Rise and fall of island butterfly diversity. Understanding population diversification and extinction in a highly diverse Archipelago.

Dapporto L.1,2*, Cini A.2*, Menchetti M.1,2*, Vodă R3*, Bonelli S1, Casacci L.P.3, Dincă V.1,4, Scalercio S.5, Hinojosa J.C.1, Biermann H.6, Forbicioni L.7, Mazzantini U.8, Venturi L.9, Zanichelli F.10, Balletto E.3, Shreeve T.G.11, Dennis R.L.H.11,12, Vila R.1

1 Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37, 08003, Barcelona, Spain
2 Dipartimento di Biologia, Università degli Studi di Firenze, Via Madonna del Piano 6, 50109, Sesto Fiorentino, Florence, Italy
3 Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università degli Studi di Torino, Via Accademia Albertina 13, 10123, Turin, Italy
4 Centre for Biodiversity Genomics, Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, N1G2W1, Canada
5 Consiglio per la ricerca in agricoltura e l’analisi dell’economia agraria, Unità di Ricerca per la Selvicoltura in Ambiente Mediterraneo, Contrada Li Rocchi, 87036 Rende, Italy
6 Markusstrasse, 17D-33014 Bad Driburg, Germany
7 Via Roma 6, 57037, Portoferraio, Italy
8 Legambiente Arcipelago Toscano, 57037, Portoferraio, Italy
9 Parco Regionale della Maremma, via del Bersagliere 7/9, 58100, Alberese (GR), Italy

10 Parco Nazionale Arcipelago Toscano, località Enfola, 57037, Portoferraio, Italy

11 Centre for Ecology, Environment and Conservation, Department of Biological and Medical Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK

12 Institute for Environment, Sustainability and Regeneration, Staffordshire University, College Road, Stoke-on-Trent ST4 2DE, UK

Dapporto L.: leondap@gmail.com; Cini A.: ciniales@gmail.com; Menchetti M.: mattiamen@gmail.com; Vodă R.: raluvoda@gmail.com; Bonelli S.: simona.bonelli@unito.it; Casacci L.P.: luca.casacci@unito.it; Dincă V.: vlad.e.dinca@gmail.com; Scalercio S.: stefano.scalercio@crea.gov.it; Hinojosa J.C.: sirsphingidae2@gmail.com; Biermann H.: biermannbd@gmx.de; Forbicioni L.: jjuniper8@gmail.com; Mazzantini U.: legambientearcipelago@gmail.com; Venturi L.: legambientearcipelago@gmail.com; Zanichelli F.: sagittarius.sagitta@libero.it; Balletto E.: emilio.balletto@unito.it; Shreeve T.G.: tgshreeve@brookes.ac.uk; Dennis R.L.H.: rlhdennis@aol.com; Vila R.: roger.vila@csic.es

Keywords
Conservation biogeography, DNA barcoding, Lepidoptera, Phylogeography, Species traits, Tuscan Archipelago

* These authors contributed equally to this work
Short running title: An integrated dataset to study island biodiversity

Corresponding author: Raluca Vodă, raluvoda@gmail.com

The number of words in the Abstract: 292

The number of words in main body of the paper, from the Introduction through the Biosketch: 4999

The number of references: 66
ABSTRACT

Aim

We described fine-scale diversity patterns of the entire butterfly fauna occurring on the Tuscan Archipelago. By assessing the traits associated with population diversification, haplotype uniqueness and extinction, we aimed to identify the factors determining the origin and the maintenance of genetic diversity and population vulnerability to environmental changes.

Location

Tuscan Archipelago, Sardinia, Tuscany (Italy) and Corsica (France).

Methods

We built a mtDNA dataset (1303 COI sequences) for the 52 butterfly species reported in the Archipelago also including specimens from neighbouring areas and compiled data on 12 species traits and on the apparent extinction of butterfly species from the main islands. We calculated different indices of genetic differentiation, and using stepwise phylogenetic regressions we evaluated the relationships between these indices and species traits. Finally, we inferred which traits are associated with apparent extinction of species on individual islands using phylogenetic regression.

Results

The overall spatial pattern of genetic diversity corresponded with the proximity of the areas, but strong contrasts were also identified between geographically close areas. Together with the island endemics, several common and widespread species revealed a high genetic
diversification among islands and mainland. Phylogenetic regressions showed that smaller-sized, more specialized species, with a preference for drier regions, displayed greater genetic structure and/or haplotype uniqueness. Capraia has particularly suffered loss of diversity, which significantly affected species with shorter flight periods.

Main conclusions

Tuscan island butterflies are characterized by strong genetic contrasts and species differ in their contribution to the overall genetic diversity. By ranking the species for their contribution to genetic diversity and by identifying the traits linked to the emergence and maintenance of diversity, we provided a valuable tool to prioritize the populations as targets for monitoring and conservation actions. The large dataset we provided represents a resource to test biogeographical hypotheses.
INTRODUCTION

The worldwide biodiversity crisis calls for the identification, prioritization and protection of biodiversity hotspots, and understanding how biodiversity is generated and lost determines the success of this endeavour (Dirzo et al., 2014; Venter et al., 2014). The wealth of information now available in the “big data era” greatly facilitates these efforts, especially the ability to work at an unprecedented resolution (Hampton et al., 2013). For conservation biogeography this includes the increased availability of molecular, occurrence and trait data for various organisms (Ladle & Whittaker, 2011; Fernández-Palacios et al., 2015).

Islands host a disproportionate fraction of global biodiversity, often characterized by distinctive faunas (Whittaker & Fernández-Palacios, 2007), and therefore represent a model system to better understand general patterns in biogeography. Insular populations are typically exposed to high risks of extinction since restricted habitats can easily become unsuitable because of stochastic events or human disturbances acting at both local (habitat fragmentation, alteration, invasion by alien species) and global scales (e.g. climate change) (Fordman et al., 2010). Understanding the factors that drive the emergence, maintenance and loss of island biodiversity is crucial for planning and implementing evidence-based conservation prioritization and protection measures (Ladle & Whittaker, 2011).

Remote oceanic islands experience infrequent colonization events followed by genetic drift and adaptive radiation (Rosindell & Phillimore, 2011). In contrast, biotas on less isolated islands are mostly assembled following frequent events of colonization and extinction, which produce nested communities of the source species with few endemics (Whittaker & Fernández-Palacios, 2007). However, a pace of extinction and colonization on islands (turnover) slower than the dynamics occurring in the surrounding regions can generate intraspecific diversification and relictuality, which creates community distinctiveness among
islands and from neighbouring mainlands (Masini et al., 2008; Dapporto et al., 2012). Such diversity is usually represented by cryptic species or genetic lineages that can be detected only after in-depth molecular and/or morphologic analyses (Hernández-Roldán et al., 2016; Vodă et al., 2015a, 2016).

The occurrence of endemic and relict populations can be the result of deterministic processes, largely affected by species characteristics. According to a widely accepted paradigm of island conservation biogeography, genetic diversification and extinction probability of island populations are inversely related to their degree of mobility and generalism (Burney & Brumfield, 2009; Dennis et al., 2011; Salisbury et al., 2012; Dawson et al., 2014 but see also Kobayashi & Sota, 2016 for different patterns). This hypothesis has profound implications for conservation ecology because populations with unique genetic fingerprints are clearly irreplaceable but potentially suffer from high extinction probability (Ricklefs, 2009). Nevertheless, this has been rarely tested (Burney & Brumfield, 2009; Vodă et al., 2016), probably because of the absence of integrative datasets combining molecular and occurrence data with species traits for entire species-rich taxonomic groups and from a large geographic area.

Here we tested if species having low dispersal capacities and restricted ecological requirements tend to colonize islands at a slower rate, experience reduced gene flow and thus have i) higher diversification rates and ii) higher likelihood of local extinction compared with more mobile and generalist species. We designed an integrated framework (Fig. 1) based on a dataset comprising: i) a revised distribution list of the 52 butterfly species reported for the Tuscan Archipelago and the apparent disappearance of several species in the last 115 years, ii) 1303 cytochrome c oxidase subunit I (COI) sequences (658 bp) for all the species, including populations from the surrounding areas (Sardinia, Corsica and Tuscany), and iii) 12 species
traits related to mobility, phenology, trophic generalism and climatic preferences. The analysis of this dataset allowed us to: i) describe the patterns of population diversification (genetic diversity and haplotype uniqueness) among islands and neighbouring areas, ii) identify the functional traits that are correlated with the emergence of diversification and species disappearance on islands. Finally, iii) we ranked the species according to their contribution to the overall island diversity. Our integrated framework (Fig. 1) allowed us to provide evidence-based guidelines for butterfly conservation in a well-known Mediterranean diversity hotspot (Dennis et al., 2008).

METHODS

Study area and occurrence data

We carried out annual surveys of the butterflies of the Tuscan islands and the neighbouring areas of Sardinia, Corsica and Tuscany between 2000 and 2015 (Fig. 1a). We sampled the main biotopes from early spring to late autumn, with different tools (mostly Malaise traps and insect nets). We compared occurrence data from our surveys with published records dating back to 1900 (Appendix S1). Sampling was not done constantly throughout the 15 years but our extensive collections represent the most intensive sampling effort done on the butterflies of these islands, therefore it is reasonable to assume that a species reported in the past but not during the last decade has either become extinct or has significantly declined. We restricted our analysis of species disappearance to the three largest islands: Elba, Capraia and Giglio, for which sufficient historical data were available.
Using standard sequencing procedures (deWaard et al., 2008), we obtained 1303 COI sequences for specimens belonging to the 52 butterfly species reported for the Tuscan islands between 1900 and 2015, including specimens from Sardinia, Corsica, Argentario and four mainland areas (Fig. 1a). We also used 2940 sequences from other parts of Europe and Asia. Information about specimens and sample size for each population is available in the "dataset.csv" file (Appendix S2). All sequences were aligned in Geneious 6.0.6 (www.geneious.com).

The butterfly species currently recognized by taxonomists, and for which trait data are available, show varying levels of intraspecific genetic divergence, including potential cases of cryptic taxa (Dincă et al., 2015). We considered as separate units most species recognized by the widely accepted checklist of the Fauna Europaea Project (Karsholt & Van Nieukerken, 2013; www.faunaeur.org) and also considered as single units taxa displaying a minimum COI p-distance lower than 3%. In butterflies, this threshold separates more than 90% of the recognized species (Hebert et al., 2003) and a recent study on Sicilian islands confirmed this observation (Vodă et al., 2016). We also repeated all the analyses without setting any distance threshold, and considered as single units only the species recognized by Fauna Europaea.

We calculated the genetic uncorrected p-distances among all sequenced specimens for each species by using the function “dist.dna” of the “ape” R package. We preferred p-distance to tree-based genetic distances because, at the intraspecific level, coalescence has not taken place and distances calculated on branch lengths of bifurcating phylogenetic trees do not properly reflect the reticulated evolutionary processes (e.g. Posada et al., 2001). Moreover, recent reviews indicate that p-distances are the best option in the analysis of COI data compared to other indices (Srivathsan and Meier, 2012). Based on p-distances, we obtained...
two measures for population differentiation: Dst and Gst (Nei, 1987). We also calculated the pairwise Gst among all pairs of populations for each species (see Appendix S1 for a description of the indices).

Based on the Gst pairwise matrices we produced for each species the mean Gst matrix, representing the degree of differentiation among areas based on all species. A Principal Coordinates Analysis (PCoA) was applied to this matrix to obtain the overall diversity pattern among areas. Subsequently, we aligned this configuration with the geographic location of the areas by using the "procrustes" analysis from the "vegan" R package and tested the correlation between the PCoA configurations of Gst and the spatial location by using the vegan function "protest". To visualize the pattern of similarity among islands in the geographic space, we projected the PCoA configuration in RGB space using the R package “recluster” (Dapporto et al., 2014). The colour resemblance of the resulting configuration is directly proportional to the genetic similarity among the communities.

Four species (*Leptidea sinapis, Aglais urticae, Nymphalis polychloros, Argynnis paphia*) recorded for Tuscan islands in the last century but not during our surveys have also been analysed by computing the Dst and Gst between Tuscany, Sardinia and Corsica to estimate the fraction of diversity that has presumably disappeared.

Island haplotype uniqueness for each species was calculated as:

\[ \text{Uni} = \frac{\frac{1}{n} \min(D_{h,m})}{\text{occ}} \]

where \( h \) are the haplotypes found on the Tuscan islands and not recorded on mainland areas of Europe and Asia, \( \min(D_{h,m}) \) is the p-distance between the \( h^{th} \) haplotype and the genetically closest haplotype from mainland, and \( \text{occ} \) is the number of islands (Tuscan islands, Sardinia and Corsica) where the \( h^{th} \) haplotype has been found. Uniqueness for a given species is higher
when: 1) there are many insular endemic haplotypes, and/or 2) they show high divergence
with respect to the closest mainland haplotype and/or 3) they are found in fewer islands.
Uniqueness values (Uni) were calculated for each species occurring on any island except for
the island endemics that do not have mainland populations (*Hipparchia aristaeus* and *H.
neomiris*).

To assess the importance of each species for the diversity of the Tuscan Archipelago we
ranked species according to Gst, Dst and Uni and subsequently summed the ranks. The
patterns of genetic variation were also analysed by inferring maximum parsimony haplotype
networks using the program TCS 1.21, with a 95% connection limit for all species except *H.
neomiris* (94% connection limit) and *Coenonympha corinna* for which we used a fixed
connection limit of 23 steps (Clement et al., 2000).

Species traits and phylogenetic regressions

Review studies suggest that only morphological, physiological or phenological features that
can be measured on individual organisms and without reference to the environment or any
other level of organization should be considered as functional traits (Violle et al., 2007;
Moretti et al., 2016). Moretti et al. (2016) identified a series of 29 functional traits to cover
the primary functions of invertebrates, divided into five major groups: morphology, feeding,
life history, physiology and behaviour. Measuring the 29 traits on individuals for all the
studied species would have entailed a major long-term effort that was beyond the scope of
this study. Thus, based on literature data and personal observations, we assessed 12 species
traits representing four of the five groups (excluding behaviour) identified by Moretti et al.
(2016): but encompassing morphology, feeding, life history and physiology. For each trait we
formulated functional hypotheses (Table 1): a) Trophic generalism (feeding), was identified
as i) the number of host plant genera reported in the literature; b) Mobility was assessed with a morphological traits represented by ii) wingspan and obtained as the average between minimum and maximum size reported in the literature; c) Phenology (life history traits) was identified as iii) the length of the flight period, iv) the first month when adults emerge, v) the last month when adults fly, and vi) voltinism. Finally, d) climatic preference and tolerance (physiology) were assessed by proxy variables for eco-physiological responses to environmental conditions. These variables have been calculated by Schweiger et al. (2014) by modelling species distribution in Europe based on occurrence data, and then by averaging temperature and precipitation among the spatial cells where each species is predicted to occur. Although these indices cannot be considered as strict functional traits since they are obtained from the geographic distribution of the species (Violle et al., 2007), they are widely recognized as proxies for the traits responsible for eco-physiological responses to climate (e.g. Devictor et al., 2012). The variables we included are: vii) mean annual temperature viii) and precipitation, ix) standard deviations of the temperature mean and x) and precipitation, xi) upper 95% confidence limit of temperature mean, and xii) lower 95% confidence limit of precipitation mean.

Butterfly traits are usually highly inter-correlated but they can be conveniently reduced to factors by using ordination methods (Carnicer et al., 2013; Dapporto & Dennis, 2013). For morphology, life history and physiology traits we applied a Principal Component Analysis (PCA), using the R function “rda” and the components with eigenvalues higher than one have been used as variables for successive analyses. Some literature sources did not report the wingspan for all the studied species and we imputed the missing values by using the “mice” function of the “mice” R package (see Appendix S2). The algorithm imputes an incomplete variable by generating plausible values based on other variables in the data by Multivariate
Imputations by Chained Equations (MICE) (Van Buuren & Groothuis-Oudshoorn, 2011).

The existence of a phylogenetic signal for the variables of each trait following the PCA ordination and for Dst, Gst and Uni was tested with Pagel’s lambda index by applying the “phylosig” R function of the “phytools” package. Gst, Dst and uniqueness have been mapped onto the phylogenetic tree by using the “contMap” function of the “phytools” package. Character mapping is accomplished by estimating states at internal nodes using maximum likelihood and then by interpolating the states along each edge (Ravell, 2013). The relationships between the variables and Dst, Gst and Uni have been assessed using phylogenetic stepwise regressions. We also employed Pagel's lambda as a model for the phylogenetic covariance of residuals and applied a two-way selection of variables based on the Akaike Information Criterion (AIC) as implemented in the function “phylostep” of the package “phylolm”. From the phylogenetic regressions we removed the species not recorded for the Tuscan Archipelago during the study period since no DNA sequences were available. We square-root transformed Dst and Uni to improve their normality and standardized the values of the traits with zeta-scores to provide a balanced contribution to the phylogenetic regression.

The importance of traits in explaining possible extinctions of butterflies on two islands (Elba and Capraia) has been assessed with a logistic phylogenetic Generalized Linear Model using the function “phyloglm” of the package “phylolm”, in which species found during our surveys and species that have not been confirmed, represented the binary response variable, and the trait variables the predictors. Logistic GLM was not performed for Giglio as only three species disappeared on this island. “Phylolm” function was used to assess if species disappeared in at least one island showed higher values of Gst, Dst and Uni.
As a reference phylogeny, we used the Maximum Likelihood (ML) phylogenetic tree based on COI sequences for all the western Mediterranean butterflies, freely available in the package “recluster” (Dapporto et al., 2013). The tree was inferred with topological constraints at family and subfamily levels following the butterfly phylogeny (see Appendix S1 for details). ML analyses were performed using RAxML BlackBox (Stamatakis et al. 2008). A GTR+Gamma+I model was selected and node supports were assessed through 100 rapid bootstrap replicates. Effect size for models has been evaluated by plots of observed vs fitted values associated with Spearman rho correlation.

RESULTS

Based on the taxonomy proposed by Fauna Europaea and by applying a 3% threshold of COI divergence, we identified 52 units among the taxa reported in literature for the Tuscan Islands (hereafter ‘species’, Table 2). During our surveys on these islands we recorded a total of 46 species. The comparison between observations during the last decade and literature data from 1900 to 2000 for Elba, Giglio and Capraia, identifies those species (see Table 2) that probably became extinct or strongly declined on these islands (respectively six, three and seven on Elba, Giglio and Capraia).

Dst was correlated with both Gst (Spearman rank test: rho 0.836, P<0.001) and Uni (rho 0.460, P<0.001), while Gst was not correlated to Uni (rho 0.192, P=0.213). Gst values showed an almost bimodal distribution (14 species with Gst<0.25 and 10 species with Gst>0.75, Table 2). Twenty-four species had haplotypes not recorded on the mainland and most of these species did not belong to endemic taxa (Table 2). Gst and Uni did not have a significant phylogenetic signal (lambda 0.218; P=0.314 and lambda <0.001; P=1.000, Fig. 2), while Dst
had a significant effect (lambda 0.334; P=0.034; Fig. 2). Ordering species by the sum of ranks of the three indices (Dst, Gst and Uni) showed that, together with endemics, several common and widespread species provided a large contribution to diversity. The first quartile of the top ranking species comprised four species/groups with endemic elements in the Tuscan islands (C. corinna, Lasiommata megera/paramegaera, H. neomiris, Aglais urticae/ichnusa) and ten widespread species, most of them ubiquitous in Europe and not included in any protection list (Table 1).

The wingspan measures reported in the four literature sources used were highly correlated (Pearson R>0.9 for all pairs) and the PCA identified only one component with an eigenvalue higher than one (Table 1, Figure S46). For life history and physiologic traits two components were considered (Table 1, Figure S47). The first phenological component was mainly linked to the length of the flight period (voltinism, number of months when adults occur), while the second was mainly linked to seasonality (first and last month of emergence). The first component for physiologic traits ordered species from those experiencing high temperatures and low precipitation to those living in colder and wetter areas, while the second component ordered species mostly according to their precipitation tolerance (Table 1, Figure S48). The six resulting variables showed a lower correlation among each other with Pearson correlation values always lower than 0.400 (Table S1). Among the six resulting variables, the number of host plants, phenology PC1, physiology PC1 and PC2 did not show a phylogenetic signal (lambda 0.282; P=0.228; lambda<0.001; P=1.000; lambda 0.062; P=0.711; lambda<0.001; P=1.000, respectively), while wingspan PC1 and phenology PC2 showed a significant effect (lambda 1.187; P<0.001; lambda 1.083; P<0.001, respectively).
The overall spatial pattern of genetic variation based on Gst corresponded to the proximity of the areas but with a rather low level of correlation (protest correlation 0.592, P=0.015, Fig. 3a,b).

The AIC procedure for the stepwise phylogenetic regression for Gst selected a model with four variables, but only mobility and trophic generalism had a significant effect (Table 3), meaning that smaller-sized and more generalist species had a higher Gst. For Dst, four variables entered the model - three were significant and showed that smaller-sized, more generalist species and those experiencing less annual precipitation had a higher variation (Table 3). Two variables entered the Uni model showing that species living in drier areas significantly had higher haplotype uniqueness (Table 3), while trophic generalism entered the model but without significant effect. Plots for observed vs fitted values of the three models (Fig. S53) showed large residuals indicating that Gst, Dst and Uni are only weakly explained by the measured traits (Spearman rho: Gst 0.537, Dst 0.430. Uni 0.368).

Species that had disappeared from at least one island showed significantly higher values of Dst, while no differences in Gst and Uni were found (Table 3).

According to the logistic phylogenetic GLM, species that disappeared from Capraia had lower values in PC1 for phenology, corresponding to shorter flight periods (Table 4). A plot for observed vs fitted values (Fig. S54) revealed a good fit for this analysis (Spearman rho 0.798). For Elba we found no significant effect explaining the disappearance of the six species (Table 4).

The analyses in which we used only the taxonomy from Fauna Europaea returned very similar results to the ones in which we used a 3% threshold for species identification (see Table S2-S5 and Figure S55 in Appendix S1).
DISCUSSION

The integration of an updated taxonomic list, species occurrence spanning across 115 years (1900-2015), mitochondrial DNA sequences and species traits, allowed us to characterize the butterfly diversity in the Tuscan Archipelago at an unprecedented resolution and to infer which species traits explain the rise and decline of butterfly diversity in this archipelago.

An area of biogeographic contrasts

The Tuscan Archipelago is an insular hotspot for butterfly diversity and stands out among European islands for hosting far more endemics than would be expected based on their geography (e.g. area and isolation) (Dennis et al., 2008), resulting in the highest priority for butterfly conservation among circum-Italian islands (Dapporto & Dennis, 2008). Its unexpected level of endemicity and richness is a consequence of its intermediate location between Tuscany and the Sardo-Corsican region, generating a double filtering effect (Dapporto & Cini, 2007; Fattorini, 2009).

We found that a main determinant for the genetic make-up of populations is island location (see the protest analysis), the same as for the community composition at species level (Dapporto & Cini, 2007). Nevertheless, a model based only on a double filtering effect is not sufficient to explain the observed degree of genetic diversity. In fact, we identified strong contrasts between geographically close areas, such as between Montecristo (very similar to the Sardo-Corsican region) and Pianosa (more similar to Elba and the Italian Peninsula), which are separated by 30 km; between Capraia and Elba (same pattern as before, distance 33
km) and between Giglio and Argentario, separated by only 14 km (Fig. 3). Striking divergence among populations from nearby areas is unexpected in butterflies characterized by a high mobility resulting in a high capacity to track suitable environments (Wilson et al., 2010; Waters, 2011; Devictor et al., 2012). However, in the western Mediterranean (Tuscan islands included), chequered distributions of sister species and genetic lineages are a common phenomenon, probably due to the combination of several historical determinants and contemporary ecological forces (e.g. the connection between land masses during the Last Glacial Maximum, density-dependent phenomena, differences in climatic and environmental features, Vodâ et al., 2015a,b, 2016). Accordingly, more than 30% of the examined species, comprising both endemic and widespread taxa had a Gst value higher than 0.5.

Endemic taxa also have unexpected patterns of intraspecific genetic diversity among islands. *Coenonympha corinna/elbana* is highly divergent among the three clades they form in: i) Tuscany, Elba and Giannutri; ii) Capraia and Corsica; and iii) Sardinia. This pattern only partially supports the current taxonomic separation into two species or subspecies (*C. corinna*: Sardinia, Corsica and Capraia; *C. elbana*: Elba, Tuscany and Giannutri) (Fig. 3c). *Hipparchia neomiris*, an endemic species from Sardinia, Corsica and Elba, also displayed notable intraspecific divergence. COI sequences from Elba and Corsica are differentiated by at least 2% compared to conspecific individuals from Sardinia. By contrast, no genetic diversification was detected in the *Plebejus idas* group, since the endemic *P. bellieri* from Sardinia and Corsica (recognized as a good species in Fauna Europea) shared COI barcodes with the population from Elba, treated as a species in some works, and with the mainland populations, elsewhere reported as as *P. idas* or *P. abetonicus* (Balletto et al. 2015).

Phylogenetic regressions suggest that the species adapted to a dry climate, the small-sized ones and the host plant specialists showed a higher degree of island uniqueness and of genetic
diversification among islands and populations. According to our hypotheses, these species
traits can facilitate the emergence and maintenance of these genetic contrasts since typical
Mediterranean species thrive on islands, while species with reduced dispersal and poor
colonization capabilities probably experience reduced gene flow.

Several species previously recorded on three Tuscan islands were not observed in the last 10
years, suggesting that they became extinct or declined considerably (Table 2). On Giglio only
three species have not been recorded during the last decade, but they include C. corinna, the
most emblematic species for the archipelago, as well as A. agestis and Polyommatus icarus,
both showing high levels of population diversification (Table 2).

Elba has apparently lost six species. In this case as well, some of the taxa have diversified
populations in the study area (Aglais urticae/ichnusa, Leptidea sinapis and Nymphalis
polychloros showed a Gst higher than 0.4). For this island we found no species traits
correlated with disappearance and the fraction of likely extinct species (11.5%) was lower
than on Capraia 30% of the species reported in the past have disappeared.

Two species disappeared from Capraia were insular endemics (H. neomiris and H. aristaeus)
and others showed divergent populations in the study area (P. cecilia, L. phlaeas, M. jurtina).
Disappeared species represented a fraction of the fauna with a short flight period. As the
length of the flight period is correlated to inter-island dispersal in the Tuscan Archipelago
(Dapporto et al., 2012), species with a short flight period have a lower probability of re-
colonizing from surrounding areas. Interestingly, there are no typical springtime species on
Capraia and taxa with short flight periods are typically monovoltine, with adults emerging at
the beginning of summer, aestivating during the hottest weeks and laying eggs in
September/October. Aestivation in the Mediterranean region is known for M. jurtina (Scali,
1971) and *Hipparchia semele* (García-Barros, 1988), a species closely related to *H. aristaeus*.

Both of them have disappeared from Capraia, together with two other Satyrinae (*P. cecilia* and *H. neomiris*), which tend to be frequent in woodlands and scrub/maquis during the hottest and driest months. Because there are barely any remaining woods on Capraia, aestivation may represent an important stress period with current temperature increases (Shreeve et al., 2009; Cerrato et al., 2016). Compared to Elba and Giglio, Capraia is more isolated and this could have hampered a rescue effect for many species. Moreover, Elba has the highest mountain peak among the small Italian islands (Monte Capanne, 1019m) and woodlands are common on both Giglio and Elba. The higher environmental heterogeneity of Elba and Giglio could have provided a wider range of suitable areas for many species under environmental stress and climatic oscillations.

Extinction events over long periods are expected on islands based on the equilibrium theory (MacArthur & Wilson, 1967), but they should be paralleled by colonization events, which was not the case for the Tuscan islands. On Elba only *Anthocharis cardamines* has been discovered after 1950; on Capraia only *P. aegeria* and *A. agestis* have been found after intensive field research between 1970-1980, when almost all the seven extinct species were present; only *C. rubi* has been recently discovered on Giglio.

**Guidelines for the conservation of the Tuscan Archipelago butterflies**

Ranking species according to their contribution to genetic diversity (Table 2) shows that butterfly diversity in the Tuscan Archipelago is encompassed both by insular endemic taxa as well as by widespread species (*A. agestis, C. pamphilus, C. alceae, Melitaea nevadensis, M. jurtina, P. aegeria* and *Zerynthia cassandra*). While these latter species are currently treated as being of ‘Least Concern’ in the European and Italian Red Lists (Van Swaay et al., 2010;
Balletto et al., 2015), some of them disappeared from some islands resulting in a loss of faunistic and genetic diversity. Some of the populations that have apparently disappeared in the last 10 years may be still rediscovered following dedicated field research, as occurred for \textit{Zerynthia cassandra}, discovered on Elba in 1932 but apparently disappeared before our intensive collection effort (Appendix S1). Species that disappeared from at least one island showed a higher overall population diversification (Dst) compared to persisting species. In fact, our analyses revealed that species accumulated genetic diversification because of their reduced migration and colonization capabilities; but these characteristics can also produce higher extinction risk due to reduced gene flow and rescue effect. We showed that one of the main peculiarities of Tuscan islands is the occurrence of strong genetic contrasts among nearby areas. If an insular relict or a genetically endemic population goes extinct, it is likely that it would be replaced by conspecific propagules from the nearest source, thus lowering the ancestral genetic diversification. There is also evidence for Mediterranean butterflies that the presence of endemic and relict island populations limits the colonization by mainland populations probably due to density-dependent phenomena (Dapporto et al. 2012, Vodă et al. 2015b). The establishment of the Tuscan Archipelago National Park in 1996 represented a fundamental step for the broad-scale protection of island communities. Nevertheless, specific conservation strategies tailored on particularly valuable species are still lacking because evidence-based information are missing. Our integrated approach, by prioritizing species according to their contribution to genetic diversity and by identifying the impact of ecological drivers on the emergence and extinction of differentiated populations, allows conservation priorities to be established, with a necessity for regular monitoring schemes. Such schemes should evaluate population consistence and health, particularly focusing on population size and trends, genetic load and persistence of habitat suitability.
Our study also suggested that reduction of shady areas might be a driver of species loss in the
Archipelago. Historically, wooded and shrub areas have been considerably reduced, mostly
on Capraia and Montecristo. It has been already reported that deforestation can be a main
driver for butterfly extinction in Mediterranean islands. For example, the strong reduction of
shady areas has been linked with post-glacial reduction of butterflies on Malta with the recent
extinction of most relict elements (Vodă et al., 2016). We thus indicate as a specific
conservation action to adopt environmental management procedures aimed at preserving and
favouring environmental heterogeneity, thus increasing resources availability and suitable
habitats for a larger number of butterfly species (Dennis, 2010). The increase of temperature
predicted by recent climate change scenarios suggests that environmental heterogeneity will
play a pivotal role in buffering increasing thermal and drought stress.

In conclusion, this study shows that the integration of molecular and trait data with long-term
occurrence records allows the identification of the eco-evolutionary processes underlying the
high butterfly diversity in this Mediterranean diversity hotspot (Dennis et al., 2008). Evidence-based priorities for future conservation actions have been provided following the
theory of conservation biogeography (Ladle & Whittaker, 2011). Public institutions, such as
the NGO Legambiente and the Tuscan Archipelago National Park that have collaborated in
this project, have already used some of these results to raise awareness for island diversity
protection and to implement key conservation measures for butterflies.

ACKNOWLEDGEMENTS

This research is dedicated to the memory of Ornella Casnati, who was fundamental in
protecting the butterflies of the Tuscan islands. Funding came from the Spanish MINECO
(CGL2013-48277-P, CGL2016-76322-P and PRX15/00305 to RoV), Generalitat de Catalunya (2014-SGR-1532), Marie Skłodowska-Curie Train to Move (T2M) to RaV (grant 609402-2020), and from the projects “Barcoding Italian Butterflies”, “Barcoding Butterflies of the Tuscan Archipelago National Park” and “Barcoding Butterflies of the Maremma Regional Park”. VD was supported by a Marie Sklodowska-Curie IOF grant (project 625997).

Biosketch

LD, RaV, AC, MM and RoV conceived the idea; LD, RaV, AC, MM, RoV, VD, LF, HB, LPC, SS, FZ, UM and LV collected the data and the specimens in the study area; LD, RaV, VD and RoV assessed the taxonomy of the study species; LD, MM, AC and SS compiled the trait data; LD, VD, JCH, RaV, MM, AC, RoV, FZ, UM and LV managed and obtained the COI sequences; LD, HB and EB gathered the occurrence data; LD, RaV, MM, JCH and AC performed the data analysis; all the authors discussed the results and participated in writing the paper.
References


**Data accessibility.** The complete dataset, together with R scripts to replicate the analyses published in the paper, is available in the Supporting Information. The COI sequences are also available in the following repositories (Genbank, BOLD and IO databases, accession codes for Genbank, XXXXX-XXXXX; BOLD, XXXXX).
Table 1 Species traits used in the study with the description of the type of trait (sensu Moretti et al. 2016), in bold and the relative functional hypothesis in italics; the trait(s) measured; a description of the trait(s), the literature sources and the weights obtained by each trait in the first two Principal Components (in the cases when the analysis has been carried out). PC1 and PC2 represent the weights of the traits in the first two components obtained after principal component analysis (PCA) of each type of trait. PCA on the single variable of host plants has not been carried out and for wingspan only the first component returned an eigenvalue higher than 1.

<table>
<thead>
<tr>
<th>Type of trait</th>
<th>Functional hypothesis</th>
<th>Trait measured and description</th>
<th>Sources</th>
<th>PC1</th>
<th>PC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeding</td>
<td>Species feeding on a larger number of plants have a wider niche, thus a higher potential to colonize islands (Dennis et al. 2012)</td>
<td>Number of host plant genera used by larvae as reported in two literature sources</td>
<td>Lafranchis (2007)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tolman &amp; Lewington (2008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphology</td>
<td>Larger species are characterized by higher mobility (Sekar, 2012), thus more probabilities to cross sea barriers (Dennis et al. 2012)</td>
<td>Wingspan: mean between minimum and maximum size reported in four main sources for European butterflies. Tshikolovets (2011) reported size for males and females</td>
<td>Higgins &amp; Riley (1970)</td>
<td>0.448</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lafranchis (2000)</td>
<td>0.448</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pamperis (2009)</td>
<td>0.446</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tshikolovets (2011) males</td>
<td>0.448</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tshikolovets (2011) females</td>
<td>0.448</td>
<td>-</td>
</tr>
<tr>
<td>Life history</td>
<td>Phenological attributes determine the length of the most mobile life stage of winged adults. The period of the year when it is expressed.</td>
<td>Month of adult first emergence. It ranges from the coldest, January (1), to the warmest, August (8). No butterfly species has a later first emergence in the</td>
<td>Authors' collection data</td>
<td>-0.293</td>
<td>-0.822</td>
</tr>
</tbody>
</table>
These characteristics affect the possibility to cross sea barriers (Dapporto et al., 2012; Dennis et al. 2012) and can interact with climatic changes in determine extinction probabilities.

| **Physiology** | study area | **Last month when adults fly**, ranging from January (1) to December (12) | Authors' collection data | 0.533 | -0.482 |
| **Length of the flight period**: number of months when the adults occur in the study area | Authors' collection data | 0.587 | -0.189 |
| **Voltinism**: Number of generations/year in the study area | Authors' collection data and Tolman & Lewington (2008) | 0.535 | 0.237 |

Mean climatic conditions of the areas inhabited by a species are considered as good proxies for their ecophysiological response to climate (Devictor et al., 2012). They can affect the probability for species persistence in the warm and dry Mediterranean climate of Tuscan islands.

| **Physiology** | **Mean temperature** occurring in the 50×50 km spatial cells where the species has been modeled to occur | Schweiger et al. (2014) | 0.334 | -0.387 |
| **Mean precipitation** in the same spatial cells as above | Schweiger et al. (2014) | -0.292 | -0.607 |
| **Maximum temperature tolerance**: Upper 95% confidence interval for temperature mean | Schweiger et al. (2014) | 0.334 | -0.387 |
| **Minimum precipitation tolerance**: Lower 95% confidence interval for precipitation mean | Schweiger et al. (2014) | -0.332 | -0.413 |
| **Overall temperature tolerance**: Standard deviation for temperature mean | Schweiger et al. (2014) | -0.332 | 0.313 |
| **Overall precipitation tolerance**: Standard deviation for precipitation mean | Schweiger et al. (2014) | -0.337 | -0.056 |
Table 2: Species ranked for Gst, Dst and Uni, with information on single island uniqueness and documented recent extinction events. Dis, species that disappeared from Elba (E), Giglio (G) or Capraia (C). The last seven columns represent the uniqueness values of each species on each island. "-" means that the species has not been reported on that particular island, "NA" means that the species has been reported but that the population was not included in this study because of lack of genetic data. Taxa endemic to the insular region (Sardinia, Corsica and Tuscan islands) are written in bold, while the six species that were not recorded on islands during our surveys in the last 10 years are highlighted in grey.

<table>
<thead>
<tr>
<th>Species</th>
<th>Gst</th>
<th>Dst</th>
<th>Uni</th>
<th>Dis</th>
<th>Gor</th>
<th>Cap</th>
<th>Elb</th>
<th>Pia</th>
<th>Mon</th>
<th>Gig</th>
<th>Gia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aglais ursica/ichnusa</td>
<td>0.937</td>
<td>0.449</td>
<td>NA</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anthocharis cardamines</td>
<td>0</td>
<td>0</td>
<td>0.051</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.051</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Argynnis pandora</td>
<td>0.171</td>
<td>0.15</td>
<td>0.076</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.076</td>
<td>-</td>
<td>-</td>
<td>0.114</td>
<td>-</td>
</tr>
<tr>
<td>Argynnis paphia</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Aricia agestis/cramera</td>
<td>0.738</td>
<td>0.984</td>
<td>0.094</td>
<td>G</td>
<td>-</td>
<td>0.152</td>
<td>0.095</td>
<td>0.046</td>
<td>-</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Callophrys rubi</td>
<td>0.846</td>
<td>0.279</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Carcharodus alceae</td>
<td>0.779</td>
<td>0.258</td>
<td>0.095</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0.456</td>
<td>NA</td>
<td>0.152</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Celastrina argiolus</td>
<td>0.114</td>
<td>0.011</td>
<td>0.076</td>
<td>0</td>
<td>0</td>
<td>0.051</td>
<td>0</td>
<td>0</td>
<td>0.076</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Charaxes jasius</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Coenonympha corinna</td>
<td>0.943</td>
<td>2.194</td>
<td>0.19</td>
<td>G</td>
<td>-</td>
<td>0.456</td>
<td>0.101</td>
<td>-</td>
<td>-</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Coenonympha pamphilus</td>
<td>0.864</td>
<td>0.86</td>
<td>0.152</td>
<td>-</td>
<td>-</td>
<td>0.152</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Colias croceus</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>-</td>
</tr>
<tr>
<td>Euchloe ausonia</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Favonius quercus</td>
<td>0.393</td>
<td>0.053</td>
<td>0.051</td>
<td>-</td>
<td>-</td>
<td>0.152</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gegenes pumilio</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glaucomysche alexis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gonepteryx cleopatra</td>
<td>0.385</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Gonepteryx rhamni</td>
<td>0.323</td>
<td>0.237</td>
<td>NA</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hipparchia aristeus</td>
<td>0.119</td>
<td>0.019</td>
<td>NA</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>NA</td>
<td>-</td>
</tr>
<tr>
<td>Hipparchia neomiris</td>
<td>0.871</td>
<td>0.968</td>
<td>End</td>
<td>C</td>
<td>-</td>
<td>NA</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hipparchia statilinus</td>
<td>0.281</td>
<td>0.053</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Iphiclides podalirius</td>
<td>0.701</td>
<td>0.063</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NA</td>
</tr>
<tr>
<td>Issoria lathonia</td>
<td>0.431</td>
<td>0.014</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Lampides boeticus</td>
<td>0.044</td>
<td>0.005</td>
<td>0.038</td>
<td>0</td>
<td>0.152</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>-</td>
</tr>
<tr>
<td>Species</td>
<td>Coverage</td>
<td>Density</td>
<td>Habitat</td>
<td>Prevalence</td>
<td>Disease</td>
<td>hosts</td>
<td>Impact</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>---------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Lasiommata megera/paramegaera</em></td>
<td>0.966</td>
<td>0.615</td>
<td>0.37</td>
<td>0.798</td>
<td>0.076</td>
<td>0.152</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Leptidea sinapis</em></td>
<td>0.501</td>
<td>0.081</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Leptotes piritous</em></td>
<td>0.298</td>
<td>0.048</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Limenitis reducta</em></td>
<td>0.185</td>
<td>0.015</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Lycaena phlaeas</em></td>
<td>0.762</td>
<td>0.098</td>
<td>0.051</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0.152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Maniola jurtina</em></td>
<td>0.603</td>
<td>0.298</td>
<td>0.085</td>
<td>0</td>
<td>0</td>
<td>0.092</td>
<td>0.076</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Melitaea cinxia</em></td>
<td>0.379</td>
<td>0.54</td>
<td>NA</td>
<td>E</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Melitaea didyma</em></td>
<td>0.344</td>
<td>0.07</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Melitaea nevadensis</em></td>
<td>0.455</td>
<td>0.492</td>
<td>0.988</td>
<td>-</td>
<td>0.988</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nymphalis polychloros</em></td>
<td>0.415</td>
<td>0.043</td>
<td>NA</td>
<td>E</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Papilio machaon</em></td>
<td>0.88</td>
<td>0.636</td>
<td>0.076</td>
<td>NA</td>
<td>0.076</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pararge aegeria</em></td>
<td>0.501</td>
<td>0.081</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0.152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pieris brassicae</em></td>
<td>0.185</td>
<td>0.015</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0.076</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pieris nappii</em></td>
<td>0.432</td>
<td>0.133</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pieris rapae</em></td>
<td>0.118</td>
<td>0.05</td>
<td>0.054</td>
<td>0.157</td>
<td>0.051</td>
<td>0</td>
<td>0.038</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Plebejus bellieri/idas</em></td>
<td>0.616</td>
<td>0.134</td>
<td>0.051</td>
<td>0</td>
<td>0.152</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Polyommatus icarus</em></td>
<td>0.216</td>
<td>0.018</td>
<td>0.076</td>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pontia edusa</em></td>
<td>0.079</td>
<td>0.015</td>
<td>0.051</td>
<td>C</td>
<td>-</td>
<td>*</td>
<td>0.076</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pyrgus armoricanus</em></td>
<td>0.26</td>
<td>0.224</td>
<td>0.091</td>
<td>C</td>
<td>NA</td>
<td>0</td>
<td>0.152</td>
<td>0.076</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pyronia cecilia</em></td>
<td>0.26</td>
<td>0.015</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pyronia tithonus</em></td>
<td>0.26</td>
<td>0.015</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Satyrium lilicis</em></td>
<td>0.417</td>
<td>0.009</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Spialia sertorius</em></td>
<td>0.432</td>
<td>0.067</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Thymelicus acteon</em></td>
<td>0.127</td>
<td>0.009</td>
<td>0.051</td>
<td>NA</td>
<td>0.051</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Vanessa atalanta</em></td>
<td>0.032</td>
<td>0.007</td>
<td>0.038</td>
<td>0</td>
<td>0</td>
<td>0.076</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Vanessa cardui</em></td>
<td>0.293</td>
<td>0.053</td>
<td>0.152</td>
<td>-</td>
<td>0.228</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- E: Difficulty sampling
- G: Disease
- NA: Data not available
- *: Data not available
Table 3 Trait variables entered in the AIC phylogenetic regressions for Gst, Dst and Uniqueness (Uni). And the differences in Gst, Dst and Uniqueness between species that have disappeared at least from one island, compared to species that have not disappeared.

<table>
<thead>
<tr>
<th>Trait variables</th>
<th>Estimate</th>
<th>StdErr</th>
<th>t.value</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gst</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host Plants</td>
<td>-0.100</td>
<td>0.045</td>
<td>-2.226</td>
<td>0.032</td>
</tr>
<tr>
<td>Mobility PC1</td>
<td>-0.141</td>
<td>0.046</td>
<td>-3.031</td>
<td>0.004</td>
</tr>
<tr>
<td>Phenol PC1</td>
<td>0.058</td>
<td>0.043</td>
<td>1.363</td>
<td>0.181</td>
</tr>
<tr>
<td>Ecophy PC2</td>
<td>0.068</td>
<td>0.040</td>
<td>1.713</td>
<td>0.095</td>
</tr>
<tr>
<td><strong>Dst</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host Plants</td>
<td>-0.110</td>
<td>0.043</td>
<td>-2.550</td>
<td>0.015</td>
</tr>
<tr>
<td>Mobility PC1</td>
<td>-0.126</td>
<td>0.047</td>
<td>-2.660</td>
<td>0.011</td>
</tr>
<tr>
<td>Phenol PC1</td>
<td>0.060</td>
<td>0.042</td>
<td>1.454</td>
<td>0.153</td>
</tr>
<tr>
<td>Ecophy PC2</td>
<td>0.119</td>
<td>0.039</td>
<td>3.042</td>
<td>0.004</td>
</tr>
<tr>
<td><strong>Uni</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host Plants</td>
<td>-0.045</td>
<td>0.028</td>
<td>-1.643</td>
<td>0.108</td>
</tr>
<tr>
<td>Ecophy PC2</td>
<td>0.102</td>
<td>0.033</td>
<td>3.112</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Table 4: The effects of trait variables in the logistic phylogenetic regressions for species disappearance in Elba and Capraia.

<table>
<thead>
<tr>
<th>Island</th>
<th>PC1/PC2</th>
<th>Estimate</th>
<th>StdErr</th>
<th>z.value</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elba</td>
<td>Host_Plants</td>
<td>-0.380</td>
<td>0.469</td>
<td>-0.810</td>
<td>0.418</td>
</tr>
<tr>
<td></td>
<td>Mobility PC1</td>
<td>0.412</td>
<td>0.482</td>
<td>0.855</td>
<td>0.393</td>
</tr>
<tr>
<td></td>
<td>Phenol PC1</td>
<td>-0.382</td>
<td>0.544</td>
<td>-0.702</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>Phenol PC2</td>
<td>-0.522</td>
<td>0.437</td>
<td>-1.194</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>Ecophy PC1</td>
<td>-0.813</td>
<td>0.584</td>
<td>-1.393</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>Ecophy PC2</td>
<td>0.766</td>
<td>0.561</td>
<td>1.367</td>
<td>0.172</td>
</tr>
<tr>
<td>Capraia</td>
<td>Host_Plants</td>
<td>-0.969</td>
<td>0.972</td>
<td>-0.997</td>
<td>0.319</td>
</tr>
<tr>
<td></td>
<td>Mobility PC1</td>
<td>1.605</td>
<td>1.206</td>
<td>1.331</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>Phenol PC1</td>
<td>-6.375</td>
<td>3.002</td>
<td>-2.124</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>Phenol PC2</td>
<td>2.835</td>
<td>1.869</td>
<td>1.517</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>Ecophy PC1</td>
<td>-1.820</td>
<td>1.196</td>
<td>-1.522</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>Ecophy PC2</td>
<td>2.753</td>
<td>1.783</td>
<td>1.544</td>
<td>0.123</td>
</tr>
</tbody>
</table>
Figure 1 (a) The study region where the islands and mainland areas are highlighted with the same colours obtained in the RGB projection of the Principal Coordinate Analyses (PCoA) in figure 3; (b) the workflow of the protocol used for the analyses.
Figure 2 Phylogenetic tree based on cytochrome c oxidase subunit 1 (COI) sequences of the butterfly species occurring in the Tuscan Archipelago, with their Gst, Dst, Uni and ecological traits. Square root Dst values are mapped over the tree, while Gst and Uni are reported as coloured squares (character mapping on the tree for Gst and Uni is available in Appendix S1). Host plants represent the number of plant genera on which the larva of a given species has been reported: one leaf - one genus; two leaves two to four genera, three leaves more than four genera. The sizes of the butterfly silhouettes are directly correlated with the species size (wingspan). Ecophy 1 represents the first PC of physiology traits mostly represent mean temperature and is reported with colours representing quartiles of values (red, preference for warm temperatures; blue, preference for cold temperatures). Ecophy 2 represents the second PC of physiology traits mostly correlated with precipitation tolerance and represented in quartile from small (high tolerance to drought) to large drops (low tolerance to drought). Phenol 1 represent quartiles the first PC of phenology correlated with the length of flying period from shortest (one black sector) to longest (four black sectors). Phenol 2, mostly linked to the period of emergence, represent spring species appearing early in the year (black sector right-top) to species with a later summer-autumn appearance (black sector left-top).
Figure 3 Overall genetic patterns obtained after comparisons among islands based on Gst. (a)
The colours obtained in the RGB projection of the Principal Coordinate Analyses (PCoA) are more similar among more genetically similar island communities. When the colours are reported on a map (b) of the studied region they show similarity and contrast among areas. The locality codes are: Arg, Argentario; Cap, Capraia; Cor, Corsica; Elb, Elba; Gia, Giannutri; Gig, Giglio; Gor, Gorgona; Mon, Montecristo; Pia, Pianosa; Sar, Sardinia; Tus, Tuscany; T_C central Tuscany coast; T_N, northern Tuscany coast; T_S, southern Tuscany coast. (c) Haplotype networks based on the COI gene for nine species that exemplify different patterns of genetic variation, obtained by comparing haplotypes from the study area with other regions of Europe (c). A molecular assessment for all the species is available in the Supporting Information (see Appendix S1).
Supporting Information

Additional Supporting Information is available in the online version of this study:

Appendix S1 {Supplementary methods and results containing the taxonomic assessment for each species, the occurrence data on Tuscan islands from 1900 to 2015 and the COI assessment for all the species. Supplementary results for PCA and Phylogenetic Models are also provided}
Appendix S2 {The dataset in the form of a fasta file (sequences_TA.fas), the information about the specimens used in the study (dataset.txt), the table containing the ecological traits for the 52 species (selected.traits.txt) and the R scripts used to carry out the analyses (Dapporto_et_al_script.R)}

Conflict of Interest:
The authors declare no conflicts of interest
Appendix S1

Rise and fall of island butterfly diversity. Understanding population diversification and extinction in a highly diverse Archipelago.

Calculation of Dst and Gst indexes (Nei, 1987).

Dst is defined as:

\[ \text{Dst} = \text{Ht} - \text{Hs} \]

where \( \text{Ht} \) represents the average intraspecific p-distances for all specimens of a given species, and \( \text{Hs} \) is the average of the intra-population p-distances. Thus, Dst represents the average genetic differentiation among populations in p-distance units.

The second measure (Gst) is a standardized index (Nei, 1987) defined as:

\[ \text{Gst} = \frac{\text{Dst}}{\text{Ht}} \]

This index ranges from negative values to 1 (complete differentiation). Negative values (intra-area differentiation higher than inter-area differentiation) can have different subtle meanings, but most often are a bias due to relatively small sample sizes; usually they are set to zero (Meirmans & Hedrick, 2011) and for several cases in our study we set the negative values to zero.

We also calculated the pairwise Gst among all pairs of populations for each species, using the following formula:

\[ \text{Gst}_{ij} = \frac{\text{Dst}_{ij}}{\text{Ht}} \]

This represents the specific fraction of the overall genetic diversity (Ht) expressed as the inter-population diversification (Dst_{ij}) between a given pair of areas (i and j).
In the following section distributional, taxonomic and historical notes for all the species are reported together with their assessment for population diversification. The genetic distances among haplotypes have been reduced to two dimension by using Principal Component Analysis. The resulting configuration has been plotted in RGB space then the resulting colours for specimens reported in the map with pie charts.

*Carcharodus alceae* (Esper, 1780)

**1900-2005**


Giglio: 1980 (Biermann & Hesch, 1982).


**2006-2015**


Capraia: 2012 (authors).

Gorgona: 2015 (authors).

This species showed some diversification among haplotypes (maximum p-distance 0.9%) and a relatively high population diversification (Dst 0.258 and Gst 0.779). Some haplotypes from Pianosa and Giglio were not found on the mainland resulting in a uniqueness of 0.095.
Figure S1. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.43% of variance.
The taxon living in Sardinia and Corsica was identified in the past as *Spialia sertorius therapne*. Recent molecular studies showed that *S. therapne* must be considered as a good species with a diversification higher than 3% respect to *S. sertorius* (Hernández-Roldán et al, 2016). For this reason we did not include Sardo-Corsican populations of *S. therapne* as comparison for *S. sertorius*.

**1900-2005**


**2006-2015**


This species showed a low diversification among haplotypes (maximum p-distance 0.2%), and a rather low population diversification (Dst 0.009 and Gst 0.417). All the haplotypes found on Elba island were also found on the mainland.
Figure S2. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100.00% of variance.
This species showed a single haplotype in the 13 specimens we analyzed from Tuscany, Elba and Corsica.

This species showed a low diversification among haplotypes (maximum p-distance 0.5%), and a medium population diversification (Dst 0.067 and Gst 0.432) mostly due to the occurrence on Elba of a single haplotype very uncommon on the Tuscan mainland.
Figure S3. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Gegenes pumilio (Hoffmannsegg, 1804)

1900-2005

2006-2015
Giglio: 2013 (authors).

This species showed a single haplotype in the 5 specimens we analyzed from Tuscany, Elba, Giglio and Sardinia.

Zerynthia cassandra (Geyer, 1828)

1900-2005
Elba: 1932 (Bryk, 1932).

2006-2015

This species showed some diversification among haplotypes (maximum p-distance 0.8%), and a rather low population diversification (Dst 0.053 and Gst 0.293) mostly due to the occurrence on Elba of a single haplotype not occurring on the Tuscan mainland. This haplotype also produced a uniqueness of 0.152.
Figure S4. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 73.84% of variance.
**Iphiclides podalirius** (Linnaeus, 1758)

**1900-2005**


Giglio: no collection year (Balletto et al. 2007).

**2006-2015**


This species showed a low diversification among haplotypes (maximum p-distance 0.3%), and a low Dst (0.063). Nevertheless it showed a high Gst of 0.701, mostly due to the occurrence on Corsica of a single haplotype not occurring in the rest of the study area. All the island haplotypes have been found in other areas of the European mainland.

---

**Figure S5.** PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Papilio machaon Linnaeus, 1758

1900-2005

Capraia: 1914 (Razzauti, 1917).


2006-2015


Giglio: 2007, 2010-2014 (authors).

Gorgona: 2015 (authors).

Pianosa: 2014 (authors).

Not recorded on Capraia in the last 10 years.

This species showed a high diversification among haplotypes (maximum p-distance 1.7%), but their distribution was not spatially structured resulting in no population diversification (Dst and Gst 0). The detection on Giglio of a haplotype not recorded on the mainland resulted in a uniqueness value of 0.038.
Figure S6. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 89.53% of variance.
Pieris brassicae (Linnaeus, 1758)

1900-2005
Capraia: 1915 (Razzauti, 1917), 2000 (authors).
Montecristo: 2001 (authors).

2006-2015
Capraia: 2006 (authors).

This species showed a diversification among haplotypes (maximum p-distance 0.9%), but their distribution was not spatially structured resulting in no population diversification (Dst and Gst 0). The detection on Giglio island of a haplotype not found on the mainland resulted in a uniqueness value of 0.076.
Figure S7. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 92.72% of variance.
**Pieris mannii** (Mayer, 1851)

**1900-2005**


**2006-2015**

Elba: 2008-2010, 2012 (authors).

This species is uncommon in the study area and we only examined five specimens. They showed a very low diversification among haplotypes (maximum p-distance 0.5%), and their distribution was not spatially structured resulting in no population diversification (Dst and Gst 0). The haplotype found on Elba was also found on the mainland.

Figure S8. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Pieris rapae (Linnaeus, 1758)

1900-2005


Giglio: 1908 (Rocci and Turati, 1925), 1980 (Biermann & Hesch, 1982), 2003, 2004 (authors).


2006-2015


Gorgona: 2015 (authors).


Giannutri: 2014 (authors).


This species showed a high diversification among haplotypes (maximum p-distance 1.4%), but their distribution was not spatially structured resulting in low population diversification (Dst 0.050 and Gst 0.112). The detection on Tuscan islands of haplotypes not found on the mainland resulted in a uniqueness value of 0.054.
Figure S9. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 75.52% of variance.
Pieris napi (Linnaeus, 1758)

1900-2005


2006-2015


Giannutri: 2014 (authors).

This species showed a diversification among haplotypes (maximum p-distance 1.1%), and a discrete population diversification (Dst 0.133 and Gst 0.432) mostly due to the occurrence on Corsica of a series of haplotypes uncommon on the Tuscan mainland, Elba and Giannutri. All island haplotypes have been also found on the mainland.
Figure S10. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 98.41% of variance.
It has been recognized since decades that *Pontia edusa* and *Pontia daplidice* represent two highly differentiated cryptic species showing 7% of COI divergence and differences in allozymes (revised by John et al. 2013). For this reason, they have been considered as distinct entities in this study. In Sardinia and Corsica only *P. daplidice* has been found, while on Tuscan islands only *P. edusa* has been identified so far. It is unknown which species was found in Capraia between 1968 and 1970 (Gross, 1970). We attributed the population to *P. edusa* for the extinction analysis and it is unlikely that this decision have affected the results of the phylogenetic regression since the traits of the two species are almost identical.

**1900-2005**


Giglio: 1908 (Rocci & Turati, 1908).


**2006-2015**

Elba: 2008-2013 (authors).


*Not recorded on Capraia in the last 10 years.*

This species showed some diversification among haplotypes (maximum p-distance 0.8%), but their distribution was not spatially structured resulting in low population diversification (Dst 0.015 and Gst 0.079). The detection on Elba of a haplotype not found on the mainland resulted in a uniqueness value of 0.051.
Figure S11. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB color space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
In Sardinia and Corsica this species is replaced by the endemic taxon *E. insularis*. *E. ausonia* and *E. insularis* are not sister species and their p-distance in COI sequences is about 5% (unpublished data). For this reason, we did not include *E. insularis* in the analysis.

1900-2005


2006-2015


This species showed a single haplotype in the 13 specimens we analyzed from Tuscany, Argentario and Elba.

1900-2005


2006-2015


This species showed a low diversification among haplotypes (maximum p-distance 0.5%), and their distribution was not spatially structured resulting in no population diversification (Dst and Gst 0). The detection on Elba of a haplotypes not found on the mainland resulted in a uniqueness value of 0.051.
Figure S12. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 73.44% of variance.

Colias croceus (Geoffroy, 1785)

1900-2005


Giglio: 1980 (Biermann & Hesch, 1982), 2003 (authors).


2006-2016
This species showed a single haplotype in the 32 specimens we analyzed from several areas.
Gonepteryx cleopatra (Linnaeus, 1767)

**1900-2005**


Giglio: 1980 (Biermann & Hesch, 1982).

**2006-2015**


Giannutri: 2014 (authors).

Gorgona: 2015 (authors).


Pianosa: 2011, 2014 (authors).

This species showed only two haplotypes differentiated for a single mutation in the 32 specimens we analyzed from several areas. One of the two haplotypes has only been found on Sardinia and on other mainland regions outside the study area. This resulted in some Gst (0.385) in a low Dst (0.004) and in a zero value for Uni.
Figure S13. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Gonepteryx rhamni (Linnaeus, 1758)

1900-2005

Elba: 1908 (Biermann & Hesch, 1982, Balletto et al., 2007).

In the study area this species showed a series of haplotypes with a high diversification (maximum p-distance 1.4%). However, they revealed to be only partially spatially structured resulting in low values of population differentiation (Dst 0.237, Gst 0.323).

Not recorded on Elba in the last 10 years.

Figure S14. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 87.53% of variance.
It has been recently showed that in Europe the sinapis group is composed by a triplet of species (\textit{L. sinapis}, \textit{L. reali} and \textit{L. juvernica}). Over the study area (Sardinia, Corsica, Tuscan islands and Tuscany mainland), only \textit{L. sinapis} has been found so far. The species was probably common on Elba island in the past and a large series of specimens is preserved in the Roger Verity collection in MZUF. The last report for this species belongs to 1980, thereafter it apparently disappeared from the island.

**1900-2005**


Not recorded on Elba in the last 10 years.

We didn’t collect any specimen of this species in Tuscan islands. Over the study area \textit{L. sinapis} showed a low diversification among haplotypes (maximum p-distance 0.3%), but their distribution revealed a good population diversification (Dst 0.081 and Gst 0.501) since the haplotypes occurring on Sardinia and Corsica differ from those detected on the Tuscan mainland.
Figure S15. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Lycaena phlaeas (Linnaeus, 1761)

1900-2005


Capraia: 1979 (Biermann & Hesch, 1982).

Giglio: 1908 (Rocci & Turati, 1925), 1980 (Biermann & Hesch, 1982).


2006-2015

Elba: 2008-2013 (authors).


Montecristo: 2014 (authors).

Pianosa: 2011 (authors).

Not recorded on Capraia in the last 10 years.

This species showed a low diversification among haplotypes (maximum p-distance 0.5%), but their distribution revealed a high population diversification (Dst 0.098 and Gst 0.762) for the occurrence on Corsica and Montecristo of haplotypes not detected on the Tuscan mainland. These haplotypes were not found elsewhere in Europe resulting in a uniqueness of 0.051.
Figure S16. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.60% of variance.
This species showed a low diversification among haplotypes (maximum p-distance 0.5%), but their distribution revealed a high population diversification (Dst 0.053 and Gst 0.393). The haplotype detected on Elba was not found elsewhere in Europe resulting in a uniqueness of 0.051.

Figure S17. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
This species showed some diversification among haplotypes (maximum p-distance 0.9%), and their distribution revealed a high population diversification (Dst 0.279 and Gst 0.846) for the occurrence on all the studied islands of a group of haplotypes not detected on the Tuscan mainland. However, since these haplotypes also occur in western Europe (Iberia) this species did not score island uniqueness.

Figure S18. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.19% of variance.
This species showed a single haplotype in the 11 specimens we analyzed from Tuscany mainland, Elba and Argentario.
Lampides boeticus (Linnaeus, 1767)

1900-2005


Giglio: 1980 (Biermann & Hesch, 1982), 2004 (authors).

2006-2015

Elba: 2008-2010, 2012 (authors).


Montecristo: 2012 (authors).


Gorgona: 2015 (authors).

Pianosa: 2014 (authors).

This species showed a low diversification among haplotypes (maximum p-distance 0.6%), and their distribution was not spatially structured resulting in low population diversification (Dst 0.005 and Gst 0.045). The detection on Capraia of a haplotype not found on the mainland resulted in a uniqueness value of 0.038.
Figure S19. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 55.88% of variance.
Leptotes pirithous (Linnaeus, 1767)

1900-2005


Capraia: 1914 (Razzauti 1917); 1979 (Biermann & Hesch, 1982), 2000 (authors).


Giglio:

2006-2015


Capraia: 2006, 2010 (authors).

Montecristo: 2012 (authors).

Giannutri: 2014 (authors).

This species showed a low diversification among haplotypes (maximum p-distance 0.6%), and their spatially structure was low (Dst 0.048 and Gst 0.298). All the haplotypes found on Tuscan islands also occurred on mainland.
Figure S20. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.79% of variance.
Celastrina argiolus (Linnaeus, 1758)

1900-2005


Giglio: 1908, 1979 (revised by Biermann & Hesch, 1982; Balletto et al. 2007).


2006-2015


Gorgona: 2015 (authors).

Montecristo: 2014 (authors).

Pianosa: 2011 (authors).

This species showed a low diversification among haplotypes (maximum p-distance 0.3%), and their distribution was not spatially structured resulting in low population diversification (Dst 0.010 and Gst 0.114). The detection on Elba and Giglio of a haplotype not found on the mainland resulted in a uniqueness value of 0.076.
Figure S21. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.98% of variance.
This species showed a high diversification among haplotypes (maximum p-distance 2.0%), but their distribution was not spatially structured resulting in no population diversification (Dst and Gst were 0). The haplotype found on Elba also occurs on the European mainland.

Figure S22. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
The taxonomy of this group is still debated. The Sardo-Corsican populations as referred as *P. bellieri*, the Elba population is referred by some authors as *P. villai* or *P. bellieri villai* and the Tuscan populations as *P. idas* or *P. abetonica*. According to COI there clear distinction between insular and mainland populations populations. For this reason, they have been analyzed together.

**1900-2005**


**2006-2015**


This species showed some diversification among haplotypes (maximum p-distance 0.8%), and their distribution showed a good spatial structure resulting in Dst equal to 0.134 and Gst equal to 0.616. The detection on Elba of a haplotype not found on the mainland resulted in a uniqueness value of 0.051.
Figure S23. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 82.89% of variance.
Aricia agestis/cramera

In Sardinia, A. agestis is replaced by the South-Western taxon A. cramera. The two taxa show a rather low COI divergence (around 2.4% depending on different haplotypes) (Sañudo-Restrepo et al., 2013). For this reason, A. agestis from Tuscany, Tuscan islands and Corsica and A. cramera from Sardinia have been used together in the analyses.

1900-2005


Giglio: 1908 (Rocci and Turati, 1925), 1979 (Biermann & Hesch, 1982).


2006-2015


Capraia: 2013 (authors).

Not recorded on Giglio in the last 10 years.

As reported above, this species showed a high diversification among two taxa (maximum p-distance 3.2%), and their distribution revealed a high population diversification (Dst 0.984 and Gst 0.738) for the occurrence on all the studied islands of a group of haplotypes not detected on the Tuscan mainland. Some haplotypes were not recorded on the mainland and this species scored a uniqueness of 0.094.
Figure S24. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 96.73% of variance.
Polyommatus icarus (Rottemburg, 1775)

It has been recently found that the populations from Sardinia belong to *P. celina*, a S-W Mediterranean taxon morphologically almost identical to *P. icarus*. The COI divergence is high (6%) and, for this reason, *P. celina* has been removed from the analysis of *P. icarus*.

1900-2005


Capraia: 1979 (Biermann & Hesch, 1982).

Giglio: 1908, 1979 (Biermann & Hesch, 1982; Balletto et al. 2007).


2006-2015


Capraia: 2013, 2014 (authors).

Pianosa: 2011 (authors).

Not recorded on Giglio in the last 10 years.

This species showed a low diversification among haplotypes (maximum p-distance 0.5%) and in a rather low population diversification (Dst 0.018 and Gst 0.216). The occurrence on Capraia of a haplotype not found on the mainland resulted in an island uniqueness of 0.076.
Figure S25. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 46.86% of variance.
**Argynnis paphia** (Linnaeus, 1758)

1900-2005

Elba: 1908, 1916, 1921 (Biermann & Hesch, 1982; Balletto et al. 2007).

Not recorded on Elba in the last 10 years.

We didn’t collect any specimen of this species in Tuscan islands. Over the study area (Tuscany, Corsica and Sardinia) this species showed a single haplotype.

**Argynnis pandora** (Denis & Schiffermüller, 1775)

1900-2005


Giglio: 1908 (Rocci & Turati, 1925), 1979 (Biermann & Hesch, 1982).

2006-2015


This species showed a high diversification among haplotypes (maximum p-distance 2.6%), but their distribution was not highly spatially structured resulting a low population Gst (0.171) and in a medium Dst (0.150). One haplotype detected on Giglio and Elba was not found on the mainland resulting in a uniqueness value of 0.076.
Figure S26. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 98.05 % of variance.
Issoria lathonia (Linnaeus, 1758)

1900-2005
Giglio: 1908 (Rocci & Turati, 1925), 1979 (Biermann & Hesch, 1982).

2006-2015
Giglio: 2014 (authors).

This species showed only three haplotypes only slightly differentiated for a single mutation. One of them has been only found in Corsica resulting in some Gst (0.431) in a low Dst (0.014) and in a zero value for island uniqueness.

Figure S27. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
In Sardinia and Corsica, *A. urticae* is replaced by the island endemic *A. ichnusa*, showing a low COI divergence (around 1.2% in our sample) (see also Vandewoestijne et al., 2004). For this reason, the two species have been analyzed together. The two taxa differ for wing pattern and the specimen collected on Elba island by Roger Verity (still preserved in his collection) revealed a typical *A. urticae* wing pattern.

**1900-2005**

Elba: 1908, 1916 (Biermann & Hesch, 1982).

*Not recorded on Elba in the last 10 years.*

According to the existence of two taxa, *A. urticae* and *A. ichnusa* showed a high diversification between Tuscany and Corsica-Sardinia (maximum divergence 1.2%, Gst (0.934) and Dst (0.449)).
Figure S28. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
**Nymphalis polychloros** (Linnaeus, 1758)

### 1900-2005

Elba: 1908, 1916 (Biermann & Hesch, 1982; Balletto et al. 2007)

In the study area this species showed a series of haplotypes with a low level of diversification (maximum p-distance 0.3%). However, they revealed some spatial structured resulting in a rather low value of Dst due to the low diversification (0.043) but in a medium values of Gst (0.415).

*Not recorded on Elba in the last 10 years.*

---

**Figure S29.** PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB color space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Vanessa atalanta (Linnaeus, 1758)

1900-2005


Giglio: 1908, 1979 (Biermann & Hesch, 1982; Balletto et al. 2007), 2003 (authors).


Pianosa: 1998 (Dapporto et al., 1999).

2006-2015


This species showed a low diversification among haplotypes (maximum p-distance 0.5%), and their distribution was not spatially structured resulting in low population diversification (Dst 0.009 and Gst 0.126). The detection on Capraia and Elba of two haplotypes not found on the mainland resulted in a uniqueness value of 0.051.
Figure S30. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 82.39% of variance.
Vanessa cardui (Linnaeus, 1758)

1900-2005


Giglio: 1908 (Rocci & Turati, 1925), 1979 (Biermann & Hesch, 1982).

Pianosa: 1998 (Dapporto et al., 1999).

Montecristo: 1979 (Fanfani & Groppali, 1979).

2006-2015


This species showed a low diversification among haplotypes (maximum p-distance 0.8%), and their distribution was not spatially structured resulting in very low population diversification (Dst 0.007 and Gst 0.032). The detection on Gorgona and Giannutri of a haplotype not found on the mainland resulted in a uniqueness value of 0.038.
Figure S31. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 76.06% of variance.
**Melitaea didyma** (Esper, 1778)

**1900-2005**


**2006-2015**

Elba: 2012, 2015 (authors).

This species showed a low diversification among haplotypes (maximum p-distance 0.6%) but their distribution revealed some spatial structure (Dst 0.070 and Gst 0.344). The haplotype from Elba has been also found on the mainland.

**Figure S32.** PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 93.13% of variance.
Melitaea nevadensis (Oberthür, 1904)

1900-2005

2006-2015
Elba: 2015 (authors).

This species revealed a high intra-specific diversification in the study area (2.1%). The existence of two closely related haplotypes from Elba not found elsewhere resulted in high population differentiation (Dst 0.492, Gst 0.454 and Uni 0.989). The haplotypes found on Elba resulted in the highest uniqueness measured in this study.

Figure S33. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 91.39% of variance.
**Limenitis reducta** Staudinger, 1901

1900-2005

2006-2015
Elba: 2006-2013 (authors).

This species showed a low diversification among haplotypes (maximum p-distance 0.3%) and their distribution showed a very low spatial structure (Dst 0.015 and Gst 0.185). The haplotypes from Elba have been also found on the mainland.

Figure S34. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 79.48% of variance.
Charaxes jasius (Linnaeus, 1767)

1900-2005
2000 (authors)
Giglio: 1908 (Rocci & Turati, 1925), 1979 (Biermann & Hesch, 1982), 2004 (authors).
Giannutri: 2014 (authors).
2006-2015
Giannutri: 2014 (authors).
This species showed three haplotypes only differentiated for a few mutations and with a maximum diversification of 0.2%. One of the haplotypes has been only found on Sardinia resulting in negative Gst and Dst values which have been set to zero. The haplotype from Tuscan islands has been also found on the mainland.
Figure S35. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 100% of variance.
Pararge aegeria (Linnaeus, 1758)

1900-2005
Elba: 1908, 1916, 1919, 1921, 1968, 1980 (Biermann & Hesch, 1982; Balletto et al. 2007),
1998 (Biermann, 1999), 2002, 2003 (authors)
Montecristo: 1979 (Fanfani & Groppali 1979).
Giglio: 1908 (Rocci & Turati 1925), 1979 (Biermann & Hesch, 1982).

2006-2015
Giglio: 2007, 2010, 2014 (authors)
Giannuutri: 2014 (authors)
Pianosa: 2011 (authors).
Montecristo: 2014 (authors).
Gorgona: 2015 (authors).

It is well known that this species shows two diverging clades between North Africa and
Europe (Weingartner et al. 2006; Vodă et al. 2016) (2.1% of maximum divergence in our
dataset). We found that the North African lineage also occurs on Sardinia and a single
specimen has been also found on Corsica. All the specimens from Tuscan islands revealed to
belong to the European clade. According to this diversification the species revealed a high
population differentiation (Dst 0.636, Gst 0.880); the occurrence on Capraia and Pianosa of a
haplotype not found on the mainland resulted in a uniqueness value of 0.076.
Figure S36.  PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 96.10% of variance.
Lasiommata megera/paramegaera

Lasiommata megera is replaced on Sardinia, Corsica and Capraia by the insular endemic taxon *L. paramegaera*, differentiated in wing pattern, genitalia shape and composition of cuticular lipids (e.g. Kudrna 1977; Dapporto 2007, 2008). To our knowledge no COI sequence was available in literature for *L. paramegaera*. Based on our sequences the genetic distance in COI between *L. megera* and *L. paramegaera* is lower than 3% (1.5% of maximum divergence). For this reason we analysed the two taxa together.

1900-2005


Montecristo: 1979 (Fanfani & Groppali 1979), 2001 (authors).

Giglio: 1908 (Rocci & Turati 1925), 1979 (Biermann & Hesch, 1982), 2003, 2004 (authors).


Pianosa: 1998 (Dapporto et al., 1999).


2006-2015


Pianosa: 2011, 2014 (authors).


Gorgona: 2015 (authors).

We revealed a strong correlation between the COI structure and the identification of populations to *L. megera* and *L. paramegaera* (Dapporto 2008) with the individuals belonging to Capraia and Montecristo clustering with those from Sardinia and Corsica.
According to this diversification this complex of taxa showed a high population differentiation (Dst 0.615, Gst 0.966, Uni 0.370).

Figure S37. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB color space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.79% of variance.
Coenonympha corinna (Hübner, 1804)

The taxonomic status of Coenonympha corinna (Sardinia, Corsica, Capraia) and Coenonympha elbana (Elba, Giglio, Giannutri, Argentario and Tuscany coast) has been highly debated in last years resulting in different authors considering them as different taxa (e.g. Kodandaramaiah et al. 2009; Balletto et al. 2007) or as a single one (e.g. Dapporto & Strumia 2008; Kudrna et al. 2015). Kodandaramaiah et al. (2009) compared specimens from Sardinia and Elba for genetic markers finding some diversification while Dapporto and Strumia (2008) did not reveal any constant morphological diversification in genitalia and a continuous cline in ocellation.

1900-2005


Giglio: 1980 (Biermann & Hesch, 1982), 2000 (Balletto et al. 2007).


Giannutri: 1994 (Baletto et al. 2007).

2006-2015


Not recorded on Giglio in the last 10 years.

The analysis of a complete COI dataset revealed a complex pattern of diversification with the existence of three main clades (I. Sardinia, II. Corsica+Capraia, III. Elba+Giannutri+Tuscan mainland). This diversification did not correlate with the supposed taxonomic diversification between Sardinia-Corsica-Capraia and Elba-Giannutri-Tuscany coast populations. The
haplotype diversification appeared to be high (with a maximum divergence of 5.2%) and it revealed a high population diversification and uniqueness (Dst 2.194, Gst 0.943, Uni 0.190).

Figure S38. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 97.41% of variance.
Coenonympha pamphilus (Linnaeus, 1758)

It has been suggested that populations from North Africa, Southern Spain, Balearics and Sardinia belong to a different species (*Coenonympha lyllus*) (Boillat, 2002). Actually, populations from North Africa and Balearics differ from those belonging to Europe but the p-distance in COI are low (Dincă et al. 2015; Vodă et al. 2016). We confirmed here that the Sardinian specimens belong to the North African clade and that the maximum divergence was 2.4%, for this reason we analysed the two taxa together.

**1900-2005**


**2006-2015**


We revealed a strong correlation between the COI structure and the identification of populations to *C. pamphilus* and *C. lyllus* (Boillat 2002) with the individuals from Sardinia representing a different genetic clade from those belonging to other localities. According to this diversification this taxon showed a high population differentiation (Dst 0.856, Gst 0.864). One haplotype from Elba was not recorded on the mainland resulted in a uniqueness of 0.152.
Figure S39. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 98.74% of variance.
**Pyronia tithonus** (Linnaeus, 1771)

**1900-2005**


**2006-2015**


This species showed a low diversification among haplotypes (maximum p-distance 0.3%), and their distribution was not spatially structured resulting in low population diversification (Dst 0.015 and Gst 0.260). The haplotype from Elba also occurs on the mainland.

Figure S40. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 97.04% of variance.
**Pyronia cecilia** (Vallantin, 1894)

### 1900-2005

- **Capraia:** 1961, 1978, 1996 (Biermann & Hesch, 1982; Balletto et al. 2007), 2001 (authors).
- **Pianosa:** 1998 (Dapporto et al., 1999).
- **Giglio:** 1980 (Biermann & Hesch, 1982), 2003, 2004 (authors).

### 2006-2015

- **Pianosa:** 2011 (authors).
- **Giglio:** 2008, 2010 (authors).

*Not recorded on Capraia in the last 10 years.*

As showed by Vodă et al. (2016) this species shows two highly diverging clades, the first belongs to North Africa, the second to the Italian mainland. The two clades are sympatric in Sicily. We confirmed the existence of two highly diverging clades also in the study area (maximum diversification 2.9%). The clades from S-W Mediterranean is the only occurring on the studied islands, while on the mainland it coexists with the clade typical of the Italian peninsula. According to this pattern the species showed a high Dst (0.224) but a medium Gst (0.260). The occurrence on Pianosa and Giglio of haplotypes not found on the mainland resulted in an island uniqueness value of 0.091.
Figure S41. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 85.40% of variance.
1912 *Maniola jurtina* (Linnaeus, 1758)

1914 **1900-2005**


1922 **2006-2015**


1927 It has been recognized since decades that this species shows two main lineages in Europe differing for morphological and genetic characters (e.g. Thomson 1987; Schmitt et al. 2005; Dapporto et al. 2009; 2014). The study area is one of the contact and hybrid zones for the two lineages (Thomson 1987; Dapporto et al. 2014). We confirmed this observation showing that this species has a maximum divergence of 1.2% in the study area and that the two lineages coexist on Elba and Corsica with similar frequencies. Accordingly, the parameters for populations differentiation were high (Dst 0.298, Gst 0.603). The occurrence on Elba and Pianosa of two haplotypes not found on the mainland resulted in an island uniqueness value of 0.085.
Figure S42. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 93.34% of variance.
1944 *Hipparchia neomiris* (Godart, 1822)

1946 **1900-2005**


1950 **2006-2015**


1953 *Not recorded in Capraia in the last 10 years.*

1955 This species, endemic of the study area, showed a large diversification (maximum p-distance 2.3%) between two lineages, the first occurring on Sardinia and the second on Corsica and Elba. A previous analysis on genitalic shape also revealed a diversification among the same populations (Dapporto 2010), suggesting the need of further studies to understand the taxonomic status of the two lineages. According to such a clear spatial pattern, the species showed a high population differentiation (Dst 0.968, Gst 0.871). The status of island endemic did not allow the calculation of uniqueness.
Figure S43. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 99.85% of variance.
The taxonomic status of the species belonging to the *H. semele* group is highly debated mostly on Mediterranean islands (e.g. Cesaroni et al. 1994; Balletto et al. 2007; Dapporto 2010; Kudrna et al. 2015; Vodă et al. 2016). For COI *Hipparchia aristaeus* represents a clearly distinct clade separated by more than 4% of divergence from *H. semele* and for this reason we analysed it separately.

**1900-2005**


**2006-2015**


This species showed a very low genetic variation over the study area (maximum p-distance 0.6%) and almost no spatial structure (Dst 0.019, Gst 0.119). The status of island endemic did not allow the calculation of uniqueness.
Figure S44. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 87.53% of variance.
1998  *Hipparchia statilinus* (Hufnagel, 1766)

1999

2000  **1900-2005**


2002  **2006-2015**


2004

2005  This species showed a low diversification among haplotypes (maximum p-distance 0.5%), and their distribution was slightly spatially structured resulting in low population diversification (Dst 0.053 and Gst 0.281). The haplotypes occurring on Elba have been also found on the mainland.

2006

2007

2008

2009

2010  Figure S45. PCoA projection (left) of genetic distances among specimens (dots) in the bidimensional RGB colour space. Individual colours have been subsequently plotted on the map (right). The first two PCoA axes explained 85.09% of variance.
Supplementary results

Figure S46. PCA scatterplot for wingspan (left). Data belong to four literature sources: WingspanP (Pamperis, 2009), WingspanL (Lafranchis, 2000), WingspanHR (Higgins & Riley, 1970), WingspanfemalesT male data from Tshikolovets (2011), WingspanmalesT male data from Tshikolovets (2011). Tags for butterflies are obtained by using the first letter of the genus name and the first three letters of the species name. The eigenvalue plot for the five components (right) shows that only the first PC (red dot) had an eigenvalue higher than 1.
Figure S47. PCA scatterplot for phenology traits (left). Tags for butterflies are obtained by using the first letter of the genus name and the first three letters of the species name. The eigenvalue plot for the five components (right) shows that the first two PCs (red dots) had an eigenvalue higher than 1.
Figure S48. PCA scatterplot for physiological traits (left). Tags for butterflies are obtained by using the first letter of the genus name and the first three letters of the species name. The eigenvalue plot for the five components (right) shows that the first two PCs (red dots) had an eigenvalue higher than 1.

<table>
<thead>
<tr>
<th>Host_Plants</th>
<th>MobilityPC1</th>
<th>PhenolPC1</th>
<th>PhenolPC2</th>
<th>EcophyPC1</th>
<th>EcophyPC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MobilityPC1</td>
<td>-0.168</td>
<td>0.277</td>
<td>0.277</td>
<td>-0.351</td>
<td>0.090</td>
</tr>
<tr>
<td>PhenolPC1</td>
<td>0.119</td>
<td>0.199</td>
<td>-0.200</td>
<td>-0.148</td>
<td>0.109</td>
</tr>
<tr>
<td>PhenolPC2</td>
<td></td>
<td></td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EcophyPC1</td>
<td></td>
<td></td>
<td></td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>EcophyPC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.073</td>
</tr>
</tbody>
</table>
Table S1. Correlations among host plants and the six traits obtained after Principal Component Analysis.
Figure S49. The topological constraints at family and subfamily levels applied to the COI phylogenetic tree based following the butterfly phylogeny published in Heikkila et al. (2012).
Figure S50. Supports for nodes in the phylogenetic tree used in phylogenetic regression. Family and subfamily nodes costrained according to Heikkila et al. (2012) are by default characterized by a maximum support.
Figure S51 Gst values for species mapped over the phylogenetic tree based on cytochrome c oxidase subunit 1 (COI) tree. Host plants represent the number of plant genera on which the larva of a given species has been reported: one leaf - one genus; two leaves two to four genera, three leaves more than four genera. The sizes of the butterfly silhouettes are directly correlated with the species size (wingspan). Ecophy 1 represents the first PC of physiology traits mostly represent mean temperature and is reported with colours representing quartiles of values (red, preference for warm temperatures; blue, preference for cold temperatures). Ecophy 2 represents the second PC of physiology traits mostly correlated with precipitation tolerance and represented in quartile from small (high tolerance to drought) to large drops (low tolerance to drought). Phenol 1 represent quartiles the first PC of phenology correlated with the length of flying period from shortest (one black sector) to longest (four black sectors). Phenol 2, mostly linked to the period of emergence, represent spring species appearing early in the year (black sector right-top) to species with a later summer-autumn appearance (black sector left-top).
Figure S52 Uniqueness values for species mapped over the phylogenetic tree based on cytochrome c oxidase subunit 1 (COI) tree. Symbols for traits as in the previous figure.
Figure S53. Plots for observed vs fitted values in phylogenetic multiple regression models for Gst, Dst and uniqueness.

Figure S54. Plot for observed vs fitted values in phylogenetic generalized linear model for extinction in Capraia.
In the main paper, we recognised as separate units most species accepted by the Fauna Europaea Project (Karsholt & Van Nieukerken, 2013; www.faunaeur.org) and considered as single unit taxa displaying a minimum COI p-distance lower than 3%. We re-performed all the analyses without setting any threshold. In this case we used the Fauna Europea taxonomy to identify units. The results turned out to be very similar and they are reported below:

<table>
<thead>
<tr>
<th></th>
<th>Gst</th>
<th>Dst</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dst</td>
<td>0.790***</td>
<td></td>
</tr>
<tr>
<td>Uni</td>
<td>0.103 n.s.</td>
<td>0.385**</td>
</tr>
</tbody>
</table>

Table S2. Spearman correlation between Gst, Dst and Uni

<table>
<thead>
<tr>
<th>Index</th>
<th>Pagel’s lambda</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gst</td>
<td>0.136</td>
<td>0.519</td>
</tr>
<tr>
<td>Dst</td>
<td>0.342</td>
<td>0.053</td>
</tr>
<tr>
<td>Uni</td>
<td>0.050</td>
<td>0.709</td>
</tr>
</tbody>
</table>

Table S3. Phylogenetic signal for Gst, Dst and Uni.
Figure S55. Overall genetic patterns obtained after comparisons among islands based on Gst. (left) The colours obtained in the RGB projection of the Principal Coordinate Analyses have been included in (right) the map of the studied region to show similarity and contrast among areas. The locality codes are: Arg, Argentario; Cap, Capraia; Cor, Corsica; Elb, Elba; Gia, Giannutri; Gig, Giglio; Gor, Gorgona; Mon, Montecristo; Pia, Pianosa; Sar, Sardinia; Tus, Tuscany; T_C central Tuscany coast; T_N, northern Tuscany coast; T_S, southern Tuscany coast. Procrustes correlation 0.641, P=0.006.
<table>
<thead>
<tr>
<th>Trait Variables</th>
<th>Estimate</th>
<th>StdErr</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gst</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host_Plants</td>
<td>-0.076</td>
<td>0.040</td>
<td>-1.880</td>
<td>0.068</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>-0.126</td>
<td>0.042</td>
<td>-2.903</td>
<td>0.006</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.066</td>
<td>0.038</td>
<td>1.769</td>
<td>0.085</td>
</tr>
<tr>
<td><strong>Dst</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host_Plants</td>
<td>-0.080</td>
<td>0.088</td>
<td>-2.520</td>
<td>0.016</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>-0.098</td>
<td>0.048</td>
<td>-2.040</td>
<td>0.048</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.122</td>
<td>0.032</td>
<td>3.860</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Uniqueness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host_Plants</td>
<td>-0.048</td>
<td>0.026</td>
<td>-1.843</td>
<td>0.073</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.093</td>
<td>0.025</td>
<td>3.680</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table S4 Trait variables entered in the AIC phylogenetic regressions for Gst, Dst and Uniqueness (Uni).
### Host_Plants

<table>
<thead>
<tr>
<th></th>
<th>Elba</th>
<th>Capraia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host_Plants</td>
<td>-0.383</td>
<td>0.969</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>0.410</td>
<td>1.628</td>
</tr>
<tr>
<td>PhenologyPC1</td>
<td>-0.401</td>
<td>-6.470</td>
</tr>
<tr>
<td>PhenologyPC2</td>
<td>-0.521</td>
<td>2.906</td>
</tr>
<tr>
<td>EcophysiolPC1</td>
<td>-0.843</td>
<td>-1.877</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.790</td>
<td>2.765</td>
</tr>
</tbody>
</table>

### MobilityPC1

<table>
<thead>
<tr>
<th></th>
<th>Elba</th>
<th>Capraia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host_Plants</td>
<td>0.467</td>
<td>0.974</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>0.481</td>
<td>1.216</td>
</tr>
<tr>
<td>PhenologyPC1</td>
<td>0.555</td>
<td>3.055</td>
</tr>
<tr>
<td>PhenologyPC2</td>
<td>0.438</td>
<td>1.905</td>
</tr>
<tr>
<td>EcophysiolPC1</td>
<td>0.622</td>
<td>1.124</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.568</td>
<td>1.789</td>
</tr>
</tbody>
</table>

### PhenologyPC1

<table>
<thead>
<tr>
<th></th>
<th>Elba</th>
<th>Capraia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host_Plants</td>
<td>-0.819</td>
<td>-0.995</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>0.851</td>
<td>1.339</td>
</tr>
<tr>
<td>PhenologyPC1</td>
<td>-0.723</td>
<td>-2.118</td>
</tr>
<tr>
<td>PhenologyPC2</td>
<td>-1.190</td>
<td>1.524</td>
</tr>
<tr>
<td>EcophysiolPC1</td>
<td>-1.355</td>
<td>-1.508</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>1.392</td>
<td>1.545</td>
</tr>
</tbody>
</table>

### PhenologyPC2

<table>
<thead>
<tr>
<th></th>
<th>Elba</th>
<th>Capraia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host_Plants</td>
<td>0.413</td>
<td>0.320</td>
</tr>
<tr>
<td>MobilityPC1</td>
<td>0.395</td>
<td>0.181</td>
</tr>
<tr>
<td>PhenologyPC1</td>
<td>0.469</td>
<td>0.034</td>
</tr>
<tr>
<td>PhenologyPC2</td>
<td>0.234</td>
<td>0.127</td>
</tr>
<tr>
<td>EcophysiolPC1</td>
<td>0.176</td>
<td>0.132</td>
</tr>
<tr>
<td>EcophysiolPC2</td>
<td>0.164</td>
<td>0.122</td>
</tr>
</tbody>
</table>

**Table S5.** Species traits associated to disappearance on Elba and Capraia based on phylogenetic GLM with logistic model.
References


Schmitt T, Röber S, Seitz A (2005) Is the last glaciation the only relevant event for the present genetic population structure of the Meadow Brown butterfly Maniola jurtina (Lepidoptera: Nymphalidae)? *Biological Journal of the Linnean Society*, 85, 419-431.


