250µl 70% ethanol, pellets were dissolved overnight at 5°C in 50 µl of ultra pure water, and stored at - 20°C until use.

PCR

Samples were amplified using Harrison laboratory mtDNA primers 'George III' 5'-TAGGTITAGCIGGAATACCTCG-3' (I=isonisol) and 'Eva' 5'-GAGACCATTACTTGCTTTCAGTCATCT-3' (Bogdanowicz et al., 1993: Fig. 2). 25µl polymerase chain reactions were carried out in a Perkin-Elmer Gene Amp 2400 PCR system 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 reaction conditions were 0.1µM each primer, 40µM each of Pharmacia Biotech Ultrapure dATP, dTTP, dCTP, and dGTP, 1X Advanced Biotechnologies Buffer AB-0289 (200mM $(NH_4)_2SO_4$, 750mM Tris-HCL, pH8.8, 0.1% Tween® 20), 3.3mM Advanced Biotechnologies or Perkin Elmer $MgCl_2$, 1 unit of Advanced Biotechnologies Thermostable DNA polymerase, and 2µl containing 40-160ng of undiluted template DNA (quantitated by spectroflurometer). 5 µl aliquots of PCR reactions were electrophoresed for one hour at 200 milliamps on 0.8% agarose gels. Marker lanes containing GibcoBRL 100 bp DNA Ladder were included on each gel. After electrophoresis, gels were soaked for twenty minutes in a 0.5µg / ml ethidium bromide solution and viewed under an ultraviolet lamp to check for amplification of single, equal sized DNA fragments from each sample.

Sequencing

Products of PCR reactions were purified with a Qiagen Quiaquick purification kit according to the manufacturers instructions, and at least 125 ng of each sample sent to Cambridge BioSciences for automated sequencing in an Abi Prism TM 377 DNA

sequencing machine. Four sequencing primers were used, (i) George III (see above), (ii) *Z. diniana* specific primer ZDF-1 5' – GCTTCTCCTTTAATAGAAC - 3', (iii) *Z. diniana* specific primer ZDF-2 5' CTTCCACCAGCTGAACAT – 3', and (iv) *Z. diniana* specific primer ZDR-1 5'-GTTCTATTAAAGGAGAGGCTCTATTTTG-3' (Fig. 2). The design of *Z. diniana* specific primers was necessary due to poor quality of several sequences obtained from George III, mainly because of multiple overlapping sequence towards the 5' end. The multiple sequencing may have been due to the presence of isonisol residues in this primer allowing priming at multiple sites, but could also have been caused by binding of George III to primer-diamers, which were faintly visible after electrophoresis. New internal primers not expected to bind to primer-diamers were subsequently designed in Web Primer (Cherry et al.) from George III primed *Z. diniana* sequences, and their suitability tested again using Amplify (version 1.2) for the Macintosh (Engles, 1993). At least 3µl of each primer, at a concentration of 5pmol/µl, was sent to Cambridge Biosciences per sequencing reaction. Sequences were received as Chromas (version 1.2) (McCarthy, 1997) chromatogram files, and edited by hand.

Alignments

Sequences were aligned using Clustal W Version 1.7 (Thompson et al., 1994), LALIGN (UK HGMP), and Sequencher Version 3.1.1 for the Macintosh (Gene Codes Corporation, 1998). A consensus sequence compiled from overlapping sequences obtained from the three primers GIII, ZDR-1, and ZDF-1 was tested for similarity to sequences deposited in GenBank (Benson et al., 2000) using BLAST version 1.4 (Altschul et al., 1990). Each fragment of the consensus sequence was shared by at least two larch and two pine samples (Figures 2 and 3). Where bases were polymorphic, the most common sequence was used.

Putative amino acid changes were determined with reference to the *Drosophila* mitochondrial genetic code (Clary and Wolstenhome, 1985).

Results

Zeiraphera diniana *mtDNA consensus sequence*

Mitochondrial DNA fragments of approximately 900bp were amplified from a total of 22 larch race and 18 pine race samples. An 804 bp mtDNA sequence compiled from overlapping sequences from the primers GIII, ZDF-1, and ZDR-1 (Figure 1) will be deposited in GenBank. The sequence corresponds to bp 2766- 3576 of the *Drosophila yakuba* reference sequence (Clary and Wolstenhome, 1985), and bp 579- 1382 of *Choristoneura rosaceana* (Lepidoptera: Tortricidae) sequence GenBank L19099 (Sperling and Hickey, 1994: Figure 2), the published sequence to which it had the greatest similarity. The *Z. diniana* sequence, by comparison with the *C. rosaceana* sequence, and other insect sequences (Lunt et al., 1996; Saraste, 1990), consists of the external loop six (E6) (partial), transmembrane helix 12 (M12), and carboxy terminus (COOH) structural regions of COI, the entire tRNA leucine gene, and the N terminal leader peptides (NH2), transmembrane helix one (M1), interval one (I1), transmembrane helix two (M2), linking region (Lnk), and copper binding (Cu) (partial) structural regions of COII (Figure three). The A-T composition of the sequence is 77.2%.

Comparison of Zeiraphera diniana *and* Choristoneura rosaceana *sequences*

An alignment of *Z. diniana* and *C. rosaceana* (Sperling and Hickey, 1994) mtDNA sequences is presented in Figure 3. The 804 bp *Z. diniana* mtDNA consensus sequence differed from the *C. rosaceana* sequence at 45 third-codon, three second-codon and 11 first-codon positions, a total of 59 sites, equal to 7.3% differentiation. Thirteen putative amino acid substitutions were observed, but no insertions or deletions had taken place. The transition:transversion ratio across the region of comparison was 0.6 (23 transitions and 36 transversions).

Nucleic acid differentiation between the two sequences was highest in COII (30 third-codon, two second-codon, eight first-codon substitutions, = 8.1%), followed by COI (14 third–codon, one second-codon, and three first-codon substitutions, = 7. 2%), and absent across the entire tRNA-leucine sequence. The level of amino acid differentiation across COI was 6.0% and, and across COII, 4.8%. Within COI, overall nucleic acid

differentiation was 10.9% across the E6 structural region, 9.8% across the M12 structural region, and 5.0% across the COOH terminus. Amino acid differentiation was 5.5% across the E6 structural region, 0% across the M12 structural region, and 8.5% across the COOH terminus. The majority of COII nucleic acid differentiation between the species occurred across the linking (10.0%) and copper binding (14.8%) regions. Levels of divergence between the *C. rosaceana* and Z. *diniana* sequences were 1-2% across all other regions of COII. Within COII, amino acid differentiation occurred only in the linking (2.0%) and copper binding (15.5%) regions

Aligned larch and pine race sequences

Aligned *Z. diniana* mtDNA sequences are presented in Figure 3. Four samples; S(L)-2, P(P)-3, and P(P)-8, extend considerably less far into 5' COI than the majority of other sequences, and nine samples; LC(L) – 1, S(L) - 1, VB(L) - 1, LA(P) - 1, LA(P) - 4, P(P) - 4, P(P) - 6, VB(P) – 1, and VB(P) - 2, considerably less far into 3' COII. There is a 46bp gap across COII (NH_2) in the sequence of sample VB(L)-2.

Intra-race mtDNA variability

In total, inspection of intra and inter race sequence differentiation was possible across an 811bp fragment of COI (partial), tRNA-leucine, and COII (partial) (Figure 3). Total nucleotide variability within the larch race was 1.0%, the result of seven third-codon and one first-codon variable sites. Total nucleotide variability within the pine race was also 1.0%, and consisted of one first-codon, one second-codon, and six third-codon variable sites. Within both races nucleic and amino acid variability was unequally distributed across the 811bp haplotype. In the case of the larch race, nucleotide substitutions occurred only in the COI E6 (one site = 1.6%), COI M12 (two sites =3.9%), COI COOH (one site=0.7%), COII M1 (2 sites = 3.5%), and COII Lnk (1 site=0.7%) structural regions. The distribution of pine race nucleotide variation across COI structural regions was similar to that within the larch race, although only a single site of differentiation (= 2.0%) was observed across the COI M12 structural region. Across COII, pine race nucleotide variation occurred only in the I1 (one site = 2.6%) Lnk (four sites = 2.7%) and Cu (two sites = 1.5%) structural regions. No variation was found in the tRNA-leu, COII NH_2, or COII M2 structural regions of either race. One first-codon substitution in the COII M1 structural region, shared by one larch (VB(L)-1) and one pine (VB(P)-3) individual resulted in a putative valine to isoleucine replacement, and one second-codon

substitution in the COII Cu structural region, observed in a single pine individual (VB(P)-3), with the putative replacement of methionine with threonine. The transition:transversion ratio within the larch race was 7.0 (seven transitions, one transversion) and within the pine race it was 9.0 (nine transitions, one transversion). Samples VB(L)-1 and VB(P)-3 shared the same transversion substitution.

Inter-race mtDNA differentiation

Nucleotide information available at all 15 variable sites detected across the 811 bp haplotype is summarized in Table 1. In total, six haplotypes were detected, the consensus haplotype, and five others hereafter referred to as 'LI', 'LII', 'LIII', 'LPa', and 'LPb'. All substitutions described below are relative to the consensus haplotype.

The consensus haplotype was common to 17 larch race and 17 pine race moths. Haplotype LI contained a single third-codon substitution within COII (M1), and was scored in two larch samples, one from Bois les Combes, France (LC(L)-1) and one from Sils, Switzerland (S(L)-1). Haplotype LII, detected in a larch sample from Montgénèvres, France (MG(L)-3) only, contained a single third-codon substitution within COI (M12), while haplotype LPa, scored in a larch sample from Val Bever, Switzerland (VB(L)-1), had the first-codon substitution in COII (M1) that resulted in the putative replacement of valine with isoleucine, and four third-codon substitutions, three within COI (E6, M12, COOH), and one within COII (I1). Haplotype LPb, scored in pine sample VB(P)-3, also from Val Bever, Switzerland, contained the first-codon and all four third-codon substitutions present in haplotype LPa, an additional four third-codon substitutions within COII (Lnk, Cu), one first-codon substitution in COII (Lnk), and one second-codon substitution in COII (Cu). The latter substitution , as described in the previous section, resulted in the putative replacement of methionine with threonine. At all five variable sites where nucleotides could be scored in both VB(L)-1 (haplotype LPa) and VB(P)-3 (haplotype LPb) the two samples were identical. Finally, haplotype LIII, scored in one larch sample from Val Bever, Switzerland (VB(L)-3), contained a single third-codon substitution (shared with haplotype LPb) in COII (Lnk).

Discussion

Comparison of Zeiraphera diniana *and* Choristoneura rosaceana *sequences*

Lepidoptera tend to have highly A-T biased mtDNA, and the A-T ratio of the *Z. diniana* mtDNA consensus sequence (77.2%) presented in this study is within the range recorded for other species (Pashley and Ke, 1992). Similarly, the absence of nucleic acid differentiation between the *Z. diniana* and *C. rosaceana* tRNA leucine genes is unsurprising, given the high level of nucleotide and amino acid conservation, and short length of this sequence (Simon et al., 1994). The slightly greater extent of nucleic acid differentiation between *Z. diniana* and *C. rosaceana* across the COII gene, relative to that across the COI gene (8.1 vs. 7.4%), mirrors the results of Caterino and Sperling (1996) in their comparison of butterfly species of the genus *Papilio,* as does the greater percentage of first and second-codon substitutions within COII. However, in contrast to their study, which revealed a 60% higher level of amino acid variability within COII than COI, and the results of other studies suggesting a high conservation of COI residues relative to other mitochondrial genes, reviewed in Simon (1994), the *Z. diniana – C. rosaceana* comparison showed a slightly higher amino acid differentiation within COI (6.0% vs. 4.8% in COII).

In a comparison of COI sequences from Orthoptera *(Chorthippus parallelus, Locusta* sp.), Hymenoptera (*Apis gambiae, A. quadrim*), and Diptera (*Drosophila yakuba, D. sechellia, D.simulans, Phormia* sp), Lunt (1996) observed that the highest level of amino acid variability within COI occurred across the COOH terminus. However, our results agree with those of Caterino and Sperling (1993), who, in their study of *Papilio,* observed higher levels of both amino acid and sequence variability within the external loops and transmembrane helices immediately preceding the COOH transmembrane terminus than within the terminus itself. Within COII, the majority of nucleic acid differentiation between the species we compared was found was found across the linking (10.0%) and copper binding (14.8%) regions, which contained 2.3% and 15.5% of COII amino acid variability, respectively. Caterino and Sperling (1999) found a similar pattern of COII sequence and amino acid variability; peaks of both within the copper binding region, and high nucleotide but low amino acid variability across the linking region. Thus,

regions of peak mitochondrial DNA nucleotide and sequence diversity may differ between insect taxa (Caterino and Sperling, 1999).

Level of Zeiraphera diniana *mtDNA diversity*

The most obvious pattern in the data was the high degree of similarity of all *Z. diniana* sequences. The low level of differentiation within and between larch and pine races of *Z. diniana* does not support a detailed comparison of regions of amino acid conservation vs. those of variability. However, a few points are noteworthy. The distribution of sites which differed between individuals within *Z. diniana* was broadly similar to that observed between the *Z. diniana* consensus sequence and the *C. rosaceana* (Sperling and Hickey, 1994) sequence. The lack of variability across the tRNA-leucine is again consistent with the results of other studies (Simon et al., 1994) and, in both comparisons, the pattern of nucleotide variation within cytochrome oxidase I was higher within the M12 than within the COOH region, a result which contrasts with those of Lunt (1996) but is consistent with the findings of Caterino and Sperling (1999) within *Papilio.*

How does the level of mtDNA diversity within the larch and pine races of *Z. diniana* (1.0%) compare with those within other Lepidoptera? Most values of intraspecific mtDNA variation recorded for Lepidoptera have been measured across several races or subspecies, defined by a variety of traits, including geographic location (e.g. Chang, 1997, Brower 1994), pheromone composition (e.g. Newcomb et al., 1998, Sperling et al. 1996), and morphological characters (e.g. Brower 1994). In many cases, the type of differentiation – polymorphism, host race, or species-level -between these forms has not been investigated in detail prior to mtDNA analysis. Only where there is evidence that intraspecific races are in fact cryptic species or host races (Chapter 1) have intraracial rather than interacial values of genetic differentiation been reported here.

In a comparison of a 472 bp fragment of the COI M3-M6, I1, I2, and E2-E6 regions obtained from two individuals of the same *Ctenopseustis obliquana* (Lepidoptera: Tortricidae) pheromotype, Newcomb (1998) detected four nucleotide substitutions (=1.3% variation). Members of several of *C. obliquana* pheromotypes differ by concordant allozyme frequencies, pheromone composition, and morphology, and are suspected sibling species (Newcomb and Gleeson, 1998). Two ‘C pheromotype’ samples of the dingy cutworm *Feltia jaculifera* (Lepidoptera: Noctuidae), whose pheromotypes also differ in allozyme frequency, and may not undergo long range cross attraction in the

wild, differed by 0.6% across an 820 bp 3' COI sequence (Sperling et al., 1996), while comparison of the COI M3, M4, M5, I1, I2, E2, and E3 structural regions across a set of seven *Plutella xylostella* (Lepidoptera: Plutellidae) of the same geographic strain revealed only a single base pair change, equal to 0.27% variation (Chang et al., 1997). Fifteen variable sites (=3.4% variation) were detected across a 441bp fragment of central COI sequenced from 26 specimens of the hemlock looper *Lambdina fiscellaria* (Lepidoptera: Geometridae), representing three subspecies with some pheromonal and larval host plant differences (Sperling et al., 1999), and, in one of the most comprehensive studies of intraspecific mtDNA variation within the Lepidoptera, a 942bp region of mtDNA spanning 3'COI, tRNA-leu, and almost all of the COII gene yielded 90 variable sites (=9.55% variation) across 49 specimens of *Heliconius erato* butterflies, representing 14 geographical colour pattern races (Brower, 1996).

Thus, levels of mtDNA diversity within *Z. diniana* host races appear to lie towards the lower end of the range of values recorded for other Lepidoptera, particularly since, in cases where we have only cited intraracial values, sample sizes are very small, and therefore some values of mtDNA diversity presented in the studies above are likely to underestimates.

It is possible that the low level of mitochondrial DNA diversity in *Z. diniana* results from a founder effect, perhaps after glaciation in the last 20,000 years. There is no evidence for a reduction in levels of nuclear DNA diversity in these moths (Emelianov et al., 1995), suggesting that, if the number of female recolonists was indeed low, male numbers remained reasonably high. *Z. diniana* males regularly migrate long distances (Baltensweiler, W., 1977), but little is known about the migratory behaviour of females. A second possible explaination for the mitochondrial uniformity is a selective sweep. Because of genetic hitchiking, selection on any mitochondrial locus, or on a maternally inherited factor, would affect all mitochondrial sites (Marcade et al., 1999; Ballard and Kreitman, 1995). Further studies of female migration habits, and of the possibility of female infection with cytoplasmic parasites could help to distinguish between these possibilities.

Differentiation between the larch and pine races.

Larch and pine races of *Zeiraphera diniana* could not be distinguished on the basis of their mitochondrial DNA sequence. A large majority of both larch (77%) and pine (94%)

sequences shared the consensus haplotype across all sites at which they were scored (Figure 3). The comparison of haplotypes is however complicated by gaps in sequence information, particularly towards the 3' end of the haplotype (Figure.4, Table 1). Across this 3' region, which includes sites 3346-3526, loss of information was due to the gradual loss of signal from the majority of ZDF-1 primed sequences. All variable sites across this region were scored only in the relatively long sequence VB(P)-3 (Figure 3). However, the general lack of differentiation between the races is supported across several hundred base pairs of mtDNA sequence from each sample.

Despite this, the possibility of interracial frequency differences in some rare haplotypes cannot be completely ruled out. Haplotypes LI, LII and LIII, for example, were scored only in larch race samples, the former in two insects from geographically separated locations (Table 1, Figure 1, Chapter 2, and Figure 3). Likewise, it is possible that haplotype LP-b is unique to pine sample VB(P)-3 (and haplotype LP-a to larch sample VB(L)-1). Sequencing of VB(L)-1 from primer ZDF-1 should resolve the latter. Likewise, sequencing of more samples may be necessary to determine whether these apparent frequency differences in rare haplotypes are real, or simply artefacts of sample size.

Gene flow between the host races

The lack of mitochondrial DNA differentiation between larch and pine *Z. diniana* races is consistent with our estimate of continuing actual gene flow of approximately 2.4% per generation (Chapter 2, Appendix) between these forms, and also suggests that actual gene flow is leading to incorporation of genetic information via backcrossing. In addition, the results suggest that at least some of this gene flow is female mediated, and reciprocal, although the relative contribution of female mediated gene flow in this system is not known.

However, it is also possible that the lack of mtDNA differentiation between the larch and pine *Z. diniana* host races can be explained by a very recent cessation of gene flow between the forms. Several studies have failed to detect, or have detected only very low levels of mtDNA divergence between species for which there is scant evidence of hybridisation, suggesting that some currently reproductively isolated species do not differ in their mitochondrial DNA. For example, among three species of ermine moths (Yponomeutidae) less than 1% divergence was detected across 2.3kb of COI and COII (Sperling et al., 1995), no divergence was detected between the two Eurasian species of

Pierid butterfly, *Colias* (Ericolias) *croceus* and C. (Neocolias) *erate poliographus* across 333bp of COI (Pollock et al., 1998) (although the authors suggest that sequencing of a longer fragment may reveal variable sites), and the distribution of variation across a 410bp sequence of 39 *Lycaeides idas* and *L. melissa* COI failed to distinguish between the species (Nice and Shapiro, 1999). Harrison (1989) reviews additional examples.

Two lines of evidence lead us to believe that, in *Z. diniana*, recent acquisition of reproductive isolation is a less likely explanation for the absence of mtDNA differences. Firstly, field and laboratory mate choice experiments suggest that hybridisation occurs in the wild (Chapter 1, Appendix). Secondly, the fact that the single larch and single pine moth that share the same rare haplotype (LPa/LPb) were collected from the same mixed stand forest is suggestive of gene flow in this area.

Finally, it is also possible that I have simply failed to detect existing mtDNA differentiation. Although 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), other areas of mtDNA protein coding sequences are likely to be as, or even more variable (Harrison, 1989; Lunt et al., 1996; Simon et al., 1994). For example, the mtDNA 'control region', which does not code for protein, and has been observed to undergo particularly rapid nucleotide evolution in many insects (Simon et al., 1994; Sperling et al., 1999), may contain sites of differentiation between the two races (but see Taylor (1993) who failed to find any control region variation among Lepidopteran species of the genus *Jalmenus*).

Conclusions

The level of mitochondrial DNA diversity observed within both larch and pine races of *Z. diniana* (1.0%) seems low in comparison with levels seen in other Lepidoptera. Comparison of an 811bp fragment of mitochondrial COI (partial), tRNA-leu, and COII (partial) between the two races failed to reveal any evidence of differentiation along host plant lines. This lack of differentiation is consistent with our hypothesis of inter-racial

actual gene flow (i.e. movement of genes via hybridisation) of approximately 2.4% per generation. The little variation that does exist also suggests that actual gene flow results in incorporation of genetic information from one biotype into the other and (at least some) is female-mediated. However, caution is always advisable when drawing conclusions about current gene flow levels from genetic data, and studies of some apparently reproductively isolated species have failed to reveal any evidence of mitochondrial DNA differentiation.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215:403-410.

Ballard, W. O., and M. Kreitman. 1995. Is mitochondrial DNA a selectively neutral marker? Trends in Ecology and Evolution. *Trends in Ecology and Evolution*, 10:485-488.

Baltensweiler, W. 1977. Colour polymorphism and dynamics of larch bud moth populations (*Zeiraphera diniana* Gn. Lep. Tortricidae). *Mitteilungen - Schweizerische Entomologische Gesellschaft*, 50:15-23.

Baltensweiler, W. 1993. A contribution to the explaination of the larch budmoth cycle, the polymorphic fitness hypothesis. *Oecologia*, 93:251-255.

Baltensweiler, W., E. Priesner, H. Arn, and V. Delucci. 1978. Untershiedliche Sexuallockstoffe bei Lärchen und Arvenform des Grauen Lärchenwicklers (*Zeiraphera diniana* Gn., Lep. Tortricidae). *Entomol. Gesellsch.*, 51:133-142.

Benson, D., I. Karsch-Mizrachi, D. Lipman, J. Ostell, B. Rapp, and D. Wheeler. 2000. GenBank. *Nucleic Acids Research*, 28:15-18.

Berryman, A. A. 1986. On the dynamics of blackheaded budworm populations. *Canadian Entomologist*, 118:775-779.

Bogdanowicz, S. M., W. E. Wallner, J. Bell, T. M. Odell, and R. G. Harrison. 1993. Asian gypsy moths (Lepidoptera: Lymantriidae) in North America: Evidence from molecular data. *Annals of the Entomological Society of America*, 86:710-715.

Bovey, P., and J. K. Maksymov. 1959. Le problème des races biologiques chèz la tordeuse grise du mélèze *Zeiraphera griseana* (Hb.). *Vierteljahresschrift der Naturforsch. Ges. Zurich.*, 104:264-274.

Bradley, J. D., W. G. Tremewan, and A. Smith. 1979. *British Tortricid Moths*. The Ray Society, London.

Caterino, M. S., and F. A. H. Sperling. 1999. *Papilo* phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Molecular phylogenetics and evolution*, 11:122-137.

Chang, W. X. Z., B. E. Tabshnik, B. Artelt, T. Malvar, V. Ballester, J. Ferre, and G. K. Roderick. 1997. Mitochondrial DNA sequence divergence among geographic strains of diamondback moth (Lepidoptera: Plutellidae). *Annnals of the Entomological Society of America*, 90:590-595.

Cherry, J. M., C. Ball, S. Dolinski, S. Dwight, M. Harris, J. Matese, G. Sherlock, G. Binkley, H. Jin, S. Weng, and D. Botstein. Web Primer, Vol. 1999.

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

Day, K. 1984. Phenology, polymorphism and insect-plant relationships of the larch budmoth, *Zeiraphera diniana* (Guenée) (Lepidoptera: Tortricidae), on alternative conifer hosts in Britain. *Bulletin of Entomological Research*, 74:47-64

Emelianov, I., W. Baltensweiler, M. Dres, and J. Mallet. in prep. Host fidelity of larch and pine race *Zeiraphera diniana* (Lepidoptera: Tortricidae).

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae): polymorphism, host races or sibling species? *Heredity*, 75:416-424.

Engles, W. R. 1993. Contributing software to the Internet: the Amplify program. *Trends in Biochemical Sciences*, 18:448-450.

Gene Codes Corporation. 1998. Sequencher Version 3.1.1 for the Macintosh.

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

Li, and Graur. 1991. *Fundamentals of Molecular Evolution*. Sinauer Publications.

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

Marcade, I., C. Souty-Grossed, D. Bouchon, T. Rigaud, and R. Raimond.1999. Mitochondrial DNA variability and Wolbachia infection in two sibling woodlice species. *Heredity*, 83:71-78.

McCarthy, C. 1997. Chromas Version 1.42. Griffith University, Brisbane.

Newcomb, R. D., and D. M. Gleeson. 1998. Pheromone evolution within the genera *Ctentopseustis* and *Planotortrix* (Lepidoptera: Tortricidae) inferred from a phylogeny based on cytochrome oxidase I gene variation. *Biochemical Systematics and Ecology*, 26:473-484.

Nice, C. C., and A. M. Shapiro. 1999. Molecular and morphological divergence in the butterfly genus *Lycaedes* (Lepidoptera: Lycaenidae) in North America: evidence of recent speciation. *Journal of Evolutionary Biology*, 12:936-950.

Pashley, D. P., and L. D. Ke. 1992. Sequence evolution in mitochondrial ribosomal and ND-1 genes in Lepidoptera: Implications for phylogenetic analyses. *Molecular Biology and Evolution*, 9:1061-1075.

Pollock, D. D., W. B. Watt, V. K. Rashbrook, and E. V. Iyengar. 1998. Molecular phylogeny for *Colias* butterflies and their relatives (Lepidoptera: Pieridae). *Annals of the Entomological Society of America*, 91:524-531.

Priesner, E., and W. Baltensweiler. 1987. A study of pheromone polymorphism in *Zeiraphera diniana* Gn (Lep, Tortricidae) .1. Male pheromonal response types in european wild populations, 1978-85. *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie*, 104:234-256.

Saraste, M. 1990. Structural features of cytochrome oxidase. *Quarterly review of Biophysics*, 23:331-366.

Saraste, M. 1999. Oxidative phosphorylation at the fin de siecle. *Science*, 5:1488-93.

Simon, C., F. Frati, A. Becenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America*, 87:651-701.

Sperling, F., R. Byers, and D. Hickey. 1996. Mitochondrial DNA sequence variation among pheromotypes of the dingy cutworm, *Feltia jaculifera* (Gn) (Lepidoptera: Noctuidae). *Canadian Journal of Zoology*, 74:2109-2117.

Sperling, F., J. Landry, and D. Hickey. 1995. DNA-based identification of introduced ermine moth species in North America (Lepidoptera: Tortricidae). *Annals of the Entomological Society of America*, 88:155-162.

Sperling, F. A., and D. A. Hickey. 1994. Mitochondrial DNA sequence variation in the spruce budworm species complex (Choristoneura: Lepidoptera). *Molecular Biology and Evolution*, 11:656-665.

Sperling, F. A. H., A. G. Raske, and I. S. Otvos. 1999. Mitochondrial DNA sequence variation among populations and host races of *Lambdina fiscellaria* (Gn.) (Lepidoptera: Geometridae). *Insect Molecular Biology*, 8:97-106.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673-4680.

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

UK Human Genetics Resource Centre. LALIGN local alignment.

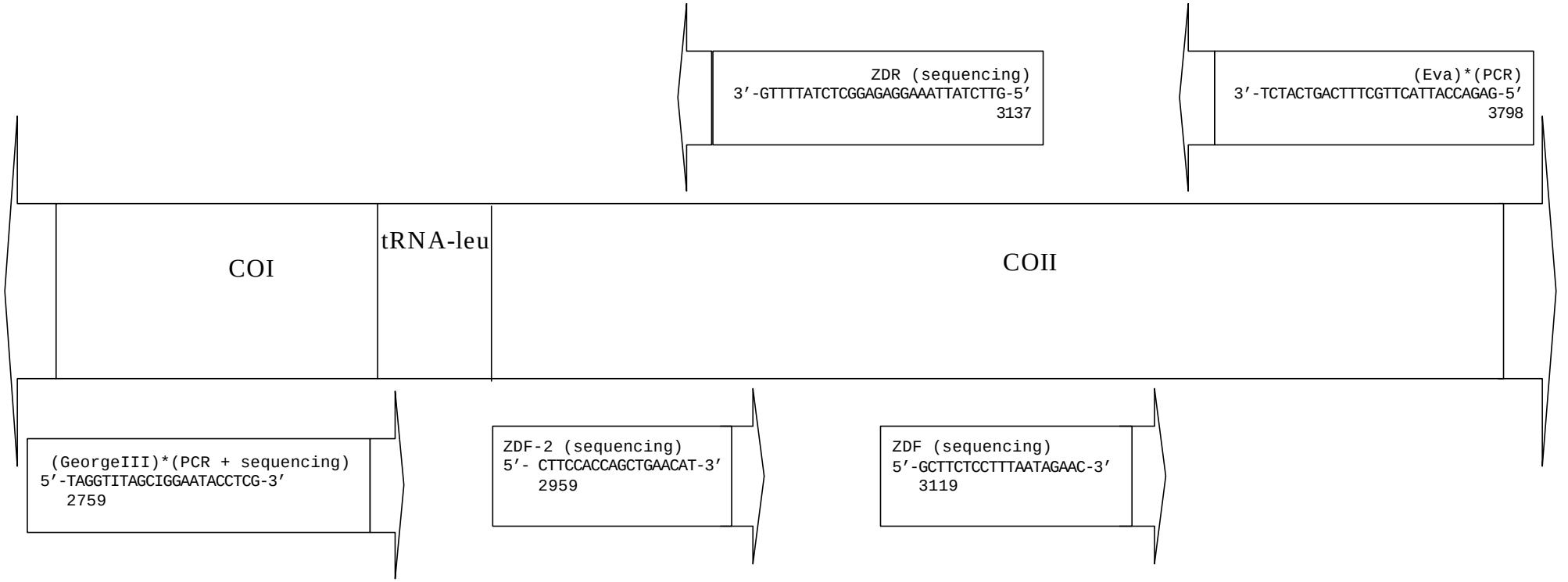

Figure 1. Direction and location of PCR and sequencing primers. Numbers indicate position of 5' bases relative to the *Drosophila yakuba* reference sequence (Clary, 1985) *Universal insect mtDNA primers (Bogdanowicz 1993) named according to the convention of Simon *et al*. (1994).

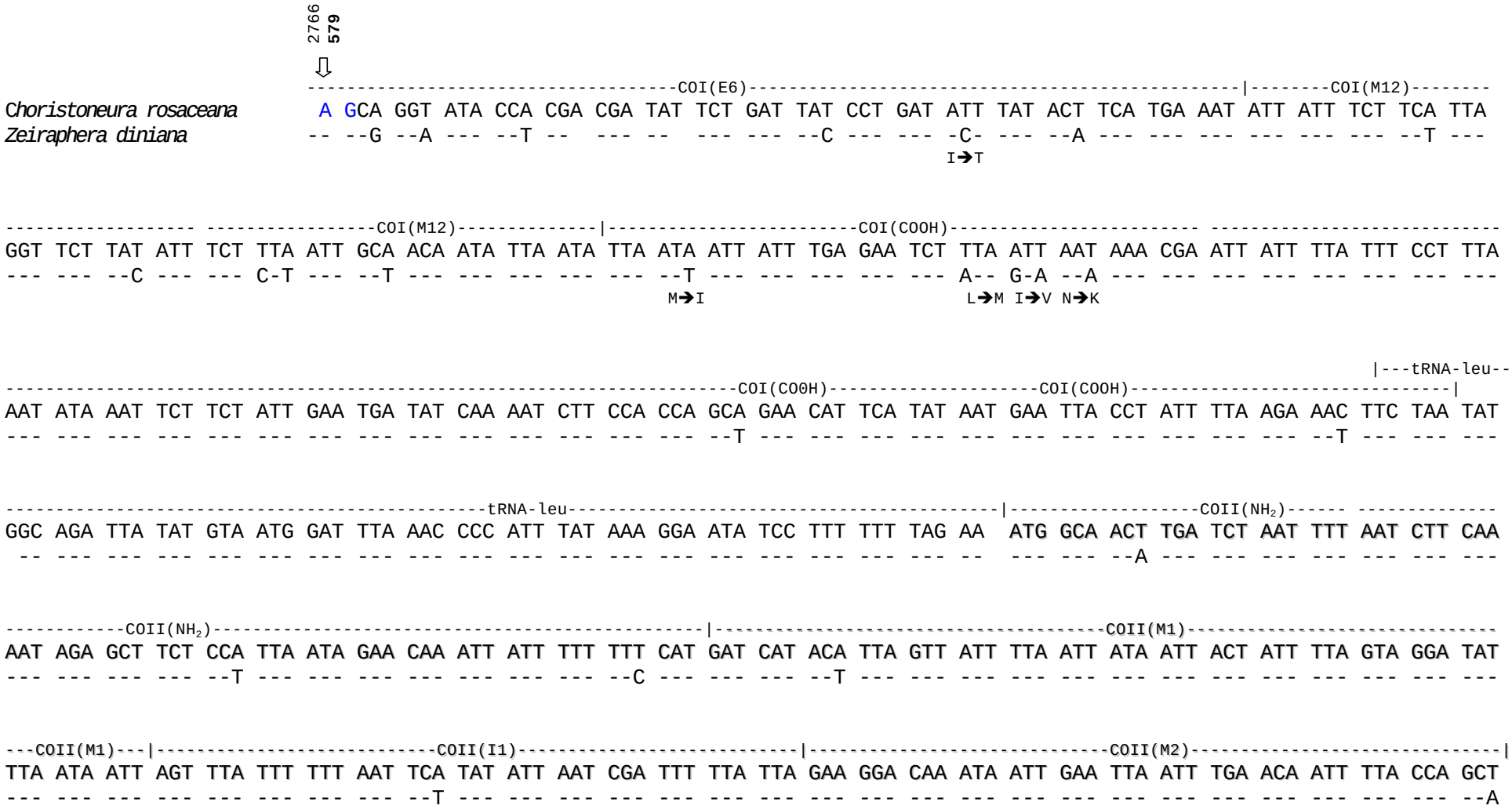

Figure 2. Comparison of *Zeiraphera diniana* consensus sequence and a published *Choristoneura rosaceana* sequence ——(Sperling and Hickey 1984). Dashes indicate identity of nucleic acid sequences. All putative amino acid differences are indicated beneath Z. *diniana* sequences - first residue is C. *rosaceana*, second is putative *Z. diniana*. Numbers above first and last Z. *diniana* bases indicate corresponding bases of the *D. yakuba* reference (first number, plain text) and *C. rosaceana* (second number, bold text) sequences

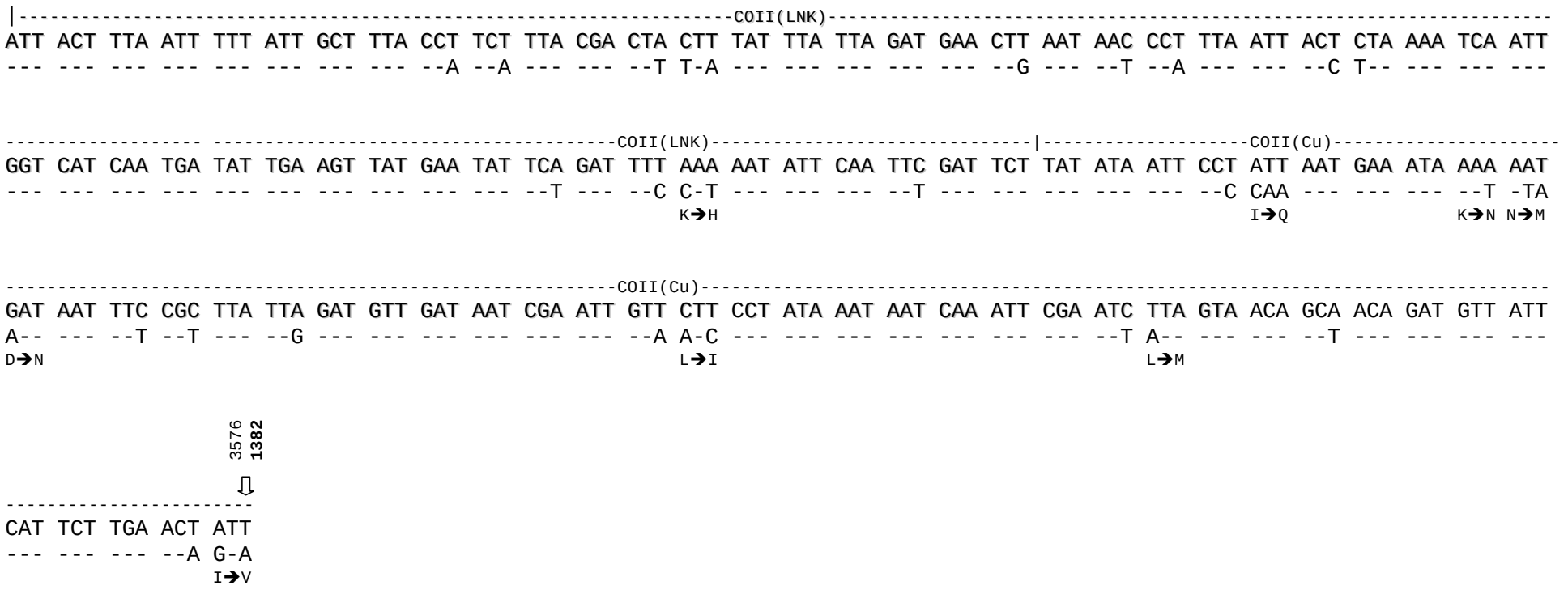

Figure 2. (continued)

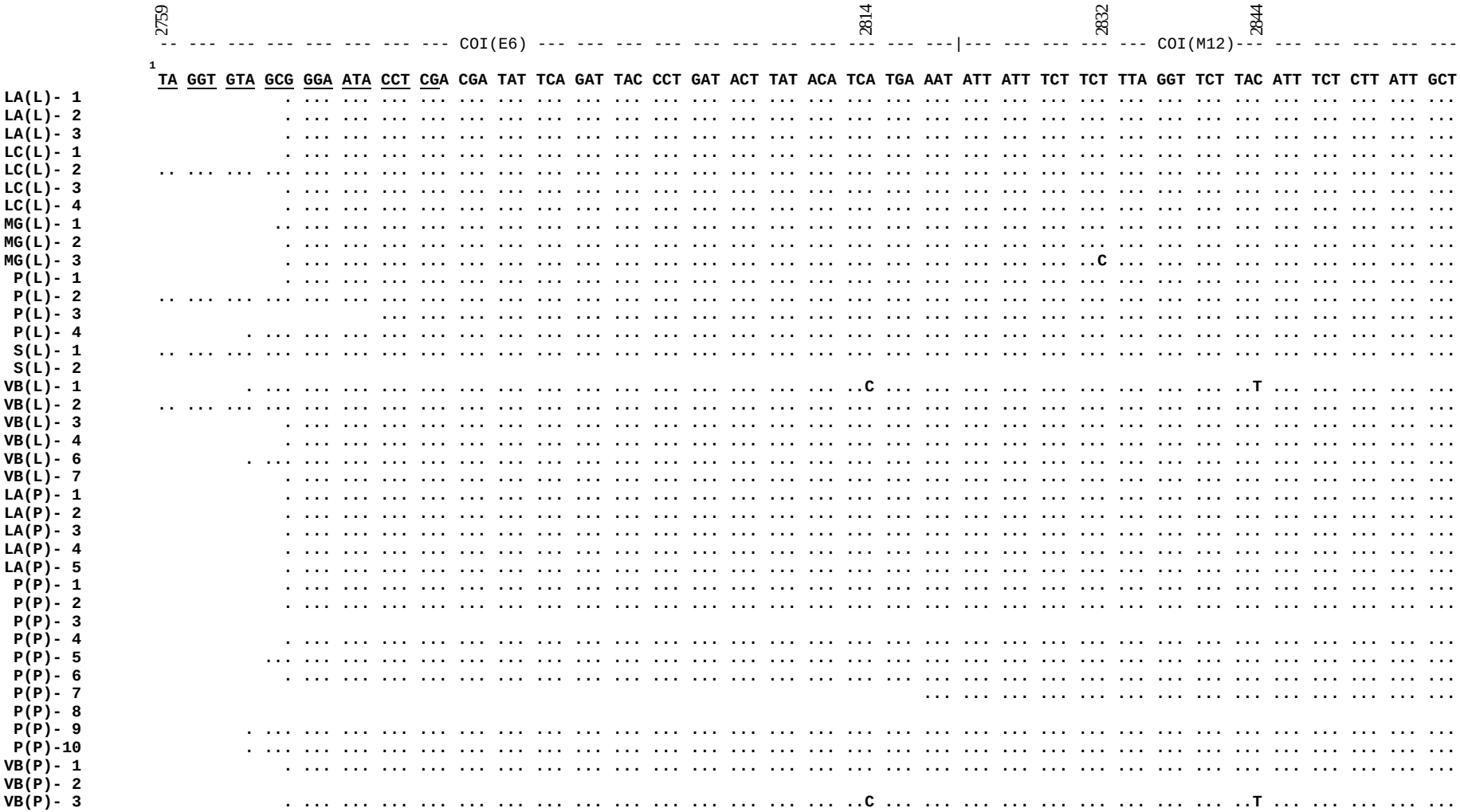

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences. Dots indicate identity of nucleotides to the sequence presented at the top of the alignment, which is divided into codons. Structural regions of COI and COIII are as described in the main text. Sites that differ between haplotypes are indicated below corresponding bases of the consensus sequence, as are any putative amino acid changes. Numbering refers to corresponding positions on the *D. yakuba* reference sequence (Clarey and Wolstenhome 1985). Letters of sample names before brackets refer to collection site, and those within brackets indicate biotype: (L) = larch race, (P) = pine race. Ns indicate failure to score.

1:Underlined bases correspond to GIII annealing site. **2**: Underlined bases correspond to ZDF-2 annealing site 3,4:Underlined bases correspond to (partially overlapping) ZDR-1 and ZDF-1 annealing sites (also see Figure 1).

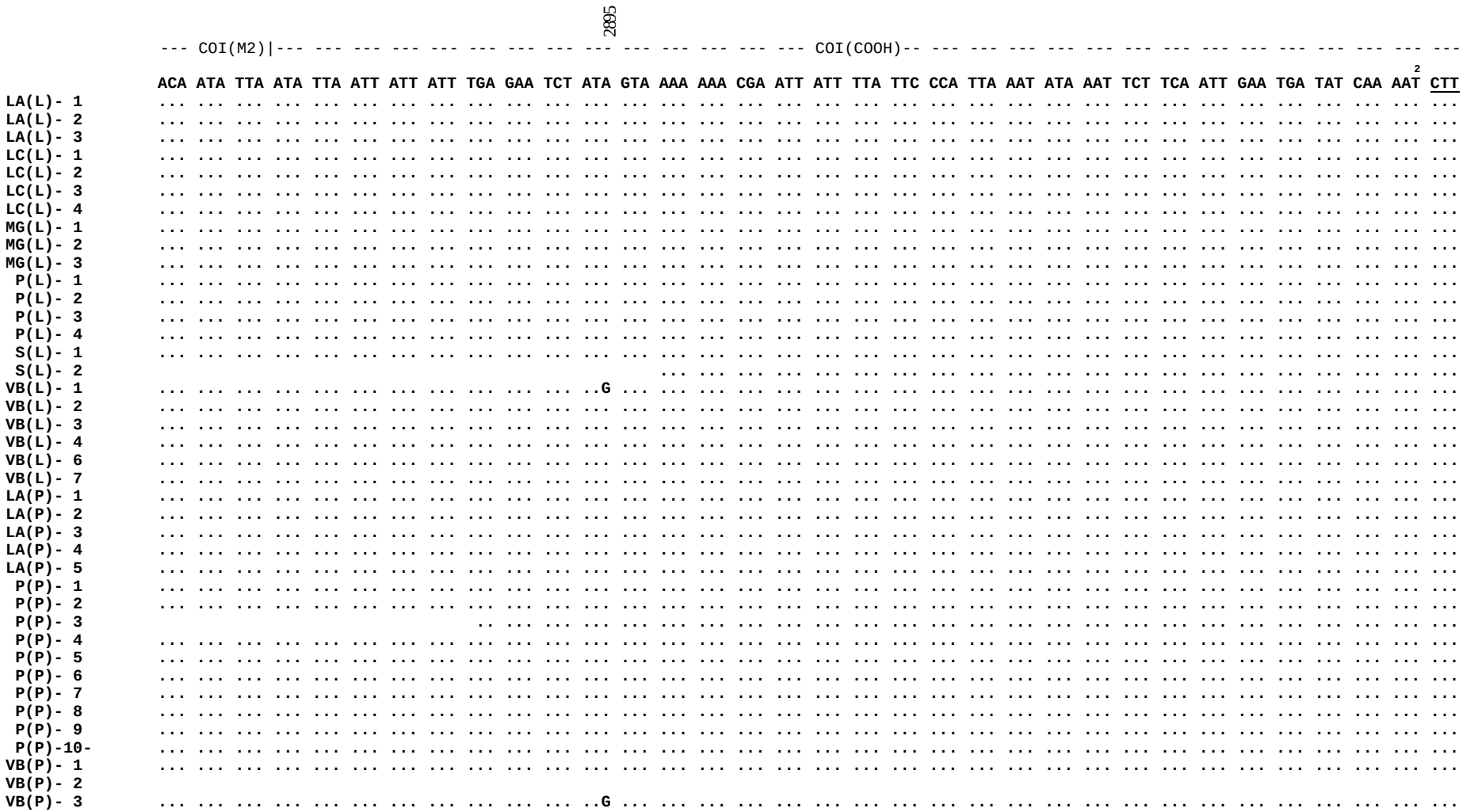

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

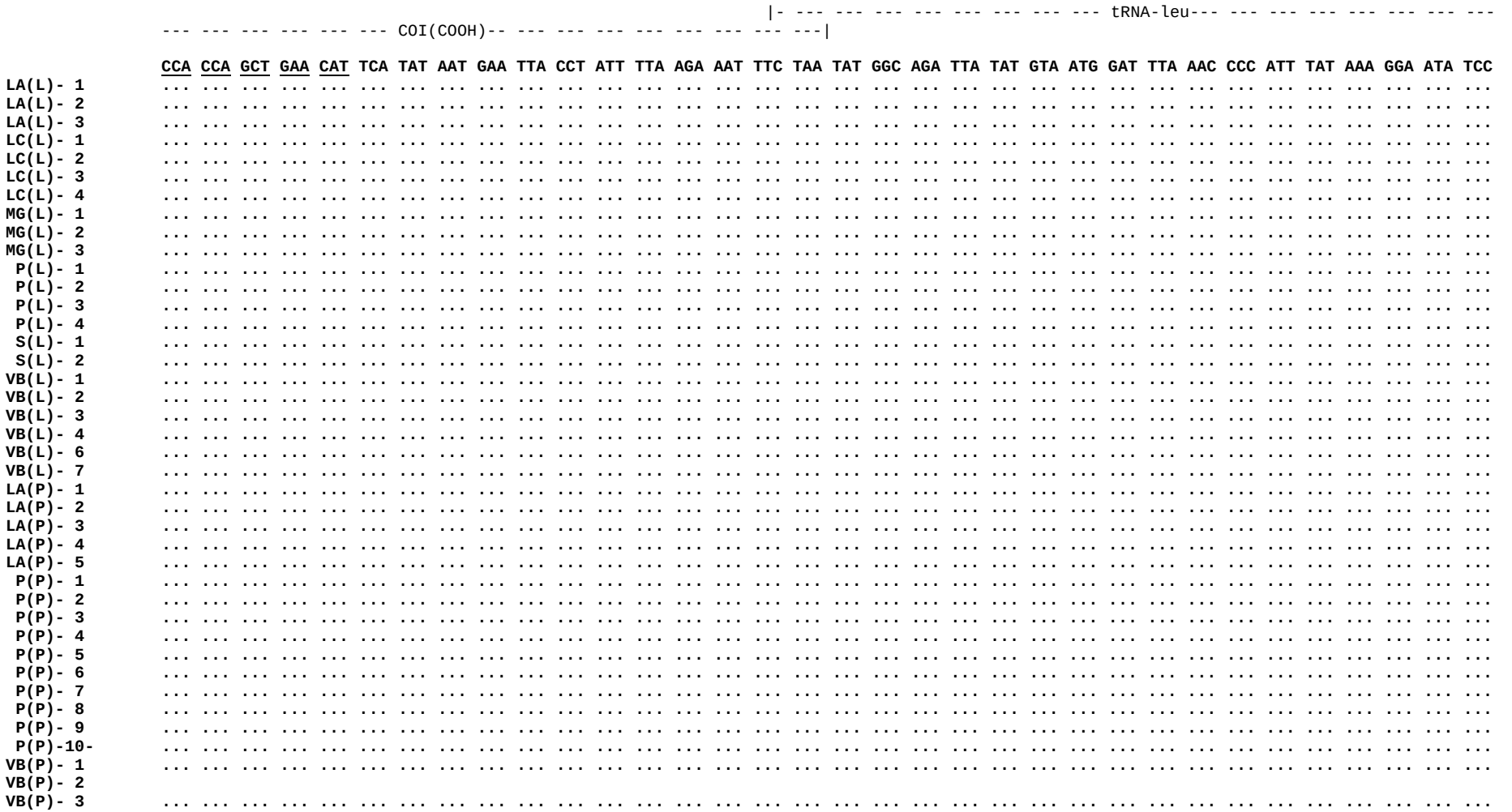

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

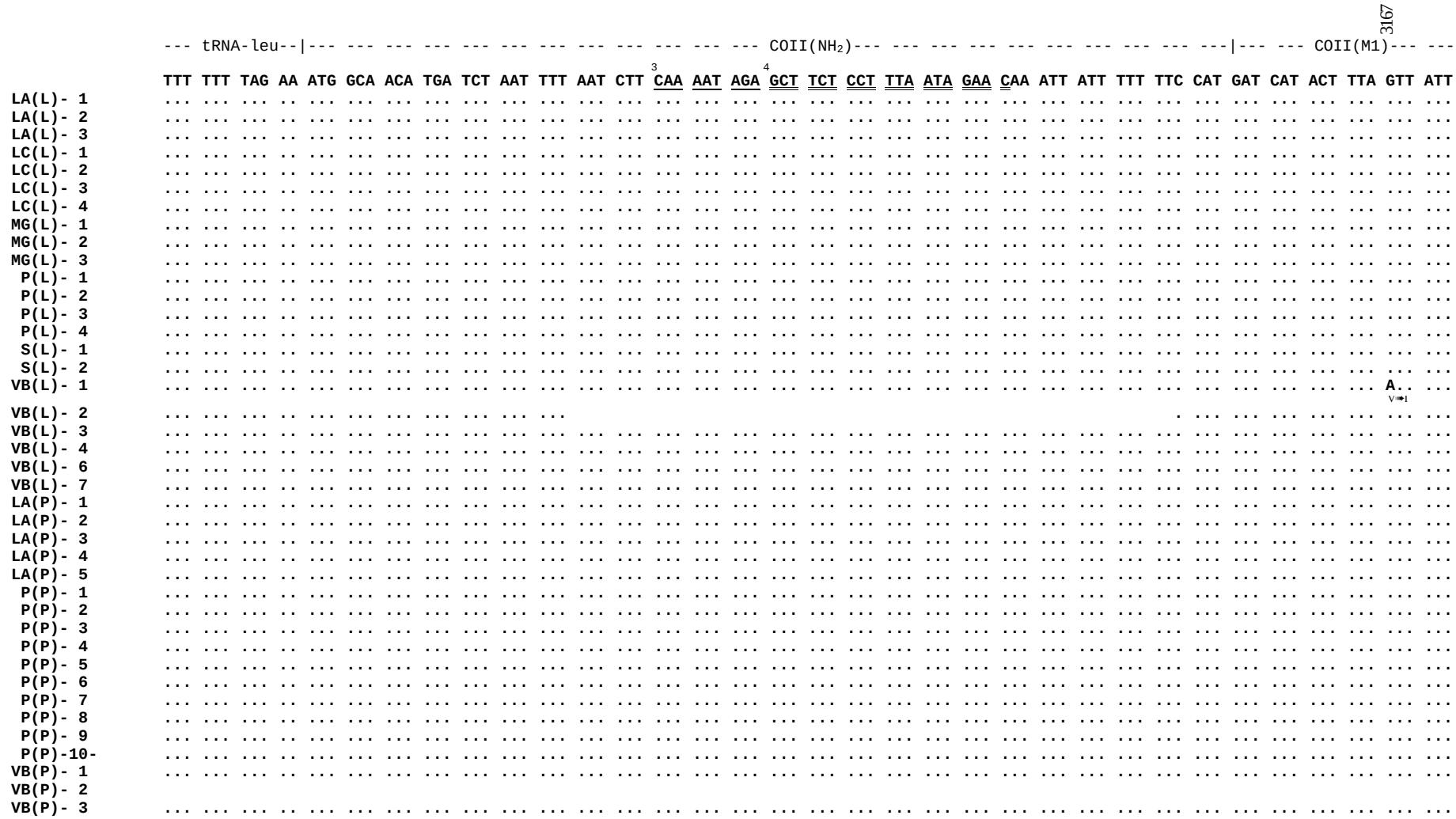

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

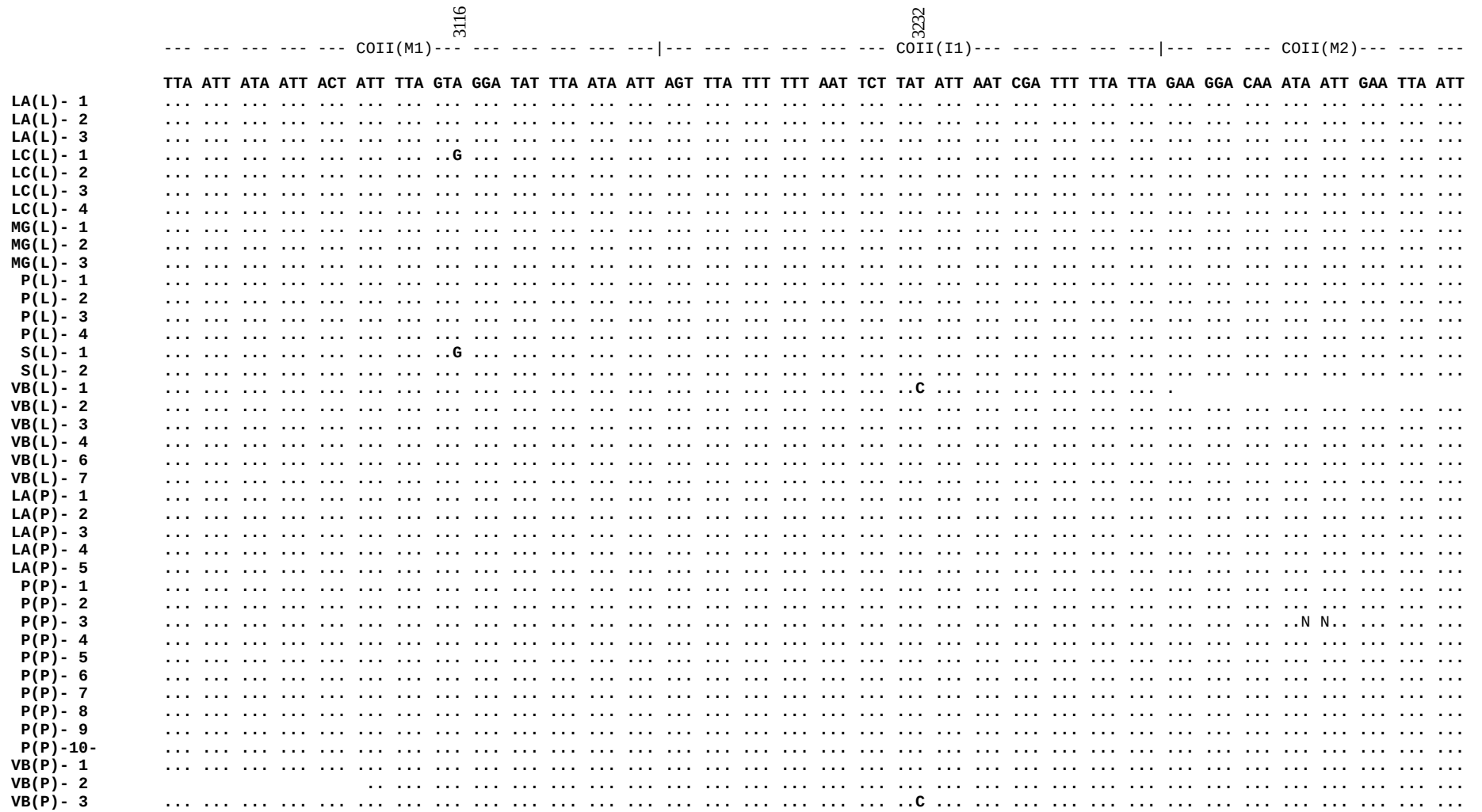

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

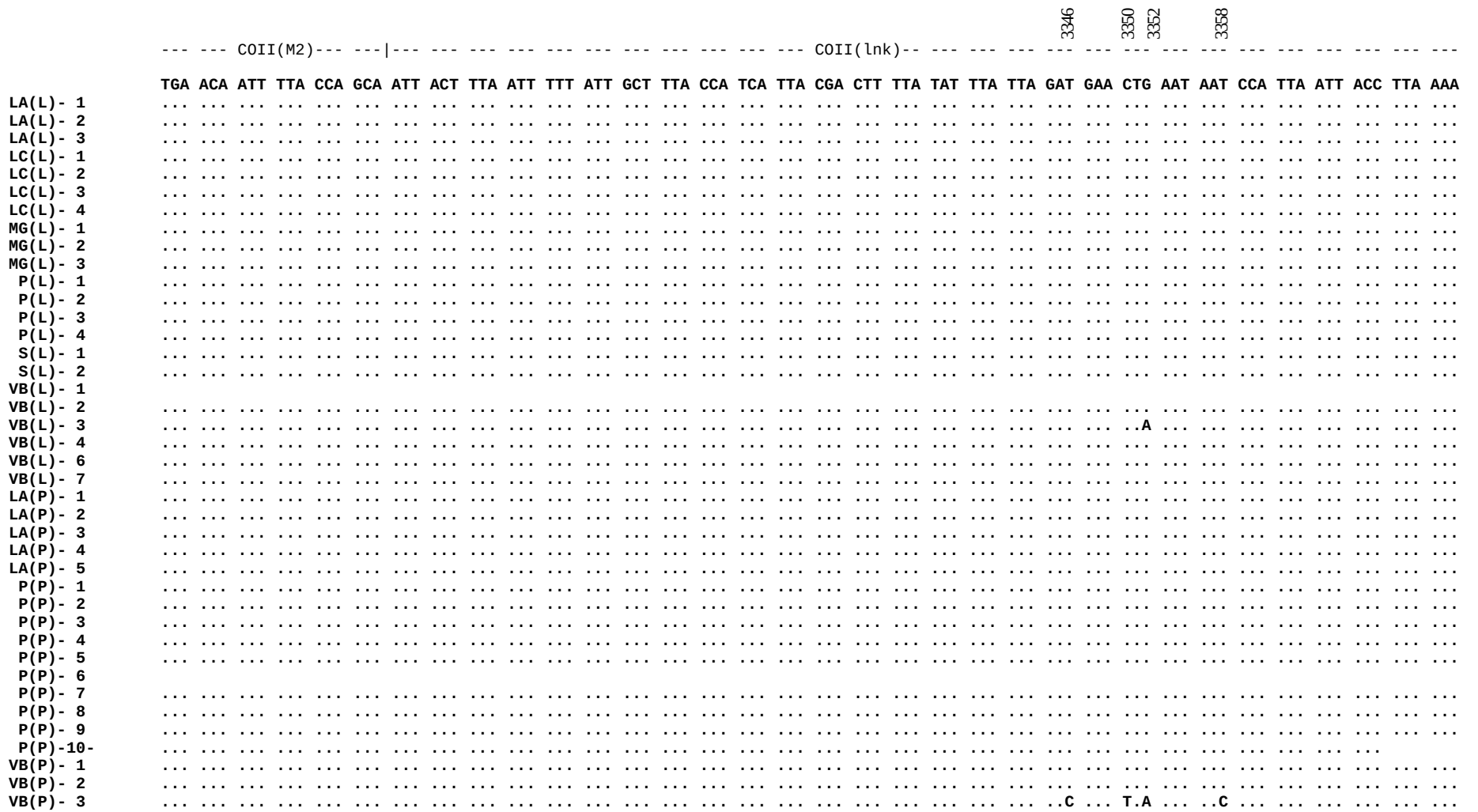

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

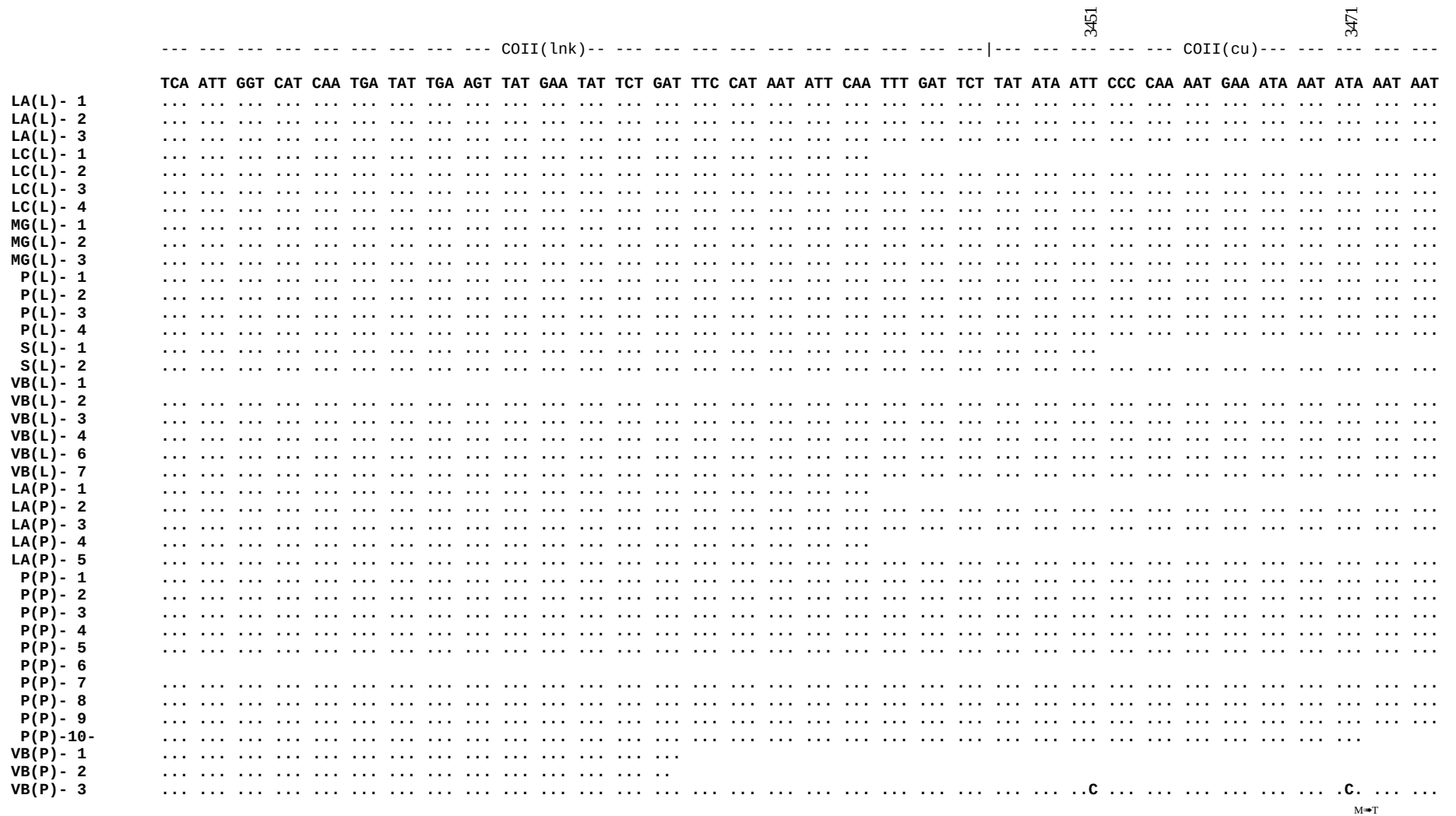

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

```
                --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- COII(cu)--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
                TTT CGT TTA TTG GAT GTT GAT AAT CGA ATT GTA ATC CCT ATA AAT AAT CAA ATT CGA ATT ATA GTA ACA GCT ACA GAT GTT ATT CAT TCT TGA ACA GTA
LA(L)- 1
LA(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LA(L)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LC(L)- 1
LC(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LC(L)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LC(L)- 4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
MG(L)- 1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..
MG(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
MG(L)- 3        ... ... ...
 P(L)- 1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(L)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(L)- 4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..
 S(L)- 1
 S(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..
 VB-L- 1
VB(L)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
VB(L)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
VB(L)- 4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
VB(L)- 6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
VB(L)- 7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LA(P)- 1
LA(P)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LA(P)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
LA(P)- 4
LA(P)- 5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)- 1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..
 P(P)- 2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)- 3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)- 4
 P(P)- 5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)- 6
 P(P)- 7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..
 P(P)- 8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)- 9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 P(P)-10        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
VB(P)- 1
VB(P)- 2
VB(P)- 3        ... ... .
```

Figure 3. Alignment of *Zeiraphera diniana* mtDNA sequences (cont'd).

Variable sites within *Zeiraphera diniana* mtDNA													
Sequence	Position relative to the *Drosophila yakuba* reference sequence (Clarey and Wolstenhome 1985)												
	2814	2832	2844	2895	3116	3167	3232	3346	3350	3352	3358	3451	3471
Consensus	A	T	C	A	G	**A**	T	T	C	G	T	T	T
LA(L)- 1	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(L)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(L)- 3	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)- 1	.	.	.	.	.	**G**	.	.	.	.	.		
LC(L)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)- 3	.	.	.	.	.	.	.	.	.	.	.	.	.
LC(L)- 4	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)- 1	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
MG(L)- 3	.	**C**	.	.	.	.	.	.	.	.	.	.	.
P(L)- 1	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)- 3	.	.	.	.	.	.	.	.	.	.	.	.	.
P(L)- 4	.	.	.	.	.	.	.	.	.	.	.	.	.
S(L)- 1	.	.	.	.	.	**G**	.	.	.	.	.	.	
S(L)- 2					.	.	.	.	.	.	.	.	.
VB(L)- 1	**C**	.	**T**	**G**	**A**	.	**C**						
VB(L)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(L)- 3	.	.	.	.	.	.	.	.	.	A	.	.	.
VB(L)- 4	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(L)- 6		.	.	.	.	.	.	.	.	.	.	.	.
VB(L)- 7	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)- 1	.	.	.	.	.	.	.	.	.	.	.		
LA(P)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)- 3	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)- 4	.	.	.	.	.	.	.	.	.	.	.	.	.
LA(P)- 5	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 1	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 2	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 3	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 4	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 5	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 6	.	.	.	.	.	.	.						
P(P)- 7		.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 8				.	.	.	.	.	.	.	.	.	.
P(P)- 9	.	.	.	.	.	.	.	.	.	.	.	.	.
P(P)- 10	.	.	.	.	.	.	.	.	.	.	.	.	.
VB(P)- 1	.	.	.	.	.	.	.	.	.	.	.		
VB(P)- 2						.	.	.	.	.	.		
VB(P)- 3	C	.	**T**	**G**	A	.	**C**	**C**	**T**	**A**	**C**	**C**	**C**

Table 1. Variable sites within *Zeiraphera diniana* mtDNA. Dots indicate identity of nucleotides to the consensus sequence presented on the top row. White spaces beneath consensus nucleotides indicate a lack of sequence information.

Chapter 5

Conclusions

The aim of this project was to clarify the host race status of larch and pine associated *Zeiraphera diniana* populations, via studies of actual gene flow, hybrid fitness, and mitochondrial DNA. Populations on the two hosts differ consistently by a number of genetically determined traits (reviewed Chapters 2-4, and in Emelianov et al. 1995). Previous studies also suggest that host race status is likely: the two forms can easily be hybridised and backcrossed in the laboratory with no sex ratio distortion (Baltensweiler, 1977; Emelianov et al., 1995), co-exist at a number of sites, and have a pattern of allozyme differentiation (considerably, but incompletely differentiated at three of 13 loci, and undifferentiated at the remaining 10, Emelianov et al., 1995) indirectly suggestive of gene flow. Genotypic clusters that are host associated and undergo appreciable actual gene flow can be considered host races (Chapter 1).

Because current (often contradictory) host race definitions are either ambigous or impractical in an empirical context, a new, experimentally testable set of criteria for host race status has been developed in Chapter 1. I concluded that actual gene flow should be considered appreciable if it is at least 1% per generation, an order of magnitude greater than that observed between taxa that hybridise but are nonetheless widely considered to be 'good' species (Grant and Grant, 1992; Mallet et al., 1998; Wang et al., 1997). Under the new definition, it must also be shown that actual gene flow (movement of genes between populations via hybridisation, leads to gene flow in the sense of the incorporation of genes from one population into another. Practically, this means that backcrossing in nature should be demonstrated.

Combined estimates of long range, pheromone-mediated cross attraction in the field (Appendix), and hybridisation in mixed populations at close quarters (Chapter 2) show that the level of actual gene flow between the larch and pine races is approximately 2.4%

each generation (and suggest that this is mainly from the larch-feeding population to the pine). Because the extent to which calling females migrate between hosts is an important determinant of actual gene flow (Appendix), and yet (because host fidelity of calling females is not yet known) it was assumed that only males migrate, the estimate of 2.4% gene flow is a cautious one that may well underestimate the total, and overestimate the directional bias. However, preliminary results suggest that female host fidelity is strong (Emelianov, pers. com.). A more precise, higher estimate of gene flow incorporating the effects of female migration would not affect the host race status of the two biotypes, as the figure is already appreciable. However, it would be desirable in its own right, and for comparison with levels observed in other host race systems.

Field measures of cross attraction at sites other than the Engadin Valley, Switzerland (Figure 1, Chapter 2) would allow us to test whether the figure obtained in the Appendix is representative of levels across the shared range of the two biotypes. The extreme variability of larch race population density, for example, may affect gene flow from year to year and site to site. The percentage of greenish-yellow pine-like larvae on larch trees increases as larch form population density decreases (Baltensweiler, 1993), and although this could be solely due to larval migration, it might also be a result of hybridisation . Numerous other factors likely to vary between places and years could also influence actual gene flow between the races (Itami et al., 1998).

Laboratory tests of intrinsic fitness suggest that hybrids do not suffer any developmental difficulties: similar proportions of F_1 hybrid and parental broods hatched and survived to the final larval instar, and there was no evidence of sex ratio distortion (Chapter 3, Baltensweiler, 1977). Hybrid females, at least, are able to successfully compete with larch race females for matings with larch males: three out of the five hybrid vs. larch mate choice experiments resulted in a mating of this type (Chapter 2). It may be that the apparently high frequency of this cross is in part due to a higher mating propensity of hybrid vs. larch females, since this trait is much stronger in pine females than in females of the opposite biotype.

F_1 hybrids are fertile, and backcross (F_1 x larch) broods failed to display any developmental difficulties or sex ratio distortion (Chapter 3 and Baltensweiler, 1977). Furthermore, the single backcross (larch female x F_1 hybrid male) vs. larch race mate choice test resulted in a backcross female x larch male mating (Chapter 2). Thus, there do not appear to be strong developmental or (mating) behavioural obstacles to

backcrossing. The lack of mitochondrial DNA differentiation between the forms provides additional, indirect support for gene flow.

However, several questions remain to be answered. The fate of hybrid forms in the wild is not known. Those laid as eggs on larch seem unlikely to suffer nutritional problems (at least when population density is moderate), given the suitability of this host as a food source for all larval types in the lab, but will have the dominant, greenish-yellow fifth instar larval colour, which may render them conspicuous to predators against the dark larch cones and bark. Likewise, those feeding on pine will have the corrrect coloration, but might experience difficulty in exploiting this food source, since pine larvae do not thrive on larch. A lack of information about the ability of hybrids to coordinate their development and emergence with host life cycles confuses the issue further. Some selection against hybrid forms is expected, given that disruptive selection is almost certainly acting to maintain differences between the larch and pine races. Their attractiveness to, and ability to attract and locate non-hybrid mates via pheromones also remains to be tested. And of course, additional mate choice experiments and measures of developmental fitness, particularly those involving pine race moths and their hybrids would be welcome.

Nonetheless, the results of this project combine with those of previous studies to support a host race designation for sympatric larch and pine races of *Zeiraphera diniana*: the populations do seem to meet all criteria of the host race definition set out in Chapter 1.

These newly identified host races represent only a tiny fraction of the diversity present in sympatric, plant-feeding insects specialized on different hosts. Nevertheless, the study of this system, and the few dozen others known to share (or be likely to share) the capacity for genetic differentiation in the face of gene flow may shed light on the origins of this diversity. Although sympatric speciation is extremely difficult to demonstrate directly, host races can, and do, provide indirect evidence that sympatric speciation by host shift is common in phytophagous insects. The existence of a continuum of unimodal hybrid zones dominated by intermediates, bimodal hybrid zones where intermediate genotypes are rare, and contact zones between taxa with no intermediates suggests that parapatric speciation is likely. Although any pair of taxa may have evolved in allopatry, the continuum between hybrid zones and species contact zones implies the existence of a stable route for purely parapatric evolution (Jiggins and Mallet, 2000). Similarly, host races are a stepping stone between host-associated polymorphisms and host-associated

species: their stable existence suggests that an evolutionary route for sympatric speciation exists. It is therefore desirable to study more examples of host races, particularly with regard to their levels of gene flow, in order to document whether there is a ‘break’ in the continuum, which might indicate that most phytophagous insect species evolved in allopatry. Otherwise, sympatric speciation seems likely.

References

Baltensweiler, W. 1977. Colour-polymorphism and dynamics of larch bud moth populations (*Zeiraphera diniana* Gn., Lep. Tortricidae). *Mitteil. Schwiz. Entomol. Gesellsh.*, 50:15-23.

Baltensweiler, W. 1993. A contribution to the explaination of the larch budmoth cycle, the polymorphic fitness hypothesis. *Oecologia*, 93:251-255.

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the larch budmoth) - polymorphism, host races or sibling species. *Heredity*, 75:416-424.

Grant, P. R., and B. R. Grant. 1992. Hybridization of bird species. *Science*, 256:193-197.

Itami, J., T. Craig, and J. Horner. 1998. Factors affecting gene flow between the host races of *Eurosta solidaginis*. In S. Mopper and S. Strauss (eds.), *Genetic structure and local adaptation in natural insect populations*, pp. 375-404. Chapman and Hall, New York.

Jiggins, C. D., and J. Mallet. 2000. Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution*, 15:250-255.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution*, 10:294-299.

Mallet, J., W. O. McMillan, and C. D. Jiggins. 1998. Estimating the mating behaviour of a pair of hybridising *Heliconius* species in the wild. *Evolution*, 52: 503-510.

Mallet, J. in press. Gene Flow. In I. P. Woiwood, D. R. Reynolds and C. D. Thomas (eds.), *Insect Movement: Mechanisms and Consequences*. CAB International, Wallingford.

Wang, R. L., J. Wakely, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close realatives. *Genetics*, 147:1091-1106.

Appendix

Host plant effects on assortative pheromone attraction by host races of the larch budmoth

Igor Emelianov[1]
Michele Drès[1]
Werner Baltensweiler[2]
James Mallet[1]

1. Galton Laboratory, Department of Biology
University College London, 4 Stephenson Way
LONDON NW1 2HE, ENGLAND
Tel: 0171-504-5100
Fax: 0171-383-2048
Email: i.emelianov@ucl.ac.uk

2. Blumenberg Strasse 9,
CH-8634 HOMBRECHTIKON, SWITZERLAND

Abstract

Assortative mating may evolve as a by-product of adaptation to different habitats despite some gene flow between diverging forms. Disruptive selection for ecological adaptation may therefore produce sympatric speciation via pleiotropy, but this idea has been deemed unlikely and virtually restricted to parasites that use the host itself as a mate-finding cue. The larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae) has sympatric larch- and pine-feeding host races which, as adults, can be identified only via molecular markers. We here show strong assortative mate attraction in the field in this species, due mainly to chemical differences in female long-range pheromones. Cross-attraction between host races occurs at a rate of 3-38%. The assortativeness of pheromone attraction (measured by numbers of attracted males belonging to the race of the calling female) depends strongly on the host substrate (larch or pine) from which females call, and also on host neighbourhood structure (proportion of larch and pine trees within 8m). Thus host choice, coupled with clumped host tree distribution creates additional assortative mating as a pleiotropic effect, even in a pair of host races that lack the requirement for mating on their own hosts. To our knowledge, this is the first time that mate location has been shown to be affected by host factors in a species with long-range mate attraction. Host shifts will often have pleiotropic effects on assortative mating such as those observed here, because host choice by both males and females and clumped host distributions are common in host-specific parasites.

Introduction

The evolution of assortative mating is arguably the most important component of speciation. In sexual populations, disruptive selection on habitat choice may cause a correlation between habitat choice and assortative mating. However, in sympatry, recombination will break up associations between loci affecting survival and those determining mate choice (Felsenstein 1981). If, on the other hand, assortative mating – or the mating microhabitat, which amounts to the same thing – is controlled by the same genes that are under disruptive selection, the antagonism between recombination and selection is reduced (Slatkin 1982), and subsequent adaptive evolution of assortative mating ("reinforcement") becomes much more likely (Rice 1984; Kelly and Noor 1996). Mate choice then becomes a pleiotropic by-product of habitat choice. This kind of pleiotropy occurs in parasites that mate only on their hosts, as in *Rhagoletis pomonella* flies (Feder et al. 1994) or *Enchenopa* treehoppers (Wood and Keese 1990).

However, pleiotropy between habitat choice and mate choice is considered unlikely in most other organisms (Tregenza and Butlin 1999). Mate location in phytophagous insects is often via long-range signalling, particularly using pheromones, rather than depending on host cues. In Lepidoptera, all major groups except butterflies use long-range pheromones, so that the food plant is not required for mating. It is known that host plant chemistry may indirectly affect both the production of and the response to pheromones (McNeil and Delisle 1989; Landolt and Phillips 1997), but no previous studies have investigated whether host plants can effect assortative mating between hybridising species or host races.

Sympatric host races of the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae) feed either on larch (*Larix decidua*) or cembran pine (*Pinus cembra*). Larch and pine races also differ in the blends of female sex pheromone (see below) and in other respects such as larval colour pattern and the allozymes *Mdh-s*, *Idh-s* and *Pgm* (reviewed by (Emelianov et al. 1995)). Females lay eggs and both sexes rest preferentially on the appropriate host tree (Bovey and Maksymov 1959; Emelianov et

al. 1995). However, none of these differences are completely diagnostic; with overlapping flight periods there is plenty of opportunity for gene exchange between the host races, and intermediates occur in natural populations (Priesner and Baltensweiler 1987). The larch race is well known for regular outbreaks which cause conspicuous defoliation of larch forests in the Alps. The population density on larch oscillates in 8-10 year cycles; at a peak, the populations may be 10^5-fold higher than during a crash (Baltensweiler 1993). In contrast, populations on pine appear more stable (Baltensweiler et al. 1977), although few data exist for this race.

Mating behaviour in the larch budmoth can be divided into two stages. In the first, long-range stage males choose and follow plumes of pheromone released by "calling" females. The female cannot choose at this stage. Both races produce two pheromone components, trans-(E)9-dodecenyl acetate (E9-12:Ac) and trans-(E)11-tetradecenyl acetate (E11-14:Ac) (Baltensweiler et al. 1978), but in opposite ratios. Most larch race females produce only E11-14:Ac, although some individuals also produce traces of E9-12:Ac. Pine race females produce both E9-12:Ac and Ell-14:Ac in the ratio of 1000:1. Males respond maximally to pheromone blends of their own race (Priesner and Baltensweiler 1987).

Males that have successfully located a signalling female enter the second, short-range stage, which may involve both male and female choice (Drès et al. 2000, in prep.). Here we experimentally investigate the first stage: long-range attraction of males by larch and pine females calling from different host tree species, and in local host neighbourhoods differing in the proportion of larch and pine trees. We ask the following questions:

-Is long-range pheromone attraction assortative between host races in natural mixed forest?
-Does the calling substrate (the host tree from which females call) affect assortative attraction?
-Does the host neighbourhood structure affect assortative attraction?
-Do host races differ significantly in diurnal timing of pheromone communication, and does this contribute to allochronic isolation and assortative attraction?

Methods

Study site

Experiments on long-range mate attraction were performed during summer 1997 on the south-facing slope of Val Bever (46°26'N, 9°50'E, alt. 1800 m), a lateral valley of the Upper Engadin, eastern Switzerland. The site has mixed larch and cembran pine forest.

Collections

Larvae were collected from larch and pine in Val Bever in 1997, the year of the experiment. Adult moths were reared and divided into two groups. The first group (40 individuals of each race) was analysed electrophoretically using procedures already described (Emelianov et al. 1995) to determine local allele frequencies for allozyme markers. Females of the second group were used in the long-range attraction experiments.

Long-range mate choice experiment and diurnal timing of pheromone attraction

A total of 39 larch and 42 pine virgin females were used to lure wild males. Each female was placed singly in a 30 cm^3 metal mesh cage, and was regularly supplied with drinking water and 5% sucrose solution. A twig of larch or pine was placed in each cage. These cages were placed individually into standard "delta" pheromone traps furnished with sticky cardboard inserts to trap males. Traps were positioned on larch or pine trees 15-20m apart, 1.5 – 2 m above the ground, along two 700 m transects aligned in parallel across the slope and separated by approximately 40 meters (Fig. 1). Four host race - host substrate combinations were used, each with approximately 20 females: (1) Larch females supplied with larch twigs in traps positioned on larch trees (LL); (2) larch females with pine twigs on pine trees (LP); (3) pine females with larch twigs on larch trees (PL); (4) pine females with pine twigs on pine trees (PP). The numbers of larch and pine trees were counted within an 8 m radius around each trap, and the fraction of pine trees was used as an index of host

neighbourhood structure (Fig. 1). The four treatments were randomly distributed along the transect. The sticky inserts were checked, and trapped males were counted at 16:00, 20:00 and 08:00 hours daily from 18 - 26 August. On 25 and 26 August the traps were additionally checked at 01:00. Every morning, inserts together with moths were removed from the traps and replaced with fresh inserts; the trapped males were frozen in liquid N_2 and shipped to the laboratory for allozyme-based host race identification. Four traps not containing females were also positioned along the transects: none of these trapped any moths, so these negative controls are not mentioned further.

Permutation test of the relationship between host neighbourhood and males attracted

We were interested in testing for correlations between the independent variable (neighbourhood host composition) and the dependent variables (numbers of larch and pine males attracted) for each host race – host substrate combination. However, each trap cannot strictly be used as an independent sample because of the possibility of spatial autocorrelation. Neighbouring trapping sites will be independent neither in terms of forest composition, which may have a spatial scale greater than 16 m (clearly visible in Fig. 1), nor in terms of moth catches, because the spatial scale of male moth density will be determined by dispersal and may also exceed the distance between neigbouring traps. The standard *Z*-transformation of the correlation coefficient may therefore give inflated significance levels. Instead, we test for this effect using a permutation test (similar to a bootstrap) to randomise the dependent variables relative to the independent variable while keeping the spatial structure of the data as constant as possible. The two transects were treated together as a 2 x 42 matrix, with each of 42 columns representing a pair of traps, one on each transect. (The lower transect in fact consists of only 39 traps; see Fig. 1. To make the data more manageable for permutation three traps with missing data were added: one before the start of the transect, and two after the end. The tree data, which are highly spatially autocorrelated (Fig. 1), were assumed to be for trap site 1 the same as the adjacent site 2, and for sites 41 and 42 to be the same as site 40; the female / tree types and numbers of males caught in these sites was assumed missing for the purpose of permutation). Product-moment correlation coefficients calculated from the actual data were compared with those obtained from data permuted while preserving

spatial relationships. The dependent variables (males trapped, female type, and calling substrate) were permuted relative to the independent variable (the tree data, which was not changed) as follows: (1) male data exchanged between transects (vertical flip); (2) male data inverted along the transects (horizontal flip); and (3) male data moved laterally along the transects by a variable number of columns (0-41), with data running off the end of the transect being added in sequence to the other end. This gives a total of 2 x 2 x 42 = 168 possible permutations, that is 167 plus the actual observations.

Trapping with synthetic pheromones

Additional delta traps baited with either E11-14:Ac or E9-12:Ac were positioned on larch and pine trees well away from the main transect to check for assortative attraction to synthetic pheromone. Three treatments (5 traps each) were performed: E11-14:Ac on larch trees, E9-12:Ac on larch trees, and E9-12:Ac on pine trees. E11-14:Ac traps were not placed on pine. The sticky inserts were checked, and trapped males were counted at 08:00, 16:00 and 20:00 daily from 18 - 26 August. We did not determine the race of males trapped in this experiment.

Host race identification using a set of semi-diagnostic allozyme markers
Cellulose acetate allozyme electrophoresis of trapped males was performed for two unlinked autosomal loci (*Mdh-s* and *Pgm*) and a sex-linked locus (*Idh-s*) according to methods given in Emelianov et al. (1995). Because there are no fixed differences, so any attracted male genotype at these loci may with finite likelihood belong to either host race. However, allozyme frequency differences allow the relative likelihood that each tree-locus genotype belongs to one of two races to be easily assessed. We determined the host race of each individual genotype as the race where expected frequency of this genotype is the highest. Consequently, the frequency of the same genotype in the less likely race is the error of identification for this genotype. Given Hardy-Weinberg and linkage equilibrium within host races, as observed in field samples (Emelianov et al. 1995), the expected frequency of each three-locus genotype in larch and pine populations can be poredicted using allelic frequencies from the reared larvae. The total expected rate of incorrect host race identification based on

these frequencies is 0.10% for the larch race and 0.39% for the pine race, an order of magnitude lower than the cross-attraction rates observed here.

Results

Allele frequencies in the larval populations.

Allele frequencies at *Mdh-s, Idh-s* and *Pgm* from samples of larvae collected in 1997 at Val Bever are shown in Table 1. Frequencies on larch and pine were similar to but significantly different from those in French and Swiss populations sampled in 1994 (Emelianov et al. 1995). Allele frequencies obtained from the same generation and site were therefore necessary in order to identify host race of adults reliably.

Effect of calling substrate

Larch females calling from pine attracted significantly fewer males of their own race (G_1=25.38, P<0.001) and overall (G_1=13.98, P<0.001) than larch females calling from larch (Table 2). Similarly, pine females on larch trees were significantly less attractive to pine males (G_1=6.70, P<0.01) and to males in general (G_1=8.62, P<0.01) than pine females on pine. The specificity of pheromone attraction by females positioned on their own hosts was high (Table 2). Cross-attraction of host races was only 3.3% and 9.1% for larch females on larch and pine females on pine, respectively. However, assortative pheromone attraction depended strongly on the host species from which the female called. Larch females calling from pine cross-attracted 37.7% alien pine males, significantly more than on pine (3.3%; G_1=34.45, P<0.001). In contrast, cross-attraction by pine females on larch (6.3%) was similar to their cross-attraction on pine (9.1%; G_1=1.42, 0.10<P<0.50).

Effect of host neighbourhood structure

For larch trees, the variance divided by the mean of numbers per 8m-radius neighbourhood (s^2 / m) was 5.6, while that of pine trees was 2.0. Both are greater than

1.0, the expectation for randomly (Poisson-) distributed trees among trapping sites, indicating that distribution of pine trees and larch trees is clumped. This clumped distribution is likely to be due to some degree of mutual exclusion of hosts, so that a greater abundance of larch coincides with reduced abundance of pine, and vice-versa. This is confirmed by a negative correlation between the local numbers of larch and pine trees (r = -0.26, n = 81, P < 0.05).

There was a significant interaction between attractiveness of females and the host neighbourhood (Fig. 2). Using the permutation test, there was a significantly negative correlation between numbers of larch males attracted by larch females calling from larch trees and the proportion of pine in the neighbourhood (r = -0.65, P = 1/168, i.e. < 0.01, Fig 2a) and a less convincing negative correlation (r = -0.35, P=15/168 ≈ 0.089, Fig. 2b) for larch females calling from pine. Similarly, there was a significantly positive correlation between the neighbourhood fraction of pine trees and the numbers of pine males attracted by pine females from pine trees (r = 0.56, P=2/168 ≈ 0.012, Fig 2c) and from larch trees (r = 0.70, P = 1/168 i.e. < 0.01, Fig 2d). There was no significant effect of host neighbourhood on the number of alien males attracted (dashed lines in Fig. 2), but the power of this test is limited by low sample sizes. However, even assuming no effect on alien male attraction, it is clear that the significant own-male effect ensures the fraction of alien males increases with the fraction of alien host trees in the neighbourhood (Fig. 2). A way of testing this overall hypothesis of assortative mating depending on the host neighbourhood is to average the correlation coefficients across all pairs, taking account of whether the expectation is positive or negative, and then performing the permutation test on the average correlation. Thus the correlation coefficients can be averaged as follows: $r_{AV} = (- r_{LLxL} + r_{LLxP} - r_{LPxL} + r_{LPxP} - r_{PLxL} + r_{PLxP} - r_{PPxL} + r_{PPxP}) / 8$ (where r_{PLxP} refers to the correlation coefficient between the count of larch males, *..xL*, attracted to pine females, *P...* calling from larch trees, *.L..*, and fraction of pine trees in the neighbourhood). The value calculated from the actual data, r_{AV} = 0.325, was not obtained in the random permutations (P = 1/168, i.e. < 0.010); in the 167 random permutations, it varied between $-0.276 \leq r_{AV} \leq 0.305$.

Diurnal timing of pheromone attraction by live females

Pheromone activity in *Zeiraphera* occurred mainly between 20:00 and 08:00 (Fig. 3a), with male attraction peaked before 01:00 (Fig. 3b). Pine females on average attracted 0.60 (0.44-0.76) males per night between 20:00 and 01:00 and only 0.25 (0.13-0.38) males per night between 01:00 and 08:00 (figures in brackets are 95% confidence limits of the mean). Larch females showed a similar pattern of attraction; they each attracted 0.23 (0.13-0.34) and 0.14 (0.06-0.21) males per night before and after 01:00 respectively.

Males attracted by synthetic pheromones: effect of host and diurnal dynamics.

Males sensitive to different pheromones differed in their diurnal cycle of attraction. Males were attracted to pine race pheromone (E9-12:Ac) almost equally by day (08:00 – 16:00) and by night. In contrast, larch race pheromone (E11-14:Ac) was attractive to males almost exclusively at night (20:00 – 08:00). The host race of males was not identified in this study.

Traps baited with pine pheromone attracted on average 7.43 (5.73-9.15) males when positioned on pine trees and 5.33 (3.33-7.30) males per trap per night when positioned on larch trees ($P < 0.05$). Traps with larch pheromone on larch each attracted 4.35 (3.05-5.65) males per night.

Discussion

The effect of calling substrate and host neighbourhood on assortative mating

The long-range pheromone attraction in the larch budmoth was highly assortative. Females positioned on their own hosts cross-attracted only 3.3 – 9.1% "alien" males. This assortative pheromone attraction depended strongly on the calling substrate (at least for larch females, Table 2), and on the host neighbourhood structure (Fig. 2). This correlation between host and male attraction is also true for artificial pheromones: pine lures attracted fewer males when positioned on larch than on pine

(see Results). We have thus shown for for the first time that host plants are strongly involved in assortative mate attraction in a sibling pair of taxa differing in host choice.

These effects of host substrate and host neighbourhood on assortative attraction might be explained in two ways:

Firstly, the results could be caused by specific host preference in males and concentration of males on their own hosts. A female will therefore have more males of her own race locally available if she signals from her own host. Similarly, there will be more locally available males of her own race if the tree from which she calls, whether her own host or not, is surrounded by other individuals of her own host species.

Secondly, the interaction between host and assortative attraction could be a direct effect of host chemistry on pheromone production, pheromone release and/or male response. Such effects are known in other Lepidoptera (Landolt and Phillips 1997), although not in the context of assortative mating between host-specific sibling forms.

Males as well as females have a strong tendency to rest on their own host, both in the field and in laboratory host choice experiments (Emelianov et al., in prep). Thus the higher densities of males on their own hosts will explain at least some of the calling substrate and neighbourhood host composition effects. The inversely correlated clumped distributions of larch and pine trees will amplify this effect of male and female host resting preference on assortative attraction still further. Finally, some additional male concentration is also likely given female landing preference, so that pheromone is released mainly from the appropriate species of host.

The importance of a potential direct effect of host chemistry on attraction versus the indirect effect of local male and female density cannot be ascertained without additional experiments. Direct effects on mate attraction have been detected in a number of other phytophagous insects (Landolt and Phillips 1997) and cannot be ruled out for *Zeiraphera* at present.

Asymmetry in calling specificity

Cross-attraction by larch females on pine was significantly greater than cross-attraction by larch females on larch. In contrast, cross-attraction by pine females did not depend on calling substrate (Table 2). This asymmetry could be due to racial differences in the width of the males' pheromone response windows. Electroantennogram experiments with synthetic pheromones (Baltensweiler and Priesner 1988) have shown that pine males respond strongly to a wide range (100:1 to less than 1:1) of E-9-12:Ac : E-11-14:Ac blends. Further increases of the fraction of E-11-14:Ac reduced but did not block the response of pine males. In contrast, even small fractions of the pine race component added to larch race lures strongly reduced their attractiveness to larch males. Pheromone blends emitted by larch females therefore appear to be within the response window of pine males, but not vice-versa. This would lead to greater cross-attraction by larch females when on pine than vice-versa, as observed here.

The potential for hybridisation between host races

We here show that assortative attraction is strongly affected both by the tree species used as a calling substrate and by the host neighbourhood structure. Provided that assortative attraction results in assortative mating, the probability of hybridisation will be strongly affected by host choice. Mate preference between host races confined at close quarters is weak: in laboratory choice tests, an average of 28% of matings was between host races (Drès et al., in prep). This compares with 3.3% cross-attraction for larch females, and 9.1% cross-attraction for pine females, calling from their own hosts. The cross-attraction rates are 37.7% for larch females calling from pine, and 6.3% for pine females calling from larch; but since own-host resting preference of both females and males is strong (80-90%; Emelianov et al. in prep.), these rates are not important. Our results therefore suggest that long-range pheromone attraction is the major determinant of assortative mating.

The diurnal cycle of pheromone communication is very similar in the two races (Fig. 3). The strong difference in timing of male attraction to artificial pheromones (Fig. 4) is puzzling, but does not apparently lead to assortative mating, because of the lack of

difference in the timing of attraction of males to actual calling females (Fig. 3). Possibly, the similarity in times of attraction may be due to a lack of differentiation in female calling time in spite of differences in the timing of male receptivity. Alternatively, the attraction times to lures may be artifactual due to unnatural concentrations or missing minor components of the artificial pheromones. In any case, there is little evidence from the experiments with natural females for any differentiated time of mate attraction.

In our experiments, assortative mating was enhanced by an own-host substrate and own-host local plant density. These host effects are almost certainly due to the density of male moths tracking the density of their host. This suggests that the strong year-to-year fluctuations in the overall population density of the larch race should lead to a similar, temporal variation in assortative mating. Peak populations on larch may lead to an increased fraction of pine females from relatively stable populations mating with abundant larch males. Similarly, a crash in density of the larch form may lead to an increasing tendency for the rare larch females to be mated by the now more abundant males of the pine race. Thus the direction of gene flow could vary depending on the phase of the larch cycle. Periodic introgression of genes of pine race, with yellow-green larval colour (a dominant allele), into the larch race, whose mature larvae are black, may partially explain the documented synchrony between the population cycles and periodic increases of frequency of yellow-green phenotypes on larch (Baltensweiler 1993). We are currently investigating the frequency and periodicity of hybridisation in this species.

Pleiotropy and the evolution of assortative mating

In theory, pleiotropy between an environmental trait under disruptive selection and mate choice can facilitate the evolution of assortative mating (Rice 1984). It has always seemed plausible that habitat adaptation had a pleiotropic effect on assortative mating in organisms such as flies that must visit their larval resource to find mates. For example, habitat preference and assortative mating both evolved rapidly during the historically documented evolution of host races of the apple maggot *Rhagoletis pomonella* (Feder et al. 1994) and in laboratory selection

experiments on *Drosophila melanogaster* (Rice and Salt 1990); in both cases, mate preferences evolved in sympatry as a result of pleiotropy with habitat choice.

It has seemed less plausible that similar habitat-induced assortative mating should evolve in insects such as butterflies and moths that find mates using long-range signals (Thomas and Singer 1998; Tregenza and Butlin 1999). Because pheromone signals released from one host usually travel far enough to attract mates associated with other hosts, host-associated assortative mating would seem especially unlikely for night-flying moths. The distance at which female moths can effectively attract males may vary from 20m to 3.8km (Gotz 1951). However, this first test of host effects on assortative mating in a pair of taxa that use long-range sex pheromones has provided strong evidence for its existence in *Zeiraphera* and likelihood in other species, in spite of its seeming *a priori* improbability.

The evolution of assortative mating, host races and speciation

The host races of the larch budmoth already have highly divergent pheromone systems that account for most of the assortative attraction measured here, but this cannot have been the case initially. We here speculate on two possibilities for the evolution of assortative mating now found in the larch budmoth:

Firstly, it is possible that novel pheromone communication system in *Zeiraphera* initially emerged due to geographic variation of pheromone blend and response, which is known in some moths (Löfstedt 1993). Their divergence may be due to a Fisherian runaway process of sexual selection, either in sympatry (Higashi et al. 1999) or allopatry, given that direct selection on pheromone blends released by females is probably weak and that male pheromone response windows are often wide (Butlin and Trickett 1997). Once the runaway has started, polymorphic equilibria for different blends and responses are unlikely to be stable (Butlin and Trickett 1997); therefore divergent populations may rapidly reach fixation for alternate pheromone systems. Ecological displacement caused by host shift by one of these populations, or host specialization by both (if both hosts were used by the ancestral form) would have led to the current association between host choice and pheromone blend.

Secondly, it is possible that divergence was initiated by adaptation to different hosts. In this case the polymorphism in host choice would have had a pleiotropic effect on mate choice because of the concentration of males and females with similar host preference within the same host clumps, as found here. The resultant assortative mating can be maintained in sympatry only by disruptive selection on host preference; sympatric speciation is unlikely to be guaranteed at this stage. Divergence in pheromone communication could then occur by sexual selection, as above, or by means of reinforcement. A recent model shows that levels of gene flow must be less than the random value of 50% between sympatric forms for reinforcement to occur (Kelly and Noor 1996), as would occur if the host-related assortative mating found here were to operate. If pheromone divergence then began to occur, assortative mating could become more stable due to additional stabilising selection acting within each race against rare pheromone blends and rare male preferences (Butlin and Trickett 1997).

Conclusions

Regardless of the exact course of evolution, it is clear that the assortative pheromone attraction between host races found here depends both on the host species used as a calling substrate and on the neighbourhood host tree composition, as well as strongly divergent pheromone communication systems. In the larch budmoth, long-range pheromone attraction is the major discrimination phase between the sexes, and close-range assortative mating, while it occurs, is weak. Thus the correlation of male attraction with host choice will cause additional assortative mating, over and above that caused by the pheromone alone. The congregation of males and females on individual trees and in clumps of trees of their own host is probably the major mechanism for this pleiotropy; however, a direct host effect on pheromone production and response is also an untested possibility. Given that long-range pheromone communication is widespread in insects, as is the clumped distribution of host plants, it seems likely that pleiotropy between host and mate choice may be much commoner in insects such as Lepidoptera than previously thought. The

Lepidoptera, mostly nocturnal moths, form around 10-15% of the described animal species on this planet. Pleiotropy between host choice and mate choice may therefore be extremely general, and the sympatric evolution of host races and speciation may occur much more readily than hitherto realized.

Acknowledgements

We thank Svetlana Potiomkina and Yvonne Graneau for help with rearing and field experiments. We also thank Laurent Dormont for help with collections. This research was supported by grants from BBSRC and a NERC research studentship to MD.

Table 1. Allele frequencies in 1997 larval samples from Val Bever*

Locus	Allele	Frequency on Larch	Frequency on Pine
Idh-s		*n* ** = 78	*n* = 82
	0.65	0.410	0.024
	2.40	0.590	0.976
Mdh-s		*n* = 84	*n* = 96
	1.00	0.833	0.010
	5.50	0.167	0.990
Pgm		*n* = 86	*n* = 96
	0.87 & 1.00	0.791	0.854
	1.12 & 1.28	0.209	0.146

* Designation of loci and alleles is the same as in (Emelianov et al. 1995).
** *n* = number of analysed genomes

Table 2. Attraction of larch and pine males by live females

Host race / host tree combinations	Number of calling females	Number of attrcted larch males	Number of attracted pine males	Cross-attraction of alien males	Total number of attracted males
Larch females on larch trees	19	118	4	3.3 (0.76-5.8)%*	122
Larch females on pine trees	20	33	20	37.7 (24.7-50.8)%	53
Pine females on larch trees	22	12	184	6.3 (2.8-9.5)%	196
Pine females on pine trees	20	26	261	9.1 (5.7-12.4)%	287

* Figures in brackets show lower and upper 95% confidence limits

Figure Legends

Figure1. Schematic of the trapping experiment using live females. This diagram shows the transects, order of trap types, and fraction of pine trees within an 8m radius of each trap.

Figure 2. Effect of local host neighbourhood structure on attractiveness of females. Product-moment correlation coefficients are shown; statistical significance is assessed by means of a permutation test (see text).

Figure 3. Diurnal dynamics of pheromone communication.

a. Male attraction by live females, 18 – 26 August.

b. Male attraction by live females, 25 and 26 August only (includes 01:00 sampling time).

Figure 4. Diurnal dynamics of pheromone attraction. Male attraction by synthetic sex attractants.

References

Baltensweiler, W. 1993. A contribution to the explanation of the larch budmoth cycle, the polymorphic fitness hypothesis. Oecologia 93:251–255.

Baltensweiler, W., and E. Priesner. 1988. A study of pheromone polymorphism in *Zeiraphera diniana* Gn. (Lep., Tortricidae). 3. Specificity of attraction to synthetic pheromone sources by different male-response types from two host races. J. Appl. Entomol. 106:217–231.

Baltensweiler, W., G. Benz, P. Bovey, and P. Delucci. 1977. Dynamics of larch budmoth populations. Annu. Rev. Entomol. 22:79–100.

Baltensweiler, W., E. Priesner, H. Arn, and V. Delucci. 1978. Untershiedliche Sexuallockstoffe bei Lärchen- und Arvenform des Grauen Lärchenwicklers (*Zeiraphera diniana* Gn., Lep. Tortricidae). Mitt. Schweiz. Entomol. Gesellsch. 51:133–142.

Bovey, P., and J.K. Maksymov. 1959. Le problème des races biologiques chèz la tordeuse grise du mélèze *Zeiraphera griseana* (Hb.). Vierteljahresschrift der Naturforsch. Ges. Zurich 104:264–274.

Butlin, R.K., and A.J. Trickett (1997) Can population genetic simulations help to interpret pheromone evolution? In: Insect Pheromone Research. New Directions. R.T. Cardé, and A.K. Minks, eds. Chapman & Hall, New York, eds. 548–562.

Emelianov, I., J. Mallet, and W. Baltensweiler. 1995. Genetic differentiation in the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae): polymorphism, host races or sibling species? Heredity 75:416–424.

Feder, J.L., S.B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. Proc. Natl. Acad. Sci., U.S.A. 91:7990–7994.

Felsenstein, J. 1981. Scepticism towards Santa Rosalia, or why are there so few kinds of animals. Evolution 35:124–138.

Gotz. 1951. Die Sexualdufstoffe an Lepidopteren. Experientia 7:406–418.

Higashi, M., G. Takimoto, and N. Yamamura. 1999. Sympatric speciation by sexual selection. Nature 402:523–526.

Kelly, J.K., and M.A.F. Noor. 1996. Speciation by reinforcement: A model derived from studies of *Drosophila*. Genetics 143:1485–1497.

Landolt, P.J., and T.W. Phillips. 1997. Host plant influences on sex pheromone behavior of phytophagous insects. Annu. Rev. Entomol. 42:371–391.

Löfstedt, C. 1993. Moth pheromone genetics and evolution. Phil. Trans. R. Soc., London, Ser. B 340:167–177.

McNeil, J.N., and J. Delisle. 1989. Are host plants important in pheromone-mediated mating systems of Lepidoptera? Experientia 45:236–240.

Priesner, E., and W. Baltensweiler. 1987. A study of pheromone polymorphism in *Zeiraphera diniana* Gn. (Lep, Tortricidae). 1. Male pheromonal response types in European wild populations, 1978-85. J. Appl. Entomol. 104:234–256.

Rice, W.R. 1984. Disruptive selection on habitat preference and the evolution of reproductive isolation - a simulation study. Evolution 38:1251–1260.

Rice, W.R., and G.W. Salt. 1990. The evolution of reproductive isolation as a correlated character under sympatric conditions - experimental-evidence. Evolution 44:1140–1152.

Thomas, C.D., and M.C. Singer. 1998. Scale-dependent evolution of specialization in a checkerspot butterfly: from individuals to metapopulations and ecotypes. In: Genetic structure and local adaptation in natural insect populations. S. Mopper, and S.Y. Strauss, eds. Chapman and Hall, New York, eds. 343–74.

Slatkin, M. 1982. Pleiotropy and parapatric speciation. Evolution 36:263–270.

Tregenza, T., and R.K. Butlin. 1999. Speciation without isolation. Nature 400:311–312.

Wood, T.K., and M.C. Keese. 1990. Host-plant-induced assortative mating in *Enchenopa* treehoppers. Evolution 44:619–628.