Received: 18 March 2022 | Revised: 25 June 2022 | Accepted: 8 July 2022

DOI: 10.1111/geb.13579

Global Ecology and Biogeography | A Journal of Macroecology

WILEY

DATA ARTICLE

The atlas of mitochondrial genetic diversity for Western Palaearctic butterflies

Leonardo Dapporto[1] | Mattia Menchetti[2] | Raluca Vodă[3] | Cecilia Corbella[2] | Sylvain Cuvelier[4] | Imed Djemadi[5] | Martin Gascoigne-Pees[6] | Joan C. Hinojosa[2] | Nok Ting Lam[7] | Maria Serracanta[2,8] | Gerard Talavera[9] | Vlad Dincă[7] | Roger Vila[2]

[1]ZEN Lab, Dipartimento di Biologia dell'Università di Firenze, Florence, Italy

[2]Institut de Biologia Evolutiva (CSIC-UPF), Barcelona, Spain

[3]Evolutionary Ecology and Genetics Group, Earth and Life Institute, UCLouvain, Louvain-la-Neuve, Belgium

[4]VVE Workgroup Butterflies, Ieper, Belgium

[5]Biodiversity and Environment Division, Environmental Research Center, Annaba, Algeria

[6]2 Barrett's Close, Stonesfield, Oxfordshire, UK

[7]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

[8]Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain

[9]Institut Botànic de Barcelona (IBB), CSIC-Ajuntament de Barcelona, Barcelona, Spain

Correspondence
Leonardo Dapporto, ZEN Lab, Dipartimento di Biologia dell'Università di Firenze, via Madonna del Piano 6, 50019, Sesto Fiorentino, Florence, Italy.
Email: leondap@gmail.com

Funding information
"la Caixa" Foundation, Grant/Award Number: LCF/BQ/DR20/11790020; 7th European Community Framework Programme, Grant/Award Number: 625997; Academy of Finland, Grant/Award Number: 328895; ESF Investing in your future; Generalitat

Abstract

Motivation: Butterflies represent a model in biology and a flagship group for invertebrate conservation. We provide four new resources for the Western Palaearctic butterflies: (1) an updated checklist comprising 552 species; (2) a curated dataset of 32,126 mitochondrial cytochrome *c* oxidase subunit I (*COI*) sequences for 532 species, including a *de novo* reference library for the Maghreb (Morocco, northern Algeria and Tunisia) and Macaronesia (Azores, Madeira and Canary Islands); (3) seven indexes of intraspecific genetic variation (IGV): observed and expected number of haplotypes, haplotype and nucleotide diversity, two fixation indexes and maximum p-distance; and (4) species-level maps illustrating the distribution of *COI* variability and haplotype networks. The updated checklist will be fundamental for any application dealing with butterfly diversity in the Western Palaearctic. The IGV indexes provide measures for genetic polymorphism and spatial structure and represent proxies for dispersal capacity. These resources will facilitate comparative studies of macrogenetics, foster integrative taxonomy and aid conservation strategies.

Main types of variables contained: A complete species checklist in table format, 32,126 mitochondrial DNA barcodes provided with metadata (species membership, WGS84 coordinates and sequence length) and a book in PDF format, including the IGV atlas and indexes, are provided.

Spatial location and grain: The checklist encompasses Europe up to the Urals in the east, north Macaronesia (the Azores, Madeira and the Canary Islands) and the Maghreb (Morocco, northern Algeria and Tunisia). *COI* sequences have been retained in the geographical interval of −31.3 to 67.5° of longitude and 27.5–71.2° of latitude.

Time period and grain: *COI* sequences originate from studies published between 1998 and 2022 and from *de novo* sequencing of 2541 specimens done between 2007 and 2022.

Leonardo Dapporto and Mattia Menchetti contributed equally as first authors.

Vlad Dincă and Roger Vila contributed equally as senior authors.

de Catalunya, Grant/Award Number: CGL2016-76322 (AEI/FEDER, UE), PID2019-107078GB-I00; MCIN/AEI/10.13039/501100011033, Grant/Award Number: PID2020-117739GA-I00 and BES-2017-080641; Ministerio de Economía y Competitividad, Grant/Award Number: CGL2010-21226/BOS and CGL2013-48277-P; Ministero della transizione ecologica, Grant/Award Number: Direttiva Biodiversità

Handling Editor: Andres Baselga

Major taxa and level of measurement: Butterflies (Lepidoptera: Papilionoidea), analysed from individual to species level.

Software format: Data and functions to manage the dataset are provided in the iodatabase R package (https://github.com/leondap/iodatabase) and in Dryad (https://doi.org/10.5061/dryad.9w0vt4bjj).

KEYWORDS

COI sequences, DNA barcoding library, intraspecific genetic variation, Papilionoidea, phylogeography

1 | INTRODUCTION

The advent of phylogeography (Avise et al., 1987) sparked the commitment of mapping the fundamental layer of biodiversity encompassed by intraspecific genetic variation (IGV). Nevertheless, mapping IGV remained a vision for decades. Recently, faster and more affordable DNA sequencing has enabled a rapid accumulation of data for a larger proportion of biodiversity. Repositories, such as GenBank and BOLD, curate $>10^6$ cytochrome *c* oxidase subunit I (*COI*) sequences (the standard mitochondrial marker for animal DNA barcoding) (Hebert et al., 2003), in many cases complemented by high-quality metadata. As *COI* sequences accumulate, comparative macrogenetic analyses become possible (Leigh et al., 2021; Theodoridis et al., 2020).

Here, we have gathered and improved large-scale biodiversity surveys to generate an IGV atlas for Western Palaearctic butterflies. The primary resources we referred to are as follows: (1) the checklists of Western Palaearctic butterflies by Wiemers et al. (2018) and Middleton-Welling et al. (2020); and (2) a series of DNA-barcode libraries of Eurasiatic butterflies (Dapporto et al., 2019; Dincă et al., 2011, 2015, 2021; Hausmann et al., 2011; Huemer & Tarmann, 2016; Litman et al., 2018; Lukhtanov et al., 2009; Menchetti et al., 2021). We revised the checklists and added 2541 specimens sequenced *de novo*, including a comprehensive DNA-barcode library for Maghreb and Macaronesia. We provided *COI* sequences and IGV assessments for 532 species at the crossroads of three continents (Africa, Europe and Asia) and compiled a complete set of networks and distribution maps for haplotypes. To our knowledge, this is the first attempt to build a complete IGV atlas at continental and intercontinental levels. We hope that this resource will pilot similar studies on other clades and/or regions and will represent a milestone in the identification, description and protection of hidden diversity.

2 | METHODS

2.1 | Checklists and geographical ranges

We revised and updated a recent butterfly checklist for Europe (Wiemers et al., 2018) and one for Europe and Maghreb (Middleton-Welling et al., 2020) (Table S1). For the taxonomic revision, we accepted the conclusions of the most updated literature combining at least quantitative morphology and different genetic markers or based on genomic approaches (including double digest restriction-site associated DNA sequencing and phylogenomics) (see Table S2 and references therein). When evidence was insufficient, we refrained from applying changes, except in cases of deeply paraphyletic *COI* trees, which we treated as separate species.

We defined the Western Palaearctic "taxonomic area" (blue line in Figure 1) as follows. In the west, we included the Azores, Madeira and the Canary Islands. To delimit the fringes of the Sahara Desert, we set the southern limit in Morocco at 27.5°, followed the political border of this country to the north-east and set the southern boundary in Algeria and Tunisia at 33°. We included the European mainland and the islands under European administration in the Mediterranean, North and Baltic seas. No obvious faunistic boundaries exist to the east, and following previous checklists (Middleton-Welling et al., 2020; Wiemers et al., 2018) we included the western slopes of the Urals and excluded the Caucasus and western Kazakhstan. Given that any geographical boundary includes taxa at their distribution limit, we used a wider "dataset area" to provide context (Figure 1). Thus, we collected *COI* sequences of species in our checklist from the area included within the coordinates of −31.3° longitude (Flores, Azores), 27.5° latitude (El Hierro, the Canary Islands), 67.5° longitude (easternmost Urals) and 71.2° latitude (North Cape).

2.2 | Data acquisition, curation and quality control

We gathered *COI* sequences (DNA-barcode region between 400 and 658 bp) from three main sources: (1) *de novo* sequencing using standard procedures (deWaard et al., 2008) for most of the Maghreb and Macaronesia, but also improving coverage in Europe; (2) published DNA-barcode libraries compiled at the regional level (Dapporto et al., 2019; Dincă et al., 2011, 2015, 2021; Hausmann et al., 2011; Huemer & Tarmann, 2016; Litman et al., 2018; Menchetti et al., 2021) and other publicly available BOLD projects (e.g., LON-NorBOL for Norway); and (3) studies providing *COI* of single species or genera (references are available in Appendix S1). In the last case, we checked whether haplotypes, instead of specimens, were included in repositories (Paz-Vinas et al., 2021). If the number of specimens sharing a given haplotype in a particular location was reported

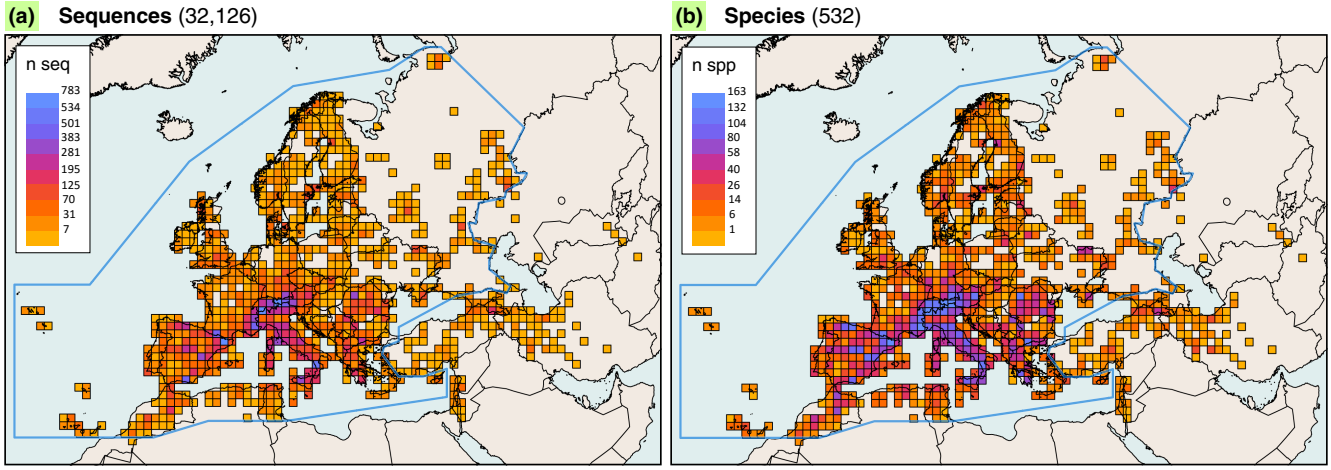

FIGURE 1 The study area with the representation of the "taxonomic area" (blue perimeter; Middleton-Welling et al., 2020) from where a complete species checklist has been assessed and updated. *COI* sequences of species also occurring outside the taxonomic area are included. (a) Number of sequences and (b) number of species for each 100 km × 100 km square cell. Variation in the number of DNA barcodes largely reflects butterfly richness across the taxonomic area.

(Ehl et al., 2021; Zinetti et al., 2013), we replicated the haplotype sequences to obtain data at the specimen level; otherwise, the sequences were excluded.

We verified species identifications by building neighbour-joining trees for each genus. When morphology could not be verified, we removed sequences not clustering within the species to which they were attributed if the mismatch did not involve one of the 74 species showing DNA-barcode sharing (updated from Dincă et al., 2021) (Table S3).

Coordinates (in decimal degrees) were retrieved from metadata and converted to the World Geodetic System 1984 (WGS84). When only an identifiable toponym was provided, we obtained approximate WGS84 coordinates at a resolution of 0.1° latitude and longitude. To obtain mainland areas of similar size, WGS84 coordinates were transformed into the Lambert azimuthal equal-area projection.

2.3 | Indexes of genetic variation

Based on *COI* sequences, we computed fundamental indexes of IGV. Initially, we obtained the number of haplotypes per species with the *haplotype* function of the *pegas* R package (Paradis, 2010). Then, for species with ≥10 sequences, we obtained the expected (asymptotic) number of haplotypes in the study area by rarefaction curves (D'Ercole et al., 2021; Dincă et al., 2021) using the *iNEXT* R package (Chao et al., 2014). Using the *hap.div* and *nuc.div* functions (*pegas*), we also calculated haplotype diversity (the probability that two sequences belong to the same haplotype) and nucleotide diversity (the average number of nucleotide differences per site between sequences). We calculated pairwise p-distances (the number of substitutions per site with pairwise deletion of missing sites) between sequences for each species and recorded the maximum p-distance among sequences.

The DST and GST fixation indexes (Meirmans & Hedrick, 2011; Nei, 1973) are defined as (HT − HS) and DST/HT, respectively, where HT is the mean p-distance among all sequences and HS is the mean p-distance among sequences within the same populations (defined below). DST varies between negative values and the maximum p-distance. Negative values (intrapopulation differentiation higher than average differentiation) often represent artefacts attributable to relatively small sample sizes and are usually set to zero (Meirmans & Hedrick, 2011), thus we applied this solution.

GST is a standardized dimensionless index ranging between negative values (set to zero) and one, with zero implying panmixia and one a complete spatial structuring. In species showing a single haplotype, DST is zero and GST is not defined. We also set GST to zero in these cases because, in the absence of any IGV, population structure can be argued to be absent.

Computing fixation indexes requires a division of specimens by populations, and the division affects the results. To minimize this effect, records were binned to different sets of areas identified by islands and by mainland grid cells of 100 km × 100 km, 200 km × 200 km, 300 km × 300 km, 400 km × 400 km and 500 km × 500 km. Furthermore, we replicated the five analyses by placing the centre of each cell on the vertex among cells in the previous division. We obtained 10 different partitions and, after computing DST and GST for each of them, we averaged their values to obtain mean indexes. Fixation indexes were computed for species having ≥10 specimens included in at least two areas with no less than three specimens each. This minimum requirement appeared to be sufficiently reliable in DST and GST computation (Scalercio et al., 2020).

Nucleotide diversity and GST are of particular interest because they express the average divergence among sequences and how much different haplotypes are segregated in different areas, respectively. For this reason, we plotted these values on bivariate bubble plots showing the distribution of GST and nucleotide diversity

Global Ecology and Biogeography A Journal of Macroecology

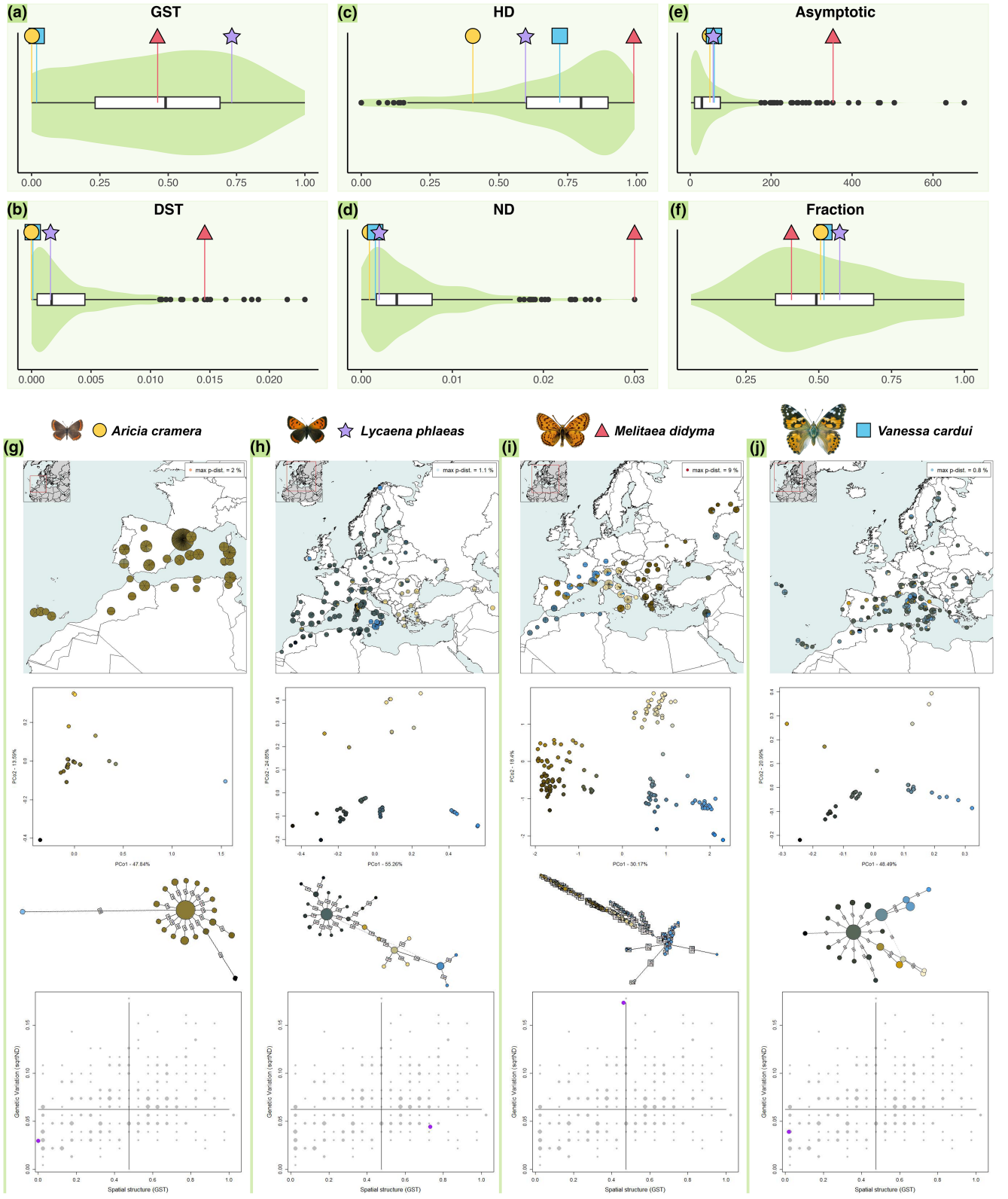

FIGURE 2 The distribution of intraspecific genetic variation (IGV) indexes: (a) GST; (b) DST; (c) haplotype diversity (HD); (d) nucleotide diversity (ND); (e) asymptotic (expected) number of haplotypes per species; (f) fraction of observed on expected number of haplotypes; and (g–j) four paradigmatic examples of the atlas representations with (from top to bottom): pie charts of sequenced specimens; the principal coordinates analysis (PCoA) in red–green–blue (RGB) space; haplotype networks, with colours for haplotypes as in pie charts, and scatterplots representing the GST and the square root-transformed values of nucleotide diversity with respect to all species in the dataset (grey bubbles with size proportional to the number of species in the intervals). The symbols for species in g–j are used to indicate their values in a–f.

(square root transformed) and their median values for all species in the dataset.

2.4 | Maps of genetic variation

We generated maps of haplotype distribution for each species with pie charts by attributing colours to sequences according to their p-distances. For each species, we subjected the p-distances matrix to a principal coordinates analysis (PCoA). We placed the two-dimensional configuration over a square having the colours yellow, blue, white and black at the corners and all possible shades in between, such that genetic similarity between two haplotypes was mirrored by colour similarity. The palette has been tested for perception by people with the most common types of colour blindness (Figure S4 in Appendix S1). The sequences were illustrated in pies grouped by areas of the same size; we used a variable cell size depending on the plotted area obtained by summing 0.3 to 1‰ of the latitudinal extension up to a maximum of 300 km × 300 km.

2.5 | Haplotype networks

For the construction of haplotype networks, we used sequences ≥600 bp. Haplotype networks were calculated with the randomized minimum spanning tree algorithm (Paradis, 2018) implemented in the *rmst* function of the R package *pegas*. Numbers of mutations among haplotypes were indicated using numbers over connections, and the haplotypes were assigned the colours obtained after PCoA.

2.6 | Script availability

The updated species checklist, *COI* sequences, metadata, the atlas and R scripts are available in Dryad (https://doi.org/10.5061/dryad.9w0vt4bjj), in the iodatabase R package designed to store and manage the dataset (https://github.com/leondap/iodatabase) and in the DS-ATLAS (https://doi.org/10.5883/DS-ATLAS), a BOLD dataset where the newly sequenced specimens are released publicly. The atlas is also available in Appendix S1.

3 | RESULTS AND DISCUSSION

We updated the latest butterfly checklists, adding 16 species and excluding four taxa recently recognized as subspecies (Table S1). Based on this checklist (Table S2), comprising 552 species present in the "taxonomic area", we obtained 32,126 geo-referenced *COI* sequences ≥400 bp from specimens within the "dataset area", including 2541 previously unpublished sequences. A total of 532 species (96.4% of the checklist) were represented, with a mean of 58.2 (±88.0 *SD*) sequences per species.

The mean number of haplotypes per species was 16.2 (±20.8 *SD*) (for IGV indexes, see Table S2). As observed by Dincă et al. (2021), the number of haplotypes detected was often much lower than the asymptotic value (mean ± *SD* completeness 52.8 ± 23.6%), owing to the occurrence of few common haplotypes and many rare ones, frequently differentiated by only one or few mutations. Accordingly, haplotype diversity was skewed towards high values (many haplotypes), whereas nucleotide diversity was skewed towards low values (low divergences) (Figure 2).

GST showed an almost normal distribution, whereas DST was skewed towards low differentiation. The species showing low GST are characterized by very low haplotype and nucleotide diversity or by relatively high haplotype and nucleotide diversity with almost complete panmixia (Figure 2a–d). Conversely, some species with low haplotype diversity showed high GST when haplotypes exhibited strong spatial structure (Figure 2h).

The maps, which are provided in a supplementary book available as a main resource of this study (Appendix S1), revealed a variety of patterns. A strong divergence between continental areas (the Maghreb, Europe and the Middle East) emerged as a major pattern (e.g., *Aporia crataegi*, *Melitaea cinxia* and *Chazara briseis*). Many species (e.g., *Thymelicus lineola*, *Pyrgus serratulae* and *Lasiommata maera*) also showed differentiation among glacial refugia in southern Europe (Iberian, Italian and Balkan peninsulas), in line with the paradigm of the genetic legacy of ice ages on European biota (Hewitt, 2000; Taberlet et al., 1998).

This dataset facilitates in-depth comparative assessments by exploring general trends in distribution patterns, although this lies beyond the aim of this paper. Such an approach is essential to avoid potential cases of pseudocongruence that occur when researchers attempt to categorize patterns visually (Soltis et al., 2006). It also represents a tool for biologists studying the butterfly species covered, because it can be a great help in the planning of sampling intended to describe IGV further, allows for testing potential eco-evolutionary correlates, and highlights potential cryptic species or other evolutionary phenomena (e.g., introgression and genetic sweeps) to be studied in greater detail. Although mitochondrial IGV can have limited power in the inference of demographic processes and encompass only female patterns (Schmidt & Garroway, 2021), indexes of mitochondrial DNA variation are used as proxies for dispersal capacity (e.g., Melero et al., 2022). When enough data for nuclear genomes become available, it will be possible to clarify whether spatial patterns reflected in both mitochondrial and nuclear genomes are generally comparable. Finally, the allopatric/parapatric haplogroups documented here could help in defining subspecific taxonomy and can be viewed as management units for conservation.

ACKNOWLEDGMENTS

We thank the many lepidopterists, too numerous to be mentioned individually, who provided samples. We thank Bernat Burriel-Carranza for the help provided with the R scripts. Open Access

Funding provided by Universita degli Studi di Firenze within the CRUI-CARE Agreement.

FUNDING INFORMATION

Support for this research was provided by the Academy of Finland (Academy Research Fellow, decision no. 328895) and by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (project no. 625997) to V.D., by "la Caixa" Foundation (ID 100010434) to M.M. (grant LCF/BQ/DR20/11790020), and by projects CGL2010-21226/BOS and CGL2013-48277-P (Spanish Ministerio de Economía y Competitividad), CGL2016-76322 (AEI/FEDER, UE), PID2019-107078GB-I00 (MCIN/AEI/10.13039/501100011033) and 2017-SGR-991 (Generalitat de Catalunya) to R.Vi., the grant BES-2017-080641 funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future" to J.C.H., the project PID2020-117739GA-I00 (MCIN/AEI/10.13039/501100011033) to G.T., and by the Direttiva Biodiversità 2019 and 2020 projects (Ministero della Transizione Ecologica) to L.D.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The R scripts and *COI* sequences are available in Dryad (https://doi.org/10.5061/dryad.9w0vt4bjj) and in the iodatabase GitHub page (https://github.com/leondap/iodatabase), where the FASTA files and the metadata comprising BOLD process IDs and GenBank accession numbers can be found. An R package (iodatabase) can also be installed from Github (https://github.com/leondap/iodatabase), where a detailed description of the new functions is provided. The atlas of mitochondrial genetic diversity for Western Palaearctic butterflies is available as a PDF file in the same repositories and as Appendix S1.

ORCID

Leonardo Dapporto https://orcid.org/0000-0001-7129-4526
Mattia Menchetti https://orcid.org/0000-0002-0707-7495
Raluca Vodă https://orcid.org/0000-0003-1970-5628
Joan C. Hinojosa https://orcid.org/0000-0002-6318-4252
Nok Ting Lam https://orcid.org/0000-0003-4774-9142
Gerard Talavera https://orcid.org/0000-0003-1112-1345
Vlad Dincă https://orcid.org/0000-0003-1791-2148
Roger Vila https://orcid.org/0000-0002-2447-4388

REFERENCES

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, *18*, 489–522.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*, 45–67.

Dapporto, L., Cini, A., Vodă, R., Dincă, V., Wiemers, M., Menchetti, M., Magini, G., Talavera, G., Shreeve, T., Bonelli, S., Casacci, L. P., Balletto, E., Scalercio, S., & Vila, R. (2019). Integrating three comprehensive data sets shows that mitochondrial DNA variation is linked to species traits and paleogeographic events in European butterflies. *Molecular Ecology Resources*, *19*, 1623–1636. https://doi.org/10.1111/1755-0998.13059

D'Ercole, J., Dincă, V., Opler, P. A., Kondla, N., Schmidt, C., Phillips, J. D., Robbins, R., Burns, J. M., Miller, S. E., Grishin, N., & Zakharov, E. V. (2021). A DNA barcode library for the butterflies of North America. *PeerJ*, *9*, e11157.

deWaard, J. R., Ivanova, N. V., Hajibabaei, M., & Hebert, P. D. N. (2008). Assembling DNA barcodes. In *Environmental genomics* (pp. 275–294). Springer.

Dincă, V., Dapporto, L., Somervuo, P., Vodă, R., Cuvelier, S., Gascoigne-Pees, M., Huemer, P., Mutanen, M., Hebert, P. D. N., & Vila, R. (2021). High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. *Communications Biology*, *4*, 1–11.

Dincă, V., Montagud, S., Talavera, G., Hernández-Roldán, J., Munguira, M. L., García-Barros, E., Hebert, P. D. N., & Vila, R. (2015). DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Scientific Reports*, *5*, 12395.

Dincă, V., Zakharov, E. V., Hebert, P. D. N., & Vila, R. (2011). Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 347–355.

Ehl, S., Ehl, S., Kramp, K., & Schmitt, T. (2021). Interglacials are driving speciation and intraspecific differentiation in the cold-adapted butterfly species group Boloria pales/napaea (Nymphalidae). *Journal of Biogeography*, *48*, 134–146.

Hausmann, A., Haszprunar, G., Segerer, A. H., Speidel, W., Behounek, G., & Hebert, P. D. (2011). Now DNA-barcoded: The butterflies and larger moths of Germany. *Spixiana*, *34*, 47–58.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B Biological Sciences*, *270*, 313–321.

Hewitt, G. (2000). The genetic legacy of the quaternary ice ages. *Nature*, *405*, 907–913.

Huemer, P., & Tarmann, G. (2016). DNA-Barcoding von Schmetterlingen Österreichs–vom nationalen Fokus zur Internationalität. *Acta ZooBot Austria*, *153*, 163–167.

Leigh, D. M., van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics*, *22*, 1–17.

Litman, J., Chittaro, Y., Birrer, S., Praz, C., Wermeille, E., Fluri, M., Stalling, T., Schmid, S., Wyler, S., & Gonseth, Y. (2018). A DNA barcode reference library for Swiss butterflies and forester moths as a tool for species identification, systematics and conservation. *PLoS One*, *13*, e0208639.

Lukhtanov, V. A., Sourakov, A., Zakharov, E. V., & Hebert, P. D. N. (2009). DNA barcoding central Asian butterflies: Increasing geographical dimension does not significantly reduce the success of species identification. *Molecular Ecology Resources*, *9*, 1302–1310.

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*, 5–18.

Melero, Y., Evans, L. C., Kuussaari, M., Schmucki, R., Stefanescu, C., Roy, D. B., & Oliver, T. H. (2022). Local adaptation to climate anomalies relates to species phylogeny. *Communications Biology*, *5*, 1–9.

Menchetti, M., Talavera, G., Cini, A., Salvati, V., Dincă, V., Platania, L., Bonelli, S., Balletto, E., Vila, R., & Dapporto, L. (2021). Two ways to be endemic. Alps and Apennines are different functional refugia

during climatic cycles. *Molecular Ecology*, *30*, 1297–1310. https://doi.org/10.1111/mec.15795

Middleton-Welling, J., Dapporto, L., García-Barros, E., Wiemers, M., Nowicki, P., Plazio, E., Bonelli, S., Zaccagno, M., Šašić, M., Liparova, J., Schweiger, O., Harpke, A., Musche, M., Settele, J., Schmucki, R., & Shreeve, T. (2020). A new comprehensive trait database of European and Maghreb butterflies, Papilionoidea. *Scientific Data*, *7*, 351. https://doi.org/10.1038/s41597-020-00697-7

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, *70*, 3321–3323.

Paradis, E. (2010). Pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419–420.

Paradis, E. (2018). Analysis of haplotype networks: The randomized minimum spanning tree method. *Methods in Ecology and Evolution*, *9*, 1308–1317.

Paz-Vinas, I., Jensen, E. L., Bertola, L. D., Breed, M. F., Hand, B. K., Hunter, M. E., Kershaw, F., Leigh, D. M., Luikart, G., Mergeay, J., Miller, J. M., van Rees, C. B., Segelbacher, G., & Hoban, S. (2021). Macrogenetic studies must not ignore limitations of genetic markers and scale. *Ecology Letters*, *24*, 1282–1284.

Scalercio, S., Cini, A., Menchetti, M., Vodă, R., Bonelli, S., Bordoni, A., Casacci, L. P., Dincă, V., Balletto, E., Vila, R., & Dapporto, L. (2020). How long is 3 km for a butterfly? Ecological constraints and functional traits explain high mitochondrial genetic diversity between Sicily and the Italian peninsula. *The Journal of Animal Ecology*, *89*, 2013–2026. https://doi.org/10.1111/1365-2656.13196

Schmidt, C., & Garroway, C. J. (2021). The conservation utility of mitochondrial genetic diversity in macrogenetic research. *Conservation Genetics*, *22*, 323–327.

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*, 4261–4293.

Taberlet, P., Fumagalli, L., Wust-Saucy, A., & Cosson, J. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*, 453–464.

Theodoridis, S., Fordham, D. A., Brown, S. C., Li, S., Rahbek, C., & Nogues-Bravo, D. (2020). Evolutionary history and past climate change shape the distribution of genetic diversity in terrestrial mammals. *Nature Communications*, *11*, 1–11.

Wiemers, M., Balletto, E., Dincă, V., Fric, Z. F., Lamas, G., Lukhtanov, V., Munguira, M. L., van Swaay, C. A., Vila, R., Vliegenthart, A., & Wahlberg, N. (2018). An updated checklist of the European butterflies (Lepidoptera, Papilionoidea). *Zookeys*, *2018*, 9–45. https://doi.org/10.3897/zookeys.811.28712

Zinetti, F., Dapporto, L., Vovlas, A., Chelazzi, G., Bonelli, S., Balletto, E., & Ciofi, C. (2013). When the rule becomes the exception. No evidence of gene flow between two *Zerynthia* cryptic butterflies suggests the emergence of a new model group. *PLoS One*, *8*, e65746. https://doi.org/10.1371/journal.pone.0065746

BIOSKETCH

Members of the research team are actively engaged in insect biogeography, systematics and conservation, with particular interest in analysing large-scale data generated by occurrence records and DNA sequencing. The main aim of their research is unravelling the historical and present-day factors responsible for species distributions across the Western Palearctic region.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dapporto, L., Menchetti, M., Vodă, R., Corbella, C., Cuvelier, S., Djemadi, I., Gascoigne-Pees, M., Hinojosa, J. C., Lam, N. T., Serracanta, M., Talavera, G., Dincă, V., & Vila, R. (2022). The atlas of mitochondrial genetic diversity for Western Palaearctic butterflies. *Global Ecology and Biogeography*, *31*, 2184–2190. https://doi.org/10.1111/geb.13579