Biochemical Systematics and Ecology 62 (2015) 36–50

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

journal homepage: www.elsevier.com/locate/biochemsyseco

DNA barcode assessment of Mediterranean mayflies (Ephemeroptera), benchmark data for a regional reference library for rapid biomonitoring of freshwaters

Simone Cardoni [a], Roberta Tenchini [a], Irene Ficulle [b], Roberta Piredda [c], Marco Cosimo Simeone [b, *], Carlo Belfiore [a]

[a] Dipartimento di Ecologia e Biologia (DEB), Università degli Studi della Tuscia, via S. Camillo de' Lellis, 01100 Viterbo, Italy
[b] Dipartimento di Agricoltura, Foreste, Natura ed Energia (DAFNE), Università degli Studi della Tuscia, via S. Camillo de' Lellis, 01100 Viterbo, Italy
[c] Stazione Zoologica Sperimentale Anton Dohrn, Villa Comunale, 80121 Napoli, Italy

ARTICLE INFO

Article history:
Received 5 May 2015
Received in revised form 21 July 2015
Accepted 26 July 2015
Available online xxx

Keywords:
Ephemeroptera
Freshwaters
Mediterranean
DNA barcoding
Biodiversity
Conservation

ABSTRACT

Accurate identification of aquatic species is fundamental to freshwater research. In this paper, we targeted Ephemeroptera, a key taxonomic insect group for biomonitoring of water bodies and present an overview on the efficacy of the DNA barcoding approach to document species identity in the Mediterranean region. We sequenced the mitochondrial cytochrome c oxidase (COI) in 39 nominal species. Sample discrimination and species identification were investigated by evaluating haplotype identity and similarity, intra-/interspecific genetic distances, optimal identification of barcoding gap thresholds, estimates of species monophyly and comparative species matches on available reference libraries. The resolving power of the obtained data was discussed in the light of statistical tools such as Spider R-package and Poisson Tree Processes. High levels of species identification were achieved with all the used methodologies, and the occurrence of cryptic species was suggested. We conclude that DNA barcoding is a powerful tool for taxonomic research in Mediterranean mayflies, with great promise to ameliorate biodiversity inventories of freshwater ecosystems and to provide the necessary accuracy for water quality assessment programs. Our results further indicated we need to upgrade the current regional mayfly diversity knowledge. The development of a Mediterranean reference library could integrate this new information system.

1. Introduction

Freshwaters are hotspots of biodiversity as well as of endangerment. Their extraordinary rich and endemic biota are now far more threatened than their marine or terrestrial counterparts (Strayer and Dudgeon, 2010). Extinctions, range reductions, population declines are mainly caused by a set of environmental transformations inherently linked with human population growth and global economy (Strayer, 2006; IUCN, 2007; Polunin, 2008), coupled with the strong sensitivity and the insular nature of many species' habitats (Dudgeon et al., 2006). We therefore need to improve our knowledge on global and local

* Corresponding author.
E-mail address: mcsimeone@unitus.it (M.C. Simeone).

http://dx.doi.org/10.1016/j.bse.2015.07.035
0305-1978/

abiotic factors, species status and threats, and explore the links between freshwater biodiversity and ecosystem function to prevent further losses and plan adequate management practices.

Mayflies (Ephemeroptera) are an order of benthic macroinvertebrates inhabiting freshwaters, known as the most primitive and ancient lineage among extant winged insects (Pterygota) (Edmunds and McCafferty, 1988; Brittain and Sartori, 2003). They are distributed worldwide and encompass over 3000 species, ca. 400 genera and 42 families (Barber-James et al., 2008). Nymphs, always aquatic, are the dominant life history stage. They colonize all types of freshwaters but are more diversified in running waters than in lakes or ponds. Length and number of life cycles per year depend largely on geographic locality and may take over 2 years to mature in temperate climates; adults usually live from a few hours to a few weeks.

According to the WFD 2000/60/CE (European Commission, 2000), mayflies are recognized as biological indicators for ecological and biomonitoring studies. Specifically, their relative species presence and abundance allow the experts to apply specific metric tools and assign values with related classes of quality to freshwater ecosystems. The importance and accuracy of species identification is therefore fundamental for multiple basic and applied subjects, such as studies of taxonomy, evolution, community structure, and for freshwater biodiversity assessment, conservation and management. However, field identification of mayfly species is a notoriously difficult task, mostly due to the time limited availability of subtle morphological diagnostic characters, to the often exceeding amount of biological material to process, and because of a shortage of adequately trained taxonomists (Jones, 2008; Monk et al., 2012). At the same time, mayfly nomenclature and taxonomy are still in progress (Barber-James et al., 2008); they have been the subject of numerous recent changes, with consistent reorganizations of major ranks, a great increase in the number of recently recognized taxa, and numerous unknown species and genera still awaiting a new description (Waltz and McCafferty, 1997; Lugo-Ortiz et al., 1999; Kluge, 2004).

Molecular technology can serve as a reliable supplemental tool for the identification of living organisms (Pennisi, 2003; Tautz et al., 2003) and DNA barcoding (Hebert et al., 2003) is the most recent and widely used tool for the generation of universal reference standards, currently promoted for a variety of bio-ecological applications, including improvement of traditional taxonomies (Dayrat, 2005; Hajibabaei et al., 2007), species discovery (Hebert et al., 2004), biodiversity inventorying (Janzen et al., 2005) and biomonitoring (Yu et al., 2012). The main actors of this innovative methodology are the DNA barcodes: globally accepted short DNA sequences whose variation allows the operator to unambiguously identify a species through the use of specific bioinformatic tools (Hebert et al., 2003). Based on the relative ease of amplification, sequencing, multi-alignment and the amount of variation displayed (sufficient to discriminate among sister species without affecting their correct assignation through intra-specific variation), the mitochondrial (mt) cytochrome c oxidase I (COI) gene region is currently used to barcode metazoan groups (Hebert et al., 2003).

In insects, DNA barcoding has been shown to be a reliable, fast and cost-effective approach for discovering new species (Smith et al., 2008), revising taxonomies (Smith et al., 2013) and for associating the sexes in dimorphic species (Sheffield et al., 2009), or immature and adult forms (Janzen et al., 2005). The latter issues were found to be particularly valuable for aquatic insects, enabling the identification of larval stages and females that often appear taxonomically ambiguous (Hajibabaei et al., 2011), including several major groups such as caddisflies (Trichoptera; Hogg et al., 2009), stoneflies (Plecoptera; Zhou et al., 2009), dragonflies (Odonata; Rach et al., 2008), midges (Diptera: Chironomidae; Ekrem et al., 2007), and blackflies (Diptera: Simuliidae; Rivera and Currie, 2009). The results obtained have stimulated growing interest in developing barcode libraries to allow comprehensive and fast surveys of regional freshwaters faunas (Hajibabaei et al., 2012).

The efficacy of DNA barcodes in Ephemeroptera was initially tested by Ball et al. (2005) on 80 taxa from the Northeastern United States and Central Canada, and by Zhou et al. (2009) on 37 morphospecies from a subartic site in Canada. Both research groups achieved high identification success (99–100%) and concluded that inclusion of DNA barcoding in future applied biomonitoring studies was realistic, given the increased taxonomic resolution delivered, especially when morphological identification was compromised by the inability to recognize early stages, damaged or fragmented specimens. Then, Webb et al. (2012) provided records for more than 4000 individuals from over 350 species from North America and extended the barcode coverage to approx. 50% of the continental known Ephemeropterans. Further studies took advantage of DNA barcodes to clear systematic relationships within a family level (Sroka, 2012), to differentiate species with morphologically indistinguishable nymphs and adult females (Elderkin et al., 2012), for understanding the local population structure (Ogitani et al., 2011), to dissect cryptic species complexes (Williams et al., 2006; Ståhls and Savolainen, 2008), and for phylogeographic inferences on endemic taxa (Gattolliat et al., 2015).

No practical applications of DNA barcoding of mayflies are currently available in the Mediterranean region, one of the 25 world hotspots of biodiversity (Myers et al., 2000) where freshwaters are among the main, most threatened, and yet poorly investigated natural components (Blondel and Aronson, 1999).

The work now being carried out aims at evaluating the DNA barcoding efficacy to document mayfly species identity in a practical study case in the Mediterranean region. Our vision is to set up the first regional reference database and further develop this to national and continental scales as a statistically sound basis for cost efficient and repeated measurements of biodiversity, with sufficient resolution to be useful in monitoring freshwater ecology, biodiversity loss, mitigation and recovery of altered habitats, and to prevent errors caused by scarcity of data in future land management. We investigated the discriminatory power of the 5' COI region sequence variation in the mayfly fauna inhabiting central Italy and nearby Tyrrhenian islands. This area was selected because of the relevant amount of occurring mayfly species and available background knowledge (Thomas and Belfiore, 2004; Belfiore, 2005). We inspected the ways in which the assessment of haplotype diversity, patterns of intra-/interspecific genetic divergences and comparison with available public reference databases can

provide rapid insights into the taxonomic identity of the sampled taxa. Obtained data were discussed in the light of the sampling regime, estimated species monophyly and current state of mayfly taxonomy to set up expedite and reliable protocols in future practical works.

2. Materials and methods

2.1. Sampling and DNA analyses

The study sites included streams and rivers of Central Italy (Latium, Tuscany) and the major Tyrrhenian islands (Sardinia, Elba and Corse) (Fig. 1). Sampling was performed throughout May, 2012 and April, 2014. Mayfly samples at the nymph stage were selected from the bulk and separately preserved in 95% ETOH. When possible, efforts were made to include at least three specimens for each species. Morphological identifications were carried out at the Laboratory of Freshwater Ecology (University of Tuscia). The final dataset included 89 samples, corresponding to 39 morphologically acknowledged species (18 genera, 8 families) (Table 1).

DNA was extracted from legs, cerci, half or whole body according to the specimen's size; the Insect DNA kit (OMEGA) was used following the manufacturer's instructions. The COI regions were amplified with both the LC01490/HC02198 (Folmer et al., 1994) and LepF1/LepR1 (Hajibabaei et al., 2006) primer pairs. PCR products were obtained with RTG PCR beads (GE Healthcare) and the thermal regime consisted of 1 cycle at 94 °C for 3 min, 35 cycle at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s and 1 cycle at 72 °C for 10 min. Amplified PCR bands were purified with Illustra GFX/PCR DNA Purification Kit (GE Healthcare); standardized aliquots were sent to Macrogen for sequencing (http://www.macrogen.com). Electropherograms were edited with CHROMAS 2.3 (http://www.technelysium.com.au) and checked visually.

2.2. Bioinformatic tools

Optimal multiple alignment and the amino acid translation to ensure that no stop codons were present in the nucleotide sequences were obtained with MAFFT and 6FrameTranslation (http://toolkit.tuebingen.mpg.de). Several analytical methods were used to assess discrimination ability of the barcode sequence data. Mothur (Schloss et al., 2009) was used to cluster sequences at 100% and 98% of distances. Clusters haplotypes in agreement with taxonomy, that is, containing only one sequence or intra-specific specimens of one morphological species, were counted as successful, whereas clusters containing sequences from different morphological species were counted as failure. We then investigated the occurrence of the 'barcoding gap' for each species, a key step to assess species discrimination (Meyer et al., 2008), i.e. the assumption that the amount of sequence divergence within species is smaller than that between species. Kimura 2P genetic distances were then computed within and among congeneric species; all the species presenting a minimum interspecific distance value higher than their maximum intraspecific distance were considered successfully discriminated.

Spider (Brown et al., 2012), an R-package for the analysis of species identity and evolution, with particular reference to DNA barcoding, was used to calculate the following parameters: K2P genetic distance matrixes (dis.dna function), the occurrence of barcoding gaps (maxInDist and nonConDist function), the threshold optimization for sample identification through the statistical evaluation of false positives (conspecific specimens misdiagnosed as 'novel' species) and false negatives (specimens from different species misdiagnosed as conspecific) occurrence (threshOpt function and threshVal function).

Species delimitation (the process of grouping 'haplotype clusters' into distinct taxonomic groups) in a Neighbor Joining tree was further evaluated under the criterium of species reciprocal monophyly (Rosenberg, 2007), that evaluates the observed branching pattern against a random branching pattern, and the nodes with p-value <0.05 were considered as significantly monophyletic (nj, read.tree, monophyly, Rosenberg's function). Alternative approaches such as bootstrap support in a Maximum Likelihood tree (1000 bootstrap replicates) built under the GTRCAT model with RAxML (Stamatakis, 2006), where all nodes achieving >70% bootstrap support were considered as monophyletic taxonomic groups, and the Poisson Tree Processes model (PTP), where differences between relationships among and within species are modelled in

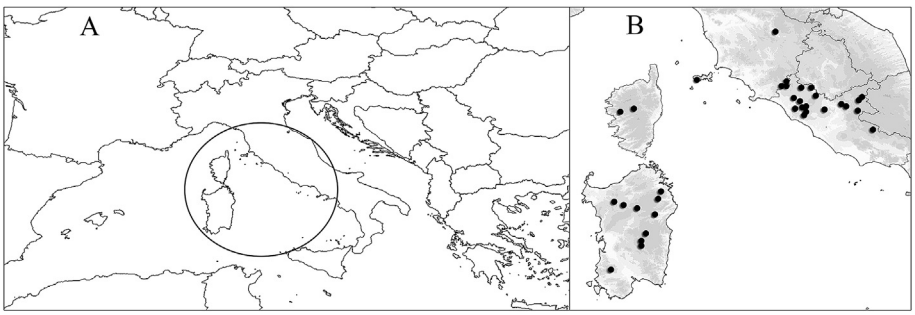

Fig. 1. Geographical area (A) and distribution map (B) of the investigated sampling sites (dotted).

Table 1
List of investigated specimens, with taxonomic nomenclature according to Fauna Europea last update 29 August 2013 | version 2.6.2, sampling sites and GenBank accession numbers.

Family	Genus	Species	Location	Accession number
Baetidae	*Baetis*	*albinatii* Sartori & Thomas	Siligo, Sardinia	LN734680
		muticus Linnaeus	Vitozza, Tuscany	LN734678
		vernus Curtis	Treja, Latium	LN734679
		cyrneus Thomas & Gazagnes	Villanova Tulo, Sardinia	LN734681
			Posada, Sardinia, specimen 1	LN734682
			Posada, Sardinia, specimen 2	LN734683
			Posada, Sardinia, specimen 3	LN734684
		lutheri Müller-Liebenau	Meleta Pitigliano, Tuscany	LN734685
			Aquarella Tuscania, Latium	LN734686
		fuscatus Linnaeus	Lunevara, Sardinia	LN734687
			Rio Mannu, Sardinia	LN734688
		alpinus Pictet	Fosso di Galantina, Latium	LN734689
			Aniene, Latium	LN734690
			Obito, Latium	LN734691
		buceratus Eaton	Mola di Oriolo, Latium, specimen 1	LN734692
			Riofi, Tuscany	LN734693
			Soriano nel Cimino, Latium	LN734694
			Mola di Oriolo, Latium, specimen 2	LN734695
			Mola di Oriolo, Latium, specimen 3	LN734696
	Centroptilum	*luteolum* Müller	Rafanello, Latium	LN734697
			Fosso Tancia, Latium	LN734698
			Fosso di Galantina, Latium	LN734699
	Procloeon	*pulchrum* Eaton	Vitozza, Tuscany	LN734700
			Ortolano, Latium	LN734701
			Rio Chiaro, Latium	LN734702
		bifidum Bengtsson	Mignone Montericcio, Latium	LN734703
			Seccheto, Elba Island	LN734704
			Torrente Leni, Sardinia	LN734705
	Cloeon	*dipterum* Linnaeus	Mignone Canale Monterano, Latium	LN734706
			Aquarella Tuscania, Latium	LN734707
		simile Eaton	Tirso, Sardinia	LN734708
Caenidae	*Caenis*	*pusilla* Navas	Mignone Montericcio, Latium	LN734709
			Aquarella Tuscania, Latium	LN734710
		beskidensis Sowa	Vitozza, Tuscany	LN734711
			Rio Chiaro, Latium	LN734712
			Aquarella Tuscania, Latium	LN734713
		martae Belfiore	Mola di Oriolo, Latium	LN734714
			Ortolano, Latium	LN734715
			Riofi, tuscany	LN734716
		sp.	Tirso, Sardinia	LN734717
	Brachycercus	*harrisella* Curtis	Vitozza, Tuscany	LN734718
			Rio Chiaro, Latium	LN734719
Heptageniidae	*Ecdyonurus*	*helveticus* Eaton	Fosso Tancia, Latium	LN734720
			Ortolano, Latium	LN734721
			Meleta Collina, Tuscany	LN734722
		corsicus Esben-Petersen	Rio Tistigliosi, Sardinia	LN734723
		venosus Fabricius	Ortolano, Latium, specimen 1	LN734724
			Introvabile, Latium	LN734725
			Mignone Canale Monterano, Latium	LN734726
			Ortolano, Latium, specimen 2	LN734727
	Epeorus	*assimilis* Eaton	Obito, Latium	LN734728
			Aniene, Latium	LN734729
			Ortolano, Latium	LN734730
	Electrogena	*grandiae* Belfiore	Introvabile, Latium	LN734731
			Rafanello, Latium	LN734732
		zebrata Hagen	Rio Mannu, Sardinia	LN734733
			Padru Rio de su Lernu, Sardinia	LN734734
	Rhithrogena	*semicolorata* Curtis	Mignone Canale Monterano, Latium	LN734735
			Fosso Tancia, Latium	LN734736
			Rio di Sesto, Trentino	LN734737
		reatina Sowa & Belfiore	Velino Antrodoco, Latium	LN734738
			Velino Canetra, Latium	LN734739
		nuragica Belfiore	Rio Aratu, Sardinia	LN734740
		insularis Esben-Petersen	Restonica, Corse	LN734741
			Murzo, Corse	LN734742
		eatoni Esben-Petersen	Murzo, Corse	LN734743
	Heptagenia	*longicauda* Stephens	Rigomero, Latium	LN734744
			Aquarella Tuscania, Latium	LN734745
			Rio Chiaro, Latium	LN734746

(*continued on next page*)

Table 1 (*continued*)

Family	Genus	Species	Location	Accession number
Leptophlebiidae	*Habroleptoides*	*confusa* Sartori & Jacob	Fosso di Galantina, Latium	LN734747
			Rafanello, Latium	LN734748
			Ortolano, Latium	LN734749
		sp.	Rio Tistigliosi, Sardinia	LN734750
	Habrophlebia	*eldae* Jacob & Sartori	Vitozza, Tuscany	LN734751
			Fosso Arlena, Latium	LN734752
			Rafanello, Latium	LN734753
	Paralaptophlebia	*submarginata* Stephens	Aquarella Tuscania, Latium	LN734754
			Vitozza, Tuscany, specimen 1	LN734755
			Vitozza, Tuscany, specimen 2	LN734756
	Choroterpes	*picteti* Eaton	Mignone Montericcio, Latium	LN734757
Ephemerellidae	*Serratella*	*ignita* Poda	Vitozza, Tuscany	LN734758
			Rafanello, Latium	LN734759
			Fosso Arlena, Latium	LN734760
			Ortolano, Latium	LN734761
Oligoneuridae	*Oligoneuriella*	*rhenana* Imhoff	Meleta Pitigliano, Tuscany	LN734762
			Rio Chiaro, Latium	LN734763
Siphlonuridae	*Siphlonurus*	*lacustris* Eaton	Vitozza, Tuscany	LN734764
			Velino Antrodoco, Latium	LN734765
Ephemeridae	*Ephemera*	*danica* Müller	Vitozza, Tuscany	LN734766

terms of numbers of substitutions as in two independent classes of Poisson processes, and used to calculate a number of entities that represent theoretical species (Zhang et al., 2013), were performed to provide multiple tests of the resolving power of the obtained data.

To simulate a practical barcode identification scenario, the obtained sequences were blasted against GenBank (http://www.ncbi.nlm.nih.gov), the global database with publicly available nucleotide sequences for over 322,000 formally described species (December 2014), and BOLD (http://barcodinglife.org), the official repository for almost 228,000 DNA barcodes of all living organisms (December 2014), to evaluate correct species identification. Before the discrimination assessment, both databases were screened for the presence of the COI sequence at the species level relatively to our dataset. A query sequence was considered as successfully discriminated if the top identity percentage obtained in the GenBank matched the name of the species. When more than one species shared the same sequence identity or the species scored lower, the result was considered an identification failure.

Finally, to evaluate the positioning of our sequences when integrated with the mayfly COI barcodes currently available in the official database, a subset of unique haplotypes longer than 500 bp was selected from BOLD to represent all the available taxonomic breadth within the same European mayfly families and genera of our study, and a RAxML dendrogram was constructed under the GTRCAT model of sequence evolution.

3. Results

3.1. Morphological and molecular dataset

Our sampling design included 34 streams and rivers of the studied area and resulted in a dataset consisting of 39 morphologically recognized species (18 genera, 8 families), representing about 12% and 37% of the total Ephemeroptera species currently listed in Europe and Italy, respectively (Thomas and Belfiore, 2004; Belfiore, 2005). Nine total endemic species (one from peninsular Italy, eight from Corse-Sardinia) were part of the dataset. Seventeen species were represented by at least three specimens sampled in different streams or rivers. No additional specimens were found in other locations for the remaining twenty-two species, that were thus represented by only one or two individuals. COI sequences were obtained from 100% of samples. The multiple alignment was unambiguous (no gaps occurred in the whole dataset). No stop codons were observed in the translated protein sequences. The total alignment was 614 bp long (primer sequences excluded).

3.2. Species discrimination

A threshold of 100% sequence identity generated 77 total haplotypes; of these, 70 were singletons and 7 were only shared among conspecific samples (*Baetis cyrneus, Baetis fuscatus,* B. buceratus, Centroptilum luteolum, Caenis beskidensis, Electrogena grandiae, Rhithrogena insularis). A 2% cut-off reduced the number of haplotypes to 48 (22 single and 26 intra-specifically shared) haplotypes. Of the seventeen species represented by three or more samples, eleven (ca. 65%) displayed multiple haplotypes (lists presented as Online Resource I). As shown in Table 2, intra-specific K2P genetic distances ranged between 0.00 and 19.6%, with three species scoring values higher than 10%: *Procleon bifidum*, *Baetis cyrneus* and *Epeorurs assimilis*. The inter-specific (intra-generic) K2P distances scored values ranging between 8.17% (*Rhithrogena reatina/R. nuragica*) and 33.02% (*Caenis beskidensis/Caenis macrura gr.*). *Procleon*, *Baetis* and *Epeorus* showed intra-specific divergences higher than the overall lower limit of the inter-specific distances observed in the entire dataset. The mean intra-specific

Table 2
Species-level summary of K2P distances for the COI sequences of 17 species with multiple specimens included in this study. Barcoding gap calculated as the difference between max. intra- and min interspecific distance.

Species	Intra-specific distance		Inter-specific distance		Barcoding gap
	Min	Max	Min	Max	
Baetis cyrneus	0.000	0.149	0.184	0.282	0.035
Baetis lutheri	0.007	0.007	0.217	0.300	0.211
Baetis fuscatus	0.000	0.000	0.200	0.296	0.200
Baetis alpinus	0.015	0.067	0.184	0.300	0.117
Baetis buceratus	0.000	0.000	0.226	0.293	0.226
Procloeon pulchrum	0.002	0.056	0.187	0.212	0.131
Procloeon bifidum	0.027	0.196	0.187	0.212	−0.009
Caenis pusilla	0.005	0.005	0.223	0.263	0.218
Caenis beskidensis	0.000	0.000	0.223	0.330	0.223
Caenis martae	0.002	0.030	0.103	0.330	0.073
Ecdyonurus helveticus	0.020	0.028	0.212	0.225	0.184
Ecdyonurus venosus	0.003	0.010	0.117	0.225	0.107
Electrogena grandiae	0.000	0.000	0.240	0.244	0.240
Electrogena zebrata	0.005	0.005	0.240	0.244	0.235
Rhithrogena semicolorata	0.015	0.018	0.180	0.205	0.162
Rhithrogena reatina	0.005	0.005	0.082	0.199	0.077
Rhithrogena insularis	0.000	0.000	0.115	0.187	0.115

divergence varied between 0.2% (*Electrogena*) and 8.73% (*Procloeon*), whereas the inter-specific values ranged between 10.33% and 33.02% (*Caenis*). When individual congeneric species were examined, the occurrence of barcoding gaps was observed in 7 genera, ranging between 3.5% (*Baetis*) and 23.5% (*Electrogena*). The only overlap in the distribution of intra- and interspecific divergences was observed within *Procloeon* and consisted of 2 individuals belonging to *Procloeon pulchrum* and *P. bifidum*.

Optimization threshold analysis using the whole dataset (Fig. 2) showed that it was not possible to define a common threshold such as to allow unambiguous discrimination of specimens; errors were minimized at a genetic distance threshold of 7% and caused a cumulative error in the 9–18% range. Analysis conducted for each genus separately and after removal of singleton species, highlighted that an optimal threshold of 11% allows unambiguous discrimination in all genus except in *Procloeon* and *Baetis* characterized by an extremely high intraspecific distance (Online Resource II).

The NJ dendrogram of the obtained COI sequences is shown in Fig. 3. All individuals clustered into distinct taxonomic species groups according to the acknowledged taxonomy, except *Procleon bifidum*; unexpectedly, five of the eight genera where multiple species were analyzed appeared as largely paraphyletic, with only *Ecdyonurus*, *Rhithrogena* and *Habroleptoides* forming cohesive clusters. The Rosenberg's probability test (p-value = 0.05) of reciprocal monophyly (Rosenberg, 2007) highlighted very few, and only the deeper nodes of the dendrogram, as significantly monophyletic. Baetidae was the only cohesive family receiving a significant monophyly support.

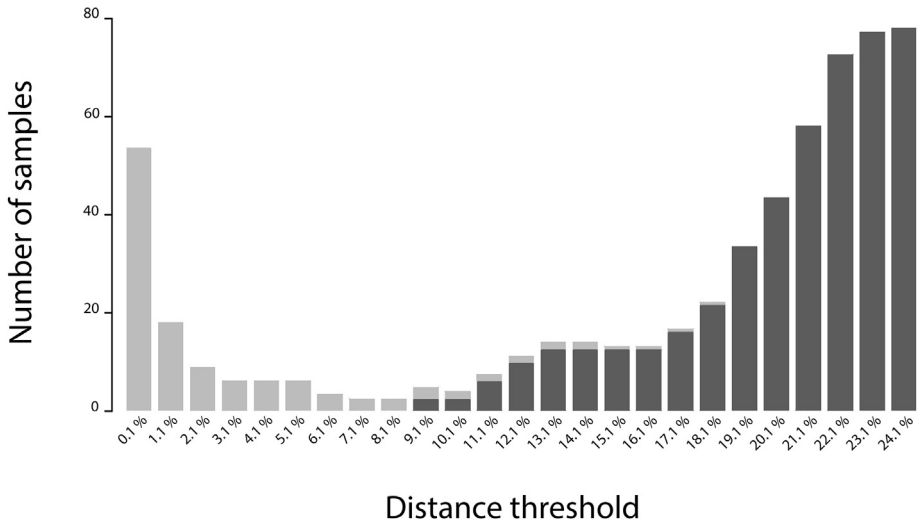

Fig. 2. Barplot showing the false positive (dark grey) and false negative (light grey) rate of mayfly identification in the present study based on COI K2P genetic distances.

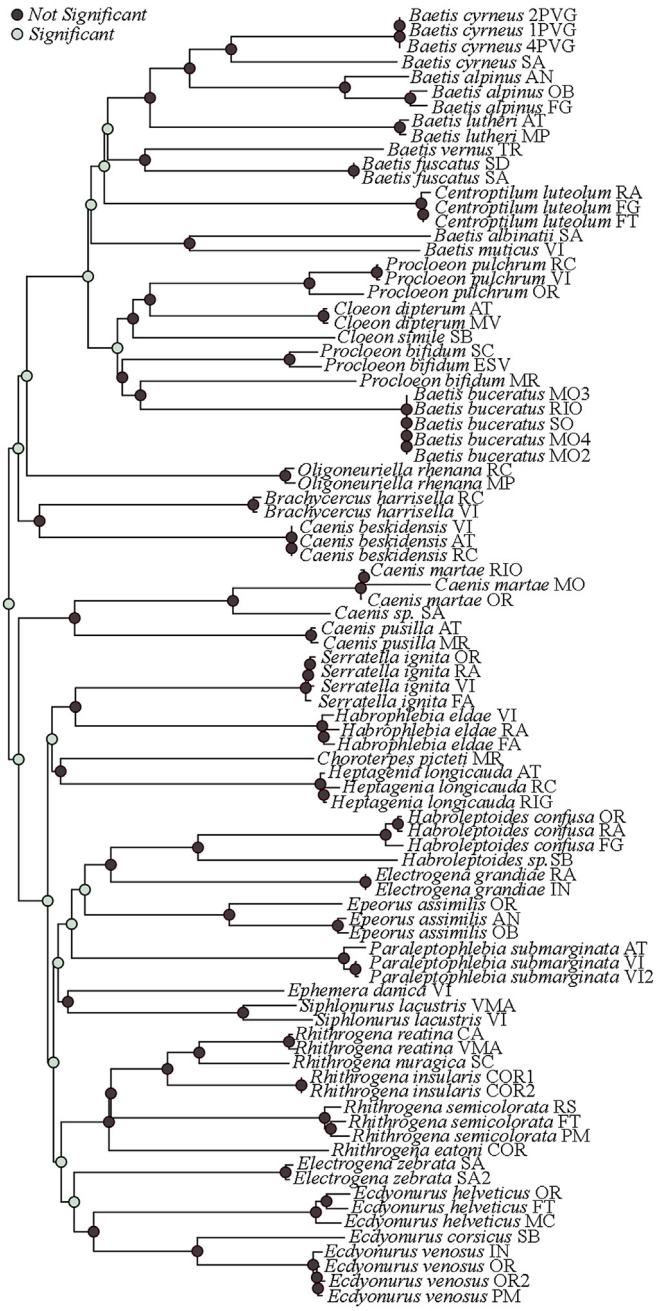

Fig. 3. Neighbor Joining dendrogram of 89 mayfly COI sequences produced with SPIDER under default options and the Rosenberg probability test of reciprocal monophyly (alpha = 0.005). Light grey nodes are significant to 0.05.

The RaxML phylogram displayed a general topology well in line with the acknowledged mayfly taxonomy, and all species clusters were highly supported, except the singleton species; the PTP method identified 45 putative species clusters (online resource III).

3.3. Identifications in a wider context

A taxonomy search (Table 3) revealed that 21 species of our dataset were present on NCBI with their corresponding COI sequence. When blasted, 48 sequences determined in our study (corresponding to 18 species) matched correctly, while 4 sequences (representing *Baetis vernus*, *Baetis lutheri* and *Ephemera danica*) scored the right genus but showed incorrect

Table 3
Taxonomic occurrence (yes/no) and relative highest species match of the obtained 77 mayfly COI haplotypes on GenBank and BOLD. Grey indicates incorrect species matches. * Synonyms of Baetis muticus (Linnaeus, 1758); ** Present only whole genome.

Sequence	Presence	NCBI	Identity	Presence	BOLD	Identity
		Species match			Species match	
[Baetis_muticus_VI]	*Yes**	*Alainites cf. muticus*	98%	*Yes**	*Alainites cf. muticus*	97.88%
[Baetis_vernus_TR]	*Yes*	*Baetis liebenauae*	94%	*Yes*	*Baetis vernus*	96.73%
[Baetis_albinatii_SA]	*No*	*Alainites cf. muticus*	99%	*No*	*Alainites cf. muticus*	99.35%
[Baetis_cyrneus_SA]	*Yes*	*Baetis cf. cyrneus*	90%	*Yes*	*Baetis cf. cyrneus*	90.03%
[Baetis_cyrneus_4PVG, 1PVG, MO]	*Yes*	*Baetis cf. cyrneus*	100%	*Yes*	*Baetis cf. cyrneus*	100%
[Baetis_lutheri_MP]	*Yes*	*Baetis alpinus*	95%	*Yes*	*Baetis alpinus*	94.61%
[Baetis_lutheri_AT]	*Yes*	*Baetis alpinus*	95%	*Yes*	*Baetis alpinus*	94.61%
[Baetis_fuscatus_SA, SD]	*Yes*	*Baetis fuscatus clone*	89%	*Yes*	*Baetis fuscatus clone*	91.34%
[Baetis_alpinus_FG]	*Yes*	*Baetis_alpinus*	95%	*Yes*	*Baetis_alpinus*	95.09%
[Baetis_alpinus_AN]	*Yes*	*Baetis alpinus*	97%	*Yes*	*Baetis alpinus*	96.79%
[Baetis_alpinus_OB]	*Yes*	*Baetis alpinus*	95%	*Yes*	*Baetis alpinus*	95.60%
[Baetis_buceratus_MO3, RIO, SO, MO2, MO4]	*Yes*	*Baetis buceratus*	89%	*Yes*	*Baetis buceratus*	90.36%
[Centroptilum_luteolum_RA]	*Yes*	*Centroptilum_luteolum*	89%	*Yes*	*Centroptilum_luteolum*	89.22%
[Centroptilum_luteolum_FT, FG]	*Yes*	*Centroptilum_luteolum*	89%	*Yes*	*Centroptilum_luteolum*	89.39%
[Procloeon_pulchrum_VI]	*No*	*Procloeon mendax*	85%	*No*	*Cloeon dipterum*	84.94%
[Procloeon_pulchrum_OR]	*No*	*Procloeon mendax*	86%	*No*	*Cloeon dipterum*	86.19%
[Procloeon_pulchrum_RC]	*No*	*Procloeon mendax*	85%	*No*	*Cloeon dipterum*	84.77%
[Procloeon_bifidum_MR]	*No*	*Procloeon mendax*	84%	*Yes*	*Procloeon bifidum*	86.21%
[Procloeon_bifidum_ES]	*No*	*Procloeon mendax*	85%	*Yes*	*Proloeon viridiculare*	86.11%
[Procloeon_bifidum_SC]	*No*	*Unclassified Ephemeroptera*	86%	*Yes*	*Proloeon viridiculare*	86.11%
[Cloeon_dipterum_MV]	*Yes*	*Cloeon dipterum*	99%	*Yes*	*Cloeon dipterum*	99.82%
[Cloeon_dipterum_AT]	*Yes*	*Cloeon dipterum*	99%	*Yes*	*Cloeon dipterum*	99.82%
[Cloeon_simile_SB]	*No*	*Procloeon sp.*	90%	*Yes*	*Procloeon sp.*	90.78%
[Caenis_pusilla_MR]	*No*	*Caenis anceps*	82%	*Yes*	*Caenis pusilla*	98.20%
[Caenis_pusilla_AT]	*No*	*Caenis sp.*	82%	*Yes*	*Caenis pusilla*	98.20%
[Caenis_beskidensis_VI, RC, AT]	*No*	*Caenis anceps*	83%	*Yes*	*Caenis beskidensis*	89.77%
[Caenis_ martae _MO]	*No*	*Leptophlebia*	77%	*Yes*	*Caenis macrura*	79.84%
[Caenis_ martae _OR]	*No*	*Caenis bajaensis*	80%	*Yes*	*Caenis macrura*	81.54%
[Caenis_ martae _RIO]	*No*	*Caenis bajaensis*	81%	*Yes*	*Caenis macrura*	81.70%
[Caenis_sp._SA98]	*No*	*Asionurus sp.*	80%	*No*	*Caenis*	82.68%
[Brachycercus_harrisella_VI]	*Yes*	*Brachycercus harrisella*	94%	*Yes*	*Brachycercus harrisella*	96.09%
[Brachycercus_harrisella_RC]	*Yes*	*Brachycercus harrisella*	94%	*Yes*	*Brachycercus harrisella*	96.24%
[Ecdyonurus_helveticus_FT]	*Yes*	*Ecdyonurus helveticus*	99%	*Yes*	*Ecdyonurus helveticus*	98.69%
[Ecdyonurus_helveticus_OR]	*Yes*	*Ecdyonurus helveticus*	98%	*Yes*	*Ecdyonurus helveticus*	98.20%
[Ecdyonurus_helveticus_MC]	*Yes*	*Ecdyonurus helveticus*	98%	*Yes*	*Ecdyonurus helveticus*	98.37%
[Ecdyonurus_corsicus_SB]	*Yes*	*Ecdyonurus corsicus*	99%	*Yes*	*Ecdyonurus corsicus*	99.67%
[Ecdyonurus_venosus_OR]	*Yes*	*Ecdyonurus venosus*	99%	*Yes*	*Ecdyonurus venosus*	99.67%
[Ecdyonurus_venosus_IN]	*Yes*	*Ecdyonurus venosus*	99%	*Yes*	*Ecdyonurus venosus*	99.84%
[Ecdyonurus_venosus_PM]	*Yes*	*Ecdyonurus venosus*	99%	*Yes*	*Ecdyonurus venosus*	99.51%
[Ecdyonurus_venosus_OR76]	*Yes*	*Ecdyonurus venosus*	100%	*Yes*	*Ecdyonurus venosus*	100%
[Epeorus_assimilis_OB]	*No*	*Epeorus sylvicola*	89%	*Yes*	*Epeorus assimilis*	89.70%
[Epeorus_assimilis_AN]	*No*	*Epeorus sylvicola*	89%	*Yes*	*Epeorus assimilis*	89.87%
[Epeorus_assimilis_OR]	*No*	*Epeorus sylvicola*	89%	*Yes*	*Epeorus assimilis*	90.70%
[Electrogena_grandie_IN, RA]	*No*	*Electrogena sp.*	98%	*No*	*Electrogena sp.*	97.55%
[Electrogena_zebrata_SA104]	*Yes*	*Electrogena zebrata*	99%	*Yes*	*Electrogena zebrata*	99.84%
[Electrogena_zebrata_SA105]	*Yes*	*Electrogena zebrata*	99%	*Yes*	*Electrogena zebrata*	99.67%
[Rhithrogena_semicolorata_PM]	*Yes*	*Rhithrogena semicolorata*	98%	*Yes*	*Rhithrogena semicolorata*	98.37%
[Rhithrogena_semicolorata_FT]	*Yes*	*Rhithrogena semicolorata*	99%	*Yes*	*Rhithrogena semicolorata*	98.69%
[Rhithrogena_semicolorata_RS]	*Yes*	*Rhithrogena semicolorata*	99%	*Yes*	*Rhithrogena semicolorata*	99.84%
[Rhithrogena_reatina_VMA79]	*No*	*Rhithrogena sp.*	93%	*No*	*Rhithrogena sp.*	92.81%
[Rhithrogena_reatina_CA]	*No*	*Rhithrogena sp.*	93%	*No*	*Rhithrogena sp.*	93.14%
[Rhithrogena_nuragica_SC]	*No*	*Rhithrogena sp.*	99%	*No*	*Rhithrogena sp.*	99.02%
[Rhithrogena_insulari_COR100, COR101]	*No*	*Rhithrogena sp.*	90%	*No*	*Rhithrogena sp.*	90.03%
[Rhithrogena_eatoni_OR103]	*No*	*Rhithrogena sp.*	90%	*No*	*Rhithrogena sp.*	89.87%
[Heptagenia_longicauda_RIG]	*No*	*Heptagenia solitaria*	89%	*Yes*	*Eptagenia longicauda*	95.42%
[Heptagenia_longicauda_AT]	*No*	*Heptagenia solitaria*	89%	*Yes*	*Eptagenia longicauda*	95.75%
[Heptagenia_longicauda_RC]	*No*	*Heptagenia sulfura*	87%	*Yes*	*Eptagenia longicauda*	94.77%
[Habroleptoides_confusa_FG]	*Yes*	*Habroleptoides confusa*	84%	*Yes*	*Habroleptoides confusa*	96.50%
[Habroleptoides_confusa_RA]	*Yes*	*Habroleptoides confusa*	84%	*Yes*	*Habroleptoides confusa*	96.83%
[Habroleptoides_confusa_OR]	*Yes*	*Habroleptoides confusa*	84%	*Yes*	*Habroleptoides confusa*	97.17%
[Habroleptoides_sp.SB]	*Yes*	*Habroleptoides modesta*	92%	*Yes*	*Habroleptoides modesta*	91.67%
[Habrophlebia_eldae_VI]	*Yes*	*Habrophlebia eldae*	99%	*Yes*	*Habrophlebia eldae*	99.35%
[Habrophlebia_eldae_FA]	*Yes*	*Habrophlebia eldae*	99%	*Yes*	*Habrophlebia eldae*	99.51%
[Habrophlebia_eldae_RA]	*Yes*	*Habrophlebia eldae*	99%	*Yes*	*Habrophlebia eldae*	99.35%
[Paraleptophlebia_submarginata_AT]	*No*	*Paraleptophlebia westoni*	82%	*Yes*	*Paraleptophlebia submarginata*	99.18%
[Paraleptophlebia_submarginata_VI]	*No*	*Siphlonurus sp.*	81%	*Yes*	*Paraleptophlebia submarginata*	98.69%
[Paraleptophlebia_submarginata_VI36]	*No*	*Siphlonurus sp.*	81%	*Yes*	*Paraleptophlebia submarginata*	98.85%

(*continued on next page*)

Table 3 (*continued*)

Sequence	Presence	NCBI	Identity	Presence	BOLD	Identity
		Species match			Species match	
[Choroterpes_picteti_MR77]	*No*	*Habroleptoides sp.*	85%	*Yes*	*Choroterpes picteti*	99.35%
[Serratella_ignita_VI]	*Yes*	*Serratella ignita*	99%	*Yes*	*Serratella ignita*	99.67%
[Serratella_ignita_RA]	*Yes*	*Serratella ignita*	99%	*Yes*	*Serratella ignita*	99.84%
[Serratella_ignita_FA]	*Yes*	*Serratella ignita*	99%	*Yes*	*Serratella ignita*	99.51%
[Serratella_ignita_OR]	*Yes*	*Serratella ignita*	99%	*Yes*	*Serratella ignita*	99.84%
[Oligoneuriella_rhenana_MP]	*No*	*Anopheles labranchiae*	85%	*Yes*	*Oligoneuriella rhenana*	89.30%
[Oligoneuriella_rhenana_RC]	*No*	*Anopheles labranchiae*	84%	*Yes*	*Oligoneuriella rhenana*	90%
[Siphlonurus_lacustris_VI]	*Yes*	*Siphlonurus lacustris*	92%	*Yes*	*Siphlonurus croaticus*	94.28%
[Siphlonurus_lacustris_VMA]	*Yes*	*Siphlonurus lacustris*	93%	*Yes*	*Siphlonurus croaticus*	95.10%
[Ephemera_danica_VI]	*Yes*	*Ephemera sachalinensis***	89%	*Yes*	*Ephemera danica*	98.36%

species matches. The remaining 37 sequences were deposited in GenBank for the first time, to represent 18 additional Ephemeroptera species. The same search showed that 31 of our species had COI barcodes present on BOLD; of these, 27 species matched correctly. Four species resulted misidentified: *Baetis lutheri* (matching *Baetis alpinus*, as in GenBank), *Siphlonurus lacustris* (matching *Siphlonurus croaticus*), *Cloeon simile* (that matched another genus, *Procloeon*) and *Procloeon bifidum*, for which one of the three determined COI sequences matched correctly whereas the other two scored another species (*P. viridoculare*). In general, a large part of the highest hits was scored with <90% identity.

Interestingly, one specimen of *Habroleptoides* (*Habroleptoides* sp.) from Rio Tistigliosi (Sardinia), damaged to the point that the field species identification was prevented, matched an endemism of Sardinia (*Habroleptoides modesta*) on both GenBank and BOLD.

Integration of our sequence data with mayfly barcodes presently deposited in BOLD produced a 2622 sequence matrix; the final multiple alignment included 570 complete sequences (>500 bp) assigned to the same families and genera investigated in this work (over 200 total species; taxa list reported as Online Resource IV).

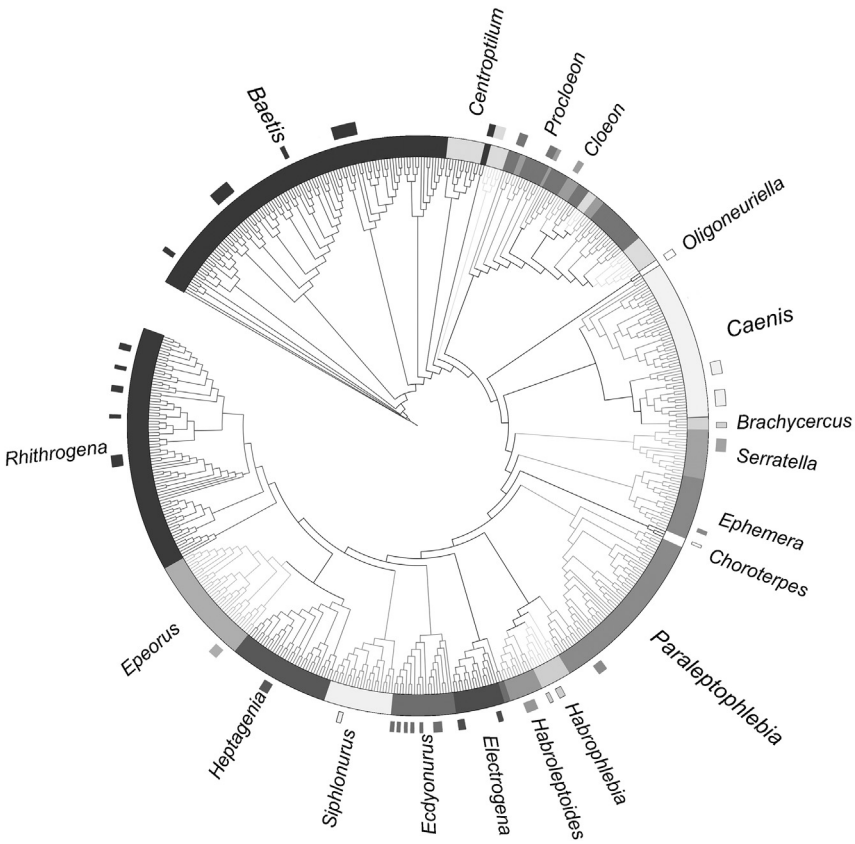

Fig. 4. Circular RaxML dendrogram of 570 mayfly COI sequences including BOLD barcodes and the investigated dataset. Bars indicate the positioning of sequences determined in this study. Different genus clusters are indicated with greyscale and the respective names.

The resulting ML profile showed a topology in agreement with the recognized taxonomy of mayflies (Fig. 4; higher resolution dendrogram with sequence details available as Online Resource V). Seven families formed cohesive groups, the only exception being family Heptageniidae, which appeared paraphyletic and included Siphlonuridae. Across the whole dataset, all genera that were represented by > 2 species formed cohesive groups, with the exception of the four genera within Baetidae. In this family, 2 single sequences of *Baetis albinatii* and *Baetis muticus* (from our dataset, out of the 30 total *Baetis* species) clustered in a paraphyletic subgroup comprising *Centroptilum* (7 species), *Cloeon* (5 species) and *Procloeon* (14 species); these three genera were also mixed, and both sequences from our dataset and BOLD infiltrated the different specific clusters. Noteworthy, *Baetis albinatii* and *B. muticus* (syn. *Alainites albinatii* and *A. muticus*) have been recently assigned to genus *Takobia* (Kluge and Novikova, 2014). Two other genera appeared non-monophyletic: *Paraleptophlebia* with four species out of 20 clustering with *Habrophlebia* and *Habroleptoides*, and *Ecdyonurus* with one species clustering with *Electrogena*. In all other cases, all species grouped correctly with their appropriate family and genus. The sequences in our dataset always formed uniform species clusters, including conspecific BOLD barcodes when these were available. However, several instances of species paraphyly were observed across all families.

4. Discussion

4.1. Potential of COI-based DNA barcoding for mayfly species discrimination

Specific sequence flags and diversity delimitation are essential prerequisites for DNA barcode-based species identification. In this work, the evaluation of DNA barcoding efficacy to achieve correct species identification was based on the detection of four main parameters: (1) haplotype specificity (the uniqueness of individual COI sequences allowing discrimination of samples and assignation of specimens to different species); (2) DNA barcoding gaps (the absence of overlaps between intra- and interspecific genetic distances within genera); (3) monophyly estimation of species clusters under different criteria; (4) correct sequence/species match on available public repositories (NCBI, BOLD).

Our results indicate that the COI-based identification system of mayflies in the study area was extremely effective to detect specific haplotypes across all species. All taxa (100%) in our dataset were discriminated based on unique sequence character states, and the haplotypes generated with a 2% distance cut-off (which is the standard threshold suggested for species delimitation; Hebert et al., 2003) were never shared among different species. At the same time, identity matches on the available reference databases (ca. 100% of genera and 90% of species correct matches, considering taxa occurrence relatively to our dataset) and the Neighbor Joining reconstruction (in which all COI barcodes grouped into species clusters according to taxonomy) provided further confirmations of the COI efficacy to discriminate specimens in large congruence with species nomenclature. Notable cases in which DNA barcoding was determinant for species identifications or discrimination of difficult taxa are those provided by a damaged sample of *Habroleptoides modesta* and by the clear separation of *Baetis albinatii* from *B. muticus*, in spite of the very few and often ambiguous diagnostic traits available (e.g., the absence of first gill; Sartori and Thomas, 1989). DNA barcoding data also supported the taxonomic split of genus *Caenis* into three morphologically highly different species groups, evidencing the high haplotype divergence of the three species (each one representing a different group). The Neighbor Joining analysis also demonstrated a good agreement of the produced COI dataset with the acknowledged mayfly taxonomy, although some discrepancies were observed at higher taxonomic levels (paraphyly of genera and families). These discrepancies were most probably due to taxonomic undersampling, to the inappropriate use of the Neighbor Joining as an analytical tool, and to the model of sequence evolution (K2P) used. Both the NJ and the K2P methods are considered “standard” in DNA barcoding (Hebert et al., 2003; Zhou et al., 2011) but probably non-optimal (Collins and Cruickshank, 2013; White et al., 2014). For instance, a parallel RaxML analysis run with the GTRCAT model on our dataset showed a highly supported topology, more congruent with mayfly major taxonomic ranks (Online Resource III). However, the Rosenberg's test of reciprocal monophyly, especially useful to assign a role to the sampling regime in relation to the production of distinct lineage groups (Rosenberg, 2007), indicated that the monophyly of the produced species clusters was probably due to chance, suggesting further sampling to assess this important (but not fundamental; see Collins and Cruickshank, 2013) requisite of species discrimination with DNA barcoding. In agreement with this, the ML reconstruction expanded with BOLD-harvested barcodes accommodated all our sequences in a well-supported, coherent representation of Ephemeroptera taxonomy. Interestingly, both our sequences and the BOLD barcodes contributed equally to evidence some paraphyly in a few genera and especially, in the family Baetidae. All these data certainly contribute to support the efficacy of the COI barcodes in the study area to (1) discriminate taxa, (2) recognize major taxonomic schemes in Ephemeroptera, as in prior studies (Ball et al., 2005; Zhou et al., 2011; Webb et al., 2012), and further highlight the importance of the intra-/interspecific sampling extent to achieve statistically sound and reliable results.

Indeed, the high efficacy of sample identification, the documented barcoding gaps, the exaggerated intra-/interspecific genetic distances and the indication of the non-significant species monophyly, in our dataset were very probably affected by the limited intra-interspecific sampling and the restricted area investigated (Meyer and Paulay, 2005; Wiemers and Fiedler, 2007). Bergsten et al. (2012) demonstrated that intraspecific genetic variation generally increases with spatial scale, and that interspecific divergences decrease with expanded samplings. As a result, the uncertainty of identifications increases with geographical and sampling coverage. This is certainly relevant for taxonomic studies based on vast DNA barcoding initiatives, with the target of detecting a representative amount of the genetic variation of the investigated species, to describe the existing diversity and facilitate DNA based species delimitation at a global level (www.ibol.org). However, for many

applications of DNA barcoding such as life-stage or gender association, biodiversity surveys and environmental monitoring, the sampling- and scale-dependency are of limited importance and rely more directly on the availability of local data reference libraries (Bergsten et al., 2012). This issue gives an incentive for regional and national barcoding campaigns, especially from understudied regions, producing reference benchmarks of relevant local species assemblages, independently of the sampling numbers, and even assigns a special relevance to singletons (Lim et al., 2012). In this view, we must note that a large part of the highest hits on the public reference databases was scored with <90% identity, and a few instances of incorrect species matches were scored. Besides the obvious possibility of accidental sample misidentification (either in our dataset or in databases), we retain that both results could be strongly related with the general availability of only extra-continental co-specific samples in the consulted libraries, a phenomenon that urges a rapid enrichment of global databases with Euro-Mediterranean samples and/or the complementary development of a regional reference library. Indeed, our work sets within this context, with an initial attempt at developing a regional database, and further provides multiple methodological tests to achieve maximal precision of query identification/species delimitation in practice.

4.2. DNA taxonomy: barcode clusters vs. morphological species

A higher number of Molecular Operational Taxonomic Units (MOTUs, Jones et al., 2011) than the morphologically recognized species was observed under the criterion of >2% sequence divergence. This is a standard threshold for species identification with barcodes in insect groups, and suggests that intra-specific sequence divergences >2% could be indicative of the existence of cryptic species (Hebert et al., 2004; Smith et al., 2013; Jackson et al., 2014). In our dataset, intra-specific divergences exceeding the 2% cut-off were found in 65% of the species, with *Procleon bifidum, Baetis cyrneus* and *Epeorus assimilis* showing extremely high values (>10%). Interestingly, a higher number of putative species than the investigated morphospecies was also predicted by the PTP approach, where apart from *Baetis cyrneus* and *Procloeon bifidum* (separated on different macroclusters), also *Epeorus assimilis*, *Baetis alpinus*, *Procloeon pulchrum*, and *Siphlonurus lacustris* were split in different species clusters.

Prior studies on mayflies (Ball et al., 2005; Webb et al., 2012) and other insect groups including Lepidoptera, Hymenoptera, Trichoptera and Diptera (Hebert et al., 2004; Smith et al., 2006; Zhou et al., 2009) have shown that such large COI divergences regularly reflect species complexes or unrecognized species. The same conclusion could be drawn for some of the taxa investigated here, although other factors such as the limited geographic and sampling scale (Bergsten et al., 2012) and population history (Theissinger et al., 2011) should be certainly explored. For instance, the uncertainty about the generic status of *Procloeon* and, more in general, the controversial taxonomy of all genera belonging to subfamily "Cloeoninae" are well known (Gillies, 1991; Kluge, 1997), and clearly require a critical revision. In our work, the absence of the barcoding gap in *Procleon* was caused by the two samples of *P. bifidum* from Sardinia and Elba islands sharing an extremely high K2P genetic distance with the specimen from mainland Italy (19.6%), that was higher than the inter-specific distance with *P. pulchrum* (18.7%). This suggests that *P. bifidum,* despite an apparent morphological uniformity, could include a cryptic species endemic to the Corse-Sardinian biogeographic region. Also the high intra-specific K2P distance (14.9%) detected in another endemism of Corse and Sardinia (*B. cyrneus*) would require further studies to check for cryptic variation, as recently suggested also by Gattolliat et al. (2015). It must be noted that *Baetis* is particularly variable (Ball et al., 2005; Webb et al., 2012), and the occurrence of cryptic species has already been achieved in specific studies (e.g., *Baetis rhodani,* Williams et al., 2006). The third case of intra-specific divergence >10% (*Epeorus assimilis*) was unexpected, and in contrast with the uniform morphological characters displayed by the specimens and the relative vicinity of the sampling sites. This finding would therefore suggest the necessity of a closer inspection of the species morphology and/or ecological traits of the sampled population/river to assess the significance of the found variation. Congruently, the large paraphyly and the high occurrence of false positives/negatives resulting from the analysis with Spider indicated *Procloeon* and *Baetis* as the genera having species with the most critical status (i.e., cryptic species or species complexes). However, we should also note the extremely high barcoding gap found in *Electrogena* (23.5–24.0%). For all the other genera in our dataset (*Procloeon* and *Baetis* excluded), Spider provided an optimal barcoding gap threshold of 11% (Online Resource II). Intra-generic barcoding gaps higher than 20% were never or only rarely described in previous mayfly barcoding studies (Ball et al., 2005; Webb et al., 2012), suggesting that *Electrogena grandiae* and *Electrogena zebrata* might belong to different genera, as previously promoted by Gaino et al. (1987) and discussed by Bauerfeind and Soldàn (2012). This would be supported by the NJ dendrogram, where the two investigated species grouped into different clusters, and in the ML reconstruction, where all the *E. zebrata* COI barcodes clustered separately from the other *Electrogena* species. Notably, egg and winged stage traits of *E. zebrata* (a Corse-Sardinian endemism) are markedly different from other members of the same genus (Gaino et al., 1987).

If confirmed with a larger species sampling, our results would therefore indicate that the taxonomy of mayflies in the study area would need a re-examination, to correlate differentiation among lineages with morpho-ecological traits; alternatively, the use of a single, arbitrary percentage discontinuity to match mayfly species concept should be re-evaluated, and the most appropriate genetic distance ranges for species discrimination in each family or genus should be carefully determined (cf. Krishnamurthy and Francis, 2012; Collins and Cruickshank, 2013). In any case, it is necessary that a general improvement of mayfly taxonomy should be pursued with a multidisciplinary approach. The discovery of new/cryptic species and nomenclatural revisions through highly divergent barcode haplogroups should correlate with additional evidences, including morphological, ecological, biogeographical descriptors and independent molecular markers (Monaghan and Sartori, 2009; Jackson et al., 2014). This is also true for the various instances of family and genera paraphyly or

incongruent groupings resulting in the ML dendrogram, in agreement with the continued revision of Ephemeroptera taxonomy on a global scale of the last decades. Besides synonymy (Meyer and Paulay, 2005) and imperfect group taxonomy (Hendrich et al., 2010), other possible reasons for the detected non-monophyly are incomplete lineage sorting and introgressive hybridization (Funk and Omland, 2003; Zhang et al., 2012).

The most outstanding result of this study leads us to suggest that the above species as noteworthy cases require additional sampling and deeper insight into taxonomical issues so as to draw more solid conclusions about their status and diversity; at the same time (see also Packer et al., 2009), collected data speak in favor of broader insights into local biodiversity patterns in mayflies and the need to re-evaluate species taxonomy with a better integration of morphological evidence and molecular data. In this sense, an optimization of the basic operational guidelines (e.g., adequate sampling, detailed reference databases, robust and/or complementary statistical tools) is certainly needed to ensure the necessary accuracy in future practical applications, where a genus or a species misidentification can be critical for biomonitoring or bio-assessment purposes (Zhou et al., 2011).

4.3. Implications for freshwater biomonitoring

Freshwater ecosystems response is particularly sensitive to the different types of stress induced by global warming and anthropogenic impact (Sala et al., 2000; Dudgeon et al., 2006). Population dynamics and biodiversity can be profoundly modified by alterations of the temperature (Wrona et al., 2006), and further stressors such as eutrophication, organic pollution, acidification, channel and bank modifications. Periodic control of the health status of water bodies is therefore of paramount importance to ensure global biodiversity and all the related services to humans (Thomas et al., 2004; Montoya and Raffaelli, 2010). Freshwaters biomonitoring is the process of evaluating the overall ecological quality of water habitats based on the composition and changes in the relative abundance of the macrobenthos community, an important part of which are mayflies (Hering et al., 2006). Taxonomic surveys are usually produced by experts and the resulting reports are used in protection programs as part of environmental monitoring or permit-required compliance assessments (WFD 2000/60/CE). Several factors may combine to prevent full access to specimen identification and slow down or affect the evaluation of habitat and water quality (Lenat and Resh, 2001; Stribling et al., 2008); moreover, limited availability of experienced taxonomists may influence the generation of incomplete assessments, restricted geographical coverage and delayed frequency of reports.

In this work, DNA barcoding of the mayfly fauna collected in a study area demonstrated to be a fast, effective and reliable tool for local species identification. It also indicated an increase of multiple species lineages than those identified based on morphology alone, providing a necessary starting point towards the establishment of a regional reference library for freshwater biodiversity. All this would support the implementation of taxonomic metrics for the activation of bio-assessment and biomonitoring programs of Mediterranean freshwaters based on DNA barcoding protocols.

Previous studies clearly showed that our knowledge of freshwater biodiversity and of its response to environmental stressors can be strongly improved by DNA barcoding. In particular, the ability to identify subtle changes in community composition (e.g., taxa abundance, cryptic species and multiple species lineages) would be fundamental to detect biodiversity loss, understand how communities respond to specific stressors, timely assess impacts to freshwaters (i.e., before a serious degradation occurs), and provide more rapid measurement of the rate and degree of habitat recovery in response to improved management or restoration practices (Sweeney et al., 2011; Jackson et al., 2014). In addition, recent technological advances allowing faster and cheaper DNA sequencing (e.g., next-generation sequencing, NGS) would certainly benefit from the established local reference libraries, and contribute to expanding the amount of information from the environment (Hajibabaei et al., 2012; Shokralla et al., 2012). This would be of great relevance especially in consideration of the increasing water quality monitoring programs in many parts of the world (including the EU), all requiring a wider number of sampling sites and a larger amount of data (Stein et al., 2014).

In view of this, DNA barcoding also proved its efficacy in the selected model area. Indeed, Central Italy, Sardinia and Corsica are rich in biodiversity, including several micro-endemic and rare species, still to be substantiated (Trizzino et al., 2014; Gattolliat et al., 2015). The local water bodies which face increasing pollution, demographic pressure and climate change, urge monitoring programs to comply with the European guidelines on water and nature protection (WFD 2000/60/CE; 92/43/EEC). In synergy with DNA-based biomonitoring, the compilation of barcode-implemented Red List species could provide useful inputs to enhance the identification of potential protected areas and assign priorities to local conservation efforts, although this approach has rarely been applied to freshwater ecosystems (Strayer and Dudgeon, 2010). We therefore believe that the development of a common DNA barcode strategy for freshwaters biomonitoring in the Mediterranean context could stimulate a successful and sustained collaboration between taxonomists and molecular biologists and corroborate conservation efforts. The establishment of a public web portal to host growing databases on species, barcodes, investigated areas and related information is essential to achieve this goal. In this sense, regional and/or national reference libraries collecting species barcodes and related ecological information are developing in Canada (Zhou et al., 2010) and the USA (Webb et al., 2012). However, our simulation of a practical barcode identification scenario on available reference libraries demonstrated the need to enrich such databases with more species and wider geographical samplings, to allow more precise identifications. Therefore, one of the major research challenges in the future will be certainly to solidify the acquired theoretical knowledge with practical experience in standard DNA barcoding protocols and workflows, with special regard to database development and/or improvement.

Acknowledgements

Funding for this research was provided by the Ph.D. School in “Ecology and Management”.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.bse.2015.07.035.

References

Ball, S., Hebert, P.D.N., Burian, S., Webb, J., 2005. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. J. North Am. Benthol. Soc. 24, 508–524. http://dx.doi.org/10.1899/0887.

Barber-James, H.M., Gattolliat, J.L., Sartori, M., Hubbard, M.D., 2008. Global diversity of mayflies (ephemeroptera, insecta) in freshwater. Hydrobiologia 595, 339–350. http://dx.doi.org/10.1007/s10750-007-9028-y.

Bauerfeind, E., Soldán, T., 2012. The Mayflies of Europe (Ephemeroptera). Apollo Books, Ollerup, p. 781.

Belfiore, C., 2005. Ephemeroptera. In: Ruffo, S., Stoch, F. (Eds.), Checklist e distribuzione della fauna italiana. Memorie del Museo Civico di Storia Naturale di Verona, pp. 127–129.

Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M.T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., Ribera, I., Nilsson, A.N., Barraclough, T.G., Vogler, A.P., 2012. The effect of geographical scale of sampling on DNA barcoding. Syst. Biol. 61, 851–869. http://dx.doi.org/10.1093/sysbio/sys037.

Blondel, J., Aronson, J., 1999. Biology and Wildlife of the Mediterranean Region. Oxford University Press, Oxford, UK.

Brittain, J.E., Sartori, M., 2003. Ephemeroptera (mayflies). In: Resh, V.H., Cardé, R.T. (Eds.), Encyclopedia of Insects. Academic Press, San Diego, pp. 373–380.

Brown, S.D.J., 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. http://dx.doi.org/10.1111/j.1755-0998.2011.03108.x.

Collins, R.A., Cruickshank, R.H., 2013. The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. 13, 969–975.

Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85, 407–415.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.-I., Knowler, D.J., Lé VeQue, C., Naiman, R.J., Prieur-Richard, A.-H., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biol. Rev. 81, 163–182. http://dx.doi.org/10.1017/S1464793105006950.

Edmunds, G.F., McCafferty, W.P., 1988. The mayfly subimago. Annu. Rev. Entomol. 33, 509–529.

Ekrem, T., Willassen, E., Stur, E., 2007. A comprehensive DNA sequence library is essential for identification with DNA barcodes. Mol. Phylogenet. Evol. 43, 530–542. http://dx.doi.org/10.1016/j.ympev.2006.11.021.

Elderkin, C., Corkum, L.D., Bustos, C., Cunningham, E.L., Berg, D.J., 2012. DNA barcoding to confirm morphological traits and determine relative abundance of burrowing mayfly species in western lake erie. J. Gt. Lakes. Res. 38, 180–186. http://dx.doi.org/10.1016/j.jglr.2011.11.010.

European Commission, 2000. Directive 2000/60/EC of the European parliament and of the council of 23 october 2000 establishing a framework for community action in the field of water policy. Off. J. Eur. Commun. L. 327, 1–72.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytpchrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3, 294–299.

Funk, D.J., Omland, K.E., 2003. Species-level monophyly and paraphyly: frequency, causes and consequences with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst 34, 397–423.

Gaino, E., Belfiore, C., Mazzini, M., 1987. Ootaxonomic investigation of the Italian species of the genus *Electrogena* (Ephemeroptera, Heptageniidae). B. Zool. 54, 169–175. http://dx.doi.org/10.1080/11250008709355578.

Gattolliat, J.L., Cavallo, E., Vuataz, L., Sartori, M., 2015. DNA barcoding of corsican mayflies (Ephemeroptera) with implications on biogeography, systematics and biodiversity. Arthropod Syst. Phylo 73 (1), 3–18.

Gillies, M.T., 1991. A diphyletic origin for the two-tailed baetid nymphs occurring in East African stony streams with a description of the new genus and species Tanzaniella spinosa gen. nov. sp. nov. In: Alba- Tercr, J., Sanchez-Ortega, A. (Eds.), Overview and Strategies of Ephemeroptera and Plecoptera. Sandhill Crane Press, Gainesville, Florida, 75–187.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W., Hebert, P.D.N., 2006. DNA barcodes distinguish species of tropical lepidoptera. Proc. Natl. Acad. Sci. U. S. A. 103, 968–971. http://dx.doi.org/10.1073/pnas.0510466103.

Hajibabaei, M., Singer, G.A.C., Clare, E.L., Hebert, P.D.N., 2007. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol. 5, 24. http://dx.doi.org/10.1186/1741-7007-5-24.

Hajibabaei, M., Spall, J.L., Shokralla, S., Van Konynenburg, S., 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through nondestructive environmental barcoding of DNA from preservative ethanol. BMC Ecol. 12, 28. http://dx.doi.org/10.1186/1472-6785-12-28.

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A., Baird, D.J., 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One 6 (4), e17497. http://dx.doi.org/10.1371/journal.pone.0017497.

Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R., 2003. Biological identifications through DNA bar codes. Proc. R. Soc. Lond 270, 313–321. http://dx.doi.org/10.1098/rspb.2002.2218.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. U. S. A. 101, 14812–14817. http://dx.doi.org/10.1073/pnas.0406166101.

Hendrich, L., Pons, J., Ribera, I., Balke, M., 2010. Mitochondrial cox1 sequence data reliably uncover patterns of insect diversity but suffer from high lineage-idiosyncratic error rates, p. e14448. Plos ONE 5.

Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K., Verdonschot, P.F.M., 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshw. Biol. 51, 1757–1785. http://dx.doi.org/10.1111/j.1365-2427.2006.01610.x.

Hogg, I.D., Smith, B.J., Banks, J.C., Dewaard, J.R., Hebert, P.D.N., 2009. Testing use of mitochondrial COI sequences for the identification and phylogenetic analysis of New Zealand caddisflies (Trichoptera). New. zeal. J. Mar. Fresh 43, 1137–1146. http://dx.doi.org/10.1080/00288330.2009.9626536.

IUCN, 2007. IUCN Red List of Threatened Species. International Union for Conservation of Nature and Natural Resources, Cambridge, UK. Available from: www.iucnredlist.org.

Jackson, J.K., Battle, J.M., White, B.P., Pilgrim, E.M., Stein, E.D., Miller, P.E., Sweeney, B.W., 2014. Cryptic biodiversity in streams: a comparison of macro-invertebrate communities based on morphological and DNA barcode identification. Freshw. Sci. 33, 312–324. http://dx.doi.org/10.1086/675225.

Janzen, D.H., Hajibabaei, M., Burns, J.M., Hallwachs, W., Remigio, E., Hebert, P.D.N., 2005. Wedding biodiversity inventory of a large and complex lepidoptera fauna with DNA barcoding. Philos. Trans. R. Soc. Lond 360. http://dx.doi.org/10.1098/rstb.2005.1715, 1835–184.

Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ. Rev. 16, 45–69.

Jones, M., Ghoorah, A., Blaxter, M., 2011. jMOTU and taxonerator: turning DNA barcode sequences into annotated operational taxonomic units. PLoS One 6, e19259. http://dx.doi.org/10.1371/journal.pone.0019259.

Kluge, N.J., 1997. Classification and phylogeny of the baetidae (Ephemeroptera) with description of the new species from the upper cretaceous resins of taimyr. In: Landolt, P., Sartori, M. (Eds.), Ephemeroptera & Plecoptera. Biology-ecology-systematics. Mauron+Tinguely & Lachat, Fribourg, pp. 527–535.
Kluge, N.J., 2004. The Phylogenetic System of Ephemeroptera (The First Experience in Consistently Non-ranking Taxonomy). In: Ephemeroptera Except for *Turbanoculata* and *Leptophlebia*/fg1, vol. 1. Kluwer Academic Publishers, Dordrecht-Hardbound, p. 456.
Kluge, N.J., Novikova, E.A., 2014. Systematics of Indobaetis Müller-Liebenau & Morihara 1982, and related implications for some other Baetidae genera (Ephemeroptera). Zootaxa 3835 (2), 209–236.
Krishnamurthy, P.K., Francis, P.A., 2012. A critical review on the utility of DNA barcoding in biodiversity conservation. Biodivers. Conserv. 21, 1901–1919. http://dx.doi.org/10.1007/s10531-012-0306-2.
Lenat, D.R., Resh, V.H., 2001. Taxonomy and stream ecology: the benefits of genus- and species-level identifications. J. North Am. Benthol. Soc. 20, 287–298.
Lim, G.S., Balke, M., Meier, R., 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. Syst. Biol. 61, 165–169.
Lugo-Ortiz, C.R., McCafferty, W.P., Waltz, R.D., 1999. Definition and reorganization of the genus *Pseudocloeon* (ephemeroptera: baetidae) with new species descriptions and combinations. T. Am. Entomol. Soc. 125, 1–37.
Meyer, C., Paulay, G., 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3, e422.
Meyer, R., Zhang, G.Y., 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. http://dx.doi.org/10.1080/10635150802406343.
Monaghan, M.T., Sartori, M., 2009. Genetic contributions to the study of taxonomy, ecology, and evolution of mayflies (Ephemeroptera): review and future perspectives. Aquat. Insects 31, 19–39. http://dx.doi.org/10.1080/01650420902734145.
Monk, W.A., Wood, P.J., Hannah, D.M., Extence, C.A., Chadd, R.P., Dunbar, M.J., 2012. How does macroinvertebrate taxonomic resolution influence ecohydrological relationships in riverine ecosystems. Ecohydrology 5, 36–45.
Montoya, J.M., Raffaelli, D., 2010. Climate change, biotic interactions and ecosystem services. Philos. Trans. R. Soc. B 365, 2013–2018. http://dx.doi.org/10.1098/rstb.2010.0114.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858. http://dx.doi.org/10.1038/35002501.
Ogitani, M., Sekine, K., Tojo, K., 2011. Habitat segregation and genetic relationship of two heptageniidae mayflies, *Epeorus latifolium* and *Epeorus l-nigrus*, in the shinano-gawa river basin. Limnology 12, 117–125. http://dx.doi.org/10.1007/s10201-010-0328-y.
Packer, L., Gibbs, J., Sheffield, C., Hanner, R., 2009. DNA barcoding and the mediocrity of morphology. Mol. Ecol. Resour. 9, 42–50. http://dx.doi.org/10.1111/j.1755-0998.2009.02631.x.
Pennisi, E., 2003. Modernizing the tree of life. Science 300, 1692–1697. http://dx.doi.org/10.1126/science.300.5626.1692.
Polunin, N.V.C., 2008. Aquatic Ecosystems: Trends and Global Prospects. Cambridge University Press, Cambridge, UK.
Rach, J., DeSalle, R., Sarkar, I.N., Schierwater, B., Hadrys, H., 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in odonata. P. Roy. Soc. Lond. B 275, 237–247. http://dx.doi.org/10.1098/rspb.2007.1290.
Rivera, J., Currie, D.C., 2009. Identification of nearctic black flies using DNA barcodes (diptera: simuliidae). Mol. Ecol. Resour. 9, 224–236. http://dx.doi.org/10.1111/j.1755-0998.2009.02648.x.
Rosenberg, N.A., 2007. Statistical tests for taxonomic distinctiveness from observations of monophyly. Evolution 61 (2), 317–323. http://dx.doi.org/10.1111/j.1558-5646.2007.00023.x.
Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., et al., 2000. Global biodiversity scenarios for the year 2100. Science 287, 1770–1774. http://dx.doi.org/10.1126/science.287.5459.1770.
Sartori, M., Thomas, A.G.B., 1989. Contribution à la connaissance du genre *Baetis Leach*, 1815 en Corse (Ephemeroptera, Baetidae). *B. albinatii* nov. sp. du groupe *muticus* (L.). Ann. Limnol. 25 (2), 131–137.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., et al., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541. http://dx.doi.org/10.1128/AEM.01541-09.
Sheffield, C.S., Hebert, P.D.N., Kevan, P.G., Packer, L., 2009. DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Mol. Ecol. Resour. 9, 196–207. http://dx.doi.org/10.1111/j.1755-0998.2009.02645.x.
Shokralla, S., Spall, J.L., Gibson, J.F., Hajibabaei, M., 2012. Next-generation sequencing technologies for environmental DNA research. Mol. Ecol. Resour. 21, 1794–1805. http://dx.doi.org/10.1111/j.1365-294X.2012.05538.x.
Smith, M.A., Rodriguez, J.J., Whitfield, J.B., et al., 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proc. Natl. Acad. Sci. U. S. A. 105, 12359–12364. http://dx.doi.org/10.1073/pnas.0805319105.
Smith, M.A., Fernandez-Triana, J.L., Eveleigh, E., Gomez, J., Guclu, C., Hallwacks, W., Hebert, P.D.N., Hrcek, J., Huber, J.T., Janzen, D., Mason, P.G., Miller, S., Quicke, D.L.J., Rodriguez, J.J., Rougerie, R., Shaw, M.R., Varkonyi, G., Ward, D.F., Whitfield, J.B., Zaldivar-Riveron, A., 2013. DNA barcoding and the taxonomy of microgastrinae wasps (hymenoptera, braconidae): impacts after 8 years and nearly 20 000 sequences. Mol. Ecol. Resour. 13, 168–176. http://dx.doi.org/10.1111/1755-0998.12038.
Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W., Hebert, P.D.N., 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (diptera: tachinidae). Proc. Natl. Acad. Sci. U. S. A. 103, 3657–3662. http://dx.doi.org/10.1073/pnas.0511318103.
Sroka, P., 2012. Systematics and phylogeny of the west palaearctic representatives of subfamily baetinae (insecta: ephemeroptera): combined analysis of mitochondrial DNA sequences and morphology. Aquat. Insect 34, 23–53. http://dx.doi.org/10.1080/01650424.2012.718081.
Ståhls, G., Savolainen, E., 2008. MtDNA COI barcodes reveal cryptic diversity in the *Baetis vernus* group (ephemeroptera, baetidae). Mol. Phylogenet. Evol. 46, 82–87. http://dx.doi.org/10.1016/j.ympev.2007.09.009.
Stamatakis, A., 2006. Raxml-vi-hpc: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22 (21), 2688–2690.
Stein, E.D., Martinez, M.C., Stiles, S., Miller, P.E., Zakharov, E.V., 2014. Is dna barcoding actually cheaper and faster than traditional morphological methods: results from a survey of freshwater bioassessment efforts in the United States? PLoS ONE 9 (4), e95525. http://dx.doi.org/10.1371/journal.pone.0095525.
Strayer, D.L., Dudgeon, D., 2010. Freshwater biodiversity conservation: recent progress and future challenges. J. North Am. Benthol. Soc. 29, 344–358. http://dx.doi.org/10.1899/08-171.1.
Strayer, D.L., 2006. Challenges for freshwater invertebrate conservation. J. North Am. Benthol. Soc. 25, 271–287. http://dx.doi.org/10.1899/0887-3593.
Stribling, J.B., Pavlik, K.L., Holdsworth, S.M., Leppo, E.W., 2008. Data quality, performance, and uncertainty in taxonomic identification for biological assessments. J. North Am. Benthol. Soc. 27, 906–919. http://dx.doi.org/10.1899/07-175.1.
Sweeney, B.W., Battle, J.M., Jackson, J.K., Dapkey, T., 2011. Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? J. North Am. Benthol. Soc. 30, 195–216. http://dx.doi.org/10.1899/10-016.1.
Tautz, D., Arctander, P., Minelli, A., Thomas, R.H., Vogler, A.P., 2003. A plea for DNA taxonomy. Trends Ecol. Evol. 18, 70–74. http://dx.doi.org/10.1016/S0169-5347(02)00041-1.
Theissinger, K., Balint, M., Haase, P., Johannesen, J., Laube, I., Pauls, S.U., 2011. Species distribution models and molecular data reveal the Pleistocene history of the cold- adapted mayfly Ameletus inopinatus (Ephemeroptera: Siphlonuridae) in Europe. Freshw. Biol. 56, 2554–2566.
Thomas, A., Belfiore, C., 2004. Fauna Europea: Ephemeroptera. Fauna Europea version 1.1. available online at: www.faunaeur.org.
Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, J.J., Collingham, Y.C., et al., 2004. Extinction risk from climate change. Nature 427, 145–148. http://dx.doi.org/10.1038/nature01096.
Trizzino, M., Bisi, F., Maiorano, L., Martinoli, A., Petitta, M., Preatoni, D.G., Audisio, P., 2014. Mapping biodiversity hotspots and conservation priorities for the Euro-mediterranean headwater ecosystems, as inferred from diversity and distribution of a water beetle lineage. Biodivers. Conserv. 24, 149–170. http://dx.doi.org/10.1007/s10531-014-0798-z.

Waltz, R.D., McCafferty, W.P., 1997. New generic synonymies in baetidae (ephemeroptera). Entomol. News 108 (2), 134–140.
Webb, J.M., Jacobus, L.M., Funk, D.H., Zhou, X., Kondratieff, B., Geraci, C.J., DeWalt, R.E., Baird, D.J., Richard, B., Phillips, I., Hebert, P.D.N., 2012. A DNA barcode library for North American ephemeroptera: progress and prospects. PLoS ONE 7, e38063. http://dx.doi.org/10.1371/journal.pone.0038063.
White, B.P., Pilgrim, E.M., Boykin, L.M., Stein, E.D., Mazor, R.D., 2014. Comparison of four species-delimitation methods applied to a DNA barcode data set of insect larvae for use in routine bioassessment. Freshw. Sci. 33, 323–335. http://dx.doi.org/10.1086/674982.
Wiemers, M., Fiedler, K., 2007. Does the barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). Front. Zool. 4, 8.
Williams, H.C., Ormerod, S.J., Bruford, M.W., 2006. Molecular systematics and phylogeography of the cryptic species complex *Baetis rhodani* (Ephemeroptera, Baetidae). Mol. Phylogenet. Evol. 40, 370–382. http://dx.doi.org/10.1016/j.ympev.2006.03.004.
Wrona, F.J., Prowse, T.D., Reist, J.D., Hobbie, J.E., Levesque, L.M.J., Vincent, W.F., 2006. Climate impacts on arctic freshwater ecosystems and fisheries: background, rationale and approach of the arctic climate impact assessment (ACIA). Ambio 35, 326–329. http://dx.doi.org/10.1579/0044-7447.
Yu, D.S.K., Van Achterberg, C., Horstmann, K., 2012. Taxapad 2012, Ichneumonoidea 2011. Database on flash-drive (Ottawa, Ontario, Canada). www.taxapad.com.
Zhang, A.B., Muster, C., Liang, H.B., 2012. A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. Mol. Ecol. 21, 1848–1863.
Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869–2876.
Zhou, X., Robinson, J.L., Geraci, C.J., Parker, C.R., Flint, O.S., Etnier, D.A., Ruiter, D., Dewalt, R.E., Jacobus, L.M., Hebert, P.D.N., 2011. Accelerated construction of a regional DNA-barcode reference library: caddisflies (Trichoptera) in the great smoky mountains national park. J. North Am. Benthol. Soc. 30, 131–162. http://dx.doi.org/10.1899/10-010.1.
Zhou, X., Jacobus, L.M., Dewalt, R.E., Adamowicz, S.J., Hebert, P.D.N., 2010. Ephemeroptera, plecoptera, and trichoptera fauna of churchill (manitoba, Canada): insights into biodiversity patterns from DNA barcoding. J. North Am. Benthol. Soc. 29, 814–837. http://dx.doi.org/10.1899/09-121.1.
Zhou, X., Adamowicz, S.J., Jacobus, L.M., DeWalt, R.E., Hebert, P.D.N., 2009. Towards a comprehensive barcode library for arctic life- ephemeroptera, plecoptera, and trichoptera of churchill, manitoba, Canada. Front. Zoology 6, 30. http://dx.doi.org/10.1186/1742-9994-6-30.