Genome-wide evidence for speciation with gene flow in Heliconius butterflies

Simon H Martin, Kanchon K Dasmahapatra, Nicola J Nadeau, et al.

Genome Res. published online September 17, 2013
Access the most recent version at doi:10.1101/gr.159426.113

Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies

Simon H. Martin[1*], Kanchon K. Dasmahapatra[2,3], Nicola J. Nadeau[1], Camilo Salazar[4], James R. Walters[5], Fraser Simpson[2], Mark Blaxter[6], Andrea Manica[1], James Mallet[2,7] & Chris D. Jiggins[1]

[1]Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK.

[2]Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, UK.

[3]Department of Biology, University of York, Wentworth Way, York YO10 5DD, UK

[4]Facultad de Ciencias Naturales y Matematicas, Universidad del Rosario, Carrera 24 No 63C-69, Bogota D.C, Colombia.

[5]Department of Biology, Stanford University, Stanford, CA, 94305, U.S.A.

[6]Institute of Evolutionary Biology, The University of Edinburgh, Edinburgh EH9 3JT, UK.

[7]Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA.

*Corresponding author: shm45@cam.ac.uk

Abstract

Most speciation events probably occur gradually, without complete and immediate reproductive isolation, but the full extent of gene flow between diverging species has rarely been characterized on a genome-wide scale. Documenting the extent and timing of admixture between diverging species can clarify the role of geographic isolation in speciation. Here we use new methodology to quantify admixture at different stages of divergence in *Heliconius* butterflies, based on whole genome sequences of 31 individuals. Comparisons between sympatric and allopatric populations of *H. melpomene*, *H. cydno and H. timareta* revealed a genome-wide trend of increased shared variation in sympatry, indicative of pervasive interspecific gene flow. Up to 40% of 100 kb genomic windows clustered by geography rather than by species, demonstrating that a very substantial fraction of the genome has been shared between sympatric species. Analyses of genetic variation shared over different time intervals suggested that admixture between these species has continued since early in speciation. Alleles shared between species during recent time intervals displayed higher levels of linkage disequilibrium than those shared over longer time intervals, suggesting that this admixture took place at multiple points during divergence and is probably ongoing. The signal of admixture was significantly reduced around loci controlling divergent wing patterns, as well as throughout the Z chromosome, consistent with strong selection for Müllerian mimicry and with known Z-linked hybrid incompatibility. Overall these results show that species divergence can occur in the face of persistent and genome-wide admixture over long periods of time.

Keywords

Hybridization, Admixture, Introgression, Sympatric Speciation,

Introduction

Ongoing hybridization between closely related species appears to be common in nature (Mallet 2005; Rieseberg 2009) and theoretical work has demonstrated a diversity of scenarios whereby species can emerge without complete geographical isolation (Kirkpatrick and Ravigné 2002; Gavrilets 2004; van Doorn et al. 2009). Despite widespread interest in these scenarios, there remains little consensus among speciation biologists regarding the extent to which ongoing gene flow actually plays a role during speciation. This is partly because it is challenging to reconstruct ancestral ranges and therefore almost impossible to know for sure the extent of historical contact between species. Fortunately, genomic approaches are now beginning to allow us to address these long-standing questions from a different angle, by documenting the extent of admixture between species on a genome-wide scale (Kulathinal et al. 2009; Dasmahapatra et al. 2012; Ellegren et al. 2012; Garrigan et al. 2012; Nosil et al. 2012). Speciation genomics therefore offers an opportunity to address long-standing questions regarding the extent to which divergence and speciation occurs in the face of ongoing gene flow.

One prediction of models of speciation with gene flow is that the level of divergence should be heterogeneous across the genome. Some loci are likely to be shared between incipient species, while selection maintains divergence at others (Turner et al. 2005; Nosil et al. 2009). Recently, considerable progress has been made in documenting patterns of genomic divergence between incipient species (Hohenlohe et al. 2010; Lawniczak et al. 2010; Michel et al. 2010; Ellegren et al. 2012; Nosil et al. 2012; Gagnaire et al. 2013). Genome-wide studies of threespine sticklebacks (Hohenlohe et al. 2010, 2012) and *Ficedula* flycatchers (Ellegren et al. 2012) revealed patterns of divergence consistent with a model of “islands” of divergence amidst a sea of gene flow. In contrast, analyses of *Anopheles gambiae* subspecies (Lawniczak et al. 2010) and *Rhagoletis* host

races (Michel et al. 2010), reported widespread divergence throughout the genome. One problem is that patterns of divergence are typically noisy, reflecting the complex interactions of selection, drift, migration, recombination, mutation and ancestral polymorphism, all of which can lead to heterogeneity in divergence (Noor and Bennett 2009; Michel et al. 2010; Nadeau et al. 2012). A key challenge is therefore to distinguish the signal of gene flow from background noise.

Analyses of genomic divergence therefore need to be complemented with more sensitive tests for gene flow between populations (Kulathinal et al. 2009; Garrigan et al. 2012; Ellegren et al. 2012; Dasmahapatra et al. 2012; Nosil et al. 2012). A widely used approach is to fit coalescent models (Pinho and Hey 2010), but this can be computationally prohibitive for genomic data sets and requires strong assumptions about population parameters. A simpler method is to compare the extent of shared variation between sympatric and allopatric populations (Grant et al. 2005). Recent gene flow should result in reduced differentiation and an excess of shared variation between sympatric populations compared to allopatric populations. This logic has been applied on a genomic scale to test for gene flow in *Drosophila* (Kulathinal et al. 2009) and hominids (Green et al. 2010). However, this approach does not account for the age of shared variation, such that recent admixture may be confounded with ancestral geographic structure (Green et al. 2010; Durand et al. 2011; Eriksson and Manica 2012). It is therefore best used in combination with other methods that can distinguish recent gene flow from ancient shared variation.

In this paper, we focus on the closely related neotropical butterfly species *Heliconius melpomene*, *Heliconius cydno* and *Heliconius timareta* (Fig. 1). These species are distasteful to predators and often involved in Müllerian mimicry with other species. All three comprise multiple distinct wing pattern races that have been considered as an early stage in speciation (Jiggins 2008). Indeed there is strong evidence that selection for Müllerian mimicry can lead to wing pattern divergence and

assortative mating without the need for geographic separation (Chamberlain et al. 2009). *Heliconius cydno* and *H. timareta* together form a clade that is sister to *H. melpomene*, estimated about two million years divergent (Bull et al. 2006; Salazar et al. 2008). *Heliconius melpomene* and *H. cydno* have distinct wing patterns and other ecological differences, and display strong assortative mating (Merrill et al. 2011a). Hybrids occur at low frequency (< 1/1000) (Mallet et al. 2007), and are female-sterile (Naisbit et al. 2002), as well as being preferentially attacked by predators due to their non-mimetic wing patterns (Merrill et al. 2012). Unlike *H. cydno*, several *H. timareta* races have *H. melpomene*-like patterns (Giraldo et al 2008, Mérot et al. 2013) and similarly show differences in host plant use and mating preferences (Giraldo et al 2008). Recent genomic studies have begun to dissect the genetic variation underlying colour pattern diversity in this genus (Nadeau et al. 2013, Supple et al. 2013, Dasmahapatra et al. 2012). One important insight is that the shared colour patterns between *H, melpomene* and *H. timareta* appear to be the have resulted from introgression (Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012). There is also evidence for exchange of other loci between *H. melpomene* and the *H. cydno*/*timareta* clade (Bull et al. 2006; Kronforst et al. 2006; Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012; Nadeau et al. 2013). RAD-tag analyses of Peruvian races of *H. melpomene* and *H. timareta* suggest that at least ~2-5% of the genome is admixed (Dasmahapatra et al. 2012). The recent completion of the *H. melpomene* genome now allows investigation of genome-wide patterns of divergence and gene flow within and between these species.

Here we take advantage of the geographic distribution of *H. melpomene*, with some populations many thousands of kilometers from the current range of the *H. cydno*/*timareta* clade (Fig. 1), and carry out a much more powerful genome-wide test for gene flow than was possible with the sequenced fragments hitherto studied. We analyzed 31 resequenced individuals (30 of which were newly sequenced in this study) from replicate sympatric species pairs of the two clades in Peru,

where they are convergent in wing pattern, and Panama, where they are divergent. We also sampled an allopatric *H. melpomene* population from French Guiana (Fig. 1). Four species of the silvaniform clade of *Heliconius* were included as outgroups. Our new methods allowed us to investigate the extent and time course of genomic admixture, both before speciation and during different time periods after speciation.

Results

Phylogenomic analysis

Five populations of *H. melpomene,* one population of *H. cydno* and one population of *H. timareta*, and four outgroup species were sequenced (Figure 1 and Table S1). Populations were represented by four wild-caught individuals (eight haploid genomes) except *H. m. melpomene* from Panama, for which three individuals were sampled (Table S1). All individuals were wild-caught except for *H. m. melpomene* specimen no. 1, which was from the inbred reference genome strain. Whole genome shotgun sequencing using the Illumina GAIIx and HiSeq2000 technology, gave an average coverage per individual of 15-62x (Table S1, SRA study accession ERP002440). Sequences were aligned to the *H. melpomene* reference genome (Dasmahapatra et al. 2012) (Version 1.1), including the complete mitochondrial scaffold. Genotyping and quality filtering (see Methods for details) produced an average of 190 million high quality genotype calls per individual (69% of the genome). Proportions of variant sites were similar across all wild-caught individuals, and the ratio of transitions to transversions with respect to the reference were similar across all taxa, indicating that there was no systematic bias in the distribution of sequencing errors among taxa.

Maximum likelihood (ML) phylogenetic reconstruction using all sites with high-quality genotype calls for all 31 individuals (60 Mb of sequence, about 25% of the genome) confirmed that the *H.*

melpomene and the *H. cydno/timareta* clades are reciprocally monophyletic (Dasmahapatra et al. 2012; Nadeau et al. 2013) (Fig. 1, S1). We here term this topology "the species tree". A tree generated from complete mitochondrial sequences produced a similar topology (Fig. S1).

Phylogenetic discordance across the genome

Although the genome-wide ML tree revealed strong support for the expected "species tree", most speciation scenarios predict discordant coalescent histories among genomic regions (Garrigan et al. 2012). To investigate this, we generated maximum-likelihood trees for non-overlapping 100 kb windows throughout the genome. To simplify the hypotheses being tested, we analyzed two sets of four taxa separately, each representing a sympatric species pair and an allopatric 'control' population. The *cydno*/*melpomene* set consisted of *H. cydno* and *H. m. rosina* (both from Panama), *H. m. melpomene* from French Guiana and outgroups, while the *timareta*/*melpomene* set consisted of *H. timareta* and *H. m. amaryllis* (both from Peru), with *H. m. melpomene* from French Guiana and outgroups. We summed the frequency of four possible topologies: species, geography, control and unresolved (Fig. 2B). Three of these we considered "resolved", meaning that two of the ingroup populations (eight individuals) formed a monophyletic clade, while the four individuals comprising the third ingroup population formed a distinct monophyletic sister clade (see Fig S3 for examples). For both datasets, the majority of genomic windows (53% and 53.2%, respectively) supported a resolved 'species tree' topology in which the *H. melpomene* populations are monophyletic (Fig. 2). Under a bifurcating topology, incomplete lineage sorting should result in similar numbers of two alternative resolved topologies, which we term the 'geography tree' (sympatric populations of different species cluster together), and the 'control tree' (allopatric populations of different species cluster together). The final possibility is an 'unresolved tree', in which the three ingroup populations are not neatly partitioned into two monophyletic clusters (Fig. 2, see Fig S3 for examples).

As expected under admixture, we found that the geography tree was far more prevalent than the control tree in both datasets: 42.2% versus 1.1% for the *cydno*/*melpomene* set and 18.5% versus 2.7% for the *timareta*/*melpomene* set; and widely distributed across the genome (Fig. 2C). While only 3.7% of trees were unresolved in the *cydno*/*melpomene* set, 25.6% were unresolved in the *timareta*/*melpomene* set. The greater fraction of unresolved trees in the second case is expected given greater shared ancestral polymorphism due to the more recent divergence between *H. m. amaryllis* and *H. m. melpomene* from French Guiana (Fig. 1, S1). These findings show that there is not only a large amount of phylogenetic discordance across the genome, but that it is strongly structured by geography, consistent with gene flow between these clades where their ranges overlap. Here we report results for 100 kb windows because linkage disequilibrium (LD) tends to break down completely within 100 kb in *Heliconius* genomes (Fig. S2), making each 100 kb block effectively independent from its neighbours. However, we also repeated the tests at various window sizes between 10 kb and 200 kb (Table S2). Although the number of unresolved trees increases at smaller window sizes, the relative ratios of resolved trees are robust to window size variation (Table S2).

Evidence of recent gene flow

Allele frequency correlations provide further evidence for recent interspecific gene flow. Our geographically structured sampling design allowed us to distinguish between ancient and recent admixture using a sensitive ‘four population’ test (Reich et al. 2009, 2012) for geographical correlations in allele frequencies. In the absence of admixture, allele frequency changes due to drift in disparate populations should not be correlated. Across all tests, there was a highly significant allele frequency correlation between *H. m. rosina* and *H. cydno* from Panama, and between *H. m. amaryllis* and *H. timareta* from Peru (Table 1). These correlations indicate recent gene flow

between these species where they occur in sympatry.

Gene flow has occurred at multiple points since early in speciation

Evidence for recent gene flow does not necessarily imply that gene flow has persisted throughout speciation. Secondary contact after allopatric speciation might be characterized by a burst of recent gene flow, while sympatric speciation should leave a signature of continuous gene flow during speciation. We estimated admixture along different branches of the phylogeny using a method devised by Green et al. (Green et al. 2010), which compares two classes of shared derived alleles, termed ABBAs and BABAs. For three populations and an outgroup, with the relationship $(((P_1,P_2,),P_3),O)$, we can test for differential admixture between P3 and either of P1 and P2 by examining the numbers of shared derived alleles between P3 and P2 (ABBAs) and between P3 and P1 (BABAs). We calculated two statistics: '*D*' used to test for a significant imbalance of ABBAs and BABAs, indicative of admixture; and '*f*', the estimated fraction of the genome that has been shared between populations (Green et al. 2010; Durand et al. 2011). These measures are robust to variation in effective population size (Durand et al. 2011).

We examined rates of gene flow between *H. timareta* and *H. m. amaryllis* across three time periods (Fig. 3): a short, recent period, subsequent to the divergence between *H. m. amaryllis* and *H. m. aglaope* (period 4 of Fig. 3A); an intermediate period, subsequent to the divergence of the Peruvian populations from French Guianan *H. m. melpomene* (periods 3 & 4 of Fig. 3A); and a long period, subsequent to the divergence between Peruvian and Panamanian populations (periods 2, 3 & 4 of Fig. 3A). Across these comparisons there was a strong trend of increasing *f* with time (Fig. 3C, Table 2). Similarly, for *H. m. rosina* and *H. cydno*, over two time periods, *f* was again much greater for the longer period (Fig. 3C, Table 2). Because this method assumes unidirectional gene flow from P3, and complete isolation between P3 and P1, the actual fraction of the genome that has been

shared may be greater than estimated here. What is important is that the relative values of *f* increase with the length of the period examined, which is consistent with gene flow having occurred during time periods 2, 3 and 4. One potential caveat is that the extent of isolation between P3 and P1 could differ between these tests, accounting for some of the variation in *f*. We therefore investigated linkage disequilibrium among these shared derived sites as an additional signal to differentiate between recent and long-term gene flow.

Linkage disequilibrium between shared derived alleles

The extent of LD between introgressed alleles carries information about the age of admixture. Recently introgressed haplotypes have had insufficient time to be broken down by recombination, and therefore closely linked introgressed alleles should occur in LD with one another (Machado et al. 2002, Sankararaman et al. 2012). By contrast, anciently introgressed alleles should display levels of LD similar to the average genomic level. We tested for this signal by examining LD at sites carrying shared derived variants (i.e. ABBA SNPs) in *H. m. amaryllis* and *H. m. rosina*. For *H. m. amaryllis*, three sets of SNPs carrying shared variants could be examined, corresponding to the three time periods described above (Fig. 3B). Likewise, in *H. m. rosina*, two sets of SNPs could be examined.

We found that the extent of LD differed dramatically between the time periods (Fig. 3D). Variants shared between *H. timareta* and *H. m. amaryllis* but absent from *H. m. aglaope* displayed the strongest LD, extending up to a megabase. This is consistent with the existence of large introgressed haplotypes that have yet to be fully broken down. Variants shared between *H. timareta* and *H. m. amaryllis* but absent from French Guyanan *H. m. melpomene* displayed weaker LD, while those shared between *H. timareta* and *H. m. amaryllis* but absent from *H. m. rosina* displayed the weakest LD, declining with distance at a similar rate to the genomic average (Fig. 3D). Thus, these two

latter comparisons include variation that appears to have been shared more anciently, giving sufficient time for introgressed haplotypes to be broken down. Similar differences were observed in the extent of LD among variants shared between *H. cydno* and *H. m. rosina* at the two time intervals examined. By exploiting a different aspect of the data, these results provide an independent line of evidence that both recent and ancient admixture has occurred between these species pairs.

Patterns of genomic divergence along the speciation continuum

We characterized patterns of divergence across the genome between populations at various levels of divergence and geographic separation using the fixation index, F_{ST}. At the earliest stage of divergence, between parapatric races, F_{ST} was low throughout the genome with just a few narrow peaks (Fig. 4), which are partly explained by known wing pattern divergence. Between *H. m. aglaope* and *H. m. amaryllis* from Peru, only two pronounced divergence peaks were present, corresponding to the known pattern loci *HmB* (red elements) and *HmYb* (yellow elements) (Baxter et al. 2010; Nadeau et al. 2012). For the Panamanian races, the level of F_{ST} was noisier but there was a small F_{ST} peak at the *HmYb* locus (Fig. 4). There was no peak at the *HmB* locus, consistent with the fact that the Panamanian races share the same red mimetic patterns. Between allopatric races, background F_{ST} was significantly higher and more heterogeneous, and colour pattern loci no longer appeared as clear outliers. Patterns of F_{ST} between species were broadly similar in mean and variance to those between allopatric races of *H. melpomene* (Fig. 4 and 5).

Reduced interspecific divergence in sympatry

Gene flow between sympatric populations should lead to reduced F_{ST} as compared to that between allopatric populations. Consistent with phylogenetic evidence for gene flow in sympatry, F_{ST} between sympatric species pairs in both Panama and Peru was significantly lower than that between either *H. timareta* or *H. cydno* and the allopatric *H. m. melpomene* from French Guiana (Table 3).

Each of the 21 chromosomes independently showed the same trend of significantly lower F_{ST} in sympatry than in allopatry (Fig. S5B). This trend is also robust to the use of different allopatric populations. Peruvian *H. m. amaryllis* can be considered as allopatric to Panamanian *H. cydno* (separated by the Andes). Likewise, *H. m. rosina* can be considered allopatric to *H. timareta*. These allopatric comparisons both display significantly higher average F_{ST} than the sympatric comparisons, although not quite as high as when the French Guianan population is used (Table S3). This variation may be partly due to differences in the extent of isolation, but probably also reflect differences in effective population size. Nevertheless no inter-species allopatric comparison shows anything close to the reduced F_{ST} observed between the species in sympatry (Fig. 5).

Plotted across individual chromosomes, the pattern of F_{ST} was highly heterogeneous in both sympatry and allopatry (Fig. S5C, S6, S7). As admixture between species is expected to be non-uniformly distributed across the genome, we predicted that there would be greater heterogeneity in F_{ST} in sympatry. Indeed the coefficient of variation was significantly greater for F_{ST} between sympatric pairs than allopatric pairs (Table 3).

Comparison of F_{ST} in sympatry relative to that in allopatry may be useful in identifying regions subject to divergent selection and hence reduced gene flow. When plotted across individual chromosomes, the trend of lower F_{ST} in sympatry was widespread but punctuated by narrow regions at which F_{ST} between the sympatric populations approached and occasionally exceeded that between allopatric populations (Fig. S5C, S6, S7). Assuming that the allopatric population pair provides a reference for expected F_{ST} in the absence of gene flow, these regions indicate putative loci at which selection has acted to eliminate introgressed alleles. This hypothesis is supported by the wing pattern loci. *H. cydno* and *H. m. rosina* have divergent colour patterns, and F_{ST} at both the *HmB* and *HmYb* patterning loci was similar in sympatry and allopatry (Fig. S5C). By contrast,

between *H. timareta* and *H. m. amaryllis*, which have convergent wing patterns due to recent introgression (Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012), there were narrow regions of reduced F_{ST} between the sympatric pair at both patterning loci.

Enhanced reproductive isolation of the Z chromosome

Multiple lines of evidence suggest that gene flow has been reduced throughout the Z chromosome compared with the rest of the genome. Estimates of the fraction of introgression for the Z chromosome were strikingly reduced compared to genome-wide estimates (Fig. 3C, Table 2). In fact, there is no evidence for significant recent Z chromosomal admixture between *H. timareta* and *H. m. amaryllis* (Table 1, 2).

Patterns of genomic differentiation were also consistent with reduced gene flow across the Z chromosome. For most population pairs, the Z chromosome had a significantly elevated level of F_{ST} compared with autosomes (Table S3, Fig. S5B and C). Higher F_{ST} on this chromosome compared to autosomes is expected given its lower effective population size. However, the ratio of sympatric/allopatric F_{ST} was closer to one on the Z chromosome (~0.9) than on autosomes (~0.65) (Table 3). This further supports the hypothesis that admixture on the Z is significantly reduced compared to that on autosomes.

Discussion

The dominant paradigm among evolutionary biologists has recently shifted from widespread belief in the virtually universal importance of allopatric speciation, towards increasing acceptance that speciation may occur in the presence of some gene flow. However, despite plenty of evidence for hybridization and gene flow between good species, we remain largely ignorant of the extent to

which speciation involves ongoing gene flow, both across the genome and through time. If gene flow is indeed common and persistent, then theoretical models of sympatric speciation might be very widely applicable, justifying the recent shift in emphasis (Wu 2001; Pinho and Hey 2010; Smadja and Butlin 2011; Feder et al. 2012). Using whole-genome resequencing combined with structured geographic sampling we now have much greater power to answer these questions. Our data indicate strong signals of admixture between species across a surprisingly large fraction of the genome. This has occurred either continuously or during multiple periods since their initial divergence. Taken together, our results indicate that species divergence can occur in the face of persistent and genome-wide admixture over long periods of time.

Quantifying gene flow through time

It has long been recognized that both incomplete lineage sorting and hybridization can lead to discordant genealogical histories across the genome. By using an allopatric population of *Heliconius melpomene* from French Guiana for comparison, we provide evidence that 20-40% of the genome in *H. melpomene* shows admixture with *H. cydno or H. timareta* in sympatry. The window-based phylogenetic approach, using 100 kb regions, averages over large numbers of sites but ensures that each 100 kb tree is statistically well-supported. Furthermore, this result was highly robust to variation in window size.

We extended the site-based ABBA/BABA method to quantify gene flow through time. A previous analysis of *H. melpomene* and *H. timareta* indicated that ~2-5% of the genome was influenced by gene flow (Dasmahapatra et al. 2012), but this comparison could detect admixture that occurred only over a short, recent time period. Our sampling design here allowed us to vary the choice of ingroup populations and examine gene flow over different time scales. Estimates of admixture increased with increasing length of the time period examined, implying continued gene flow during

speciation as opposed to a recent burst. Furthermore, LD between derived alleles that were shared during the recent time period was strongest, indicating the existence of introgressed haplotype blocks that are yet to be broken down fully by recombination. This signal was most extensive for alleles shared between *H. timareta* and *H. m. amaryllis* but absent from *H. m. aglaope*, with LD extended up to 1 Mb. This is consistent with extremely recent gene flow, as *H. m. amaryllis* and *H. m. aglaope* coalesce very recently. By contrast, LD between variants shared over longer time periods was weaker, and declined with physical distance at a rate similar to the genome-wide average, implying that most of these admixed variants were shared very long ago. Thus two independent lines of evidence suggest that gene flow extends from early in speciation to the present. While we cannot rule out periods of allopatry during this time, particularly very early during the species divergence, our results imply that admixture has been a major influence on the genome throughout most of the speciation process.

Genomic divergence through time and space

There has been mixed support for the verbal model of islands of divergence amidst a sea of gene flow (Nosil et al. 2009; Noor and Bennett 2009; Feder et al. 2012). Here we examined this model by comparing patterns of genomic divergence at different stages of speciation and different levels of geographical separation. Parapatric races that are known to hybridize in nature, and in particular *H. m. amaryllis* and *H. m. aglaope* from Peru, displayed patterns of differentiation strongly congruent with this islands of divergence model, with strong differentiation at known wing patterning loci. Nonetheless, patterns of divergence are likely to be heterogeneous regardless of gene flow (Noor and Bennett 2009; Michel et al. 2010). For example, between allopatric populations of *H. melpomene*, subject to isolation by distance and biogeographic barriers such as the Andes, there is a higher average F_{ST} but also considerable heterogeneity across the genome, including divergence peaks at the colour pattern loci. This probably reflects the fact that strong selection, and various

other demographic factors, can cause certain regions to become outliers for population differentiation, even in the absence of homogenizing gene flow at other loci. The presence of 'islands of divergence' alone does not provide sufficient evidence for that divergence occurred with ongoing gene flow.

F_{ST} between sympatric species was highly heterogeneous, and was not congruent with an idealized scenario of islands of divergence against an otherwise homogenized genome. Nevertheless, interspecific F_{ST} between sympatric species was generally lower, and more variable (Table 3) than between the corresponding allopatric populations, as expected under a model of admixture with variable selection against introgressing alleles. The trend of lower F_{ST} in sympatry was widespread across all chromosomes, consistent with pervasive admixture across the whole genome. Despite this widespread signal, the rate of effective gene flow between *H. melpomene* and the *H. cydno/timareta* clade is apparently insufficient to completely abolish differentiation across most of the genome (Nosil et al. 2009; Feder et al. 2012).

Comparisons of sympatric and allopatric populations also permit detection of outlier loci using the joint distribution of F_{ST} in sympatry and allopatry. Loci at which interspecific F_{ST} is similar in sympatry and allopatry could indicate putative targets of divergent selection where the effective rate of gene flow is reduced. In effect, the allopatric population provides a reference for the expected divergence value in the absence of gene flow, controlling for the inherent heterogeneity in rates of divergence across the genome. This is conceptually similar to an approach applied in hybrid zones, where allopatric populations are used as a control to detect introgressed loci (Gompert and Buerkle 2010). Loci known to be under selection offer a test of this logic. *H. cydno* and *H. melpomene* from Panama have distinct wing patterns and both the *HmB* and *HmYb* pattern loci fall under peaks at which F_{ST} is similar in sympatry and allopatry. The Peruvian pair has convergent wing patterns and

narrow tracts of the genome have here introgressed at both colour pattern loci (Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012). Indeed, at both loci, there is a narrow trough of low F_{ST} between these populations. The relatively high levels of F_{ST} surrounding these troughs may be remnants of hitchhiking following initial divergence in wing pattern. Although we are here mostly interested in the genome-wide patterns of divergence and admixture, we believe that in the future such joint distributions of F_{ST} are likely to provide a powerful method for detection of genomic regions subject to selection.

The Z chromosome is at a more advanced stage of speciation

There is both theoretical and empirical evidence for a disproportionate role of the sex chromosomes in speciation (Qvarnström and Bailey 2009). Sex-linked genes are expected to diverge more rapidly (Coyne and Orr 2004), and in the Lepidoptera species differences have been shown to map disproportionately to the Z chromosome (Prowell 1998). In our data there was a significantly reduced signal of admixture on the Z chromosome compared to autosomes. The discrepancy between the Z and autosomal F_{ST} was also considerably greater in sympatry than in allopatry (Table 3). Thus the difference cannot be explained solely by reduced effective population size of sex chromosomes. Numbers of shared derived alleles suggest that ancient gene flow did occur on the Z, but that the contemporary migration rate for this chromosome is very low. This can be explained, in part, by Z-autosome incompatibilities known to cause female hybrid sterility (Naisbit et al. 2002; Jiggins et al. 2001). These sex chromosome vs. autosome patterns are similar to those seen in the genomes of the *Drosophila simulans* group and in *Ficedula* flycatchers (Garrigan et al. 2012; Ellegren et al. 2012), providing general support for the hypothesis that sex chromosomes play a major role in speciation.

Conclusions

Genomic methods offer the opportunity to address the ongoing debate between recent proponents of sympatric speciation and the classical wisdom of ubiquitous allopatric speciation. It is unlikely that genomic data from extant species will ever rule out brief periods of allopatry during speciation. Nonetheless, it is clear from our results that admixture between *H. melpomene* and the *H. cydno/timareta* lineage has taken place on a large scale throughout much of their divergence history. To some extent, our findings fit with verbal ideas of speciation with gene flow (Wu 2001; Feder et al. 2012), in which a progression from narrow islands leads to more genomically widespread divergence. Indeed, despite increasing genome-wide divergence later on, the effects of gene flow remain pervasive throughout the genome. Up to 40% of the genome shows a discordant phylogenetic pattern consistent with admixture in sympatry. Our results imply that the recent focus on mechanisms that permit speciation-with-gene-flow in the literature is not misguided (Smadja and Butlin 2011; Servedio et al. 2011). In the case of *H. melpomene* and *H. cydno*, wing patterns have a relatively simple genetic basis (Naisbit et al. 2007), and the loci that affect male mate preference and hybrid sterility are associated with colour pattern loci (Merrill et al. 2011b), both of which should make speciation easier. Genomics therefore has provided empirical data that helps answer thorny questions about the relative importance of allopatric isolation in speciation, which have hitherto proved to be among the most intractable debates in evolutionary biology.

Methods

Whole-genome resequencing and genotype calling

Samples were preserved in NaCl-saturated DMSO solution at -20°C and DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen). Illumina paired-end libraries were generated according to the manufacturer's protocol (Illumina Inc.). These were shotgun sequenced on either Illumina's Genome Analyzer IIx system or Illumina's HiSeq 2000 system, according to the manufacturer's

protocol (Illumina Inc.).

Quality-filtered, paired-end sequence reads were mapped to the *H. melpomene* genome scaffolds (version 1.1) (Dasmahapatra et al. 2012) using Stampy v1.0.13 (Lunter and Goodson 2011). Defaults were used for all parameters with the exception of the expected substitution rate, which was set to 0.03 for *H. melpomene* samples (0.001 for the individual from the reference genome strain), 0.04 for *H. cydno*/*timareta* samples and 0.05 for outgroup silvaniform samples to allow mapping of reads from divergent species. To minimize false SNPs due to inconsistent mapping around indels, base alignment quality (BAQ) was considered during mapping, and then local realignment around indels was performed using the Genome Analysis Tool Kit (GATK) v1.6 (DePristo et al. 2011). SAM/BAM file conversion, analysis and filtering were performed using SAMtools (Li et al. 2009) and Picard (http://picard.sourceforge.net). PCR-duplicate reads were removed using Picard.

Genotypes were called using the GATK v1.6 UnifiedGenotyper (DePristo et al. 2011). Individuals from the same population were genotyped simultaneously. Default parameters were used, except expected heterozygosity was set to 0.01, and BAQ calculation was performed where necessary to optimize calls around indels. For a genotype call to be considered high quality, it had to meet the following criteria: Quality (QUAL) ≥ 30, 10 ≤ depth ≤ 200 (the upper bound was imposed to avoid false SNPs due to mis-mapping in repetitive regions), and for variant (non-reference) calls, genotype quality (GQ) ≥ 30. Only these “high quality” genotype calls were used in downstream analyses. Genotyping summary statistics for each sample are provided in Table S1.

Assigning scaffolds to chromosomes

Several analyses involved comparisons among chromosomes. Scaffolds were assigned to

chromosomes based on the *Heliconius melpomene* linkage map (Dasmahapatra et al. 2012), version 1.1, which has ~80% of the genome assigned to chromosomes. An important focus of this study was the comparison between autosomal and Z-linked regions. We therefore performed extra tests to confirm Z-linkage of mapped scaffolds and identify additional Z-linked scaffolds among those previously unmapped (see Appendix A for details). This procedure also identified several miss-assembled scaffolds that were Z/autosome chimeras. Using the most likely breakpoints identified, we removed Z-linked regions from autosomes and also removed autosomal regions from the Z-linked scaffolds.

Phylogenomic analysis

A whole-genome maximum-likelihood tree was generated using only sites in the genome with high-quality genotype calls for all 31 individuals, resulting in an alignment of 60 Mb. RAxML (Stamatakis 2006, Ott et al. 2007, Stamatakis et al. 2008) was used with the GTRGAMMA model, and 100 bootstrap replicates were performed. A separate tree was constructed for the mitochondrial genome (alignment of 9.5 kb), using the same procedure, but with 1000 bootstraps. To investigate phylogenetic discordance across the genome, independent maximum-likelihood trees were generated for non-overlapping 100 kb windows. To minimize artifacts of data quality, only sites with a high-quality genotype call for all 31 genomes were used, and windows that contained fewer than 10000 sites were rejected.

Four-population tests for admixture

To test for admixture between pairs of heterospecific populations, we used the four-population test (Reich et al. 2009, 2012). This test is based on the fact that genetic drift should be uncorrelated in unadmixed populations. Given the populations A, B, C and D, with the unrooted relationship ((A,B),(C,D)), the f_4 statistic, f_4(A,B;C,D), allows a test for whether allele frequency differences

between A and B are correlated with differences between C and D, thus indicative of admixture (either between A and C, or between B and D, or both). We calculated the f_4 statistic (Equation S6.1 of Reich et al. (2012) using all informative sites (i.e. biallelic sites at which both pairs of populations differ in allele frequency). The mean and variance in f_4 was then estimated using a block jack-knifing approach (Reich et al. 2009), which controls for LD among sites. We used a block size of 1 Mb, far greater than the extent of LD in the *Heliconius* genomes studied here (Dasmahapatra et al. 2012, Fig S2). This allowed us to test whether f_4 deviated significantly from zero. Such deviations would indicate that the allele frequency differences between the two population pairs are significantly correlated, indicating gene flow.

Quantifying gene flow over specific time periods

To quantify gene flow along a specific branch of the phylogeny, we used a method based on the relative abundance of two classes of polymorphic sites called "ABBAs" and "BABAs" (Green et al. 2010; Durand et al. 2011). Given four populations, P_1, P_2, P_3 and an outgroup O, with the relationship (((P1,P2),P3),O), ABBAs are SNPs at which P_2 and P_3 share a derived allele "B", while P_1 retains the ancestral allele "A", as inferred from the outgroup (i.e. $P_2 = P_3 \neq P_1 = O$). Similarly, BABAs are SNPs at which P_1 and P_3 share a derived allele "B", while P_2 retains the ancestral allele "A" (i.e. $P_1 = P_3 \neq P_2 = O$). Under the null hypothesis of no gene flow, ABBA and BABA patterns can only arise via incomplete lineage sorting, and should be equally infrequent (assuming no recurrent mutation, and random mating in the ancestral population). However, if there has been gene flow between P_3 and P_2 since the split between P_1 and P_2, there should be an over-representation of ABBA patterns. The relative abundance of ABBA and BABA patterns throughout the genome was compared using the *D* statistic (Equation 2 of (Durand et al. 2011)), based on allele frequencies at each SNP. Only sites at which the four outgroup genomes were homozygous for the same allele were considered to ensure confident assignment of the ancestral

and derived states. We used a 1 Mb block jack-knifing approach to calculate the mean and variance of *D*, allowing a test for whether *D* differed significantly from zero.

We then estimated *f*, the fraction of the genome that is admixed. In the example described above, the fraction of the genome that is admixed between P_3 and P_2 subsequent to the split between P_1 and P_2 can be estimated by comparing the observed difference in abundance of ABBA and BABA patterns to that which would be expected under a scenario of 100% admixture between P_3 and P_2 (Equation 8 of (Durand et al. 2011)). As above, we used a 1 Mb block jack-knife approach to calculate the mean and variance of the *f* value.

Estimating the extent of linkage disequilibrium (LD)

Linkage disequilibrium (LD) was estimated using all pairs of biallelic sites with high-quality genotype calls in all 31 genomes and a minor allele count of at least five. We estimated r^2 within *H. melpomene* populations using the maximum likelihood estimator (Clayton and Leung 2007), implemented in the R package "snpstats", which does not require phased haplotypes. To investigate how LD breaks down with distance, r^2 values were binned according to distance in logarithmically increasing bin sizes, to account for small numbers of SNP pairs at large distances. Only SNP pairs on the same scaffold were considered. To obtain an estimate of background LD between unlinked sites, subsets of 500 SNPs were randomly selected and r^2 was estimated for all pairs for which the two SNPs were on separate chromosomes. This procedure was repeated 100 times and a 95% confidence interval was calculated.

We investigated the rate of decline in LD between shared derived alleles in *H. m. amaryllis* and *H. m. rosina*. Following the definition of an ABBA site above, all sites at which P_1 was fixed for the ancestral state while P_2 and P_3 carried a derived allele, were considered. r^2 values were binned

according to distance as described above.

Patterns of genetic differentiation between populations

We estimated levels of genetic differentiation between populations by calculating F_{ST} for 100 kb genomic windows. Nadeau et al. (2012) showed that averaging over large numbers of sites in this way provides highly repeatable F_{ST} estimates from small samples. F_{ST} was calculated using the EggLib Python module (De Mita and Siol 2012). To minimize variation due to stochasticity and genotyping errors, windows were rejected if they contained fewer than 2500 variant sites genotyped with high quality for all individuals from the two populations being analyzed. Windows were restricted to single scaffolds (i.e. they did not cross scaffold boundaries). To plot F_{ST} across chromosomes, scaffolds were arranged according to the *Heliconius melpomene* linkage map (Dasmahapatra et al. 2012), version 1.1, having corrected for the several Z/autosome chimeric scaffolds identified as described in Appendix A.

Data Access

Paired-end fastq files are available from the European Nucleotide Archive (http://www.ebi.ac.uk/ena/, study accession: ERP002440). The following files have been deposited in the Data Dryad repository (http://datadryad.org/, DOI: XXX): All processed VCF files, site-based allele frequency data used for the four-population tests and ABBA BABA analyses, all pairwise F_{ST} values for 100 kb windows , maximum likelihood trees for each 100 kb window for the two four-taxon datasets analysed (Newick format), data files providing the topology supported by each window, a list of scaffolds and scaffold regions designated as Z-linked and custom Python and R scripts used for data analyses.

565 **Acknowledgements**

566

567 We are grateful to Richard Merrill and Simon Baxter for field work. Judith Mank and Jamie Walters
568 kindly contributed funds from their Fell Fund grant. We also thank Anders Eriksson for his input in
569 discussions of this study. Lastly, we are most grateful to the three anonymous referees, whose
570 thoughtful and in-depth reviews greatly improved the standard of this work. This study was funded
571 by the BBSRC (G006903/1, G008841/1 and H01439X/1), the Leverhulme Trust (F/09364/E) and
572 The Fell Fund.

573

574

575

Tables

Table 1. Results of the four population test for recent gene flow

test[1]	f_4 +- std err	Z score	p value
Whole Genome			
f_4(*cydno,timareta*; *rosina,melp.* [FG])	0.0764 +- 0.0013	60.14***	<0.0001
f_4(*cydno,timareta*; *amaryllis,melp.* [FG])	-0.0370 +- 0.0016	-22.70***	<0.0001
f_4(*cydno,timareta*; *rosina,melp.* [Pan])	0.0056 +- 0.0005	10.95***	<0.0001
f_4(*cydno,timareta*; *amaryllis,aglaope*)	-0.0039 +- 0.0006	-6.91***	<0.0001
f_4(*cydno,timareta*; *rosina,amaryllis*)	0.0883 +- 0.0015	58.51***	<0.0001
Z chromosome			
f_4(*cydno,timareta*; *rosina,melp.* [FG])	0.0256 +- 0.0064	4.00**	0.0001
f_4(*cydno,timareta*; *amaryllis,melp.* [FG])	-0.0235 +- 0.0100	-2.34*	0.0192
f_4(*cydno,timareta*; *rosina,melp.* [Pan])	0.0021 +- 0.0027	0.77	0.4418
f_4(*cydno,timareta*; *amaryllis,aglaope*)	-0.0108 +- 0.0104	-1.04	0.2997
f_4(*cydno,timareta*; *rosina,amaryllis*)	0.0394 +- 0.0085	4.64***	<0.0001

For f_4(A,B; C,D), a significantly positive Z score implies gene flow between A and C, or B and D, or both. A significantly negative Z score implies gene flow between A and D, or B and C, or both.

* Indicates f_4 significantly different from 0, p < 0.05, ** p < 0.01, *** p < 0.0001.

"aglaope": H. m. aglaope, "amaryllis": H. m. amaryllis, "rosina": H. m. rosina, "melp.": H. m. melpomene, "cydno": H. c. chioneus, "timareta": H. t. thelxinoe, "Pan": Panama, "FG": French Guiana

Table 2. Results of ABBA BABA tests to quantify gene flow over specific time periods

P_1	P_2	P_3	time period[1]	D[2]	Z[3]	p value [3]	*f (%)* [4]
Whole Genome							
aglaope	*amaryllis*	*timareta*	4	0.039 +- 0.006	6.58***	<0.0001	2.1 +- 0.3
melp. [FG]	*amaryllis*	*timareta*	4 + 3	0.197 +- 0.009	22.60***	<0.0001	11.2 +- 0.7
rosina	*amaryllis*	*timareta*	4 + 3 + 2	0.209 +- 0.013	16.09***	<0.0001	14.6 +- 1.2
melp. [Pan]	*amaryllis*	*timareta*	4 + 3 + 2	0.229 +- 0.013	17.38***	<0.0001	15.6 +- 1.2
melp. [Pan]	*rosina*	*cydno*	4	0.073 +- 0.005	14.71***	<0.0001	4.4 +- 0.3
melp. [FG]	*rosina*	*cydno*	4 + 3 + 2	0.493 +- 0.009	53.08***	<0.0001	27.6 +- 0.8
amaryllis	*rosina*	*cydno*	4 + 3 + 2	0.490 +- 0.009	56.83***	<0.0001	29.3 +- 0.8
aglaope	*rosina*	*cydno*	4 + 3 + 2	0.501 +- 0.009	55.49***	<0.0001	29.7 +- 0.9
Z chromosome							
aglaope	*amaryllis*	*timareta*	4	0.092 +- 0.098	0.94	0.3460	1.1 +- 1.1
melp. [FG]	*amaryllis*	*timareta*	4 + 3	0.204 +- 0.074	2.77**	0.0056	2.5 +- 1.1
rosina	*amaryllis*	*timareta*	4 + 3 + 2	0.140 +- 0.057	2.47*	0.0136	2.2 +- 1.1
melp. [Pan]	*amaryllis*	*timareta*	4 + 3 + 2	0.155 +- 0.059	2.64**	0.0082	2.4 +- 1.1
melp. [Pan]	*rosina*	*cydno*	4	0.050 +- 0.012	4.27***	<0.0001	0.7 +- 0.2
melp. [FG]	*rosina*	*cydno*	4 + 3 + 2	0.191 +- 0.025	7.52***	<0.0001	4.0 +- 0.8
amaryllis	*rosina*	*cydno*	4 + 3 + 2	0.103 +- 0.008	12.46***	<0.0001	2.3 +- 0.4
aglaope	*rosina*	*cydno*	4 + 3 + 2	0.120 +- 0.030	3.94**	0.0001	2.7 +- 0.9

P_1, P_2 and P_3 refer to the three populations used for the ABBA BABA tests (see Methods for details).

[1] Period over which gene flow is measured, as shown in Fig. 3A.

[2] D statistic, to test for an over-representation of ABBA vs. BABA patterns, +- standard error.

[3] Z score and P value for the block-jackknife test of whether D differs significantly from zero.

[4] Estimated fraction of introgression, given as a percentage +- standard error.

* Indicates D significantly different from 0, $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$.

"*aglaope*": *H. m. aglaope*, "*amaryllis*": *H. m. amaryllis*, "*rosina*": *H. m. rosina*, "*melp.*": *H. m. melpomene*, "*cydno*": *H. c. chioneus*, "*timareta*": *H. t. thelxinoe*, "Pan": Panama, "FG": French Guiana

Table 3. F_{ST} between sympatric and allopatric populations

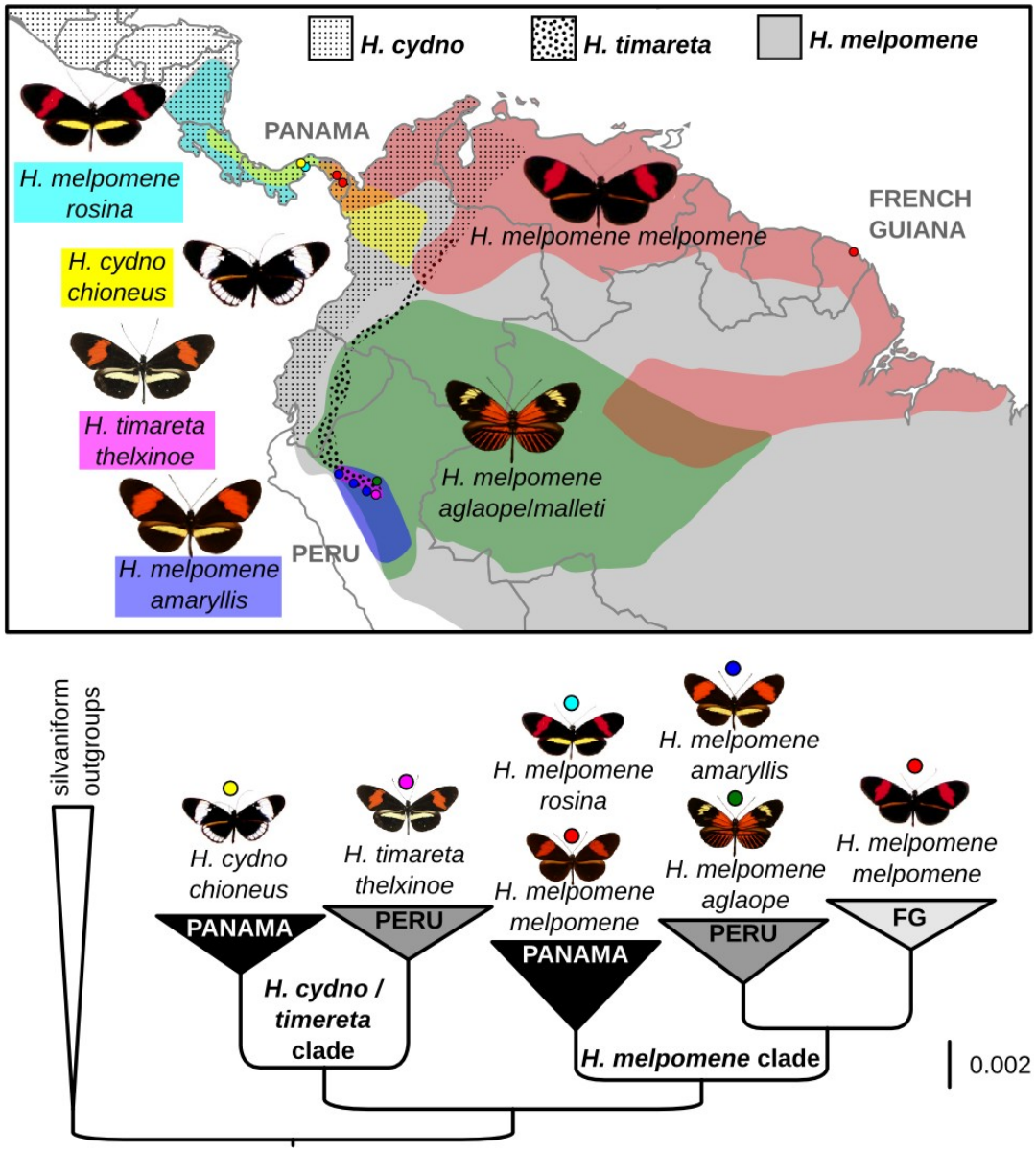
population pair[1]	F_{ST} (WG)	F_{ST} (autosomes)	F_{ST} (Z chrom.)	Coeff. Var. (WG)
cydno : *rosina* (sympatric)	0.292 +-0.003	0.286 +-0.001	0.515 +-0.004	0.252*
cydno : *melpomene* (FG) (allopatric)	0.439 +-0.002*	0.440 +-0.001*	0.540 +-0.003*	0.048
timareta : *amaryllis* (sympatric)	0.287 +-0.003	0.282 +-0.002	0.672 +-0.004	0.470*
timareta : *melpomene* (FG) (allopatric)	0.419 +-0.003*	0.415 +-0.002*	0.716 +-0.004*	0.175

Mean F_{ST} for all non-overlapping 100 kb windows is presented, +- standard error. The coefficient of variation is a standardised measure, representing the standard deviation divided by the mean.

* Indicates significantly greater F_{ST} in allopatry (t-test on arcsine square-root transformed values, $p << 0.0001$), and significantly greater coefficient of variation in sympatry (F-test, $p < 0.001$).

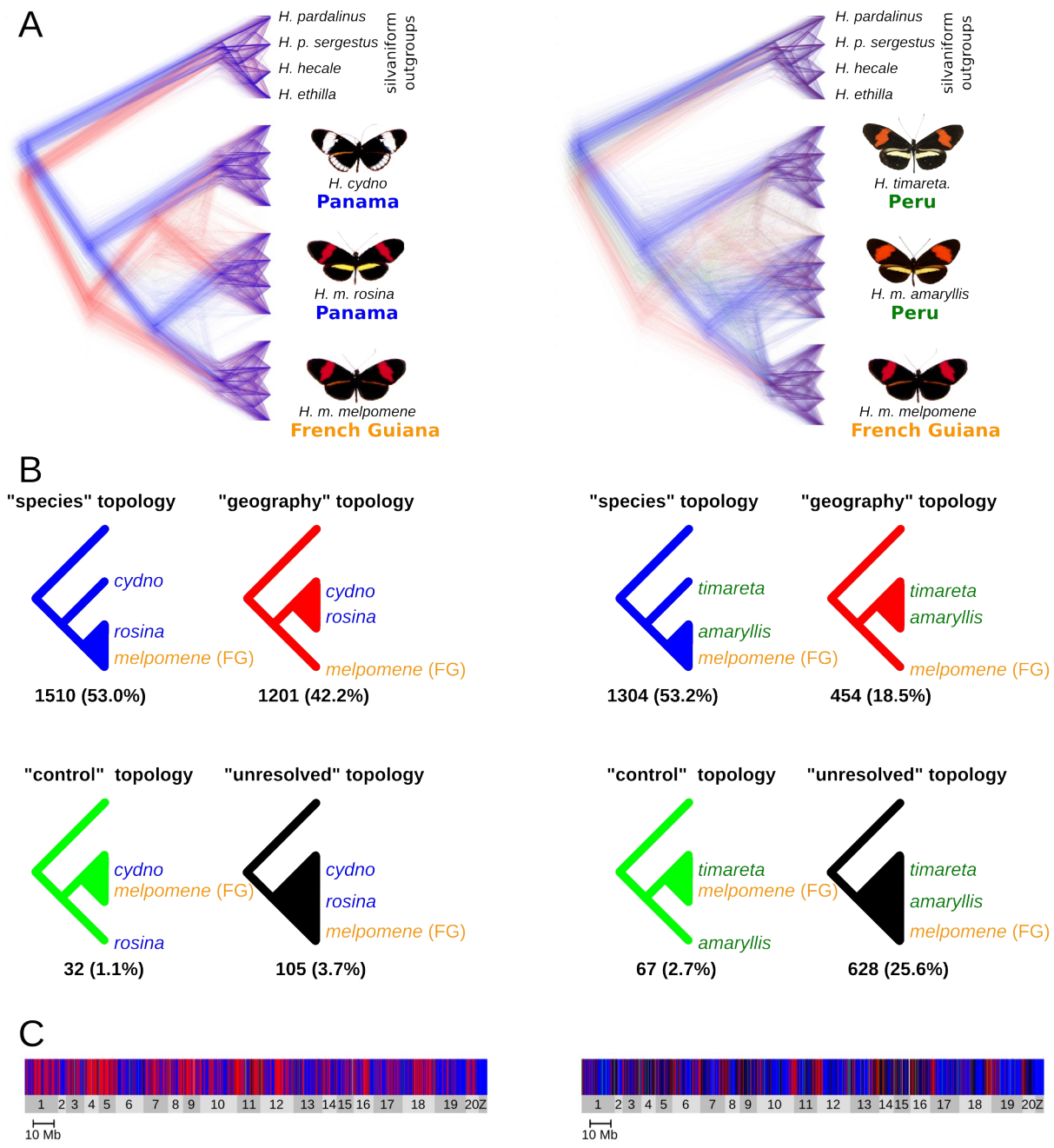
"*amaryllis*": *H. m. amaryllis*, "*rosina*": *H. m. rosina*, "*melpomene*": *H. m. melpomene*, "*cydno*": *H. c. chioneus*, "*timareta*": *H. t. thelxinoe*, "Pan": Panama, "FG": French Guiana

608 **Figures**



610 **Figure 1. Populations sampled and their phylogenetic relationships**

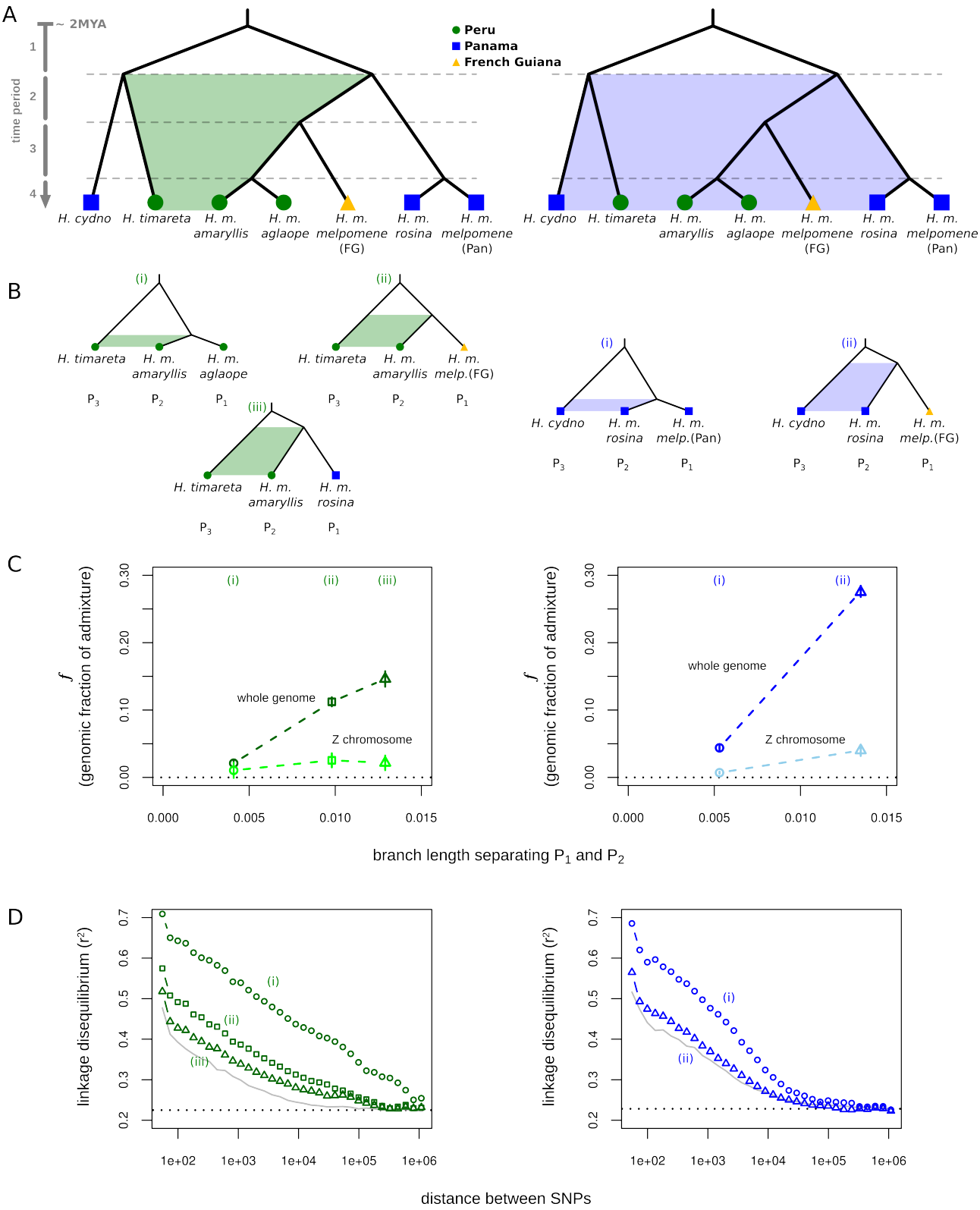
611 The entire distribution of *H. melpomene* is shown in grey. The entire distribution of the *H.*
612 *cydno*/*timareta* clade is shown with dots (Rosser et al. 2012). Colours depict distributions of races
613 used in this study, with dots indicating the sampling locations, and correspond to the coloured dots
614 on the tree. The tree is a compressed version of the whole genome ML tree (Fig. S1). The three
615 general sampling locations, Panama, Peru and French Guiana are indicated. The scale bar refers to
616 the number of substitutions per site.



617 **Figure 2. Four-taxon maximum-likelihood trees for 100 kb windows**

618 **(A)** Trees were superimposed using Densitree (Bouckaert 2010). There were 2848 trees for the *H.*
619 *cydno – H. melpomene* dataset (left) and 2453 for the *H. timareta – H. melpomene* dataset (right).
620 Tree-lengths were equalized so that all trees could be superimposed, and then a random jitter was
621 added to all branch lengths to show density. Trees supporting each of the four possible topologies
622 are coloured accordingly: blue for the species tree, red for the geography tree, green for the control
623 tree and black for unresolved trees. **(B)** The four topologies scored, along with the number and

624 percentage of trees supporting each. See Fig. S3 for examples of trees assigned to each topology.
625 **(C)** The distribution of the four topologies across the genome. Chromosomes are shaded light and
626 dark grey. See figure S4 for an enlarged version.



627 **Figure 3. Measuring admixture at different phylogenetic scales**

628 **(A)** We can distinguish between admixture in different time periods as follows. If gene flow was

629 ancient only (i.e. period 1), then *H. m. amaryllis* and *H. m. rosina* should both be equally admixed

with *H. timareta* and *H. cydno*. However if gene flow is more recent (i.e. period 2, 3 or 4), then *H. m. amaryllis* should be more admixed with Peruvian *H. timareta* (green shading), and *H. m. rosina* should be more admixed with Panamanian *H. cydno* (blue shading). The same logic applies when quantifying admixture for a specific branch: if *H. timareta* shares more derived alleles with *H. m. amaryllis* than with *H. m. aglaope*, this skew must reflect gene flow between *H. timareta* and *H. m. amaryllis* that is more recent than the coalescence between *H. m. amaryllis* and *H. m. aglaope* (i.e. during period 4). **(B)** Our sampling allowed us to quantify admixture at three time scales between *H. timareta* and *H. m. amaryllis*, and two time scales between *H. cydno* and *H. m. rosina*. **(C)** The estimated fraction of admixture (*f*), plotted for the whole genome and the Z chromosome specifically against the estimated length of the time period being analysed, calculated as the average branch length separating populations P_1 and P_2 in the genomic ML phylogeny (Fig. S1). Vertical lines depict standard errors. **(D)** Linkage disequilibrium (r^2) between shared-derived alleles in the P_2 population (*H. m. amaryllis* left, *H. m. rosina* right), plotted as a function of distance on a logarithmic scale. The SNPs used to estimate LD were those carrying a shared derived allele in P_2 and P_3, while P_1 was fixed for the ancestral state (i.e. an ABBA pattern, where the B alleles are not necessarily fixed). The grey line represents the average genomic LD level and the dashed line shows the average LD among unlinked sites.

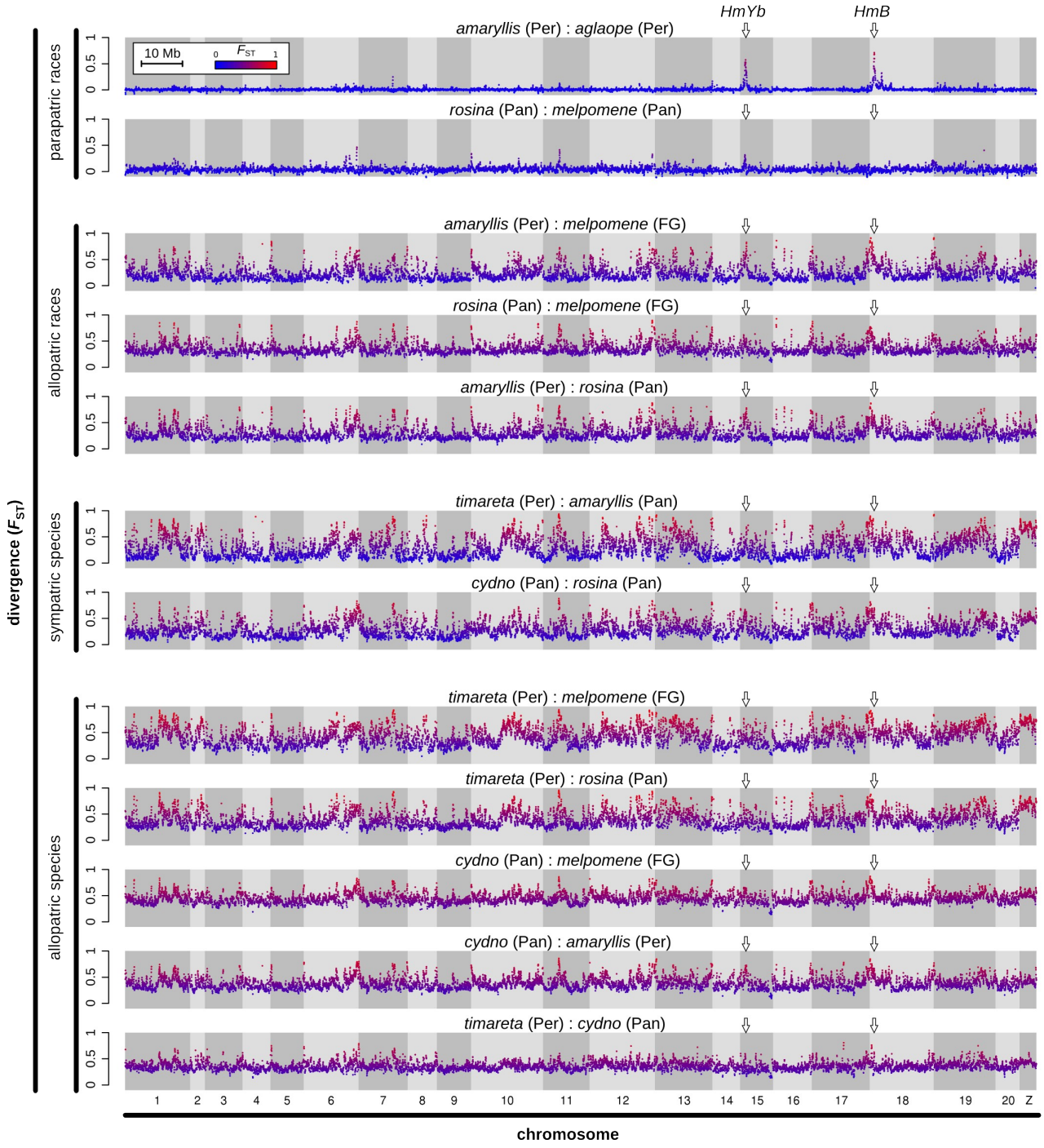
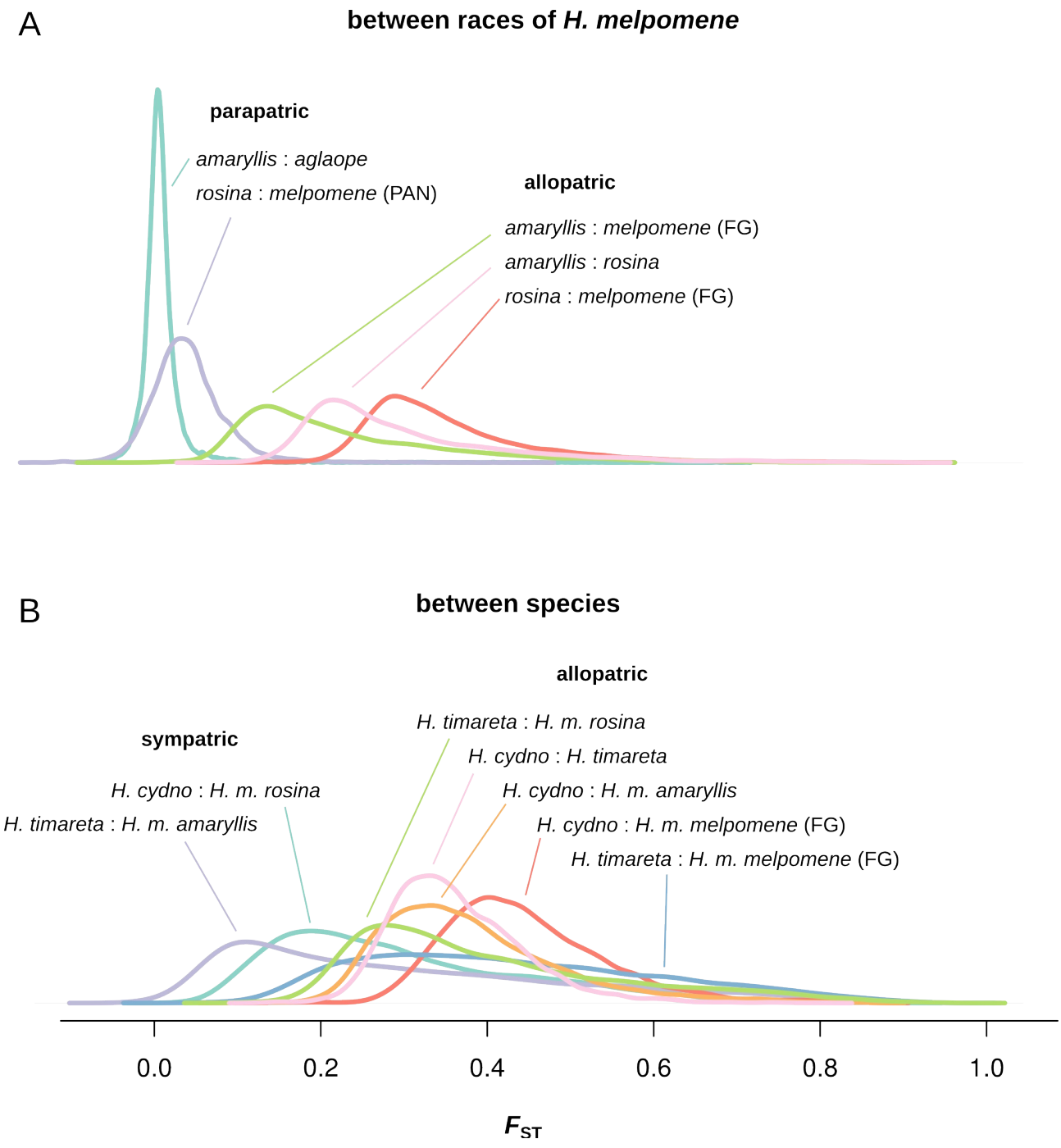


Figure 4. Genomic divergence across the genome at different levels of divergence

F_{ST} values were calculated for 100 kb windows sliding in increments of 20 kb. Chromosomes are shown with alternating light and dark shading. Point colours reflect the absolute level of F_{ST} to allow for comparison between plots. The locations of the wing pattern loci *HmYb* and *HmB* are indicated by arrows. "*amaryllis*": *H. m. amaryllis*, "*rosina*": *H. m. rosina*, "*melpomene*": *H. m. melpomene*, "*cydno*": *H. c. chioneus*, "*timareta*": *H. t. thelxinoe*, "Pan": Panama, "Per": Peru, "FG":

653 French Guiana



655 **Figure 5. Density plots of pairwise F_{ST} values for non-overlapping 100 kb windows**

656 All pairwise comparisons, corresponding to the plots in Fig. 4, between races of *H. melpomene* (A),

657 and between species (B).

References

Baxter SW, Nadeau NJ, Maroja LS, Wilkinson P, Counterman BA, Dawson A, Beltran M, Perez-Espona S, Chamberlain N, Ferguson L, et al. 2010. Genomic Hotspots for Adaptation: The Population Genetics of Müllerian Mimicry in the Heliconius melpomene Clade ed. M.W. Nachman. *PLoS Genet* **6**: 12.

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**: 1372–3.

Bull V, Beltrán M, Jiggins CD, McMillan WO, Bermingham E, Mallet J. 2006. Polyphyly and gene flow between non-sibling Heliconius species. *BMC biol* **4**: 11.

Chamberlain NL, Hill RI, Kapan DD, Gilbert LE, Kronforst MR. 2009. Polymorphic butterfly reveals the missing link in ecological speciation. *Science* **326**: 847–50.

Clayton D, Leung H-T. 2007. An R package for analysis of whole-genome association studies. Human heredity 64: 45–51.

Coyne JA, Orr HA. 2004. *Speciation*. Sinauer, Sunderland, MA.

Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin A V, Hughes DST, Ferguson LC, Martin SH, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**: 94–8.

DePristo M a, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, Philippakis A a, Del Angel G, Rivas M a, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genet* **43**: 491–8.

Van Doorn GS, Edelaar P, Weissing FJ. 2009. On the origin of species by natural and sexual selection. *Science* **326**: 1704–7.

Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Mol biol evol* **28**: 2239–2252.

Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A, et al. 2012. The genomic landscape of species divergence in Ficedula flycatchers. *Nature* 1–5.

Eriksson A, Manica A. 2012. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *P Natl Acad Sci USA* **109**: 13956–60.

Feder J, Egan S, Nosil P. 2012. The genomics of speciation-with-gene-flow. *Trends Genet* 1–9.

Gagnaire P-A, Pavey SA, Normandeau E, Bernatchez L. 2013. The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by rad-sequencing. *Evolution* **In press**.

Garrigan D, Kingan SB, Geneva AJ, Andolfatto P, Clark AG, Thornton K, Presgraves DC. 2012. Genome sequencing reveals complex speciation in the Drosophila simulans clade. *Genome Res* **22**: 1499–1511.

Gavrilets S. 2004. *Fitness landscapes and the origin of species*. Princeton University Press.

Giraldo N, Salazar C, Jiggins CD, Bermingham E, Linares M. 2008. Two sisters in the same dress: Heliconius cryptic species. *BMC Evol Biol* **8**: 324.

Gompert Z, Buerkle AC. 2010. Introgress: a Software Package for Mapping Components of Isolation in Hybrids. *Mol ecol resour* **10**: 378–84.

Grant PR, Grant BR, Petren K. 2005. Hybridization in the recent past. *Am nat* **166**: 56–67.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-Y, et al. 2010. A draft sequence of the Neandertal genome. *Science* **328**: 710–22.

Hohenlohe P a, Bassham S, Etter PD, Stiffler N, Johnson E a, Cresko W a. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genet* **6**: e1000862.

Hohenlohe PA, Bassham S, Currey M, Cresko WA. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philos T Roy Soc B* **367**: 395–408.

Jiggins CD, Linares M, Naisbit RE, Salazar C, Yang ZH, Mallet J. 2001. Sex-linked hybrid sterility in a butterfly. Evolution: International Journal of Organic *Evolution* **55**: 1631–1638.

Jiggins C. 2008. Ecological speciation in mimetic butterflies. *Bioscience* **58**: 541–548.

Kirkpatrick M, Ravigné V. 2002. Speciation by natural and sexual selection: models and experiments. *Am nat* **159 Suppl**: S22–35.

Kronforst MR, Young LG, Blume LM, Gilbert LE. 2006. Multilocus analyses of admixture and introgression among hybridizing Heliconius butterflies. *Evolution* **60**: 1254–68.

Kulathinal RJ, Stevison LS, Noor M a F. 2009. The genomics of speciation in Drosophila: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS genet* **5**: e1000550.

Lawniczak MKN, Emrich SJ, Holloway a K, Regier a P, Olson M, White B, Redmond S, Fulton L, Appelbaum E, Godfrey J, et al. 2010. Widespread divergence between incipient Anopheles gambiae species revealed by whole genome sequences. *Science* **330**: 512–4.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–9.

Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome res* **21**: 936–9.

Machado CA, Kliman RM, Markert JA, Hey J. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of Drosophila pseudoobscura and close relatives. *Mol Biol Evol* **19**: 472–488.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends ecol evol* **20**: 229–37.

Mallet J, Beltrán M, Neukirchen W, Linares M. 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC evol biol* **7**: 28.

Mérot C, Mavárez J, Evin A, Dasmahapatra KK, Mallet J, Lamas G, Joron M. 2013. Genetic differentiation without mimicry shift in a pair of hybridizing Heliconius species (Lepidoptera : Nymphalidae). *Biol J Linn Soc* **In Press**.

Merrill RM, Gompert Z, Dembeck LM, Kronforst MR, McMillan WO, Jiggins CD. 2011a. Mate preference across the speciation continuum in a clade of mimetic butterflies. *Evolution* **65**: 1489–500.

Merrill RM, Van Schooten B, Scott J a, Jiggins CD. 2011b. Pervasive genetic associations between

741 traits causing reproductive isolation in Heliconius butterflies. *Proc Roy Sci B* **278**: 511–8.

742 Merrill RM, Wallbank RWR, Bull V, Salazar PC a, Mallet J, Stevens M, Jiggins CD. 2012.
743 Disruptive ecological selection on a mating cue. *Proc Roy Sci B* **279**: 4907–13.

744 Michel APA, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL. 2010. Widespread genomic
745 divergence during sympatric speciation. *P Natl Acad Sci USA* **107**: 9724–9729.

746 De Mita S, Siol M. 2012. EggLib: processing, analysis and simulation tools for population genetics
747 and genomics. *BMC genet* **13**: 27.

748 Nadeau NJ, Martin SH, Kozak KM, Salazar C, Dasmahapatra KK, Davey JW, Baxter SW, Blaxter
749 ML, Mallet J, Jiggins CD. 2013. Genome-wide patterns of divergence and gene flow across a
750 butterfly radiation. *Mol ecol* **22**: 814–26.

751 Nadeau NJ, Whibley A, Jones RT, Davey JW, Dasmahapatra KK, Baxter SW, Quail M a., Joron M,
752 Ffrench-Constant RH, Blaxter ML, et al. 2012. Genomic islands of divergence in hybridizing
753 Heliconius butterflies identified by large-scale targeted sequencing. *Philos Trans Roy Soc B*
754 **367**: 343–353.

755 Naisbit RE, Jiggins CD, Linares M, Salazar C, Mallet J. 2002. Hybrid Sterility, Haldane's Rule and
756 Speciation in. *Race* **1526**: 1517–1526.

757 Naisbit RE, Jiggins CD, Mallet J. 2007. Mimicry: developmental genes that contribute to
758 speciation. *Evol dev* **5**: 269–80.

759 Noor MAF, Bennett SM. 2009. Islands of speciation or mirages in the desert? Examining the role of
760 restricted recombination in maintaining species. *Heredity* **103**: 439–444.

761 Nosil P, Funk DJ, Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic
762 divergence. *Mol ecol* **18**: 375–402.

763 Nosil P, Gompert Z, Farkas TE, Comeault AA, Feder JL, Buerkle CA, Parchman TL. 2012.
764 Genomic consequences of multiple speciation processes in a stick insect. *Proc Roy Sci B* **In**
765 **Press**.

766 Ott M, Zola J, Stamatakis A, Aluru S. 2007. Large-scale maximum likelihood-based phylogenetic
767 analysis on the IBM BlueGene/L. Proceedings of the 2007 ACM/IEEE conference on
768 Supercomputing - SC '07 1.

769 Pardo-Diaz C, Salazar C, Baxter SW, Mérot C, Figueiredo-Ready W, Joron M, McMillan WO,
770 Jiggins CD. 2012. Adaptive introgression across species boundaries in Heliconius butterflies.
771 *PLoS genet* **8**: e1002752.

772 Pinho C, Hey J. 2010. Divergence with Gene Flow: Models and Data. *Annu Rev Ecol Evol S* **41**:
773 215–230.

774 Prowell DP. 1998. Sex linkage and speciation in Lepidoptera. In *Endless Forms. Species and*
775 *Speciation* (ed. D.J. Howard), pp. 309–319, Oxford University Press, New York.

776 Qvarnström a, Bailey RI. 2009. Speciation through evolution of sex-linked genes. *Heredity* **102**:
777 4–15.

778 Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra M V, Rojas W, Duque C,
779 Mesa N, et al. 2012. Reconstructing Native American population history. *Nature* **488**: 370–4.

780 Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population
781 history. eds. G. Bjerknes, P. Ehn, and M. Kyng. *Nature* **461**: 489–94.

782 Rieseberg LH. 2009. Evolution: replacing genes and traits through hybridization. *Current Biol* **19**:
783 R119–R122.

784 Rosser N, Phillimore AB, Huertas B, Willmott KR, Mallet J. 2012. Testing historical explanations
785 for gradients in species richness in heliconiine butterflies of tropical America. *Biol Journal of*
786 *Linn Soc* **105**: 479–497.

787 Salazar C, Jiggins CD, Taylor JE, Kronforst MR, Linares M. 2008. Gene flow and the genealogical
788 history of Heliconius heurippa. *BMC evol biol* **8**: 132.

789 Sankararaman S, Patterson N, Li H, Pääbo S, Reich D. 2012. The date of interbreeding between
790 Neandertals and modern humans. *PLoS genet* **8**: e1002947.

791 Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P. 2011. Magic traits in speciation:
792 “magic” but not rare? *Trends ecol evol* **26**: 389–97.

793 Smadja CM, Butlin RK. 2011. A framework for comparing processes of speciation in the presence
794 of gene flow. *Mol ecol* **20**: 5123–40.

795 Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with
796 thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–90.

797 Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web
798 servers. *Syst Biol* **57**: 758–71.

799 Supple M a, Hines HM, Dasmahapatra KK, Lewis JJ, Nielsen DM, Lavoie C, Ray D a, Salazar C,
800 McMillan WO, Counterman B a. 2013. Genomic architecture of adaptive color pattern
801 divergence and convergence in Heliconius butterflies. *Genome Res*. **In Press**.

802 Turner TL, Hahn MW, Nuzhdin S V. 2005. Genomic islands of speciation in Anopheles gambiae.
803 *PLoS biol* **3**: e285.

804 Wu C. 2001. The genic view of the process of speciation. *Science* **14**: 851–865.

805

806