Article (Discoveries)

Complex introgression history of the *erato-sara* clade of *Heliconius* butterflies

Yuttapong Thawornwattana[1], Fernando A. Seixas[1], Ziheng Yang[2*], James Mallet[1*]

1. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

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

Corresponding authors: z.yang@ucl.ac.uk, jmallet@oeb.harvard.edu

Abstract (230/250 words)

Introgression plays a key role in adaptive evolution and species diversification in many groups of species including *Heliconius* butterflies. However, frequent hybridization and subsequent gene flow between species makes estimation of the species phylogeny challenging. Here, we infer species phylogeny and introgression events from whole-genome sequence data of six members of the *erato-sara* clade of *Heliconius* using a multispecies coalescent model with introgression (MSci) and an isolation-with-migration (IM) model. These approaches probabilistically capture the genealogical heterogeneity across the genome due to introgression and incomplete lineage sorting in a full likelihood framework. We detect robust signals of introgression across the genome, and estimate the direction, timing and magnitude of each introgression event. The results clarify several processes of speciation and introgression in the *erato-sara* group. In particular, we confirm ancestral gene flow between the *sara* clade and an ancestral population of *H. telesiphe*, a hybrid origin of *H. hecalesia*, and gene flow between the sister species *H. erato* and *H. himera*. The ability to confidently infer the presence, timing and magnitude of introgression events using genomic sequence data is helpful for understanding speciation in the presence of gene flow and will be useful for understanding the adaptive consequences of introgressed regions of the genome. Our analysis serves to highlight the power of full likelihood methods under the MSci model to the history of species divergence and cross-species introgression from genome-scale data.

Keywords: *Heliconius*, *erato*, introgression, multispecies coalescent

Introduction

Gene flow between species (or hybridization and introgression) is increasingly recognized as an important process affecting a variety of species and potentially contributes to adaptive evolution and species diversification (Mallet et al. 2016; Taylor and Larson 2019). Thanks to increasing availability of genomic data and recent advances in analytical methods (Sousa and Hey 2013; Payseur and Rieseberg 2016), introgression has been detected in many groups of species including

Anopheles mosquitoes (Fontaine et al. 2015), *Panthera* cats (Figueiró et al. 2017) and cichlid fishes (Malinsky et al. 2018), as well as in *Heliconius* butterflies (Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012; Kozak et al. 2018; Edelman et al. 2019).

Heliconius is a group of highly diverse butterflies in tropical America. They are unpalatable to predators and are perhaps best known for mimicry in which a single species can have many different wing coloration patterns across its geographical range (classified as subspecies) while multiple unrelated sympatric species have converged to share the same wing pattern in local regions as a common warning sign to deter predators (Bates 1862; Müller 1879). The genus *Heliconius* comprises two major clades, the *erato-sara* clade and the *melpomene*-silvaniform clade, which diverged around 10–12 million years ago (Ma) in the Miocene (Kozak et al. 2015). Natural hybridization and introgression between species are well-documented within both major clades (Mallet et al. 2007). However, the prevalence of introgression between species coupled with rapid radiation and diversification of species and geographic races makes estimation of the species phylogeny challenging. As a result, our understanding of the full speciation history of *Heliconius* remains limited. This work aims to resolve the history of key speciation events and cross-species introgression in the *erato-sara* clade.

Most previous studies of the genus *Heliconius*, including the *erato* group, have focused on evolutionary relationships and gene flow at specific regions of the genome, especially the color pattern loci responsible for phenotypic variation in the mimetic wing patterns, typically in a few species, and mainly in the *melpomene*-silvaniform clade where gene flow appears to be more prevalent (Dasmahapatra et al. 2012; Pardo-Diaz et al. 2012; Nadeau et al. 2013; Martin and Van Belleghem 2017; Martin et al. 2019). Other studies have focused on wing-color pattern loci between two species with comimic races, *H. erato* (*erato-sara* clade) and *H. melpomene* (*melpomene*-silvaniform clade) (Hines et al. 2011; Reed et al. 2011).

Earlier molecular phylogenetic studies of *Heliconius* were based on a small number of nuclear and mitochondrial loci (Brower 1994; Brower and Egan 1997; Beltrán et al. 2002; Beltrán et al. 2007; Kozak et al. 2015), revealing variation in gene genealogies among loci. In particular, Kozak et al. (2015) employed coalescent-based approaches including BUCKy (Larget et al. 2010) and *BEAST (Heled and Drummond 2010) to account for the heterogeneity of gene genealogies across loci. Hybridization and introgression were acknowledged but were not directly accounted for in the analytical methods. Kozak et al. (2018) analyzed genome-wide coding loci of over hundred individuals from forty *Heliconius* species to estimate the species tree using a range of analysis methods to account for introgression. These include summary statistics, such as D (Patterson et al. 2012) and f_4 (Reich et al. 2009), and methods that approximate the multispecies coalescent (MSC) such as ASTRAL (Mirarab et al. 2014) and MP-EST (Liu et al. 2010), as well as InferNetwork MPL in PhyloNet (Yu and Nakhleh 2015), an extension of MP-EST to infer phylogenetic networks. These methods rely on estimated gene trees as input, typically estimated from non-overlapping sliding windows along the genome, and thus did not account for the uncertainty of those estimated gene trees. Van Belleghem et al. (2017) estimated the species phylogeny of the *erato* clade from concatenated SNP data from a subset of chromosomes, focusing on wing color pattern loci among different geographic races of *H. erato*.

Recently, Edelman et al. (2019) carried out extensive phylogenetic analyses with sixteen new genome assemblies of *Heliconius* species using concatenation with sliding-windows and an approximate method implemented in ASTRAL (Mirarab et al. 2014). Neither approach accounts for gene flow between species which can mislead species tree estimation (Jiao et al. 2020). In Edelman et al. (2019), introgression events were inferred using the *D* statistic (Patterson et al. 2012), a phylogenetic network model implemented in PhyloNet-SEQ (Wen and Nakhleh 2018), and a new statistical test to distinguish between introgression and incomplete lineage sorting (ILS) called QuIBL. However, summary statistics such as *D* and QuIBL ignore much information in the multi-locus sequence data such as genealogical variation across the genome, and as a result are unable to provide reliable inference of the direction, timing, and intensity (as measured by the migration rate or introgression probability) of gene flow. The PhyloNet-SEQ method was computationally demanding, only applicable to a small subset of the full genome. In this case, it did not seem to produce reliable inference of the species tree and introgression events, with considerable uncertainty surrounding the tree topology and the timing and direction of introgression events. Finally, Massardo et al. (2020) analyzed the genome data from Edelman et al. (2019) together with three additional genomes in both *erato-sara* and *melpomene*-silvaniform clades and inferred species trees using the same approach as in Edelman et al. (2019).

Here, we infer the species phylogeny and introgression history of six members of the *erato-sara* clade using genome data from Edelman et al. (2019). Unlike previous attempts, we used two full-likelihood coalescent-based phylogenetic approaches that explicitly account for introgression and ILS as sources of genealogical variation across the genome. One approach is based on a multispecies coalescent model with introgression (MSci) (Degnan 2018; Wen and Nakhleh 2018; Zhang et al. 2018), implemented in the program BPP (Flouri et al., 2020). In this approach, introgression is modelled as discrete events that occur at single time points in the past. Another is based on an isolation-with-migration (IM) model (Hey and Nielsen 2004; Hey 2010) implemented in the program 3s (Zhu and Yang 2012; Dalquen et al. 2017). This complementary approach allows for continuous migration at a constant rate between a pair of species after their divergence. Advantages of using full likelihood methods over approximate coalescent methods or summary statistics include making full use of information in the sequence data and properly accounting for uncertainty in local gene genealogies (Xu and Yang 2016). The full likelihood methods allow us to obtain reliable estimates of the direction, timing and magnitude of introgression, as well as effective population sizes and species divergence times for the first time. Estimation of such important parameters from genome-scale sequence data can provide powerful insights into the speciation history of these species, and a basis for further investigations of the evolution of adaptive traits of interest.

Results

Genealogical variation across the genome

Whole genome re-sequencing data of six *Heliconius* species from the *erato-sara* clade - *H. erato demophoon*, *H. himera*, *H. hecalesia formosu*s, *H. telesiphe telesiphe*, *H. demeter* and *H. sara magdalena* (**Table S1**) - were aligned to the *H. erato demophoon* assembly to produce unphased sequence alignments of 74,999 coding loci and 92,966 noncoding loci (**Table S2**). The loci were at least 2,000 bp apart and their lengths ranged between 100-2,000 bp for noncoding loci. We did not constrain the maximum length of coding loci. Mapping to a reference genome allows us to retain information for heterozygous sites (represented using IUPAC codes, e.g. Y for T/C), as opposed to *de novo* assembly of haploid consensus sequence which tends to distort information in the phase of heterozygote sites in the diploid genome and lead to bias in downstream inferences (Andermann et al. 2019). The genotype phase of multiple heterozygous sites is accommodated by the analytical integration method of Gronau et al. (2011), implemented in BPP (Yang 2015; Flouri et al. 2018).

We first used BPP to infer species trees under the multispecies coalescent (MSC) model. This is the A01 analysis in Yang (2015), which explicitly accounts for ILS but assumes no gene flow. We grouped loci into blocks of 100 consecutive loci along each of the 21 chromosomes. Loci within known inversion regions on chromosomes 2 and 15, denoted 2b and 15b, were analysed separately. There were 25 chromosome regions in total, with 749 coding blocks and 933 noncoding blocks (**Table S3**). There is considerable variation in the tree topology across the genome (**Figure 1**, **Tables S4–S5**). Ten species trees were maximum *a posteriori* probability (MAP) trees in at least one block, denoted Trees (i)–(x), although Trees (i) and (ii) accounted for over 95% of the blocks. These two trees differ only in the position of *H. hecalesia* within the *erato* clade: in Tree (i), *H. hecalesia* is sister to the clade (*H. erato*, *H. himera*) while in Tree (ii), it is sister to *H. telesiphe*.

Three regions of the genome have species tree topologies distinct from most of the rest of the genome, consistent with Edelman et al. (2019). First, chromosome 21 (Z chromosome) is the only chromosome for which Tree (i) is a MAP tree for almost all blocks (**Figure 1**, **Table S4**). Among the autosomes excluding inversions, Tree (i) is found in only about 40% and 46% of noncoding and coding blocks, respectively (**Table S4**). The remaining fraction of autosomal blocks consist almost entirely of Tree (ii) (58% in the noncoding loci and 47% in the coding loci), making Tree (ii) the autosome-majority tree. Second, the chromosome 2 inversion (denoted 2b, 1.95 Mb) has an unusual history in the genome whereby *H. telesiphe* is more closely related to the *sara* clade (Tree iii). Note that in this inversion region, *H. erato*, *H. himera* and *H. hecalesia* share a derived inverted rearrangement relative to *H. melpomene, H. sara* and *H. demeter* (Van Belleghem et al. 2017; Davey et al. 2017; Edelman et al. 2019), consistent with Trees (i) and (iii) where these species are clustered together. Finally, the chromosome 15 inversion (denoted 15b, 580 kb) supports Tree (iv) in both coding and noncoding regions, whereby (*H. telesiphe*, *H. hecalesia*) is sister to the *sara* clade instead of being more closely related to other members of the *erato* clade as in Tree (i) or (ii). This grouping strongly suggests that *H. telesiphe*, *H. hecalesia*, *H. demeter* and *H. sara* share the derived inverted rearrangement of this region relative to *H. erato* (Edelman et al. 2019). This 15b inversion contains the *cortex* gene that controls mimetic wing color patterning across *Heliconius* species (Nadeau et al. 2016; Van Belleghem et al. 2017). Tree (iv) also appears as a MAP tree sporadically in other parts of the autosomes, sometimes with high posterior probabilities (**Figure 1**, **Table S5**). These include regions outside the inversion on chromosome 15 as well as on chromosomes 4, 5 and 11. Highly similar results were obtained when the loci were grouped into larger blocks of 200 instead of 100 loci (**Figure S1, Tables S6–S8**).

A possible explanation for the different patterns observed between the Z chromosome and the autosomes is that the Z chromosome tree (Tree i) represents the true species tree while the alternative autosomal trees can be attributable to large amounts of introgression. In fact, Tree (i) and Tree (ii) (autosome-majority tree) can be related via gene flow events. Given Tree (i), gene flow between *H. telesiphe* and *H. hecalesia* in either direction results in Tree (ii). Conversely, given Tree (ii), gene flow between the common ancestor of (*H. erato*, *H. himera*) and *H. telesiphe* results in Tree (i). Moreover, the chromosome 2b inversion-majority tree (Tree iii) can be obtained from Tree (i) via gene flow between *H. telesiphe* and the common ancestor of the *sara* clade (*H. demeter*, *H. sara*), while it requires more complex gene flow scenarios if Tree (ii) is the true species tree.

We also observe that shorter autosomes (e.g. chromosomes 5, 9, 14 and 16) tend to contain predominantly Tree (ii), while longer ones (e.g. chromosomes 1, 10, 12 and 13) tend to have higher proportions of Tree (i). This pattern, also discussed by Edelman et al. (2019), may be explained by the fact that longer autosomes tend to have lower recombination rates, resulting in stronger linkage (Davey et al., 2017; Martin et al., 2019). Since we expect the removal of introgressed loci due to their linkage to deleterious introgressed loci to be more effective in regions of lower recombination, this pattern may thus provide further support for the scenario in which Tree (i) is the true species tree.

Finally, Tree (iv) in the chromosome 15b inversion may be explained by ancient introgression on the autosome-majority tree (Tree ii) between the common ancestors of (*H. telesiphe*, *H. hecalesia*) and (*H. demeter*, *H. sara*). This scenario was supported by phylogenetic network analysis, *D* statistic and an internal branch length test (Edelman et al. 2019). Another possibility is that *H. hecalesia* could result from hybrid speciation that involved *H. telesiphe* and the common ancestor of (*H. erato*, *H. himera*). In this case, both Trees (i) and (ii) would be partially correct in representing the species history. We discuss the possibility of a hybrid origin of *H. hecalesia* in more detail in Discussion.

Pairwise gene flow rates

Gene flow between particular pairs of species could reconcile different species trees from the BPP A01 analysis reported above. We investigated this possibility by explicitly estimating the migration rates between each species pair under the isolation-with-migration (IM) model implemented for species triplets in the program 3S (Zhu and Yang 2012; Dalquen et al. 2017). In this model, migration is allowed between the two species after their divergence at constant rates per generation in the two directions. Note that the migration rate in the IM model, measured in the expected number of migrants per generation ($M = Nm$), represents the long-term effect of continuous gene flow, after the filtering of introgressed alleles by natural selection, affected by local recombination rate (Martin and Jiggins 2017). We used *H. melpomene* as an outgroup in all triplets to improve parameter estimation (**Table S2**). No migration involving the outgroup is allowed in this model.

Based on the 3S analysis, we find evidence for bidirectional gene flow between *H. telesiphe* and *H. hecalesia*, consistent with a scenario in which Tree (ii) (autosome-majority tree) is related to Tree (i) (Z chromosome tree) through introgression between these two species (**Figure 2**). Significant gene flow between these species is found only in the autosomal loci (both coding and noncoding) but not in the Z chromosome. Furthermore, we find no evidence of gene flow in the Z chromosome between the common ancestor of (*H. erato*, *H. himera*) and *H. hecalesia*, which is required to explain Tree (i) as resulting from introgression if Tree (ii) is the true species tree. Together, these results suggest that Tree (i) most likely reflects the true species history.

There were other notable patterns in our estimates of migration rates (**Figure 2**, **Table S9**). First, the Z chromosome, particularly the noncoding loci, is almost devoid of gene flow, in sharp contrast with the autosomes. Second, in the autosomes, the highest estimates of migration rates were found between the two sister species *H. erato* and *H. himera*, with gene flow occurring in both directions (at the rate of 0.085–0.121 migrants per generation from *H. himera* to *H. erato* and 0.048–0.067 in the opposite direction) (**Figure 2**). Third, there was evidence for gene flow into *H. erato*, *H. himera* and *H. demeter* from every other species, and gene flow from *H. sara* to all other species. We also observe consistent patterns when individual autosomes were analyzed separately (**Figure S2**). Moreover, noncoding regions of several chromosomes such as chromosomes 2, 3, 4, 7 and 15 show no evidence of gene flow. There is no clear association between these chromosomes and the chromosome size, the number of loci or the proportion of the Z chromosome tree.

The complex pattern of gene flow inferred using 3S may reflect truly complex introgression in this group of species as well as a limitation of the triplet approach and the difficulty of using pairwise migration rates to reconstruct the full history for all species. One possible scenario for the detected gene flow between most pairs of species is that there was extensive ancestral introgression between the two common ancestors of the *sara* group and the erato group. However, the model assumes a continuous gene flow from the time of divergence between the ingroup species S_1 and S_2 to the present time. If gene flow occurred after the species split but stopped some time later, the 3S analysis under the misspecified model may still be expected to infer gene flow between the ingroup species, but with distorted estimates of migration rates and other parameters.

Construction of a full history of species divergences and cross-species introgression

We used the results from the previous two analyses to formulate a plausible species history of the *erato-sara* clade, incorporating both species divergences and cross-species introgression events (**Figure 3**). The Z chromosome tree from the BPP A01 analysis was used as a base species tree while introgression events were informed by migration rate estimates from the 3S analysis. We then used this species-tree model with introgression to fit the MSci model to estimate introgression probabilities (φ), population size parameters (θ) and species divergence times (τ) using all coding or noncoding loci in each chromosomal region. This is the A00 analysis under MSci in BPP (Flouri et al. 2020).

The estimates of the introgression probability from *H. telesiphe* into *H. hecalesia* ($1 - \varphi_c$) are consistently high (>0.5) across the genome (**Figure 4**) and gene flow between these species is estimated to have occurred right after *H. hecalesia* split from (*H. erato*, *H. himera*) (**Figure 5**), supporting the hypothesis that *H. hecalesia* is a hybrid species. Even though the model assumes different times for species divergence and for introgression (with $\tau_{ehc} > \tau_c$ or model B in Flouri et al. (2020)), the posterior estimates strongly suggest that the times actually coincided (with $\tau_{ehc} \approx \tau_c$ or model C in Flouri et al. (2020)). This introgression is virtually absent in two regions of the genome, the chromosome 2b inversion and the Z chromosome. This pattern is consistent with our blockwise estimates of the species trees (BPP A01 analysis; **Figure 1**) where autosomal regions, with the exception of the chromosome 2b inversion, are dominated by Tree (ii) as a result of *H. telesiphe* → *H. hecalesia* introgression on Tree (i), the Z chromosome tree.

We infer substantial introgression probabilities between *H. telesiphe* and the common ancestor of *H. demeter* and *H. sara*, across the entire genome. Introgression is more prevalent from *H. telesiphe* to (*H. demeter*, *H. sara*) ($1 - \varphi_{ds2}$ approximately 0.5) than in the reverse direction ($1 - \varphi_{tc2}$ ~0.1) (**Figure 4**, **Tables S10–S11**). One exception is the chromosome 2b inversion, which has the opposite pattern, with higher introgression from (*H. demeter*, *H. sara*) to *H. telesiphe*. Note that introgression in either direction between this pair, and assuming Tree (i) as the base tree, would result in Tree (iii), which is mostly restricted to the chromosome 2b inversion.

Other introgression events have relatively low probabilities across the genome despite strong evidence from the 3s analysis using the IM model. First, bidirectional gene flow between *H. erato* and *H. himera* had probabilities of about 15% for *H. erato* → *H. himera* ($1 - \varphi_h$) and about 5% for *H. himera* → *H. erato* ($1 - \varphi_e$), consistently throughout the genome (**Figure 4**). For comparison, the continuous migration rate per generation from the IM model is relatively larger in the *H. himera* → *H. erato* direction (**Figure 2**). The differences in the predominant direction of gene flow between *H. erato* and *H. himera* might reflect different episodes of gene flow. During an early phase, there might be more *H. himera* → *H. erato* gene flow as the range of *H. himera* contracted. The present-day range of *H. himera* is confined to a small area surrounded by several races of *H. erato*, so more *H. erato* → *H. himera* may be expected. We caution that the introgression probabilities in the MSci model and migration rates in the IM model are not directly comparable since they capture different aspects of gene flow (Jiao et al. 2020).

Second, the *H. sara* → *H. demeter* introgression probability ($1 - \varphi_d$) is small (<1%), except for the coding loci in the Z chromosome and in a part of the chromosome 2 adjacent to the inversion 2b, labelled 2c. Note that in both cases, the introgression is between two sister species so we may expect the IM model to well capture the long-term effective migration rates between the species pair, provided that there is no gene flow involving the outgroup species (*H. melpomene*).

The estimates of population sizes (θ) and species divergence or introgression times (τ) are highly consistent across chromosomal regions, as well as between coding and noncoding loci (**Figure S3**). The divergence time estimates from noncoding loci tend to be higher, often almost twice as high, than those from coding loci. This is mainly because both θ and τ are measured by the expected number of mutations per site and noncoding loci have a high mutation rate. Exceptions are the two introgression times, one between *H. erato* and *H. himera* and the other for *H. sara* → *H. demeter*, both of which are estimated to be recent, with $\tau < 0.001$ for most chromosomes. We estimate the age of the base of the *erato-sara* clade (τ_r) to be about 0.027 for noncoding loci on average (**Figure S3**, **Table S11**). This translates to about 2.3 million years ago (Ma), assuming 2.9×10^{-9} neutral mutations per site per generation and 4 generations per year (Keightley et al. 2015). The oldest τ_r estimates of about 0.029 (~2.5 Ma) are from noncoding loci in three chromosomes: 2, 15 and Z. As noted previously by Keightley et al. (2015), estimates based on mutation rate are much younger than previous estimates of ~5 Ma based on few coding loci and fossil calibration or ~4 Ma under the MSC model with relaxed molecular clock (Kozak et al. 2015). However, because few details are provided in Kozak et al. (2015), the estimates are not readily compared.

We also estimate that the introgression *H. telesiphe* → *H. hecalesia* (τ_{tc1}) occurred soon after *H. hecalesia* diverged from the common ancestor of *H. erato* and *H. himera* (τ_{ehc}) at around 1.1–1.2 Ma, based on estimates of τ_{tc1} ~0.0128 and τ_{ehc} ~0.0140 from the noncoding loci. Exceptions are chromosomes 2, 15 and Z, where there is a more noticeable gap between these two events (**Figures 5** and **S4**).

The population size parameters (θ) show considerable variation across chromosomes, suggesting that they are more difficult to estimate reliably (**Figure S3**, **Tables S10–S11**). Population sizes for ancestral species represented by short internal branches on the tree are in particular hard to estimate due to lack of coalescent events during the short time interval. Most populations have θ

estimates in the order of 0.01 while a few have more extreme estimates. *H. erato* (θ_{Era}) has small population size estimates of orders 0.0001–0.001 in most chromosomes, except for chromosome 2b inversion where θ_{Era} is much larger, ~0.05, in both coding and noncoding regions. We note that *H. erato* and *H. himera* genome sequences were obtained from partially inbred individuals to aid assembly. In the other extreme, four populations appear to have large population size estimates (>0.01) with large intervals. These are *H. sara*, ds2 (the parent population of the *sara* clade), e (the parent population of *H. erato*) and tc1 (the parent population of *H. telesiphe*) (see **Figure 3**). The large uncertainties may reflect the lack of information to estimate the parameters accurately.

Discussion

Comparisons with previous studies

In this study, we used whole-genome sequence data of the *erato-sara* clade of *Heliconius* to infer the speciation history under the MSci model in a full-likelihood framework. This is an improvement over previous approaches, which rely mainly on summary statistics or approximate methods that do not make full use of sequence information. We find that the Z chromosome is more likely to reflect the true species relationships, consistent with previous findings that the Z chromosome is more resistant to gene flow (Van Belleghem et al. 2018). We also find evidence supporting *H. hecalesia* as a hybrid species, gene flow from the common ancestor of *H. telesiphe* into the *sara* clade, and ongoing gene flow between *H. erato* and *H. himera*.

The sliding-window analysis of Edelman et al. (2019) using non-overlapping 50-kb windows, while not directly comparable, is largely consistent with our analysis. In both analyses, longer chromosomes have a greater proportion of windows or blocks of Tree (i), which is thought to be caused by a reduced recombination rate per base of the larger chromosomes. However, in our analysis the contrast between long and short chromosomes seems sharper (**Figure 1**). Also, Trees (i) and (ii) in our analysis, which are MAP trees in over 95% of the genomic blocks, were represented by only about 70% of sliding windows in Edelman et al. (2019, their Figure 2). One sliding-window tree, Tree (xi) in **Figure 1**, never appears as a MAP tree in our analysis. Among the rare trees (v)–(x) from our analysis, Tree (v), with *H. demeter* forming a clade with *H. hecalesia*, *H. erato* and *H. himera*, is not in the top eight trees considered in Edelman et al. (2019), while it is a MAP tree with probability 0.95 in a single block in the middle of chromosome 17 (**Figure 1**). Tree (vi), which is similar to Tree (i) but with the branching order of *H. telesiphe* and *H. hecalesia* swapped, is a MAP tree in ~1% of the blocks (both coding and noncoding), with generally low posterior probabilities (**Table S5**).

One reason for those differences is that our block covers a much larger region along the chromosome, about 400 kb and 800 kb for a block of 100 and 200 loci, respectively, than the 50-kb window of Edelman et al. (2019). Another important factor is that in our analysis, gene-tree conflicts due to the coalescent process are filtered out by the MSC model whereas they contribute to the fluctuations among the sliding windows of Edelman et al. (2019). In the BPP analysis, there are two major reasons for the fluctuations among the blocks: sampling errors due to only 100 or 200 loci in the block, and cross-species gene flow with intensity varying among the chromosomal regions or along the genome. These factors also contribute to the fluctuations among sliding windows, as well as the conflicts among gene trees due to the coalescent process. While concatenation is known to produce anomalous gene trees (Roch and Steel 2015) so that the estimated gene tree from each sliding window may differ from the underlying species phylogeny, we expect this to be rare and may not contribute to the differences observed between the analyses. If we consider species divergences and cross-species introgression events, which affect whole genomes or whole autosomes, as the species history and local genealogical fluctuations as stochastic noise from the coalescent process, the BPP A01 analysis (**Figure 1**) may be seen to be describing the history, while the sliding-window analysis will be describing the noise as well as the history. Note that it is difficult to separate the noise from the history in a sliding-window analysis. For example, the frequencies of (gene) Trees (i) and (ii) in Edelman et al. (2019) comprise 25.8% and 24.3% of 10kb sliding windows (their Fig. S78), but 34.8% and 35.5% of 50kb windows (their Fig. 2B) when mapped to the *H. erato* reference, and 45.4% and 36.4% of 50 kb windows when mapped to the more distant *H. melpomene* reference (their Fig. S77). In our BPP analysis, the frequencies of (species) Trees (i) and (ii) are 42.8% and 54.8% among the 100-loci noncoding blocks, 48.3% and 44.2% among the 100-loci coding blocks, 40.2% and 58.9% among the 200-loci noncoding blocks, and 49.5% and 46.9% among the 200-loci coding blocks. While our MSci analysis, by using relatively large blocks, may smear local genealogical history variation, the narrower scale sliding window approach suffers by incorporating noise and incomplete lineage sorting. It is currently hard to adjudicate between these competing problems, but the overlapping results of both approaches provides some confidence that neither is very inaccurate.

Previous estimates of migration rates and timings in *Heliconius* under models of gene flow such as the IM model have been limited to a few loci (Bull et al. 2006; Kronforst et al. 2006; Kronforst 2008; Salazar et al. 2008; Pardo-Diaz et al. 2012) or a few species (Kronforst et al. 2013; Van Belleghem et al. 2020), most of which focused on the *melpomene* clade. Using joint site-frequency spectrum and a secondary-contact model, Van Belleghem et al. (2020) estimated migration rates between *H. himera* and two races of *H. erato* (*H. erato favorinus* and *H. erato emma*) to be around 0.5–0.6 migrants per generation from *H. erato* to *H. himera*, and 0.07–0.13 in the opposite direction. This asymmetric gene flow between the two species is consistent with our estimates of introgression probabilities under the MSci model (**Figure 4**). No estimates of the other introgression probabilities in **Figure 3** have been previously reported.

Evidence of introgression in the wild

Using IM and MSci models, our analyses support four introgression events (**Figure 3**): (1) between *H. telesiphe* and the common ancestor of the *sara* clade, (2) from *H. telesiphe* into *H. hecalesia*, (3) between *H. erato* and *H. himera*, and (4) from *H. sara* into *H. demeter*. The first two are consistent with previous genomic studies (Kozak et al. 2018; Edelman et al. 2019). Ancient introgression like (1) may be difficult to confirm with empirical evidence. The recent introgression (3) is well-documented in natural populations and in mating experiments (Jiggins et al. 1997; McMillan et al. 1997; Mallet et al. 2007). Introgression (4) has been reported neither from genomic studies nor from natural populations.

However, since the six species we included in this study constitute only a small fraction of the species and geographic races in the *erato-sara* clade, interpretation of the estimated introgression requires some caution. In particular, detected signals for introgression may be indirect, involving related species or subspecies unsampled in the data.

For example, the gene flow signal *H. telesiphe* → *H. hecalesia* from the genome data likely does not imply that introgression actually occurred between these two species. *H. telesiphe telesiphe* and *H. hecalesia formosus* (the subspecies used in this study) do not currently overlap in their geographical distributions (Rosser et al. 2012). However, the range of *H. hecalesia* does overlap with that of *H. clysonymus* and its sister species *H. hortense* in West Ecuador, the Colombian Andes, and Central America (Rosser et al. 2012), and there are documented natural hybrids between *H. hecalesia* and both *H. clysonymus* and *H. hortense* (Mallet et al. 2007). Previous phylogenetic studies showed that *H. telesiphe* forms a well-supported clade with *H. clysonymus* and *H. hortense*, nested within the *erato* clade (Kozak et al. 2015; Massardo et al. 2020). Therefore, it is likely that the signal from introgression loci in *H. hecalesia* might come from *H. clysonymus*, *H. hortense*, or any ancestral population of these three species. These sources of introgression, *H. clysonymus* and *H. hortense*, are supported by a phylogenetic network analysis and *D* statistic in Kozak et al. (2018).

Another case is the possible indirect gene flow signal between *H. erato* and *H. himera*. *H. himera* is considered an incipient species within *H. erato* (*sensu lato*). It is restricted to middle elevations (800–2,000 metres) of the Andes in South America; in contrast, the subspecies of *H. erato* used in this study, *H. erato demophoon*, is found in Central America (Rosser et al. 2012). However, *H. himera* is parapatric with several subspecies of *H. erato* (such as *cyrbia*, *favorinus*, *lativitta* and *emma*) (Rosser et al. 2012), with narrow contact zones where natural hybrids can be found at high frequencies around 5–10% (Jiggins et al. 1996; Jiggins et al. 1997; Mallet et al. 1998). Thus the introgression signal could come from these other subspecies of *H. erato* instead of *H. erato demophoon*.

Challenges in Bayesian inference of species tree with introgression

Bayesian inference under the MSci model implemented in the program BPP (Flouri et al. 2020) provides a powerful tool for estimating the timing and intensity of introgression on a species phylogeny from genome-scale data while accounting for ILS and genealogical heterogeneity across the genome. Our approach also retains heterozygous sites in the genome data and averages over all possible heterozygote phase resolutions (Gronau et al. 2011), rather than using haploid consensus sequences. Other likelihood-based methods incorporating cross-species gene flow are under active development (Hey et al. 2018; Wen and Nakhleh 2018; Zhang et al. 2018; Jones 2019). However, these are computationally demanding and currently impractical for data of 50 or 100 loci.

Here we discuss several limitations of our approach in this study. First, the introgression model we formulated (**Figure 3**) relies on two main sources of information: the species tree from the BPP analysis under the MSC without introgression and pairwise migration rate estimates. However, integrating such information into a single species tree model with introgression is nontrivial. Ideally, one would like to infer the species phylogeny and introgression events simultaneously as well as the associated parameters such as introgression and divergence times. However, this is a difficult trans-model inference problem and is beyond the limit of current implementation of the MSci model. We attempted to use several heuristic methods to infer the introgression model, which take estimated gene trees for individual loci as input data, but the results do not appear to be plausible, with in particular apparently spurious introgression events inferred around the root of the species tree. Instead we have relied on the 3S program, which is a full likelihood method applied to sequence alignments directly and can accommodate tens of thousands of loci but is limited to only 3 species and 3 sequences per locus.

Second, even with a fixed MSci model (**Figure 3**), inference of demographic parameters is computationally challenging. Some parameters appear more difficult to estimate reliably, such as population sizes associated with short branches in the species tree (e.g. θ_{Sar}, θ_e, θ_{ds1}; see **Figure S3**). Contributing factors include unidentifiability of the model, multimodal posterior distribution, poor mixing of Markov chain Monte Carlo (MCMC) sampling, and potentially model misspecification. The unidentifiability problem arises because of the symmetry in the likelihood induced by bidirectional gene flow such as the one between *H. erato* and *H. himera*. In this case, swapping their two parent populations e and h and flipping the introgression probabilities, so that θ_e and θ_h are interchanged, with φ_e becoming $1 - \varphi_e$ and ϕ_h becoming $1 - \varphi_h$, yields equivalent likelihood values. This is a label-switching problem, and is discussed in Flouri et al. (2020). When this occurs, we process the MCMC samples to collapse the two equivalent sets of parameters into a single set (e.g. with $\varphi_e > 0.5$) before calculating posterior summaries. However, in a complex model with multiple bidirectional introgression edges, identifying parameters involved in label-switching is less trivial. We relied on manual inspection of MCMC samples to identify the issue. An example of MCMC outputs and processing are provided in **Figure S5**.

Third, we avoided explicit modeling of recombination by analyzing short genomic loci (~200 bp on average) that are far apart (at least 2 kb), assuming that sites within each locus have zero recombination, whereas different loci are free to recombine. Sites are expected to be approximately independent at physical distances of ~10 kb apart, based on linkage disequilibrium decay estimated in the *H. melpomene* group species (Dasmahapatra et al. 2012). Thus, some correlation of gene genealogies between consecutive loci may be expected. The unrealistic assumption of free recombination is expected to have the effect of exaggerating the information content in the data, leading to too narrow credibility intervals in parameter estimates. The assumption of no recombination among sites of the same locus may be more concerning since recombination causes different parts of the sequence to have different histories while the model assumes one history. However, the loci used in this study are short, and the assumption of no recombination within locus appears to hold approximately. In general the effects of recombination on many inference problems under the MSC model with and without gene flow are poorly understood. Previously Lanier and Knowles (2012) found that species tree estimation under MSC is robust to moderate and realistic levels of recombination. The impacts of recombination on estimates of the migration rate or introgression probability are unknown, in particular when the recombination rate varies across the genome. Variable recombination rates across the genome are known to differentially affect the rate of gene flow because introgressed loci are accepted in the recipient population with different probabilities due to linkage to introgressed deleterious loci, as found in *Heliconius* (Edelman et al. 2019; Martin et al. 2019) and other organisms (Burri et al. 2015; Schumer et al. 2018).

Fourth, the migration (IM) model assumes in 3S and the introgression (MSci) model in BPP may be the simplest model of gene flow incorporating a species phylogeny. Currently we have little knowledge of the behavior of inferences when the model is misspecified, for example, when the MSci model is fitted to data generated under the IM model or vice versa (Jiao et al. 2020, their Figure 7). In reality one may expect the intensity of gene flow to vary across the genome and over time. It will be interesting to examine the performance of estimation methods such as 3S and BPP under such complex scenarios of gene flow, and to develop methods that may account for such variation.

Methods

Genome sequence data and genotyping

We used raw whole-genome sequencing data of seven *Heliconius* species (**Table S1**) previously generated (Edelman et al. 2019): six species in the *erato-sara* clade and one outgroup (*H. melpomene*). Sequencing reads were filtered for Illumina adapters using cutadapt v1.8.1 (Martin, 2011) and then mapped to each of the chromosome-level genome assembly of *H. erato demophoon* v1 available from lepbase.org, using BWA mem v0.7.15 (Li 2013) with default parameters and marking short split hits as secondary. Mapped reads were sorted and duplicate reads removed using sambamba v0.6.8 (Tarasov et al. 2015). Realignment around indels was performed with the Genome Analysis Toolkit (GATK) v3.8 *RealignerTargetCreator* and *IndelRealigner* modules (McKenna et al. 2010; DePristo et al. 2011), in order to reduce the number of indel miscalls. Read depth and other relevant read alignment quality control metrics were computed using QualiMap v2.2.1 (Okonechnikov et al. 2016).

The main program for inference of cross-species introgression we used, BPP (Yang 2015; Flouri et al. 2018), implements the analytical integration method of genotype phase resolution, averaging over all possible phase resolutions, weighting them using the likelihood based on the multispecies alignment at each locus (Gronau et al. 2011). The program takes the unphased diploid genomic sequence data as input. Thus we generate the unphased diploid sequences from the reads, instead of the haploid consensus sequences commonly produced in genome sequencing projects. The latter in effect resolves the genotype phase at random and produces chimeric sequences that may not exist in nature, potentially introducing systematic biases in downstream inference (Andermann et al. 2019). We analyzed the genotype calling error rate as a function of the base-calling error rate (or per-base sequencing error rate) and the read depth to help determine cutoffs for data filtering (see next section).

Genotype calling was performed for each individual separately with bcftools v1.5 (Li et al. 2009) *mpileup* and *call* modules (Li 2011), using the multiallelic-caller model (*call -m*) and requiring a minimum base and mapping qualities of 20. Genotypes were then filtered using the bcftools *filter* module. Both invariant and variant sites were required to have a minimum quality score (QUAL) of 20. Furthermore, we required that each individual genotype had a genotype quality score (FMT/GQ) >= 20 and a depth of coverage (FMT/DP) between half and twice the mean depth of coverage. For individuals for which half of the mean depth of coverage was below 20X a minimum FMT/DP filter of 20 was used instead. The latter filter was chosen to minimize genotyping error rates (see next section) while retaining a sufficiently large number of loci across the genome. This choice of filter achieved the genotype calling error rate under 0.05% (**Figure 6**), assuming the base-calling error rate of 0.1% (**Table S12**). For female Z chromosome, only one chromosome copy is present, so the minimum required depth was half of that used for the autosomes. All genotypes that did not fulfill these requirements or were located within 5-bp of an indel (--SnpGap) were recoded as missing data.

Analysis of base calling and genotype calling error rates as a guide for data compilation

The base-calling error rate for our dataset was estimated from the genome data using the proportion of non-matching bases in homozygote genotype calls when mapping to the *H. melpomene* reference genome. First, only positions with homozygous genotype calls with a minimum depth of coverage (FMT/DP) ≥ 50 were retained. Because mapping errors can bias our estimates of base call error rate, sites overlapping repetitive regions were excluded. Homozygous sites matching the reference allele (GT = 00) or alternative allele (GT = 11) were analyzed separately. For sites passing the above filters we recorded both the number of reads passing quality control covering that site (FMT/DP) and the number of reads supporting each of the reported alleles (FMT/AD). We considered the difference between DP and AD (i.e. the number of reads not supporting the allele in the homozygote call) as the number of erroneous base calls. An estimate of the base-calling error rate was then obtained by calculating the ratio between the total number of erroneous base calls and total depth, both summed across all sites passing filters. The estimate was about 0.08% for the homozygous reference allele while it was higher for the homozygous alternative allele, ~0.20%, with the difference depending on the genome (**Table S12**). This difference may be due to an increase in alignment errors in more divergent parts of the genome. Different filtering conditions yielded similar estimates (**Table S13**).

We assessed the genotype calling error rate for a given base-calling error rate and the read depth (**Figure 6**), with the genotype at any single site called using a maximum likelihood (ML) method (Li 2011). Let ε be the given base-calling error rate, which is assumed to be the same among reads and independent of the true nucleotide, and let n be the read depth. Given the data of k 1s and $(n-k)$ 0s among the n reads, the likelihoods for the three genotypes (00, 01, and 11) are given by the binomial probabilities as

$$\begin{aligned} L(00|k) &= \Pr(k|\mathrm{GT}=00) = \binom{n}{k}(1-\varepsilon)^{n-k}\varepsilon^{k}, \\ L(01|k) &= \Pr(k|\mathrm{GT}=01) = \binom{n}{k}\left(\tfrac{1}{2}\right)^{n}, \\ L(11|k) &= \Pr(k|\mathrm{GT}=11) = \binom{n}{k}(1-\varepsilon)^{k}\varepsilon^{n-k}. \end{aligned} \tag{1}$$

The genotype achieving the highest likelihood is the called (inferred) genotype. The genotype calling error rate (for the true genotype 00, say) is an average over the possible read outcomes (i.e., over k = 0, …, n)

$$\begin{aligned} e_{\mathrm{GT}=00} &= \textstyle\sum_{k=0}^{n} Pr(\,k|\mathrm{GT}=00) \times I_{\widehat{\mathrm{GT}}\neq 00}\,, \\ e_{\mathrm{GT}=01} &= \textstyle\sum_{k=0}^{n} Pr(\,k|\mathrm{GT}=01) \times I_{\widehat{\mathrm{GT}}\neq 01}, \end{aligned} \tag{2}$$

where $I_{\widehat{\mathrm{GT}}\neq 00}$ is either 1 or 0 depending on whether or not the called genotype $\widehat{\mathrm{GT}}$ (from data k) is in error.

The calculations are shown in **Figure 6** for different base-calling error rates. Note that even at very low base-calling error rate (ε), the genotype calling error can be very high, especially at low read depth. Furthermore, the error rate for heterozygotes (GT = 01) is much higher than for homozygotes. Given the read depth n, the genotype calling error does not necessarily decrease when ε decreases. For example, when $n = 7$, the genotype calling error for a homozygote is 0.0020 at $\varepsilon = 0.01$, but rises to 0.0345 when ε is reduced to 0.005 (**Figure 6**). This is due to the discrete nature of the read outcome. At $\varepsilon = 0.01$, the called genotype is 00 (the true GT) when k = 0 or 1, is 01 when k = 2-5 and is 11 when k = 6 or 7, while at $\varepsilon = 0.005$, the called genotype is 00 when k = 0, is 01 when k = 1-6 and is 11 when k = 7. Similarly given the base-calling error rate ε, the genotype calling error does not necessarily decrease when the read depth increases. For example, when $\varepsilon = 0.01$ the genotype calling error for a homozygote is 0.0199 at $n = 2$, but rises to 0.0585 when n increases to 6 (and then drops to 0.00269 when $n = 7$) (**Figure 6**).

Multilocus datasets for BPP and 3S analyses

We prepared two datasets, one for two BPP analyses (under the MSC and MSci models) and the other for 3S analysis (under the IM model). The 'BPP dataset' has six species in the *erato-sara* clade. The '3S dataset' has an extra species, *H. melpomene*, as an outgroup (**Table S2**). Both datasets were prepared in the same way. For each dataset, we defined coding (exonic) and noncoding regions based on the gene annotation of *H. erato demophoon* v1 reference assembly. For the noncoding regions, we extracted the genotype calls for small genomic segments, referred to as loci, with sizes between 100 and 2,000 bp, and with the further requirement that any two consecutive loci must be at least 2,000 bp apart. For each of these loci we then produced sequence alignments from the genotype calls (heterozygous genotype calls were coded using IUPAC codes), removing positions with missing genotypes or overlapping repetitive regions (based on the repeat annotation of *H. erato demophoon* v1 reference assembly, available at http://download.lepbase.org/v4/repeatmasker/). We excluded loci with 10 or fewer sites and loci with more than 50% gaps. For the coding regions, the same filters were applied except that there was no constraint on the maximum size of loci. There were 74,999 coding loci (median length of 165, and median informative sites of 3) and 92,966 noncoding loci (median length of 237, and median informative sites of 5) (**Table S2**).

Inferring species divergence history across the genome using multispecies coalescent model (BPP A01 analysis)

We inferred species trees using the Bayesian method under the MSC model implemented in BPP v4.1.4 (Yang and Rannala 2014; Rannala and Yang 2017; Flouri et al. 2020). This A01 analysis under the MSC accounts for deep coalescence and the resulting gene tree heterogeneity along the genome but assumes no gene flow. We performed one analysis for each set of coding and noncoding regions of autosomes and the Z chromosome (four sets in total). For each set, we grouped the loci into blocks of 100 loci and inferred one species tree per block (**Table S3**). Two inversion regions in chromosomes 2 and 15 were analyzed separately, denoted 2b and 15b, respectively. The 2b region contained the Herato0211 scaffold from position 1434133 to the end and the entire scaffold of Herato0212, Herato0213 and Herato0215, with the total size of 1.95 Mb. The 15b region corresponded to Herato1505:1977997-2558395 (580 kb), which included the *cortex* gene (Herato1505:2074108–2087841, ~13.7 kb). Thus, there were 25 chromosomal regions in total (21 chromosomes, with chromosomes 2 and 15 split into three regions). Blocks with fewer than 40 loci were excluded from analysis due to limited information that could result in unreliable inference, except when there was only one block in the chromosomal region.

We assigned an inverse gamma prior (InvG) to the root age τ_0 and population size parameters θ with parameters based on rough estimates from preliminary runs. We used $\theta \sim$ InvG(3, 0.04) for all populations. This prior has mean 0.02, which translated to about 1.7 million individuals. For the divergence times τ, we used $\tau_0 \sim$ InvG(a, b) for the root age, where $a = 3$, $b = 0.06$ for the coding loci and a = 3, b = 0.12 for the noncoding loci, with mean 0.03 (about 2.6 million years) and 0.06 (about 5.2 million years), respectively. The shape parameter a = 3 means that the priors are fairly diffuse, and the prior means are chosen to be roughly consistent with a previous estimate of the divergence time of the clade (Kozak et al. 2015). Given τ_0, the species divergence times for nonroot nodes were assigned a uniform distribution on (0, τ_0) generated from the symmetric Dirichlet distribution. Population-size parameters (θ) are integrated out analytically to improve mixing of the MCMC (Yang 2015).

The MCMC was run for 2×10^6 iterations after 10^5 iterations of burn-in. Samples were recorded every 20 iterations. For each 100-locus block, ten independent runs were performed using different starting species trees. Convergence of the MCMC runs was assessed by checking for consistency of among independent runs. Non-convergent runs were discarded. The samples were then combined to produce the posterior summary such as the maximum a posteriori (MAP) tree (**Tables S4–S5**).

For checking the robustness to the choice of block size, we also performed the same analysis using larger blocks of 200 loci (**Tables S6–S8**).

Both species divergence times (τ) and population size parameters (θ) are in the units of the expected number of mutations per site. To convert them into actual times (before present) and actual population sizes, we used the mutation rate estimate for *H. melpomene* of 2.9×10^{-9} mutations per site per generation and 4 generations per year (Keightley et al. 2015).

Exploring gene flow scenarios using the IM model for species triplets (3s analysis)

In order to formulate hypotheses about gene flow between species for inference of species introgression histories using the MSci model, we estimated pairwise gene flow rates among the species within the *erato* clade using the program 3s (Dalquen et al. 2017). This is a maximum likelihood implementation of the isolation-with-migration (IM) model for three species S_1, S_2 and S_3, assuming the species tree ((S_1, S_2), S_3). It accommodates both ILS and migration and can be applied to large datasets with >10,000 loci, but is limited to only three species and three sequences per locus. Gene flow is only allowed between the two ingroup species (S_1 and S_2) while the third species (S_3) is used as an outgroup to improve parameter estimation. *H. melpomene* was used as an outgroup for all pairs. The outgroup species was added to the dataset by mapping it to the *H. erato demophoon* reference genome.

Since the sequence data were unphased, and 3s requires haploid data, we first phased the loci using PHASE v2.1.1 (Stephens et al. 2001), resulting in two phased diploid sequences per species. The program uses three sequences but allows multiple sequences per species. For each locus, we sampled three sequences of configurations 123, 113 and 223 with probabilities 0.5, 0.25 and 0.25, respectively. Here, 123 means one sequence from each species, etc.

As before, coding and noncoding regions in each chromosomal region were analyzed separately. We also performed analysis using all of the autosomal loci together. For each analysis, we fitted two models: MSC model without gene flow (M0) and with gene flow (M2; isolation-with-migration). There are two divergence time parameters (τ_1 and τ_0) and four effective population sizes (θ_1, θ_2, θ_4, and θ_5) in model M0, and additionally two gene flow rates M_{12} and M_{21} in model M2. For each model, we performed ten independent runs, and the run with the highest log-likelihood value was used. We compared the two models (M0 and M2) using a likelihood ratio test (LRT). Only gene flow scenarios that passed the LRT at the 1% level were considered later.

Strong evidence for pairwise gene flows from this 3S analysis was used together with the patterns of local species/introgression trees from the previous BPP A01 analysis to generate a species tree model with introgression.

Inferring a species tree model with introgression (BPP A00 analysis under the MSci model)

We added likely gene flow events as inferred from the IM model (3S analysis) to the candidate species tree obtained from the MSC model (BPP A01 analysis) to formulate a speciation-tree model with introgression (or the MSci model). There were two main difficulties in this approach. First, the IM model allows for continuous gene flow since the split of the ingroup species S_1 and S_2. This may be unrealistic for many of the species pairs; for example, gene flow might occur immediately after species split but then stop as the two species became more diverged. Furthermore, the MSci model assumes introgression events at specific time points. Second, most pairs of species had strong statistical evidence of gene flow between them. To keep the number of introgression events small and the model manageable, we placed introgression edges between ancestral populations that seemed most likely, for example, when the 3S analysis suggested gene flow between most of their descendant species. In addition, we included introgression that could explain the relationships between different species trees across the genome from the previous MSC model results.

Given a species tree model with introgression, we ran BPP v4.3.0 (Flouri et al. 2020) to estimate the population sizes (θ), species divergence times (τ) and in particular the introgression probabilities (φ). The priors for θ were assigned inverse-gamma priors, InvG(3, 0.01), with mean 0.005. The prior for τ_0 was InvG(3, 0.04). The priors for φ were Beta(4, 2), with mean 0.75. Initial values of φ were set to 0.8 or 0.9. These parameter settings were based on their rough estimates from preliminary runs. We performed inference for each of the 25 chromosomal regions, using either all coding or noncoding loci in each region (the number of loci in each of the regions is listed in **Table S3**).

The MCMC was ran for 10^6 million iterations after a burnin of 10^6 iterations. Ten independent runs were performed and convergence was assessed by examining consistency between runs. Non-convergent runs were discarded. For parameters with a label-switching issue caused by unidentifiability in the model (Flouri et al. 2020), post-processing of the MCMC samples was applied before the samples were combined for posterior summaries.

Acknowledgements

We thank the Broad Institute of Harvard and MIT for funding and carrying out the original sequencing of the *Heliconius* genomes studied here. We thank Nathaniel Edelman, Paul Frandsen, Michael Miyagi, and Bernardo Clavijo for their inputs to this overall project. The current work was supported by Harvard University and by Biotechnology and Biological Sciences Research Council grant (BB/P006493/1 to Z.Y.) and a BBSRC equipment grant (BB/R01356X/1).

Competing interest statement

The authors declare no competing interests.

Data availability

The two multilocus sequence datasets for bpp and 3s analyses are available in Zenodo, at https://doi.org/10.5281/zenodo.4520845.

References

Andermann T, Fernandes AM, Olsson U, Töpel M, Pfeil B, Oxelman B, Aleixo A, Faircloth BC, Antonelli A. 2019. Allele Phasing Greatly Improves the Phylogenetic Utility of Ultraconserved Elements. Syst Biol. 68(1):32–46. doi:10.1093/sysbio/syy039. https://doi.org/10.1093/sysbio/syy039.

Bates HW. 1862. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconinæ. Trans Linn Soc London. 23:495–566.

Van Belleghem SM, Baquero M, Papa R, Salazar C, McMillan WO, Counterman BA, Jiggins CD, Martin SH. 2018. Patterns of Z chromosome divergence among Heliconius species highlight the importance of historical demography. Mol Ecol. 27(19):3852–3872. doi:10.1111/mec.14560. https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14560.

Van Belleghem SM, Cole JM, Montejo-Kovacevich G, Bacquet CN, McMillan WO, Papa R, Counterman BA. 2020. Selection and gene flow define polygenic barriers between incipient butterfly species. bioRxiv.:2020.04.09.034470. doi:10.1101/2020.04.09.034470. http://biorxiv.org/content/early/2020/04/11/2020.04.09.034470.abstract.

Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J, Lewis JJ, Hines HM, et al. 2017. Complex modular architecture around a simple toolkit of wing pattern genes. Nat Ecol Evol. 1(3):52. doi:10.1038/s41559-016-0052. https://doi.org/10.1038/s41559-016-0052.

Beltrán M, Jiggins CD, Brower AVZ, Bermingham E, Mallet J. 2007. Do pollen feeding, pupal-mating and larval gregariousness have a single origin in Heliconius butterflies? Inferences from multilocus DNA sequence data. Biol J Linn Soc. 92(2):221–239. doi:10.1111/j.1095-8312.2007.00830.x. https://doi.org/10.1111/j.1095-8312.2007.00830.x.

Beltrán M, Jiggins CD, Bull V, Linares M, Mallet J, McMillan WO, Bermingham E. 2002. Phylogenetic Discordance at the Species Boundary: Comparative Gene Genealogies Among Rapidly Radiating Heliconius Butterflies. Mol Biol Evol. 19(12):2176–2190. doi:10.1093/oxfordjournals.molbev.a004042. https://doi.org/10.1093/oxfordjournals.molbev.a004042.

Brower AVZ. 1994. Phylogeny of Heliconius Butterflies Inferred from Mitochondrial DNA Sequences (Lepidoptera: Nymphalidae). Mol Phylogenet Evol. 3(2):159–174. doi:https://doi.org/10.1006/mpev.1994.1018. http://www.sciencedirect.com/science/article/pii/S1055790384710189.

Brower AVZ, Egan MG. 1997. Cladistic analysis of Heliconius butterflies and relatives (Nymphalidae: Heliconiiti): a revised phylogenetic position for Eueides based on sequences from mtDNA and a nuclear gene. Proc R Soc London Ser B Biol Sci. 264(1384):969–977. doi:10.1098/rspb.1997.0134. https://doi.org/10.1098/rspb.1997.0134.

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(1):11. doi:10.1186/1741-7007-4-11. https://doi.org/10.1186/1741-7007-4-11.

Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, Suh A, Dutoit L, Bureš S, Garamszegi LZ, et al. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. Genome Res. 25(11):1656–1665. doi:10.1101/gr.196485.115. http://genome.cshlp.org/content/25/11/1656.abstract.

Dalquen DA, Zhu T, Yang AZ. 2017. Maximum likelihood implementation of an isolation-with-migration model for three species. Syst Biol. 66(3):379–398. doi:10.1093/sysbio/syw063.

Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin A V., Salazar C, Ferguson LC, Martin SH, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature. 487(7405):94–98. doi:10.1038/nature11041. https://doi.org/10.1038/nature11041.

Davey JW, Barker SL, Rastas PM, Pinharanda A, Martin SH, Durbin R, McMillan WO, Merrill RM, Jiggins CD. 2017. No evidence for maintenance of a sympatric Heliconius species barrier by chromosomal inversions . Evol Lett. 1(3):138–154. doi:10.1002/evl3.12. https://onlinelibrary.wiley.com/doi/abs/10.1002/evl3.12.

Degnan JH. 2018. Modeling hybridization under the network multispecies coalescent. Syst Biol. 67(5):786–799. doi:10.1093/sysbio/syy040.

DePristo MA, Banks E, Poplin R, Garimella K V., Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43(5):491–501. doi:10.1038/ng.806. http://dx.doi.org/10.1038/ng.806.

Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-Accinelli G, Van Belleghem SM, Patterson N, Neafsey DE, et al. 2019. Genomic architecture and introgression shape a butterfly radiation. Science (80-). 366(6465):594–599. doi:10.1126/science.aaw2090. https://science.sciencemag.org/content/366/6465/594.

Figueiró H V, Li G, Trindade FJ, Assis J, Pais F, Fernandes G, Santos SHD, Hughes GM, Komissarov A, Antunes A, et al. 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci Adv. 3(7):e1700299. doi:10.1126/sciadv.1700299. http://advances.sciencemag.org/content/3/7/e1700299.abstract.

Flouri T, Jiao X, Rannala B, Yang Z. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol Biol Evol. 35(10):2585–2593. doi:10.1093/molbev/msy147.

Flouri T, Jiao X, Rannala B, Yang Z. 2020. A bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. Mol Biol Evol. 37(4):1211–1223. doi:10.1093/molbev/msz296. https://doi.org/10.1093/molbev/msz296.

Fontaine MC, Pease JB, Steele A, Waterhouse RM, Neafsey DE, Sharakhov I V., Jiang X, Hall AB, Catteruccia F, Kakani E, et al. 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science (80-). 347(6217):1258524. doi:10.1126/science.1258524. http://www.sciencemag.org/content/347/6217/1258524.abstract.

Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nat Genet. 43(10):1031–1035. doi:10.1038/ng.937. http://dx.doi.org/10.1038/ng.937.

Heled J, Drummond AJ. 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol Biol Evol. 27(3):570–580. doi:10.1093/molbev/msp274. http://mbe.oxfordjournals.org/content/27/3/570.abstract.

Hey J. 2010. Isolation with migration models for more than two populations. Mol Biol Evol. 27(4):905–920. doi:10.1093/molbev/msp296.

Hey J, Chung Y, Sethuraman A, Lachance J, Tishkoff S, Sousa VC, Wang Y. 2018. Phylogeny estimation by integration over isolation with migration models. Mol Biol Evol. 35(11):2805–2818. doi:10.1093/molbev/msy162.

Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D. persimilis. Genetics. 167(2):747–760. doi:10.1534/genetics.103.024182. http://www.genetics.org/content/167/2/747.abstract.

Hines HM, Counterman BA, Papa R, De Moura PA, Cardoso MZ, Linares M, Mallet J, Reed RD, Jiggins CD, Kronforst MR, et al. 2011. Wing patterning gene redefines the mimetic history of Heliconius butterflies. Proc Natl Acad Sci U S A. 108(49):19666–19671. doi:10.1073/pnas.1110096108. http://www.pnas.org/content/108/49/19666.abstract.

Jiao X, Flouri T, Rannala B, Yang Z. 2020. The Impact of Cross-Species Gene Flow on Species Tree Estimation. Syst Biol. 69(5):830–847. doi:10.1093/sysbio/syaa001. https://doi.org/10.1093/sysbio/syaa001.

Jiggins CD, McMillan WO, King P, Mallet J. 1997. The maintenance of species differences across a Heliconius hybrid zone. Heredity (Edinb). 79(5):495–505. doi:10.1038/hdy.1997.189. https://doi.org/10.1038/hdy.1997.189.

Jiggins CD, McMillan WO, Neukirchen W, Mallet J. 1996. What can hybrid zones tell us about speciation? The case of Heliconius erato and H. himera (Lepidoptera: Nymphalidae). Biol J Linn Soc. 59(3):221–242. doi:10.1111/j.1095-8312.1996.tb01464.x. https://doi.org/10.1111/j.1095-8312.1996.tb01464.x.

Jones GR. 2019. Divergence Estimation in the Presence of Incomplete Lineage Sorting and Migration. Syst Biol. 68(1):19–31. doi:10.1093/sysbio/syy041. https://doi.org/10.1093/sysbio/syy041.

Keightley PD, Pinharanda A, Ness RW, Simpson F, Dasmahapatra KK, Mallet J, Davey JW, Jiggins CD. 2015. Estimation of the Spontaneous Mutation Rate in Heliconius melpomene. Mol Biol Evol. 32(1):239–243. doi:10.1093/molbev/msu302. https://doi.org/10.1093/molbev/msu302.

Kozak KM, Owen McMillan W, Joron M, Jiggins CD. 2018. Genome-wide admixture is common across the Heliconius radiation. bioRxiv.:414201. doi:10.1101/414201. http://biorxiv.org/content/early/2018/09/11/414201.abstract.

Kozak KM, Wahlberg N, Neild AFE, Dasmahapatra KK, Mallet J, Jiggins CD. 2015. Multilocus species trees show the recent adaptive radiation of the mimetic heliconius butterflies. Syst Biol. 64(3):505–524. doi:10.1093/sysbio/syv007. https://doi.org/10.1093/sysbio/syv007.

Kronforst MR. 2008. Gene flow persists millions of years after speciation in Heliconius butterflies. BMC Evol Biol. 8(1):98. doi:10.1186/1471-2148-8-98. https://doi.org/10.1186/1471-2148-8-98.

Kronforst MR, Hansen MEB, Crawford NG, Gallant JR, Zhang W, Kulathinal RJ, Kapan DD, Mullen SP. 2013. Hybridization Reveals the Evolving Genomic Architecture of Speciation. Cell Rep. 5(3):666–677. doi:10.1016/j.celrep.2013.09.042. http://www.sciencedirect.com/science/article/pii/S2211124713005652.

Kronforst MR, Young LG, Blume LM, Gilbert LE. 2006. Multilocus Analyses of Admixture and Introgression Among Hybridizing Heliconius Butterflies. Evolution (N Y). 60(6):1254–1268. doi:10.1111/j.0014-3820.2006.tb01203.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0014-3820.2006.tb01203.x.

Lanier HC, Knowles LL. 2012. Is recombination a problem for species-tree analyses? Syst Biol. 61(4):691–701. doi:10.1093/sysbio/syr128.

Larget BR, Kotha SK, Dewey CN, Ané C. 2010. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. Bioinformatics. 26(22):2910–2911. doi:10.1093/bioinformatics/btq539. http://bioinformatics.oxfordjournals.org/content/26/22/2910.abstract.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 27(21):2987–2993. doi:10.1093/bioinformatics/btr509.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv e-prints. http://arxiv.org/abs/1303.3997.

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(16):2078–2079. doi:10.1093/bioinformatics/btp352.

Liu L, Yu L, Edwards S V. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol Biol. 10(1):302. doi:10.1186/1471-2148-10-302. http://www.biomedcentral.com/1471-2148/10/302.

Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nat Ecol Evol. 2(12):1940–1955. doi:10.1038/s41559-018-0717-x. https://doi.org/10.1038/s41559-018-0717-x.

Mallet J, Beltrán M, Neukirchen W, Linares M. 2007. Natural hybridization in heliconiine butterflies: The species boundary as a continuum. BMC Evol Biol. 7(1):28. doi:10.1186/1471-2148-7-28. https://doi.org/10.1186/1471-2148-7-28.

Mallet J, Besansky N, Hahn MW. 2016. How reticulated are species? BioEssays. 38(2):140–149. doi:10.1002/bies.201500149. http://dx.doi.org/10.1002/bies.201500149.

Mallet J, McMillan WO, Jiggins CD. 1998. Estimating the Mating Behavior of a Pair of Hybridizing Heliconius Species in the Wild. Evolution (N Y). 52(2):503–510. doi:10.2307/2411085. http://www.jstor.org/stable/2411085.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 17(1):10. doi:10.14806/ej.17.1.200. http://journal.embnet.org/index.php/embnetjournal/article/view/200.

Martin SH, Van Belleghem SM. 2017. Exploring evolutionary relationships across the genome using topology weighting. Genetics. 206(1):429–438. doi:10.1534/genetics.116.194720. https://www.genetics.org/content/206/1/429.

Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17(2):e2006288. doi:10.1371/journal.pbio.2006288. https://doi.org/10.1371/journal.pbio.2006288.

Martin SH, Jiggins CD. 2017. Interpreting the genomic landscape of introgression. Curr Opin Genet Dev. 47:69–74. doi:10.1016/j.gde.2017.08.007. http://www.sciencedirect.com/science/article/pii/S0959437X17300357.

Martin SH, Möst M, Palmer WJ, Salazar C, McMillan WO, Jiggins FM, Jiggins CD. 2016. Natural selection and genetic diversity in the butterfly heliconius melpomene. Genetics. 203(1):525–541. doi:10.1534/genetics.115.183285. http://www.genetics.org/content/203/1/525.

Massardo D, Vankuren NW, Nallu S, Ramos RR, Ribeiro PG, Silva-Brandão KL, Brandão MM, Lion MB, Freitas AVL, Cardoso MZ, et al. 2020. The roles of hybridization and habitat fragmentation in the evolution of Brazil's enigmatic longwing butterflies, Heliconius nattereri and H. hermathena. BMC Biol. 18(1):84. doi:10.1186/s12915-020-00797-1. https://doi.org/10.1186/s12915-020-00797-1.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303. doi:10.1101/gr.107524.110. http://genome.cshlp.org/content/20/9/1297.abstract.

McMillan WO, Jiggins CD, Mallet J. 1997. What initiates speciation in passion-vine butterflies? Proc Natl Acad Sci. 94(16):8628–8633. doi:10.1073/pnas.94.16.8628. https://www.pnas.org/content/94/16/8628.

Mirarab S, Reaz R, Bayzid MS, Zimmermann T, S. Swenson M, Warnow T. 2014. ASTRAL: Genome-scale coalescent-based species tree estimation. Bioinformatics. 30(17):i541--i548. doi:10.1093/bioinformatics/btu462.

Müller F. 1879. Ituna and Thyridia; a remarkable case of mimicry in butterflies (transl. by Ralph Meldola from the original German article in Kosmos, May 1879, p. 100). Trans Entomol Soc London.:xx–xxix.

Nadeau NJ, Martin SH, Kozak KM, Salazar C, Dasmahapatra KK, Davey JW, Baxter SW, Blaxter ML, Mallet J, Jiggins CD. 2013. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Mol Ecol. 22(3):814–826. doi:10.1111/j.1365-294X.2012.05730.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05730.x.

Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko S V., Wallbank RWR, Wu GC, Maroja L, Ferguson L, Hanly JJ, et al. 2016. The gene cortex controls mimicry and crypsis in butterflies and moths. Nature. 534(7605):106–110. doi:10.1038/nature17961. https://doi.org/10.1038/nature17961.

Okonechnikov K, Conesa A, García-Alcalde F. 2016. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 32(2):292–294. doi:10.1093/bioinformatics/btv566.

Pardo-Diaz C, Salazar C, Baxter SW, Merot C, Figueiredo-Ready W, Joron M, McMillan WO, Jiggins CD. 2012. Adaptive introgression across species boundaries in Heliconius butterflies. PLoS Genet. 8(6):e1002752. doi:10.1371/journal.pgen.1002752. https://doi.org/10.1371/journal.pgen.1002752.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics. 192(3):1065–1093. doi:10.1534/genetics.112.145037. http://genetics.org/content/192/3/1065.

Payseur BA, Rieseberg LH. 2016. A genomic perspective on hybridization and speciation. Mol Ecol. 25(11):2337–2360. doi:10.1111/mec.13557. https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13557.

Rannala B, Yang Z. 2017. Efficient Bayesian species tree inference under the multispecies coalescent. Syst Biol. 66(5):823–842. doi:10.1093/sysbio/syw119.

Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, Jiggins CD, Chamberlain NL, Kronforst MR, Chen R, et al. 2011. optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. Science (80-). 333(6046):1137–1141. doi:10.1126/science.1208227. https://science.sciencemag.org/content/333/6046/1137.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. Nature. 461(7263):489–494. doi:10.1038/nature08365. http://dx.doi.org/10.1038/nature08365.

Roch S, Steel M. 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theor Popul Biol. 100:56–62. doi:10.1016/j.tpb.2014.12.005.

Rosser N, Phillimore AB, Huertas B, Willmott KR, Mallet J. 2012. Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America. Biol J Linn Soc. 105(3):479–497. doi:10.1111/j.1095-8312.2011.01814.x. https://doi.org/10.1111/j.1095-8312.2011.01814.x.

Salazar C, Jiggins CD, Taylor JE, Kronforst MR, Linares M. 2008. Gene flow and the genealogical history of Heliconius heurippa. BMC Evol Biol. 8(1):132. doi:10.1186/1471-2148-8-132. https://doi.org/10.1186/1471-2148-8-132.

Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, Blazier JC, Sankararaman S, Andolfatto P, Rosenthal GG, et al. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. Science (80-).

360(6389):656–660. doi:10.1126/science.aar3684. https://science.sciencemag.org/content/360/6389/656.

Sousa V, Hey J. 2013. Understanding the origin of species with genome-scale data: Modelling gene flow. Nat Rev Genet. 14(6):404–414. doi:10.1038/nrg3446. http://dx.doi.org/10.1038/nrg3446.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 68(4):978–989. doi:10.1086/319501. http://www.sciencedirect.com/science/article/pii/S0002929707614244.

Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: Fast processing of NGS alignment formats. Bioinformatics. 31(12):2032–2034. doi:10.1093/bioinformatics/btv098.

Taylor SA, Larson EL. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. Nat Ecol Evol. 3(2):170–177. doi:10.1038/s41559-018-0777-y. https://doi.org/10.1038/s41559-018-0777-y.

Wen D, Nakhleh L. 2018. Coestimating reticulate phylogenies and gene trees from multilocus sequence data. Syst Biol. 67(3):439–457. doi:10.1093/sysbio/syx085.

Xu B, Yang Z. 2016. Challenges in species tree estimation under the multispecies coalescent model. Genetics. 204(4):1353–1368. doi:10.1534/genetics.116.190173. http://www.genetics.org/content/204/4/1353.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. Curr Zool. 61(5):854–865. doi:10.1093/czoolo/61.5.854.

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol Biol Evol. 31(12):3125–3135. doi:10.1093/molbev/msu279. http://mbe.oxfordjournals.org/content/31/12/3125.abstract.

Yu Y, Nakhleh L. 2015. A maximum pseudo-likelihood approach for phylogenetic networks. BMC Genomics. 16(Suppl 10):1–10. doi:10.1186/1471-2164-16-S10-S10. http://dx.doi.org/10.1186/1471-2164-16-S10-S10.

Zhang C, Ogilvie HA, Drummond AJ, Stadler T. 2018. Bayesian inference of species networks from multilocus sequence data. Mol Biol Evol. 35(2):504–517. doi:10.1093/molbev/msx307.

Zhu T, Yang Z. 2012. Maximum likelihood implementation of an isolation-with-migration model with three species for testing speciation with gene flow. Mol Biol Evol. 29(10):3131–3142. doi:10.1093/molbev/mss118.

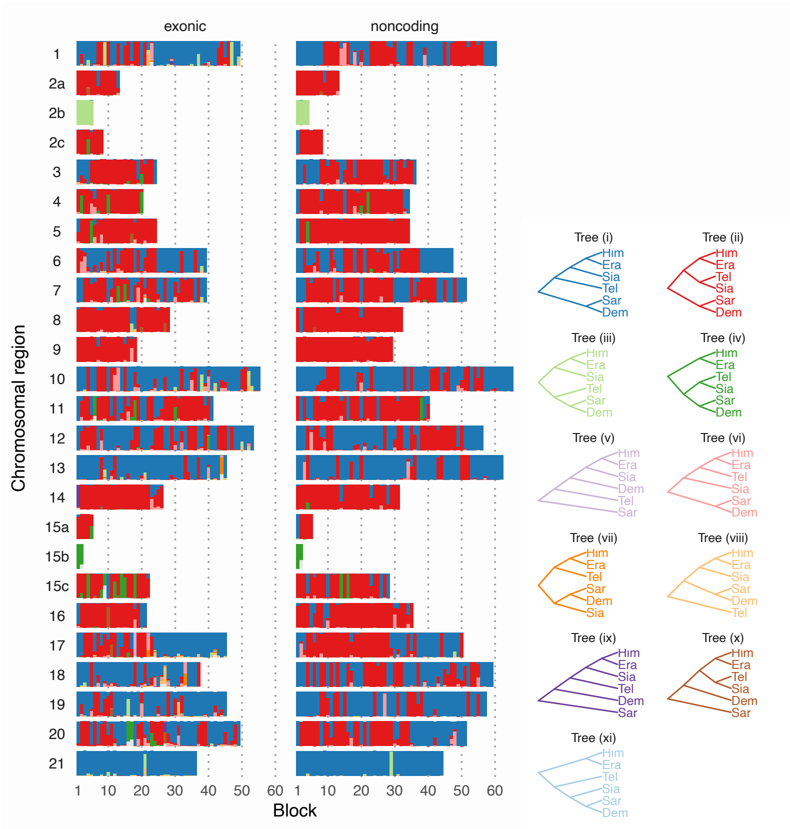

Figure 1. Estimated species trees from blocks of 100 loci across the genome under the multispecies coalescent model. The y-axis represents posterior probability (range from 0 to 1). Trees (i)–(x) in the legend are MAP trees in at least one block. Tree (xi) appeared as one of the top eight trees in the sliding-window analysis of Edelman et al. (2019) but not in our analysis. Era: *H. erato*, Him *H. himera*, Sia: *H. hecalesia*, Tel: *H. telesiphe*, Dem: *H. demeter*, Sar: *H. sara*.

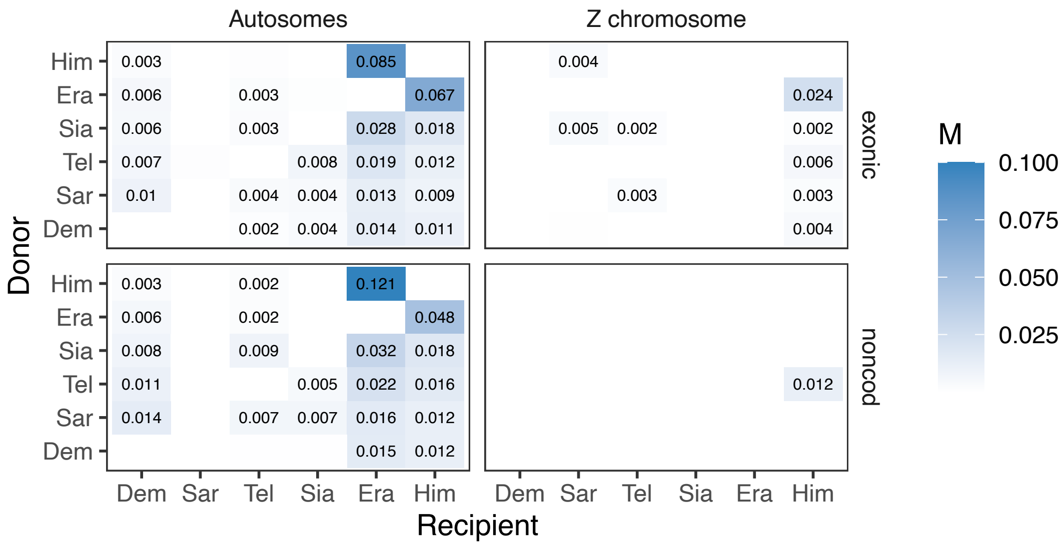

Figure 2. Pairwise migration rate estimates ($M = Nm$) under the IM model from 3s using coding and noncoding loci in the autosomes (using all autosomal loci) and Z chromosome. For each pairwise comparison, the donor and recipient species are given in the y- and x-axis, respectively. *H.*

melpomene was used as an outgroup. Only significant migration rate estimates, with likelihood ratio test at $p < 0.01$ and with values of $M > 0.001$ are shown. See **Table S9** for complete results.

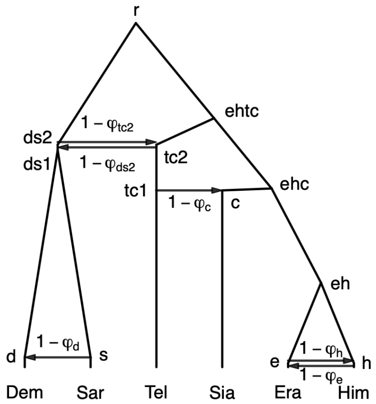

Figure 3. Introgression model. Each horizontal arrow represents a unidirectional introgression event. Timescale drawn here is based on the estimates from all 6,030 noncoding loci in chromosome 1.

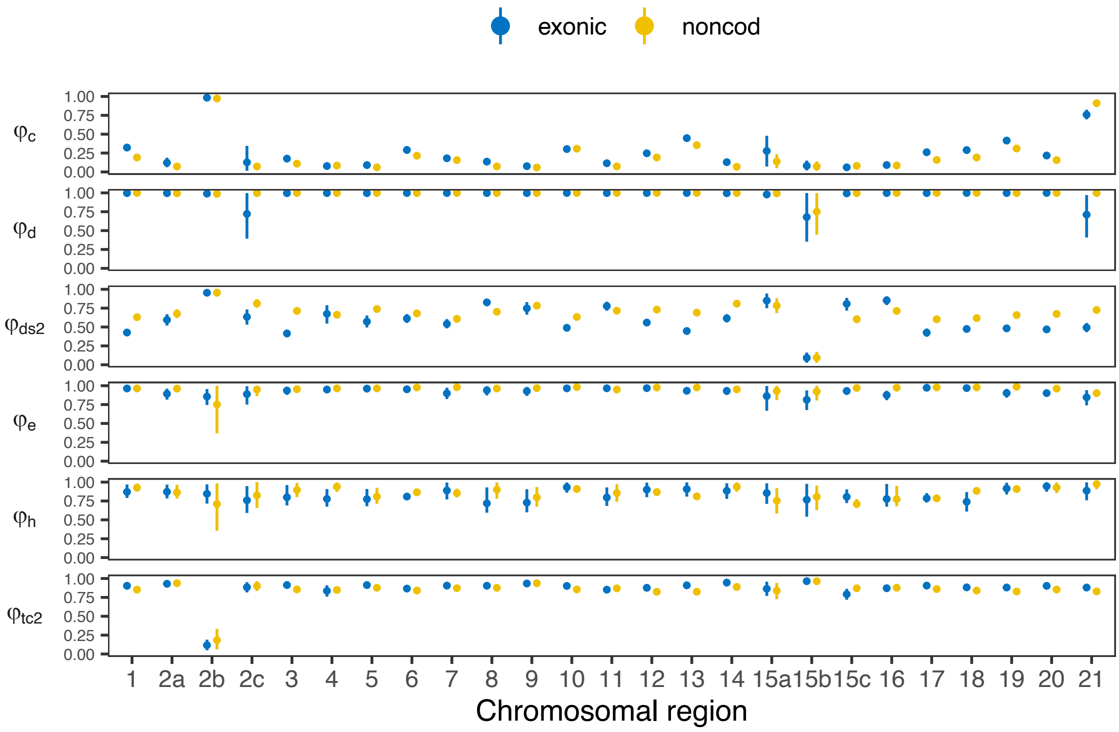

Figure 4. Estimates (posterior mean with 95% HPD interval) of introgression probabilities (φ) under the model of **Figure 3**.

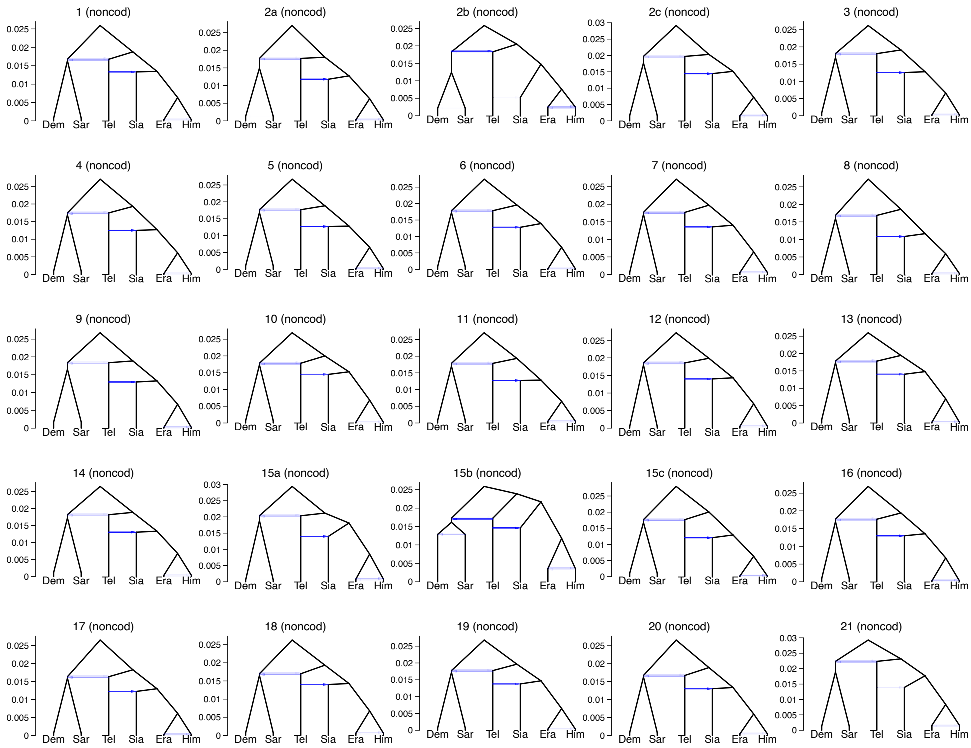

Figure 5. Estimated introgression history of each chromosomal region from noncoding loci. Thickness of the horizontal edges is proportional to the introgression probability estimate. The y-axis is the divergence time. See **Figure S4** for the results for coding loci.

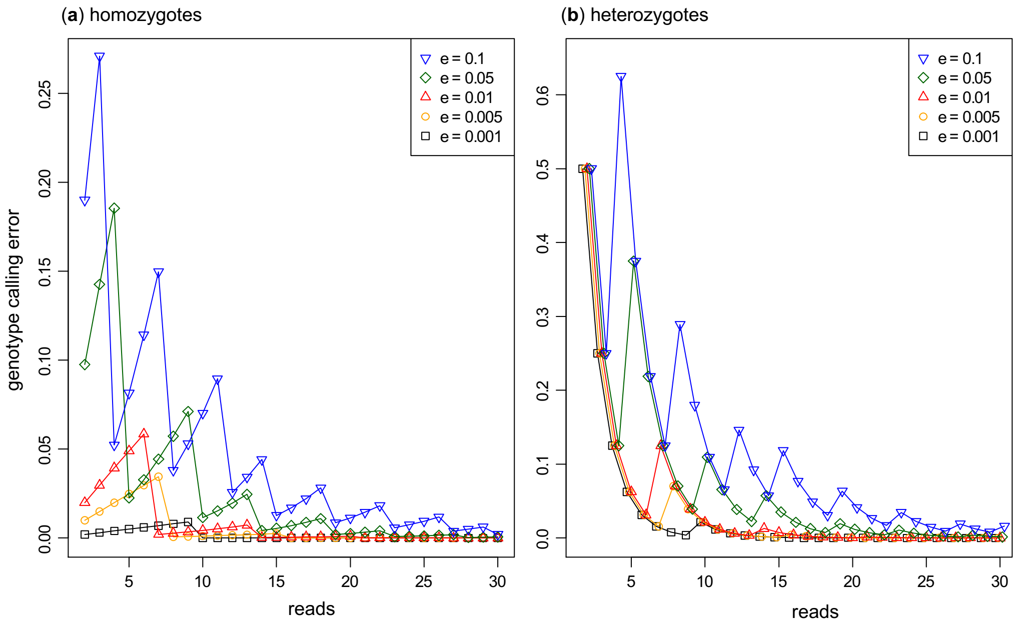

Figure 6. The genotype calling error rate for given base-calling error and read depth when (a) the true genotype is a homozygote and (b) when the true genotype is a heterozygote. Note that for a given base-call error rate, the genotype calling error rate does not increase monotonically with the read depth. The calculation is performed using a C program, implementing the ML method of (Li 2011).