Genetics of behavioral evolution in *Heliconius* butterflies

Dissertation der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Vorgelegt von
Matteo Rossi
2021

A prince was dining with his mother. While slicing white cheese, he cut his finger.
A drop of blood fell on the cheese, so he said to his mother:
"Mother, I would like to marry a woman who is as white as milk, and as red as blood."
"My son", replied his mother, "She who is white is not red, and she who is red is not white.
But go search for her if you want"
[...]

The love of the three pomegranates, from Italo Calvino's *Italian folktales*

Diese Dissertation wurde unter der Leitung von Prof. Dr. **Richard Merrill** angefertigt.

1. Gutachter: Prof. Dr. **Richard Merrill**
2. Gutachter: Prof. Dr. **Jochen Wolf**

Tag der Einreichung: 18.11.2021
Tag der mündlichen Prüfung: 12.5.2022

Table of Contents

Statutory declaration and statement……I

List of publications……...II

Declaration of contributions……...III

Summary……....1

Introduction……...3

1. Speciation, genetics and the evolution of species-specific mating preferences…….3

1.1. Speciation and behavioral isolation…….3

1.2. Why the genetics of behavior? ……4

2. Genomic complexity, location and signatures of divergent mate preference behaviors……5

2.1. How many genetic changes underlie variation in preference behaviors? ……....5

2.2. Genomic associations between cues and preference loci…….7

2.3. Role of sex chromosomes in the evolution of mate preferences……...8

2.4. Genomic signatures of (behavioral) barrier loci…….....9

3. How do genes act within sensory and neural systems to alter mating preferences?…….11

3.1. Chemosensation……...……..11

3.2. Hearing……...……...13

3.3. Vision.……...…….14

3.4. Multimodal integration and neuromodulation……...…….....14

3.5. The locus of behavioral evolution……...……...16

4. Development of behavioral differences……..17

4.1. What genetic changes underlie behavioral evolution?……....17

4.2 When during development do behavioral genes act?……...18

4.3 What genetic mechanisms are responsible for sex differences in behavior?……....20

5. Concluding remarks and study system……...…….21

5.1 The study system……...……..22

5.2 *Heliconius* butterflies for studying the evolution of visual mate preferences…….....22

Results…….……...……..25

Paper I: Visual mate preference evolution during butterfly speciation is linked to neural processing genes…….……25

Manuscript II: Signatures of adaptive introgression implicate a regucalcin in the evolution of visual mate preference behaviors……....26

Paper III: Neural divergence and hybrid disruption between ecologically isolated *Heliconius* butterflies……....67

Discussion…….....68

1. Neural divergence contributes to speciation in *Heliconius* …….....68

2. Visual mate preference is independent of adaptation to the broader environment……...69

3. Visual preference loci control divergent mating preferences in the *H. melpomene-cydno* group…….....69

4. Do *Heliconius* butterflies exchange adaptive behavioral alleles?…….....71

5. Changes in visual integration or processing likely underlie mate preferences.…….....71

6. Genetic mechanisms associated with the QTL on chromosome 18…….....73

7. Functionally validating the link between genes and behavior…….....75

8.Conclusion…….....76

References……...…….....77

Acknowledgments…….....87

Statutory declaration and statement

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 17.11.2021 Matteo Rossi

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Ich habe nicht versucht, anderweitig eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

München, den 17.11.2021 Matteo Rossi

List of publications

1. **Rossi, M.**, Hausmann, A. E., Thurman, T., Montgomery, S. H., Papa, R., Jiggins, C. D., McMillan, W. O., & Merrill, R. M. (2020). Visual mate preference evolution during butterfly speciation is linked to neural processing genes. *Nature Communications*, 11(1):1-10.

2. Montgomery, S. H., **Rossi, M.**, McMillan, W. O., & Merrill, R. M. (2021). Neural divergence and hybrid disruption between ecologically isolated *Heliconius* butterflies. *PNAS*, 118 (6): e2015102118.

Declaration of contributions

Paper I

Rossi, M., Hausmann, A. E., Thurman, T., Montgomery, S. H., Papa, R., Jiggins, C. D., McMillan, W. O., & Merrill, R. M. (2020). Visual mate preference evolution during butterfly speciation is linked to neural processing genes. *Nature Communications*, 11(1):1-10.

Matteo Rossi helped conceive the study and design experiments, contributed to rearing butterflies, dissecting neural tissues and extracting RNA, performed most of the analyses (expression and sequence data), and wrote the manuscript together with Richard Merrill.

Other contributions

Alexander Hausmann: analyzed behavioral data

Timothy Thurman: reared butterflies

Stephen Montgomery: contributed to dissecting neural tissue

Riccardo Papa: contributed resources

Chris Jiggins, Owen McMillan: secured funding, contributed resources and provided supervision

Richard Merrill: conceived the study and designed the experiments, acquire behavioral and expression data, secured funding and provided supervision

Signature of the supervisor:

Prof. Dr. Richard Merrill

Manuscript II

Rossi, M., Hausmann, A. E., Moest, M., Melo, L., Kuo, C., Wright, D. S., Borrero, J., Linares, M., McMillan, W. O., Pardo-Diaz, C., Salazar, C., & Merrill, R. M. (manuscript). Signatures of adaptive introgression implicate a *regucalcin* in the evolution of visual mate preference behaviors.

Matteo Rossi helped conceive the study and designed experiments, contributed to rearing butterflies, acquired expression and sequence data, analyzed sequence and expression data, designed and supervised the acquisition of genotype data, designed and performed CRISPR experiments, and wrote the manuscript with input from Richard Merrill.

Other contributions

Alexander Hausmann: helped conceive the study and designed experiments, collected and helped to rear butterflies, supervised the acquisition of and analyzed the behavioral data, provided input on the manuscript.

Markus Moest: Performed selective sweeps analyses.

Lina Melo: reared butterflies and acquired behavioral data.

Chi-Yun Kuo: helped designed experiments, acquired reflectance measurements and performed color vision analyses.

Shane Wright and José Borrero: acquired reflectance measurements and performed color vision analyses.

Mauricio Linares: collected butterflies and contributed resources.

Carolina Pardo-Diaz, Camilo Salazar: contributed resources and acquired part of the sequence data.

Richard Merrill: conceived the study and designed experiments, acquired funding, provided supervision, provided input on the manuscript

Signature of the supervisor:

Prof. Dr. Richard Merrill

Paper III

Montgomery, S. H., Rossi, M., McMillan, W. O., & Merrill, R. M. (2021). Neural divergence and hybrid disruption between ecologically isolated *Heliconius* butterflies. *PNAS*, 118 (6): e2015102118.

Matteo Rossi contributed to rearing butterflies, acquired part of the expression data, analyzed sequence and expression data, and contributed to write the manuscript.

Other contributions:
Stephen Montgomery (first author): conceived the study and designed experiments, collected and reared butterflies, acquired and analyzed brain morphometric data, acquired funding and wrote the manuscript
Owe McMillan: contributed resources, provided supervision, contributed to write the manuscript.
Richard Merrill: helped to conceive and design experiments, acquired expression data, provided supervision and contributed to write the manuscript.

Signature of the supervisor:
Prof. Dr. Richard Merrill

Summary

Changes in the way animals perceive and respond to the environment are key to adaptation and speciation. However, we still know little of the genetic mechanisms underlying shifts in behavior. Identifying these genetic changes would provide an important route towards understanding how behavior is generated, both during development and across evolutionary time. To begin to fill this gap, in this thesis I investigate the genetics of behavioral evolution in *Heliconius* butterflies. In particular, I investigate the genetics of visual mate preferences as well as broader visual adaptations across *Heliconius* species.

Heliconius butterflies display a striking diversity of warning patterns, which they also use as mating cues to recognize conspecifics. Preferences for conspecific warning patterns have a strong genetic component, but unlike the warning pattern cues, the exact genes responsible remain unknown. In chapter 1, I analyse a causative genomic region for such divergent visual behaviors in two *Heliconius* species: red *H. melpomene* and white *H. cydno*. I couple population genomic and gene expression analyses of neural tissue of these species and their hybrids across development, to identify five genes that are strongly associated with divergent visual preferences. The functions of these candidate genes suggest shifts in behavior involve changes in visual integration or processing, which would allow mate preferences to evolve without altering perception of the wider environment.

In chapter 2, I expand my analyses to include another species: red *H. timareta*, a co-mimic of *H. melpomene*. There is substantial evidence that *H. timareta* acquired its red wing pattern coloration through hybridization (adaptive introgression) with *H. melpomene*. In this chapter, together with my colleagues, we test the hypothesis that *H. timareta* also acquired alleles for visual mate preference from *H. melpomene*. We first show that the same causative region associated with the divergent visual preferences of *H. melpomene* and *H. cydno*, also controls visual divergence between *H. timareta* and *H. cydno*. I then find genomic signatures of adaptive introgression at the level of candidate behavioral genes identified in chapter 1. One of these candidates, *regucalcin1*, also shows patterns of gene expression strongly linked to visual preference across *Heliconius* species and their hybrids. Overall, I find evidence that visual

preference alleles have crossed species barriers to facilitate adaptive shifts in behavior in *Heliconius*.

Finally, chapter 3 goes beyond divergence in mate preferences to investigate broad-scale neural divergence associated with speciation in the *H. melpomene*/*H. cydno* group. Species within this group are separated across an ecological gradient of open to closed forest. We find evidence that species have adapted to this ecological transition at the neural level, through heritable, volumetric expansion of visual processing regions of the brain. We find that these same visual structures show intermediate morphologies in F1 hybrids, which likely disrupt their behavioural function. I then show that these brain volumetric changes are mirrored by adaptive divergence in gene expression level in the neural tissue of these species. Finally, I find evidence for selection against the introgression of alleles (with distinct neural expression level) between species, further indicating that neural divergence contributes to reproductive isolation. Overall, we show that broad-scale sensory/neural adaptations to the visual environment, at both morphological and gene expression level, contribute substantially to behavioral isolation and speciation across *Heliconius*.

Introduction

Animals often have a preference to mate with only a subset of potential partners (Rosenthal 2017), and variation in these mating preferences is observed both within and between species. Mate preferences direct mate choice, which is fundamental to individual fitness, because animals should "choose" mates that will maximize their fitness, and is also a key evolutionary force, because it will influence which alleles are passed on to the next generation.

Here I review species-specific mating preferences and how they arise at the genetic level. I briefly consider why species-specific mating preferences are thought to evolve in the first place, and why studying their underlying genetics is important. I then review genetic patterns and mechanisms underlying the evolution of mate preference behaviors. I explore these genetic changes through the lenses of i) evolutionary genomic processes, ii) how they act in neural systems to produce behavioral change, and iii) how they unfold within a developmental program. Finally, I introduce *Heliconius* butterflies, the study system of the thesis.

1. Speciation, genetics and the evolution of species-specific mating preferences.

1.1. Speciation and behavioral isolation.

To understand why species-specific mating preferences evolve in the first place, it is useful to consider speciation, the process during which new species evolve. Speciation is often initiated through the action of divergent selection, as populations adapt to different ecological niches (Schluter 2009, Nosil 2012). If these diverging lineages come in contact however, reproductive barriers between them must evolve for speciation to proceed. These reproductive barriers can stem from a number of biological mechanisms, for example hybrids between individuals from these divergent populations could show developmental abnormalities, be inviable or infertile (intrinsic barriers). Also, hybrids could be maladapted to either parental environmental

condition (extrinsic barriers). Nevertheless, one barrier that often emerges between incipient species involves preferences to mate with conspecifics.

In fact, behavioral isolation is often key to speciation (Mayr 1963). Differences in courtship and mating behaviors can evolve rapidly, including for example, male courtship song (Ding et al. 2019), and song perception by females (Wang et al. 2020). In particular, mating discrimination is often among the first reproductive barriers to arise during speciation with gene flow (Coyne & Orr 1989, Mendelson et al. 2003, Coyne & Orr 2004, Jiggins et al. 2004, but see Matute & Cooper 2020). The evolution of divergent mating preferences is thought to largely occur by means of natural selection, as a by-product of divergent selection on populations, or due to selection against the production of unfit hybrids between populations (i.e. reinforcement and similar mechanisms) (Servedio & Noor 2003, Coyne & Orr 2004, Nosil et al. 2006). These rapid changes in mating behaviors, leading to the suppression of mating between diverging populations, raise a series of interesting mechanistic questions. In particular: how can evolution tinker with genetic programs to generate such diverse mating preferences?

1.2. Why the genetics of behavior?

Identifying the genetic variation underlying the evolution of species mating preferences is important to understand how both behavioral diversity and new species are generated, a central question in evolutionary biology.

Variation in behavioral outputs is the result of neural processes, and differences in innate behaviors are ultimately genetically encoded. Although behavior is shaped by the repeated interaction between neural systems and the environment (Bendesky & Bargmann 2011), many behavioral functions are innate and have a heritable component on which selection can act (Niepoth & Bendesky 2020). This applies to a wide range of mating preference behaviors. There is an increasing appreciation that complex behavioral differences between species can be hardwired in the genome. In recent years, for example, complex motor pattern differences in natural populations have been mapped to (relatively few) genomic regions (e.g. burrowing behavior, Weber et al. 2013; schooling behavior, Greenwood et al. 2013; wing display, Massey et al. 2020).

Here, I focus on those genetic changes specifically linked to variation in mate preference/choice. I have not reviewed genetic mechanisms more broadly involved in mating behaviors. Although important for understanding how overall mating behaviors are genetically constructed, and with full appreciation that genes act in concert with many other genes to allow the expression of behavior, these are beyond the scope of this review. Finally, I have also not included studies analyzing gene expression changes associated with variation in learned mate preferences (for instance, Declos al. 2020), or those associated with priming mate preference behavior (for instance, Bloch et al. 2018). Although an exciting avenue for understanding how mate preference behaviors can be executed at the molecular level (and although learned mate preferences might play an important role in speciation, Verzijden et al. 2012), these do not necessarily inform us of the genetic changes responsible for establishing variation in mate preference.

2. Genomic complexity, location and signatures of divergent mate preference behaviors.

2.1. How many genetic changes underlie variation in preference behaviors?

How complex is the genetic control of behavioral differences? Can a genetic change at a single locus trigger a shift in mate preference, or are multiple, perhaps interacting loci necessary? Behavioral differences between populations or species are often reported to be controlled by multiple loci of small effects (for example in sticklebacks, Peichel & Marques 2017), although seemingly complex behavioral patterns, like differences in courtship song, have been mapped to single loci (Ding et al. 2016).

Differences in mate preference/choice behaviors between species have often been mapped to 1-3 loci of large effect (**Table 1**). Nonetheless, they have also been mapped to many loci of small effects, for example in some *Drosophila* species (possibly linked to larger sample sizes, **Table 1**). Note that it is possible that, by using proxies for preference behavior, including for example mating success or time spent courting a mate, some mapped loci might not reflect behavioral differences, but rather morphological or physiological traits affecting behavioral measures (Hu & Hoekstra 2016, Massey et al. 2019).

Table 1. Mapping studies (QTL or deletion mapping) of divergent species mate preference behaviors (previous studies have been reviewed in Coyne & Orr 2004 Table 6.2). Note that this is meant as a comprehensive but incomplete list.

Study	Species of	#Loci*	Technique	Sample size
Moehring et al. 2004	*Drosophila* (fruit flies)	7	QTL	1005
Moehring et al. 2006	*Drosophila*	2 + 3**	QTL	535 + 539
Kronforst et al. 2006	*Heliconius* (butterflies)	1	QTL	29
Shaw & Lesnick 2009	*Lapaula* (crickets)	1	QTL	26
Laturney & Moehring 2012	*Drosophila*	5	deletion mapping	not applicable
Koutroumpa et al. 2016	*Ostrinia* (moths)	1	QTL	470
Bay et al. 2017	*Gasterosteus* (sticklebacks)	2	QTL	302
Merrill et al. 2019	*Heliconius*	3	QTL	146
Blankers et al. 2019a	*Lapaula*	2	QTL	143
Xu & Shaw 2019	*Lapaula*	1	QTL	89
Shahandeh et al. 2020	*Drosophila*	1	deletion mapping	not applicable
Shahandeh & Turner 2020	*Drosophila*	3	QTL	382

* = (minimum), ** = different backcross directions

Do distinct loci influence different aspects of mating behaviors? Or do they interact together to specify complex behavioral patterns? Quantitative trait loci (QTL) mapping studies have thus far suggested mostly additive effects between loci rather than epistatic interactions, though most QTL mapping studies will not have the power to detect epistatic interactions and/or small effect loci, and can implicate only large genomic regions. For instance, Shahandeh & Turner 2020 further dissected a QTL region of large effect underlying differences in pheromone preference between *Drosophila* species, to find that it contained at least two distinct epistatic loci controlling male mate choice. Again, Shahandeh & Turner 2020 tried to disentangle the contribution of different QTLs/loci to behavioral differences. Interestingly, they found that different QTLs control different behavioral aspects of gustatory-guided male choice in two *Drosophila* species: attraction in one species and aversion in the other species, for the same pheromone. Overall, divergent species-specific mating preferences are often controlled by multiple loci, which might control different aspects or sensory modalities of behavioral preferences, or interact to specify behavior, and even single QTL might contain multiple causal, interacting genes.

Although mapping divergent behaviors to genomic regions might be *relatively* straightforward, further dissecting their genetic basis is particularly challenging. As of yet, the resolution of actual genetic changes underlying the evolution of divergent mate preference behaviors remain low, and very few studies have identified likely candidate genes within QTL regions. Fewer still have identified more specific genetic changes linked to divergent mating preferences (and most are restricted to olfactory-guided preferences, Leary et al. 2012, Brand et al. 2020, Unbehend et al. 2021). Although genetic control of behavioral differences is likely complex, these few studies suggest that changes at single loci can explain a large proportion of variation in behavior between species, and possibly trigger a change in mate preference behavior.

2.2. Genomic associations between cues and preference loci.

Mate preference behaviors often evolve in concert with the cues/signals they are based on. The co-evolution of traits under selection, including ecological (cues) and mating (preference) traits, can produce predictable underlying genetic architectures (McKinnon & Pierotti 2010, Svensson et al. 2021). These have been hypothesized to facilitate the evolution of divergent cues and preferences during speciation with gene flow, by reducing/impeding breakdown of adaptive associations between the two by recombination (Kopp et al. 2018). Such architectures include i) pleiotropic alleles governing both cue and preference and ii) physical linkage (proximity on the same chromosome) between cue and preference loci and iii) one allele mechanisms, i.e. where the spread of a single allele results in isolation between populations (Servedio 2009). These include phenotype matching, whereby an allele would "instruct" individuals to mate with those that look like themselves.

Distinguishing these is difficult. To my knowledge no single pleiotropic allele has yet been demonstrated to control divergence in both preference and cue (although for example a desaturase gene was found to affect both sex pheromone production and perception in *Drosophila melanogaster*, Bousquet et al. 2012). Also, there are no clear examples of one-allele mechanisms (Ortiz-Barrientos & Noor 2005 found results indicative of such mechanism in *Drosophila* species, but failed to replicate their findings with different populations/lines, Rosenthal 2017). Nevertheless, physical linkage between preference and cue loci, and possibly pleiotropy, has been reported multiple times (Kronforst et al. 2006, Saether et al. 2007, Pryke 2009, Wiley et al. 2012, McNiven & Moehring 2013, Bay et al. 2017, Blankers et al. 2019a,

Merrill et al. 2019, Xu & Shaw 2019, Xu & Shaw 2021). This suggests that physical linkage or pleiotropy has often been favored by natural or sexual selection to facilitate the evolution of these divergent (co-adapted) traits (although studies also report an absence of physical linkage between preference and cue loci, Limousin et al. 2012, Koutroumpa et al. 2016, Shanandeh & Turner 2020).

2.3. Role of sex chromosomes in the evolution of mate preferences.

Sexes are often dimorphic in form and behavior, including sexual signals and preferences. In many of the studies cited, specific cues and preference are expressed in either one sex. However, sexes are largely constructed from a shared genetic sequence. It is thought that these differences mostly result from changes in gene regulation (Ellegren & Parsch 2007), that unfold during development through diverse mechanisms (Williams & Carroll 2009, Galouiz & Prud'homme 2021), and are eventually controlled/triggered by a difference in sex chromosome composition (for taxa where sex is chromosomally determined). Nevertheless, it is not yet clear if genes underlying mate preferences, or sexual dimorphism more generally (Dean & Mank 2014) are more frequently found on sex chromosomes. For example, in principle, genes across the genome could co-opt sex-determination regulatory machineries.

Nevertheless, given their special mode of inheritance, sex chromosomes are affected differently by forces like selection and drift compared to the autosomes (Ellegren 2011), which might in turn influence where in the genome mate preferences are preferentially encoded. Also, in many taxa, sex chromosomes have reduced rates (or suppression) of recombination, which might "lock" more effectively phenotypic divergence between lineages. Importantly, males and females often differ in optimal (fitness) trait value, and given that they largely share the same genetic sequence, this can lead to conflict between them, i. e. sexual conflict (Arnqvist & Rowe 2013). Theory predicts that sexually antagonistic alleles (where selection acts in opposing directions in the two sexes) will tend to accumulate on sex chromosomes (Dean & Mank 2014). Reviewing the literature, (sex-linkage of mate preference loci) Rosenthal (2017) found a trend for over-representation of female preference loci on the sex chromosome in ZW systems (e.g. in birds Saether et al. 2007, Pryke 2009); however, these findings did not extend to XY sex-determining systems. More generally, evidence for frequent sex-linkage of divergent mating preferences remains mixed (for example Koutruompa et al. 2016 vs. Bay et

al. 2017, Merrill et al. 2019, Blankers et al. 2019a), and how these might be affected by sexual antagonism requires further investigation. Overall, the genetic architecture of divergence in mating preferences will be shaped by multiple evolutionary factors, possibly including for example the location of cue loci.

2.4. Genomic signatures of (behavioral) barrier loci.

Speciation can be seen as a process of increasing genetic divergence between lineages. This process is often not homogenous across the genome. In fact, in the presence of gene flow between lineages, this process tends to be localized to relatively few regions during the early stages of speciation (Nosil et al. 2021). As speciation proceeds, these regions are expected to expand across the genome as isolating mechanisms between lineages accumulate (Wolf & Ellegren 2017). Genetic elements underlying reproductive isolation between populations can produce distinct genomic signatures (Seehausen et al. 2014, Ravinet et al. 2017). These can be leveraged to map the emergence of behavioral (isolating) differences in the genome of speciating lineages.

It is important to note, however, that these genomic patterns (e. g. genetic divergence) can be generated by factors other than barrier loci (CruickShank & Hahn 2014, Wolf & Ellegren 2017, Ravinet et al. 2017). Also, they might not necessarily be detectable if, for example, traits are highly polygenic or epistatic, or if loci effect are small or under indirect selection (Barton & Bengtsson 1986, Jiggins & Martin 2017, Lewis et al. 2020). Overall, the predictions that follow are still crude. Note that these signatures are generally more likely to stand out in whole-genome scans when comparing recently diverged lineages, or lineages with high levels of hybridization (Jiggins & Martin 2017).

Given these caveats, one expected signature of a locus controlling divergent mating preferences under selection, is elevated genetic differentiation (F_{ST}) and divergence (d_{xy}) between populations. This should result from decreased fitness for backcross hybrids inheriting mis-matched mate preference alleles, who will either be less likely to mate/survive or more likely to leave maladapted hybrid offspring. In contrast, other loci not under divergent selection, are expected to flow more freely (recombine) between divergent populations. For example, Brand et al. (2020) identified an odorant receptor tuned to conspecific perfume compounds, associated

with divergent female mate preference in orchid bees, by first identifying loci with pronounced (genome-wide) genetic differentiation and divergence. In parallel to high genetic divergence, a reduction in gene flow (between divergent populations) is expected at the level of barrier loci, and can be observed at the level of candidate mate preference loci (Rossi et al. 2020). Although more direct links between genomic islands of divergence and behavioral mating preferences are still missing, these associations are found in multiple lineages (Poelstra et al. 2014, Hench et al. 2019, Turbek et al. 2021), and constitute a promising approach to identify candidate behavioral barrier loci.

Relatedly, a selective sweep (marked reduction in genetic diversity) might be expected at barrier loci controlling divergent mating preferences. This would either stem from direct selection within species to find/mate with locally adapted individuals, or, in sympatric species, from indirect selection against hybridization between populations (i.e. reinforcement-like processes) (Garner et al. 2018). As an example, both Smadja et al. (2015) and Brand et al. (2020) found signatures of selective sweeps at odorant receptors associated with sexual isolation in hybridizing species.

Finally, given that cue and preferences often co-evolve to produce assortative mating, patterns of linkage disequilibrium (i.e. the non-random statistical association of alleles, Lewontin & Kojima 1960) between preference and cue loci might be expected (Kirkpatric 1982). In a noticeable example, Unbehend et al. 2021 found pronounced linkage disequilibrium between *bab*, the transcription factor they identify controlling for divergence in male mate preference in moth strains, and a *reductase* (Lassance et al. 2010, on a different chromosome) controlling the ratio of pheromone blends that guides the divergent male response. In another example, with incipient fish species that have divergent pigmentation patterns and exhibit visually-based assortative mating, Hench et al. 2019 found marked patterns of linkage disequilibrium between primary photoreceptor genes and pigmentation genes, implicating them as visual preference and cue loci (although this might reflect a more general genetic association between co-evolving visual traits).

3. How do genes act within sensory and neural systems to alter mating preferences?

To understand which genes might control the evolution of mating preferences, it may be useful to put them in context of how they can alter the neural computation of "mate preference". Mate recognition starts with the transduction of signals (e.g. color, odor, etc.) at specialized sensory organs into patterns of neuron firing. In a simplified version, these firing patterns are then passed to the brain for further processing, and integrated and combined with information relating to the animal's internal state and previous experience, to activate the 'correct' motor responses (or update the internal state). Within this network, genes may act at different nodes, to produce a change in mate sensory perception.

Below, I explore these processes by focusing on mate preference evolution mediated through three major sensory modalities: chemosensation, hearing and vision. These play a prominent role in mate preference evolution across taxa, and are the most common modalities studied in relation to mate choice (Ryan & Cummings 2013, Rosenthal 2017). I put a particular focus on where genes might act on the axis between the sensory periphery to higher-processing and decision-making centers in the brain.

3.1. Chemosensation.

Progress in understanding the genetics of mate preference evolution has largely been limited to chemosensory-guided preferences, mediated by changes in chemoreceptors. These mate preferences are mediated by smell (olfaction) and/or taste (gustation), conveyed through sex-pheromones, diverse chemical compounds (volatile and involatile) that are detected by chemoreceptors housed in chemosensory neurons (Touhara & Vosshall 2009, Sato & Yamamoto 2020a). Sex-pheromones can provide a wealth of information, such as mating status (whether an individual has already mated or not) or species identity. Both pheromone and chemoreceptor composition evolve rapidly and often have species-specific profiles (Khallaf et al. 2021, Zhang & Löfsted 2015) that can guide mate recognition.

Odor- and gustatory- guided mate preference behaviors can be tuned by changes in chemoreceptor gene sequences. This type of genetic change can alter what chemical

compounds receptors can bind (del Màrmol et al. 2021), potentially leading to different behavioral responses. As a striking example, engineering olfactory neurons of *Drosophila* males to express a (divergent) moth pheromone receptor makes them respond to moth pheromones (Kurtovic et al. 2007).

Attraction or aversion to chemical compounds can also be altered through changes in chemoreceptor gene expression, including broad changes in expression level, how the expression of different receptors is combined, and which olfactory neurons receptors are expressed in. As an illustration of the latter, expressing an odorant receptor tuned to normally attractant (food) compounds in *C. elegans* in a different olfactory neuron (whose firing is associated with avoidance behaviors), will mediate avoidance instead of attraction behavior towards the same (food) resource (Troemel et al. 1997). This also illustrates that the nervous system can have an innate map for interpreting what the firing of a particular chemosensory neuron corresponds to.

The possibility of flexible tuning of behavioral responses to chemical compounds through changes in chemoreceptor sequence and/or expression has been paralleled by the empirical observation that chemosensory-guided behavioral change often involves an alteration of the peripheral (chemosensory) nervous system (Cande et al. 2013, Zhao & McBride 2020). Multiple studies have linked changes at chemosensory receptors to the evolution of divergent mating preferences between species (Leary et al. 2012, Fan et al. 2013, Ahmed et al. 2019, Brand et al. 2020).

Despite this, divergent neural circuits, downstream of chemosensory neurons (without changes in chemoreceptor response), have also been shown to be involved in the evolution of chemosensory-guided mating preferences (Seeholzer et al. 2018, Khallaf et al. 2020). The genetics underlying these circuitry changes is mostly unknown (but see Shanandeh et al. 2020), except for the transcription factor *bab*, which likely modulates male mate preference in moth populations through an alteration of the olfactory system development and eventually olfactory processing (Unbehend et al 2021). More generally, other divergent chemosensory-guided behavioral preferences (for example for host-use) can involve changes at chemoreceptors (McBride et al. 2014, Prieto-Godino et al. 2017), but also a combination of receptors and downstream circuitry (Auer et. al 2020), or downstream circuitry (including divergent neural activity rather than connectivity, Tait et al. 2021). Overall, it remains unclear if the frequent

associations observed between changes at sensory receptors and the evolution of chemosensory-guided behaviors reflects actual developmental constraints on central circuit evolution, or if they are simply more easily detected (Cande et al. 2013, Zhao & McBride 2020).

3.2. Hearing.

Animals can distinguish different properties of sound (e. g. frequency) through mechanosensory neurons and auditory processing (Theunissen & Elie 2014). Preferences for divergent mating acoustic signals, for example a courtship song or call, have been found in many species and populations (Searcy & Andersson 1986). In this sensory modality, the neural mechanisms underlying divergent (acoustic) preferences (see Wilczynski and Ryan 2010 for anurans) have been suggested to involve mechanoreceptor tuning (to respond only to certain frequencies, for example) or changes in processing/central circuitry.

As the complexity of the divergent acoustic signal increases (e.g. temporal pattern of a song/call), it seems likely that the corresponding neural mechanisms that would produce divergent preference would be encoded in auditory higher-processing centers, where different song- and call-selective neurons have been identified (Theunissien & Elie 2014). For example, Wang et al. (2020) identified descending neurons controlling vaginal plate opening in *Drosophila melanogaster* (virgin) females that respond almost only to conspecific male pulse songs (compared to other seven hetero-specific songs). These descending neurons receive input from central auditory neurons that are partly tuned to conspecific song features, although the neural mechanism that allow females to discriminate conspecific-songs is otherwise unknown.

While specific neurons involved in discriminating species-specific acoustic signals have been identified, hearing is the sensory modality for which we have least insight in the genetics underlying its evolution. To my knowledge only three studies, all in species of cricket, have identified putative candidate genes for divergence in acoustic (female) mate preference: specifically, Blankers et al. (2019a) and Xu & Shaw (2019, 2021) suggest candidate genes involved in neural connectivity and activity/identity (in particular a nucleotide-gated ion channel gene). Interestingly, given that both divergent male song production and female preferences map to the same region, Xu & Shaw (2019) suggest that a shared molecular

mechanism (controlled by pleiotropic alleles) might regulate both male song production and female song detection circuits.

3.3. Vision.

For many animals, an important component of mate attraction is based on visual cues, for example color patterns (Seehausen & van Alphen 1998, Jiggins et al. 2001, Finkbeiner et al. 2014). Vision requires the extraction of visual information through neural processing. In color vision, although some spectral information can already be extracted at the photoreceptors' synapse (Schnaitmann et al. 2018, Yoshimatsu et al. 2021), "color" and its behavioral significance is largely encoded downstream of photoreceptors in the retina and/or brain (by comparing the output of different photoreceptor types that harbor different opsins/respond to different wavelengths, Longden 2016, Song & Lee 2018, Schnaitmann et al. 2020). For example, in *Drosophila*, specific neurons in the optic lobes (visual processing centers) and higher-order processing centers mediate innate spectral preferences for specific wavelengths (Gao et al. 2008, Karuppudurai et al. 2014, Otsuna et al. 2014).

Although one could argue that "most of vision happens in the brain", most research has focused on how variation in opsin genes (which mediate phototransduction) can drive the evolution of divergent color-guided mate preferences. This is likely because the sensory periphery has been more experimentally tractable. In speciation research, this is often studied in the context of sensory bias or drive (Ryan & Cummings 2013, Price 2017, Cummings & Endler 2018), or how perception of signals, including mating signals, is influenced by broader visual adaptation to different light environments. Variation in opsin gene sequence/expression (and more broadly in photoreceptor/retinal mosaics) has been shown to correlate with divergent light environments and divergent mating signals/mate preferences in multiple instances (Hort 2007, for example in birds, Price 2017, in fish, Cummings 2007, Seehausen et al. 2008, Sandkam et al. 2014, or insects Lienard et al. 2021). However, it is not yet clear if opsin genes are a frequent locus for the evolution of visual mate preferences, and more direct links between opsin variation and visual preference behaviors are missing. For instance, while different opsin alleles are associated with female mate preferences in cichlid fish species/populations (Seehausen et al. 2008), their segregation in hybrid crosses do not correlate (with statistical significance for QTL mapping) with mating preference behavior, suggesting that also other

downstream or peripheral mechanisms contribute to divergence in female preferences (Rosenthal 2017). The study of opsin genes, photoreceptor and retinal variation (including variation in filtering pigments and opsin distribution in photoreceptors/retina, van der Kooi 2021) continues to provide important insights into the evolution of mate choice (Ryan & Cummings 2013). Nevertheless, it seems likely that color pattern-guided visual mate preferences could also evolve through genetic changes that alter visual processing or integration (Rossi et al. 2020). More generally, color is only one of many visual features that could guide mating preference, which require computations downstream of photoreceptors to be detected (e.g. motion, Borst et al. 2020, for example, some neurons might fire only in response to conspecific movement types). Furthermore, specific visual features are conveyed distinctly from visual processing to higher centers to activate specific behaviors (Wu et al. 2016). It seems likely therefore that genetic changes that would alter visual processing and/or integration might more specifically tune visually-guided mating preferences (albeit through yet unknown genetic mechanisms).

3.4 Multimodal integration and neuromodulation.

Mate preferences and mating decisions often rely on the integration of signals from different sensory channels (Auer & Benton 2016). For example, courtship initiation in *Drosophila* males is modulated by a combination of gustatory and olfactory cues (Clowney et al. 2015), female *tungara* frogs are more attracted to visual and acoustic stimuli when combined than if stimuli are presented independently (Taylor & Ryan 2013), and in swordtail fish female mating decisions correlate with early gene expression (neural activity) in brain regions associated with multimodal sensory integration (Wong et al. 2012). Although divergence in how mating cues are integrated might constitute an important neural mechanism underlying divergent mating preference, genes that might underlie such a mechanism are virtually unknown. To my knowledge, only a single study has identified a gene (the transcription factor *fruitless*) that could influence female preference (rejection) behaviors in *Drosophila* by modulating the integration of signals across multiple sensory modalities (Chowdhury et al. 2020).

Finally, how sensory inputs are processed can depend on the state/activity of other neurons (including downstream neurons), that can actively alter sensory processing (through the exchange of molecules between neurons). This type of feedback mechanism, known as

neuromodulation, is another candidate mechanism underlying the evolution of behavior (Katz & Warrick 1999). Through neuromodulation, it is possible, for example, to reconfigure the receptive fields of sensory neurons (what sensory inputs they will respond to), including those of photoreceptors (Cheng & Frye 2020). In this way, the same neural circuits can flexibly modulate the expression of a behavior depending on sensory context (Bargmann 2012, Siju et al. 2021). *Drosophila* males provide an example of sensory processing modulation. Here, the detection of pheromonal cues from conspecific females activates a central circuit (P1 neurons) in the male brain controlling arousal state. This circuit in turn enhances the gain (responsiveness) of upstream visual processing neurons, so that males will track/pursuit any moving object (including a previously unattractive rotating magnet) (Sten et al. 2021). Although neuromodulation is often studied in relation to how behavior can be flexibly expressed in individuals depending on sensory context, this could also underlie differences in decision-making between populations/species. For instance, regulatory mutations of a tyramine-receptor (expressed in olfactory neurons) underlie differences in exploratory behavior in *C. elegans* strains, likely by encoding a differential sensitivity to arousal states / neuromodulatory mechanisms (Bendesky et al. 2011). Neuromodulation might also have relevance for mate preference behaviors, as it is involved, for example, in modulating visual mate preferences for familiar over unfamiliar individuals in Medaka fish (through hormonal regulation, Okuyama et al. 2014).

3.5. The locus of behavioral evolution.

Genes underlying behavioral variation and evolution are often associated with functions modulating neuronal activity. These behavioral genes include: a protein kinase (signalling protein) for foraging behavior in *Drosophila* (larval) strains (Osborne et al. 1997), a neuropeptide receptor for foraging/aggregative behaviors in *C. elegans* (DeBono & Bargmann 1998), a *syntaxin* (mediating synaptic vesicle-release) for social behavior in halictid bees (Kocher et al. 2018), and an estrogen receptor for aggressive behavior in sparrows (Merritt et al. 2020). Although these include very diverse behavioral patterns (and taxa) this seems to suggest that genetic changes altering neuronal activity or wiring, besides sensory receptors, constitute a likely route for behavioral evolution, including mate attraction/aversion behaviors. Mating decisions are often based on integration across sensory modalities, and likely take place

in higher-processing centers (DeAngelis & Hoffman 2020, Ryan 2020). Therefore, genetic alterations of central circuitries are another candidate mechanism for the evolution of mate preferences. Nevertheless, alterations of central circuitry might also be pleiotropic, and displace adaptive links between perception and action in the brain. Overall, mate preference alleles could act at different places between detection at the sensory periphery and implementation of motor output in the brain. Where they will act on this axis, will probably be the result of evolutionary forces that will favor alleles with minimal pleiotropic effects (affecting perception of the wider environment), while maximizing shifts/effectiveness in (adaptive) mate preferences.

4. Development of behavioral differences.

Organisms are constructed through genetic programs, that unfold sequentially during development. During this process, the function of a gene is defined by where and when it is acting, and so by which other genes are acting together with it in that cell/tissue type, in that particular moment, i.e. its regulatory network (Levine & Davidson 2005). Many advances in our understanding of the genetics of morphological diversity have emerged from within a developmental biology framework (Carroll 2008). Most notably perhaps, it was found that animal taxa possess similar genetic toolkits (genes), and that phenotypic evolution is therefore often caused by a change in how these genes interact during development (expression), rather than changes in the toolkit genes themselves. A developmental approach might also help conceptualize behavioral diversification. In this context, I address three questions: What genetic changes underlie behavioral evolution? When during development do behavioral genes act? And, what genetic mechanisms are responsible for sex differences in behavior?

4.1. What genetic changes underlie behavioral evolution?

Morphological evolution is most often linked to changes in gene regulation during development (e. g. of transcription factors and signalling molecules) (Prud'homme et al. 2007). This is thought to be favored by selection because gene regulatory changes are often less likely to cause fitness penalties due to pleiotropic effects compared to changes in protein-coding

sequences. Because it is still unknown whether morphological and behavioral changes share similar ontogenetic principles (i. e. encoded during embryonic development, see later), it is premature to say whether these genetic principles are also broadly applicable to behavioral evolution. Nevertheless, the available data suggest that if behavioral evolution relies on changes at sensory receptors (arguably not very pleiotropic in the case of chemosensory systems), the underlying genetic changes can be protein coding (Brand et al. 2020, Auer et al. 2020), or a combination of coding and regulatory changes (McBride et al. 2014, Prieto-Godino et al. 2017) (note that these are all studies on olfactory-guided behaviors). As behavioral change involves alterations of neural activity/connectivity pattern, it might instead rely on changes in gene regulation (Bendesky et al. 2017, Kocher et al. 2018, Merritt et al. 2020, Unbehend et al. 2021).

Among the regulatory changes, alternative splicing, which generates transcript variation with functional relevance to neural processes (Li & Black 2007), has been linked to behavioral evolution (Ding et al. 2016), including divergent mate preferences (Chowdury et al. 2020). Also, non-coding regulatory RNAs, which have important functions in neural development (Cochella & Hobert 2012), including sexual behavior (Keshavarz & Tautz 2021), might constitute another genetic route to behavioral evolution. Finally, it seems likely that both protein-coding or regulatory changes underlying behavioral evolution might follow gene duplication events, for example of sensory receptors (Hort 2007), although to my knowledge gene duplications (and non-coding RNAs) have not yet been specifically linked to the evolution of mate preference behaviors.

Further research is needed to establish whether mate preferences evolve more often through gene/splicing regulatory changes or protein-coding changes. More importantly, identifying these changes is a key step towards understanding how behavior is genetically constructed.

4.2. When during development do behavioral genes act?

Behavior emerges from neural (and sensory) systems, that are constructed through embryonic and adult development. But when are innate behavioral changes established in this process? Earlier in development, when a specific neuronal cell type is specified and neural circuits assembled (Perry et al. 2017, Tosches 2017)? Or in the adult form, for example through the

refinement of neuronal activity within neural circuits? Identifying the developmental timing of behavioral change is challenging. Nevertheless, it is important for understanding the developmental principles guiding behavioral evolution, such as whether there are any key developmental transitions (e.g. assembly of neural circuits, onset of the behavioral pattern in adults) associated with it.

Gene regulatory changes underlying behavioral variation/evolution have often been revealed by studying gene expression differences in adults (McBride et al. 2014, Bendesky et al. 2017, Kocher et al. 2018). Therefore, genetic alteration of sensory and neural functioning in adults would appear an important candidate mechanism for behavioral change. Despite this, it remains unclear if gene expression differences found in adults might also be found/exert effects earlier during development. More generally, with the exception of changes at sensory receptor genes, it is largely unknown how genetic changes act on neural systems to modulate behavior. For instance, in a noticeable example, specific mutations linked to alternative splicing of an ion channel gene have been shown to control differences in a motor pattern (song) in *Drosophila* species (Ding et al. 2016). However, ion channels could modulate behavior either by mediating circuit activity in adults, or through axon guidance during development (or both), among other processes. In another example, divergent male moth pheromone preferences have been found to be encoded in the introns of the transcription factor *bric a brac* (*bab*), but *bab* expression in brains and antennae of divergent moth strains oscillates considerably throughout pupal and adult stages (Unbehend et al. 2021). Although it seems likely that *bab* exerts a behavioral effect by influencing the development of olfactory neuron identity or connectivity, the developmental timing of the associated behavioral change remains unclear.

As a last note, it is perhaps worth noting that the same allele might have diverse behavioral effects depending on which developmental stage it is acting in (and therefore which other genes are acting together with it). For example, in some species of *Peromyscus* mice, the burrow (extended phenotype) of juvenile and adult mice individuals can be quite different (i.e. length). Nevertheless, a shared genetic region modulates both juvenile and adult burrowing behavior (Metz et al. 2017). Of course, juvenile and adult behavior might be modulated by different genetic elements within the same genetic region, but this also raises the possibility that the same pleiotropic allele could produce different behavioral outputs depending on when it acts during development.

4.3. What genetic mechanisms are responsible for sex differences in behavior?

As discussed above, sexual dimorphism is often observed in mating behavior and preferences. Because sexes share most of the genome, this phenotypic variation is thought to be mostly controlled by gene regulatory differences. Understanding how these regulatory mechanisms specify sex-specific behaviors can inform us of the genetic changes underlying mate preference evolution.

One of the best examples of regulatory control of sex-specific mating behavior is that of the transcription factor *fruitless* in *Drosophila*. *fruitless* is spliced differently in male and female flies as a result of sex-determination regulatory cascades, and eventually determines much of sex-specific mating behavior (Sato & Yamamoto 2020b). Remarkably, engineering females flies to express male-splice variants of *fruitless* is sufficient to make them perform aspects of the elaborate male courtship ritual (Manoli et al. 2005). Some mutations of the *fruitless* locus also affect sexual preference (Sato & Yamamoto 2020b). In fact, sex differences in mating behavior are often experimentally reversible. For example, female moths acquire male mate-seeking behavior (pheromone-guided) upon grafting male antennae on them (Schneiderman et al. 1986).

Even complex central circuitry for sex-specific behaviors can be latent in the opposite sex (Rezával et al. 2016). For example, female mice deficient for an ion channel (*Trpc2*) (that gates odor-guided behaviors) display male-like sexual behaviors, including male-like vocalizations (Kimchi et al. 2007). If a gene regulatory switch can control sex-specific behavioral repertories (and often sexes differ in mating preference), does this suggest mate preference evolution might evolve by analogous gene regulatory changes? A short answer is not necessarily: genes that control mating behavior differences between the sexes are not necessarily those that control variation between species (for example *fruitless* Cande et al. 2014). Also, if they (again *fruitless*) do influence divergent (female) species preference behaviors, this might not be modulated by sex-specific regulatory (splicing) changes (Chowdury et al. 2020). In fact, it is possible that the same regulatory changes, as well as protein-coding functional effects, could be integrated differently in the sexes.

Nevertheless, gene regulatory changes and their functional effects can often be sex-specific, including for genes involved in behavioral evolution (for example *vasopressin* for male parental care in *Peromyscus* mice, Bendesky et al. 2017). Therefore, the genetic switches that

govern the evolution of mate preferences are likely to be linked to those that specify differences between the sexes. To cite another example concerning fruitless, the mate preferences of two Drosophila male species are controlled by a neuronal population (P1) in their central brain (Seeholzer et al. 2018). Because the functioning of this neuronal cluster in males is specified by (splicing of) fruitless, it would make sense that the genetic change that govern this behavioral shift is a male-specific regulatory change that co-opted (downstream of) fruitless (Shanandeh et al. 2020) or its network.

Furthermore, experimental designs based on predictions that gene regulatory differences will be detectable only in one sex (displaying the sexually dimorphic phenotype of interest), have proven fruitful in identifying genes underlying phenotypic differences between species (Combs et al. 2018). Similar designs might assist the identification of gene expression changes linked to sex-limited mate preference evolution. To conclude, the genetic logic behind the evolution of sex and species differences in behavior might show similarities (genetic mode) and differences, as the potential for expressing alternative sexes, but not species, must remain latent in the same genome.

5. Concluding remarks and study system.

Although often shaped by experience and context-dependent (for example depending on environmental conditions or physiological state), all complex behaviors, including mating preference behaviors, require "pre-established" wiring of neural systems by genetic programs. In this light, I have reviewed emerging genetic patterns and mechanisms underlying the evolution of mate preferences. Overall, identifying the genetic mechanisms underlying the diversification of species-specific mating preferences will further our understanding of how new species and their diverse behaviors can be generated during development, through the activity of neural circuits, and across evolutionary time. Despite this, the genetics underlying the evolution of species mate preference behaviors remain largely unknown. In particular, almost nothing is known about how visual mate preferences (and divergence in visually-guided behavior more broadly) can be encoded in the DNA. In this thesis, I address this gap by tackling the genetics underlying shifts in mating behavior between species of *Heliconius* butterflies.

5.1. The study system.

Heliconius is a genus of Neotropical butterflies. Its species are known for both their striking diversity and mimicry of warning color patterns. In particular, distantly related *Heliconius* species often mimic each other to share the cost of advertising predators of their distastefulness (i.e. Müllerian mimicry). Among more closely related species and within species however, there is often striking diversity of color patterns (Jiggins 2017). *Heliconius* have been a window into our understanding of the genetics of adaptation (especially of adaptive coloration) and speciation (Merrill et al. 2015, Van Bellenghem et al. 2021). Their aposematic color patterns are under strong, frequency-dependent selection, where higher predation rates are observed for both non-local patterns and intermediate hybrid patterns (Mallett & Barton 1989, Merrill et al. 2012**,** Chouteau et al. 2016). Therefore, divergent selection on color patterns imposes strong reproductive barriers between species, and the evolution of divergent colorations is often thought to be key for driving speciation in *Heliconius* (McMillan et al. 1997).

Beside shifts in color patterns between closely related species, speciation in *Heliconius* is also often associated with shifts in habitat use (McMillan et al. 1997, Estrada & Jiggins 2002). These distinct habitats often pose contrasting sensory environments that have to be met by behavioral/neural adaptation (plastic or heritable) (Merrill et al. 2013, Montgomery & Merrill 2017). In turn, heritable adaptation of neural systems to local conditions can contribute to reproductive isolation, if species hybrids experience a disruption of neural function, mismatched to the local sensory conditions of either parental species (Montgomery et al. 2021). Therefore, as it is expanded upon in chapter 3, neural divergence between lineages can also contribute to speciation in *Heliconius*.

5.2. *Heliconius* butterflies for studying the evolution of visual mate preferences.

Heliconius butterflies exhibit complex visually-guided recognition and behaviors. For example, they regularly visit the same flowers and host plants (for feeding and oviposition), by following the same route across complex forest environments (i. e. trap-lining behavior, Jiggins 2017, implying considerable visual memory). *Heliconius* possess color vision, and can distinguish artificial flowers/food resources of different colors, ranging from the ultraviolet to the red spectrum (Swihart 1971, Zaccardi et al. 2006, Finkbeiner & Briscoe 2021). Color vision can be achieved by comparing the output of two or more chromatic channels (e.g.

photoreceptors with different photosensitive-pigments/opsins), which in *Heliconius* can comprise as much as four or five channels, with some variation within the genus, and in some species, between sexes (McCulloch et al. 2017, McCulloch et al. 2021).

The aposematic cues on the wing of *Heliconius* are not only used to signal distastefulness to predators, but they are also used during mate recognition (Crane 1955). This is particularly clear for males, which almost invariably prefer to court and attempt to mate live females and artificial models that share their own warning pattern (e.g. Jiggins et al. 2004, Chamberlain et al. 2009, Merrill et al. 2011a, Merot et al. 2017). This has led to varying degrees of premating isolation between taxa (with divergent patterns), ranging from polymorphic populations (Chamberlain et al. 2009), to incipient species (Hausmann et al. 2021, Jiggins et al. 2004), to species arguably at the later stages of divergence (Jiggins et al. 2001). These isolating behaviors are probably among the first reproductive barriers to evolve during speciation in *Heliconius* (Jiggins et al. 2004, Merrill et al. 2011a), and together with differences in habitat use, they constitute crucial pre-mating barriers between species (Jiggins 2008). Other cues, transmitted through other sensory modalities, likely also contribute to differences in male attraction, but color patterns often play a major role (Jiggins 2017).

Despite an impressive literature on color-based preferences in *Heliconius*, many mechanistic questions about these mating preferences remain unanswered. For example, the exact visual cues used by males during mate recognition remain poorly understood (although forewing color probably plays an important role for some population and species, Kronforst et al. 2006, Merrill et al. 2011b, Finkbeiner et al. 2014). Finkbeiner et al. 2014 found that color followed by pattern (how colors are arranged), is the best predictor of whether *Heliconius erato* males will court a model of the female wings, but it is not known if this extends to other species/populations as well. Importantly, although genomic regions associated with these behavioral shifts have been identified (Kronforst et al. 2006, Chamberlain et al. 2009, Merrill et al. 2019), the specific genes involved and their mode of action remain unknown. Also, although the duplication of an UV opsin gene correlates at the genus level with the presence of UV cues on the wings (Briscoe et al. 2010), it is unknown whether behavioral shifts between *Heliconius* species and populations are mediated by changes at the sensory periphery (e. g. opsin genes) or in visual processing or integration.

Overall, the rapid diversification of these sophisticated visual behaviors across *Heliconius*, coupled with state-of-the-art genomic resources (Davey et al. 2016) and possibilities for genome editing (for example, Concha et al. 2019), makes *Heliconius* a great system to investigate the genetic principles underlying the diversification of visual attraction behaviors.

Results

Paper I

Visual mate preference evolution during butterfly speciation is linked to neural processing genes.

Matteo Rossi, Alexander E. Hausmann, Timothy J. Thurman, Stephen H. Montgomery, Riccardo Papa, Chris D. Jiggins, W. Owen McMillan & Richard M. Merrill

Nature communications (2020), 11, 4763
https://doi.org/10.1038/s41467-020-18609-z

Manuscript II

Signatures of adaptive introgression implicate a *regucalcin* in the evolution of visual mate preference behaviours.

Matteo Rossi, Alexander E. Hausmann, Markus Moest, Lina Melo, Chi-Yun Kuo, D. Shane Wright, José Borrero, Mauricio Linares, W. Owen McMillan, Carolina Pardo-Diaz, Camilo Salazar, Richard M. Merrill

Abstract

Animals can evolve convergent behaviours to meet similar demands imposed by the environment. However, the genetic mechanisms underlying behavioural evolution remain largely unknown. Here, we study the molecular underpinnings of convergent, adaptive visual behaviours in *Heliconius* butterflies. These tropical butterflies have repeatedly evolved mimetic colour patterns, which are also used as mating cues. In the *H. melpomene* group, there is evidence that red pattern mimicry in different species has evolved via adaptive introgression. In this study, we provide evidence that adaptive introgression also underlies the evolution of visual preference behaviours. We first show that divergent visual mate preferences across this butterfly group are associated with the same genomic location (tightly linked to the colour cue locus). We then find signatures of adaptive introgression between red-preferring butterflies in this same region, at the level of previously identified candidate behavioural genes. These include a *regucalcin*, whose gene regulation is strongly linked to preference behaviours. Overall, we find evidence that behavioural alleles crossed the species barriers to facilitate adaptation.

Introduction

Understanding the genetic causes of convergent evolution has been a long-standing goal of evolutionary biology (Stern 2013). In recent decades, it has been appreciated that convergent phenotypes in distinct lineages often evolve through independent reuse of the same genes (Stern & Orgogonzo 2008, Stern 2013, Martin & Orgogonzo 2013). However, convergent phenotypes may also arise non-independently through adaptive introgression (i.e. the exchange of alleles between otherwise separately evolving lineages). While both these scenarios have been documented for a number of morphological traits (e.g. Prud'homme et al. 2006, Jones et al. 2018, Semenov et al. 2021), much less is known about the genetics underlying convergent behavioural phenotypes, or behavioural evolution more broadly (Arguello & Benton 2017).

Incomplete reproductive barriers between species can allow the exchange of adaptive alleles through hybridization, which can facilitate both adaptation and speciation (Abbott et al. 2013, Marques et al. 2019). For example, the shuffling of allelic variants of an opsin (primary photoreceptor) gene through hybridization may have facilitated diversification in cichlid fish by allowing them to adapt to different light environments (Meier et al. 2018). However, key behavioural data are still missing. In particular, almost nothing is known about the genetics underlying the evolution of mate preference behaviours, especially that of visually-guided ones, and a direct link between genes, behavioural phenotypes and adaptive introgression has not yet been demonstrated.

Heliconius butterflies are well known for their diversity of warning colour patterns. These are often associated with Müllerian mimicry, where distantly related species converge on the same warning signals. Colour pattern variation in this genus is modulated by a few loci of major effect (Reed et al. 2011, Martin et al. 2013, Nadeau et al. 2016, Westerman et al. 2018), which have been independently reused across the genus (although through co-option of different genetic networks, Concha et al. 2019), leading to near-perfect mimicry of colour pattern signals between distantly related species. Among more closely related species, however, there is substantial evidence that mimicry alleles have often evolved via adaptive introgression of colour pattern alleles across the species boundaries (Dasmahapatra et al. 2014, Jay et al. 2018, Edelman et al. 2019, Kozak et al. 2021).

In addition to being under selection due to mimicry, warning colour patterns are used as mating cues (Crane 1955). *Heliconius* males almost invariably prefer to court live females and artificial models that share their own colour pattern (e.g. Jiggins et al. 2001, Jiggins et al. 2004, Kronfrost et al. 2006, Mérot et al. 2017, Hausmann et al. 2021). We have started to resolve the genetic bases of these shifts in visual behaviours. In particular, the divergent mating

behaviours of red *Heliconius melpomene* and white *H. cydno* are modulated (at least in part) by three major quantitative trait locus (QTL) (Merrill et al. 2019), and a few candidate genes have now been identified (Rossi et al. 2020). Notably, the best supported of these behavioural QTLs is tightly linked to the colour pattern gene *optix* (Merrill et al. 2019), which is responsible for the forewing colour switch between *H. melpomene* and *H. cydno*, and more generally the presence of red colour pattern elements in *Heliconius* (Reed et al. 2011). Physical linkage of genes controlling associated ecological (cue) traits under divergent selection and mating (preference) traits is expected to facilitate speciation in the face of gene flow (Felsenstein 1981).

Here, we expand genetic analyses of mate preference behaviours to *H. timareta*, a close relative to white/yellow *H. cydno* (and with a similar ecological niche, Montgomery et al. 2021), but often sharing convergent red colour pattern elements with the more distantly related *H. melpomene*. There is considerable evidence that taxa within the *H. timareta* group have repeatedly acquired the red colour pattern alleles of local *H. melpomene* populations (east of the Andes) through adaptive introgression of *optix* alleles (Pardo- Diaz et al. 2012, Dasmahapatra et al. 2014, Wallbank et al. 2016, Martin et al. 2019, Kozak et al. 2021). Here, we hypothesize that *H. timareta* also acquired an attraction to red butterflies through adaptive introgression from *H. melpomene*. In particular, we test two key predictions of this hypothesis: first, that the same genomic location on chromosome 18 (responsible for variation in mating preferences between *H. cydno* and *H. melpomene*) also contributes to differences in mate preference phenotypes between *H. cydno* and *H. timareta*; and second, that patterns of genetic variation between these populations suggests a history of adaptive introgression at the preference QTL.

Results

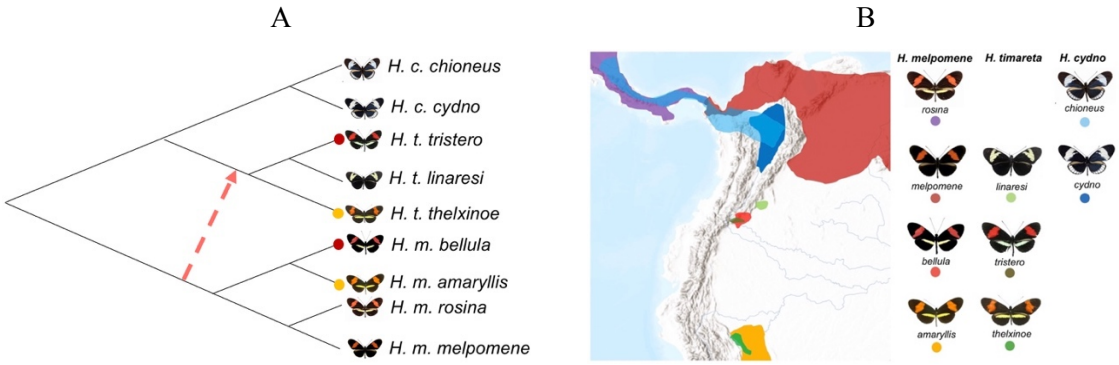
Convergent visual mate preference behaviours in *H. melpomene* and *H. timareta*.

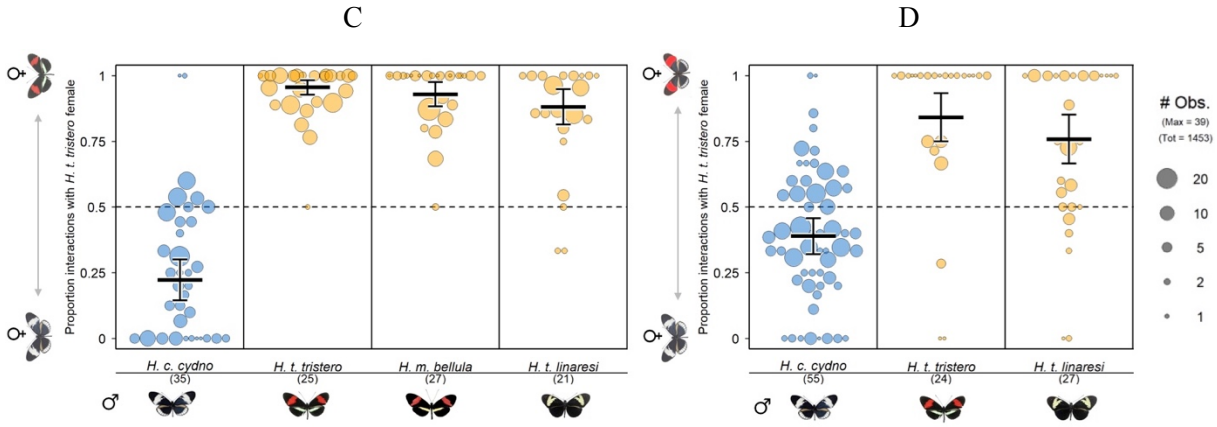
We first assayed male mate preference behaviours across three Colombian *Heliconius* populations within the *melpomene-cydno* group (**Figure 1A, Figure 1B**) in standardized choice trials (Merrill et al. 2019): *H. cydno cydno*, which has a white (or yellow) forewing band and is found on the western slopes of the Andes, as well as *H. melpomene bellula* and its sympatric co-mimic *H. timareta tristero*, which both have a red forewing band and are found on the eastern slopes of the Andes. When simultaneously presented with a white virgin *H. cydno* and a virgin *H. timareta* female, we found that both red *H. t. tristero* and *H. melpomene bellula*

have a strong preference for red *H. t. tristero* females over white *H. c. cydno* females, as compared to *H. c. cydno* males (differences in proportion of time spent courting the red female/effect size: *H. t. tristero* - *H. c. cydno* = 0.73 [0.65 – 0.82], z = 17.11, p <0.001; *H. m. bellula* - *H. c. cydno* = 0.71 [0.61 – 0.80], z = 14.89, p <0.001). These results closely mirror those of our previous studies investigating male preference in Panamanian populations of *H. cydno* and *H. melpomene* (Merrill et al. 2019, Rossi et al. 2020). Notably, we were unable to detect any difference in preference between either of the two red Colombian species (*H. t. tristero* - *H. m. bellula* = 0.03 [-0.03 – 0.08], z = 0.98, p = 1.000, **Figure 1C**).

In order to confirm that these preferences are visually-guided, we repeated these experiments, but this time with two *H. c. cydno* females, one of which was manipulated with a Copic R05 red marker pen so that her dorsal forewing band matched the red colouration of *H. t. tristero*, taking into account *Heliconius* colour vision models (based on *Heliconius* photoreceptor cell sensitivities, **Supplementary figure 1**). The dorsal forewing band of the other *H. cydno* was painted with a Copic 0 transparent marker pen, which incorporates the same chemical solvent (pers. comm. Copic Ciao). Once again, the *H. t. tristero* males had a stronger preference for females with a red forewing than the *H. c. cydno* males (effect size: *H. t. tristero* - *H. c. cydno* = 0.45 [0.34 – 0.57], z = 7.59, p <0.001) (**Figure 1D**). Although differences in male preference are reduced compared to the previous experiment (perhaps because we only manipulated the forewing colour, see figure 1D), these results suggest that the divergent mating preferences of *H. c. cydno and H. timareta* are largely driven by the presence of red on the forewing.

Figure 1





(A) phylogeny of *Heliconius* subspecies analyzed in this study (after Mérot et al 2013 and Arias et al. 2017), circles of the same colour indicate sympatric co-mimics. Dashed arrow indicates putative introgression event of red phenotypes (note that gene flow between sympatric co-mimics is still ongoing). **(B)** Geographical ranges of *H. melpomene*, *H. timareta* and *H. cydno* subspecies in this study (adapted from Rosser et al. 2012). **(C)** Proportion of courtship time directed towards red female types over white female types by males of different species (blue = *H. cydno cydno*, orange = *H. timareta tristero*/*linaresi* or *H. melpomene bellula*). Males allowed to choose between red *H. t. tristero* female and white *H. c. cydno* female. **(D)** Males allowed to choose between red painted *H. c. cydno* female and white (transparently painted) *H. c. cydno* female. In graphs C and D dot size is scaled to the number of total minutes a male responded (chase or courtship) to either female type (a custom swarmplot was used to distribute dots horizontally). Estimated marginal means (EMMs, from generalized linear mixed models) and their 95% confidence intervals are displayed in black for each species. A horizontal, black, dotted line has been drawn at a y-value=0.5, indicating an equal amount of courtship time directed towards the two females.

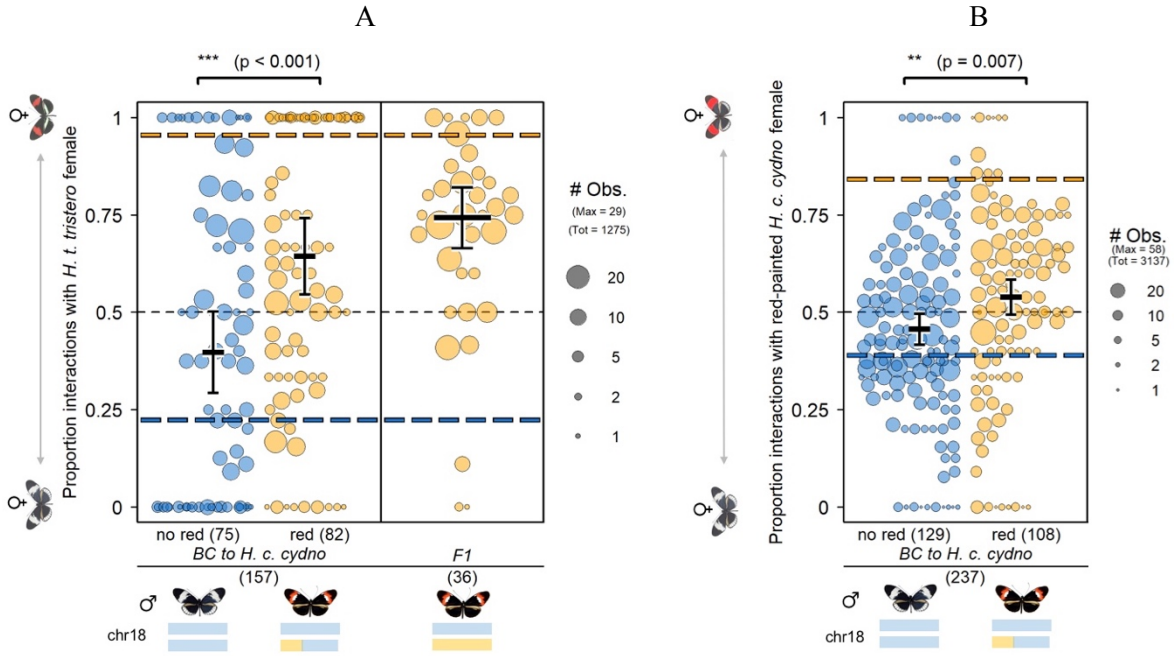
A shared genetic region underlies visual preference behaviours of *H. melpomene* and *H. timareta*.

Our previous work identified a major effect QTL on chromosome 18 for differences in male preference behaviour between *H. cydno* and *H. melpomene*, which segregates with the presence of the red forewing band (due to tight linkage with the major colour pattern gene *optix*) (Merrill et al. 2019). To determine whether this genomic region also contributes to differences in male mate preference behaviours between the more closely related (but allopatric) species *H. cydno* and *H. timareta* – as predicted by our hypothesis of adaptive introgression – we tested whether forewing colour is also associated with male preference behaviours in crosses between these species.

In choice trials with paired red *H. timareta* and white *H. cydno* females we found that F1 (*H. c. cydno* x *H. t. tristero*) hybrid males prefer to court red *H. t. tristero* females, similarly to *H. t. tristero* males (**Figure 2A**), implying that the *H. timareta* preference alleles are largely dominant over the *H. cydno* alleles. This mirrors the results of our previous experiments between with *H. cydno* and *H. melpomene*, where the F1 hybrid males also prefer to court the red parental phenotype (Merrill et al. 2019, **Supplementary Figure 2**). Based on these results we generated backcross-to-*cydno* broods and assayed the resulting males (n=157) in the same way. Because the presence of the red band is dominant over its absence in the *melpomene-cydno* group, we could infer genotype at the *optix* locus, and tightly linked regions on chromosome 18, by inspecting the forewing band colour in these hybrids. Specifically, hybrid individuals with a red band are heterozygous for *H. t. tristero/H. c. cydno* alleles (*Bb*), and individuals lacking it are homozygous for the *H.* c. *cydno* allele (*bb*). This allows a test of whether this genomic region influences variation in male preference based on wing pattern phenotype alone (Merrill *et al.* 2011).

As in our previous crosses between *H. cydno* and *H. melpomene* (Merrill et al 2011; Merrill 2019), we found that genotype at the *optix* colour pattern locus is a strong predictor of the relative time hybrid males spent courting red *H. t. tristero* or white *H. c. cydno* females (difference in proportion of time spent courting the red female: backcross with red – backcross without red = 0.25 [0.10 – 0.39], z = 3.33, p <0.001) **(Figure2A)**. Notably the effect size observed here is almost identical to the one seen in backcross hybrids between *H. cydno* and *H. melpomene* (i.e. ~0.25 in absolute terms, Merrill *et al*. 2019, **Supplementary Figure 2**). Given the tight linkage of *optix* (1.2cM) and the previously identified QTL peak (in *H. melpomene-H. cydno* crosses), we would expect just ~2 recombination events between these two loci across the 157 hybrids tested. Nevertheless, to confirm the genotype status of backcross hybrids at candidate behavioural genes in the QTL peak on chromosome 18, we PCR-amplified an intronic/exonic segment of *regucalcin1* in the same hybrid (and parental) individuals. The presence of indels between *H. c. cydno* and *H. t. tristero* in this segment, allowed us to track the allele status at the QTL peak by PCR-fragment size (**Supplementary Figure 3**), to confirm that *H. t. tristero* alleles segregated with red colour phenotype, and to find 3 recombinants between *optix* (red colour) and the QTL peak (*regucalcin1*) among backcross hybrid individuals.

Figure 2



A) Proportion of courtship time directed towards the red *H. t. tristero* female by backcross to *H. c. cydno* males and F1 hybrid males. Orange points represent individuals that are heterozygous (*i.e. cyd*/*tri*) and blue points represent individuals that are homozygous for *H. cydno* alleles at optix on chromosome 18 (*i.e. cyd*/*cyd*). Dot size is scaled to the number of total minutes a male responded to either female type (a swarmplot was used to distribute dots horizontally). Estimated marginal means (EMMs) from GLMM models and their 95% confidence intervals are displayed by horizontal and vertical black bars. EMMs of the parental species are displayed as dashed horizontal lines in blue (*H. c. cydno*) and orange (*H. t. tristero*), for reference. **B)** Backcross hybrid males, segregating at chromosome 18, allowed to choose between red painted *H. c. cydno* female and white (transparently painted) *H. c. cydno* female.

To confirm that the QTL on chromosome 18 modulates visual mate preference, we next assayed mate preference behaviours of backcross-to-*cydno*-hybrids (n=237) towards transparent- and red-painted *H. c. cydno* females (originally white, as described above). Once again, we found that males heterozygous for *H. t. tristero* and *H. c. cydno* alleles at the *optix* locus (i.e. *Bb*) court red-painted females (relative to white control) more frequently than their brothers homozygous for the *H.* c. *cydno* allele (i.e. *bb*) (effect size = 0.08 [0.02 – 0.14], z = 2.71, p = 0.007). Although the difference between backcross hybrids in courtship time directed towards the (artificial) red *H. c. cydno* vs. white *H. c. cydno* females (0.08) is reduced compared to that observed towards red *H. t. tristero* vs. white *H. c. cydno* females (0.25), together these

results suggest that the same genomic region on chromosome 18 modulates mate preferences across the *melpomene-cydno* group, and that it has an effect on visual preference behaviour.

Increased admixture between *H. melpomene* and red *H. timareta* at the behavioural QTL. If variation in mate preferences between *H. cydno* and *H. timareta* associated with the QTL on chromosome 18 is the result of introgression from *H. melpomene* to *H. timareta*, we would expect an increase of shared alleles between these two species at the QTL. To test this, we analysed admixture proportion using the f_d statistic (based on the ABBA-BABA test, Martin el. al 2015) between *H. m. bellula* and H. t. *tristero* in 20kb windows across the QTL candidate region (previously determined in *H. melpomene-H. cydno* crosses in Merrill et al. 2019). We found that the two highest admixture peaks at the QTL region (top 5% quantile across chromosome 18, **Supplementary figure 4)** are located in the region upstream of *optix*, where *optix* regulatory activity is putatively encoded, and within the QTL peak (i.e. the region of greatest statistical association with male preference). Notably the latter was coincident with three (of a total of 5) candidate genes we have previously hypothesised to influence divergent preference behaviours between *H. melpomene* and *H. cydno* (Rossi et al. 2020) **(Figure 3)**, based on either expression or protein coding differences.

It is possible that our results are explained by variation in recombination rate across the genome, as increased local recombination rates can drive higher admixture proportions (by facilitating the removal of deleterious variants or retention of beneficial variants that introgress between species, Martin et al. 2019). However, recombination rate (estimated from linkage maps from hybrid crosses *H. melpomene* x *H. cydno*, Davey et al. 2017) decreases uniformly towards the end of chromosome 18 along the QTL (**Supplementary figure 5**), suggesting that the patterns we see are independent of local recombination rate.

Admixture patterns are also repeatable across geographical populations, as we also observed admixture peaks at both the behavioural QTL peak and *optix* in comparisons between the sympatric (red) Peruvian subspecies *H. melpomene amaryllis* and *H. timareta thelxinoe* (**Figure 1D,** top 1% quantile across autosomes, **Supplementary figure 4**). These subspecies are co-mimetic, and males do not *visually* distinguish females of either species during courtship (Mérot et al. 2015).

To further corroborate whether *H. t. tristero* and *H. m. bellula* share a recent history of gene flow at the behavioural QTL peak (and *optix*), we used *Twisst* (Martin & Van Belleghem 2017), which quantifies the proportion of different topological (phylogenetic) relationships

among individuals, in 50 SNPs windows across the genome. In particular, we quantified the frequency of the "introgression" tree (*i.e.* clustering by phenotype) as opposed to "species" tree (*i.e.* clustering by expected species relationships where *H. timareta* is more closely related to *H. cydno*), across the QTL candidate region. The "introgression" topology, where *H. timareta* clusters with its sympatric *H. melpomene* co-mimic (and allopatric *H. melpomene*) is again predominantly supported at *optix* and at the level of previously identified behavioural candidate loci (**Figure 3**, topology abundances in **Supplementary figure 6** and raw weightings in **Supplementary figure 7**). This topology suggests introgression from *H. melpomene* into *H. timareta* (Martin et al. 2019), in line with the current theory of *H. timareta* acquiring the colour pattern alleles of the local *H. melpomene* populations as it expanded east of the Andes (Pardo-Diaz et al. 2012).

Increased admixture at the behavioural QTL, but not at *optix*, is associated with preference for red across populations.

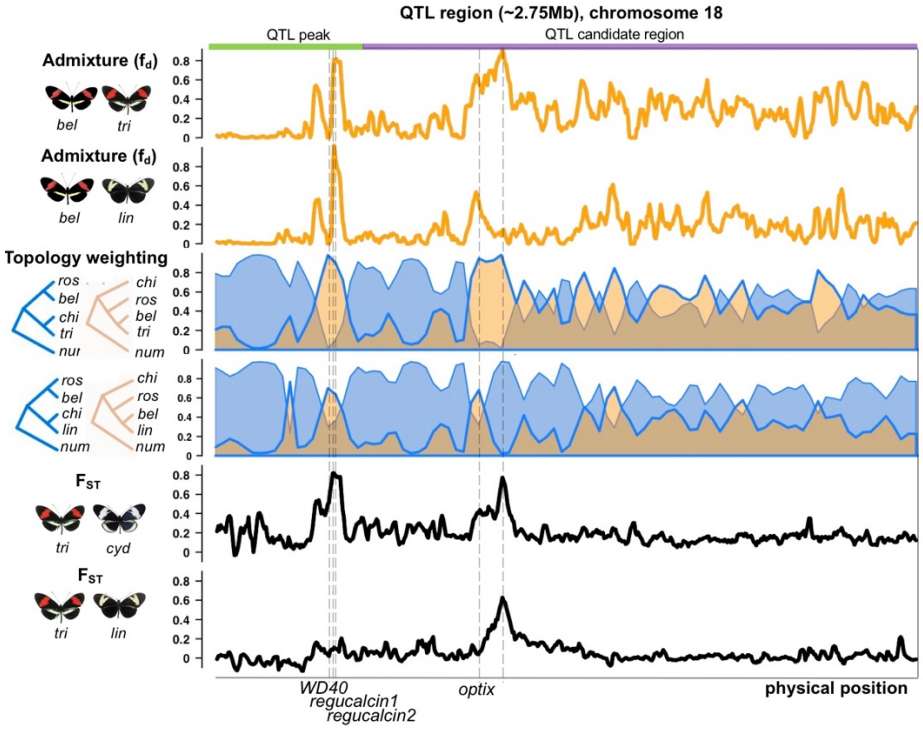
We next broadened our admixture analyses to include an additional population of *H. timareta:* the yellow *H. timareta linaresi,* which is devoid of red colour pattern elements. As expected, in comparisons between *H. t. linaresi* and the geographically adjacent (and likely sympatric) red *H. melpomene bellula,* we found no signal of admixture at (the putative regulatory regions upstream of) the colour patterning gene *optix*. Surprisingly, however, we did observe substantially increased levels of admixture at the behavioural QTL peak (top 1% quantile across chromosome 18), once again coincident with the candidate genes identified previously. We also found increased support for the topology that suggests a recent history of gene flow between *H. t. linaresi* and *H. melpomene* at the same genomic location **(Figure 3)**.

To test whether *H. t. linaresi* might prefer red over white females – despite the absence of a red forewing bar – we assayed mate preferences of yellow-banded *H. t. linaresi* males in the same choice trials as described above. We found that *H. t. linaresi* males did indeed prefer to court red *H. t. tristero* females over white *H. c. cydno* females (as compared to *H. cydno* males: 0.66 [0.55 – 0.76], z = 12.41, p <0.001, **Figure 1C**). We found qualitatively the same result with artificially red-coloured *H. c. cydno* females (*H. t. linaresi* - *H. c. cydno* = 0.37 [0.25 – 0.49], z = 6.20, p <0.001) (**Figure 1D**), suggesting that *H. t. linaresi* has similar red (over white) visual mating preference as *H. t. tristero*.

We also found elevated levels of F_{ST} and absolute divergence (d_{xy}) between red-preferring red populations and white-preferring populations at the colour pattern locus and at candidate behavioural genes (top 1% quantile across autosomes, **Supplementary figure 8)**,

but only at the colour pattern locus when comparing the red-preferring yellow *H. t. linaresi* and red *H. t. tristero* **(Figure 3)**. Together these results suggest that *optix* (only) controls the colour pattern switch, further implicating candidate behavioural genes.

Figure 3



In the top two panels, admixture proportion values, estimated in 20kb windows, between *H. melpomene bellula* and *H. t. tristero* (first panel) or *H. t. linaresi* (second panel), at the behavioural QTL region on chromosome 18 (x-axis indicates physical position). In the two middle panels, topology weightings (proportions of a particular phylogenetic tree over all possible rooted trees), inferred from 50 SNPs windows, for the "species" (blue) and "introgression" (orange) trees. Species/clades used: *H. melpomene rosina* (ros), *H. melpomene bellula (bel), H. c. cydno chioneus (chi), H. t. tristero (tri) or H. t. linaresi* (lin), *and H. numata* (num) as outgroup. A *loess* (smoothing) function has been applied on the topology weightings across 150kb windows. In the bottom two panels, fixation index (F_{ST}), a measure of genetic differentiation, estimated in 20kb windows, between *H. t. tristero and H. c. cydno* (top), *and between H. t. tristero and H. t. linaresi* (bottom). The gene coordinates of three candidates for behavioural differences (*WD40*, *regucalcin1* and *regucalcin2*), as well as *optix* and its putative regulatory region/s, are highlighted by vertical, grey dotted lines.

A selective sweep at the behavioural QTL in *H. timareta*.

An allele encoding "a preference for red butterflies" might have conferred a selective advantage to red *H. timareta* individuals through a reinforcement-like mechanism (selection/predation against misfit, i.e. non-mimetic hybrids, Merrill et al. 2012) with other non-red *H. timareta* individuals. Additionally, alleles that increase male attraction to receptive females – as we expect for *Heliconius* alleles that shift male preferences towards the locally most abundant conspecific pattern – will be under sexual selection (i.e. "scramble competition", Andersson 1994).

To test for signatures of selection associated with the evolution of male preferences in *H. timareta*, we scanned chromosome 18 using *Sweepfinder2* (DeGiorgio et al. 2016). We found that the genomic region with the highest support (top 1% quantile) for a selective sweep across chromosome 18 (using the site-frequency-spectra of chromosome 18 as neutral background) in *H. timareta thelxinoe* is coincident with the region showing a peak of admixture between red *H. melpomene* and *H. timareta* co-mimics, containing our candidate behavioural genes (**Figure 4**). We found no evidence of a selective sweep at the colour pattern locus.

Differences in *regucalcin1* expression are linked to visual preference behaviour.

To determine whether consistent differences in gene expression are associated with the behavioural QTL across red and white preferring populations, we generated RNA-seq libraries for combined eye and brain tissue from adult males (10 days of age) for all subspecies tested in our preference assays, in addition to first generation hybrids, to complement those for *H. melpomene rosina*, *H. cydno chioneus* and their hybrids reported previously (Rossi et al. 2020). We sampled at the adult stage reasoning that if the neural mechanism underlying divergent behaviours involves for example a change in neuronal activity, this might need sustained transcription/protein replenishment. We conducted differential gene expression analyses for these populations in a pairwise manner, comparing only butterfly specimens that were raised in common garden conditions (to avoid environmental effects, **Supplementary Figure 9**), with DESeq2 (Love et al. 2014).

Only a single gene within the QTL candidate region, *regucalcin1*, was consistently differentially expressed across all species comparisons. Specifically, *regucalcin1* shows significantly lower expression in the eye and brain tissue of *H. m. rosina, H. m. bellula, H. t. tristero* and *H. t. linaresi*, all which we have shown to have a red (over white) preference, compared to *H. c. cydno* and *H. c. chioneus*, which lack a preference for red patterns. Although two previously identified candidate genes *Grik2* and *regucalcin2* (Rossi et al. 2020) were

differentially expressed in *H. m. rosina* and *H. m. melpomene* compared to *H. c. chioneus* and *H. c. cydno*, they were not consistently differentially expressed across all species comparisons (defined as genes showing a >2-fold change in expression, at $p < 0.05$ adjusted for false discovery rate 5%, Wald test). We also found no consistent differences in exon usage between red and white preferring populations (with permissive mapping parameters, see Methods), suggesting that if the causal change is in exon usage (through splicing regulation), it is not detectable from whole-brain and eyes at the adult stage.

Behavioural experiments suggest that the allele for red preference (over white) observed in *H. melpomene* and *H. timareta* are largely dominant over alleles for a lack of red preference as seen in populations of *H. cydno*. Therefore, we previously hypothesised that any candidates underlying shifts in preference should be differentially expressed both in species and F1 hybrids populations with red preference relative to white preferring populations. In support of this hypothesis, *regucalcin1* was differentially expressed in the eye-brain tissue of *H. c. cydno* X *H. t. tristero* F1 hybrid males as compared to *H. c. cydno* males. However, we did not find this to be the case when comparing *H. m. rosina X H. c. chioneus* F1 hybrid males with *H. c. chioneus* males (and consequently we did not report expression differences as support for *regucalcin1* as a candidate in our previous analysis (see Rossi et al 2020)). Nevertheless, we did find significant 2-fold up-regulation of the *H. c. chioneus* allele relative to the *H. m. rosina* allele in *H. m. rosina X H. c. chioneus* F1 hybrids ($p < 0.001$, Wald test), and of the *H. c. cydno* allele relative to the *H. t. tristero* allele in *H. c. cydno* X *H. t. tristero* F1 hybrids ($p < 0.001$, Wald test). This indicates that there are cis-regulatory changes of *regucalcin1* linked to male preference behaviour.

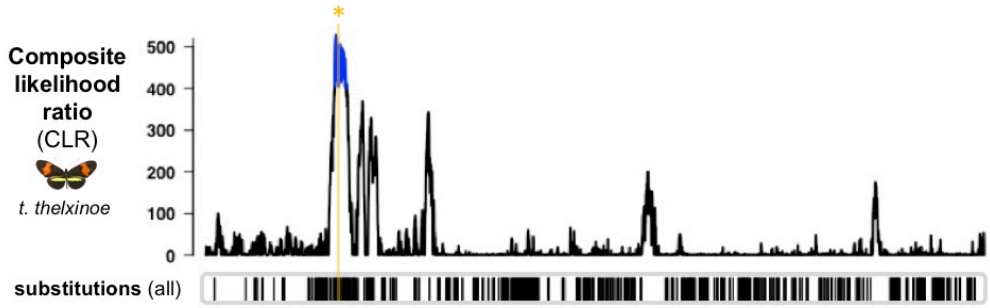
Mutations associated with male preference behaviour.

To begin to identify causative mutations underlying shifts in preference behaviour, we searched for "fixed" variants across the behavioural QTL candidate region shared by red-preferring but not white-preferring populations. We first compared species of the *H. melpomene-cydno* clade, which we have shown to have a red (over white) preference in this and previous (Merrill et al. 2019) studies (i.e. *H. m. rosina, H. t. tristero, H. m. bellula* and *H. t. linaresi* vs. *H. c. chioneus* and *H. c. cydno*) and identified 922 fixed variants differentiated these populations across the QTL region. We then extended our search to include subspecies with a visual preference for red butterflies (*H. m. melpomene* (Jiggins et al. 2001) as estimated with analogous mate choice assays, or that do not discriminate between red co-mimics based on visual cues (*H. m. amaryllis* and *H. t. thelxinoe* (Mérot et al. 2015), which we hypothesize share the same QTL-linked alleles

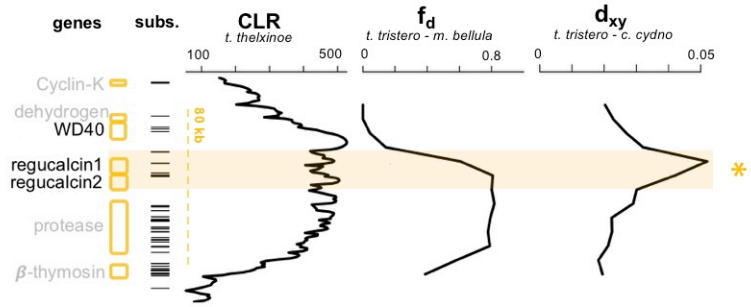
on chromosome 18 modulating preference behaviour. 734 fixed variants differentiating these red-preferring *H. melpomene*/*H. timareta* populations from white-preferring *H. cydno* were again spread across the QTL region **(Figure 4A),** including at the level of candidate genes (**Figure 4B**).

Figure 4

A



B



(A) Composite likelihood ratio (CLR) of a selective sweep in *H. t. thelxinoe*, estimated for 50bp windows, across the QTL region on chromosome 18. 1% upper quantile CLR (>400) values are highlighted in blue. Below, fixed mutations (substitutions) shared by all red-preferring subspecies relative to white-preferring subspecies, indicated by vertical black bars (note the low-resolution results in partially overlapping bars). **B)** Zooming into the region containing candidate behavioural genes (rotated by 90 degrees compared to A). From right to left are displayed: the coordinates of the annotated genes in the region (orange boxes), fixed mutations between red- and white preferring butterflies, the CLR of a selective sweep in *H. t. thelxinoe*, admixture proportions (f_d) between *H. t. tristero* and *H. m. bellula***,** and d_{xy} (a measure of absolute nucleotide divergence) between *H. t. tristero* and *H. c. cydno.* Highlighted in orange (*) the region that is roughly centred across the CLR, f_d and d_{xy} peaks, containing candidate *regucalcin* genes.

Overall, peak estimates of sequence divergence and genetic differentiation between red- and white-preferring populations, admixture proportion and phylogenetic clustering between red-preferring populations, and selection in "recipient" red-preferring *H. timareta* populations, all converge on a genomic region of less than 80kb, that is roughly centred on *regucalcin1* and *regucalcin 2* (**Figure 4b**).

Discussion

Changes in the way animals perceive and respond to the environment are key to adaptation and to the formation of new species, but their genetic basis is largely unknown. Behavioural adaptations might be exchanged between species through hybridization, as has been suggested for morphological traits (Hedrick 2013, Mallet et al. 2016), but direct links between specific introgressed loci and behavioural phenotypes are missing. Here, we provide strong evidence that a genetic mechanism underlying adaptive visual mate preferences was exchanged through hybridization between species of *Heliconius* butterfly. In particular, we have shown that: *H. melpomene* and *H. timareta* have "convergent" (collateral, sensu Stern et al. 2013) visual preferences for mimetic red butterflies, that these behaviours are modulated by the same region on chromosome 18 (linked to the colour patterning gene *optix*), and that candidate behavioural genes in this region show signatures of adaptive introgression. The preference locus we identify has a considerable behavioural effect, and its acquisition in *H. timareta* probably contributed to determine a shift in mate preference to facilitate adaptation.

We found that different subspecies of *H. timareta*, with yellow and red pattern elements, share the same preference for red butterflies over white ones. Admixture at candidate preference loci between red *H. melpomene* and yellow *H. timareta* suggests an unexpected adaptive acquisition of a "red-preference" by a yellow subspecies. Although we have not yet directly tested whether *H. t. linaresi* alleles segregate with preference behaviour at chromosome 18, this might indicate that the preference mechanism acquired by *H. timareta* subspecies from *H. melpomene* on chromosome 18 conferred a more general attraction towards butterflies with red-shifted colour pattern elements, including yellow. In line with this, divergent preferences towards artificial female wing models are more pronounced in red vs. white subspecies compared to red vs. yellow ones in the *H. melpomene* - *H. cydno* clade (Mérot et al. 2017). We also note that other sympatric forms of *H. melpomene* and *H. timareta* (*H. m.* malleti *and H. t. florencia*), with differently shaped (rayed) red and yellow pattern elements, show high admixture at both *optix* and at the putative preference locus (admixture estimates in

Martin et al. 2019), suggesting that the preference locus might encode a colour-guided preference independent of pattern shape. Nevertheless, further studies are needed to establish what exactly constitutes the cue that guides the preference mechanisms coded on chromosome 18, for example by studying whether *H. melpomene*/*H. timareta* alleles might be associated with a preference for yellow over white *H. cydno* females.

The reduced behavioural effect of the locus on chromosome 18 in experiments with painted females could suggest that our wing manipulation did not fully mimic visual cues used by *Heliconius* species in mate recognition (for example white hindwings colour patterns were not manipulated), and in particular the cue that guides the preference encoded on chromosome 18. Another possibility is that the QTL on chromosome 18 could encode a partially multimodal mate preference, where for example visual preference is enhanced depending on the presence of species-specific pheromonal cues.

Other preference mechanisms separate *H. timareta* subspecies, as *H. t. linaresi* males prefer *H. linaresi* over *H. t. tristero* females, and *H. t. tristero* prefers *H. t. tristero* over *H. linaresi* females (**Supplementary figure 10**). Further experiments are needed to disentangle the contribution of odour and visual cues to these differences, for example assaying male choice for yellow vs. red-painted females of the same subspecies. Nevertheless, the two subspecies are possibly able to distinguish colour patterns at finer scales, similarly to other *H. timareta* subspecies comparisons (Hausmann et al. 2021). Reproductive isolation between co-mimetic *H. melpomene and H. timareta* populations is likely maintained through isolating mechanisms other than colour-guided mating preferences, including pheromonal cues (Gonzáles-Rojas et al. 2020) and divergent ecological niches (Montgomery et al. 2021).

It remains unclear if the putative preference alleles have been acquired by different *H. timareta* subspecies through distinct introgression events with local *H. melpomene* populations, as it has been suggested for the evolution of mimicry (*optix*) alleles (Pardo-Diaz et al. 2012), or through a single introgression event. However, the independence of introgression signals at the candidate preference locus and cue (*optix*) (i.e. only preference in *H. t. linaresi*) suggests that preference and colour alleles could follow distinct evolutionary trajectories, and might have introgressed at different times. This is also hinted by the detection of a selective sweep at the putative preference locus but not at the colour locus. It is possible that the introgression of the red colour pattern alleles was immediately advantageous given the selective pressure for mimicry of most abundant local pattern, whereas preference for red/red-shifted butterflies was not, due to possible collateral deleterious effects of hybridization with *H. melpomene,* or

because preferring red butterflies would have been advantageous only in a (mostly) red population.

We found that *regucalcin1* is differentially expressed in adult brain and eye tissue across all eight red- and white-preferring subspecies comparisons, and shows allele-specific expression in (red-preferring) F1 hybrids, thereby being strongly linked to divergence in visual preference behaviours. Protein-coding substitutions identified here do not match previously identified mutations (between *H. m. rosina* and *H. c. chioneus*) at candidate genes predicted to alter protein function (Rossi et al. 2020), suggesting that the previously identified substitutions do not control the behavioural shift. Overall, it remains unclear which genetic change/s, protein-coding or regulatory, underlie the behavioural shift. However, signatures of genetic differentiation, selection and admixture, that should encompass the causative genetic element, are centered on two *regucalcin* genes, including the differentially expressed *regucalcin1*, that therefore is the strongest candidate for encoding a change in mate attraction. *Regucalcin* seems to be involved in brain calcium signaling (Yamaguchi 2012), but its function has not been well characterized. Nevertheless, calcium signaling modulates, among many biological processes, neuronal excitability (Berridge 1998), thereby being in line with controlling for behavioural phenotype.

The genetic basis of behavioural evolution remains largely unknown. As of 2021, of the >2000 entries for genotype-phenotype associations on *Gephebase* (Courtier-Orgogozo et al. 2019), only ~1% are linked to variation in behavioural traits. In particular, except a few cases of olfactory-guided mating preferences (Leary et al. 2012, Brand et al. 2020, Unbehend et al. 2021), specific genes linked to the evolution of visually-guided mating behaviours remain virtually unknown. Here we have shown that a < 80 kb genomic region, and in particular a *regucalcin*, is strongly associated with the evolution of visual attraction behaviours. Of course, a study of the functional effects of candidate genes would be desirable in the future, in order to elucidate the link between the action of individual genes and behaviour. In this perspective, we have used CRISPR-Cas9 technology to produce a handful of *regucalcin1* knock-out individuals (**Supplementary figure 11**), which sets the stage for studying the functional effects of candidate genes. In conclusion, we have provided strong evidence that a locus encoding visual mate preference was exchanged between hybridizing species lineages. This provides a link between the genetic code and variation in innate sensory preferences important for speciation, and expands the range of adaptive traits that could cross species boundaries to facilitate adaptation.

Methods

Butterfly collection and rearing. *H. melpomene rosina*, *H. m. amaryllis* and *H. cydno chioneus* stocks were reared concurrently, under the same conditions, at the Smithsonian Tropical Research Institute insectaries in Gamboa, Panama. Wild-caught individuals used to establish the stocks were originally from the Soberania and San Lorenzo National Parks, Panama (*H. m. rosina, H. c. chioneus*) and from the vicinity of Tarapoto and Cordillera Escalera, Peru (*H. m. amaryllis*). F1 hybrids were obtained by crossing a *H. c. chioneus* male to a *H. m. rosina* female.

H. m. bellula, *H. m. melpomene*, *H. c. cydno*, *H. timareta tristero* and *H. t. linaresi* stocks were reared concurrently, in common garden conditions, at the Universidad del Rosario insectaries in La Vega, Colombia. F1 hybrid broods were obtained by crossing a *H. t. tristero* male to a *H. c. cydno* female. Wild individuals used to establish insectary stocks were caught respectively in the area of La Vega (*H. c. cydno*), Mocoa (*H. m. bellula* and *H. t. tristero*), Buenavista (*H. m. melpomene*) and Guayabal (*H. t. linaresi*), Colombia. In both insectaries in Panama and Colombia, larvae were reared on *Passiflora* leaves until pupation and butterflies were provided with fresh *Psiguria* flowers and ~ 20% sugar solution daily.

Mate preference analyses. We assayed male attraction behaviours of *H. t. tristero*, *H. c. cydno,* their first generation (F1) hybrids and backcross hybrids to *H. c. cydno*, as well as of *H. m. bellula* and *H. t. linaresi* in standardized choice trials. We followed the same experimental design as in Merrill et al. 2019, where males were introduced into outdoor experimental cages (2x2x2m) with a virgin female of each species. 15-minutes trials were divided into 1-min. intervals, where courtship (sustained hovering or chasing) directed toward females was scored as having occurred or not, and the number of "courtship minutes" directed towards *H. cydno* or *H. timareta* females counted. If a male courted the same female twice during a minute interval, it was recorded only once; if courtship continued into a second minute, it was recorded twice. If a male courted both female species within a minute interval, both events were counted. Whenever possible, trials were repeated 5 times for each male (median number of trials with male responses to females among backcross hybrids = 2). From these choice trials we extracted a measure of male preference behaviour, "relative courtship time", as the "proportion of courting minutes directed toward *H. timareta* / courting minutes directed toward *H. timareta* + courting minutes directed toward toward *H. cydno*". We fitted generalized linear mixed models (GLMM) of family binomial (transformed with logit-link function), using the R package *lme4*

(Bates et al. 2007). For saturated models, we included male type (*i.e.* species or hybrid type) or male genotype at *regucalcin1* (only for backcross hybrid data) as fixed effect. All models included male ID as random factor. Pairwise z-tests between estimated marginal means, as implemented in the R package *emmeans*, were used to compare between male types or hybrid genotypes. p-values were corrected for multiple testing with the Bonferroni method whenever more than one comparison was performed.

Genotyping of backcross hybrids at the QTL peak/candidate genes. To confirm the segregation of *optix* alleles with red-colour pattern in hybrid crosses, and assay more specifically the genotype of hybrids at tightly linked candidate genes in the QTL peak, we performed PCR amplification of a *regucalcin1* segment. WGS alignment data indicated that fixed indels differentiate red-preferring vs. white-preferring subspecies in this region, so we designed primers to encompass putative *regucalcin1* indels (forward primer (5'-3') = GACATGCCAGGCTTCATAAT, reverse primer (5'-3') = TGAATTACCTGAGAGCCATC). gDNA was extracted from thorax tissue of grandparents (*H. c. cydno* and *H. t. tristero*), parents and backcross hybrid progeny, using a DNAeasy Blood & Tissue kit with RNAase A treatment (Qiagen, Valencia, CA, USA), that were previously stored in 20 % DMSO, 0.5 M EDTA (pH 8.0) solution. Then, we conducted PCR-amplification using a Taq polymerase (ThermoFisher) at melting temperature (T_m) 54 °C. We found that *H. c. cydno* and *H. t. tristero* individuals consistently differed in size of PCR-amplified fragment, allowing us to track alleles/genotype hybrid progeny.

Colour pattern-guided mate preference analyses and vision models. In a second behavioural assay, we recorded male attraction behaviours, this time towards two artificially coloured virgin *H. c. cydno* females. One female had the dorsal side of the (white) forewing band painted with a red marker pen (R05, Copic Ciao, Tokyo, Japan), the other with a control transparent pen (Ciao 0, Copic Ciao). The red marker pen had been previously chosen among others (R14, R17, R27, R29, R35, R46 and RV29, Copic Ciao) to best mimic the forewing colour pattern of *H. t. tristero*. For this, we took photographs of both red painted wings of *H. c. cydno* and of *H. t. tristero*, with a Nikon Nikkor D7000 camera (Nikon, Melville NY, USA) with a visible light (380-750nm range allowed) and a UV (100 – 380 nm) filter, in RAW format. A 40% grey standard was included in each photograph for colour calibration. The visible light and UV images of each wing were combined to generate a multispectral image, using the "Image calibration and analysis toolbox" (Troscianko & Stevens 2015), in ImageJ (Schneider

et al. 2012). The reflectance spectra of the forewing bands were extracted from the images and converted to cone catch models (Troscianko & Stevens 2015), based on cone sensitivities of *H. erato* (McCulloch et al. 2016) and relative abundance of cone receptors for species in the *H. melpomene/cydno* clade (McCulloch et al. 2017). Note that *Heliconius* can discriminate in the red-range even though they have only one long-wavelength (LW) opsin with peak sensitivity at 560nm, due to the presence of red-filtering pigments in some ommatidia (Zaccardi et al. 2006), that shifts the peak absorbance of some cones to ~600nm (McCulloch et al. 2016). However, this was not modelled in a first instance (because the relative abundance of this cone receptor remains unknown).

We calculated pairwise "just noticeable differences' (JND), using a tetrachromatic (*H. erato*) colour vision model (with noise-limited opponent colour channels, after Voroybev & Osorio 1998), between the forewing band of *H. t. tristero* and the red-coloured *H. c. cydno* band, using a Weber fraction of 0.05 (noise-to-signal ratio). The marker R05 had the lowest pairwise JND (0.89) and was therefore the marker we used to manipulate the forewing colours in experimental *H. cydno* females. A JND value less than 1 is considered to be undistinguishable by visual systems (Siddiqi et al. 2004).

Statistical analyses of behavioural data were conducted as (above) for the male choice experiments between female species (substituting *H. t. tristero* females with artificially red-coloured *H. c. cydno* females and *H. c. cydno* females with transparently-coloured *H. c. cydno* females). Genotype at the *optix* locus was inferred by the presence/absence of the wing red band (Reed et al. 2011), where hybrid individuals with a red band are heterozygotes for *H. timareta/H. cydno* alleles, and individuals lacking it are homozygous for the *H. cydno* allele.

Reflectance spectra and colour space analyses. We acquired reflectance spectra of the artificial and natural red and white (artificial = clear) pattern elements using an Ocean Optics FLAME-T-XR1-ES spectrometer, a UV/Vis bifurcated fiber, and a PX-2 Pulsed Xenon Lamp. A spectralon white standard (Ocean Optics WS-1) was used to calibrate the spectrometer. Each colour pattern (i.e. the forewing bar) was measured at three different locations, and the mean of the three measurements was used for further analyses. The reflectance data was analysed through a tetrachromatic colour vision model, incorporating recently published *H. melpomene* photoreceptor cell sensitivities (McCulloch et al. 2021). This differs from the model above in that we added the fourth (red-shifted; λ_{max} = 590) chromatic channel, linked to the presence of red filtering pigments (McCulloch et al. 2021). Cone abundances are not available for this newly classified photoreceptor type, so we were unable to calculate *JND* values. However, as

seen in Supplementary figure 1A and 1B, the artificial and natural red patterns have similar reflectance spectra and group closely when plotted in the *H. melpomene* tetrahedral colour space. Also, the reflectance spectra of the white forewing bar and those painted with the clear marker overlap in shape (Suppl. fig. 1A) and colour space (Suppl. fig. 1B).

Sampling, gDNA extraction and genotyping. gDNA was extracted from thorax tissue of 4 *H. m. bellula*, 9 *H. t tristero* and 4 *H. t linaresi* individuals, as well as parents of F1 hybrid crosses (see later), that were previously stored in 20 % DMSO, 0.5 M EDTA (pH 8.0) solution, using a DNAeasy Blood & Tissue kit, with RNAase treatment (Qiagen). Illumina, whole-genome resequencing libraries were prepared and sequenced at Novogene (Hong Kong, China) at either 125bp or 150bp paired-end. Previously compiled whole-genome resequencing data or variant calling (vcf) files were retrieved from (Martin et al. 2013, Nadeau et al. 2016, Jay et al. 2018, Martin et al. 2019), for a total of 5 *Heliconius numata*, 4 *H. m. bellula*, 10 *H. c. chioneus*, 10 *H. c. zelinde*, 10 *H. m. rosina*, 10 *H. m. amaryllis* and 10 *H. t thelxinoe*. Whole-genome resequencing reads were mapped to the *H. melpomene* genome version 2 (Davey et al. 2016) with BWA mem v.0.7.15 (Li & Durbin 2010). Duplicate reads were marked with Picard (https://broadinstitute.github.io/picard/), and variant (and invariant) calling was performed with GATK v3.7 HaplotypeCaller (McKenna et al. 2010), with default parameters except heterozygosity set to 0.02 (parameters as in Martin et al. 2019, for comparable analyses). Individual genomic records were combined and jointly genotyped (GATK's GenotypeGVCFs) separately for each subspecies.

Admixture proportions, F_{ST} and d_{xy} calculation. We calculated f_d (Martin el a. 2015), an admixture proportion estimate based on the ABBA-BABA test, between *H. melpomene* and *H. timareta* populations, as in Martin et al. 2019, implementing python scripts available at https://github.com/simonhmartin/. For this, variant sites had to be biallelic SNPs (no indels), with Quality (Q) >30 and read depth (DP) >8. Furthermore, variant sites were filtered out if > 30% of individuals had missing genotype calls and if > 75% of individuals had heterozygous calls (again for comparable analyses with Martin et al. 2019). The following populations were used to estimate admixture proportions: *H. cydno chioneus* and *H. cydno zelinde* as a (combined) allopatric control population, H. *timareta tristero* (or *H. t. linaresi*) and *H. melpomene bellula* (or *H. timareta thelxinoe* and *H. m. amaryllis*) as the two sympatric species, and *H. numata bicoloratus* as the outgroup. f_d was calculated in 20kb sliding windows (step =

5kb). For f_d estimates, only sites where >60% of individuals had a genotype were considered, and f_d values had to be based on >100 ABBA-BABA informative sites per window.
We calculated sequence divergence (d_{xy}) (Nei & Jin 1989) and the fixation index (F_{ST}) (Hudson et al. 1992), in sliding 20kb windows (step = 5kb, 2000 genotyped sites required per window) with the script "popgenWindows.py' (again available at https://github.com/simonhmartin/).

Topology weighting. To quantify phylogenetic relationships between species in genomic intervals along the QTL region associated with visual preferences, we used *Twisst* (Martin & Van Bellenghem 2017). We used the same invariant and variant sites filtered as above (for f_d estimation), with the further requirement that no more than 10 individuals were allowed to have missing genotypes at each site. Genotypes were phased and imputed using Beagle (Browning & Browning 2007). Neighbor-joining trees (Gascuel 1997) were inferred using PhyML (Guidon et al. 2013) (substitution model = GTR), as implemented in *Twisst*. Weightings for 15 possible topologies (rooted with *H. numata* as the outgroup) were estimated for non-overlapping 50 SNPs windows, using the following subspecies: *H. c. chioneus*, *H. m. rosina*, *H. t. tristero* (or *H. t. thelxinoe*) and *H. m. bellula* (or *H. m. amaryllis*).

Selective sweeps. Variant sites were filtered for genotype quality (GQ) > 30 and read depth (DP)>10, and were required to be biallelic SNPs (no indels). For downstream analyses, variant sites had to be called in 8 individuals out of 10 for the focal population, and in 3 individuals out of 5 for the outgroup. Sites were polarized (ancestral vs. derived) using *H. numata* as an outgroup. The site-frequency-spectrum (SFS) was computed for chromosome 18. We used SweepFinder2 (DeGiorgio et al. 2016), to estimate the composite likelihood ratio (CLR) of a sweep model compared to a neutral model (neutrality is represented by the background SFS of chromosome 18) in 50bp steps using polymorphic sites and substitutions (Huber et al. 2015). We considered those regions with top 1% quantile CLR values as having undergone a selective sweep. SweepFinder2 has been previously used to detect introgressed sweeps at colour pattern loci in *Heliconius* (Moest et al. 2020).

Fixed variants. To find fixed variants between red- and white-preferring *Heliconius* populations, we retrieved biallelic variant sites across these populations (called from whole-genome resequencing data), with genotype quality (GQ) > 30 and read depth DP > 10. To be considered fixed, variants were required to have allele frequency (AF) > 0.8 (present in 80% of individuals with a valid genotype call), and to be in homozygous state in 6/9 or 7/10 samples

of the same subspecies. We used *bcftools intersect* (Li et al. 2009) to extract those fixed variants sites that differed between red-preferring and white-preferring populations

Brain tissue collection, RNA extraction and sequencing. Brain (optic lobes and central brain) and eye (ommatidia) tissue were dissected out of the head capsule (as a single combined tissue), of sexually naïve, 10-days old males, in cold (4 ºC) 0.01M PBS. We sampled a total of: 5 *H. m. rosina*, 5 *H. m. amaryllis*, 5 *H. m. bellula,* 5 *H. m. melpomene,* 5 *H. t. tristero,* 5 *H. t. linaresi,* 5 *H. c. chioneus,* 4 *H. c. cydno, 6 F1 hybrids H. m. rosina-H. c. chioneus, 4 F1 hybrids H. c. cydno-H. t. tristero*. Tissues were stored in RNAlater (Thermo Fisher, Waltham, MA, USA) at 4 ºC for 24 hours, and subsequently at -20 ºC (Colombian samples) or -80 ºC (Panamanian samples) until RNA extraction. RNA was extracted and purified using TRIzol Reagent (Thermo Fisher) and a PureLink RNA Mini Kit with PureLink DNase digestion on column (Thermo Fisher). Illumina 150bp paired-end RNA-seq libraries were prepared and sequenced (in a single batch) at Novogene.

Differential gene expression and exon usage. After trimming adaptor and low-quality bases from raw reads using TrimGalore v.0.4.4 (www.bioinformatics.babraham.ac.uk/projects), Illumina RNA-seq reads were mapped to the *H. melpomene* 2 genome (Davey et al. 2016) / *H. melpomene* 2.5 annotation (Pinharanda et al. 2019) using STAR v.2.4.2a in 2-pass mode (Dobin et al. 2013), with default parameters (at first). Only reads that mapped in 'proper pairs' were kept for further analysis, using Samtools (Li et al. 2009). For gene expression analyses, the number of reads mapping to each annotated gene was estimated with HTseq v. 0.9.1 (model = union) (Anders et al. 2015). For exon usage analyses, the number of reads mapping to each annotated exon was estimated using the python script "dexseq_counts.py" from the DEXSeq package (Anders et al. 2012). Differential gene expression analyses were conducted with DESeq2 (Love et al. 2014), differential exon usage analyses with DEXSeq (Anders et al. 2012). Pairwise transcriptomic comparisons were conducted only between species raised in the same insectary locations (either Panama or Colombia), to avoid the confounding effect of environmentally-induced gene expression changes. We considered to be differentially expressed those genes and exons showing a 2-fold change in expression level, at adjusted (false discovery rate 5%) p-values < 0.05 (Wald test).

An initial finding that all red-preferring subspecies showed a significantly higher expression of the last exon (5) of *regucalcin1* (HMEL013551g4) compared to white preferring species, prompted us to study in more detail whether the highly divergent sequence of red-preferring

(including the *H. m. melpomene* reference genome) vs. white-preferring subspecies in this region might have affected this. In fact, when using more permissive parameters than the default parameters in STAR v.2.4.2a (Supplementary methods, parameters set 1), differential usage of exon 5 of *regucalcin1* disappeared in many comparisons. Given that i) with these permissive parameters there is uniform RNA-seq reads coverage of exon 5 in *H. cydno* subspecies ii) when using even more permissive parameters (parameters set 2 in Supplementary methods) the results remain unchanged, and that iii) when using PacBio RNA long-read data from *H. cydno* to assemble the *regucalcin1* transcript, exon 5 is included (see Supplementary Methods), we concluded that the more permissive parameters are likely more appropriate. Therefore, the initial finding of consistent differential exon 5 usage is likely an artifact of too stringent (default) mapping parameters.

Allele-specific expression analyses. The 8 parental individuals of the *6 F1 hybrids H. m. rosina-H. c. chioneus* and *4 F1 hybrids H. c. cydno-H. t. tristero* (two broods for each F1 hybrid type), were genotyped as above using GATK v3.7 HaplotypeCaller. Individual genomic records were filtered with "hard-filters" following the GATK's best practices. We extracted variant sites with opposite alleles between each parental brood pair with *bcftools intersect* for further analysis. We marked duplicate (F1 hybrid) RNA-seq reads with Picard v.1.8 (https://broadinstitute.github.io/picard/), applied the GATK's SplitNCigarReads function and genotyped RNA-reads with HaplotypeCaller. We filtered out variant sites that had quality by depth (QD) < 2 and strand bias (FS) >30. We kept only biallelic, heterozygous SNPs for further analysis (allele-informative sites should be heterozygous for the parental alleles). Finally, we used GATK's ASEReadCounter (without deduplicating RNA reads) to count how many RNA-reads mapped to either parental allele. We tested for differential allele specific expression for each gene with the model "~0 + individual + allele" in DESeq2 (setting sizeFactors = 1, i. e. without library size normalization between samples).

CRISPR/Cas9 mediated mutagenesis of r*egucalcin1*. *Heliconius melpomene rosina* individuals were obtained from Costa Rica Entomological Supply and used to establish a stock in an external greenhouse at LMU Munich. To sample genetic variation in the stock population at the level of potential CRISPR/Cas9 target sites, we extracted gDNA from 6 individuals as described above (DNAeasy Blood & Tissue kit), amplified fragments of *regucalcin1* including exon 1 and exon 2 with Taq polymerase, purified products with an ExoSAP kit (Thermo Fisher), and Sanger-sequenced with a BigDye v1.1 kit (Thermo Fisher) with the Genomics

Service Unit of LMU Munich. We used *GeneiousPrime* v2021.1 to design 4 guide RNAs corresponding to N_{20}NGG (on either strand), targeting exon1 and exon2 of *regucalcin1* (**Supplementary figure 11**), considering the gRNA efficiency scores predicted from Doench et al. 2016, favouring GC-rich regions close to the PAM (NGG) sequence, and avoiding polymorphic sites in our butterfly stock. N_{20}NGG sequences were screened for off-targets in the *H. melpomene* 2.5 genome with the BLAST function of Lepbase v4. Only guide RNAs that had unique seed regions 12bp upstream of the PAM were considered further (to avoid off-targets, following Livraghi et al. 2021, Hsu et al. 2013). Synthetic sgRNAs were ordered from *Synthego* (Redwood City, CA, US) and resuspended in TE (0.1mM EDTA, pH 8.0) buffer (Sigma Aldrich, St. Louis, MO, US). Cas9 protein (CP01, PNAbio) was reconstituted in nuclease-free water and 5% Phenol Red Solution (Sigma Aldrich), following the guidelines in Martin et al. 2020. A mix of 4 gRNAs (and later 2gRNAs) and Cas9 protein (250:500ng/µl) was injected in eggs between 2 and 4 hours after laying. To genotype individuals, we extracted gDNA from a pull of 2-4 caterpillar spikes, by squishing the spikes with a filter tip in 9 µl NaOH solution (50mM), incubating at 95°C for 15 minutes, cooling the reaction on ice for 2 min, and adding 1 µl of Tris-HCl (1M) (modified from Meeker et al. 2007, Nicolas Gompel and Luca Livarghi personal communication). We PCR-amplified a region of *regucalcin1* (forward primer (5'-3') = GACATGCCAGGCTTCATAAT, reverse primer (5'-3') = ATCGATATCCACCTCCATCA), to screen for CRISPR/Cas9 mediate deletions as a result of non-homologous end-joining following multiple double-strand breaks. To confirm CRISPR/Cas9 mediated deletions suggested by PCR-fragments sizes, we purified DNA from gel bands of the allele carrying the predicted deletion (Supplementary figure 11) with a MinElute Gel Extraction Kit (QIAGEN) and EXOSap. Sanger-sequencing indicated that the same 2 gRNAs ((5'-3') AAGCAGUCACUGAGCCGGUG, GUAGUGGUCGUACAGUGGGA) consistently mediated the introduction of double-strand breaks. CRISPR/Cas9 mediated deletion efficiency was ~9% (4/45) when injecting 4sgRNAs and ~18% (6/34) when injecting 2 gRNAs.

References

Abbott, R. *et al*. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229-246.

Anders S, Reyes A, Huber W (2012). "Detecting differential usage of exons from RNA-seq data." *Genome Research*, 22, 4025.

Anders, S., Pyl, P. T. & Huber, W (2015) HTSeq- a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169.

Andersson, M. (1994). Sexual selection. *Princeton University Press.*

Arguello, J. R., & Benton, R. (2017). Open questions: Tackling Darwin's "instincts": the genetic basis of behavioral evolution. *BMC biology*, *15*(1), 1-3.

Arias, C. F. *et al*. (2017). A new subspecies in a Heliconius butterfly adaptive radiation (Lepidoptera: Nymphalidae). *Zoological Journal of the Linnean Society*, *180*(4), 805-818.

Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R package version*, *2*(1), 74.

Berridge, M. J. (1998). Neuronal calcium signaling. *Neuron*, *21*(1), 13-26.

Brand, P. *et al*. (2020). The evolution of sexual signaling is linked to odorant receptor tuning in perfume-collecting orchid bees. *Nature communications*, *11*(1), 1-11.

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, *81*(5), 1084-1097.

Chowdhury, T., Calhoun, R. M., Bruch, K., & Moehring, A. J. (2020). The fruitless gene affects female receptivity and species isolation. *Proceedings of the Royal Society B*, *287*(1923), 20192765.

Courtier-Orgogozo, V., Arnoult, L., Prigent, S. R., Wiltgen, S., & Martin, A. (2020). Gephebase, a database of genotype–phenotype relationships for natural and domesticated variation in Eukaryotes. *Nucleic acids research*, *48*(D1), D696-D703.

Crane, J. (1955). Imaginal behavior of a Trinidad butterfly, Heliconius erato hydara Heiwitson, with special reference to the social use of color. *Zoologica (NY)*, *40*, 167-196.

Dasmahapatra, K. K. *et al*. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*(7405), 94.

Davey, J. *et al.* (2016). Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3* **6**, 695–708.

DeGiorgio, M. *et al*. (2016). SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics*, *32*(12), 1895-1897.

Ding, Y., Berrocal, A., Morita, T., Longden, K. D., & Stern, D. L. (2016). Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature*, *536*(7616), 329-332.

Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).

Doench, J. G., *et al*. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature biotechnology*, *34*(2), 184-191. (2016).

Edelman, N. B. *et al*. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science*, *366*(6465), 594-599.

Fandino, R. A., *et al*. (2019). Mutagenesis of odorant coreceptor Orco fully disrupts foraging but not oviposition behaviors in the hawkmoth Manduca sexta. *Proceedings of the National Academy of Sciences*, *116*(31), 15677-15685.

Felsenstein, J. (1981). Skepticism Towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35, 124–138.

Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular biology and evolution*, *14*(7), 685-695.

González-Rojas, M. F. *et al*. (2020). Chemical signals act as the main reproductive barrier between sister and mimetic Heliconius butterflies. *Proceedings of the Royal Society B*, *287*(1926), 20200587.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, *59*(3), 307-321.

Hausmann, A. E. *et al*. (2021). Light environment influences mating behaviours during the early stages of divergence in tropical butterflies. *in press, Proceedings of the Royal Society B: Biological Sciences*

Hedrick, P. W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular ecology*, *22*(18), 4606-4618.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., ... & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*, *31*(9), 827-832.

Hudson, R. R., Slatkin, M. & Maddison, W. P. Estimation of levels of gene flow from DNA sequence data. Genetics 132, 583–589 (1992).

Jay, P. *et al.* (2018). Supergene evolution triggered by the introgression of a chromosomal inversion. *Current Biology*, *28*(11), 1839-1845.

Jiggins, C. D. *et al.* (2001). Reproductive isolation caused by colour pattern mimicry. *Nature* 411, 302–305.

Jiggins, C. D., Estrada, C. & Rodrigues, A. (2004). Mimicry and the evolution of premating isolation in *Heliconius melpomene Linnaeus*. *J. Evol. Biol.* 17, 680–691.

Jiggins, C. D. (2017). The ecology and evolution of *Heliconius* butterflies. *Oxford University Press.*

Jones, M. R. *et al.* (2018). Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science*, *360*(6395), 1355-1358.

Kozak, K. M., Joron, M., McMillan, W. O., & Jiggins, C. D. (2021). Rampant genome-wide admixture across the *Heliconius* radiation. *Genome Biology and Evolution*, in press

Kronfrost, M. R *et al.* (2006). Linkage of butterfly mate preference and wing color preference cue at the genomic location of *wingless*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6575-6580.

Leary, G. P. *et al.* (2012). Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proceedings of the National Academy of Sciences*, *109*(35), 14081-14086.

Li, Q., Lee, J. A., & Black, D. L. (2007). Neuronal regulation of alternative pre-mRNA splicing. *Nature Reviews Neuroscience*, *8*(11), 819-831.

Livraghi, L. *et al.* (2021). Cortex cis-regulatory switches establish scale colour identity and pattern diversity in Heliconius. *Elife*, *10*, e68549.

Love, M. I. Huber, W. & Anders, S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21.

Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species?. *BioEssays*, *38*(2), 140-149.

Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in ecology & evolution*, *34*(6), 531-544.

Martin, A. *et al*. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences*, *109*(31), 12632-12637

Martin, A., & Orgogozo, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution*, *67*(5), 1235-1250.

Martin, S. H. *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828.

Martin, S. H., Davey, J. W. & Jiggins, C. D. (2015). Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32, 244–257.

Martin, S. H., & Van Belleghem, S. M. (2017). Exploring evolutionary relationships across the genome using topology weighting. *Genetics*, *206*(1), 429-438.

Martin, S. H. *et al*. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* 17, 1–28.

Martin, A., Wolcott, N. S., & O'Connell, L. A. (2020). Bringing immersive science to undergraduate laboratory courses using CRISPR gene knockouts in frogs and butterflies. *Journal of Experimental Biology*, *223*(Suppl_1), jeb208793.

McCulloch K. J. et *a*. (2016). Sexual dimorphism in the compound eye of Heliconius erato: a nymphalid butterfly with at least five spectral classes of photoreceptor. *Journal of Experimental Biology*, *219*(15), 2377-2387.

McCulloch K. J. et *a*. (2017). Sexual dimorphism and retinal mosaic diversification following the evolution of a violet receptor in butterflies. Molecular Biology and Evolution 34: 2271-2284.

McCulloch, K. J., Macias-Munoz, A., Mortazavi, A., & Briscoe, A. D. (2021). Multiple mechanisms of photoreceptor spectral tuning following loss of UV color vision in Heliconius butterflies. *bioRxiv*.

McKenna, A. *et al.* (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303.

Meier, J. I. *et al*. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature communications*, *8*(1), 1-11.

Meeker, N. D., Hutchinson, S. A., Ho, L., & Trede, N. S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Biotechniques*, *43*(5), 610-614.

Mérot, C. *et al*. (2013). Genetic differentiation without mimicry shift in a pair of hybridizing Heliconius species (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, *109*(4), 830-847.

Mérot, C., Frérot, B., Leppik, E., & Joron, M. (2015). Beyond magic traits: Multimodal mating cues in Heliconius butterflies. *Evolution*, *69*(11), 2891-2904.

Mérot, C. *et al*. (2017). What shapes the continuum of reproductive isolation? Lessons from Heliconius butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856), 20170335.

Merrill, R. M. *et al*. (2011). Pervasive genetic associations between traits causing reproductive isolation in *Heliconius* butterflies. *Proc. R. Soc. B Biol. Sci.* 278, 511–518.

Merrill, R. M. *et al.* (2012). Disruptive ecological selection on a mating cue. *Proc. R. Soc. B Biol. Sci.* 279, 4907–4913.

Merrill, R. M. *et al*. (2015). The diversification of Heliconius butterflies: what have we learned in 150 years?. *Journal of Evolutionary Biology*, *28*(8), 1417-1438.

Merrill, R. M. *et al.* (2019). Genetic dissection of assortative mating behaviour. *PLoS Biol 17*, e2005902.

Moest, M. *et al*. (2020). Selective sweeps on novel and introgressed variation shape mimicry loci in a butterfly adaptive radiation. *PLoS biology*, *18*(2), e3000597.

Montgomery, S. H. *et al*. (2021). Neural divergence and hybrid disruption between ecologically isolated Heliconius butterflies. *Proceedings of the National Academy of Sciences*, *118*(6).

Monteiro, A. (2012). Gene regulatory networks reused to build novel traits. *BioEssays*, *34*(3), 181-186.

Nadeau, N. J. *et al*. (2016). The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature*, *534*(7605), 106-110.

Nei, M. & Jin, L. (1989). Variances of the average numbers of nucleotide substitutions within and between populations. Mol. Biol. Evol. 6, 290–300.

Li, H. *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, H. & Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Livraghi, L. *et al*. (2021). *Cortex cis*-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*. Elife, *in press*

Pardo-Diaz, C. *et al*. (2012). Adaptive introgression across species boundaries in Heliconius butterflies. *PLoS Genet*, *8*(6), e1002752.

Pinharanda, A. *et al.* (2019) Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*. *J. Evol. Biol.* **32**, 194–204.

Prud'Homme, B *et al.* (2006). Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature*, *440*(7087), 1050-1053.

Reed, R. D. *et al.* (2011). Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, *333*(6046), 1137-1141.

Rosenberg, S. S., & Spitzer, N. C. (2011). Calcium signaling in neuronal development. *Cold Spring Harbor perspectives in biology*, *3*(10), a004259.

Rosser, N., Phillimore, A. B., Huertas, B., Willmott, K. R., & Mallet, J. (2012). Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America. *Biological journal of the linnean society*, *105*(3), 479-497.

Rossi, M. *et al.* (2020). Visual mate preference evolution during butterfly speciation is linked to neural processing genes *Nature communications.* 11, 2743.

Semenov, G. A. *et al.* (2021). Asymmetric introgression reveals the genetic architecture of a plumage trait. *Nature Communications*, *12*(1), 1-9.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, *9*(7), 671-675.

Siddiqi, A. et *al*. (2004). Interspecific and intraspecific views of color signals in the strawberry poison frog Dendrobates pumilio. *Journal of Experimental Biology*, *207*(14), 2471-2485.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, *14*(11), 751-764.

Stern, D. L., & Orgogozo, V. (2008). The loci of evolution: how predictable is genetic evolution?. *Evolution*, *62*(9), 2155-2177.

Troscianko, J., & Stevens, M. (2015). Image calibration and analysis toolbox–a free software suite for objectively measuring reflectance, colour and pattern. *Methods in Ecology and Evolution*, *6*(11), 1320-1331.

Unbehend, M. *et al.* (2021). bric à brac controls sex pheromone choice by male European corn borer moths. *Nature communications*, *12*(1), 1-11.

Vorobyev, M., & Osorio, D. (1998). Receptor noise as a determinant of colour thresholds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1394), 351-358.

Wallbank, R. W. *et al.* (2016). Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. *PLoS biology*, *14*(1), e1002353.

Westerman, E. L. *et al.* (2018). Aristaless controls butterfly wing color variation used in mimicry and mate choice. *Current Biology*, *28*(21), 3469-3474.

Yamaguchi, M. (2012). Role of regucalcin in brain calcium signaling: involvement in aging. *Integrative Biology*, *4*(8), 825-837.

Zaccardi, G., Kelber, A., Sison-Mangus, M. P. & Briscoe, A. D. (2006). Colour discrimination in the red range with only one long-wavelength sensitive opsin. *J. Exp. Biol.* **209**, 1944–1955.

Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S., & Yang, S. H. (2015). Off-target effects in CRISPR/Cas9-mediated genome engineering. *Molecular Therapy-Nucleic Acids*, *4*, e264.

Acknowledgments.

We thank Daniela and Geraldine for assistance and acquiring behavioral data at the insectaries of the Universidad del Rosario. We thank Morgan Oberweiser for assistance at the Smithsonian Tropical Institute insectaries, and Mathieu Choteau for help with fieldwork. We thank Christine Rottenberger for technical assistance. We thank Luca Livraghi and Joseph H. Hanly for advice on molecular techniques. We thank Francesco Cicconardi for sharing a pipeline for Iso-Seq analysis and Simon H. Martin for sharing vcf files. We are grateful to Jochen B. Wolf and Chris D. Jiggins who kindly provided computational resources. We thank Nicolas Gompel for providing access to a Micropipette Puller, as well for his input on the project together with Ricardo Pereira. We thank John Parsch for access to a FemtoJet. We thank the Graduate School Life Science Munich (LSM) for a LSM Travel Fund to Matteo Rossi. This work was supported by an Emmy Noether grant awarded to Richard M. Merrill by the Deutsche Forschungsgemeinschaft (DFG) (Grant number: GZ:ME 4845/1-1).

Supplementary methods

RNA-seq mapping parameters. The default mapping parameters in STAR v.2.4.2a (Dobin et al. 2013) were changed to more permissive ones (parameters set 1): --outFilterMismatchNmax 15 --outFilterMismatchNoverReadLmax 0.1 --outFilterMismatchNoverLmax 0.1 --outFilterScoreMinOverLread 0.5 --outFilterMatchNminOverLread 0.5.
We also conducted the same analyses with yet more permissive parameters (parameters set 2): --outFilterMismatchNmax 20 --outFilterMismatchNoverReadLmax 2 --outFilterMismatchNoverLmax 0.2 --outFilterScoreMinOverLread 0.33 --outFilterMatchNminOverLread 0.33

PacBio isoform sequencing. Brain (optic lobes and central brain) and eye (ommatidia) tissue were dissected out of the head capsule (as a single combined tissue), of sexually naïve, 10-days old males, in cold (4 ºC) 0.01M PBS. Tissues were stored in RNAlater (Thermo Fisher,

Waltham, MA, USA) at 4 ºC for 24 hours, and subsequently at -20 ºC or -80 ºC until RNA extraction. RNA was extracted and purified using TRIzol Reagent (Thermo Fisher) and a PureLink RNA Mini Kit with PureLink DNase digestion on column (Thermo Fisher), from a pull of whole-brain and eye tissue of the same subspecies (4 *H. m. rosina*, 4 *H. t. tristero* and 2. *H. c. chioneus* male individuals) for a total of 3 libraries, one for each subspecies. Single molecule real-time (SMRTbell) libraries were prepared and sequenced at Novogene (Hong Kong, China), on a PacBio RSII platform (Pacific Biosciences, Menlo Park, CA, USA).

Isoform assembly/discovery and transcript-guided annotation. Following the custom IsoSeq v3 pipeline (https://github.com/PacificBiosciences/IsoSeq/), Iso-Seq subreads from each library were used to generate circular consensus sequences (ccs). Primers (5' AAGCAGTGGTATCAACGCAGAGTACATGGG, 3' GTACTCTGCGTTGATACCACTGCTT), polyA tails and artificial concatemers were removed. Bam files were transformed into fastq format using Samtools (Li et al. 2009). Reads were mapped to the *H. melpomene* 2 (Davey et al. 2016) genome using *minimap2* (Li 2018) with default parameters for PacBio Iso-seq (-ax splice:hq). Stringtie2 (Kovaka et al. 2019) was used to assemble de-novo transcripts, in order to conduct between-species comparison of isoform expression. However, gene coverage of Iso-Seq reads was low, and the resulting de-novo annotation incomplete, not permitting reliable inference of differential isoform expression between species.

Supplementary references

Davey, J. *et al.* (2016). Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3* **6**, 695–708.

Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2, *Genome Biology* 20, 278 (2019), doi:10.1186/s13059-019-1910-1.

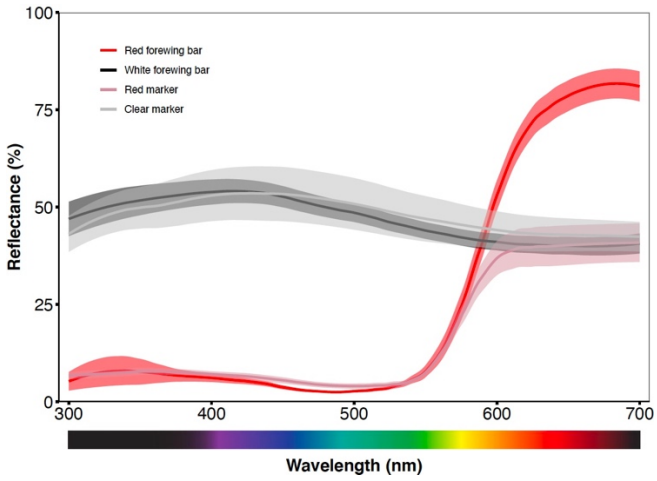
Li, H. *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34:3094-3100.

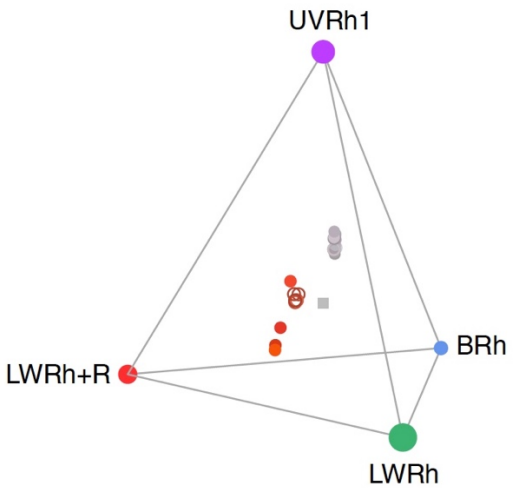
SUPPLEMENTARY FIGURES

Supplementary figure 1. (A) Reflectance spectra of the natural red (averaged across 4 *Heliconius timareta tristero* female samples), red- painted (4 *H. cydno cydno* female samples) forewing bars, white (9 *H. c. cydno* female samples) and transparently painted (4 clear-painted *H. c. cydno* female samples) forewing bars **(B)** Colour space (i. e. stimulation of cone types) for the different forewing reflectances, using a tetrachromatic model with *H. melpomene* photoreceptor cell sensitivities (McCulloch et al. 2021). Maximum sensitivity: UV-Rhodopsin1 (360 nm), blue-Rhodopsin (470nm), long wavelength-Rhodopsin without (570nm) and with red filtering pigments (+R) (590nm). Solid circles indicate unmanipulated forewings (n=5), open circles indicate painted forewings (n=5).

A



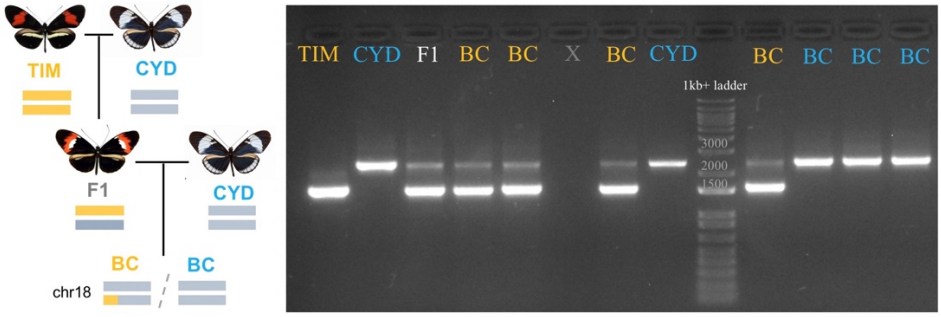
B



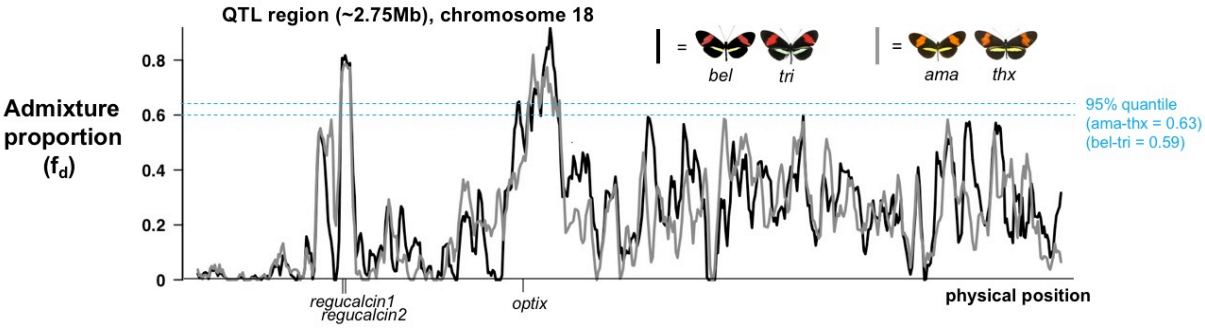
Supplementary Figure 2. Figure and text from Merrill et al. 2019. (A) Proportion of courtships directed toward *H. melpomene* (as opposed to *H. cydno*) females for CYD, MEL, their F1, and BC and BM. Values in parentheses indicate total number of individuals with behavioural data. Solid colored boxes represent expected average genome contribution of each generation. BC, backcross hybrid to *H. cydno*; BM, backcross hybrid to *H. melpomene*; CYD, *H. cydno*; MEL, *H. melpomene*; F1, first-generation hybrid. (*H. melpomene* = *H. m. rosina*, *H. cydno* = *H. c. chioneus*) (B) Proportion of time males court *H. melpomene* (as opposed to *H. cydno*) females for each of the two genotypes for the QTL on chromosome 18 (homozygous = CYD:CYD and heterozygous = CYD:MEL). Error bars represent 95% confidence intervals. Lower dashed blue and upper orange bars represent mean phenotypes measured in *H. cydno* and *H. melpomene*, respectively. Circle size depicts total number of "courtship minutes" for each male. Vertical black bars indicate the percentage of the difference measured in the parental species explained.

Figure 1D and 2B at: https://doi.org/10.1371/journal.pbio.2005902

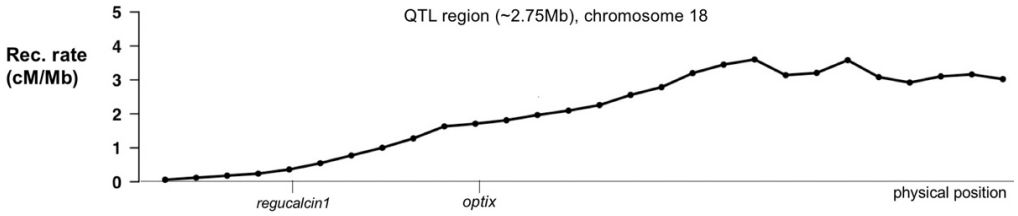
Supplementary figure 3. On the left: crossing design for producing backcross hybrid individuals segregating at *optix*/the behavioural QTL region on chromosome 18. On the right: PCR-amplification of an exonic-intronic segment of *regucalcin1* (in the QTL region) for grandparents, parents and a few backcross hybrid individuals, showing the expected segregation of parental alleles (differing by size/indels). TIM= *H. t. tristero,* CYD = *H. c. cydno,* F1 = F1 hybrid *H. c. cydno* x *H. t. tristero,* BC = backcross to *H. c. cydno hybrid*. For all parental species individuals *H. t. tristero* PCR-amplified fragments were ~1500bp, *H. c. cydno* were ~2400bp.



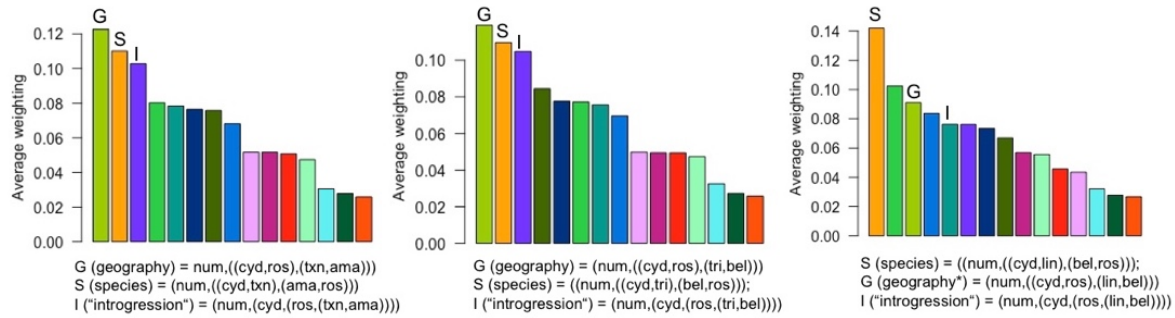
Supplementary figure 4. Admixture proportion (f_d) values (estimated in 20kb sliding windows) between *H. melpomene bellula* and *H. timareta tristero* (black line) and between *H. melpomene amaryllis* and *H. timareta thelxinoe* (grey line), at the behavioural QTL region on chromosome 18 (x-axis indicates physical position). f_d 95% quantiles for chromosome 18 for the two populations are represented by horizontal blue dotted lines. The two sympatric populations show consistent high admixture proportions at candidate behavioural genes and in the regions upstream of *optix* (colour pattern locus), indicated by vertical bars.



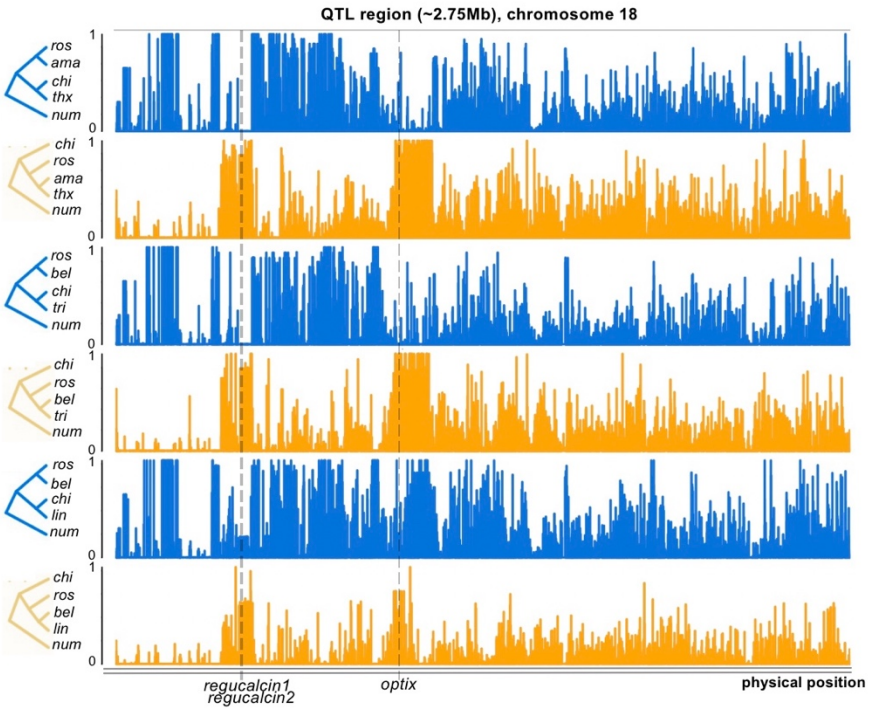
Supplementary figure 5. Recombination rate (cM/Mb) for 100kb windows across the QTL region, as estimated (from hybrid crosses pedigree) in Davey et al. 2017. The *x*-axis represents physical position. Two grey bars indicate the gene coordinates of *regucalcin1* and *optix*, displayed for reference.



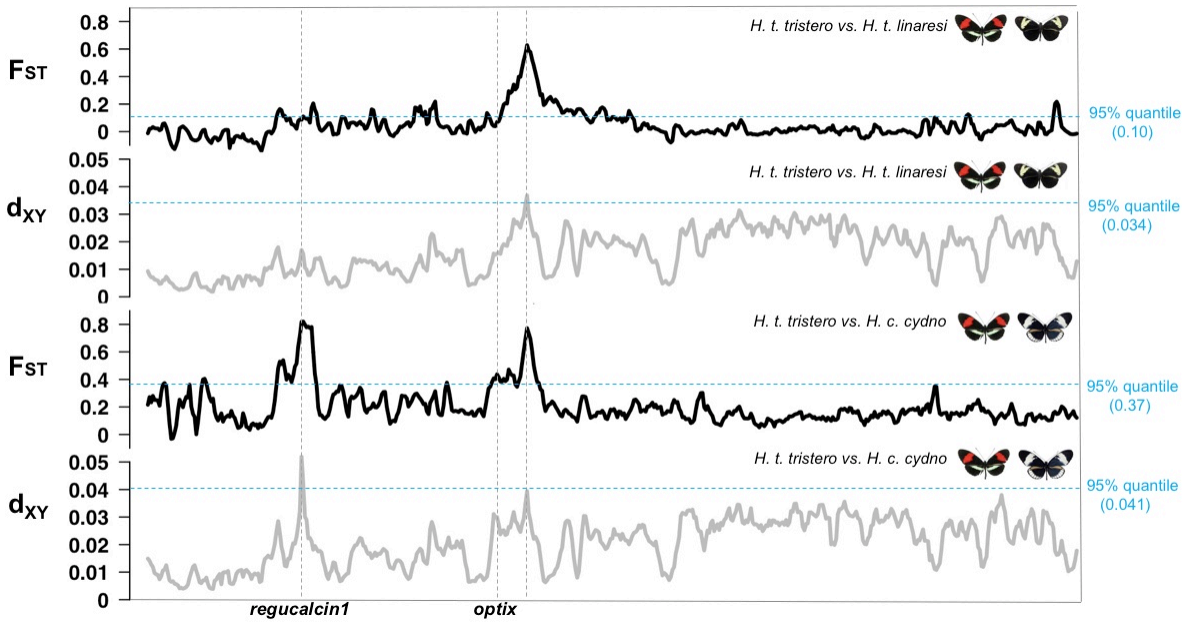
Supplementary figure 6. Average topology weightings for all possible 15 rooted topologies (colour coded) across chromosome 18 (50 SNPs windows) for 3 combinations of 5 different *Heliconius* subspecies. These combinations include 3 different *H. melpomene* and *H. timareta* subspecies pairs that show increased admixture at the putative preference locus. Three topological conformations are highlighted: G = geography (topology where sympatric/geographically close subspecies pairs cluster together), S= species topology (following expected phylogenetic relationships), I = introgression topology (*H. timareta* clusters with its sympatric/geographically close *H. melpomene* subspecies, nested within the *H. melpomene* clade). Subspecies sampled: *H. melpomene rosina* (ros), *H. m. amaryllis* (ama), *H. melpomene bellula (bel), H. c. cydno chioneus (chi), H. t. thelxinoe* (thx), *H. t. tristero (tri), H. t. linaresi* (lin), *H. numata* (num) – outgroup.



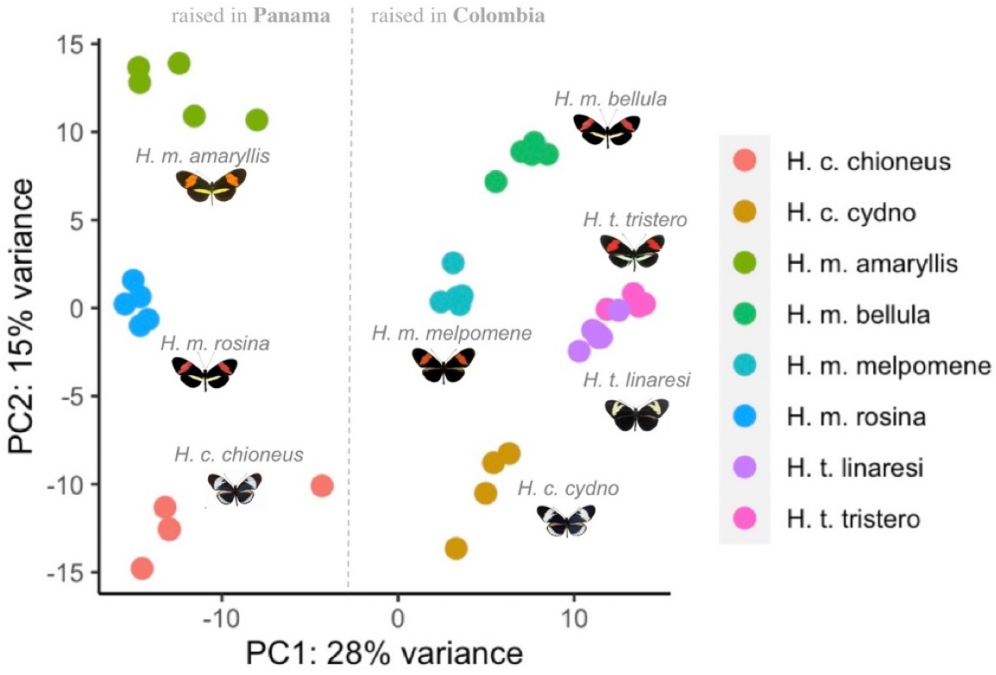
Supplementary figure 7. Topology weightings (proportion of a particular phylogenetic tree over all possible 15 rooted trees), ranging from 0 to 1 (*y*-axis), inferred from 50 SNPs windows along the QTL region (*x*-axis represent physical position). The "species" tree (expected species relationships: *H. timareta* more closely related to *H. cydno* than *H. melpomene*) is represented in blue, the "introgression" tree (where *H. timareta* clusters with its sympatric *H. melpomene* co-mimic) in orange. Species/clades used: *H. melpomene rosina* (ros), *H. m. amaryllis* (ama), *H. melpomene bellula (bel), H. c. cydno chioneus (chi), H. t. thelxinoe* (thx), *H. t. tristero (tri), H. t. linaresi* (lin), *H. numata* (num) as outgroup. Gene coordinates of two candidate genes for behavioural differences (*regucalcin1* and *regucalcin2*) and *optix* protein-coding region are highlighted by vertical, grey dotted lines.



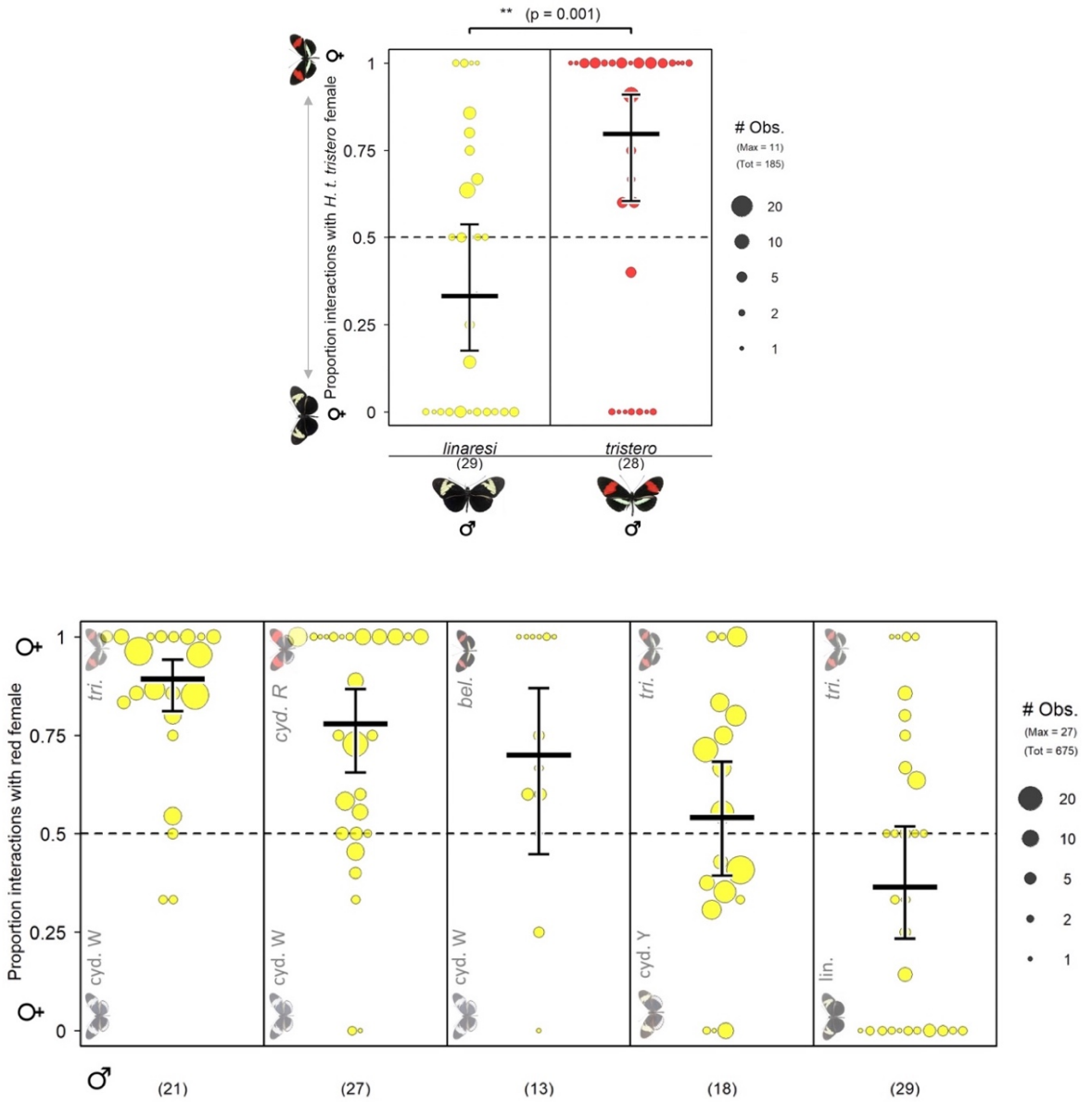
Supplementary figure 8. Fixation index (F_{ST}), estimated in 20kb windows (2000 genotyped sites were required per window), *between H. t. tristero and H. t. linaresi* (top panel), and between *H. t. tristero and H. c. cydno* (third panel). d_{xy}, estimated in 20kb windows, between *H. t. tristero and H. t. linaresi* (second panel), and between *H. t. tristero* and *H. c. cydno* (bottom panel). Horizontal dashed blue lines indicate the 95% quantile of F_{ST} and d_{xy} values across autosomes. The gene coordinates of *regucalcin1* (candidate behavioural gene), as well *optix* and its putative regulatory region/s are indicated by vertical grey dotted lines.



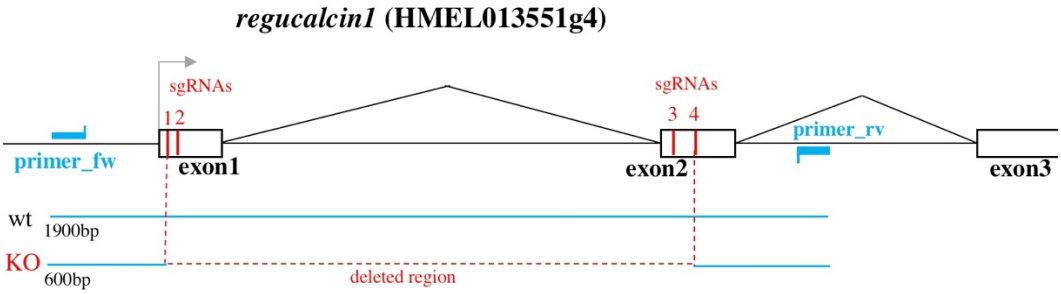
Supplementary Figure 9. Principal component analysis (PCA) of neural gene expression for the 500 genes with most variable expression level across samples. *Heliconius* male samples are colour-coded by subspecies. A vertical dotted line has been drawn to indicate the division (PC1) between individuals that were raised in Panama and in Colombia.

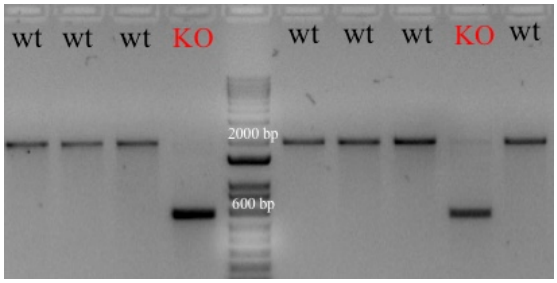


Supplementary figure 10. Top panel: Proportion of courtship time directed towards red *H. t. tristero* females as opposed to *H. t linaresi* females, by *H. t linaresi* males and by *H. t. tristero* males. Dot size is scaled to the number of total minutes a male responded to *H. t. tristero* females (a swarmplot was used to distribute dots horizontally). Estimated marginal means (EMMs) from GLMM models and their 95% confidence intervals are displayed by horizontal and vertical black bars. Bottom panel: Proportion of courtship time directed towards different red *Heliconius* female types as opposed to yellow or white *Heliconius* female types by *H. t linaresi* males. Note that *H. c. cydno* populations are polymorphic for the yellow and white forewing band. tri. = *H. t. tristero,* cyd. W= white *H. c. cydno*, cyd. R= red *H. c. cydno* (artificial), cyd. Y= yellow *H. c. cydno,* lin. = *H. t. linaresi*.



Supplementary Figure 11. On top, schematic representation of the *regucalcin1* locus with the target sites of the 4 injected sgRNAs (sgRNA "1" and "4" mediated the CRISPR cuts), and the primers for genotyping the CRISPR/Cas9 mediated deletion. In the middle, gel electrophoresis of PCR-amplified *regucalcin1* fragments from individuals without (wt = wild-type predicted length, 1900bp) - and with (KO = knock-out, 600bp) deletion. At the bottom, a snapshot of the corresponding sequencing results. PAM (NGG) sequences adjacent to sgRNA "1" and "4" are highlighted in bold and red (Cas9-cleavage is predicted 3bp upstream of the PAM sequence).





wt (..)GCAGTCACTGAGCCGGTG**TGG**C(1343bp)AGTGGGA**CGG**GCTGGGAGAGGCC(..)
wt (..)GCAGTCACTGAGCCGGTG**TGG**C(1343bp)AGTGGGA**CGG**GCTGGGAGAGGCC(..)
KO (..)GCAGTCACTGAGCC----------GGA**CGG**GCTGGGAGAGGCC(..)
KO (..)GCAGTCACTGAGC----------GGA**CGG**GCTGGGAGAGGCC(..)

Paper III

Neural divergence and hybrid disruption between ecologically isolated *Heliconius* butterflies.

Stephen H. Montgomery, Matteo Rossi, Owen McMillan & Richard M. Merrill

Proceedings of the National Academy of Sciences (2021), 118(6), e2015102118
https://doi.org/10.1073/pnas.2015102118

Discussion

Neural systems need to adapt to local conditions to maintain adaptive behavioral function. During species divergence, populations often undergo ecological transitions, that impose new sensory challenges. These challenges can be met by plasticity (e. g. learning) or heritable change/behavioral evolution. Although behaviors can be transmitted across generations without changes in the DNA sequence (for example through epigenetic changes, Dias & Ressler 2014), and although developmental programs can be influenced by environmental variation, variation in behavior is often hardwired in the genome. In particular, species-specific variation in behaviors, which are important for survival or reproduction, is likely to be genetically encoded (Baker et al. 2001). As such, distinct genetic programs can guide the assembly of neural circuits, controlling the expression of adaptive behavior across evolutionary time. However, we still know little of how genes, brain, behavior and evolution are connected. Below I discuss how my results, relating to the genetics of behavioral evolution in *Heliconius* butterflies, contribute to fill this gap.

Neural divergence contributes to speciation in *Heliconius*. In chapter 3, my colleagues and I report that species divergence between *Heliconius melpomene* and *H. cydno/timareta* has involved an ecological transition from open to closed forest habitats, and that this is associated with changes in their neural systems. In particular, species that inhabit closed forest habitats (i.e. *H. cydno* and *H. timareta*) consistently show larger visual processing areas of the brain as compared to *H. melpomene,* which is found in more open forest habitats. We hypothesized that this heritable, adaptive volumetric expansion of visual neuropils in *H. cydno* and *H. timareta* reflects adaptation to dim light environments, increasing their visual sensitivity and maintaining basic behavioral functions. These differences are similar to heritable differences in neuropil investment associated with another ecological transition from low to high altitude (and humid to dry environment) in another *Heliconius* species pair, *H. erato cyrbia* and *H. himera* (Montgomery & Merrill 2017). However, our data additionally show that the same visual structures that differ between *H. melpomene* and *H. cydno*, also show intermediate morphologies in their F1 hybrids, which may disrupt hybrid behavioral function. A potential mismatch of the hybrid neural systems to either parental habitat could act as a reproductive barrier and play a role in maintaining the two taxa as separate species.

In addition to brain volumetric changes, *H. melpomene* and *H. cydno* also differ extensively in their neural gene expression profiles. I revealed that these changes in gene expression are – at least in part – adaptive, likely reflecting alterations of neuronal wiring or activity for optimized sensory/neural function in the two distinct habitats. Importantly, the same genes showing divergent expression in the neural tissue of *H. melpomene* and *H. cydno*, and in particular those with intermediate expression in their F1 hybrids, are located in regions of the genome resistant to gene flow between the two species. This implies selection against these alleles when they introgress between species. Overall, our data, and those of the previous study by Montgomery & Merrill (2017), suggest that broad-scale neural adaptations may contribute substantially to behavioral isolation and speciation across *Heliconius*, and other taxa more generally.

Visual mate preference is independent of adaptation to the broader environment. Sensory drive models of speciation posit that adaptation to different sensory niches can drive the evolution of divergent mating preferences. In a notable example, closely related species of cichlid fish living at different depths in Lake Victoria have visual systems and nuptial male colorations tuned to their local light environment (Seehausen et al. 2008). Importantly, the females of these species show preferences based on these divergent male colorations (blue vs. red), that disappears if the light environment is disrupted / the differences in male coloration are masked (Seehausen & van Alphen 1998). This example and others suggest that divergent mating preferences can emerge as a by-product of divergent selection imposed by sensory environments.

However, this seems not to be the case in our *Heliconius* species. In particular, *H. timareta* inhabits a different sensory environment (closed forest habitat) compared to *H. melpomene*, and seems to have adapted to it at the neural level (*H. timareta* brain morphometrics are not significantly different from *H. cydno*). However, as seen in chapter 2, different subspecies of *H. timareta* display visual mate preference for red butterflies akin to their *H. melpomene* co-mimics. Therefore, *H. timareta* seems to have evolved a parallel (collateral, *sensu* Stern 2013) visual mate preference to *H. melpomene*, independently of adaptation to the broader light environment.

Visual preference loci control divergent mating preferences in the *H. melpomene-cydno* group. *H. melpomene* and *H. cydno* remain reproductively isolated largely because of assortative mating. A series of previous studies have shown that these marked differences in

male mating behavior between *H. melpomene* and *H. cydno* are i) largely driven by visual cues (Jiggins et al. 2001), ii) not the result of phenotype-matching (Merrill et al. 2011) and iii) largely innate/genetically hardwired (Jiggins et a. 2004, Merrill et al. 2011, Merrill et al. 2019). In particular, the segregation of alleles at three genomic regions in hybrid crosses explains a high proportion of the difference in male preference behavior between the two species (~60%), i. e. whether males will spend their time courting *H. melpomene* or *H. cydno* females (Merrill et al. 2019).

In chapter 1, we have provided more evidence that two of these QTL regions (on chromosome 1 and 18) might encode visually-guided preferences. In particular, we found that these regions influence whether males will initiate their courtship towards *H. melpomene* or *H. cydno* females, which likely reflects a response to (visual) cues used in the medium-range (2x2x2m insectary cage). Importantly, this also shows that the influence of these two regions on male relative courtship time is not an artifact of female attraction/rejection behaviors (unpublished experiments suggest that female *H. cydno* do not discriminate male species based on visual cues, Chi-Yun Kuo & Richard Merrill personal communication). In chapter 2, we have provided further evidence that the QTL on chromosome 18 harbors alleles for visual attraction, by showing that it segregates with male preference for *H. cydno* females with red-painted forewings over *H. cydno* females with white (transparently-painted) ones. Together with previous evidence that this region also associates with divergent male preferences for red over white female wing models in *H. melpomene* x *H. cydno* crosses (Merrill et al. 2011), this indicates that chromosome 18 harbors alleles for visual mate preference across the *H. melpomene* group.

Divergence in male visual preference, modulated by the QTL on chromosome 18, may be guided by attraction towards females with a red pattern or aversion for females with a white pattern. However, data from Merrill et al. 2019 suggests that the QTL on chromosome 18 influences the number of courtships directed towards *H. melpomene* females, but not towards *H. cydno females*, implying that *H. melpomene* alleles on chromosome 18 encode attraction towards red butterflies.

In the future, it would be interesting to further investigate how the overall mating preference difference between *H. melpomene* and *H. cydno* is behaviorally and genetically constructed. That is, investigating how the different QTLs (which seem to have mostly additive effects)

contribute together to specify variation in mating behavior between species. For example, do different QTLs modulate behavioral responses towards the same visual cues (e. g. forewing color) or different ones?

Do *Heliconius* butterflies exchange adaptive behavioral alleles? There is considerable evidence that *H. timareta* has repeatedly acquired its warning red pattern elements from *H. melpomene*, through adaptive introgression of *optix* alleles (Pardo-Diaz et al. 2012, Dasmahapatra et al. 2014, Martin et al. 2019, Kozak et al. 2021). The evolution of these "new" mimetic patterns probably required the evolution of the corresponding visual preferences. In chapter 2, we have provided evidence that red *Heliconius timareta* acquired visual preference alleles from red *H. melpomene* (although possibly in a second introgression event after the acquisition of the color patterning alleles).

Intriguingly, we found that the yellow subspecies *H. timareta linaresi* might have also acquired alleles for visual preference from red *H. melpomene*, or at least that these alleles are shared by the adjacent *H. m. bellula* and *H. t. linaresi* populations. This raises the question of why would a yellow species maintain preference alleles for red butterflies? We hypothesized that this might be because these alleles actually encode a visual preference tuned more broadly to red-shifted pattern elements, including yellow. This could be tested by studying whether segregation of red (*H. t. tristero*)/white (*H. c. cydno*) species alleles at chromosome 18 associates also with a preference for yellow over white *H. cydno* females (a natural occurring color polymorphism in some *H. cydno* populations). An alternative explanation is that the locus might encode aversion for white patterns, but this seems unlikely given data from Merrill et al. 2019 discussed above.

In the future, it would also be interesting to test whether the divergent behavior of *H. timareta* and *H. cydno* share a broader genetic architecture with that of *H. melpomene* and *H. cydno* (segregating at chromosome 1 and 17 as well). This might reveal a broader parallel genetic architecture, and might suggest that multiple visual preference alleles were acquired by *H. timareta* from *H. melpomene*, to drive an adaptive shift in visual mate preference.

Change in visual integration or processing likely underlie mate preferences. In chapter 1, we identified five candidate genes for the mate preference shift between *H. melpomene* and *H. cydno*. Based on the functions associated with these genes, we argued that it seems more likely

that divergent visual preferences are mediated by changes in processing or integration of visual information, rather than changes in detection at the sensory periphery (eye). Among the candidates, we had identified an ionotropic glutamate receptor gene, which seemed a particularly promising candidate, as this type of ion channel is associated for example with mediating preference for UV over green light in *Drosophila melanogaster* (Karuppudurai et al. 2014). Later on (in chapter 2), we found that two candidate genes with predicted "regucalcin" function are more strongly associated with visual preference shifts across the *H. melpomene* group. Regucalcin seems to be involved in brain calcium signaling (Yamaguchi 2012), and could alter behavior by, for example, modulating neuronal excitability/activity. Arguably, this could influence the excitability of photoreceptors (Krizaj & Copenaghen 2002) and possibly sensory reception in the eye, but other behavioral considerations might suggest this is not the case. In fact, both species feed on red flowers and probably use color information for foraging. A change altering the detection of visual information at the periphery would probably affect visual perception more broadly, with possible detrimental effects. It seems likely therefore, that the neural mechanism underlying this behavioral shift instead tunes more specifically how visual information is processed or integrated in the brain.

The neural basis of these divergent visual preferences remains unknown. Nevertheless, some other considerations are possible. Although it is not clear which visual cues elicit *Heliconius* male attraction, waving a red rag is enough to elicit attraction from *H. melpomene* males (which will approach the rag to "inspect" it), whereas waving a white rag attracts *H. cydno* (Jiggins 2017). It seems possible therefore, that color and motion information alone could trigger male attraction behaviors. Color would probably need to be "patterned" or "be seen in motion" as "*Heliconius*-like" in order to elicit chasing or courtship behavior from males (as for example many other butterflies might display red/white patterns, and to my knowledge *Heliconius* males do not court the rags). In *Drosophila*, integration of color and motion circuitry starts in the major visual processing centers of the optic lobe, the medulla and lobula (Pagni et al. 2021). From these processing centers, and most notably from the lobula, different visual projection neurons convey distinct visual information to the central brain (optic glomeruli), to activate distinct behavioral programs (Wu et al. 2016) (with more processing in the central brain nonetheless). This might suggest that visual information/cues conveyed from the medulla/lobula to the central brain might be channeled differently between species, to activate chasing or courtship behavior only towards conspecifics.

These visual cues might be particularly relevant in the medium range as *Heliconius* males in the wild are reported to approach different species of *Heliconius* females (or other butterflies) from afar, for then "giving up" the chase. In fact, visually-guided mate recognition could gate a "continue pursuing a female or not" response. In other words, the behavioral shift could involve central neural circuitry controlling motivation/male sexual arousal. In *Drosophila*, a cluster of 20 neurons (P1) in the central brain of males controls sexual arousal (Sten et al. 2021), and functions as a gate for species-specific behavior in some species, where for example it is activated by the perception of conspecific rather than heterospecific female pheromones (Seeholzer et al. 2018). *Heliconius* might possess analogous central circuitry, whose activity is gated by visual cues (and possibly integrating other cues, including for example pheromones). In particular, a central circuit could integrate color pattern information differently between species, to control male sexual arousal towards conspecifics and heterospecifics. Overall, neural pathways between the medulla/lobula (visual processing centers) and central circuits (controlling motivation/sexual arousal), could be likely substrates for encoding divergent visual mate preferences.

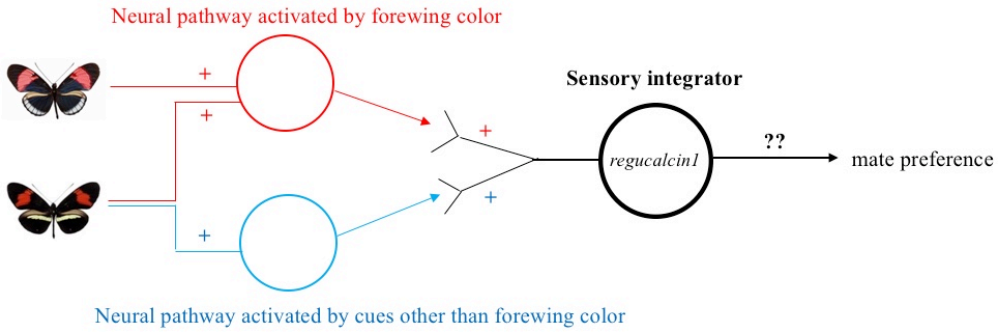
This of course remains speculative. Nonetheless, in the future it could be possible to gain more insight into the neural basis of divergent mating preferences from immunostainings. Revealing where and with which intensity the protein products of candidate genes localize in the brains of different species (with different mating preferences), could highlight specific neuronal populations that might mediate the behavioral shift.

Genetic mechanisms associated with the QTL on chromosome 18. I have identified a *regucalcin* as the strongest candidate gene contributing to shifts in visual mate preference. This gene (*regucalcin1*) is differentially expressed across six red- vs. white-preferring subspecies comparisons. In particular, it is down-regulated in all red-preferring subspecies. This might seem surprising, as alleles for red-preference seem largely dominant, and dominant alleles are perhaps more often associated with up-regulation, or better, expression in a new cell/tissue type where before there was none (Stern 2011). Nevertheless, there could be a repressor element of *regucalcin1* with a dominant effect on behavior.

I detected this change in *regucalcin1* expression in adult eyes and brain tissue. This might indicate that *regucalcin1* alters mate preference by affecting the activity of neural circuits in adults, rather than affecting earlier developmental processes (the assembly of neural circuitry

for example). Nevertheless, time-series expression data of *regucalcin1* (RT-qPCR) is needed to elucidate when this gene expression difference arises in the brain/eyes of red- and white preferring species.

It is perhaps worth noting that the behavioral alleles of the red-preferring species (*H. melpomene*, *H. timareta*) are not completely dominant over those of *H. cydno*. In particular, the behavioral effect of the QTL on chromosome 18 is reduced when hybrid males are presented with differently painted *H. cydno* females compared to when they are presented with different species (*H. cydno* and *H. timareta*) females. This might mean that not all mating cues that guide the preference mechanism encoded at this QTL are present in painted females. For example, the yellow hindwing bar of H. *timareta* is missing in painted females, which might otherwise contribute to trigger a preference for females with red-shifted pattern elements. We haven't directly tested F1 hybrid behavior in choice trials with differently painted females, but it is possible that "red-preferring" alleles would show less complete dominance in this case (compared to when tested with species females). Another option in this case could be that *regucalcin1*/the causal gene affects molecular mechanisms within a sensory integrator circuit (figure below).



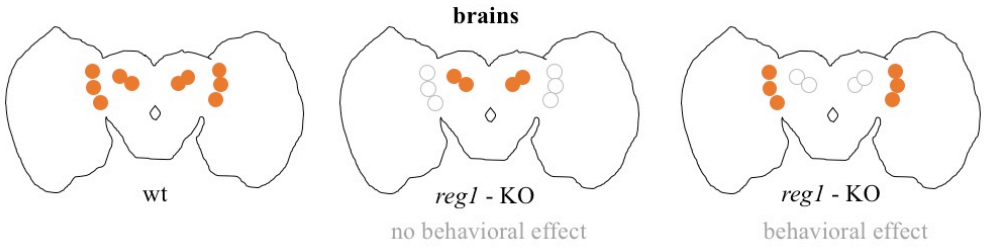
In this hypothetical scenario, *regucalcin1* alleles could affect the excitability of a neuron involved in sensory integration. For example, the *H. timareta* allele might make this neuron more likely to fire an action potential compared to the *H. cydno* allele, and this might scale up depending on allele dose. If only the forewing color cue is presented to males (painted females), then the sensory integrator neuron would be activated with less probability than if multiple "conspecific" cues are present (*H. timareta* female). In this way, the *regucalcin1 H. timareta* allele might appear to show incomplete dominance, in that only one of its copies might often fail to trigger the expression of a preference for red butterflies.

Functionally validating the link between genes and behavior. I have produced a handful of *regucalcin1* knock-outs (KO) individuals, which sets the stage for functionally validating the link between genes and behavior. Of the 10 *regucalcin1*-KO butterflies generated, 6 reached the pupal stage, and 3 adults eclosed**.** Two of these mosaic (female) individuals did not show evidence of a severe impairment of behavior (they could fly and land on cage walls for example), however one (a male) was not able to fly (and seemed to have difficulty feeding). This suggests that either *regucalcin1*-KO might produce a range of behavioral effects (perhaps linked to in which neural cells *regucalcin1*-KO has occurred), or that behavioral effects are sex-specific and KO impairs only male behavior. In case the former is true, in principle we could perform behavioral assays in the future, that might validate the link between *regucalcin1* and mating behavior.

In particular, to test the possible effect of *regucalcin1* on mating behavior/preference, we could present *H. melpomene regucalcin1*-KO males and wild-types (wt) with a *H. melpomene* female. *H. melpomene regucalcin1*-KO males could show reduced attraction (less amount of chasing and courtship behavior) towards *H. melpomene* females relative to wild-types males. Mate preference assays could be coupled with other behavioral assays to ensure that color vision or visual sensorimotor skills of *regucalcin1*-KO individuals are not more generally impaired. For example, after choice/no-choice trials with females, KOs and wild-type males could be tested with color conditioning experiments, i.e. if they can be trained on associating color with reward/food source (as has been done for example with *H. erato* males, Zaccardi et al. 2006). Another possibility would be to assay the optomotor response of both males and females to test their visual sensorimotor skills. Ideally, we would test the link between both *regucalcin1* and *regucalcin2* with behavior. If for example *regucalcin2*-KO should not affect mating behavior in these assays, this might further suggest that *regucalcin1* is the gene within the QTL on chromosome 18 modulating mate preference.

To potentially shed light on the neural basis of divergent mating behaviors, immunostainings in knock-out males could be performed after the behavioral assays. Inactivation of genes through a knock-out generally results in protein product not being produced/being degraded. Nevertheless, given that we would generate mosaic-KO individuals (with CRISPR-mediated deletion occurring only in some cells), the gene might not be inactivated in all neural cells it would normally be expressed in. This might eventually associate with the behavioral phenotype of *regucalcin1*-KO individuals. Note that this experiment would be informative only if

behavior/the expression of mate preference does not rely on the causal gene being transcribed in all neural cells it is normally transcribed in. Nevertheless, with enough replicates, this experiment has the potential to highlight neuronal populations required to mediate the behavioral shift (figure below).



In this hypothetical scenario, the protein product of *regucalcin1* is found in only a subset of neural cells (protein found = full orange circles, protein not found = gray circles) in knock-out (KO) individuals compared to wild-types (wt). This might associate with the behavioral phenotype in KO individuals. In this example, only KO in a subset of neural (central) cells has an effect on mating behavior.

Eventually, if the knock-out of *regucalcin1* disrupted for example male attraction behaviors towards *H. melpomene* females (or disrupted behavior more generally), an ideal follow-up experiment would be to knock the *H. melpomene*/*H. timareta regucalcin1* allele in *H. cydno*. If *H. cydno* acquired increased attraction towards red butterflies this would perhaps unequivocally establish the link between *regucalcin1* and behavior. This experiment would probably need to be preceded by other analyses, to identify more specific genetic elements linked to behavior, i.e. possible regulatory elements, to inform on which genetic region to swap between species.

Conclusion. How can complex differences in innate behaviors between species be transmitted across generations? How can for example an elaborate courtship ritual, or perhaps the shape of a spider's web, be encoded in a string of "A","C", "G" and "T"s? We don't really know, especially when it comes to visually-guided behaviors. Nevertheless, with the expansion of genetic resources and tools to non-traditional laboratory organisms, there are new possibilities to tackle this question across taxa. In particular, one could exploit the evolution of behavioral differences between closely related-species (e.g. divergent mating preferences), to map the links between genes and behavior. Color pattern-guided preferences across species of *Heliconius* butterflies is an excellent system to start probing these links. With this thesis, I hope to have resolved in more detail, perhaps down to the gene level, how variation in the DNA can generate this striking diversity of visually-guided behaviors.

References

Ahmed, O. M., Avila-Herrera, A., Tun, K. M., Serpa, P. H., Peng, J., Parthasarathy, S., ... & Shah, N. M. (2019). Evolution of mechanisms that control mating in Drosophila males. *Cell reports*, *27*(9), 2527-2536.

Arnqvist, G., & Rowe, L. (2013). *Sexual conflict*. Princeton University Press.

Auer, T. O., & Benton, R. (2016). Sexual circuitry in Drosophila. *Current opinion in neurobiology*, *38*, 18-26.

Auer, T. O., Khallaf, M. A., Silbering, A. F., Zappia, G., Ellis, K., Álvarez-Ocaña, R., ... & Benton, R. (2020). Olfactory receptor and circuit evolution promote host specialization. *Nature*, *579*(7799), 402-408.

Baker, B. S., Taylor, B. J., & Hall, J. C. (2001). Are complex behaviors specified by dedicated regulatory genes? Reasoning from Drosophila. *Cell*, *105*(1), 13-24.

Bay, R. A., Arnegard, M. E., Conte, G. L., Best, J., Bedford, N. L., McCann, S. R., ... & Peichel, C. L. (2017). Genetic coupling of female mate choice with polygenic ecological divergence facilitates stickleback speciation. *Current Biology*, *27*(21), 3344-3349.

Bargmann, C. I. (2012). Beyond the connectome: how neuromodulators shape neural circuits. *Bioessays*, *34*(6), 458-465.

Barton, N., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, *57*(3), 357-376.

Bendesky, A., & Bargmann, C. I. (2011). Genetic contributions to behavioural diversity at the gene–environment interface. *Nature Reviews Genetics*, *12*(12), 809-820.

Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L., & Bargmann, C. I. (2011). Catecholamine receptor polymorphisms affect decision-making in C. elegans. *Nature*, *472*(7343), 313-318.

Bendesky, A., Kwon, Y. M., Lassance, J. M., Lewarch, C. L., Yao, S., Peterson, B. K., ... & Hoekstra, H. E. (2017). The genetic basis of parental care evolution in monogamous mice. *Nature*, *544*(7651), 434-439.

Blankers, T., Berdan, E. L., Hennig, R. M., & Mayer, F. (2019a). Physical linkage and mate preference generate linkage disequilibrium for behavioral isolation in two parapatric crickets. *Evolution*, *73*(4), 777-791.

Blankers, T., Oh, K. P., & Shaw, K. L. (2019b). Parallel genomic architecture underlies repeated sexual signal divergence in Hawaiian Laupala crickets. *Proceedings of the Royal Society B*, *286*(1912), 20191479.

Bloch, N. I., Corral-López, A., Buechel, S. D., Kotrschal, A., Kolm, N., & Mank, J. E. (2018). Early neurogenomic response associated with variation in guppy female mate preference. *Nature ecology & evolution*, *2*(11), 1772-1781.

Borst, A., Haag, J., & Mauss, A. S. (2020). How fly neurons compute the direction of visual motion. *Journal of Comparative Physiology A*, *206*(2), 109-124.

Bousquet, F., Nojima, T., Houot, B., Chauvel, I., Chaudy, S., Dupas, S., ... & Ferveur, J. F. (2012). Expression of a desaturase gene, desat1, in neural and nonneural tissues separately affects perception and emission of sex pheromones in Drosophila. *Proceedings of the National Academy of Sciences*, *109*(1), 249-254.

Brand, P., Hinojosa-Díaz, I. A., Ayala, R., Daigle, M., Obiols, C. L. Y., Eltz, T., & Ramírez, S. R. (2020). The evolution of sexual signaling is linked to odorant receptor tuning in perfume-collecting orchid bees. *Nature communications*, *11*(1), 1-11.

Briscoe, A. D., Bybee, S. M., Bernard, G. D., Yuan, F., Sison-Mangus, M. P., Reed, R. D., ... & Chiao, C. C. (2010). Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in Heliconius butterflies. *Proceedings of the National Academy of Sciences*, *107*(8), 3628-3633.

Cande, J., Prud'homme, B., & Gompel, N. (2013). Smells like evolution: the role of chemoreceptor evolution in behavioral change. *Current opinion in neurobiology*, *23*(1), 152-158.

Cande, J., Stern, D. L., Morita, T., Prud'homme, B., & Gompel, N. (2014). Looking under the lamp post: neither fruitless nor doublesex has evolved to generate divergent male courtship in Drosophila. *Cell reports*, *8*(2), 363-370.

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, *134*(1), 25-36.

Chamberlain, N. L., Hill, R. I., Kapan, D. D., Gilbert, L. E., & Kronforst, M. R. (2009). Polymorphic butterfly reveals the missing link in ecological speciation. *Science*, *326*(5954), 847-850.

Cheng, K. Y., & Frye, M. A. (2020). Neuromodulation of insect motion vision. *Journal of Comparative Physiology A*, *206*(2), 125-137.

Chowdhury, T., Calhoun, R. M., Bruch, K., & Moehring, A. J. (2020). The fruitless gene affects female receptivity and species isolation. *Proceedings of the Royal Society B*, *287*(1923), 20192765.

Chouteau, M., Arias, M., & Joron, M. (2016). Warning signals are under positive frequency-dependent selection in nature. *Proceedings of the national Academy of Sciences*, *113*(8), 2164-2169.

Clowney, E. J., Iguchi, S., Bussell, J. J., Scheer, E., & Ruta, V. (2015). Multimodal chemosensory circuits controlling male courtship in Drosophila. *Neuron*, *87*(5), 1036-1049.

Cochella, L., & Hobert, O. (2012). Diverse functions of microRNAs in nervous system development. *Current topics in developmental biology*, *99*, 115-143.

Combs, P. A., Krupp, J. J., Khosla, N. M., Bua, D., Petrov, D. A., Levine, J. D., & Fraser, H. B. (2018). Tissue-specific cis-regulatory divergence implicates eloF in inhibiting interspecies mating in Drosophila. *Current biology*, *28*(24), 3969-3975.

Concha, C., Wallbank, R. W., Hanly, J. J., Fenner, J., Livraghi, L., Rivera, E. S., ... & McMillan, W. O. (2019). Interplay between developmental flexibility and determinism in the evolution of mimetic Heliconius wing patterns. *Current Biology*, *29*(23), 3996-4009.

Coyne, J. A., & Orr, H. A. (1989). Patterns of speciation in Drosophila. *Evolution*, *43*(2), 362-381.

Coyne, J. A., Orr, H. A. *Speciation* (Sunderland, MA: Sinauer 2004).

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular ecology*, *23*(13), 3133-3157.

Cummings, M. E. (2007). Sensory trade-offs predict signal divergence in surfperch. *Evolution*, *61*(3), 530-545.

Cummings, M. E., & Endler, J. A. (2018). 25 Years of sensory drive: the evidence and its watery bias. *Current Zoology*, *64*(4), 471-484.

Davey, J. W., Chouteau, M., Barker, S. L., Maroja, L., Baxter, S. W., Simpson, F., ... & Jiggins, C. D. (2016). Major improvements to the Heliconius melpomene genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3: Genes, Genomes, Genetics*, *6*(3), 695-708.

Dasmahapatra, K. K., Walters, J. R., Briscoe, A. D., Davey, J. W., Whibley, A., Nadeau, N. J., ... & Heliconius Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*(7405), 94.

Dean, R., & Mank, J. E. (2014). The role of sex chromosomes in sexual dimorphism: discordance between molecular and phenotypic data. *Journal of Evolutionary Biology*, *27*(7), 1443-1453.

De Bono, M., & Bargmann, C. I. (1998). Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. *Cell*, *94*(5), 679-689.

Delclos, P. J., Forero, S. A., & Rosenthal, G. G. (2020). Divergent neurogenomic responses shape social learning of both personality and mate preference. *Journal of Experimental Biology*, *223*(6).

Dias, B. G., & Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nature neuroscience*, *17*(1), 89-96.

Ding, Y., Berrocal, A., Morita, T., Longden, K. D. & Stern, D. L. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* **536**, 329–332 (2016).

Ding, Y., Lillvis, J. L., Cande, J., Berman, G. J., Arthur, B. J., Long, X., ... & Stern, D. L. (2019). Neural evolution of context-dependent fly song. *Current biology*, *29*(7), 1089-1099.

Ellegren, H., & Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, *8*(9), 689-698.

Ellegren, H. (2011). Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics*, *12*(3), 157-166.

Estrada, C., & Jiggins, C. D. (2002). Patterns of pollen feeding and habitat preference among Heliconius species. *Ecological Entomology*, *27*(4), 448-456.

Fan, P., Manoli, D. S., Ahmed, O. M., Chen, Y., Agarwal, N., Kwong, S., ... & Shah, N. M. (2013). Genetic and neural mechanisms that inhibit Drosophila from mating with other species. *Cell*, *154*(1), 89-102.

Finkbeiner, S. D., Briscoe, A. D., & Reed, R. D. (2014). Warning signals are seductive: relative contributions of color and pattern to predator avoidance and mate attraction in Heliconius butterflies. *Evolution*, *68*(12), 3410-3420.

Finkbeiner, S. D., & Briscoe, A. D. (2021). True UV color vision in a female butterfly with two UV opsins. *Journal of Experimental Biology*, *224*(18), jeb242802.

Galouzis, C. C., & Prud'homme, B. (2021). Transvection regulates the sex-biased expression of a fly X-linked gene. *Science*, *371*(6527), 396-400.

Gao, S., Takemura, S. Y., Ting, C. Y., Huang, S., Lu, Z., Luan, H., ... & Lee, C. H. (2008). The neural substrate of spectral preference in Drosophila. *Neuron*, *60*(2), 328-342.

Garner, A. G., Goulet, B. E., Farnitano, M. C., Molina-Henao, Y. F., & Hopkins, R. (2018). Genomic signatures of reinforcement. *Genes*, *9*(4), 191.

Greenwood, A. K., Wark, A. R., Yoshida, K., & Peichel, C. L. (2013). Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Current Biology*, *23*(19), 1884-1888.

Hausmann, A. E., Kuo, C. Y., Freire, M., Rueda-M, N., Linares, M., Pardo-Diaz, C., ... & Merrill, R. M. (2021). Light environment influences mating behaviours during the early stages of divergence in tropical butterflies. *Proceedings of the Royal Society B*, *288*(1947), 20210157.

Hench, K., Vargas, M., Höppner, M. P., McMillan, W. O., & Puebla, O. (2019). Inter-chromosomal coupling between vision and pigmentation genes during genomic divergence. *Nature ecology & evolution*, *3*(4), 657-667.

Horth, L. (2007). Sensory genes and mate choice: evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics*, *90*(2), 159-175.

Hu, C. K., & Hoekstra, H. E. (2017). Peromyscus burrowing: a model system for behavioral evolution. In *Seminars in cell & developmental biology* (Vol. 61, pp. 107-114). Academic Press.

Jiggins, C. D., Naisbit, R. E., Coe, R. L., & Mallet, J. (2001). Reproductive isolation caused by colour pattern mimicry. *Nature*, *411*(6835), 302-305.

Jiggins, C. D., Estrada, C., & Rodrigues, A. (2004). Mimicry and the evolution of premating isolation in Heliconius melpomene Linnaeus. *Journal of evolutionary biology*, *17*(3), 680-691.

Jiggins, C. D. (2008). Ecological speciation in mimetic butterflies. *BioScience*, *58*(6), 541-548.

Jiggins, C. D., & Martin, S. H. (2017). Glittering gold and the quest for Isla de Muerta. *Journal of evolutionary biology*, *30*(8), 1509-1511.

Jiggins, C. D. (2017). *The ecology and evolution of Heliconius butterflies*. Oxford University Press.

Karuppudurai, T., Lin, T. Y., Ting, C. Y., Pursley, R., Melnattur, K. V., Diao, F., ... & Lee, C. H. (2014). A hard-wired glutamatergic circuit pools and relays UV signals to mediate spectral preference in Drosophila. *Neuron*, *81*(3), 603-615.

Katz, P. S., & Harris-Warrick, R. M. (1999). The evolution of neuronal circuits underlying species-specific behavior. *Current opinion in neurobiology*, *9*(5), 628-633.

Keshavarz, M., & Tautz, D. (2021). The imprinted lncRNA Peg13 regulates sexual preference and the sex-specific brain transcriptome in mice. *Proceedings of the National Academy of Sciences*, *118*(10).

Khallaf, M. A., Cui, R., Weißflog, J., Erdogmus, M., Svatoš, A., Dweck, H. K., ... & Knaden, M. (2021). Large-scale characterization of sex pheromone communication systems in Drosophila. *Nature Communications*, *12*(1), 1-14.

Khallaf, M. A., Auer, T. O., Grabe, V., Depetris-Chauvin, A., Ammagarahalli, B., Zhang, D. D., ... & Knaden, M. (2020). Mate discrimination among subspecies through a conserved olfactory pathway. *Science advances*, *6*(25), eaba5279.

Kimchi, T., Xu, J., & Dulac, C. (2007). A functional circuit underlying male sexual behaviour in the female mouse brain. *Nature*, *448*(7157), 1009-1014.

Kirkpatrick, M. (1982). Sexual selection and the evolution of female choice. *Evolution*, 1-12.

Kocher, S. D., Mallarino, R., Rubin, B. E., Douglas, W. Y., Hoekstra, H. E., & Pierce, N. E. (2018). The genetic basis of a social polymorphism in halictid bees. *Nature communications*, *9*(1), 1-7.

Kopp, M., Servedio, M. R., Mendelson, T. C., Safran, R. J., Rodríguez, R. L., Hauber, M. E., ... & Van Doorn, G. S. (2018). Mechanisms of assortative mating in speciation with gene flow: connecting theory and empirical research. *The American Naturalist*, *191*(1), 1-20.

Koutroumpa, F. A., Groot, A. T., Dekker, T., & Heckel, D. G. (2016). Genetic mapping of male pheromone response in the European corn borer identifies candidate genes regulating neurogenesis. *Proceedings of the National Academy of Sciences*, *113*(42), E6401-E6408.

Kozak, K. M., Joron, M., McMillan, W. O., & Jiggins, C. D. (2021). Rampant genome-wide admixture across the Heliconius radiation. *Genome Biology and Evolution*.

Krizaj, D., & Copenhagen, D. R. (2002). Calcium regulation in photoreceptors. *Frontiers in bioscience: a journal and virtual library*, *7*, d2023.

Kronforst, M. R., Young, L. G., Kapan, D. D., McNeely, C., O'Neill, R. J., & Gilbert, L. E. (2006). Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences*, *103*(17), 6575-6580.

Kurtovic, A., Widmer, A., & Dickson, B. J. (2007). A single class of olfactory neurons mediates behavioural responses to a Drosophila sex pheromone. *Nature*, *446*(7135), 542-546.

Lassance, J. M., Groot, A. T., Liénard, M. A., Antony, B., Borgwardt, C., Andersson, F., ... & Löfstedt, C. (2010). Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature*, *466*(7305), 486-489.

Laturney, M., & Moehring, A. J. (2012). Fine-scale genetic analysis of species-specific female preference in Drosophila simulans. *Journal of evolutionary biology*, *25*(9), 1718-1731.

Leary, G. P., Allen, J. E., Bunger, P. L., Luginbill, J. B., Linn, C. E., Macallister, I. E., ... & Wanner, K. W. (2012). Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proceedings of the National Academy of Sciences*, *109*(35), 14081-14086.

Levine, M., & Davidson, E. H. (2005). Gene regulatory networks for development. *Proceedings of the National Academy of Sciences*, *102*(14), 4936-4942.

Lewis, J. J., Van Belleghem, S. M., Papa, R., Danko, C. G., & Reed, R. D. (2020). Many functionally connected loci foster adaptive diversification along a neotropical hybrid zone. *Science advances*, *6*(39), eabb8617.

Lewontin, R. C., & Kojima, K. I. (1960). The evolutionary dynamics of complex polymorphisms. *Evolution*, 458-472.

Li, Q., Lee, J. A., & Black, D. L. (2007). Neuronal regulation of alternative pre-mRNA splicing. *Nature Reviews Neuroscience*, *8*(11), 819-831.

Liénard, M. A., Bernard, G. D., Allen, A., Lassance, J. M., Song, S., Childers, R. R., ... & Pierce, N. E. (2021). The evolution of red color vision is linked to coordinated rhodopsin tuning in lycaenid butterflies. *Proceedings of the National Academy of Sciences*, *118*(6).

Longden, K. D. (2016). Central brain circuitry for color-vision-modulated behaviors. *Current Biology*, *26*(20), R981-R988.

Mallet, J., & Barton, N. H. (1989). Strong natural selection in a warning-color hybrid zone. *Evolution*, *43*(2), 421-431.

Manoli, D. S., Foss, M., Villella, A., Taylor, B. J., Hall, J. C., & Baker, B. S. (2005). Male-specific fruitless specifies the neural substrates of Drosophila courtship behaviour. *Nature*, *436*(7049), 395-400.

Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS biology*, *17*(2), e2006288.

Massey, J. H., Chung, D., Siwanowicz, I., Stern, D. L., & Wittkopp, P. J. (2019). The yellow gene influences Drosophila male mating success through sex comb melanization. *Elife*, *8*, e49388.

Massey, J. H., Rice, G. R., Firdaus, A. S., Chen, C. Y., Yeh, S. D., Stern, D. L., & Wittkopp, P. J. (2020). Co-evolving wing spots and mating displays are genetically separable traits in Drosophila. *Evolution*, *74*(6), 1098-1111.

Matute, D. R., & Cooper, B. S. (2021). Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution*, *75*(4), 764-778.

Mayr, E. (1963). *Animal species and evolution*. Harvard University Press.

McBride, C. S., Baier, F., Omondi, A. B., Spitzer, S. A., Lutomiah, J., Sang, R., ... & Vosshall, L. B. (2014). Evolution of mosquito preference for humans linked to an odorant receptor. *Nature*, *515*(7526), 222-227.

McCulloch, K. J., Yuan, F., Zhen, Y., Aardema, M. L., Smith, G., Llorente-Bousquets, J., ... & Briscoe, A. D. (2017). Sexual dimorphism and retinal mosaic diversification following the evolution of a violet receptor in butterflies. *Molecular biology and evolution*, *34*(9), 2271-2284

McKinnon, J. S., & Pierotti, M. E. (2010). Colour polymorphism and correlated characters: genetic mechanisms and evolution. *Molecular ecology*, *19*(23), 5101-5125.
McMillan, W. O., Jiggins, C. D., & Mallet, J. (1997). What initiates speciation in passion-vine butterflies?. *Proceedings of the National Academy of Sciences*, *94*(16), 8628-8633.
McNiven, V. T., & Moehring, A. J. (2013). Identification of genetically linked female preference and male trait. *Evolution*, *67*(8), 2155-2165.
Mendelson, T. C. (2003). Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: Etheostoma). *Evolution*, *57*(2), 317-327.
Merot, C., Salazar, C., Merrill, R. M., Jiggins, C. D., & Joron, M. (2017). What shapes the continuum of reproductive isolation? Lessons from Heliconius butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856), 20170335.
Merrill, R. M., Gompert, Z., Dembeck, L. M., Kronforst, M. R., McMillan, W. O., & Jiggins, C. D. (2011a). Mate preference across the speciation continuum in a clade of mimetic butterflies. *Evolution: International Journal of Organic Evolution*, *65*(5), 1489-1500.
Merrill, R. M., Van Schooten, B., Scott, J. A., & Jiggins, C. D. (2011b). Pervasive genetic associations between traits causing reproductive isolation in Heliconius butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1705), 511-518.
Merrill, R. M., Wallbank, R. W., Bull, V., Salazar, P. C., Mallet, J., Stevens, M., & Jiggins, C. D. (2012). Disruptive ecological selection on a mating cue. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4907-4913.
Merrill, R. M., Naisbit, R. E., Mallet, J., & Jiggins, C. D. (2013). Ecological and genetic factors influencing the transition between host-use strategies in sympatric Heliconius butterflies. *Journal of evolutionary biology*, *26*(9), 1959-1967.
Merrill, R. M., Rastas, P., Martin, S. H., Melo, M. C., Barker, S., Davey, J., ... & Jiggins, C. D. (2019). Genetic dissection of assortative mating behavior. *PLoS biology*, *17*(2), e2005902.
Merritt, J. R., Grogan, K. E., Zinzow-Kramer, W. M., Sun, D., Ortlund, E. A., Soojin, V. Y., & Maney, D. L. (2020). A supergene-linked estrogen receptor drives alternative phenotypes in a polymorphic songbird. *Proceedings of the National Academy of Sciences*, *117*(35), 21673-21680.
Metz, H. C., Bedford, N. L., Pan, Y. L., & Hoekstra, H. E. (2017). Evolution and genetics of precocious burrowing behavior in Peromyscus mice. *Current Biology*, *27*(24), 3837-3845.
Moehring, A. J., Li, J., Schug, M. D., Smith, S. G., DeAngelis, M., Mackay, T. F., & Coyne, J. A. (2004). Quantitative trait loci for sexual isolation between Drosophila simulans and D. mauritiana. *Genetics*, *167*(3), 1265-1274.
Moehring, A. J., Llopart, A., Elwyn, S., Coyne, J. A., & Mackay, T. F. (2006). The genetic basis of prezygotic reproductive isolation between Drosophila santomea and D. yakuba due to mating preference. *Genetics*, *173*(1), 215-223.
Montgomery, S. H., & Merrill, R. M. (2017). Divergence in brain composition during the early stages of ecological specialization in Heliconius butterflies. *Journal of Evolutionary Biology*, *30*(3), 571-582.
Montgomery, S. H., Rossi, M., McMillan, W. O., & Merrill, R. M. (2021). Neural divergence and hybrid disruption between ecologically isolated Heliconius butterflies. *Proceedings of the National Academy of Sciences*, *118*(6).
Niepoth, N., & Bendesky, A. (2020). How natural genetic variation shapes behavior. *Annual review of genomics and human genetics*, *21*, 437-463.
Nosil, P., Crespi, B. J., Gries, R., & Gries, G. (2007). Natural selection and divergence in mate preference during speciation. *Genetica*, *129*(3), 309.
Nosil, P. (2012). *Ecological speciation*. Oxford University Press.
Nosil, P., Feder, J. L., & Gompert, Z. (2021). How many genetic changes create new species?. *Science*, *371*(6531), 777-779.

Okuyama, T., Yokoi, S., Abe, H., Isoe, Y., Suehiro, Y., Imada, H., ... & Takeuchi, H. (2014). A neural mechanism underlying mating preferences for familiar individuals in medaka fish. *Science*, *343*(6166), 91-94.

Ortíz-Barrientos, D., & Noor, M. A. (2005). Evidence for a one-allele assortative mating locus. *Science*, *310*(5753), 1467-1467.

Osborne, K. A., Robichon, A., Burgess, E., Butland, S., Shaw, R. A., Coulthard, A., ... & Sokolowski, M. B. (1997). Natural behavior polymorphism due to a cGMP-dependent protein kinase of Drosophila. *Science*, *277*(5327), 834-836.

Otsuna, H., Shinomiya, K., & Ito, K. (2014). Parallel neural pathways in higher visual centers of the Drosophila brain that mediate wavelength-specific behavior. *Frontiers in neural circuits*, *8*, 8.

Pagni, M., Haikala, V., Oberhauser, V., Meyer, P. B., Reiff, D. F., & Schnaitmann, C. (2021). Interaction of "chromatic" and "achromatic" circuits in Drosophila color opponent processing. *Current Biology*, *31*(8), 1687-1698.

Pardo-Diaz, C., Salazar, C., Baxter, S. W., Merot, C., Figueiredo-Ready, W., Joron, M., ... & Jiggins, C. D. (2012). Adaptive introgression across species boundaries in Heliconius butterflies. *PLoS genetics*, *8*(6), e1002752.

Peichel, C. L., & Marques, D. A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1713), 20150486.

Perry, M., Konstantinides, N., Pinto-Teixeira, F., & Desplan, C. (2017). Generation and evolution of neural cell types and circuits: insights from the Drosophila visual system. *Annual review of genetics*, *51*, 501-527.

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Müller, I., ... & Wolf, J. B. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, *344*(6190), 1410-1414

Price, T. D. (2017). Sensory drive, color, and color vision. *The American Naturalist*, *190*(2), 157-170.

Prieto-Godino, L. L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A. F., ... & Benton, R. (2017). Evolution of acid-sensing olfactory circuits in Drosophilids. *Neuron*, *93*(3), 661-676.

Prud'homme, B., Gompel, N., & Carroll, S. B. (2007). Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences*, *104*(suppl 1), 8605-8612

Pryke, S. R. (2010). Sex chromosome linkage of mate preference and color signal maintains assortative mating between interbreeding finch morphs. *Evolution: International Journal of Organic Evolution*, *64*(5), 1301-1310.

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., ... & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of evolutionary biology*, *30*(8), 1450-1477.

Rezával, C., Pattnaik, S., Pavlou, H. J., Nojima, T., Brüggemeier, B., D'Souza, L. A., ... & Goodwin, S. F. (2016). Activation of latent courtship circuitry in the brain of Drosophila females induces male-like behaviors. *Current Biology*, *26*(18), 2508-2515.

Rosenthal, G. G. *Mate choice* (Princeton University Press, 2017).

Rossi, M., Hausmann, A. E., Thurman, T. J., Montgomery, S. H., Papa, R., Jiggins, C. D., ... & Merrill, R. M. (2020). Visual mate preference evolution during butterfly speciation is linked to neural processing genes. *Nature communications*, *11*(1), 1-10.

Ryan, M. J., & Cummings, M. E. (2013). Perceptual biases and mate choice. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 437-459.

Ryan, M. J. (2021). Darwin, sexual selection, and the brain. *Proceedings of the National Academy of Sciences*, *118*(8).

Sæther, S. A., Sætre, G. P., Borge, T., Wiley, C., Svedin, N., Andersson, G., ... & Qvarnström, A. (2007). Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *science*, *318*(5847), 95-97.

Sandkam, B., Young, C. M., & Breden, F. (2015). Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (Poecilia reticulata). *Molecular ecology*, *24*(3), 596-609.

Sato, K., & Yamamoto, D. (2020a). Contact-Chemosensory Evolution Underlying Reproductive Isolation in Drosophila Species. *Frontiers in Behavioral Neuroscience*, *14*.

Sato, K., & Yamamoto, D. (2020b). The mode of action of Fruitless: Is it an easy matter to switch the sex?. *Genes, Brain and Behavior*, *19*(2), e12606.

Schnaitmann, C., Haikala, V., Abraham, E., Oberhauser, V., Thestrup, T., Griesbeck, O., & Reiff, D. F. (2018). Color processing in the early visual system of Drosophila. *Cell*, *172*(1-2), 318-330.

Schnaitmann, C., Pagni, M., & Reiff, D. F. (2020). Color vision in insects: insights from Drosophila. *Journal of Comparative Physiology A*, *206*(2), 183-198.

Schneiderman, A. M., Hildebrand, J. G., Brennan, M. M., & Tumlinson, J. H. (1986). Trans-sexually grafted antennae alter pheromone-directed behaviour in a moth. *Nature*, *323*(6091), 801-803.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*(5915), 737-741.

Searcy, W. A., & Andersson, M. (1986). Sexual selection and the evolution of song. *Annual Review of Ecology and Systematics*, *17*(1), 507-533.

Seehausen, O., & van Alphen, J. J. (1998). The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (Haplochromis nyererei complex). *Behavioral Ecology and Sociobiology*, *42*(1), 1-8.

Seehausen, O., Terai, Y., Magalhaes, I. S., Carleton, K. L., Mrosso, H. D., Miyagi, R., ... & Okada, N. (2008). Speciation through sensory drive in cichlid fish. *Nature*, *455*(7213), 620-626.

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., ... & Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176-192.

Seeholzer, L. F., Seppo, M., Stern, D. L., & Ruta, V. (2018). Evolution of a central neural circuit underlies *Drosophila* mate preferences. *Nature*, *559*(7715), 564-569.

Servedio, M. R., & Noor, M. A. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 339-364.

Servedio, M. R. (2009). The role of linkage disequilibrium in the evolution of premating isolation. *Heredity*, *102*(1), 51-56.

Shahandeh, M. P., Pischedda, A., Rodriguez, J. M., & Turner, T. L. (2020). The genetics of male pheromone preference difference between Drosophila melanogaster and Drosophila simulans. *G3: Genes, Genomes, Genetics*, *10*(1), 401-415.

Shahandeh, M. P., & Turner, T. L. (2020). The complex genetic architecture of male mate choice evolution between *Drosophila* species. *Heredity*, *124*(6), 737-750.

Shaw, K. L., & Lesnick, S. C. (2009). Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proceedings of the National Academy of Sciences*, *106*(24), 9737-9742.

Siju, K. P., De Backer, J. F., & Kadow, I. C. G. (2021). Dopamine modulation of sensory processing and adaptive behavior in flies. *Cell and Tissue Research*, 1-19.

Smadja, C. M., Loire, E., Caminade, P., Thoma, M., Latour, Y., Roux, C., ... & Boursot, P. (2015). Seeking signatures of reinforcement at the genetic level: a hitchhiking mapping and candidate gene approach in the house mouse. *Molecular ecology*, *24*(16), 4222-4237.

Song, B. M., & Lee, C. H. (2018). Toward a mechanistic understanding of color vision in insects. *Frontiers in Neural Circuits*, *12*, 16.

Sten, T., Li, R., Otopalik, A., & Ruta, V. (2021). Sexual arousal gates visual processing during Drosophila courtship. *Nature*, *595*(7868), 549-553.

Stern, D. L. (2011). *Evolution, development, & the predictable genome*. Roberts and Co. Publishers.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, *14*(11), 751-764.

Svensson, E. I., Arnold, S. J., Bürger, R., Csilléry, K., Draghi, J., Henshaw, J. M., ... & Runemark, A. (2021). Correlational selection in the age of genomics. *Nature ecology & evolution*, *5*(5), 562-573.

Swihart, C. A. (1971). Colour discrimination by the butterfly Heliconius charitonius Linn. *Animal Behaviour*, *19*(1), 156-164.

Tait, C., Kharva, H., Schubert, M., Kritsch, D., Sombke, A., Rybak, J., ... & Olsson, S. B. (2021). A reversal in sensory processing accompanies ongoing ecological divergence and speciation in Rhagoletis pomonella. *Proceedings of the Royal Society B*, *288*(1947), 20210192.

Taylor, R. C., & Ryan, M. J. (2013). Interactions of multisensory components perceptually rescue túngara frog mating signals. *Science*, *341*(6143), 273-274.

Theunissen, F. E., & Elie, J. E. (2014). Neural processing of natural sounds. *Nature Reviews Neuroscience*, *15*(6), 355-366.

Tosches, M. A. (2017). Developmental and genetic mechanisms of neural circuit evolution. *Developmental biology*, *431*(1), 16-25.

Touhara, K., & Vosshall, L. B. (2009). Sensing odorants and pheromones with chemosensory receptors. *Annual review of physiology*, *71*, 307-332.

Troemel, E. R., Kimmel, B. E., & Bargmann, C. I. (1997). Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in C. elegans. *Cell*, *91*(2), 161-169.

Turbek, S. P., Browne, M., Di Giacomo, A. S., Kopuchian, C., Hochachka, W. M., Estalles, C., ... & Campagna, L. (2021). Rapid speciation via the evolution of pre-mating isolation in the Iberá Seedeater. *Science*, *371*(6536).

Unbehend, M., Kozak, G. M., Koutroumpa, F., Coates, B. S., Dekker, T., Groot, A. T., ... & Dopman, E. B. (2021). bric à brac controls sex pheromone choice by male European corn borer moths. *Nature Communications*, *12*(1), 1-11.

Van Belleghem, S. M., Lewis, J. J., Rivera, E. S., & Papa, R. (2021). Heliconius butterflies: a window into the evolution and development of diversity. *Current Opinion in Genetics & Development*, *69*, 72-81.

van der Kooi, C. J., Stavenga, D. G., Arikawa, K., Belušič, G., & Kelber, A. (2021). Evolution of insect color vision: from spectral sensitivity to visual ecology. *Annual review of entomology*, *66*, 435-461.

Verzijden, M. N., Ten Cate, C., Servedio, M. R., Kozak, G. M., Boughman, J. W., & Svensson, E. I. (2012). The impact of learning on sexual selection and speciation. *Trends in ecology & evolution*, *27*(9), 511-519.

Wang, K., Wang, F., Forknall, N., Yang, T., Patrick, C., Parekh, R., & Dickson, B. J. (2021). Neural circuit mechanisms of sexual receptivity in Drosophila females. *Nature*, *589*(7843), 577-581.

Weber, J. N., Peterson, B. K., & Hoekstra, H. E. (2013). Discrete genetic modules are responsible for complex burrow evolution in Peromyscus mice. *Nature*, *493*(7432), 402-405.

Wilczynski, W., & Ryan, M. J. (2010). The behavioral neuroscience of anuran social signal processing. *Current opinion in neurobiology*, *20*(6), 754-763.
Wiley, C., Ellison, C. K., & Shaw, K. L. (2012). Widespread genetic linkage of mating signals and preferences in the Hawaiian cricket Laupala. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1731), 1203-1209.
Williams, T. M., & Carroll, S. B. (2009). Genetic and molecular insights into the development and evolution of sexual dimorphism. *Nature Reviews Genetics*, *10*(11), 797-804.
Wolf, J. B., & Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*, *18*(2), 87.
Wong, R. Y., Ramsey, M. E., & Cummings, M. E. (2012). Localizing brain regions associated with female mate preference behavior in a swordtail. *PloS one*, *7*(11), e50355.
Wu, M., Nern, A., Williamson, W. R., Morimoto, M. M., Reiser, M. B., Card, G. M., & Rubin, G. M. (2016). Visual projection neurons in the Drosophila lobula link feature detection to distinct behavioral programs. *Elife*, *5*, e21022.
Xu, M., & Shaw, K. L. (2019). Genetic coupling of signal and preference facilitates sexual isolation during rapid speciation. *Proceedings of the Royal Society B*, *286*(1913), 20191607.
Xu, M., & Shaw, K. L. (2021). Extensive Linkage and Genetic Coupling of Song and Preference Loci Underlying Rapid Speciation in Laupala Crickets. *Journal of Heredity*, *112*(2), 204-213.
Yamaguchi, M. (2012). Role of regucalcin in brain calcium signaling: involvement in aging. *Integrative Biology*, *4*(8), 825-837.
Yoshimatsu, T., Bartel, P., Schröder, C., Janiak, F. K., St-Pierre, F., Berens, P., & Baden, T. (2021). Ancestral circuits for vertebrate colour vision emerge at the first retinal synapse. *Science Advances, 7*(42).
Zaccardi, G., Kelber, A., Sison-Mangus, M. P., & Briscoe, A. D. (2006). Color discrimination in the red range with only one long-wavelength sensitive opsin. *Journal of Experimental Biology*, *209*(10), 1944-1955.
Zhang, D. D., & Löfstedt, C. (2015). Moth pheromone receptors: gene sequences, function, and evolution. *Frontiers in Ecology and Evolution*, *3*, 105.
Zhao, Z., & McBride, C. S. (2020). Evolution of olfactory circuits in insects. *Journal of Comparative Physiology A*, *206*(3), 353-367.

Acknowledgments

I am grateful to Richard Merrill for the opportunity to be part of this project and for his advice and guidance throughout my PhD journey. Much of this thesis work benefited from collaboration with Alex Hausmann: many fruitful ideas were born during our walks, and I look forward to many more. I thank Stephen Montgomery and Laura Hebberecht-Lopez for teaching me how to dissect butterfly brains, and Stephen also for our very enjoyable collaboration. Special thanks also to my advisory committee, Nicolas Gompel and Ricardo Pereira, for valuable discussions that stirred the project in productive ways. Many thanks to Owen McMillan for his input, for hosting me at STRI, and for the best barbecues in the whole Caribbean. I am grateful to Luca Livraghi and Joe Hanly for their advice on molecular techniques, to Morgan Oberweiser and Sophie Smith for assistance with insectary work, and to Jonah Walker for delivering the sequencing samples. Thanks to the Life Science Munich (LSM) graduate school for support. I thank Chi, Lucie, Shane, and the members of the Merrill lab, past and present, and the members of the Division of Evolutionary Biology for valuable discussions. Going further back, I am grateful to my MSc advisors, Helene Hinaux and Paulina Bolivar, for their guidance and advice, which set me on the path toward this PhD project. Finally, as a proper Italian, I am indebted to many in my famiglia. To my uncle Tommaso, who inspired me to get into science. To my sister Elisabetta and my brothers Giovanni and Francesco for brightening up my days. I dedicate this thesis to the memory of my mother Maria, and to my father Mauro who, as a classic physician still doubts whether mine was a good career choice, and yet continues to support me unconditionally, and to answer my phone calls at any time of day and night.