



OPEN

Wolbachia affects mitochondrial population structure in two systems of closely related Palaearctic blue butterflies

Alena Sucháčková Bartoňová^{1,2}, Martin Konvička^{1,2}, Jana Marešová^{1,2}, Martin Wiemers³, Nikolaj Ignatev^{1,2}, Niklas Wahlberg⁴, Thomas Schmitt^{3,5} & Zdeněk Faltýnek Fric¹

The bacterium *Wolbachia* infects many insect species and spreads by diverse vertical and horizontal means. As co-inherited organisms, these bacteria often cause problems in mitochondrial phylogeny inference. The phylogenetic relationships of many closely related Palaearctic blue butterflies (Lepidoptera: Lycaenidae: Polyommatainae) are ambiguous. We considered the patterns of *Wolbachia* infection and mitochondrial diversity in two systems: *Aricia agestis*/*Aricia artaxerxes* and the *Pseudophilotes baton* species complex. We sampled butterflies across their distribution ranges and sequenced one butterfly mitochondrial gene and two *Wolbachia* genes. Both butterfly systems had uninfected and infected populations, and harboured several *Wolbachia* strains. *Wolbachia* was highly prevalent in *A. artaxerxes* and the host's mitochondrial structure was shallow, in contrast to *A. agestis*. Similar bacterial alleles infected both *Aricia* species from nearby sites, pointing to a possible horizontal transfer. Mitochondrial history of the *P. baton* species complex mirrored its *Wolbachia* infection and not the taxonomical division. *Pseudophilotes baton* and *P. vicrama* formed a hybrid zone in Europe. *Wolbachia* could obscure mitochondrial history, but knowledge on the infection helps us to understand the observed patterns. Testing for *Wolbachia* should be routine in mitochondrial DNA studies.

Wolbachia is an intracellular bacterium massively infecting arthropods and filarial nematodes. As a maternally inherited symbiont, *Wolbachia* facilitates its spread by various methods of reproductive manipulation, such as cytoplasmic (sperm-egg) incompatibility, male killing, feminization, or parthenogenesis^{1–3}. *Wolbachia* can be transferred also horizontally, using diverse vectors such as shared parasitoids, predators, or host plants^{4–6}.

Wolbachia is passed to the next generation in a similar way as mitochondria. This fact causes selective sweeps in haplotypes due to genetic hitchhiking, followed by reduced mitochondrial diversity^{7–9}. The subsequent effects are mito-nuclear discordance¹⁰ and deep divergences in mitochondrial phylogenies^{11,12}. These may result in failures of mitochondrial DNA barcoding (i.e. using a short standardized DNA sequence) in the identification of species^{1,13}.

Through these means, *Wolbachia* could interfere in diversification processes, but its role in arthropod evolution is still unclear. Speciation processes arise via the evolution of genetic barriers between populations¹⁴. When two different *Wolbachia* strains infect a host population, bidirectional cytoplasmic incompatibility may arise, leading to divergence¹⁵. On the other hand, when an infected and uninfected population meet, it can lead to a more common unidirectional cytoplasmic incompatibility. In such a case, the bacteria are predicted to spread faster than a gene flow barrier could evolve¹⁶. However, in theory, under specific circumstances of infection rate and migration among populations, even the unidirectional cytoplasmic incompatibility could lead to host genetic divergence^{17,18}.

The existence of common horizontal transmission of *Wolbachia* between host species is supported by a general lack of congruence between the bacterial and host phylogenies, where either similar strains appear in phylogenetically distinct taxa, or closely related host taxa harbour distinct strains^{19,20}.

¹Biology Centre CAS, Institute of Entomology, České Budějovice, Czech Republic. ²Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic. ³Senckenberg German Entomological Institute, Müncheberg, Germany. ⁴Department of Biology, Lund University, Lund, Sweden. ⁵Faculty of Natural Sciences I, Institute of Biology, Zoology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany. ✉email: al.bartonova@gmail.com

Lycaenid butterflies (Lepidoptera: Lycaenidae) are known to be infected by the highest number of *Wolbachia* strains among butterfly families²¹. Within the lycaenids, Polyommatae (the blues) often harbour unresolved phylogenetic relationships on lower taxonomical levels^{22,23}, which may be further complicated by *Wolbachia* infection. *Wolbachia* may play an important role in butterfly mitochondrial structure, possible hybridization and even evolution. While its importance is becoming evident^{24,25}, studies targeting patterns of *Wolbachia* infection across larger distribution areas are still scarce.

To shed some light on the complexity of infection by *Wolbachia* in Polyommatae, we studied two systems of closely related blue butterflies, the widely distributed Palaearctic genera *Aricia* Reichenbach, 1817, and *Pseudophilotes* Beuret, 1958. The taxonomy of these systems relies on often subtle differences in morphology, mitochondrial DNA data, life histories, or habitat diversification^{26–30}. The few nuclear markers classically used in butterfly phylogeny often fail to distinguish these species^{28,29}. We focused on the taxa co-existing in Central Europe: *Aricia agestis* (Denis & Schiffermüller, 1775) and *A. artaxerxes* (Fabricius, 1793), as well as *Pseudophilotes baton* (Bergsträsser, 1779) and *P. vicrama* (Moore, 1865), and, for the latter genus, their closest relatives from other areas (i.e. the *Pseudophilotes baton* species complex).

We investigated the presence of *Wolbachia* infection in these butterflies using samples across their whole distributional ranges. We connected the *Wolbachia* infection patterns to the butterfly population structures inferred by a co-inherited marker, i.e. mitochondrial DNA. Our study will provide background information for future analyses of speciation, not only in the highly diverse group of Polyommatae.

Methods

Study models. *Aricia agestis* and *A. artaxerxes* are cryptic species, which can be distinguished by mtDNA barcoding, but not by adult morphology in most of the area of their overlapping ranges^{26,30}. Minor differences are in larval and pupal morphology³¹. *Aricia agestis* is a western Palaearctic species, inhabiting a wide range of mesic to xeric grasslands in lowlands of the Mediterranean and temperate regions. *Aricia artaxerxes* is a Palaearctic species reaching Northern Europe and inhabiting calcareous short-turf grasslands at higher elevations in the south. Their elevational ranges overlap in Central Europe³⁰. They differ in the number of annual generations (multiple in *A. agestis*, single in *A. artaxerxes*). Host plants of both species are various Geraniaceae and *Helianthemum* Mill. spp. Several of their close relatives occur throughout Eurasia^{28,32}, with west-Mediterranean *Aricia montensis* (Verity, 1928) being sister either to *A. artaxerxes*²⁸ or *A. agestis*³³.

Pseudophilotes baton and *P. vicrama* from the *P. baton* species complex are cryptic vicariant species: *P. baton* inhabits Western Europe (including the Alps and Italy) and *P. vicrama* occurs from Central Europe to Altai, western China and northern Mongolia. Both species are (or were in the recent past) present in Central Europe (Germany, Czech Republic, and Austria). They differ in the shape of male genitalia, which is used as a diagnostic trait³⁴, but they share a similar mtDNA barcode in some cases²⁹. Both inhabit xerothermic short turf grasslands, steppe-like sites in the case of *P. vicrama*, and pastures in mountains and their foothills in *P. baton*. Host plants are various species of *Thymus* L. There are several other taxa in the *Pseudophilotes baton* complex. *P. baton jacuticus* Korshunov et Viidalepp, 1980 inhabits steppe areas from the eastern and northern Baikal Lake surroundings to central and southern Yakutia, Russia. *Pseudophilotes panoptes* (Hübner, 1813) inhabits dry scrubby places in Iberia. *Pseudophilotes sinaicus* Nakamura, 1975 is present at a small area in the Sinai Peninsula Mountains (Egypt).

Sampling and sequencing. To get sequences of mitochondrial DNA and examine the patterns of *Wolbachia* infection, we collected tissue samples of 115 individuals of *A. agestis*, 68 *Aricia artaxerxes*, 23 *Pseudophilotes baton*, 88 *P. vicrama* and 10 individuals of other *Pseudophilotes* (*P. sinaicus*: 4, *P. b. jacuticus*: 4, *P. panoptes*: 2) (Supplementary Dataset S1). In the case of *Pseudophilotes*, we dissected four male individuals from populations with mitochondrial sequences assigning them to a different species (Croatia, France), and 20 male individuals from all parts of the species complex's distributional range to illustrate their diversity. The abdomens of male individuals were boiled for 5 min in 10% KOH to get the genital apparatus.

We extracted DNA with the Genomic DNA Mini Kit—Tissue (Geneaid) following the manufacturer's protocols. To check the quality of extracted DNA and compare the mitochondrial structure with *Wolbachia* infection, we amplified, using polymerase chain reaction (PCR), the 5' part of *cytochrome c oxidase subunit I* gene (COI), which is the most frequently used barcode in animals and hence a large reference library is available. We used the primer pair LCO/HCO and, when not successful, two pairs: LCO/K699 and Ron/HCO (primers:^{35,36}). Only samples with a clear COI bands on agarose gels and obtained full COI sequence were included in the study.

We screened for the presence of *Wolbachia* in the host DNA samples by amplifying the *Wolbachia surface protein* gene (wsp) with the primer pair 81F and 691R³⁷. Wsp is the common marker used to detect the presence of the bacteria (e.g.³⁸). We included a negative control (PCR mixture with PCR H₂O instead of DNA sample) and a positive control (a sample of *Pseudophilotes bavius* (Eversmann, 1832) with already sequenced wsp gene) in each PCR. To ascertain the presence or absence of *Wolbachia*, we ran 5 µl of the PCR product on 1.5% agarose gels. In the case of presence of a band ~ 700 bp long, the sample was scored as a positive and sequenced; otherwise it was scored as a negative. All samples were screened twice with wsp and sequenced once. In positive DNA samples, the *Wolbachia* gene *cytochrome c oxidase subunit I* (coxA) from *Wolbachia* multilocus sequence typing was amplified using primers coxA_F1 and coxA_R1³⁹.

For all markers, we prepared PCR in 12.5 µl volume (6.25 µl Bioline 2 × MyTaq HS Red Mix, 4 µl PCR H₂O, 0.625 + 0.625 µl primers; 1 µl DNA). The thermal cycling profile was 95 °C for 5 min; then 40 cycles of 94 °C for 30 s, 50 °C (COI)/55 °C (wsp, coxA) for 30 s, 72 °C for 90 s; and final extension 72 °C for 10 min. All the forward primers had T7 promoter and the reverse primers T3 universal tails attached.

We cleaned PCR products with enzymes FastAP and ExoI (ThermoFisher) and sequenced them in one direction in Macrogen Inc. on ABI3730XL DNA analysers. We checked obtained sequences visually and aligned

them in Geneious v. 8.0.5⁴⁰. Wsp and coxA strains were defined using the reference sequences of the *Wolbachia* MLST database (³⁹; <https://pubmlst.org/wolbachia/>). A possible double infection of the two arthropod *Wolbachia* supergroups in one individual³⁹ may have been overlooked, but this should be rare within individuals (cf.⁴¹) and should not have any consequences on the conclusions of this study.

Molecular data analyses. To ascertain the mitochondrial structure throughout the distribution ranges, in addition to our sequences, we used mitochondrial sequences of the target taxa and other species in these genera from BOLD⁴², publicly available in November 2019 (Supplementary Dataset S1). From the COI sequences collated in this study and in BOLD, we prepared three datasets: *Aricia agestis* (381 samples), *A. artaxerxes* (163), and *Pseudophilotes baton* complex (228) (Supplementary Dataset S1). While the barcodes of *Aricia agestis* and *A. artaxerxes* displayed nucleotide differences of about 2%²⁶, the species *P. baton* and *P. vicrama*, together with the western Mediterranean *P. panoptes*, shared haplotypes²⁹, and thus we treated them together as *P. baton* species complex.

To explore the butterfly population structure, we constructed the COI haplotype network by TCS statistical parsimony algorithm⁴³ in POPART⁴⁴. To cluster the COI sequences, we employed the hierarchical Bayesian Analysis of Population Structure (hierBAPS)⁴⁵ in the package 'rhierbaps'⁴⁶ using R version 3.6.0⁴⁷, without prior information on the geographic origin of each sample. To estimate the genetic diversity and to test the neutrality of the datasets (Tajima's D, Fu and Li's D* tests), we used DNASP v.5⁴⁸.

To display the patterns of butterfly interpopulation genetic distances as genetic landscapes, we used the GDisPAL function of SPADS v. 1.0⁴⁹ in R. First, we merged the sequences into 94, 43, and 69 populations (*A. agestis*, *A. artaxerxes* and *P. baton* complex, respectively) based on the geographic proximity of their localities (Supplementary Dataset S1). Second, we generated the genetic and geographic distance matrices in SPADS v. 1.0. The GDisPAL function is based on inverse distance-weighted interpolation (IDW)⁵⁰. Here, the interpolation depends on distance values assigned at mid-points of each edge of a connectivity network built among the sampling localities (i.e., the Delaunay triangulation). Interpolation surfaces, visualised as heat maps, are affected by a distance weighting parameter *a*. We calculated the surface for *a* = 1–7. Considering potential correlation between genetic and geographic distances, we performed the distance interpolations using residual distances derived from the linear regression of genetic vs. geographic distances (inter-individual distances 1 of SPADS v. 1.0)⁵¹.

We utilized Bayesian analysis in Beast 1.10.4⁵² to unravel the mitochondrial genetic history in the two butterfly species groups and their closely related species. We transformed the samples of each species into haplotypes. Haplotypes shared among different taxonomic species entered the analysis in a number of copies corresponding to the number of these species. As outgroups, we used sequences of *Aricia anteros* (Freyer, 1838), *A. crassipuncta* (Christoph, 1893), and *A. morronensis* (Ribbe, 1910) in the *Aricia* tree, and sequences of *Pseudophilotes bavius* (Eversmann, 1832) and *P. fatma* (Oberthür, 1890) in the *Pseudophilotes* tree. This resulted in 118 samples in the *Aricia* and 71 samples in the *Pseudophilotes* input files. Beast analysis was run for 20 million MCMC generations, sampled every 5000 generations, for three independent runs. We employed the Coalescent Constant Size model with strict clock as the tree prior. Prior to the analysis, substitution models were selected in jModelTest⁵³ by the lowest BIC, i.e. GTR + I + G for *Aricia* and GTR + G for *Pseudophilotes*. The outputs of the three runs were inspected in Tracer 1.6⁵⁴. The first 10% of trees were discarded as burn-in. The trees from the three runs were combined and a single best ultrametric tree was obtained.

To explore the relationships among *Wolbachia* alleles, we used MrBayes 3.2.6⁵⁵. We prepared four datasets with all sequences available: *Aricia* wsp (91 samples), *Aricia* coxA (81 samples), *Pseudophilotes* wsp (32 samples) and *Pseudophilotes* coxA (29 samples). As outgroups we used sequences of *Wolbachia* extracted from the nematode *Litomosoides sigmodontis* (Chandler, 1931) (GenBank codes: wsp: AF409112, coxA: FJ390246). Substitution models were GTR + G in *Aricia* wsp, and HKY in *Pseudophilotes* wsp, *Aricia* coxA and *Pseudophilotes* coxA. MrBayes was run for 20 million MCMC generations, sampled every 5000 generations, with temperature = 0.2, four simultaneous chains, four independent runs. The first 10% of trees were discarded as burn-in. We inspected the convergence of the four runs and the sufficient sampling by the *sump* command based on the log likelihood and high effective sample sizes. The trees from the four runs were summarized under the 50% majority-rule consensus. We inferred Bayesian trees in the CIPRES Science Gateway⁵⁶.

Results

***Aricia*: butterfly mitochondrial structure.** In the dataset of 381 sequences of *Aricia agestis*, the TCS analysis revealed 54 haplotypes (A1–A54, 35 segregating sites) belonging to six BAPS clusters (Fig. 1a). The first cluster BAPSa1 harboured the majority of all samples, showing a star-like topology typical for recent population expansions; this cluster was found all over the distribution range of the species, except for Central Asia. The second cluster BAPSa2 was restricted to Western, Central and Southern Europe plus one sample from the Levant. The clusters BAPSa3 and BAPSa5 were found in the Mediterranean fringe of Europe, but both were relatively rare; in addition, BAPSa3 was also found from Alborz (northern Iran) to Central Asia. The BAPSa4 cluster was mostly restricted to Corsica and adjoining islands, plus one sample from Alatau (Central Asia). The BAPSa6 cluster was obtained for Sicily, the Aeolian Islands and one sample from the mainland side of the strait of Messina (haplotype A49). This genetic structure was mirrored by the genetic landscape analysis (Fig. 2a), showing the highest genetic distances along the Mediterranean fringe of Europe from southern France to the northern Balkans. The dataset's nucleotide diversity was 0.0044, Tajima's D was not significant (–1.634; *p* > 0.05), Fu and Li's D* test was significant (–3.020; *p* < 0.05). The latter test suggested the excess of singleton haplotypes.

In the dataset of 163 sequences *Aricia artaxerxes*, the TCS analysis revealed 16 haplotypes (X1–X16, 15 segregating sites) belonging to a single BAPSx1 cluster (Fig. 1b). The genetic landscape showed the highest distances among the samples from the Western Alps, in the Caucasus and in Central Asia (Fig. 2b). A barrier appeared also

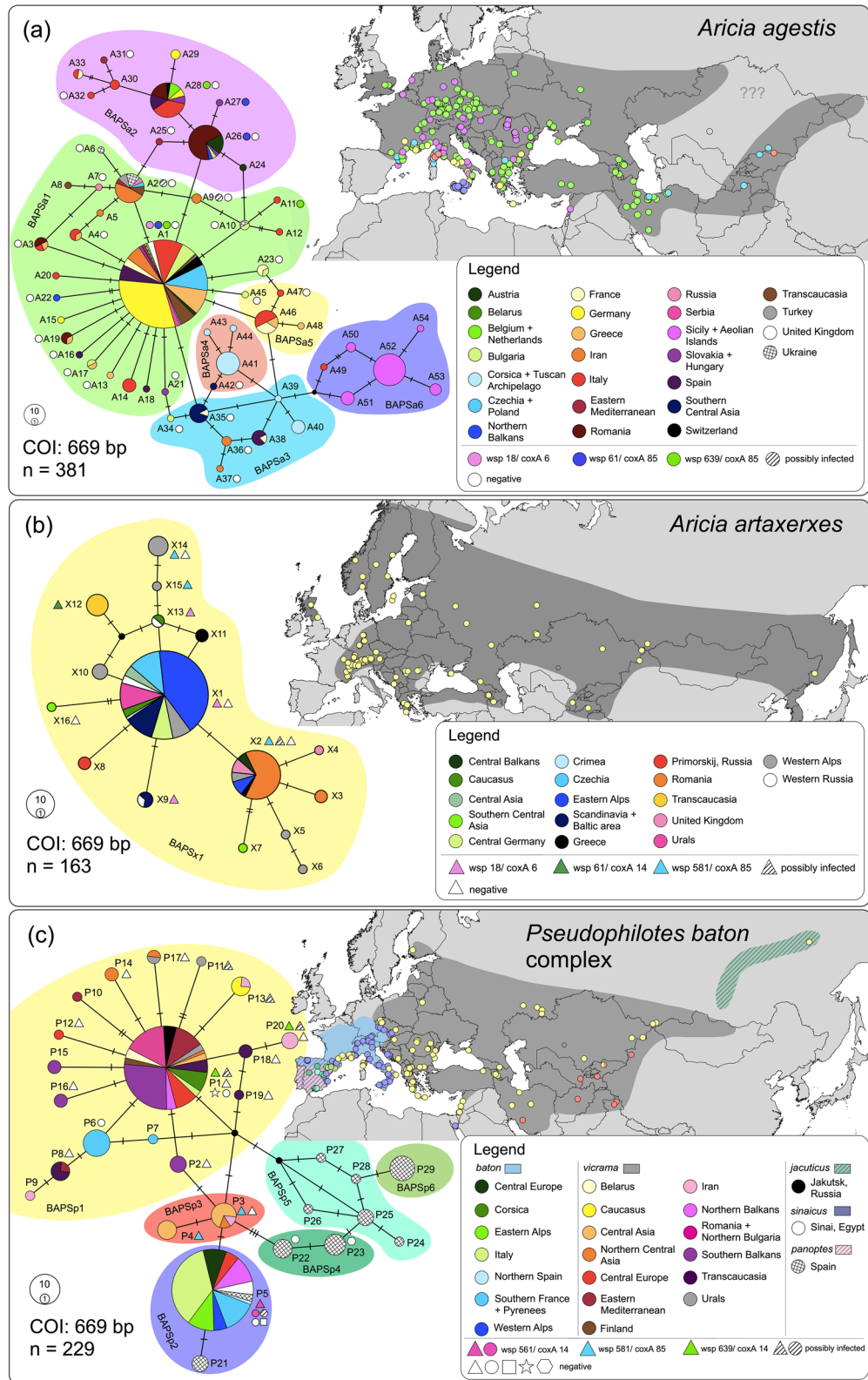


Figure 1. TCS haplotype networks of (a) *Aricia agestis*, (b) *Aricia artaxerxes* and (c) *Pseudophilotes baton* species complex, based on COI. The colours of map points correspond to BAPS clusters (coloured network background). The shades on the maps represent the distribution of each taxon. The specimens in the networks were allocated to geographical or political regions, for readers' orientation. The haplotypes, in which some of the individuals were inspected for *Wolbachia* infection, are indicated, with a different mark for each bacterial strain. These marks correspond to Fig. 3; and in (c), circle = *P. baton*, triangle = *P. vicrama*, square = *P. sinaicus*, star = *P. b. jacuticus*, hexagon = *P. panoptes*. Note that COI sequences not marked with any symbol were obtained in databases and hence not tested for *Wolbachia*. Maps were created in QGIS v. 2.18 (<http://qgis.org>) and the graphics was compiled in Graphic for Mac v. 3.1 (<https://www.graphic.com/mac/>).

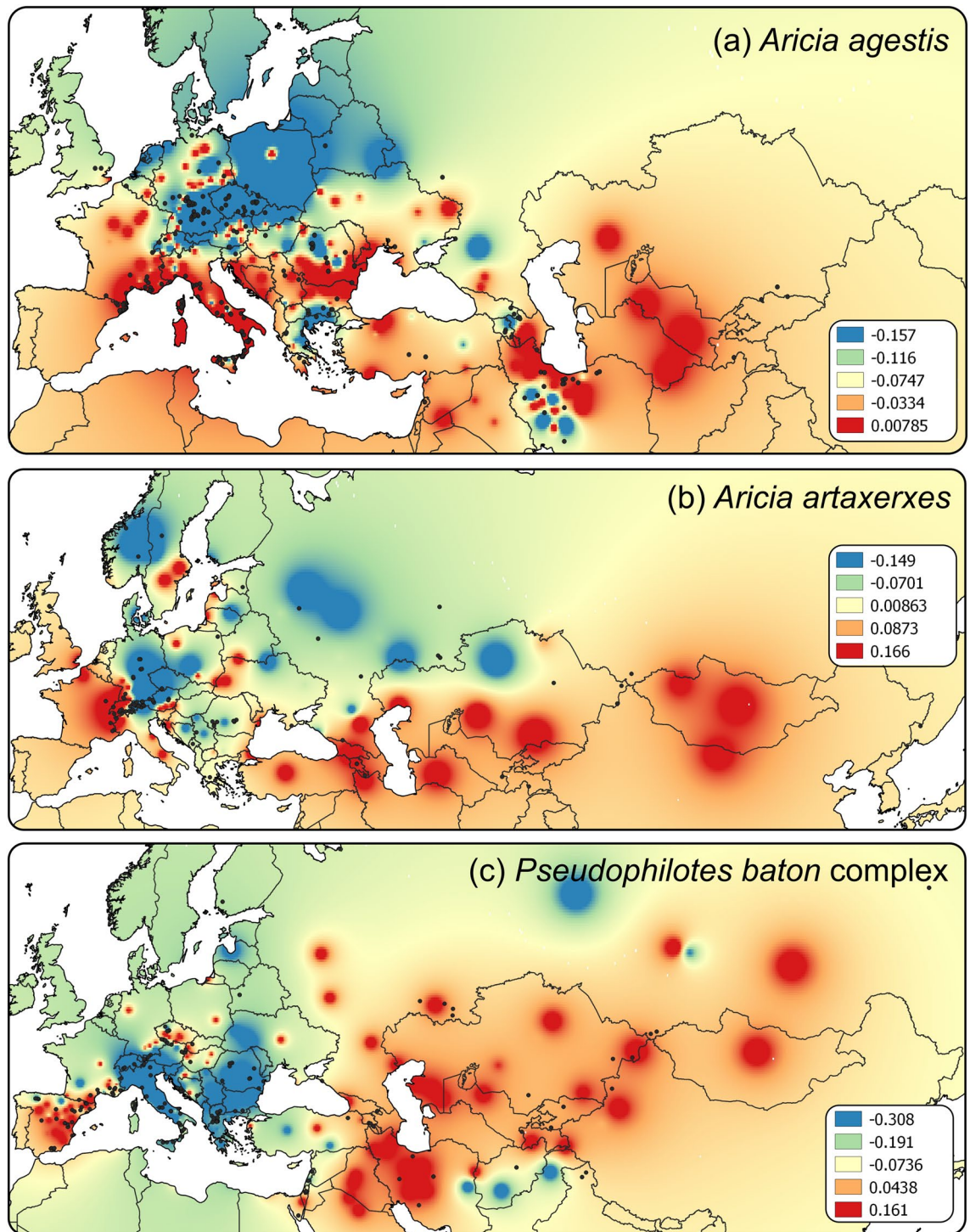


Figure 2. Genetic landscapes based on COI genetic distances among individual populations of (a) *Aricia agestis*, (b) *Aricia artaxerxes* and (c) *Pseudophilotes baton* species complex, created by the GDisPAL function in SPADS. The distance weighting parameter $a = 3$ was used as a presentable visualization. The map colours represent residuals of genetic distances. Maps were created in QGIS v. 2.18 (<http://qgis.org>) and the graphics was compiled in Inkscape v. 1.0 (<https://inkscape.org/>).

between south-eastern and Central Europe. The dataset's nucleotide diversity was 0.0019, which is 2.3 times lower than in *A. agestis*. Neither Tajima's D (-1.706 ; $p > 0.05$) nor Fu and Li's D^* (-1.796 ; $p > 0.1$) tests were significant.

In the mitochondrial Beast tree, haplotypes of *A. agestis* and *A. artaxerxes* formed well-defined monophyletic groups. The closest relative of *Aricia artaxerxes* was the Iberian *A. montensis*, with *A. agestis* being sister to this clade (Supplementary Fig. S1). Within *A. agestis*, Sicilian haplotypes (cluster BAPSa6) formed a well-supported lineage. A part of the haplotypes from cluster BAPSa2 (A28-A33) formed another separate lineage. Within *A.*

Species	wsp allele	coxA allele	Host haplotype	Host mtDNA BAPS cluster	Prevalence in host haplotypes	Geographical distribution (positive samples)
<i>A. agestis</i>	18	6	A1	BAPSa1	11/64	CZE + W RUS + SVK
<i>A. agestis</i>	61	85	A1, A26, A27	BAPSa1, BAPSa2	2/64, 5/7, 1/1	IRN, ROU + AUT, HUN
<i>A. agestis</i>	639	85 and 14	A28, A1, A11	BAPSa2, BAPSa1	1/3, 6/64, 1/1	ITA, GEO + IRN, IRN
<i>A. artaxerxes</i>	18	6	X1, X9, X13	BAPSp1	49/51, 1/1, 1/1	CZE + RUS + POL + AUT + ITA + CHE, W RUS, RUS: Cauc
<i>A. artaxerxes</i>	581	85	X2, X14, X15	BAPSp1	5/7, 1/2, 1/1	ROU + AUT, FRA, FRA
<i>A. artaxerxes</i>	61	14	X12	BAPSp1	5/5	ARM
<i>P. baton</i>	561	14	P5	BAPSp2	11/13 in <i>P. baton</i> P5	CZE + ITA + AUT
<i>P. vicrama</i>	561	14	P5	BAPSp2	7/7 in <i>P. vicrama</i> P5	HRV
<i>P. vicrama</i>	581	85	P3, P4	BAPSp3	5/6, 4/4	IRN + AFG + CHN + IND, TJK + KGZ
<i>P. vicrama</i>	639	14	P1, P20	BAPSp1	2/41, 2/3	GEO, IRN

Table 1. *Wolbachia* wsp and coxA alleles detected in samples of *Aricia* spp. and *Pseudophilotes* spp. Country codes accord to ISO-3 standard; W RUS = western Russia, RUS: Cauc = Russia, Caucasus Mts. In *A. agestis* wsp 639, two samples were recognized as coxA 14 (host haplotype A1, from eastern Iran), one as coxA 85 (host haplotype A1, from Georgia), and in four cases, the sequencing of coxA was not successful.

artaxerxes, samples from Kyrgyzstan formed a well-supported clade probably representing a distinct lineage within the species⁽³²⁾; these samples were not used in further analyses).

Aricia: Wolbachia. Our analyses revealed that 29 *Aricia agestis* individuals out of 115 tested were positive for *Wolbachia* (prevalence: 22%). Of these, 27 were successfully sequenced for wsp and 17 for coxA. All positive host samples were from the widely distributed groups BAPSa1 and BAPSa2. Sequencing revealed three wsp alleles and three coxA alleles, all already known (Table 1). The wsp allele 61 was congruent with coxA allele 85, wsp allele 639 accorded to coxA alleles 85 and 14, and wsp allele 18 to coxA allele 6. Two different wsp alleles were discovered in the same locality in Iran. Three different wsp alleles and three coxA alleles were discovered in the COI butterfly haplotype A1 (Table 1). However, the *Wolbachia* alleles were separated geographically in this case (Fig. 3a): Wsp allele 18/coxA allele 6 (together referred as a strain) were distributed northerly of the Alps and wsp alleles 61 and 639/coxA 85 were distributed southeast of the Alps.

In *A. artaxerxes*, 64 out of 69 tested individuals were positive for *Wolbachia* (prevalence: 93%); 63 were successfully sequenced for wsp, 62 for coxA. The sequencing revealed three wsp alleles and three coxA alleles, all already known. Wsp allele 18 (and sequences with 1–2 differing bases) corresponded with coxA allele 6, wsp allele 581 with coxA allele 85, and wsp allele 61 with coxA allele 14. Alleles were distributed geographically (Fig. 3a). The strain wsp 18/coxA 6 was distributed north of the Alps (similarly as in *A. agestis*) where it met with wsp 581/coxA 85 in the Alps. The strain wsp 61/coxA 14 was only detected in Transcaucasia.

In the *Wolbachia* wsp tree (Fig. 3b, Supplementary Fig. S1), the widely distributed wsp allele 18 split first, containing samples of the bacteria from *A. artaxerxes* across Europe and several samples from *A. agestis* from the Czech Republic, Slovakia and western Russia. The rest of the samples formed a monophyletic branch, within which (1) *Wolbachia* from *Aricia agestis* from Europe and Iran formed a clade together with that from Armenian *A. artaxerxes* (allele 61), (2) material from *Aricia agestis* from Iran, Georgia and one individual from Italy formed a second clade (allele 639), sister to (3) the third clade, detected in *A. artaxerxes* from France, Romania, and Austria (allele 581). The *Wolbachia* coxA tree (Fig. 3c) split into two branches: (1) the widely distributed allele 6 found in both *A. agestis* and *A. artaxerxes*, and (2) the closely related alleles 14 and 85, found in samples of both butterfly species from Iran and Transcaucasia, but also from Eastern Europe and the Alps.

Pseudophilotes: butterfly mitochondrial structure. In the 228 sequences of *Pseudophilotes* spp., the TCS analysis revealed 29 haplotypes (P1–P29, 23 segregating sites) belonging to six BAPS clusters (Fig. 1c). In several cases, the assignment to a BAPS cluster did not correspond to the taxonomical species revealed by genitalia dissection: BAPSp1 (P1–2, P6–20) contained mostly *P. vicrama* from Europe, but also the north-Asian *P. b. jacuticus* and *P. baton* from France, Corsica, and eastern Italy (the latter with “*baton*” valva type, Supplementary Fig. S2); this cluster had a star-like topology typical for recent population expansions. BAPSp2 (P5, P21) contained mostly *P. baton*, but also *P. vicrama* from adjacent distributional regions of Austria and Croatia (“*vicrama*” valva type: Supplementary Fig. S2 and ABOLD478-16), *P. sinicus* and *P. panoptes*. BAPSp3 (P3–4) included samples of *P. vicrama* (“*vicrama*” valva type) from Central Asia (Afghanistan, westernmost China, northernmost India, Iran, Kyrgyzstan, Tajikistan). BAPSp4–6 harboured only samples of *P. panoptes*. The dataset’s nucleotide diversity was 0.0046. Neither Tajima’s D (−1.106; $p > 0.1$) nor Fu and Li’s D* (−1.729; $p > 0.1$) tests were significant. The genetic landscape (Fig. 2c) showed the highest differentiation in western Iran and Central Asia and then in the Iberian Peninsula and southern France. The distribution of the mitochondrial species delimitation is deviating from the one based on male genital structures with the first running a bit more to the east than the

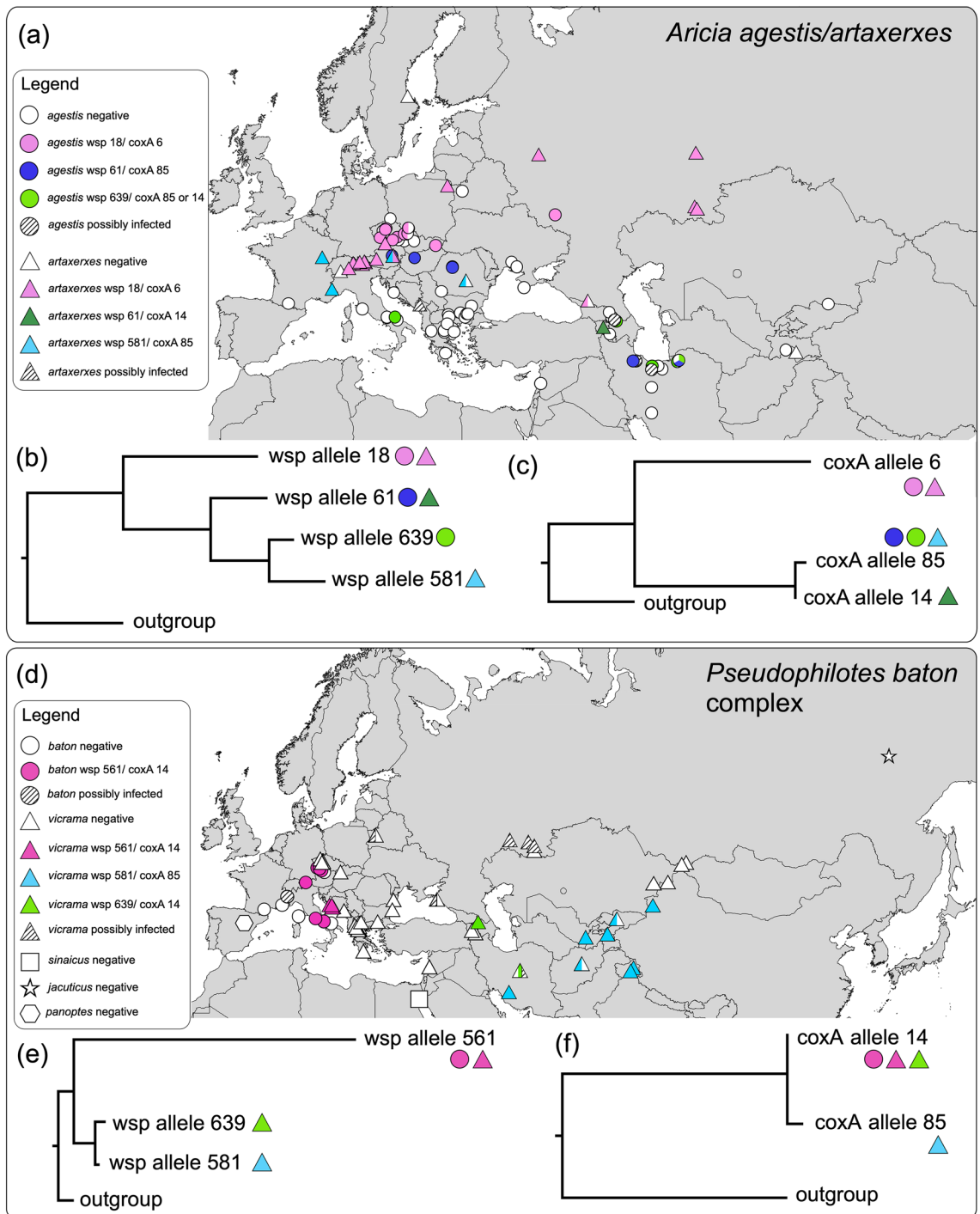


Figure 3. *Wolbachia* infection in two groups of closely related lycaenid butterflies: (a–c) *Aricia agestis/A. artaxerxes*: (a) distribution map, (b) Bayesian phylogenetic tree (MrBayes) of *wsp*, and (c) of *coxA* alleles. d–f *Pseudophilotes baton* species complex, (d) distribution map, (e) Bayesian phylogenetic tree of *wsp*, and (f) of *coxA* alleles. For a part of the samples, only the *wsp* gene was obtained, such samples were categorized by their *wsp* alleles. Maps were created in QGIS v. 2.18 (<http://qgis.org>) and the graphics was compiled in Graphic for Mac v. 3.1 (<https://www.graphic.com/mac/>).

second, i.e. crossing the Czech Republic, running through the border regions between eastern Austria and Slovakia, as well as between Croatia and Bosnia and Herzegovina.

The Beast tree (Fig. 4) revealed two clades: the first clade contained two lineages: BAPSp3 (P3–4, *P. vicrama* from Central Asia) and BAPSp2 (*P. baton*, *P. sinaicus*, part of *P. panoptes*, *P. vicrama* from the contact zone). The second clade split into two lineages: the first one consisted of *P. panoptes* haplotypes (except for those in BAPSp2) and the second of *P. vicrama* from BAPSp1, *P. b. jacuticus* and *P. baton* from southern France.

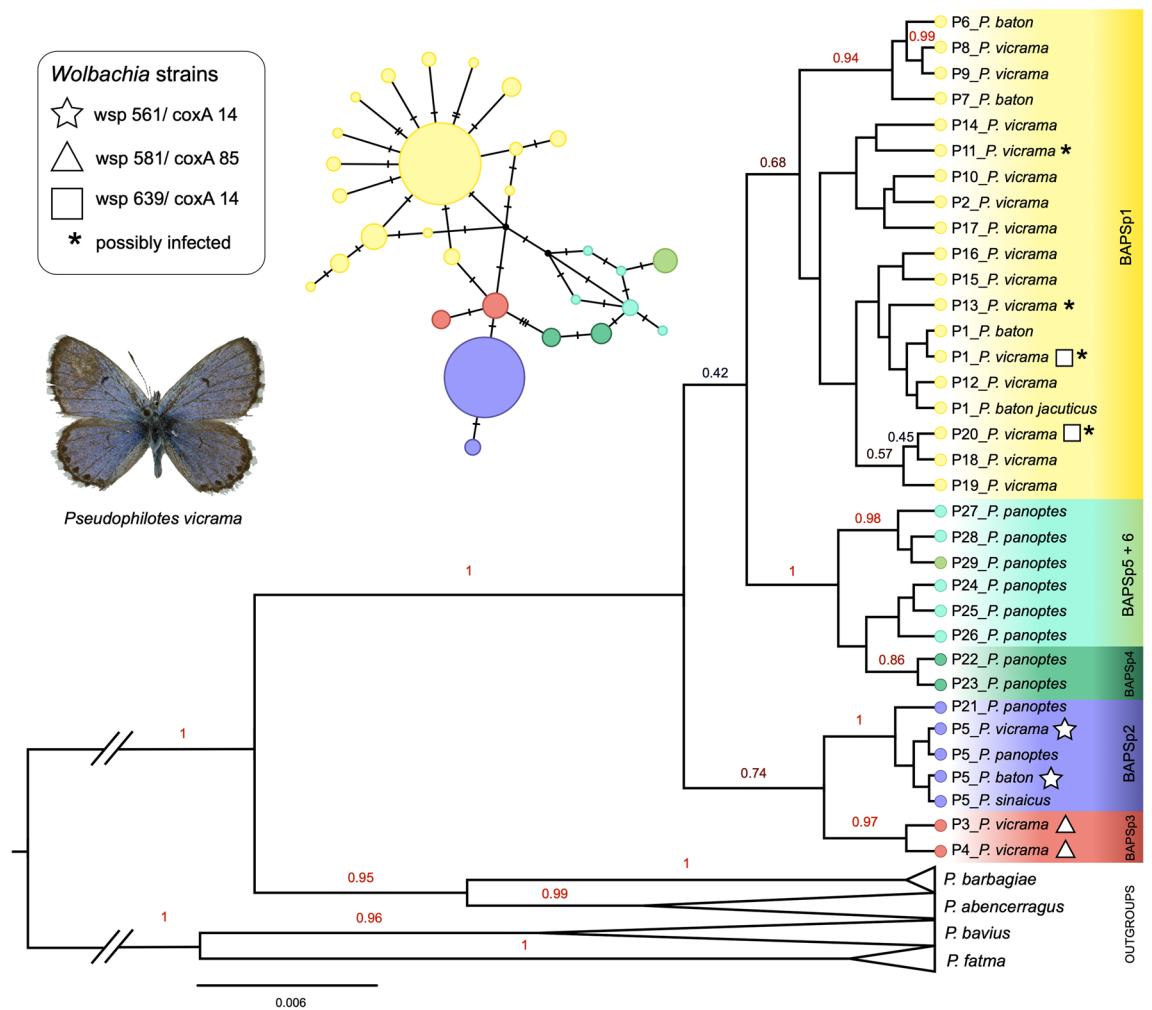


Figure 4. Mitochondrial Bayesian phylogenetic tree of *Pseudophilotes* spp. COI haplotypes. The branch labels represent posterior probabilities > 0.4 . Haplotypes positive for *Wolbachia* are marked. The tree was visualized in FigTree v. 1.3.1 (<https://github.com/rambaut/figtree>). Graphics was compiled in Graphic for Mac v. 3.1 (<https://www.graphic.com/mac>).

Pseudophilotes: Wolbachia. Looking for *Wolbachia* presence, 37 out of 108 tested individuals were positive (prevalence: 34%): 12 out of 17 in *P. baton* (71%), 25 out of 81 in *P. vicrama* (31%), 0 out of 4 in *P. sinaicus* (0%), 0 out of 4 in *P. b. jacuticus* (0%), and 0 out of 2 in *P. panoptes* (0%). The samples positive for *Wolbachia* followed geographic structures (Fig. 3d): 9 out of 36 in BAPSp1 (25%, positive samples from Iran, Georgia, western Russia and Belarus, but none from Western, Central and south-eastern Europe), 19 out of 24 in BAPSp2 (79%, Central Europe, Italy, Croatia; the majority of negative samples were *P. sinaicus*), 9 out of 10 in BAPSp3 (90%, Central Asia; one negative sample from Afghanistan), and 0 out of 2 in BAPSp4. BAPSp5 and BAPSp6 were based on database samples only and thus the *Wolbachia* presence could not be tested. Thirty-one samples were successfully sequenced for *wsp*, and 28 for *coxA*. We revealed three *wsp* alleles and two *coxA* alleles, all previously known. *Wsp* allele 561 corresponded with *coxA* allele 14 and was restricted to BAPSp2 (haplotype P5; *P. baton* and *P. vicrama* from the contact zone). *Wsp* allele 639 also accorded to *coxA* allele 14 and this combination was found in two Georgian and one Iranian samples (P1, BAPSp1). *Wsp* allele 581 in combination with *coxA* allele 85 was only obtained for Central Asia (BAPSp3, haplotypes P3–4).

The *Wolbachia* *wsp* tree (Fig. 3e, Supplementary Fig. S1) showed that *wsp* allele 561 present in BAPSp2 (*P. baton* and *P. vicrama* from the contact zone) formed a separate lineage from the closely related alleles 639 and 581 (Central Asian *P. vicrama* from BAPSp3 and three samples from Georgia and Iran from BAPSp1). The two *coxA* alleles (Fig. 3f, Supplementary Fig. S1) differed in a single mutation. *Pseudophilotes vicrama* from Georgia hence harbour *Wolbachia* related to those from Central Asian *P. vicrama* (BAPSp3) in *wsp* and those from *P. baton* (BAPSp2) in *coxA*.

Discussion

Two closely related groups of Polyommata butterflies displayed a composite pattern of *Wolbachia* infection, containing both infected and uninfected populations and several different *Wolbachia* strains, which were linked both to their mitochondrial history and to their geographical distribution.

***Wolbachia* infection in *Aricia agestis* and *A. artaxerxes*.** *Wolbachia* was highly prevalent in *Aricia artaxerxes* throughout its whole distribution range, in contrast to *A. agestis*. *Aricia agestis* and *A. artaxerxes* are believed to represent well-defined species. In the COI barcode, there is a constant minimum p-distance of about 2% between them, sufficient to distinguish two species^{26,28}. They are not distinguished by some nuclear markers (e.g. internal transcribed spacer 2²⁸) but are well separated by others as allozymes⁵⁷. Lineages distinguished by mtDNA and allozymes also differ in voltinism. A possibly analogous situation was observed in a cryptic species pair of Hesperidae butterflies, where *Spialia rosae* Hernández-Roldán, Dapporto, Dincă, Vicente & Vila, 2016 differed from *S. sertorius* (Hoffmannsegg, 1804) by the COI barcode, specific cuticular hydrocarbons, their host plants, and the presence of *Wolbachia*⁵⁸. *Wolbachia* presence thus might serve as a further distinguishing trait of the cryptic species pair *A. agestis* and *A. artaxerxes*.

The mitochondrial diversity of the highly infected *A. artaxerxes* was low compared to much less affected *A. agestis* (16 vs. 54 haplotypes, respectively). This represents indirect evidence for a selective sweep in *A. artaxerxes*. Such a probable selective sweep was also observed in the Chinese hesperid butterfly *Polytremis nascens* Leech, 1893, with *Wolbachia* infected populations being genetically less diverse than uninfected ones⁵⁹.

Overlapping areas of *Aricia agestis* and *A. artaxerxes* and transfer of *Wolbachia*. A proportion of the examined *Aricia agestis* individuals were also positive for *Wolbachia*. In both *Aricia* species, we discovered several *Wolbachia* strains. Those samples of *A. agestis* from lowland Central and north-eastern Europe, which were positive for *Wolbachia*, were infected by the wsp 18/coxA 6 strain, also found in *A. artaxerxes* in this region (Fig. 3a). At Hochschwab, a mountain massif within the north-eastern limestone Alps (Austria), the barcodes of both species co-occur, infected by a similar coxA allele 85 and two wsp alleles (581 and 61).

Although *Aricia artaxerxes* inhabits higher elevations in Southern Europe than *A. agestis*, their elevational distribution overlaps in Central and Eastern Europe^{26,30}. Our specific Hochschwab locality (Seebergsattel pass) is situated at 1260 m a.s.l., and a sequence of an *A. agestis* specimen from an even higher elevation (about 1600 m a.s.l.), 30 km away from our record, is published (Hochkar, Austria: ABOLD655-17).

The occurrence of similar *Wolbachia* alleles found in the two *Aricia* species in lowland Central Europe suggests a transfer of this bacterium between these species. Hybridization of the two species was documented alongside a contact zone established during Holocene in Great Britain⁶⁰. However, sharing the same *Wolbachia* allele between the two butterfly species is unlikely a product of hybridization, as the host's mitochondrial DNA would be transferred together with *Wolbachia*. In this context, Sintupachee et al.⁵ observed phylogenetically distant insect groups infected by the same *Wolbachia* strain utilizing a similar host plant, suggesting a horizontal transfer. Thus, hemipteran insects, injecting their saliva containing bacteria to plant cells, could have infected leaf-chewing insects via this means. Larvae of leaf-mining lepidopterans (Gracillariidae) were also found to introduce *Wolbachia* into plant tissue, and evidence of horizontal transfers was detected in the comparison of their phylogeny and *Wolbachia* infection⁶¹. Larvae of *Aricia agestis* and *A. artaxerxes* share their host plants, which may have served as vector for the infection. Alternatively, *Wolbachia* might be transferred horizontally through a shared parasitoid⁶².

Situation in Hochschwab, where a population contained the barcodes of both species, and two different wsp alleles observed also in other parts of their ranges, could be product of a past hybridization and would require further attention.

***Aricia agestis* and *A. artaxerxes* biogeographic history.** The two examined *Aricia* species strongly differ in their biogeographic patterns. On the one hand, *Aricia artaxerxes* is a boreo-montane species, with decreased mitochondrial diversity and thus with its mitochondrial history obscured by *Wolbachia* infection. However, *Wolbachia* infection could also give us some hints to the butterfly's phylogeographic history. The different *Wolbachia* strains associated with spatially separated *A. artaxerxes* haplotypes suggest the existence of two host haplogroups (not distinguished in BAPS analysis): (1) a Central European group reaching Northern Europe and the Russian forest-steppe belt (infected by *Wolbachia* wsp allele 18/coxA allele 6, host haplotypes X1, X9, X13), and (2) a south-western European group (*Wolbachia* wsp allele 581 + two closely related coxA alleles, host haplotypes X2, X15, X14). Both groups are present in the Alps and in the Balkan Peninsula. The current boreo-montane range with such longitudinally distributed haplogroups could be a legacy of a widespread distribution in the Pleistocene mammoth steppe biome, with remnants of such distribution existing in the Holocene⁶³, the result of postglacial expansion from refugia located north of the Mediterranean region⁶⁴, or a mixture of both.

Aricia agestis, on the other hand, is a temperate grassland species with the highest genetic diversity in the Italian and the Balkan peninsulas (Fig. 2b). The observed genetic patterns call for glacial retreats in peninsular Italy and the Balkan Peninsula as often observed in warm-adapted species in Europe^{65,66}, with both expanding from these retreats postglacially⁶⁷. During this postglacial expansion through Europe, the populations stemming from both refugia were mixing in a low density hybridisation system (cf. ⁶⁶) and, in addition, occasionally hybridize with *A. artaxerxes*⁶⁰ and also might have been infected by *Wolbachia* from this species. The mitochondrial clusters BAPSa4 and BAPSa6 restricted to Corsica and adjoining islands, as well as Sicily, respectively, indicated further independent differentiation centres on these two islands due to their rear edge position⁶⁸. Whereas the biogeographically isolated position of Corsica (and Sardinia) has been known for some time⁶⁵, Sicily, separated by just 3 km of sea from the southernmost tip of peninsular Italy, has been recovered as being biogeographically

particular only recently^{69–71}. In addition, the isolated occurrence of the eastern branch BAPSa3 from north-east Iran to Central Asia calls for a separate retreat in this area; however, with its glacial-interglacial dynamics still being ambiguous.

Wolbachia distribution in the *Pseudophilotes baton* species complex. The patterns of *Wolbachia* infection in the *Pseudophilotes baton* species complex correspond to the mitochondrial history of the butterflies, which does not agree with the described taxa (Figs. 1c, 3d). The phylogenetics of closely related *Pseudophilotes* species from the *Pseudophilotes baton* complex is traditionally problematic. These species were not distinguished sufficiently by four nuclear genes²⁹. *Pseudophilotes vicrama* is the most widely distributed member of the complex, present from Central Europe to Mongolia, distinguished by the shape of its valvae from the putative taxa occurring at the margins of its distribution, i.e. *P. baton*, *P. b. jacuticus*, *P. sinicus*, and *P. panoptes*. The COI barcodes of the butterflies do not correspond to this taxonomic classification. Thus, *P. b. jacuticus* shares haplotype P1 with *P. vicrama* from Europe; and *P. sinicus* shares haplotype P5 with the majority of *P. baton*. With the data available, it therefore seems impossible to decide whether the complex represents completely separated species with secondary contacts, a speciation process under way, or just a single species with polymorphic male genitalia (cf. bad species concept⁷²).

Similarly to the situation in *Aricia agestis* and *A. artaxerxes*, the *Pseudophilotes* mitochondrial clusters with high *Wolbachia* prevalence rates (BAPSp2, BAPSp3) harboured considerably lower genetic diversity (two haplotypes each) than the mostly uninfected BAPSp1 cluster (17 haplotypes), again pointing to a selective sweep in the heavily infected clusters.

Three different *Wolbachia* strains were detected in this butterfly species complex, found in different parts of its distribution area (Fig. 3d). *Wolbachia* was highly prevalent in *P. baton* (host mtDNA cluster BAPSp2), except for the samples from southern France, Corsica, and two samples from Italy, which were negative in *Wolbachia* and clustered with *P. vicrama* in the host mtDNA. Another *Wolbachia* strain was found in *P. vicrama* from Central Asia, which mtDNA also formed a separate cluster (BAPSp3). The cluster BAPSp1 was mostly uninfected, except for some samples from western Russia (samples from which we failed to obtain *Wolbachia* sequences, Supplementary Dataset S1) and samples from Iran and Georgia. These latter were infected by a strain related to the one found in BAPSp3 in *wsp*, and to the one in BAPSp2 in *coxA* (Fig. 3d).

In the host mitochondrial Bayesian tree, the two clusters infected by *Wolbachia*, BAPSp3 (Central Asian *P. vicrama*) and BAPSp2 (mostly *P. baton*), formed a clade separated from the mostly uninfected samples. However, BAPSp2 and BAPSp3 were infected by two different *Wolbachia* strains. As these two strains are genetically differentiated from each other and geographically located on opposite angles of the range (i.e. south-central Europe vs. Central Asia), two independent infections with similar consequences is a likely scenario.

All in all, the *Pseudophilotes* species complex follows the butterfly paradigm of glacial refugia and postglacial range dynamics in Europe⁷³ with a non-expansive taxon in Iberia (i.e. *P. panoptes*) and expansive lineages out of the Adriatic- (i.e. *P. baton*) and the Pontic-Mediterranean (i.e. *P. vicrama*) refugia⁶⁷. Similar to *A. agestis*, *P. vicrama* seems to be polycentric with at least one western centre of differentiation (i.e. BAPSp1) from the Balkans to Iran (i.e. Pontic-Mediterranean-Iranian elements, cf.⁷⁴) and one Central Asian element (BAPSp3) with still unknown range dynamics, also supported by a *Wolbachia* type restricted in *P. vicrama* to this genetic lineage. Additionally, *P. b. jacuticus* might represent a Siberian element with a Siberian centre of dispersal (cf.⁶⁵), while *P. sinicus* might be a rear edge population with in situ differentiation and adaptation to the prevailing eremic conditions.

Hybrid zone between *Pseudophilotes baton* and *P. vicrama*. Sequences of *Pseudophilotes vicrama* from Croatia and lowland eastern Austria are assigned to haplotype P5 (BAPSp2), mostly found in *P. baton*; furthermore, the Croatian samples also harboured the same *Wolbachia* strain as *P. baton*. However, their genital structures correspond to the “*vicrama*” type (Supplementary Fig. S2). In addition, two samples of *P. baton* from south-eastern Italy were assigned as the *P. vicrama* haplotype P1 and were negative for *Wolbachia*. This implies the existence of a hybrid zone between these two taxa crossing Central Europe. This region is a known contact zone between eastern and western lineages of many invertebrates and vertebrates, e.g., in the lycaenid butterfly *Polyommatus coridon*⁷⁵, mice *Mus musculus* and *M. domesticus*⁷⁶, and toads *Bombina bombina* and *B. variegata*⁷⁷.

Pseudophilotes baton samples from southern France and Corsica also lacked *Wolbachia* and were assigned to the BAPSp1 cluster, otherwise confined to *P. vicrama*, suggesting that hybridization and population replacement might be multiple within the species complex. Furthermore, the distribution of this mitochondrial cluster suggests remarkable range dynamics in *P. baton* in the past glacial and interglacial phases, with most of their results being overwritten later on, but with geographically small rear edge remnants, hereby supporting glacial refugia in southern France⁷⁸ and Corsica⁶⁵.

In several cases, *Wolbachia* infection was observed to match mitochondrial phylogenies in butterflies, but the nuclear information coincided with life history traits (⁷⁹ for *Eurema hecabe* (Linnaeus, 1758), Pieridae³⁸ for *Hypolimnas* Hübner, 1819 spp., Nymphalidae). In these examples, *Wolbachia* is suspected to mediate interspecific introgression. In our case of *Pseudophilotes*, however, the hybridization probably happened regardless of *Wolbachia*. *Wolbachia* presence (in Croatian samples) or absence (in French, Italian, and Corsican *P. baton*) rather provides additional evidence for hybridization.

Limitations and future directions. We addressed the patterns and diversity in *Wolbachia* infection across a broad geographic range in several species. However, the exact mechanisms whether and how *Wolbachia* affects these butterfly hosts remains undiscovered. Routes to discovering them may include breeding experiments and antibiotic treatments of infected populations. Such strategy would identify cytoplasmic incompatibility, male

killing or feminisation (e.g.⁸⁰). Males of the butterflies inspected are not notably rare in natural conditions and our dataset contained males bearing all *Wolbachia* alleles detected here (Supplementary Dataset S1). To identify the exact *Wolbachia* strains, the whole MLST typing³⁹ for each *wsp* allele would be necessary. As the two used *Wolbachia* markers originate from different parts of *Wolbachia* genome and the information, which they deliver, is congruent, we believe that this additional information would not change our conclusions, but rather refine them.

The exact distribution of *Aricia artaxerxes* in temperate Europe is still obscured, and all putative populations should be examined using DNA barcoding. Using high-throughput sequencing of samples from the Eastern Alps, where *A. agestis* and *A. artaxerxes* barcodes coexist at the same site, would distinguish between past hybridization and horizontal transfer of *Wolbachia*.

Synthesis. Taken together, in the two analysed Polyommatae groups, we observed several different relations between *Wolbachia* infection and mitochondrial DNA of their hosts. First, presence of *Wolbachia* caused impoverishment in mitochondrial haplotypes, when compared to uninfected relatives. Second, within a species (or species complex), different strains of *Wolbachia* usually infect different mitochondrial groups (clusters). Third, a possible horizontal transfer of *Wolbachia* occurred between sympatric populations. Fourth, *Wolbachia* presence, absence or allele identity served as an additional marker of hybridization in a contact zone.

Specifically, in the case of *Aricia*, the two taxa with diverging mitochondrial structures mirroring *Wolbachia* prevalence were probably more spatially separated in the past, and both a hybridization⁶⁰ and a horizontal transmission of *Wolbachia* likely occurred during Holocene in the contact areas of both species. In the case of *Pseudophilotes*, we observed several poorly defined taxa, with different mitochondrial clusters infected by different *Wolbachia* strains, or remaining uninfected by the bacteria. Alongside the contact zone, *Pseudophilotes* species hybridized, which was revealed by the discrepancy between the morphological trait (genital structure) and the mitochondrial signal corresponding to *Wolbachia* infection.

Palaeartic Polyommatae butterflies were exposed to the late Cenozoic cooling of the Earth, habitat opening and climate alternations. Their host plants are often herbs occurring in grasslands, chemically protected from ungulate grazing. Many taxa rapidly diversified under such conditions. The population divergence leading to speciation is often not completed, and we observe different levels of its completeness in various taxa⁸¹. *Aricia* spp. and *Pseudophilotes* spp. represent two different levels of such diversification. Species-rich radiations seem to be common in other Polyommatae (e.g. the *Polyommatus coridon* group^{82,83} and *Agrodiaetus* species with chromosomal fissions⁸⁴). *Wolbachia* and other symbiotic bacteria are suspected drivers of arthropod evolution⁸⁵. The species might even lose the infection after some time⁸⁶, so that we could not reveal the culprit of diversification. We suggest that *Wolbachia* infection plays an important role in the mitochondrial diversity of butterflies and this should be taken into account in all phylogeographic and cryptic diversity studies involving mitochondrial DNA. *Wolbachia* infection may prevent inferring the host species' mitochondrial history, but knowledge on the distribution of *Wolbachia* alleles can be helpful in understanding the mitochondrial patterns and biogeographic scenarios.

Data availability

All sequences obtained in this study are stored in GenBank (Accession numbers: butterfly COI *Aricia agestis*/*A. artaxerxes*: MT883793–MT883959; COI *Pseudophilotes baton* species complex: MT878244–MT878363; *Wolbachia* *wsp* from *A. agestis* and *A. artaxerxes*: MT950442–MT950530, and *coxA*: MT950363–MT950441; *wsp* from *P. baton* species complex: MT890972–MT891001, and *coxA*: MT891002–MT891029). The codes of each specimen are in Supplementary Dataset S1. The Nexus alignments are stored at the Figshare repository <https://doi.org/10.6084/m9.figshare.12917426.v1>.

Received: 4 September 2020; Accepted: 19 January 2021

Published online: 04 February 2021

References

- Jiggins, F. M. Male-killing *Wolbachia* and mitochondrial DNA: Selective sweeps, hybrid introgression and parasite population dynamics. *Genetics* **164**, 5–12 (2003).
- Poinsot, D., Charlat, S. & Merçot, H. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: Confronting the models with the facts. *BioEssays* **25**, 259–265 (2003).
- Werren, J. H., Baldo, L. & Clark, M. E. *Wolbachia*: Master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**, 741–751 (2008).
- Vavre, F., Fleury, F., Lepetit, D., Fouillet, P. & Bouléreau, M. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Mol. Biol. Evol.* **16**, 1711–1723 (1999).
- Sintupachee, S., Milne, J. R., Poonchaisri, S., Baimai, V. & Kittayapong, P. Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. *Microb. Ecol.* **51**, 294–301 (2006).
- Li, S.-J. *et al.* Plant-mediated horizontal transmission of *Wolbachia* between whiteflies. *ISME J.* **11**, 1019–1028 (2017).
- Turelli, M. & Hoffmann, A. A. Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**, 440 (1991).
- Oliveira, D. C. S. G., Raychoudhury, R., Lavrov, D. V. & Werren, J. H. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol. Biol. Evol.* **25**, 2167–2180 (2008).
- Raychoudhury, R. *et al.* Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial-*Wolbachia* sweep in North America. *Heredity* **104**, 318–326 (2010).
- Telschow, A., Gadau, J., Werren, J. H. & Kobayashi, Y. Genetic incompatibilities between mitochondria and nuclear genes: Effect on gene flow and speciation. *Front. Genet.* **10**, 62 (2019).

11. Kodandaramaiah, U., Simonsen, T. J., Bromilow, S., Wahlberg, N. & Sperling, F. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecol. Evol.* **3**, 5167–5176 (2013).
12. Ritter, S. *et al.* *Wolbachia* infections mimic cryptic speciation in two parasitic butterfly species, *Phengaris teleius* and *P. nausithous* (Lepidoptera: Lycaenidae). *PLoS ONE* **8**, e78107 (2013).
13. Whitworth, T. L., Dawson, R. D., Magalon, H. & Baudry, E. DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). *Proc. R. Soc. B Biol. Sci.* **274**, 1731–1739 (2007).
14. Barton, N. & Bengtsson, B. O. The barrier to genetic exchange between hybridising populations. *Heredity* **57**, 357–376 (1986).
15. Vavre, F., Fleury, F., Varaldi, J., Fouillet, P. & Boulétreau, M. Infection polymorphism and cytoplasmic incompatibility in Hymenoptera-*Wolbachia* associations. *Heredity* **88**, 361–365 (2002).
16. Jaenike, J., Dyer, K. A., Cornish, C. & Minhas, M. S. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* **4**, e325 (2006).
17. Telschow, A., Flor, M., Kobayashi, Y., Hammerstein, P. & Werren, J. H. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: Mainland-island model. *PLoS ONE* **2**, e701 (2007).
18. Flor, M., Hammerstein, P. & Telschow, A. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and the stability of infection polymorphism in parapatric host populations. *J. Evol. Biol.* **20**, 696–706 (2007).
19. Graham, R. I. & Wilson, K. Male-killing *Wolbachia* and mitochondrial selective sweep in a migratory African insect. *BMC Evol. Biol.* **12**, 204 (2012).
20. Ahmed, M. Z., Breinholt, J. W. & Kawahara, A. Y. Evidence for common horizontal transmission of *Wolbachia* among butterflies and moths. *BMC Evol. Biol.* **16**, 118 (2016).
21. Salunkhe, R. C., Narkhede, K. P. & Shouche, Y. S. Distribution and evolutionary impact of *Wolbachia* on butterfly hosts. *Indian J. Microbiol.* **54**, 249 (2014).
22. Talavera, G., Lukhtanov, V. A., Pierce, N. E. & Vila, R. Establishing criteria for higher-level classification using molecular data: The systematics of *Polyommatus blue* butterflies (Lepidoptera, Lycaenidae). *Cladistics* **29**, 166–192 (2013).
23. Espeland, M. *et al.* A Comprehensive and dated phylogenomic analysis of butterflies. *Curr. Biol.* **28**, 770–778.e5 (2018).
24. 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 Biol. Sci.* **286**, 20191311 (2019).
25. Gaunet, A. *et al.* Two consecutive *Wolbachia*-mediated mitochondrial introgressions obscure taxonomy in Palearctic swallowtail butterflies (Lepidoptera, Papilionidae). *Zool. Scr.* **48**, 507–519 (2019).
26. Dinca, 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.* **278**, 347–355 (2011).
27. Ugelvig, L. V., Vila, R., Pierce, N. E. & Nash, D. R. A phylogenetic revision of the *Glaucopsyche* section (Lepidoptera: Lycaenidae), with special focus on the *Phengaris-Maculinea* clade. *Mol. Phylogenet. Evol.* **61**, 237–243 (2011).
28. Sañudo-Restrepo, C. P., Dincă, V., Talavera, G. & Vila, R. Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). *Mol. Phylogenet. Evol.* **66**, 369–379 (2013).
29. Todisco, V. *et al.* Molecular phylogeny of the Palearctic butterfly genus *Pseudophilotes* (Lepidoptera: Lycaenidae) with focus on the Sardinian endemic *P. barbagiae*. *BMC Zool.* **3**, 4 (2018).
30. Sucháčková Bartoňová, A., Beneš, J., Fric, Z. F. & Konvička, M. Genetic confirmation of *Aricia artaxerxes* (Fabricius, 1793) (Lepidoptera, Lycaenidae) in the Czech Republic, its conservation significance and biogeographic context. *Nota Lepidopterol.* **42**(2), 163–176 (2019).
31. Kames, P. Die Aufklärung des Differenzierungsgrades und der Phylogenese der beiden *Aricia*-Arten *agestis* Den. et Schiff. und *artaxerxes* Fabr. (allous G.-Hb.) mit Hilfe von Eizuchten und Kreuzungsversuchen (Lep., Lycaenidae). *Mitt. Entomol. Ges. Basel, N. F.* **26**, 7–13, 29–64 (1976).
32. Korb, S., Faltýnek Fric, Z. & Bartonova, A. On the status of *Aricia* cf. *scythissa* (Nekrutenko, 1985) (Lepidoptera: Lycaenidae) based on molecular investigations. *Euroasian Entomol. J.* **14**, 237–240 (2015).
33. Wiemers, M., Chazot, N., Wheat, C. W., Schweiger, O. & Wahlberg, N. A complete time-calibrated multi-gene phylogeny of the European butterflies. *ZooKeys* **938**, 97–124 (2020).
34. Settele, J., Steiner, R., Reinhardt, R., Feldmann, R. & Hermann, G. *Schmetterlinge: Die Tagfalter Deutschlands* (Ulmer Eugen Verlag, Stuttgart, 2015).
35. Monteiro, A. & Pierce, N. E. Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1 α gene sequences. *Mol. Phylogenet. Evol.* **18**, 264–281 (2001).
36. Wahlberg, N. & Wheat, C. W. Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic DNA extractions of lepidoptera. *Syst. Biol.* **57**, 231–242 (2008).
37. Braig, H. R., Zhou, W., Dobson, S. L. & O'Neill, S. L. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J. Bacteriol.* **180**, 2373–2378 (1998).
38. Sahoo, R. K. *et al.* Evolution of *Hypolimnas* butterflies (Nymphalidae): Out-of-Africa origin and *Wolbachia*-mediated introgression. *Mol. Phylogenet. Evol.* **123**, 50–58 (2018).
39. Baldo, L. *et al.* Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol.* **72**, 7098–7110 (2006).
40. Kearse, M. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
41. Kajtoch, Ł *et al.* Using host species traits to understand the *Wolbachia* infection distribution across terrestrial beetles. *Sci. Rep.* **9**, 847 (2019).
42. Ratnasingham, S. & Hebert, P. D. N. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* **7**, 355–364 (2007).
43. Clement, M., Posada, D. & Crandall, K. A. TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659 (2000).
44. Leigh, J. W. & Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
45. Cheng, L., Connor, T. R., Sirén, J., Aanensen, D. M. & Corander, J. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol. Biol. Evol.* **30**, 1224–1228 (2013).
46. Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W. & Corander, J. RhiereBAPS: An R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res.* **3** (2018).
47. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/> (2019).
48. Librado, P. & Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
49. Dellicour, S. & Mardulyn, P. SPADS 1.0: A toolbox to perform spatial analyses on DNA sequence data sets. *Mol. Ecol. Resour.* **14**, 647–651 (2014).
50. Watson, D. F. & Philip, G. M. A Refinement of inverse distance weighted interpolation. *Geoprocessing* **2**, 315–327 (1985).
51. Manni, F., Guérard, E. & Heyer, E. Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Hum. Biol.* **76**, 173–190 (2004).

52. Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, 016 (2018).
53. Posada, D. jModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
54. Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. *Tracer v1.6*. <http://tree.bio.ed.ac.uk/software/tracer/>. (2014).
55. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
56. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (2010).
57. Aagaard, K. *et al.* Phylogenetic relationships in brown argus butterflies (Lepidoptera: Lycaenidae: *Aricia*) from northwestern Europe. *Biol. J. Linn. Soc.* **75**, 27–37 (2002).
58. Hernández-Roldán, J. L. *et al.* Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Mol. Ecol.* **25**, 4267–4284 (2016).
59. Jiang, W. *et al.* *Wolbachia* infection status and genetic structure in natural populations of *Polytremis nascens* (Lepidoptera: Hesperidae). *Infect. Genet. Evol.* **27**, 202–211 (2014).
60. Mallet, J., Wynne, I. R. & Thomas, C. D. Hybridisation and climate change: Brown argus butterflies in Britain (*Polyommatus* subgenus *Aricia*). *Insect Conserv. Divers.* **4**, 192–199 (2011).
61. Gutzwiller, F., Dedeine, F., Kaiser, W., Giron, D. & Lopez-Vaamonde, C. Correlation between the green-island phenotype and *Wolbachia* infections during the evolutionary diversification of Gracillariidae leaf-mining moths. *Ecol. Evol.* **5**, 4049–4062 (2015).
62. Mascarenhas, R. O., Prezotto, L. F., Perondini, A. L. P., Marino, C. L. & Selivon, D. *Wolbachia* in guilds of Anastrepha fruit flies (Tephritidae) and parasitoid wasps (Braconidae). *Genet. Mol. Biol.* **39**, 600–610 (2016).
63. Sucháčková Bartoňová, A. *et al.* Recently lost connectivity in the Western Palearctic steppes: The case of a scarce specialist butterfly. *Conserv. Genet.* **21**, 561–575 (2020).
64. Schmitt, T. & Varga, Z. Extra-Mediterranean refugia: The rule and not the exception?. *Front. Zool.* **9**, 1–12 (2012).
65. de Lattin, G. *Grundriss der Zoogeographie* (VEB Gustav Fischer Verlag, Jena, 1967).
66. Hewitt, G. M. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276 (1996).
67. Schmitt, T. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front. Zool.* **4**, 11 (2007).
68. Hampe, A. & Petit, R. J. Conserving biodiversity under climate change: The rear edge matters. *Ecol. Lett.* **8**, 461–467 (2005).
69. Heiser, M., Dapporto, L. & Schmitt, T. Coupling impoverishment analysis and partitioning of beta diversity allows a comprehensive description of Odonata biogeography in the Western Mediterranean. *Org. Divers. Evol.* **14**, 203–214 (2014).
70. Vodá, R. *et al.* Historical and contemporary factors generate unique butterfly communities on islands. *Sci. Rep.* **6**, 28828 (2016).
71. Scalercio, S. *et al.* How long is 3 km for a butterfly? Ecological constraints and functional traits explain high mitochondrial genetic diversity between Sicily and the Italian Peninsula. *J. Anim. Ecol.* **89**, 2013–2026 (2020).
72. Descimon, H. & Mallet, J. Bad species. in *Ecology and Evolution of European Butterflies* (Oxford University Press, Oxford, 2009).
73. Habel, J. C., Schmitt, T. & Müller, P. The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: The phylogeography of the Marbled White butterfly (Satyridae, Lepidoptera). *J. Biogeogr.* **32**, 1489–1497 (2005).
74. Varga, Z. Das Prinzip der areal-analytischen Methode in der Zoogeographie und die Faunenelement-Einteilung der europäischen Tagmetterlinge (Lepidoptera: Diurna). *Acta Biol. Debrecina* **14**, 223–285 (1977).
75. Schmitt, T. & Zimmermann, M. To hybridize or not to hybridize: What separates two genetic lineages of the Chalk-hill Blue *Polyommatus coridon* (Lycaenidae, Lepidoptera) along their secondary contact zone throughout eastern Central Europe?. *J. Zool. Syst. Evol. Res.* **50**, 106–115 (2012).
76. Janoušek, V. *et al.* Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. *Mol. Ecol.* **21**, 3032–3047 (2012).
77. Nürnberg, B., Lohse, K., Fijarczyk, A., Szymura, J. M. & Blaxter, M. L. Para-allopatry in hybridizing fire-bellied toads (*Bombina orientalis* and *B. variegata*): Inference from transcriptome-wide coalescence analyses. *Evolution* **70**, 1803–1818 (2016).
78. Vitali, F. & Schmitt, T. Ecological patterns strongly impact the biogeography of western Palearctic longhorn beetles (Coleoptera: Cerambycoidea). *Org. Divers. Evol.* **17**, 163–180 (2017).
79. Narita, S., Nomura, M., Kato, Y. & Fukatsu, T. Genetic structure of sibling butterfly species affected by *Wolbachia* infection sweep: Evolutionary and biogeographical implications. *Mol. Ecol.* **15**, 1095–1108 (2006).
80. Kageyama, S. *et al.* Feminizing *Wolbachia* endosymbiont disrupts maternal sex chromosome inheritance in a butterfly species. *Evol. Lett.* **1**, 232–244 (2017).
81. Nosil, P., Harmon, L. J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol. (Amst.)* **24**, 145–156 (2009).
82. Talavera, G., Lukhtanov, V. A., Rieppel, L., Pierce, N. E. & Vila, R. In the shadow of phylogenetic uncertainty: The recent diversification of *Lysandra* butterflies through chromosomal change. *Mol. Phylogenet. Evol.* **69**, 469–478 (2013).
83. Kühne, G., Kosuch, J., Hochkirch, A. & Schmitt, T. Extra-Mediterranean glacial refugia in a Mediterranean faunal element: The phylogeography of the chalk-hill blue *Polyommatus coridon* (Lepidoptera, Lycaenidae). *Sci. Rep.* **7**, srep43533 (2017).
84. Wiemers, M., Keller, A. & Wolf, M. ITS2 secondary structure improves phylogeny estimation in a radiation of blue butterflies of the subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: Polyommatus). *BMC Evol. Biol.* **9**, 300 (2009).
85. Duron, O. & Hurst, G. D. Arthropods and inherited bacteria: From counting the symbionts to understanding how symbionts count. *BMC Biol.* **11**, 45 (2013).
86. Bailly-Bechet, M. *et al.* How long does *Wolbachia* remain on board?. *Mol. Biol. Evol.* **34**, 1183–1193 (2017).

Acknowledgements

We would like to thank O. Adam, J. Beneš, K. Dovgailo, M. Fišer, P. Herman, T. Kadlec, A. Kairouz, S. Korb, O. Kudrna, A. V. Kurmaev, A. Laštůvka, Z. Laštůvka, D. Leština, J. Lipárová, D. V. Morgun, O. Pak, M. Pálka, A. Pavlíčko, N. Rubin, D. Shovkoon, L. Spitzer, J. Vávra, P. Vrba and V. Zaritský. The research was supported by the University of South Bohemia, Ceske Budejovice (Grant No. 152/2016/P), and the Swedish Research Council (Grant No. 2015-04441).

Author contributions

A.S.B. and Z.F.F. designed the study. A.S.B., M.K., J.M., M.W., N.W., T.S. and Z.F.F. collected the material. A.S.B., J.M., M.W. and N.I. performed laboratory analyses. A.S.B. and Z.F.F. analysed the data. J.M. produced the graphics. A.S.B., T.S. and M.K. drafted the manuscript and all authors contributed to its final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-82433-8>.

Correspondence and requests for materials should be addressed to A.S.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021