

EVOLUTIONARY BIOLOGY

A macroevolutionary role for chromosomal fusion and fission in *Erebia* butterflies

Hannah Augustijnen^{1*}, Livio Bätcher¹, Martin Cesanek², Tinatin Chkhartishvili³, Vlad Dincă⁴, Giorgi Iankoshvili³, Kota Ogawa^{5,6}, Roger Vila⁷, Seraina Klopstein^{8,9}, Jurriaan M. de Vos¹, Kay Lucek^{1,10*}

The impact of large-scale chromosomal rearrangements, such as fusions and fissions, on speciation is a long-standing conundrum. We assessed whether bursts of change in chromosome numbers resulting from chromosomal fusion or fission are related to increased speciation rates in *Erebia*, one of the most species-rich and karyotypically variable butterfly groups. We established a genome-based phylogeny and used state-dependent birth-death models to infer trajectories of karyotype evolution. We demonstrated that rates of anagenetic chromosomal changes (i.e., along phylogenetic branches) exceed cladogenetic changes (i.e., at speciation events), but, when cladogenetic changes occur, they are mostly associated with chromosomal fissions rather than fusions. We found that the relative importance of fusion and fission differs among *Erebia* clades of different ages and that especially in younger, more karyotypically diverse clades, speciation is more frequently associated with cladogenetic chromosomal changes. Overall, our results imply that chromosomal fusions and fissions have contrasting macroevolutionary roles and that large-scale chromosomal rearrangements are associated with bursts of species diversification.

INTRODUCTION

The evolution of barriers to gene flow is a critical requirement for the progress of speciation (1). Although several barriers may contribute to the process, their relative importance often remains unknown, especially at a macroevolutionary scale (2). Chromosomal speciation theory suggests that large-scale chromosomal rearrangements, such as fusions and fissions, are able to promote speciation. They may allow for the buildup of genetic incompatibilities between lineages either by causing hybrid dysfunction (3, 4) or by suppressing recombination in rearranged sections of the genome (5–7). The relevance of chromosomal speciation has been criticized because of the expected “underdominance” of chromosomal rearrangements, whereby strong hybrid fitness disadvantages ensure that fixation of novel karyotypes is difficult, rendering barrier formation unlikely (5, 6). Conversely, if the effects of chromosomal rearrangements on hybrids were minor, then fixation would be possible, but the resulting barriers would remain shallow (5, 6). These theories were developed for monocentric chromosomes, whereas the chromosomes of several major organismal groups, such as butterflies and sedges (8), are holocentric, i.e., they have centromere-like structures spread across their chromosomes rather than concentrated in a single centromere (9).

Holocentric chromosomes may be more likely to overcome the aforementioned underdominance paradox, as rearranged chromosomes can retain kinetochore functionality and so lead to the creation

of meiotic multivalents that are only slightly deleterious in hybrids (9–11). Crosses between closely related holocentric species with different karyotypes may remain viable (11) and do not necessarily result in reproductive isolation (10, 12), although this may not be the case when sex chromosomes are rearranged (13). In addition, some holocentric clades have evolved mechanisms to facilitate proper chromosome segregation even when chromosomes are rearranged (14, 15), which has been suggested to promote chromosomal speciation (9). Empirical evidence for a link between speciation and chromosomal rearrangements, especially for chromosomal fusions and fissions, is sparse for both mono- and holocentric clades (8, 16). However, the fact that many holocentric groups within plants and invertebrates are very species-rich suggests that chromosomal rearrangements could have driven diversification in some of them (9, 17).

Lepidoptera is one of the largest taxonomic groups with holocentric chromosomes, comprising more than 160,000 species of butterflies and moths (18). While some genera within Lepidoptera are extremely diverse in chromosome numbers, sometimes differing by a count of more than 200 even within a single genus (19, 20), most others have conserved chromosome numbers, often close to the inferred ancestral karyotype [haploid number (n) = 31]. Comparative phylogenetic analyses indicate a positive association between the rate of speciation and karyotype evolution for several of the most karyotypically diverse butterfly genera (16).

Erebia is one of the most speciose of all Palearctic butterfly genera, consisting of around 90 to 100 species that mainly inhabit cold mountainous regions, with the majority of diversity found in Europe, where closely related species often form narrow zones of secondary contact with little gene flow (21, 22). Notably, *Erebia* is also one of the genera with the highest known karyotype diversity among butterflies (23), although this diversity differs between clades within the genus. Most karyotypic variation can be found in the comparatively young *tyndarus* clade (n = 8 to 51) (Fig. 1B and table S1), where phylogenetic relationships have remained unclear (24). As *tyndarus* clade species are mostly found in mountainous regions (23), glacial range expansions

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

¹Department of Environmental Science, University of Basel, 4056 Basel, Switzerland. ²Slovak Entomological Society, Slovak Academy of Sciences, Bratislava 1, Slovakia. ³Institute of Ecology, Ilia State University, Tbilisi 0162, Georgia. ⁴Ecology and Genetics Research Unit, University of Oulu, 90570 Oulu, Finland. ⁵Faculty of Social and Cultural Studies, Kyushu University, Fukuoka 819-0395, Japan. ⁶Insect Sciences and Creative Entomology Center, Kyushu University, Fukuoka 819-0395, Japan. ⁷Institut de Biologia Evolutiva (CSIC-Univ. Pompeu Fabra), 08003 Barcelona, Spain. ⁸Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland. ⁹Life Sciences, Natural History Museum Basel, 4051 Basel, Switzerland. ¹⁰Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland.

*Corresponding author. Email: hannah.augustijnen@unifr.ch (H.A.); kay.lucek@unine.ch (K.L.)

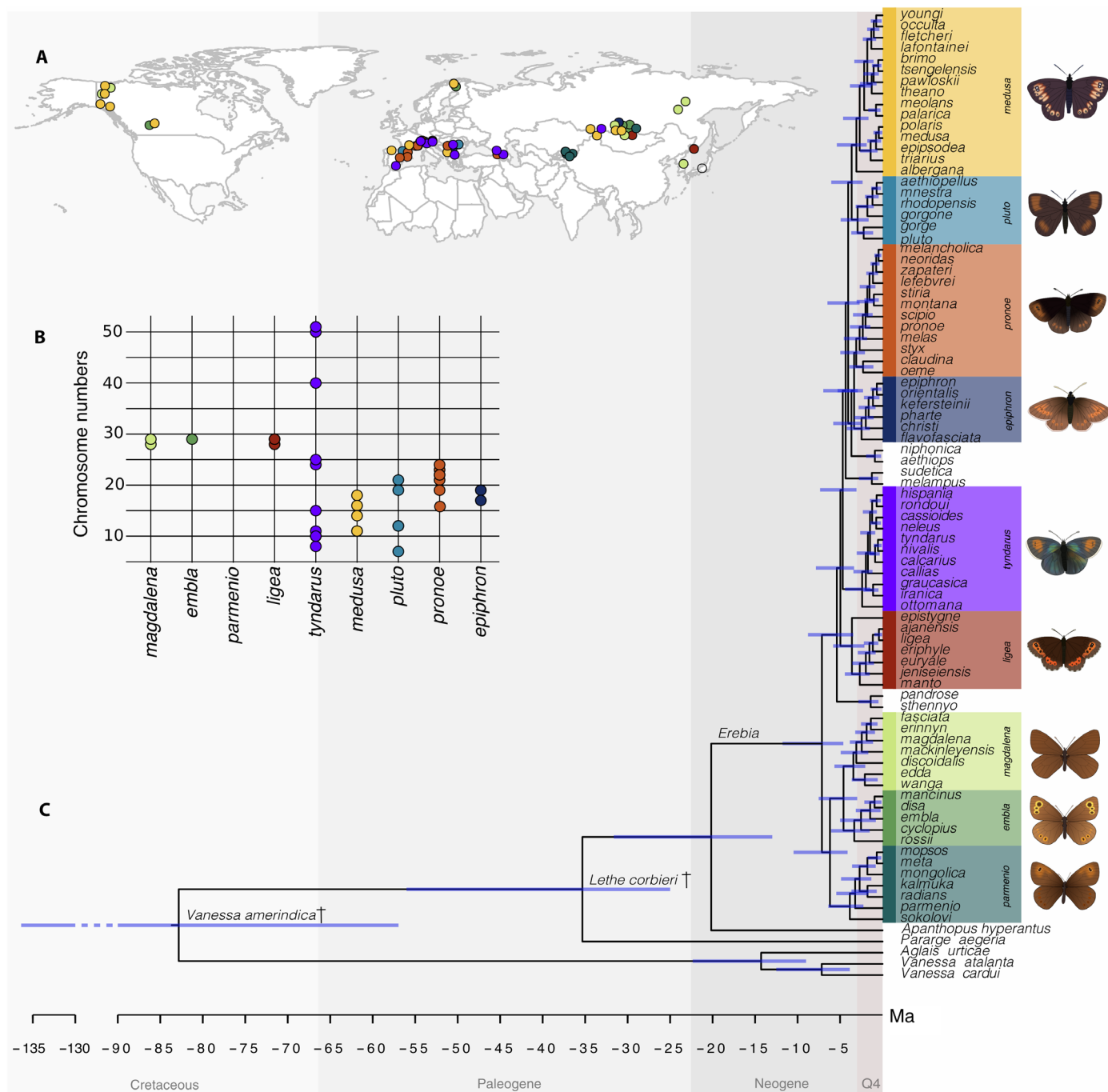


Fig. 1. Sample distribution and relationships within the Palearctic genus *Erebia*. (A) Map of the Northern Hemisphere indicating sampling locations of *Erebia* specimens used, colored by clade. (B) Known chromosome numbers of *Erebia* species, grouped by clade [from (21)]. (C) Time-calibrated phylogeny of *Erebia* calculated in MrBayes. The fossil *V. amerindica* was used to calibrate the root of the tree (i.e., stem lineage of Satyrini), while the fossil *L. corbieri* was placed at the crown node of Satyrini. Clade names are based on (21) with the exception of *medusa* and *pluto*. For each clade a representative phenotype is shown. From top to bottom, these are *Erebia medusa*, *Erebia pluto*, *Erebia pronoe*, *Erebia epiphron*, *Erebia tyndarus*, *Erebia ligea*, *Erebia magdalena*, *Erebia embra*, and *Erebia parmenio*.

and contractions [e.g., (25)] may have led to population bottlenecks and so promoted the fixation of the clade's large variety of rearranged karyotypes through drift, as has been found for other butterflies (26).

Here, we leveraged karyotype diversity across *Erebia* to test for its role in species diversification. Specifically, we first quantified the overall impact of chromosomal fusion and fission on diversification

in *Erebia* using phylogenomic inference and Bayesian state-dependent birth-death models. We then assessed the association between diversification and chromosomal fusion and fission across *Erebia* clades of different ages and karyotype diversity. We hypothesized that chromosomal speciation has played a substantial role in the diversification of *Erebia* and that more karyotypically diverse clades show an

increased signal of chromosomal changes and higher associated speciation rates.

RESULTS

The diversification of *Erebia*

Although approximately 90 to 100 *Erebia* species are recognized, the phylogenetic resolution of former studies on *Erebia* was limited, especially for evolutionary younger taxa, as these studies included either few genes (21) or few species (27). Using whole-genome resequencing data for 82 *Erebia* species, 57% of which are karyotyped (Fig. 1, A and B, and table S1), we constructed a nearly fully resolved species-level coalescent-based phylogeny based on 2920 individual maximum likelihood (ML) gene trees (fig. S1). Branch support was very high overall (>0.9 ASTRAL consensus for 96.5% of all nodes), and we further validated the relationships among taxa following (28) (figs. S2 to S6). We used the resulting topology to constrain a molecular clock dating analysis in MrBayes using fossil calibrations for the stem and crown nodes of Satyrini and a subset of genes selected for minimal missing data, especially among outgroups (Fig. 1C). We confirmed the monophyly of previously defined (21) clades *tyndarus* [2.41 million years (Ma) ago; 95% highest posterior density (HPD) interval, 1.18 to 3.43 Ma ago], *epiphron* (2.47 Ma ago; 95% HPD, 1.49 to 4.21 Ma ago), and *pronoe* (3.10 Ma ago; 95% HPD, 2.07 to 4.98 Ma ago), as well as the classic taxonomic clades (29) *ligea* (3.60 Ma ago; 95% HPD, 2.14 to 5.81 Ma ago), *medusa* (3.08 Ma ago; 95% HPD, 1.87 to 5.38 Ma ago), and *pluto* (2.94 Ma ago; 95% HPD, 1.64 to 4.93 Ma ago). We estimated the age of *Erebia* to be 20.16 Ma old (95% HPD, 12.98 to 31.63 Ma), with the first major split between the mostly non-European *embla*, *magdalena*, and *parmenio* clades and all other *Erebia* at 7.14 Ma ago (95% HPD, 4.62 to 11.77 Ma ago).

Cladogenesis and chromosomes

We fitted a ChromoSSE model (30) to decompose rates of chromosomal fusion and fission into their anagenetic (chromosomal change along a branch) and cladogenetic (chromosomal change at a speciation event) components (Figs. 2 and 3). ChromoSSE infers anagenetic parameters through a continuous-time Markov process based on a Q matrix, which describes instantaneous rates of change in terms of chromosome numbers (30). Cladogenetic parameters, defined as rates of cladogenesis (i) with chromosomal fusions, (ii) with chromosomal fissions, and (iii) without chromosomal change, are estimated via a birth-death process (30). Our models, unless otherwise specified, are based on the hypothesis that chromosomal fusions and fissions evolve both ana- and cladogenetically.

We found that chromosomal changes through anagenetic fusion (0.636 events per species per million years; 95% HPD, 0.134 to 1.118) occurred at a higher rate than anagenetic fission (0.212 events per species per million years; 95% HPD, 0.002 to 0.550; Fig. 3 and table S2). However, most inferred speciation (cladogenetic) events in *Erebia* coincide with chromosomal change, either with cladogenetic chromosomal fusion (0.196 events per species per million years; 95% HPD, 0.074 to 0.340) or with cladogenetic fission (0.328 events per species per million years; 95% HPD, 0.056 to 0.561; Fig. 3), while the speciation rate without chromosomal change was lower (0.127 events per species per million years; 95% HPD, 0.001 to 0.320). The relative extinction rate across *Erebia* was 0.273 events per species per million years (95% HPD: 0.006–0.511) and total speciation (summed speciation rates of cladogenetic fusion, cladogenetic fission, and without

chromosomal change) was 0.651 events per species per million years (95% HPD, 0.471 to 0.839).

We inferred that the karyotype at the root of *Erebia* was likely $n = 22$ or 23 (Fig. 2). To explore the influence of the root value, we repeated this analysis constraining the root karyotype to $n = 29$, which is the modal karyotype of *Erebia* and the chromosome number of its closest relatives (31). We obtained very similar parameter values whether we estimated or constrained the root value, although fixing the root at $n = 29$ led to a higher estimated rate of anagenetic fusion (table S3). Similar results were obtained when constraining the root value to $n = 29$ to 33 (table S3), thereby allowing it to vary around the ancestral karyotype of all butterflies (32). Consequently, by estimating the root value, we obtained conservative estimates.

To examine the scenario where chromosomal fusions and fissions would not contribute to speciation in *Erebia*, we fitted an alternative ChromoSSE model where speciation (cladogenesis) was constrained to be unrelated to chromosomal change, with fusions and fissions evolving only by anagenesis (table S3). Our initial model provides a better fit to the data than this alternative (table S4), again confirming that fusions and fissions evolve both cladogenetically and anagenetically.

Clade-specific chromosomal changes

To test whether ana- or cladogenetic chromosomal change rates would be higher in clades that show a higher karyotype diversity (Fig. 1B), we ran ChromoSSE for the six clades with sufficient chromosome count data and species (Figs. 1B and 4A and table S5). We found that the rates of both ana- and cladogenetic chromosomal change differed across clades under ChromoSSE: Anagenetic fusions differed significantly among clades (Kruskal-Wallis test, $\chi^2_5 = 3303.5$, $P < 0.001$; permutation test with 1000 iterations, $P < 0.001$; Fig. 4B), with posterior distribution means ranging from comparatively low, 0.084 (*ligea*), to moderate, 0.296 (*pronoe*), events per species per million years. This was also true for anagenetic fissions ($\chi^2_5 = 1582.6$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4C), with mean rates between 0.065 (*ligea*) and 0.168 (*pronoe*) events per species per million years. Total anagenetic chromosomal change, defined as the sum of rates for anagenetic fusions and fissions, ranged from 0.148 (*ligea*) to 0.464 (*pronoe*) events per species per million years, which, for all clades, is less than the 0.848 (95% HPD, 0.420 to 1.295) found for the overall ChromoSSE analysis of *Erebia* (table S2).

The rate of cladogenetic chromosomal fusion showed a higher degree of differentiation among clades than other ana- or cladogenetic parameters ($\chi^2_5 = 6958.5$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4F), ranging from 0.327 (*epiphron*) to 1.263 (*tyndarus*) events per species per million years. The rate of cladogenetic chromosomal fission was likewise variable between clades ($\chi^2_5 = 2365.5$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4G), ranging between 0.113 (*ligea*) and 0.594 (*tyndarus*) events per species per million years. The summed rates of cladogenetic fusion and fission ranged from 0.538 (*ligea*) to 1.856 (*tyndarus*) events per species per million years. These rates were generally higher than for overall *Erebia* (0.524 events per species per million years), but the posterior distributions for most clades overlap with the 95% HPD (0.304 to 0.783) of the overall parameter (table S2). The exception was the karyotypically diverse *tyndarus* clade (95% HPD, 0.906 to 2.898), whose rate exceed that of *Erebia*. While the parameter values were very similar between some clades, their proportional contributions to rates of speciation differ (Fig. 4). For example, *medusa*, *pluto*, and *epiphron* show similar rates

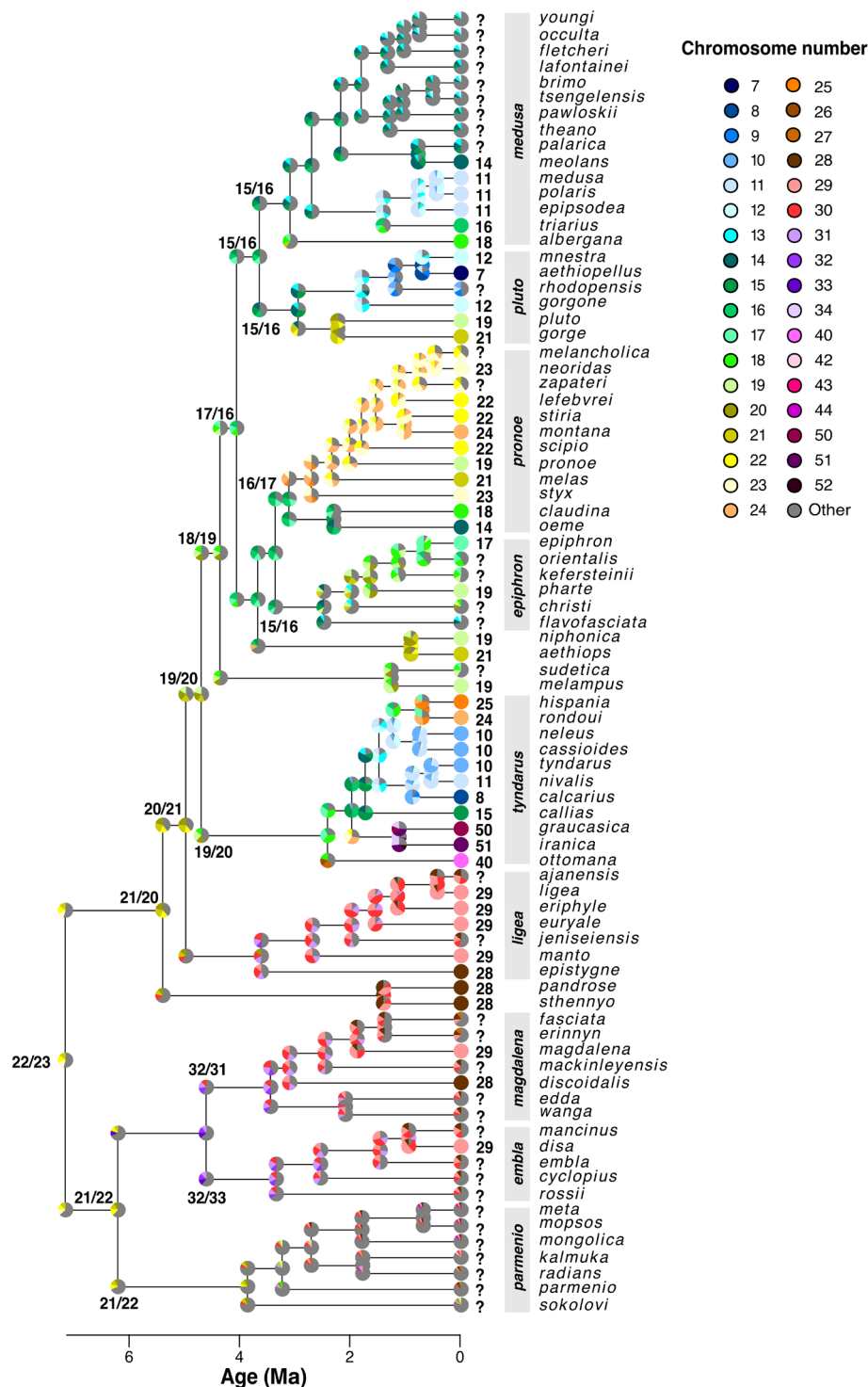


Fig. 2. Summary of the chromosome evolution model for *Erebia*, implemented in ChromoSSE. Estimated ancestral chromosome numbers for *Erebia*, inferred using the phylogenomic topology of fig. S1. Chromosome numbers are indicated proportionally by the color of the pie charts at the branch nodes. The pie charts at the “shoulders” of each node represent the inferred chromosomal state immediately after a speciation event. The two inferred chromosome numbers with highest posterior distribution of the ancestor of each clade are depicted atop or below the node that starts that clade. For deeper nodes within the tree, the reconstructed ancestral character states are likewise presented. For each extant species in the phylogeny, the karyotype is shown before the name. When the karyotype of a species is not known, the species is denoted with a “?”. Clades are named as in Fig. 1. The posterior probabilities of reconstructed ancestral chromosome states are visualized in fig. S7.

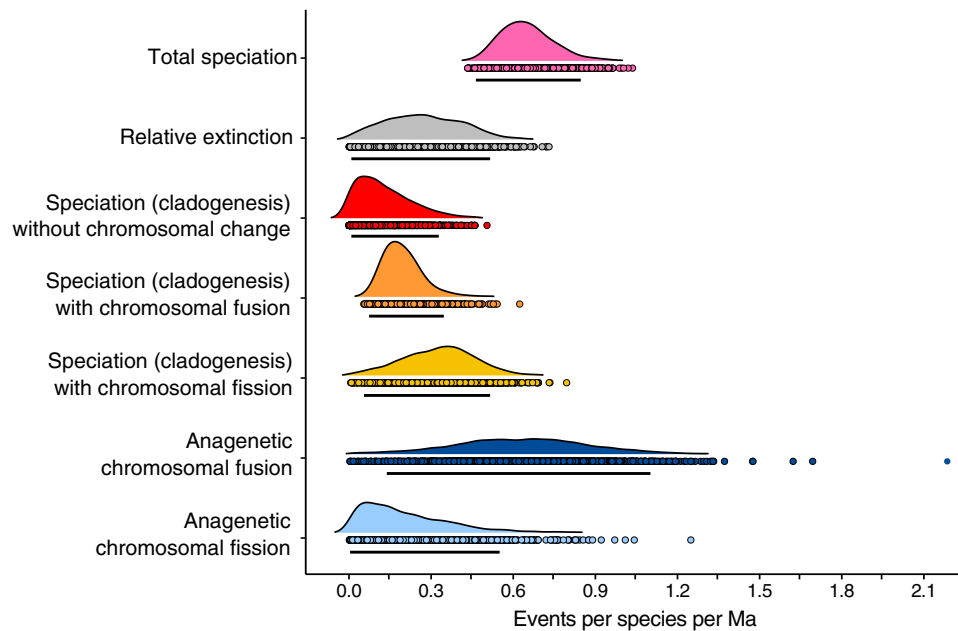


Fig. 3. A summary of the posterior distributions of inferred ChromoSSE parameters for the overall *Erebia* analysis. Shown are the posterior densities per parameter, the raw data points, and the 95% HPD interval as presented by the bold black line underneath each density plot (table S2). All parameter rates are expressed in events per species per million years (Ma).

of cladogenetic chromosomal fissions but they account for 19.8, 18.9, and 30.0% of the total cladogenetic events, respectively.

Speciation rates (cladogenesis) without chromosomal change differed among clades ($\chi^2_5 = 4144.1$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4E) ranging from 0.145 (*pronoe*) to 0.423 (*ligea*) events per species per million years but never accounted for a majority of the total speciation in a clade (though nearly so in *ligea*; Fig. 4, E and F, and table S2).

Relative extinction rates were highest for *tyndarus* at 0.745 and varied between clades ($\chi^2_5 = 5089.6$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4D), with the lowest rates for *pronoe* (0.197) and *ligea* (0.252). Total speciation related directly to rates of cladogenetic events and likewise varied between clades ($\chi^2_5 = 5807.1$, $P < 0.001$; permutation test, $P < 0.001$; Fig. 4H), with *tyndarus* having the highest degree of total speciation (2.191; 95% HPD, 1.283 to 3.363 events per species per million years). The correlation between relative extinction and total speciation was 0.38 ($t_{1,14998} = 50.7$, $P < 0.001$).

We also repeated this analysis while equalizing the maximum chromosome number allowed in each clade ($n = 56$) to minimize the potential effects of differing Q matrices between models. The resulting parameter estimates are largely similar to the models where maximum chromosome numbers were allowed to be tailored to each clade, with all 95% HPD overlapping (table S6). However, running all clades with 56 maximum chromosomes does lead to lower estimates of total speciation (table S6) and to lower estimates of cladogenetic chromosomal fusion. Nonetheless, in all cases, the model with 56 maximum chromosomes was not preferred over the original one (table S7).

As for the overall phylogeny, we fitted alternative ChromoSSE models for each clade, where fusions and fissions were constrained to be only anagenetic. For *epiphron*, *medusa*, *pluto*, and, especially, *tyndarus*, the overall model with combined cladogenetic and anagenetic evolution of chromosome numbers was preferred (table S8). In

pronoe, where anagenetic changes are most frequent of all clades (table S2), but cladogenetic fusions are likewise common, the evidence is mixed, with Akaike's information criterion through Markov chain Monte Carlo (AICM) preferring the model where speciation does not relate to chromosomal change and Bayes factors (BFs) finding no difference between model fits (table S8). For *ligea*, the clade with the least chromosomal variation (Fig. 1B and table S1), the alternative model was preferred (table S8).

A tale of two models

To assess to which degree the combined inference of speciation and extinction rates in ChromoSSE might affect our estimations of cladogenetic and anagenetic parameters, we validated the results of ChromoSSE by carrying out similar analyses in ChromEVOL (33) following (30), which allows to model fusion- and fission-associated cladogenetic changes but does not include an extinction parameter. We expected potential differences between the models in clades with higher extinction and speciation rates. Although the results of ChromoSSE and ChromEvol were largely similar for nearly all clades (table S9, Fig. 4, and fig. S8), the marginal likelihoods and AICMs of ChromEvol models were generally higher, likely owing to a larger parameter space of ChromoSSE compared to ChromEVOL. The difference in marginal likelihood was more pronounced for the overall tree than for individual clades (difference in BFs ranging from 10.002 (*pluto*) to 40.411 (*medusa*); tables S4 and S10). In contrast to the other *Erebia* clades, we found that the ChromoSSE model leads to a better relative fit to the data than the ChromEVOL one (BF = 8.857, AICM = 34.025) for *tyndarus*, the clade with the highest inferred extinction rates (Fig. 4D). An explanation for the difference between ChromoSSE and ChromEVOL could be that simpler models with fewer parameters (ChromEVOL) result in higher marginal likelihoods than more complex models

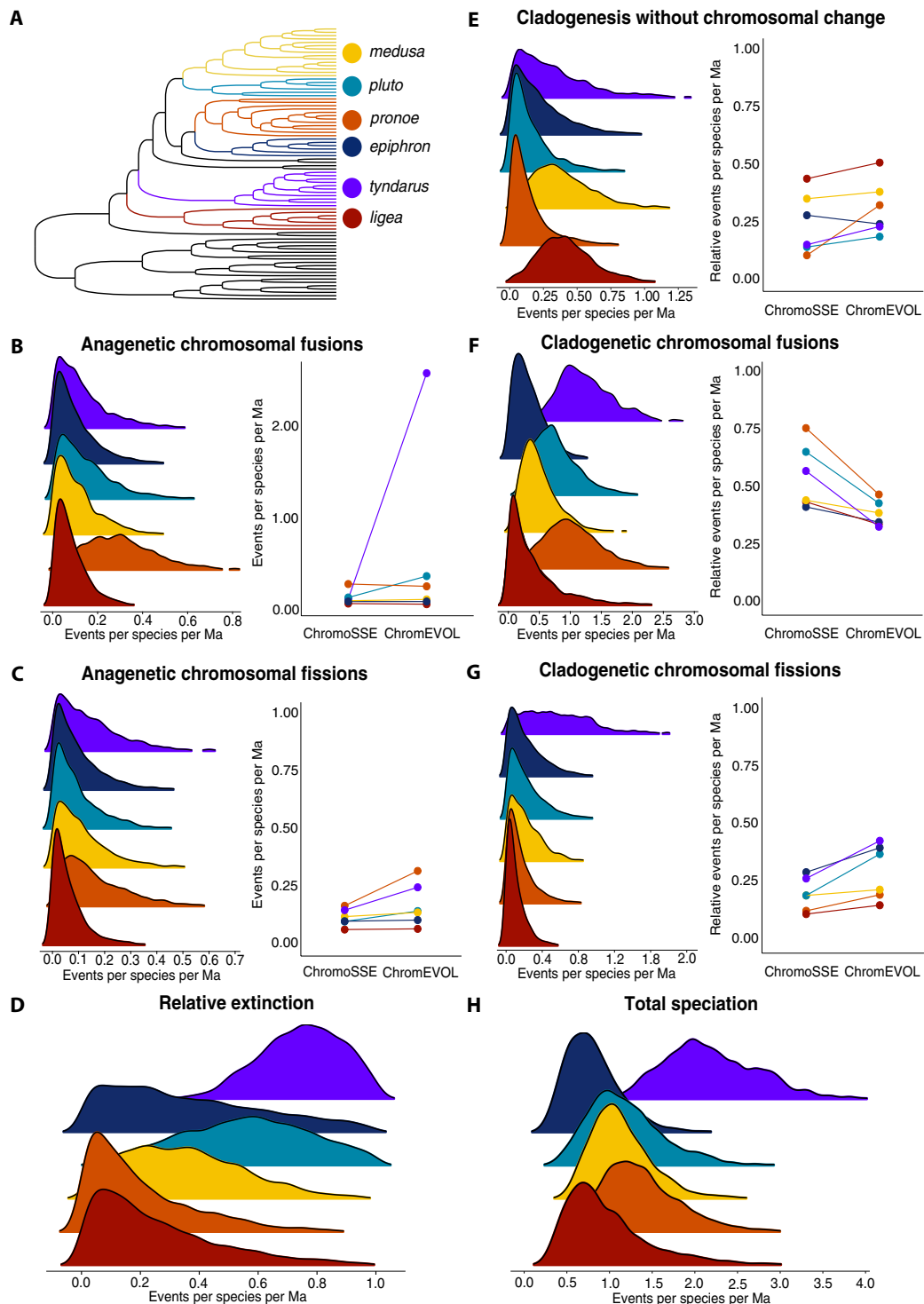


Fig. 4. Summary of ana- and cladogenetic ancestral chromosome estimations for six *Erebia* clades. (A) Overview of the phylogeny of *Erebia* with relevant clades highlighted. (B to H) Distributions of the posterior probability estimates for (B) anagenetic fusions, (C) anagenetic fissions, (D) relative extinction rates, (E) cladogenesis changes unrelated to chromosomal change, (F) cladogenetic chromosomal fusions, (G) cladogenetic chromosomal fissions, and (H) total speciation rates. For (B) and (C) and (E) to (G), comparisons between standardized parameter estimations of clade specific state-dependent birth-death models (ChromoSSE) and birth-death independent models (ChromEVOL) are shown. Each dot indicates the mean of the posterior probability space. All parameter rates are expressed in events per species per million years. An alternative visualization, per clade, not per parameter, is given in fig. S8.

which may suffer from overparameterization, particularly in relatively small datasets. In clades with low to medium inferred relative extinction rates (e.g., *medusa*), the more complex ChromoSSE model was not preferred over the relatively simpler ChromEvol. In *tyndarus*, by contrast, it appears that modeling extinction in tandem with chromosomal evolution as ChromoSSE was designed to do was sufficient to overcome the potential issue of overparameterization.

Because ChromEvol estimates cladogenetic parameters as proportions of the total rate of cladogenetic change, cladogenetic parameters from the ChromoSSE analysis were divided by the mean total speciation rate (Fig. 4G), allowing us to compare rates between models. For the entire tree, rates of cladogenesis without chromosomal change, cladogenetic fusions and cladogenetic fissions overlapped in their posterior distributions, but posterior means differed between the models (Kruskal-Wallis and permutation tests, all $P < 0.001$; fig. S9), in congruence with the relatively high difference between BFs. However, the rates of chromosomal fission-associated speciation remained higher than those for chromosomal fusion under both models (table S9). Rates for anagenetic fissions showed a higher overlap in posterior distribution and did not differ significantly in means ($\chi^2_5 = 0.017$, $P = 0.896$; permutation test, $P = 0.208$; fig. S9). Rates of anagenetic fusions did differ between models ($\chi^2_5 = 2925.2$, $P < 0.001$; permutation test, $P < 0.001$; fig. S9). This difference was most likely driven by *tyndarus* clade (see Fig. 4B) and may be related to its high relative extinction rates.

To assess the effects of missing karyotype data on our inferences, we ran ChromoSSE and ChromEvol again, using similar parameters, on a dataset limited to taxa for which chromosome numbers are available (fig. S10). The results of ChromEvol and ChromoSSE for this dataset were highly congruent with our overall *Erebia* analyses (table S3). For ChromoSSE, however, chromosomal fissions tended to be assigned more as anagenetic (table S3), which is consistent with analyses using less complete phylogenies (34).

DISCUSSION

Compared to other rearrangements, such as inversions, much more remains unknown about the role of chromosomal fusion and fission in speciation (8). Theory predicts that fusions and fissions could have major impacts on the speciation process (3, 5), particularly in species whose chromosomes are holocentric and therefore may not initially be as strongly affected by the underdominance paradox (8, 11, 15). Here, we found phylogenomic evidence for an association of chromosomal fusion and fission with higher speciation rates at macroevolutionary scale, providing indirect evidence for their involvement during speciation. Chromosome-associated cladogenetic events were prevalent in *Erebia* and were associated with higher rates of speciation than speciation without chromosomal change (Fig. 2B), outperforming alternative models without chromosomal change (table S4). Such cladogenetic events coincide with splits between lineages and, in theory, represent cases where chromosomal rearrangements may have a more causative involvement in speciation (30). Alternatively, cladogenetic chromosomal events could represent a direct consequence of speciation if the buildup of reproductive isolation separates diverging karyotypic lineages within the same population (8). However, a causal role for chromosomal fusions and fissions may be especially likely when they themselves act as intrinsic barriers to gene flow, e.g., by resulting in hybrid dysfunction (3, 4), by physically bringing together sites

under selection (35), or if sex chromosomes are involved (13). Chromosomal fusions and fissions may therefore accompany rapid diversification of lineages into species radiations, e.g., (36), for instance, here, in *Erebia* (21).

From a phylogenetic perspective, chromosome-associated speciation should mainly concern cladogenesis, i.e., the splitting of an ancestral species into two new lineages. However, recent theory predicts that anagenetic chromosomal changes could also contribute to speciation by gradually building up differentiation in karyotypes over time, potentially leading to increased reproductive isolation (8). Most changes in chromosome number in the overall *Erebia* phylogeny appear to have occurred anagenetically (Fig. 2). Chromosomal fusion and fission could therefore act as barriers to gene flow, similar to other rearrangements (35), including inversions, e.g., (37). Fusion and fission could suppress recombination within or around rearranged sections of the genome between hybrids (38) and so prevent the breakup of linkage disequilibrium between locally co-adapted genes (5), promoting the gradual accumulation of differences through time (39, 40). Speciation may then be completed through additional processes (41), such as reinforcement upon secondary contact, as has been suggested for other butterflies (42).

Our analyses suggest that the evolutionary impacts of chromosomal fusions and fissions differ, as fusions were more likely to be anagenetic across the entire *Erebia* phylogeny. Conversely, fissions were more likely to be associated with cladogenetic events, leading to higher cladogenetic speciation rates when fissions are involved (Fig. 2). Reductions in chromosome number through fusion events were more common than fissions in *Erebia*, as well as other groups of butterflies (16) and holocentric organisms, e.g., (43). Chromosomal fissions, although rarer, could therefore be associated with higher speciation rates if they more often result in somewhat deleterious meiotic multivalents (16, 44). More studies are needed to confirm this hypothesis, and this effect may be partially mitigated in Lepidoptera due to their holocentric chromosomes (15, 45). Shorter chromosomes have also been shown to have an increased likelihood to be involved in fusions (32), suggesting that longer chromosomes may be more evolutionarily stable. The apparent higher evolutionary stability of fused chromosomes and the instability of fissions could explain why clades with higher chromosome numbers show higher cladogenetic speciation rates in *Erebia* (Figs. 2 and 4) and other butterflies (16). The higher overall degree of anagenetic fusions observed in this study might contribute to the buildup of reproductive isolation by gradually reducing recombination rates, as has been found in mice (46) and other butterflies (47). Genome-wide crossing-over rates correspondingly seem to be considerably higher when chromosomes are short or evolve to become shorter (48), but see (49).

While there are some examples of studies focusing on the macroevolutionary impact of chromosomal fusions and fissions, as well as other rearrangements (50–52), understanding patterns of chromosome-associated speciation may require additional analysis at a finer taxonomic scale (43). Here, we assessed the relative contributions of chromosomal fusion and fission between clades of differing ages and karyotype diversity, to assess the impact of these rearrangements on speciation. We identified a continuum of chromosome-associated speciation rates, ranging from the young, karyotypically very diverse *tyndarus* clade that showed the highest rates of speciation related to chromosomal change (Fig. 4 and fig. S8) to karyotypically more conserved clades, such as *ligea*, where the proportion of speciation unrelated to chromosomal change was much higher (Fig. 4E) and a model

without chromosomal speciation was even preferred (table S8). Other clades fell along this continuum, where rates of speciation with chromosomal fusion were generally high and anagenesis appears to be less important at the subclade level than for the overall tree (Fig. 4). Notably, the clades *pluto* and *pronoe*, which are second to *tyndarus* in chromosome variability (Fig. 1B), also show the lowest proportion of speciation that is not driven by chromosomal fusion and fission (Fig. 4E), although they remain behind *tyndarus* in absolute values (table S2).

The high karyotype diversity and associated high speciation rates of the *tyndarus* clade may, in part, be explained by their ecology. Species of the *tyndarus* clade occur almost exclusively in Alpine areas (23), whereas other *Erebia* clades are ecologically more diverse (24, 53). During glacial cycles, repeated range expansions and contractions across relatively small geographic areas have caused population subdivisions of many *Erebia* species, e.g., (54, 55), including the *tyndarus* clade (25). For the latter, drift and other stochastic processes could have promoted the fixation of novel chromosomal rearrangements (56). Natural selection may also have played a role in the fixation of fusions, although it is not known to which degree (26). Further investigation of butterfly lineages with widely varying chromosome numbers (16, 57) will be required to determine the underlying mechanisms of the fusions and fissions that affect speciation (8).

While state-dependent speciation and extinction (SSE) models allow for unprecedented phylogenetic insights into macroevolutionary aspects of speciation (58–60), their reliability has been partially questioned (61). They may, for example, suffer from excess false-positive rates because shifts in diversification rates across the phylogeny may be assigned to a studied trait in an SSE model even when these shifts are caused by an undetected and unrelated trait (61, 62). This could lead to an overestimation of inferred parameters. However, simulations have indicated that ChromoSSE is more likely to underestimate, rather than overestimate, cladogenetic changes, implying that our inferences are instead rather conservative (30).

Furthermore, as the model cannot detect “cryptic” chromosomal rearrangements, e.g., fusion and fission events that counterbalance each other and, therefore, do not lead to changes in chromosome number, our data may even represent an underestimation of the effects of chromosomal rearrangements on speciation. For example, Mackintosh *et al.* (63) found nine possible fusion and fission events that may have contributed to speciation between two *Brenthis* butterflies that otherwise differ little in their chromosome numbers. Other rearrangements, such as inversions, may similarly affect chromosome evolution and even contribute to speciation (64). However, chromosome-level assemblies for all *Erebia* would be required to study their impact in this system.

We also highlight that a solid taxonomic framework, as presented here, is highly preferable for the correct interpretation of SSE-based inferences (34). Limiting the dataset to only karyotyped species returns similar values for most parameters (table S3) but tends toward higher uncertainty (larger 95% HPD intervals) and overestimation of anagenetic parameters, particularly fissions, which is expected for incomplete phylogenies (34). However, given the taxonomic complexity of *Erebia* (21), it is likely that some cryptic species may exist, again indicating that our estimates of diversification may be rather conservative. In addition, hybridization in *Erebia* is not common, and closely related species tend to form very narrow contact zones with little to no gene flow, e.g., (22, 65), as is the case for other European butterflies (66). Although the possibility for gene flow among

our sequenced individuals in the distant past cannot be fully excluded, such gene flow is unlikely and would not have affected the overall topology of our phylogeny.

We further confirmed the validity of our inferences by fitting ChromEvol (33) models that do not estimate extinction and may thus not suffer the same potential pitfalls as SSE type models. We found very similar results as for the SSE-based analyses (Figs. 2B and 4), with the exception of inferred rates of anagenetic fusions in the *tyndarus* clade, which were estimated to be much higher for ChromEvol. This difference may be due to the influence of the high relative extinction rates in this clade (Fig. 4H), as ChromEvol does not consider any unobserved speciation that may have resulted in extinction, whereas ChromoSSE does (30). We further estimated hidden background speciation and extinction rates without considering chromosomal change using MiSSE (59) and found that, for extant species, these rates do not appear to vary much (Supplementary Methods and figs. S12 and S13). Consequently, our reconstruction of chromosomal change across the deeper *Erebia* tree is unlikely to have been influenced by hidden speciation.

Here, we used state-of-the-art phylogenomic models in one of the most karyologically diverse groups of butterflies, providing evidence for a macroevolutionary impact of major chromosomal rearrangements that equally occur in many other animal, e.g., (51, 67) and plant (50, 52) groups. Overall, we provide evidence that speciation rates are higher with increased chromosomal changes. Similar inferences of the impacts of chromosomal rearrangements are often carried out at higher taxonomic levels or across vast evolutionary timescales, e.g., (68, 69), potentially masking fine-scaled patterns in younger clades. Our study bridges these former investigations and microevolutionary studies that focus on one species or compare sibling species, e.g., (10, 51, 63, 68), by demonstrating within-genus differences of chromosomal fusion- and fission-related speciation. We highlight that chromosomal speciation may be more relevant in clades with more diversity in chromosome numbers. In this genomic era, high-quality reference genomes can be generated and used to build phylogenomic frameworks, which will enable us to further unravel the complexities of chromosomal evolution and speciation.

MATERIALS AND METHODS

Data collection

Adult specimens of 83 *Erebia* species were collected between 2009 and 2021 (table S1). Whenever possible, species were not sampled from regions where they might hybridize with other, closely related *Erebia* species. Bodies were either stored in ethanol at -20°C ($n = 53$) with wings separated or pinned at room temperature ($n = 30$). For the latter, the wings were cut and stored separately before DNA extraction. DNA was extracted from thorax tissue using a Qiagen Blood and Tissue kit (Qiagen AG, Hombrechtikon, Switzerland) following the standard manufacturer's protocol. Paired-end sequencing libraries were constructed at the Department of Biosystems Science and Engineering of ETH Zürich in Basel, followed by sequencing on an Illumina NovaSeq 6000. Samples were sequenced on two S1 flow cells. Species identity of each specimen was confirmed using an *in silico* DNA barcoding approach. In short, we used the pipeline of (70) with standard settings, to extract reads that showed a high k -mer similarity to *Erebia* mitochondrial barcode sequences. Barcode reference sequences for all available species were obtained from the Barcode of Life Database (www.boldsystems.org). For each individual, we then performed a

de novo assembly for the filtered reads with SPAdes 3.13.0 (71) and mapped the contigs back to the barcode database with Yass 1.14 (72). On the basis of the assembled barcode sequence, the species identity of all individuals in our study was correctly assigned.

Demultiplexed raw sequence reads were processed using fastp (73), trimming poly-G tails. Retained reads were mapped to the chromosome-resolved *Erebia ligea* reference genome (74), using bwa v0.7.17 (75). Average mapping coverage was 38% but varied among species, reflecting phylogenetic distance to the reference genome (fig. S11). One species (*Erebia atramentaria*) was omitted because of very low coverage (2.5%). SAMtools v.1.13 (76) was then used to remove unmapped, unpaired, or duplicated reads. A pileup file for each sample was generated with BCFtools v.1.12 (77) mpileup, followed by variant calling in BCFtools call (78). Individual Variant Call Format (VCF) files were then merged and subsequently filtered to remove (i) non-biallelic single-nucleotide polymorphisms (SNPs), (ii) insertions and deletions and adjacent SNPs within 5 base pairs (bp), (iii) SNPs with quality score < 30, (iv) SNPs with more than 80% missing data, (v) SNPs with depths < 4 or > 25, (vi) SNPs with minor allele frequencies (MAFs) < 0.03, and (vii) SNPs falling within repetitive parts of the genome as identified by RepeatMasker 4.0.9 (79). This resulted in a dataset containing 1.87 million SNPs from 82 *Erebia* samples. The raw reads for these samples have been deposited on National Center for Biotechnology Information (NCBI) under the accession number PRJNA1000734.

Outgroup data were taken from chromosome-resolved assemblies of other Nymphalid butterflies that were publicly available at the time of analysis: the Satyrini *Pararge aegeria* (80), *Aphantopus hyperantus* (81), and more distant outgroups *Aglaia urticae* (82), *Vanessa atalanta* (83), and *Vanessa cardui* (84). To obtain single copy orthologs (SCOs) among all assemblies, each assembly including *E. ligea* was annotated with WebAugustus (85, 86) in two steps: First, gene prediction was run using the standard Augustus species parameters for *Heliconius melpomene*. Then, pairwise SCOs between *E. ligea* and all other species were identified with Orthofinder (87). These SCOs were subsequently used for gene prediction in a second run of WebAugustus.

A total of 4505 SCOs among *E. ligea* and our selected outgroups were identified, of which 2920 SCOs with ≥ 20 SNPs were retained for downstream analyses. Exons of each SCO were concatenated into a single coding region to create 2920 separate VCF files extracted from the *Erebia* dataset with BEDtools (88). VCF files were then transposed into FASTA format, and the same exons were extracted for each outgroup species and aligned to the *Erebia* sequence files. Each resulting gene sequence file was aligned with MAFFT v.7.467 (89).

Phylogenomic analyses

ML gene trees were estimated in IQTREE2 (90), which first estimates the optimal substitution model for a gene alignment through ModelFinder (91). Gene trees were estimated with ultrafast bootstrap approximation for 1000 iterations. A species consensus tree was then inferred using coalescent methods in ASTRAL-III (92). Before this analysis, branches with <10% bootstrap support were collapsed to improve the accuracy of ASTRAL-III (92). The resulting species tree had a normalized quartet score of 0.64, indicating 64% agreement between gene tree topologies (reflecting an intermediate amount of incomplete lineage sorting) and high overall support (fig. S1).

The robustness of the phylogeny was further validated following (28), i.e., selecting loci based on phylogenetically informative parameters to reduce incongruency (28, 93, 94). Five validation datasets

were created, each containing the 600 gene trees that (i) had the highest average bootstrap in IQTREE; (ii) showed the highest clocklikeness, indicating how well a gene tree approaches an ultrametric tree formation; (iii) had the lowest CG versus AT contents; (iv) had the highest proportion of parsimony informative sites; and (v) showed the lowest mutation saturation potential. Datasets (i), (ii), and (v) were generated with a modified tree_props.R script from (95). Datasets (iii) and (iv) were generated with FASconCAT (96). The saturation potential was calculated through regression slopes, with lower mutational saturation potential indicating a lower degree of amino acid substitutions and a lower sensitivity to differences in evolutionary rates (93, 97). Consensus topologies for each dataset based on IQTREE gene trees were generated with ASTRAL-III. The average normalized quartet score was similar to the overall dataset with 0.68 (range, 0.61 to 0.77). Comparisons between topologies of the overall dataset and the validation topologies indicated only minor differences concerning relationships within clades and concerned the placement of *Erebia niponica* and *Erebia aethiops*, which are not used in the clade-level analyses. All clades used for further analyses were strongly supported in all validation datasets (figs. S2 to S6). The phylogeny based on the overall dataset, hereafter referred to as the *All gene* topology, was therefore used in downstream analyses to constrain dating analyses using fossil calibration.

Divergence time estimation

Most state-dependent birth-death (SSE) models require ultrametric trees because rates are expressed in relative or absolute units of time. The latter allows for the expression of state-dependent rates of diversification in terms of events per species per million years and is therefore more informative to interpret. We inferred an ultrametric time tree based on the 56 genes with <1% missing data, for a total of 82,923 bp, with the topology constrained according to the *All gene* topology. We chose to minimize missing data for the dating analysis as it occurred primarily in outgroups, which could bias branch lengths and relationships among outgroups, which would, in turn, affect the calibrations, which were placed among the outgroups.

No fossils are known for *Erebia* and few for Satyrini overall. *Lethe corbieri* (98), which has previously been used for dating Satyrini phylogenies (99), was therefore selected. As the two outgroup species within Satyrini, *P. aegeria* and *A. hyperantus*, span the breadth of the clade, their split from *Erebia* coincides with the crown age of Satyrini. Thus, the age of *L. corbieri* (25.0 Ma ago) was taken as a conservative minimal bound for the age of Satyrini, with the median age set at 41.8 Ma ago, following the inference of (99). The root of the tree represents the age of Nymphalidae and stem age of Satyrini, previously estimated to be 69.4 Ma ago (59.0 to 80.2 Ma ago) (100), which was taken as the median age for this node. The minimum age of Nymphalidae was bounded using the fossil *Vanessa amerindica*, as its placement within *Vanessa* is debated (101), and we considered placing this fossil closer to the base of Nymphalidae to be the more conservative approach.

Calculations of the divergence times were carried out by node dating in MrBayes v.3.2.7 (102). PartitionFinder (103) was used to partition the dataset and select the per-partition best substitution models. The invgamma model for among-site rate variation was the best model for all partitions, and we further allowed for integration over the full generalized time-reversible (GTR) substitution model space. Hard constraints were placed on all nodes based on the *All gene* topology, with the clade Satyrini constrained to *Erebia*, *P. aegeria*, and *A. hyperantus* to estimate its node age based on an offset

exponential prior with 25.0 Ma ago as minimum and 41.8 Ma ago as median, while 33.7 Ma ago and 69.4 Ma ago served as the minimum and median bounds of the root. A rooted birth-death process was used to inform the prior probability distribution of branch lengths, strategy of extant taxa was set to random, and sampling probability was set to 0.0145 (87 species of the approximately 6000 Nymphalidae). MrBayes was run for 10,000,000 Markov chain Monte Carlo (MCMC) generations in two independent runs, with four chains each with temperature set to 0.3 and a burn-in of 50%. All model parameters converged to an effective sampling size (ESS) > 200, and potential scale reduction factors approached 1.00.

The root and outgroups were subsequently pruned in R version 3.6.3 (104) using the `extractClade` function of `addTaxa` (105). To allow for comparisons of chromosome-related diversification rates, the six clades of *Erebia* for which chromosome data were available for at least two species (*ligea*, *pronoe*, *medusa*, *pluto*, *epiphron*, and *tyndarus*) were similarly extracted from the overall time-calibrated tree with their internal branch lengths and node ages preserved.

Inference of ancestral chromosome numbers

ChromoSSE (30) was used to jointly infer ancestral chromosome evolution and the phylogenetic birth-death process within the Bayesian framework of RevBayes (106). Chromosome numbers are available for 47 of the 82 included *Erebia* species (57%; table S1), which makes *Erebia* one of the best karyotyped genera of Lepidoptera (16). Species with unknown chromosome numbers were treated as missing data (see Fig. 2). The maximum chromosome number allowed for inference was set at 56, five more than the known maximum of $n = 51$ for *Erebia iranica* to allow for computationally optimal parameter exploration. Polyploidy is rare, if not absent, in Lepidoptera (19, 57, 107); the polyploidization and demi-polyploidization probabilities were set to zero for both the anagenetic and cladogenetic parts of the model. The frequency of each possible chromosome number at the root of *Erebia* was assumed to be unknown and was estimated alongside all other parameters to ensure a more accurate assessment of the root number. To this end, all possible root chromosome numbers were given an equal weight of possibility at the start of the analysis. For the overall analysis, taxon sampling probability (ρ_{bd}) was set to 0.82 as our ultrametric tree included 82 species of the approximately 90 to 100 recorded for *Erebia* (21, 108). Stochastic character mapping of ancestral chromosome numbers was added to the model, with sampling every 10 generations. The inference of ancestral states and the model parameters were run in two independent runs of 10,000 MCMC generations each, sampled every 10 generations, with a burn-in of 20%. These runs were then combined to calculate the final parameter estimates. Convergence of the runs was assessed visually in Tracer v.1.6.0 (109), and the posterior parameter spaces of the runs were combined before further analyses (all ESS > 200 for the combined runs). Priors for anagenetic parameters were established from an exponential distribution centered on 10.0, while priors for cladogenetic parameters were calculated from the number of taxa involved and the root age of the phylogeny following the standard ChromoSSE analysis (30). For our analyses of clades, the same model setup was used as above, with the maximum chromosome number adjusted to the maximum count within that clade +5, but analyses were carried out as single runs, with more generations (25,000 MCMC generations for the *tyndarus* clade, and between 50,000 and 100,000 MCMC generations for all other clades, always with a burn-in of 50%) depending on the size of the clade and

the maximum chromosome count, as these affect computational intensity. ρ_{bd} was set to the percentage of known extant *Erebia* species that was sampled for each clade (between 0.75 and 1.00), and the details of each parameter and run are recorded in table S5. The convergence of each chain in each clade was assessed in Tracer, ensuring an ESS > 200 for all parameters.

To validate the results of ChromoSSE and assess the impact of modeling the birth-death process along with chromosome evolution, a similar analysis was carried out in ChromEvol (33) as modified according to (30), allowing to model cladogenetic changes without extinction. Parameter and burn-in setup were similar to the ChromoSSE setup (table S5).

We ran ChromEvol and ChromoSSE again with the same parameters for only species for which karyotype information is available (47 species; table S1) to assess the impact of missing data in our overall inferences. We pruned species with missing data from the phylogeny with the `drop.tip` function in the package Ape (110) and adjusted the sampling fraction (ρ) of our models to 0.47.

The closest relatives of the genus *Erebia* show a karyotype of $n = 29$ in the few karyotyped species that are known (31). As such, we complemented our overall analyses in which the root chromosome number of *Erebia* was inferred, with two additional ChromoSSE analyses that constrain the root to either $n = 29$ or allow it to vary between $n = 29$ and 33 (table S3). All other parameters were the same as for the overall analysis.

To assess the impact of differing Q matrices as a result of different maximum chromosome numbers in each clade, we repeated our clade-specific ChromoSSE analyses with the maximum chromosome number set to 56 for each clade, in two runs of 10,000 MCMC generations with a burn-in of 20%. The results of both runs were combined for each clade (see table S7).

Last, we also examined the alternative hypothesis that all speciation might be unrelated to chromosomal change and compared this to our full model with cladogenetic fusions and fissions. To this end, we ran a ChromoSSE model in which speciation (cladogenesis) was constrained to be unrelated to chromosomal change, with no cladogenetic fusions or fissions allowed, and changes in chromosome number being always anagenetic. We ran this model for the overall *Erebia* tree in two runs of 10,000 MCMC generations and once for each clade for 25,000 generations, allowing a burn-in of 50%.

Model assessment and clade comparisons

We assessed the difference in model fit between the ChromoSSE and ChromEvol models for the overall *Erebia* tree and, for each clade, by calculating AICM and log BFs in Tracer v.1.6.0 and interpreting their values following (111) (tables S4 and S10). The alternative hypothesis models, in which all cladogenesis is unrelated to chromosomal change, were likewise compared to the full models using these estimators (tables S4 and S8). Path sampling or stepping stone sampling, although known to improve the estimation of model fit as compared to AICM and harmonic mean estimates of BFs (112), was not implemented here due to computational limitations.

To compare the estimated chromosomal parameters between clades, posterior parameter means were first computed from the burned-in posterior parameter spaces of the ChromoSSE and ChromEvol analyses separately, along with the 95% HPDs of each parameter. Posterior parameter distributions from ChromoSSE were then compared among clades using nonparametric Kruskal-Wallis tests, as well as permutation independence tests with 1000 iterations using the R package `coin`

(113) As ChromoSSE and ChromEvol calculate most of the same parameters for the same phylogeny, we further quantified the differences between posterior distribution spaces of these models using the same tests. However, as ChromEvol estimates cladogenetic parameters as proportions of the total amount of cladogenetic change, cladogenetic parameters from the ChromoSSE analysis were first divided by the mean of the total speciation rate, thereby obtaining standardized parameters, allowing us to compare events per species per million years across models.

Supplementary Materials

This PDF file includes:

Supplementary Methods

Figs. S1 to S13

Legends for tables S1 to S12

References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S12

REFERENCES AND NOTES

1. J. A. Coyne, H. A. Orr, *Speciation*. (Sinauer Associates, 2004).
2. L. Zhang, R. Reifová, Z. Halenková, Z. Gompert, How important are structural variants for speciation? *Genes* **12**, 1084 (2021).
3. M. J. D. White, Models of speciation. *Science* **159**, 1065–1070 (1968).
4. M. King, *Species Evolution: The Role of Chromosome Change* (Cambridge Univ. Press, 1995).
5. R. Faria, A. Navarro, Chromosomal speciation revisited: Rearranging theory with pieces of evidence. *Trends Ecol. Evol.* **25**, 660–669 (2010).
6. A. Navarro, N. H. Barton, Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. *Evolution* **57**, 447–459 (2003).
7. L. H. Rieseberg, Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**, 351–358 (2001).
8. K. Lucek, H. Augustijn, M. Escudero, A holocentric twist to chromosomal speciation? *Trends Ecol. Evol.* **37**, 655–662 (2022).
9. D. P. Melters, L. V. Paliulis, I. F. Korf, S. W. L. Chan, Holocentric chromosomes: Convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Res.* **20**, 579–593 (2012).
10. K. H. Hora, F. Marec, P. Roessingh, S. B. J. Menken, Limited intrinsic postzygotic reproductive isolation despite chromosomal rearrangements between closely related sympatric species of small ermine moths (Lepidoptera: Yponomeutidae). *Biol. J. Linn. Soc.* **128**, 44–58 (2019).
11. V. A. Lukhtanov, V. Dincă, M. Friberg, R. Vila, C. Wiklund, Incomplete sterility of chromosomal hybrids: Implications for karyotype evolution and homoploid hybrid speciation. *Front. Genet.* **11**, 583827 (2020).
12. M. Escudero, J. A. Weber, A. L. Hipp, Species coherence in the face of karyotype diversification in holocentric organisms: The case of a cytogenetically variable sedge (*Carex scoparia*, Cyperaceae). *Ann. Bot.* **112**, 515–526 (2013).
13. L. Z. Carabajal Paladino, I. Provazniková, M. Berger, C. Bass, N. S. Aratchige, S. N. López, F. Marec, P. Nguyen, Sex chromosome turnover in moths of the diverse superfamily Gelechioidea. *Genome Biol. Evol.* **11**, 1307–1319 (2019).
14. P. S. Maddox, K. Oegema, A. Desai, I. M. Cheeseman, "Holo"er than thou: Chromosome segregation and kinetochore function in *C. elegans*. *Chromosome Res.* **12**, 641–653 (2004).
15. V. A. Lukhtanov, V. Dincă, M. Friberg, J. Šíchová, M. Olofsson, R. Vila, F. Marec, C. Wiklund, Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E9610–E9619 (2018).
16. J. M. de Vos, H. Augustijn, L. Batscher, K. Lucek, Speciation through chromosomal fusion and fission in Lepidoptera. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190539 (2020).
17. M. Mandrioli, G. C. Manicardi, Holocentric chromosomes. *PLOS Genet.* **16**, e1008918 (2020).
18. J. F. Barrera, J. L. Capinera, Eds., *Encyclopedia of Entomology* (Springer, ed. 2, 2008).
19. V. A. Lukhtanov, The blue butterfly *Polyommatus (Plebicula) atlanticus* (Lepidoptera, Lycaenidae) holds the record of the highest number of chromosomes in the non-polyploid eukaryotic organisms. *Comp. Cytogenet.* **9**, 683–690 (2015).
20. M. Wiemers, "Chromosome differentiation and the radiation of the butterfly subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: *Polyommatus*): A molecular phylogenetic approach," thesis, Bonn University, Bonn, Germany (2003).
21. C. Peña, H. Wittthauer, I. Klečková, Z. Fric, N. Wahlberg, Adaptive radiations in butterflies: Evolutionary history of the genus *Erebia* (Nymphalidae: Satyrinae). *Biol. J. Linn. Soc.* **116**, 449–467 (2015).
22. H. Augustijn, T. Patsiou, K. Lucek, Secondary contact rather than coexistence—*Erebia* butterflies in the Alps. *Evolution* **76**, 2669–2686 (2022).
23. J. Albre, C. Gers, L. Legal, Taxonomic notes on the species of the *Erebia tyndarus* group (Lepidoptera, Nymphalidae, Satyrinae). *Lépidoptères* **17**, 12–28 (2008).
24. J. Albre, C. Gers, L. Legal, Molecular phylogeny of the *Erebia tyndarus* (Lepidoptera, Rhopalocera, Nymphalidae, Satyrinae) species group combining *CoxII* and *ND5* mitochondrial genes: A case study of a recent radiation. *Mol. Phylogenet. Evol.* **47**, 196–210 (2008).
25. T. Schmitt, D. Louy, E. Zimmermann, J. C. Habel, Species radiation in the Alps: Multiple range shifts caused diversification in Ringlet butterflies in the European high mountains. *Organ. Divers. Evol.* **16**, 791–808 (2016).
26. A. Mackintosh, R. Vila, S. H. Martin, D. Setter, K. Lohse, Do chromosome rearrangements fix by genetic drift or natural selection? Insights from *Brenthis* butterflies. *Mol. Ecol.* 10.1111/mec.17146, (2023).
27. K. Lucek, Evolutionary mechanisms of varying chromosome numbers in the radiation of *Erebia* butterflies. *Genes* **9**, 166 (2018).
28. F. V. Freitas, M. G. Branstetter, T. Griswold, E. A. Almeida, Partitioned gene-tree analyses and gene-based topology testing help resolve incongruence in a phylogenomic study of host-specialist bees (Apidae: Eucerinae). *Mol. Biol. Evol.* **38**, 1090–1100 (2021).
29. B. C. Warren, *Monograph of the Genus Erebia* (British Museum of Natural History, 1936).
30. W. A. Freyman, S. Höhna, Cladogenetic and anagenetic models of chromosome number evolution: A Bayesian model averaging approach. *Syst. Biol.* **67**, 195–215 (2018).
31. M. Wiemers, N. Chazot, C. W. Wheat, O. Schweiger, N. Wahlberg, A complete time-calibrated multi-gene phylogeny of the European butterflies. *Zookeys* **938**, 97–124 (2020).
32. V. Ahola, R. Lehtonen, P. Somervuo, L. Salmela, P. Koskinen, P. Rastas, N. Välimäki, L. Paulin, J. Kvist, N. Wahlberg, J. Tanskanen, E. A. Hornett, L. C. Ferguson, S. Luo, Z. Cao, M. A. de Jong, A. Duploup, O.-P. Smolander, H. Vogel, R. C. Mc Coy, K. Qian, W. S. Chong, Q. Zhang, F. Ahmad, J. K. Haukka, A. Joshi, J. Salojärvi, C. W. Wheat, E. Grosse-Wilde, D. Hughes, R. Katainen, E. Pitkänen, J. Ylinen, R. M. Waterhouse, M. Turunen, A. Vähärautio, S. P. Ojanen, A. H. Schulman, M. Taipale, D. Lawson, E. Ukkonen, V. Mäkinen, M. R. Goldsmith, L. Holm, P. Auvinen, M. J. Frilander, I. Hanski, The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat. Commun.* **5**, 4734 (2014).
33. L. Glick, I. Mayrose, ChromEvol: Assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Mol. Biol. Evol.* **31**, 1914–1922 (2014).
34. P. Mynard, A. C. Algar, L. T. Lancaster, G. Bocedi, F. Fahri, C. Gubry-Rangin, P. Lupiyaningdyah, M. Nangoy, O. G. Osborne, A. S. T. Papadopoulos, I. M. Sudiana, B. Juliandi, J. M. J. Travis, L. Herrera-Alsina, Impact of phylogenetic tree completeness and mis-specification of sampling fractions on trait dependent diversification models. *Syst. Biol.* **72**, 106–119 (2023).
35. M. Joron, L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M. Becuwe, S. W. Baxter, L. Ferguson, P. A. Wilkinson, C. Salazar, C. Davidson, R. Clark, M. A. Quail, H. Beasley, R. Glithero, C. Lloyd, S. Sims, M. C. Jones, J. Rogers, C. D. Jiggins, R. H. French-Constant, Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–206 (2011).
36. J. Auvinet, P. Graça, L. Belkadi, L. Petit, E. Bonnard, A. Dettai, W. H. Detrich, C. Ozouf-Costaz, D. Higuier, Mobilization of retrotransposons as a cause of chromosomal diversification and rapid speciation: The case for the Antarctic teleost genus *Trematomus*. *BMC Genomics* **19**, 339 (2018).
37. A. D. Twyford, J. Friedman, Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* **69**, 1476–1486 (2015).
38. K. Yoshida, C. Rödelberger, W. Rösler, M. Riebesell, S. Sun, T. Kikuchi, R. J. Sommer, Chromosome fusions repatterned recombination rate and facilitated reproductive isolation during *Pristionchus* nematode speciation. *Nat. Ecol. Evol.* **7**, 424–439 (2023).
39. J. L. Feder, P. Nosil, S. M. Flaxman, Assessing when chromosomal rearrangements affect the dynamics of speciation: Implications from computer simulations. *Front. Genet.* **5**, 295 (2014).
40. S. M. Flaxman, A. C. Wacholder, J. L. Feder, P. Nosil, Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Mol. Ecol.* **23**, 4074–4088 (2014).
41. J. Kulmuni, R. K. Butlin, K. Lucek, V. Savolainen, A. M. Westram, Towards the completion of speciation: The evolution of reproductive isolation beyond the first barriers. *Philosop. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190528 (2020).
42. V. A. Lukhtanov, N. P. Kandul, J. B. Plotkin, A. V. Dantchenko, D. Haig, N. E. Pierce, Reinforcement of pre-zygotic isolation and karyotype evolution in *Agrodiaetus* butterflies. *Nature* **436**, 385–389 (2005).
43. J. I. Márquez-Corro, S. Martín-Bravo, D. Spalink, M. Luceño, M. Escudero, Inferring hypothesis-based transitions in clade-specific models of chromosome number evolution in sedges (Cyperaceae). *Mol. Phylogenet. Evol.* **135**, 203–209 (2019).

44. R. J. Baker, J. W. Bickham, Speciation by monobrachial centric fusions. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8245–8248 (1986).
45. E. G. Pringle, S. W. Baxter, C. L. Webster, A. Papanicolaou, S. F. Lee, C. D. Jiggins, Synteny and chromosome evolution in the Lepidoptera: Evidence from mapping in *Heliconius melpomene*. *Genetics* **177**, 417–426 (2007).
46. C. J. Bidau, M. D. Giménez, C. L. Palmer, J. B. Searle, The effects of Robertsonian fusions on chiasma frequency and distribution in the house mouse (*Mus musculus domesticus*) from a hybrid zone in northern Scotland. *Heredity* **87**, 305–313 (2001).
47. F. Cicconardi, J. J. Lewis, S. H. Martin, R. D. Reed, C. G. Danko, S. H. Montgomery, Chromosome fusion affects genetic diversity and evolutionary turnover of functional loci but consistently depends on chromosome size. *Mol. Biol. Evol.* **38**, 4449–4462 (2021).
48. Q. Haenel, T. G. Laurentino, M. Roesti, D. Berner, Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol. Ecol.* **27**, 2477–2497 (2018).
49. A. P. I. Torres, L. Höök, K. Näsvall, D. Shipilina, C. Wiklund, R. Vila, P. Pruischer, N. Backström, The fine-scale recombination rate variation and associations with genomic features in a butterfly. *Genome Res.* **33**, 810–823 (2023).
50. A. L. Hipp, P. E. Rothrock, R. Whitkus, J. A. Weber, Chromosomes tell half of the story: The correlation between karyotype rearrangements and genetic diversity in sedges, a group with holocentric chromosomes. *Mol. Ecol.* **19**, 3124–3138 (2010).
51. S. Potter, J. G. Bragg, R. Turakulov, M. D. Eldridge, J. Deakin, M. Kirkpatrick, R. J. Edwards, C. Moritz, Limited introgression between rock-wallabies with extensive chromosomal rearrangements. *Mol. Biol. Evol.* **39**, msab333 (2022).
52. M. Escudero, M. Hahn, B. H. Brown, K. Lueders, A. L. Hipp, Chromosomal rearrangements in holocentric organisms lead to reproductive isolation by hybrid dysfunction: The correlation between karyotype rearrangements and germination rates in sedges. *Am. J. Bot.* **103**, 1529–1536 (2016).
53. I. Klečková, J. Klečka, Z. F. Fric, M. Česánek, L. Dutoit, L. Pellissier, P. Matos-Maraví, Climatic niche conservatism and ecological diversification in the Holarctic cold-dwelling butterfly Genus *Erebia*. *Insect Syst. Divers.* **7**, 2 (2023).
54. F. Cupedo, Reproductive isolation and intraspecific structure in Alpine populations of *Erebia euryale* (Esper, 1805) (Lepidoptera, Nymphalidae, Satyrinae). *Nota Lepidopterol.* **37**, 19–36 (2014).
55. J. Besold, T. Schmitt, More northern than ever thought: Refugia of the Woodland Ringlet butterfly *Erebia medusa* (Nymphalidae: Satyrinae) in Northern Central Europe. *J. Zool. Syst. Evol. Res.* **53**, 67–75 (2015).
56. P. A. Martinez, U. P. Jacobina, R. V. Fernandes, C. Brito, C. Penone, T. F. Amado, C. R. Fonseca, C. J. Bidau, A comparative study on karyotypic diversification rate in mammals. *Heredity* **118**, 366–373 (2017).
57. C. J. Wright, L. Stevens, A. Mackintosh, M. Lawniczak, M. Blaxter, Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera. *Nat. Ecol. Evol.* 10.1038/s41559-024-02329-4, (2024).
58. M. G. Harvey, G. A. Bravo, S. Claramunt, A. M. Cuervo, G. E. Derryberry, J. Battilana, G. F. Seeholzer, J. S. McKay, B. C. O'Meara, B. C. Faircloth, The evolution of a tropical biodiversity hotspot. *Science* **370**, 1343–1348 (2020).
59. T. Vasconcelos, B. C. O'Meara, J. M. Beaulieu, A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution* **76**, 1420–1433 (2022).
60. D. S. Caetano, B. C. O'Meara, J. M. Beaulieu, Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* **72**, 2308–2324 (2018).
61. B. C. O'Meara, J. M. Beaulieu, Past, future, and present of state-dependent models of diversification. *Am. J. Bot.* **103**, 792–795 (2016).
62. D. L. Rabosky, E. E. Goldberg, Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* **64**, 340–355 (2015).
63. A. Mackintosh, R. Vila, D. R. Laetsch, A. Hayward, S. H. Martin, K. Lohse, Chromosome fissions and fusions act as barriers to gene flow between *Brenthis* fritillary butterflies. *Mol. Biol. Evol.* **40**, msad043 (2023).
64. E. A. Pazhenkova, V. A. Lukhtanov, Whole-genome analysis reveals the dynamic evolution of holocentric chromosomes in Satyrine butterflies. *Genes* **14**, 437 (2023).
65. T. Schmitt, P. Müller, Limited hybridization along a large contact zone between two genetic lineages of the butterfly *Erebia medusa* (Satyrinae, Lepidoptera) in Central Europe. *J. Zool. Syst. Evol. Res.* **45**, 39–46 (2007).
66. S. Ebdon, D. R. Laetsch, L. Dapporto, A. Hayward, M. G. Ritchie, V. Dincă, R. Vila, K. Lohse, The Pleistocene species pump past its prime: Evidence from European butterfly sister species. *Mol. Ecol.* **30**, 3575–3589 (2021).
67. K. Wellband, C. Mérot, T. Linnansaari, J. A. K. Elliott, R. A. Curry, L. Bernatchez, Chromosomal fusion and life history-associated genomic variation contribute to within-river local adaptation of Atlantic salmon. *Mol. Ecol.* **28**, 1439–1459 (2019).
68. J. Hill, P. Rastas, E. A. Hornett, R. Neethiraj, N. Clark, N. Morehouse, M. de la Paz Celorio-Mancera, J. C. Cols, H. Dirksen, C. Meslin, N. Keehn, P. Pruischer, K. Sikkink, M. Vives, H. Vogel, C. Wiklund, A. Woronik, C. L. Boggs, S. Nylin, C. W. Wheat, Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Sci. Adv.* **5**, eaau3648 (2019).
69. S. N. Ruckman, M. M. Jonika, C. Casola, H. Blackmon, Chromosome number evolves at equal rates in holocentric and monocentric clades. *PLOS Genet.* **16**, e1009076 (2020).
70. M. Genete, V. Castric, X. Vekemans, Genotyping and de novo discovery of allelic variants at the Brassicaceae self-incompatibility locus from short-read sequencing data. *Mol. Biol. Evol.* **37**, 1193–1201 (2020).
71. A. Pribelski, D. Antipov, D. Meleshko, A. Lapidus, A. Korobeynikov, Using SPAdes de novo assembler. *Curr. Protoc. Bioinformatics* **70**, e102 (2020).
72. L. Noé, G. Kucherov, YASS: Enhancing the sensitivity of DNA similarity search. *Nucleic Acids Res.* **33**, W540–W543 (2005).
73. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
74. K. Lohse, A. Hayward, D. R. Laetsch, R. Vila, K. Lucek; Wellcome Sanger Institute Tree of Life programme; Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective; Tree of Life Core Informatics collective; Darwin Tree of Life Consortium, The genome sequence of the Arran brown, *Erebia ligea* (Linnaeus, 1758). *Wellcome Open Res.* **7**, 259 (2022).
75. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
76. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
77. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
78. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. M. Carthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
79. A. F. A. Smit, R. Hubley, P. Green, RepeatMasker Open-4.0. 2013–2015 (2015).
80. K. Lohse, E. Taylor-Cox; Darwin Tree of Life Consortium, The genome sequence of the speckled wood butterfly, *Pararge aegeria* (Linnaeus, 1758). *Wellcome Open Res.* **6**, 287 (2021).
81. D. Mead, I. Saccheri, C. J. Yung, K. Lohse, C. Lohse, P. Ashmole, M. Smith, C. Corton, K. Oliver, J. Skelton, E. Betteridge, M. A. Quail, J. Dolucan, S. A. McCarthy, K. Howe, J. Wood, J. Torrance, A. Tracey, S. Whiteford, R. Challis, R. Durbin, M. Blaxter, The genome sequence of the ringlet, *Aphantopus hyperantus* Linnaeus 1758. *Wellcome Open Res.* **6**, 165 (2021).
82. G. Bisschop, S. Ebdon, K. Lohse, R. Vila; Darwin Tree of Life Consortium, The genome sequence of the small tortoiseshell butterfly, *Aglais urticae* (Linnaeus, 1758). *Wellcome Open Res.* **6**, 233 (2021).
83. K. Lohse, A. García-Berro, G. Talavera; Darwin Tree of Life Consortium, The genome sequence of the red admiral, *Vanessa atalanta* (Linnaeus, 1758). *Wellcome Open Res.* **6**, 356 (2021).
84. L. Zhang, R. A. Steward, C. W. Wheat, R. D. Reed, High-quality genome assembly and comprehensive transcriptome of the painted lady butterfly *Vanessa cardui*. *Genome Biol. Evol.* **13**, evab145 (2021).
85. M. Stanke, O. Keller, I. Gunduz, A. Hayes, S. Waack, B. Morgenstern, AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).
86. K. J. Hoff, M. Stanke, WebAUGUSTUS—A web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res.* **41**, W123–W128 (2013).
87. D. M. Emms, S. Kelly, OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
88. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
89. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
90. L.-T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
91. S. Kalyaanamoorthy, B. Q. Minh, T. K. Wong, A. Von Haeseler, L. S. Jermiin, ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
92. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).
93. M. L. Borowiec, E. K. Lee, J. C. Chiu, D. C. Plachetzki, Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. *BMC Genomics* **16**, 987 (2015).
94. M.-Y. Chen, D. Liang, P. Zhang, Selecting question-specific genes to reduce incongruence in phylogenomics: A case study of jawed vertebrate backbone phylogeny. *Syst. Biol.* **64**, 1104–1120 (2015).
95. M. L. Borowiec, Convergent evolution of the army ant syndrome and congruence in big-data phylogenetics. *Syst. Biol.* **68**, 642–656 (2019).

96. P. Kück, K. Meusemann, FASconCAT: Convenient handling of data matrices. *Mol. Phylogenet. Evol.* **56**, 1115–1118 (2010).
97. H. Philippe, P. Forterre, The rooting of the universal tree of life is not reliable. *J. Mol. Evol.* **49**, 509–523 (1999).
98. A. Nel, J. Nel, C. Balme, Un nouveau Lépidoptère Satyrinae fossile de l'Oligocene du sud-est de la France (Insecta, Lepidoptera, Nymphalidae). *Linneana Belg.* **14**, 20–36 (1993).
99. C. Peña, S. Nylin, N. Wahlberg, The radiation of Satyrini butterflies (Nymphalidae: Satyrinae): A challenge for phylogenetic methods. *Zool. J. Linn. Soc.* **161**, 64–87 (2011).
100. A. Y. Kawahara, D. Plotkin, M. Espeland, K. Meusemann, E. F. A. Toussaint, A. Donath, F. Gimnich, P. B. Frandsen, A. Zwick, M. D. Reis, J. R. Barber, R. S. Peters, S. Liu, X. Zhou, C. Mayer, L. Podsiadlowski, C. Storer, J. E. Yack, B. Misof, J. W. Breinholt, Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22657–22663 (2019).
101. R. De Jong, Fossil butterflies, calibration points and the molecular clock (Lepidoptera: Papilionoidea). *Zootaxa* **4270**, 1–63 (2017).
102. F. Ronquist, M. Teslenko, P. Van Der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, J. P. Huelsenbeck, MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
103. R. Lanfear, P. B. Frandsen, A. M. Wright, T. Senfeld, B. Calcott, PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**, 772–773 (2017).
104. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria (2017).
105. A. R. Mast, P. M. Olde, R. O. Makinson, E. Jones, A. Kubes, E. T. Miller, P. H. Weston, Paraphyly changes understanding of timing and tempo of diversification in subtribe Hakeinae (Proteaceae), a giant Australian plant radiation. *Am. J. Bot.* **102**, 1634–1646 (2015).
106. S. Höhna, M. J. Landis, T. A. Heath, B. Boussau, N. Lartillot, B. R. Moore, J. P. Huelsenbeck, F. Ronquist, RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* **65**, 726–736 (2016).
107. Y. Nakatani, A. McLysaght, Macrosyteny analysis shows the absence of ancient whole-genome duplication in lepidopteran insects. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1816–1818 (2019).
108. W. J. Tennent, A checklist of the satyrine genus *Erebia* (Lepidoptera) (1758–2006). *Zootaxa* **1900**, 1–109 (2008).
109. A. Rambaut, A. Drummond, Tracer: A program for analysing results from Bayesian MCMC programs such as BEAST & MrBayes. University of Edinburgh, UK (2003).
110. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
111. R. E. Kass, A. E. Raftery, Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
112. G. Baele, P. Lemey, T. Bedford, A. Rambaut, M. A. Suchard, A. V. Alekseyenko, Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29**, 2157–2167 (2012).
113. T. Hothorn, K. Hornik, M. A. Van De Wiel, A. Zeileis, Implementing a class of permutation tests: The coin package. *J. Stat. Softw.* **28**, 1–23 (2008).
114. J. M. Beaulieu, B. C. O'Meara, Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601 (2016).
115. R. Robinson, *Lepidoptera Genetics*. (Pergamon Press, 1971).
116. H. de Lesse, Signification supraspécifique des formules chromosomiques chez les Lépidoptères. *Bull. Soc. Entomol. Fr.* **66**, 71–83 (1961).
117. H. de Lesse, Spéciation et variation chromosomique chez les Lépidoptères Rhopalocères. *Ann. Sci. Nat. Zool. Biol. Anim.* **12**, 1–223 (1960).
118. K. Saitoh, Y. Kumagai, I. Tateyama, M. Kawata, A chromosome study of *Erebia nipponica* Janson, 1877 (Lepidoptera, Satyridae) from Hokkaido, Japan. *Tyô To Ga* **42**, 11–15 (1991).

Acknowledgments: We thank C. Praz for advice on validation of phylogenetic inferences, as well as A. Filippov and P. Jakubek for aid in specimen collection. Analyses were performed at sciCORE (<http://scicore.unibas.ch/>) the scientific computing center at University of Basel.

Funding: H.A. was supported by the Burckhardt-Bürgin Foundation and the Swiss National Science Foundation (SNSF) project: Genomic rearrangements and the origin of species (310030_184934) awarded to K.L. K.L. was further supported by the SNSF Eccellenza project: The evolution of strong reproductive barriers toward the completion of speciation (PCEFP3_202869). J.M.d.V. was supported, in part, by SNSF grant 310030_185251. R.V. was supported by grant 2021_SGR_00420 from the Departament de Recerca i Universitats, Generalitat de Catalunya, and by Project PID2022-139689NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF, EU. V.D. was supported by the Academy of Finland (Academy Research Fellow, decisions nos. 324988 and 352652). K.O. was supported by JSPS KAKENHI grant number JP21K15165. **Author contributions:** Conceptualization: J.M.d.V., K.L., R.V., and H.A. Data curation: K.L. and H.A. Formal analysis: S.K., J.M.d.V., and H.A. Funding acquisition: K.L. Investigation: G.I., K.L., T.C., K.O., and H.A. Methodology: S.K. and J.M.d.V. Project administration: K.L. Resources: S.K., J.M.d.V., K.L., R.V., V.D., K.O., H.A., and M.C. Software: S.K., J.M.d.V., and H.A. Supervision: K.L. Validation: K.L. and H.A. Visualization: H.A. Writing—original draft: H.A. Writing—review and editing: G.I., S.K., J.M.d.V., K.L., T.C., R.V., V.D., K.O., H.A., and L.B. **Competing interests:** The authors declare that they have no competing interest. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The raw read data can be found on NCBI under the accession number PRJNA1000734. Code to generate the phylogeny, as well as ChromoSSE and ChromEVOL scripts and output, and R codes for analysis can be found on Zenodo under the accession number 10.5281/zenodo.10568537.

Submitted 28 September 2023

Accepted 14 March 2024

Published 17 April 2024

10.1126/sciadv.adl0989