Genetica **116:** 291–300, 2002.

Divergence and reproductive isolation in the early stages of speciation

Tom Tregenza
Ecology and Evolution Group, School of Biology, University of Leeds, Leeds, LS2 9JT, UK (E-mail: T.Tregenza@leeds.ac.uk)

Received 17 October 2001 Accepted 8 April 2002

Key words: Chorthippus parallelus, grasshopper, mating signals, postmating, premating, reproductive isolation

Abstract

To understand speciation we need to identify the factors causing divergence between natural populations. The traditional approach to gaining such insights has been to focus on a particular theory and ask whether observed patterns of reproductive isolation between populations or species are consistent with the hypothesis in question. However, such studies are few and they do not allow us to compare between hypotheses, so often we cannot determine the relative contribution to divergence of different potential factors. Here, I describe a study of patterns of phenotypic divergence and premating and postmating isolation between populations of the grasshopper *Chorthippus parallelus*. Information on the phylogeographic relationships of the populations means that *a priori* predictions from existing hypotheses for the evolution of reproductive isolation can be compared with observations. I assess the relative contributions to premating isolation, postmating isolation and phenotypic divergence of long periods of allopatry, adaptation to different environments and processes associated with colonisation (such as population bottlenecks). Likelihood analysis reveals that long periods of allopatry in glacial refugia are associated with postmating reproductive isolation, but not premating isolation, which is more strongly associated with colonisation. Neither premating nor postmating isolation is higher between populations differing in potential environmental selection pressures. There are only weak correlations between patterns of genetic divergence and phenotypic divergence and no correlation between premating and postmating isolation. This suggests that the potential of a taxon to exercise mate choice may affect the types of factor that promote speciation in that group. I discuss the advantages and disadvantages of the general approach of simultaneously testing competing hypotheses for the evolution of reproductive isolation.

Introduction

One of the longest running debates in evolutionary biology has been over the relative importance of different ecological and demographic factors in causing speciation. There remains plenty of room for speculation and opinion on this subject because there is a scarcity of hard evidence. This deficiency is a symptom of the fact that with the exception of a few processes confined to particular groups (such as chromosomal speciation in plants), speciation takes longer than the lifetime of a scientist, let alone a research grant.

The difficulty of observing speciation in nature has left biologists with three main approaches to understanding the evolution of reproductive isolation:

1. Laboratory experiments in which population sizes and selection regimes are artificially imposed, frequently with extreme values.
2. Observations of related species.
3. Observations of populations of a single species, and hybrid zones between populations.

These approaches have different advantages and disadvantages. Laboratory experiments (reviewed by Rice & Hostert, 1993; Odeen & Florin, 2000) are valuable in determining whether divergence can proceed under particular conditions, and have shown that reproductive isolation can develop rapidly between allopatric populations, and between parapatric or sympatric populations providing divergent selection is strong relative to gene flow. However, most involve radically different selection regimes between populations and very few can distinguish between divergence as a result of pleiotropic or linked effects of adaptation to differing environments and divergence due to random accumulation of mutations (Rice & Hostert, 1993). Also, artificially imposed selection and demographic regimes do not throw a great deal of light on the relative importance of these processes in nature.

Comparisons of related species, such as Coyne and Orr's study of *Drosophila* species pairs (Coyne & Orr, 1989, 1997) have shown that in sympatric species pairs postzygotic isolation accumulates slowly while premating isolation increases more rapidly and less predictably. Interspecific studies have the benefit that we know that observed differences between species include the divergence that prevents them exchanging genes. However, such studies have the disadvantage that we cannot distinguish between divergence contributing to speciation and divergence occurring afterwards.

Hybrid zones provide the opportunity to investigate pairs of populations, and have provided numerous insights into the nature of differences between populations and species (Hewitt, 1988). However, hybrid zones are typically the result of secondary contact between already divergent taxa, so they do not allow us to discover what caused the divergence in the first place. To be sure that we are observing patterns of divergence associated with speciation itself, we need to examine patterns of differences between multiple populations that are not completely reproductively isolated. This has been done for very few species. In the salamander *Desmognathus ochrophaeus* (Tilley, Verrell & Arnold, 1990) and the tungara frog *Physalaemus pustulosus* (Ryan, Rand & Weigt, 1996) divergence appears to be the result of gradual genetic divergence in allopatry since geographically distant populations are genetically more distant. In contrast, the rapid evolution of differences in mating signals between races of the brown planthopper *Nilaparvata lugens* (Butlin, 1996) and races of *Drosophila willistoni* (Gleason & Ritchie, 1998), and the lack a correlation between geographic and genetic distances in the copepod *Eurytemora affinis* (Lee, 2000) suggest either that population bottlenecks induced rapid genetic drift or that selection on traits contributing to reproductive isolation were responsible for divergence. In sticklebacks (Rundle et al., 2000) there is good evidence that natural selection has driven reproductive isolation between populations that have experienced different ecological conditions. This isolation is either due to pleiotropic effects of loci involved in environmental adaptation or to genetic hitch hiking of genes linked with these loci.

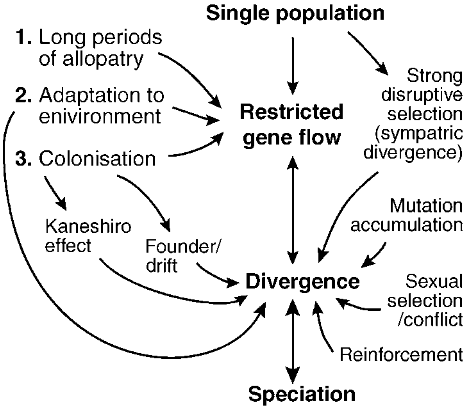

Figure 1. Factors likely to be important in the evolution of reproductive isolation between populations. There is broad agreement over the theoretical potential for all these factors to play a role in speciation, but little empirical data regarding their relative importance. I describe a study examining the relative importance of the factors labelled 1–3 in causing divergence between populations of the meadow grasshopper *C. parallelus*.

Overall, although there is broad agreement on the main factors likely to be important in causing speciation (Figure 1), their relative importance remains unclear. Here, I describe a project designed to examine the relative contribution to reproductive isolation and phenotypic divergence of three factors hypothesised to be important in speciation: (a) long periods of allopatry in refugia, (b) natural selection imposed by differing environments and (c) effects associated with colonisation including reduced population size and possible founder events. These factors are thought to affect population divergence for different reasons (Figure 1): Allopatry restricts gene flow, allowing mutations to accumulate in a random fashion over a long period of time, adaptation to differing environments can restrict gene flow, but also contributes directly to divergence, and processes associated with colonisation affect divergence through various effects including low population size and repeated founder events. There are other hypotheses for divergence not tested in this study. For instance, sexual selection might drive the evolution of reproductive isolation, but as a theory is difficult to test (Panhuis et al., 2001) since it makes few unique predictions.

Comparing the contribution of factors hypothesised to influence divergence using an insect model system

Chorthippus parallelus is a flightless gomphocerine grasshopper inhabiting meadows and smaller patches of grass throughout Europe and Asia. Its current distribution (like that of many European species) is the result of a northerly recolonisation following the last glaciation when European populations were confined to refugia in the Iberian Peninsula, Italy and the Balkans. The distribution of non-coding nuclear and mitochondrial DNA haplotypes (Cooper, Ibrahim & Hewitt, 1995; Lunt, Ibrahim & Hewitt, 1998) indicates that rather than there having been a general movement northwards, present-day populations north of the Mediterranean are descendants of the Balkan refuge (Figure 2). Populations with glacial refugia in Spain and Italy appear to have been confined to these regions following the end of the last ice age, presumably because the Pyrenean and Alpine mountain chains presented a barrier to dispersal. Because of this pattern of postglacial range expansion, modern populations have differing evolutionary histories. Those in Spain, Italy and Greece are descendants of local refugial populations, and have been isolated from one another for several hundreds of thousands of years (Hewitt, 1996; Lunt, Ibrahim & Hewitt, 1998). Populations in central and northern Europe have colonised that region in the last 10,000 years and are descendants of a single glacial refuge, therefore they have not experienced long periods of allopatry. However, they are likely to have experienced frequent population bottlenecks, since the rate of recolonisation (reaching Britain before the closure of the land bridge 7000 years ago) exceeds by orders of magnitude the typical dispersal distances of individuals in contemporary populations (Virdee & Hewitt, 1990), and because haplotypes are less variable than in refugial populations. Additionally, within the descendants of the Balkan refuge, there are populations at altitudes varying between sea level and over 1500 m. There are also populations sympatric with a closely related species, *C. montanus*, which has very similar mating signals, and often co-occurs within a single meadow with *C. parallelus*. The risk of mating with *C. montanus* is a potentially major environmental selection pressure with the potential to cause reproductive character displacement.

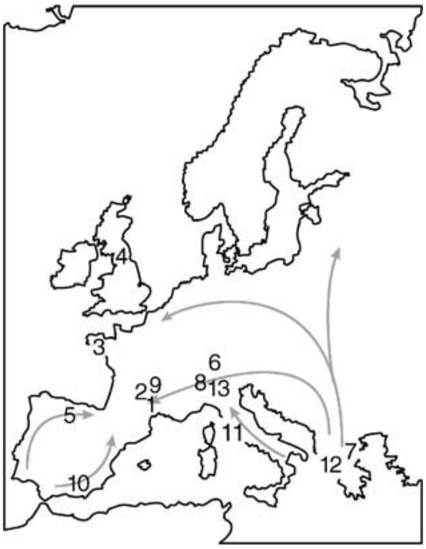

Figure 2. C. parallelus populations sampled (numbered 1–13) and postglacial recolonisation patterns (arrows) inferred from non-coding nuclear and mitochondrial markers (Cooper et al., 1995; Lunt et al., 1998).

The differing histories of these populations allowed us to compare patterns of divergence with models based on competing theories for the evolution of reproductive isolation. Our approach was to collect females from 13 wild populations representing a range of evolutionary histories (Figure 2). Eggs from these females underwent an obligatory diapause and hatchlings from all populations were reared simultaneously under standardised conditions. At adulthood, all individuals were marked with a spot of paint and individuals were isolated according to their sex to preserve virginity. To examine divergence in phenotype (Tregenza, Pritchard & Butlin, 2000c) we examined three sets of traits: morphology (in 1217 individuals), and two putative sexual signals: contact pheromones (in 882 individuals), and calling song (in 260 males). Morphological measurements consisted of four body part lengths in both sexes and two additional measurements in males of the number and density of pegs (used in song production) on the inside of the femur. The existence of contact pheromones, known in other orthopterans (Howard & Blomquist, 1982; Bell & Carde, 1984; Tregenza & Wedell, 1997) has been inferred in *C. parallelus* by the existence of mate choice in the absence of male song (Ritchie, 1990) and by the finding that males will court dead females but cease to do so if their cuticular hydrocarbons are removed (Butlin, 1998). Cuticular composition was quantified according to the relative proportions of 14 cuticular components differentiated using gas chromatography (Tregenza et al., 2000a). Male calling song is implicated in mate attraction and discrimination (Butlin &

Hewitt, 1985; Ritchie, 1990) and was recorded in the laboratory, standardised in relation to temperature, digitised and four temporal characteristics of the song were measured five times in each individual.

To directly examine reproductive isolation, we conducted assortative mating trials using seven of our populations with at least 20 crosses between all pair-wise combinations (Tregenza, Pritchard & Butlin, 2000b). In each trial a male and a female from each of two populations were placed in a 40 × 30 × 20 cm mating arena and the identity of the first pair to mate was recorded. This design allowed us to independently examine differences in mating propensity and differences in assortative mating across populations. After mating, females were allowed to lay eggs and the following year these eggs were reared to adulthood. We assessed postmating isolation between populations by measuring male testes follicle length (Tregenza, Pritchard & Butlin, 2002). Follicle length has previously been shown to correlate strongly with other measures of testis function in crosses between populations from either side of the *C. parallelus* Pyrenean hybrid zone (Hewitt, Butlin & East, 1987; Virdee & Hewitt, 1992) and in other grasshoppers (Hewitt & East, 1978), with shorter follicles associated with more complete sterility.

Results and analytical methodologies

We used different statistical approaches to examine trait divergence and reproductive isolation; full details of methodologies are given in Tregenza, Pritchard and Butlin (2000b, c, 2002). Below I summarise the results of these studies and describe additional analyses comparing patterns of divergence in genetic markers, traits and premating and postmating reproductive isolation.

Phenotypic divergence

To examine patterns of phenotypic divergence, we first used discriminant function analysis to combine all measurements within each of three suites of traits (morphology, cuticular composition and male song) to produce combined measures that maximise variance among populations. Dividing data into three trait suites is a compromise necessitated by the substantial non-independence of characters within trait suites (e.g. femur length/width are non-independent within 'morphology'), which mean that it is not meaningful to analyse each character separately. However, this approach means that possible associations between characters in different trait suites are not considered. Between population analysis of variance was then performed using each trait suite, and planned comparisons representing the different potential causes of divergence were carried out by dividing populations into groups expected to be most divergent under a particular hypothesis: (1) gradual accumulation of mutations during extended allopatry in glacial refugia: three groups, the descendants of the three glacial refugia; (2) processes associated with colonisation: two groups, the Greek refugial populations and their descendants in northern Europe; (3) adaptation to environment (altitude): three pairs of populations at high altitude with neighbouring populations at low altitude; (4) adaptation to environment (character displacement): two populations sympatric with the closely related *C. montanus* and neighbouring populations outside the range of *C. montanus*. The *F*-value associated with the percentage of the overall variance explained by each of these comparisons indicates whether phenotypic divergence has proceeded as would be expected if each of the potential sources of divergence were predominant. The differing numbers of populations in the different contrasts does not bias the ability of the different hypotheses to explain variance, as confirmed by simulation in Tregenza, Pritchard and Butlin (2000c).

Morphology

The vast majority of inter-population variance in female morphology was explained by a comparison of the Greek populations with their descendants in northern Europe (71% of total variance $F_{1,3} = 39.1$, $P = 0.008$). There was a similar pattern for male morphology (36% of total variance $F_{1,3} = 15.1$, $P = 0.03$), although there was also a lot of variation in male morphology between the Spanish populations and all other populations (46% of total variance $F_{1,3} = 19.6$, $P = 0.02$). Comparisons of populations at different altitudes, or between those exposed to potential reproductive character displacement due to sympatry with *C. montanus* and those allopatric with this species did not explain a significant proportion of variation.

Male calling song

Although there were significant differences in calling song between populations, none of the divisions between populations suggested by factors likely to drive divergence explained a significant proportion of between population variance.

Cuticular composition

There were substantial differences between populations in cuticular composition for both sexes. Comparing patterns of divergence with the hypotheses discussed above revealed significant differences between the Greek population and its descendants for both sexes ($F_{1,2} > 6.7$, $P < 0.05$). Additionally, in females, there were significant differences between Spanish refugial populations and those from other areas ($F_{1,2} = 7.6$, $P = 0.04$), and between Italian refugial populations and those from the Greek refuge and their descendants ($F_{1,2} = 7.4$, $P = 0.04$). Note: there is a typographical error in Table 5 of Tregenza, Pritchard and Butlin (2000c). The data for the sexes in the analysis of cuticular composition are transposed although significance tests are not.

Premating and postmating isolation

To compare the pattern of premating and postmating isolation with different hypotheses for the origin of reproductive isolation, we used a likelihood-based approach (Davies et al., 1997). Likelihood analysis allows us to compare the observed pattern of premating and postmating isolation with models that attempt to predict this pattern, using the 21 independent crosses between populations (7 × 7 populations). Beginning with the simplest model possible, increasingly complex models were compared with the data, including those based on existing theories for which aspects of evolutionary history are likely to drive population divergence. For each new model, a log-likelihood value was computed; this can be tested for an improvement in fit because twice the increase in log-likelihood asymptotically follows the χ^2 distribution with degrees of freedom equal to the number of additional parameters.

Premating isolation

See Tregenza, Pritchard and Butlin (2000b) for details. There was substantial premating isolation between the seven populations, and much of this variation can be explained by fitting a sequence of increasingly complex models. Starting with a null model of random mating between populations, significant increases in log-likelihood were achieved using a model that allowed for differences in male and female mating propensity between populations and a uniform degree of isolation between all populations (an increased likelihood of assortative mating). This intermediate model was modified to reflect the predictions of the potential sources of reproductive isolation discussed above. Adding a parameter to the model that allowed populations descended from different refugia to be more isolated from one another failed to improve the model's explanatory power. Modification of this model to allow greater isolation between populations sympatric with *C. montanus* also failed to improve the fit. However, a large improvement in the fit of the model to the data was achieved through addition of a parameter allowing isolation between populations that have colonised northern Europe and their ancestral Balkan population.

Postmating isolation

See Tregenza, Pritchard and Butlin (2002) for details. In contrast to premating isolation, the model that best explained the pattern of variation in follicle lengths was one that allowed variation between populations from different refugia but not within populations from the same refuge ($P < 0.001$, 60% of the among cross variance). This is because crosses between Spain and the other populations generally produced males with greatly reduced follicle lengths. However, there is clearly differentiation between descendants of the Balkan refuge since a model in which populations are allowed to vary such that their isolation from one another can be expressed by separation along a single axis provided a better fit (more than 90% of the among-cross variance) than the refuge model.

Comparisons between patterns of divergence in genetic markers, traits and premating and postmating reproductive isolation

The likelihood approach allowed us to describe isolation between populations using a model in which populations were arranged on a linear axis with greater distances between more isolated populations. There was no correlation between models of this type for premating and postmating isolation ($r = -0.011$) indicating that the characters underlying the two forms of isolation have evolved independently. The strongest postzygotic isolation was between Spain and the other populations, indicating slow divergence in refugia, while the strongest premating isolation was between Greece and populations derived from the Balkan refuge, indicating an effect of colonisation generating relatively rapid change.

The potential for genetic distance to predict premating isolation was assessed by comparing two models in which male and female mating propensity varied between populations. In the first model, there was a uniform degree of isolation between populations, in the second, isolation increased exponentially with genetic distance (as suggested by Arnold, Verrell & Tilley, 1996). There was no improvement in fit of the second model over the first, indicating that genetic distance and premating isolation were not correlated, primarily because the Spanish population is genetically divergent but not strongly reproductively isolated. We did not test alternative relationships between isolation and genetic distance since there are no *a priori* reasons to think that they would more realistic.

To compare postmating isolation with genetic distance, we constructed models that assume follicle length declines (according to several different functions) with the genetic distance between populations and compared the fit of these models to the data. Genetic distances were taken from K_{st} distances from Cooper, Ibrahim and Hewitt (1995) who combined groups of sites into regions before calculating K_{st} values according to Hudson, Boos and Kaplan (1992). Although these models explained more variation than a null model, they explained less than a model in which there was simply more postmating isolation between than within populations, indicating that there is not a strong correlation between postmating isolation and genetic distance.

Premating isolation versus divergence in phenotypic traits including sexually selected traits

Since premating isolation is an emergent property of other phenotypic traits, patterns of variation in characters involved in mate recognition or choice are expected to be associated with patterns of premating reproductive isolation. We tested this hypothesis using the same approach described above for comparing genetic divergence and premating isolation. Instead of using genetic distance, we used Mahalanobis distance matrices between populations for each trait and followed the suggestion of Arnold, Verrell and Tilley (1996) that, as populations diverge in a signal character and associated preference, the probability of mating between individuals from different populations will decline exponentially. Using this approach, only two of the six trait groups produced models with better fits than the uniform isolation model: those based on female morphology and male cuticular composition.

Discussion

Phenotypic divergence

There is substantial divergence in phenotypic traits between populations that have been isolated from one another for long periods of time; however, there is also a lot of variance among populations derived from the same glacial refugium. This indicates that both long periods of allopatry in refugia and processes associated with range expansion have contributed to the current pattern of phenotypic divergence. The difference between the pattern of divergence in male and female morphology is a reflection of the two additional characters considered as part of male morphology – stridulatory peg number and density. These characters show more divergence between refugia, whilst the other characters have a similar pattern to female morphology. It is tempting to treat peg characters independently of other morphological characters, but this would be a *post hoc* decision of questionable justification.

The fact that none of our hypotheses explains the pattern of variation in male calling song may reflect the limited across population variance in this trait, but may also indicate that song divergence is the result of a combination of factors that are not encapsulated by any of the population divisions we used. For instance, songs may be subjected to selection in regard to ambient temperature, whilst also being influenced by genetic drift.

The presence of sexual dimorphism and the extent of differences between populations in cuticular composition suggest that sexual selection acting through the cuticle's role as a contact pheromone has driven divergence. However, those compounds that differ most between the sexes are not the same as those that vary the most between populations, suggesting that sexual selection alone is not responsible for causing divergence in cuticular composition. The large differences between populations descended from the Greek refugium show that cuticular composition can diverge rapidly without extended periods of allopatry.

Reproductive isolation

Processes associated with rapid range expansion increased premating isolation, but it is not clear which processes. Such rapid range expansion in a substantially sessile species must have created population bottlenecks and founder events, but also would have involved exposure to differing environments. Additionally, low population densities may relax sexual selection since females cannot afford to be choosy about their mates when encounters with males are scarce (Kaneshiro, 1989). It is very difficult to disentangle these factors. The 'Kaneshiro effect' would predict greater divergence in sexual signalling traits in colonising populations, which is not the case. The evidence from comparisons of populations living at different altitudes and sympatric or otherwise with a closely related species suggests that adaptation to environment does not drive premating isolation, but these are only two aspects of the environment. Additionally, although the colonisation model explains much of the pattern of premating isolation there is still considerable unexplained variation indicating that other factors are also important.

Populations from different refugia showed substantial postmating isolation. There was no general effect of colonisation although there was substantial differentiation between descendants of the Balkan refuge. Overall, long periods of allopatry do appear to promote the evolution of postmating isolation, although clearly it can also evolve more rapidly.

Correlations between forms of divergence

The lack of correspondence between patterns of premating and postmating isolation implies that speciation is highly dependent on mating systems. In species where there is less scope for mate choice, speciation may tend to be associated with long periods of allopatry, whereas factors such as population bottlenecks and natural selection may be more important in species with mate selection. Comparative analyses indicate that taxa with elaborate secondary sexual traits have higher net rates of speciation (see Panhuis et al., 2001 for review), consistent with the idea that species with mate choice have an additional route to evolve reproductive isolation, which may accelerate speciation. Our study does not explicitly test the possibility that changes in arbitrary mating signals and preferences as a result of sexual selection might have driven divergence. However, this process predicts greater divergence in sexually selected than non-sexual traits, which does not appear to have occurred.

Genetic divergence is at best weakly correlated with premating reproductive isolation, a pattern also seen in taxa such as the Hawaiian *Drosophila* Coyne and Orr (1989, 1997), and similar to the complete lack of correlation seen in the *D. willistoni* group (Gleason & Ritchie, 1998) the brown planthopper (*N. lugens*) (see Butlin & Tregenza, 1998) and in butterfly fish (Chaetodontidae) (Mcmillan, Weigt & Palumbi, 1999). Indeed, examples of strong correlations between genetic distance and premating isolation are rare (but see Tilley, Verrell & Arnold, 1990) suggesting that our observations may reflect common patterns. Postmating isolation is similarly only weakly associated with genetic divergence, which may reflect the short time scale in which postmating isolation has evolved in this system.

The overall pattern of trait divergence is not correlated with either premating or postmating divergence, however, two traits, female morphology and male cuticular composition are correlated with premating isolation. However, it is unlikely that female morphology had a causative role in mating isolation; its effect was largely driven by the large size of females from the Greek refuge, and an analysis of individual mating trials (Tregenza & Butlin, in preparation) revealed that neither the absolute nor the relative sizes of male and female grasshoppers had a detectable influence on outcomes. Cuticular compounds on the other hand do appear to function as contact pheromones (Ritchie, 1990; Buckley, 1998; Butlin, 1998), they were sexually dimorphic, and varied considerably among populations, with their effect on isolation not dominated by a single population. Song, which is normally considered a major component of the mate recognition system in gomphocerine grasshoppers, did not predict premating isolation, presumably because variation among populations was weak relative to within, giving it little power to explain the pattern of assortment.

To summarise, we found no correlation between postmating and premating isolation. There was only a weak correlation between genetic distance and premating isolation and no correlation with postmating isolation. Most traits, including male calling song were uncorrelated with premating isolation. Variation in female morphology and male cuticular composition were correlated with premating isolation, but female morphology is unlikely to be causative. Our overall finding that patterns of genetic divergence, phenotypic divergence and premating and postmating isolation are not strongly correlated indicates that examining a single form of divergence may give a misleading picture – populations that are genetically or phenotypically more divergent may not be closer to speciation.

Limitations of the intraspecific population comparison approach

Previous studies have examined the relationship between different forms of divergence and reproductive isolation by calculating correlations among factors, particularly between genetic distance and measures of isolation and sexually selected traits. Our study takes this approach a step further by directly comparing the influence of different aspects of evolutionary history. We believe that this approach will be essential to further progress in understanding the evolution of reproductive isolation, but inevitably there are weaknesses and factors that need to be borne in mind.

1. In all studies of reproductive isolation, independent units are ultimately the populations examined. There are a larger number of independent measures of distance (between each pair of populations), providing more degrees of freedom for analysis, but this does not resolve the problem that single populations can have a major influence on the overall pattern. If a population is divergent for reasons that are not explicitly tested in the analysis a misleading picture may be painted. The large influence of the strongly isolated Greek populations in our study is an example of the potential for a single population to have a large influence. Much (but by no means all) of the evidence for the importance of processes associated with colonisation rests on differences between the Balkan refugial populations and their descendants in northern Europe. This makes it tempting to speculate that there is something 'odd' about the Greek population that might undermine our conclusions. However, as with any *post hoc* explanation this hypothesis can only be tested in future studies. It would only be possible to avoid large effects of single populations by using a very large number of replicate populations to reduce the influence of each individual population. Because of the time involved, it seems unlikely that a large number of dramatically more extensive studies will be carried out in the near future. What are needed are replicate studies that also test clear *a priori* predictions.
2. Choice of populations for study is a balance between independence and the scale of variation in reproductive isolation and phenotypic traits. If populations are too closely related they will not be statistically independent, but if they are too distant there may be a lot more variation at a smaller scale than can be revealed by the sample of populations considered. Independent contrasts are possible, but require a large number of populations.
3. Any analysis based on the evolutionary history of populations is dependent on our confidence in the accuracy of that history. In our study, a cause for concern is the possibility there may have been multiple independent glacial refugia in Spain, Italy and the Balkans. If this is the case, our Greek population may not be the ancestral population for the populations re-colonising Europe, although we believe this is unlikely on the basis of shared nuclear and mitochondrial DNA haplotypes between refugial and descendant populations (Cooper, Ibrahim & Hewitt, 1995; Lunt, Ibrahim & Hewitt, 1998). A related issue is the overall reliability of the phylogenies we used. Although the two phylogenies agree in relation to the inferred histories upon which our analysis was based, each uses only a single gene (respectively a 281–286 bp region of the anonymous non-coding nuclear DNA segment *cpnl*-1 and a 300 bp portion of the mitochondrial cytochrome oxidase subunit I gene). The time frame we are interested in, a few hundred thousand years, may not be enough to produce unambiguous patterns of sequence divergence among genes among populations, casting doubt over the reliability of our inferred evolutionary histories.
4. Because approaches based on observations of divergence in natural populations do not allow us to control selection and demographic factors, it is not possible to test completely non-overlapping hypotheses. In our study, colonisation may involve exposure to different environments, and it is not clear to what extent gene flow between colonising populations may have been restricted. Therefore, we could not assess the relative contributions of small population size and adaptation to novel environments in driving divergence between colonising populations. Similarly, populations allopatric for long periods may also have experienced different environmental conditions that have also promoted divergence.
5. The hypothesis that natural selection has driven reproductive isolation through adaptation to environment is difficult to falsify given the multifarious potential sources of environmental selection pressures. The best that can be done is to attempt to identify likely sources of natural selection and to look for effects of those. Similarly, sexual selection does not predict a particular pattern of

isolation in relation to evolutionary history, and so cannot be tested using this type of approach.

6. Two interrelated problems with all studies of variation between populations are that the impact of divergence can only be assessed in a laboratory environment and that the future potential of populations is unknown. We cannot be certain that the premating reproductive isolation measured in our mating experiments mirrors the degree of reproductive isolation that might prove to exist if the same populations were to come into contact in nature. Similarly, populations that are currently more reproductively isolated may not actually be the first to become full species. Interspecific studies would be valuable for comparison with patterns of divergence within species. Across species there is very little information on the correspondence between assortative mating in laboratory and field populations. This is presumably because those species in which it is possible to observe mating patterns in the wild tend to be large vertebrates which are difficult to use in laboratory mating experiments. In *Drosophila*, there are examples of potentially major sources of discrepancy between wild and laboratory studies, such as the dependence of premating isolation on larval substrate in *D. mojavensis* (Etges, 1992). However, the general pattern is that where differences between laboratory and wild mating patterns have been observed they tend to be conservative, in that hybridizations are more frequent in the laboratory, so isolation is underestimated, rather than overestimated (Coyne & Orr, 1989).

Conclusions

The main conclusions that can be drawn from our study are:

1. Long periods of allopatry are not necessary for substantial phenotypic and premating isolation to evolve.
2. The lack of divergence between populations in different environments and the weakness of the correlation between phenotypic traits and premating and postmating isolation suggests that natural selection has not been the primary force driving the evolution of reproductive isolation.
3. Mating signals have not diverged more rapidly than other traits, suggesting that sexual selection is probably not responsible for the evolution of premating isolation.
4. Patterns of genetic divergence, phenotypic divergence and premating and postmating isolation are at best only weakly correlated.
5. Comparisons of the relative importance of long periods of allopatry, natural selection and processes associated with colonisation, show that the latter hypothesis provides a much better fit to the data for premating isolation than either of the other two, but for postmating isolation, long periods of allopatry are more important.

Our finding that certain pairs of populations show strong postmating reproductive isolation, whilst others have higher premating isolation illustrates the value of examining as many populations as possible. For instance, had we simply examined Spanish and French populations we would have found that there was strong among-population postmating isolation, but rather weak premating isolation. Alternatively a study of only northern European populations would have led to the opposite conclusion.

It is clear that the question of the relative importance of different ecological and demographic factors in driving speciation will only be answered by further studies which allow us to gradually build up a picture of how speciation occurs in nature. An important methodology will be to examine populations with known and differing evolutionary histories and to compare patterns of divergence with the predictions of different hypotheses for the origins of reproductive isolation.

Acknowledgements

This work was carried out in collaboration with R.K. Butlin and V.L. Pritchard. It was supported by a NERC grant to R.K.B., and a NERC fellowship to T.T. Thanks to J. Mallet, Jeff Feder, Mohamed Noor and an anonymous referee for comments on the manuscript.

References

Arnold, S.J., P.A. Verrell & S.G. Tilley, 1996. The evolution of asymmetry in sexual isolation: a model and a test case. Evolution 50: 1024–1033.

Bell, W.J. & R.T. Carde, 1984. Chemical Ecology of Insects. Sinauer, Sunderland, Massachusetts.

Buckley, S.H., 1998. Cuticular hydrocarbons and the *Chorthippus parallelus* hybrid zone. PhD Thesis. University of Leeds, Leeds.

Butlin, R.K., 1996. Co-ordination of the sexual signalling system and the genetic basis of differentiation between populations in the brown planthopper, *Nilaparvata lugens*. Heredity 77: 369–377.

Butlin, R.K., 1998. What do hybrid zones in general, and the *Chorthippus parallelus* zone in particular, tell us about speciation?, pp. 367–378 in Endless Forms: Species and Speciation, edited by D.J. Howard & S. Berlocher. Oxford University Press, New York.

Butlin, R.K. & G.M. Hewitt, 1985. A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera, Acrididae) – behavioural characters. Biol. J. Linn. Soc. 26: 287–299.

Butlin, R.K. & T. Tregenza, 1998. Levels of genetic polymorphism: marker loci versus quantitative traits. Phil. Trans. R. Soc. Lond. B 353: 187–198.

Cooper, S.J.B., K.M. Ibrahim & G.M. Hewitt, 1995. Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. Mol. Ecol. 4: 49–60.

Coyne, J.A. & H.A. Orr, 1989. Patterns of speciation in *Drosophila*. Evolution 43: 362–381.

Coyne, J.A. & H.A. Orr, 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 51: 295–303.

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

Etges, W.J., 1992. Premating isolation is determined by larval substrates in cactophilic *Drosophila mojavensis*. Evolution 46: 1945–1950.

Gleason, J.M. & M.G. Ritchie, 1998. Evolution of courtship song and reproductive isolation in the *Drosophila willistoni* species complex: do sexual signals diverge the most quickly? Evolution 52: 1493–1500.

Hewitt, G.M., 1988. Hybrid zones – natural laboratories for evolutionary studies. Trends Ecol. Evol. 3: 158–167.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58: 247–276.

Hewitt, G.M., R.K. Butlin & T.M. East, 1987. Testicular dysfunction in hybrids between parapatric subspecies of the grasshopper *Chorthippus parallelus*. Biol. J. Linn. Soc. 31: 25–34.

Hewitt, G.M. & T.M. East, 1978. Effects of B chromosomes on development in grasshopper embryos. Heredity 41: 347–356.

Howard, R.W. & G.J. Blomquist, 1982. Chemical ecology and biochemistry of insect hydrocarbons. Annu. Rev. Entomol. 27: 149–172.

Hudson, R.R., D.D. Boos & N.L. Kaplan, 1992. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. 9: 138–151.

Kaneshiro, K.Y., 1989. The dynamics of sexual selection and founder effects in species formation, pp. 279–296 in Genetics, Speciation and the Founder Principle, edited by L.V. Giddings, K.Y. Kaneshiro & W.W. Anderson. Oxford University Press, New York.

Lee, C.E., 2000. Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". Evolution 54: 2014–2027.

Lunt, D.H., K.M. Ibrahim & G.M. Hewitt, 1998. MtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*. Heredity 80: 633–641.

Mcmillan, W.O., L.A. Weigt & S.R. Palumbi, 1999. Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (chaetodontidae). Evolution 53: 247–260.

Odeen, A. & A.B. Florin, 2000. Effective population size may limit the power of laboratory experiments to demonstrate sympatric and parapatric speciation. Proc. R. Soc. Lond. B 267: 601–606.

Panhuis, T.M., R. Butlin, M. Zuk & T. Tregenza, 2001. Sexual selection and speciation. Trends Ecol. Evol. 16: 364–371.

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

Ritchie, M.G., 1990. Are differences in song responsible for assortative mating between subspecies of the grasshopper *Chorthippus parallelus* (Orthoptera, Acrididae). Anim. Behav. 39: 685–691.

Rundle, H.D., L. Nagel, J.W. Boughman & D. Schluter, 2000. Natural selection and parallel speciation in sympatric sticklebacks. Science 287: 306–308.

Ryan, M.J., A.S. Rand & L.A. Weigt, 1996. Allozyme and advertisement call variation in the Tungara frog, *Physalaemus pustulosus*. Evolution 50: 2435–2453.

Tilley, S.G., P.A. Verrell & S.J. Arnold, 1990. Correspondence between sexual isolation and allozyme differentiation – a test in the salamander *Desmognathus ochrophaeus*. Proc. Natl. Acad. Sci. USA 87: 2715–2719.

Tregenza, T., S.H. Buckley, V.L. Pritchard & R.K. Butlin, 2000a. Inter- and intrapopulation effects of sex and age on epicuticular composition of meadow grasshopper, *Chorthippus Parallelus*. J. Chem. Ecol. 26: 257–278.

Tregenza, T., V.L. Pritchard & R.K. Butlin, 2000b. The origins of premating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. Evolution 54: 1687–1698.

Tregenza, T., V.L. Pritchard & R.K. Butlin, 2000c. Patterns of trait divergence between populations of the meadow grasshopper, *Chorthippus parallelus*. Evolution 54: 574–585.

Tregenza, T., V.L. Pritchard & R.K. Butlin, 2002. Testing hypotheses for the origins of reproductive isolation: hybrid sterility in the grasshopper *Chorthippus parallelus*. Pop. Ecol. 44: (in press).

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

Virdee, S.R. & G.M. Hewitt, 1990. Ecological components of a hybrid zone in the grasshopper *Chorthippus parallelus* (Zetterstedt) (Orthoptera: Acrididae). Bol. San. Veg. Pla. 20: 299–309.

Virdee, S.R. & G.M. Hewitt, 1992. Postzygotic isolation and Haldane's rule in a grasshopper. Heredity 69: 527–538.