





Article

Genomics Reveal Admixture and Unexpected Patterns of Diversity in a Parapatric Pair of Butterflies

Mohadeseh Sadat Tahami ^{1,*}, Vlad Dincă ¹, Kyung Min Lee ¹, Roger Vila ², Mukta Joshi ¹, Maria Heikkilä ³, Leonardo Dapporto ⁴, Sarah Schmid ⁵, Peter Huemer ⁶ and Marko Mutanen ¹

¹ Ecology and Genetics Research Unit, University of Oulu, P.O. Box 3000, 90014 Oulu, Finland; vlad.e.dinca@gmail.com (V.D.); kyung.m.lee@helsinki.fi (K.M.L.); muktajoshi3@gmail.com (M.J.); marko.mutanen@oulu.fi (M.M.)

² Institut de Biologia Evolutiva (CSIC—Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta, 37, 08003 Barcelona, Spain; roger.vila@csic.es

³ Zoology Unit, Finnish Museum of Natural History, University of Helsinki, P.O. Box 17, 00014 Helsinki, Finland; maria.heikkila@helsinki.fi

⁴ Numerical and Experimental Zoology Laboratory (ZEN Lab), Dipartimento di Biologia, Dell' Università di Firenze, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy; leonardo.dapporto@unifi.it

⁵ Departement de Biologie Computationnelle, Faculte de Biologie et Medecine, Universite de Lausanne, 1015 Lausanne, Switzerland; sarah.schmid@unil.ch

⁶ Tiroler Landesmuseen Betriebsges.m.b.H., Naturwissenschaftliche Sammlungen, Krajnc-Str. 1, A-6060 Hall, Austria; p.huemer@tiroler-landesmuseen.at

* Correspondence: tahami.m@hotmail.com



Citation: Tahami, M.S.; Dincă, V.; Lee, K.M.; Vila, R.; Joshi, M.; Heikkilä, M.; Dapporto, L.; Schmid, S.; Huemer, P.; Mutanen, M. Genomics Reveal Admixture and Unexpected Patterns of Diversity in a Parapatric Pair of Butterflies. *Genes* **2021**, *12*, 2009. <https://doi.org/10.3390/genes12122009>

Academic Editor: Petr Kotlík

Received: 5 November 2021

Accepted: 14 December 2021

Published: 17 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: We studied the evolutionary relationship of two widely distributed parapatric butterfly species, *Melitaea athalia* and *Melitaea celadussa*, using the ddRAD sequencing approach, as well as genital morphology and mtDNA data. *M. athalia* was retrieved as paraphyletic with respect to *M. celadussa*. Several cases of mito-nuclear discordance and morpho-genetic mismatch were found in the contact zone. A strongly diverged and marginally sympatric clade of *M. athalia* from the Balkans was revealed. An in-depth analysis of genomic structure detected high levels of admixture between *M. athalia* and *M. celadussa* at the contact zone, though not reaching the Balkan clade. The demographic modelling of populations supported the intermediate genetic make-up of European *M. athalia* populations with regards to *M. celadussa* and the Balkan clade. However, the dissimilarity matrix of genotype data (PCoA) suggested the Balkan lineage having a genetic component that is unrelated to the *athalia-celadussa* group. Although narrowly sympatric, almost no signs of gene flow were found between the main *M. athalia* group and the Balkan clade. We propose two possible scenarios on the historical evolution of our model taxa and the role of the last glacial maximum in shaping their current distribution. Finally, we discuss the complexities regarding the taxonomic delimitation of parapatric taxa.

Keywords: ddRAD sequencing; parapatry; paraphyletic species; genomic admixture; *Melitaea*; species delimitation

1. Introduction

Difficulties to delimit species under many circumstances are severely undermining attempts to catalogue global biodiversity and have an adverse effect on conservation efforts [1]. These difficulties stem from both biological and operational causes. As speciation is not an event but a process, defining the time of birth of a new species remains inherently arbitrary [2]. Ontological difficulties, in turn, are largely due to the lack of consensus over the species concept and critical properties that a population should bear in order to be considered as a valid species [3]. Additionally, the critical properties of species can emerge at different time and order during the course of divergence [4]. Although attempts have been made, a widely accepted consensus over the criteria of species is evidently hard to

reach. Traditionally, insect species have largely been circumscribed based on morphological traits, assuming that morphological differences, particularly those in genitalia, play a key role in preventing gene flow between populations undergoing speciation [5]. The “lock-and-key hypothesis”, i.e., an idea that genital differences between species have evolved in order to prevent insemination across species [6], was long popular particularly among entomologists, although it suffers from both theoretical weaknesses [7,8] and a lack of empirical support [2,9,10].

The parapatric model of distribution, in which ranges of closely related taxa confront but show only narrow areas of overlap, presents particular challenges for species delimitation. The parapatric distribution pattern may result from parapatric speciation itself, but also from secondary contact of formerly separated and genetically diverged populations [2]. Parapatric taxa represent an appealing condition to study gene flow among emerging species as it is likely a frequently occurring phase in the course of speciation. To obtain a better perspective of relationships of parapatric species, the spatial component and historical dimension of isolation should be taken into account [11]. Indeed, parapatry could have been achieved very recently, notably in the last interglacial, between taxa that remained in allopatry during most of their evolutionary history, with no or only casual contacts.

Populations with a contact zone can be used as a natural experiment to investigate features that hinder or fully prevent gene flow between them. Parapatry is not uncommon, among the best documented, being that of the Hooded Crow (*Corvus cornix*) and the Carrion Crow (*Corvus corone*), species which show broad genetic admixture, but also strong assortative mating [12]. As the evolutionary trajectory of natural populations is traceable in populations’ gene pool, an in-depth genomic study enables obtaining accurate information about the genetic structure and the demographic history of those populations. Emergent genomics tools also provide high potential for better-standardized and quantifiable criteria to species delimitation under para- and allopatry [13,14].

We aimed to provide genomic insight on the evolutionary dynamics of parapatric populations of the butterflies *M. athalia* (Rottemburg, 1775) and *M. celadussa* Fruhstorfer, 1910 (Nymphalidae). These taxa represent one of the most remarkable cases of parapatry among European butterflies. They are widely distributed in most parts of Europe and western parts of Asia and exhibit a narrow contact zone in France, Switzerland, Italy and Austria, with *M. celadussa* occurring in south and west, and *M. athalia* in north and east of the contact zone [15–18]. Distinguishing between these two taxa is difficult as they are extremely similar in their external morphology, being only diagnosable by a set of characters in their male genitalia. However, specimens with intermediate male genital characters have been documented in the contact zone, for which reason *M. celadussa* was long considered as a subspecies of *M. athalia* [15,19]. *M. celadussa* was recently proposed as a separate species based on molecular evidence [20,21] and it is currently regarded as such in the latest checklist of European butterflies [22]. However, molecular data for this taxon are based on either a single specimen [20] or restricted to DNA barcodes [18]. We used a high-throughput genomics approach to address the genetic structure and evolutionary history of this parapatric pair of butterflies.

2. Materials and Methods

2.1. Sample Data

The sampling is representative for the distribution of *M. athalia* and *M. celadussa* in Europe, including their contact zone. DNA extraction was performed for 61 specimens of *M. athalia* and *M. celadussa*, as well as for 10 specimens representing outgroup species; *M. caucasogenita/athalia* (see comment in the Supplementary Material), *M. britomartis* Assmann, 1847, *M. deione* (Geyer, [1832]), *M. aurelia* Nickerl, 1850. We assigned individuals to putative species based on established differences in DNA barcodes (the 5’ region of mitochondrial cytochrome c oxidase subunit 1—COI) since it was the only method of identification that was available for all specimens. Additionally, whenever possible, we used genital

structure for species attribution (i.e., males with intact genitalia). For a detailed description of morphological, and mitochondrial data preparation and analyses, please refer to the supplementary material (SM). Additionally, data of 418 specimens of *M. athalia* and *M. celadussa* collected across their European distribution and sequenced for DNA barcodes were used to complement and optimize our study (Table S5). The COI tree (Figure S1) and details of voucher specimens (Table S1) are available in SM.

2.2. Genomic Assembly and Bioinformatics

After the library preparation and sequencing (please refer to the supplementary material for details), raw demultiplexed Illumina (San Diego, CA, USA) paired-end reads were merged using PEAR [23] with default parameter values. Merging paired-end reads to single-end minimizes false-positive results and generates highly accurate single-end reads. To rule out the potential effect of bacterial contamination and maternally inherited mitochondrial genes, merged reads that were mapped to *Wolbachia pipientis* (GenBank: NZ_JQAM01000001) and *Melitaea cinxia* mitochondrial genome (GenBank: CM002851) were removed using Geneious Prime 2019.2.1 (<https://www.geneious.com> (accessed on 16 September 2019)). Merged reads were processed as input data for the ipyrad pipeline version 0.9.31 [24].

Cleaned reads were loaded and filtered from adapters and primers through step 1 and 2 of ipyrad. To explore the ddRAD data, several assembly branches were created to test for various levels of consensus quality by applying parameters change. Clustering threshold ($c = 80, 85, 88, 90, 93, 94, 95$) and minimum number of samples per locus ($m = 4, 8, 10, 12, 20$) were tested independently and as combinations with each other using both de novo and reference assembly methods. The latter method assembles loci by mapping the paired reads onto the *M. cinxia* (Linnaeus, 1758) complete genome (GCA_00071638). After clustering and aligning reads (step 3), consensus allele sequences were estimated (steps 4 and 5). Samples with consensus coverage lower than 5% were excluded from downstream analyses. Subsequently, other parameters that can affect the quality of the final consensus alignment were also tested as follows: maximum number of SNPs per locus ($\text{max_SNPs_locus} = 12, 15, 20, 22, 25$), maximum number of indels per final locus ($\text{max_Indels_locus} = 4, 6, 8$), maximum fraction of uncalled bases in the consensus sequences ($\text{max_Ns_consens} = 3, 5, 7$), maximum fraction of heterozygous bases allowed in the consensus sequences ($\text{max_Hs_consens} = 6, 8, 10$), and maximum number of low-quality base calls in a read ($\text{max_low_qual_bases} = 3\text{--}5$). The latter parameter was kept below the recommended value of 5 as higher values will increase the final number of ambiguous (N) sites which will affect the accuracy and reliability of downstream clustering. Before clustering the concatenated consensus sequences, one allele was randomly called per locus so that ambiguous characters have a lesser effect on clustering, but the resulting data retained information for heterozygotes (step 6) [24]. The clustered sequences were then aligned and filtered according to the given “m”, “max_SNPs_locus” and other parameters and saved in multiple output formats (step 7).

2.3. Phylogenetic Analysis

To study the phylogenetic relationships among taxa, we conducted a maximum likelihood (ML) analysis. We first searched for the best fitting model of sequence evolution using ModelFinder implemented in IQ-TREE version 1.5.4 [25]. The result recovered TVM + F + I + G4 as the best fitting substitution model for our data. The phylogenetic tree was constructed with an ultrafast bootstrap approximation model (1000 replicates) using the same IQ-TREE version [26,27]. The consensus tree was visualized in FigTree v1.4.4 (<https://github.com/rambaut/figtree/releases> (accessed on 25 November 2019)) and was rooted on *M. aurelia*.

2.4. Analyses of Genomic Diversity and Admixture

To understand the genomic patterns and the level of admixture in parapatry, we performed STRUCTURE [28] analysis using Bayesian algorithm. The admixture model was used on SNP data shared between a minimum number of 20 samples (*m20* and 18,383 generated SNPs) with no prior population assignment. To determine the optimal number of genetic clusters (*K*), we used the ΔK method in STRUCTURE HARVESTER [29,30] with 500,000 iterations for MCMC and 100,000 as pre-burn in. We tested five putative numbers of clusters, *K* = 1–5, with 10 iterations for each *K* out of which the optimum of *K* = 2 was estimated. We then aligned the cluster assignments of 2–4 across replicate analyses in CLUMPP [31] and used DISTRUCT [32] for visual representations of the aligned clusters. To have a better understanding of the geographic patterns of genomic admixture at *K* = 2 and *K* = 3, we mapped the membership coefficient matrix obtained from the Bayesian clustering model to the matrix of geographical coordinates in Jupyter lab (<https://jupyter.org> (accessed on 23 April 2020)) using the Basemap Matplotlib Toolkit v1.2.0 implemented in the Matplotlib package v3.0.3 [33].

To further investigate levels of divergence within and between clusters, we calculated pairwise F_{st} and F_{is} in R 3.6.1 [34] using the function “basic.stat” of package Hierfstat [35].

2.5. Population History

The history of populations was investigated using approximate Bayesian computations in DIYABC v2.1.0 [36] based on SNP data (*m20*). Four historical scenarios were assumed as follows: (1) main *M. athalia* are diverged from the Balkan lineage (see phylogenetic result), (2) both main *M. athalia* and Balkan lineage are diverged from *M. celadussa*, (3) main *M. athalia* is an admixture of the Balkan lineage and *M. celadussa*, and (4) main *M. athalia*, the Balkan lineage and *M. celadussa* split at the same time considering the Balkan lineage as the ancestral population. All scenarios were assumed uniform with priors for population effective size (*N_e*) and times (*t*) between 10–10,000, *t*₂ was considered bigger than *t*₁, and the admixture rate (*r*) between 0.001–0.999. The analysis was performed with 1,000,000 simulations on a subset of 2000 loci out of 10,724.

2.6. Isolation by Distance and Geographic Representation of Genetic Differentiation

As genomic data sometimes tend to be clustered geographically, we also tested for the genomic isolation by geographical distance (IBD) based on individual level data. Mantel correlation between genomic and geographic distance was assessed with the “mantel” function of R package adegenet [37] with 999 permutations. Genomic Edwards’ distance matrices were generated using the “dist” function of R. A geographical distance matrix encompassing the minimum path distance among specimens over land was produced by calculating the length of the minimum over-ground path among sequenced specimens using the “costDistance” function as implemented in the gdistance R package (v. 1.2-1) [38]. The local density was measured using a 2-dimensional kernel density estimation within the R package adegenet. As done with COI (see Supplementary Materials for details), we applied PCoA to the ddRADseq dissimilarity matrix using the “cmdscale” R function and minimized the location and rotation differences with respect to the COI PCoA, then we projected it in the RGB colour space the specimens and mapped the resulting colours.

3. Results

3.1. Patterns of COI Variability

The neighbor joining (NJ) tree based on DNA barcodes (Figure S1) grouped specimens of *M. athalia*–*M. celadussa* into three main well-supported (Bootstrap Support (BS) > 0.95) clusters. Based on their established distribution [22], one cluster was assigned to *M. celadussa*, and the other two to *M. athalia*. One of the *M. athalia* clusters consisted of six specimens originating from the Balkans and from here on we refer to it as “the Balkan lineage”, and the rest of *M. athalia* will be referred to as “main *M. athalia*”. The minimum *p*-distances between the three main clusters were: main *M. athalia*–*M. celadussa* = 2.75%;

Balkan lineage–main *M. athalia*: 3.92%; Balkan lineage–*M. celadussa*: 3.92%. It should be noted that the main *M. athalia* cluster also included the specimen that we tentatively identified as *M. caucasogenita/athalia* (see comment in the supplementary material) (Figure S1).

3.2. Morphometrics

The measured elements of the male genitalia (sub-unci and posterior process of the valva) (Figure S2) revealed a morphological continuum with an overlap between *M. celadussa* and *M. athalia* (Figures 1c and S3, Table S2), while the degree of bifurcation of the posterior process of valva showed a higher concordance with the DNA barcodes. The length of the posterior process of valva showed considerable overlap between taxa, although specimens attributable to *M. celadussa*, based on DNA barcodes, tended to have a longer posterior process (Table S2, Figure S3a). *M. celadussa* also had absent or vestigial sub-unci, as opposed to the longer sub-unci in main *M. athalia* and the Balkan lineage (Figure S3). However, this pattern was broken by six specimens with intermediate length of sub-unci (north-eastern Spain, south-eastern France and Switzerland), as well as by six specimens (Switzerland and northern Italy) which had a genitalia morphology of *M. athalia* type, but belonged to *M. celadussa* (Table S2, Figure S3b). The three measured specimens belonging to the Balkan COI lineage had genitalia morphology of *M. athalia* type (i.e., well-developed sub-unci and shorter posterior process of the valva) (Figure S3).

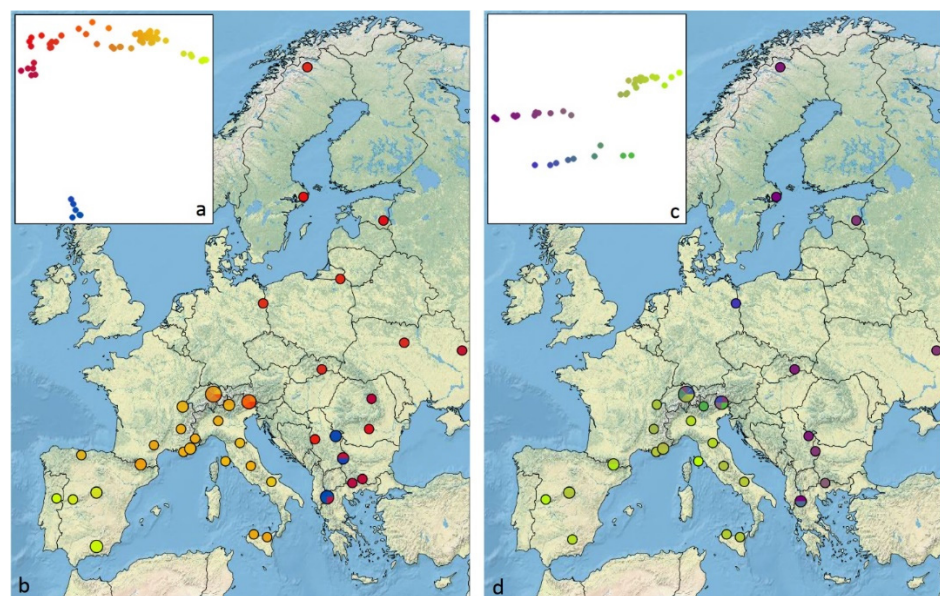


Figure 1. Patterns of genetic and morphological differentiation in *M. athalia* and *M. celadussa*: (a) The configuration obtained after PCoA of Edwards' distances calculated on ddRAD data and projected in RGB space where similar individuals are characterised by similar colours; (b) The geographic location of each specimen with the colour obtained in the PCoA; (c) The configuration obtained after PCA on genitalia traits; (d) The geographic location of specimens.

The presence or absence of a bifurcation on the posterior process of the valva (Table S2, Figure S3a) was better correlated with COI identification than the length of the sub-unci. Indeed, all specimens with COI of *M. celadussa* showing intermediate sub-unci had no bifurcation except for one specimen showing a small bifurcation. All specimens with genitalia of *M. athalia* type had a bifurcation, including all those attributable to *M. celadussa* based on COI and those belonging to the Balkan COI lineage. Accordingly, a PCA identified two components with eigenvalue higher than one (Figure S4) where longer sub-unci were highly correlated with the presence of a larger bifurcation. The projection of this configuration on RGB colour and the location of obtained colours on the map revealed

that the individuals with intermediate genitalic morphology and/or incongruent mitomorphology are only seen in the contact zone of the Alps (Figure 1d).

3.3. Overview of ddRAD Data

An average of 1.8 million raw reads were obtained from Illumina sequencing (Table S3). After checking results from different combinations of parameter values for clustering using both de novo and reference assembly methods, we chose 85% ($c0.85$) of sequence similarity to retain loci shared across more than four samples ($m4$) for phylogenetic analysis along with the default values for other parameters. This approach recovered a robust sequence alignment with confident summary statistics. The tree topologies reconstructed from either of the assembly methods with the selected parameter values were generally congruent. Although the de novo assembly could recover a higher number of loci and SNPs, we chose the reference-based assembly as it yielded a more robust tree. Generally, it has been recommended to map restriction site associated genomic data to a reference genome to avoid potential collecting of repeated regions and non-informative loci [39]. We retrieved a total of 3,892,196 base pairs (bp) with 30,696 (0.8%) identical sites, 37,507 putative orthologous loci and 204,430 SNPs of which 71,774 are parsimony informative. The minimum percentage of sample coverage was 6.23% and the maximum was 24.13%, while the proportion of missing data was 85.9% over all loci. For genomic analyses other than phylogeny, we used the $c0.85$ - $m20$ data assembly, therefore significantly lowering the proportion of missing data to below 50% (41%) and increasing the average sample coverage to 77.81%.

3.4. ddRAD Phylogeny

The ML tree reconstructed based on concatenated ddRAD loci revealed a grouping of specimens that is discordant to the taxonomic assignments based on COI and morphology (Figure 2). In a tree rooted to *M. aurelia*, the putative specimen of *M. caucasogenita/athalia* was found as a sister lineage to the well supported (BS = 100) *M. athalia/celadussa* clade. Specimens of *M. athalia* formed a highly paraphyletic grade with respect to *M. celadussa*. Remarkably, a lineage consisting of six Balkan specimens attributed to *M. athalia* was recovered as a sister lineage to all other specimens of *M. athalia* and *M. celadussa* (BS = 100). *M. celadussa* formed a near-monophyletic clade, albeit with the following two exceptions. The specimen RVcoll13U296 recognized as *M. athalia* by both genitalia morphology and COI was nested within this clade. Similarly, one specimen of *M. celadussa* COI (RVcoll15I495) was recovered among two specimens of *M. athalia* (RVcoll13U438 and RVcoll15I360). Within the clade of *M. celadussa*, five Swiss specimens, each showing mitomorphological discordance (Figures S1 and S3, Table S2), formed a relatively divergent sister clade to the rest of *M. celadussa* (plus RVcoll13U296). This Swiss clade did not show a clear separation from most specimens of *M. celadussa* in the COI tree (Figure S1). Other cases of intermediate genitalia were placed within *M. celadussa*.

3.5. Genomic Diversity and Admixture

The STRUCTURE analysis estimated the highest likelihood for the existence of two clusters in our dataset (Figure S5a). All STRUCTURE results from $K = 2$ to 4 (Figure S5b) showed frequent genomic admixture at the contact zone. However, the geographic distribution of genomic admixture within *M. athalia* varied depending on the K . At $K = 2$ admixture from *M. celadussa* extended into many of the analysed specimens of *M. athalia*, (all Northern and Eastern Europe) and to a lesser extent in the Balkans (Figure 3a). The pattern was different in results for $K = 3$ and $K = 4$ (Figure S5b), in which admixture between *M. celadussa*-main *M. athalia* was limited to the contact zone; in $K = 3$ there is also admixture in the Balkans, between the main *M. athalia* and the Balkan lineage, as well as slight admixture between Iberian *M. celadussa* and the Balkan lineage of *M. athalia*. (Figure 3b).

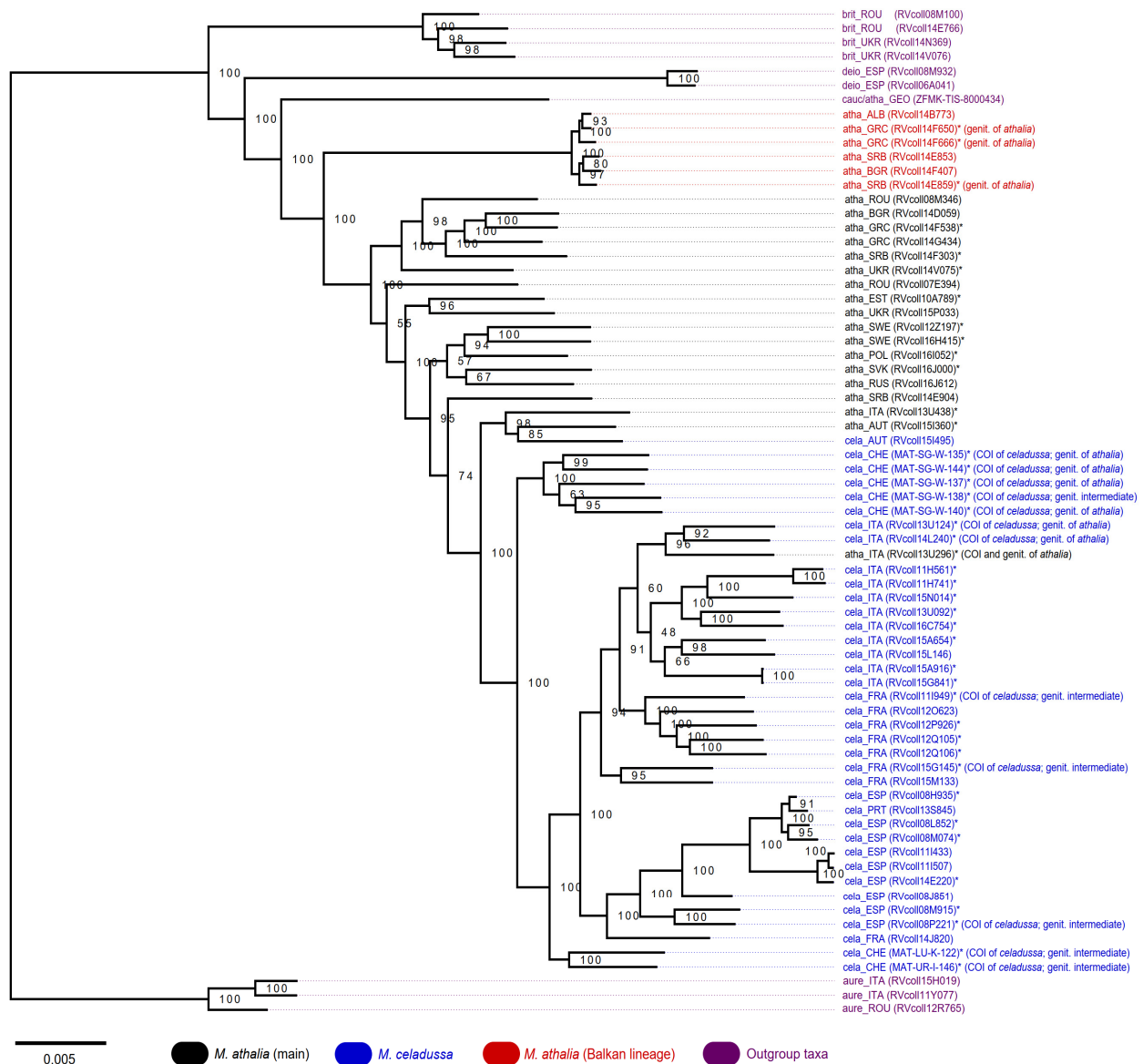


Figure 2. Maximum likelihood tree of *M. athalia*–*M. celadussa* inferred based on ddRAD data using reference assembly. Bootstrap supports are given at each node. The tree is rooted on *M. aurelia*. Asterisks indicate specimens examined for morphometrics.

A PCoA visualization of the dissimilarity matrix (Figure 1a) confirmed this result since two main clusters were obtained (differentiated by PC2); the Balkan specimens as one cluster and a continuous group formed by main *M. athalia*–*M. celadussa* specimens as the other (which is spread along PC1). When plotted on the maps, the intermediate specimens between extreme *M. celadussa* and main *M. athalia* were located in the contact zone of the Alps (Figure 1b). Iberian *M. celadussa* was also recovered as a slightly differentiated group, thus confirming COI data (Figures 1a,b and S6).

Analyses of genetic distance based on $K = 2$ clusters obtained $F_{st} = 0.1967$ and the diversity within populations was $F_{is} = 0.6108$. When applying a clustering model equivalent to morphological species, the values did not significantly change: $F_{st} = 0.0964$ and $F_{is} = 0.6112$, indicating that gene flow is generally high between and within populations.

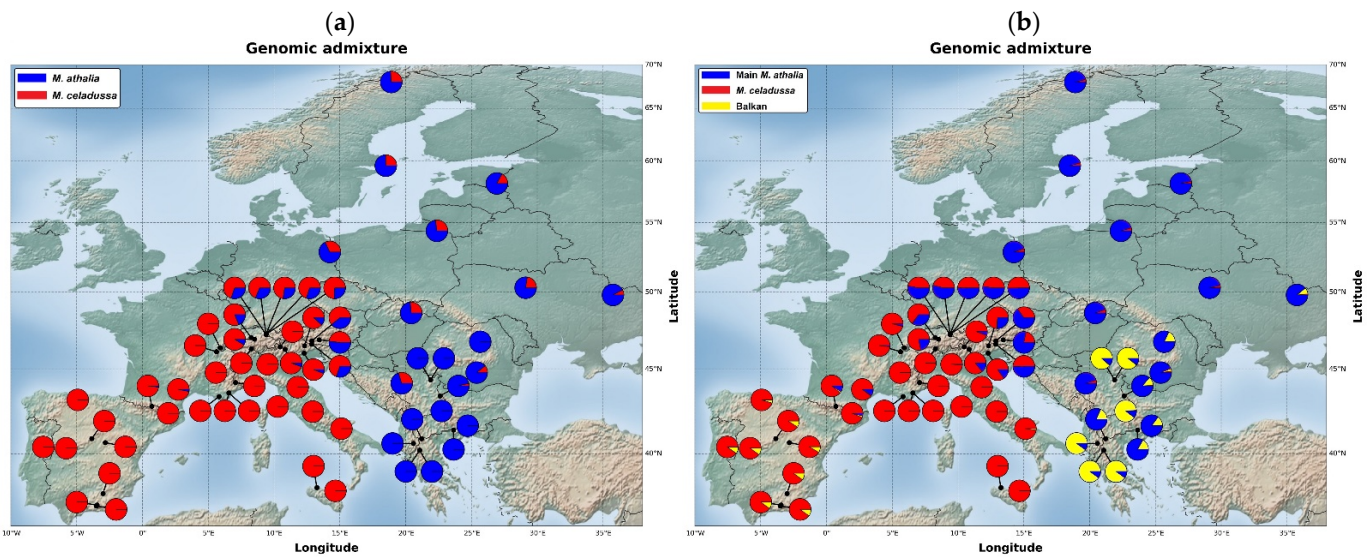


Figure 3. Pie chart map following STRUCTURE results of $K = 2$ (a) and $K = 3$ (b) based on ddRAD data, showing individual genomic admixture of *M. celadussa*–*M. athalia* according to geographic location.

3.6. Population History

Of four tested scenarios, scenario 3 had the highest logistic regression of posterior probability (0.9986 [0.9770, 1.0000]) with 2% of simulated data (Figure 4), which supports past hybridization. Testing for the model fitness based on principal component analysis confirmed that the admixture scenario fits well the empirical data because the observed data was located within the posterior probability distribution of the dataset (Table S4, Figure S7).

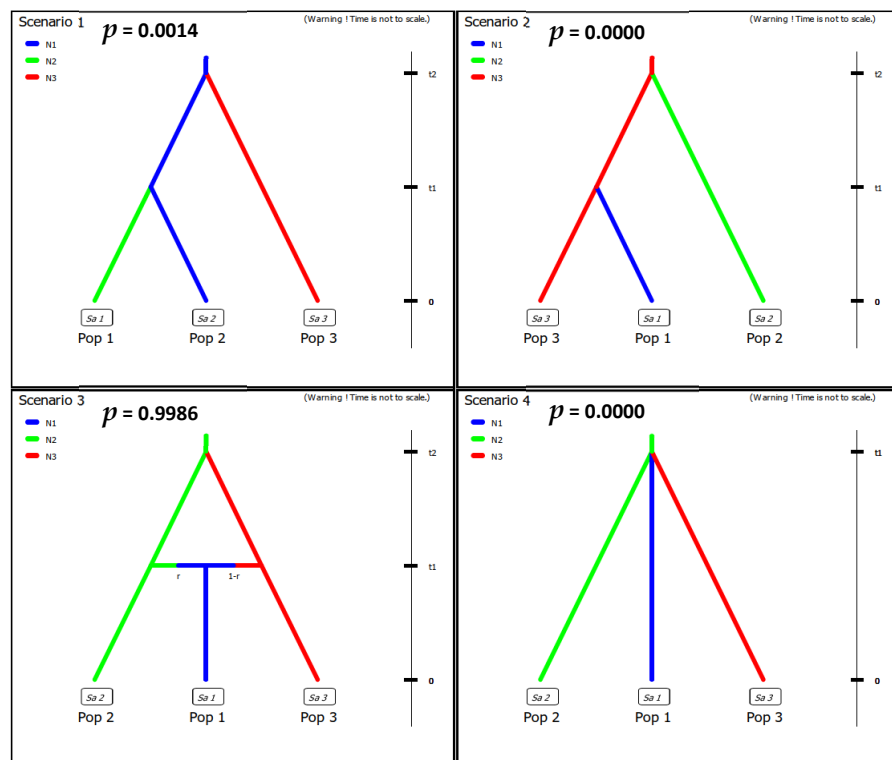


Figure 4. Demographic scenarios designed for DIYABC analysis. Pop 1: main *M. athalia*, Pop 2: Balkan lineage, Pop 3: *M. celadussa*. The logistic regressions of posterior probabilities are given for each scenario suggesting the third scenario as the fittest one.

3.7. Isolation by Distance

The isolation by distance (IBD) analysis for *M. athalia*–*M. celadussa* displayed one cloud of points indicating a continuous cline of genetic variation correlating with geographic distance, an exception being the Balkan lineage, displaying a higher level of differentiation given the geographic distance (smaller cloud above the main gene pool) (Figure 5). The Mantel test revealed a significant positive correlation of $r^2 = 0.504$ ($p = 0.001$) between the two matrices indicating that genomic differentiation increases with geographical distance.

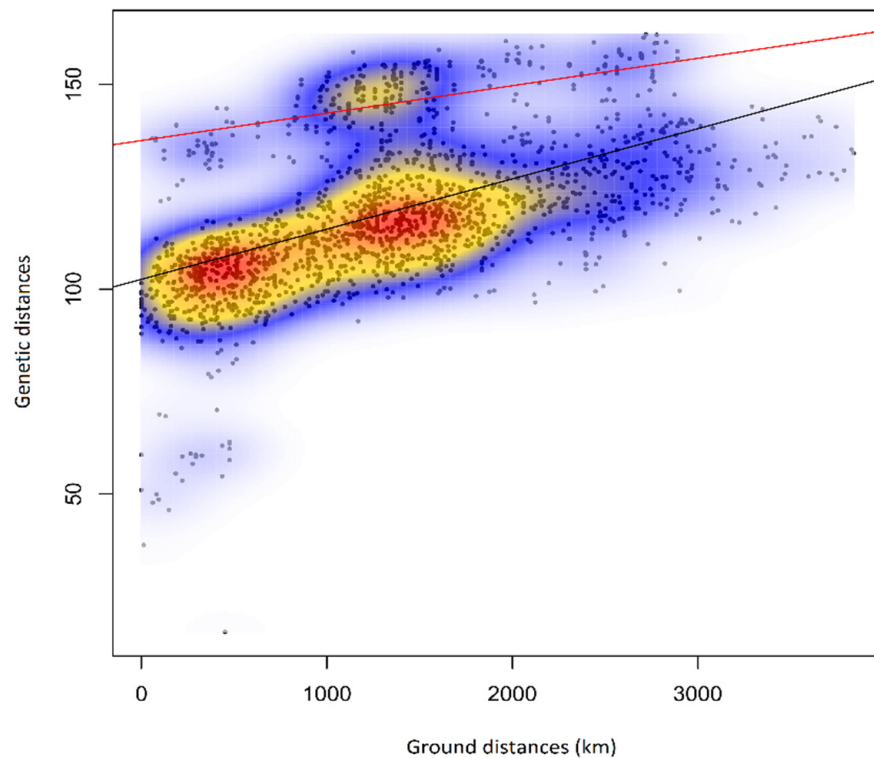


Figure 5. Isolation by distance plot illustrating the pattern of genetic differentiation in the *M. athalia*–*M. celadussa* species group ($r^2 = 0.536$, $p = 0.001$) in relation to pairwise geographic distances. The black line represents the overall trend, while the red line represents the relationship between the Balkan clade and the group composed by main *M. athalia* + *M. celadussa*.

4. Discussion

Melitaea athalia and *M. celadussa* are among the most widespread and common European butterfly species [18,20,21,40,41]. However, their relationship has been the subject of considerable interest [15,16,42] due to their intermediate feature displays and occasional mito-morphological discordance at the contact zone. Here, we contributed a genome-wide assessment and interpreted results considering the spatial distribution of the observed diversity and the paleoclimatic reconstructions for Europe [43,44].

4.1. Genetic Structure and Levels of Admixture in *M. athalia* and *M. celadussa*

The topology of the ML tree based on ddRAD data largely reflects the pattern of COI differentiation, but only partially supports the reciprocal monophyly of *M. athalia* and *M. celadussa*. In the ddRAD tree, most of the *M. athalia* appeared as paraphyletic suggesting a certain degree of admixture with *M. celadussa*. Although ML phylogenetic reconstruction does not consider gene flow, the numerous potentially admixed specimens, especially from the contact zone, seem to generate the paraphyly (Figure 2).

Most specimens identified as *M. celadussa*, based on COI and genitalia, are also recovered as a monophyletic clade in the ML tree, with two exceptions showing mito-nuclear discordance in the area of contact (Figure 2). Mito-nuclear discordance is a relatively common

phenomenon at contact zones, where it may be produced by events of hybridization [45]. The analyses of genomic structure and isolation by distance uncovered a continuum across the geographic range of *M. celadussa* and *M. athalia* due to the admixture at the contact zone (Alps). The degree of admixture declines to central and northern Europe, is at its lowest in the Balkans, and reflects a non-symmetric introgression north-east of the contact zone (Figure 3a) [46]. This could be an outcome of adaptive introgression of a set of *M. celadussa* genes into *M. athalia* eastward or, alternatively, the remnants of an ancient contact zone that may have been shifting to the west, being spatially established in the Alps e.g., [47]. The spread of strongly admixed individuals to the north in areas climatically available to these species only after the onset of the last interglacial could, in essence, reflect a natural phenomenon where recolonization happened by chance (founder effect). Alternatively, the current pattern could agree with the first scenario involving a selective advantage for admixed specimens and allowing for the fast introgression of selected genes, possibly through traits that are related to temperature, hibernation or changes in the larval host plants (hybrid fitness) [48–51]. Such a selection is likely to stabilize the geographic location of introgressed populations [52].

Several taxa of the genus *Melitaea* display notable levels of intraspecific genetic differentiation and have a challenging taxonomy [18,53]. We documented a well-differentiated clade in the Balkans that has only recently been reported based on DNA barcodes [18]. This lineage, with a long basal branch, is sister to the rest of the group for ddRAD data. An Approximate Bayesian Computations (ABC) analysis suggested a scenario indicating the hybrid nature for main European populations of *M. athalia* with regards to *M. celadussa* and the Balkan clade as more likely compared to scenarios with strict bifurcations among taxa (i.e., without gene flow) (Figures 4 and S7, Table S4). However, it is still possible that the solution of ABC is only driven by the introgression between the main *M. athalia* and *M. celadussa* and that, as suggested by structure analyses with $K = 3$ (Figure 3b), and by comparing the ML tree with COI structure, the main *M. athalia* and the Balkan clade show only little genomic admixture. The dissimilarity matrix of genotype data (PCoA) and isolation by distance suggest that the Balkan lineage has a genetic component distinct from the *celadussa-athalia* group. Interestingly, samples of *M. athalia* representing the main genotypes have been collected almost syntopically with specimens representing the Balkan genotype without evidence of mito-nuclear discordance or strong genomic admixture. Further research with additional sampling is needed to better assess potential barriers to gene flow in the Balkans and to determine whether this lineage represents a distinct species [54,55].

4.2. Patterns of Variability in Male Genital Morphology

When the three variables measured for male genital morphology are combined in a PCA (Figure S4), two morphotypes clearly emerged corresponding to the genital variation described in previous studies [15,16]. Overall, morphometrics reveal a continuum of genital shape, but specimens with intermediate traits are only found in the contact zone (mainly in the Alps, but one instance in the Spanish Pyrenees, the western extreme of the putative contact zone), where also mito-nuclear discordance is found. Moreover, comparison of genital morphology and COI revealed cases of mito-morphological discordance, i.e., presence of mtDNA of *M. celadussa* in specimens with genital morphology of *M. athalia*. All cases of morphologically intermediate specimens and mito-morphologically mismatch specimens are attributed to *M. celadussa* by COI which suggests discordant mitochondrial and nuclear clines. In addition to stochastic reasons, sex-biased dispersal, assortative mating or asymmetric reproductive fitness can be among the causes for this discrepancy [56,57]. Similar cases of mito-morphological mismatch have been reported previously [45,55,58,59]. Possibly, a similar condition can be found all along the contact zone in central/southern France and reaching the Pyrenees [15,16,19] as suggested by one specimen from the Spanish Pyrenees (RVcoll08P221). Outside the contact zone, the

populations do not display intermediate morphology, and genitalic measures for the three male specimens of the Balkan lineage fell within the variation of the typical *M. athalia*.

4.3. The Evolutionary Origin of the *M. athalia*–*M. celadussa* Distribution

Theoretically, secondary contact of diverged populations may result either in speciation through reinforcement or merging of the populations [60,61]. However, an unstable situation can be maintained at least for a certain time in which hybridization is limited to a narrow contact zone without the involvement of most populations. Under some circumstances, i.e., when hybrids have a lower fitness than pure individuals, populations close to the contact zone can behave as bi-stable systems that tend to be composed by individuals belonging to one species of the interacting pair [62]. Delayed sympatry and the creation of narrow contacts are relatively common phenomena in European butterflies [63–65] and mostly occur among closely related species because of reproductive interference and competition [66–68]. The mito-morphological discordance suggests that, despite the structural mismatch in genitalia, mating may result in successful sperm transfer and egg fertilization. Nevertheless, this phenomenon is restricted to a narrow contact zone where discordances between morphology, COI and nuclear markers can be found.

The divergence time estimated for *M. athalia* and *M. celadussa* (~2.2 Ma) [20] coincides nearly to onset of the Pleistocene glaciations (~2.6 Ma) and the repeated glacial cycles covering Northern Europe and most of the Alps with ice sheets have probably played a fundamental role in shaping the current distribution of these species. Presumably, the region was recently (re)colonized by *M. athalia* and *M. celadussa* post-glacially [18,69]. This evidence favours the secondary contact hypothesis, indicating that the initial differentiation likely happened in allopatry, although the process of differentiation may have included other contacts in previous interglacial periods. In the most probable scenario, the taxa examined here were fragmented into three main refugia in Europe along several glacial cycles: (i) Iberia and the Italian Peninsula for *M. celadussa*; (ii) the Balkan Peninsula for the Balkan lineage; and (iii) an eastern refugium for the main *M. athalia*, as suggested by the large distribution in Russia and Asia of this COI lineage (Figure S6). In agreement to this hypothesis, a recent study comparing the climatic niches of a series of European cryptic taxa showed that *M. athalia* has a climatic niche significantly shifted toward a colder and drier climate, typical of steppe areas, compared to *M. celadussa* [17]. At the end of the last glacial period, the Asiatic *M. athalia* could have been favoured in colonizing the temperate plains of central and northern Europe, reaching Scandinavia, due to the climatic similarity with its original habitat and, possibly, also due to a stronger propagule pressure. Concurring with the post-glacial range expansion, these populations of *M. athalia* came into secondary contact with *M. celadussa* and hybridized, although the existence of previously admixed populations cannot be ruled out. They also came into secondary contact with the Balkan lineage, apparently with a lower degree of hybridization despite morphological similarity, although here the limited sample size needs to be taken into account. Nevertheless, the rapid post-glacial expansion of the eastern *M. athalia* seems to have produced contact zones and blocked the expansion of the other two lineages out of their glacial refugia, namely *M. celadussa* in western Europe and the Balkan lineage in eastern Europe.

4.4. Species Delimitation

Morphological differentiation of genitalia often seems to have taken place along the process of speciation in European species pairs with highly similar external morphology [70–72], although cases having virtually indistinguishable genitalia are also known [70,73,74]. The biological species concept, in particular, largely relies on criteria that do not allow splitting populations into separate species until their genetic differentiation is significant enough to impede gene flow and consequent merging of populations [75]. On the other hand, natural hybridization between closely related species has been reported to occur in up to 10% of animal species [76] and in about 16% of European butterflies [77]. Hybridization and resulting introgression is particularly frequent between parapatric taxa. An increasing

body of evidence suggests that introgression plays an important role in the process of speciation e.g., [63,78]. This observation led to the introduction of alternative definitions of the species concept. For example, the differential fitness concept identifies species as “groups of individuals that are reciprocally characterized by features that would have negative fitness effects in other groups and that cannot be regularly exchanged between groups upon contact” [79]. Therefore, critical questions to address are: (1) have the taxa developed barriers to gene flow, (2) if present, are hybrids having reduced fitness, and (3) what kind of patterns are expressed by other parapatric and sympatric European butterfly species?

According to our data, current hybridization is detected in the contact zone of main *M. athalia* and *M. celadussa*. While no mito-morphological discordance was detected outside the contact zone, STRUCTURE suggested widespread admixture in the nuclear genome of main *M. athalia* and *M. celadussa*, albeit to a lesser extent in the latter species for $K = 2$ and much lower admixture for $K = 3$. Should we therefore consider *M. athalia* and *M. celadussa* as valid at species or subspecies level, or consider them synonymous? By strictly relying on the reproductive isolation or on reciprocal phylogenetic monophyly, it would be hard to consider *M. celadussa* as a distinct species from *M. athalia*. On the other hand, the two taxa appear as distinct ecological entities (species) because the two main lineages are adapted to different climatic niches [17,80]. The answer is semantic in nature and depends upon the species concept applied. Whatever the solution over the species delimitation of parapatric and allopatric taxa will be, it is bound to be inherently arbitrary, for which reason more efforts should be made to find a consensus over the standards of delimitation of species under various evolutionary circumstances.

The Balkan lineage shows a different pattern. Despite the genetic divergence (Figure 2) and lack of introgression with main *M. athalia* (genetic structure at $K = 2$), we did not find differences in morphological traits in our samples. Moreover, due to the near sympatry with main *M. athalia*, their climatic niches likely largely overlap. While the potential drift could have played a role in the differentiation of the Balkan lineage, it might also represent another case of cryptic species among European butterflies. However, the sampling for this lineage is insufficient and this question will be addressed in a dedicated study.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12122009/s1>, Figure S1: Neighbour joining tree based on COI of specimens used in this study, Figure S2: Examples of male genitalia of *M. athalia* and *M. celadussa*, Figure S3: Scatterplot showing the relation between genitalic features measurements, Figure S4: PCA biplot showing the distribution of specimens and variables, Figure S5: STRUCTURE results, Figure S6: PCoA of COI data, Figure S7. The principal component analyses (PCAs) of summary statistics from DIYABC analysis, Table S1: *Melitaea* specimens used in this study, Table S2: Measurements of diagnostic male genitalia features in *M. athalia*-*M. celadussa*, Table S3: Summary statistics of ddRAD analysis in *M. athalia*-*M. celadussa*, Table S4: Posterior probability values of estimated parameters for the fittest scenario by DIYABC, Table S5: The specimens used for the COI analysis.

Author Contributions: Conceptualization, M.M., V.D. and R.V.; Data curation, V.D. and M.S.T.; Formal analysis, M.S.T., V.D. and L.D.; Funding acquisition, M.M.; Investigation, M.S.T., M.J. and M.H.; Methodology, M.S.T., V.D., K.M.L., R.V. and L.D.; Project administration, M.S.T.; Resources, V.D., R.V., L.D., S.S., P.H. and M.M.; Supervision, V.D., K.M.L. and M.M.; Validation, K.M.L. and L.D.; Visualization, M.S.T.; Writing—original draft, M.S.T.; Writing—review and editing, M.S.T., V.D., K.M.L., R.V., M.J., M.H., L.D., S.S., P.H. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Academy of Finland, grant numbers 314702 and 328895; the Kvantum Institute; the Ministerio de Ciencia e Innovación of Spain, grant number PID2019-107078GB-I00/AEI/10.13039/501100011033; and supported by the project “Ricerca e conservazione sugli Impollinatori dell’Arcipelago Toscano e divulgazione sui Lepidotteri del parco”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The *Melitaea* demultiplexed fastq data are archived in NCBI SRA: PRJNA638526. All 71 COI sequences are available in DS-MELITAEA on BOLD (dx.doi.org/10.5883/DS-MELITAEA) at <https://www.boldsystems.org/> (accessed on 29 November 2021). The Jupyter cookbook script for Figure 3 is publicly available at <https://doi.org/10.5281/zenodo.5762221> (accessed on 6 December 2021).

Acknowledgments: The authors wish to acknowledge CSC–IT Centre for Science, Finland, for providing computational resources. Some post processing analyses (i.e., jupyter, DIYABC) were carried out on the personal server SG-2 (Ubuntu 18.04.5 LTS) authorized by Markku Stenberg. We thank Marianne Espeland for providing us a specimen of *M. caucasogenita/athalia* for this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Sei, M.; Porter, A.H. Delimiting species boundaries and the conservation genetics of the endangered maritime ringlet butterfly (*Coenonympha nipisiquit* McDunnough). *Mol. Ecol.* **2007**, *16*, 3313–3325. [CrossRef]
- Abbott, R.; Albach, D.; Ansell, S.; Arntzen, J.W.; Baird, S.J.E.; Bierne, N.; Boughman, J.; Brelsford, A.; Buerkle, C.A.; Buggs, R.; et al. Hybridization and speciation. *J. Evol. Biol.* **2013**, *26*, 229–246. [CrossRef]
- Fišer, C.; Robinson, C.T.; Malard, F. Cryptic species as a window into the paradigm shift of the species concept. *Mol. Ecol.* **2018**, *27*, 613–635. [CrossRef] [PubMed]
- De Queiroz, K. Species Concepts and Species Delimitation. *Syst. Biol.* **2007**, *56*, 879–886. [CrossRef]
- Eberhard, W.G. *Sexual Selection and Animal Genitalia*; Harvard University Press: Cambridge, MA, USA, 1985; Volume 244.
- Mikkola, K. The lock-and-key mechanisms in the internal genitalia of the noctuid and geometrid moths (Lepidoptera) in relation to the speciation concepts. *Folia Baeriana* **1993**, *6*, 149–157.
- Ennos, R.A.; French, G.C.; Hollingsworth, P.M. Conserving taxonomic complexity. *Trends Ecol. Evol.* **2005**, *20*, 164–168. [CrossRef] [PubMed]
- Wheeler, Q.D.; Raven, P.H.; Wilson, E.O. Taxonomy: Impediment or Expedient? *Science* **2004**, *303*, 285. [CrossRef]
- Sites, J.W.; Marshall, J.C. Delimiting species: A Renaissance issue in systematic biology. *Trends Ecol. Evol.* **2003**, *18*, 462–470. [CrossRef]
- Mutanen, M.; Kaitala, A.; Mönkkönen, M. Genital variation within and between three closely related Euxoa moth species: Testing the lock-and-key hypothesis. *J. Zool.* **2006**, *268*, 109–119. [CrossRef]
- Flantua, S.G.A.; Payne, D.; Borregaard, M.K.; Beierkuhnlein, C.; Steinbauer, M.J.; Dullinger, S.; Essl, F.; Irl, S.D.H.; Kienle, D.; Kreft, H. Snapshot isolation and isolation history challenge the analogy between mountains and islands used to understand endemism. *Glob. Ecol. Biogeogr.* **2020**, *29*, 1651–1673. [CrossRef]
- Knief, U.; Bossu, C.M.; Saino, N.; Hansson, B.; Poelstra, J.; Vijay, N.; Weissensteiner, M.; Wolf, J.B.W. Epistatic mutations under divergent selection govern phenotypic variation in the crow hybrid zone. *Nat. Ecol. Evol.* **2019**, *3*, 570–576. [CrossRef]
- Leaché, A.D.; Fujita, M.K.; Minin, V.N.; Bouckaert, R.R. Species delimitation using genome-wide SNP Data. *Syst. Biol.* **2014**, *63*, 534–542. [CrossRef] [PubMed]
- Eberle, J.; Ahrens, D.; Mayer, C.; Niehuis, O.; Misof, B. A plea for standardized nuclear markers in metazoan DNA taxonomy. *Trends Ecol. Evol.* **2020**, *35*, 336–345. [CrossRef] [PubMed]
- Higgins, L.G. A descriptive catalogue of the genus *Mellicta* Billberg (Lepidoptera: Nymphalidae) and its species, with supplementary notes on the genera *Melitaea* and *Euphydryas*. *Trans. R. Entomol. Soc. Lond.* **1955**, *106*, 1–127. [CrossRef]
- Van Oorschot, H.; Coutsis, J.G. *The Genus Melitaea Fabricius, 1807: Taxonomy and Systematics with Special Reference to the Male Genitalia (Lepidoptera, Nymphalidae, Nymphalinae)*; Tshikolovets Publications: Pardubice, Czech Republic, 2014; ISBN 8090490042.
- Platania, L.; Menchetti, M.; Dincă, V.; Corbella, C.; Kay-Lavelle, I.; Vila, R.; Wiemers, M.; Schweiger, O.; Dapporto, L. Assigning occurrence data to cryptic taxa improves climatic niche assessments: Biodecrypt, a new tool tested on European butterflies. *Glob. Ecol. Biogeogr.* **2020**, *29*, 1852–1865. [CrossRef]
- Dincă, V.; Dapporto, L.; Somervuo, P.; Vodă, R.; Cuvelier, S.; Gascoigne-Pees, M.; Huemer, P.; Mutanen, M.; Hebert, P.D.N.; Vila, R. High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. *Commun. Biol.* **2021**, *4*, 1–11. [CrossRef] [PubMed]
- Higgins, L.G. *Classification of European Butterflies*; Collins: New York, NY, USA, 1975; ISBN 0002196247.
- Leneveu, J.; Chichvarkhin, A.; Wahlberg, N. Varying rates of diversification in the genus *Melitaea* (Lepidoptera: Nymphalidae) during the past 20 million years. *Biol. J. Linn. Soc.* **2009**, *97*, 346–361. [CrossRef]
- Dincă, V.; Montagud, S.; Talavera, G.; Hernández-Roldán, J.; Munguira, M.L.; García-Barros, E.; Hebert, P.D.N.; Vila, R. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Sci. Rep.* **2015**, *5*, 12395. [CrossRef] [PubMed]
- Wiemers, M.; Balletto, E.; Dincă, V.; Fric, Z.F.; Lamas, G.; Lukhtanov, V.; Munguira, M.L.; van Swaay, C.A.M.; Vila, R.; Vliegenthart, A. An updated checklist of the European butterflies (Lepidoptera, Papilionoidea). *Zookeys* **2018**, *811*, 9–45. [CrossRef]

23. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2014**, *30*, 614–620. [[CrossRef](#)] [[PubMed](#)]
24. Eaton, D.A.R.; Overcast, I. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* **2019**, *36*, 2592–2594. [[CrossRef](#)]
25. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; Von Haeseler, A.; Jermini, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [[CrossRef](#)]
26. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2014**, *32*, 268–274. [[CrossRef](#)]
27. Hoang, D.T.; Chernomor, O.; Von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522. [[CrossRef](#)] [[PubMed](#)]
28. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945–959. [[CrossRef](#)] [[PubMed](#)]
29. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)] [[PubMed](#)]
30. Earl, D.A. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
31. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [[CrossRef](#)]
32. Rosenberg, N.A. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* **2004**, *4*, 137–138. [[CrossRef](#)]
33. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
34. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
35. Goudet, J. Hierstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186. [[CrossRef](#)]
36. Estoup, A.; Cornuet, J.-M.; Dehne-Garcia, A.; Loire, E. *New Version of the Computer Program DIYABC (DIYABC v2. 1.0): A User-Friendly Approach to Approximate Bayesian Computation for Inference on Population History Using Molecular Markers*; ResearchGate: Berlin, Germany, 2015.
37. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [[CrossRef](#)]
38. Van Etten, J. R Package gdistance: Distances and Routes on Geographical Grids. *J. Stat. Softw.* **2017**, *76*, 1–21. [[CrossRef](#)]
39. Shafer, A.B.A.; Peart, C.R.; Tusso, S.; Maayan, I.; Brelsford, A.; Wheat, C.W.; Wolf, J.B.W. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol. Evol.* **2017**, *8*, 907–917. [[CrossRef](#)]
40. Huemer, P.; Wiesmair, B. DNA-Barcoding der Tagfalter (Lepidoptera, Papilionoidea) Österreichs. Unbekannte genetische Vielfalt im Zentrum Österreichs. *Wiss. J. Tirol. Landesmuseen* **2017**, *10*, 8–33.
41. Litman, J.; Chittaro, Y.; Birrer, S.; Praz, C.; Wermeille, E.; Fluri, M.; Stalling, T.; Schmid, S.; Wyler, S.; Gonseth, Y. A DNA barcode reference library for Swiss butterflies and forester moths as a tool for species identification, systematics and conservation. *PLoS ONE* **2018**, *13*, e0208639. [[CrossRef](#)]
42. Ahtelik, G. Molekularbiologische Analyse der genetischen Diversität des Melitaea Athalia/Celadussa-Komplexes (Lepidoptera: Nymphalidae) unter Anwendung der ISSR-PCR auf Art-, Unterart- und Populationsebene. Ph.D. Thesis, Ruhr-Universität, Bochum, Germany, 2006.
43. Taberlet, P.; Fumagalli, L.; Wust-Saucy, A.; Cosson, J. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* **1998**, *7*, 453–464. [[CrossRef](#)]
44. Hewitt, G. The genetic legacy of the Quaternary ice ages. *Nature* **2000**, *405*, 907–913. [[CrossRef](#)]
45. Toews, D.P.L.; Brelsford, A. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **2012**, *21*, 3907–3930. [[CrossRef](#)]
46. Dapporto, L.; Vodá, R.; Dincă, V.; Vila, R. Comparing population patterns for genetic and morphological markers with uneven sample sizes. An example for the butterfly *Maniola jurtina*. *Methods Ecol. Evol.* **2014**, *5*, 834–843. [[CrossRef](#)]
47. Rosser, N.; Shirai, L.T.; Dasmahapatra, K.K.; Mallet, J.; Freitas, A.V.L. The Amazon river is a suture zone for a polyphyletic group of co-mimetic heliconiine butterflies. *Ecography* **2021**, *44*, 177–187. [[CrossRef](#)]
48. De La Torre, A.; Ingvarsson, P.K.; Aitken, S.N. Genetic architecture and genomic patterns of gene flow between hybridizing species of *Picea*. *Heredity* **2015**, *115*, 153–164. [[CrossRef](#)]
49. Mavárez, J.; Salazar, C.A.; Bermingham, E.; Salcedo, C.; Jiggins, C.D.; Linares, M. Speciation by hybridization in *Heliconius* butterflies. *Nature* **2006**, *441*, 868–871. [[CrossRef](#)] [[PubMed](#)]
50. Salazar, C.; Jiggins, C.D.; Taylor, J.E.; Kronforst, M.R.; Linares, M. Gene flow and the genealogical history of *Heliconius heurippa*. *BMC Evol. Biol.* **2008**, *8*, 132. [[CrossRef](#)]
51. Pritchard, V.L.; Dimond, L.; Harrison, J.S.; Velázquez, C.C.S.; Zieba, J.T.; Burton, R.S.; Edmands, S. Interpopulation hybridization results in widespread viability selection across the genome in *Tigriopus californicus*. *BMC Genet.* **2011**, *12*, 54. [[CrossRef](#)]
52. Nadeau, N.J.; Salazar, P.; Morrison, A.; Jiggins, C.D.; Ruiz, M.; Papa, R.; Counterman, B.; Medina, J.A.; Ortiz-Zuazaga, H.; McMillan, W.O. Population genomics of parallel hybrid zones in the mimetic butterflies, *H. melpomene* and *H. erato*. *Genome Res.* **2014**, *24*, 1316–1333. [[CrossRef](#)]

53. Dincă, V.; Lee, K.M.; Vila, R.; Mutanen, M. The conundrum of species delimitation: A genomic perspective on a mitogenetically super-variable butterfly. *Proc. R. Soc. B* **2019**, *286*, 20191311. [[CrossRef](#)] [[PubMed](#)]
54. Koren, T.; Beretta, S.; Črne, M.; Verovnik, R. On the distribution of *Pyrgus malvoides* (Elwes & Edwards, 1897)(Lepidoptera: Hesperidae) at the eastern part of its range. *Entomol. Gaz.* **2013**, *64*, 225–234.
55. Lin, D.; Bi, K.; Conroy, C.J.; Lacey, E.A.; Schraiber, J.G.; Bowie, R.C.K. Mito-nuclear discordance across a recent contact zone for California voles. *Ecol. Evol.* **2018**, *8*, 6226–6241. [[CrossRef](#)] [[PubMed](#)]
56. Larmuseau, M.H.D.; Raeymaekers, J.A.M.; Hellemans, B.; Van Houdt, J.K.J.; Volckaert, F.A.M. Mito-nuclear discordance in the degree of population differentiation in a marine goby. *Heredity* **2010**, *105*, 532–542. [[CrossRef](#)] [[PubMed](#)]
57. Marshall, J.C.; Sites, J.W., Jr. A comparison of nuclear and mitochondrial cline shapes in a hybrid zone in the *Sceloporus grammicus* complex (Squamata; Phrynosomatidae). *Mol. Ecol.* **2001**, *10*, 435–449. [[CrossRef](#)] [[PubMed](#)]
58. Ivanov, V.; Lee, K.M.; Mutanen, M. Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Mol. Ecol.* **2018**, *27*, 1681–1695. [[CrossRef](#)] [[PubMed](#)]
59. Mullen, S.P.; Dopman, E.B.; Harrison, R.G. Hybrid zone origins, species boundaries, and the evolution of wing-pattern diversity in a polytypic species complex of North American admiral butterflies (Nymphalidae: Limenitis). *Evolution* **2008**, *62*, 1400–1417. [[CrossRef](#)] [[PubMed](#)]
60. Gavrilets, S.; Li, H.; Vose, M.D. Patterns of parapatric speciation. *Evolution* **2000**, *54*, 1126–1134. [[CrossRef](#)]
61. Lukhtanov, V.A.; Kandul, N.P.; Plotkin, J.B.; Dantchenko, A.V.; Haig, D.; Pierce, N.E. Reinforcement of pre-zygotic isolation and karyotype evolution in *Agrodiaetus* butterflies. *Nature* **2005**, *436*, 385–389. [[CrossRef](#)]
62. Barton, N.H.; Turelli, M. Spatial waves of advance with bistable dynamics: Cytoplasmic and genetic analogues of Allee effects. *Am. Nat.* **2011**, *178*, E48–E75. [[CrossRef](#)]
63. Dincă, V.; Wiklund, C.; Lukhtanov, V.A.; Kodandaramaiah, U.; Norén, K.; Dapporto, L.; Wahlberg, N.; Vila, R.; Friberg, M. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. *J. Evol. Biol.* **2013**, *26*, 2095–2106. [[CrossRef](#)]
64. Habel, J.C.; Vila, R.; Vodă, R.; Husemann, M.; Schmitt, T.; Dapporto, L. Differentiation in the marbled white butterfly species complex driven by multiple evolutionary forces. *J. Biogeogr.* **2017**, *44*, 433–445. [[CrossRef](#)]
65. Platania, L.; Vodă, R.; Dincă, V.; Talavera, G.; Vila, R.; Dapporto, L. Integrative analyses on Western Palearctic *Lasiommata* reveal a mosaic of nascent butterfly species. *J. Zool. Syst. Evol. Res.* **2020**, *58*, 809–822. [[CrossRef](#)]
66. Pigot, A.L.; Tobias, J.A. Species interactions constrain geographic range expansion over evolutionary time. *Ecol. Lett.* **2013**, *16*, 330–338. [[CrossRef](#)]
67. Waters, J.M.; Fraser, C.I.; Hewitt, G.M. Founder takes all: Density-dependent processes structure biodiversity. *Trends Ecol. Evol.* **2013**, *28*, 78–85. [[CrossRef](#)]
68. Vodă, R.; Dapporto, L.; Dincă, V.; Vila, R. Why do cryptic species tend not to co-occur? A case study on two cryptic pairs of butterflies. *PLoS ONE* **2015**, *10*, e0117802. [[CrossRef](#)]
69. Schmitt, T.; Louy, D.; Zimmermann, E.; Habel, J.C. Species radiation in the Alps: Multiple range shifts caused diversification in Ringlet butterflies in the European high mountains. *Org. Divers. Evol.* **2016**, *16*, 791–808. [[CrossRef](#)]
70. Dincă, V.; Zakharov, E.V.; Hebert, P.D.N.; Vila, R. Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 347–355. [[CrossRef](#)] [[PubMed](#)]
71. Gratton, P.; Trucchi, E.; Trasatti, A.; Riccarducci, G.; Marta, S.; Allegrucci, G.; Cesaroni, D.; Sbordoni, V. Testing classical species properties with contemporary data: How “bad species” in the brassy ringlets (*Erebia tyndarus* complex, Lepidoptera) turned good. *Syst. Biol.* **2016**, *65*, 292–303. [[CrossRef](#)]
72. Zinetti, F.; Dapporto, L.; Vovlas, A.; Chelazzi, G.; Bonelli, S.; Balletto, E.; Ciofi, C. When the rule becomes the exception. No evidence of gene flow between two *Zerynthia* cryptic butterflies suggests the emergence of a new model group. *PLoS ONE* **2013**, *8*, e65746. [[CrossRef](#)]
73. Dainou, K.; Blanc-Jolivet, C.; Degen, B.; Kimani, P.; Ndiade-Bourobou, D.; Donkpegan, A.S.L.; Tosso, F.; Kaymak, E.; Bourland, N.; Doucet, J.-L. Revealing hidden species diversity in closely related species using nuclear SNPs, SSRs and DNA sequences—A case study in the tree genus *Milicia*. *BMC Evol. Biol.* **2016**, *16*, 259. [[CrossRef](#)]
74. Hernández-Roldán, J.L.; Dapporto, L.; Dincă, V.; Vicente, J.C.; Hornett, E.A.; Šichová, J.; Lukhtanov, V.A.; Talavera, G.; Vila, R. Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Mol. Ecol.* **2016**, *25*, 4267–4284. [[CrossRef](#)] [[PubMed](#)]
75. Zachos, F.E. An Annotated List of Species Concepts. In *Species Concepts in Biology*; Springer International Publishing: Cham, Switzerland, 2016; pp. 77–96.
76. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **2005**, *20*, 229–237. [[CrossRef](#)] [[PubMed](#)]
77. Descimon, H.; Mallet, J. Bad Species. In *Ecology of Butterflies in Europe*; Settele, J., Shreeve, T., Konvič, M., Van Dyck, H., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 219–249.
78. Mallet, J.; Beltrán, M.; Neukirchen, W.; Linares, M. Natural hybridization in heliconiine butterflies: The species boundary as a continuum. *BMC Evol. Biol.* **2007**, *7*, 1–16. [[CrossRef](#)]
79. Hausdorf, B. Progress toward a general species concept. *Evol. Int. J. Org. Evol.* **2011**, *65*, 923–931. [[CrossRef](#)] [[PubMed](#)]
80. Rundle, H.D.; Nosil, P. Ecological speciation. *Ecol. Lett.* **2005**, *8*, 336–352. [[CrossRef](#)]