Correspondence: Roger Vila, Institute of Evolutionary Biology, Passeig Marítim de la Barceloneta 37, 08003, Barcelona, Spain; Tel: +34 610925825; E-mail: roger.vila@csic.es

Two consecutive *Wolbachia*-mediated mitochondrial introgressions obscure taxonomy in Palearctic swallowtail butterflies (Lepidoptera, Papilionidae)

Running title: Introgression in swallowtail butterflies.

AURÉLIEN GAUNET[1*], VLAD DINCĂ[1,2*], LEONARDO DAPPORTO[3*], SERGIO MONTAGUD[4], RALUCA VODĂ[5], SÄMI SCHÄR[1], ARNAUD BADIANE[6,7], ENRIQUE FONT[7] & ROGER VILA[1]

[*] These authors contributed equally to the paper.

[1] Institute of Evolutionary Biology, Passeig Marítim de la Barceloneta 37, 08003, Barcelona, Spain

[2] Department of Ecology and Genetics, PO Box 3000, 90014, University of Oulu, Oulu, Finland

[3] Dipartimento di Biologia, Università degli Studi di Firenze, Via Madonna del Piano 6, 50109, Sesto Fiorentino, Florence, Italy

[4] Museu [UV] Història Natural. Universitat de València. C/. Dr. Moliner, 50. E-46100, Burjassot (Valencia), Spain

[5] Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università degli Studi di Torino, Via Accademia Albertina 13, 10123, Turin, Italy

[6] Department of Biological Sciences, Macquarie University, Sydney, NSW2109, Australia

[7] Ethology Lab, Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Valencia APDO 22085, 46071, Spain

Gaunet et al. (2019) Two consecutive *Wolbachia*-mediated mitochondrial introgressions obscure taxonomy in Palearctic swallowtail butterflies (Lepidoptera, Papilionidae). *Zoologica Scripta*, 00, 000-000.

Swallowtail butterflies (Papilionidae) are among the most spectacular and well known Lepidoptera in the European fauna, but their systematics is not fully elucidated. A notable case is that of *Iphiclides feisthamelii* which, after more than 180 years since description, still has a debated status, being often considered as a subspecies of *I. podalirius*. To elucidate the relationship between the two taxa and the evolutionary processes that led to their separation, we combine mitochondrial and nuclear DNA (mtDNA and nDNA) data, *Wolbachia* screening, genitalia morphology and wing UV reflectance. Our results show that the two taxa clearly differ in male and female genital morphology, male wing UV reflectance and nDNA. Two *Wolbachia* strains were found to widely infect the studied samples, apparently explaining the phylogeographic pattern displayed by mtDNA. The available data point towards a historical *Wolbachia* infection that spread from *I. podalirius* to *I. feisthamelii* and produced a mitochondrial introgression. Currently, a new *Wolbachia* strain is spreading across mainland populations of *I. podalirius*, mediating once more a mitochondrial genetic sweep, which has already infected and introgressed *I. feisthamelii* populations in south-eastern France. We conclude that, given the marked differences in morphology and nDNA between the two taxa, and the apparent restriction of hybridization to a narrow contact area where non-hybrid specimens are common, the taxon *feisthamelii* should be considered as a separate species. Within this species, two well-differentiated nDNA lineages that represent European and Maghrebian populations are documented, here proposed as subspecies. The case of, presumably, two consecutive *Wolbachia*-mediated mitochondrial introgression events, further supports the view that infection by this endosymbiont may be frequently related to mito-nuclear discordance in insects.

Introduction

Although butterfly species have been described in the last 250 years mostly based on morphological characters, the development of molecular techniques in the last decades has provided additional and highly reliable tools for species discovery. One of the most used markers for molecular assessments is mitochondrial DNA (mtDNA) (Brown et al., 1979). However, this maternally inherited marker may not reveal by itself an accurate view of the history of species. For instance, cases of introgression, hybridization or incomplete lineage sorting are only discovered by analysing nuclear DNA (nDNA) (Toews et al., 2012). Mito-nuclear discordance is thus a relatively common phenomenon (e.g. Gompert et al., 2008; Wahlberg et al., 2009; Tóth et al., 2017) and, in some instances, it may be related to infection by bacteria such as *Wolbachia* (Kodandaramaiah et al., 2013). *Wolbachia* is a bacterial endosymbiont of arthropods and it is extremely widespread among insects (Zug et al., 2012). As it is maternally inherited, infection and different strains of *Wolbachia* tend to co-vary with mitochondrial lineage.

Butterflies of the family Papilionidae comprise some of the largest and most spectacular butterflies in the world. They also include emblematic insects for the European fauna, with five species protected at European level (Annexes II and/or IV of the Habitats Directive 92/43/EEC) and two, *Papilio hospiton* (Géné, 1839) and *Parnassius apollo* (Linnaeus, 1758), included in the CITES Appendices I and II, respectively (the only European butterflies listed in CITES Appendices). Despite their conspicuous appearance, popularity among researchers and the general public, and relevance for conservation, the species composition of European Papilionidae is still not fully clarified. Indeed, a new species endemic to Italy has been documented only recently (Dapporto, 2010; Zinetti et al., 2013), raising the number of species to 14. Furthermore, taxonomic uncertainty still exists regarding the genus *Iphiclides*, traditionally regarded as comprising just two species worldwide: *I. podalirius* (Linnaeus, 1758), widely distributed from north-western Africa to central Asia, and *I. podalirinus* (Oberthür, 1890), restricted to parts of Tibet (Racheli & Cotton, 2009). The taxon *feisthamelii* (Duponchel, 1832), distributed in north-western Africa and the Iberian Peninsula (e.g. Tennent, 1996; Lafranchis, 2003; Tarrier & Delacre, 2008; Tshikolovets, 2011), is regarded as either a subspecies of *I. podalirius* (e.g. De Prins & Iversen, 1996; Tolman & Lewington, 2008; García Barros et al., 2013; Kudrna et al., 2015), or as a distinct species (e.g. Tshikolovets, 2011; Vila et al., 2018). Currently, in the checklist of the Fauna Europaea Project (Karsholt & van Nieukerken, 2013; http://fauna-eu.org/), the taxon *feisthamelii* is

considered as a subspecies of *I. podalirius*. The controversy probably exists because these two taxa are ecologically, behaviourally and morphologically very similar, (Lafranchis, 2015a; Leraut, 2016). Both usually have two generations per year (with sometimes a third partial generation in the southern populations, at the end of the summer), they occupy a wide variety of open habitats, ranging from sea level to over 2000 m.a.s.l. and they apparently prefer open meadow areas with a few scattered trees, where nectar is abundantly available. The males are territorial, and their dispersal is mainly limited to the closest favourable hill-topping areas. Females are more mobile and appear able to disperse and lay their eggs over long distances. Most of the time, eggs are laid on the underside of leaves, mainly on trees of the genus *Prunus*. However, *Crataegus monogyna* and other trees in the Rosaceae family are sometimes used as well (Lafranchis, 2015a). No morphological differences were found in caterpillars of both species (Lafranchis, 2015a). Caterpillars spend most of their time immobile on a leaf tip, where they spin a silk cushion. In order to be more mimetic, larvae growing at the end of summer have more brown spots on their dorsal surface than larvae growing in spring. Both species spend winter as pupae (Lafranchis, 2015a), and the chrysalides that hibernate are brown while the others remain green.

Although various morphological differences between the two taxa have been noted, intermediate individuals are reported in the contact zone in south-eastern France (Lafranchis, 2015b). Moreover, local variants exist, as is the case in some third-generation individuals of Greek *I. podalirius,* which have a habitus closer to that of *I. feisthamelii* (Lafranchis, 2015b). The uncertainty surrounding the status of the taxon *feisthamelii* has led to a series of studies attempting to clarify its relationship with *I. podalirius*. Wiemers & Gottsberger (2010) sequenced three specimens of *I. podalirius* and three of *I. feisthamelii* and reported a discordant pattern between mitochondrial data on one hand, and nuclear and morphological data on the other hand. Coutsis & Van Oorschot (2011) reported differences in the genitalia of both sexes but their results were not based on morphometrics and the studied material included a small number of individuals from just a few localities. Dincă et al. (2015) confirmed that genital and genetic differences exist between the two taxa, but their study was not focused on these taxa and consequently the sampling was limited. Nevertheless, the authors identified mtDNA and nDNA patterns that suggest mitochondrial introgression from *I. podalirius* into Iberian *I. feisthamelii*. Recently, Lafranchis et al. (2015b) studied the contact zone between the two taxa in France and concluded that they represent two incompletely separated species that still experience a certain exchange of genetic material. However, Lafranchis et al. (2015b) did not rely on molecular data and their study was based only on

morphology and ecology. Thus, no comprehensive phylogenetic analysis and/or morphometric studies exist for the *Iphiclides* genus that describe their evolutionary history and clarify whether *I. podalirius* and *I. feisthamelii* represent two distinct species.
In this study, we use a combination of genetic (mt and nDNA markers), morphological (morphometrics and wing UV reflectance patterns) and microbiological data (*Wolbachia* screening) to investigate the phylogeography and taxonomic status of *I. podalirius* and *I. feisthamelii*, and the processes that led to the apparent mito-nuclear discordance between them. Our study is also the first comprehensive analysis covering a large part of the distribution range of the two taxa.

Materials and methods

Data collection

Locality and data analysis information for the specimens used in this study are listed in Table S1. All samples are deposited in the Butterfly Diversity and Evolution Lab collection at the Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain. For the genetic analyses, we also included all *Iphiclides* sequences available in GenBank, most of which were published by Wiemers & Gottsberger (2010) or Dincă et al. (2015). We were not able to obtain specimens of *I. podalirinus* and no sequences were available in GenBank.

Morphometric analyses

Elements of the male and female genitalia were measured for morphometric analyses. We selected 139 male and 61 female specimens from 149 localities to comprehensively sample the distribution of the two taxa (from southern Maghreb to eastern Kazakhstan), including the contact zones from northern Iberia and southern France (Fig. S3; detailed sampling information is provided in Table S1). We measured three elements of the male phallus (Fig. 1A), and two elements of the female ductus bursae (Fig. 1B). For males, we normalized ostium length and width by phallus length, as this provides a better separation. We also measured wing size for all individuals (Fig. S2; further details in the supplementary material).

UV reflectance of wings

We employed two methods in order to measure wing UV reflectance. The first method, more technically accessible, was to measure UV reflectance from digital photographs. With this method we analysed the differences between both taxa in UV wing pattern reflectance for both sexes (between 320-380 nm). We analysed a total of 132 fresh specimens (88 males and 44 females), whose genitalia had been previously extracted (Fig. S5; detailed sampling information is provided in Table S1). All specimens were photographed with a Nikon D70 and a Nikon lens (AF Micro Nikkor 60mm) coupled to two flashes and using a Baader U-Filter (60nmHBW/320-380nm fully blocked VIS & IR). We assessed UV reflectance using the software Adobe Photoshop in two squares of the upperside of the hind-wing: one location in the blue crescent-shaped markings (lunules) and the other location on the pale yellow wing background (Fig. 2A; further details in the supplementary material). To validate the first methodology we also took measurements using a reflectance spectrophotometer. For this we obtained the reflectance spectra for both the blue lunules and the pale yellow wing background on both hindwings of 14 individuals belonging to each taxon. We then analysed spectral data in R v.3.3.2 (R Development Core Team, 2008) using the software packages *pavo* and *lme4* (Bates et al., 2015). The detailed protocol is provided in the supplementary material.

DNA sequencing and analyses

The mitochondrial marker cytochrome *c* oxidase subunit I (*COI*) was sequenced from 117 individuals. Additionally, a total of 73 sequences were obtained from GenBank. The nuclear internal transcribed spacer 2 (*ITS2*) was amplified from the same specimens as for *COI* but we were only able to obtain good sequences from 57 specimens. DNA extraction, amplification, sequencing and alignment protocols are described in the supplementary material. For each DNA marker, we inferred phylogenetic relationships with Bayesian inference in BEAST v1.8.0 (Drummond & Rambaut, 2007) (details in the supplementary material). All sequences obtained in this study are available in GenBank (accession numbers: MK587175-MK587438), and in the dataset DS-IPHICLID (DOI: dx.doi.org/10.5883/DS-IPHICLID) from the Barcode of Life Data Systems (http://www.boldsystems.org/).

Delimitation of specimens to two possible taxa

In order to objectively assign the specimens to one of the two taxa, either *I. feisthamelii* or *I. podalirius* and visualize their spatial pattern for all analysed markers, we applied the k-means

clustering method and forced two groups for phenotypic markers and two, three or four groups for genotypic markers. For *COI* and *ITS2* we calculated dissimilarity matrices using p-distances among our *n* specimens and projected them in *n-1* dimensions, employing a Principal Coordinate Analysis (PCoA) with the 'cmdscale' R function. Then, we applied k-means to the coordinates. For genital morphology and UV reflectance data, we applied k-means separately for the measurements made on male and female specimens. The specimens belonging to the groups obtained for each marker were plotted with different colours on a map. Specimens belonging to the same grid square of 2° for latitude and longitude were grouped, and their colours were plotted on a map using pie charts. For the contact zone (near the Pyrenees) we grouped specimens to squares of 0.2° of latitude and longitude in order to get a better resolution.

Presence and identification of Wolbachia *strains*

A total of 66 specimens, covering most of the distribution of the two *Iphiclides* were surveyed for the presence of the maternally inherited bacterial endosymbiont *Wolbachia,* using primers that amplify the markers *wsp* and *coxA*. In case of infection, the fragments amplified were sequenced in order to assess the strain. For details on the methods, see the supplementary material.

Results

Morphometric analyses

Application of k-means forcing the formation of two clusters based on the measured genital elements of males and females produced a spatial distribution of specimens belonging to either one of the two groups that highly corresponds to the distribution of the taxa *podalirius* and *feisthamelii* (Fig. S3A, C). Indeed, for both sexes we found constant characteristics that differentiated the two taxa based on the measured elements (Fig. 1C): *I. podalirius* have larger genitalia than *I. feisthamelii,* and the second generation *I. podalirius* tend to have smaller genitalia than the first, particularly notable in females. The genitalia of *I. feisthamelii* did not display notable differences between generations (Fig. 1C). Regarding the valva, no constant difference was found between the two taxa. In the supplementary material we illustrate

examples of the variability of this part of the genitalia throughout the distribution range of the two taxa (Fig. S4).
(Fig. 1 here)

Analyses of wing UV reflectance

A k-means analysis for UV reflectance patterns based on digital photographs showed high congruence with the spatial distribution of the two taxa for males (Fig. S5A, B), but not for females (Fig. S5C, D). Based on the k-means identifications, we demonstrate that, in males, the pale background on the upper side of the forewing reflects significantly more UV in *I. feisthamelii* than in *I. podalirius* ($W = 0$, p-value < 0.001). However, in *I. feisthamelii* the UV reflectance of blue lunules is lower than that of pale background ($t = -8.2337$, $df = 42.327$, p-value < 0.001), while in *I. podalirius* males, it's the blue lunules that have a higher intensity ($W = 2915$, p-value < 0.001) (Fig. 2B).
In male specimens identified as *I. podalirius*, we also found significant differences between the first and the second generation both for blue lunules ($t = -2.8242$, $df = 53$, p-value = 0.007) and for pale background ($t = -4.0479$, $df = 49.146$, p-value < 0.001). In both cases, the second generation had a higher UV reflectance than the first. In the supplementary material we illustrate examples of the variability of UV reflectance patterns displayed by the two taxa throughout their distribution range (Fig. S6).
In the case of the females, it was not possible to assign the specimens to one of the two taxa based on the k-means method using UV reflectance. Thus, we used k-means assignment based on female genital measurements and corroborated by the nuclear *ITS2* marker when available. The resulting scatter plot (Fig. 2B) shows considerable overlap in the UV reflectance of the two species for females, while for males no overlap exists.
(Fig. 2 here)

Regarding the analyses of wing full spectrum reflectance using a spectrophotometer (Fig. 3), we found no significant inter-taxa differences for the blue lunules in terms of luminance, hue or chroma for males or for females (Table S2). In the case of the pale background however, we found that male *I. podalirius* had higher green and yellow chroma but lower UV and violet chroma than male *I. feisthamelii* (Fig. 3C; Table S2). Similarly, the pale background of female *I. podalirius* had a higher yellow chroma but lower UV and violet chroma than that of female *I. feisthamelii* (Fig. 3D; Table S2). The differences in the UV spectrum between both

sexes are much more noticeable in males than females. We did not detect any significant differences between the two taxa in terms of luminance, or hue of the pale background. These results are in partial agreement with the results obtained using the photographic method, which also suggested inter-taxa differences in the coloration of the pale background but not the blue lunules. The spectrophotometer has also revealed important differences between males and females of both species, outside the UV range that has not been analysed by the photographic method, especially in the yellow and green spectrum.

Further, there is a statistically significant correlation between the luminance of the pale background as measured with reflectance spectrophotometry ($R^2 = 0.7196$, $p < 0.0001$; Fig. S7) and the scores obtained with the photographic method, which suggests that both types of measurements produce similar results.

(Fig. 3 here)

Genetic analyses

COI. The Bayesian analysis based on *COI* sequences (Fig. 4, Fig. S9) recovered two well-supported main clades, displaying a minimum uncorrected p-distance of 2.0% between sequences originating from each clade. The estimated time of divergence between the two was 1.25 million years (Ma) (0.7-1.94 Ma, CI 95%). One clade consists exclusively of specimens of *I. feisthamelii* from the Maghreb, while the other includes all other *Iphiclides* specimens analysed, ranging from Iberia (mostly attributable to *I. feisthamelii* based on morphology and nDNA) to Kazakhstan, and also including islands - Corsica, Sardinia, the Tyrrhenian islands, Sicily and Crete. Within this widespread clade, specimens from Crete formed a lineage that was well supported as sister to all other individuals and displayed a minimum genetic p-distance of 0.6% to the closest conspecific. Its split from the rest of the Eurasian clade was estimated to 0.4 Ma (0.18-0.66, CI 95%). Finally, within the Eurasian clade, specimens from Sicily were monophyletic with relatively good support (pp 0.96). When imposing two clusters in the k-means analysis, we obtained a division between North African and European specimens, when forcing three clusters the Crete lineage is separated from European specimens, and when forcing four clusters a western clade of the European lineage (Iberia and S. France) that also occurs in Kazakhstan, Turkey, Corsica and Sicily emerged (Fig. 5A, B).

(Fig. 4 here)

ITS2. Bayesian phylogenetic analysis based on *ITS2* sequences recovered three main well-supported clades (Fig. 4, Fig. S10): one included all the specimens attributable to *I. podalirius* based on morphological data, another included all European specimens attributable to *I. feisthamelii*, while the third comprised all Maghrebian specimens of *I. feisthamelii*. The sister-group relationship between the two clades of *I. feisthamelii* was not supported (pp 0.81). The European and Maghrebian clades of *I. feisthamelii* displayed a minimum uncorrected p-distance of 0.9% and 1.3%, respectively with respect to *I. podalirius*. The two clades of *I. feisthamelii* displayed a minimum p-distance of 1.1%. The Cretan clade of *I. podalirius* recovered by *COI* did not form a clade when *ITS2* was used.

When imposing two clusters in the k-means analysis a main division between the supposed *I. podalirius* and *I. feisthamelii* taxa was recovered, and when forcing three clusters, the Iberian specimens were separated from North African ones (Fig. 5C, D). A solution with four clusters separated the European clade without any spatial coherence. A closer look at the contact zone (Fig. 5D) emphasized highly congruent results to those obtained for male and female genital morphology and for male UV reflectance.

(Fig. 5 here)

Presence and identification of Wolbachia *strains*

Wolbachia infection tests showed a generalised infection of Eurasian *I. podalirius* and *I. feisthamelii* (Fig. 6C). All the individuals from Eurasia, including Mediterranean islands, were positive for the two targeted markers (except for *coxA* in Crete). On the contrary, none of the Maghrebian individuals showed signs of infection by *Wolbachia* (Fig. 6C). We detected the presence of two main combinations of sequences (wsp1-coxA1 and wsp2-coxA1, which we will refer to as “strains”, see below) that follow closely, but not exactly, the European mainland distributions of *I. podalirius* and *I. feisthamelii*. One of them (wsp2-coxA1) is present in Western Europe (except for the Iberian Peninsula), Central Europe and extends to mainland Greece and Turkey. The other one (wsp1-coxA1) is located in the Iberian Peninsula, and three Mediterranean islands: Corsica, Sardinia and Sicily. Interestingly, despite being separated by a great geographical distance, this strain is also present in eastern Kazakhstan. Samples from Crete appear to be infected by a third strain because *coxA* was not successfully amplified while *wsp* worked very well (sequence identical to wsp1), which suggests that this strain mutated in the regions targeted by the *coxA* primers used.

Our wsp1 and wsp2 alleles were highly divergent, displaying a pairwise distance of 18.3%. When compared with sequences available in the *Wolbachia* MLST database (https://pubmlst.org/wolbachia/) wsp1 was closest to wsp 593, from which it differed by seven mutations; wsp2 was identical to wsp 10 and coxA1 was identical to coxA 14. All three alleles belong to *Wolbachia* supergroup B. According to the data available in the *Wolbachia* MLST database, wsp 593 was detected in *Cotesia* (*Apanteles*) *chilonis* (Hymenoptera, Braconidae). Allele wsp 10 was detected in various species of Lepidoptera belonging to several families (e.g. Papilionidae, Lycaenidae, Pieridae, Nymphalidae, Tortricidae, Pyralidae etc.), as well as in *Culex* species (Diptera). Allele coxA 14 was reported from various species of Lepidoptera belonging to several families (e.g. Papilionidae, Lycaenidae, Pieridae, Nymphalidae, Geometridae, Crambidae etc.), as well as from a few species of Coleoptera, Diptera, Hymenoptera, Hemiptera and Thysanoptera.

Discussion

Morphometric analyses suggest constant differences between the two taxa

Since the description of *I. feisthamelii* by Duponchel more than 180 years ago, several morphological characters that differentiate it from *I. podalirius* have been proposed, either related to wing colour patterns, genitalia, or the larval stage. However, several authors, including Duponchel, Oberthür and Verity, oscillated between treating *I. feisthamelii* as a species and a subspecies or variety of *I. podalirius*. More recently, analyses of the genital apparatus of both males and females suggested that there are constant differences between the two taxa, but the sampling of those studies was too limited to draw a statistically supported conclusion (Coutsis & Van Oorschot, 2011). Our analyses, which involved a large number of specimens representative for the distribution range of the taxa, show that both male and female specimens can be divided into two coherent groups on the basis of their genital shape (Fig. 1) that coincide with the presumed distribution of the two taxa (Fig. S3). *Iphiclides podalirius* has ~~a~~ larger genitalia compared to *I. feisthamelii*, and the differences are apparently constant across their entire distribution range (Fig. 1). Moreover, these dissimilarities were also maintained in specimens studied from the contact zone, and thus it seems that no hybrid was included in our dataset. Lafranchis et al. (2015b) discuss the existence of potential hybrids, both natural and obtained in experimental crosses. In any case, hybrids seem to be rare and limited to the contact zone.

The analysis of the wing UV reflectance also suggested a differentiation between males of the two taxa, with the pale background of the wings of *I. feisthamelii* males reflecting more UV light than those of *I. podalirius*. Since for our study we used a large number of specimens from many localities across the entire distribution range of the two taxa, we were able to exclude potential wing patterns caused by environmental variables such as sun exposure, specimen preservation and date of collection. For example, *I. podalirius* individuals originating from the Mediterranean – Corsica, Sardinia, Greece or Turkey, show the same UV reflectance pattern as those that originate from mountains or from northern areas. We have not been able to apply spectrophotometry for all samples during the main analysis, so we used a simple photographic method focused on the UV spectrum in order to analyse the high number of samples. However, we have selected a set of specimens on which we have used spectrophotometry to ensure that the first methodology was adequate. The photographic and spectrophotometry analyses led to convergent results. Both methodologies reveal significant differences in terms of male UV reflectance. Furthermore, the photographic data highlight an overlap between females from both taxa, which was also observed using the spectrophotometric approach. Interestingly, spectrophotometry revealed differences between both species not only in UV chroma, but also throughout the entire spectrum, especially in yellow and green chroma. It would be interesting to analyse these differences in future studies. The marked UV reflectance differentiation of the two taxa found exclusively in males suggests that mate choice is done mostly by the females, as is frequently the case in butterfly species (Friberg et al*.,* 2008; Dincă et al*.,* 2013, Southcott & Kronforst, 2018).

Mito-nuclear discordance presumably mediated by Wolbachia

Our study confirms the discordant pattern found between the mitochondrial (*COI*) and nuclear (*ITS2*) markers (Fig. 4, Fig. S9, Fig. S10) (Wiemers & Gottsberger, 2010, Dincă et al., 2015), which supports the hypothesis of mitochondrial introgression from *I. podalirius* to Iberian *I. feisthamelii*. Thus, information based exclusively on the mtDNA variation among European *I. podalirius* and *I. feisthamelii* is not sufficient by itself to separate the two species. The *COI* gene recovered two main clades: one formed by *Iphiclides* specimens from Africa and the other from Eurasia (including east Kazakhstan). The *ITS2* marker recovered three main clades, including two clades that match perfectly with the presumed distribution of the taxa and with morphological results (genitalia and male UV reflectance), even at fine scale at the contact zone. The *I. feisthamelii* clade is further divided into two well-differentiated lineages,

one Iberian and one North African that apparently diverged soon after the split of *podalirius* and *feisthamelii*, more than 1 Ma.

Exploration of the infection pattern by the endosymbiont *Wolbachia* across the range of both taxa demonstrates that all Eurasian *Iphiclides* surveyed (Kazakhstan included) were infected by *Wolbachia,* while none of the North African individuals were. Differences in *Wolbachia* sequences suggest the presence of two main strains in mainland Europe that approximate the distribution of *I. podalirius* and *I. feisthamelii*, except for south-eastern France, where *I. feisthamelii* is infected by the wsp2-coxA1 strain.

Overall, characteristics such as male and female genitalia, male UV reflectance of the wings and the *ITS2* nuclear marker confirm the existence of two divergent parapatric entities, presumably possessing at most a poor hybridisation capacity. These elements lead us to recognize the existence of two species: *Iphiclides podalirius* and *Iphiclides feisthamelii*. We suggest that *I. podalirius* is a monotypic species since genetic and morphological differences across its range are not pronounced (albeit more detailed analyses on Sicilian and Cretan populations may shed light on morphological/ecological differences here unnoticed). Regarding *I. feisthamelii*, the differentiation observed for the *ITS2* nuclear marker between Iberian and North African populations suggests that their divergence is considerably old. Moreover, the apparent absence of mitochondrial gene flow among these populations demonstrates that the Mediterranean represents a stable geographical barrier. However, beyond the larger size of North African specimens (Fig. S8), we did not observe any strong morphological difference in the characters here studied. As a consequence, we recognize two subspecies for *I. feisthamelii*: one Iberian and one North African. *Iphiclides feisthamelii feisthamelii* (Duponchel, 1832) is present in Portugal, Spain and in south-eastern France, whereas *Iphiclides feisthamelii lotteri* (Oberthür, 1879), which appears to be the valid name (Leraut, 2016), is present in Morocco, Algeria and Tunisia.

Evolutionary history

The endosymbiotic bacterium *Wolbachia* is widespread in insects (Zug et al., 2012) and infection has diverse consequences for the host, such as male-killing or cytoplasmic incompatibility (Werren et al., 2008). As *Wolbachia* and the mitochondria are both maternally inherited, unusual patterns in mtDNA may be explained by the effects of *Wolbachia* on insect populations. Thus, it is advisable to compare patterns of infection by this endosymbiont to those of mtDNA, especially in cases where mito-nuclear discordance occurs.

Several cases of highly divergent mitochondrial intraspecific lineages have been hypothesised to arise because of *Wolbachia*-mediated genetic sweeps and cytoplasmic incompatibility (e.g. Ritter et al., 2013). The transfer of *Wolbachia* infection across species through occasional hybridisation may also lead to mitochondrial introgression (Dumas et al., 2013, Hernández-Roldán et al., 2016).

(Fig. 6 here)

Although there is a generally good correlation between the distribution of *COI* haplogroups (Fig. 5A) and the distribution of *Wolbachia* strains (Fig. 6C), the hypothesis explaining the current patterns in *Iphiclides podalirius* and *I. feisthamelii* is necessarily complex (Fig. 6). We hypothesise that strain wsp1-coxA1 infected *I. podalirius* and spread throughout Europe, including the Mediterranean islands and Kazakhstan, but possibly not Crete, which may harbour a relict *Wolbachia* strain and *COI* lineage (Fig. 6A). Subsequently, this strain spread from *I. podalirius* to Iberian *I. feisthamelii* and caused a genetic sweep in the latter, which led to the apparently fixed mitochondrial introgression that we documented in the Iberian Peninsula (Fig. 6B). Note that the asymmetrical nature of cytoplasmic incompatibility typically confers advantage to *Wolbachia*-infected females, which would explain the rapid expansion of the infections and their associated mtDNA.

Then a new infection by the wsp2-coxA1 strain occurred in *I. podalirius*, which spread across continental Europe replacing wsp1-coxA1, but was not able to spread to geographically isolated island populations (Fig. 6C). This is not surprising because prevailing winds appear to underlie the reproductive isolation of many butterfly species in the Messina Strait (Dapporto et al., 2010). In the same way, geographical isolation explains the absence of *Wolbachia* in North Africa (Strait of Gibraltar). Finally, geographical distance may best explain the persistence of the wsp1-coxA1 ancestral strain of *Wolbachia* in the far east of Kazakhstan, and it is possible that the new wsp2-coxA1 strain is still spreading eastwards. Additional sampling in the eastern part of the range will help testing this hypothesis. The spread of the wsp2-coxA1 strain westwards replacing wsp1-coxA1, on the contrary, has reached the current limits of *I. podalirius*, and it has even spread into *I. feisthamelii* in south-eastern France (Fig. 6C).

Interestingly, the spread of wsp2-coxA1 *Wolbachia* strain within *I. podalirius* and into *I. feisthamelii* is apparently causing a mitochondrial genetic sweep, because a *COI* genetic lineage (highlighted in k = 4 in Fig 5A, B) correlates closely with the infection. Only at the geographical limits of this new infection (in southern France and in Turkey), some specimens of *I. podalirius* are infected by the wsp2-coxA1 strain but with the presumably ancestral *COI*

lineage. On the other hand, the only specimen analysed from Sardinia presented the *Wolbachia* strain wsp1-coxA1 and the *COI* lineage typical of mainland Europe. It is worth noting that in Sardinia only sporadic specimens have been reported, suggesting a potential allochtonous origin. Thus, the *Wolbachia*-mediated mitochondrial genetic sweep in *I. podalirius* and the introgression to *I. feisthamelii* are apparently still ongoing, and could represent an interesting system for the study of the mechanisms and consequences of this phenomenon.

Conclusion

This is the first study to combine genetic, morphological and microbiological data for a large sample of European and North African *Iphiclides*. Our results confirm the species status of *Iphiclides feisthamelii* and *Iphiclides podalirius*, which differ clearly in nDNA, male and female genitalia, and male UV wing reflectance. We further divide *Iphiclides feisthamelii* into two subspecies: *I. f. feisthamelii* in Portugal, Spain and in south-eastern France, and *I. f. lotteri* in Morocco, Algeria and Tunisia. We document a mitochondrial introgression event from *I. podalirius* to *I. feisthamelii* that was probably mediated by the endosymbiont *Wolbachia*. An ongoing replacement of *Wolbachia* strains within *I. podalirius* (already progressing into *I. feisthamelii* as well) is also observed, which again mediates a mitochondrial genetic sweep. The complex temporal and spatial patterns for the interaction between *Wolbachia* and the *Iphiclides* sister species revealed in this study render this system as a potential new model for the study of the effects of this endosymbiont on the evolution and speciation of insects.

Acknowledgements

We thank Alex Hayward, Zdeněk Faltýnek ~~Zdenek Faltynek~~ Fric and an anonymous reviewer for suggestions and language editing. Funding for this research was provided by project CGL2016-76322 (AEI/FEDER, UE) to RVi, Marie Skłodowska-Curie Train2Move to RVo (grant 609402-2020), Marie Skłodowska-Curie International Outgoing Fellowship within the 7th European Community Framework Programme (project 625997) to VD and from the project “Barcoding-Italian-Butterflies”. We are grateful to all the colleagues who provided samples used in this study.

Data, code and materials

The *Iphiclides* (*COI* and *ITS2*) and *Wolbachia* (*wsp* and *coxA*) sequences generated for this study are available in GenBank (accession numbers: MK587175-MK587438), and are also publicly available in the dataset DS-IPHICLID (DOI: dx.doi.org/10.5883/DS-IPHICLID) from the Barcode of Life Data Systems (http://www.boldsystems.org/).

References

Arnqvist, G. (1997) The evolution of animal genitalia: distinguishing between hypotheses by single species studies. *Biological Journal of the Linnean Society*, 60(3), 365–379.

Austaud, J.L. (1879) Lépidoptères nouveaux d'Algérie. *Petites Nouvelles Entomologiques*, *2*(212), 293.

Bates, D., Mächler, M. Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Brunton, C. F. & Majerus, M. E. (1995) Ultraviolet colours in butterflies: intra-or inter-specific communication? *Proceedings of the Royal Society of London B: Biological Sciences*, 260(1358), 199–204.

Cong, Q., Shen, J., Borek, D., Robbins, R.K., Opler, P.A., Otwinowski, Z. & Grishin, N.V. (2017) When *COI* barcodes deceive: complete genomes reveal introgression in hairstreaks. *Proceedings of the Royal Society B: Biological Sciences*, 284.

Coutsis, J. & Van Oorschot, H. (2011) Differences in the male and female genitalia between *Iphiclides podalirius* and *Iphiclides feisthamelii*, further supporting species status for the latter (Lepidoptera: Papilionidae). *Phegea*, 39, 12–22.

Dapporto, L. (2010) Speciation in Mediterranean refugia and post-glacial expansion of *Zerynthia polyxena* (Lepidotera, Papilionidae). *Journal of Zoological Systematics and Evolutionary Research*, 48, 229–237.

Dincă, V., Wiklund, C., Lukhtanov, V.A., Kodandaramaiah, U., Norén, K., Dapporto, L., Wahlberg, N., Vila, R. & Friberg, M. (2013) Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. *Journal of Evolutionary Biology,* 26, 2095–2106. doi: 10.1111/jeb.12211

Dincă, V., Montagud, S., Talavera, G., et al. (2015) DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Scientific Reports*, 5, 12395.

Descimon, H. & Mallet, J. (2009) Bad species. *Ecology of butterflies in Europe*, vol. 500(C), p. 219.

Dumas, E., Atyame, C.M., Milesi, P., Fonseca. D.M., Shaikevich, E.V., Unal, S., Makoundou, P., Weill, M. & Duron, O. (2013) Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evolutionary Biology*, 13, 181.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214. https://doi.org/10.1186/1471-2148-7-214.

Friberg, M., Vongvanich, N., Borg-Karlson, A.K., Kemp, D.J., Merilaita, S. & Wiklund, C. (2008) Female mate choice determines reproductive isolation between sympatric butterflies. *Behavioral Ecology and Sociobiology*, 62: 6, 873–886.

García-Barros, E., Munguira, M.L., Stefanescu, C. & Vives Moreno, A. (2013) Lepidoptera Papilionoidea. Fauna Ibérica, vol. 37. Ramos, M.A, et al. (Eds). – Museo Natural de Ciencias Naturales, CSIC, Madrid, 1213 pp.

Gompert, Z., Forister, M.L, A Fordyce, J. & Nice, C.C. (2008) Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. *Molecular Ecology*, 17(24), 5231–5244.

Hernández-Roldán, J. L., et al. (2016) Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Molecular Ecology,* 25, 4267–4284.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., et al. (2008) How many species are infected with *Wolbachia*? A statistical analysis of current data. *FEMS Microbiology Letters.* 281(2), 215–220.

Kageyama, D., Nishimura, G., Hoshizaki, S. & Ishikawa, Y. (2002) Feminizing Wolbachia in an insect, *Ostrinia furnacalis* (Lepidoptera: Crambidae). *Heredity*, 88(6), 444.

Karsholt, O. & van Nieukerken, E.J. (2013) Fauna Europaea: Lepidoptera. *Fauna Europaea* version 2017.06. Available at https://fauna-eu.org.

Kodandaramaiah, U. (2009) Effects of Wolbachia on Butterfly Life History and Ecology. *Advances in Medicine and Biology*, Volume 16. Nova Publishers, New York.

Kodandaramaiah, U., Simonsen, T.J., Bromilow, S. et al. (2013) Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution*, 3, 5167–5176.

Lafranchis, T., Jutzeler, D., Guillosson, J. Y., Kan, P. & Kan, B. (2015a) La vie des papillons. Ecologie, Biologie et Comportement des Rhopalocères de France. *Diatheo*, 130–136.

Lafranchis, T., Mazel, R. & Delmas, S. (2015b) Le contact *Iphiclides feisthamelii – I. podalirius.* Statut de ces deux taxons. *Revue de l'Association Roussillonnaise d'Entomologie*, 24 (3), 111–132.

Leraut, P. (2016) Papillons de jour d'Europe et des contrées voisines. *NAP editions*: 1100 p.

Narita, S., Nomura, M., Kato, Y. & Fukatsu, T. (2006) Genetic structure of sibling butterfly species affected by *Wolbachia* infection sweep: evolutionary and biogeographical implications. *Molecular Ecology*, 15(4), 1095–1108.

Racheli, T. & Cotton, A.M. (2009) Papilionidae part I. *In* Bozano, G. C. [ed.], *Guide to the Butterflies of the Palaearctic region*. Omnes Artes, Milano.

Ritter, S., Michalski, S.G., Settele, J., Wiemers, M., Fric, Z.F., Sielezniew, M., et al. (2013) *Wolbachia* infections mimic cryptic speciation in two parasitic butterfly species, *Phengaris teleius* and *P. nausithous* (Lepidoptera: Lycaenidae). *PLoS one*, 8(11): e78107. https://doi.org/10.1371/journal.pone.

Robertson, K. A. & Monteiro, A. (2005) Female *Bicyclus anynana* butterflies choose males on the basis of their dorsal UV-reflective eyespot pupils. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1572), 1541–1546.

Sasaki, T., & Ishikawa, H. (1999) *Wolbachia* infections and cytoplasmic incompatibility in the almond moth and the Mediterranean flour moth. *Zoological Science*, 16(5), 739–744.

Southcott, L. & Kronforst, M.R. (2018) Female mate choice is a reproductive isolating barrier in *Heliconius* butterflies. *Ethology*, 124, 12, 862–869.

Toews, D.P. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16), 3907–3930.

Tóth, J.P., Varga, Z., Verovnik, R., Wahlberg, N., Váradi, A. & Bereczki, J. (2017) Mito-nuclear discordance helps to reveal the phylogeographic patterns of *Melitaea ornata* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, 121(2), 267–281. https://doi.org/10.1093/biolinnean/blw037

Tshikolovets, V. (2011) Butterflies of Europe and the Mediterranean area. *Tshikolovets publications*, Pardubice: 544 p.

Van Opijnen, T., & Breeuwer, J.A.J. (1999) High temperatures eliminate *Wolbachia*, a cytoplasmic incompatibility inducing endosymbiont, from the two-spotted spider mite. *Experimental and Applied Acarology*, *23*(11), 871–881.

Vila, R., Stefanescu, C. & Sesma, J.M. (2018) Guia de les papallones diürnes de Catalunya. *Lynx Edicions*, Bellaterra: 509 p.

Wahlberg, N., Weingartner, E., Warren, A.D. & Nylin, S. (2009) Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini). *BMC Evolutionary Biology* 9(1), 92. doi:10.1186/1471-2148-9-92.

Werren, J. H., Baldo, L. & Clark, M.E. (2008) *Wolbachia*: master manipulators of invertebrate biology. *Nature Reviews Microbiology*, 6, 741–751.

Wiemers, M., & Gottsberger, B. (2010) Discordant patterns of mitochondrial and nuclear differentiation in the Scarce Swallowtail *Iphiclides podalirius feisthamelii* (Duponchel, 1832) (Lepidoptera: Papilionidae). *Entomologische Zeitschrift*, 120, 111–115.

Zinetti, F., Dapporto, L., Vovlas, A., Chelazzi, G., Bonelli, S., Balletto, E. & Ciofi, C. (2013) 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*, 8(6), p.e65746.

Zug, R. & Hammerstein, P. (2012) Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PloS one*, 7, e38544.

Figure legends

Fig. 1. (A) Comparison of male genitalia in specimens identified by k-means analyses as *I. podalirius* (above) and *I. feisthamelii* (below), and elements measured: (a) ostium length, (b) phallus length and (c) ostium width. (B) Comparison of female genitalia of *I. podalirius* (above) and *I. feisthamelii* (below), and elements measured: (d) ductus bursae length and (e) ductus bursae width. (C) Scatter plot of genitalia measurements in males (left) and females (right) with k-means identification and a supplementary division by generation and region: yellow = European *I. feisthamelii*, red = Maghrebian *I. feisthamelii*, blue = *I. podalirius*; circles = first generation (1G), crosses = second generation (2G).

Fig. 2. A) Comparison of the UV reflectance in male specimens identified as *I. podalirius* and *I. feisthamelii*. Inserts show the differences in pale background (above) and blue lunules (below) on the upperside of the hindwing; B) Scatter plot of corrected UV reflectance average in males (left) and females (right): yellow = European *I. feisthamelii*, red = Maghrebian *I. feisthamelii,* blue = *I. podalirius*, circles = first generation (1G), crosses = second generation (2G). Identifications were obtained by k-means analyses based on UV reflectance data for males and on genital morphology for females. Only male wing UV reflectance does not overlap between *I. podalirius* and *I. feisthamelii*.

Fig. 3. Spectra of the blue lunules in males (A) and females (B) and the pale background of males (C) and females (D) measured for seven *I. podalirius* specimens (three females from Turkey, Greece and Switzerland; and four males from Italy, Bulgaria and Kazakhstan) and seven *I. feisthamelii* specimens (three females from Spain, Tunisia and Portugal; and four males from Spain and Morocco). Note the difference in the spectra produced by the pale background of males of the two species in C.

Fig. 4. Schematic comparison of the genetic relationships displayed by mitochondrial (*COI*) and nuclear (*ITS2*) DNA. European specimens attributable to *I. feisthamelii* based on morphology and nuclear DNA clustered with *I. podalirius* when the *COI* gene was used. Posterior probabilities are shown above the recovered nodes. Detailed Bayesian trees are provided as supplementary material.

Fig. 5. Results of k-means clustering on *COI* genetic distances, when forcing k = 2, 3 and 4 clusters (A, B) and when forcing k = 2 and 3 clusters for the *ITS2* genetic distances (C, D) (specimens from east Kazakhstan, not shown on the map, belong to the red cluster in *COI* and to the red cluster in *ITS2*). Magnified maps (B, D) illustrate with better resolution the contact zone between the two taxa.

Fig. 6. Hypothesis for the evolutionary history of the two *Iphiclides* species (chronologically from A to C, events within each subfigure numerated, see text for explanation). C illustrates results of the *Wolbachia* infection assessment for *I. podalirius* (squares) and *I. feisthamelii* (triangles), with the various strains indicated (yellow = wsp1-coxA unknown, because it could not be amplified).

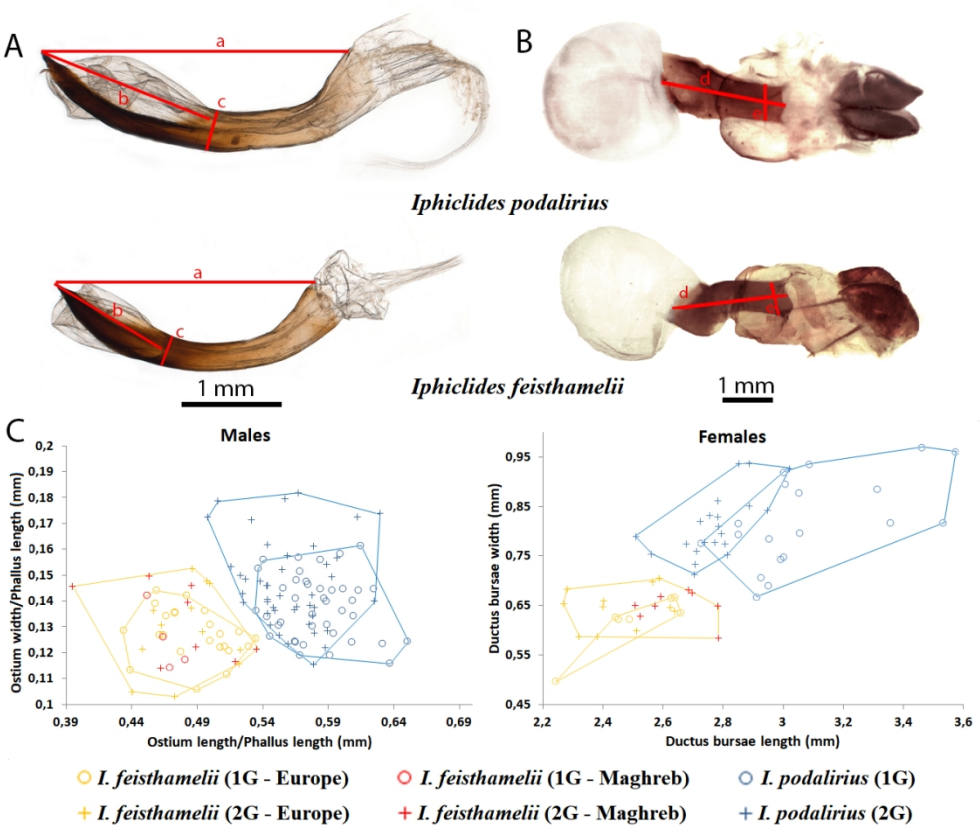
A
a
b
c
B
d
e
Iphiclides podalirius
a
b
c
d
e
1 mm
Iphiclides feisthamelii
1 mm
C
Males
Ostium width/Phallus length (mm)
Ostium length/Phallus length (mm)
Females
Ductus bursae width (mm)
Ductus bursae length (mm)
I. feisthamelii (1G - Europe)
I. feisthamelii (1G - Maghreb)
I. podalirius (1G)
I. feisthamelii (2G - Europe)
I. feisthamelii (2G - Maghreb)
I. podalirius (2G)

Figure 1

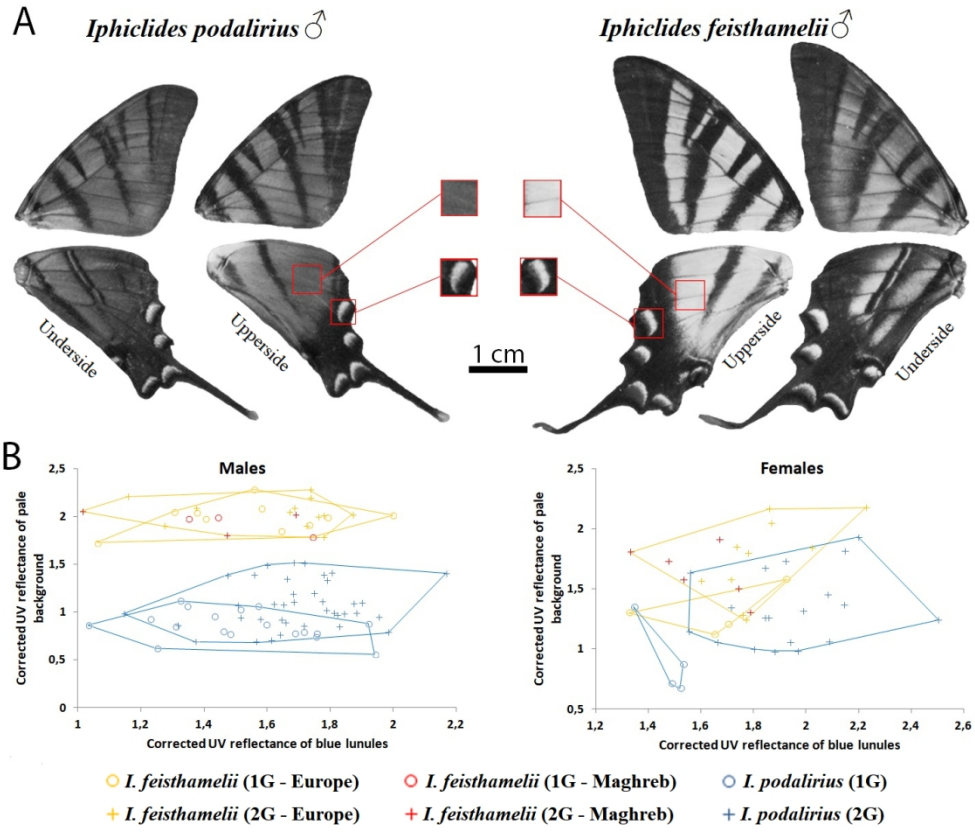
A
Iphiclides podalirius ♂
Iphiclides feisthamelii ♂
Underside
Upperside
1 cm
Upperside
Underside
B
Males
Females
Corrected UV reflectance of pale background
Corrected UV reflectance of blue lunules
I. feisthamelii (1G - Europe)
I. feisthamelii (1G - Maghreb)
I. podalirius (1G)
I. feisthamelii (2G - Europe)
I. feisthamelii (2G - Maghreb)
I. podalirius (2G)

Figure 2

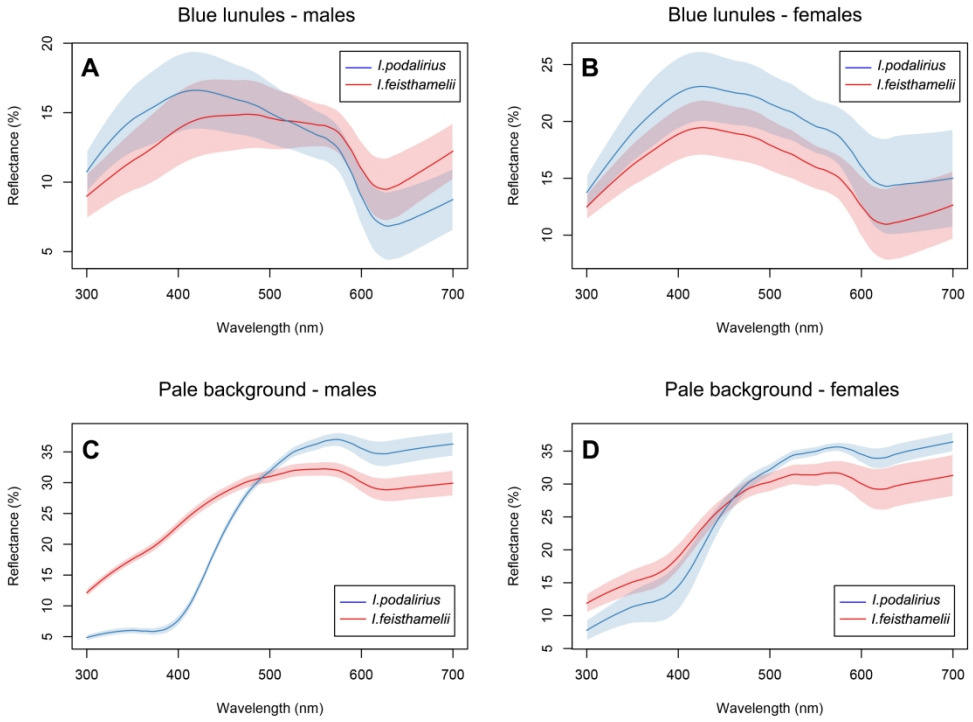

Figure 3

250x190mm (300 x 300 DPI)

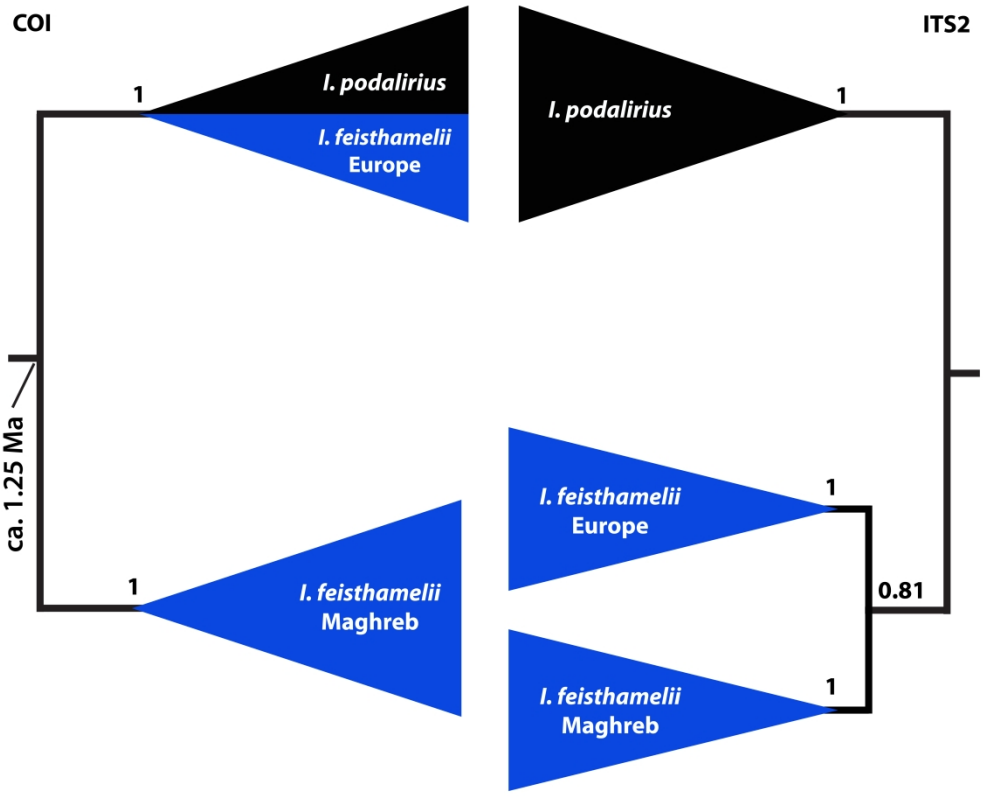

Figure 4

1652x1359mm (72 x 72 DPI)

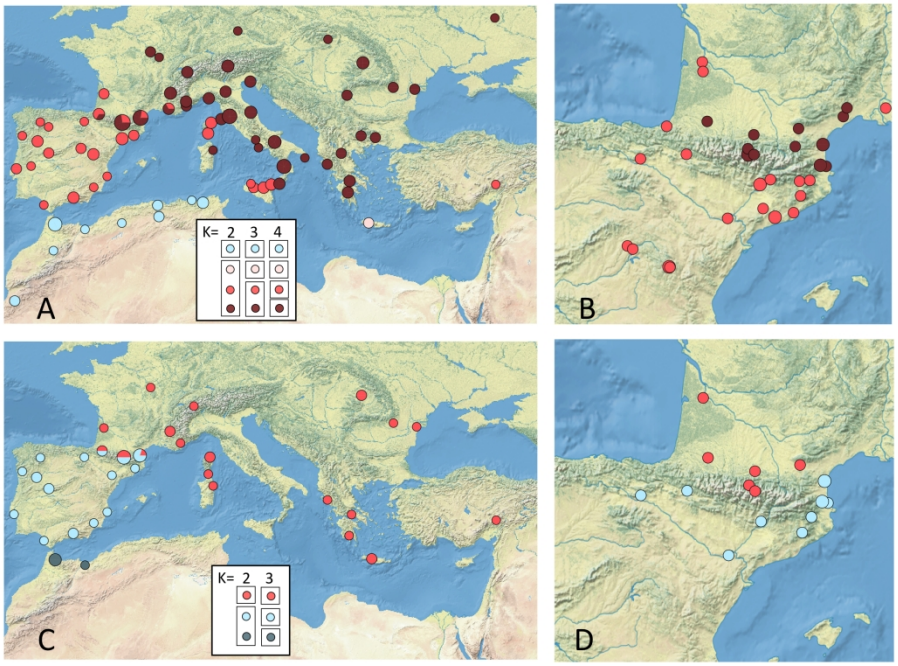

Figure 5

1057x793mm (72 x 72 DPI)

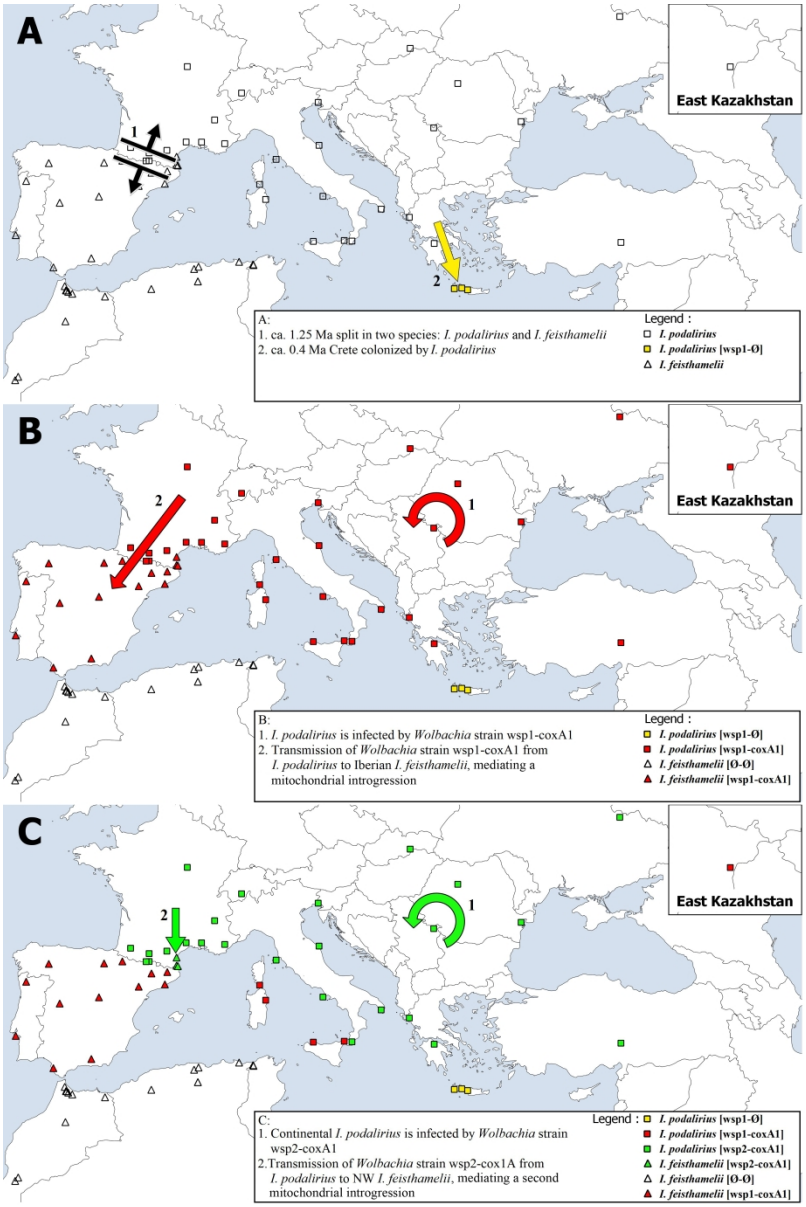

Figure 6

862x1290mm (96 x 96 DPI)