Characterising reproductive barriers between three closely related *Heliconius* butterfly taxa.

Lucie M. Queste

MSc by Research

University of York,

Biology

November 2015

Abstract

Debates about the possibility of divergence in the face of gene flow have been an ongoing feature in the field of speciation. However, recent theoretical studies and examples in nature have demonstrated evidence for such a process. Much research now focuses on finding more evidence of reinforcement such as stronger isolation in sympatric populations. Genomic studies have also been investigating the role of gene flow in sympatric speciation and the formation of islands of divergence. *Heliconius* butterflies offer extensive opportunities to answer such questions. Here, I test whether male colour pattern preference and female host plant preference act as reproductive barriers in three Heliconius taxa with varying degrees of geographic overlap. Further experiments on the F2 hybrids of two of these taxa aimed to identify the underlying genomic architecture of these traits. My results suggest that male colour pattern preference and host preference are acting as reproductive barriers. Stronger differences between the sympatric species were found demonstrating evidence for reinforcement and divergence with gene flow. Initial analyses of the F2 hybrid phenotypes suggest that several loci control these traits and pave the way for future genetic analyses to further understand the role of gene flow in speciation.

Table of Contents

Abstract **2**

Table of Contents **3**

List of Figures **4**

List of Tables **5**

Acknowledgements **6**

Author’s Declaration **7**

Chapter 1 – Introduction **8**

1. Speciation 8
2. *Heliconius* 15
3. Aims and Objectives 20

Chapter 2 – Identifying traits acting as reproductive barriers between three taxa of *Heliconius* with varying levels of gene flow. **21**

Introduction 21
Methods 24
Results 27
Discussion 32

Chapter 3 – Quantifying the phenotype of traits involved in reproductive isolation in F2 hybrids for genomic analysis. **39**

Introduction 39
Methods 41
Results 44
Discussion 53

Chapter 4 – Concluding remarks **58**

References **60**

List of Figures

Figure 1. Pictures of *Heliconius elevatus*, *H. p. butleri* and *H. p. sergestus*. 17

Figure 2. Map of Northern South America showing the ranges of the three taxa. 17

Figure 3. Phylogenetic relationship of species and subspecies in the *melpomene*–silvaniform clade of *Heliconius*. 18

Figure 4. Representation of the experimental set up for the male colour pattern preference experiment. 24

Figure 5. Likelihood of courtship behaviours of the pure species in the male colour pattern experiment. 28

Figure 6. Likelihood of eggs being laid on the four *Passiflora* species by the pure species in the host plant preference experiment. 29

Figure 7. Average number of eggs laid on *P. laurifolia* and *P. riparia* by the three taxa in 31their respective stock cages. 31

Figure 8. Pictures of a hybrid mating between *H. elevatus* and *H. p. butleri*, a hybrid cross, F1 generation and a hybrid cross, F2 generation . 41

Figure 9. F2 individual proportions of courtship behaviours in the male colour pattern preference experiment. 44

Figure 10. Spread of F2 individual preferences and corresponding variance values over K clusters for the male colour pattern experiment. 4746

Figure 11. Likelihoods of eggs laid on *P. laurifolia* against *P. riparia* for the pure species and their F1 and F2 hybrid crosses 48

Figure 12. F2 individual proportions of eggs laid in the host plant experiment 48

Figure 13. Spread of F2 individual preferences for egg laying and corresponding variance values over K clusters for the host plant preference experiment. 49

Figure 14. Proportion of courtship behaviours towards the different colour patterns by the three genotypes in the male colour pattern preference. 51

Figure 15. Proportion of eggs laid on the two hosts by the three genotypes in the host plant preference experiment. 52

List of Tables

Table 1. Significance values from the X^2-tests testing for associations between each colour pattern locus with male colour pattern preference (for approaches and hovers) and female host plant preference. 510

Acknowledgements

I would firstly like to acknowledge Dr. Kanchon Dasmahapatra who gave me the opportunity to carry out my MSc by Research in his group. His continuous guidance (with Skype or in person) throughout the year has helped drive the project forward to new and exciting ends. I have learnt a lot this past year and look forward to four more years working alongside him.

I would then like to thank my unofficial supervisor, Dr. Neil Rosser, who not only put up with me for eleven months in the field but also taught me the ins and outs of rearing *Heliconius* and living in Peru. His input in the different aspects of the project, from data collection to writing up, has been indispensable.

Dr. Julia Ferrari and Dr. Colin Beale also deserve my thanks. They formed an excellent TAP panel, available even outside of mandatory meeting times to give helpful insight to improve my work.

I would then like to thank the people I have met this past year in the context of my MSc. These people have had various forms of input on my work, either through scientific discussions relating to *Heliconius* and speciation, through data collection on the few days off, through help in the field, advice or just through new friendships. These people are (in order in which I met them) Jake Morris, Monica Arias, Ronald Mori, Dr. Mathieu Chouteau, Dr. Melanie McClure, Patricia Velado, Carolina Segami, Dr. Stephen Montgomery, Nate Edelman, Jarreth Caldwell and Christian Alessandro Perez.

Author’s Declaration

I, Lucie Queste, confirm that all the material presented in this thesis is my own, except for the following. Dr. Neil Rosser was involved in the design of the host plant experiment. This work has not been submitted for any other academic award at this, or any other, institution. Information derived from other sources is appropriately referenced.

Chapter 1

Introduction

1. Speciation

Speciation is the formation of new species. Understanding this process can help explain the generation of biodiversity, species' adaptions to specific niches and their responses to changes in their environment. Speciation must be considered in the context of species definitions (Templeton, 1989) but several definitions use different criteria to delimit groups of organisms. The biological species concept (BSC) separates two species as two groups that exhibit complete reproductive isolation (Mayr, 2000) making this the final stage of speciation. Another definition, the genotypic clusters species concept, where two species are genetically clustered groups demonstrating little or no hybrids and no strict delimitation with other overlapping groups, considers reproductive isolation as a means to speciation (Mallet, 2007). This is consistent with observations in nature of varying degrees of isolation between populations. The genome is not considered a single unit involved in reproductive isolation but within the genome itself there can be degrees of divergence due to differences in strength of selection (Wu, 2001). The genotypic clusters species concept will be used to define species in this thesis.

1.1 Modes of speciation

Speciation can occur in different geographic contexts. A geographical classification defines allopatric speciation as speciation in two populations which are physically separate and sympatric speciation where the two populations share the same distribution (Mayr, 1963). These are two extremes of a continuum with various degrees of geographical overlap and consequently speciation is now being considered more in terms of gene flow (Fitzpatrick et al., 2008; Mallet et al., 2009).

Sympatric speciation was until recently a controversial idea. This is because recombination breaks down genetic associations favouring divergence (Felsenstein, 1981). On the other hand, populations that evolved separately could easily diverge through selection in their separate environment or by drift. Another argument against sympatric speciation was that there was very little evidence of sister species with overlapping distributions. A review of studies looking at the current distribution of sister species in different taxa found that out of 309 speciation events, less than 10 % resulted in a shared distribution of more than 90 %. However, over 70 % of sister species had no overlap (Bolnick and Fitzpatrick, 2007). This study failed to include important taxa such as phytophagous insects. Another study looking specifically at birds found that less than 2 % of pairs exhibited complete geographic overlap (Phillimore et al., 2008). However, a study of heliconiine butterflies found that 32-40 % of sister species had over 95 % overlap and 50-65 % shared over 50 % of their distributions (Rosser et al., 2015). The species definition used and the taxa studied will therefore strongly influence evidence for and against sympatric speciation as what might be considered overlapping sister species under one definition would be considered a single species under another (Rosser et al., 2015). However, it seems undeniable today that sympatric speciation has played some role in the formation of species. Furthermore, theoretical (Dieckmann and Doebeli, 1999; Kondrashov and Mina, 1986; Maynard Smith, 1966) and empirical evidence have led to a resurgence in research on sympatric speciation.

1.2 Sympatric speciation

Various scenarios have been shown to lead to conditions that would allow divergence with gene flow. Kondrashov (1986) describes these conditions. Initially, we imagine a single population with polymorphisms where intermediates are maladapted compared to marginal individuals. Two environments are considered; in the first, resources are distributed homogenously. Here, divergent selection occurs through frequency dependent selection as extreme phenotypes are less likely to compete (Dieckmann and Doebeli, 1999). In the second, the population occupies two sub-niches where different phenotypes have higher fitness in different sub-niches (Maynard Smith, 1966). Local adaption occurs and intermediate phenotypes are once again selected against and the polymorphism is maintained.

Once disruptive selection is established, reproductive isolation needs to develop. Reproductive isolation can be caused by pre or post-zygotic barriers which are traits that stop two divergent populations from mixing gametes (Wu, 2001). Post-zygotic barriers arise after the zygote is formed, such as hybrid inviability or maladaptation. Pre-zygotic barriers occur before the formation of the zygote and can involve behaviour, pheromones, host plant, and any other trait that is used by individuals to find and recognise mates.

There are different possibilities for how reproductive isolation can occur, with one allele or two. In the two allele model, a genetic association links a trait causing assortative mating to the trait under divergent selection (Maynard Smith, 1966). If the phenotypes have a higher fitness for a particular resource then the advantage of mating with similar individuals and avoiding intermediates would select for this association. However, how easily this association would evolve is dependent on certain conditions, such as the strength of selection and migration (Maynard Smith, 1966); this is called the selection-recombination antagonism (Coyne and Orr, 2004). The one allele model does not need such an association to evolve as the trait under divergent selection also causes assortative mating (Kondrashov and Mina, 1986). However, there is a lot of controversy around the existence and prevalence of such traits (Maynard Smith, 1966). The two models have been shown to lead to reproductive isolation (Dieckmann and Doebeli, 1999) and it is likely that the two interact in the process of sympatric speciation.

Theoretical studies have shown that once divergent selection and assortative mating have been established, then sympatric speciation can occur (Kondrashov and Mina, 1986). This will depend on many factors such as the number of loci controlling the trait under divergent selection, the strength of selection against hybrids, the strength of sexual selection on assortative mating and the interaction between these factors. Nevertheless, sympatric speciation is theoretically possible and new evidence in nature supports this. Debates about sympatric speciation now focus on its prevalence.

The most famous evidence of sympatric speciation in nature comes from the apple maggot fly *Rhagoletis pomonella*. Evidence suggests that the apple feeding race arose from the hawthorne feeding race (Bush, 1969). The change in host led to reproductive isolation due to host associated mating and shifts in diapause of the different races. Host shift driven divergence has often been studied in the context of sympatric speciation as host shifts are considered to occur within the range of the initial population (Bolnick and Fitzpatrick, 2007). Arguments against this example of sympatric speciation arose after evidence suggested that the genes involved in host shift introgressed from an allopatric population (Feder et al., 2003). Nevertheless, there is still evidence of divergence occurring in sympatry even though initial variation did not and the apple maggot fly remains an example supporting early stages of sympatric speciation.

Other examples of sympatric speciation are found in cichlid fish. In crater lakes of Cameroon, cichlids showing evidence of assortative mating with few intermediates have a monophyletic origin suggesting divergence in sympatry (Schliewen et al., 1994). More recent work using next generation sequencing however, suggests that introgression from nearby rivers was involved in some of these speciation events (Martin et al., 2015). Another more recent study by Barluenga et al. (2006) found evidence for sympatric speciation in cichlids in a Lake in Nicaragua. F-statistics and mate choice experiments have demonstrated assortative mating and evidence from mitochondrial DNA suggest a single colonisation event in the lake. There is again, controversy over these results; according to Schliewen et al. (2006) this example does not take into account other species in the lake. It also describes results from nuclear DNA which disagree with the conclusions from the mitochondrial DNA, suggesting that there may have been more than one colonisation. Identifying examples of species which have evolved in sympatry is not straightforward, these studies underline the difficulty in determining past conditions and understanding the selective forces at play in speciation.

1.3 Reinforcement

Due to the difficulty of disproving any period of allopatry a lot of research now focuses on finding evidence that divergence can occur in the face of gene flow, regardless of whether an allopatric period has initiated it. Many studies have therefore looked for evidence of reinforcement. Reinforcement is the process by which natural selection increases reproductive isolation to avoid maladaptive hybrids (Noor, 1999), which implies ongoing gene flow. This process was traditionally considered to occur during secondary contact after a period of allopatry (Coyne and Orr, 2004; Noor, 1999) as it requires pre-existing pre and post-zygotic isolation (Liou and Price, 1994). It has since been shown that a similar process arises in the later stages of sympatric speciation (Servedio and Noor, 2003).

Originally, reinforcement was not widely accepted as a plausible process of divergence. The same arguments used against sympatric speciation about the selection-recombination antagonism apply to reinforcement. However, the same mechanisms as in the one and two allele models can allow selection to occur in the face of gene flow during reinforcement (Coyne and Orr, 2004). It was also argued that if two populations were to come into contact, one population being much smaller in size, than individuals of this population would by chance encounter more heterospecific individuals. The smaller population would, relatively to its size, be producing many more hybrids than the bigger population. In doing so, the small population would decrease in numbers (Coyne and Orr, 2004); either because the hybrids had an advantage creating a “hybrid swarm” or because they were unfit, therefore driving the smaller population to extinction (Liou and Price, 1994). Populations of similar sizes with symmetrical migration from either side are therefore more likely to survive long enough for reinforcement to occur (Servedio and Kirkpatrick, 1997). There is also the “swamping effect” where genes from the allopatric population “swamp” new alleles being selected for in the area of contact which inhibits reinforcement (Coyne and Orr, 2004). However, if the populations overlap over most of their range (like in the stages of sympatric speciation) this effect is not as strong.

One of the predictions under reinforcement is that there will be stronger pre-zygotic isolation between populations in sympatry than in allopatry (Coyne and Orr, 2004; Liou and Price, 1994; Noor, 1999; Rice and Hostert, 1993); no difference is expected for post-zygotic barriers. This is because post-zygotic isolation will be costly to individuals involved in hybrid mating, therefore, selection will act upon traits causing pre-zygotic isolation, avoiding the cost of post-zygotic isolation.

Evidence of stronger pre-zygotic isolation in sympatry has been observed in many taxa (Hillis, 1981; Waage, 1975), sparking a resurgence of studies on the theoretical potential of reinforcement (Coyne and Orr, 2004). An example in insects is found in crickets which, in sympatry, have stronger differences in song (chirp rate) and evidence for stronger female discrimination against heterospecifics, compared to populations in allopatry (Jang and Gerhardt, 2006). Song is also a reproductive barrier in certain species of frogs. Blair (1964) studied several Anuran species pairs and found evidence of stronger differences in mating call in sympatry than in allopatry, which was accompanied by female discrimination between calls in the sympatric populations. A famous example is also found in the sticklebacks where females from sympatric populations mated less readily with heterospecific males while allopatric females showed no difference (Rundle and Schluter, 1998). These cases are only a fraction of the examples of stronger isolation in sympatry, evidence extends to other taxa such as plants and fungi. In all these cases however, we observe the outcome of a process and sometimes reinforcement is not the only explanation. In Darwin’s finches, for example, males in sympatric populations have a stronger preference for conspecific females compared to allopatric males (Ratcliffe and Grant, 1983) but the effect of learning could not be disproved as having a role in this pattern (Coyne and Orr, 2004).

Coyne and Orr (2004) carried out a comparative analysis of 171 species pairs in *Drosophila* to look at the level of isolation in sympatry and allopatry, while controlling for age of the speciation event (as estimated by genetic distance). The results showed much higher isolation in sympatric taxa compared to allopatric taxa, and no difference in post zygotic isolation. Although this data is restricted to one genus, it demonstrates strong evidence for reinforcement. Alternatives to reinforcement exist, such as differential fusion (where only populations with sufficient pre-existing prezygotic barriers manage to remain distinct in sympatry and so isolation does not evolve as a response to contact) or runaway sexual selection over a gradient (causing different traits to be selected for at the extremes). However, these do not necessarily account for the lack of post-zygotic isolation in sympatry or cases where the population is not distributed over a gradient (Coyne and Orr, 2004).

Therefore, reinforcement has become a widely accepted process in sympatric populations; with the identification of stronger isolation in sympatry compared to allopatry being a strong indicator of divergence with gene flow. However, there are also cases where no difference is found between allopatric and sympatric populations and cases where pre-zygotic barriers are stronger in allopatry (Marshall et al., 2002). Nevertheless, evidence suggests reinforcement occurs and can plays a role in speciation. Work is now focused on finding more evidence, determining the extent of its role in shaping diversity and understanding which conditions lead to it. For example, a study by Nosil et al. (2003)

found empirical evidence demonstrating the importance of the balance between population size and levels of gene flow (which must be strong enough to allow selection to act upon hybrids but weak enough to avoid the effects or recombination).

1.4 Genomic structure of speciation: islands of divergence

It has therefore been established that evidence for reinforcement suggests divergence with gene flow. However, with the development of new techniques for genetic analysis, the cost and time required for these types of studies have decreased. This has allowed an increase in genome wide studies, in a wider range of organisms, to identify genomic patterns of sympatric speciation and divergence with gene flow (Nosil, 2012). Divergence in allopatry and sympatry occurs under different selective pressures and are expected to create different patterns of divergence across the genome. Identifying and describing these patterns can shed light upon the speciation process.

In allopatry, gene flow does not constrain divergence, which occurs through selection and drift. Therefore divergence is not restricted to areas of the genome with lower recombination. With time, divergence is distributed randomly across the genome (Coyne and Orr, 2004; Nosil, 2012). In sympatry, two forces will act upon the genome: recombination, which has a homogenising effect, and divergent selection creating a heterogeneous genome (Martin et al., 2013). This may result in the formation of a mosaic, or “islands of divergence” (Via and West, 2008).

Islands of divergence are areas of the genome where differentiation is higher than expected under neutral selection (Nosil, 2012). Islands occur as loci which are physically linked to a locus under strong divergent selection will “hitchhike”, extending the effect of this selection (Feder et al., 2012; Nosil and Feder, 2012). This can allow for less strongly selected loci to diverge in association with others (Via and West, 2008).These regions therefore become target areas for potential loci controlling traits involved in reproductive isolation, (Nadeau et al., 2012). The recombination-selection balance is an important factor in these islands as selection needs to be strong enough to maintain associations between loci, and migration weak enough to avoid breaking them (Nosil, 2012). Different structures can help reduce recombination in areas under divergent selection, such as genetic inversions which are chromosomal rearrangements which prevent recombination (Noor et al., 2001) and therefore maintain associations between loci.

When traits which cause reproductive isolation are “taken up” in these regions under strong divergent selection then reproductive isolation can increase. This occurs, as stated before, through tight physical linkage or inversions. Pleiotropy, where a single locus controls two seemingly unrelated phenotypic traits (Nosil, 2012), can also play an important role. If a gene under divergent selection also controls assortative mating than the associated trait is called a “magic trait” (Servedio et al., 2011). The existence of magic traits has been discussed extensively, and a growing number of examples have appeared in the past few years (Servedio et al., 2011). A convincing example is found in *Heliconius* butterflies which have aposematic colour patterns. These patterns are under divergent selection as intermediate patterns are more vulnerable to predators, and experiments have demonstrated assortative mating based on these patterns (Merrill et al., 2015). Whether these traits are common and what role they play in speciation is yet to be determined (Servedio et al., 2011); but their potential in establishing reproductive isolation is important.

The identification of islands of divergence and the role of the genes within them is strong evidence of divergence with gene flow. It is important to note however, that with time increased reproductive isolation due to divergent selection at one or several islands of divergence can reduce gene flow on a genomic scale, called “genomic hitchhiking” (Feder et al., 2012). This creates conditions similar to allopatric conditions with no gene flow where divergence can occur by drift, called “isolation by adaptation” (Nosil et al., 2009; Via, 2012), removing the pattern of islands of divergence. Therefore, the lack of evidence for islands of divergence does not rule out the possibility of divergence with gene flow and sympatric speciation.

Evidence for such islands has been found in studies carrying out genomic analyses on populations such as the oceanic and freshwater sticklebacks (Hohenlohe et al., 2012) where several islands, forming an “archipelago”, of higher linkage disequilibrium were found. Another study in the flycatchers found 50 peaks of high divergence, also fitting the pattern of islands of divergence, though no function was attributed to these peaks (Ellegren et al., 2012). However, there are also counter examples. Genomic analyses of *Rhagoletis* flies in the early stages of speciation did not find evidence of islands of divergence, but rather large areas with varying levels of divergence facilitated by pre-existing variation (Michel et al., 2010). These large regions were associated with loci controlling host preferences, latitudinal clines and eclosion time, important for isolation. This suggests a role of selection acting on many loci instead of a few islands. Interactions with different factors can also influence the genomic pattern, such as migration, recombination, and ancestral variation, making it difficult to identify evidence of gene flow (Martin et al., 2013).

Current research is focusing on differentiating ancestral variation from recent gene flow, and the timing of gene flow with the establishment of reproductive isolation. An analysis in *Heliconius* carried out a test for admixture in several species using single nucleotide sites and determined that introgression had shaped patterns of shared variation (Heliconius

Genome Consortium, 2012). In the study of flycatchers, peaks of divergence were shown to have low nucleotide diversity which was not due to a low mutation rate and not attributed to ancestral polymorphism (Ellegren et al., 2012). Therefore, gene flow can be distinguished from other factors that result in shared variation. A study by Martin (2013) in *Heliconius* looked at gene flow at different periods of divergence to determine its effects over different time scales. This suggested several periods of ongoing gene flow throughout divergence.

Therefore, genomics can provide a more thorough understanding of past speciation events by looking at patterns of divergence. It is important to be aware of factors which can influence these patterns, but studies have been able to disentangle these. Finally, although strong evidence of divergence in the face of gene flow can be found, it is difficult to prove that speciation occurred without periods of allopatry. It is probable that both modes interact in nature.

2. *Heliconius*

Heliconius butterflies are the most species rich genus in the *Heliconiini* Tribe (Nymphalidae: Heliconiinae).The genus arose 10.5-13.4 Ma after which it underwent a rapid adaptive radiation, which mainly occurred on the eastern slope of the Andes and in the upper and middle Amazon basin, and coincided with the rise of the Andes (Kozak et al., 2015; Rosser et al., 2012). Today, *Heliconius* are found throughout the Neotropics from northern Argentina to the Southern United States, with species richness highest in the eastern slopes of the Andes (Rosser et al., 2012). *Heliconius* have been the subject of extensive research for over a century, which is due in part to their abundance and potential to be reared in laboratory conditions, but mostly to the variety of evolutionary questions they can answer.

2.1 A study system in evolutionary biology

The recent and rapid adaptive radiation of *Heliconius* has been the focus of much speciation research (Supple et al., 2014). *Heliconius* offer the possibility to study speciation at different stages of divergence from species races to fully reproductively isolated species (Supple et al., 2013). The geography of *Heliconius* species, which demonstrates extensive overlap of race and species, also allow investigation into the role of geographic barriers and gene flow in divergence (Supple et al., 2015). Evidence of hybrid speciation and horizontal transfer of adaptive traits have also shaped the

Heliconius phylogeny which has increased our understanding of the mechanisms underlying these processes (Heliconius Genome Consortium, 2012).

Heliconius butterflies are known for their bright aposematic colour patterns, which are used to signal their unpalatability to predators (Bates, 1862). Species across the phylogeny converge in colour patterns to form mimicry rings with other species within *Heliconius* and from other genera (Merrill et al., 2015) making them an excellent example of Müllerian mimicry. The coexistence of these mimicry rings is somewhat surprising as species would be expected to converge to a single model. This may be explained in part by differences in predators but a role of habitat and geography is likely (Mallet and Gilbert, 1995).

This mimicry is not restricted to colour pattern and evidence has found convergence in traits relating to wing shape (Jones et al., 2013; Mérot et al., 2013) and flight (Srygley, 1999). Colour pattern also plays an important role in mate choice, with evidence of male discrimination against heterospecific colour patterns (Jiggins et al., 2001; Merrill et al., 2011b). This dual function of colour pattern is strong evidence for its role as a magic trait (Servedio et al., 2011).

The striking colour patterns of *Heliconius* are not the only source of interest for evolutionary biologists. The adaptive radiation is explained in part by the genus' specialisation on *Passiflora* used for oviposition and larval rearing (Benson et al., 1975). The co-evolution between *Passiflora* and *Heliconius* has sparked interest in research looking at host adaptation and its role in diversification. It is also known that cyanogenic compounds are sequestered by the larvae from *Passiflora* (Merrill et al., 2015). Adaptations in the *Passiflora* also seem to reflect this co-evolution with species evolving nodules which mimic butterfly eggs to avoid oviposition and extra-floral nectaries which attract ants, predators of the eggs (Merrill et al., 2015). *Helconius* are also the only known genus in Lepidoptera to feed on pollen as adults (Gilbert, 1972). Pollen provides a source of amino acids to which their surprising longevity (up to six months) is attributed and is crucial to the reproductive success (Gilbert, 1972).

2.2 A model system for speciation genomics

Heliconius have also been the subject of extensive genetic research, which led to the sequencing of the *H. melpomene* genome (Heliconius Genome Consortium, 2012) and the complete taxonomical phylogeny (Kozak et al., 2015). The recent radiation across the Amazon and current geographical distribution of species coexisting does not fit models of speciation by geographic isolation (Dasmahapatra et al., 2010). Therefore, genetic studies have investigated gene flow between species and its role in speciation (Heliconius

Genome Consortium, 2012; Martin et al., 2013). There is also evidence of hybridisation and hybrid speciation in *Heliconius*, *H. heurippa* is thought to have evolved through hybrid speciation of *H*. melpomene and *H. cydno* (Mavárez et al., 2006).

Genomic analyses have been able to study the architecture of adaptive traits in *Heliconius*. Adaptive introgression has played an important role in their diversity, for example the wing pattern loci *B/D* and *Y/b* seems to have introgressed from the *melpomene* clade into the silvaniform clade (Heliconius Genome Consortium, 2012). A single origin of these colour pattern trait was also found within *H. erato*. However, genetic analyses suggest these traits evolved independently between *H. melpomene* and *H. erato* (Supple et al., 2013).

The genomics of wing pattern have also been extensively studied in *H. numata.* Several polymorphisms exist within *H. numata* which mimic species of the genus *Melinaea* (Ithominae)*.* The entire wing pattern is controlled within a supergene P within a chromosomal inversion which prevents recombination (Joron et al., 2006). The different patterns demonstrate a dominant relationships to avoid intermediates. *Heliconius* research has therefore evolved with technology and offers the potential of answering complex genomic questions in the field of evolutionary biology.

2.3 Heliconius elevatus and the Heliconius pardalinus sub-species

This thesis focusses on three taxa: *H. elevatus* and two sub-species, *H. pardalinus butleri* and *H. p. sergestus* (fig 1). The geographic ranges of *H. elevatus* and *H. p. butleri* extend over the northern half of the Amazon basin and are largely overlapping (fig 2). *Heliconius p. sergestus* is found in a restricted area at the edge of the Amazon basin in Peru (fig 2); it is allopatric with respect to *H. p. butleri* and parapatric with *H. elevatus*.

Figure 1. *Heliconius elevatus* (on the left), *H. p. butleri* (centre) and *H p. sergestus* (on the right).

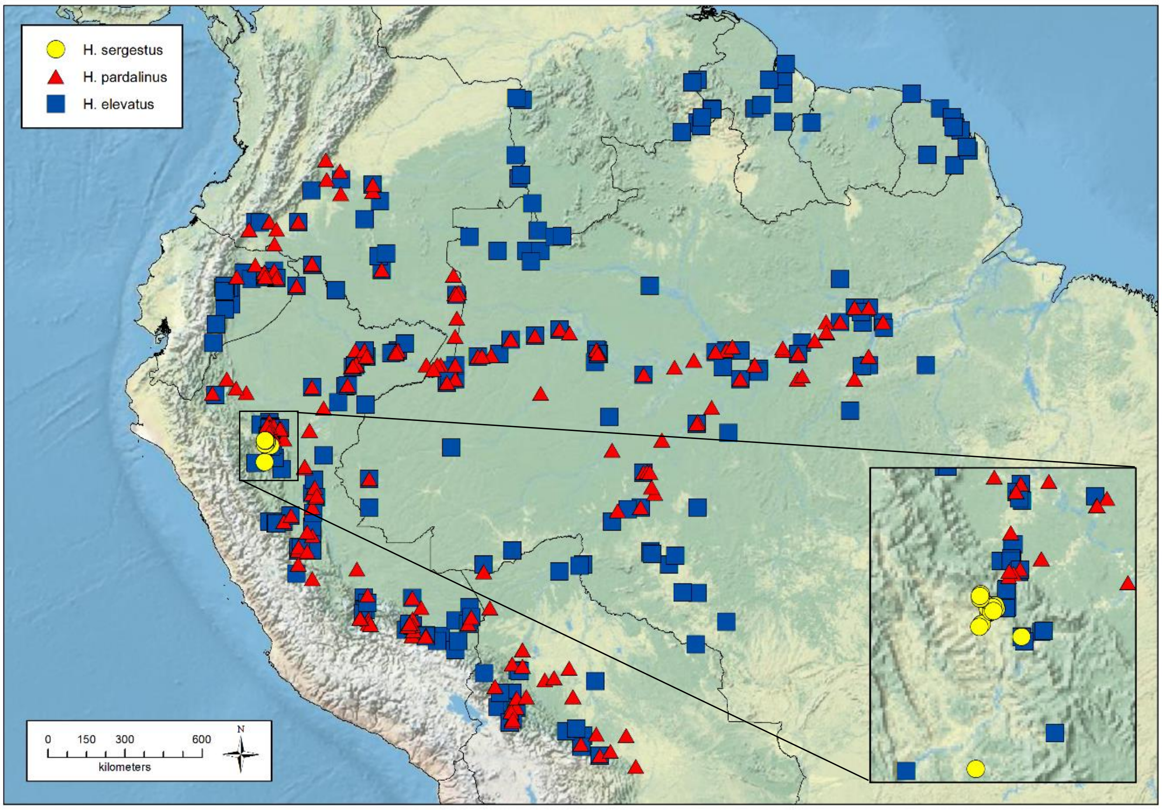

Figure 2. Map of Northern South America showing the ranges of the three taxa *H. elevatus* in blue, *H. p. butleri* in red and *H. p. sergestus* in yellow. Data points were collected from museum samples and field collections (Rosser, 2012).

Heliconius elevatus and *H. p. butleri* are sympatric with peaks of divergence comprising ~5% of the genome; the remaining 95 % show evidence for gene flow (Dasmahapatra *et al*., In prep.). Two of the peaks of divergence between the species are known to contain loci controlling colour pattern, for which they are phenotypically divergent. Although part of the silvaniform clade, *H. elevatus* belongs to a separate mimicry ring, sharing its "Dennis rayed" colour pattern with other species like *H. melpomene* (fig 3).

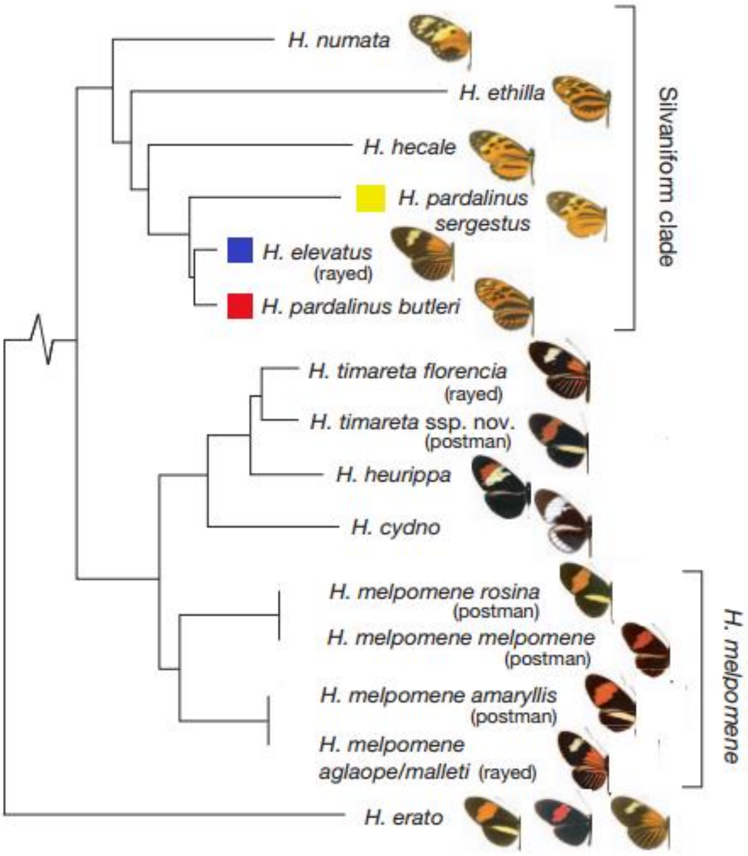

Figure 3. Phylogenetic relationship of species and subspecies in the *melpomene*–silvaniform clade of *Heliconius* (Heliconius Genome Consortium, 2012). The coloured squares refer to the geographical distributions of the taxa in fig. 2.

Evidence suggests *H. elevatus* could be the result of adaptive introgression of colour pattern genes from *H. melpomene* into the silvaniform clade. Although it is a sub-species of *H. pardalinus, H. p. sergestus* is more distantly related to *H. p. butleri* than *H. elevatus* is, making *H. pardalinus* paraphyletic (fig 3) The paraphyly of *H. pardalinus* is presumed to be due to gene flow within Amazonian *H. p. butleri* and *H. elevatus*.

3. Aims and Objectives

The second chapter of this thesis will aim to identify traits causing reproductive isolation between the three taxa. Male colour pattern preference and female host plant preference will be assessed for their role as reproductive barriers. The relative strength of this isolation will then be compared in light of the geographical overlap of the taxa. If traits acting as reproductive barriers exhibit stronger differences in sympatry this would suggest a role for reinforcement in the speciation process of these taxa.

The third chapter will aim to quantify the traits identified as reproductive barriers in the F2 hybrid generation of *H. p. butleri* and *H. elevatus.* This will allow an initial assessment of the genomic architectures underlying these traits, and pave the way for further genetic studies to understand the forces of selection acting upon these traits and assess the role of sympatric speciation in the formation of these taxa.

Chapter 2

Identifying traits acting as reproductive barriers between three taxa of *Heliconius* with varying levels of gene flow.

Introduction

The possibility of two populations diverging despite ongoing gene flow has been a controversial topic in the field of speciation (Coyne and Orr, 2004). The main argument put forward against divergence with gene flow is that recombination will break down any associations between traits under divergent selection and those causing assortative mating (Felsenstein, 1981). Very high levels of selection against intermediate phenotypes would be necessary to overcome recombination between diverging populations. However, new theories and a growing body of empirical evidence have shown that, under certain conditions, species can arise and be maintained with continuous gene flow (Kondrashov and Mina, 1986). Associations between "speciation" traits can be maintained by mechanisms which reduce recombination, such as chromosomal inversions, tight physical linkage or pleiotropy (Nosil, 2012).The debate has now moved on to whether divergence in the presence of gene flow is a common process in speciation (Papadopulos et al., 2011). Many studies have been able to demonstrate historical or ongoing gene flow between divergent populations currently exhibiting strong reproductive isolation (Heliconius Genome, 2012; Martin et al., 2013), though none have been able to disprove any period of allopatry.

Reproductive isolation can arise through reinforcement, whereby prezygotic isolation evolves in response to selection against maladapted hybrids (Coyne and Orr, 2004; Noor, 1999). Reinforcement was originally considered to occur between populations with renewed gene flow after a period of allopatry (Coyne and Orr, 2004). The process was contested for several reasons. Among the arguments were that selection would break down associations between loci, that one of the populations would be driven to extinction before reinforcement could occur, or that alleles from the allopatric population would swamp the areas in contact and prevent new alleles from spreading (Coyne and Orr, 2004). However, evidence of stronger prezygotic isolation in sympatric populations compared to allopatric populations, expected under reinforcement, has led to a resurgence of support (Coyne and Orr, 2004). Reinforcement is now considered to play a role in modes of speciation other than allopatric speciation, as similar conditions to secondary contact are expected in in the later stages of sympatric and parapatric speciation (Servedio and Noor, 2003).

Much research now focuses on identifying examples of divergence with gene flow in various taxa regardless of whether or not it was initiated in sympatry. This should further our understanding of the underlying mechanisms that maintain species differences. Evidence of reinforcement, i.e. stronger pre-zygotic isolation between sympatric sister species than between allopatric sister species (Coyne and Orr, 2004), can be used to determine whether species have diverged with ongoing gene flow. Prezygotic barriers are traits which stop individuals from finding, or mating with, individuals from different populations or species (Wu, 2001). These traits can be involved in behaviour, morphology or physiology. In this thesis, I will be focusing on female host plant preference for oviposition and male colour pattern preference in mate choice.

Male colour pattern preference as a reproductive barrier

Colour pattern is an important trait in the *Heliconius* genus. *Heliconius* butterflies are unpalatable to their predators and signal this using vibrant colours on their wings (Merrill et al., 2015). They participate in Mullerian mimicry rings, in which different species of different genera form groups that exhibit similar colour patterns. Co-mimics are not necessarily closely phylogenetically related. An example can be seen in the sister species *H. melpomene* and *H. cydno*. *H. melpomene* exhibits a red, yellow and black pattern and mimics *H. erato*, whereas *H. cydno* exhibits a blue and white pattern and mimic *H. sapho*. Hybrids displaying an intermediate pattern suffer increased mortality because predators do not recognise them as belonging to a mimicry ring (Merrill et al., 2015).

Males of certain species of*Heliconius* are also known to use colour pattern for mate choice. Many experiments, using both live females and models, have shown evidence of assortative mating based on colour pattern, with males more attracted to conspecifics (Jiggins et al., 2001; Merrill et al., 2011b). Therefore colour pattern can act as a ‘magic trait’, where a trait under divergent natural selection is also involved in mate choice (Jiggins et al., 2001; Merrill et al., 2015; Servedio et al., 2011). In this chapter, I will attempt to identify differences in male colour pattern preference in my system which, if present, could be acting as pre-zygotic barriers; and thereby assess the potential role of colour pattern as a magic trait in our system.

Host plant preference as a reproductive barrier

The interaction between host and phytophagous insects has been extensively studied, with many cases of tight coevolution between the two. *Heliconius* are famously known for their coevolution with their Passifloracea hosts (Benson et al., 1975). Although the available evidence is not as strong as for colour pattern, this close interaction between host and *Heliconius* can have a direct role in reproductive isolation. In populations of phytophagous insects using several host plants, selection may favour individuals using rarer hosts due to competition for feeding or oviposition (Berlocher and Feder, 2002). Different hosts may also offer different advantages such as higher predator deterrence and higher nutritional value or other compounds used by the insects. Plants can also develop defences against phytophagous insects, leading to further host specific adaptations. Therefore, plant host can be a strong driver of divergent selection in phytophagous insects.

This divergent selection can have a direct effect on reproductive isolation. Evidence of host associated adaptations causing isolation have been found in the apple maggot fly, which is temporally isolated from other races of maggot fly because individuals tend to mate on the host they developed on (Feder et al., 1994). In aphids, positive genetic associations between traits controlling specialisation (measured as fecundity) on host and host acceptance (or preference) which is associated to mate choice were found, which facilitates divergence (Hawthorne and Via, 2001). Assortative mating is caused by host preference as mating occur on the host. Males of some species of Heliconius are known to visit different host plants sequentially looking for females (Mallet, 1986). *Heliconius erato* even performs “pupal mating” where males guard pupae waiting for the females to emerge to mate (Deinert et al., 1994), although no evidence of this behaviour was found in the species studied here. Even without this specific behaviour, individuals using the same hosts might demonstrate a higher chance of encountering conspecifics (Rosser et al., 2015). Therefore, host plant use can lead to reproductive isolation if the loci controlling for this trait are under divergent selection.

System

The studied system comprises three taxa, *H. elevatus*, *H. p. butleri* and *H. p. sergestus*, which have varying levels of taxonomic, genetic and geographic distance between them (see fig 2 and 3, Chapter 1). *Heliconius p. sergestus* and *butleri*, are allopatric, *H. p. butleri* and *H. elevatus* are sympatric, and *Heliconius p. sergestus* and *H. elevatus* are parapatric. Although *H. p. butleri* and *H. p. sergestus* are currently classed as the same species, *H. p. butleri* and *H. elevatus* are more closely related to each other than either is to *H. p. sergestus,* over 95 % of their genome shows no divergence (Dasmahapatra et al., In prep).

Using this system, the differences in phenotypic divergence between different degrees of gene flow can be investigated. Under reinforcement, stronger pre-zygotic isolation between taxa with higher levels of gene flow is expected. This was observed in assortative mating experiments where a single female of one of the three taxa was presented to five males of each taxon simultaneously (Segami Marzal, 2015; Velado, 2015). In these experiments, matings occurred between *H. p. butleri* and *H. p. sergestus* but *H. elevatus* only ever mated with conspecifics. I therefore aim to identify evidence for reproductive barriers causing the observed isolation in two traits, host preference in egg laying and male colour pattern preference, and test whether these differences are stronger between the sympatric species, *H. elevatus* and *H. p. butleri.*

Methods

Collection and rearing

All experiments were carried out in Tarapoto, San Martin, in Peru, on *H. elevatus*, *H. p. butleri* and *H. p. sergestus*. Wild butterflies were collected in the regions of San Martin and Loreto in Peru. Butterflies were kept in large insectaries, separated by taxon into “stock” populations containing both males and females. Cages were supplied with a 10% sugar solution containing pollen and provided *Lantana* flowers for additional pollen. Females were presented with shoots of *Passiflora edulis*, *P. laurifollia*, *P. riparia* and *P. serrato-digitata* in bottles for oviposition. Rearing of the larvae was performed in pots with usually one to five larvae, or in large rearing cages where larvae were left to develop on shoots. On emergence, new butterflies were placed in with the stock; virgin females were occasionally kept isolated for experimental purposes.

Male colour pattern preference

In *Heliconius*, wing colour pattern has been shown to be used in mate selection, males preferring females with the same pattern (Jiggins et al., 2001; Merrill et al., 2011b). Colour pattern therefore has an important role in mate choice and potentially as a reproductive barrier. Male colour pattern preference was investigated measuring courtship towards female models. Models were made using real female wings to control for any cryptic sexual dimorphism. Wings were collected and cleaned with dichloromethane to remove any potential effect of pheromones on male preference and glued to pieces of paper cut to the same shape. Using a small piece of adhesive tape, wings were fixed to a thin black straw, imitating the body of the butterfly. The butterfly models were glued onto a cable tie fixed to a horizontal tube in the cage (see fig 4). This allowed the wings to be manipulated from outside the cage to simulate wing movement without disturbing the butterflies.

Two taxa were tested at a time; two males, one of each taxon, were put into 2x2x2 m cages and left to acclimate for at least one day. Courtship towards two models of the same taxon as the males was measured over 15 minute trials. Males were measured four times throughout the day (weather permitting) with at least 15 minutes between each trial. Male colour pattern was measured using this setup for the three taxa in pairwise comparisons (*elevatus*/*butleri*, *elevatus*/*sergestus* and *butleri*/*sergestus*).

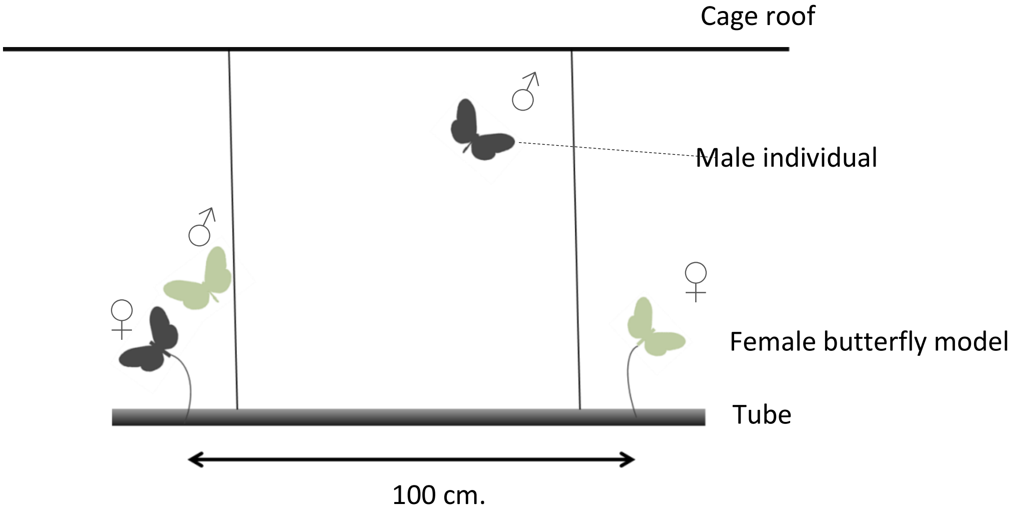

Figure 4. Representation of the experimental set up for the male colour pattern preference experiment. The different coloured butterflies represent two different taxa being tested simultaneously. The female models are fixed to a hanging tube in the experimental cage.

Courtship was measured as number of approaches (direct flight towards the model), hovers (static hovering within a 10 cm radius of the model) and alightments (alighting on the model) performed by the individual males towards the models (Klein and De Araújo, 2010).

Host plant experiment

Heliconius are known for their close association with the *Passiflora* species used for adult oviposition and larval feeding (Benson et al., 1975). Host plants differences can cause associated reproductive isolation, which has been shown in many species of phytophagous insects (Berlocher and Feder, 2002). In order to identify the host plant preference of the different taxa and potential differences between them, a preliminary experiment was carried out testing eight different species of *Passiflora*. Host selection was based upon availability and wild and captive reports of use by the taxa; these were P. *edulis*, *P. laurifolia*, *P. riparia*, *P. serrato-digitata, P. vitifolia*, *P. coccinea* and two undescribed species “misteriosa” and “linda”. The two unnamed species are likely part of the Granadilla section along with the *P. laurifolia* and *P. riparia* (Killip, 1938; Rosser, personal communications). No evidence of use of these unnamed species has been recorded but they are very commonly found within the taxa’s range (Rosser, personal communications). *Passiflora edulis* is the only species not found within the range of at least one of the three taxa. However, it is native to other regions of South America, is widely cultivated in Peru and was one of the plants used for rearing the larvae.

Fresh individual shoots of each *Passiflora* species were placed in separate bottles and spread out on a table in each taxon’s stock cage. Eggs were counted at the end of each day for 3 days (data not presented). *P. vitifolia*, *P. coccinea*, “misteriosa” and “linda” had a tendency to wilt in water and were not available in sufficient numbers for experimental use. No eggs were laid on these during this preliminary experiment (perhaps due to the wilting) and they were therefore removed from the experiment. The four other species (*P. edulis*, *P. laurifolia*, *P. riparia* and *P. serrato-digitata)* were used for oviposition by the three *Heliconius* taxa and consequently used for the host plant preference experiment.

Host preference was measured by counting the number of eggs laid on the four species tested. Fresh shoots considered of similar sizes and of high quality (presence of growing tip, young leaves and tendrils) were set up in separate bottles of water hung at each corner of the 1x2x2 m cage. Ants were stopped from reaching the eggs by placing a water trap on the wire from which the bottles were hung. Each day, a single female was placed in the cage with the four hosts. Eggs were counted and collected from each shoot without damaging the leaves or tendrils before putting in a new individual. Several females were tested simultaneously in different cages. Females were never tested in the same cage two consecutive days to remove variation from shoot quality and cage location. Some individuals were left in a cage without plants for oviposition for a day before being tested to increase the number of eggs collected in a single trial. To avoid egg laying preference being affected by the quality of shoots kept in water, new shoots were cut every 6 days. Three individuals, two from *H. elevatus* and one from *H. p. sergestus,* only laid one egg and were removed from the analysis as a single egg was not deemed to be representative of a preference.

During the experiment a trend appeared suggesting a strong difference in preference for *P. laurifolia* and *P. riparia* between the three taxa. For future experiments (see Chapter 3) I wanted to confirm this trend, without the variation caused by other hosts present. Whole plants of *P. laurifolia* and *P. riparia* hosts were therefore placed in each taxon's stock cage for 13 days. Each plant had more than one growing shoot to ensure quality did not influence results. Eggs were counted and removed from the plants without damaging it every day.

Results

Male colour pattern preference

To test for differences in male preference for colour pattern pairwise comparisons of the three taxa were performed measuring courtship behaviours (approach, hover and alightment) towards either the conspecific or heterospecific model. The total number of each courtship behaviour recorded and the number of individuals tested for each taxon are shown in figure 5. A likelihood analysis, used by Merrill et al. (2011b) in a similar experiment, was used to determine the likelihoods of each taxon courting their own pattern against the other two patterns separately in using Solver in MS Excel. The likelihood function:

$$\ln(L) = \Sigma\, m_i\, ln\left(P_j^1\right) + c_i \ln(1 - P_j^1)$$

Where m_i is the number of courtship behaviour performed towards one of the patterns and c_i the number towards the other pattern and P_j, the probability of the courtship behaviour being performed towards the first pattern.

A Generalised Linear Model (GLM) with binomial errors, was carried out to obtain values of significant differences in preferences for each pair (significance values in fig 5). The function cbind was used for numbers of the different courtship behaviours towards the two patterns as a response variable and male taxon as an explanatory variable. Differences within geographical distributions were also tested against each other in pairwise comparisons to determine whether sympatric taxa demonstrated more assortative than allopatric or parapatric taxa. The same was also done between the parapatric and allopatric taxa. This was done in a separate test where taxa were assigned the arbitrary names of "taxa 1" and "taxa 2" in each pair. Doing so, the interaction between species and pair of each geographic distribution, the two explanatory variables, could be tested.

We observed a significant difference for all three behaviours between the two sympatric species, *H. elevatus* and *H. p. butleri* (fig 5, a). While *H. elevatus* clearly demonstrates stronger preferences towards its own colour pattern in all behaviours, *H. p. butleri* only demonstrates a preference for its own pattern when alighting. In the parapatric comparison between *H. elevatus* and *H. p. sergestus* (fig 5, b), *H. elevatus* retains its conspecific preference mainly when hovering but shows no preference in alightments and only weak preference when approaching. *Heliconius p. sergestus* demonstrated no preference for either pattern. The two only differ significantly when hovering. Finally, the allopatric sub species demonstrate significantly different preferences when approaching and hovering (fig 5, c) however, no discrimination between the patterns is seen for *H. p. butleri* when approaching or for *H. p. sergestus* when hovering. No differences were found in the alighting behaviour, however, the sample size for each taxon is very small.

Testing for significance between different geographic pairs did not find strong evidence for more assortative mating in sympatric species pair. The only significant differences were found between the allopatric against the parapatric pairs ($p < 0.005$) and the parapatric and sympatric pairs ($p < 0.005$) when approaching.

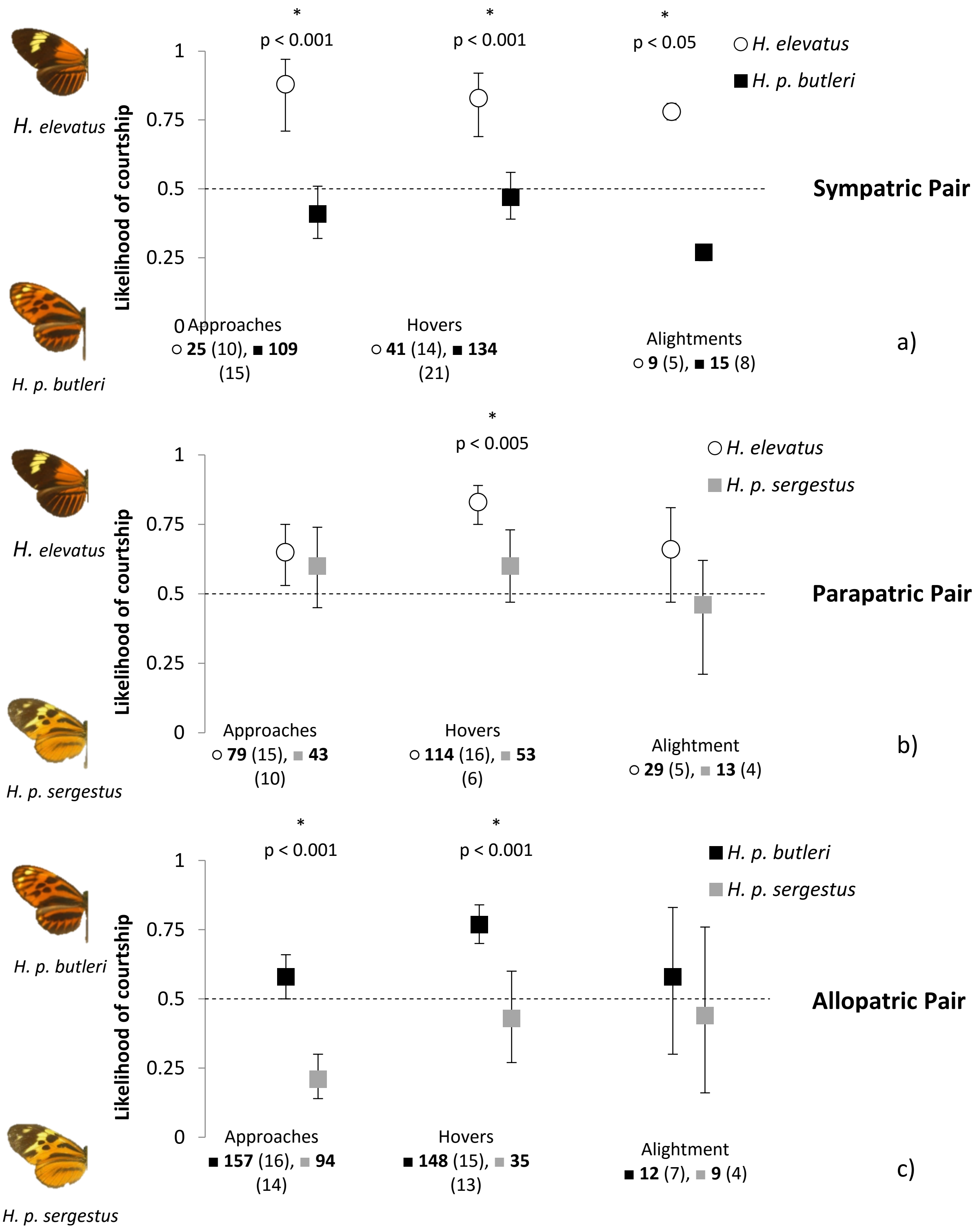

Figure 5. Likelihood of courtship behaviours towards a. the *H. elevatus* colour pattern by the two sympatric species, *H. elevatus* and *H. p. butleri*; b. the *H. elevatus* pattern by the two parapatric species, *H. elevatus* and *H. p. sergestus*; and c. the H. p. butleri pattern by the two allopatric sub-species, *H. p. butleri* and *H. p. sergestus*. The dashed line represents no preference between the two patterns. Error bars correspond to the 95 % confidence interval of the courtship likelihoods and p values were obtained using the GLM. Number of recordings of each courtship behaviour by the different taxa is shown in bold; the numbers in brackets are the numbers of individuals performing the courtship.

Host plant experiment

The host plant experiment aimed to identify differences in preference for host plants between the three taxa with differences in gene flow to identify evidence of reinforcement. The total number of eggs collected for *H. elevatus*, *H. p. butleri* and *H. p. sergestus* were 170, 106 and 150 respectively, with 9, 10 and 10 individuals respectively laying at least eight eggs.

In figure 6, the likelihoods of laying on each host by each taxon are presented. *Heliconius elevatus* demonstrated a strong preference for *P. laurifolia* compared to the other hosts, especially *P. riparia*, on which very few eggs were laid. On the other hand, *H. p. butleri* eggs were laid more evenly across the four hosts and there was no difference between number of eggs laid on *P. laurifolia and P. riparia*. *H. p. sergestus* had a similar pattern of preference to *H. elevatus*, though the difference in preference between *P. laurifolia* and *P. riparia* was less marked.

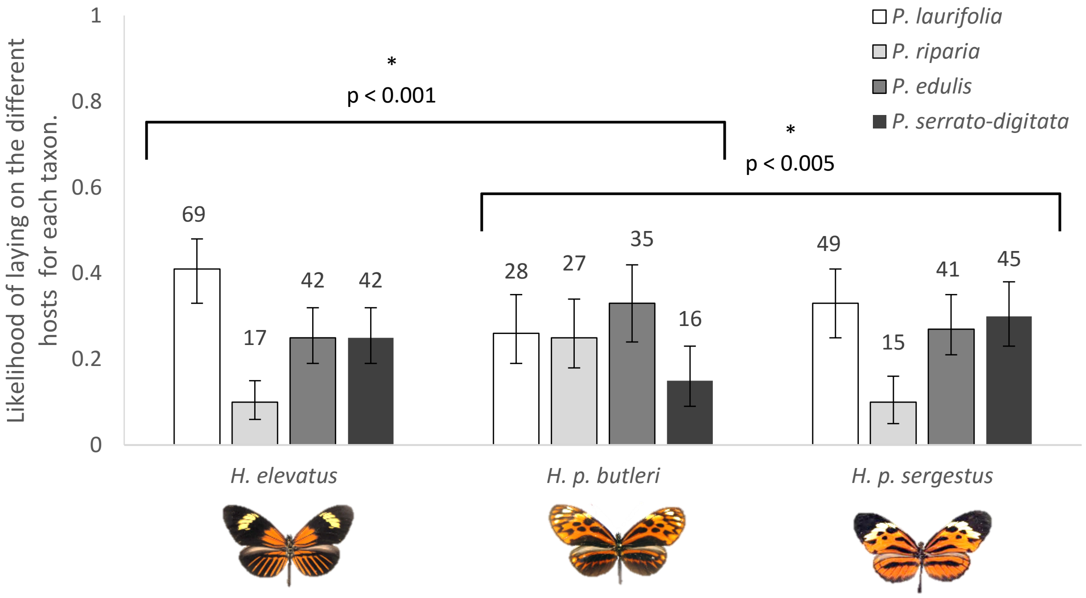

Figure 6. Likelihood of eggs being laid on the four *Passiflora* species by the three taxa. Error bars correspond to the 95 % confidence interval and the labels above are the total numbers of eggs laid on each host by each taxon. These data represent the oviposition preferences of 9, 10 and 10 individuals for *H. elevatus*, *H. p. butleri* and *H. p. sergestus* respectively. The p values show a significant effect in the interaction between taxon and host species on the number of eggs laid determined with the Generalised Linear Mixed Effect Model (there were no significant differences between *H. elevatus* and *H. p. sergestus*).

The number of eggs laid on the different hosts was statistically analysed in R to identify differences in host preferences between each pair of taxa. Pairwise comparisons of the taxa were carried out using the total number of eggs laid per individual as the response variable. A Generalized Linear Mixed Effect Model (GLMM) for negative binomial (to control for overdispersion) fit by maximum likelihood (Bates et al., 2015) was used to analyse the differences in number of eggs laid on each plant, using individuals as a random factor. Two nested models were fitted for each of the three pairwise comparisons (*elevatus*/*pardalinus*, *pardalinus*/*sergestus* and *elevatus*/*sergestus*). In Model 1 (M1), host species and butterfly taxon were used as fixed effects. In Model 2 (M2) the interaction between these fixed effects, which shows whether the butterfly taxa being tested differ in their host plant preferences, was included. Models were tested against one another using ANOVA and the Akaike information criterion (AIC) was used for model selection to identify the significance of the interaction.

Once the best model had been determined (including the interaction between host plant and butterfly taxa), the same method was used to determine whether the level of difference between taxa was different between the pairs of taxa from different geographic distributions. As for the colour pattern preference experiment, this was done in pairwise comparisons naming the taxa in each pair “taxa 1” and “taxa 2”. Pairs from two different geographic distributions were added as an explanatory variable and two models were tested against each other, M1 without the interaction of pairs and M2 including the interaction.

In the sympatric comparison between *H. elevatus* and *H. p. butleri* a significant difference was found between the two models ($p < 0.001$; $AIC_{(M1)} = 541{,}62$; $AIC_{(M2)} = 522{,}97$), suggesting the interaction between butterfly taxa and which host plant they lay on is important to the model. No difference was observed between the two models comparing the parapatric species *H. elevatus* and *H. p. sergestus* ($p = 0.48$, $AIC_{(M1)} = 569{,}27$; $AIC_{(M2)} = 565{,}73$). We found that the interaction also had a significant effect in the allopatric sub-species pair, *H. sergestus* and *H. p. butleri* ($p < 0.005$, $AIC_{(M1)} = 558{,}66$; $AIC_{(M2)} = 549{,}21$). Significant interactions between taxa are highlighted in figure 6.

The interaction between pairs from different geographic distributions and between taxa host preference was found to be significant between the sympatric and allopatric pairs ($p < 0.001$; $AIC_{(M1)} = 1089.7$; $AIC_{(M2)} = 1073.7$) and between the sympatric and parapatric pairs ($p < 0.05$; $AIC_{(M1)} = 1109.1$; $AIC_{(M2)} = 1107.5$). No significance was found for the interaction between the allopatric and parapatric pairs ($p = 0.095$; $AIC_{(M1)} = 1111.9$; $AIC_{(M2)} = 1113.9$).

A side experiment was carried out to test the difference between *P. laurifolia* and *P. riparia,* without variation from other hosts, to determine if this preference can be used for genomic analysis (see chapter 3). Eggs were laid by females present in the taxa's stock cages; a t-test was used to compare the number of eggs between the hosts within each butterfly taxon. The results (fig 7) indicate that *H. elevatus* lays significantly more eggs on *P. laurifolia* than on *P. riparia* (t = -4,61, d.f. = 21.21, p <0.001) as does *H. p. sergestus* (t = -3.3, d.f. = 13.21, p <0.01). No difference was found for *H. p. butleri* (t = 0.71, d.f. = 23.78, p <0.48).

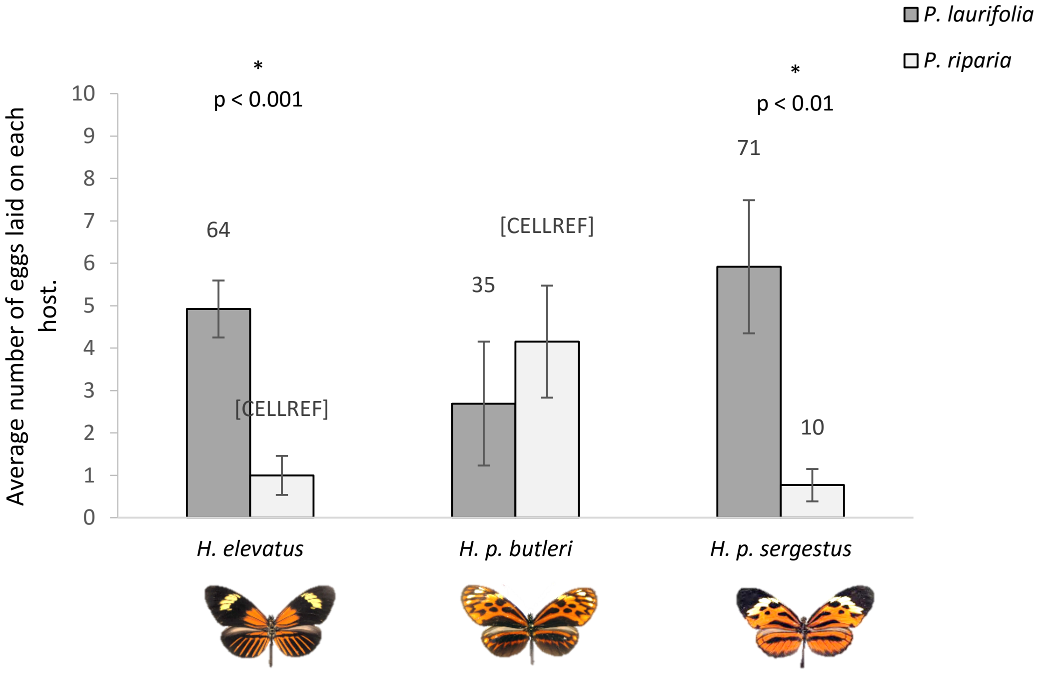

Figure 7. Average number of eggs laid on *P. laurifolia* and *P. riparia* by the three taxa in their respective stock cages, over 13 days. The error bars represent the standard error and the labels above are the total numbers of eggs laid on each host. The p values were determined using a simple t-test.

Discussion

Much of the debate on species divergence and maintenance despite ongoing gene flow now focuses on finding examples of this in nature. In sympatry, reinforcement will increase the pre-zygotic barriers between populations, races or species in response to selection against maladapted hybrids. Therefore, stronger differences in traits acting as reproductive barriers in sympatric populations than in allopatric populations can be used as evidence for reinforcement and in turn for gene flow.

Male colour pattern preference

The male colour pattern preference experiment identified evidence for assortative mating between the three taxa. The strongest difference in colour pattern preferences were observed between the sympatric species, *H. elevatus* and *H. p. butleri.* This suggests stronger reproductive barriers have evolved in response to gene flow, supporting the role of reinforcement in the speciation of these taxa. Although there is a pattern of stronger reproductive barriers in sympatry, this was not supported by the test investigating a difference between pairs from different geographic distributions as only the parapatric pair was found to be different to both the allopatric and sympatric pairs. This could in part be due to small sample sizes and large standard deviation which exacerbate the effect of an already weak trend. Some evidence still remains with *H. elevatus* showing overall stronger preferences for its own colour pattern in the sympatric comparison. Furthermore, other evidence for stronger assortative mating in sympatry using male colour pattern preference has previously been found between the sympatric species *H. melpomene* and *H. cydno* from Panama and an allopatric population of *H. melpomene* from Guiana (Jiggins et al., 2001). Evidence for reinforcement was also observed in assortative mating experiments using live females in three taxa *H. pachinus*, *H. melpomene* and *H. cydno* (Kronforst et al., 2007).

The data therefore suggests a stronger conspecific preference from a taxon in a sympatric comparison compared to an allopatric comparison. Other processes, aside from reinforcement, could have led to a similar pattern, such as differential fusion and runaway sexual selection (see Chapter 1). These do not account for a lack of post zygotic barriers (Coyne and Orr, 2004), and fertile hybrids of *H. elevatus* and *H. p. butleri* are obtained after crosses (see chapter 3). However, for differential fusion, it can be argued that a lack of post-zygotic isolation arises after homogenisation due to gene flow in sympatry or that pre-mating isolation evolves quicker and therefore strong post-zygotic isolation is not as necessary (Boake and Gavrilets, 1998). To disprove differential fusion the levels of isolation in sympatry would have to be demonstrated as being distinct to the levels observed in allopatry (Noor, 1999). This was found in the study on 119 pairs of *Drosophila* carried by Coyne and Orr (1989). Therefore, to correctly disprove differential fusion a study of similar magnitude would be necessary, though not feasible in the context of this thesis. Even so, it is possible that even this study may fail to accurately determine whether isolation in sympatry is a subset of isolation in allopatry if fusion of populations with low levels of isolation is common and allopatric populations with strong pre-zygotic isolation rare (Noor, 1999). Nevertheless, *H. elevatus* and the two *pardalinus* sub species belong to separate mimicry rings, the first to the dennis ray and the other two to the silvaniform mimicry ring. Colour pattern is under strong natural selection, therefore hybrids of *H. elevatus* and *H. p. butleri* with intermediate colour patterns would be selected against which is a necessary component of reinforcement; and although we cannot fully disprove an effect of differential fusion, it seems unlikely this would be the only effect causing the levels of higher levels of pro-zygotic isolation observed (Noor, 1999).

In this experiment, the difference observed between the sympatric species, *H. elevatus* and *H. p. butleri*, is due to a very strong *H. elevatus* preference for conspecifics. The *H. p. butleri* males' likelihood of courtship (for approaches and hovers, alightments are excluded due to the small sample size) do not differ significantly from the 0.5 line which could suggest that there is asymmetrical assortative mating, and that gene flow could still occur through *H. p. butleri* males. In other assortative mating experiments on *Heliconius*, strong, symmetrical assortative mating (both species in a pair show conspecific preference) in a sympatric and a parapatric species pair was found (Merrill et al., 2011a). However, these experiments were not focusing specifically on the effect of colour pattern and live females were used, which could elicit stronger responses from males due to other factors (such as pheromones or behaviour).

In colour pattern experiments on *H. melpomene*/*cydno* (Jiggins et al., 2001), both species in the pair demonstrated preference for the conspecific colour pattern. However, *H. cydno* approached both patterns equally and while it courted (hovered) its own colour pattern preferentially, there was a lot of variation in male preference. Therefore, the pattern observed here, in which only *H. elevatus* has a strong preference, does not necessarily imply an incomplete barrier for colour pattern. Furthermore, some evidence suggests females might exert some choice on males for mating (Merrill et al., 2011b). Therefore, even though *H. p. butleri* males does not seem to demonstrate a preference, a barrier could still occur through female choice in *H. elevatus*.

Heliconius butterflies are also known for being attracted to red, the colour of their preferred flowers for pollen feeding, *Psiguria* and *Gurania* (Merrill et al., 2015). In the colour pattern study carried out by Jiggins et al. (2001), it was suggested that the red colour on models could create a similar attraction. This could conceivably influence male attraction to the *H. elevatus* colour pattern which has a prominent red colouration in my own experiment. This general attraction to the colour red could potentially affect butterflies in the wild, causing inter-specific courting. However, other cues would influence the mating outcome, such as pheromones or behaviour.

There are also issues with data collection which could have affected the results, especially when measuring courtship behaviour. For example, approaches can sometimes be difficult to discriminate from other flights which can be very erratic in cages. Hovering was also difficult as hovers over the model could last anywhere between a couple of seconds to over a minute, but each was counted as a single hover. Measuring the time spent courting as opposed to the number of courtships could be a good alternative (Jiggins et al., 2001). Many variables can also affect butterflies being tested such as sunlight or wind forcing the individuals to one side of the cage. Nevertheless, the experiment found some evidence of preference for conspecific patterns, which is consistent with other studies.

The conspecific preference observed in *H. elevatus* could have occurred through adaptive introgression of *H. melpomene* alleles from which *H. elevatus* inherited its colour pattern (Heliconius Genome Consortium, 2012). Assortative mating in *H. m. rosina/amaryllis* (sub species of *H. melpomene* with postman colour pattern) has been shown to be strongly associated to the *B* locus (Merrill et al., 2011b) which is linked to the *D* locus controlling the dennis pattern (Mallet, 1989) found in *H. elevatus*. Therefore, there could have been genetic hitchhiking of a locus controlling a preference trait along with the colour pattern. Natural selection on the colour pattern could therefore have driven this preference in *H. elevatus*. Sympatry could have driven the stronger conspecific preference observed *H. p. butleri* when tested against *H.* elevatus, compared to the conspecific preference in the parapatric *H. p. sergestus*, when tested against *H.* elevatus*.*

Colour pattern has also been suggested to function as a magic trait (Servedio et al., 2011). The data from these experiments support this idea as assortative mating was found in response to colour pattern, which itself is likely under strong divergent selection. Associations between these traits have been found, though whether this is due to tight physical linkage or pleiotropy is still unsure (Merrill et al., 2015). Magic traits can play an important role in speciation as they are more resistant to recombination. If colour pattern is indeed a magic trait, this could have played an important role in the preference observed in *H. elevatus*, as it would remove the need for genetic hitchhiking mentioned above.

Although the data suggests some role for colour pattern as a reproductive barrier, other barriers are necessary to create isolation needed to maintain species differences, especially as this role was not shown to be very strong. Therefore, other traits must be considered to explain the maintenance of the species' identity in sympatry.

Host plant experiment

Many phytophagous insects and their host plants are known to have tight coevolutions, which is the case for *Heliconius* and species of *Passifllora* (Benson *et al.,* 1975). Males of some species of *Heliconius,* are known for visiting their *Passiflora* hosts sequentially in search for mates (Mallet, 1986), and others even uses hosts to find pupae to perform pupal mating (Deinert et al., 1994). Though it is not known whether the species studied in this thesis share these behaviours it does suggest that host plant can play an important role as a reproductive barrier. Furthermore, whether hosts are explicitly used for mate searching, individuals using the same hosts are more like to encounter one another by chance (Rosser *et al.*, 2015).

The results of the GLMM demonstrated a strong significant difference in the number of eggs laid on the different hosts in the sympatric (*elevatus*/*butleri*) pair, a weaker significant difference in the allopatric (*butleri*/*sergestus*) pair and no difference between the parapatric (*elevatus*/*sergestus*) pair. The second GLMM testing for differences between pairs of different geographic distributions found the strongest significant difference between the sympatric and allopatric pairs, a significant difference between the allopatric and parapatric pairs and no significant difference between the parapatric and allopatric pairs which fits the pattern of stronger isolation in sympatry expected under reinforcement. The taxa's preference for two of the hosts was also investigated in the second experiment. *Heliconius elevatus* has a strong significant preference for *P. laurifolia* over *P. riparia*, which was not observed for *H. p. butleri; H. p. sergestus* demonstrates the same preference as *H. elevatus*, however the difference is less pronounced. Therefore evidence suggests stronger differences in the sympatric species, consistent with predictions of reinforcement.

Stronger differences in the parapatric (*elevatus*/*sergestus*) pair compared to the allopatric (*butleri*/*sergestus*) could have been expected as there should be intermediate levels of gene flow. However, the preference for *P. laurifolia* could be an ancestral trait from which *H. p. butleri* has diverged from, due to sympatry with *H. elevatus*. This could have led *H. p. butleri* to also diverge from *H. p. sergestus,* a process known as "reinforcement cascade" (Nosil, 2012), explaining the stronger difference in allopatry than in parapatry. The cascade effect illustrates how this divergence between two sympatric taxa due to reinforcement can cause divergence with a third allopatric taxon as a result. The similarities between *H. elevatus* and *H. p. sergestus* could also explain the exclusion of *H. p. sergestus* from the geographical distribution of the other taxa if there was competition for resources.

Assessing the role of a diverging trait between two taxa as a reproductive barrier is not straightforward. In the case of host plant use in *Heliconius,* populations within a species may vary in their host preference, making the determination of species preference difficult. Furthermore, the species do not demonstrate strong specialisation in host use, particularly *H. p. butleri*, which did not show a strong preference for any particular host. This amount of variation weakens the effect host plant can have as a reproductive barrier. Diversification of phytophagous insects however, is often associated to this plasticity in host use. A study of the phylogeny of nymphalid butterflies looked at the occurrence of polyphagy and suggested that polyphagous stages could be transient to more specialisation (Nylin et al., 2014) and maybe speciation. More host species would also need to tested; in my experiment the availability of plants was limiting and some species could only be kept in soil. The species tested here may therefore not have been the most appropriate hosts for this system. However, data collection on host plant use of Heliconius species in the wild is very difficult. My experiment remains a reliable and rigorous method to gather information on host preference. Furthermore, it is assumed reproductive isolation is not achieved with a single barrier, but several will interact to reduce gene flow.

Other factors may have influenced the differences observed in the host preference. As described in Merrill et al. (2013), which did not attribute host difference in *H. melpomene* and *H. cydno* to reinforcement; competition with other *Heliconius* can also select for resource partitioning. Competition for resources, cannibalism of eggs by larvae and competition for oviposition can have an influence on host selection (Gilbert et al., 1991). However, this study only investigated preferences for a sympatric pair without the comparison to an allopatric population. The pattern of stronger differences in sympatry than in allopatry in the taxa studied here is still consistent with the predictions of reinforcement, though this may be interacting with other factors. Furthermore, different conditions for different species pairs will influence the outcome of contact with other species. Genetic associations were found between host plant and colour pattern traits in hybrids of *H. melpomene* and *H. cydno* (Merrill et al., 2013). If similar associations are present in *H. elevatus* and *H. p. butleri* then differences in host plant use could have evolved in response to selection on the colour pattern

Evidence therefore suggests that host plant is acting as a reproductive barrier with stronger differences in the sympatric pair compared to the allopatric and parapatric pairs. This is evidence of reinforcement with suggests divergence with gene flow.

Conclusion

The results of the experiments looking at differences in traits as evidence for reproductive barriers suggest that divergence has occurred, and the evidence for divergence is stronger between the sympatric species. This is evidence for reinforcement therefore implying that divergence has occurred in the face of gene flow, although further work on more taxa verifying that levels of isolation in sympatry are distinct to levels observed in allopatry would be necessary to disprove an effect of differential fusion. There is also some divergence observed between the allopatric sub-species, which can be attributed to their allopatric distributions or to a cascade effect. These reproductive barriers are expected to interact to increase overall isolation between the species. Studies investigating the association between different loci have found that increasing loci that cause divergence can lead to an overall increase in genomic divergence without necessarily being physically linked (Flaxman *et al.*, 2014). However, it has been argued that migration modification, which in this case could occur due to a host change, could limit assortative mating through direct competition in a sympatric context (Yukilevich and True, 2006) which, if the case in this system, could explain the weaker trend observed for the colour pattern preference.

A cryptic effect of small scale habitat segregation could also be causing the stronger differences in sympatry. *Heliconius elevatus* and *H. p. butleri* share the same geographic distribution, however, gene flow could be limited if the species remain in separate habitats, essentially creating allopatric conditions or “micro-allopatry” (Berlocher and Feder, 2002). However, butterflies have a wide dispersal ability and there are known sites where both species can be found, therefore it seems unlikely that the observed differences would be caused only by micro-allopatry in parts of their range, if it is present.

The colour pattern preference and host plant preference have been identified as reproductive barriers between the sympatric species. However, very strong reproductive isolation is observed between the sympatric species determined by the assortative mating experiment (Segami Marzal, 2015; Velado, 2015) and the limited number of hybrids found in the wild. Therefore, other traits would need to be investigated as potential barriers, such as pheromones, flight pattern or behaviour, to assess their role and interactions in reproductive isolation. Another important aspect of these barriers is their underlying genetic architecture. Understanding which genes express traits causing reproductive isolation and where they are relative to each other can help understand the mechanisms of selection acting upon these traits and whether they have played a role in the speciation process of the taxa.

Chapter 3

Quantifying the phenotype of traits involved in reproductive isolation in F2 hybrids for genomic analysis.

Introduction

A major focus of evolutionary biology has been to understand the modes of speciation, and their relative contributions, which have led to the incredible diversity observed today. The main limitation to this however, is that speciation is a very slow and gradual process and cannot be observed in the wild in a short timescale. Studies have had to focus on genomic, ecological and behavioural patterns of sister species to understand the past processes involved in their divergence (Nosil, 2012). Research has also been carried out on divergent populations, taxa supposedly in “the process” of speciation. However, without confirmation that speciation will indeed occur, these studies can only go so far towards demonstrating the importance of different modes of divergence in the speciation process (Coyne and Orr, 2004).

To study speciation it is therefore important to define what a species is to accurately determine when speciation has occurred (Templeton, 1989). There are various definitions of species used in different fields (Howard and Berlocher, 1998), however, I will focus on two. The Biological Species Concept (BSC), as defined by Mayr (2000), delimits species into cohesive groups of interbreeding organisms, reproductively isolated from other such groups. This concept, though widely used, has been disputed as it considers reproductive isolation as the end point of speciation (Mallet, 2005). The Genotypic Species Cluster (GSC) therefore defines species as “distinguishable groups of genotypes that remain distinct in the face of potential or actual hybridization and gene flow” (Mallet, 2007). Species are no longer discrete and constant units of diversity but continuous groups with varying degrees of gene flow. The BSC also does not consider the genome to be a single diverging unit, which allows for introgression (Wu 2001). This follows the Darwinian view of species and species adaptation, where reproductive isolation occurs as a response to divergence and not as a pre-requisite (Mallet, 2007; Wu, 2001). The GSC is widely adopted in *Heliconius* research (Merrill et al., 2015) and is the definition used in this thesis.

Speciation was originally considered in the context of a continuum of geographic range overlap (Mayr, 1963). The two extremes of this continuum are allopatric speciation, where two populations evolve isolated by geographic barriers, and sympatric speciation, where speciation occurs in two populations with overlapping distributions. Parapatric speciation occurs when two populations are in contact, but do not overlap (Coyne and Orr, 2004). However, speciation research has moved on to consider these different modes in the context of gene flow, where allopatric speciation occurs with no gene flow and where sympatric speciation occurs in a population with complete gene flow (Nosil, 2012).

Allopatric speciation is the most widely accepted geographic mode of speciation, in which geographically isolated groups diverge through drift and adaptation in different habitats (Coyne and Orr, 2004). Though there is theoretical and empirical evidence that allopatric speciation can lead to reproductive isolation through adaptation (Rice and Hostert, 1993) and sometimes drift, some argue that it should not be considered a default against which other modes of speciation can be tested against (Coyne and Orr, 2004; Bolnick and Fitzpatrick, 2007). The reasons for this are that complete isolation, instead of a continuum of levels of isolation, is probably not common and allopatric speciation is difficult to falsify, therefore its prevalence may be overestimated.

Theoretical studies (Kondrashov and Mina, 1986) and evidence in nature (Nosil, 2012) have sparked a new interest in sympatric speciation. The main argument against sympatric speciation is that recombination will break down favourable co-adapted gene combinations. Doing so, recombination will inhibit correlations between adaptive traits under divergence and prevent the formation of species (Felsenstein, 1981). However, mechanisms such as pleiotropy, tight physical linkage and chromosomal inversions can reduce the amount of recombination in parts of the genome (Nosil, 2012). These associations between adaptive traits can be selected for, further reducing the effects of recombination.

In the early stages of sympatric speciation, adaptive regions of the genome are therefore under divergent selection while gene flow may homogenises the non-adaptive regions, creating a “genetic mosaic” (Via and West, 2008), also known as “islands of divergence”. In allopatry, the entire genome is isolated by a geographic barrier therefore loci causing divergence can accumulate uniformly (Via, 2012). Looking at the underlying genetic structure of traits under disruptive selection and the overall genomic patterns in sympatric species can therefore help identify whether loci for reproductive barriers diverged in the presence of gene flow.

System

The system comprises of two species, *Heliconius elevatus* and *Heliconius pardalinus butleri*, which have overlapping geographic distributions (see fig 2, Chapter 1). Over 95 % of their genome shows no evidence of divergence but the species do exhibit different colour patterns (Heliconius Genome, 2012). Evidence suggests that *H. elevatus* received its colour pattern from another species, *H. melpomene* (see fig 3, Chapter 1), with the same dennis ray pattern through hybridisation (Heliconius Genome, 2012). Two loci known for controlling colour pattern have been identified within peaks of divergence in the genome.

Despite their shared genome and geographic distribution, there appears to be pre-zygotic isolation between the two species as there have been very few records of natural hybrids in the wild. Furthermore, a no choice experiment was carried out to quantify this isolation testing 14 *H. elevatus* and 16 *H. p. butleri* individual females. Individual females were left for three days (or until mating) with males of the two species and did not result in any hybrid mating (Segami Marzal, 2015; Velado, 2015). Therefore, strong pre-zygotic barriers exist between the species.

In chapter two of this thesis I demonstrated evidence of strong differences between the two sympatric species for two traits, male colour pattern preference and female host plant choice, which act as reproductive barriers. In this chapter, I aim to accurately measure these traits in their F2 hybrid cross to identify qualitative patterns of inheritance and associations with colour pattern traits. The measurement of F2 individuals can then be used in the future for Quantitative Trait Loci (QTL) analysis to determine whether these traits are found in the known peaks of strong divergence of the genome. This would strongly support the pattern of islands of divergence in the presence of gene flow which could be used as evidence for sympatric speciation.

Methods

Collecting, rearing and crossing

Butterflies were kept, collected and reared in the same manner as described in Chapter two. Crosses between species were obtained by mixing males and virgin females from each taxon in insectaries and checking for matings every hour. Additionally, some matings were obtained by “hand-pairing” which involves holding a male and female together until the male claspers have a tight grip on the female and then leaving the pair to mate. Females from the F1 generation were mated to F1 males to generate broods of F2 hybrids (fig 8).

Figure 8. Hybrid mating between *H. elevatus* and *H. p. butleri* (to the left); hybrid cross, F1 generation (centre) and hybrid cross, F2 generation (to the right).

<u>Quantifying colour pattern preference in the F2 cross.</u>

To analyse the genetic structure of a trait, the phenotype needs to be precisely quantified in the pure species and their F2 cross. After determining colour pattern preferences between *H. elevatus* and *H. p. butleri* (see Chapter 2), the same colour pattern experiment was carried out on their F2 cross for genomic analysis. Individual males were placed in 2x2x2 m cages and left to acclimate for at least one day. Models made of cleaned female wings of the pure species were presented to the males simulating flight. Approaches, hovers and alightments were recorded for each male to the different colour patterns as done for the experiment testing the pure species (details in Chapter 2).

<u>Quantifying the host plant preferences in the F2 cross</u>

Genomic analyses require clear, quantifiable phenotypic traits differing in the species studied. We therefore decided to simplify the four host plant experiment described in Chapter 2 by only using two hosts, as a binary trait is easier to phenotype and analyse statistically. In the previous experiment testing four *Passiflora* species, the strongest difference in preference was found for *P. laurifolia* and *P. riparia* between the *H. elevatus* and *H. p. butleri*. This pattern was confirmed with a second experiment using only these two species where a specimen of each was left in the butterfly species' stock cages, therefore results do not show individual preferences. *Heliconius elevatus* and *H. p. butleri* individuals laid 77 and 89 eggs respectively, and *H. elevatus* individuals laid significantly more on *P. laurifolia* than on *P. riparia*, no difference was observed in *H. p. butleri* (see Chapter 2).

We therefore carried out an experiment to quantify individual preferences for these two hosts for *H. elevatus*, *H. p. butleri*, and their F1 and F2 hybrid crosses. Although unfertilised females can lay eggs, females from the F2 generation were mated with closely related species of the *Heliconius* genus depending on availability to increase egg laying rate. Potted plants with a single growing shoot of both *P. laurifolia* and *P. riparia* were placed at opposite corners of a 1x2x2 m cage. Plants were manipulated to have a single growing shoot considered of "good" quality (both new leaves and tendrils were available for oviposition). Individual females were put in one of the three experimental cages and left to lay. Each day, females were either removed or moved to another cage if more eggs were needed. Regardless of the number of eggs already laid, each female was tested until at least one egg was laid in the three different cages to control for inter-cage plant quality. Eggs were collected from each host without damaging leaves or tendrils. A total of 106, 269, 120 and 268 eggs were laid by 12, 21, 11 and 19 females of *H. elevatus*, *H. pardalinus*, the F1 and F2 generations respectively.

Identifying genetic association between traits

I was interested in identifying genetic associations between traits under divergent selection (through pleiotropy, physical linkage, inversion or other genetic mechanisms). However, without genomic data to identify regions controlling these traits, other methods using phenotypic traits controlled by known loci can provide preliminary insight into these associations. Genetic data was inferred from colour pattern phenotypes from wings of the pure species and their F2 cross (Cama, 2015). Three loci controlling for the dennis pattern, the hindwing rays and the forewing yellow band respectively were identified as having a single Mendelian segregation pattern (Cama, 2015). The two pure species' traits are controlled by homozygous alleles with the recombinant demonstrating incomplete dominance and an intermediate pattern for each trait. Hybrid F2 individuals were therefore recorded as having the *H. elevatus* homozygous pattern (EE), the *H. p. butleri* pattern (PP) or the heterozygous pattern (EP) for each locus.

Results

Quantifying colour pattern preference in the F2 cross.

The experiment aimed to accurately quantify individual preference for the two pure species colour patterns in their F2 hybrid cross. A total of 44 F2 individuals were tested resulting in 37, 39 and 11 individuals performing at least one approach, hover and alightment, leading to a total of 580, 606 and 88 behaviours recorded, respectively. The proportions of courtship behaviours were calculated and plotted on graphs for each individual (fig 9) with the likelihoods of courtship and confidence intervals determined for the pure species in Chapter 2. Data for alightments was not plotted due to the small number of individuals performing this behaviour.

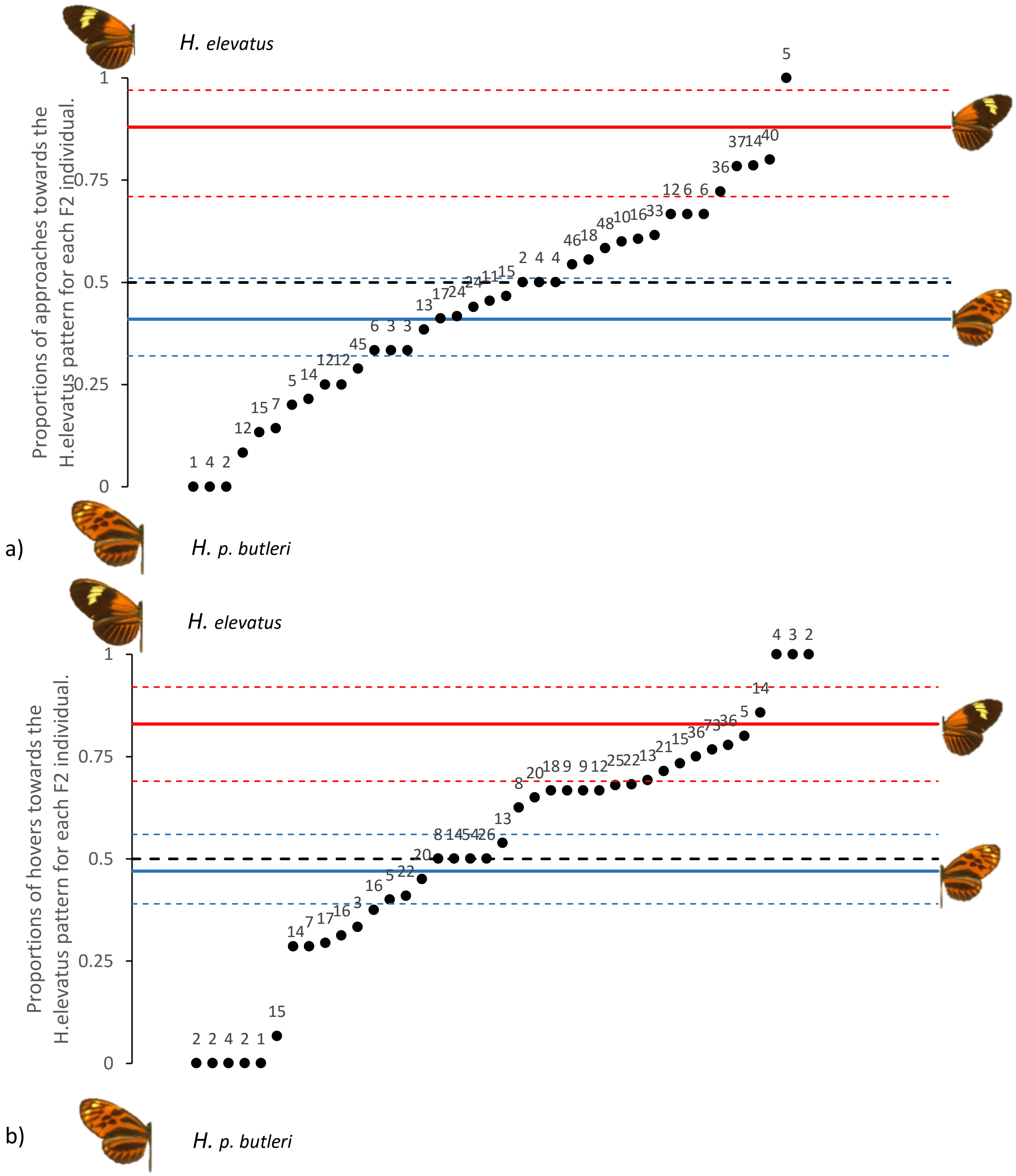

Figure 9. F2 individual proportions of (a) approaches and (b) hovers performed towards H. elevatus (value of 1) and H. p. pardalinus (value of 0). Each data point represents an F2 individual and figures above each point are the number of behaviours performed for that individual. The likelihoods of courtship towards the two colour patterns for each behaviour determined in Chapter 2 for *H. elevatus* (in red) and *H. p. butleri* (in blue) are represented by the bold lines, the dotted lines represent the 95 % confidence intervals.

To acquire an initial understanding of the underlying genetic mechanism of a trait we looked at the phenotypes present in the F2 generation. Some variation will occur due to variation in the pure species, that do not demonstrate complete preferences for a single pattern, and to brood specific effects. If a trait is controlled by a single locus then under a simple model of Mendelian inheritance two or three groups are expected within the F2 population. These would express either of the pure species' phenotype or an intermediate depending on the dominance or co-dominance of the alleles. However, if the trait is controlled by many loci than the phenotypes observed in the F2 cross will likely form a continuum between the two pure phenotypes. The spread of F2 individual preferences is presented in figure 10 (a and c). In order to identify underlying groups within the F2 population we carried out a cluster analysis on the individuals' phenotypes (alightments were not analysed to do the small number of data points).

The cluster analysis will give the total variance (V) explained by separating the data into k non-overlapping clusters. We therefore calculated V for 1 to 10 clusters to identify the cluster k where k+1 does not explain more variance (Hothorn and Everitt, 2014), known as the "elbow method". This was carried out on the full dataset for approaches and hovers and separately for the datasets including only proportions measured from at least 5 observations as data points with less are not accurate (fig 10, c and d).

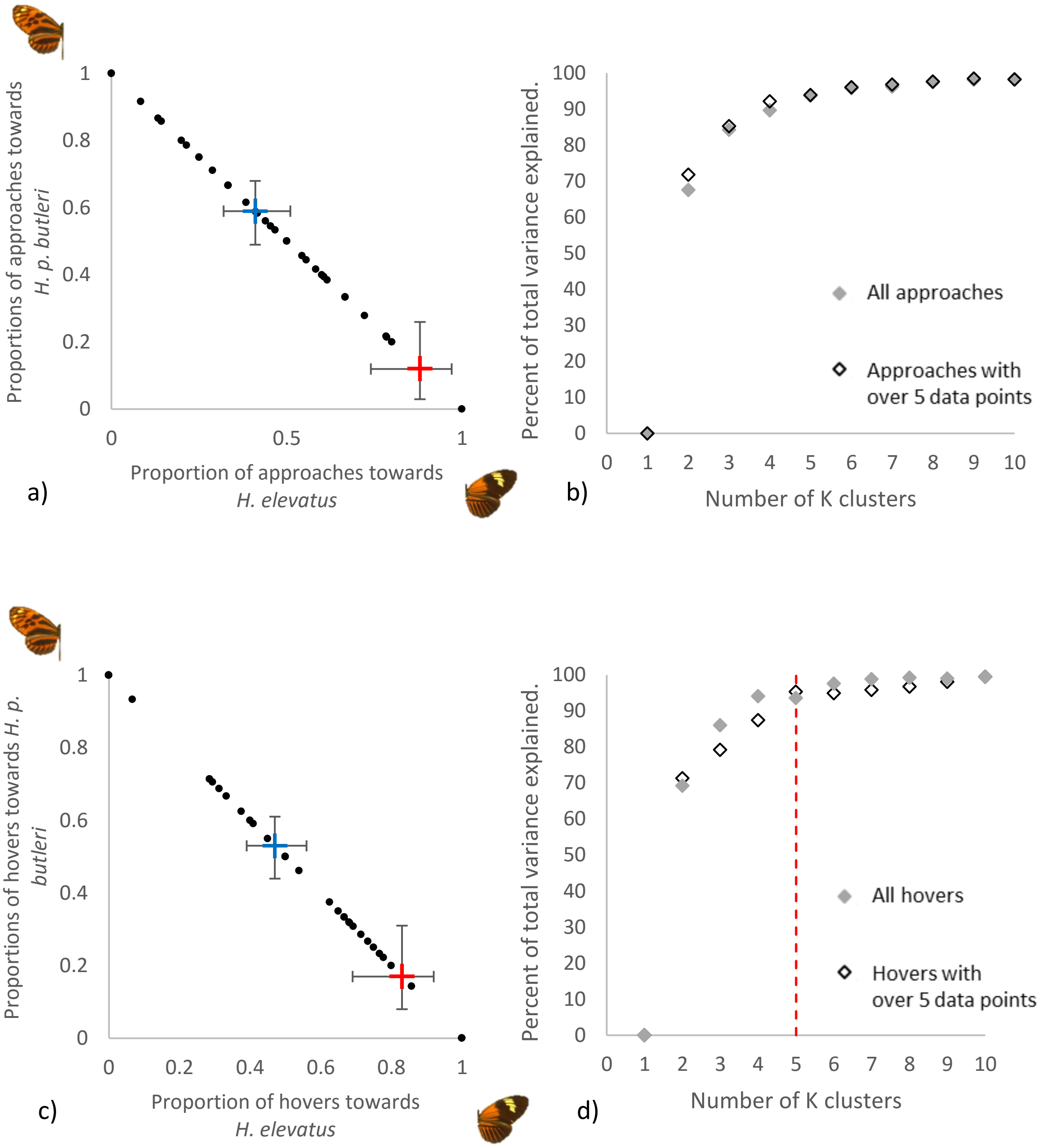

Figure 10. Spread of F2 individual preferences for (a) approaches and (c) hovers. Each data point represents an individual. The crosses represent the likelihoods for each behaviour determined in Chapter 2 for *H. elevatus* (in red) and *H. p. butleri* (in blue); the error bars are determined from the 95 % confidence intervals. The corresponding variance values over K clusters (b. approaches and d. hovers) are presented for all data and data with over five data points. The dashed red line demonstrates the number of clusters determined from the cluster analysis.

When looking at the spread of the proportions for approaches there is no obvious cluster (fig 10, a). When testing the V over k for approaches, no obvious "elbow" is observed even when removing data points with a small sample size (fig 10, b). For hovers however, there seems to be a less even spread of proportions (fig 10, c). When removing data points with less than 5 observations a trend appears suggesting $k = 5$ is adequate to explain the data (fig 10, d). Therefore, there is no evidence for clustering for approaches and some evidence for 5 distinct clusters for hovers which suggests a non Mendelian inheritance of the traits.

Quantifying the host plant preferences in the F2 cross

The host plant experiment from Chapter 1 was repeated using only two hosts identified in previous experiments, *P. riparia* and *P. laurifolia*, for an accurate identification of a difference between the pure species. This difference was measured for in the pure species and their hybrids F1 and F2 crosses. The likelihoods of laying on *P. laurifolia* for each female were calculated along with the 95 % confidence intervals (Merrill et al., 2011b) using MS Excel. These are presented in figure 11 along with the total number of eggs laid by each. Data for the pure species tested for the two plant experiment in the taxas' stock cages containing multiple individuals (experiment Chapter 2) was included as a comparison.

Looking at the likelihoods (fig 11), no differences were found between the pure species' preferences tested in this experiment with individual data and in the experiment in the stock cages, confirming the consistency of the species' preferences. However, a difference was found between the pure species in each experiment confirming the divergence observed in Chapter 2. The F1 population's preferences seem different to both pure species suggesting the trait is not controlled by a single locus with a simple dominant allele but possibly codominant alleles or multiple loci. The same was observed for the F2.

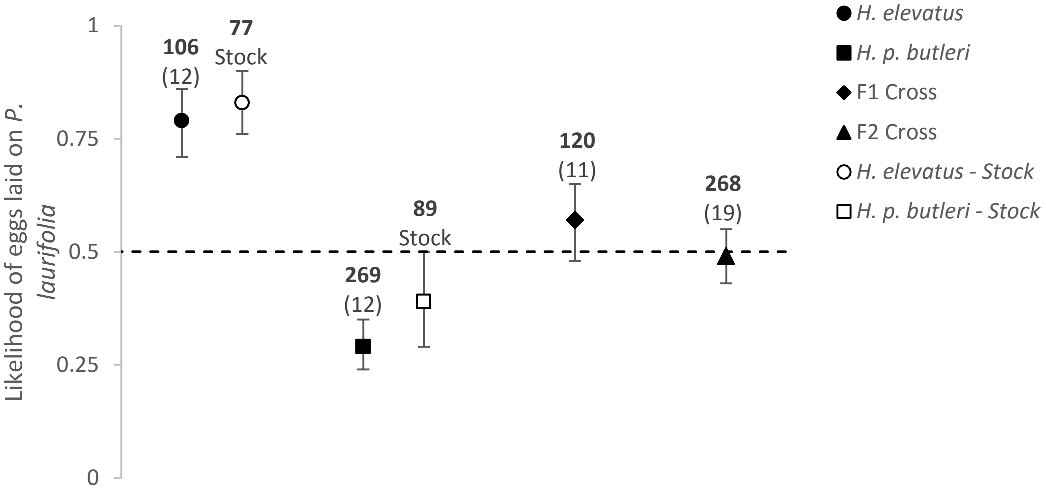

Figure 11. Likelihoods of eggs laid on *P. laurifolia* (value of 1) against *P. riparia* (value of 0) for the different taxa measured. Error bars represent the 95 % confidence intervals, the total numbers of eggs for each are in bold with the number of individuals laying for each taxon in brackets.

The proportions of eggs laid by each F2 individual is presented in figure 12 with the total number of eggs laid by each.

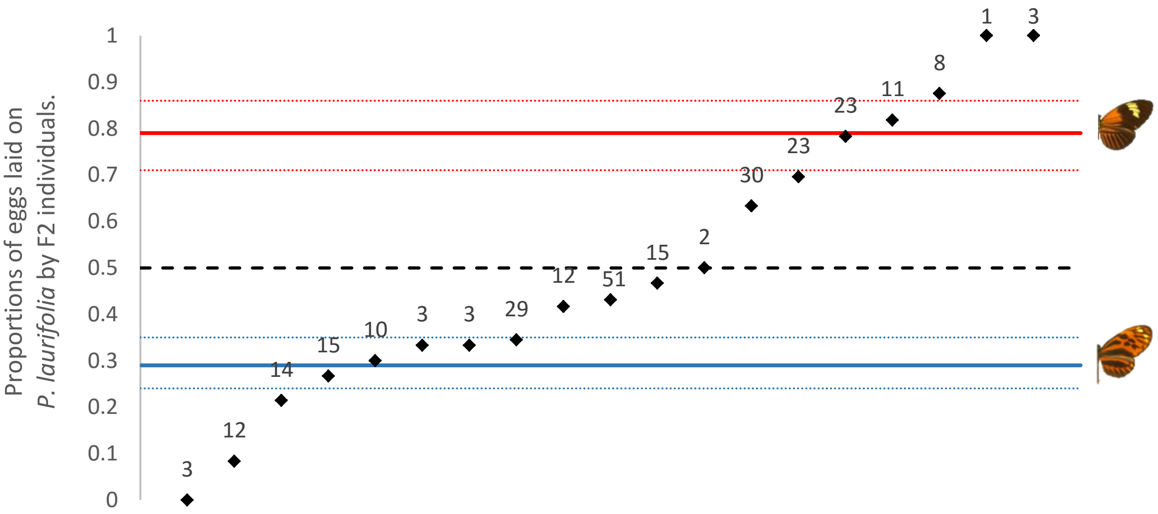

Figure 12. F2 individual proportions of eggs laid on *P. laurifolia* (value of 1) against *P. riparia* (value of 0). Each data point represents an F2 individual and figures above each point are the number of eggs laid for that individual. The likelihoods of laying on the different hosts for *H. elevatus* (in red) and *H. p. butleri* (in blue) are represented by the bold lines, the dotted lines represent the 95 % confidence intervals.

The spread of proportions of eggs laid on the two hosts (fig 13, a) was analysed using the same cluster analysis as for colour pattern to identify any underlying pattern in the F2 dataset. The analysis was performed on the entire dataset and separately excluding individuals that had laid less than 5 eggs. Looking at V over k (fig 13, b) no obvious "elbow" appears in the data suggesting that there are no clusters, even for the dataset excluding data points with small sample sizes. This suggests a non Mendelian inheritance of the trait.

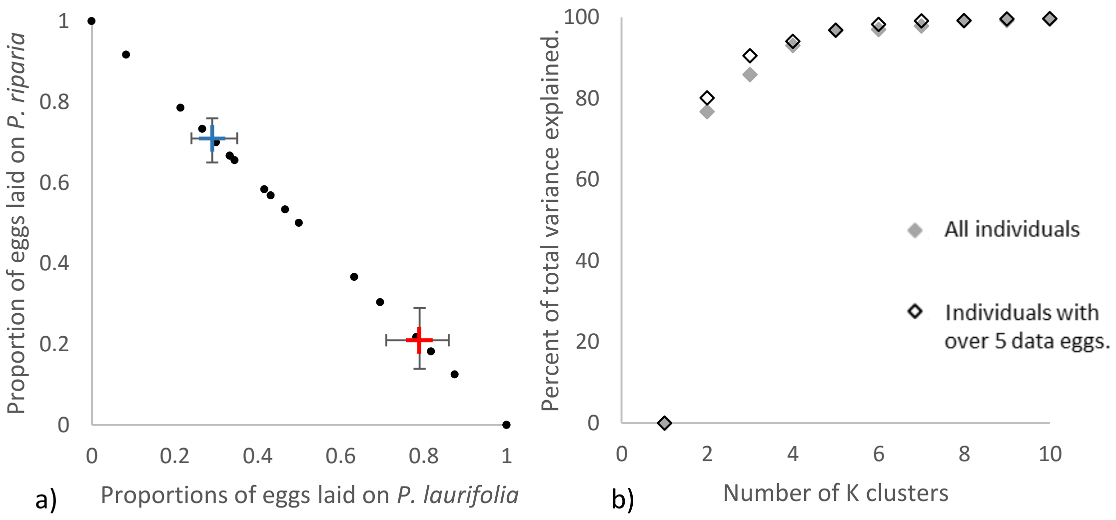

Figure 13. Spread of F2 individual preferences for egg laying on the different hosts (a). Each data point represents an individual. The crosses represent the likelihoods of laying on the different host for *H. elevatus* (in red) and *H. p. butleri* (in blue); the error bars are determined from the 95 % confidence intervals. The corresponding variance values over K clusters are presented for all data and data with over five data points (b).

Identifying genetic association between traits

The traits for male colour pattern preference and female host preference were tested for genetic associations with loci controlling colour pattern traits. Genotypic data was inferred from wing pattern traits from the F2 individuals.

Genetic associations were then analysed using a GLM using binomial families, overdispersion was corrected for using quasiGLM. Significance was tested with X^2-tests.

For the male colour pattern preference, the numbers of approaches and hovers towards the *H. elevatus* and *H. p. butleri* were used as a response variable removing individuals with a single observation. Number of eggs was used as a response variable for the host plant experiment. A total of 36, 38 and 17 individuals were used for the analyses of approaches, hovers and egg laying respectively; one female was lost before collecting the wings and was removed from the analysis. The effect of each locus was then analysed for the three response variables separately to identify an effect of colour pattern genotype on traits involved in reproductive barriers. The “dennis” and “rays” loci had no effect on male courtship or egg laying preference (table 1). However, there were significant effects of the forewing band genotype on the three measured traits (table 1).

Table 1. Significance values from the X^2-tests testing for associations between each colour pattern locus with male colour pattern preference (for approaches and hovers) and female host plant preference.

	Male colour pattern preference		**Host plant preference**
	Approach (n = 36)	**Hover (n=38)**	**Egg laying (n = 17)**
"Dennis" locus	N.S. (p = 0.5896)	N.S. (p = 0.5994)	N.S. (p = 0.6647)
"Rays" locus	N.S. (p = 0.5597)	N.S. (p = 0.5978)	N.S. (p = 0.4961)
"Band" locus	**$X^2_{(2)} = 17.34$, $p < 0.05$**	**$X^2_{(2)} = 20.88$, $p < 0.05$**	**$X^2_{(2)} = 18.65$, $p < 0.05$**

The proportions of courtship behaviours towards the *H. elevatus* pattern by the F2 individuals are shown in figure 14 (a and b). EP individuals have a significantly higher proportion of hovers towards the *H. elevatus* pattern than EE homozygous individuals. Both EP and PP individuals demonstrate higher proportions of hovers towards the *H. elevatus* colour pattern than EE individuals.

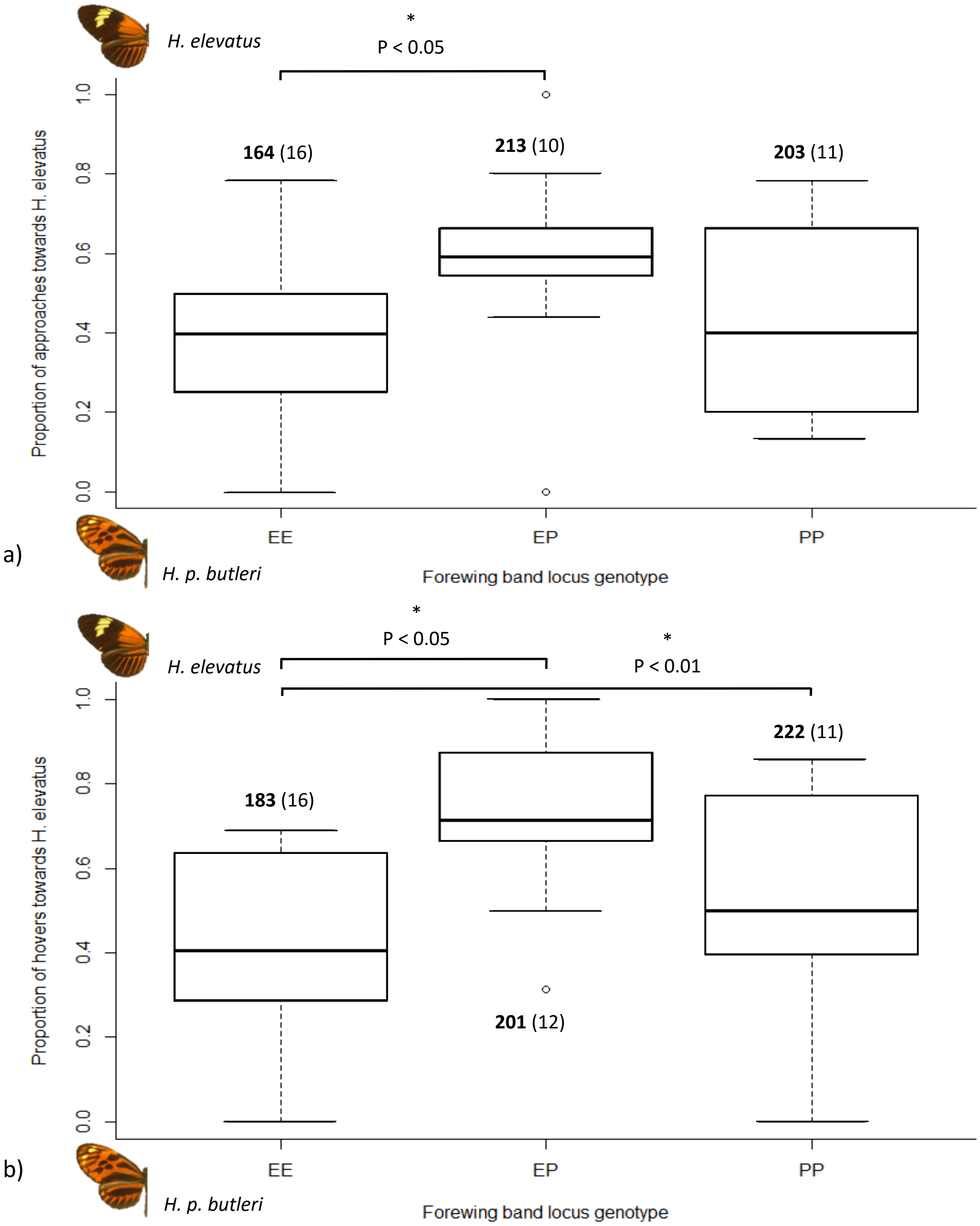

Figure 14. Proportion of a) approaches and b) hovers performed towards the *H. elevatus* colour pattern (value of 1) against the *H. p. butleri* colour pattern (value of 0) for the three genotypes of the F2 individuals. The labels show the number of behaviours recorded for each genotype in bold and the number of individuals with the different genotypes in brackets. The p values are determined from the X^2-tests.

The effect of the forewing band locus on egg laying preference is shown in figure 15. Homozygous individuals showed no difference in proportion of eggs laid on *P. riparia,* however, the heterozygote individuals for this locus demonstrated significantly higher proportions than both EE and PP individuals.

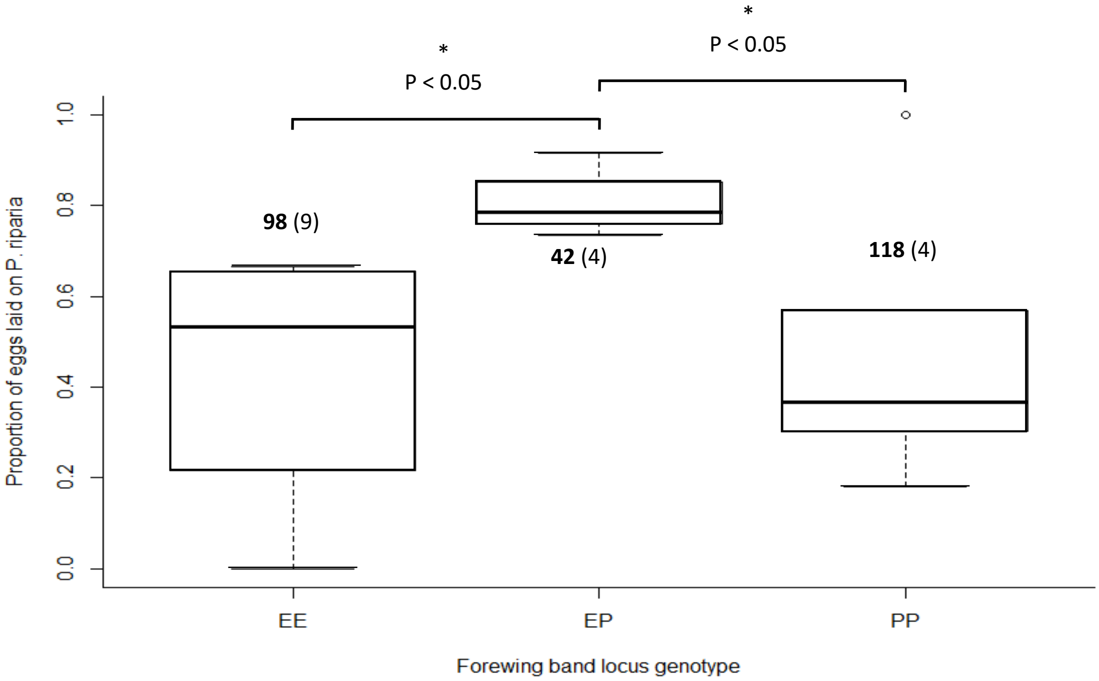

Figure 15. Proportion of eggs laid on *P. riparia* (value of 1) against *P. laurifolia* (value of 0) for the three genotypes of the F2 individuals. The labels show the number of eggs laid by each genotype in bold and the number of individuals with the different genotypes in brackets. The p values are determined from the X^2-tests.

Discussion

Understanding the genetic basis of traits involved in reproductive isolation is vital to understanding the forces of selection acting upon these traits. My experiments aimed to accurately quantify the phenotypes of the pure species and their F2 cross for two traits acting as reproductive barriers in a sympatric pair. Using this, I aim to identify patterns of inheritance and associations with colour pattern traits. This data will also be used for QTL analyses using Restricted site Associated DNA (RAD) markers. At this stage I can only make crude conclusions about the genetic structure of these traits based on observations made in the F2 generation hybrids.

In the colour pattern experiment, the pure species preference was already determined in the previous chapter. Although there is some variation in the species' preferences, results are comparable to results from similar studies in other *Heliconius* species (Jiggins et al., 2001) and strong differences were identified in the pure species. Therefore, the experiment was carried out on the F2 population. Phenotypes measured in the F2 population ranged from complete preference towards *H. elevatus* to complete preference for *H. p. butleri*. Evidence of five clusters was found for hovers. This does not suggest a simple Mendelian pattern of inheritance, rather that the trait for colour pattern preference is controlled by several loci. However, the small sample size and potential brood variation mean that this conclusion should only be taken as an initial observation. More individuals need to be measured, while controlling for brood. The variation in the parental species and difficulty of accurately measuring behaviours can also influence results. Environmental effects could also have caused the variation observed in the data. No measure of heritability was carried out on the F1 population to estimate the relative contributions of genes and environment. This was in part due to the fact that F1 individuals were veryimportant for laying and testing butterflies can potentially cause stress or damage that reduces their productivity.For host plant preference, pure species were tested again on two hosts before testing the F2 population; pure species preferences found were consistent with the results in Chapter two. No clustering was found in the F2 phenotypes suggesting the effect of several loci involved in the trait. However, the same factors influencing results in the colour pattern preference are valid for the host plant data. This result goes against many examples of insect which demonstrate host preference traits that seemed to be controlled by a few genes, such as the aphid *Cryptomyzus galeopsidis* (Guldemond, 1990) though other species have been found to demonstrate similar polygenic host preference traitssuch as *Timema cristinae* (Gompert *et al.*, 2014). This demonstrates how little is known still of the genetic determination of ecological traits and the importance of studying this in a large range of taxa.

The analysis looking for genetic associations in male colour pattern preference and female host preference traits found a significant association with the forewing band locus. Homozygotes are expected to demonstrate the same preferences as measured in the pure species for courtship behaviours or host preference, with heterozygotes displaying either a pure species or an intermediate phenotype. This was not observed in any of the traits analysed. This is perhaps not too surprising considering the small sample size of F2 individuals phenotyped. Furthermore, other studies have investigated genetic associations between colour pattern traits and host plant preference (Merrill et al., 2013) and male colour pattern preference (Merrill et al., 2011b).

Host plant use was shown to be associated to the *Ac* locus controlling for the forewing band in *H. cydno* (Merrill et al., 2013). In this study however, the differences in preference were much more marked, with less variation, due to *H. melpomene* specialising on *P. menispermifolia*. This makes the determination of genetic association more robust. Male colour pattern preference was measured in 183 males and was found to be associated to the B locus controlling for forewing colouration in *H. melpomene* (Merrill et al., 2011b). Therefore, other studies do not agree on genetic associations found in this thesis. Although different associations could occur in the taxa studied in this thesis, the sample size is too small to make any definitive conclusions, reinforcing the importance of collecting more data and carrying out genomic analyses.

Evidence of strong associations between colour pattern traits and other traits involved in reproductive isolation in species of *Heliconius* (Merrill et al., 2013; Merrill et al., 2011b) is promising for similar studies in other species. Furthermore, we already know of two highly divergent regions between *H. elevatus* and *H. p. butleri* which contain loci controlling for colour pattern (Heliconius Genome Consortium, 2012). Colour pattern is a very important trait in *Heliconius* as it is used as an aposematic signal to predators, therefore we expect these loci to be under strong selection as intermediates will have sub-optimal signalling (Merrill et al., 2015). Colour pattern has the potential of a magic trait due to its dual role in mating cues (Servedio et al., 2011) resulting in strong divergence from potentially both natural and sexual selection on these two genomic regions. Identifying other loci controlling for traits involved in reproductive isolation within these regions could generate strong evidence for patterns of islands of divergence. This can be done using QTL analyses.

One of the main limitations of QTL analyses, however, is the need for very large sample sizes (Slate, 2005). With smaller sample sizes it can be difficult to identify loci with smaller effects. Added variation due to inaccuracies in measuring phenotypes, genotypes and the effects of the environment can also result in an underestimation of the number of loci involved (Erickson, 2005). A sample size of 300 individuals is considered to be adequate for a thorough analysis (Erickson, 2005). This can be difficult for studies of natural populations to obtain due to the difficulty to cross or generate large numbers of individuals. This number is just a guideline however, as traits controlled by fewer loci of larger effect sizes will require fewer individuals.

Studies carrying out QTL analyses in natural populations have nevertheless succeeded in identifying QTLs using relatively “small” sample sizes. The study of genetic linkage in pea aphids was performed on 194 F2 individuals (Hawthorne and Via, 2001), the study on the genetic architecture in threespine sticklebacks used 92 back crossed individuals (Peichel et al., 2001) and the genetic mapping in monkeyflowers was carried out on 93 F2 plants (Bradshaw Jr et al., 1995). However, small sample sizes can also lead to an overestimation of effect size of individual QTLs, known as the Beavis effect (Erickson, 2005; Slate, 2005). In contrast, our sample size of 39 F2 individuals performing hovers is insufficient. More work is currently being carried out to increase sample size though the difficulty of rearing, crossing and collecting behavioural data are limiting factors.

Another limiting factor to the identification of QTLs is the number of markers used. The three studies mentioned above used markers such as Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and microsatellites. However, Single Nucleotide Polymorphisms (SNP) obtained by RADseq have been increasingly used as they are present in high density and can therefore allow better coverage of the genome (Vignal et al., 2002). These will be used in the future analysis of these traits.

The use of a natural population can also make QTL analyses more difficult. Slate (2005) defines natural populations as “recently sampled individuals of a non-domesticated organism”, which is the case for the *H. elevatus* and *H. p. butleri* populations used in these experiments. Natural populations, as opposed to selected, inbred lines, are not necessarily completely homozygous (Mauricio, 2001), which increases variation after recombination in the hybrid individuals. Not only will the characterisation of the phenotype be less accurate but linkage mapping of natural populations is more difficult. However, with decreasing costs and easier techniques for genomic studies, the number of QTL studies in natural populations have increased (Slate et al., 2009).

Although there are limitations in QTL analyses, studies have used this technique to understand the genetics of traits involved in reproductive isolation. For example, major effect loci were found for traits that contribute to reproductive isolation in two sympatric species of monkeyflowers (Bradshaw Jr et al., 1995). Similar findings are demonstrated in rapidly diverged sticklebacks for traits controlling armour plate (Peichel et al., 2001). Traits are generally considered to be controlled by different loci of varying effect magnitudes; with genes of major effects increasing the speed of divergence and facilitating speciation (Coyne, 1992). Another study in pea aphids found evidence of pleiotropy or close physical linkage between QTLs linked to host acceptance (affecting mating outcome) and host performance (Hawthorne and Via, 2001) supporting the importance of such mechanisms for divergent selection.

Genomic studies and QTL analyses can therefore be used to identify potential patterns of islands of divergence. It is important to note however, that finding such patterns does not prove that the species in questions diverged in sympatry. A homogenous genome with peaks of divergence is unlikely to evolve under allopatric conditions (Via, 2012; Via and West, 2008). However, it is difficult to disprove that a period of allopatry or parapatry did not initiate divergence before secondary contact (Rosser et al., 2015). One can, however, demonstrate periods of divergence with gene flow and speciation being completed with gene flow (Heliconius Genome, 2012). Indeed, a study carried out by Martin *et al.* (2013) was able to show evidence of gene flow over different time points of speciation across the genome in species of *Heliconius*. However, period of reduced gene flow may have also occurred. The recent and rapid adaptive radiation in *Heliconius* however, does reduce the likelihood of species diverging in allopatry before undergoing rapid range movements. Furthermore, given the distribution of *Heliconius* sister species today, unless very rapid range movements have occurred, simulations suggest sympatric speciation has driven 32 % to 95 % of speciation events (Rosser et al., 2015).

Finally, if the colour pattern observed in *H.elevatus* did evolve through introgression from *H. melpomene* then this suggests a role of gene flow as hybridisation occurs within the range of the parent species (Bolnick and Fitzpatrick, 2007). Although this may not be evidence for sympatric speciation as parent species already demonstrate reproductive isolation when hybridising it nevertheless underlines the plausibility of divergence in the face of gene flow. However, whether speciation has occurred in complete sympatry or allopatry should not be the sole focus of speciation studies. It is more likely that several modes of speciation will come into play in the diversification of species. It is important however, to look at the prevalence of these different modes, their relative contributions, the underlying mechanisms and the roles of natural and sexual selection in the generation of species diversity. For our system, more data needs to be collected on the divergent traits in the F2 populations for the QTL analysis to start answering these questions.

Chapter 4

Concluding Remarks

Speciation is the process of species formation. Traditionally, it was considered in the context of geographic isolation. Two opposite modes were discussed: allopatric speciation, with zero geographic isolation between two populations; and sympatric speciation; with complete distribution overlap (Coyne and Orr, 2004). These modes comprise the ends of a continuum, connected by parapatric speciation, where two species come into contact but do not overlap. Speciation is now more often considered in terms of gene flow. In allopatry, there is no gene flow and in sympatry there is complete gene flow (Nosil, 2012).

Allopatric speciation has been, and still is to some extent today, considered as a "default", i.e. all speciation is assumed to have occurred in allopatry unless proved otherwise (Bolnick and Fitzpatrick, 2007). Sympatric speciation was not considered possible as gene flow would break down favourable associations between loci under divergent selection (Felsenstein, 1981). However, theoretical evidence has demonstrated under which conditions sympatric speciation could occur (Kondrashov and Mina, 1986); many examples of putative sympatric speciation support this idea (Nosil, 2012).

The focus of research is now centred on finding more examples of sympatric speciation to assess its contribution to speciation events and understand the underlying mechanisms. However, it can be difficult to determine past processes, and without detailed historical geographical distributions of species it is difficult to prove that divergence did not initially occur in a past period of allopatry (Rosser et al., 2015). Therefore, a lot of research now focuses on identifying evidence of divergence in the face of gene flow. Many studies have investigated the role of reinforcement in species formation. Identifying stronger pre-zygotic barriers in sympatry than in allopatry can be evidence for divergence with gene flow (Coyne and Orr, 2004).

In the first chapter I demonstrated differences in male colour pattern preference. Colour pattern is an important trait in *Heliconius*. Not only is it involved in Müllerian mimicry, making it subject to strong divergent selection against intermediate phenotypes (Merrill et al., 2015), but it is also used for assortative mating (Jiggins et al., 2001; Merrill et al., 2011b). Therefore, male colour pattern preference acts as a strong reproductive barrier. This dual role as a trait under divergent selection involved in assortative suggests colour pattern could be acting as a magic trait (Servedio et al., 2011), which can facilitate speciation as magic traits are more resistant to recombination.

Differences in preference in host plants for female oviposition were also identified between the three taxa. *Heliconius* are known for their tight co-evolution with species of the *Passiflora* genus (Benson et al., 1975) which can lead to host specific adaptions in the *Heliconius*. Competition for resources between races or species (Berlocher and Feder, 2002) can also lead to frequency and density dependent selection, potentially leading to divergent selection. This can cause reproductive barriers in different ways. If matings occur on or around the host then differences in host plant preference can cause assortative mating. This can also arise if loci controlling for traits involved in assortative mating are genetically associated to loci controlling for host plant preference under divergent selection. However, it is also possible that individuals that use the same hosts are more likely to encounter individuals with the same preferences (Rosser et al., 2015).

Male colour pattern preference and female oviposition preference are therefore acting as reproductive barriers in the studied taxa. This difference is stronger between the sympatric species *Heliconius elevatus* and *H. pardalinus butleri*. The results therefore suggest a pattern of reinforcement, which in turn is evidence for divergence with gene flow. This is consistent with evidence of gene flow between these two species (Heliconius Genome Consortium, 2012). However, total pre-zygotic barriers between the two species are strong (Segami, 2015; Velado 2015) so other reproductive barriers must also be interacting with colour pattern and host plant preference. Such barriers could comprise differences in various traits such as pheromones, flight or behaviour.

The third chapter aimed to investigate the underlying genetic architecture of these reproductive barriers. Phenotyping the F2 hybrids from the *H. elevatus* and *H. p. butleri* crosses did not show patterns of simple Mendelian inheritance. This suggests several loci might be involved in these traits though only crude conclusions can be made when analysing small samples. Evidence for genetic associations between male colour pattern and host plant preference traits was found, but the genotypes did not exhibit the expected phenotypes. This was probably an effect of the small sample size. Other studies have demonstrated such associations in other *Heliconius* species (Merrill et al., 2013; Merrill et al., 2011b) which confirms the potential of finding such associations.

More work will be carried out to increase sample size and increase the robustness of the data. Work will then focus on using QTL analyses to identify the underlying genomic architecture of these traits. This may detect patterns of islands of divergence, where loci controlling traits acting as reproductive barriers will be identified within known peaks of divergence in the genome (Heliconius Genome Consortium, 2012). If so, the *H. elevatus* and *H. p. butleri* system could greatly increase our understanding of the processes of divergence with gene flow and shed light on the contribution and mechanisms of sympatric speciation.

References

Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, and A. Meyer, 2006, Sympatric speciation in Nicaraguan crater lake cichlid fish: Nature, v. 439, p. 719-723.

Bates, H. W., 1862, XXXII. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconidæ: Transactions of the Linnean Society of London, v. 23, p. 495-566.

Bates D, Maechler M, Bolker BM and Walker S, 2015, Fitting Linear Mixed-Effects Models using lme4. ArXiv e-print; in press, Journal of Statistical Software.

Benson, W. W., K. S. Brown Jr, and L. E. Gilbert, 1975, Coevolution of plants and herbivores: passion flower butterflies: Evolution, v. 29, p. 659-680.

Berlocher, S. H., and J. L. Feder, 2002, Sympatric speciation in phytophagous insects: moving beyond controversy?: Annual Review of Entomology, v. 47, p. 773-815.

Blair, W. F., 1964, Isolating Mechanisms and Interspecies Interactions in Anuran Amphibians: The Quarterly Review of Biology, v. 39, p. 334-344.

Bolnick, D. I., and B. M. Fitzpatrick, 2007, Sympatric speciation: models and empirical evidence: Annual Review of Ecology, Evolution, and Systematics, v. 38, p. 459-487.

Bradshaw Jr, H. D., M. Wilbert, and K. G. Otto, 1995, Genetic mapping of floral traits associated with reproductive isolation: Nature, v. 376, p. 762-765.

Boake, C. R. B., and S. Gavrilets, 1998, On the evolution of premating isolation after a founder event: The American Naturalist, v.152, p. 706-716.

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

Cama, B., 2015, A phenotypic analysis of the warning colour pattern in two species of *Heliconius* butterflies: putative loci and evidence for linkage. Undergraduate dissertation, University of York.

Coyne, J. A., and H. A. Orr, 1989, Patterns of speciation in Drosophila: Evolution, v. 43, p. 362-381.

Coyne, J. A., 1992, Genetics and speciation: Nature, v. 355, p. 511-515.

Coyne, J. A., and H. A. Orr, 2004, Speciation, v. 37, Sunderland, Massachusetts: Sinauer Associates.Dasmahapatra, K. K., G. Lamas, F. Simpson, and J. Mallet, 2010, The anatomy of a ‘suture zone’in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation: Molecular Ecology, v. 19, p. 4283-4301.

Deinert, E. I., J. T. Longino, and L. E. Gilbert, 1994, Mate competition in butterflies: Nature, v. 370, p. 23-24.

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

Ellegren, H., L. Smeds, R. Burri, P. I. Olason, N. Backström, T. Kawakami, A. Künstner, H. Mäkinen, K. Nadachowska-Brzyska, and A. Qvarnström, 2012, The genomic landscape of species divergence in *Ficedula* flycatchers: Nature, v. 491, p. 756-760.

Erickson, D., 2005, Quantitative trait loci: mapping the future of QTL's: Heredity, v. 95, p. 417-418.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja, 2003, Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*: Proceedings of the National Academy of Sciences, v. 100, p. 10314-10319.

Feder, J. L., S. P. Egan, and P. Nosil, 2012, The genomics of speciation-with-gene-flow: Trends in Genetics, v. 28, p. 342-350.

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, v. 91, p. 7990-7994.

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

Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrilets, 2008, What, if anything, is sympatric speciation?: Journal of Evolutionary Biology, v. 21, p. 1452-1459.

Flaxman, S. M., A. C. Wacholder, J. L. Feder, and P. Nosil, 2014, Theoretical models of the influence of genomic architecture on the dynamics of speciation: Molecular Ecology, v. 23, p. 4074-4088.

Gilbert, L. E., 1972, Pollen feeding and reproductive biology of *Heliconius* butterflies: Proceedings of the National Academy of Sciences, v. 69, p. 1403-1407.

Gilbert, L. E., P. W. Price, T. M. Lewinsohn, G. W. Fernandes, and W. W. Benson, 1991, Plant-animal Interactions: Evolutionary Ecology in Tropical and Temperate Regions: John Wiley and Sons, Inc., 1991.

Gompert, Z., A.A. Comeault, T.E. Farkas, J. L. Feder, T. L. Parchman, C. A. Buerkle and P. Nosil, 2014, Experimental evidence for ecological selection on genome variation in the wild: Ecology Letters, v. 17, p. 369-379.

Guldemond, J.A., 1990, 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, v. 57, p.65-76.

Hawthorne, D. J., and S. Via, 2001, Genetic linkage of ecological specialization and reproductive isolation in pea aphids: Nature, v. 412, p. 904-907.

Heliconius Genome Consortium, 2012, Butterfly genome reveals promiscuous exchange of mimicry adaptations among species: Nature, v. 487, p. 94-98.

Hillis, D. M., 1981, Premating isolating mechanisms among three species of the *Rana pipiens* complex in Texas and southern Oklahoma: Copeia, v. 2, p. 312-319.

Hohenlohe, P. A., S. Bassham, M. Currey, and W. A. Cresko, 2012, Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes: Philosophical Transactions of the Royal Society of London B: Biological Sciences, v. 367, p. 395-408.

Hothorn, T., and B. S. Everitt, 2014, A handbook of statistical analyses using R. CRC press.

Howard, D. J., and S. H. Berlocher, 1998, Endless forms: species and speciation, Oxford: Oxford University Press.

Jang, Y., and H. C. Gerhardt, 2006, Divergence in female calling song discrimination between sympatric and allopatric populations of the southern wood cricket *Gryllus fultoni* (Orthoptera: Gryllidae): Behavioral Ecology and Sociobiology, v. 60, p. 150-158.

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

Jones, R. T., Y. L. Poul, A. C. Whibley, C. Mérot, and M. Joron, 2013, Wing shape variation associated with mimicry in butterflies: Evolution, v. 67, p. 2323-2334.

Joron, M., R. Papa, M. Beltrán, N. Chamberlain, J. Mavárez, S. Baxter, M. Abanto, E. Bermingham, S. J. Humphray, and J. Rogers, 2006, A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies: PLoS Biol, v. 4, p. e303.

Killip, E. P., 1938, The American species of Passifloraceae: The American species of Passifloraceae, Chicago: University of Chicago.

Klein, A. L., and A. M. De Araújo, 2010, Courtship behavior of *Heliconius erato phyllis* (Lepidoptera, Nymphalidae) towards virgin and mated females: conflict between attraction and repulsion signals?: Journal of Ethology, v. 28, p. 409-420.

Kondrashov, A. S., and M. V. Mina, 1986, Sympatric speciation: when is it possible?: Biological Journal of the Linnean Society, v. 27, p. 201-223.

Kozak, K. M., N. Wahlberg, A. F. E. Neild, K. K. Dasmahapatra, J. Mallet, and C. D. Jiggins, 2015, Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies: Systematic Biology, v. 64, p. 505-524.

Kronforst, M. R., L. G. Young, and L. E. Gilbert, 2007, Reinforcement of mate preference among hybridizing *Heliconius* butterflies: Journal of Evolutionary Biology, v. 20, p. 278-285.

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

Mallet, J., 1986, Dispersal and gene flow in a butterfly with home range behavior: *Heliconius erato* (Lepidoptera: Nymphalidae): Oecologia, v. 68, p. 210-217.

Mallet, J., 1989, The genetics of warning colour in Peruvian hybrid zones of *Heliconius erato* and *H. melpomene*: Proceedings of the Royal Society of London B: Biological Sciences, v. 236, p. 163-185.

Mallet, J., 2005, Hybridization as an invasion of the genome: Trends in Ecology & Evolution, v. 20, p. 229-237.

Mallet, J., 2007, Hybrid speciation: Nature, v. 446, p. 279-283.

Mallet, J., and L. E. Gilbert, 1995, Why are there so many mimicry rings? Correlations between habitat, behaviour and mimicry in *Heliconius* butterflies: Biological Journal of the Linnean Society, v. 55, p. 159-180.

Mallet, J., A. Meyer, P. Nosil, and J. L. Feder, 2009, Space, sympatry and speciation: Journal of Evolutionary Biology, v. 22, p. 2332-2341.

Marshall, J. L., M. L. Arnold, and D. J. Howard, 2002, Reinforcement: the road not taken: Trends in Ecology & Evolution, v. 17, p. 558-563.

Martin, C. H., J. S. Cutler, J. P. Friel, C. T. Dening, G. Coop, and P. C. Wainwright, 2015, Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation: Evolution, v. 69, p. 1406-1422.

Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins, 2013, Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies: Genome Research, v. 23, p. 1817-1828.

Mauricio, R., 2001, Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology: Nature Reviews Genetics, v. 2, p. 370-381.

Mavárez, J., C. A. Salazar, E. Bermingham, C. Salcedo, C. D. Jiggins, and M. Linares, 2006, Speciation by hybridization in *Heliconius* butterflies: Nature, v. 441, p. 868-871.

Maynard Smith, J., 1966, Sympatric speciation: American Naturalist, v. 100, p. 637-650.

Mayr, E., 1963, Animal species and evolution, v. 797, Cambridge, Massachusetts: Belknap Press of Harvard University Press.

Mayr, E., 2000, The biological species concept: Species concepts and phylogenetic theory: a debate. New York: Columbia University Press.

Merrill, R. M., K. K. Dasmahapatra, J. W. Davey, D. D. Dell'Aglio, J. J. Hanly, B. Huber, C. D. Jiggins, M. Joron, K. M. Kozak, and V. Llaurens, 2015, The diversification of *Heliconius* butterflies: what have we learned in 150 years?: Journal of Evolutionary Biology, v. 28, p. 1417-1438.

Merrill, R. M., Z. Gompert, L. M. Dembeck, M. R. Kronforst, W. O. McMillan, and C. D. Jiggins, 2011a, Mate preference across the speciation continuum in a clade of mimetic butterflies: Evolution, v. 65, p. 1489-1500.

Merrill, R. M., R. E. Naisbit, J. Mallet, and C. D. Jiggins, 2013, Ecological and genetic factors influencing the transition between host-use strategies in sympatric *Heliconius* butterflies: Journal of Evolutionary Biology, v. 26, p. 1959-1967.

Merrill, R. M., B. Van Schooten, J. A. Scott, and C. D. Jiggins, 2011b, Pervasive genetic associations between traits causing reproductive isolation in *Heliconius* butterflies: Proceedings of the Royal Society B: Biological Sciences, v. 278, p. 511-518.

Michel, A. P., S. Sim, T. H. Q. Powell, M. S. Taylor, P. Nosil, and J. L. Feder, 2010, Widespread genomic divergence during sympatric speciation: Proceedings of the National Academy of Sciences, v. 107, p. 9724-9729.

Mérot, C., J. Mavárez, A. Evin, K. K. Dasmahapatra, J. Mallet, G. Lamas, and M. Joron, 2013, Genetic differentiation without mimicry shift in a pair of hybridizing *Heliconius* species (Lepidoptera: Nymphalidae): Biological Journal of the Linnean Society, v. 109, p. 830-847.

Nadeau, N. J., A. Whibley, R. T. Jones, J. W. Davey, K. K. Dasmahapatra, S. W. Baxter, M. A. Quail, M. Joron, M. L. Blaxter, and J. Mallet, 2012, Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing: Philosophical Transactions of the Royal Society B: Biological Sciences, v. 367, p. 343-353.

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

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland, 2001, Chromosomal inversions and the reproductive isolation of species: Proceedings of the National Academy of Sciences, v. 98, p. 12084-12088.

Nosil, P., 2012, Ecological speciation, Oxford: Oxford University Press.

Nosil, P., B. J. Crespi, and C. P. Sandoval, 2003, Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement: Proceedings of the Royal Society of London B: Biological Sciences, v. 270, p. 1911-1918.

Nosil, P., and J. L. Feder, 2012, Genomic divergence during speciation: causes and consequences: Philosophical Transactions of the Royal Society B: Biological Sciences, v. 367, p. 332-342.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos, 2009, Divergent selection and heterogeneous genomic divergence: Molecular Ecology, v. 18, p. 375-402.

Nylin, S., J. Slove, and N. Janz, 2014, Host plant utilization, host range oscillations and diversification in nymphalid butterflies: a phylogenetic investigation: Evolution, v. 68, p. 105-124.

Papadopulos, A. S. T., W. J. Baker, D. Crayn, R. K. Butlin, R. G. Kynast, I. Hutton, and V. Savolainen, 2011, Speciation with gene flow on Lord Howe Island: Proceedings of the National Academy of Sciences, v. 108, p. 13188-13193.

Peichel, C. L., K. S. Nereng, K. A. Ohgi, B. L. E. Cole, P. F. Colosimo, C. A. Buerkle, D. Schluter, and D. M. Kingsley, 2001, The genetic architecture of divergence between threespine stickleback species: Nature, v. 414, p. 901-905.

Phillimore, A. B., C. D. L. Orme, G. H. Thomas, T. M. Blackburn, P. M. Bennett, K. J. Gaston, and I. P. F. Owens, 2008, Sympatric speciation in birds is rare: insights from range data and simulations: The American Naturalist, v. 171, p. 646-657.

Ratcliffe, L. M., and P. R. Grant, 1983, Species recognition in Darwin's finches (Geospiza, Gould). II. Geographic variation in mate preference: Animal Behaviour, v. 31, p. 1154-1165.

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

Rosser, N., K. M. Kozak, A. B. Phillimore, and J. Mallet, 2015, Extensive range overlap between heliconiine sister species: evidence for sympatric speciation in butterflies?: BMC Evolutionary Biology, v. 15, p. 125.

Rosser, N., A. B. Phillimore, B. Huertas, K. R. Willmott, and J. Mallet, 2012, Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America: Biological Journal of the Linnean Society, v. 105, p. 479-497.

Rundle, H. D., and D. Schluter, 1998, Reinforcement of stickleback mate preferences: sympatry breeds contempt: Evolution, v. 52, p. 200-208.

Schliewen, U. K., T. D. Kocher, K. R. McKaye, O. Seehausen, and D. Tautz, 2006, Evolutionary Biology: Evidence for sympatric speciation?: Nature, v. 444, p. E12-E13.

Schliewen, U. K., D. Tautz, and S. Pääbo, 1994, Sympatric speciation suggested by monophyly of crater lake cichlids: Nature, v. 368, p. 629-632.

Segami Marzal, C., 2015, Mate choice and hybrid viability in two species of Heliconius. MSc Thesis, Ludwig-Maximilians University Munich.

Servedio, M. R., and M. Kirkpatrick, 1997, The effects of gene flow on reinforcement: Evolution, v. 51, p. 1764-1772.

Servedio, M. R., and M. A. F. Noor, 2003, The role of reinforcement in speciation: theory and data: Annual Review of Ecology, Evolution, and Systematics, v. 34, p. 339-364.

Servedio, M. R., G. S. Van Doorn, M. Kopp, A. M. Frame, and P. Nosil, 2011, Magic traits in speciation:'magic'but not rare?: Trends in Ecology & Evolution, v. 26, p. 389-397.

Slate, J., J. Gratten, D. Beraldi, J. Stapley, M. Hale, and J. M. Pemberton, 2009, Gene mapping in the wild with SNPs: guidelines and future directions: Genetica, v. 136, p. 97-107.

Slate, J. O. N., 2005, INVITED REVIEW: Quantitative trait locus mapping in natural populations: progress, caveats and future directions: Molecular Ecology, v. 14, p. 363-379.

Srygley, R. B., 1999, Locomotor mimicry in *Heliconius* butterflies: contrast analyses of flight morphology and kinematics. Philosophical Transactions of the Royal Society B: Biological Sciences, v. 354, p. 203-214.

Supple, M., R. Papa, B. Counterman, and W. O. McMillan, 2014, The genomics of an adaptive radiation: insights across the Heliconius speciation continuum in:Ecological Genomics, Springer, p. 249-271.

Supple, M. A., H. M. Hines, K. K. Dasmahapatra, J. J. Lewis, D. M. Nielsen, C. Lavoie, D. A. Ray, C. Salazar, W. O. McMillan, and B. A. Counterman, 2013, Genomic architecture of adaptive color pattern divergence and convergence in Heliconius butterflies:Genome Research, v. 23, p. 1248-1257.

Supple, M. A., R. Papa, H. M. Hines, W. O. McMillan, and B. A. Counterman, 2015, Divergence with gene flow across a speciation continuum of Heliconius butterflies: BMC Evolutionary Biology, v. 15, p. 204.

Templeton, A. R., 1989, The meaning of species and speciation: a genetic perspective, in: The units of evolution: Essays on the nature of species, p. 159-183.

Velado, P., 2015, Species status of the neotropical butterfly Heliconius pardalinus sergestus. MSc Thesis, Ludwig-Maximilians University Munich.

Via, S., 2012, Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow: Philosophical Transactions of the Royal Society B: Biological Sciences, v. 367, p. 451-460.

Via, S., and J. West, 2008, The genetic mosaic suggests a new role for hitchhiking in ecological speciation: Molecular Ecology, v. 17, p. 4334-4345.

Vignal, A., D. Milan, M. SanCristobal, and A. Eggen, 2002, A review on SNP and other types of molecular markers and their use in animal genetics: Genetics Selection Evolution, v. 34, p. 275-306.

Waage, J. K., 1975, Reproductive isolation and the potential for character displacement in the damselflies, *Calopteryx maculata* and *C. aequabilis* (Odonata: Calopterygidae): Systematic Biology, v. 24, p. 24-36.

Wu, C. I., 2001, The genic view of the process of speciation: Journal of Evolutionary Biology, v. 14, p. 851-865.

Yukilevich, R. and J.R.True, 2006, Divergent outcomes of reinforcement speciation: the relative importance of assortative mating and migration modification: The American Naturalist, v. 167, p.638-654.