BRILL

Amphibia-Reptilia 44 (2023): 399-413

brill.com/amre

# COI barcoding provides reliable species identification and pinpoints cryptic diversity in Western Palearctic amphibians

Guillermo Velo-Antón[1,2,3,*], E. Anne Chambers[4,5], Nikolay A. Poyarkov Jr.[6], Daniele Canestrelli[7], Roberta Bisconti[7], Borislav Naumov[8], María José Fernández Benéitez[9], Alex Borisenko[10], Iñigo Martínez-Solano[9,11,*]

1 - Universidad de Vigo, Facultad de Biología, Edificio de Ciencias Experimentales, Bloque B, Planta 2, Laboratorio 39 (Grupo GEA), E-36310, Vigo, Spain

2 - CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

3 - BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal

4 - Department of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley, CA 94720, USA

5 - Museum of Vertebrate Zoology, University of California Berkeley, Berkeley, CA 94720, USA

6 - Faculty of Biology, Department of Vertebrate Zoology, Moscow State University, Moscow, 119234, Russia

7 - Department of Ecological and Biological Sciences, University of Tuscia, 01100 Viterbo, Italy

8 - Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 2 Gagarin Street, 1113 Sofia, Bulgaria

9 - Instituto de Investigación en Recursos Cinegéticos (IREC-CSIC-UCLM-JCCM), Ronda de Toledo, s/n, 13071 Ciudad Real, Spain

10 - Canadian Centre for DNA Barcoding, Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, N1G 2W1, Canada

11 - Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain

*Corresponding authors; e-mails: guillermo.velo@uvigo.es; inigomsolano@mncn.csic.es

ORCID iDs: Velo-Antón: 0000-0002-9483-5695; Chambers: 0000-0002-7369-0108; Poyarkov Jr.: 0000-0002-7576-2283; Canestrelli: 0000-0001-9351-4972; Bisconti: 0000-0002-0600-7436; Naumov: 0000-0003-2146-208X; Benéitez: 0000-0003-3797-4805; Borisenko: 0000-0002-3061-3057; Martínez-Solano: 0000-0002-2260-226X

Received 3 October 2022; final revision received 15 March 2023; accepted 19 June 2023; published online 11 July 2023; published in print 31 November 2023

Associate Editor: Ariel Rodríguez

**Abstract.** Assembling DNA barcode reference libraries for various taxonomic groups allows researchers to use metabarcoding or environmental DNA approaches to gain a rapid understanding of diversity in given environments. However, our ability to use reference libraries depends on how accurately DNA barcodes are able to recover taxonomic boundaries and identify species, which is rarely considered. We constructed an extensive COI barcoding library for amphibians of the Western Palearctic and successfully recovered barcodes from 60 urodele and 73 anuran species (representing 94% and 98% of the nominal anuran and urodele species in the Western Palearctic, respectively), covering the intraspecific diversity of the majority of species in this region. We tested the effectiveness of our assembled DNA barcode dataset for species identification using barcoding gap, efficiency analyses, and two phylogenetic species delimitation methods. We obtained DNA barcodes for 1251 specimens (691 anurans and 560 urodeles) with a high success rate (92-96%) of species identification. The absence of a barcoding gap in a number of samples was linked to species misidentifications, which suggest incipient speciation or cryptic diversity, or previously described mitochondrial introgression events. The phylogenetic species delimitation methods

Published with license by Koninklijke Brill NV | DOI: 10.1163/15685381-bja10148
 | ISSN: 0173-5373 (print) 1568-5381 (online)

resulted in substantial oversplitting of currently accepted taxonomy. This *COI* barcoding library provides an almost complete and reliable reference library for Western Palearctic amphibians. We highlight the importance of generating comprehensive and well curated reference libraries that include intra- and interspecific genetic variability and the need of detailed taxonomic revision when ambiguous or incorrect DNA barcodes exist.

*Keywords*: barcoding gap, DNA barcoding, mitochondrial DNA, reference library, species delimitation, specimen identification.

## Introduction

The development of next-generation sequencing technology and the subsequent emergence of environmental DNA (eDNA) metabarcoding approaches have provided novel and powerful toolboxes for ecologists (Taberlet et al., 2018). Detecting species in a given environment using metabarcoding is now possible by obtaining DNA barcodes for samples from a wide variety of different sources, such as water, sediment, or feces (Taberlet et al., 2012), with a broad range of applications, including ecosystem biomonitoring (Stat et al., 2017), the identification of invasive species (Ficetola et al., 2008), paleoecological studies (Bálint et al., 2018), and dietary analyses (e.g., De Barba et al., 2014; Marques et al., 2022), among others. However, a major limitation of these DNA barcode-based methods is the need for curated reference databases, which are crucial for successful species detection and discovery.

DNA barcoding (Hebert et al., 2003) is rapidly progressing towards completing a global DNA barcode database to help identify and characterize diversity in a wide variety of taxonomic groups. DNA barcoding studies targeting groups with robust taxonomy typically sequence DNA barcodes from vouchered specimens, which have been reliably identified to the species level typically using morphological characters (Stoeckle and Hebert, 2008). Such studies then assemble comprehensive DNA barcode reference libraries, which are crucial for accurate species identification, as well as for recognizing undescribed species. Several taxonomic initiatives within the DNA barcoding framework have been established to assemble reference libraries for specific taxonomic groups (e.g., fungi, http://www.fungalbarcoding.org/; insects, http://lepbarcoding.org/; fishes, http://www.fishbol.org/; mammals; or marine organisms, http://www.marinebarcoding.org/) and numerous others are in progress.

Amphibians are an old (extant families date back to the mid-Cretaceous, 110-120 million years ago; Wake and Koo, 2018), and morphologically conserved group (often due to homoplasy; Wake, Wake and Specht, 2011) of vertebrates. Traditional species identification based on diagnostic morphological characters is not always straightforward, especially at early developmental stages, making it a particularly challenging vertebrate group for taxonomists and ecologists, but also when ranges of morphologically similar, closely related taxa overlap, or when species identification relies solely on characters present only in adults. Thus, genetic information is critical for species identification in amphibians and some species are only reliably identified using molecular methods. Exploring global diversity is particularly pertinent for amphibians, as they are the most threatened vertebrate group on Earth, with approximately 40 percent of species being of conservation concern (Stuart et al., 2004). Thus, now more than ever before, understanding amphibian diversity across the globe using DNA barcodes is critical.

Previous studies have demonstrated the efficiency of DNA barcoding and its applicability for identifying invasive species (Ficetola et al., 2008), cryptic lineages (Smith, Poyarkov and Hebert, 2008; Crawford et al., 2013), and in eDNA studies (Lopes et al., 2017; Peixoto et al., 2020; 2023). While early amphibian DNA barcoding studies used a mitochondrial fragment of the *16S* gene as a reference to estimate diversity (e.g., Vences et al., 2005), as DNA barcoding

using *COI* spread, other initiatives such as Cold Code (Murphy et al., 2013) were later launched to DNA barcode amphibians and reptiles using a different region of the mitochondrial genome, a fragment of the cytochrome oxidase subunit I (*COI*) gene, which is the marker of choice across many other taxonomic groups. The use of the *COI* region as a standard barcode marker was initially debated due to the lower amplification success and apparent lack of a clear "barcoding gap" in amphibians (Vences et al., 2005; Smith et al., 2008). However, the subsequent development of amphibian-specific *COI* primers (Che et al., 2012; Xia et al., 2012), as well as the higher genetic variability in *COI* sequences compared to other mtDNA markers (Xia et al., 2012; Jeong et al., 2013; Grosjean et al., 2015; Deichmann et al., 2017), has led to the general consensus that efforts must be directed towards building a reference *COI* barcode database for amphibians (Murphy et al., 2013; Chambers and Hebert, 2016). While *COI* has been used in some phylogenetic and phylogeographic studies in amphibians (e.g., Recuero et al., 2007; Smith et al., 2008; Ahmadzadeh et al., 2020), the effectiveness of *COI* as a DNA barcode marker for species identification in this group has not been thoroughly assessed. In DNA barcoding studies, a useful preliminary investigation into the reliability of DNA barcodes for species identification is the establishment of the barcoding gap, which is the gap between maximum intraspecific and minimum interspecific genetic distances within a group of organisms (e.g., Hebert et al., 2004), the assumption being that interspecific divergence should exceed intraspecific divergence by an order of magnitude. However, a proper evaluation of whether a reliable barcoding gap exists in a given taxonomic group requires a thorough examination of both interspecific and intraspecific variation.

Although amphibian DNA barcoding reference libraries have been assembled for large geographic regions including Central America (Crawford et al., 2013), East Asia (Che et al., 2012; Xia et al., 2012; Jeong et al., 2013), Madagascar (Perl et al., 2014), and Africa (Deichmann et al., 2017; Sampaio et al., 2021), there are still substantial geographic gaps across the globe where the lack of curated DNA barcode reference databases impedes the application of eDNA and metabarcoding approaches in amphibian ecology and conservation. The Western Palearctic (WP) is one of the eight global biogeographic realms, comprising Europe, North Africa, and parts of the Arabian Peninsula and temperate Asia, and is home to 138 species of amphibians in 34 genera (61 urodeles and 77 anurans; supplementary table S1). Amphibian taxonomy in this region has been thoroughly studied, and for many species, patterns of intraspecific diversity have also been investigated in detail in phylogeographic studies, identifying major mitochondrial lineages within species (e.g., Dinis et al., 2019; Dufresnes et al., 2020). However, an extensive, curated reference *COI* barcode database for Western Palearctic amphibians is currently lacking.

Creating a comprehensive *COI* reference database will help with preliminary identification of species, morphologically cryptic lineages, and specimens in scientific collections, as well as forensic applications (Saats et al., 2016). Here, we aim to: i) construct a comprehensive *COI* reference database for the amphibians of the Western Palearctic, thereby providing a solid basis for species identification for amphibians in this region; ii) evaluate the use of the *COI* barcoding gene region in species delimitation methods to identify mitochondrial lineages that might constitute cryptic species, and iii) investigate the existence of a barcoding gap in urodeles (Caudata dataset) and anurans (Anura dataset).

## Materials and methods

### *Samples and species coverage*

For the purposes of this study, we defined the Western Palearctic region as depicted in fig. 1, based on previous studies (Duellman and Trueb, 1986; Borkin, 1999; Borkin

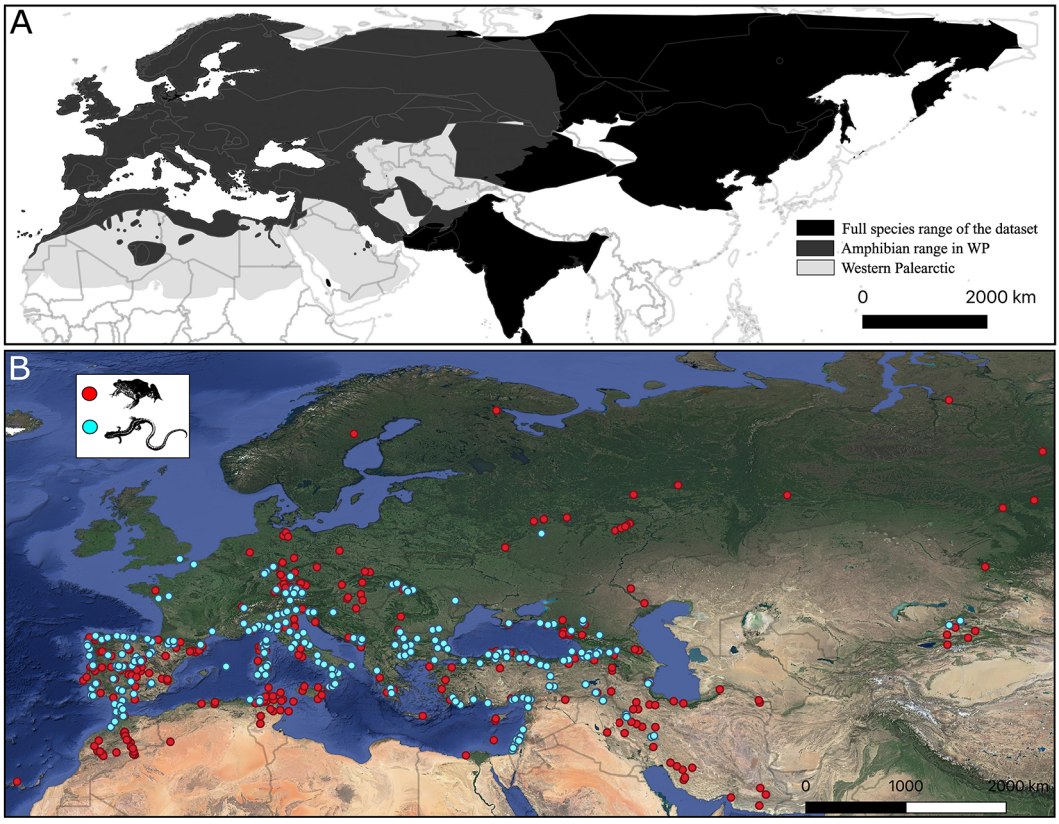


**Figure 1.** The distribution of samples for Anura and Caudata datasets included in this study with the limits of the Western Palearctic region and georeferenced samples indicated.

and Litvinchuk, 2013, 2014). We compiled a total of 1251 sequences, which are organised in BOLD projects DS-AMWPA and DS-AMWPC at www.boldsystems.org for Anura and Caudata, respectively, and includes 60 urodele species and 73 anuran species (belonging to 18 and 16 genera, respectively; supplementary table S1). DS-AMWPA includes 691 sequences, of which 525 (76%) were generated for this study and previously organised in separate BOLD subprojects managed by coauthors on this manuscript (ABAAP, ABCAP, AMWP, IMSAM, ZHABI), as well as 166 sequences (24%) mined from GenBank and other BOLD projects (GBAP, FBHER; supplementary table S2). DS-AMWPC includes 560 sequences, of which 452 (81%) were generated by our team and previously organised in separate BOLD subprojects (ABCAP, AMWP, IMSAM, ZHABI, CABM), as well as 108 sequences (19%) mined from GenBank and other BOLD projects (GBAP, FBHER; supplementary table S3). New sequences were obtained through sequenced DNA barcodes from the collection of tissue samples from different field projects and scientific collections. From the sequences mined from GenBank, we removed those sequences that clearly represent taxonomic misidentifications. Our sampling paid particular focus to the southern climatic refugia in the Western Palearctic (southern European peninsulas and Mediterranean ecosystems in North Africa and the Near East; fig. 1), which constitute biodiversity hotspots for this taxonomic group and in which most of the intraspecific diversity occurs. The taxonomic framework used followed Speybroeck et al. (2020) as a reference. For further details on species ranges and samples used in this reference library, see supplemental material.

*DNA extraction, amplification, and sequencing*

DNA was extracted, amplified, and sequenced using standard DNA barcoding protocols and primers (see details in BOLD datasets DS-AMWPA and DS-AMWPC; Ivanova, DeWaard and Hebert, 2006; Smith et al., 2008; Vences et al., 2012). Sequences were aligned by eye in Mesquite v3.61 (Maddison and Maddison, 2019) and checked for internal stop codons indicative of low-quality reads or pseudogenes; these sequences were subsequently removed from the dataset. BOLD datasets are publicly available at www.boldsystems.org.

*Gene tree reconstruction*

We used BEAST v1.10 (Suchard et al., 2018) to reconstruct two separate gene trees for the Anura and Caudata datasets. We selected the best nucleotide substitution models for

these datasets using jModeltest v2 (Guindon and Gascuel, 2003; Darriba et al., 2012) and ran analyses under the Yule coalescent model, assuming a strict molecular clock and fixing the substitution rate to 1. Chains were run for 20 million generations, and 10 000 trees were sampled after an initial burn-in of 10%. We ensured consistency of results and adequate effective sample sizes to estimate parameters of interest using Tracer version 1.7.1 (Rambaut et al., 2018). A maximum clade credibility consensus tree for each taxonomic group (Anura and Caudata) was then constructed for subsequent analyses with TreeAnnotator v1.10 (distributed as part of the BEAST software).

*Species delimitation*

To identify mitochondrial lineages, or MOTUs (Molecular Operational Taxonomic Units), we selected two methods developed for automated species delineation that have been used successfully in other taxonomic groups (e.g., Monaghan et al., 2009): the Generalized Mixed Yule Coalescent (GMYC) model (Fujisawa and Barraclough, 2013; Pons et al., 2006), and the Poisson tree processes (PTP) model (Zhang et al., 2013). Both methods use single genes but differ in the use of the phylogenetic trees to identify putative species. The GMYC relies on the expected abrupt change in branching events between a pure birth (among-species branching events) and a neutral coalescent process (within-species branching events). This method identifies the transition between inter- and intra-species branching rates using a time-calibrated ultrametric tree by maximizing the likelihood score of the model. The PTP, by contrast, infers putative species boundaries on a given phylogenetic input tree relying on the branch lengths that represent the mean expected number of substitutions per site between two branching events, assuming that the number of substitutions between species is significantly higher than the number of substitutions within species. As such, ultrametric trees are only required for the GMYC method, as the speciation model requires time-calibrated phylogenies, while the PTP method models speciation rates directly by using the number of substitutions (Zhang et al., 2013).

We ran the GMYC analysis for each dataset using the R package *SPLITS* (Ezard, Fujisawa and Barraclough, 2009). We selected the single-threshold GMYC model because the multiple-threshold GMYC, which allows for several thresholds, results in a smaller number of "true" species correctly identified (Fujisawa and Barraclough, 2013). We then used the multi-rate Poisson Tree Processes (mPTP) approach (Kapli et al., 2017), which fits branch lengths of each putative species into distinct exponential distributions (speciation and coalescent). We ran the mPTP analysis with the default parameters using the webserver (https://mptp.h-its.org/#/tree).

*Barcoding gap*

Identifying the existence of a barcoding gap (Meyer and Paulay, 2005) is important to assess which species identification method may be most appropriate for the dataset in question. However, it is difficult to determine the existence of a barcoding gap when frequency distributions of both intra- and interspecific divergences for pooled species are constructed, as they could represent either specimen identification or species discovery (Collins and Cruickshank, 2013). A more accurate representation of the barcoding gap involves a histogram in which, for each individual, the distance to the furthest conspecific is plotted against the distance to the nearest non-conspecific (Meyer and Paulay, 2005) using the smallest, rather than the mean interspecific distances (Meier, Zhang and Ali, 2008). To evaluate the presence of a barcoding gap in our two datasets, we first calculated a pairwise distance matrix between sequences using the Kimura 2-parameter model (K2P) (Kimura, 1980) and then used the statistics *maxInDist* (furthest intraspecific distance) and *nonConDist* (smallest interspecific distance) implemented in the R package *spider* v.1.3 (Brown et al., 2012).

*Specimen identification*

To test how effective barcodes were at identifying samples to the species level (barcoding efficiency) in our datasets, we used two query-based identification analyses that determine genetic distance thresholds, both of which are implemented in the *spider* package in R. First, we determined the BOLD identification criterion (hereafter, referred to as BOLDi), which mimics the "species identification" method used by BOLD. This method uses a threshold-based criterion considering all specimens within the threshold of the query. For this, we used the threshID() function as implemented in the R package *spider* (Brown et al., 2012). Second, we used Meier's best close match function (Meier's BCM, Meier et al., 2006). Meier's BCM is another distance-based analysis that identifies the best barcode match of a query based on direct sequence comparison and a given threshold, but only assigns the species name of the nearest-neighbour match, rather than all matches within the threshold (as with threshID()); to calculate this value we used the bestCloseMatch() function, also implemented in *spider*. Both BOLDi and Meier's BCM identify whether a sequence corresponds to a: i) "correct match" (when the name of the closest match is the same than the specimen considered), ii) "incorrect match" (when the name of the closest match is different from the specimen considered), iii) "ambiguous match" (when there are more than one equally close matches of different species), or (iv) "no identification", if no species is identified within the given threshold. To better explore specimen identification based on several threshold values and not simply the default threshold of 1%, we used a range of threshold values (0.5%-2.5%). Preliminary results indicated that 1.1% and 2.1% were the threshold values that minimized the cumulative error (false negative + false positive) for Anura and Caudata, respectively (supplementary fig. S1), and thus we applied the corresponding optimized threshold value criterion to both BOLDi and Meier's BCM and each dataset. Since rampant gene flow between the *Pelophylax ridibundus* and *P. lessonae* (Dubey et al., 2019) would confound specimen identification using DNA barcodes, we removed these two species, and their hybrids,

from the barcoding efficiency analyses; their correct identification must rely on multilocus datasets. Species represented by only one individual (singletons) were removed from these analyses, which reduced the Anura and Caudata datasets to 64 and 48 species, respectively.

## Results

### *DNA barcode reference libraries*

We compiled *COI* barcodes for 1251 specimens. The Anura dataset (BOLD ID: DS-AMWPA) comprised 691 sequences (with lengths ranging from 105-654 bp; mean length = 579 bp) belonging to 73 species in 16 genera, while the Caudata dataset (BOLD ID: DS-AMWPC) contained 560 sequences (lengths ranging from 255-654 bp; mean length = 612 bp) belonging to 60 species in 18 genera. Within the Anura dataset, 54.0% of barcodes were of a length considered barcode compliant (>487 bp), while the Caudata dataset contained 58.4% barcode compliant length sequences. A total of 11 anuran and 14 urodele species were represented by singletons.

### *Gene tree reconstruction*

We obtained well-resolved gene trees for Anura and Caudata, and TN93 + I + G was the best nucleotide substitution model for both groups. Most nodes had Bayesian posterior probabilities greater than 95%, with the exception of some of the basal nodes (figs. 2 and 3). Overall, the gene tree relationships that were recovered were largely consistent with previous phylogenetic

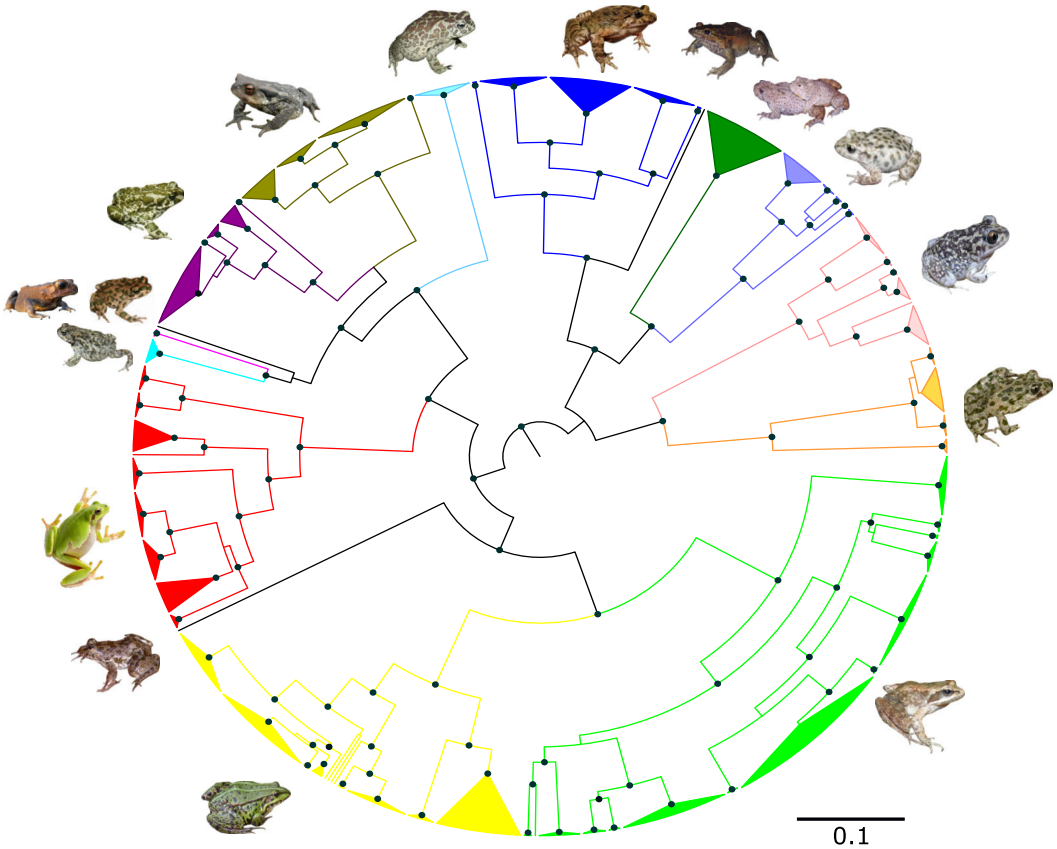


**Figure 2.** Gene tree constructed using the DNA barcode fragment of the *COI* mitochondrial gene tree for Western Palearctic anurans, estimated under a Bayesian framework. Nodes are collapsed at the genus level and colours correspond to genera; images are for one representative species for each genus. Black dots on nodes indicate posterior probabilities > 0.90.

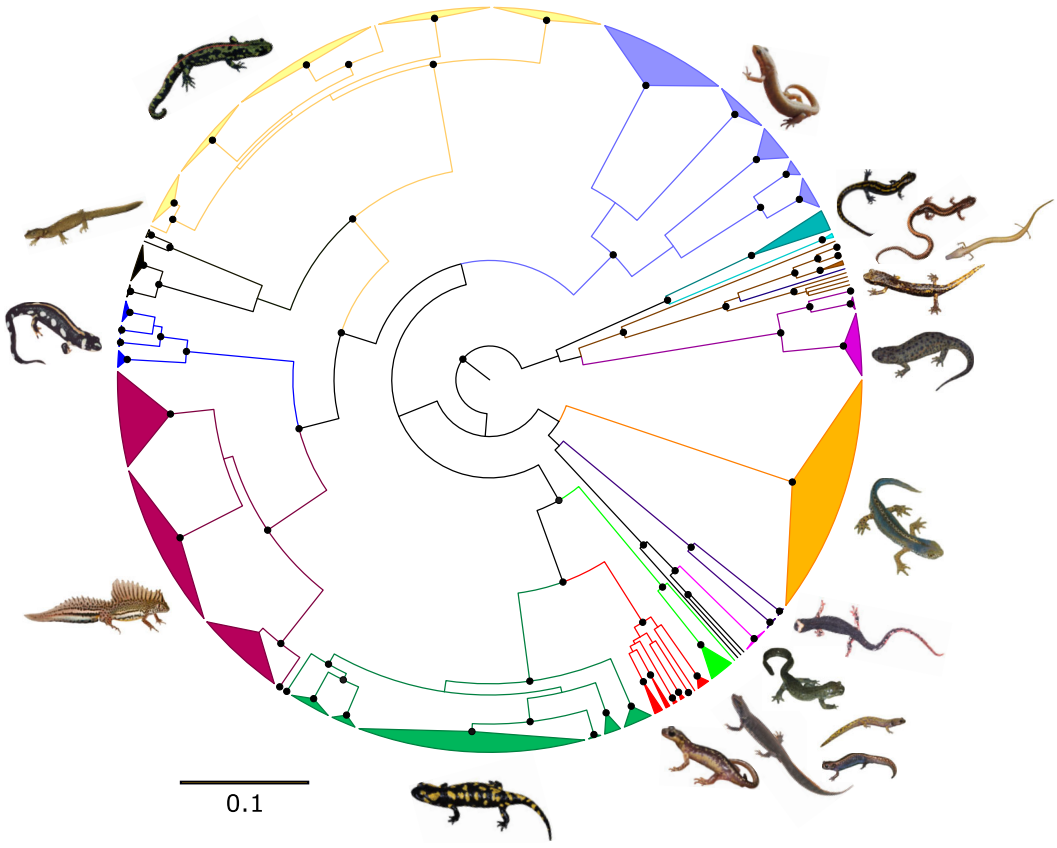


**Figure 3.** Gene tree constructed using the DNA barcode fragment of the *COI* mitochondrial gene tree for Western Palearctic urodeles, estimated under a Bayesian framework. Nodes are collapsed at the genus level and colours correspond to genera; images are a representative species for each genus. Black dots on nodes indicate posterior probabilities > 0.90.

studies and with currently accepted taxonomy (Speybroeck et al., 2020). Exceptions included recently diverged species (*Alytes almogavarii* nested within *A. obstetricans*, *Pelodytes atlanticus* nested within *P. punctatus*), and species complexes pending detailed taxonomic revision (*Pelophylax*, which includes several hybrid taxa that make mtDNA identification unreliable, and *Lissotriton vulgaris*).

*Species delimitation*

The number of putative species determined using both methods (GMYC and mPTP) exceeded the species list based on current taxonomy in both datasets. Within Anura, GMYC recovered 102 hypothesized species (Confidence Interval [C.I.]: 91-110; supplementary table S4), while mPTP results identified 108 tentative species (supplementary table S5). Results from GMYC and mPTP thus represent an increase in species diversity of 32% and 40% from current taxonomy, respectively. Within Caudata, GMYC identified 101 hypothesized species (C.I.: 91-108; supplementary table S6), while mPTP hypothesized 87 (supplementary table S7), which represents an increase of 60% and 38% from current taxonomy for GMYC and mPTP, respectively. The majority of the discordance between current taxonomy and results from the two PSC-based delimitation methods resulted in an overestimation of species counts (supplementary fig. S2). For instance, *Discoglossus galganoi* was split into two (GMYC) and seven (mPTP) putative

species, and *Ichthyosaura alpestris* was split into six (GMYC) and seven (mPTP). However, there were also several cases in which species numbers were underestimated, wherein valid separate species were collapsed into single units in both Anura (e.g., *Bufotes* spp., *Pelophylax* spp.) and Caudata (e.g., *Speleomantes* spp.). Both methods produced a similar partition but dissimilarities between species delimitation methods were found in 15% and 30% of the species in Anura and Caudata, respectively.

### *Barcoding gap*

A barcoding gap was present among most species and individuals (fig. 4), but it was lacking in some species in which the observed maximum intraspecific distances exceeded the interspecific distances of their nearest neighbours. We could not evaluate the existence of a barcoding gap for 11 anuran and 14 urodele species because there was only a single representative sequence for each of these species (supplementary tables S8 and S9). Within Anura, we did not observe a barcoding gap in *Alytes* (three *A. almogavarii* and nine *A. obstetricans*), *Bufo* (four *B. verrucosissimus*), *Pelodytes* (four *P. punctatus*), *Pelophylax* (four *P. lessonae* and 34 *P. ridibundus*) and *Rana* (19 *R. temporaria*) (fig. 4A, supplementary table S8). Within Caudata, a barcoding gap was only missing in *Lissotriton vulgaris* (fig. 4B, supplementary table S9).

### *Specimen identification*

The results from the two barcoding efficiency methods (BOLDi and Meier's BCM) differed slightly, although most of the specimens were correctly assigned (fig. 5, supplementary tables S8 and S9). For the Anura dataset, BOLDi and Meier's BCM correctly assigned 92% and 94% of the individuals, respectively, while for Caudata, barcode efficacy was slightly higher, with 95% of individuals correctly identified using BOLDi and 96% using Meier's BCM. The percentage of species assigned to the "ambiguous" or "no identification" categories was lower than

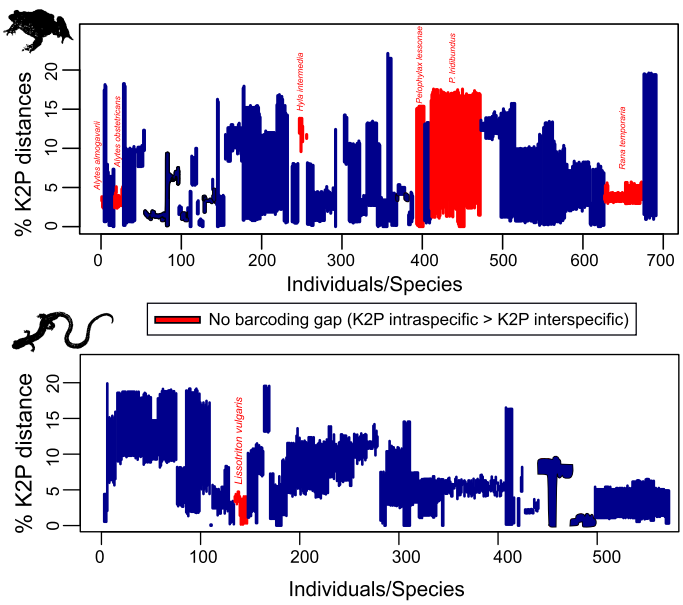


**Figure 4.** Representation of the barcoding gap for the Anura (upper panel) and Caudata (bottom panel) datasets. Each individual in the dataset is represented by a vertical line in: blue, when a barcoding gap exists (the bottom of the line representing the maximum intraspecific distance, and the top of the line representing the minimum interspecific distance); or red, when the barcoding gap is not present (the maximum intraspecific distance is higher than the minimum interspecific distance). Individuals showing no barcoding gap due to excess of missing data are not represented.

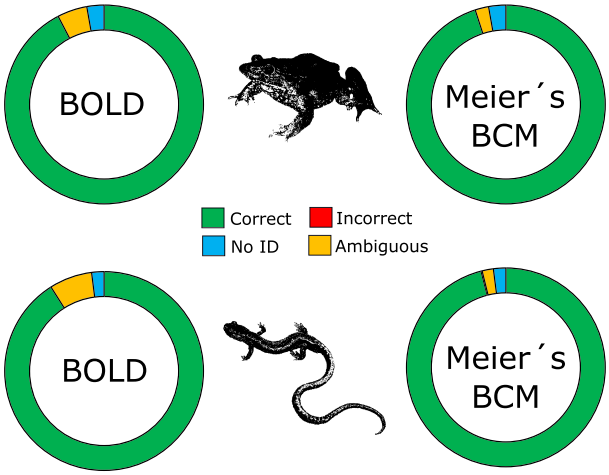


**Figure 5.** Results from barcoding efficiency methods (BOLDi and Meier's BCM) to determine the consistency of DNA barcodes with currently accepted taxonomy for Anura (top) and Caudata (bottom) datasets.

5% (except for BOLDi for Caudata, with up to 7%). "Incorrect" specimens were only observed in Meier's BCM for Caudata (0.2%).

Maps depicting species ranges and location of samples from which COI sequences were obtained for this study are available in supplementary figs. S3-S34.

## Discussion

### *A COI barcode library for Western Palearctic amphibians*

The WP is a vast area located among the Eurasian, African, and Arabian tectonic plates and covers many distinct ecoregions, such as boreal, broadleaf and Mediterranean forests, desert and xeric shrublands, and montane grasslands across an environmentally heterogeneous region (Dinerstein et al., 2017), which has resulted in the diversification of many taxonomic groups at different temporal and spatial scales (e.g., Sanmartín, 2003; Ehl, Vences and Veith, 2019; Martínez-Freiría et al., 2020; Päckert et al., 2020). Here, we provide a comprehensive, curated and publicly available DNA barcode library of a highly diverse and ancient vertebrate group, the amphibians of the WP, including 132 species (72 anurans and 60 urodeles) representing 94% and 98% of the nominal anuran and urodele species, respectively, in this region (Speybroeck et al., 2020). This reference database is, to the best of our knowledge, amongst the most complete *COI* reference libraries for any taxonomic group in the WP (see also studies in wasps, Ács et al., 2010; Schmid-Egger et al., 2019; and beetles, Jordal and Kambestad, 2014). Our study also contributes to the rapidly growing *COI* barcode repository for herpetofauna (e.g., Che et al., 2012; Xia et al., 2012; Crawford et al., 2013; Jeong et al., 2013, Perl et al., 2014; Grosjean et al., 2015; Chambers and Hebert, 2016; Hawlitschek et al., 2016; Deichmann et al., 2017; Zangl et al., 2020; Sampaio et al., 2021; Velo-Antón et al., 2022).

Quantifying intraspecific diversity is a crucial step to evaluate the effectiveness of DNA barcode reference libraries (e.g., van den Burg et al., 2020), and thus part of the robustness of our study rests in the inclusion of several representatives per species to capture the high levels of intraspecific variation observed in many amphibian species. The high levels of intraspecific variation in our datasets might be partly explained by their generally limited dispersal abilities, physiological

characteristics (ectothermy) and their complex phylogeographic histories during Pleistocene climatic fluctuations, which induced cycles of population retreat and isolation in glacial refugia. These were mostly located in the Iberian (Gómez and Lunt, 2007; Hanson et al., 2021; Paúl et al., 2023), Italian (Canestrelli, Cimmaruta and Nascetti, 2008; Chiocchio et al., 2017) and Balkan peninsulas (Pabijan et al., 2015; Jovanović et al., 2020), and also include Mediterranean ecosystems in North Africa (Husemann et al., 2014; Dinis et al., 2019) and the Near East (van Riemsdijk et al., 2017; Veith et al., 2020). Isolation in refugia boosted allopatric differentiation and diversification processes in many amphibians, with subsequent population expansions that created a mosaic of inter- and intraspecific diversity across the WP and contributed to the high levels of diversity and endemicity in the Mediterranean global biodiversity hotspot (Myers et al., 2000). Our study paid particular attention to these regional hotspots of amphibian diversity, where well-differentiated lineages can be found within many of the studied anurans (e.g., *Alytes* spp., *Discoglossus* spp., *Hyla* spp., *Rana* spp.) and urodeles (e.g., *Ichthyosaura alpestris*, *Ommatotriton* spp., *Salamandra salamandra*, *Triturus* spp.) (fig. 1 and supplementary table S1). Thus, we tried to account for shallower diversification events (i.e., within the species level and recently diverged taxa) by considering intra-specific diversity while covering a large spatial scale and a high number of amphibian species, which alleviates issues derived from neglecting intraspecific diversity in DNA barcoding studies at large geographic scales (Lukhtanov et al., 2009).

### *The discordance between mtDNA species identification and taxonomy*

Although typically described as a limitation of DNA barcoding data, the absence of barcoding gaps in certain species can provide insights about evolutionary processes. Most species were also correctly assigned (>92%) using two identification methods; however, no barcoding gap was found in eight species (seven anurans and one urodele), highlighting ongoing evolutionary processes and the inherent caveats of using mtDNA barcodes for species identification. The absence of a barcoding gap in these eight species can be explained by the following, not mutually exclusive, scenarios. Firstly, there is evidence from RAD sequencing data that some of the species in this study have recently diverged (e.g., *Alytes almogavarii*, Dufresnes and Martínez-Solano, 2020; *Rana parvipalmata*, Dufresnes et al., 2020). In these cases, incipient species would be expected to have low mtDNA interspecific divergence with their closest congeners while maintaining similar levels of intraspecific divergence at the mtDNA level (Dufresnes et al., 2021). We found support for these hypotheses in that nearest neighbour distance in these two species (*A. obstetricans*, 2-3%; *R. temporaria*, 3-5%) was similar to levels of intraspecific divergence (∼3-5%).

Other scenarios in which a barcoding gap would be absent is in the presence of hybridization and/or introgression (Dufresnes et al., 2021); perhaps the best-known example in our dataset is large-scale hybridization between *Pelophylax ridibundus* and *P. lessonae* (Dubey et al., 2019). Alternatively, absent barcoding gaps could be explained by high levels of intraspecific distance in some species. For instance, levels of intraspecific diversity within *Hyla intermedia* (12-13%) and *Hyla savignyi* (8-12%) suggest the possible existence of cryptic taxa, requiring further investigation. In fact, a northern lineage of *H. intermedia* was recently described as a distinct species (*H. perrini*; Dufresnes et al., 2018) based on deep levels of genetic differentiation; however, due to the broad geographic extent of genetic admixture with its sister lineage (*H. intermedia*), *H. perrini* was not considered as a valid species in the recent taxonomic update by the Taxonomic Committee of the Societas Europaea Herpetologica (Speybroeck et al., 2020). Relatively deep levels of mtDNA differentiation

were also observed in *Bufo spinosus* (8-12%), associated with two geographically isolated lineages (Iberian and North African; Recuero et al., 2012), which may warrant independent species status. In contrast, we found a lack of barcoding gap in cases where low interspecific mtDNA divergence combined with moderate levels of intraspecific divergence, as in *Bufo verrucossisimus* and *Lissotriton vulgaris*. These discrepancies are a reminder that the success in species identification using DNA barcode libraries relies on detailed taxonomic revisions that integrate other sources of biological information (e.g., morphology, ecology, bioacoustics, genomic DNA, patterns of hybridization), and that researchers should be cautious to not overinterpret conclusions drawn from DNA barcode data alone.

### *COI phylogenetic inference and species delimitation*

Although our phylogenetic analyses relied on a single gene tree based on a fragment of the mitochondrial *COI* gene, our results are largely congruent with previous studies in recovering relationships among closely related species, with statistical support decreasing towards the root of the trees (as is typical for data with limited phylogenetic signal). Additionally, as described above, incipient speciation and complex patterns of introgression and hybridization between amphibian taxa can blur the picture obtained with mtDNA data.

Species delimitation methods inferred the existence of substantially more species-level diversity than currently recognized, suggesting the need for taxonomic revision in many taxa. While this a common finding when applying GMYC and mPTP methods to phylogenetic datasets (Satler, Carstens and Hedin, 2013), the reasons behind such overestimated species numbers are diverse, including limited sampling and spatial sampling bias. However, cryptic speciation in amphibian diversification is also common, and several taxa have been described based on divergent mtDNA differentiation among intra-specific linages that were subsequently supported using genomic data.

### *Conclusions*

Our comprehensive *COI* reference library for WP amphibians represents a foundation for future studies targeting these species and provides insight into some of the evolutionary processes that shape intra- and inter-specific diversity in the region. With an increase in eDNA approaches, numerous studies from multiple research fields (e.g., taxonomy, biogeography, macroecology, conservation biology) are now highly dependent on available genetic database repositories; however, they are frequently constrained by the lack of curated DNA databases or subject to errors or misidentifications in global platforms such as GenBank (Harris, 2003) that affect downstream analyses, or imply to construct *de novo* local reference libraries (e.g., Peixoto et al., 2020). Thus, well curated and open access genetic data repositories are essential for studies relying on their availability for successful species identification and discovery.

DNA barcoding has inherent advantages as a cost-effective method for specimen identification and species discovery, particularly in taxonomic groups with extensive morphologic conservatism across life history stages (i.e., eggs, larvae and tadpoles; Wake et al., 2011). However, the use of DNA barcoding can fail when incipient speciation is taking place, hybridization is occurring between species, or when cryptic species are present. Thus, while it is important to not overinterpret results from DNA barcodes, this method can be helpful in detecting groups in which these processes may be occurring, highlighting areas for future investigation. While most WP urodeles can be identified using DNA barcodes, relying on mtDNA posits a problem for some WP anurans, for which additional markers are necessary. Thus, other studies (e.g., Dupont et al., 2016; Sampaio et al., 2021) have relied on multilocus barcoding

to avoid the well-known limitations of mtDNA and single-locus phylogenetic inference (Degnan and Rosenberg, 2009). For such studies, the main challenge is to find reliable nuclear barcodes with sufficient informative phylogenetic signal for robust species identification in closely related taxa; in this context, next-generation sequencing approaches can largely increase phylogenetic resolution in DNA barcoding studies (Mallo and Posada, 2016; Liu et al., 2017).

Finally, our results identified likely cryptic diversity that warrants further molecular and taxonomic evaluation, highlighting the utility of DNA barcoding approaches in uncovering cryptic diversity in poorly studied and/or complex taxonomic groups. While we cannot rely on mtDNA markers alone to delimit and describe new species, uncovering high levels of intraspecific mtDNA diversity provides important clues for targeting putative cryptic taxa that should be further investigated using genomic-level data.

**Acknowledgements.** We thank Valentina F. Orlova, Evgeniy A. Dunayev and Roman A. Nazarov (ZMMU, Moscow, Russia) for support and access to specimens under their care, and Sergius L. Kuzmin (IPEE, Moscow, Russia) for longstanding support. Wouter Beukema and Sergé Bogaerts provided pictures. We thank Nicolas Hubert for his thorough review of a previous draft of this manuscript. The associate editor and two anonymous reviewers provided valuable suggestions to improve the manuscript. Specimen collection and molecular analyses were conducted with the financial support of the Spanish Ministerio de Ciencia e Innovación (Ref.: CGL2008-04271-C02-01/ BOS) and Junta de Comunidades de Castilla la Mancha (Ref.: PPII10-0097- 4200) to IMS, the Portuguese National Funds through FCT – Foundation for Science and Technology (FCT grant No. PTDC/BIA- EVL/28475/2017), to G.V.-A., the Russian Science Foundation (RSF grant No. 22-14-00037) to N.A.P., and the Italian Ministry of University and Research (grants No. 2012FRHYRA and 2017KLZ3MA) to D.C. G.V.-A. was supported by FCT (IF/01425/2014 and CEECIND/00937/2018) and by a Ramón y Cajal research grant funded by the Spanish State Research Agency (RYC2019-026959-I). Author contributions: G.V.-A. and I.M.-S. designed the study; G.V.-A., N.P., D.C., R.B., B.N., M.J.F. and I.M.-S. contributed with samples and/or data; G.V.-A. and I.M.-S. analyzed the data; G.V.-A. prepared the first draft of the manuscript with contributions from all other authors. COI datasets are accessible on BOLD (www.boldsystems.org; see details in BOLD projects DS-AMWPA and DS-AMWPC).

**Supplementary material.** Supplementary material is available online at:
https://doi.org/10.6084/m9.figshare.23544729

## References

Ács, Z., Challis, R.J., Bihari, P., Blaxter, M., Hayward, A., Melika, G., Csóka, G., Pénzes, Z., Pujade-Villar, J., Nieves-Aldrey, J.L., Schönrogge, K., Stone, G.N. (2010): Phylogeny and DNA barcoding of inquiline oak gallwasps (Hymenoptera: Cynipidae) of the Western Palaearctic. Mol. Phylogenet. Evol. **55**: 210-225.

Ahmadzadeh, F., Shahrokhi, G., Saberi-Pirooz, R., Oladi, M., Taati, M., Poyarkov, N.A., Rödder, D. (2020): Alborz Heritage: geographic distribution and genetic differentiation of Iranian *Paradactylodon* (Amphibia: Hynobiidae). Amphibia-Reptilia **2020**: 1-16.

Bálint, M., Pfenninger, M., Grossart, H.P., Taberlet, P., Vellend, M., Leibold, M.A., Englund, G., Bowler, D. (2018): Environmental DNA time series in ecology. Trends Ecol. Evol. **33**: 945-957.

Borkin, L.J. (1999): Distribution of amphibians in North Africa, Europe, Western Asia, and the former Soviet Union. In: Patterns of Distribution of Amphibians: a Global Perspective. Duellman, W.E., Ed., The Johns Hopkins University Press, Baltimore and London, p. 329-420.

Borkin, L.J., Litvinchuk, S.N. (2013): Amphibians of the Palearctic: taxonomic composition. Proceedings ZIN **317**: 494-541.

Borkin, L.J., Litvinchuk, S.N. (2014): Zoogeography of the Northern Hemisphere and amphibians: the Palearctic and the Nearctic vs. the Holarctic?. Proceedings ZIN **318**: 433-485.

Brown, S.D., Collins, R.A., Boyer, S., Lefort, M.C., Malumbres-Olarte, J., Vink, C.J., Cruickshank, R.H. (2012): Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. **12**: 562-565.

Canestrelli, D., Cimmaruta, R., Nascetti, G. (2008): Population genetic structure and diversity of the Apennine endemic stream frog, *Rana italica* – insights on the Pleistocene evolutionary history of the Italian peninsular biota. Mol. Ecol. **17**: 3856-3872.

Chambers, E.A., Hebert, P.D. (2016): Assessing DNA barcodes for species identification in North American reptiles and amphibians in natural history collections. PLoS One **11**: e0154363.

Che, J., Chen, H.M., Yang, J.X., Jin, J.Q., Jiang, K.E., Yuan, Z.Y., Murphy, R.W., Zhang, Y.P. (2012): Universal COI primers for DNA barcoding amphibians. Mol. Ecol. Resour. **12**: 247-258.

Chiocchio, A., Bisconti, R., Zampiglia, M., Nascetti, G., Canestrelli, D. (2017): Quaternary history, population genetic structure and diversity of the cold-adapted Alpine newt *Ichthyosaura alpestris* in Peninsular Italy. Sci. Rep. **7**: 1-12.

Collins, R.A., Cruickshank, R.H. (2013): The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. **13**: 969-975.

Crawford, A.J., Cruz, C., Griffith, E., Ross, H., Ibáñez, R., Lips, K.R., Driskell, A.C., Bermingham, E., Crump, P. (2013): DNA barcoding applied to ex situ tropical amphibian conservation program reveals cryptic diversity in captive populations. Mol. Ecol. Resour. **13**: 1005-1018.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D. (2012): jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods **9**: 772.

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., Taberlet, P. (2014): DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. Mol. Ecol. Resour. **14**: 306-323.

Degnan, J.H., Rosenberg, N.A. (2009): Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. **24**: 332-340.

Deichmann, J.L., Mulcahy, D.G., Vanthomme, H., Tobi, E., Wynn, A.H., Zimkus, B.M., McDiarmid, R.W. (2017): How many species and under what names? Using DNA barcoding and GenBank data for west Central African amphibian conservation. PLoS One **12** (11): e0187283.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., et al., Hansen, M. (2017): An ecoregion-based approach to protecting half the terrestrial realm. BioScience **67**: 534-545.

Dinis, M., Merabet, K., Martínez-Freiría, F., Steinfartz, S., Vences, M., Burgon, J.D., Elmer, K.R., Donaire, D., Hinckley, A., Fahd, S., Joger, U., Fawzi, A., Slimani, T., Velo-Antón, G. (2019): Allopatric diversification and evolutionary melting pot in a North African Palearctic relict: the biogeographic history of *Salamandra algira*. Mol. Phylogenet. Evol. **130**: 81-91.

Dubey, S., Maddalena, T., Bonny, L., Jeffries, D.L., Dufresnes, C. (2019): Population genomics of an exceptional hybridogenetic system of *Pelophylax* water frogs. BMC Evol. Bio. **19**: 1-13.

Duellman, W.E., Trueb, L. (1986): Biology of Amphibians. McGraw-Hill Book Company, New York-St. Louis-San Francisco, 670 p.

Dufresnes, C., Martínez-Solano, Í. (2020): Hybrid zone genomics supports candidate species in Iberian *Alytes obstetricans*. Amphibia-Reptilia **41**: 105-112.

Dufresnes, C., Mazepa, G., Rodrigues, N., Brelsford, A., Litvinchuk, S.N., Sermier, R., Lavanchy, G., Betto-Colliard, C., Blaser, O., Borzée, A., Cavoto, E., Fabre, G., Ghali, K., Grossen, C., Horn, A., Leuenberger, J., Phillips, B.C., Saunders, P.A., Savary, R., Maddalena, T., Stöck, M., Dubey, S., Canestrelli, D., Jeffries, D.L. (2018): Genomic evidence for cryptic speciation in tree frogs from the Apennine Peninsula, with description of *Hyla perrini* sp. nov. Front. Ecol. Evol. **6**: 144.

Dufresnes, C., Nicieza, A.G., Litvinchuk, S.N., Rodrigues, N., Jeffries, D.L., Vences, M., Perrin, N., Martínez-Solano, Í. (2020): Are glacial refugia hotspots of speciation and cytonuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. Mol. Ecol. **29**: 986-1000.

Dufresnes, C., Brelsford, A., Jeffries, D.L., Mazepa, G., Suchan, T., Canestrelli, D., Nicieza, A., Fumagalli, L., Martínez-Solano, I., Litvinchuk, S.N., Vences, M., Perrin, N., Crochet, P.A. (2021): Mass of genes rather than master genes underlie the genomic architecture of amphibian speciation. Proc. Natl. Acad. Sci. U.S.A. **118**: e2103963118.

Dupont, L., Porco, D., Symondson, W.O.C., Roy, V. (2016): Hybridization relics complicate barcode-based identification of species in earthworms. Mol. Ecol. Resour. **16**: 883-894.

Ehl, S., Vences, M., Veith, M. (2019): Reconstructing evolution at the community level: a case study on Mediterranean amphibians. Mol. Phylogenet. Evol. **134**: 211-225.

Ezard, T., Fujisawa, T., Barraclough, T.G. (2009): Splits: species' limits by threshold statistics. R package version, 1, r29.

Ficetola, G.F., Miaud, C., Pompanon, F., Taberlet, P. (2008): Species detection using environmental DNA from water samples. Bio. Lett. **4**: 423-425.

Fujisawa, T., Barraclough, T.G. (2013): Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. Syst. Biol. **62**: 707-724.

Gómez, A., Lunt, D.H. (2007): Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Phylogeography of Southern European Refugia, p. 155-188. Springer, Dordrecht.

Grosjean, S., Ohler, A., Chuaynkern, Y., Cruaud, C., Hassanin, A. (2015): Improving biodiversity assessment of anuran amphibians using DNA barcoding of tadpoles. Case studies from Southeast Asia. Comptes Rendus Biologies **338**: 351-361.

Guindon, S., Gascuel, O. (2003): A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**: 696-704.

Hanson, J.O., Veríssimo, A., Velo-Antón, G., Marques, A., Camacho-Sanchez, M., Martínez-Solano, Í., Gonçalves, H., Sequeira, F., Possingham, H., Carvalho, S.B. (2021): Evaluating surrogates of genetic diversity for conservation planning. Cons. Biol. **35**: 634-642.

Harris, D.J. (2003): Can you bank on GenBank?. Trends Ecol. Evol. **18**: 317-319.

Hawlitschek, O., Morinière, J., Dunz, A., Franzen, M., Rödder, D., Glaw, F., Haszprunar, G. (2016): Comprehensive DNA barcoding of the herpetofauna of Germany. Mol. Ecol. Resour. **16**: 242-253.

Hebert, P.D., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M. (2004): Identification of birds through DNA barcodes. PLoS Biol. **2**: e312.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R. (2003): Biological identifications through DNA barcodes. Proc. Royal Soc. B. **270**: 313-321.

Husemann, M., Schmitt, T., Zachos, F.E., Ulrich, W., Habel, J.C. (2014): Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. J. Biogeogr. **41**: 81-94.

Ivanova, N.V., DeWaard, J., Hebert, P.D.N. (2006): An inexpensive, automation friendly protocol for recovering high quality DNA. Mol. Ecol. Notes. **6**: 998-1002.

Jeong, T.J., Jun, J., Han, S., Kim, H.T., Oh, K., Kwak, M. (2013): DNA barcode reference data for the Korean herpetofauna and their applications. Mol. Ecol. Resour. **13**: 1019-1032.

Jordal, B.H., Kambestad, M. (2014): DNA barcoding of bark and ambrosia beetles reveals excessive NUMT s and consistent east-west divergence across Palearctic forests. Mol. Ecol. Resour. **14**: 7-17.

Jovanović, M., Bisbal-Chinesta, J.F., Đurić, D., Bogićević, K., Nenadić, D., Agustí, J., Blain, H.A. (2020): Pleistocene herpetofaunal studies in Serbia (Balkan Peninsula, SE Europe): state of the art and perspectives. Quat. Sci. Rev. **233**: 106235.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., Flouri, T. (2017): Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics **33**: 1630-1638.

Kimura, M. (1980): A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**: 111-120.

Liu, J., Jiang, J., Song, S., Tornabene, L., Chabarria, R., Naylor, G.J., Li, C. (2017): Multilocus DNA barcoding – species identification with multilocus data. Sci. Rep. **7**: 1-12.

Lopes, C.M., Sasso, T., Valentini, A., Dejean, T., Martins, M., Zamudio, K.R., Haddad, C.F. (2017): eDNA metabarcoding: a promising method for anuran surveys in highly diverse tropical forests. Mol. Ecol. Resour. **17**: 904-914.

Lukhtanov, V.A., Sourakov, A., Zakharov, E.V., Hebert, P.D. (2009): DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. Mol. Ecol. Resour. **9**: 1302-1310.

Maddison, W.P., Maddison, D.R. (2019): Mesquite: a modular system for evolutionary analysis. Version 3.61. http://www.mesquiteproject.org.

Mallo, D., Posada, D. (2016): Multilocus inference of species trees and DNA barcoding. Philos. Trans. R. Soc. B: Biol. Sci. **371**: 20150335.

Marques, A.J., Mata, V.A., Velo-Antón, G. (2022): COI metabarcoding provides insights into the highly diverse diet of a generalist salamander, *Salamandra salamandra* (Caudata: Salamandridae). Diversity **14**: 89.

Martínez-Freiría, F., Freitas, I., Zuffi, M.A., Golay, P., Ursenbacher, S., Velo-Antón, G. (2020): Climatic refugia boosted allopatric diversification in western Mediterranean vipers. J. Biogeogr. **47**: 1698-1713.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K. (2006): DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. **55**: 715-728.

Meier, R., Zhang, G., Ali, F. (2008): The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. Syst. Biol. **57**: 809-813.

Meyer, C.P., Paulay, G. (2005): DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. **3**: 2229-2238.

Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J., Lees, D.L., Ranaivosolo, R., Eggleton, P., Barraclough, T., Vogler, A.P. (2009): Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst. Biol. **58**: 298-311.

Murphy, R.W., Crawford, A.J., Bauer, A.M., Che, J., Donnellan, S.C., Fritz, U., Haddad, C.F.B., Nagy, Z.T., Poyarkov, N.A., Vences, M., Wang, W., Zhang, Y. (2013): Cold code: the global initiative to DNA barcode amphibians and nonavian reptiles. Mol. Ecol. Resour. **13**: 161-167.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A., Kent, J. (2000): Biodiversity hotspots for conservation priorities. Nature **403**: 853-858.

Pabijan, M., Zieliński, P., Dudek, K., Chloupek, M., Sotiropoulos, K., Liana, M., Babik, W. (2015): The dissection of a Pleistocene refugium: phylogeography of the smooth newt, *Lissotriton vulgaris*, in the Balkans. J. Biogeogr. **42**: 671-683.

Päckert, M., Bader-Blukott, M., Künzelmann, B., Sun, Y.H., Hsu, Y.C., Kehlmaier, C., Albrecht, F., Illera, J.C., Martens, J. (2020): A revised phylogeny of nuthatches (Aves, Passeriformes, *Sitta*) reveals insight in intra-and interspecific diversification patterns in the Palearctic. Vertebr Zool. **70**: 241-262.

Paúl, M.J., Rosauer, D., Tarroso, P., Velo-Antón, G., Carvalho, S.B. (2023): Environmental and topographic drivers of amphibian phylogenetic diversity and endemism in the Iberian Peninsula. Ecol. Evol. **13**: e9666.

Peixoto, S., Chaves, C., Velo-Antón, G., Beja, P., Egeter, B. (2021): Species detection from aquatic eDNA: assessing the importance of capture methods. Environmental DNA **3**: 435-448.

Peixoto, S., Mota-Ferreira, M., Chaves, C., Velo-Antón, G., Beja, P., Egeter, B. (2023): Multi-species occupancy modeling reveals methodological and environmental effects on eDNA detection of amphibians in temporary ponds. Environ. DNA.

Perl, R.B., Nagy, Z.T., Sonet, G., Glaw, F., Wollenberg, K.C., Vences, M. (2014): DNA barcoding Madagascar's amphibian fauna. Amphibia-Reptilia **35**: 197-206.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P. (2006): Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. **55**: 595-609.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A. (2018): Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. **67**: 901-904.

Recuero, E., Iraola, A., Rubio, X., Machordom, A., García-París, M. (2007): Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern. J. Biogeogr. **34**: 1207-1219.

Recuero, E., Canestrelli, D., Vörös, J., Szabó, K., Poyarkov, N.A., Arntzen, J.W., Crnobrnja-Isailovic, J., Kidov, A.A., Cogălniceanu, D., Caputo, F.P., Nascetti, G., Martínez-Solano, I. (2012): Multilocus species tree analyses resolve the radiation of the widespread *Bufo bufo* species group (Anura, Bufonidae). Mol. Phylogenet. Evol. **62**: 71-86.

Saats, M., Arulandhu, A.J., Gravendeel, B., Holst-Jensen, A., Scholtens, I., Peelen, T., et al., Kok, E. (2016): Advances in DNA metabarcoding for food and wildlife forensic species identification. Analytical and Bioanalytical Chemistry **408** (17): 4615-4630.

Sampaio, M., Velo-Antón, G., Martínez-Freiría, F., Sánchez-Vialas, A., Pleguezuelos, J.M., Geniez, P., Crochet, P.-A., Brito, J.C. (2021): Beyond the comfort zone: Amphibian diversity and distribution in the West Sahara-Sahel using mtDNA and nuDNA barcoding and spatial modelling. Conserv. Genet. **22**: 233-248.

Sanmartín, I. (2003): Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). J. Biogeogr. **30**: 1883-1897.

Satler, J.D., Carstens, B.C., Hedin, M. (2013): Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). Syst. Biol. **62**: 805-823.

Schmid-Egger, C., Straka, J., Ljubomirov, T., Blagoev, G.A., Morinière, J., Schmidt, S. (2019): DNA barcodes identify 99 per cent of apoid wasp species (Hymenoptera: Ampulicidae, Crabronidae, Sphecidae) from the Western Palearctic. Mol. Ecol. Resour. **19**: 476-484.

Smith, M.A., Poyarkov Jr., N.A., Hebert, P.D. (2008): DNA barcoding: CO1 DNA barcoding amphibians: take the chance, meet the challenge. Mol. Ecol. Resour. **8**: 235-246.

Speybroeck, J., Beukema, W., Dufresnes, C., Fritz, U., Jablonski, D., Lymberakis, P., Martínez-Solano, I., Razzetti, E., Vamberger, M., Vences, M., Vörös, J., Crochet, P.-A. (2020): Species list of the European herpetofauna – update by the Taxonomic Committee of the Societas Europaea Herpetologica. Amphibia-Reptilia **41**: 139-189.

Stat, M., Huggett, M.J., Bernasconi, R., DiBattista, J.D., Berry, T.E., Newman, S.J., Harvey, E.S., Bunce, M. (2017): Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Sci. Rep. **7**: 1-11.

Stoeckle, M.Y., Hebert, P.D. (2008): Barcode of life. Sci. Am. **299**: 82-86.

Stuart, S.N., Chanson, J.S., Cox, N.A., Young, B.E., Rodrigues, A.S., Fischman, D.L., Waller, R.W. (2004): Status and trends of amphibian declines and extinctions worldwide. Science **306**: 1783-1786.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A. (2018): Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. **4**: vey016.

Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L.H. (2012): Environmental DNA. Mol. Ecol. **21**: 1789-1793.

Taberlet, P., Bonin, A., Coissac, E., Zinger, L. (2018): Environmental DNA: for Biodiversity Research and Monitoring. Oxford University Press.

van den Burg, M.P., Herrando-Pérez, S., Vieites, D.R. (2020): ACDC, a global database of amphibian cytochrome-b sequences using reproducible curation for GenBank records. Sci. Data **7**: 1-6.

van Riemsdijk, I., Arntzen, J.W., Bogaerts, S., Franzen, M., Litvinchuk, S.N., Olgun, K., Wielstra, B. (2017): The Near East as a cradle of biodiversity: a phylogeography of banded newts (genus *Ommatotriton*) reveals extensive inter-and intraspecific genetic differentiation. Mol. Phylogenet. Evol. **114**: 73-81.

Veith, M., Göçmen, B., Sotiropoulos, K., Eleftherakos, K., Lötters, S., Godmann, O., Karış, M., Oğuz, A., Ehl, S. (2020): Phylogeographic analyses point to long-term survival on the spot in micro-endemic Lycian salamanders. PloS One **15**: e0226326.

Velo-Antón, G., Henrique, M., Liz, A.V., Martínez-Freiría, F., Pleguezuelos, J.M., Geniez, P., Crochet, P.A., Brito, J.C. (2022): DNA barcode reference library for the West Sahara-Sahel reptiles. Sci. Data. **9**: 1-8.

Vences, M., Thomas, M., Van der Meijden, A., Chiari, Y., Vieites, D.R. (2005): Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. Front. Zoo. **2**: 5.

Vences, M., Nagy, Z.T., Sonet, G., Verheyen, E. (2012): DNA barcoding amphibians and reptiles. In: DNA Barcodes, p. 79-107. Humana Press, Totowa, NJ.

Wake, D.B., Koo, M.S. (2018): Amphibians. Curr. Biol. **28**: R1237-R1241.

Wake, D.B., Wake, M.H., Specht, C.D. (2011): Homoplasy: from detecting pattern to determining process and mechanism of evolution. Science **331**: 1032-1035.

Xia, Y.U.N., Gu, H.F., Peng, R.U.I., Chen, Q.I.N., Zheng, Y.C., Murphy, R.W., Zeng, X.M. (2012): COI is better than 16S rRNA for DNA barcoding Asiatic salamanders (Amphibia: Caudata: Hynobiidae). Mol. Ecol. Resour. **12**: 48-56.

Zangl, L., Daill, D., Schweiger, S., Gassner, G., Koblmüller, S. (2020): A reference DNA barcode library for Austrian amphibians and reptiles. PloS One **15**: e0229353.

Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A. (2013): A general species delimitation method with applications to phylogenetic placements. Bioinformatics **29**: 2869-2876.