



UNIVERSITÀ DEGLI STUDI DELLA TUSCIA DI VITERBO
DIPARTIMENTO DI SCIENZE ECOLOGICHE E BIOLOGICHE

CORSO DI DOTTORATO DI RICERCA IN
ECOLOGIA E GESTIONE DELLE RISORSE BIOLOGICHE - XXVIII CICLO

**Il DNA *Barcoding* come strumento preliminare e speditivo per
approfondimenti sistematici, valutazione della biodiversità e
biomonitoraggio**
(s.s.d. BIO/05)

Tesi di dottorato di:
Dott. Simone Cardoni

Coordinatore del corso
Prof. Daniele Canestrelli

Tutor
Prof. Carlo Belfiore (DEB)

Co-tutor
Dott. Marco Cosimo Simeone
(DAFNE)

Anno 2016

INDICE

1. INTRODUZIONE	- 1 -
1.1 DNA <i>barcoding</i>	- 1 -
1.1.1 Il Progetto Barcode of life	- 6 -
1.1.2 Principi applicativi del DNA <i>barcoding</i>	- 7 -
1.1.3 Valutazione dei dati nel metodo <i>barcoding</i>	- 10 -
1.2 Next Generation Sequencing	- 13 -
1.2.1 Applicazioni della tecnica NGS al <i>barcoding</i> ambientale	- 16 -
2. PRIMO CASO DI STUDIO: Biomonitoraggio delle acque dolci	- 18 -
2.1 Macroinvertebrati bentonici	- 20 -
2.2 Ephemeroptera	- 21 -
2.3 Applicazione della tecnica DNA <i>barcoding</i> agli Efemerotteri	- 23 -
3. SECONDO CASO DI STUDIO: La pesca marittima	- 25 -
3.1 La situazione degli stock di Alici e Sarde in Italia	- 26 -
3.2 Biologia delle specie studiate	- 26 -
3.2.1 <i>Engraulis encrasicolus</i> (Linnaeus, 1758)	- 27 -
3.2.2 <i>Sardina pilchardus</i> (Walbaum, 1792)	- 28 -
3.3 Il “bianchetto”	- 30 -
4. SCOPO DEL LAVORO	- 33 -
5. MATERIALI E METODI	- 35 -
5.1 Campionamento acque dolci	- 35 -
5.1.1 Prima fase di campionamento	- 35 -
5.1.2 Seconda fase di campionamento	- 38 -
5.2 Campionamento acque marine	- 40 -
5.3 Analisi del DNA	- 42 -
5.3.1 Estrazione del DNA	- 42 -
5.3.2 Amplificazione del frammento COX1 (Citocromo ossidasi subunità 1)	- 45 -
5.3.3 Amplificazione del frammento COX1 per l’analisi NGS	- 46 -

5.3.4 Preparazione dei campioni per il sequenziamento <i>NGS</i>	- 49 -
5.4 Analisi bioinformatiche	- 51 -
6. RISULTATI	- 55 -
6.1 DNA <i>barcoding</i> degli Efemerotteri	- 55 -
6.1.1 Dataset morfologico e molecolare degli Efemerotteri	- 55 -
6.1.2 Discriminazione delle specie	- 56 -
6.1.3 Identificazione tassonomica nel contesto Europeo	- 65 -
6.1.4 <i>Barcoding</i> ambientale delle acque dolci	- 74 -
6.2 DNA <i>barcoding</i> di <i>Engraulis encrasicolus</i> e <i>Sardina pilchardus</i>	- 85 -
6.2.1 Dataset molecolare	- 85 -
6.2.2 Correlazione con le banche dati di riferimento (NCBI e BOLD)	- 85 -
6.2.3 Analisi bioinformatiche	- 86 -
6.2.4 <i>Barcoding</i> ambientale delle acque marine	- 86 -
7. DISCUSSIONI	- 88 -
7.1 Potenzialità del DNA <i>barcoding</i> nella discriminazione delle specie italiane di Efemerotteri	- 88 -
7.2 Tassonomia basta su DNA: raggruppamenti <i>barcode</i> vs. specie morfologiche	- 91 -
7.3 Potenzialità del metabarcoding nelle acque dolci	- 94 -
7.3.1 library Efemerotteri	- 94 -
7.3.2 Library Bulk e campioni ambientali di sedimento ed acqua	- 95 -
7.4 Potenzialità del metabarcoding nelle acque marine	- 100 -
7.5 Implicazioni nel biomonitoraggio	- 102 -
8. CONCLUSIONI	- 105 -
BIBLIOGRAFIA	- 107 -
APPENDICI	- 131 -

1. INTRODUZIONE

1.1 DNA *barcoding*

La sistematica è la scienza che studia la diversità degli organismi. Da quando Carl Nilsson Linnaeus (1758) formalizzò la classificazione biologica con il suo sistema di nomenclatura binomiale cercando di raggruppare gli individui in base alle loro caratteristiche morfologiche condivise, l'identificazione degli organismi assunse un ruolo sempre maggiore nella descrizione della biodiversità, nello studio degli organismi viventi sul nostro pianeta, nelle relazioni intessute tra questi e l'ambiente fino al più recente utilizzo come strumento nel monitoraggio degli effetti biologici del cambiamento climatico globale (CDB, 1992). Tutt'ora, la classificazione tassonomica tradizionale linneana rimane un fondamento essenziale per le ricerche in ambito biologico, in particolar modo per tutela della biodiversità animale e vegetale a fronte delle veloci accelerazioni nella distruzione degli habitat naturali.

La conservazione della biodiversità è infatti uno degli argomenti più trattati e discussi degli ultimi anni nel settore scientifico, basti pensare che la nuova stima del tasso globale di estinzione annuo determinato dalle attività umane è di 1000 volte superiore al tasso di estinzione naturale (De Vos et al. 2015). Questo significa che annualmente vengono perse specie animali e vegetali, la maggior parte delle quali non ancora identificate. La tassonomia classica infatti non riesce a far fronte a questa continua corsa contro il tempo; le metodologie utilizzate risultano ancora lente, laboriose e dispendiose rispetto all'entità dei sistemi biologici e dei fenomeni coinvolti. I campioni raccolti, ad esempio, indipendentemente dal regno di appartenenza, di norma necessitano di trattamenti particolari in modo da mantenere intatte le loro caratteristiche distintive. La corretta identificazione richiede il giudizio di un esperto in quanto spesso complicata dalle sottili differenze morfologiche o anatomiche tra specie strettamente correlate (*cryptic species*) o da chiavi tassonomiche incomplete o spesso funzionali solo per un determinato stadio di vita o per un genere (ad es. i Ditteri, in cui la discriminazione morfologica è basata prevalentemente sui genitali maschili; Hennig 1976).

Chiaramente, le problematiche relative a questi “impedimenti tassonomici” costituiscono un ostacolo alla conoscenza della biodiversità ed alla sua gestione più appropriata.

Considerando inoltre che i programmi di biomonitoraggio e sulla conservazione sono alla base di politiche regionali, nazionali ed internazionali per contrastare la perdita di biodiversità e per la gestione e compilazione di liste rosse delle specie a rischio, il riconoscimento delle

unità tassonomiche di base acquista un peso rilevante. Di contro, la difficoltà nella conoscenza in dettaglio delle specie, in quanto entità biologiche dinamiche in continua evoluzione poiché soggette alle pressioni ambientali, ecologiche ed antropiche, ne rende ulteriormente problematico l'utilizzo come unità tassonomiche operative.

Risulta quindi evidente la necessità di effettuare valutazioni rapide per un esame preliminare in grado di fornire una visione d'insieme della situazione al momento dello studio, affiancando l'approfondimento tassonomico di dettaglio sia per ampliare la conoscenza scientifica che per migliorare la disponibilità delle informazioni utili per l'applicazione nei programmi di monitoraggio biologico.

In questo contesto il DNA *barcoding* si propone come strumento veloce, pratico ed economico in grado di far fronte alle esigenze sopra descritte, ponendo i presupposti per la differenziazione delle suddette unità tassonomiche su base molecolare in maniera indipendente dalle caratteristiche morfologico/anatomiche dei campioni e partendo da materiale biologico anche danneggiato, disponibile in piccole quantità o industrialmente trasformato.

Tale metodo è stato proposto come procedura standard per l'identificazione delle specie per la prima volta da Paul D.N. Hebert (2003), ricercatore presso la Università di Guelph in Canada. La tecnica è incentrata sull'impiego di brevi sequenze di DNA la cui variazione consente all'operatore di distinguere in modo univoco una specie dalle altre attraverso l'uso di specifici strumenti bioinformatici (Hebert et al. 2003). La caratteristica innovativa di queste sequenze, esistenti in ogni essere vivente e facilmente analizzabili, è quella di permettere la creazione di un "codice a barre del DNA" universale, finalizzato a catalogare e definire gli organismi allo stesso modo in cui il modello del codice a barre EAN (*European Article Number*) identifica ogni prodotto di consumo.

In realtà, il valore sistematico del DNA era già noto da tempo. Un primo approccio molecolare agli studi di classificazione fu effettuato da Carl Woese et al. (1990) che attraverso analisi condotte sulla parte codificante per l'RNA ribosomale 16s riuscirono a discriminare gli Archea come organismi procarioti differenti dai batteri. Successivamente numerose altre ricerche incentrate maggiormente sulla genetica di popolazione sono state prodotte utilizzando la variabilità di differenti marcatori molecolari come le varie regioni del DNA ribosomale (ITS, ETS, 18 e 25 S), gli allosimi, i microsattelliti, gli AFLP, ecc.. Sulla base di questa proliferazione di lavori sull'argomento, la tassonomia ha trovato nuovi impulsi aprendo nuove e promettenti frontiere nelle ricerche in questo settore (Tautz et al. 2003). Tuttavia, tali strumenti hanno dimostrato un limite applicativo principalmente sulla base dell'universalità dei metodi (cioè la possibilità di applicare le stesse procedure ad ogni essere vivente o almeno a ciascun regno per

poter meglio operare confronti a diversi livelli tassonomici), della laboriosità dei procedimenti, dei costi applicativi, delle caratteristiche necessarie al materiale di partenza delle analisi in termini di quantità e qualità dei tessuti biologici, e infine della diversa sensibilità degli strumenti nel rilevare differenze anche minime tra organismi simili.

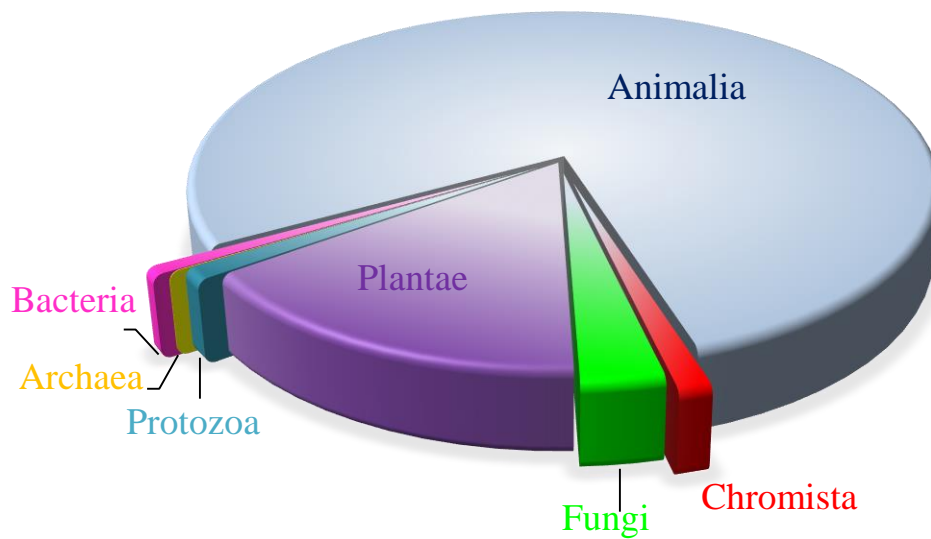
Il concetto di DNA *barcoding*, invece, permette di supportare le metodologie tassonomiche classiche per fronteggiare queste difficoltà, come ben riassunto dalla definizione ufficiale CBOL (*Consortium for the Barcoding of Life*, di cui parleremo nelle pagine successive), permettendo di:

1. *Lavorare con frammenti di tessuto.* Il *barcoding* consente di identificare una specie anche partendo da piccoli pezzi o frammenti di individui. Pertanto, esso fornisce la possibilità di discriminare individui qualora il riconoscimento con la tassonomia classica non riesca per l'insufficienza materiale di caratteri distintivi.
2. *Lavorare su individui a diversi stadi di sviluppo.* Il *barcoding*, infatti, basa la sua potenzialità applicativa su una molecola (il DNA) presente in ogni singola cellula; esso potrà quindi identificare una specie nei suoi molteplici stadi di sviluppo, sia in forme adulte che giovanili, da uova o seme di animali o piante fino ai tessuti morti.
3. *Smascherare i "simili".* La metodologia promette di poter evidenziare anche differenze "minime" tra individui (mutazioni singole), spesso non immediatamente rintracciabili a livello morfologico o ecologico; essa sarà quindi in grado di distinguere anche le specie fortemente somiglianti, consentendo una visione più precisa della biodiversità.
4. *Ridurre le ambiguità.* Il *barcode* è infatti espresso come una sequenza di quattro nucleotidi discreti - *ACTG* - appartenente ad una regione uniforme sul genoma che potrebbe fornire un'identificazione digitale per ogni specie. Questo, associato ad una banca dati digitale di codici a barre genetici di tutte le specie, offrirà uno strumento di riferimento inequivocabile per confronti a livello nazionale e internazionale, che potrebbe consentire una identificazione univoca.
5. *Potenziare le competenze scientifiche.* La biodiversità di circa 2 milioni di specie ufficialmente descritte sulla terra a confronto con la stima totale di oltre cento milioni di specie viventi, dimostra come le attuali potenzialità scientifiche abbiano un limite reale (Fig.1 a,b). Gli scienziati potrebbero utilizzare il *barcoding* per agevolare la discriminazione iniziale delle specie, per individuare le problematiche tassonomiche da approfondire successivamente.
6. *Ottenere una maggiore democratizzazione della conoscenza.* Una libreria pubblica di codici a barre di riferimento consentirà, anche a persone non esperte (cittadini,

amministratori della cosa pubblica), di poter identificare le specie con le quali vengono a contatto; la conoscenza, valorizzazione e la gestione della biodiversità locale e globale verrebbero in questo modo favorite, attraverso l'immediato riconoscimento di specie sia rare che abbondanti, sia autoctone che invasive.

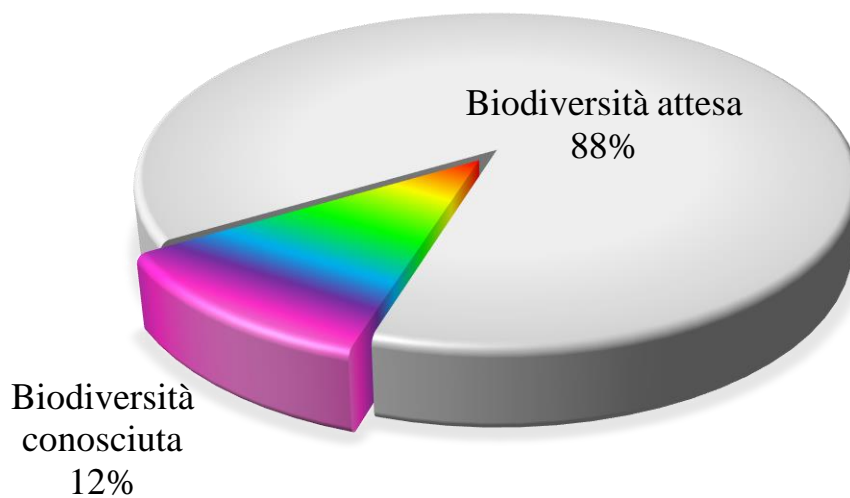
7. *Potenziare il valore delle collezioni.* La compilazione della banca dati genetica avrà un fondamentale campo di studio e approfondimento negli esemplari già presenti nelle collezioni di musei, erbari, zoo e giardini botanici. Questo comporterà una rivalutazione di queste importanti collezioni, finalizzata ad una migliore conoscenza della biodiversità passata e presente.
8. *Compilare velocemente l'enciclopedia della vita.* Come già detto, dall'epoca di Darwin i biologi sono sempre stati alla ricerca di un sistema di classificazione naturale, approntando anche un albero genealogico che rappresentasse l'evoluzione della vita. Il *barcoding* potrebbe indubbiamente partecipare a migliorare il metodo di classificazione basato su somiglianze e differenze, apportando un maggiore livello di dettaglio nella caratterizzazione delle specie, sia per quelle già descritte che per quelle future. L'associazione del campione conservato alla sequenza in banca dati faciliterà l'accesso collettivo alla conoscenza biologica, contribuendo a creare una libreria online della vita sulla terra con la digitalizzazione delle informazioni su animali e piante.

Biodiversità conosciuta



1a

Biodiversità stimata



1b

Fig.1 - (a) Biodiversità stimata, circa 10 milioni di specie, (b) Biodiversità conosciuta, approssimativamente 1.4 milioni di specie (escluse le comunità microbiche in entrambi i casi).

1.1.1 Il Progetto Barcode of life

Il progetto *Barcode of life*, patrocinato dal consorzio internazionale CBOL (*Consortium for the Barcode of life*), nasce con l'obiettivo di creare una banca dati di riferimento contenente le sequenze *Barcode* di tutti gli organismi viventi.

Il *Consortium for the Barcoding of life* è un'iniziativa di collaborazione internazionale alla quale prendono parte attualmente più di 130 organizzazioni di ricerca, tra le quali giardini botanici, musei, università e istituti nazionali di ricerca, operanti in 43 paesi, attraverso 6 continenti (www.barcoding.si.edu/CBOLMembers_details.htm) (Fig.2).

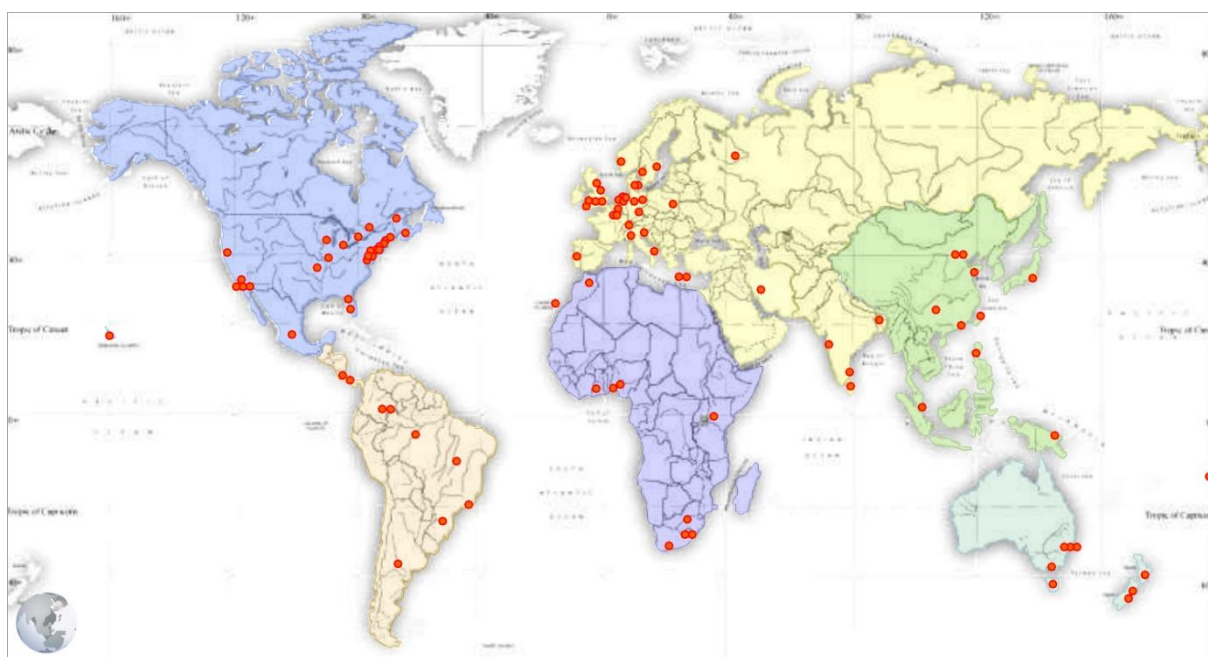


Fig.2 - Organizzazioni aderenti al “*Consortium for the barcoding of life*”

Scopo principale dell'organizzazione è quello di incentivare lo sviluppo e l'utilizzo del DNA *barcoding* come procedura standard per la discriminazione delle specie. Il procedimento suggerito per la riuscita del progetto prevede diverse condizioni, come la realizzazione di archivi di alta qualità dei vari *barcode* organizzati in banche dati di sequenze genomiche, oppure lo sviluppo di nuove metodologie che permetteranno prestazioni sempre più rapide, economiche, trasferibili e portatili, oltre che la complementarietà di partecipazione con tassonomi e gruppi di ricerca tassonomica.

In questo contesto si inserisce il progetto iBOL (*International Barcode of life*), la più grande iniziativa di ricerca sulla conservazione della biodiversità attualmente in atto. Quest'ultima ha come finalità estendere la copertura geografica e tassonomica delle banche dati di riferimento

per le sequenze *barcode* e promuovere la creazione di nuovi dispositivi in modo da garantire alla comunità la completa fruibilità a questo tipo di informazioni.

Attualmente sono impegnati più di 100 ricercatori dislocati in 25 paesi, che nella prima fase di lavoro (2010-2015) hanno come obiettivo l'analisi di cinque milioni di esemplari rappresentanti 500.000 specie riconosciute. Il risultato fino ad ora ottenuto ed in continuo aggiornamento è una banca dati pubblica generata appositamente dall'iniziativa al sito www.boldsystems.org nella quale sono presenti le sequenze, le informazioni tassonomiche dei campioni con le relative foto, le date di raccolta e conservazione, ed i luoghi di campionamento. Inoltre numerosi progetti legati all'acquisizione di sequenze *barcode* sono nati al fine di integrare la banca dati suddetta, alcuni di questi sono mostrati in Tab.1.

Tab.1 - Alcuni dei progetti *Barcoding* attualmente in atto con i relativi dati aggiornati.

<i>Group</i>	<i>N° of Target Species</i>	<i>Species barcoded</i>	<i>Specimens barcoded</i>	<i>Species Remaining</i>
<i>iBOL Formicidae</i>	12205	1359	15753	10846
<i>iBOL Trichoptera</i>	13165	2779	19243	10386
<i>FishBOL</i>	32257	11063	110888	21194
<i>iBOL Lepidoptera</i>	1166319	84327	982678	1081992
<i>iBOL Mammal</i>	5426	858	19862	4568
<i>iBOL Marine</i>	55451	6199	37182	49252
<i>iBOL Polar</i>	30871	3279	23953	27592

1.1.2 Principi applicativi del DNA *barcoding*

Nell'applicazione della metodologia del DNA *barcoding*, il frammento genomico da utilizzare deve avere le caratteristiche necessarie alla soddisfazione delle condizioni descritte, soprattutto tali da poter funzionare bene per tutte le specie.

Nel regno animale, il genoma mitocondriale possiede specifiche caratteristiche che lo favoriscono rispetto al genoma nucleare. Innanzitutto è provvisto di un elevato numero di copie per cellula (da 100 a 10000) rispetto alle 2 copie di quello nucleare, ha una modalità di eredità uniparentale (citoplasmatica), assenza di ricombinazione, e soprattutto un tasso evolutivo 10 volte superiore rispetto al genoma nucleare a causa di un sistema di riparazione assente o poco efficiente che gli conferisce un maggior potenziale evolutivo.

Il DNA mitocondriale degli animali, a parte rare eccezioni, contiene il medesimo set di geni: 13 proteine codificanti, 2 RNA ribosomali, 22 transfer RNA, tuttavia l'ordine e la polarità dei geni è abbastanza variabile. Per questo motivo la scelta del frammento da utilizzare come

potenziale *barcode* è ricaduta all'interno di un singolo gene. I geni mitocondriali sono stati valutati secondo determinati standard quali: la lunghezza, il tasso evolutivo, la presenza di inserzioni o delezioni (*indels*) e la possibilità di costruire, ai margini, i primer (inneschi) necessari all'amplificazione via PCR che siano robusti e universali. Il risultato degli screening iniziali ha portato alla selezione di un frammento, la subunità 1 della Citocromo Ossidasi (COX1), avente una lunghezza di 648 nucleotidi, normalmente affiancata all'estremità 5' dal gene tRNATyr (un RNA transfer che serve al trasferimento dei residui di tirosina) e all'estremità 3' dal gene tRNA Leu (un RNA transfer che serve al trasferimento della laucina) (Fig.3). La potenzialità del frammento è stata testata (Hebert et al. 2003a, 2003b) utilizzando le sequenze disponibili nelle banche dati (NCBI) e il risultato dei saggi ha dimostrato l'effettiva applicabilità del metodo, rilevando un'elevata variabilità interspecifica (maggiore del 2% di distanza genetica nel 98% dei casi studio).

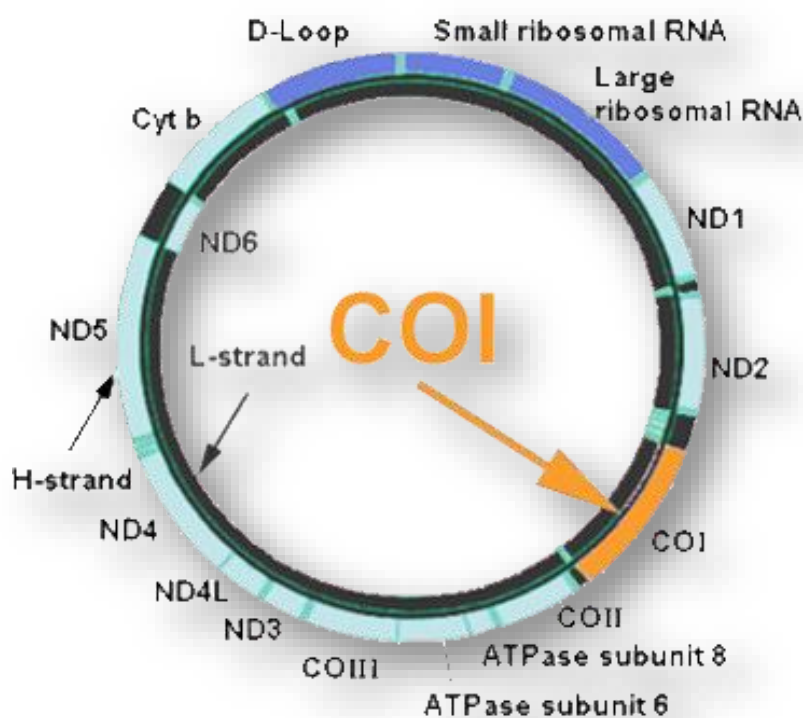


Fig.3 - Genoma mitocondriale e posizione del gene COX1.

L'opportunità aperta in questo senso ha suscitato reazioni favorevoli (Janzen 2004, Pfenninger et al. 2007; Packer et al. 2009; Porter et al 2014) e numerose critiche (Seberg et al. 2003; Lipscomb et al. 2003; Moritz & Cicero 2004; Lee 2004), evidenziando alcuni limiti all'applicazione del metodo. Tra questi, ad esempio: a) non per tutti i gruppi di organismi è

valida la soglia di divergenza pari al 2%; esistono infatti gruppi tassonomici con bassa divergenza interspecifica tale da risultare inferiore a tale soglia (ad esempio, come sottolineato da France & Hoover 2002 e Shearer et al. 2002, nella classe Anthozoa del phylum Cnidaria); b) nel caso di specie separate da poco tempo (inferiore al milione di anni), le mutazioni accumulate potrebbero non essere sufficienti alla loro distinzione su base molecolare. Assumendo infatti un tasso di mutazione pari al 2% per milione di anni, un frammento di 600 paia di basi mostrerebbe solamente 12 siti diagnostici, probabilmente inefficace per una robusta discriminazione; c) la corretta assegnazione di un individuo ad una specie potrebbe essere sfavorita nel caso di individui ibridi; molte specie, inoltre, derivano da ibridazione tra specie diverse o poliploidia, e potrebbero pertanto non evidenziare variabilità nei DNA citoplasmatici. Infine, bisogna anche rilevare l'eventuale coamplificazione di NumtS (presenza di frammenti mitocondriali traslocati nel nucleo) e la possibilità di introgressione di geni mitocondriali nel nucleo che possono certamente causare una sovrastima delle specie reali (Song et al. 2008; Gaziev & Shaikhaev 2010; Moulton et al. 2010; Jordal et al. 2014; Ermakov et al. 2015).

Per una corretta applicazione della metodologia, occorre pertanto che vengano utilizzati diversi accorgimenti (bio-ecologici, statistici, genetici), in modo da poter utilizzare tale tecnica come un valido strumento complementare alla tassonomia classica. Infatti, le premesse di base per un efficace utilizzo di questa tecnica (incentrate sulla molecola che è alla base dell'evoluzione degli organismi viventi), quali la sicurezza, l'oggettività e la confrontabilità del dato, rendono le sequenze di DNA tali da poter essere considerate il possibile sistema di riferimento per la tassonomia moderna (Tautz et al. 2003).

1.1.3 Valutazione dei dati nel metodo *barcoding*

L'analisi dei dati ottenuti col metodo *barcoding* ha lo scopo di verificare la corrispondenza tra la sequenza ottenuta con i marcatori standard previsti dal protocollo ed un elemento tassonomico definito dalla tassonomia classica e nomenclatura linneana.

Uno dei primi criteri di analisi utilizzati è quello basato sui “caratteri” (*character-based approach*); questo prevede il raggruppamento e l'identificazione delle sequenze generate in base alla presenza di caratteri (o combinazione di essi) condivisi, nello specifico nucleotidi. Presupponendo quindi che membri dello stesso gruppo tassonomico abbiano in comune determinati attributi caratteristici (CAs) assenti in altri gruppi (come ad es. un certo polimorfismo) (Rach et al., 2008).

L'approccio filogenetico (*tree-based approach*) è un'altra metodologia ampiamente utilizzata nella quale il risultato è la costruzione di un albero. Tuttavia questo tipo di rappresentazione non attenua le critiche mosse nei confronti della tecnica, accentuando invece la confusione tra *DNA barcoding* e *DNA taxonomy*. Infatti, nonostante la chiara differenza tra i due approcci, il loro reale significato può perdere di concretezza a causa dell'uso improprio o ambiguo che ne viene fatto. Il *DNA barcoding*, infatti, nasce come strumento “semplicemente” identificativo, con lo scopo principale di discriminare “tutte” le specie partendo da dati basati sulla variabilità molecolare di un unico frammento di DNA che può essere analizzato rapidamente in ogni essere vivente. A questa tecnica, in un secondo tempo può venire associata la tassonomia classica, per possibili conferme/smentite o ulteriori approfondimenti. Differente è il caso della *DNA taxonomy* in cui la principale prerogativa è quella di circoscrivere e delineare gruppi di sequenze che possano concordare con la struttura gerarchica propria della tassonomia classica. In aggiunta le sequenze generate possono provenire da marcatori appartenenti a differenti regioni genomiche non legate a determinati loci standard come nel caso del *DNA barcoding*. Pertanto nell'approccio filogenetico vengono prodotti dendrogrammi attraverso l'impiego di differenti criteri come Neighbour Joining, Maximum Likelihood e metodi Bayesiani, mentre l'identificazione dei campioni viene ottenuta attraverso l'inclusione delle sequenze prodotte in cladi comuni. La robustezza del risultato viene quindi valutata in base ai valori di *bootstrap* calcolati per ciascun clado; questi, di norma, devono risultare maggiori del 70%. Il posizionamento dei membri di una specie relativamente a quelli di un'altra possono inoltre fornire un'indicazione della monofilia della specie stessa; secondo questo metodo, solo le specie con monofilia supportata da oltre il 70% di valori *bootstrap* possono ritenersi correttamente identificate (Fazekas et al 2008).

Un'altra metodologia ampiamente utilizzata presuppone l'esistenza di un database completo di sequenze di riferimento relativo a tutte le specie, all'interno del quale effettuare una ricerca della corrispondenza con la nostra sequenza incognita. Questo tipo di ricerca viene normalmente eseguito su GeneBank (<http://www.ncbi.nlm.nih.gov>), una banca dati pubblica contenente sequenze nucleotidiche per oltre 322,000 specie formalmente descritte (aggiornata a Dicembre 2014). Su GenBank la ricerca viene eseguita in automatico attraverso determinati algoritmi di similarità (es. Blastn o MegaBlast), tuttavia il valore che quest'ultimi associano alle relative sequenze (BIT score) può risultare condizionato dalle dimensioni del database preso a riferimento (cioè il numero di sequenze disponibili per il frammento in esame). Con lo scopo di fornire maggior dettaglio per i risultati ottenibili con il confronto di sequenze, negli ultimi anni è stata sviluppata la banca dati BOLD (Barcode of Life Data Systems; <http://www.barcodinglife.org/>), unicamente dedicata alla raccolta delle sequenze *barcode* e che si prefigge di rappresentare lo strumento di verifica fondamentale per le ricerche condotte attraverso questo approccio.

BOLD, inizialmente costituito da sole sequenze COX1, successivamente è stato ampliato con sequenze relative all'identificazione dei funghi (marcatore ITS) e delle piante (marcatori *rbcL* e *matK*). Gli algoritmi di ricerca utilizzati da BOLD sono tutt'ora in continua sperimentazione e vengono aggiornati conseguentemente.

Un ulteriore sviluppo di questa tecnica basata sul confronto tra sequenze prevede la creazione di un database di riferimento sul gruppo tassonomico (o sulla regione geografica) oggetto di studio. In questo caso, è necessario l'affiancamento dello studio tassonomico classico a quello molecolare, al fine di predisporre una "propria" banca dati di riferimento sulla quale condurre i confronti di similarità. In questo ambito, anche l'utilizzo di algoritmi di similarità più adatti (ad esempio, BLAST, BLAT, megaBLAST) può essere determinante in quanto, attraverso il settaggio di alcuni parametri relativi al grado di similarità cercato, si può individuare una soglia che consenta non solo di discriminare le specie ma anche di formare complessi di individui appartenenti alla stessa specie ma di aree geografiche diverse.

Infine, un ulteriore criterio utilizzato per raggiungere l'identificazione attraverso il DNA *barcoding* è quello basato sull'esame delle distanze genetiche esibite dai campioni analizzati (attraverso il calcolo delle misure *p-distance* o *Kimura-2 parametri*), con le quali si possono evidenziare le caratteristiche di variabilità del *locus barcode* utilizzato (ovvero la variabilità nucleotidica tra ed entro le specie). Nello specifico, l'identificazione sicura di un campione può essere ottenuta quando il minor valore di distanza genetica riscontrata tra le specie risulta maggiore della massima distanza riscontrata entro ciascuna specie, generando così una

differenza tra le due distanze, definita “*barcoding-gap*” (Fig.4a). Al contrario, una sovrapposizione “*overlap*” delle due distanze genetiche non permette una chiara discriminazione delle specie in esame (Fig.4b).

E’ possibile quindi affermare che non esiste un “metodo migliore” che permetta di ottimizzare le potenzialità di identificazione del *locus* analizzato ma, piuttosto, è possibile utilizzare differenti procedure in funzione delle caratteristiche del set di dati da analizzare.

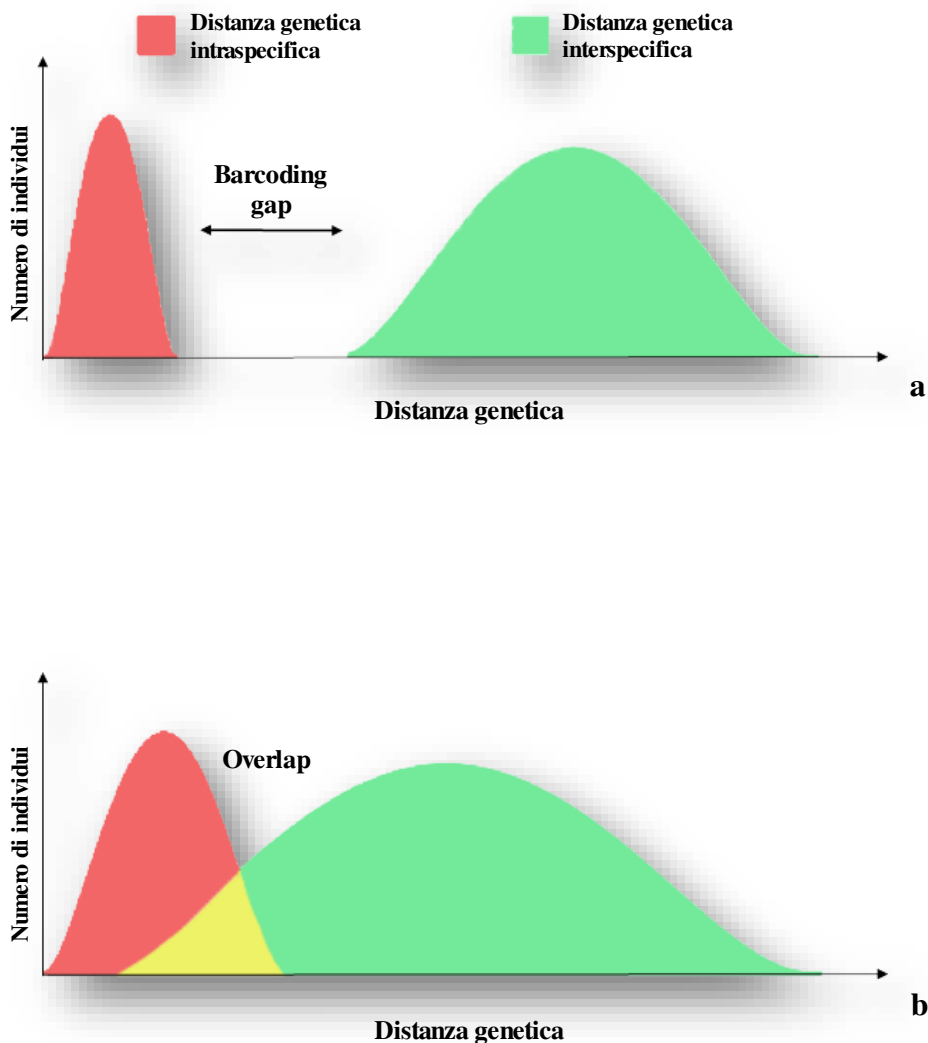


Fig.4 a,b - (a) *Barcoding gap* generato dalla differenza tra le distanze genetiche intraspecifiche ed interspecifiche, (b) *Overlap* generato dalla mancata differenza tra le stesse due distanze genetiche.

1.2 Next Generation Sequencing

Negli ultimi decenni i metodi di sequenziamento del DNA si sono basati sul metodo messo a punto nel 1977 da Frederick Sanger (o metodo a terminatori di catena), (Sanger et al. 1977). Questo metodo ha aperto le possibilità di ottenere le sequenze nucleotidiche di singoli frammenti del DNA fino al sequenziamento di interi genomi dando un immenso impulso alla ricerca scientifica moderna in diversi settori di applicazione (medico, bio-ecologico, evolutivo, etc.). Tuttavia, la condizione necessaria di funzionamento del metodo Sanger si basa sulla possibilità di analizzare un solo frammento di DNA per volta. Nei primi anni del 2000 sono state sviluppate delle tecnologie di sequenziamento del DNA di nuova generazione (*Next Generation Sequencing*, *NGS* o *Second Generation Sequencing*), chiamate anche “*high-throughput sequencing*” (sequenziamento ad alta resa), le quali stanno gradualmente soppiantando il metodo Sanger grazie alla loro capacità di analizzare contemporaneamente frammenti multipli di DNA. Tale innovazione ha permesso di ottenere elevate prestazioni ad alta velocità, riducendo drasticamente i tempi di attesa ed i costi del sequenziamento, soprattutto nel caso di “matrici complesse” di DNA. Queste ultime, infatti, precedentemente non potevano essere analizzate con il metodo Sanger, se non previa esecuzione di processi lunghi, costosi e difficoltosi come il clonaggio

In generale, i metodi *NGS* permettono di analizzare frammenti di DNA con minor lunghezza rispetto al metodo Sanger (<200-250 bp contro ca. 1000 bp), tuttavia, la capacità di effettuare più letture in parallelo (nell'ordine di migliaia con una singola analisi), permette di accelerare notevolmente i tempi di analisi e quindi di ottenere una quantità di dati enormemente superiore. Per questo motivo, l'output generato da una corsa *NGS* richiede anche l'utilizzo di sistemi bioinformatici molto avanzati sia per l'elaborazione dei dati che per la comprensione dell'informazione ottenuta, in modo da avere una corretta ed immediata interpretazione biologica dell'esperimento condotto.

Attualmente sono disponibili diversi metodi (o piattaforme) *NGS*, sviluppate da differenti ditte commerciali e tutte caratterizzate da diverse efficienze lavorative relativamente alla lunghezza dei frammenti analizzabili, al numero di analisi parallele e alla riduzione dei margini di errore. La prima tecnica sviluppata fu il pirosequenziamento e veniva effettuato con la piattaforma 454 delle Roche. Attraverso questo sistema si riuscivano ad ottenere frammenti di circa 500 bp (Fig.5). Tuttavia, questa tecnica si caratterizzava da costi e margini d'errore molto elevati. Successivamente questo sistema è stato sostituito dalla piattaforma Solexa dell'Illumina con la quale è possibile ottenere un numero superiore di frammenti più corti (circa 200 bp) detti

“reads” ad un costo inferiore e con una maggiore precisione di sequenziamento (Fig.6). Tutte le piattaforme comunque condividono i processi principali del metodo: la preparazione e immobilizzazione del DNA (cioè la preparazione della cosiddetta *sequencing library*), la reazione di amplificazione e la reazione di sequenziamento.

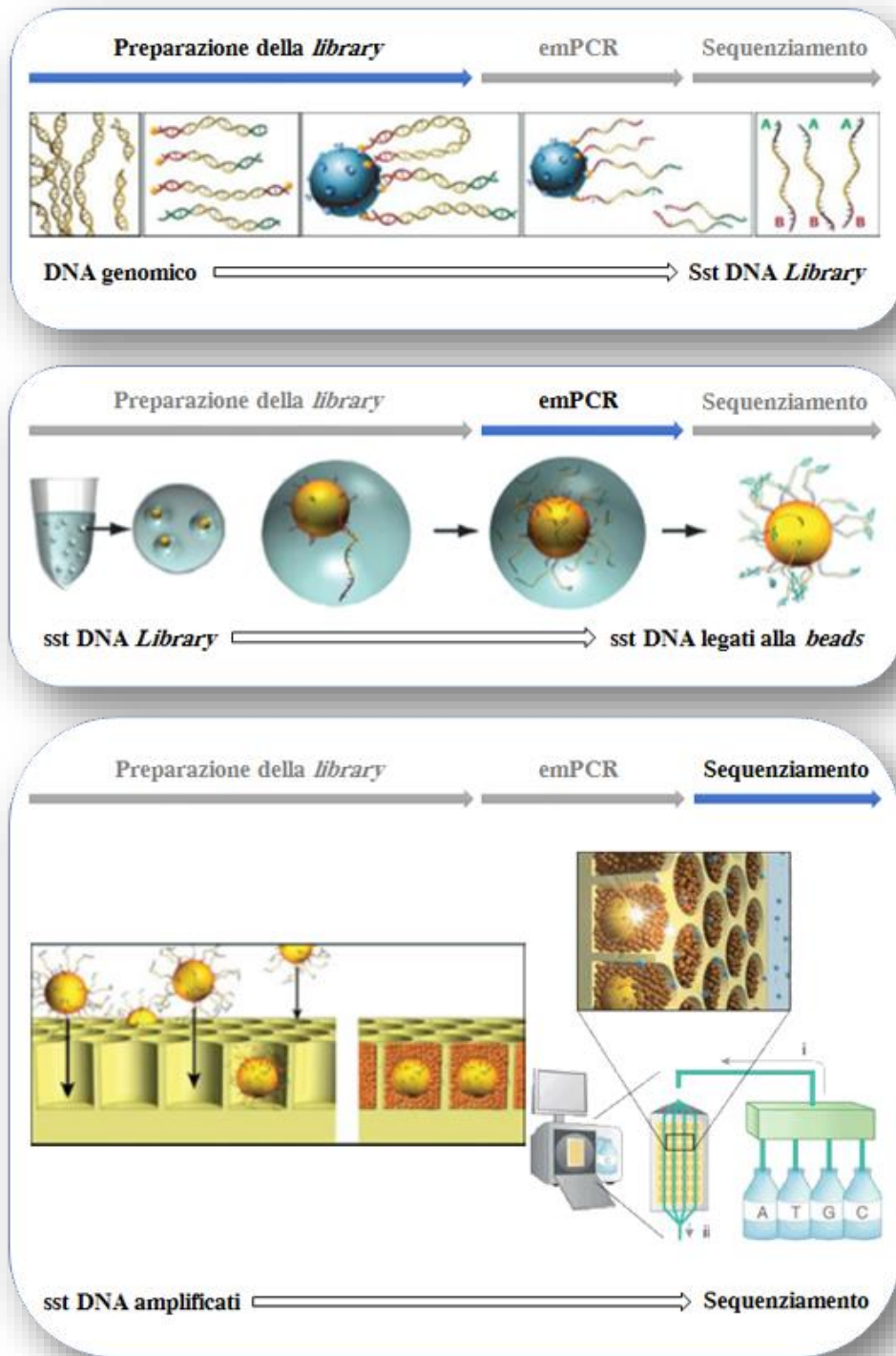


Fig.5 - Fase PCR metodo La Roche 454

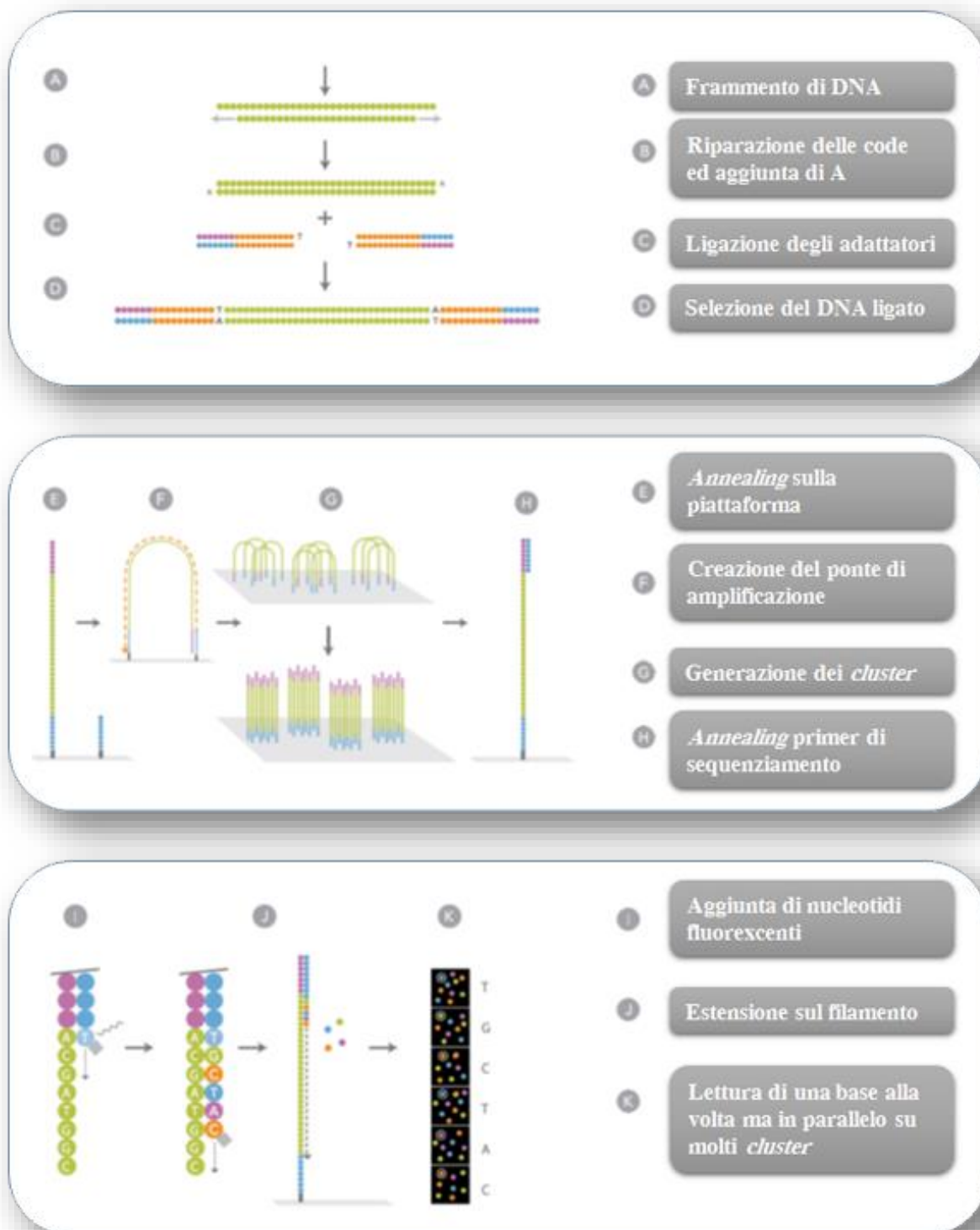


Fig.6 - Fase di PCR metodo Illumina

1.2.1 Applicazioni della tecnica *NGS* al *barcoding* ambientale

Attualmente le tecnologie *NGS* svolgono un ruolo fondamentale in differenti ambiti della ricerca, dall'epigenetica (Taylor et al. 2007) al sequenziamento de novo e ri-sequenziamento di genomi completi (Burgess 2011; Nyaga et al. 2016), al sequenziamento di RNA (Ramsingh et al. 2010; Khalifa et al. 2015), fino ad arrivare al largo impiego che hanno avuto nelle ricerche sul genoma umano. Secondo alcune recenti stime, infatti, più di 100000 genomi di individui sono stati sequenziati, consentendo uno studio approfondito sulle diversità umane e sulle malattie che ci colpiscono (Roach et al. 2010; Tennessen et al. 2012; Yang et al. 2013; Fromer et al. 2014; Lawrence et al. 2014).

Tra i numerosi campi applicativi offerti da questa metodologia vi è anche la valutazione della biodiversità di un ecosistema. Alcuni esempi in questo senso sono gli studi effettuati sulla comunità microbica e fungina, dai funghi agli Archea, di vari ambienti di suolo agricolo e forestale (Leininger et al. 2006; Fierer et al. 2007; Acosta-Martinez et al. 2008; Buée et al. 2009; Jumpponen et al. 2010; Rousk et al. 2010; Colombo et al. 2016).

Altre applicazioni sono state effettuate anche nel settore alimentare. Qui, l'applicazione di tali tecnologie ha permesso, ad esempio, la verifica dei prodotti commercializzati e il rilevamento di patogeni presenti (Leonard et al. 2015). Queste nuove tecnologie hanno anche permesso l'identificazione di patogeni associati alla distruzione di colonie di api mellifere (Cox-Foster et al. 2007), di virus responsabili delle infezioni sulle piante di pomodori (Adams et al. 2009) o semplicemente il riconoscimento di piante impollinate e degli insetti impollinatori a partire da 1 ml di miele (Schnell et al. 2010).

Molti altri studi hanno avuto come oggetto l'analisi delle diete, ad esempio quelle di alcuni erbivori effettuate sulle componenti intestinali utilizzando il marcatore plastidiale *trnL* (Pegard et al. 2009; Soininen et al. 2009; Valentini et al. 2009; Kowalczyk et al. 2011). Oppure studi condotti sul microbioma intestinale dei gatti attraverso l'analisi di sequenze di 16 rDNA (Turnbaugh et al. 2008, 2009; Murphy et al. 2010; Ravussin et al. 2011; Serino et al. 2011).

Anche l'ambiente marino è stato oggetto di ricerche condotte con l'*NGS* sulla comunità batterica attraverso il gene ribosomale 18S (Huber et al. 2007), il gene 16S (Sogin et al. 2006) l'espressione genica attraverso l'analisi del trascrittoma del cDNA (Frias-Lopez et al. 2008) oppure, sempre attraverso l'analisi del 18S rDNA, il microbiota eucariotico.

In ambito ecologico, le metodologie *NGS* si stanno diffondendo soprattutto per eseguire analisi su "miscele di DNA" sconosciuti, in combinazione alla metodologia del DNA *barcoding*. Tale tecnica permette il riconoscimento degli organismi presenti all'interno di un campione ambientale (*Environmental barcoding*), il quale può esser caratterizzato da materiale

organico e inorganico proveniente da suolo, aria, acqua o composti organici (tessuti, alimenti, feci, etc.) al cui interno si possono trovare organismi animali, vegetali o microbici, o anche tracce organiche di questi. L'impiego simultaneo di queste nuove tecnologie (capaci di amplificare ogni singolo filamento di DNA presente in un campione) e di marcatori universali del tipo *barcode* (aventi come bersaglio una breve sequenza caratteristica), hanno permesso l'identificazione chiara e precisa di tutti i costituenti organici del campione in esame.

In questo senso molti lavori incentrati sullo studio della diversità biologica attraverso l'utilizzo di queste tecniche hanno permesso, ad esempio, l'analisi di una grande varietà di microhabitat terrestri nelle foreste pluviali del Costa Rica (suolo, lettiera ed epifite: Creer et al. 2010; Porazinska et al. 2010), oppure, sempre in campo forestale, sia la discriminazione di antiche specie forestali attraverso l'analisi di pollini presenti nel permafrost (Sønstebø et al. 2010) sia l'identificazione di specie forestali, sempre da pollini, ma recuperati dagli insetti impollinatori (Sickel et al. 2015).

Allo stesso modo recenti studi sulle acque dolci con i metodi *NGS* e *DNA barcoding* hanno sottolineato l'applicabilità di queste tecniche al settore del biomonitoraggio: Ficetola et al. (2008) hanno analizzato il citocromo b combinando *NGS* e metodo Sanger per riscontrare la presenza della rana toro (*Bullfrog*) in campioni ambientali provenienti dalle acque correnti. Ulteriori esempi applicativi nel settore delle acque correnti hanno riguardato la diversità di specie rare di pesci, anfibi, mammiferi, crostacei ed insetti utilizzando campioni ambientali di acqua dolce prelevati in laghi stagni e ruscelli (Thomsen et al. 2011) e la diversità delle comunità di protisti provenienti da campioni ambientali di acque correnti (Medinger et al. 2010). Recentemente, Hajibabaei et al. (2011) hanno utilizzato porzioni di COX1 (la subunità 1 del gene codificante per la citocromo ossidasi) per fornire l'identificazione a livello di specie dei macroinvertebrati presenti in campioni bentonici.

Tuttavia, quando si effettuano studi di biodiversità su campioni ambientali, dovrebbero essere considerati alcuni potenziali errori, ad esempio la possibilità di amplificazione di artefatti o eventuali identificazioni tassonomiche errate (Degnan & Ochman 2011).

2. PRIMO CASO DI STUDIO: Biomonitoraggio delle acque dolci

Le acque dolci sono *hotspot* di biodiversità particolarmente a rischio di forti alterazioni. Estinzioni, riduzione dell'habitat e diminuzione delle popolazioni presenti sono principalmente causati da una serie di trasformazioni ambientali intrinsecamente legate alla crescita della popolazione umana, all'economia globale (Strayer 2006; IUCN 2007; Polunin 2008) ed alla forte sensibilità e natura degli habitat di molte specie (Dudgeon et al. 2006). Per millenni infatti gli esseri umani hanno sfruttato le acque dolci e l'ambiente circostante attraverso l'irrigazione, smaltimento rifiuti, trasporto, pesca, produzione energia, attività agricole e industriali.

Allo stesso tempo le frammentazioni degli habitat dettate dalle dighe, dal rinforzo degli argini, dalle rettificazioni degli alvei per la navigazione ed il controllo delle inondazioni insieme alla diffusione nelle trasformazioni delle zone ripariali, hanno modificato gli equilibri negli apporti di nutrienti, sostanza organica e sedimenti. Di conseguenza, un eccessivo carico di inquinanti e nutrienti derivati da questi cambi di uso dei suoli hanno eutrofizzato molte acque dolci causando l'insostenibilità delle loro biocenosi naturali (Polunin 2008; Smol 2008). Ne consegue una maggior vulnerabilità di questi ecosistemi per la riduzione di habitat idonei per specie endemiche e biocenosi strutturate rispetto a quelli marini e terrestri (Strayer & Dudgeon 2010).

Occorre quindi migliorare le conoscenze sullo stato delle specie con le loro relative minacce e dei fattori abiotici locali e globali che entrano in gioco, esplorando i legami tra biodiversità d'acqua dolce e funzione dell'ecosistema in modo da prevenire ulteriori perdite e pianificare adeguate pratiche di gestione.

In questo contesto si inseriscono i quadri legislativi previsti dalla Comunità Europea come la WFD 2000/60/CE (*Water Framework Directive* - European Commission, 2000), che istituisce un quadro di azione comunitaria in materia di acque.

Lo stato di salute delle acque interne, infatti, è un argomento di evidente rilevanza socio-economica, che interessa tante figure pubbliche e private: gli organi di gestione del territorio che devono garantirne lo stato di qualità, rispondendo appropriatamente alle normative nazionali e alle direttive EU, il settore primario (agricoltori, allevatori, pescatori) che fruisce di corsi d'acqua naturali o artificiali per le sue attività e che allo stesso tempo ne influenza la qualità e quantità, i cittadini che ne dipendono per usi e consumi, il turismo di aree protette o naturali, che ne può risultare incentivato o sfavorito, e più in generale lo stato di salute della biodiversità naturale stessa, intesa come insieme delle componenti faunistiche e vegetali che caratterizzano e denotano paesaggio, storia, tradizioni e cultura locali. Tutto questo ha enormi

ripercussioni nel settore economico, sociale e della bio-sicurezza delle comunità agricole e cittadine.

Lo stato ecologico delle acque interne viene attualmente definito secondo la WFD (*Water Framework Directive*, 2000/60/CE) come “l’espressione della qualità della struttura e del funzionamento degli ecosistemi acquatici associati alle acque superficiali”. Tale concetto assume rilevante importanza in quanto, un ambiente idrologico ben conservato è in grado di fornire molteplici servizi ecologici ed economici che verrebbero meno in caso di alterazione ambientale. Dopo l’indice IBE (Indice Biotico Estesio; Ghetti 1995, 1997), introdotto con la 152/99, e la successiva emanazione della Direttiva 2000/60/CE (recepita in Italia con la Legge Quadro sulle Acque 152/06), diversi nuovi indici basati sulla presenza ed abbondanza di specifici organismi animali definiti “bio-indicatori” sono oggi previsti dalla legislazione per ottenere valori indicativi di una classe di qualità delle acque analizzate (ad es. l’indice di intercalibrazione STAR_ICMi (*STAR Intercalibration Common Metric Index*, Buffagni A. & Erba S. 2007) e l’indice accessorio MAS (*Mayflies Average Score*) per i macroinvertebrati dei fiumi, l’indice multimetrico di intercalibrazione ICMi (*Intercalibration Common Metrix Index*, ISS 2009) per le diatomee, l’indice IBMR per le macrofite, ISECI per la fauna ittica, ICF per il fitoplancton dei corpi idrici lacustri, LFI per la fauna ittica dei laghi, M-AMBI per i macroinvertebrati bentonici delle acque marine costiere, il CARLIT per le macroalghe, PREI per le praterie a *Poseidonia* oceanica). Tra questi, quelli più comunemente utilizzati per le acque correnti prendono in considerazione i macroinvertebrati bentonici (STAR_ICMi e MAS). Elemento essenziale nell’applicazione di questi indici è il riconoscimento tassonomico di insetti (principalmente Efemerotteri, Ditteri, Plecotteri e Tricotteri) ed altri organismi acquatici (Anellidi, Molluschi, etc.) che, in virtù della loro sensibilità alle condizioni ecologiche del corpo idrico, popolano in maniera più o meno copiosa i corsi d’acqua.

2.1 Macroinvertebrati bentonici

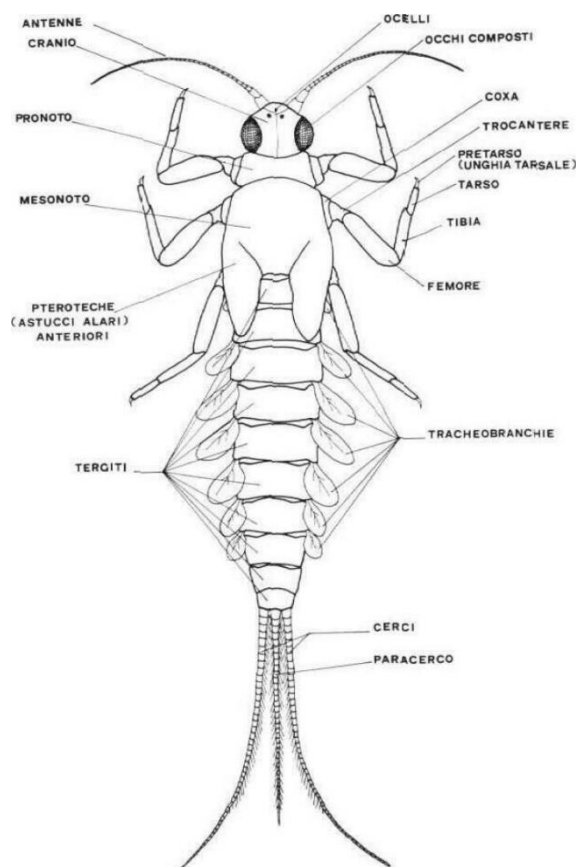
I macroinvertebrati bentonici sono organismi acquatici utilizzati per la valutazione della qualità delle acque. Il loro impiego in questo settore risale all'inizio del secolo scorso (Kolkwitz & Marsson 1909), ma soltanto negli ultimi quarant'anni hanno acquisito una notevole rilevanza all'interno dei programmi di monitoraggio. Si è visto pertanto che il biomonitoraggio, a differenza dell'utilizzo esclusivo delle analisi chimico-fisiche e microbiologiche (effettuate fino agli anni '70), fornisce un contesto dinamico della situazione presente e passata nella stazione analizzata, dando una misura del grado di salute di un ecosistema in termini sia spaziali che temporali.

Numerose sono infatti le caratteristiche che rendono i macroinvertebrati bentonici ottimi bioindicatori, innanzitutto la loro presenza in un'ampia varietà di ambienti, con adattamenti specifici a condizioni diverse e con elevato numero di individui e biomassa. La consistente quantità di specie che compone le differenti comunità e che occupano specifiche nicchie ecologiche, dà un'ampia visione delle risposte agli stress ambientali. La peculiare stanzialità di questi organismi è un'altra caratteristica importante che permette di correlare un determinato sito ad una specifica comunità, consentendo di effettuare valutazioni spaziali delle perturbazioni. I caratteristici cicli vitali, che possono durare più di un anno per alcune specie, aiutano anche nelle analisi delle alterazioni ambientali sia continue che intermittenti causate sia da agenti singoli che dalla complementarietà di alcuni di questi.

La totalità di queste informazioni fornisce la possibilità di poter effettuare indagini sull'intera comunità, su come questa venga alterata, su come i taxa che la compongono si succedono e si sostituiscono fornendo uno scenario riassuntivo sul grado di alterazione ambientale e sulla risposta dei taxa ai gradi di inquinamento.

2.2 Ephemeroptera

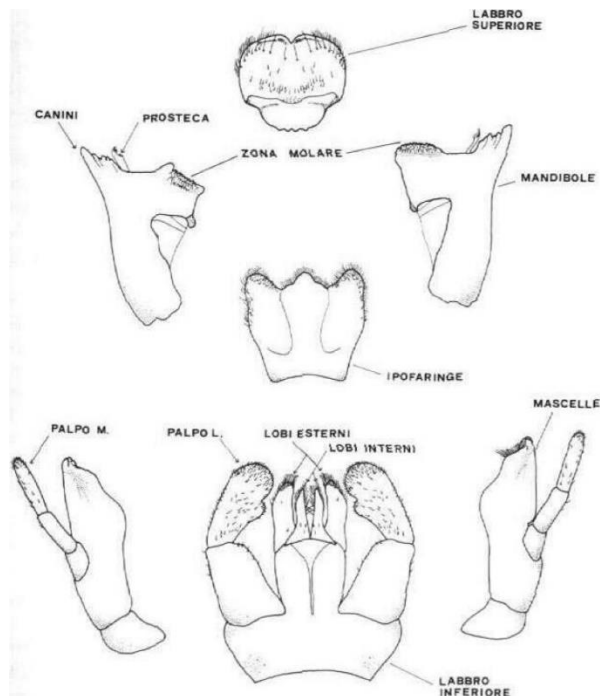
Gli Efemerotteri (Ephemeroptera Hyatt & Arms, 1891) sono un ordine primitivo di Insetti Pterigoti emimetaboli conosciuti come uno dei più antichi lignaggi esistenti degli Insetti alati (Pterygota) (Edmunds & McCafferty 1988; Brittain & Sartori 2003). Tali organismi hanno un ampio areale di distribuzione (diffusi in gran parte del globo, eccetto che in Artide e in Antartide) e comprendono più di 3000 specie, raggruppate in circa 400 generi e 42 famiglie (Barber-James et al. 2008). Essi colonizzano tutti i tipi di acque dolci ma sono più diversificati in quelle correnti rispetto a quelle dei laghi o degli stagni. Sono aericoli allo stato adulto e dotati di larve acquatiche, una loro peculiare caratteristica è la presenza di uno stadio alato per lo più incapace di riprodursi, la subimmagine, che subisce una definitiva muta prima di trasformarsi in immagine. La prerogativa dell'apparato boccale, involuto e non funzionante, rende ben riconoscibili e caratterizzati gli adulti. I maschi sono ben distinguibili dalla femmina per il maggior sviluppo degli occhi e delle zampe anteriori, per la presenza dell'organo copulatore nell'ultimo segmento addominale e per un paio di appendici (i gonostili) che hanno la funzione di presa dell'addome della femmina. Le immagini hanno vita breve, che può durare da poche ore a qualche settimana, la brevità della vita adulta è all'origine del nome Efemerotteri (dal greco ephemeros = vivere un giorno). Esse compiono voli nuziali spesso in grandi masse di individui, prevalentemente maschi, in genere presso laghi o corsi d'acqua.



Il ciclo di sviluppo prevede 4 stadi di crescita: uovo, neanide, ninfa e adulto. Il primo stadio, l'uovo, è deposto secondo metodologie diverse a seconda dei generi, dal rilascio in volo sulla superficie dell'acqua (es. *Serratella ignita*) alla deposizione direttamente in acqua con immersioni ripetute dell'addome (es. *Ephemera danica* e alcune specie della famiglia Heptageniidae), in alcuni generi le femmine emettono le uova rimanendo sulla superficie con

le ali distese (es. gen. *Cloeon*), altre invece, come nel caso del genere *Baetis*, ripiegano le ali in modo da trattenere bolle d'aria e si immergono in acqua, deponendo le uova sotto le pietre.

Il numero di uova che viene deposto varia da poche centinaia a parecchie migliaia ed il loro sviluppo embrionale dura da qualche giorno a diversi mesi. Sono stati studiati anche casi di



sviluppo partenogenico in cui però il ciclo è molto più lungo. Le uova hanno inoltre strutture adesive che gli permettono di rimanere ancorati al substrato senza essere trasportati dalla corrente. Nello stadio successivo, neanide, l'individuo è caratterizzato da apparato boccale, occhi rudimentali e assenza di astucci alari e tracheobranchie, in questa fase subisce numerose mute. Lo stadio larvale compie lo sviluppo quando gli astucci alari cominciano a svilupparsi, in questa fase l'apparato boccale diventa masticatore, la respirazione è aiutata

dalle tracheobranchie addominali che con il loro movimento aumentano l'ossigenazione, presentano cerci ben sviluppati (Belfiore 1983). E' possibile distinguere differenti tipi di cicli vitali in relazione alla ciclicità e alla loro durata: cicli univoltini, polivoltini e semivoltini. I primi, sono caratteristici di specie che hanno una sola generazione all'anno (es.: generi *Oligoneuriella*, *Rhithrogena*), i secondi proprio di specie con due o più generazioni all'anno (es. generi *Ecdyonurus*, *Baetis*), in questi gli adulti delle seconde generazioni risultano più piccoli in dimensioni rispetto alle generazioni precedenti, gli ultimi, i semivoltini sono tipici di specie (es. *Ephemera*, *Ephoron*) in cui lo stadio larvale si protrae per un periodo intorno ai due anni, e che pertanto hanno una generazione ogni due anni. Le ninfe, acquatiche e fitofaghe hanno un'alimentazione basata principalmente di detriti in decomposizione di natura animale o vegetale, prediligono acque non inquinate, per cui sono considerate indicatori biologici nelle indagini ecologiche e di biomonitoraggio. Sono adattate a diverse nicchie ecologiche e condizioni ambientali, distinguendosi in quattro forme principali: camminatrici, nuotatrici, striscianti e scavatrici.

In accordo con la WFD, la loro abbondanza relativa e presenza permette agli esperti di applicare specifici strumenti metrici (gli indici sopra citati) ed assegnare valori corrispondenti alle classi di qualità correlate per gli ecosistemi acquatici.

2.3 Applicazione della tecnica DNA *barcoding* agli Efemerotteri

Il metodo DNA *barcoding* è stato fin da subito applicato all'ordine degli Insetti, per la quale si è rivelato un approccio affidabile, veloce e conveniente, ad esempio per la scoperta di nuove specie (es. Parassiti della famiglia Braconidae, Hymenoptera; Smith et al. 2008), per revisioni nomenclaturali e compilazione di liste nazionali (Coleoptera, Hendrich et al. 2014) e per l'associazione sia dei sessi nelle specie dimorfiche (Hymenoptera: Apoidea, Sheffield et al. 2009) che tra individui immaturi e adulti (Lepidoptera; Janzen et al. 2005). Questo approccio è risultato particolarmente utile anche nelle ricerche condotte sulle acque dolci, permettendo l'identificazione degli individui allo stadio larvale e negli esemplari di sesso femminile per le quali si riscontra spesso una tassonomia ambigua (Hajibabaei et al. 2011). Numerosi sono i lavori effettuati sui maggiori gruppi di insetti d'acqua dolce esistenti; Tricotteri (Hogg et al. 2009), Plecotteri (Zhou et al. 2009), Odonati (Rach et al. 2008), Ditteri Chironomidi (Ekrem et al. 2007) e Ditteri Simulidi (Rivera & Currie 2009). I risultati ottenuti hanno stimolato un crescente interesse per lo sviluppo di database di “codici a barre” che consentano indagini complete e veloci sulla fauna regionale delle acque dolci (Hajibabaei et al. 2012).

Nell'ordine degli Efemerotteri, l'efficacia del DNA *barcoding* è stato inizialmente testato da Ball et al. (2005) con 80 taxa campionati nel Nord-Est degli Stati Uniti e nel Canada centrale, seguito poi da Zhou et al. (2009) con 37 morfospesie provenienti dalle zone subartiche del Canada. Entrambe le ricerche hanno ottenuto un elevato livello di identificazione (99-100%) e data la consistente risoluzione ottenuta, soprattutto nei casi in cui l'identificazione morfologica era compromessa a causa della difficoltà di riconoscimento delle fasi iniziali o nel caso di campioni danneggiati o frammentati, hanno confermato la reale possibilità di inserire la suddetta tecnica nelle future applicazioni per studi di biomonitoraggio.

Successivamente Webb et al. (2012) hanno fornito sequenze *barcode* per più di 4000 individui, raggruppati in circa 350 specie, di Efemerotteri Nord Americani rappresentanti più del 50% della fauna nota nel continente.

Ulteriori studi hanno utilizzato tale metodica per chiarire le relazioni sistematiche a livello di famiglia (Baetidae, Sroka 2012), per differenziare le specie morfologicamente indistinguibili sia allo stadio di ninfe che di femmine adulte (Genere *Hexagenia*; Elderkin et al. 2012), per la comprensione delle strutture di popolazione locali (Genere *Epeorus*; Ogitali et al. 2011), per analizzare complessi di specie criptiche (*Baetis rhodani*; Williams et al. 2006; *Baetis vernus* Ståhls & Savolainen 2008) o per inferenze filogeografiche di taxa endemici (Gattolliat et al. 2015).

Nella regione mediterranea, non è attualmente disponibile alcun tipo di applicazione pratica del DNA *barcoding* agli Efemerotteri, nonostante questa sia uno dei 25 *hotspot* di biodiversità mondiali (Myers et al. 2000) e nella quale le acque dolci sono tra le principali componenti naturali più a rischio ed allo stesso tempo scarsamente studiate (Blondel & Aronson 1999).

3. SECONDO CASO DI STUDIO: La pesca marittima

La pesca marittima è considerata una delle principali attività economiche praticate su scala mondiale. Secondo i dati FAO nel 2012 la produzione mondiale di pescato ha raggiunto i 91 milioni di tonnellate annue (FAO 2014). Le precedenti stime effettuate da Santojanni (2001) affermavano un incremento esponenziale della produzione da circa 4 milioni tonnellate annue all'inizio del '900 a circa 70 milioni tonnellate annue negli anni '70. Questo processo è stato senza dubbio incentivato dallo sviluppo tecnologico che ha permesso l'utilizzo di battelli sempre più veloci ed attrezzati (impianti di refrigerazione e dotazioni elettroniche capaci di individuare banchi di pesce).

Lo sfruttamento eccessivo degli stock ha portato la comunità europea all'emanazione di un regolamento europeo relativo alle misure di gestione per lo sfruttamento sostenibile delle risorse della pesca nel Mar Mediterraneo (n°1967 del 2006).

Nei primi anni del 2000 l'attività di pesca ha subito un calo in funzione proprio degli incentivi per la riduzione della pressione di pesca e il mantenimento degli stock ma, nonostante gli sforzi effettuati in questo senso, i dati FAO del 2009 relativi al 2007 registrarono un sovrasfruttamento per il 28% degli stock, un impoverimento per l'8% degli stessi, mentre solamente l'1% appariva in recupero. Alla luce del quadro mostrato è stata messa in evidenza la necessità di effettuare politiche gestionali mirate a mantenere e preservare la fauna ittica.

In questo senso, proprio perché il settore della pesca offre opportunità lavorative ad innumerevoli aziende piccole, medie e grandi operanti in Italia, sono nati diversi piani di gestione per la pesca in differenti zone del Mediterraneo allo scopo di mantenere le condizioni sia per il sostentamento degli stock entro i limiti biologici di sicurezza che per quello delle attività produttive stesse (vedi ad esempio il Piano Di Gestione Nazionale Per La Sciabica E Reti A Circonazione Senza Chiusura (SEINES) Ex art. 19 e art. 15 del Reg (CE) n. 1967/2006, Piano Di Gestione (ex art.24 del Reg. (CE) n.1198/2006), GSA 10 Mar Tirreno Centro-Meridionale Strascico).

3.1 La situazione degli stock di Alici e Sarde in Italia

I dati IREPA del 2009 hanno messo in evidenza come in Italia la cattura di queste specie ha rappresentato circa il 26% del pescato totale, mettendo in evidenza come esse siano tra le più importanti risorse alieutiche in diverse zone del Mediterraneo.

Le valutazioni sulle dinamiche degli stock di Alice e Sarde effettuate in differenti parti del mondo sono concordi nell'affermare che le condizioni ambientali presenti negli areali di distribuzione di queste specie incidono sull'espansione o la contrazione sia dell'abbondanza che dell'areale stesso (Lindegren et al. 2013; MacCall et al. 2015; Zhou et al. 2015). Nello specifico è importante sottolineare come le dinamiche biologiche ed ambientali incidano sulla sopravvivenza dei primi stadi vitali di queste specie e come esse diventano quindi essenziali nel determinare gli incrementi e le contrazioni annuali dello stock adulto (Cuttitta et al. 2003, 2006; Coombs et al. 2004; Bachiller et al. 2015).

Ne consegue che per il corretto sfruttamento delle risorse alieutiche diventa di notevole importanza lo studio dei cicli ittioplantonici in relazione alle componenti ambientali ed agli altri organismi.

3.2 Biologia delle specie studiate

Le specie oggetto di questo studio sono l'alice (*Engraulis encrasicolus*) e la sardina (*Sardina pilchardus*). Entrambe sono pesci ossei appartenenti alla classe degli Actinopterygii, ordine Clupeiformes.

Questo ordine è caratterizzato da pesci marini pelagici, contraddistinti da pinna dorsale, pinne ventrali o pelviche posizionate nella zona addominale, raggi di pinne molli, grosse squame cicloidi che ricoprono tutto il corpo ad eccezione della testa. La vescica natatoria si prolunga fino alla testa ed è comunicante con il tratto gastro-esofageo (fisostomi).

Sono generalmente gregari, distribuiti nelle acque costiere poco profonde ed in acque salmastre delle zone temperate e tropicali.

Nel Mediterraneo sono presenti anche le specie *Sardinella aurita*, *Alosa fallax* e la sottospecie agone (*Alosa fallax lacustris*), ma sono le alici e sarde ad avere il più grande interesse dal punto di vista commerciale.

Dal punto di vista della catena trofica i piccoli pelagici hanno un ruolo fondamentale nell'ecosistema catturando energia dai livelli più bassi per trasferirla ai grossi predatori (Coll et al. 2007). Allo stesso tempo queste popolazioni sono soggette a consistenti oscillazioni nello

spazio e nel tempo dovute principalmente alle variazioni ambientali, al loro breve ciclo vitale e dalla pressione della pesca (Azzali et al. 2002; Santojanni et al. 2003, 2005 e 2006; Leonori 2007).

3.2.1 *Engraulis encrasicolus* (Linnaeus, 1758)

Appartenente alla famiglia Engraulidae, questa specie è l'unica rappresentante nel Mediterraneo (presente anche nel Mar Nero). Ha un'areale di distribuzione che si sviluppa lungo le coste dell'Atlantico orientale, dalla Norvegia al Sud Africa. Predilige acque marine e salmastre, sopporta bene la salinità e durante il periodo riproduttivo può arrivare fino ad estuari, laghi e lagune.

Ha il corpo sottile e slanciato, il muso prolungato con mascella anteriore che oltrepassa la mascella inferiore, la livrea è di colore blu, verde, tendente al grigio, con una striscia argento che cammina lungo i fianchi ma che tende a scomparire con l'avanzare del ciclo biologico (Fischer et al. 1987) (Fig.7).

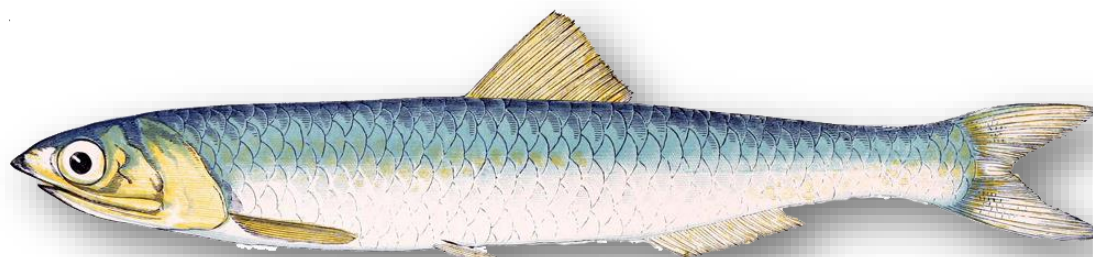


Fig.7 - *Engraulis encrasicolus*

La specie è fortemente pelagica con una taglia che si aggira tra i 15 ed i 20 cm di lunghezza, è gregaria e vive in grossi banchi a profondità abbastanza elevate (intorno ai 150 m nel Mar Mediterraneo), in primavera-estate effettua grandi migrazioni portandosi alla costa dove è maggiore l'approvvigionamento di cibo grazie all'apporto fluviale (Fischer et al. 1987).

Si nutre principalmente nelle ore diurne di zooplancton (principalmente copepodi e larve di molluschi), uova di pesce e larve, il periodo di riproduzione avviene in un lasso temporale che va da aprile a novembre con un picco nei periodi estivi (Gramitto 2001).

Ogni femmina depone circa 40.000 uova che rimangono in ambiente pelagico (circa 50 m) dove vengono fecondate. La schiusa di quest'ultime avviene dopo circa due giorni e le larve di

circa 2 mm cominciano subito la loro vita gregaria (Fischer et al. 1987); i giovanili e gli immaturi costituiscono insieme ai giovanili di sarde il “bianchetto”.

L’area di deposizione delle uova (anche detta spawning) è molto ampia comprende tutto il Mediterraneo con alte incidenze nell’Adriatico Nord, nel Gargano, Nel Tirreno Nord Occidentale, il Golfo di Biscaglia di Tunisi e in Sicilia (Zarrad et al. 2006; Giannoulaki et al. 2013; Bonanno et al. 2014).

Gli stadi larvali sono quelli maggiormente soggetti a mortalità causata in gran parte da predazione. Infatti, sono noti i ritrovamenti di uova e parti di larve negli stomaci di sardine (Vučetić 1963) o di meduse (*Pelagia noctiluca*; Zavodnik 1991).

Predatori di questa specie sono anche merluzzo (*Merluccius merluccius*; Frogia 1973), lanzardo (*Scomber japonicus*), tonnidi (tonno rosso, *Thunnus thynnus*; Orsi Relini et al. 1999), ma anche uccelli marini e delfini (Coll et al. 2007).

Altro fattore fondamentale consiste nell’intensità di pesca effettuata che, colpendo in maniera preoccupante gli stadi giovanili di questa specie, le impedisce lo sviluppo fino allo stadio adulto. Essa rappresenta la specie commercialmente più appetibile e di valore, la cattura in questo periodo del ciclo biologico provoca un notevole impatto sia sulla catena alimentare marina che sull’economia della filiera di pesca (Sanfilippo et al. 2011).

3.2.2 *Sardina pilchardus* (Walbaum, 1792)

Questa specie appartiene alla famiglia Clupeidae, nel Mediterraneo è composta da altre tre specie pelagiche; alaccia (*Sardinella aurita*), alosa o cheppia (*Alosa fallax*) e spratto (*Sprattus sprattus*).

E’ presente in tutte le coste atlantiche delle regioni europee, nel Mediterraneo ha differenti abbondanze tra le zone orientali (dove è meno presente) e le zone occidentali (dove la presenza è più consistente).

Questa specie è caratterizzata da un corpo affusolato leggermente compresso lateralmente, ha un muso acuto, le mascelle terminano sotto l’occhio con quella inferiore più pronunciata di quella superiore. Possiede una livrea color oliva con punteggiature nerastre laterali e una fascia bluastra che attraversa il corpo longitudinalmente (Fig.8).

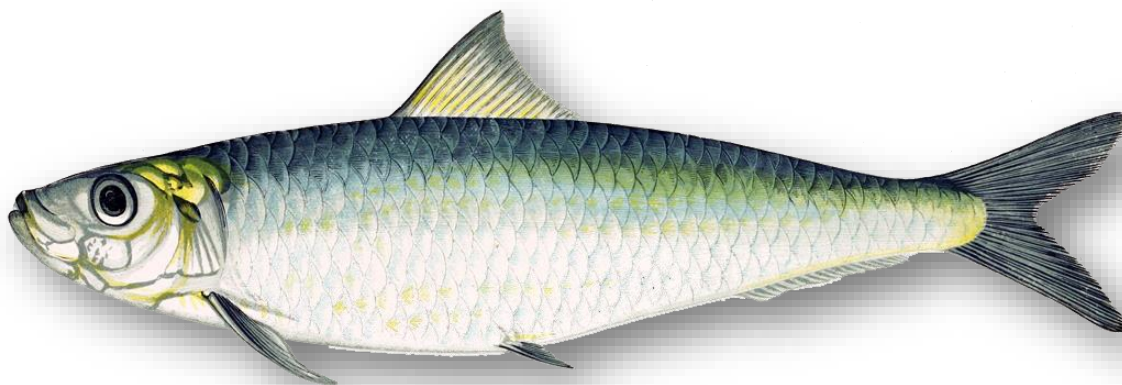


Fig.8 - *Sardina pilchardus*

Le dimensioni che possono raggiungere sono leggermente maggiori della specie sopra descritta (circa 21 cm).

Ha la tendenza ad aggregarsi in banchi di notevoli dimensioni che possono essere caratterizzati sia da una sola specie che da più specie e percorre grandi distanze soprattutto sotto l'influenza della temperatura (Morello & Arneri 2009).

Raggiunge profondità di circa 180 m in inverno, mentre nel periodo primaverile estivo raggiunge la costa disponendosi a profondità tra i 20 ed i 55 m.

L'alimentazione è prevalentemente caratterizzata da crostacei o di organismi più grandi. La sardina cattura le prede una per volta a differenza delle altre specie della stessa famiglia che utilizzano le branchiospine per trattenere il cibo incanalandolo verso lo stomaco (Gramitto 2001). Ha una buona tolleranza alle variazioni di temperature e salinità.

L'attività riproduttiva può avvenire più volte durante il ciclo vitale (specie iteropara), rilascia gruppi di uova ("batch") che maturano in maniera differente durante il periodo riproduttivo (Ganias et al. 2003). E' una specie ad elevata fecondità, produce mediamente intorno alle 55.000 uova per individuo. Questa caratteristica è favorita sia dal lungo periodo di riproduzione che parte ad inizio autunno (Ottobre) per terminare quasi a fine primavera (maggio) sia dalle condizioni ambientali che mantengono continua la gametogenesi (Sinovčić 2009). I diversi picchi di deposizione sembrano essere influenzati dalla temperatura, dalla salinità e dalla disposizione di cibo (Regner et al. 1987).

La mortalità della sardina è causata principalmente dalla predazione, le larve sono spesso cacciate dal lanzardo (*Scomber japonicus*), mentre gli adulti da merluzzo, tonni e pesce spada, ma anche delfini ed uccelli marini (Coll et al. 2007).

3.3 Il “bianchetto”

Con il termine “bianchetto” si intende la neonata o novellame di alici e sarde (pesce azzurro). Questo prodotto molto famoso commercialmente ha differenti nomi a seconda delle regioni e dei dialetti. Si presentano come una massa gelatinosa traslucida composta di organismi di lunghezza poco superiore ai 2 cm (Fig.9).



Fig.9 - Particolare del “bianchetto”

Riguardo la ‘raccolta’ di questo prodotto si hanno cenni che risalgono alla fine dell’800, provenienti da un libro pubblicato nel 2002, “San Vincenzo e la pesca del pesce azzurro” di Vincenzo Biagi. Nel libro vi sono racconti dei vecchi della zona che parlavano di grandi quantità di bianchetto, soprattutto in periodi di alta marea, tali da “attappare i fossi”.

La pesca di questo pesce è quindi effettuata da secoli lungo le coste italiane ma l’intensa attività degli ultimi decenni, accentuata dalla grande richiesta di questo prodotto, ha portato ad un intenso impoverimento degli stock di alici e sarde, gli organismi infatti, vengono prelevati in stadi ancora non maturi rendendo impossibile la riproduzione degli stessi.

Questo processo è avvenuto a causa dell’eccessivo sfruttamento delle risorse ittiche di grandi dimensioni che hanno diminuito le quantità presenti e concentrato gli sforzi di pesca su individui più piccoli, portando ad una riduzione sensibile delle femmine che raggiungono età fertile e minacciando così la sopravvivenza delle popolazioni.

L’attuale situazione ha portato all’intervento degli organismi preposti alla tutela con regolamentazioni in materia di pesca.

Appositi Decreti Ministeriali regolamentano la pesca commerciale italiana. Nello specifico il D.M. del 1° dicembre 1989 e il successivo D.M. del 3 ottobre 1994 permettono una raccolta annuale dal 1 dicembre al 30 aprile per un periodo consecutivo di massimo due mesi.

Con l'entrata in vigore del Regolamento Europeo CE n. 1967/2006, che prende in considerazione le misure di gestione per lo sfruttamento sostenibile delle risorse alieutiche nel Mar Mediterraneo, viene vietato questo tipo di pescato. Nello specifico l'articolo 15 del suddetto impone le taglie minime degli organismi marini vietandone la vendita, la tenuta a bordo, il trasbordo, lo sbarco, il trasferimento, l'immagazzinamento e l'esposizione.

Esistono però delle deroghe che possono essere richieste alla Comunità Europea. Queste devono riguardare esclusivamente il novellame di sardine per consumo umano a condizione che tali pratiche siano effettuate con sciabiche da natante o da spiaggia, che le relative attività siano regolate da disposizioni nazionali stabilite in un piano di gestione approvato dall'Unione Europea e che non incidano sugli stock di sardine mantenendoli nei limiti biologici di sicurezza.

Il periodo dell'anno in cui la deroga è consentita deve assicurare quindi l'esclusiva (o perlomeno la maggiore incidenza) di *Sardina pilchardus* rispetto *Engraulis encrasicolus*. Avendo queste due specie un ciclo biologico che si accavalla, risulta necessario uno studio periodico del bianchetto e la valutazione della presenza delle rispettive specie al fine di evitare multe e sequestri di materiale ittico dovuti ad inottemperanze alla legge.

Molte delle regioni italiane non hanno avuto i permessi di deroga ma perseverano comunque nelle loro pratiche di pesca poiché il bianchetto è considerato un'importante fonte di reddito per le comunità di pescatori locali ed una tradizione da mantenere. Questo comporta un grosso danno alle comunità locali in quanto periodicamente ingenti quantità di prodotto vengono sequestrate dai banchi di pesca locali ed i pescatori soggetti a multe salatissime.

L'esclusiva presenza di *Sardina pilchardus* nel bianchetto dipende dal periodo in cui esso viene pescato ma soprattutto dal riconoscimento derivato dallo studio dei dati morfometrici acquisiti dagli esemplari catturati (Cutitta et al. 2015). Quest'ultimo è un procedimento che deve essere effettuato da personale esperto con attrezzature specializzate e può risultare comunque lento, laborioso e dispendioso; basti pensare che gli esemplari da analizzare oltre ad avere una grandezza di circa 2 cm (Fig.10) sono molto fragili ed i caratteri morfologici discriminanti possono non essere visibili a causa di danneggiamenti.



Fig.10 - Neonata di *Engraulis encrasicolus*

Metodiche di discriminazione basate sulle differenze di sequenza genomica possono aiutare notevolmente questo tipo di lavoro. Lo sviluppo tecnologico ha permesso in questo senso un abbattimento dei costi, permettendo un accesso più libero a questo tipo di approcci che si sono rivelati efficienti, economici e veloci.

L'applicabilità della tecnica del DNA *barcoding* a questo prodotto, unito alla tecnologia di sequenziamento basata sulla *Next Generation Sequencing*, potrebbe risolvere i problemi di difficile discriminazione delle due specie all'interno della massa di novellame pescato. Questo permetterebbe di comprendere in maniera scientifica il giusto periodo dell'anno in cui la presenza di alici è tale da rientrare nei limiti dei regolamenti.

4. SCOPO DEL LAVORO

La presente ricerca ha avuto come obiettivo finale quello di valutare l'efficacia della tecnica DNA *barcoding* applicata a due problematiche tassonomiche ed ecologiche riguardanti le acque dolci e quelle marine.

Nel primo caso di studio, il DNA *barcoding* è stato applicato agli Efemerotteri, con lo scopo di ampliare le conoscenze di questo gruppo tassonomico, tutt'ora in continua evoluzione sistematica, e proporre prospettive applicative riguardanti le metodologie di biomonitoraggio. In una prima fase è stata testata la capacità discriminatoria di tale tecnica ponendo le basi per un primo database di sequenze *barcode* di diverse specie campionate in un'area della regione mediterranea. Questo, sviluppato ulteriormente su scala nazionale e continentale, potrebbe costituire in futuro una banca dati di riferimento sugli Efemerotteri, per effettuare misure ripetute della biodiversità con una risoluzione sufficiente per essere utile in ecologia e nel biomonitoraggio delle acque dolci, per studi sulla perdita di biodiversità, per la mitigazione e recupero di ambienti alterati e per evitare errori causati dalla scarsità di dati nella gestione del territorio. Questo studio è stato oggetto di una pubblicazione in corso di dottorato (Cardoni et al 2015, vedere Appendice IV)

Successivamente, nella seconda fase della ricerca è stata testata l'applicabilità del *metabarcoding* al biomonitoraggio delle acque dolci e marine attraverso l'utilizzo della piattaforma *Next Generation Sequencing (NGS)*.

Il secondo caso studiato (acque marine) ha riguardato lo sviluppo e la verifica di protocolli applicativi atti a determinare l'efficacia del metodo e la capacità risolutiva a livello di genere e specie delle forme giovanili di *Engraulis encrasicolus* e *Sardina pilchardus*. Ai fini legali e commerciali, una pronta e sicura determinazione dei rapporti percentuali di queste due specie nel bianchetto permetterebbe alle istituzioni regionali di chiedere la deroga per questo tipo di pesca secondo le normative vigenti.

Uno schema riassuntivo delle diverse fasi che hanno caratterizzato questo lavoro è riportata in Fig.11.

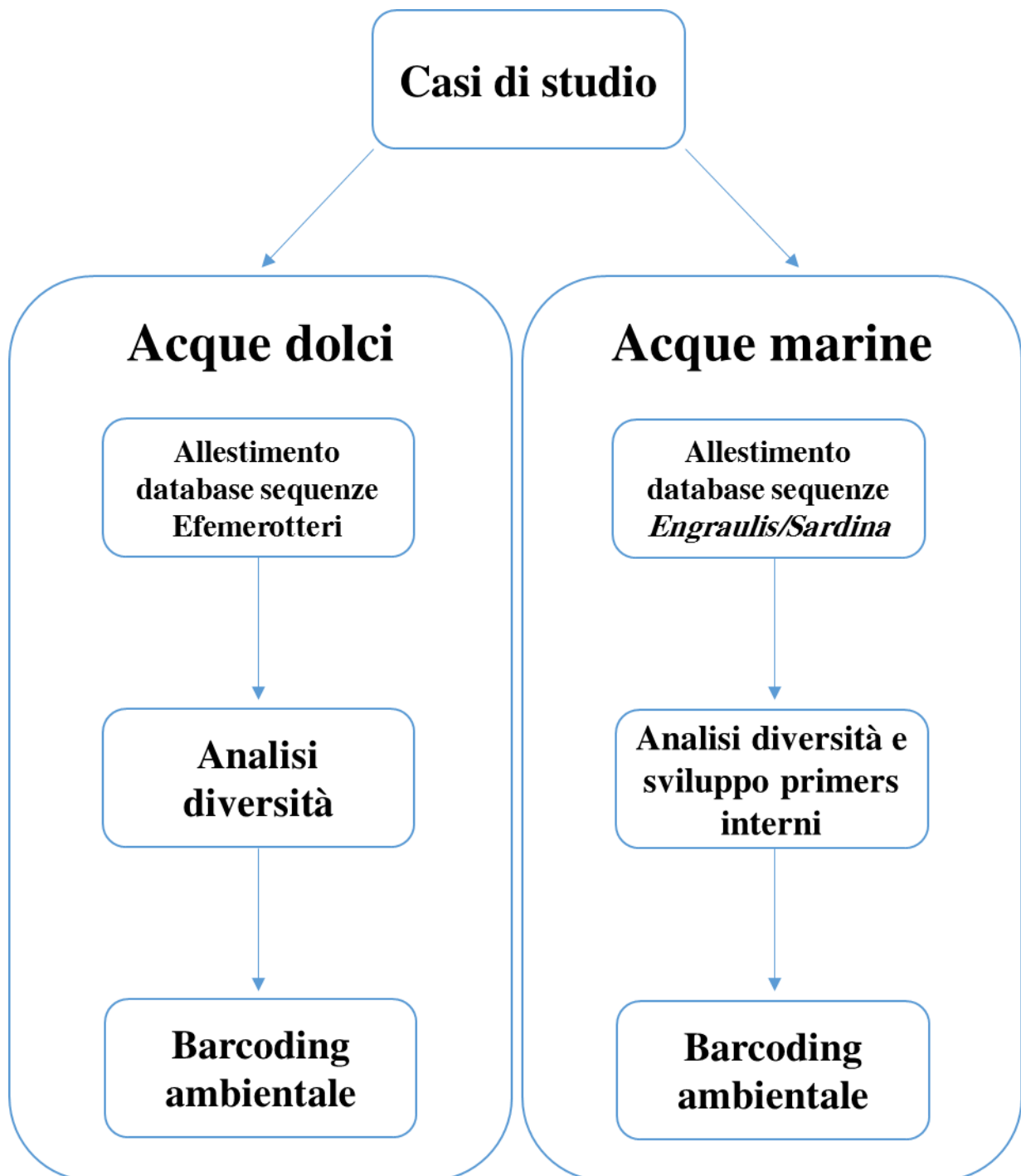


Fig.11 - Mappa concettuale che riassume le fasi di lavoro

5. MATERIALI E METODI

5.1 Campionamento acque dolci

5.1.1 Prima fase di campionamento

La prima fase di campionamento, mirata alla raccolta della maggior quantità possibile di specie di Efemerotteri, è stata effettuata tra Maggio 2012 e Aprile 2014 in fiumi e torrenti dell'area mediterranea. I siti oggetto dello studio sono stati considerati nelle regioni del centro Italia (Toscana e Lazio) e nelle principali isole tirreniche (Sardegna, Elba e Corsica) (Fig.12).

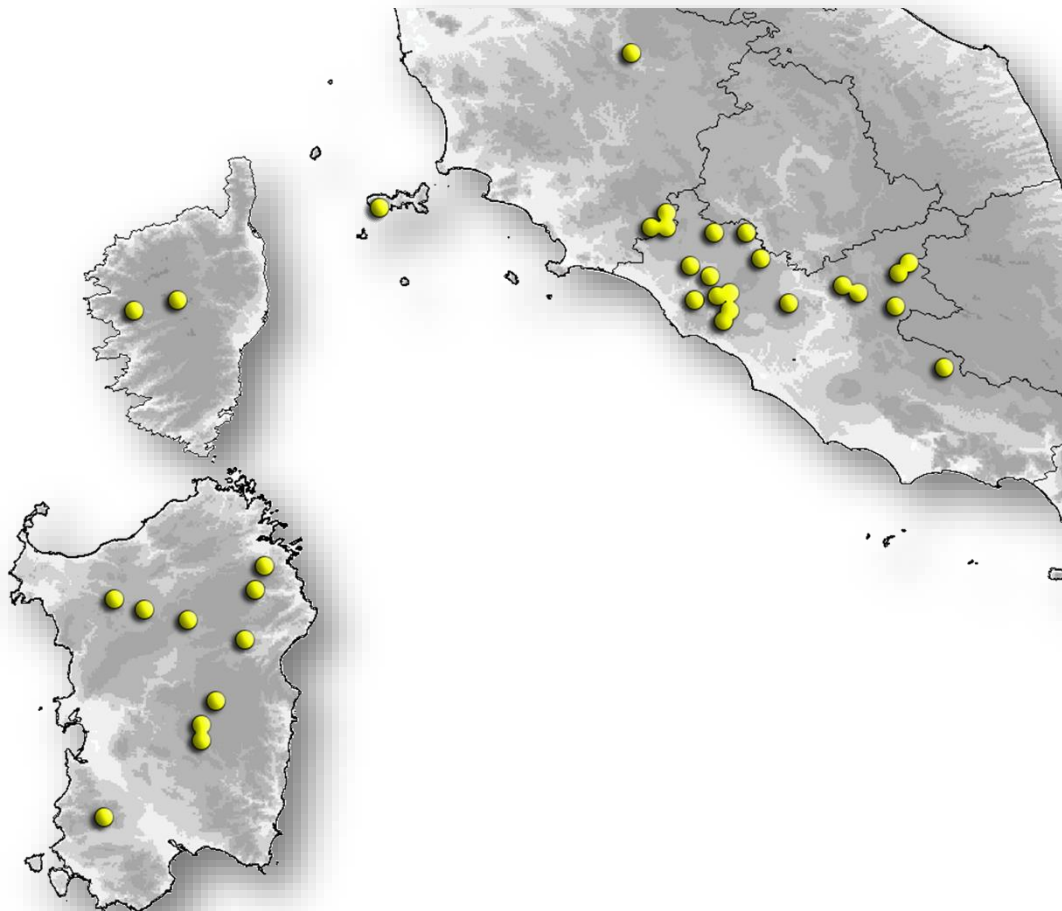


Fig.12 - Area geografica (A) e di distribuzione (B) dei siti di campionamento investigati.

La raccolta è stata effettuata con un retino immanicato effettuando la tecnica del “kick sampling” (ovvero smuovendo il substrato di fondo direttamente con il piede per circa due minuti) e si è concentrata esclusivamente sulle larve di Efemerotteri. Il materiale raccolto è stato smistato in sito attraverso l’utilizzo di apposite bacinelle e successivamente conservato in provette contenenti alcol etilico non denaturato al 95%. Quando possibile si è cercato di recuperare tre esemplari per ciascuna specie.

La fase di identificazione morfologica è stata effettuata presso il laboratorio di ecologia delle acque dolci (Università Degli studi della Tuscia) attraverso uno stereoscopio Leica S8APO o un microscopio ottico OLYMPUS CH, utilizzando apposite chiavi di identificazione (ad esempio Belfiore 1983). Il dataset finale conta 91 individui corrispondenti a 39 specie morfologicamente riconosciute (18 generi, 8 famiglie) (Tab.2).

Tab.2 - Lista dei campioni analizzati con la relativa nomenclatura tassonomica in accordo con Fauna Europea (aggiornata al 29 Agosto 2013 | versione 2.6.2), siti campionati e numero di accessione in GeneBank delle sequenze COX1 prodotte in questo studio.

Famiglia	Genere	Specie	Località	N° Access.	
Baetidae	<i>Baetis</i>	<i>albinatii</i> Sartori & Thomas	Siligo, Sardinia	LN734680	
		<i>muticus</i> Linnaeus	Vitozza, Tuscany	LN734678	
		<i>vernus</i> Curtis	Treja, Latium	LN734679	
		<i>cyrneus</i> Thomas & Gazagnes	Villanova Tulo, Sardinia	LN734681	
			Posada, Sardinia, specimen 1	LN734682	
			Posada, Sardinia, specimen 2	LN734683	
			Posada, Sardinia, specimen 3	LN734684	
		<i>lutheri</i> Müller-Liebenau	Meleta Pitigliano, Tuscany	LN734685	
			Aquarella Tuscania, Latium	LN734686	
		<i>fuscatus</i> Linnaeus	Lunevara, Sardinia	LN734687	
			Rio Mannu, Sardinia	LN734688	
		<i>alpinus</i> Pictet	Fosso di Galantina, Latium	LN734689	
			Aniene, Latium	LN734690	
			Obito, Latium	LN734691	
		<i>buceratus</i> 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	
		<i>Centroptilum</i>	<i>luteolum</i> Müller	Rafanello, Latium	LN734697
				Fosso Tancia, Latium	LN734698
				Fosso di Galantina, Latium	LN734699
		<i>Procloeon</i>	<i>pulchrum</i> Eaton	Vitozza, Tuscany	LN734700
		Ortolano, Latium	LN734701		
		Rio Chiaro, Latium	LN734702		
	<i>bifidum</i> Bengtsson	Mignone Montericcio, Latium	LN734703		
		Seccheto, Elba Island	LN734704		
		Torrente Leni, Sardinia	LN734705		
<i>Cloeon</i>	<i>dipterum</i> Linnaeus	Mignone Canale Monterano, Latium	LN734706		
		Aquarella Tuscania, Latium	LN734707		
	<i>simile</i> Eaton	Tirso, Sardinia	LN734708		

Caenidae	<i>Caenis</i>	<i>pusilla</i> Navas	Mignone Montericcio, Latium	LN734709	
			Aquarella Tuscania, Latium	LN734710	
		<i>beskidensis</i> Sowa	Vitozza, Tuscany	LN734711	
			Rio Chiaro, Latium	LN734712	
		<i>martae</i> Belfiore	Aquarella Tuscania, Latium	LN734713	
			Mola di Oriolo, Latium	LN734714	
			Ortolano, Latium	LN734715	
		<i>sp.</i>	Riofi, tuscany	LN734716	
			Tirso, Sardinia	LN734717	
		<i>Brachycercus</i>	<i>harrisella</i> Curtis	Vitozza, Tuscany	LN734718
				Rio Chiaro, Latium	LN734719
Heptageniidae	<i>Ecdyonurus</i>	<i>helveticus</i> Eaton	Fosso Tancia, Latium	LN734720	
			Ortolano, Latium	LN734721	
			Meleta Collina, Tuscany	LN734722	
		<i>corsicus</i> Esben-Petersen	Rio Tistigliosi, Sardinia	LN734723	
			<i>venosus</i> Fabricius	Ortolano, Latium, specimen 1	LN734724
		<i>Epeorus</i>	<i>assimilis</i> Eaton	Introvabile, Latium	LN734725
				Mignone Canale Monterano, Latium	LN734726
				Ortolano, Latium, specimen 2	LN734727
				Obito, Latium	LN734728
				Aniene, Latium	LN734729
	Ortolano, Latium			LN734730	
	<i>Electrogena</i>	<i>grandiae</i> Belfiore	Introvabile, Latium	LN734731	
			Rafanello, Latium	LN734732	
		<i>zebrata</i> Hagen	Rio Mannu, Sardinia	LN734733	
	<i>Rhithrogena</i>	<i>semicolorata</i> Curtis	Padru Rio de su Lenu, Sardinia	LN734734	
			Mignone Canale Monterano, Latium	LN734735	
			Fosso Tancia, Latium	LN734736	
			Rio di Sesto, Trentino	LN734737	
			<i>reatina</i> Sowa & Belfiore	Velino Antrodoco, Latium	LN734738
			Velino Canetra, Latium	LN734739	
			<i>nuragica</i> Belfiore	Rio Aratu, Sardinia	LN734740
			<i>insularis</i> Esben-Petersen	Restonica, Corse	LN734741
				Murzo, Corse	LN734742
			<i>eatoni</i> Esben-Petersen	Murzo, Corse	LN734743
	<i>Heptagenia</i>	<i>longicauda</i> Stephens	Rigomero, Latium	LN734744	
			Aquarella Tuscania, Latium	LN734745	
			Rio Chiaro, Latium	LN734746	
	Leptophlebiidae	<i>Habroleptoides</i>	<i>confusa</i> Sartori & Jacob	Fosso di Galantina, Latium	LN734747
				Rafanello, Latium	LN734748
				Ortolano, Latium	LN734749
<i>sp.</i>			Rio Tistigliosi, Sardinia	LN734750	
<i>Habrophlebia</i>			<i>eldae</i> Jacob & Sartori	Vitozza, Tuscany	LN734751
		Fosso Arlena, Latium		LN734752	
<i>Paralaptophlebia</i>		<i>submarginata</i> Stephens	Rafanello, Latium	LN734753	
			Aquarella Tuscania, Latium	LN734754	
			Vitozza, Tuscany, specimen 1	LN734755	
			Vitozza, Tuscany, specimen 2	LN734756	
<i>Choroterpes</i>		<i>picteti</i> Eaton	Mignone Montericcio, Latium	LN734757	
Ephemerellidae	<i>Serratella</i>	<i>ignita</i> Poda	Vitozza, Tuscany	LN734758	
			Rafanello, Latium	LN734759	
			Fosso Arlena, Latium	LN734760	
			Ortolano, Latium	LN734761	
Oligoneuridae	<i>Oligoneuriella</i>	<i>rhenana</i> Imhoff	Meleta Pitigliano, Tuscany	LN734762	
			Rio Chiaro, Latium	LN734763	
Siphonuridae	<i>Siphonurus</i>	<i>lacustris</i> Eaton	Vitozza, Tuscany	LN734764	
			Velino Antrodoco, Latium	LN734765	
Ephemeridae	<i>Ephemer</i>	<i>danica</i> Müller	Vitozza, Tuscany	LN734766	

5.1.2 Seconda fase di campionamento

Nella successiva fase, gli sforzi di campionamento si sono concentrati sulla raccolta di macroinvertebrati bentonici, utilizzati come bioindicatori di stato ecologico delle acque. Sono stati campionati tre siti: sul torrente Biedano in località Barbarano Romano (VT), Monteromano e sul torrente Traponzo provincia di Viterbo, nel mese di Ottobre 2015. I siti di campionamento (Fig.13) sono stati selezionati sulla base del gradiente qualitativo dell'acqua atteso, il primo trovandosi nel tratto alto, non sottoposto a particolari pressioni antropiche, l'ultimo posto alla confluenza di un affluente soggetto agli scarichi della città di Viterbo e il secondo in situazione intermedia.

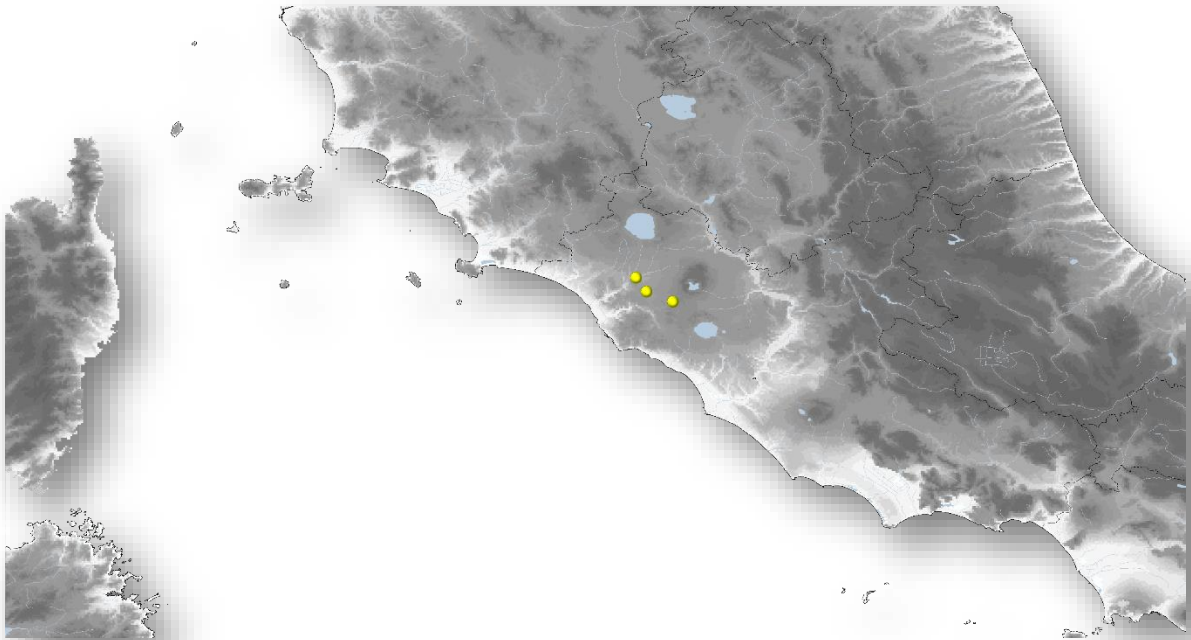


Fig.13 - Siti dei 3 campionamenti per l'analisi NGS

In questa fase è stato raccolto tutto il materiale di macrobentos presente, conservato immediatamente con le stesse modalità riportate per la fase precedente. Successivamente il materiale è stato identificato in laboratorio, con gli strumenti utilizzati nella fase precedente e chiavi di determinazione apposite (Belfiore 1983; Sansoni 1988; Campaioli et al. 1994, 1999; Tachet et al. 2006). Ogni individuo è stato misurato su carta millimetrata durante l'identificazione, in modo da avere un riferimento per la normalizzazione dei risultati una volta ottenute le sequenze.

Il dataset finale conta 455 individui raggruppati in 26 famiglie, 13 generi, in 115 esemplari si è arrivati alla specie come mostra la tabella riportata in Appendice II.

Contemporaneamente, sono stati raccolti dei campioni di sedimento nella zona di pool della prima stazione campionata al fine di verificarne un'eventuale presenza di DNA e posizionati in 3 tubi Falcon da 50 ml posti in congelatore a -20°C entro un paio di ore dal campionamento.

Infine, nella medesima zona di pool, sono stati raccolti campioni di acqua e conservati in una tanica da 5L. In seguito, questi sono stati sottoposti ad una procedura di filtraggio utilizzando una pompa a vuoto. Un primo filtraggio è stato effettuato con un filtro “Glass-Microfibre Discs” Di 4.7 cm di diametro ed una maglia di 2.7 micron, successivamente lo stesso campione di acqua è stato filtrato con un filtro del medesimo diametro ma con una maglia di 1.6 micron ed infine la stessa procedura con un filtro a maglia di 0.45 micron. Tutti i filtri con gli altri campioni ricavati sono stati conservati in congelatore a -20°C fino al momento della fase di estrazione del DNA.

5.2 Campionamento acque marine

Le strategie di campionamento attuate nella fase di studio delle acque marine si sono concentrate sul recupero di esemplari di *Engraulis encrasicolus* e *Sardina pilchardus*. La raccolta è stata effettuata nel periodo 11 marzo 2015 / 22 settembre 2015 cercando di rappresentare tutte le zone battute dai pescherecci della flotta italiana (Fig.14)

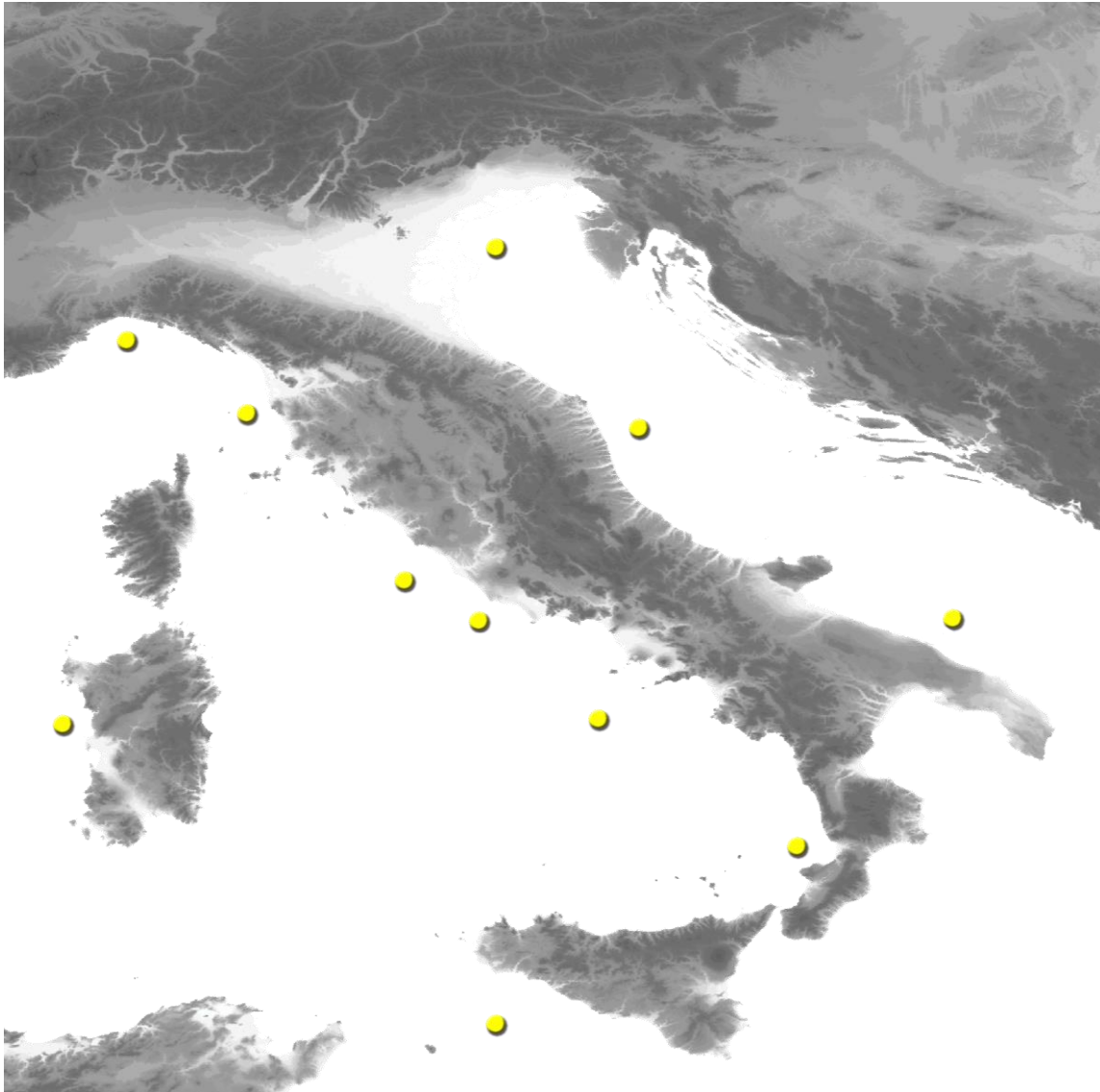


Fig.14 - Mappa dei siti campionati nel recupero degli esemplari di alici e sardine.

I campioni sono stati prelevati freschi direttamente dalle pescherie o dai pescherecci approdati al porto. La lista finale conta 44 individui di cui 13 appartenenti a *Sardina pilchardus* e 31 appartenenti a *Engraulis encrasicolus* (Tab.3) ed un campione di bianchetto della zona di Vibo Valentia.

Tab.3- Lista degli esemplari campionati con la data e la relativa zona di recupero.

Ordine	Species	Provenienza	Data
CLUPEIFORMES	<i>Sardina pilchardus 1</i>	Civitavecchia	11.03.15
	<i>Sardina pilchardus 2</i>	Civitavecchia	11.03.15
	<i>Sardina pilchardus 3</i>	Civitavecchia	11.03.15
	<i>Sardina pilchardus 4</i>	Civitavecchia	11.03.15
	<i>Engraulis encrasicolus 1</i>	Civitavecchia	11.03.15
	<i>Engraulis encrasicolus 2</i>	Civitavecchia	11.03.15
	<i>Engraulis encrasicolus 3</i>	Civitavecchia	11.03.15
	<i>Engraulis encrasicolus 4</i>	Civitavecchia	11.03.15
	<i>Sardina pilchardus 5</i>	S.Benedetto/Ancona	08.06.15
	<i>Sardina pilchardus 6</i>	S.Benedetto/Ancona	08.06.15
	<i>Sardina pilchardus 7</i>	S.Benedetto/Ancona	08.06.15
	<i>Engraulis encrasicolus 5</i>	S.Benedetto/Ancona	08.06.15
	<i>Engraulis encrasicolus 6</i>	S.Benedetto/Ancona	08.06.15
	<i>Engraulis encrasicolus 7</i>	S.Benedetto/Ancona	08.06.15
	<i>Engraulis encrasicolus 8</i>	Chioggia	19.06.15
	<i>Engraulis encrasicolus 9</i>	Chioggia	19.06.15
	<i>Engraulis encrasicolus 10</i>	Chioggia	19.06.15
	<i>Sardina pilchardus 8</i>	Chioggia	19.06.15
	<i>Sardina pilchardus 9</i>	Chioggia	19.06.15
	<i>Sardina pilchardus 10</i>	Chioggia	19.06.15
	<i>Engraulis encrasicolus 11</i>	Isola di Gorgona	30.07.15
	<i>Engraulis encrasicolus 12</i>	Isola di Gorgona	30.07.15
	<i>Engraulis encrasicolus 13</i>	Isola di Gorgona	30.07.15
	<i>Engraulis encrasicolus 14</i>	Genova	22.08.15
	<i>Engraulis encrasicolus 15</i>	Genova	22.08.15
	<i>Engraulis encrasicolus 16</i>	Genova	22.08.15
	<i>Engraulis encrasicolus 17</i>	Anzio	12.09.15
	<i>Engraulis encrasicolus 18</i>	Anzio	12.09.15
	<i>Engraulis encrasicolus 19</i>	Anzio	12.09.05
	<i>Engraulis encrasicolus 20</i>	Porto Palo	08.09.15
	<i>Engraulis encrasicolus 21</i>	Porto Palo	08.09.15
<i>Engraulis encrasicolus 22</i>	Porto Palo	08.09.15	
<i>Engraulis encrasicolus 23</i>	Bari	8.15	
<i>Engraulis encrasicolus 24</i>	Bari	8.15	
<i>Engraulis encrasicolus 25</i>	Bari	8.15	
<i>Engraulis encrasicolus 26</i>	Napoli	9.15	
<i>Engraulis encrasicolus 27</i>	Napoli	9.15	
<i>Engraulis encrasicolus 28</i>	Napoli	9.15	
<i>Engraulis encrasicolus 29</i>	Sardegna Occ	22.09.15	
<i>Engraulis encrasicolus 30</i>	Sardegna Occ	22.09.15	
<i>Engraulis encrasicolus 31</i>	Sardegna Occ	22.09.15	
<i>Sardina pilchardus 14</i>	Sardegna Occ	22.09.15	
<i>Sardina pilchardus 15</i>	Sardegna Occ	22.09.15	
<i>Sardina pilchardus 16</i>	Sardegna Occ	22.09.15	

5.3 Analisi del DNA

L'analisi del DNA ha previsto una prima fase di estrazione nella quale sono state seguite differenti procedure a seconda del tipo di campione in esame. Per le fasi di sequenziamento, del frammento COX1 sono stati utilizzati il metodo Sanger, previa amplificazione PCR del frammento investigato nella prima parte della ricerca, ed una analisi NGS nella seconda. Infine, le analisi bioinformatiche delle relative sequenze è stata effettuata diversamente in base alle due fasi di ricerca precedentemente elencate.

5.3.1 Estrazione del DNA

Il DNA è stato estratto utilizzando il kit “E.Z.N.A. Insect DNA kit” (OMEGA), in quanto considerato quello con il rapporto migliore resa/prezzo (Fig.15).

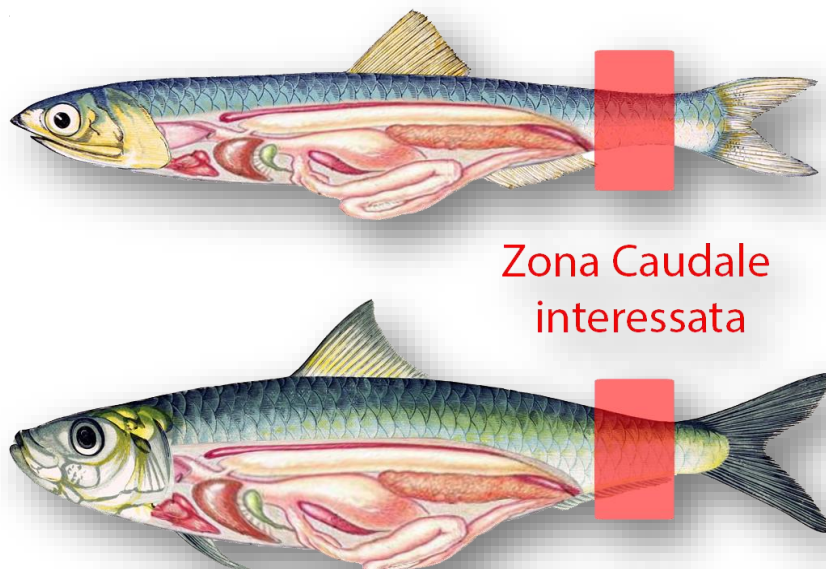


Fig.15 - Particolare del Kit di estrazione utilizzato

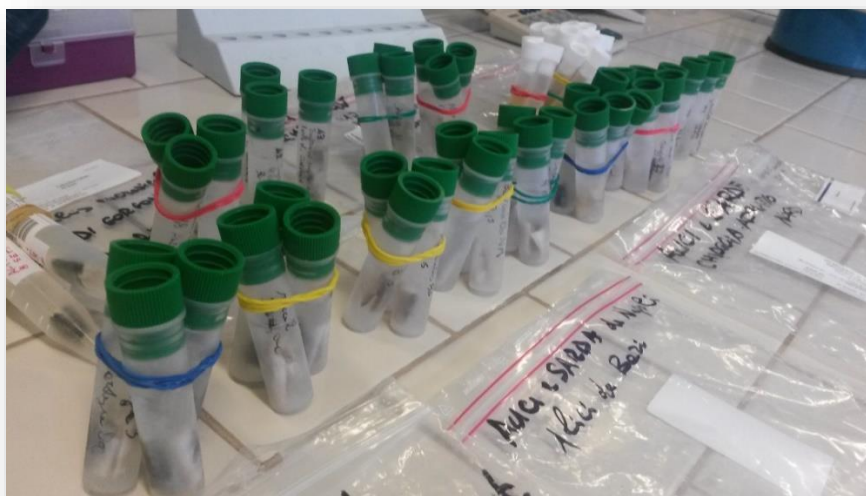
Seguendo il protocollo del kit sopra citato, i campioni di macroinvertebrati (precedentemente riconosciuti, etichettati e fissati in alcol non denaturato al 95%) sono stati sottoposti alla procedura di estrazione utilizzando zampe, cerci o l'intero corpo dell'individuo.

La preparazione dei campioni per l'estrazione del DNA dal materiale marino è stata invece effettuata asportando da ciascun esemplare una parte della zona caudale dell'individuo

(Fig.16a). Quest'ultima è stata scelta al fine di escludere l'apparato digerente ed evitare eventuali inquinanti apportati da residui di digestione. Successivamente ogni porzione è stata etichettata e fissata in alcol non denaturato al 95%.



a



b

Fig.16 - (a) Particolare della zona caudale utilizzata, (b) Conservazione del tessuto prelevato in alcol non denaturato al 95%

Per quanto riguarda i campioni di bianchetto, è stata selezionata un porzione del prodotto casuale, i campioni separati su piastre petri sono stati essiccati sotto cappa per far evaporare l'alcool e successivamente liofilizzati con azoto liquido. Successivamente il liofilizzato ottenuto è stata sottoposta alla fase di estrazione del DNA con il medesimo kit di estrazione (Fig.17 a, b)



a



b

Fig.17 - (a) Fase di essiccamento dei campioni su piastre petri, (b) Risultato della liofilizzazione con azoto liquido.

Nella seconda fase di indagine (*NGS*), le procedure di estrazione hanno previsto l'utilizzo di differenti kit, ognuno specifico per il tipo di campione da analizzare.

Per quanto riguarda l'estrazione di DNA dalla massa di macroinvertebrati, il kit utilizzato è stato il medesimo della precedente fase di ricerca. Per ogni provetta, contenente la massa di macroinvertebrati bentonici appartenente ad un determinato sito, è stata messa in pratica la

stessa procedura. Inizialmente l'agglomerato è stato separato dall'alcol etilico facendolo asciugare su carta bibula sotto cappa, successivamente pesato con una bilancia di precisione è stato posizionato in un mortaio e immerso in azoto liquido per farlo deidratare completamente. La massa secca ottenuta è stata liofilizzata all'interno del mortaio, ulteriormente pesata e posizionata in eppendorf da 1,5 ml per la fase di estrazione.

Per i campioni di sedimento è stato utilizzato il "PowerSoil® DNA Isolation Kit" seguendo il relativo protocollo di estrazione.

L'estrazione del DNA dai filtri ha previsto l'utilizzo del kit "E.Z.N.A. Insect DNA kit" (OMEGA) e del kit "PowerWater® DNA Isolation Kit" (MO BIO) specifico per l'estrazione da filtro per campioni di acqua. Quest'ultimo è stato utilizzato con il filtro da 0.45 micron, per i filtri da 1,6 micron e 2,7 micron è stato utilizzato l'altro kit. Nel primo caso la procedura svolta è stata quella indicata dal protocollo del kit, nei secondi casi invece i filtri sono stati prima sminuzzati poi posizionati in 2 eppendorf ed infine è iniziata la procedura standard prevista dal protocollo.

5.3.2 Amplificazione del frammento COX1 (Citocromo ossidasi subunità 1)

L'amplificazione della regione COX1, per i campioni di acqua dolce, è stata ottenuta mediante PCR con le seguenti coppie di primer:

LC01490: 5'-GGTCAACAAATCATAAAGATATTGG-3'

HC02198: 5'-TAAACTTCAGGGTGACCAAAAATCA-3' (Folmer et al., 1994)

LepF1: 5'-ATTCAACCAATCATAAAGATATTGG-3'

LepR1: 5'-TAAACTTCTGGATGTCCAAAAATCA-3' (Hajibabaei et al., 2006).

Il frammento COX1 dei campioni marini è stato invece amplificato valutando due coppie di primer:

Fish F1: 5'-TCAACCAACCACAAAGACATTGGCAC-3'

Fish R1: 5'-TAGACTTCTGGGTGGCCAAAGAATCA-3'

Fish F2: 5'-TCGACTAATCATAAAGATATCGGCAC-3'

Fish R2: 5'-ACTTCAGGGTGACCGAAGAATCAGAA-3' (Ward et al., 2005)

Il termocicizzatore utilizzato è stato il MJ Mini™ Personal Thermo Cycler (Bio-Rad). Il prodotto di PCR è stato ottenuto tramite RTG PCR beads (GE Healthcare) con un regime termico che ha previsto vari passaggi illustrati in Tab.4.

Tab.4 - Regime termico PCR

Fasi	Temperatura e tempo
Step 1	94° x 3'
Step 2	94° x 30''
Step 3	53°/55° x 30''
Step 4	72° x 1'0 30''
Step 5	Torna allo step 2 x 35 volte
Step 6	72° x 10'
Step 7	4° forever

L'amplificazione delle bande è stata verificata attraverso una corsa su gel di agarosio ed un successivo controllo su transilluminatore (Fig.18).

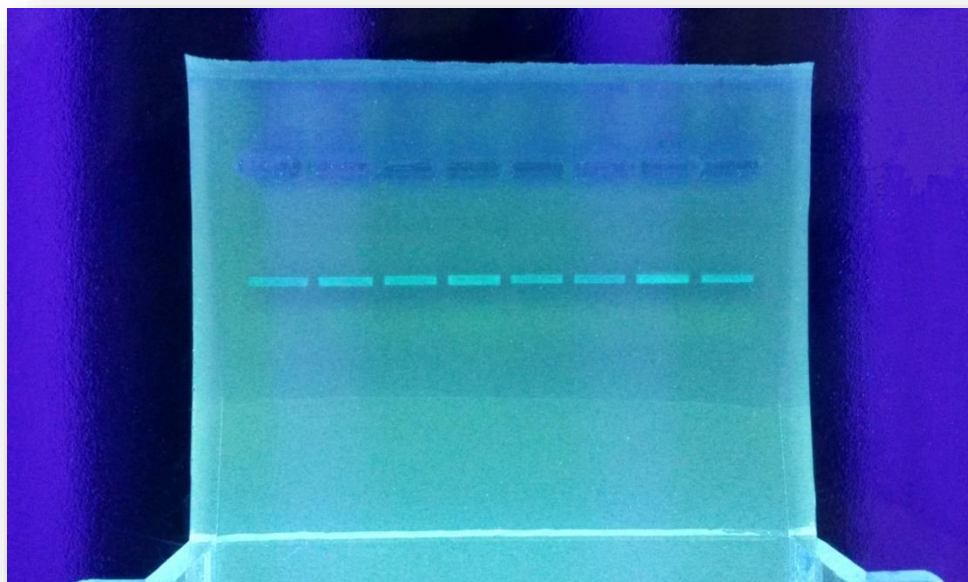


Fig.18 - Corsa su gel di agarosio

Le bande amplificate sono state successivamente purificate con Illustra GFX/PCR DNA Purification Kit (GE Healthcare), quindi delle aliquote sono state spedite alla ditta Macrogen (<http://www.macrogen.com>) per il sequenziamento.

Infine i cromatogrammi risultanti dal sequenziamento sono stati controllati visivamente ed editati con CHROMAS 2.3 (<http://www.technelysium.com.au>).

5.3.3 Amplificazione del frammento COX1 per l'analisi NGS

Come accennato precedentemente, la tecnica NGS consente un sequenziamento limitato a circa 250/300 bp (rispetto alle circa 1000 bp possibili con il metodo Sanger). Pertanto, nel

presente lavoro, l'amplificazione del frammento COX1 (circa 690 pb) sarà limitata entro queste dimensioni.

L'obiettivo per la fase di studio relativo alle acque marine è stato quello di ottenere il maggior numero di informazioni possibili dal frammento, ovvero il maggior numero di paia di basi sequenziate che consentissero la discriminazione sicura delle due specie in studio. A questo scopo è stato sviluppato un primer degenerato interno (*Forward*) in una zona abbastanza conservata del multiallineamento (Fig.19) che, in combinazione con il primer FISH_R2 (*Reverse*) consentisse: (a) l'amplificazione di una porzione di circa 450 paia di basi totali del frammento COX1 per entrambe le specie, (b) *l'overlapping* parziale delle *reads Forward* con quelle *Reverse* in modo da ottenere il sequenziamento di circa 450 paia di basi totali (sommando le 250/300 bp previste con il sequenziamento ottenibile con ciascuno dei due primer) (Fig.20). La sequenza identificata per tale primer interno è la seguente:
GAYATGGCATTCCCMCGAATGAAYAA

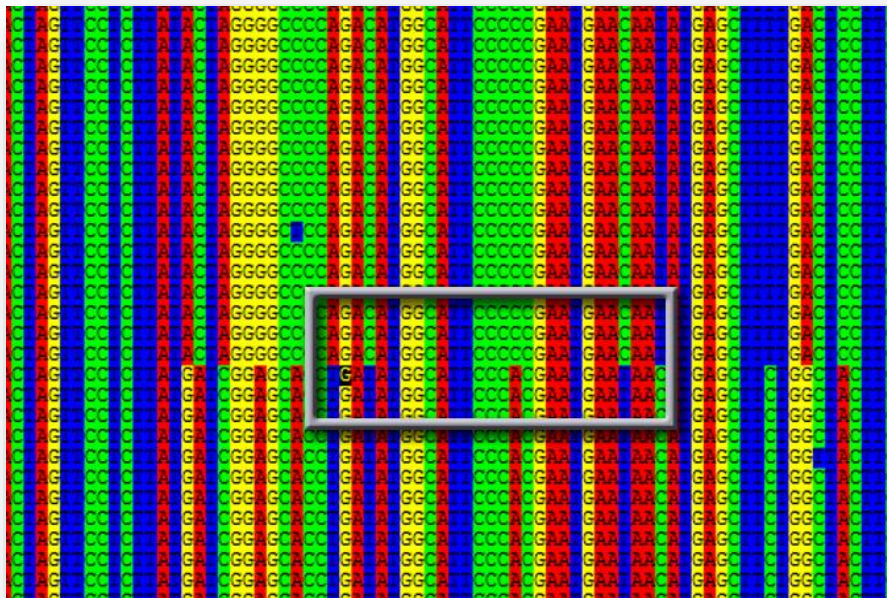


Fig.19 - Zona del multiallineamento designata per l'individuazione del primer *Forward* interno.

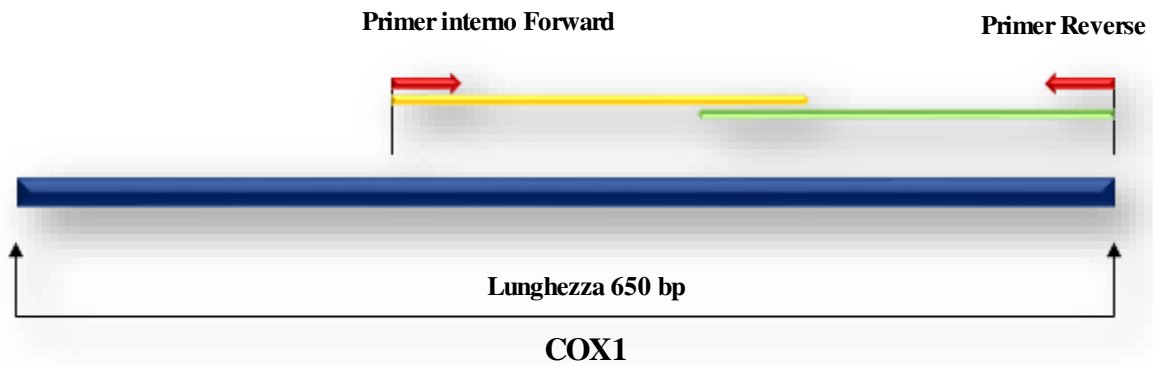


Fig.20 - Schema del gene COX1, primer interno e sovrapposizione delle letture *Forward* e *Reverse*

Nel caso dei macroinvertebrati (Efemerotteri), l'elevata variabilità del multiallineamento non ha permesso l'individuazione di una zona conservata per tutte le specie studiate, non consentendo quindi l'identificazione di primer interni che permettessero l'amplificazione parziale del COX1 per tutti i taxa. In questo caso l'analisi *NGS* è stata effettuata con dei primer degenerati basati sulla fusione delle due coppie utilizzate nella parte sperimentale del DNA *barcoding* (LCO/HCO e LepF1/LepR1). Le sequenze di questi nuovi primer sono le seguenti:

Primer_degenerato_Fw: RKTCAACMAATCATAAAGATATTGG

Primer_degenerato_Rw: TAAACTTCWGGRTGWCCAAAAAATCA

I primer, sequenziando circa 250 bp dalla posizione 5' e 250 bp dalla posizione 3', non consentiranno una sovrapposizione delle due *reads* (*Forward* e *Reverse*). La *library* dei risultati così generata sarà pertanto formata da doppie *reads* per ciascun filamento di DNA non direttamente sovrapponibili (Fig.21); questo si potrebbe tradurre in un minor numero di informazioni da poter utilizzare nella fase di identificazione dei campioni, anche se il numero di mutazioni riscontrato in ciascuna regione e la combinazione dei risultati ottenuti con ciascun primer dovrebbero essere sufficienti a discriminare ogni specie presente nell'esperimento.

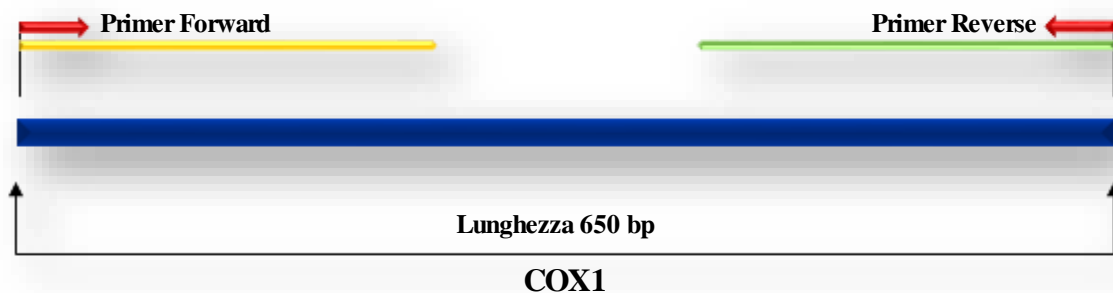


Fig.21 - Schema del gene COX1 e la mancata sovrapposizione delle letture *Forward* e *Reverse*

5.3.4 Preparazione dei campioni per il sequenziamento *NGS*

Il sequenziamento *NGS* prevede la possibilità di analisi simultanea di numerose “*library*” corrispondenti a differenti campioni.

I campioni da noi testati sono stati suddivisi in base al numero di *library* a disposizione per la corsa come illustrato nella Tab.5.

Tab.5 - Descrizione delle *library* acqua dolce (Efemerotteri) per la corsa *NGS*

Library	Codice	Campione	Descrizione	Concentrazione del DNA
1	33343	Efemerotteri_1	DNA estratto da 91 individui mescolati con lo stessa concentrazione di DNA	(~30 ng/μl)
2	33344	Efemerotteri_2	DNA estratto da 91 individui mescolati con lo stesso volume (1μl ciascuno)	(~30 ng/μl)
3	33345	Bulk_1	DNA estratto da Aggregato di Macroinvertebrati (135 individui)	(>250 ng/μl)
4	33346	Bulk_2	DNA estratto da Aggregato di Macroinvertebrati (138 individui)	(>500 ng/μl)
5	33347	Bulk_3	DNA estratto da Aggregato di Macroinvertebrati (182 individui)	(>500 ng/μl)
6	33348	eDNA Sedimento	DNA estratto da sedimento	(~5 ng/μl)
7	33349	eDNA Acqua	DNA estratto da filtrato di acqua doce	(~7 ng/μl)

Per la prima *library* è stata inserita una mix dei campioni risultanti dalla precedente fase di ricerca. In quest’ultima, i campioni opportunamente quantificati con saggi fluorimetrici attraverso il Multiskan™ GO Microplate Spectrophotometer (Thermo Scientific), sono stati immessi nella miscela in proporzione alle relative concentrazioni (ng/μl) in modo che ciascuno di essi fosse ugualmente rappresentato. Scopo della prova è stato proporzionare i DNA all’interno del pozzetto e verificare se il relativo output di sequenziamento attraverso tecnica *NGS* riconducesse ai campioni inseriti.

Allo stesso modo nella seconda sono stati inseriti i medesimi DNA estratti ma in concentrazioni differenti al precedente, è stato infatti inserito il medesimo voume (1μl) prelevato direttamente dallo stock madre.

Le successive tre *library* sono state utilizzate immettendo il DNA ricavato dall’estrazione della massa di macroinvertebrati sopra citata per ogni stazione di campionamento. Scopo di questa prova è stato testare se l’identificazione delle sequenze COX1 ottenute dal sequenziamento, per ogni sito di campionamento, riconducessero le specie riconosciute morfologicamente.

Le due *library* seguenti sono state utilizzate con il DNA ricavato rispettivamente dall’estrazione del filtrato di acqua dolce e dal sedimento, con lo scopo di testare l’eventuale presenza di DNA extracellulare che riconducesse alla presenza degli organismi campionati e analizzati mediante *NGS*, oltre alla eventuale presenza di ulteriori organismi non campionati.

Le analisi riguardanti gli organismi marini hanno previsto l'utilizzo di altre 7 *library* contenenti, in differenti modalità e concentrazioni, i DNA delle stesse due specie analizzate (*Engraulis encrasicolus* e *Sardina pilchardus*) come illustrato in Tab.6.

Tab.6 - Descrizione delle *library* “acqua marina” (*Engraulis/Sardina*/bianchetto) per la corsa NGS

Library	Codice	Campione	Descrizione	Concentrazione del DNA
8	33350	<i>Engraulis</i>	DNA estratto da 1 individuo di <i>Engraulis encrasicolus</i> (20 mg di tessuto)	(~70 ng/μl)
9	33351	<i>Sardina</i>	DNA estratto da 1 individuo di <i>Sardina pilchardus</i> (20 mg di tessuto)	(~100 ng/μl)
10	33352	50%-50%	DNA estratto da 50% di <i>Engraulis encrasicolus</i> - 50% di <i>Sardina pilchardus</i> (stessa concentrazione in DNA estratto)	(~30 ng/μl)
11	33353	5%-95%	DNA estratto da 5% di <i>Engraulis encrasicolus</i> - 95% di <i>Sardina pilchardus</i> (concentrazione in DNA estratto)	(~30 ng/μl)
12	33354	eDNA acqua	DNA estratto da filtrato di acqua marina	(~3 ng/μl)
13	33355	2.5%-97.5%	DNA estratto da 2.5% <i>Engraulis encrasicolus</i> - 97.5% <i>Sardina pilchardus</i> (proporzione in quantità di tessuto estratto)	(~100 ng/μl)
14	33356	Bianchetto	DNA estratto da 2.5 g di giovanili di <i>Engraulis encrasicolus</i> e <i>Sardina pilchardus</i>	(~80 ng/μl)

Nelle prime due è stato inserito DNA estratto rispettivamente da un campione di *Engraulis encrasicolus* e da uno di *Sardina pilchardus*. In questo caso è stato pesato 20 mg di tessuto (ricavato sempre dalla zona caudale) di ciascun individuo, il DNA estratto è stato quantificato con il medesimo strumento sopra descritto e inserito nel pozzetto proporzionalmente alla concentrazione risultante in modo da avere entrambi gli esemplari ugualmente rappresentati (ca. 30 ng/μl).

Nelle successive due sono stati inseriti gli stessi DNA dei sopra descritti (*Engraulis encrasicolus* e *Sardina pilchardus*) ma combinati, sempre in base alla concentrazione precedente, ottenendo rispettivamente il 50% e il 50% per un pozzetto e rispettivamente il 5% e il 95% per l'altro pozzetto.

Le ultime tre *library* sono state utilizzate rispettivamente con: a) DNA estratto dal filtrato di acqua marina, b) DNA estratto unendo 1mg di tessuto proveniente da alicce con 49 mg di tessuto ottenendo una combinazione in percentuale pari a 2.5% e 97,5% rispettivamente, c) DNA estratto da 2.5g di bianchetto.

5.4 Analisi bioinformatiche

Nella fase di indagine del DNA *barcoding* degli Efemerotteri, le sequenze ottenute sono state multi-allineate con MAFFT 6FrameTranslation (<http://toolkit.tuebingen.mpg.de>). Successivamente, per verificare l'assenza di codoni di stop nelle sequenze stesse, è stata effettuata una traduzione amminoacidica con il programma DnaSP 5.10 DNA Sequence Polymorphism (Rozas et al. 2010). Diversi modelli analitici sono stati utilizzati successivamente per valutare la capacità di discriminazione dei dati di sequenze *barcode*. Mothur (Schloss et al., 2009) è stato utilizzato per creare dei cluster di sequenze (aplotipi) con una soglia del 100% e del 98% di distanza genetica. I cluster di aplotipi formati in accordo con la classificazione tassonomica classica, contenenti un solo campione o più campioni intraspecifici, sono stati considerati correttamente identificati, al contrario i cluster contenenti aplotipi di specie differenti sono stati considerati errati.

Successivamente è stata verificata la presenza del “*barcoding gap*” per ogni specie analizzata. Questa è un'analisi chiave per valutare la discriminazione delle specie (Meyer et al., 2008) e presuppone un intervallo tra la quantità di divergenza genetica delle sequenze all'interno delle specie e la quantità di divergenza genetica tra le sequenze di specie diverse.

Kimura 2P è stato il modello utilizzato per il calcolo delle distanze genetiche tra specie congeneriche e tra le specie stesse, tutte quelle che presentavano la massima distanza genetica intraspecifica minore della minima distanza genetica interspecifica, venivano considerate discriminate con successo.

Spider (Brown et al., 2012), un pacchetto di R specifico per analisi sulle identità delle specie e l'evoluzione, con particolare riferimento al DNA *barcoding*, è stato utilizzato per calcolare alcuni parametri come: a) una matrice delle distanze genetiche con il modello Kimura 2P attraverso la stringa di comando “dis.dna function”, b) la presenza del *barcoding gap* attraverso la stringa di comando “maxInDist and nonConDist function”, c) l'ottimizzazione di una soglia di distanza genetica per la discriminazione delle specie attraverso la valutazione statistica dei falsi positivi (esemplari conspecifici identificati invece come specie nuove) e dei falsi negativi (esemplari appartenenti a specie differenti identificati invece come la stessa specie) attraverso la stringa di comando “threshOpt function and threshVal function”.

La delimitazione delle specie (il processo di raggruppamento dei cluster aplotipici all'interno di distinti gruppi tassonomici), è stata ulteriormente valutata secondo il criterio della monofilia reciproca della specie (Rosenberg, 2007), producendo un cladogramma con l'algoritmo Neighbor Joining; quest'ultima valutazione si basa sul modello di ramificazione osservato nel

cladogramma confrontandolo con un modello di ramificazione casuale, e i valori dei nodi con p-value <0.05 sono considerati significativamente monofiletici (nj, read.tree, monophyly, Rosenberg's function).

Un altro approccio utilizzato per testare la monofilia delle specie ha preso in considerazione i valori di supporto *bootstrap* in un albero di Maximum Likelihood (con 1000 repliche) generato con il modello GTRCAT dal programma RAxML (Stamatakis, 2006), dove tutti i cladi generati con nodi aventi un valore di bootstrap >70% sono stati considerati monofiletici. Inoltre anche il modello PTP (*Poisson Tree Processes*) (<http://sco.h-its.org/exelixis/web/software/PTP/>) è stato eseguito per fornire ulteriori prove del potere risolutivo dei dati ottenuti. In quest'ultimo le differenze delle relazioni tra specie e all'interno della stessa specie sono modellate in termini di numero di sostituzioni come in due classi indipendenti del processo di Poisson, queste utilizzate poi per calcolare una serie di entità che rappresentano le specie teoriche (Zhang et al., 2013).

Per simulare un pratico scenario di simulazione di identificazione dei *barcode*, le sequenze ottenute sono state confrontate con quelle presenti su due banche dati genetiche: N.C.B.I. (<http://www.ncbi.nlm.nih.gov>), il database globale che mette a disposizione oltre 322.000 sequenze nucleotidiche pubbliche di specie formalmente descritte (Dicembre 2014), e BOLD Systems (<http://barcodinglife.org>), che costituisce il deposito ufficiale di quasi 228.000 sequenze *barcode* di tutti gli organismi viventi, per valutare la corretta identificazione delle specie. Prima della valutazione della discriminazione, entrambi i database sono stati sottoposti ad un controllo preliminare per verificare l'effettiva presenza delle sequenze COX1 relative al nostro dataset (a livello di specie). L'assegnazione tassonomica di una sequenza è stata considerata positiva, quindi discriminata correttamente, se la percentuale di identità più alta ottenuta nella banca dati corrispondeva al nome attribuitogli dopo il riconoscimento morfologico. Al contrario quando una sequenza risultava avere un basso valore di identità oppure la corrispondenza risultava differente dalla specie morfologicamente riconosciuta, l'esito veniva valutato negativo.

Infine, per valutare il posizionamento delle nostre sequenze, è stato generato un dendrogramma con RAxML secondo il modello evolutivo GTRCAT integrando il nostro dataset con i COX1 *barcode* degli Efemerotteri attualmente disponibili nel database ufficiale. E' stato quindi selezionato da BOLD Systems un sottoinsieme di aplotipi unici di lunghezza superiore a 500 bp in modo da rappresentare tutta l'ampiezza tassonomica disponibile all'interno delle famiglie e dei generi di Efemerotteri europei del nostro studio.

5.4.1 Analisi bioinformatiche NGS

Le sequenze in formato grezzo (fastq) ottenute dall'output della piattaforma Illumina sono state analizzate con il programma Mothur (REF).

Lo schema di lavoro seguito per il processamento delle sequenze fastq è illustrato in Fig.22.

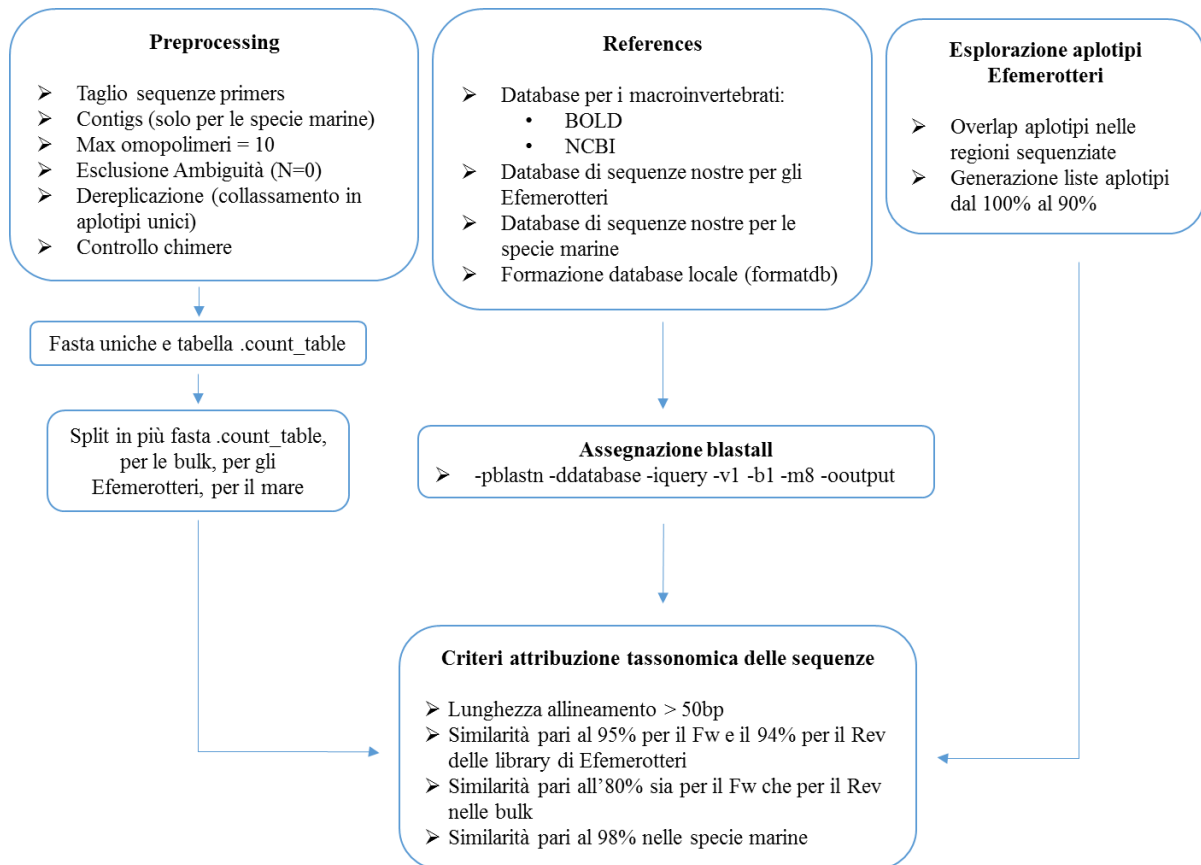


Fig.22 - Schema di lavoro seguito per le analisi bioinformatiche delle sequenze ottenute dalla corsa NGS

La fase di *preprocessing* inizia con il taglio delle sequenze dei *primer*; quindi, per le sequenze provenienti dalle *library* marine, sono stati formati i *contigs*, ovvero l'unione di ciascuna sequenza forward con la sua rispettiva reverse. Per tutte le *library*, inoltre, sono state eliminate le sequenze con omopolimeri maggiori di 10 (ovvero sequenze ripetute di A o C o T o G) ed escluse le sequenze contenenti ambiguità (N=0). Infine le sequenze uguali sono state clusterizzate in aplotipi unici (dereplicate) e controllate per verificare la presenza di chimere (ovvero false sequenze che si generano durante il processo di sequenziamento, nelle quali sono presenti pezzi di sequenze appartenenti a individui diversi unite insieme) che generano falsi livelli di diversità.

Il database di riferimento locale è stato generato a seconda delle *library* da confrontare; per le bulk, il campione di acqua e di sedimento sono state scaricate dalle banche dati mondiali (NCBI e BOLD) le sequenze *reference* dei taxa di macroinvertebrati riconosciuti morfologicamente (a livello di famiglia o genere). Per le *library* di Efemerotteri e delle specie marine (*Engraulis encrasicous* e *Sardina pilchardus*) sono stati utilizzati i nostri database di riferimento, prodotti nella prima fase di questo lavoro.

Al fine di verificare quale fosse la soglia di similarità che ci consentisse di separare specie differenti, sono state calcolate le soglie d'identità genetica rispetto alle lunghezze delle sequenze *forward* e *reverse* ottenute (Fig.23). I 91 *barcode* reference sono stati allineati con un frammento forward e reverse per identificare la regione (206 bp) su cui calcolare le soglie (variabili dal 100% al 90%).

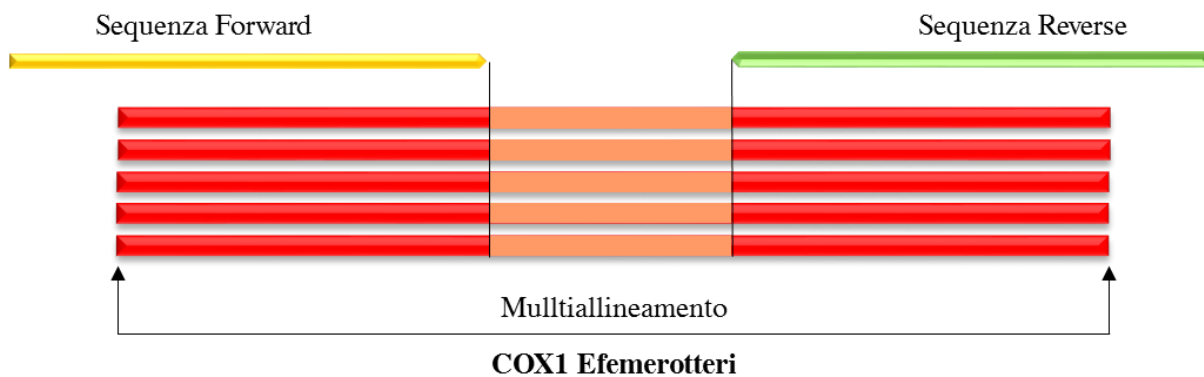


Fig.23 - Allineamento di una sequenza ambientale (forward e reverse) con il nostro multi-allineamento di efemerotteri per evidenziare la regione di interesse (206 bp)

Parallelamente, il dataset di riferimento è stato formattato (formatdb) per la successiva fase di interrogazione (assegnazione tassonomica, blast locale).

Al termine della fase di *preprocessing* viene generato un file delle sequenze uniche (.fasta) e una tabella (.count_table) dove è riportata l'abbondanza totale e la distribuzione di ciascuna sequenza unica nelle varie *library*.

La fase di assegnazione delle sequenze con i rispettivi database di riferimento è stata effettuata con la stringa di comando riportata in Fig.22. Il file generato è stato unito al file ".count_table" e ad essi sono stati applicati i criteri di accettazione dell'attribuzione tassonomica (Fig.22).

6. RISULTATI

6.1 DNA *barcoding* degli Efemerotteri

6.1.1 Dataset morfologico e molecolare degli Efemerotteri

Il piano di campionamento ha incluso 34 corsi d'acqua e fiumi guadabili presenti nell'area studiata. La raccolta effettuata ha consentito di ottenere una collezione di dati comprensiva di 39 specie morfologicamente identificate, raggruppate in 18 generi e 8 famiglie, che rappresentano il 12% delle specie Europee ed il 37% delle specie italiane (Thomas & Belfiore, 2004; Belfiore, 2005). Il dataset finale (89 individui) è risultato contenere membri di 9 specie endemiche, di cui una presente in Italia peninsulare e 8 presenti nel blocco Sardo-Corso. Un totale di 17 specie sono state rappresentate da almeno 3 individui per ciascuna, campionati in differenti corsi d'acqua. Le restanti 22 specie sono rappresentate da uno o due individui in quanto non è stato possibile trovare campioni aggiuntivi in corsi d'acqua differenti.

Il kit per l'estrazione di DNA utilizzato ha mostrato una buona resa per tutti i campioni, in gran parte dipendente dalla grandezza dell'individuo analizzato. Come dimostrato dalla successiva quantificazione del DNA estratto, è stato identificato un intervallo di concentrazione compreso tra 1.43 ng/μl e 1916.02 ng/μl. I limiti del range sono rappresentati da due individui (*Caenis martae* per il limite inferiore e *Ecdyonurus venosus* per il limite superiore)

Sono state ottenute sequenze COX1 per il 100% dei campioni analizzati, con regimi termici di PCR per la temperatura di annealing compresi in un range tra 50° e 53° C.

Le sequenze ottenute dal sequenziamento del prodotto di PCR sono risultate prive di ambiguità in tutti i casi. La standardizzazione del protocollo nella fase di spedizione delle aliquote per il sequenziamento è stata testata inviando sia i primer *Forward* che *Reverse*. Questi ultimi hanno mostrato una migliore resa nel sequenziamento e l'elettroferogramma ottenuto non ha mostrato doppi picchi o rumori di fondo permettendo una lettura chiara e precisa della sequenza. Il multiallineamento generato non ha evidenziato gap per l'intero dataset e la traduzione proteica effettuata non ha mostrato codoni di stop.

Il risultato finale è stato un allineamento di 89 sequenze per un totale di 614 paia di basi.

6.1.2 Discriminazione delle specie

Una prima verifica nella discriminazione delle specie è stata effettuata generando una lista degli aplotipi con una soglia d'identità pari al 100%. Il risultato ottenuto ha mostrato un raggruppamento delle sequenze in 77 aplotipi, di cui 70 formati da elementi singoli e 7 condivisi con campioni della stessa specie (*Baetis cyrneus*, *B. fuscatus*, *B. buceratus*, *Centroptilum luteolum*, *Caenis beskidensis*, *Electrogena grandiae*, *Rhithrogena insularis*).

Una riduzione della soglia (*cut-off*) del 2% (identità del 98%) mostra invece una diminuzione del numero di aplotipi generati, evidenziando che, dei 48 cluster ottenuti, 22 sono composti da elementi singoli e 26 condivisi da individui della stessa specie. Inoltre, delle 17 specie comprendenti 3 o più individui analizzati, 11 (circa il 65%) mostrano aplotipi multipli (Tab.7 a,b).

Tab.7a - Lista degli aplotipi (sequenze con il 100% di identità)

Hap 1: 5	<i>Baetis buceratus</i> MO3, <i>Baetis buceratus</i> RIO, <i>Baetis buceratus</i> SO, <i>Baetis buceratus</i> MO2, <i>Baetis buceratus</i> MO4	Hap 41: 1	<i>Epeorus assimilis</i> OB
Hap 2: 3	<i>Baetis cyrneus</i> PVG4, <i>Baetis cyrneus</i> PGI, <i>Baetis cyrneus</i> MO	Hap 42: 1	<i>Epeorus assimilis</i> AN
Hap 3: 2	<i>Baetis fuscatus</i> SA, <i>Baetis fuscatus</i> SD	Hap 43: 1	<i>Epeorus assimilis</i> OR
Hap 4: 1	<i>Baetis muticus</i> VI	Hap 44: 2	<i>Electrogena grandiae</i> IN, <i>Electrogena grandiae</i> RA
Hap 5: 1	<i>Baetis vernus</i> TR	Hap 45: 1	<i>Electrogena zebrata</i> SA1
Hap 6: 1	<i>Baetis albinatii</i> SA	Hap 46: 1	<i>Electrogena zebrata</i> SA2
Hap 7: 1	<i>Baetis cyrneus</i> SA	Hap 47: 2	<i>Rhithrogena insularis</i> COR1, <i>Rhithrogena insularis</i> COR2
Hap 8: 1	<i>Baetis lutheri</i> MP	Hap 48: 1	<i>Rhithrogena semicolorata</i> PM
Hap 9: 1	<i>Baetis lutheri</i> AT	Hap 49: 1	<i>Rhithrogena semicolorata</i> FT
Hap 10: 1	<i>Baetis alpinus</i> FG	Hap 50: 1	<i>Rhithrogena semicolorata</i> RS
Hap 11: 1	<i>Baetis alpinus</i> AN	Hap 51: 1	<i>Rhithrogena reatina</i> VMA
Hap 12: 1	<i>Baetis alpinus</i> OB	Hap 52: 1	<i>Rhithrogena reatina</i> CA
Hap 13: 2	<i>Centroptilum luteolum</i> FT, <i>Centroptilum luteolum</i> FG	Hap 53: 1	<i>Rhithrogena nuragica</i> SCC
Hap 14: 1	<i>Centroptilum luteolum</i> RA	Hap 54: 1	<i>Rhithrogena eatoni</i> COR
Hap 15: 1	<i>Procloeon pulchrum</i> VI	Hap 55: 1	<i>Heptagenia longicauda</i> RIG
Hap 16: 1	<i>Procloeon pulchrum</i> OR	Hap 56: 1	<i>Heptagenia longicauda</i> AT
Hap 17: 1	<i>Procloeon pulchrum</i> RC	Hap 57: 1	<i>Heptagenia longicauda</i> RC
Hap 18: 1	<i>Procloeon bifidum</i> MR	Hap 58: 1	<i>Habroleptoides confusa</i> FG
Hap 19: 1	<i>Procloeon bifidum</i> ESV	Hap 59: 1	<i>Habroleptoides confusa</i> RA
Hap 20: 1	<i>Procloeon bifidum</i> SC	Hap 60: 1	<i>Habroleptoides confusa</i> OR
Hap 21: 1	<i>Cloeon dipterum</i> MV	Hap 61: 1	<i>Habroleptoides</i> sp. SB
Hap 22: 1	<i>Cloeon dipterum</i> AT	Hap 62: 1	<i>Habrophlebia eldae</i> VI
Hap 23: 1	<i>Cloeon simile</i> SB	Hap 63: 1	<i>Habrophlebia eldae</i> FA
Hap 24: 3	<i>Caenis beskidensis</i> VI, <i>Caenis beskidensis</i> RC, <i>Caenis beskidensis</i> AT	Hap 64: 1	<i>Habrophlebia eldae</i> RA
Hap 25: 1	<i>Caenis pusilla</i> AT	Hap 65: 1	<i>Paraleptophlebia submarginata</i> AT
Hap 26: 1	<i>Caenis pusilla</i> MR	Hap 66: 1	<i>Paraleptophlebia submarginata</i> VII
Hap 27: 1	<i>Caenis macrura</i> gr MO	Hap 67: 1	<i>Paraleptophlebia submarginata</i> VI2
Hap 28: 1	<i>Caenis macrura</i> gr OR	Hap 68: 1	<i>Choroterpes picteti</i> MR
Hap 29: 1	<i>Caenis macrura</i> gr RIO	Hap 69: 1	<i>Serratella ignita</i> VI
Hap 30: 1	<i>Caenis martaе</i> SA	Hap 70: 1	<i>Serratella ignita</i> RA
Hap 31: 1	<i>Brachycercus harrisella</i> VI	Hap 71: 1	<i>Serratella ignita</i> FA
Hap 32: 1	<i>Brachycercus harrisella</i> RC	Hap 72: 1	<i>Serratella ignita</i> OR
Hap 33: 1	<i>Ecdyonurus helveticus</i> FT	Hap 73: 1	<i>Oligoneuriella rhenana</i> RC
Hap 34: 1	<i>Ecdyonurus helveticus</i> OR	Hap 74: 1	<i>Oligoneuriella rhenana</i> MP
Hap 35: 1	<i>Ecdyonurus helveticus</i> MC	Hap 75: 1	<i>Siphonurus lacustris</i> VMA
Hap 36: 1	<i>Ecdyonurus corsicus</i> SB	Hap 76: 1	<i>Siphonurus lacustris</i> VI
Hap 37: 1	<i>Ecdyonurus venosus</i> ORI	Hap 77: 1	<i>Ephemera danica</i> VI
Hap 38: 1	<i>Ecdyonurus venosus</i> IN		
Hap 39: 1	<i>Ecdyonurus venosus</i> PM		
Hap 40: 1	<i>Ecdyonurus venosus</i> OR2		

Tab.7b - Lista degli aplotipi (sequenze con il 98% di identità)

Hap 1: 5	<i>Baetis buceratus</i> MO3, <i>Baetis buceratus</i> RIO, <i>Baetis buceratus</i> SO, <i>Baetis buceratus</i> MO2, <i>Baetis buceratus</i> MO4	Hap 25: 4	<i>Ecdyonurus venosus</i> IN, <i>Ecdyonurus venosus</i> OR, <i>Ecdyonurus venosus</i> PM, <i>Ecdyonurus venosus</i> OR
Hap 2: 3	<i>Baetis cyrneus</i> PVG4, <i>Baetis cyrneus</i> PVG1, <i>Baetis cyrneus</i> PVG2	Hap 26: 2	<i>Ecdyonurus helveticus</i> FT, <i>Ecdyonurus helveticus</i> OR
Hap 3: 2	<i>Baetis fuscatus</i> SA, <i>Baetis fuscatus</i> SD	Hap 27: 1	<i>Ecdyonurus helveticus</i> MC
Hap 4: 2	<i>Baetis lutheri</i> MP, <i>Baetis lutheri</i> AT	Hap 28: 1	<i>Ecdyonurus corsicus</i> SB
Hap 5: 1	<i>Baetis alpinus</i> FG, <i>Baetis alpinus</i> OB	Hap 29: 2	<i>Epeorus assimilis</i> OB, <i>Epeorus assimilis</i> AN
Hap 6: 1	<i>Baetis cyrneus</i> SA	Hap 30: 1	<i>Epeorus assimilis</i> OR
Hap 7: 1	<i>Baetis alpinus</i> AN	Hap 31: 2	<i>Electrogena zebrata</i> SA1, <i>Electrogena zebrata</i> SA2
Hap 8: 1	<i>Baetis vernus</i> TR	Hap 32: 2	<i>Electrogena grandiae</i> IN, <i>Electrogena grandiae</i> RA
Hap 9: 1	<i>Baetis muticus</i> VI	Hap 33: 3	<i>Rhithrogena semicolorata</i> PM, <i>Rhithrogena semicolorata</i> FT, <i>Rhithrogena semicolorata</i> RS
Hap 10: 1	<i>Baetis albinatii</i> SA	Hap 34: 2	<i>Rhithrogena reatina</i> VMA, <i>Rhithrogena reatina</i> CA
Hap 11: 3	<i>Centropitulum luteolum</i> RA, <i>Centropitulum luteolum</i> FT, <i>Centropitulum luteolum</i> FG	Hap 35: 2	<i>Rhithrogena insularis</i> COR1, <i>Rhithrogena insularis</i> COR2
Hap 12: 2	<i>Proclaeon pulchrum</i> VI, <i>Proclaeon pulchrum</i> RC	Hap 36: 1	<i>Rhithrogena eatoni</i> COR
Hap 13: 1	<i>Proclaeon pulchrum</i> OR	Hap 37: 1	<i>Rhithrogena nuragica</i> SC
Hap 14: 1	<i>Proclaeon bifidum</i> MR	Hap 38: 3	<i>Heptagenia longicauda</i> RIG, <i>Heptagenia longicauda</i> AT, <i>Heptagenia longicauda</i> RC
Hap 15: 1	<i>Proclaeon bifidum</i> ESV	Hap 39: 3	<i>Habroleptoides confusa</i> FG, <i>Habroleptoides confusa</i> RA, <i>Habroleptoides confusa</i> OR
Hap 16: 1	<i>Proclaeon bifidum</i> SC	Hap 40: 1	<i>Habroleptoides</i> sp. SB
Hap 17: 2	<i>Cloeon dipterum</i> MV, <i>Cloeon dipterum</i> AT	Hap 41: 3	<i>Habrophlebia eldae</i> VI, <i>Habrophlebia eldae</i> FA, <i>Habrophlebia eldae</i> RA
Hap 18: 1	<i>Cloeon simile</i> SB	Hap 42: 3	<i>Paraleptophlebia submarginata</i> AT, <i>Paraleptophlebia submarginata</i> VI, <i>Paraleptophlebia submarginata</i> VI
Hap 19: 3	<i>Caenis beskidensis</i> VI, <i>Caenis beskidensis</i> RC, <i>Caenis beskidensis</i> AT	Hap 43: 1	<i>Choroterpes picteti</i> MR
Hap 20: 2	<i>Caenis martae</i> OR, <i>Caenis martae</i> RIO	Hap 44: 4	<i>Serratella ignita</i> FA, <i>Serratella ignita</i> VI, <i>Serratella ignita</i> RA, <i>Serratella ignita</i> OR
Hap 21: 2	<i>Caenis pusilla</i> MR, <i>Caenis pusilla</i> AT	Hap 45: 2	<i>Oligoneuriella rheana</i> MP, <i>Oligoneuriella rheana</i> RC
Hap 22: 1	<i>Caenis martae</i> MO	Hap 46: 1	<i>Siphonurus lacustris</i> VI
Hap 23: 1	<i>Caenis</i> sp. SA	Hap 47: 1	<i>Siphonurus lacustris</i> VMA
Hap 24: 2	<i>Brachycercus harrisella</i> VI, <i>Brachycercus harrisella</i> RC	Hap 48: 1	<i>Ephemera danica</i> VI

Come mostrato nella tabella successiva (Tab.8), gli intervalli di distanze genetiche (K2P) rilevate hanno valori compresi tra 0.00% e 19.6%, con tre specie che registrano valori più alti del 10.0%, *Proclaeon bifidum*, *Baetis cyrneus* e *Epeorus assimilis*.

Tab.8 - Sintesi a livello di specie delle distanze genetiche K2P per le sequenze COX1 di 17 specie con più esemplari inclusi in questo studio. Il *barcoding gap* è stato calcolato come differenza tra la minima distanza interspecifica e la massima distanza intraspecifica.

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

Gli intervalli delle distanze genetiche K2P intraspecifiche, calcolate all'interno dello stesso genere, hanno valori che vanno da 8.17% (*Rhithrogena reatina/R. nuragica*) e 33.02% (*Caenis beskidensis/C. martae*).

I generi *Proclleon*, *Baetis* ed *Epeorus* hanno mostrato valori di distanze genetiche intraspecifiche più alti del complessivo valore inferiore di distanze interspecifiche registrate sull'intera collezione di dati.

La media delle distanze intraspecifiche per l'intero set di dati è compresa tra i valori 0.2% (*Electrogena*) e 8.73% (*Proclleon*), mentre la media dei valori per le distanze interspecifiche varia tra 10.33% e 33.02% (*Caenis*).

La presenza del *barcoding gap* è stata valutata analizzando sia l'intero dataset che individui congenerici. Degli 8 generi analizzati (quelli con più di una specie), 7 hanno evidenziato la presenza di *barcoding gap* con un range compreso tra il 3.5% per *Baetis* e il 23.5% per

Nell'analisi sulla ricerca di una soglia ottimale per l'intero dataset (Fig.25) si nota una sovrapposizione dei falsi positivi e falsi negativi che generano un errore cumulativo in un intervallo di soglie compreso tra il 9% e il 18% di divergenza genetica K2P. Da questo si desume che non è possibile definire un limite comune tale da consentire la discriminazione netta dei campioni. L'errore infatti non è annullato, ma minimizzato per una soglia di circa il 7% di divergenza genetica.

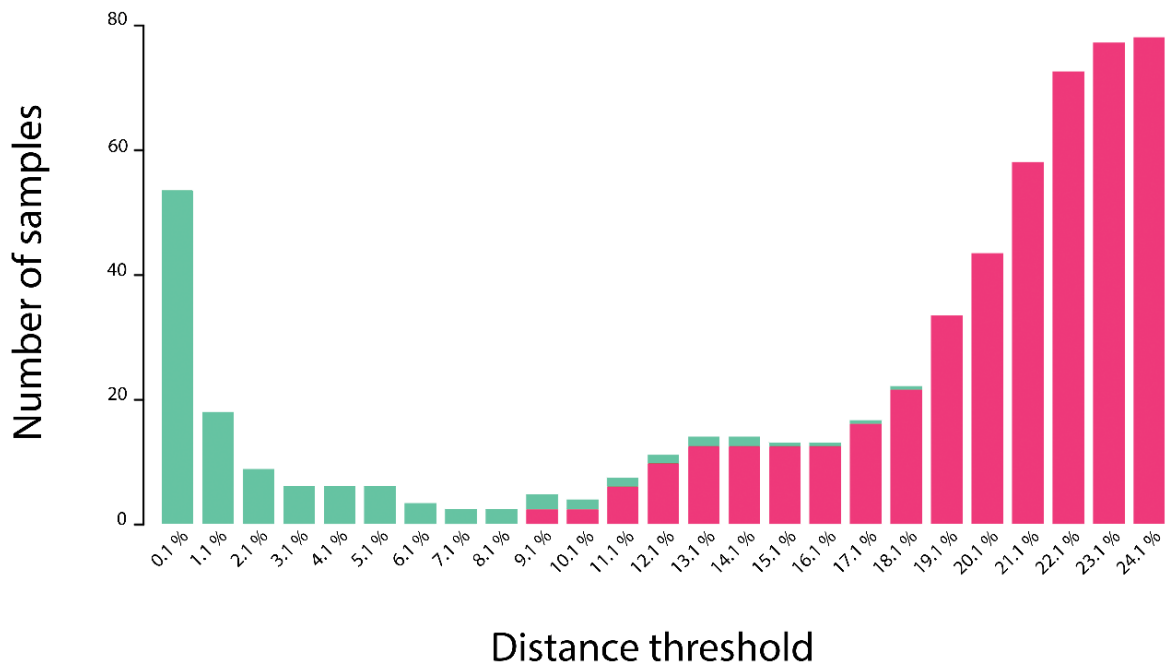


Fig.25 - Grafico a barre del tasso di falsi positivi (verde) e falsi negativi (rosso) dell'identificazione a livello di specie degli Efemerotteri basato sulle distanze genetiche K2P dell'intero set di dati.

Una successiva analisi condotta su tutti i generi escludendo le specie con elementi singoli ed i generi (*Baetis* e *Procloeon* caratterizzati da una elevata distanza genetica intraspecifica), ha prodotto una soglia ottimale dell'11% capace di discriminare nettamente in tutti i generi analizzati (Fig.26).

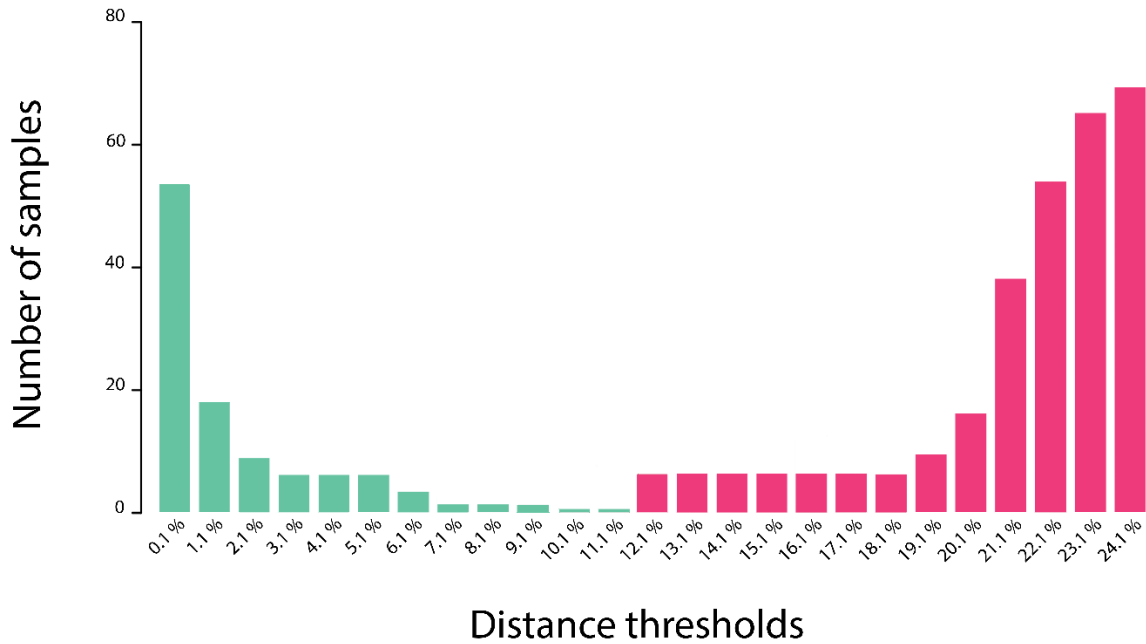


Fig.26 - Grafico a barre del tasso di falsi positivi (verde) e falsi negativi (rosso) dell'identificazione a livello di specie per tutti i generi in esame esclusi *Baetis*, *Proclleon* e le specie con elementi singoli.

Il dendrogramma NJ ottenuto con le sequenze COX1 mostrato in Fig.27 evidenzia come tutti gli individui sono raggruppati all'interno di cluster tassonomici di specie ben distinti in accordo con la tassonomia classica, con l'unica eccezione riscontrata nella specie *Proclleon bifidum*. Inoltre, degli 8 generi analizzati con più specie al loro interno, 5 appaiono largamente parafiletici (*Baetis*, *Caenis*, *Cloeon*, *Electrogena*, *Proclleon*) mentre 3 (*Eclyonurus*, *Rhithrogena* e *Habroleptoides*) formano gruppi coesi.

Il test di Rosenberg sulla probabilità della reciproca monofilia (p-value = 0.05) (Rosenberg, 2007) mostra che i nodi del dendrogramma con significatività monofiletica sono pochi e solamente i più profondi. Di questi solamente la famiglia Baetidae forma un raggruppamento coeso con un buon supporto monofiletico.

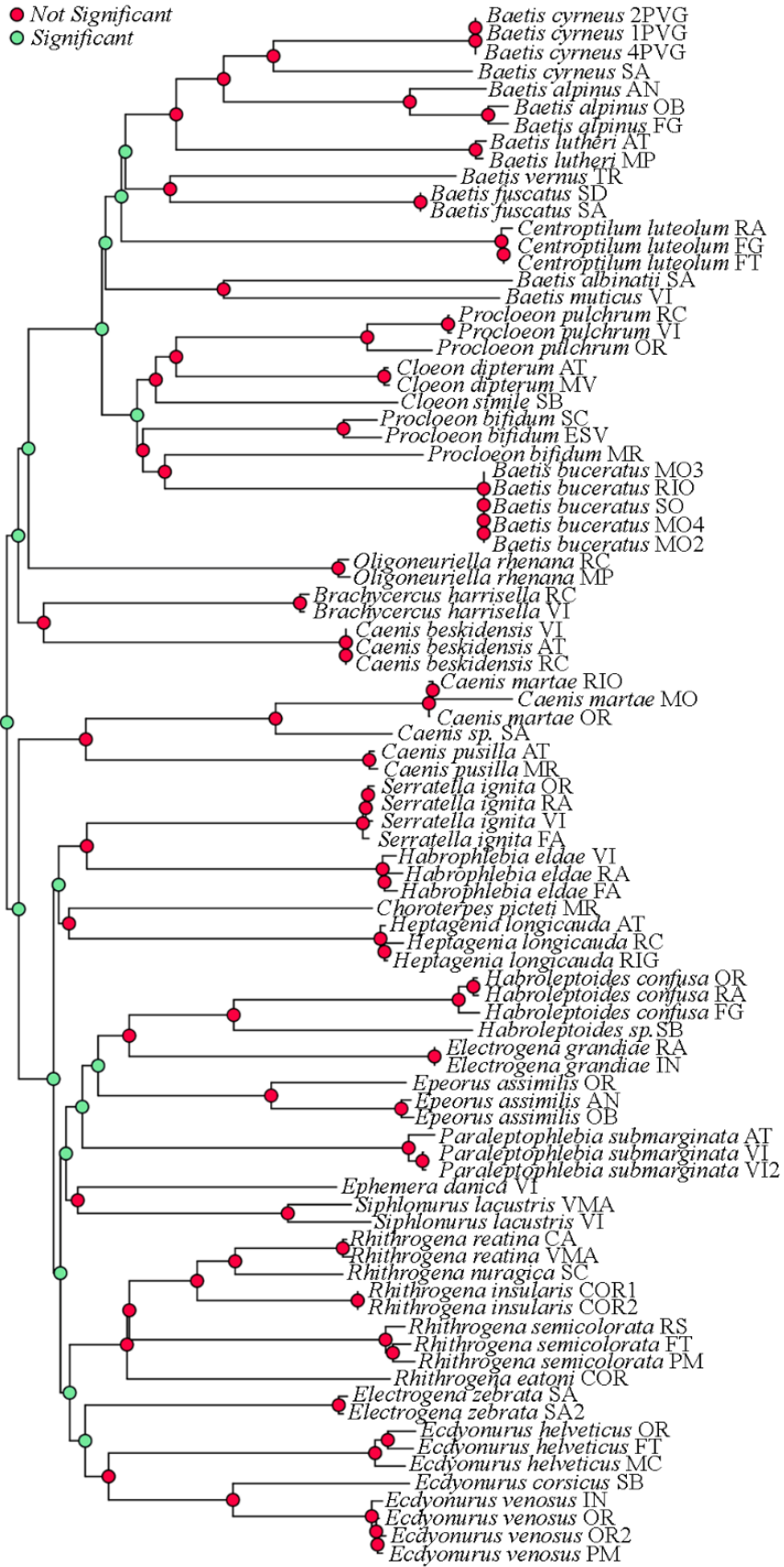


Fig.27 - Dendrogramma Neighbor Joining delle 89 sequenze di Efemeroterteri prodotto con Spider utilizzando i parametri di default e il test di Rosenberg sulla probabilità della reciproca monofilia ($\alpha = 0.005$).

I nodi in verde sono significativi.

Il filogramma di maximum likelihood generato con RaxML (Fig.28) ha esibito una topologia generale in linea con la tassonomia classica e tutte le specie raggruppate sono supportate statisticamente da alti valori di bootstrap eccetto per le specie con elementi singoli.

Uno stesso tipo di analisi effettuata con il bPTP server che utilizza un'implementazione Bayesiana al metodo "Poisson Tree Process", configurata con i parametri di default, ha mostrato un filogramma che identificava 45 cluster di specie putative (Fig.29).

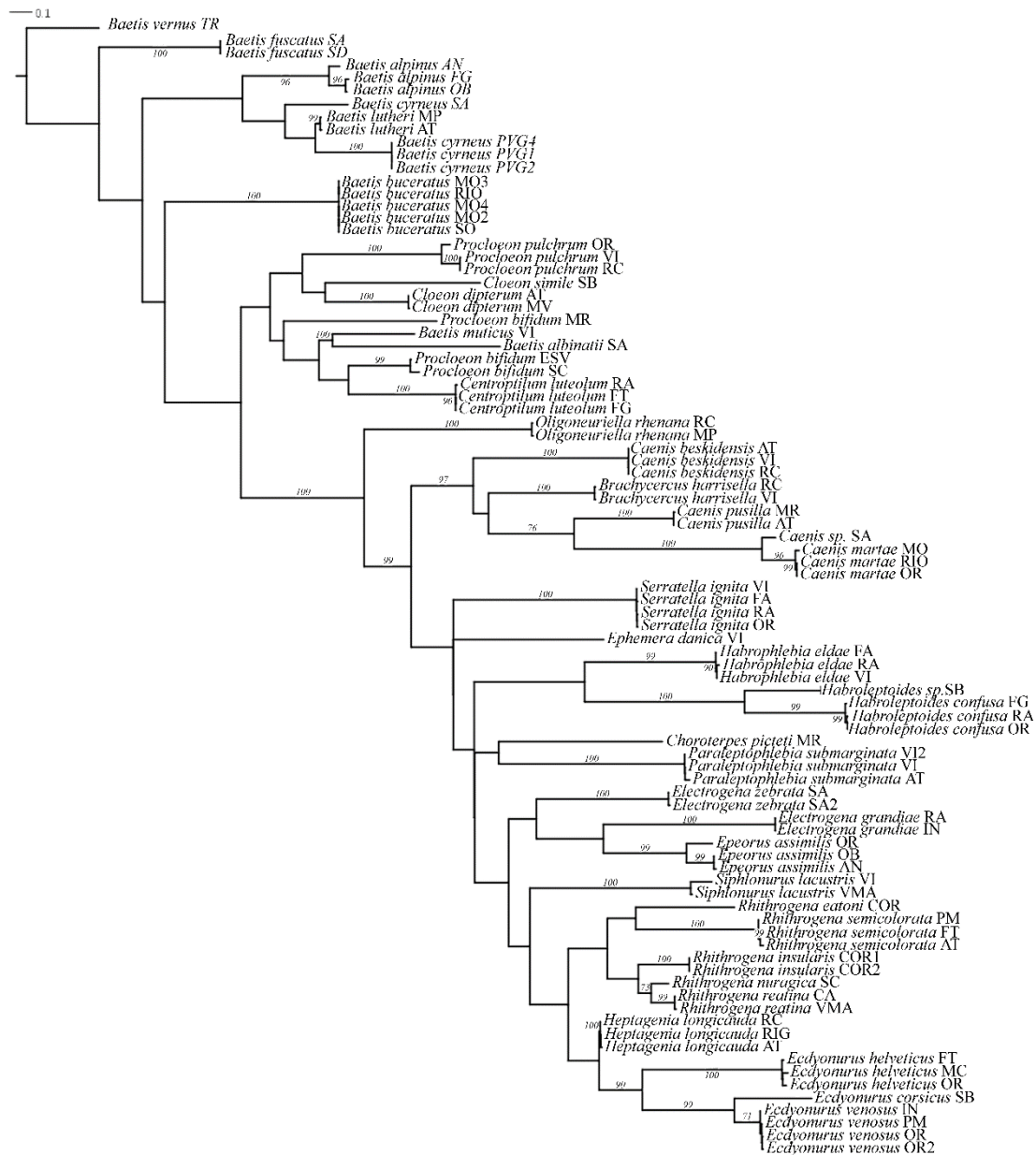


Fig.28 - Dendrogramma generato con RaxML delle 89 sequenze COX1 analizzate dove sono indicati i valori di bootstrap (1000 repliche).

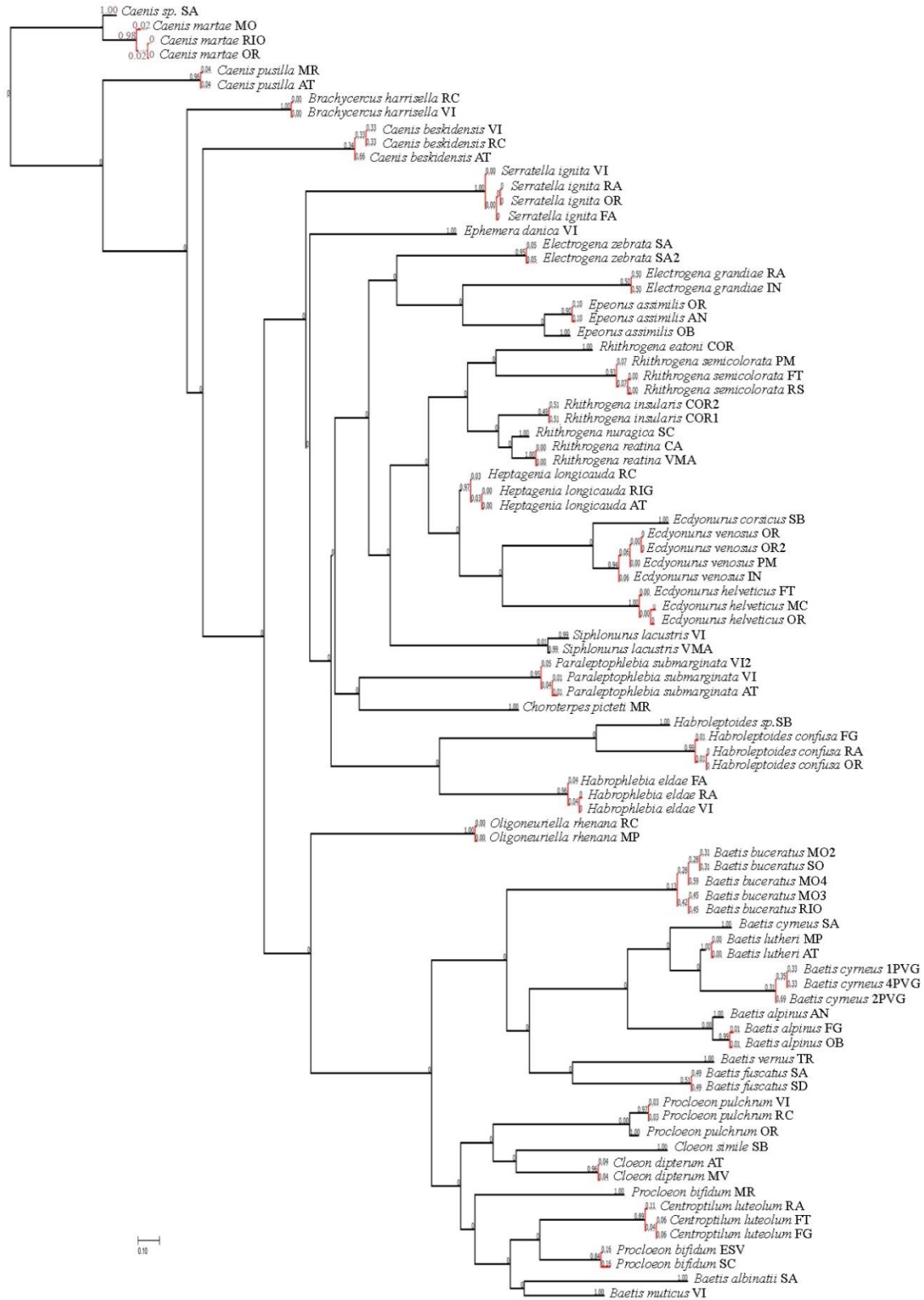


Fig.29 - Dendrogramma delle 89 sequenze COX1 generate con il metodo PTP in cui vengono desunte 45 specie putative.

6.1.3 Identificazione tassonomica nel contesto Europeo

La ricerca tassonomica effettuata sulle principali banche dati disponibili sul web (Tab.9) ha rivelato che 21 delle specie analizzate nel nostro dataset sono presenti in NCBI con la relativa sequenza COX1. Quando le nostre sequenze sono state caricate per verificare la correlazione tra specie identificata e la relativa sequenza assegnata, 48 sequenze presenti nel nostro dataset (corrispondenti a 18 specie) hanno riscontrato un abbinamento positivo, mentre 4 (relative alle specie *Baetis vernus*, *Baetis lutheri* e *Ephemera danica*) hanno rilevato una corrispondenza positiva con il genere ma non con la specie. Le rimanenti 37 sequenze sono state depositate per la prima volta in banca dati ampliando il database di 18 nuove specie di Efemeroteri.

La stessa indagine effettuata su BOLD ha evidenziato la presenza di 31 dei nostri COX1 *barcodes*, di questi ultimi 27 hanno mostrato un'esatta corrispondenza a livello di specie tra identificazione tassonomica assegnata e relativa sequenza, 4 specie invece, risultano avere un incongruente corrispondenza in banca dati. *Baetis lutheri* (viene associato a *Baetis alpinus*, lo stesso risultato è stato ottenuto con il blast in *GeneBank*), *Siphonurus lacustris* (associato a *Siphonurus croaticus*), *Cloeon simile* (ha una correlazione con un altro genere, *Procloeon*) e *Procloeon bifidum* di cui una sequenza ha rivelato una corrispondenza esatta a livello di specie mentre due hanno ottenuto un riscontro solamente a livello di genere (prendendo un'altra specie *P. viridoculare*). In generale la gran parte dei valori di identificazione riscontrati sono stati maggiori del 90% di identità.

Di particolare interesse è stato il caso di un campione di *Habroleptoides* proveniente da Rio Tistigliosi in Sardegna, identificato morfologicamente come *Habroleptoides sp.* in quanto danneggiato al punto che non è stato possibile il riconoscimento a livelli tassonomici più bassi. La corrispondenza tassonomica ottenuta con il *match* in GenBank e BOLD ha dato il medesimo risultato, rivelando una correlazione molto alta con un endemismo Sardo (*Habroleptoides modesta*).

Tab.9 - Presenza tassonomica (si/no) con la relativa migliore correlazione ottenuta in GeneBank e BOLD dei 77 aplotipi COX1. In rosso le correlazioni errate riscontrate nelle banche dati. * Sinonimo di *Baetis muticus* (Linnaeus, 1758); ** Presente solamente l'intero genoma.

Sequenza	NCBI			BOLD		
	Presenza	Specie correlata	Identità	Presence	Specie correlata	Identità
<i>Baetis muticus VI</i>	Yes*	<i>Alainites cf. muticus</i>	98%	Yes*	<i>Alainites cf. muticus</i>	97.88%
<i>Baetis vernus TR</i>	Yes	<i>Baetis liebenauae</i>	94%	Yes	<i>Baetis vernus</i>	96.73%
<i>Baetis albinatii SA</i>	No	<i>Alainites cf. muticus</i>	99%	No	<i>Alainites cf. muticus</i>	99.35%
<i>Baetis cyrneus SA</i>	Yes	<i>Baetis cf. cyrneus</i>	90%	Yes	<i>Baetis cf. cyrneus</i>	90.03%
<i>Baetis cyrneus 4PVG, 1PVG, MO</i>	Yes	<i>Baetis cf. cyrneus</i>	100%	Yes	<i>Baetis cf. cyrneus</i>	100%
<i>Baetis lutheri MP</i>	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	94.61%
<i>Baetis lutheri AT</i>	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	94.61%
<i>Baetis fuscatus SA, SD</i>	Yes	<i>Baetis fuscatus clone</i>	89%	Yes	<i>Baetis fuscatus clone</i>	91.34%
<i>Baetis alpinus FG</i>	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	95.09%
<i>Baetis alpinus AN</i>	Yes	<i>Baetis alpinus</i>	97%	Yes	<i>Baetis alpinus</i>	96.79%
<i>Baetis alpinus OB</i>	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	95.60%
<i>Baetis buceratus MO3, RIO, SO, MO2, MO4</i>	Yes	<i>Baetis buceratus</i>	89%	Yes	<i>Baetis buceratus</i>	90.36%
<i>Centroptilum luteolum RA</i>	Yes	<i>Centroptilum luteolum</i>	89%	Yes	<i>Centroptilum luteolum</i>	89.22%
<i>Centroptilum luteolum FT, FG</i>	Yes	<i>Centroptilum luteolum</i>	89%	Yes	<i>Centroptilum luteolum</i>	89.39%
<i>Procloeon pulchrum VI</i>	No	<i>Procloeon mendax</i>	85%	No	<i>Cloeon dipterum</i>	84.94%
<i>Procloeon pulchrum OR</i>	No	<i>Procloeon mendax</i>	86%	No	<i>Cloeon dipterum</i>	86.19%
<i>Procloeon pulchrum RC</i>	No	<i>Procloeon mendax</i>	85%	No	<i>Cloeon dipterum</i>	84.77%
<i>Procloeon bifidum MR</i>	No	<i>Procloeon mendax</i>	84%	Yes	<i>Procloeon bifidum</i>	86.21%
<i>Procloeon bifidum ES</i>	No	<i>Procloeon mendax</i>	85%	Yes	<i>Proloeon viridiculare</i>	86.11%

<i>Procloeon bifidum</i> SC	No	Unclassified Ephemeroptera	86%	Yes	<i>Procloeon viridiculare</i>	86.11%
<i>Cloeon dipterum</i> MV	Yes	<i>Cloeon dipterum</i>	99%	Yes	<i>Cloeon dipterum</i>	99.82%
<i>Cloeon dipterum</i> AT	Yes	<i>Cloeon dipterum</i>	99%	Yes	<i>Cloeon dipterum</i>	99.82%
<i>Cloeon simile</i> SB	No	<i>Procloeon</i> sp.	90%	Yes	<i>Procloeon</i> sp.	90.78%
<i>Caenis pusilla</i> MR	No	<i>Caenis anceps</i>	82%	Yes	<i>Caenis pusilla</i>	98.20%
<i>Caenis pusilla</i> AT	No	<i>Caenis</i> sp.	82%	Yes	<i>Caenis pusilla</i>	98.20%
<i>Caenis beskidensis</i> VI, RC, AT	No	<i>Caenis anceps</i>	83%	Yes	<i>Caenis beskidensis</i>	89.77%
<i>Caenis martae</i> MO	No	<i>Leptophlebia</i>	77%	Yes	<i>Caenis macrura</i>	79.84%
<i>Caenis martae</i> OR	No	<i>Caenis bajaensis</i>	80%	Yes	<i>Caenis macrura</i>	81.54%
<i>Caenis martae</i> RIO	No	<i>Caenis bajaensis</i>	81%	Yes	<i>Caenis macrura</i>	81.70%
<i>Caenis</i> sp. SA98	No	<i>Asionurus</i> sp.	80%	No	<i>Caenis</i>	82.68%
<i>Brachycercus harrisella</i> VI	Yes	<i>Brachycercus harrisella</i>	94%	Yes	<i>Brachycercus harrisella</i>	96.09%
<i>Brachycercus harrisella</i> RC	Yes	<i>Brachycercus harrisella</i>	94%	Yes	<i>Brachycercus harrisella</i>	96.24%
<i>Ecdyonurus helveticus</i> FT	Yes	<i>Ecdyonurus helveticus</i>	99%	Yes	<i>Ecdyonurus helveticus</i>	98.69%
<i>Ecdyonurus helveticus</i> OR	Yes	<i>Ecdyonurus helveticus</i>	98%	Yes	<i>Ecdyonurus helveticus</i>	98.20%
<i>Ecdyonurus helveticus</i> MC	Yes	<i>Ecdyonurus helveticus</i>	98%	Yes	<i>Ecdyonurus helveticus</i>	98.37%
<i>Ecdyonurus corsicus</i> SB	Yes	<i>Ecdyonurus corsicus</i>	99%	Yes	<i>Ecdyonurus corsicus</i>	99.67%
<i>Ecdyonurus venosus</i> OR	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.67%
<i>Ecdyonurus venosus</i> IN	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.84%
<i>Ecdyonurus venosus</i> PM	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.51%
<i>Ecdyonurus venosus</i> OR76	Yes	<i>Ecdyonurus venosus</i>	100%	Yes	<i>Ecdyonurus venosus</i>	100%
<i>Epeorus assimilis</i> OB	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	89.70%
<i>Epeorus assimilis</i> AN	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	89.87%

<i>Epeorus assimilis</i> OR	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	90.70%
<i>Electrogena grandie</i> IN, RA	No	<i>Electrogena</i> sp.	98%	No	<i>Electrogena</i> sp.	97.55%
<i>Electrogena zebrata</i> SA104	Yes	<i>Electrogena zebrata</i>	99%	Yes	<i>Electrogena zebrata</i>	99.84%
<i>Electrogena zebrata</i> SA105	Yes	<i>Electrogena zebrata</i>	99%	Yes	<i>Electrogena zebrata</i>	99.67%
<i>Rhithrogena semicolorata</i> PM	Yes	<i>Rhithrogena semicolorata</i>	98%	Yes	<i>Rhithrogena semicolorata</i>	98.37%
<i>Rhithrogena semicolorata</i> FT	Yes	<i>Rhithrogena semicolorata</i>	99%	Yes	<i>Rhithrogena semicolorata</i>	98.69%
<i>Rhithrogena semicolorata</i> RS	Yes	<i>Rhithrogena semicolorata</i>	99%	Yes	<i>Rhithrogena semicolorata</i>	99.84%
<i>Rhithrogena reatina</i> VMA79	No	<i>Rhithrogena</i> sp.	93%	No	<i>Rhithrogena</i> sp.	92.81%
<i>Rhithrogena reatina</i> CA	No	<i>Rhithrogena</i> sp.	93%	No	<i>Rhithrogena</i> sp.	93.14%
<i>Rhithrogena nuragica</i> SC	No	<i>Rhithrogena</i> sp.	99%	No	<i>Rhithrogena</i> sp.	99.02%
<i>Rhithrogena insulari</i> COR100, COR101	No	<i>Rhithrogena</i> sp.	90%	No	<i>Rhithrogena</i> sp.	90.03%
<i>Rhithrogena eatoni</i> OR103	No	<i>Rhithrogena</i> sp.	90%	No	<i>Rhithrogena</i> sp.	89.87%
<i>Heptagenia longicauda</i> RIG	No	<i>Heptagenia solitaria</i>	89%	Yes	<i>Heptagenia longicauda</i>	95.42%
<i>Heptagenia longicauda</i> AT	No	<i>Heptagenia solitaria</i>	89%	Yes	<i>Heptagenia longicauda</i>	95.75%
<i>Heptagenia longicauda</i> RC	No	<i>Heptagenia sulfura</i>	87%	Yes	<i>Heptagenia longicauda</i>	94.77%
<i>Habroleptoides confusa</i> FG	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	96.50%
<i>Habroleptoides confusa</i> RA	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	96.83%
<i>Habroleptoides confusa</i> OR	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	97.17%
<i>Habroleptoides</i> sp.SB	Yes	<i>Habroleptoides modesta</i>	92%	Yes	<i>Habroleptoides modesta</i>	91.67%
<i>Habrophlebia eldae</i> VI	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.35%
<i>Habrophlebia eldae</i> FA	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.51%
<i>Habrophlebia eldae</i> RA	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.35%

<i>Paraleptophlebia submarginata</i> AT	No	<i>Paraleptophlebia westoni</i>	82%	Yes	<i>Paraleptophlebia submarginata</i>	99.18%
<i>Paraleptophlebia submarginata</i> VI	No	<i>Siphonurus</i> sp.	81%	Yes	<i>Paraleptophlebia submarginata</i>	98.69%
<i>Paraleptophlebia submarginata</i> VI36	No	<i>Siphonurus</i> sp.	81%	Yes	<i>Paraleptophlebia submarginata</i>	98.85%
<i>Choroterpes picteti</i> MR77	No	<i>Habroleptoides</i> sp.	85%	Yes	<i>Choroterpes picteti</i>	99.35%
<i>Serratella ignita</i> VI	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.67%
<i>Serratella ignita</i> RA	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.84%
<i>Serratella ignita</i> FA	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.51%
<i>Serratella ignita</i> OR	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.84%
<i>Oligoneuriella rhenana</i> MP	No	<i>Anopheles labranchiae</i>	85%	Yes	<i>Oligoneuriella rhenana</i>	89.30%
<i>Oligoneuriella rhenana</i> RC	No	<i>Anopheles labranchiae</i>	84%	Yes	<i>Oligoneuriella rhenana</i>	90%
<i>Siphonurus lacustris</i> VI	Yes	<i>Siphonurus lacustris</i>	92%	Yes	<i>Siphonurus croaticus</i>	94.28%
<i>Siphonurus lacustris</i> VMA	Yes	<i>Siphonurus lacustris</i>	93%	Yes	<i>Siphonurus croaticus</i>	95.10%
<i>Ephemera danica</i> VI	Yes	<i>Ephemera sachalinensis</i> **	89%	Yes	<i>Ephemera danica</i>	98.36%

L'integrazione delle nostre sequenze con le COX1 *barcode* degli Efemerotteri europei presenti su BOLD ha prodotto una matrice di 2622 sequenze. La successiva selezione di queste rispetto a determinati canoni quali la lunghezza (>500 bp), l'assenza di gap, gli aplotipi identici e la relativa appartenenza rispetto alle famiglie presenti nel nostro dataset, ne ha ridotto il numero a 570 (appartenenti a più di 200 specie e 7 famiglie) (Tab.10).

Il risultato dell'analisi con RaxML è un dendrogramma che ha mostrato una topologia in accordo con la tassonomia riconosciuta degli Efemerotteri (Fig.30, il dendrogramma in alta risoluzione con tutti i nomi delle specie è riportato in Appendice I).

Le sette famiglie analizzate formano gruppi coesi, con l'unica eccezione per la famiglia Heptageniidae che appare parafiletica includendo al suo interno la famiglia Siphonuridae.

All'interno del dataset completo, tutti generi rappresentati da più di due individui formano gruppi coesi; gli unici che mostrano un segnale poco chiaro risultando mescolati al loro interno sono gli individui appartenenti a 4 generi della famiglia Baetidae. In quest'ultima due sequenze del nostro dataset *Baetis albinatii* e *Baetis muticus* risultano esterne al cluster *Baetis* raggruppandosi invece in un sottogruppo parafiletico comprendente *Centroptilum* (7 specie) *Cloeon* (5 specie) e *Procloeon* (14 specie). Questi tre generi appaiono a loro volta confusi, con le sequenze del nostro database e quelle provenienti da BOLD che si includono in differenti raggruppamenti specifici.

Anche altri due generi appaiono parafiletici, il genere *Paraleptophlebia* con 4 specie su 20 (*P. submarginata* appartenente al nostro campionamento e altre 3 specie integrate da BOLD) che si raggruppano con *Habrophlebia* e *Habroleptoides*, ed il genere *Ecdyonurus* con una specie (*E. criddlei*) che si inserisce nel cluster del genere *Electrogena*.

In tutti gli altri casi le sequenze si raggruppano correttamente con i loro generi e famiglie di appartenenza. Le sequenze del nostro dataset formano anche esse gruppi coesi includendo i *barcode* conspecifici provenienti da BOLD (quando disponibili). Tuttavia, numerosi casi di specie parafiletiche sono stati rilevati in quasi tutte le famiglie analizzate.

Tab.10 - Lista tassonomica delle sequenze COX1 scaricate da BOLD ad integrazione del nostro dataset usate nelle analisi di RaxML

<i>Famiglia</i>	<i>Genere</i>	<i>specie</i>							
Baetidae	<i>Baetis</i>	<i>bicaudatus</i>	<i>tricaudatus</i>	<i>magnus</i>	<i>phoebus</i>	<i>bundyae</i>	<i>adonis</i>	<i>intercalaris</i>	<i>foemina</i>
		<i>persecutor</i>	<i>brunneicolor</i>	<i>hudsonicus</i>	<i>rhodani</i>	<i>vernus</i>	<i>subalpinus</i>	<i>macani</i>	<i>liebenauae</i>
		<i>notos</i>	<i>flavistriga</i>	<i>rusticans</i>	<i>pluto</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>
		<i>sp7</i>	<i>sp8</i>	<i>sp9</i>	<i>sp10</i>	<i>sp44</i>	<i>sp41</i>	<i>sp38</i>	<i>sp30</i>
		<i>sp36</i>	<i>sp34</i>	<i>sp33</i>	<i>sp35</i>				
	<i>Cloeon</i>	<i>cognatum</i>	<i>praetextum</i>	<i>smaeleni</i>					
	<i>Procloeon</i>	<i>fragile</i>	<i>rufostrigatum</i>	<i>viridoculare</i>	<i>rivulare</i>	<i>mendax</i>	<i>pennulatum</i>	<i>rubropictum</i>	<i>intermediale</i>
		<i>Prru6</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>	<i>sp5</i>	<i>sp6</i>	<i>sp7</i>
		<i>sp8</i>	<i>sp9</i>	<i>sp10</i>	<i>sp11</i>	<i>sp12</i>			
	<i>Centroptilum</i>	<i>triangulifer</i>	<i>minor</i>	<i>alamance</i>	<i>luteolum</i>	<i>victoriae</i>	<i>bifurcatum</i>		
Caenidae	<i>Brachycercus</i>	<i>harrisella</i>							
	<i>Caenis</i>	<i>youngi</i>	<i>latipennis</i>	<i>bajaensis</i>	<i>diminuta</i>	<i>punctata</i>	<i>amica</i>	<i>anceps</i>	<i>horaria</i>
		<i>eglinensis</i>	<i>hilaris</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>	<i>sp5</i>	<i>sp6</i>
		<i>sp7</i>	<i>sp48</i>	<i>sp89</i>	<i>sp81</i>	<i>sp83</i>	<i>sp84</i>		
Heptageniidae	<i>Ecdyonurus</i>	<i>venosus</i>	<i>corsicus</i>	<i>criddlei</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>	<i>sp5</i>
		<i>sp6</i>	<i>sp7</i>	<i>sp8</i>	<i>sp9</i>	<i>sp10</i>			
	<i>Electrogena</i>	<i>fallax</i>	<i>lateralis</i>	<i>zebrata</i>	<i>sp2</i>				
	<i>Epeorus</i>	<i>longimanus</i>	<i>fragilis</i>	<i>pleuralis</i>	<i>vitreus</i>	<i>longimanus</i>	<i>grandis</i>	<i>album</i>	<i>nr. Dulciana</i>
		<i>sylvicola</i>	<i>sylvicola</i>	<i>subpallidus</i>	<i>deceptivus</i>	<i>sp1</i>	<i>sp2</i>	<i>sp14</i>	<i>sp22</i>
	<i>Heptagenia</i>	<i>solitaria</i>	<i>pulla</i>	<i>marginalis</i>	<i>dalecarlica</i>	<i>flavescens</i>	<i>solitaria</i>	<i>culacantha</i>	<i>nr. Townesi</i>
		<i>nr. Dolosa</i>	<i>whitingi</i>	<i>marginalis</i>	<i>julia</i>	<i>adaequata</i>	<i>elegantula</i>	<i>sp1</i>	<i>sp2</i>
	<i>Rhithrogena</i>	<i>amica</i>	<i>robusta</i>	<i>manifesta</i>	<i>nr. rubicunda</i>	<i>uhari</i>	<i>anomal</i>	<i>impersonata</i>	<i>sanguinea</i>

Università Degli Studi della Tuscia di Viterbo
 Corso di Dottorato di Ricerca – Ecologia e Gestione delle Risorse Biologiche

	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>	<i>sp5</i>	<i>sp6</i>	<i>sp7</i>	<i>sp8</i>	
	<i>sp9</i>	<i>sp10</i>	<i>sp11</i>	<i>sp12</i>	<i>sp13</i>	<i>sp14</i>	<i>sp15</i>	<i>sp16</i>	
	<i>sp17</i>	<i>sp18</i>	<i>sp19</i>	<i>sp20</i>	<i>sp21</i>	<i>sp22</i>	<i>sp23</i>	<i>sp24</i>	
	<i>sp25</i>	<i>sp26</i>	<i>sp27</i>	<i>sp28</i>	<i>sp29</i>	<i>sp30</i>	<i>sp31</i>	<i>sp32</i>	
	<i>sp33</i>	<i>sp34</i>	<i>sp35</i>	<i>sp36</i>	<i>sp37</i>	<i>sp38</i>	<i>sp39</i>	<i>sp40</i>	
	<i>sp41</i>	<i>sp42</i>	<i>sp43</i>	<i>sp44</i>	<i>sp45</i>	<i>sp46</i>	<i>sp47</i>	<i>sp48</i>	
	<i>sp49</i>	<i>sp50</i>	<i>sp51</i>						
Ephemerellidae	<i>Serratella</i>	<i>micheneri</i>	<i>serratoides</i>	<i>serrata</i>	<i>frisoni</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>
		<i>sp5</i>							
Siphonuriidae	<i>Siphonurus</i>	<i>quebecensis</i>	<i>phyllis</i>	<i>occidentalis</i>	<i>alternatus</i>	<i>barbaroides</i>	<i>rapidus</i>	<i>lacustris</i>	<i>minnoi</i>
		<i>typicus</i>	<i>marshalli</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>	<i>sp4</i>		
Ephemeridae	<i>Ephemer</i>	<i>simulans</i>	<i>vulgata</i>	<i>orientalis</i>	<i>blanda</i>	<i>guttulata</i>			
Leptophlebiidae	<i>Habrophlebia</i>	<i>vibrans</i>	<i>lauta</i>	<i>eldae</i>					
	<i>Habroleptoides</i>	<i>budtzi</i>	<i>modesta</i>	<i>confusa</i>	<i>sp1</i>				
	<i>Paraleptophlebia</i>	<i>assimilis</i>	<i>adoptiva</i>	<i>adoptiva</i>	<i>mollis</i>	<i>vaciva</i>	<i>debilis</i>	<i>heteronea</i>	<i>volitans</i>
		<i>guttata</i>	<i>cf. debilis</i>	<i>cf. debilis</i>	<i>praepedita</i>	<i>strandii</i>	<i>ontario</i>	<i>jeanae</i>	<i>gregalis</i>
		<i>kirchneri</i>	<i>heteronea</i>	<i>memorialis</i>	<i>guttata</i>	<i>adoptiva</i>	<i>sp1</i>	<i>sp2</i>	<i>sp3</i>
		<i>sp4</i>	<i>sp5</i>						
	<i>Choroerpes</i>	<i>basalis</i>							

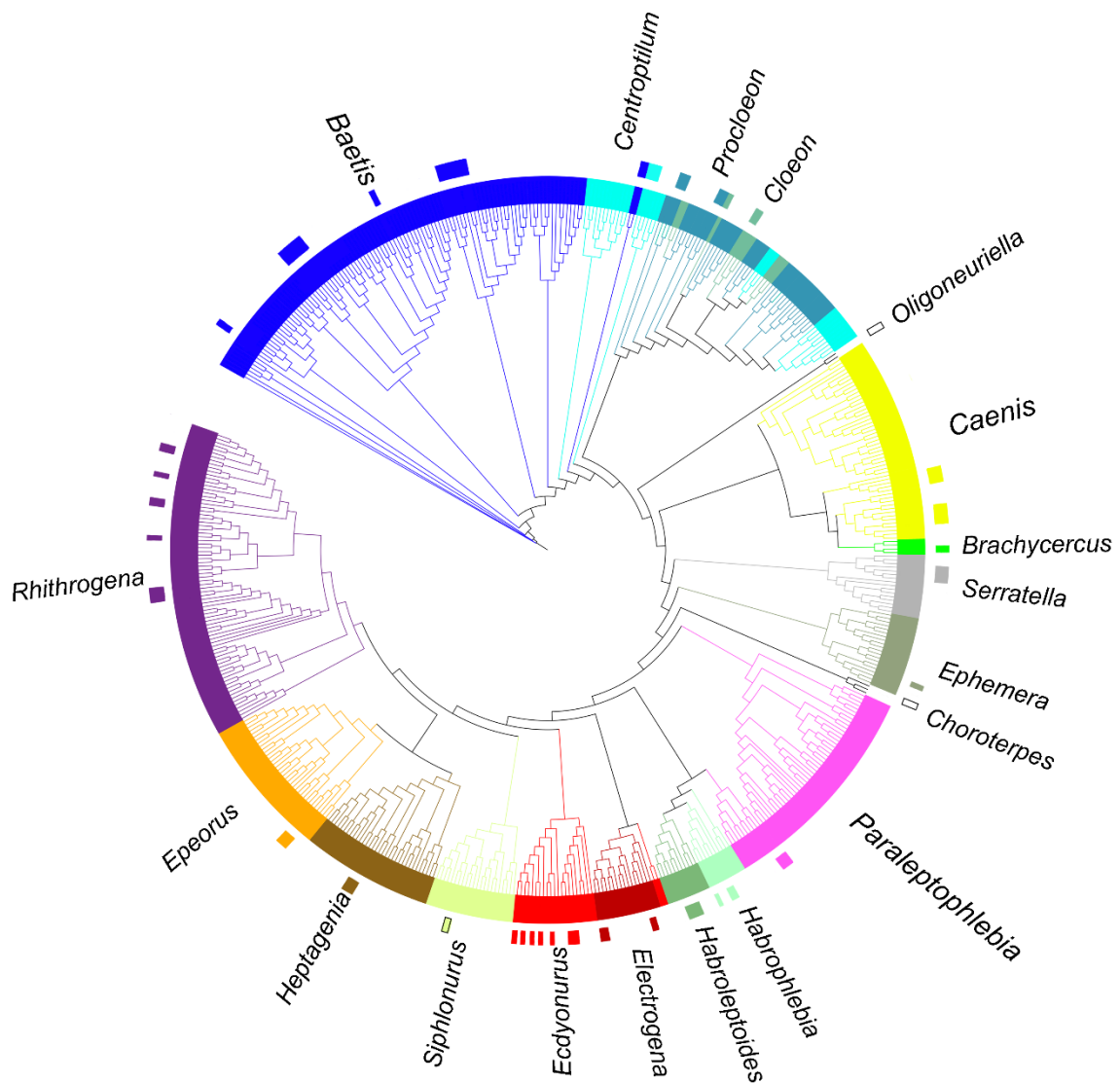


Fig.30 - Dendrogramma circolare RaxML di 570 sequenze COX1 di Efemerotteri del presente dataset inclusi i *barcode* provenienti da BOLD. I segmenti indicano la posizione delle sequenze provenienti dai campioni analizzati in questo studio. I differenti generi sono indicati con colori differenti ed il relativo nome.

6.1.4 *Barcoding* ambientale delle acque dolci

Una descrizione generale del dataset è mostrata in Tab.11.

Tab.11 - Conteggio totale delle sequenze grezze ottenute dalla corsa *NGS* e conteggio totale delle sequenze pulite utilizzate per l'analisi dei dati nei campioni di macroinvertebrati.

<i>Library</i>		N° Sequenze grezze	N° Sequenze pulite
Efemerotteri 1	Fw	83784	82279
	Rev	83784	82158
Efemerotteri 2	Fw	36925	36471
	Rev	36925	36427
Bulk 1	Fw	29241	29201
	Rev	29241	29178
Bulk 2	Fw	37128	37046
	Rev	37128	37074
Bulk 3	Fw	60455	60334
	Rev	60455	60285
eDNA Sedimento	Fw	123	121
	Rev	123	122
eDNA acqua	Fw	2838	2771
	Rev	2838	2809

➤ *Analisi library* Efemerotteri

L'analisi preliminare condotta sul multiallineamento totale dei nostri frammenti barcode di Efemerotteri (39 specie con 91 individui vedi Appendice III), ha evidenziato che le prime e le ultime 206 bp della regione COX1 erano in grado di identificare tutte le specie senza che si evidenziasse un overlap tra di esse sino ad una soglia di similarità del 95%. Sulla base di quest'ultima le sequenze assegnate con una similarità inferiore sono state eliminate. Le restanti sequenze (95-100%) sono state raggruppate per specie.

I risultati ottenuti sono riportati nelle Tab.12 - a,b,c.

Tab.12 - a) Comparazione tra i taxa inseriti nella library (Efemerotteri 1) e le sequenze NGS, b) Comparazione tra i taxa inseriti nella library (Efemerotteri 2) e le sequenze NGS, c) Tabella riassuntiva delle specie riconosciute.

a)

Taxa	N° sequenze ottenute		Presenza/assenza Fw+Rev
	Efemerotteri 1		
	Fw	Rev	
<i>Baetis alpinus FG</i>	162	138	√
<i>Baetis alpinus OB</i>	23	19	√
<i>Baetis alpinus AN</i>	0	0	X
<i>Baetis cyrneus SA</i>	5	3	√
<i>Baetis cyrneus 4 PVG 1 PVG 2 PVG</i>	0	0	X
<i>Baetis vernus TR80</i>	0	0	X
<i>Baetis muticus VI</i>	0	0	X
<i>Baetis albinatii SA</i>	0	0	X
<i>Baetis lutheri MP62 AT72</i>	0	0	X
<i>Baetis fuscatus SA SD</i>	0	0	X
<i>Baetis buceratus MO3 RIO SO MO2 MO4</i>	0	0	X
<i>Brachycercus harrisella RC</i>	68290	16102	√
<i>Brachycercus harrisella VI</i>	0	42191	√
<i>Caenis beskidensis VI</i>	388	342	√
<i>Caenis martae RIO</i>	18	9	√
<i>Caenis martae OR</i>	0	2	√
<i>Caenis pusilla AT</i>	1	1	√
<i>Centroptilum luteolum FT</i>	156	124	√
<i>Centroptilum luteolum RA</i>	15	15	√
<i>Choroterpes picteti MR77</i>	605	486	√
<i>Cloeon dipterum</i>	0	0	X
<i>Cloeon simile</i>	0	0	X
<i>Ecdyonurus corsicus SB</i>	21	20	√
<i>Ecdyonurus helveticus FT</i>	12	8	√
<i>Ecdyonurus helveticus MC</i>	56	45	√
<i>Ecdyonurus helveticus OR</i>	4	4	√
<i>Ecdyonurus venosus IN</i>	0	280	√
<i>Ecdyonurus venosus OR</i>	0	139	√
<i>Ecdyonurus venosus OR76</i>	19	29	√
<i>Ecdyonurus venosus PM</i>	533	0	√
<i>Ephemera danica</i>	0	0	X
<i>Electrogena zebrata SA105</i>	56	2	√
<i>Electrogena zebrata SA105</i>	0	47	√
<i>Electrogena grandiae</i>	0	0	X
<i>Epeorus assimilis AN</i>	103	129	√
<i>Epeorus assimilis OB</i>	61	919	√
<i>Epeorus assimilis OR</i>	1191	0	√
<i>Habroleptoides confusa OR</i>	193	168	√
<i>Habroleptoides sp.</i>	0	0	X
<i>Habrophlebia eldae FA</i>	259	172	√
<i>Habrophlebia eldae RA</i>	82	66	√
<i>Habrophlebia eldae VI</i>	0	41	√
<i>Heptagenia longicauda AT</i>	39	31	√
<i>Heptagenia longicauda RC</i>	1	0	√
<i>Oligoneuriella rhenana</i>	0	0	X
<i>Paraleptophlebia submarginata AT</i>	280	227	√
<i>Paraleptophlebia submarginata VI36</i>	1506	1380	√
<i>Paraleptophlebia submarginata VI</i>	0	1	√
<i>Procloeon bifidum SC</i>	0	1	√
<i>Procloeon bifidum ESV</i>	1	0	√
<i>Procloeon pulchrum</i>	0	0	X
<i>Procloeon pulchrum</i>	0	0	X
<i>Procloeon pulchrum</i>	0	0	X

<i>Rhithrogena eatoni cor103</i>	5	4	✓
<i>Rhithrogena nuragica 3</i>	1	18	✓
<i>Rhithrogena nuragica SC</i>	22	0	✓
<i>Rhithrogena reatina CA</i>	13	13	✓
<i>Rhithrogena semicolorata FT</i>	117	96	✓
<i>Rhithrogena semicolorata RS</i>	5	4	✓
<i>Rhithrogena semicolorata PM</i>	0	4	✓
<i>Rhithrogena insularis</i>	0	0	X
<i>Serratella ignita FA</i>	66	11	✓
<i>Serratella ignita OR</i>	35	45	✓
<i>Serratella ignita RA</i>	0	24	✓
<i>Serratella ignita VI</i>	0	47	✓
<i>Siphonurus lacustris VI</i>	122	108	✓
Aplotipi comunità artificiale			66
Aplotipi sequenziati			48
Correlazione			73%

b)

Taxa	N° sequenze ottenute Efemerotteri 2		Presenza/assenza Fw+Rev
	Fw	Rev	
<i>Baetis alpinus FG</i>	99	81	✓
<i>Baetis alpinus OB</i>	0	0	X
<i>Baetis alpinus AN</i>	0	0	X
<i>Baetis cyrneus SA</i>	0	0	X
<i>B.cyrneus 4PVG 1 PVG 2 PVG</i>	0	0	X
<i>Baetis vernus TR80</i>	1	1	✓
<i>B.muticus vi</i>	0	0	X
<i>B.albinatii sa</i>	0	0	X
<i>B.lutheri MP62 AT72</i>	0	0	X
<i>B.fuscatus SA SD</i>	0	0	X
<i>B.buceratus MO3 RIO SO MO2 MO4</i>	0	0	X
<i>Brachycercus harrisella RC</i>	31135	18776	✓
<i>Brachycercus harrisella VI</i>	0	7442	✓
<i>Caenis beskidensis VI</i>	485	416	✓
<i>Caenis martae RIO</i>	10	7	✓
<i>Caenis martae OR</i>	0	1	✓
<i>Caenis pusilla MR78</i>	2	2	✓
<i>Centroptilum luteolum FT</i>	20	15	✓
<i>Centroptilum luteolum RA</i>	13	11	✓
<i>Choroterpes picteti MR77</i>	109	78	✓
<i>Cloeon dipterum</i>	0	0	X
<i>Cloeon simile</i>	0	0	X
<i>Ecdyonurus corsicus SB</i>	6	4	✓
<i>Ecdyonurus helveticus FT</i>	5	3	✓
<i>Ecdyonurus helveticus MC</i>	0	1	✓
<i>Ecdyonurus helveticus OR</i>	8	5	✓
<i>Ecdyonurus venosus IN</i>	0	0	X
<i>Ecdyonurus venosus OR</i>	0	1	✓
<i>Ecdyonurus venosus OR76</i>	4	3	✓
<i>Ecdyonurus venosus PM</i>	2	0	✓
<i>Ephemera danica</i>	0	0	X
<i>Electrogena zebrata SA104</i>	3	0	✓
<i>Electrogena zebrata SA105</i>	0	3	✓
<i>Electrogena grandiae</i>	0	0	X
<i>Epeorus assimilis AN</i>	10	16	✓
<i>Epeorus assimilis OR</i>	18	144	✓
<i>Epeorus assimilis OB</i>	193	0	✓
<i>Habroleptoides confusa OR</i>	77	65	✓

<i>Habroleptoides sp.</i>	0	0	X
<i>Habrophlebia eldae FA</i>	53	20	√
<i>Habrophlebia eldae RA</i>	2	2	√
<i>Habrophlebia eldae VI</i>	0	23	√
<i>Heptagenia longicauda AT</i>	2	2	√
<i>Heptagenia longicauda RC</i>	0	0	X
<i>Oligoneuriella rhenana</i>	0	0	X
<i>Paraleptophlebia submarginata AT</i>	182	145	√
<i>Paraleptophlebia submarginata VI36</i>	106	92	√
<i>Paraleptophlebia submarginata VI</i>	0	0	X
<i>Proclleon bifidum SC</i>	0	0	X
<i>Proclleon bifidum ESV</i>	0	0	X
<i>Proclleon pulchrum</i>	0	0	X
<i>Proclleon pulchrum</i>	0	0	X
<i>Proclleon pulchrum</i>	0	0	X
<i>Rhithrogena eatoni cor103</i>	0	0	X
<i>Rhithrogena nuragica 3</i>	0	0	X
<i>Rhithrogena nuragica SC</i>	0	0	X
<i>Rhithrogena reatina CA</i>	0	0	X
<i>Rhithrogena semicolorata FT</i>	16	14	√
<i>Rhithrogena semicolorata RS</i>	0	0	X
<i>Rhithrogena semicolorata PM</i>	0	0	X
<i>Rhithrogena insularis</i>	0	0	X
<i>Serratella ignita FA</i>	32	2	√
<i>Serratella ignita OR</i>	6	7	√
<i>Serratella ignita RA</i>	0	2	√
<i>Serratella ignita VI</i>	0	30	√
<i>Siphonurus lacustris VI</i>	1	0	√
Aplotipi comunità artificiale			37
Aplotipi sequenziati			66
Correlazione			56%

c)

Famiglia	Specie	Presenza Fw+Rw
Baetidae	<i>Baetis alpinus</i>	√
	<i>Baetis cyrneus</i>	√
	<i>Baetis vernus</i>	√
	<i>Baetis muticus</i>	X
	<i>Baetis albinatii</i>	X
	<i>Baetis lutheri</i>	X
	<i>Baetis fuscatus</i>	X
	<i>Baetis buceratus</i>	X
	<i>Centroptilum luteolum</i>	√
	<i>Cloeon dipterum</i>	X
	<i>Cloeon simile</i>	X
	<i>Proclleon bifidum</i>	√
	<i>Proclleon pulchrum</i>	X
	Caenidae	<i>Brachycercus harrisella</i>
<i>Caenis beskidensis</i>		√
<i>Caenis martae</i>		√
<i>Caenis pusilla</i>		√
<i>Caenis sp.</i>		X
Ephemerellidae	<i>Serratella ignita</i>	√
Ephemeridae	<i>Ephemera danica</i>	X
Heptageniidae	<i>Ecdyonurus corsicus</i>	√
	<i>Ecdyonurus helveticus</i>	√
	<i>Ecdyonurus venosus</i>	√
	<i>Electrogena zebrata</i>	√

	<i>Electrogena grandiae</i>	X
	<i>Epeorus assimilis</i>	✓
	<i>Heptagenia longicauda</i>	✓
	<i>Rhithrogena eatoni</i>	✓
	<i>Rhithrogena nuragica</i>	✓
	<i>Rhithrogena reatina</i>	✓
	<i>Rhithrogena semicolorata</i>	✓
	<i>Rhithrogena insularis</i>	X
Leptophlebiidae	<i>Choroterpes picteti</i>	✓
	<i>Habroleptoides confusa</i>	✓
	<i>Habroleptoides sp.</i>	X
	<i>Habrophlebia eldae</i>	✓
	<i>Paraleptophlebia submarginata</i>	✓
Oligoneuridae	<i>Oligoneuriella rhenana</i>	X
Siphonuridae	<i>Siphonurus lacustris</i>	✓
	Aplotipi comunità artificiale	25
	Aplotipi sequenziati	39
	Correlazione	64%

Nelle tabelle mostrate risultano percentuali di successo lievemente differenti in base al disegno sperimentale (modalità con cui è stata costruita la comunità artificiale). Nella *library* Efemerotteri 1, dove le concentrazioni di DNA erano le stesse (ng/μl) per tutti i campioni, la percentuale di identificazione degli aplotipi è stata del 73% mentre nella *library* Efemerotteri 2, in cui è stato mescolato lo stesso volume (1μl) di DNA, la percentuale di identificazione è calata al 56%.

Sommando tutte le corrispondenze ottenute sia con le sequenze forward che reverse, le percentuali di successo a livello di specie risultano essere del 64% (Tab. 12c). Quest'ultima tabella mette anche in evidenza come famiglia Baetidae risulti essere il gruppo con le più basse percentuali di successo. Al contrario, per entrambe le *library* (Efemerotteri 1 e 2) si può notare una sovrastima per *Brachycercus harrisella* rispetto alle altre specie presenti nella comunità artificiale.

In entrambe le *library* non sono stati riscontrati taxa differenti da quelli inseriti all'interno del campione ambientale.

➤ Analisi *library* bulk, campioni ambientali di sedimento e acqua.

L'attribuzione tassonomica delle sequenze ottenute è stata effettuata a livello di famiglia, pertanto il criterio di accettazione dell'assegnazione tassonomica è stato impostato su una soglia del 80% d'identità genetica. Tutte le sequenze sotto la soglia indicata sono state scartate mentre le altre sono state clusterizzate secondo il livello tassonomico stabilito morfologicamente. Ad un'identificazione di maggior dettaglio (es. genere, specie) l'attribuzione tassonomica delle sequenze è stata effettuata manualmente utilizzando soglie di similarità più alte (es. 95-100%).

I risultati ottenuti sono riportati nelle Tab.13 a,b,c.

Tab.13 - Confronto tra le identificazioni ottenute con approcci morfologici e molecolari, a) Fiume Ortolano, b) Fiume Traponzo, c) Fiume Biedano.

I numeri tra parentesi rappresentano il numero di individui contati morfologicamente

a)

Ortolano (Bulk 1)						
Ordine	Morfologia			Sequenze ottenute		
	Famiglia	Genere	Specie	Forward	Reverse	
Amphipoda	Gammaridae 15			350	310	
Coleoptera	Dryopidae Ad 1			0	0	
	Elmidae Ad 6			80	557	
	Helodidae L 21			0	0	
	Hydraenidae Ad 1			80	233	
Diptera	Athericidae 5			23	2	
Ephemeroptera	Baetidae	<i>Baetis</i>	<i>Baetis muticus</i> 3	0	0	
			<i>Baetis rhodani</i> gr.	0	0	
			<i>Procloeon</i>	<i>Procloeon pennulatum</i> 2	0	0
	Ephemeridae	<i>Ephemera</i>	<i>Ephemera danica</i> 3	0	0	
	Heptageniidae juv 15				7	10
			<i>Ecdyonurus</i> 8		27	26
			Electrogena	<i>Electrogena grandiae</i> 9	0	0
	Leptophlebiidae	<i>Habroleptoides</i>	<i>Habroleptoides</i> sp 8	0	2	
		<i>Habrophlebia</i>	<i>Habrophlebia eldae</i> 1	3	3	
Lumbriculida	Lumbriculidae 2			1	1	
Megaloptera	Sialidae			14	1	
Neotaenioglossa	Hydrobiidae	<i>Potamopyrgus</i> 2		17	21	
Plecoptera	Leuctridae	<i>Leuctra</i> 4 juv		9	16	
	Nemouridae	<i>Protonemura</i> 2		263	22	
Trichoptera	Hydropsychidae 18			24354	22721	
	Philopotamidae 2			3270	1122	
Tricladida	Dugesiiidae	<i>Dugesia</i> 5		0	0	

b)

Traponzo (Bulk 2)							
Ordine	Famiglia	Morfologia		Sequenze ottenute			
		Genere	Specie	Forward	Reverse		
Coleoptera	Dryophidae	2 Ad		0	0		
	Hydraenidae	Ad		1028	929		
Diptera	Chironomidae	6		170	91		
	Simuliidae	11		24617	24667		
Ephemeroptera	Baetidae	<i>Baetis</i>	<i>Baetis buceratus</i>	18	0	125	
			<i>Baetis fuscatus</i>	8	0	0	
			<i>Baetis lutheri</i>		30	0	
			<i>Baetis muticus</i>		0	0	
			<i>Baetis rhodani gr.</i>	3	0	0	
Ephemeroptera	Caenidae	<i>Caenis</i>	<i>Caenis pseudorivolorum gr.</i>		0*	0*	
			<i>Caenis pusilla</i>	19	0	0	
	Heptageniidae	<i>Heptagenia</i>	<i>Heptagenia longicauda</i>	2	0	0	
Odonata	Gomphidae	<i>Onychogomphus</i>		12	0	1	
Plecoptera	Leuctridae	<i>Leuctra</i>		4	117	216	
Trichoptera	Hydropsychidae				59	93	52
Tricladida	Dugesiidae	<i>dugesia</i>		2	0	0	
	Tubificidae				0	11	

c)

Biedano							
Ordine	Famiglia	Morfologia		Sequenze ottenute			
		Genere	Specie	Forward	Reverse		
Amphipoda	Athericidae	1			2	14	
Coleoptera	Dryopidae	ad 24			0	36	
	Elmidae	6 ad			35	322	
	Helodidae	1			0	0	
	Hydraenidae	ad 5			806	568	
Diptera	Simulidae	48			5197	4203	
	Tabanidae	1			0	0	
	Tipulidae				1881	6284	
Ephemeroptera	Baetidae	<i>Baetis</i>	<i>Baetis buceratus</i>	9	3	69	
			<i>Baetis fuscatus</i>	37	0	0	
			<i>Baetis rhodani gr</i>	13	0	0	
			<i>Caenis pusilla</i>	1	562	3914	
	Caenidae	<i>Caenis</i>	<i>Caenis pusilla</i>	1			
	Heptageniidae	<i>Ecdyonurus</i>	<i>Ecdyonurus venosus</i>	1	5	1	
Neotaenioglossa	Hydrobiidae	16			9	37	
Odonata	Gomphidae	<i>Onychogonphus</i>		4	1	3	
Plecoptera	Leuctridae	<i>Leuctra</i>		30	18	56	
Pulmonata	Ancylidae	<i>Ancylus</i>		1	0	0	
Trichoptera	Hydropsychidae				2	267	4763
	Sericostomatidae	juv 2			0	0	
Tricladida	Dugesiidae	<i>Dugesia</i>		8	0	0	

Il confronto tra l'identificazione con approccio NGS e identificazione con approccio morfologico ha mostrato una percentuale di corrispondenza pari al 65% per i fiumi Biedano e Ortolano (corrispondenti alle bulk 1 e 3) ed una percentuale pari al 58% per il fiume Traponzo (bulk 2).

A differenza delle precedenti library dedicate agli Efemerotteri, tutte e tre le bulk analizzate hanno evidenziato la presenza di sequenze appartenenti a taxa non evidenziati morfologicamente, spesso anche con elevati numeri di sequenze, ad esempio: *Tipula* (bulk 1; 1.930 sequenze), *Potamopyrgus*, *Tipula* e *Sialis* (bulk 2; 3.848, 2.481 e 2.125 sequenze, rispettivamente), e altri generi presenti con più di 2000 sequenze nella bulk 3 (ad es., *Chimarra* e *Rhithrogena* con 37.921 e 18.243 sequenze, rispettivamente). L'elenco completo di questi generi è riportato in Tab.14.

Tab.14 - Lista dei taxa (a livello di genere) rilevati dalla corsa NGS per le tre bulk ma non presenti nella lista riconosciuta morfologicamente.

Famiglia	Genere	Bulk 1	Bulk 2	Bulk 3
Anisogammaridae	<i>Jesogammarus</i>	√		
Athericidae	<i>Atherix</i>		√	
Baetidae	<i>Baetis</i>		√	√
	<i>Cloeon</i>	√	√	√
	<i>Procloeon</i>	√	√	
Caenidae	<i>Brachycercus</i>	√	√	√
	<i>Caenis</i>	√		
Chironomidae	<i>Ablabesmyia</i>	√		√
	<i>Arctopelopia</i>	√		√
	<i>Clinotanypus</i>	√		√
	<i>Conchapelopia</i>	√		√
	<i>Kiefferulus</i>			√
	<i>Macropelopia</i>			√
	<i>Paracladopelma</i>			√
	<i>Psectrotanypus</i>			√
Corydalidae	<i>Corydalidae</i>		√	√
	<i>Protohermes</i>		√	√
Crangonyctidae	<i>Sialis</i>	√	√	√
Curculionidae	<i>Homarus</i>	√	√	√
Dytiscidae	<i>Nebrioporus</i>		√	√
Elmidae	<i>Elmis</i>		√	√
	<i>Limnius</i>		√	√
	<i>Ordobrevia</i>		√	√
	<i>Oulimnius</i>		√	
	<i>Riolus</i>		√	
	<i>Stenelmis</i>		√	√
	<i>Zaitzevia</i>		√	
Ephemerellidae	<i>Serratella</i>	√		
Gammaridae	<i>Echinogammarus</i>			√
	<i>Gammarus</i>		√	√
	<i>Rhipidogammarus</i>		√	
Gomphidae	<i>Ophiogomphus</i>		√	√
Heptageniidae	<i>Ecdyonurus</i>		√	
	<i>Electrogena</i>			√
	<i>Epeorus</i>		√	
	<i>Heptagenia</i>		√	
	<i>Rhithrogena</i>		√	√
Hydraenidae	<i>Aulacochthebius</i>		√	
Hydrobiidae	<i>Hydrobia</i>		√	
	<i>Potamopyrgus</i>		√	

	<i>Pseudamnicola</i>		√	
Hydroptilidae	<i>Agraylea</i>	√		
	<i>Ithytrichia</i>			√
Leptophlebiidae	<i>Choroterpes</i>			√
	<i>Habrophlebia</i>			√
Lumbriculidae	<i>Stylodrilus</i>	√	√	√
	<i>Trichodrilus</i>		√	
Nemouridae	<i>Protonemura</i>		√	√
Nymphalidae	<i>Taygetis</i>	√		√
	<i>Taygetomorpha</i>	√	√	√
Oligoneuridae	<i>Oligoneuriella</i>			√
Philopotamidae	<i>Chimarra</i>		√	√
	<i>Chimarrhodella</i>			√
	<i>Dolophilodes</i>		√	√
	<i>Philopotamus</i>			√
	<i>Wormaldia</i>		√	√
Pholcidae	<i>Leptopholcus</i>		√	√
Rhyacophilidae	<i>Rhyacophila</i>			√
Scirtidae	<i>Contacyphon</i>	√	√	√
	<i>Cyphon</i>		√	√
	<i>Elodes</i>	√		√
	<i>Hydrocyphon</i>	√	√	√
	<i>Microcara</i>	√		√
	<i>Prionocyphon</i>	√	√	√
	<i>Scirtes</i>	√	√	√
Sialidae	<i>Sialis</i>		√	√
Simuliidae	<i>Greniera</i>	√		
	<i>Simulium</i>	√		
Siphonuridae	<i>Siphonurus</i>	√	√	√
Stenopsychidae	<i>Stenopsyche</i>			√
	<i>Stenopsychodes</i>	√		
Tephritidae	<i>Didacus</i>		√	
Tipulidae	<i>Ctenophora</i>	√	√	
	<i>Lunatipula</i>		√	
	<i>Nephrotoma</i>	√	√	
	<i>Pterelachisus</i>		√	
	<i>Tipula</i>	√	√	
Tubificidae	<i>Limnodrilus</i>	√		√
	<i>Potamothrix</i>	√		√
	<i>Psammoryctides</i>	√		√
	<i>Rhyacodrilus</i>	√		√
	<i>Tubifex</i>	√		√

Per quanto riguarda i campioni ambientali di sedimento e di acqua, il numero di sequenze pulite ottenute dal sequenziamento NGS sono risultate notevolmente inferiori rispetto alle altre bulk analizzate (Tab.11). I taxa riconosciuti a livello di genere sono riportati nelle Tab.15 e 16. Il numero di sequenze per i campioni attribuiti, all'interno di entrambe le library, appare proporzionato. Nessun taxon supera le 210 sequenze attribuite, fatta eccezione per Rhithrogena all'interno del campione ambientale di acqua (>2000 sequenze). In quest'ultimo è interessante notare la presenza del genere Cloeon, assente all'interno delle library Efemerotteri 1 e 2.

Tab.15 - Lista dei taxa riconosciuti a livello di genere nel campione ambientale di sedimento

eDNA Sedimento		
Famiglia	Genere	N° Seq Fw+Rev
Chironomidae	<i>Ablabesmyia</i>	1
	<i>Clinotanypus</i>	2
	<i>Conchapelopia</i>	1
	<i>Kiefferulus</i>	1
	<i>Paracladopelma</i>	1
Crangonyctidae	<i>Crangonyx</i>	1
	<i>Sialia</i>	5
Curculionidae	<i>Homarus</i>	8
Elmidae	<i>Elmis</i>	3
	<i>Oulimnius</i>	1
	<i>Stenelmis</i>	2
Ephemerellidae	<i>Serratella</i>	5
Gammaridae	<i>Gammarus</i>	6
Gomphidae	<i>Ophiogomphus</i>	1
Heptageniidae	<i>Ecdyonurus</i>	1
	<i>Rhithrogena</i>	1
Hydraenidae	<i>Limnebius</i>	1
	<i>Corrosella</i>	3
	<i>Pseudammicola</i>	1
Hydropsychidae	<i>Hydropsyche</i>	1
Hydroptilidae	<i>Ithytrichia</i>	3
Lumbriculidae	<i>Stylodrilus</i>	6
Philopotamidae	<i>Chimarra</i>	2
	<i>Wormaldia</i>	1
Scirtidae	<i>Hydrocyphon</i>	1
	<i>Prionocyphon</i>	1
Simuliidae	<i>Simulium</i>	43
Tipulidae	<i>Tipula</i>	8
Tubificidae	<i>Aulodrilus</i>	1
	<i>Limnodrilus</i>	1
	<i>Rhyacodrilus</i>	1

Tab.16 - Lista dei taxa riconosciuti a livello di genere nel campione ambientale di acqua

eDNA Acqua		
Famiglia	Genere	N° Seq Fw+Rev
Baetidae	<i>Baetis</i>	13
	<i>Cloeon</i>	14
	<i>Cloeon dipterum</i>	17
	<i>Cloeon simile</i>	3
	<i>Procloeon</i>	165
Caenidae	<i>Brachycercus</i>	1
	<i>Caenis</i>	8
Chironomidae	<i>Ablabesmyia</i>	95
	<i>Arctopelopia</i>	9
	<i>Clinotanypus</i>	6
	<i>Conchapelopia</i>	27
	<i>Kiefferulus</i>	140
	<i>Macropelopia</i>	15
	<i>Paracladopelma</i>	9
<i>Psectrotanypus</i>	1	
Corydalidae	<i>Corydalidae</i>	5
	<i>Protohermes</i>	3
Crangonyctidae	<i>Sialia</i>	11
Curculionidae	<i>Homarus</i>	7

Dryopidae	<i>Dryops</i>	3
DugesIIDae	<i>Dugesia</i>	2
Elmidae	<i>Elmis</i>	4
	<i>Esolus</i>	1
	<i>Macronychus</i>	10
	<i>Optioservus</i>	1
	<i>Oulimnius</i>	1
	<i>Riolus</i>	54
EphemereIIidae	<i>Serratella</i>	11
Gammaridae	<i>Dikerogammarus</i>	1
	<i>Echinogammarus</i>	50
	<i>Gammarus</i>	66
	<i>Rhipidogammarus</i>	2
Gomphidae	<i>Ophiogomphus</i>	3
Heptageniidae	<i>Ecdyonurus</i>	10
	<i>Heptagenia</i>	4
	<i>Rhithrogena</i>	2097
Hydraenidae	<i>Aulacochthebius</i>	2
	<i>Hydraena</i>	176
	<i>Limnebius</i>	90
	<i>Ochthebius</i>	96
	<i>Bythinella</i>	3
Hydrobiidae	<i>Corrosella</i>	14
	<i>Hydrobia</i>	85
	<i>Islamia</i>	6
	<i>Mercuria</i>	19
	<i>Potamopyrgus</i>	26
	<i>Pseudammicola</i>	9
Hydropsychidae	<i>Hydropsyche</i>	24
Hydroptilidae	<i>Ithytrichia</i>	39
Leptophlebiidae	<i>Habrophlebia</i>	3
Leuctridae	<i>Leuctra</i>	7
Lumbriculidae	<i>Lumbriculus</i>	1
	<i>Stylodrilus</i>	18
Lymnaeidae	<i>Stagnicola</i>	9
Nemouridae	<i>Protonemura</i>	1
Oligoneuridae	<i>Oligoneuriella</i>	10
Philopotamidae	<i>Chimarra</i>	75
	<i>Chimarrhodella</i>	2
	<i>Dolophilodes</i>	63
	<i>Philopotamus</i>	2
	<i>Wormaldia</i>	57
Pholcidae	<i>Leptopholcus</i>	2
Rhyacophilidae	<i>Rhyacophila</i>	39
Scirtidae	<i>Contacyphon</i>	1
	<i>Cyphon</i>	2
	<i>Microcara</i>	1
	<i>Prionocyphon</i>	23
	<i>Scirtes</i>	1
Sialidae	<i>Sialis</i>	45
Simuliidae	<i>Simulium</i>	93
Symbiidae	<i>Symbion</i>	3
Tipulidae	<i>Nephrotoma</i>	5
	<i>Pterelachisus</i>	83
	<i>Tipula</i>	204
	<i>Triplicitipula</i>	1
Tubificidae	<i>Limnodrilus</i>	1
	<i>Potamothrrix</i>	1
	<i>Rhyacodrilus</i>	55
	<i>Tubifex</i>	2

6.2 DNA *barcoding* di *Engraulis encrasicolus* e *Sardina pilchardus*

6.2.1 Dataset molecolare

La strategia di campionamento effettuata ha visto la raccolta di campioni in 10 aree portuali distribuite lungo tutta la costa italiana consentendo di ottenere una collezione di 44 esemplari. Per ogni sito campionato si è cercato di ottenere 3 esemplari di entrambe le specie in esame anche se le differenti disponibilità commerciali per questi prodotti ne ha influenzato il numero prefissato. Sono stati infatti raccolti più esemplari di *Engraulis encrasicolus* rispetto a *Sardina pilchardus* come mostrato in Tab.4.

Il Kit di estrazione utilizzato ha mostrato un'ottima resa: tutti i campioni estratti, sottoposti a quantificazione, hanno registrato una concentrazione di DNA superiore ai 50 ng/μl.

Sono state ottenute sequenze COX1 per il 100% dei campioni analizzati.

Sia i primer *Forward* che *Reverse* hanno manifestato un'ottima resa nel sequenziamento e gli elettroferogrammi ottenuti non hanno mostrato doppi picchi o rumori di fondo permettendo una lettura chiara e precisa della sequenza.

Il multiallineamento generato non ha presentato *gap* e la traduzione proteica effettuata non ha evidenziato codoni di stop consentendo di ottenere un allineamento finale di lunghezza pari a 638 paia di basi.

6.2.2 Correlazione con le banche dati di riferimento (NCBI e BOLD)

La ricerca tassonomica effettuata sulle principali banche dati disponibili sul *web* ha rivelato che entrambe le specie analizzate sono presenti sia su NCBI che su BOLD con le relative sequenze COX1.

Tutte le nostre sequenze hanno ottenuto una correlazione positiva con le rispettive specie. I valori di identità risultati dai “match” sono stati tutti superiori al 99%, mentre i campioni 3, 29, 30 e 31, ottenuti sotto forma di frammenti di tessuto provenienti rispettivamente da pescherie di Civitavecchia (un campione) e Sardegna occidentale (tre campioni), hanno avuto una corrispondenza con *Spicara meana*, *Sardina pilchardus*, e *Sardinella aurita* rispettivamente, anche queste con valori superiori al 99.8% di identità genetica. Queste ultime, inoltre, hanno mostrato distanze genetiche con le due specie studiate comprese tra l'11% ed il 16%.

6.2.3 Analisi bioinformatiche

Gli intervalli di distanze genetiche K2P risultanti dalla matrice delle distanze ha messo in evidenza un'elevata, ma attesa, distanza genetica interspecifica tra le due specie in esame compresa in un intervallo tra il 20.8% e il 27.7% di divergenza genetica.

I livelli di divergenza genetica intraspecifica mostrano invece una variabilità interna differente tra *Engraulis encrasicolus* (range compreso tra 0% e 4.7%) e *Sardina pilchardus* (range compreso tra 0% e 0.5%).

Infine, un'ulteriore informazione sull'applicabilità della tecnica è risultata dalla possibilità di identificare efficacemente due specie (*Sardinella aurita* e *Spicara meana*) erroneamente incluse nel campionamento di *Sardina* e *Engraulis*, grazie alle analisi condotte sui confronti di sequenza e sulle banche dati di riferimento, avvalorando l'utilità della tecnica del DNA *barcoding* anche per fini commerciali.

6.2.4 *Barcoding* ambientale delle acque marine

I risultati ottenuti dal sequenziamento NGS per i campioni marini hanno mostrato un'elevato numero di sequenze per ogni *library* analizzata (Tab.17).

Tab.17 - Numero totale delle sequenze grezze ottenute dal sequenziamento NGS e numero totale delle sequenze pulite utilizzate per l'analisi dei dati nei campioni marini.

Campione	N° sequenze ottenute	N° sequenze pulite
<i>Engraulis encrasicolus</i> 100% (20mg di tessuto)	210420	167460
<i>Sardina pilchardus</i> 100% (20mg di tessuto)	221780	139746
50% <i>Engraulis encrasicolus</i> - 50% <i>Sardina pilchardus</i> (concentrazione DNA estratto)	250240	200823
5% <i>Engraulis encrasicolus</i> - 95% <i>Sardina pilchardus</i> (concentrazione DNA estratto)	91904	72564
eDNA acqua (5L di acqua)	54695	38033
2.5% <i>Engraulis encrasicolus</i> - 97.5% <i>Sardina pilchardus</i> (% in tessuto)	227265	171202
Bianchetto (individui ~ 60 individui)	142114	109182

La *library* che ha ottenuto più sequenze è stata quella in cui è stato inserito il DNA delle due specie analizzate nella stessa concentrazione, mentre la *library* che ha ottenuto meno resa è stata quella in cui il DNA di partenza proveniva dall'estrazione del filtro d'acqua.

Le percentuali di sequenze attribuite dopo la fase di pulizia è mostrata in Tab.18.

Tab.18 - Percentuale di sequenze attribuite nell'analisi delle *library*.

Campione	% di sequenze attribuite
<i>Engraulis encrasicolus</i> 100% (20mg di tessuto)	100% <i>Engraulis encrasicolus</i>
<i>Sardina pilchardus</i> 100% (20mg di tessuto)	98.2% <i>Sardina pilchardus</i> - 1.2% <i>Engraulis encrasicolus</i>
50% <i>Engraulis encrasicolus</i> - 50% <i>Sardina pilchardus</i> (concentrazione DNA estratto)	99% di <i>Engraulis encrasicolus</i> - 1% di <i>Sardina pilchardus</i>
5% <i>Engraulis encrasicolus</i> - 95% <i>Sardina pilchardus</i> (concentrazione DNA estratto)	82% di <i>Engraulis encrasicolus</i> - 18% di <i>Sardina pilchardus</i>
eDNA acqua (5L di acqua)	70% di <i>Engraulis encrasicolus</i> - 30% di <i>Sardina pilchardus</i>
2.5% <i>Engraulis encrasicolus</i> - 97.5% <i>Sardina pilchardus</i> (% in tessuto)	82% di <i>Engraulis encrasicolus</i> - 18% di <i>Sardina pilchardus</i>
Bianchetto (individui ~ 60 individui)	89% di <i>Engraulis encrasicolus</i> - 11% di <i>Sardina pilchardus</i>

Nel dettaglio si può notare che le *library* contenenti *Engraulis* e *Sardina* al 100% hanno mostrato una congruenza tra il risultato atteso ed il risultato osservato, la percentuale minima (1.2%) di *Engraulis* all'interno del campione ambientale di *sardina* può essere dovuta ad una leggera contaminazione tra le due specie quando sono state raccolte o dal rumore di fondo prodotto durante la corsa.

L'osservazione delle percentuali di attribuzione ottenute quando sono state testate differenti percentuali delle due componenti evidenziano un forte bias/distorsione che sembra favorire fortemente l'amplificazione di *Engraulis* rispetto a *Sardina*.

Il risultato ottenuto dal sequenziamento del DNA estratto dal filtro dell'acqua mostra un'evidente presenza delle due specie nel tratto di mare testato, un totale di più di 38000 sequenze attribuite per un 70% ad *Engraulis* ed un 30% a *Sardina*. Questo evidenzia la possibilità di ottenere una percentuale di presenza delle due specie anche dal campione ambientale.

Allo stesso modo le attribuzioni delle sequenze ottenute per il campione di bianchetto analizzato mostrano un'elevata presenza di *Engraulis* (89%) rispetto a *Sardina* (11%).

7. DISCUSSIONI

7.1 Potenzialità del DNA *barcoding* nella discriminazione delle specie italiane di Efemerotteri

Sequenze specifiche e delimitazione della diversità intra- ed interspecie sono le condizioni essenziali per una corretta identificazione a livello specifico attraverso il DNA *barcoding*. In questo studio, la valutazione dell'efficacia di tale tecnica per verificare la corretta discriminazione delle specie è stata basata su quattro parametri:

a) Lista degli aplotipi (l'unicità delle singole sequenze permette la discriminazione dei campioni e l'attribuzione degli esemplari a differenti specie);

b) DNA *barcoding gap* (l'assenza di sovrapposizione tra le distanze genetiche intra- ed interspecifiche all'interno di ciascun genere);

c) Stima della monofilia tra i raggruppamenti di specie secondo differenti criteri;

d) Verifica della giusta correlazione sequenza/specie sulle banche dati pubbliche disponibili (NCBI, BOLD);

I nostri risultati indicano che il sistema di discriminazione degli Efemerotteri basata sulle sequenze COX1 nella presente area di studio ha mostrato una notevole efficacia nell'individuazione di aplotipi specifici all'interno di tutte le specie studiate.

Tutti i taxa del nostro studio (100%) sono stati discriminati basandosi esclusivamente sullo stato dei caratteri unici delle sequenze. Gli aplotipi generati con una riduzione della soglia di identità da 100% a 98% (ovvero la soglia standard di delimitazione delle specie proposta da Hebert et al. 2003) non hanno presentato in nessun caso una condivisione di sequenza con specie differenti. Allo stesso tempo la corrispondenza delle identità sulle banche dati di riferimento disponibili (ca. il 100% dei generi ed il 90% delle specie hanno evidenziato un'estatta corrispondenza) e la rappresentazione del dendrogramma Neighbor Joining (nella quale tutte le sequenze *barcode* si raggruppano in specifici cluster in accordo con la tassonomia) forniscono ulteriori conferme sull'efficacia del frammento indagato nella discriminazione degli esemplari con un'ampia congruenza rispetto alla nomenclatura delle specie.

Degni di nota sono i casi di (a) *Habroleptoides modesta*, in cui il DNA *barcoding* è stato determinante nella identificazione della specie poichè il campione danneggiato non permetteva

una sicura analisi morfologica e (b) della netta separazione evidenziata tra *Baetis albinatii* e *B. muticus* malgrado i pochi e spesso ambigui tratti diagnostici disponibili (ad esempio l'assenza della prima tracheobranchia; Sartori & Thomas 1989).

Il dendrogramma di Neighbor Joining ottenuto dalle sequenze COX1 ha evidenziato anche una buona congruenza con la tassonomia tradizionale di questo ordine nonostante siano state riscontrate alcune discrepanze a livelli tassonomici più elevati (ovvero, diversi casi di parafilia per alcuni genere e famiglie). Tali discrepanze possono senz'altro essere dovute al campionamento incompleto, all'errato utilizzo del Neighbor Joining come strumento analitico e al modello evolutivo (K2P) utilizzato. Questi ultimi due sono comunque considerati degli standard per il DNA *barcoding* (Hebert et al. 2003, 2004; Zhou et al. 2009, 2011) ma probabilmente non ottimali ai fini della risoluzione (Collins & Cruickshank 2013; White et al. 2014). Ad esempio, un'analisi parallela effettuata sul nostro dataset con RaxML e basata sul modello evolutivo GTRCAT ha mostrato una topologia altamente supportata e molto congruente con le maggiori classificazioni tassonomiche (Appendice I).

Allo stesso tempo il test sulla reciproca monofilia di Rosenberg, particolarmente utile per assegnare una valutazione al regime di campionamento utilizzato in relazione alla formazione delle distinte linee evolutive dei gruppi (Rosenberg 2007), ha indicato che la monofilia dei gruppi di specie prodotti è stata probabilmente dovuta al caso, suggerendo ulteriori campionamenti per la valutazione di questo importante (ma non fondamentale; vedi Collins & Cruickshank 2013) requisito nella discriminazione delle specie attraverso la tecnica del DNA *barcoding*.

In accordo con quanto appena affermato è la ricostruzione di Maximum likelihood effettuata con il dataset ampliato di sequenze *barcode* provenienti da BOLD dove tutte le nostre sequenze si posizionano in una rappresentazione ben supportata e coerente con la tassonomia di questo ordine. È interessante notare che sia le nostre sequenze che quelle provenienti da BOLD contribuiscono allo stesso modo nell'evidenziare alcuni casi di parafilia in alcuni generi ed in particolare nella famiglia Baetidae (Fig.30).

Tutti questi dati contribuiscono certamente a sostenere l'efficacia del *barcoding* nell'area di studio per (1) la discriminazione dei taxa, (2) la distinzione dei principali schemi tassonomici degli Efemerotteri, come affermato anche in precedenti studi (Ball et al. 2005; Zhou et al. 2011; Webb et al. 2012), ed (3) per evidenziare ulteriormente quei generi o specie per i quali è necessario ampliare il campionamento inter-/intraspecifico al fine di ottenere risultati statisticamente validi ed affidabili.

Non è da escludere, infatti, che gli elevati livelli di distanza genetica intra-/interspecifici registrati ed i casi di non monofilia riscontrati nel nostro dataset possano essere stati influenzati dal limitato campionamento intra-/interspecifico e dalla limitata area di studio saggiata (Meyer & Paulay 2005; Wiemers & Fiedler 2007). Bergsten et al. (2012) hanno dimostrato come la variazione genetica intraspecifica in genere aumenti con la scala spaziale, e la divergenza interspecifica diminuisca con l'ampliamento dei campionamenti. Ciò significa che l'incertezza di identificazione aumenta con la copertura geografica e di campionamento. Questo aspetto è certamente rilevante per gli studi tassonomici basati su vaste iniziative di DNA *barcoding* con l'obiettivo di rilevare una quantità rappresentativa della variabilità genetica delle specie esaminate, per descrivere la diversità esistente e agevolare la delimitazione delle specie basata sul DNA in un contesto globale (www.ibol.org).

Tuttavia, per molte applicazioni riguardanti il DNA *barcoding* come l'associazione ad un genere o ad un particolare stadio biologico di vita oppure indagini sulla biodiversità e sul monitoraggio ambientale, il campionamento e l'ampiezza dell'area studiata sono di limitata importanza in quanto tali studi si affidano in maniera più diretta alla disponibilità delle *library* di riferimento locale (Bergsten et al. 2012). Questo problema dà un incentivo alle iniziative di *barcoding* regionali e nazionali, in particolare per le aree poco studiate, generando sequenze di riferimento delle relative raccolte di specie locali, indipendentemente dai numeri del campionamento ed anche assegnando una speciale rilevanza ai campioni singoli (Lim et al. 2012). In questo contesto è importante notare che la maggior parte delle corrette assegnazioni sulle banche dati genetiche hanno riscontrato un valore di identità genetica superiore al 90%, ma sono stati anche registrati casi di errata correlazione con le specie.

Oltre alla ovvia possibilità di un'accidentale errata identificazione del campione (sia nel nostro set di dati che nelle banche dati stesse), riteniamo che entrambi questi risultati possono essere fortemente correlati con la generale disponibilità di campioni conspecifici appartenenti a zone fuori continente disponibili nelle banche dati consultate. Quest'ultimo fenomeno richiede certamente un rapido arricchimento di campioni nei database Euro-Mediterranei e contemporaneamente lo sviluppo di librerie regionali.

Infatti in questo contesto il lavoro presentato tenta di sviluppare una prima banca dati regionale e fornisce ulteriori test multi-metodologici per cercare di raggiungere, nella pratica applicativa, la massima precisione nella richiesta di identificazione o delimitazione delle specie.

7.2 Tassonomia basta su DNA: raggruppamenti *barcode* vs. specie morfologiche

Dalle analisi ricavate sulla base della distanza genetica > del 2% è stato osservato un numero maggiore di unità tassonomiche operative (MOTUs, Jones et al. 2011) rispetto a quelle riconosciute morfologicamente. Tale soglia, considerata standard per l'identificazione delle specie con il DNA *barcoding* nella classe degli insetti, suggerisce che una distanza genetica maggiore del 2% potrebbe indicare l'esistenza di specie criptiche (Hebert et al. 2004; Ball et al. 2005; Zhou et al. 2009; Park et al. 2011; Smith et al. 2013; Jackson et al. 2014).

In questo studio, divergenze intraspecifiche che eccedono il 2% di distanza genetica sono state riscontrate nel 65% delle specie in esame, con *Procloeon bifidum*, *Baetis cyrneus* ed *Epeorus assimilis* che registrano valori superiori al 10%.

E' interessante notare come un numero maggiore di specie putative è stato evidenziato anche dall'analisi con il metodo PTP, dove oltre a *Baetis cyrneus* e *Procloeon bifidum* (separati su differenti microcluster), anche *Epeorus assimilis*, *Baetis alpinus*, *Procloeon pulchrum* e *Siphonurus lacustris* sono stati separati in raggruppamenti di specie diversi.

Precedenti studi su questo ordine (Ball et al. 2005; Zhou et al. 2011; Webb et al. 2012) e su altri gruppi di insetti come Lepidotteri, Tricotteri, Imenotteri e Ditteri (Hebert et al. 2004, Smith et al. 2006, 2008, Zhou et al., 2009) hanno dimostrato che divergenze genetiche di una certa rilevanza regolarmente riflettono complessi di specie o specie non ancora riconosciute. La medesima conclusione potrebbe essere tratta per alcuni dei taxa indagati in questo studio, anche se fattori come ad esempio la limitata scala geografica e di campionamento (Bergsten et al. 2012) o una particolare storia evolutiva delle popolazioni studiate (Theissinger et al. 2011) dovrebbero essere ulteriormente considerati ed approfonditi. Per esempio, l'incertezza tassonomica sullo status di alcuni taxa come il genere *Procloeon* e, più in generale, la controversa tassonomia di tutti i generi appartenenti alla sottofamiglia "Cloeoninae" sono ben noti (Gillies 1991; Kluge 1997) e richiederebbero chiaramente una critica revisione tassonomica.

In questo lavoro, l'assenza del *barcoding gap* all'interno del genere *Procloeon* è determinato da due campioni di *P. bifidum* provenienti dalla Sardegna e dall'isola d'Elba che condividono una distanza genetica K2P estremamente elevata rispetto al campione proveniente dall'Italia centrale (19.6%), più alta anche di quella condivisa con un'altra specie appartenente allo stesso genere, *P. pulchrum* (18.7%). Questo suggerisce che *P. bifidum*, nonostante l'apparente uniformità dei caratteri morfologici, potrebbe includere specie criptiche endemiche della regione biogeografica Sardo-Corsa.

Stesso discorso potrebbe essere effettuato nei confronti di un altro endemismo Sardo-Corso (*B. cyrneus*) in cui una notevole divergenza genetica rilevata (14.9%) richiederebbe ulteriori studi di approfondimento per verificare l'eventuale presenza di variazioni criptiche come recentemente suggerito anche da Gattolliat et al. (2015) all'interno del gruppo *Baetis alpinus* (al quale appartiene *B. cyrneus*). E' importante anche sottolineare che il genere *Baetis* è particolarmente variabile (Ball et al. 2005; Zhou et al. 2009; Webb et al. 2012) e la presenza di specie criptiche al suo interno è già stata evidenziata con studi specifici (e.g., *Baetis rhodani*; Williams et al. 2006; Bisconti et al. 2016).

Il terzo esempio, in cui è stata registrata un'anomala distanza genetica intraspecifica superiore al 10% (*Epeorus assimilis*), è stato completamente inaspettato in quanto in contrasto con le caratteristiche morfologiche uniformi visualizzate nei diversi individui analizzati e la relative prossimità dei siti di campionamento. Questo risultato suggerirebbe quindi un esame più attento della morfologia della specie e/o dei tratti ecologici delle popolazioni o dei fiumi per valutare con maggiori informazioni la variabilità riscontrata.

Congruentemente, l'ampia parafilia e l'elevata presenza di falsi positivi e negativi risultati dall'analisi con Spider indicano che i generi *Procloeon* e *Baetis* sono quelli con la presenza più elevata di specie aventi status critici (es. specie criptiche o complessi di specie).

Allo stesso modo è importante notare anche l'elevato valore di *barcoding gap* riscontrato tra le specie di *Electrogena* (23.5-24.0%). Per tutti gli altri generi nel dataset (esclusi *Procloeon* e *Baetis*) Spider fornisce un valore ottimale di *barcoding gap* ad una soglia dell'11%.

Livelli di *barcoding gap* intragenerici superiori al 20% negli Efemerotteri sono stati solo raramente descritti in precedenti lavori (Ball et al. 2005; Zhou et al. 2011; Webb et al. 2012), suggerendo che *Electrogena grandiae* ed *E. zebrata* potrebbero appartenere a generi differenti, come precedentemente sostenuto da Gaino et al (1987) e discusso da Bauerfeind & Soldan (2012). Questa separazione sarebbe anche supportata sia dal dendrogramma NJ generato, in cui le specie investigate sono raggruppate in differenti lignaggi, sia dalla ricostruzione ML in cui *E. zebrata* si posiziona separatamente rispetto alle altre specie del genere. A supporto di questa tesi, anche i caratteri morfologici delle uova e degli stadi alati di *E. zebrata* (endemismo Sardo-Corso) mostrano nette differenze rispetto a quelli delle altre specie del medesimo genere (Gaino et al. 1987).

Su queste basi, e se confermati da un campionamento più ampio, i nostri dati suggerirebbero una revisione tassonomica degli Efemerotteri nella presente area di studio in modo da ottenere un'adeguata correlazione tra le linee evolutive ed i tratti morfologici ed ecologici. In aggiunta, l'utilità di un'unica, e arbitraria, soglia percentuale di discontinuità genetica per stabilire una

correlazione al concetto di specie in questo ordine deve essere certamente rivalutata, dopo aver attentamente determinato il più appropriato *range* di distanza genetica per la discriminazione delle specie in ogni famiglia e genere (cf. Krishnamurthy & Francis 2012; Collins & Cruickshank 2013).

In ogni caso, un generale miglioramento della tassonomia degli Efemerotteri dovrebbe essere necessariamente perseguito attraverso un approccio multidisciplinare in cui, la scoperta di nuove specie (o specie criptiche) e le revisioni nomenclaturali suggerite attraverso gli aplogruppi *barcode* altamente divergenti saranno corredate da prove aggiuntive come quelle morfologiche, ecologiche, descrittori biogeografici e marcatori molecolari indipendenti (Monaghan & Satori 2009; Jackson et al. 2014). Questo è vero anche per i vari esempi di parafilìa nei generi o famiglie e per le incongruenze risultanti nel dendrogramma ML in accordo con la continua revisione tassonomica che sta avvenendo negli ultimi anni su scala globale per questo ordine. In aggiunta, oltre alla sinonimia (Meyer & Paulay 2005) o l'imperfetta tassonomia dei gruppi (Hendrich et al. 2010) ci sono altre potenziali ragioni per l'assenza di monofilia riscontrata in alcuni taxa come l'incompleta suddivisione nei lignaggi o l'ibridazione introgressiva (Funk & Homland 2003; Zhang et al. 2012).

Come risultato principale di questa fase dello studio è pertanto evidente la necessità di un ampliamento del campionamento e di un approfondimento tassonomico dei gruppi discussi, in modo da trarre conclusioni più solide circa il loro stato e la loro biodiversità; allo stesso tempo (vedi Packer et al. 2009), i dati raccolti suggeriscono modelli di biodiversità locali più ampi per quanto riguarda questo gruppo di insetti e la necessità di una rivalutazione tassonomica di alcune specie con una migliore integrazione di dati morfologici e molecolari.

In questo senso è fondamentale una ottimizzazione delle linee guida operazionali (un adeguato campionamento, banche dati di riferimento dettagliate, strumenti statistici robusti e/o complementari) per assicurare la necessaria accuratezza nelle future applicazioni pratiche, dove un genere o una specie male identificata può vanificare gli sforzi di biomonitoraggio o biovalutazione.

7.3 Potenzialità del metabarcoding nelle acque dolci

7.3.1 library Efemerotteri

L'efficienza della tecnica *NGS* per le due comunità artificiali degli Efemerotteri (1 e 2) ha mostrato una percentuale di identificazione degli aplotipi pari al 73% ed al 56% rispettivamente (Tab.12 a e b), corrispondenti al 65% ed al 54% delle specie studiate, rispettivamente. Questa differenza tra i risultati delle due library, a favore della library Efemerotteri_1 evidenzia come le diverse concentrazioni iniziali di DNA dei diversi campioni incidano sulle successive fasi del lavoro (PCR e sequenziamento).

La percentuali totali (Forward e Reverse) ottenute per queste due *library* (Tab.12 c) mostrano inoltre un quadro generale del successo nelle varie famiglie di Efemerotteri. Appare evidente come la famiglia Baetidae risulti avere la percentuale più bassa di attribuzioni tassonomiche (38.5% rispetto a 80% delle altre famiglie). I generi *Cloeon* e *Procloeon* risultano quasi assenti nel sequenziamento per le due *library*, il primo con entrambe le specie *C. dipterum* e *C. simile*, il secondo presente solo con la specie *P. bifidum* (1 sequenza per il forward ed 1 per il reverse). Il genere *Baetis* è presente con 3 specie (*B. alpinus*, *B. cyrneus*, *B. vernus*) su 8. Questi risultati sono in accordo con precedenti studi di metabarcoding effettuati utilizzando primers universali degenerati, i quali hanno evidenziato una difficoltà nel sequenziamento del genere *Baetis* (Elbrecht & Leese 2015); la stessa difficoltà è stata riscontrata anche utilizzando le parti più interne del gene COXI, mediante un approccio “minibarcode” (150bp circa) (Hajibabaei et al. 2011), mentre non sono noti in letteratura studi di *barcoding* ambientale che hanno analizzato i generi *Cloeon* e *Procloeon*.

L'elevata diversità genetica presente all'interno della famiglia, già discussa nella prima parte di questo studio, potrebbe spiegare una elevata divergenza del locus COXI ed una conseguente minore affinità dei primers per tali organismi, i quali risulterebbero quindi svantaggiati durante la fase di amplificazione in presenza di altri campioni con maggiore compatibilità per i primers utilizzati. Ad esempio, è interessante notare come *Brachycercus harrisella* abbia prodotto un numero di sequenze estremamente elevato (31.135 sequenze forward e 18.776 sequenze reverse) sottolineando una probabile maggiore affinità per i primers utilizzati.

Anche nelle altre famiglie analizzate, le specie che hanno evidenziato un più alto fallimento di sequenziamento (*Caenis* sp., *Ephemera danica*, *Electrogena grandiae*, *Rhithrogena insularis*, *Habroleptoides* sp., *Oligoneuriella rhenana*) corrispondono ai taxa per i quali erano state riscontrate alcune criticità nella fase di amplificazione della prima parte dello studio, per le

quali una sola coppia di primers (LEP o LCO/HCO) aveva mostrato di funzionare efficacemente, probabilmente a causa di uno o più mismatch nelle sequenze di appaiamento.

Queste diverse affinità possono aver creato uno sbilanciamento nella fase di amplificazione del campione avvantaggiando gli organismi con affinità più alta per i primers degenerati utilizzati. . L'utilizzo di mix di primer universali può infatti portare a differenti efficienze di amplificazioni tra gli organismi (Polz & Cavanaugh 1998; Piñol et al. 2014) evidenziando la necessità di valutare opportunamente le temperature di annealing (Shokralla et al. 2012) e disegnare primer specifici per i gruppi tassonomici più difficili, al fine di bilanciare i risultati ottenuti e permettere una più corretta quantificazione delle specie (Gibson et al. 2014).

7.3.2 Library Bulk e campioni ambientali di sedimento ed acqua

Le *library* prodotte in questo studio possono essere definite matrici complesse, in quanto rappresentative delle comunità naturali realmente presenti nei luoghi di campionamento. Gli organismi analizzati inoltre non sono stati suddivisi in base alle taglia (come, ad esempio, fatto da Elbrecht & Leese, 2015, al fine di massimizzare le efficienze di sequenziamento) ma unicamente in base ai siti di raccolta.

La comparazione tra le sequenze *NGS* ottenute e l'identificazione morfologica mettono in evidenza un successo del 65% nelle bulk 1 e 3 e del 58% nella bulk 2, evidenziando problematiche di efficienza dei primers e (probabilmente) legate anche alle dimensioni relative dei campioni, riscontrate nelle *library* dedicate agli Efemerotteri. Allo stesso tempo è stata messa in evidenza la presenza di ulteriori taxa non evidenziati morfologicamente, spesso con un numero di sequenze attribuite molto elevato.

In dettaglio, è interessante notare come nella bulk 1 (fiume Ortolano), la famiglia Hydropsychidae ha ottenuto più di 20.000 sequenze (sia per il forward che per il reverse) rispetto alle circa 29.000 totali (sia per il forward che per il reverse) ottenute per tutta la *library*. Tale taxon è infatti rappresentata da 18 individui tra i più grandi rinvenuti per questa stazione. Il predominio di questi organismi durante la fase di amplificazione potrebbe aver generato una sottostima degli altri taxa che hanno ottenuto un numero di sequenze notevolmente inferiore (ad es. Gammaridae, 15 individui di media grandezza con circa 300 sequenze, sia per il forward che per il reverse, oppure Helodidae, 21 individui di piccole dimensioni con nessuna sequenza ottenuta).

Nella bulk 2 (Fiume Traponzo), invece, la famiglia Simuliidae (rappresentata da 11 individui di media grandezza) ha ottenuto più di 24.000 sequenze sulle 37.000 totali della *library*. Anche in questo caso la dominanza nella fase di amplificazione del taxon potrebbe aver generato una

sottostima degli altri taxa (ad es. Hydropsychidae, presente con 59 individui di medie dimensioni ha ottenuto solo 93 sequenze per il forward e 53 per il reverse; Gomphidae, presente con 12 individui di grandi dimensioni ha ottenuta solo 1 sequenza con il reverse e nessuna per il forward).

Un caso particolare è stato riscontrato per la bulk 3 (Fiume Biedano) dove i taxa che hanno ottenuto circa il 90% delle sequenze sono state le famiglie Philopotamidae (più di 36.000 sequenze su circa 60.000) e la famiglia Heptageniidae (con il genere *Rhithrogena*, più di 18.000 sequenze su 60.000) entrambe non evidenziati morfologicamente. Anche per questa *library* la prevalenza di queste due famiglie nella fasi di amplificazione potrebbe aver portato ad una distorsione nei risultati in termini di quantificazione delle sequenze attribuite ai diversi taxa (es. Hydrobiidae, rappresentato da 16 individui di medie dimensioni che ha ottenuto 9 sequenze forward e 37 reverse, oppure Baetidae, presente con 37 individui di *Baetis fuscatus* di medie dimensioni che non ha ottenuto nessuna sequenza).

Per tutte e tre i siti di campionamento quindi le famiglie che hanno ottenuto il più alto numero di sequenze attribuite non corrispondono alle famiglie con il più alto numero di individui riconosciuti morfologicamente.

La discrepanza quantitativa tra l'approccio morfologico e molecolare è dovuta probabilmente a fattori differenti nelle 3 bulk. Nella prima, la distorsione sembra essere dovuta alla biomassa della famiglia Hydropsychidae e di conseguenza alla sua maggior presenza di DNA all'interno della library che potrebbe aver favorito la sua amplificazione, generando uno sbilanciamento nel numero di sequenze amplificate a favore di questo taxon. Nella seconda, la maggiore affinità dei primer per la famiglia Simuliidae, unita alla sua biomassa relativamente abbondante, potrebbe aver causato lo stesso effetto riscontrato nella Bulk precedente. Nella terza, invece, tale sbilanciamento ha riguardato taxa assenti; malgrado essi non siano stati trovati durante il riconoscimento morfologico, l'elevato numero di sequenze ad essi attribuito esclude la possibilità di un artefatto. La possibile presenza di questi organismi potrebbe essere dovuta infatti a frammenti di tessuti o uova raccolti inavvertitamente durante la fase di campionamento, mentre, il predominante numero di sequenze attribuite potrebbe essere dovuto esclusivamente alla migliore affinità dei primers per le famiglie in questione.

In generale i risultati ottenuti con le bulk sottolineano sia le potenzialità dell'approccio NGS sia la necessità di ulteriori esperimenti. Un interessante spunto è quello di Elbrecht & Leese (2015), nel quale la comunità macrobentonica viene frazionata sulla base della classe di taglia degli individui (piccole, medie, grandi). Inoltre, al fine di aumentare la relativa efficienza di amplificazione, potrebbero essere testati nuovi primers, oppure potrebbero esserne sviluppati

di nuovi, specifici per i gruppi più problematici (ad es., Baetidae). Dopo averne accertata l'efficienza, queste diverse coppie di primers potrebbero essere utilizzate in contemporanea nei diversi esperimenti *NGS*; in uno studio recente (Dowle et al. 2015) è stata testata la possibilità di un approccio che non richieda lo step di PCR e primer prima del sequenziamento *NGS*. Tale approccio, denominato *Gene Enrichment*, ha mostrato risultati interessanti e potrebbe costituire un protocollo valido nello studio di comunità target già ben definite come quelle dei biomonitoraggi.

E' interessante notare come i risultati della library relativa al campione ambientale d'acqua (eDNA) (recuperata nella stessa stazione di campionamento della bulk 1) hanno mostrato un'attribuzione delle sequenze in grado di rilevare 32 taxa (Tab.16), nonostante le concentrazioni iniziali del DNA estratto fossero molto basse (circa 7ng/μl) le sequenze hanno individuato 32 famiglie e 79 generi, corrispondenti a 21 dei 23 taxa riconosciuti morfologicamente nella medesima stazione ottenendo una percentuale di successo di oltre il 92% rispetto al 68% ottenuto con la bulk 1. Anche in questo caso i numeri delle sequenze attribuiti sottolineano la prevalenza di un genere (nello specifico *Rhithrogena*, Famiglia *Heptageniidae*) rispetto agli altri, ottenendo il maggior numero di sequenze (2.097 su circa 2800 tra forward e reverse).

Solamente due taxa, la cui presenza era stata effettivamente riscontrata nel sito di campionamento della Bulk 1, non sono stati riscontrati con l'analisi del campione d'acqua. Questi appartengono alla famiglia *Athericidae* (che anche nella bulk 1 aveva mostrato un numero basso di sequenze =25) ed alla famiglia *Ephemeridae* (*Ephemera danica*). Quest'ultimo taxon aveva già mostrato una notevole difficoltà di amplificazione per la generazione della sequenza barcode reference, come discusso precedentemente. Anche in questo caso, probabilmente, una bassa affinità con i primers e la diversa temperatura di annealing utilizzata durante la fase *NGS* (55° rispetto ai 53° utilizzati nella prima parte del lavoro) possono aver contribuito alla bassa resa di sequenziamento. Nonostante questo, all'interno dei dati grezzi sono presenti sequenze attribuibili a *Ephemera danica* con similarità superiore al 95% ma non sono state considerate ai fini dell'attribuzione in quanto troppo corte (<50 bp).

Confrontando le due library (bulk 1 e filtrato acqua) è di notevole interesse notare come i taxa attribuiti in quella del filtrato sono identici a quelli attribuiti nella bulk 1 con l'aggiunta di 9 taxa (famiglie) in più. Tuttavia, è anche importante sottolineare che 5 generi (appartenenti a 3 famiglie: *Anisogammaridae*, *Nymphalidae*, *Siphonuridae*) sono presenti nella bulk 1 ma non nelle sequenze del filtrato. Come per *Ephemera danica*, anche in questo caso nei dati grezzi è possibile rintracciare 1 sequenza (solo per il forward) per tutti e tre le famiglie; tuttavia, queste

non superano l'88% di similarità ed hanno una lunghezza di circa 120 bp e pertanto sono stati considerati artefatti.

Seppur condotto su un numero molto basso di taxa bersaglio (*Baetis buceratus*, *Ancyclus fluviatis* e *Gammarus pulex*) e mediante primers specifici, l'unico studio precedente su campioni ambientali di acqua di fiume ha dimostrato che esiste una correlazione tra volume di acqua analizzato e il rilevamento dei taxa (Mächler et al. 2016). Quest'ultima tende infatti ad aumentare con bassi volumi indagato per alcuni taxa, mentre per altri taxa esiste una correlazione opposta. Tuttavia, per ottenere una concentrazione di eDNA tale da riuscire a rilevare il maggior numero di taxa bersaglio il volume minimo di acqua campionata non dovrebbe mai essere minore 1L.

E' anche importante sottolineare che la corrente dei fiumi potrebbe influenzare i dati ottenuti sovrastimando la diversità biologica della stazione campionata. Infatti la capacità di drift (trasporto verso valle) della corrente d'acqua può trasferire organismi, parti di essi o DNA extracellulare presenti a monte del sito di campionamento, producendo falsi positivi durante la fase di elaborazione dati.

Su queste basi, anche l'esito di quest'ultima library sottolinea la necessità di effettuare ulteriori esperimenti, valutando differenti volumi di acqua analizzata e più siti sottoposti a diverse velocità di corrente, in modo da approfondire la capacità di rilevamento rispetto alle comunità realmente presenti. Studi su questo argomento sono stati recentemente effettuati da Deiner et al. (2014) che hanno monitorato la presenza di due taxa (un pesce e un anfibio) in un fiume naturale attraverso l'eDNA. Le conclusioni a cui sono giunti mettono in evidenza la potenzialità di trasporto dei fiumi, avendo rilevato tracce del DNA delle due specie ad una distanza compresa tra i 5 ed i 9 km dal sito di rilevamento della comunità. Altri autori suggeriscono che la persistenza del eDNA possa dipendere da molteplici fattori quali le caratteristiche del fiume, la sua posizione geografica ed il periodo di campionamento, oltre che da fattori già noti quali la presenza di nucleasi, il livello di ossigenazione dell'acqua, l'irraggiamento solare e l'azione di batteri e funghi (Shapiro 2008).

Inoltre, anche in questo caso andrebbero testati nuovi primers in base alla diversa affinità che può essere mostrata dai differenti organismi.

Lo stesso tipo di confronto è stato effettuato tra la bulk 1 ed il campione ambientale di sedimento (anch'esso recuperato nella medesima stazione). La percentuale di attribuzioni delle sequenze per questa *library* rispetto ai taxa morfologicamente riconosciuti è stata di 13 taxa su 23, corrispondente al 56%. Il taxon che prevale rispetto agli altri è la famiglia Simuliidae con 43 sequenze ottenute sulle circa 122 totali (forward più reverse) mentre gli altri si attestano su

un numero di sequenze attribuite inferiore a 10. Anche in questo caso è possibile notare l'attribuzione di 11 famiglie in più rispetto al dato morfologico. Non vi sono studi in letteratura che possono essere utilizzati come riferimento, in quanto esperimenti di questo tipo per le comunità macrobentoniche dei fiumi non sono stati finora tentati. Tuttavia, la minore resa di questo esperimento è sicuramente dovuta al basso quantitativo di DNA estratto dal campione (la concentrazione in DNA per questa library era infatti la più bassa). Questo è dipeso principalmente dal limitato quantitativo di materiale inerte utilizzato durante la fase di estrazione. L'indicazione principale di tale risultato, oltre agli accorgimenti discussi fin qui discussi, riguarda sicuramente la necessità di utilizzare una maggiore quantità di materiale di partenza.

Infine, bisogna notare che la mancata presenza di alcuni taxa nel confronto tra riconoscimento morfologico e molecolare potrebbe dipendere anche dall'inadeguatezza degli attuali database di riferimento globali per gli organismi analizzati, le quali possono aver "fallito" la corretta attribuzione delle sequenze prodotte per motivi legati alla mancanza di sequenze per tutti i taxa studiati, alla mancanza di sequenze di riferimento locali (cioè italiane o europee), nonché ad eventuali errate identificazioni morfologiche per i campioni presenti nei data base.

E' importante sottolineare, infatti, come l'identificazione tassonomica legata ad approcci molecolari di questo tipo è strettamente legata a differenti fattori, tra i quali: 1) le banche dati di riferimento in termini di qualità delle sequenze reference, di affidabilità nel riconoscimento tassonomico degli esperti, di copertura geografica e tassonomica; 2) i marcatori utilizzati in termini di potere discriminante e di affinità con gli organismi analizzati; 3) dalle procedure bioinformatiche e di laboratorio che possono includere distorsioni o sbilanciamenti dovuti ad errore umano (Haase et al. 2010, Kermarrec et al. 2013). Infine, da un punto di vista strettamente tecnico, è essenziale porre l'attenzione sul fatto che questo tipo di problemi sono comuni in tutti gli studi di metabarcoding e che dipendono principalmente dallo step di PCR e non dall'efficienza della tecnica *NGS*.

7.4 Potenzialità del metabarcoding nelle acque marine

7.4.1 *Library* campioni ambientali marini (*Engraulis encrasicolus* e *Sardina pilchardus*)

I risultati ottenuti per le 7 *library* di campioni ambientali marini hanno mostrato una notevole efficienza di amplificazione per i primer degenerati utilizzati (Tab.17). Il numero di sequenze prodotte, infatti, è risultato notevolmente maggiore rispetto alle *library* dedicate alle acque dolci.

Gli esperimenti condotti separatamente sul DNA estratto dai tessuti prelevati dalle due specie (20 mg), ha effettivamente evidenziato la presenza esclusiva dei taxa indagati. Il campione ambientale di *Engraulis encrasicolus* (esclusivo al 100%) ha generato più di 167.000 sequenze (pulite) attribuite alla specie in oggetto evidenziando la capacità specifica dei primer. Il campione ambientale di *Sardina pilchardus* ha prodotto invece più di 139.000 sequenze (pulite) attribuite per il 98.8% al taxa esaminato e per il 1.2% ad *Engraulis encrasicolus*. Tale alterazione dei risultati potrebbe essere dovuta ad una contaminazione avvenuta durante la fase di raccolta del materiale. Gli individui, infatti, collocati indistintamente nelle pescherie o nei pescherecci dai quali sono stati recuperati, potrebbero essere stati a contatto per un tempo sufficientemente lungo da aver permesso lo scambio di secrezioni, squame o tessuti. In aggiunta, un'affinità maggiore di *Engraulis encrasicolus* ai primer utilizzati giustificerebbe il numero di sequenze riscontrate con la presenza anche minima della specie.

E' interessante notare come nelle *library* in cui il DNA dei due taxa è stato combinato in differenti percentuali, i risultati ottenuti hanno messo in evidenza sbilanciamenti (nel numero di sequenze attribuite) sempre a favore di una delle due specie (*Engraulis encrasicolus*).

Quando le percentuali in DNA estratto sono state inserite nella medesima proporzione (50% e 50%) le sequenze ottenute (circa 200.000 sequenze pulite) hanno avuto un'attribuzione per il 99% a *Engraulis encrasicolus* e per l'1% a *Sardina pilchardus*. Questa dominanza della prima specie, avvenuta probabilmente durante la fase di amplificazione del DNA, ha ridotto la presenza dell'altra provocando una distorsione nei risultati. In maniera analoga gli esperimenti effettuati nelle altre *library* hanno mostrato lo stesso squilibrio, infatti, anche quando le percentuali di DNA inserito erano a favore di *Sardina pilchardus* (es. 5% *Engraulis encrasicolus* - 95% *Sardina pilchardus*) il numero di sequenze attribuite è stato notevolmente inferiore (circa 6000 sequenze per *Sardina* rispetto le circa 30.000 per *Engraulis*), con una percentuale pari a 18%. Nell'esperimento condotto sul DNA estratto da tessuto in percentuali pari a 2.5% *Engraulis encrasicolus* e 97.5% *Sardina pilchardus*, il numero di sequenze attribuite all'una ed all'altra specie rispecchia perfettamente le percentuali (18% e 82%)

ottenute nella precedente *library*, con uno squilibrio a vantaggio di *Engraulis* rispetto a *Sardina* (circa 57.000 sequenze dell'una rispetto a circa 17.000 dell'altra).

Il campione ambientale di bianchetto ha ottenuto più di 100.000 sequenze delle quali circa l'89% attribuite a *Engraulis* e l'11% attribuite a *Sardina*.

Sulla base di quanto detto, l'ipotesi sulla migliore affinità dei primer con *Engraulis encrasicolus* risulta la più plausibile ma non sufficiente per spiegare il forte squilibrio tra le proporzioni generate artificialmente e le sequenze ottenute. L'incidenza di questi eventi impedisce la corretta quantificazione degli individui presenti in un campione ambientale composto da queste due specie. A tal proposito, una soluzione potrebbe consistere nel continuare ad effettuare prove con lo stesso marcatore (COX1) su materiale di partenza ambientale e "puro", al fine di trovare un fattore di correzione che consenta la quantificazione delle specie in oggetto; alternativamente, essendo la comunità in oggetto composta da solo due taxa, si potrebbe dirigere la ricerca futura verso l'identificazione e l'utilizzo di un nuovo marcatore, selezionato tra i geni a singola copia, in modo da evitare l'alterazione prodotta dalle differenti copie presenti nei differenti taxa.

Di notevole interesse il campione ambientale di acqua che ha ottenuto, inaspettatamente, più di 38.000 sequenze (delle quali circa il 70% attribuite ad *Engraulis encrasicolus* e circa il 30% a *sardina pilchardus*). E' stato dimostrato che la presenza di DNA extracellulare in acqua persiste per un periodo breve (meno di un mese), un lasso di tempo che varia in funzione delle condizioni ambientali in cui si trova, dai microorganismi presenti, dall'irraggiamento solare e dalle temperatura presenti (Dejean et al., 2011). E' quindi importante sottolineare le potenzialità applicative di questo approccio già utilizzato, ad esempio, per monitorare le specie aliene (che utilizzano differenti vettori di dispersione come le zavorre di acqua delle navi IMO, 2004) oppure semplicemente nel monitoraggio dei mammiferi marini (Foote et al., 2012).

7.5 Implicazioni nel biomonitoraggio

Gli ecosistemi di acque dolci rispondono sensibilmente alle sollecitazioni indotte dal riscaldamento globale e dall'impatto antropico (Sala et al. 2000; Dudgeon et al. 2006). Dinamica di popolazioni e biodiversità possono essere infatti fortemente modificate dalle alterazioni di temperatura (Wrona et al. 2006) e altri fattori di stress come eutrofizzazione, inquinamento organico, acidificazione, modifiche degli alvei e dei canali. Il controllo periodico dello stato di salute dei corpi idrici è quindi di fondamentale importanza per garantire la biodiversità globale e tutti i suoi servizi per l'uomo (Thomas et al. 2004; Montoya & Raffaelli 2010).

Il biomonitoraggio delle acque dolci è il processo di valutazione della qualità ecologica complessiva degli habitat di acqua dolce in base alla composizione e alle variazioni in abbondanza relativa della comunità di macrobenthos, in gran parte rappresentata anche dagli Efemeroteri (Hering et al. 2006).

Le indagini tassonomiche sono generalmente effettuate da esperti del settore e le relative relazioni sono utilizzate in programmi di protezione nell'ambito del biomonitoraggio ambientale o valutazioni di conformità richieste dalla direttiva (WFD 2000/60/CE).

Diversi fattori possono combinarsi impedendo il completo accesso all'identificazione dei campioni, rallentando ed inficiando la valutazione dell'habitat e della qualità delle acque (Lenat & Resh 2001; Gresens et al. 2007; Stribling et al. 2008); allo stesso modo, la limitata disponibilità di esperti tassonomi può aggravare la situazione causando valutazioni incomplete, coperture geografiche ristrette e ritardi nelle consegne delle relazioni tecniche.

In questo lavoro, il DNA *barcoding* degli Efemeroteri collezionati nella presente area di studio ha dimostrato di essere uno strumento veloce, affidabile ed efficace per la discriminazione delle specie locali, ha indicato inoltre un numero maggiore di specie presenti rispetto a quelle morfologicamente identificate ed ha fornito un punto di partenza per la costruzione di una banca dati di riferimento regionale per la biodiversità delle acque dolci. Tutto questo potrebbe supportare l'implementazione delle attuali metriche tassonomiche per l'attivazione di programmi di bio-valutazione e bio-monitoraggio delle acque dolci nell'area mediterranea basati sul DNA *barcoding*.

Studi precedenti hanno mostrato chiaramente come l'attuale conoscenza della biodiversità delle acque dolci e la sua risposta agli stress ambientali può essere ulteriormente migliorata da questa tecnica. In particolare la capacità di individuare i sottili cambiamenti nella composizione delle comunità (es. abbondanza dei taxa, specie criptiche o complessi di specie) potrebbe essere

fondamentale per rilevare perdite di biodiversità, capire come le comunità rispondono a determinati fattori specifici di stress, valutare gli impatti che si hanno sulle acque dolci in maniera più tempestiva (es prima che si verifichi il degrado) e fornire una misura più rapida della velocità e del grado di reupero degli habitat in risposta ai miglioramenti generate dalle pratiche di gestione e restauro dell'ambiente (Sweeney et al. 2011; Yu et al. 2012; Jackson et al. 2014).

In questa prospettiva, il DNA *barcoding* ha dimostrato la sua efficienza nell'area modello esaminata. Infatti l'Italia centrale, la Sardegna e la Corsica sono zone ricche di biodiversità tra cui molti endemismi e specie rare non ancora verificate (Blondel & Aronson 1999, Trizzino et al. 2014, Gattolliat et al. 2015).

Inoltre i recenti progressi in ambito tecnologico consentendo un più rapido ed economico sequenziamento del DNA (es. *Next Generation Sequencing, NGS*) beneficerebbero delle banche dati di riferimento costituite contribuendo ad ampliare la quantità di informazioni sull'ambiente (Hajibabaei et al. 2012; Shokralla et al. 2012). Questo potrebbe avere una enorme rilevanza in relazione alla crescente quantità di programmi sul controllo delle acque dolci in atto in molte parti del mondo (compresa l'Unione Europea) e che richiedono un numero sempre maggiore di siti da campionare ed una quantità sempre maggiore di dati da valutare (Stein et al. 2014).

I corpi idrici locali necessitano di programmi di biomonitoraggio per far fronte al continuo aumento delle sostanze inquinanti, alla pressione demografica, ai cambiamenti climatici e soprattutto per rispettare le direttive europee in materia di acque dolci e protezione della natura (WFD 2000/60/CE; 92/43/EEC).

In sinergia con il bio-monitoraggio basato sul DNA, la compilazione di liste rosse implementate dalle sequenze *barcode* potrebbe fornire input utili per migliorare l'individuazione di potenziali aree protette e assegnare delle priorità agli sforzi di conservazione locali.

La creazione di un portale web pubblico che consenta di ospitare la continua crescita di database sulle specie, di sequenze *barcode* e delle aree indagate con le relative informazioni è essenziale per il raggiungimento di questo obiettivo. In questo senso, banche dati di questo tipo, contenenti sequenze e relative informazioni ecologiche ed ambientali si stanno sviluppando in Canada (Zhou et al. 2010) e Stati Uniti (Webb et al. 2012).

Tuttavia, la nostra simulazione di identificazione DNA *barcoding* sulle banche dati di riferimento disponibili ha evidenziato la necessità di arricchire tali database con nuove specie

ed un campionamento geografico più ampio al fine di permettere una migliore e corretta identificazione.

Pertanto, una delle maggiori sfide della ricerca del futuro sarà certamente consolidare le conoscenze teoriche acquisite con l'esperienza pratica nei protocolli DNA *barcoding* con particolare riguardo nello sviluppo e miglioramento delle banche dati di riferimento.

8. CONCLUSIONI

I risultati ottenuti dalla presente ricerca di dottorato confermano l'applicabilità della tecnica del DNA *barcoding* nella discriminazione preliminare delle specie esaminate.

I dati risultanti dallo studio effettuato sugli Efemerotteri hanno ampliato la conoscenza di questo gruppo tassonomico tutt'ora in continua evoluzione e dimostrato l'importanza dello sviluppo di una banca dati di riferimento locale. Quest'ultima, in continuo aggiornamento e sviluppata ulteriormente su scala nazionale e continentale, agevolerà studi sulla biodiversità permettendo misure e valutazioni ripetibili nel tempo con una risoluzione tale da poter essere utile in ecologia e nel biomonitoraggio delle acque dolci.

Inoltre, lo sviluppo di una comune strategia *barcoding* per lo studio della biodiversità macrobentonica delle acque può sicuramente favorire una collaborazione positiva e costante tra tassonomi e biologi molecolari per corroborare gli sforzi di conservazione. Ad esempio, la compilazione di Liste Rosse implementate dalle indicazioni provenienti da studi sul DNA *barcoding* potrebbero migliorare l'individuazione di potenziali aree protette focalizzando gli sforzi per la conservazione locale.

La successiva fase di questo lavoro ha evidenziato come lo studio del *metabarcoding* ambientale, applicato a tecniche di sequenziamento del DNA di nuova generazione (*NGS*) ed opportunamente impostato, può permettere una maggiore risoluzione dei dati ottenibili, specialmente se coadiuvato da analisi ecologiche complementari, e di trarre un notevole risparmio in tempo e denaro una volta messi a punto in maniera definitiva i protocolli. Tuttavia ulteriori esperimenti sono necessari allo scopo di eliminare le distorsioni riscontrate.

Tali metodologie potrebbero avere una grande rilevanza a livello nazionale ed internazionale soprattutto in considerazione dell'aumentata necessità di condurre programmi di biomonitoraggio della qualità delle acque in tutto il mondo, traducibile in un aumento della quantità di siti da campionare e studiare, una scadenza ripetuta e regolare delle analisi e una notevole massa di dati da elaborare.

Il secondo caso di studio, effettuato su due specie alieutiche (*Engraulis encrasicolus* e *Sardina pilchardus*) considerate tra le più importanti risorse del Mar Mediterraneo, ha dimostrato come l'elevata capacità risolutiva del metodo a livello di genere e specie delle loro forme giovanili possa essere un punto di partenza per lo sviluppo di Piani di gestione riguardanti i prodotti commerciali da esse derivati. Ulteriori verifiche ed approfondimenti degli studi effettuati con la tecnica *NGS* potrebbero permettere la quantificazione delle due specie in termini

di frequenza consentendo sia un migliore sfruttamento degli stock di alici e sarde che deroghe alla pesca di questi prodotti che altrimenti rimarrebbe limitata a pochi periodi dell'anno e suscettibile di sanzioni con grande impatto sulle comunità di pescatori delle varie zone italiane.

BIBLIOGRAFIA

- Acosta-Martínez V., Dowd S., Sun Y., Allen V., 2008. Tagencoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biology and Biochemistry*, 40: 2762–2770.
- Adams I.P., Glover R.H., Monger W.A. et al., 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular Plant Pathology*, 10: 537–545.
- Armani A., Guardone L., La Castellana R., Gianfaldoni D., Guidi A., Castigliero L., 2015. DNA barcoding reveals commercial and health issues in ethnic seafood sold on the Italian market. *Food Control*, 55: 206-214, ISSN 0956-7135, <http://dx.doi.org/10.1016/j.foodcont.2015.02.030>.
- Azzali M., De Felice A., Luna M., Cosimi G., Parmiggiani F., 2002. The state of the Adriatic Sea centered on the small pelagic fish populations. *Pubblicazioni della Stazione Zoologica di Napoli I: Marine Ecology*, 23 (1): 78-91.
- Bachiller E., Cotano U., Ibaibarriaga L., Santos M., Irigoien X., 2015. Intraguild predation between small pelagic fish in the Bay of Biscay: impact on anchovy (*Engraulis encrasicolus* L.) egg mortality. *Marine Biology*, 162 (6): 1351-1369.
- Ball S., Hebert P.D.N., Burian S., Webb J., 2005. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society*, 24: 508–524. doi: 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. doi: 10.1007/s10750-007-9028-y
- Basilone G., Guisande C., Patti B., Mazzola S., Cuttitta A., Bonanno A., Vergara A.R., Maneiro I., 2006. Effect of habitat conditions on reproduction of the European anchovy (*Engraulis encrasicolus*) in the Strait of Sicily. *Fisheries Oceanography*, 15: 271-28.

- Bauerfeind E. & Soldán T., 2012. *The Mayflies of Europe (Ephemeroptera)*. Apollo Books, Ollerup, 781 pp.
- Belfiore C., 1983. *Efemerotteri (Ephemeroptera)*. Consiglio nazionale delle ricerche AQ/1/201 Guide per il riconoscimento delle specie animali delle acque interne Italiane. 24: 1-113.
- 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, 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. *Systematic Biology* 61: 851-869. DOI:10.1093/sysbio/sys037.
- Bisconti R., Canestrelli D., Tenchini R., Belfiore C., Buffagni A. & Nascetti G., *in press*. Cryptic diversity and multiple origins of the widespread mayfly species group *Baetis rhodani* on northwestern Mediterranean islands. *Scientific Reports*.
- Blondel J., Aronson J., 1999. *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, Oxford, UK.
- Bonanno A., Giannoulaki M., Barra M., Basilone G., Machias A., Genovese S., Goncharov S., Popov S., Rumolo P., Di Bitetto M., Aronica S., Patti B., Fontana I., Giacalone G., Ferreri R., Buscaino G., Somarakis S., Pyrounaki M.M., Tsoukali S., Mazzola S., 2014. Habitat Selection Response of Small Pelagic Fish in Different Environments. Two Examples from the Oligotrophic Mediterranean Sea. *PLoS One*, 9(7): e101498. doi:10.1371/journal.pone.0101498
- Brittain J. E., Sartori M., 2003. *Ephemeroptera (Mayflies)*. In: V. H. Resh & R. T. Cardé (Eds), *Encyclopedia of Insects*. Academic Press, San Diego: 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. doi: 10.1111/j.1755-0998.2011.03108.x.
- Bue´e M., Reich M., Murat C. et al., 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist*, 18, 449–456.
- Buffagni A. & Erba S., 2007. Intercalibrazione e classificazione di qualità ecologica dei fiumi per la 2000/60/EC (WFD): l'indice STAR_ICMI. IRSA-CNR. *Notiziario dei Metodi Analitici*, 1: 94-100.
- Burgess D. J., 2011. Human disease: Next-generation sequencing of the next generation. *Nature Reviews Genetics*, 12: 78-79.
- Campaioli S., Ghetti P.F., Minelli A. & S. Ruffo, 1994. Manuale per il riconoscimento dei macroinvertebrati delle acque dolci italiane. Vol. 1°. Provincia Autonoma di Trento: 356 pp.
- Campaioli S., Ghetti P.F., Minelli A. & S. Ruffo, 1999. Manuale per il riconoscimento dei macroinvertebrati delle acque dolci italiane. Vol. 2°. Provincia Autonoma di Trento: 127 pp.
- Cardoni S., Tenchini R., Ficulle I., Piredda R., Simeone M.C., Belfiore C., 2015. DNA barcode assessment of Mediterranean mayflies (Ephemeroptera), benchmark data for a regional reference library for rapid biomonitoring of freshwaters. *Biochemical Systematics and Ecology*, 62: 36-50.
- Chaves B.R.N., Chaves A.V., Nascimento A.C.A., Chevitarese J., Vasconcelos M.F. and Santos F.R. 2014. Barcoding Neotropical birds: assessing the impact of nonmonophyly in a highly diverse group. *Molecular Ecology Resources*, 15: 921–931. doi: 10.1111/1755-0998.12344.
- Coll M., Santojanni A., Palomera I., Tudela S., Arneri A., 2007. An ecological model of the Northern and Central Adriatic Sea: Analysis of eco system structure and fishing impacts. *Journal of Marine System*, 67: 119-154.

- Collins R.A., Cruickshank R.H., 2013. The seven deadly sins of DNA barcoding. *Mol. Ecol. Resour.* 13, 969-975.
- Colombo F., Macdonald C.A., Jeffries T.C., Powell J.R., Singh B.K., 2016. Impact of forest management practices on soil bacterial diversity and consequences for soil processes. *Soil Biology and Biochemistry*, 94: 200-210. doi:10.1016/j.soilbio.2015.11.029.
- Coombs S.H., Boyra G., Rueda L.D., Uriarte A., Santos M., Conway D.V.P., Halliday N.C., 2004. Buoyancy measurements and vertical distribution of eggs of sardine (*Sardina pilchardus*) and anchovy (*Engraulis encrasicolus*). *Marine Biology*, 145 (5): 959-970.
- Cox-Foster D.L., Conlan S., Holmes E.C. et al., 2007. A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, 318, 283–287.
- Creer S., Fonseca V.G., Porazinska D.L. et al., 2010. Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Molecular Ecology*, 19 (1): 4–20.
- Cuttitta A., Patti B., Maggio T., Quinci E.M., Pappalardo A.M., Ferrito V., De Pinto V., Torri M., Falco F., Nicosia A., Musco M., Armeri G.M., Placenti F., Tranchida G., Mifsud R., Bonanno A., Mazzola S., 2015. Larval population structure of *Engraulis encrasicolus* in the Strait of Sicily as revealed by morphometric and genetic analysis. *Fisheries Oceanography*, 24: 135–149. doi: 10.1111/fog.12098
- Degnan P.H., Ochman H., 2011. Illumina-based analysis of microbial community diversity. *The ISME Journal*, 2011: 1–12.
- Dejean T., Valentini A., Duparc A., Pellier-Cuit S., Pompanon F., Taberlet P. et al., 2011. Persistence of Environmental DNA in Freshwater Ecosystems. *PLoS ONE* 6(8): e23398. doi:10.1371/journal.pone.0023398
- Deiner K. & Altermatt F., 2014. Transport Distance of Invertebrate Environmental DNA in a Natural River. *PLoS ONE* 9(2): e88786. doi:10.1371/journal.pone.0088786

- De Vos J.M., Joppa L.N., Gittleman J.L., Stephens P.R., Pimm S.L., 2015. Estimating the normal background rates of species extinction. *Conservation Biology*, 29: 452–462.
- Doña J., Ruiz-Ruanob F. J., Jovania R., 2014. DNA barcoding of Iberian Peninsula and North Africa Tawny Owls *Strix aluco* suggests the Strait of Gibraltar as an important barrier for phylogeography. *Mitochondrial DNA*, 14:1-4. Doi: 10.3109/19401736.2015.1089573.
- Dowle E.J., Pochon X., Banks J. C., Shearer K., Wood S. A., 2015. Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. *Molecular Ecology Resources*, 10.1088/1751-8113/44/8/085201.
- Dudgeon D., Arthington A.H., Gessner M.O., Kawabata Z.-I., Knowler D.J., Le' Ve^Que 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. *Biological Reviews*, 81:163-182. doi: 10.1017/S1464793105006950.
- Edmunds G.F., McCafferty W.P., 1988. The mayfly subimago. *Annual Review of Entomology*, 33: 509–529.
- Ekrem T., Willassen E., Stur E., 2007. A comprehensive DNA sequence library is essential for identification with DNA barcodes. *Molecular Phylogenetics and Evolution* 43: 530–542. doi: 10.1016/j.ympev.2006.11.021.
- Elbrecht V. & Leese F., 2015. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass - sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, 10, e0130324.
- 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. *Journal of Great Lakes Research*, 38: 180-186. doi: 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.

FAO, 2014. The State of World Fisheries and Aquaculture 2014. Rome, 223 pp.

Fazekas A.J., Burgess K.S., Kesanakurti P.R., Graham S.W., Newmaster S.G., Husband B.C., Percy D.M., Hajibabaei M., and Barrett S.C.H., 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 3(7): e2802.

Ficetola G.F., Miaud C., Pompanon F., Taberlet P., 2008. Species detection using environmental DNA from water samples. *Biology Letters*, 4: 423–425.

Fierer N., Breitbart M., Nulton J., Salamon P., Lozupone C., Jones R., Robeson M., Edwards R.A., Felts B., Rayhawk S., Knight R., Rohwer F., Jackson R.B., 2007. Metagenomic and small-subunit rRNA analyses of the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Applied and Environmental Microbiology*, 73: 7059–7066.

Fischer W., Bauchot M.L., Schneider M., 1987. Fiches FAO d'identification des espèces pour les besoins de la pêche (Révision 1) , Méditerranée et mer Noire. Zone de pêche 37. 1, Vegetaux et Invertébrés, Publication préparée par la FAO (Project GCP/INT/422/EEC), Rome, FAO: 760 pp.

Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5): 294-299

Footo A.D., Thomsen P.F., Sveegaard S., Wahlberg M., Kielgast J., Kyhn L.A. et al., 2012. Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *PLoS ONE* 7(8): e41781. doi:10.1371/journal.pone.0041781

France S.C., Hoover I.L., 2002. DNA sequences of the mitochondrial COI gene have low levels of divergence among deep-sea octocorals (Cnidaria: Anthozoa). *Hydrobiology* 471: 149–155.

- Frias-Lopez J., Shi Y., Tyson G.W., Coleman M.L., Schuster S.C., Chisholm S.W., DeLong E.F., 2008. Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences*, 105: 3805–3810.
- Froggia C., 1973. Osservazioni sull'alimentazione del merluzzo (*Merluccius merluccius*) del Medio Adriatico. In *Atti V Congresso della Società Italiana di Biologia Marina*. Nardò (Italy): Società Italiana di Biologia Marina, 327-341.
- Fromer M., Pocklington A. J., Kavanagh D. H., Williams H. J., Dwyer S., Gormley P., et al. , 2014. De novo mutations in schizophrenia implicate synaptic networks. *Nature* 506, 179–184. doi: 10.1038/nature12929
- Funk D.J., Omland K.E., 2003. Species-level monophyly and paraphyly: frequency, causes and consequences with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 34: 397e423.
- Gaino E., Belfiore C., Mazzini M., 1987. Ootaxonomic investigation of the Italian species of the genus *Electrogena* (Ephemeroptera, Heptageniidae). *Bolletino di zoologia*, 54 (2): 169-175. doi: 10.1080/11250008709355578.
- Ganias K., Somarakis S., Machias A., Theodorou A. J., 2003. Evaluation of spawning frequency in a Mediterranean sardine population (*Sardina pilchardus*). *Marine Biology* 14: 1169-1179.
- Ganias K., 2014. *Biology and ecology of sardines and anchovies*. CRC Press
- Gattolliat J.L., Cavallo E., Vuataz L. & Sartori M., 2015. DNA barcoding of Corsican mayflies (Ephemeroptera) with implications on biogeography, systematics and biodiversity. *Arthropod Systematics & Phylogeny*, 73 (1): 3–18.
- Ghetti P.F., 1995. *Indice Biotico Esteso (I.B.E.) (Metodi di analisi per ambienti di acque correnti)*. IRSA-CNR *Notiziario dei Metodi Analitici*, 1-24.

- Ghetti P.F., 1997. Indice Biotico Esteso (I.B.E.). I macroinvertebrati nel controllo della qualità degli ambienti di acque correnti. Provincia Autonoma di Trento, pp. 222.
- Giannoulaki M., Iglesias M., Tugores M.P., Bonanno A., Patti B., et al., 2013. Characterizing the potential habitat of European anchovy *Engraulis encrasicolus* in the Mediterranean Sea, at different life-stages. *Fisheries Oceanography* 22: 69–89. doi: 10.1111/fog.12005
- Gibson J., Shokralla S., Porter T.M., King I., van Konynenburg S., Janzen D.H. et al., 2014. Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasytematics. *Proceedings of the National Academy of Sciences*, 111: 8007–8012. doi: 10.1073/pnas.1406468111 PMID: 24808136
- Gomez-Alvarez V., King G.M., Nüsslein K., 2007. Comparative bacterial diversity in recent Hawaiian volcanic deposits of different ages. *FEMS Microbiology Ecology*, 60: 60–73
- Gramitto M.E., 2001. La gestione della pesca marittima in Italia. Fondamenti tecnico-biologici e normativa vigente. Marchesi grafiche Editoriali Spa Monografie scientifiche CNR, 319 pp.
- Gresens S.E., Belt K.T., Tang J.A., Gwinn D.C., Banks P.A., 2007. Temporal and spatial responses of Chironomidae (Diptera) and other benthic invertebrates to urban stormwater runoff. *Hydrobiologia*, 575: 173-190. doi 10.1007/s10750-006-0366-y
- Haile J., Froese D.G., MacPhee R.D.E. Roberts R.G., Arnold L.J., Reyes A.V., Rasmussen M., Nielsen R., Brook B.W., Robinson S., Demuro M., Gilbert M.T.P., Munche K., Austin J.J., Cooper A., Barnesh I., Mölleri P., Willerslev E., 2009. Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences*, 106: 22352–22357.
- Haase P., Pauls S.U., Schindehütte K., Sundermann A., 2010. First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. *Journal of the North American Benthological Society*, 29: 1279–1291. doi: 10.1899/09-183.1

- Hajibabaei M., Janzen D.H., Burns J.M., Hallwachs W., Hebert P.D.N., 2006. DNA barcodes distinguish species of tropical lepidoptera. *Proceedings of the National Academy of Sciences* 103: 968-971. <http://dx.doi.org/10.1073/pnas.0510466103>.
- Hajibabaei M., Shokralla S., Zhou X., Singer G.A.C., Baird D.J., 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, 6, e17497.
- 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 Ecology*, 12: 28. <http://dx.doi.org/10.1186/1472-6785-12-28>.
- Hebert P.D.N., Cywinska A., Ball S.L. and deWaard J.R., 2003a. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B*, 270: 313-321 doi: 10.1098/rspb.2002.2218.
- Hebert P.D.N., Ratnasingham S., deWaard J.R., 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B*, 270: 596-599.
- 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 *Astrartes fulgerator*. *Proceedings of the National Academy of Sciences*, 101: 14812–14817. doi: 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 lineageidiosyncratic error rates. *PloS ONE* 5(12): e14448. doi:10.1371/journal.pone.00114448
- Hendrich L., Morinière J., Haszprunar G. et al., 2014. A comprehensive DNA barcode database for Central European beetles with a focus on Germany: adding more than 3500 identified species to BOLD. *Molecular Ecology Resources*, 15: 795–818.

- Hennig W., 1976. Anthomyiidae. In *Die Fliegen der Palaearktischen Region* (Lindner, E., ed.), pp. 329–376, Schweizerbart
- 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. *Freshwater Biology*, 51: 1757-1785. <http://dx.doi.org/10.1111/j.1365-2427.2006.01610.x>.
- Herrera A., Hérya M., Stach J.E.M., Jaffré T., Normanda P., Navarroa E., 2007. Species richness and phylogenetic diversity comparisons of soil microbial communities affected by nickel-mining and revegetation efforts in New Caledonia. *European Journal of Soil Biology*, 43 (2): 130– 139.
- 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 Zealand Journal of Marine and Freshwater Research*, 43 (5): 1137-1146. <http://dx.doi.org/10.1080/00288330.2009.9626536>.
- Huber J.A., Welch D.B.M., Morrison H.G., Huse S.M., Neal P.R., Butterfield D.A., Sogin M.L., 2007. Microbial population structures in the deep marine biosphere. *Science*, 318: 97– 100.
- IMO, 2004. *International Convention for the Control and Management of Ships' ballast water and sediments*. IMO.
- IREPA, 2009. *Il settore ittico in Italia. Check-up 2008*.
- 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 macroinvertebrate communities based on morphological and DNA barcode identification. *Freshwater Science* 33: 312-324.

- 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. *Philosophical Transactions of the Royal Society B*, 360 (1462): 1835-45.
- Jones M., Ghoorah A., Blaxter M., 2011. jMOTU and taxonator: turning DNA barcode sequences into annotated operational taxonomic units. *PLoS ONE* 6e19259. doi: 10.1371/journal.pone.0019259.
- Kermarrec L., Franc A., Rimet F. et al., 2014. A next-generation sequencing approach to river biomonitoring using benthic diatoms. *Freshwater Science*, 33: 349–363.
- Khalifa M.E., Varsani A., Ganley A.R.D., Pearson M.N., 2015. Comparison of Illumina de novo assembled and Sanger sequenced viral genomes: A case study for RNA viruses recovered from the plant pathogenic fungus *Sclerotinia sclerotiorum*. *Virus Research*, In Press.
- 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.
- Kolkwitz R., Marsson M., 1909. *Okologie der tierische Saprobien. Beitrage zur Lehre von des biologischen Gewasserbeurteilung. Internationale Revue der Gesamten Hydrobiologie und Hydrographie*, 2:126-152.
- Kowalczyk R., Taberlet P., Coissac E., Valentini A., Miquel C., Kamińska T., Wójcika J.M., 2011. Influence of management practices on large herbivore diet—Case of European bison in Białowieża Primeval Forest (Poland). *Forest Ecology and Management*, 261: 821–828.
- Krishnamurthy P.K., Francis P.A., 2012. A critical review on the utility of DNA barcoding in biodiversity conservation. *Biodiversity and Conservation*, 21: 1901-1919. doi: 10.1007/s10531-012-0306-2.

- Lawrence M.S., Stojanov P., Mermel C.H., Robinson J.T., Garraway L.A., Golub T.R., Meyerson M., Gabriel S.B., Lander E.S., Getz G., 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*, 505: 495–501. doi: 10.1038/nature12912
- Lee M.S.Y., 2004. The molecularisation of taxonomy. *Invertebrate Systematics*, 18: 1-6.
- Leininger S., Urich T., Schloter M., Schwark L., Qi J., Nicol G.W., Prosser J.I., Schuster S.C., Schleper C., 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*, 442: 806–809.
- Lenat D.R., Resh V.H., 2001. Taxonomy and stream ecology: the benefits of genus- and species-level identifications. *Journal of the North American Benthological Society*, 20: 287-298.
- Leonard S.R., Mammel M.K., Lacher D.W., Elkins C.A., 2015. Application of metagenomic sequencing to food safety: detection of Shiga toxin-producing *Escherichia coli* on fresh bagged spinach. *Applied and Environmental Microbiology*, 81:8183–8191. doi:10.1128/AEM.02601-15.
- Leonorì I., 2007. Valutazione di quelle caratteristiche dei piccolo pelagici e del loro ambiente utili a programmare l'attività di pesca in Adriatico. Relazione Finale Programma di ricerca n. 6A67, VI Piano triennale L. N. 41/82. Ministero per le Politiche Agricole e Forestali, Roma: 133 pp.
- Librado P., Rozas J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Lim G.S., Balke M., Meier R., 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. *Systematic Biology*, 61: 165-169.
- Lindegren M., Checkley D.M.J., Rouyer T., MacCall A.D., Stenseth N.C., 2013. Climate, fishing, and fluctuations of sardine and anchovy in the California Current. *Proceedings of the National Academy of Sciences*, 110: 13672–13677.
- Lipscomb D., Platnick N., Wheeler Q., 2003. The intellectual content of taxonomy: a

comment on DNA taxonomy. *Trends in Ecology and Evolution*, 18 (2): 65- 66.

MacCall A.D., Sydeman W.J., Davison P.C., Thayer J.A., 2015. Recent collapse of northern anchovy biomass off California. *Fisheries Research*, 175: 87-94. doi:10.1016/j.fishres.2015.11.013

Mächler E., Deiner K., Spahn F., 2016. Fishing in the water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. *Environmental Science & Technology*, 50: 305 - 312. DOI: 10.1021/asc.est.5b04188.

Mancini L., Puccinelli C., Marcheggiani S., Scanu G., Aste F., Bernabei S., Martone C., Sollazzo C., 2009. Metodo per la valutazione dello stato ecologico delle acque correnti: comunità diatomiche. Roma: Istituto Superiore di Sanità (Rapporti ISTISAN 09/19).

Meyer C., Paulay G., 2005. DNA barcoding: error rates based on comprehensive sampling. *PLOS Biology*, 3:e422.

Minhós T., Wallace E., Da Silva M.J.F., Sá R.M., Carmo M., Barata A., Bruford M.W., 2013. DNA identification of primate bushmeat from urban markets in Guinea-Bissau and its implications for conservation. *Biological Conservation*, 167: 43–49.

Montoya J.M., Raffaelli D., 2010. Climate change, biotic interactions and ecosystem services. *Philosophical Transactions of the Royal Society B*, 365: 2013-2018. <http://dx.doi.org/10.1098/rstb.2010.0114>.

Morello E.B., Arneri E., 2009. Anchovy and sardine in the Adriatic sea - An ecological review. In *Oceanography and Marine Biology*, 47: 209-255.

Moritz C., Cicero C., 2004. DNA barcoding: promise and pitfalls. *PLOS Biology*, 2 (10): 354 pp.

Murphy E.F., Cotter P.D., Healy S., Marques T.M., O'Sullivan O., Fouhy F., Clarke S.F., O'Toole P.W., Quigley E.M., Stanton C., Ross P.R., O'Doherty R.M., Shanahan F., 2010.

Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut*, 59: 1635–1642.

Myers N., Mittermeier R.A., Mittermeier C.G., da Fonseca G.A.B., Kent J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853e858. <http://dx.doi.org/10.1038/35002501>.

Nacke H., Thürmer A., Wollherr A., Will C., Hodac L., Herold N., Schöning I., Schrumpf M., Daniel R., 2011. Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. *PLoS ONE*, 6(2) e17000.

Nyaga M.M., Peenze I., Potgieter C.A., Seheri L.M., Page N.A., Yinda C.K., Steele A.D., Matthijnssens J., Mphahlele M.J., 2016. Complete genome analyses of the first porcine rotavirus group H identified from a South African pig does not provide evidence for recent interspecies transmission events. *Infection, Genetics and Evolution*, 38: 1–7

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. doi: 0.1007/s10201-010-0328-y.

Orsi Relini L., Cima C., Palandri G., Relini M., Garibaldi F., 1999. Alimentazione del tonno giovanile nell'ecosistema del largo del Mar Ligure. *Biologia Marina Mediterranea* 6 (2): 295-302.

Packer L., Gibbs J., Sheffield C., Hanner R., 2009. DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources*, 9: 42-50. <http://dx.doi.org/10.1111/j.1755-0998.2009.02631.x>.

Park D.S., Footitt R., Maw E., Hebert P.D.N., 2011. Barcoding Bugs: DNA-Based Identification of the True Bugs (Insecta: Hemiptera: Heteroptera). *PLoS ONE*, 6(4) e18749. doi: 10.1371/journal.pone.0018749.

- Pegard A., Miquel C., Valentini A., Coissac E., Bouvier F., François D., Taberlet P., Engel E., Pompanon F., 2009. Universal DNAbased methods for assessing the diet of grazing livestock and wildlife from feces. *Journal of Agricultural and Food Chemistry*, 57: 5700–5706.
- Piñol J., Mir G., Gomez-Polo P., Agustí N., 2014. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology*, 1–12. doi: 10.1111/1755-0998.12355
- Pimm S.L., Russell G.J., Gittleman J.L., Brooks T.M., 1995. The future of biodiversity. *Science*, 269: 347–350.
- Polunin N.V.C., 2008. *Aquatic ecosystems: trends and global prospects*. Cambridge University Press, Cambridge, UK.
- Polz M.F. & Cavanaugh C.M., 1998. Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, 64, 3724–3730.
- Porazinska D.L., Giblin-Davis R.M., Esquivel A., Powers T.O., Sung W., Thomas W.K., 2010. Ecometagenetics confirms high tropical nematode diversity. *Molecular Ecology*, 19: 5521–5530.
- Rach J., De Salle R., Sarkar I.N., Schierwater B., Hadrys H., 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in odonata. *Proceedings of the Royal Society of London B*, 275: 237-247. <http://dx.doi.org/10.1098/rspb.2007.1290>.
- Ramsingh G., Koboldt D.C., Trissal M., Chiappinelli K.B., Wylie T., Koul S., Chang L.W., Nagarajan R., Fehniger T.A., Goodfellow P., Magrini V., Wilson R.K., Ding L., Ley T.J., Mardis E.R., Link D.C., 2010. Complete characterization of the microRNAome in a patient with acute myeloid leukemia. *Blood*, 116: 5316–5326,.
- Ravussin Y., Koren O., Spor A., LeDuc C., Gutman R., Stombaugh J., Knight R., Ley R.E., Leibel R.L., 2011. Responses of gut microbiota to diet composition and weight loss in lean and obese mice. *Obesity*, 20 (4):738-747. doi: 10.1038/oby.2011.111.

- Regner S., Regner D., Marasović I., Kršinić F., 1987. Spawning of sardine, *Sardina pilchardus* (Walbaum, 1972), in the Adriatic under upwelling conditions. *Acta Adriatica* 28: 161-198.
- Regolamento (CE) 1967/2006 del Consiglio del 21 dicembre 2006, relativo alle misure di gestione per lo sfruttamento sostenibile delle risorse della pesca nel Mar Mediterraneo e recante modifica del regolamento (CEE) 2847/1993 e che abroga il regolamento (CE) 1626/1994.
- Rivera J., Currie D.C., 2009. Identification of nearctic black flies using DNA barcodes (diptera: simuliidae). *Molecular Ecology Resources*, 9 : 224-236. 10.1111/j.1755-0998.2009.02648.x.
- Roach J.C., Glusman G., Smit A.F., Huff C.D., Hubley R., Shannon P.T., Rowen L., Pant K.P., Goodman N., Bamshad M., Shendure J., Drmanac R., Jorde L.B., Hood L., Galas D.J., 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, 328: 636–639. doi: 10.1126/science.1186802
- Rosenberg N.A., 2007. Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution* 61 (2): 317-323. doi: 10.1111/j.1558-5646.2007.00023.x.
- Rousk J., Bååth E., Brookes P.C., Lauber C.L., Lozupone C., Caporaso J.G., Knight R., Fierer N., 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal*, 4: 1340–1351.
- Rusch D.B., Halpern A.L., Sutton G., Heidelberg K.B., Williamson S., Yooseph S., Wu D., Eisen J.A., Hoffman J.M., Remington K., Beeson K., Tran B., Smith H., Baden-Tillson H., Stewart C., Thorpe J., Freeman J., Andrews-Pfannkoch C., Venter J.E., Li K., Kravitz S., Heidelberg J.F., Utterback T., Rogers Y.H., Falcón L.I., Souza V., et al., 2007. The Sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biology*, 5 (3): e77.
- 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>.

- Sanfilippo M., Pulicanò G., Costa F., Manganaro A., 2011. Juvenile fish populations in two areas of the Sicilian coast. *Natura Rerum* 1: 43-50.
- Sanger F., Nicklen S., Coulson A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74: 5463-7.
- Sansoni G., 1988. Atlante per il riconoscimento dei macroinvertebrati dei corsi d'acqua italiani. Provincia autonoma di Trento, Trento, 191 pp.
- Santojanni A., 2001. La dinamica di popolazione: problematiche e metodiche di studio. In: La gestione della pesca marittima in Italia – Fondamenti tecnico-biologici e normativa vigente, pp. 33-40. Monografie scientifiche, Serie Scienze e tecnologie dell'ambiente, Consiglio Nazionale delle Ricerche, 319 pp.
- Santojanni A., Arneri E., Barry C., Belardinelli A., Cingolani N., Giannetti G., Kirkwood G., 2003. Trends of anchovy (*Engraulis encrasicolus*, L.) biomass in the northern and central Adriatic Sea. *Scientia Marina* 67: 327-340.
- Santojanni A., Cingolani N., Arneri E., Kirkwood G., Belardinelli A., Giannetti G., Colella S., Donato F., Barry C., 2005. Stock assessment of sardine (*Sardina pilchardus*, Walb.) in the Adriatic Sea, with an estimate of discards. *Scientia Marina* 69 (4): 603-617.
- Santojanni A., Arneri E., Bernardini V., Cingolani N., Di Marco M., Russo A., 2006. Effects of environmental variables on recruitment of anchovy in the Adriatic Sea. *Climate Research* 31: 181-193.
- Sari A., Duran M., Sen A., Bardakci F. (2015). Investigation of Chironomidae (Diptera) relationships using mitochondrial COI gene. *Biochemical Systematics and Ecology*, 59: 229-238. <http://dx.doi.org/10.1016/j.bse.2015.01.005>.
- 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.). *Annales de Limnologie*, 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. *Applied and Environmental Microbiology*, 75: 7537–7541. doi: 10.1128/AEM.01541-09.
- Schnell I.B., Fraser M., Willerslev E., Thomas M., Gilbert P., 2010. Characterisation of insect and plant origins using DNA extracted from small volumes of bee honey. *Arthropod Plant Interact*, 4: 107–116.
- Seberg O., Humphries C.J., Knapp S., Stevenson D.W., Petersen G., Scharff N. & Andersen N.M., 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. *Trends in Ecology and Evolution*, 18 (2): 63-65.
- Serino M., Luche E., Gres S., Baylac A., Bergé M., Cenac C., Waget A., Klopp P., Iacovoni J., et al., 2011. Metabolic adaptation to a high-fat diet is associated with a change in the gut microbiota. *Gut*, 61: 543–553.
- Shapiro B., 2008. Engineered polymerases amplify the potential of ancient DNA. *Trends in biotechnology*, 26: 285–287.
- Shearer T.L., Van Oppen M.J.H., Romano S.L., Worheide G., 2002. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, 11: 2475–2487.
- 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. *Molecular Ecology Resources*, 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. *Molecular Ecology Resources*, 21: 1794-1805. <http://dx.doi.org/10.1111/j.1365-294X.2012.05538.x>.
- Sickel W., Ankenbrand M.J., Grimmer G., Holzschuh A., Härtel S., Lanzen J., Steffan-Dewenter I. & Keller A., 2015. Increased efficiency in identifying mixed pollen

- samples by meta-barcoding with a dual-indexing approach. *BMC Ecology*, 15: 20. DOI 10.1186/s12898-015-0051-y.
- Sinovčić G., Zorica B., Čikeš Keč V., Mustać B., 2009. Inter-annual fluctuations of the population structure, condition, length-weight relationship and abundance of sardine, *Sardina pilchardus* (Walb., 1792), in the nursery and spawning ground (coastal and open sea waters) of the eastern Adriatic Sea (Croatia). *Acta Adriatica*, 50 (1): 11-22.
- 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). *Proceedings of the National Academy of Sciences*, 103: 3657–3662. doi: 10.1073/pnas.0511318103.
- 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. *Proceedings of the National Academy of Sciences*, 105: 12359-12364. <http://dx.doi.org/10.1073/pnas.0805319105>.
- Smith M.A., Fernandez-Triana J.L., Eveleigh E., Gomez J., Guclu C., Hallwachs W., Hebert P.D.N., Hrcak 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 20000 sequences. *Molecular Ecology Resources*, 13: 168-176. <http://dx.doi.org/10.1111/1755-0998.12038>.
- Smol J.P., 2008. *Pollution of lakes and rivers: a paleoecological perspective*. 2nd edition. Blackwell, Malden, Massachusetts.
- Sogin M.L., Morrison H.G., Huber J.A., et al., 2006. Microbial diversity in the deep sea and the underexplored ‘rare biosphere’. *Proceedings of the National Academy of Sciences*, 103: 12115–12120.

- Soininen E.M., Valentini A., Coissac E. et al., 2009. Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, 6: 16.
- Sønstebø J.H., Gielly L., Brysting A.K., et al., 2010. Using nextgeneration sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources*, 10: 1009–1018.
- Sroka P., 2012. Systematics and phylogeny of the West Palaearctic representatives of subfamily Baetinae (Insecta: Ephemeroptera): combined analysis of mitochondrial DNA sequences and morphology. *Aquatic Insect*, 34: 23-53. doi: 10.1080/01650424.2012.718081.
- Ståhls G., Savolainen E., 2008. MtDNA COI barcodes reveal cryptic diversity in the *Baetis vernus* group (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution*, 46: 82–87. doi: 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., 2006. Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society*, 25: 271–287. doi: 10.1899/0887-3593.
- Strayer D.L., Dudgeon D., 2010. Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, 29: 344-358. doi: 10.1899/08-171.1.
- Stribling J.B., Pavlik K.L., Holdsworth S.M., Leppo E.W., 2008. Data quality, performance, and uncertainty in taxonomic identification for biological assessments. *Journal of the North American Benthological Society*, 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? *Journal of the North American Benthological Society*, 30: 195-216. <http://dx.doi.org/10.1899/10-016.1>.
- Tachet H., Richoux P., Bournaud M. & Usseglio-Polatera P., 2006. *Invertébrés d'eau douce (systematique, biologie, ecologie)*. CNRS Editions, Paris. 587 pp
- Taylor K.H., Kramer R.S., Davis J.W., Guo J., Duff D.J., Xu D., Caldwell C.W., Shi H., 2007. Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Research*, 67: 8511–8518,
- Tautz D., Arctander P., Minelli A., Thomas R.H. & Vogler A.P., 2003. A plea for DNA taxonomy. *Trends in Ecology and Evolution*, 18 (2): 70-74.
- Tennesen J.A., Bigham A.W., O'connor T.D., Fu W., Kenny E.E., Gravel S., et al., 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*, 337: 64–69. doi: 10.1126/science.1219240
- Theissinger K., Ba'lint M., Haase P., Johannesen J., Laube I., Pauls S.U., 2011. Species distribution models and molecular data reveal the Pleistocene history of the coldadapted mayfly *Ameletus inopinatus* (Ephemeroptera: Siphonuridae) in Europe. *Freshwater Biology*, 56: 2554–2566
- Thomas A., Belfiore C., 2004. *Fauna Europea: Ephemeroptera*. Fauna Europea version 1.1. available online at www.faunaeur.org.
- Thomsen P.F. et al., 2011. Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology Resources*, 21: 2565–2573.
- 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

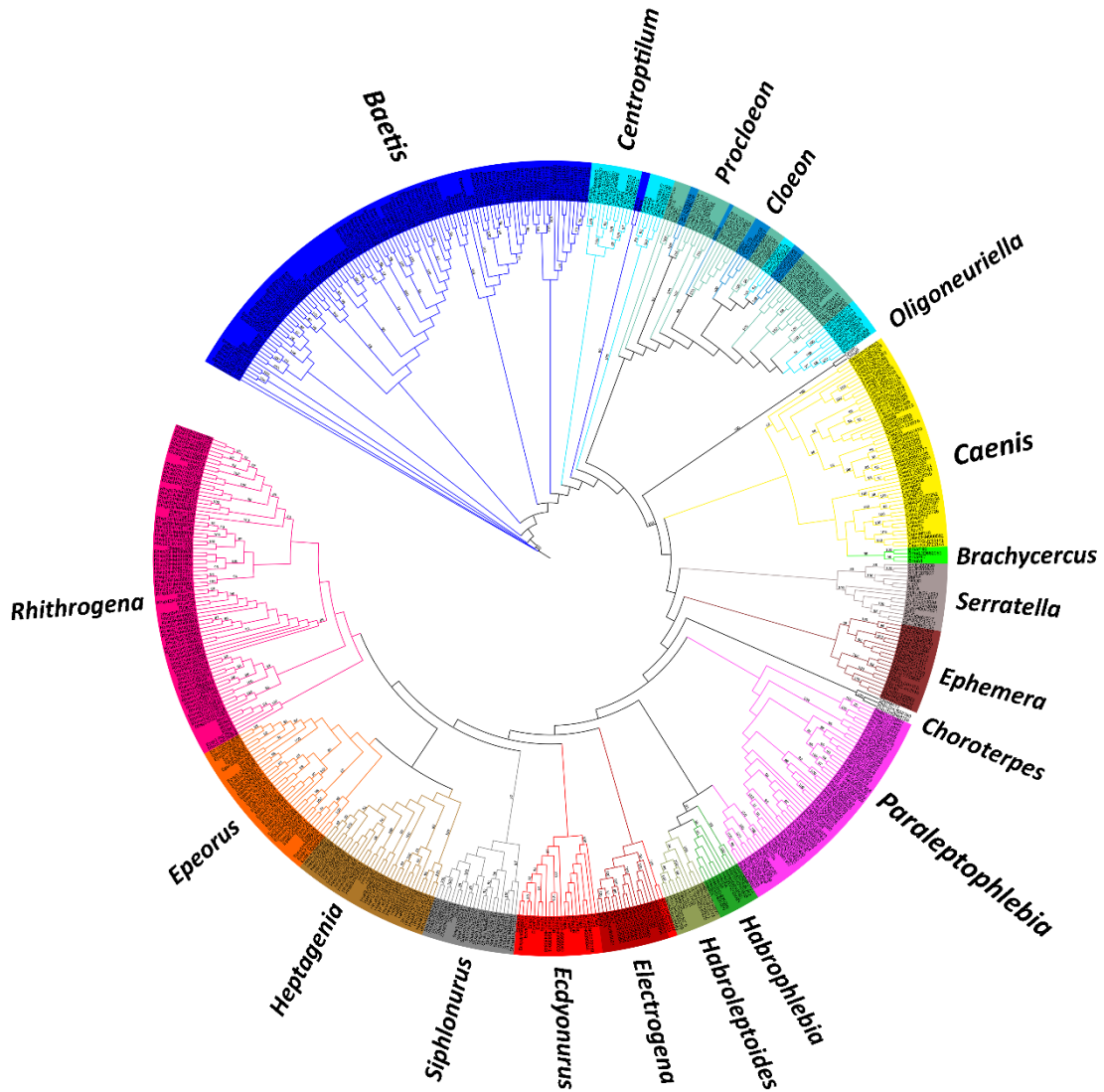
- headwater ecosystems, as inferred from diversity and distribution of a water beetle lineage. *Biodiversity and Conservation*, 24: 149e170. <http://dx.doi.org/10.1007/s10531-014-0798-z>.
- Turnbaugh P.J., Backhed F., Fulton L., Gordon J.I., 2008. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host & Microbe*, 3: 213–223.
- Turnbaugh P.J., Ridaura V.K., Faith J.J., Rey F.E., Knight R., Gordon J.I. 2009. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine*, 1: 6ra14.
- Valentini A., Miquel C., Nawaz M.A. et al., 2009. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources*, 9: 51–60.
- Vučetić T., 1963. O ishrani odrasle srdele (*Sardina pilchardus* Walb.) u srednjem Jadranu. *Acta Adriatica*, 10 (2): 3-50.
- Ward R.D., Zemplak T.S., Innes B.H., Last P.R., Hebert P.D.N., 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, 360: 1847-1857.
- 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. doi: 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. *Freshwater Science*, 33: 323-335. doi: 10.1086/674982.
- Wiemers M., Fiedler K., 2007. Does the barcoding gap exist? - A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology*, 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). *Molecular Phylogenetics and Evolution*, 40: 370–382. doi: 10.1016/j.ympev.2006.03.004.
- Woese C.R., Kandler O., Wheelis M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, 87 (12): 4576-4579. doi:10.1073/pnas.87.12.4576
- 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: 326e329. <http://dx.doi.org/10.1579/0044-7447>.
- Yang Y., Muzny D.M., Reid J.G., Bainbridge M.N., Willis A., Ward P.A., et al., 2013. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *The New England Journal of Medicine*, 369: 1502–1511. doi: 10.1056/NEJMoa1306555
- Yu D.S.K., Van Achterberg C., Horstmann K., 2012. *Ichneumonidea 2011*. Database on flash-drive, Taxapad, (Ottawa, Ontario, Canada). www.taxapad.com.
- Zarrad R., Missaoui H., Alemany F., Mohamed Salah R., Garcia A., Ridha M., Othman J., Amor E.A., 2006. Spawning areas and larval distributions of anchovy *Engraulis encrasicolus* in relation to environmental conditions in the Gulf of Tunis (Central Mediterranean Sea). *Scientia Marina*, 70(2): 137–146.
- Zavodnik D., 1991. On the food and feeding in the northern Adriatic of *Pelagia noctiluca* (Scyphoa). In *Proceedings of the Workshop on Jellyfish Blooms in the Mediterranean*. Athens: UNEP: 212-216.
- Zhang A.B., Muster C., Liang H.B., 2012. A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. *Molecular Ecology*, 21: 1848-1863.
- Zhang J., Kapli P., Pavlidis P., Stamatakis A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29 (22): 2869–2876.

- 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. *Frontiers in Zoology*, 6: 30. <http://dx.doi.org/10.1186/1742-9994-6-30>.
- 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. *Journal of the North American Benthological Society*, 30: 131– 162. doi: 10.1899/10-010.1.
- Zhou X., Sun Y., Huang W., Smol J.P., Tang Q., Sun L. 2015. The Pacific decadal oscillation and changes in anchovy populations in the Northwest Pacific. *Journal of Asian Earth Sciences*, 14 (3): 204-511. doi:10.1016/j.jseaes.2015.06.027
- Zinger L., Gury J., Alibeu O., Rioux D., Gielly L., Sage L., Pompanon F., Geremia R.A., 2008. CE-SSCP and CE-FLA, simple and high- throughput alternatives for fungal diversity studies. *Journal of Microbiological Methods*, 72: 42–53.

APPENDICI

Appendice I - Dendrogramma circolare RaxML di 570 sequenze COX1 di Efemerotteri del presente dataset inclusi i barcode provenienti da BOLD. I differenti generi sono indicati con colori differenti ed il relativo nome.



Appendice II - Lista Macroinvertebrati bentonici campionati nelle 3 stazioni studiate.

Ortolano		Traponzo		Biedano	
Organismo	mm	Organismo	mm	Organismo	Mm
Gammaridae_1	7	Simuliidae_1	6	Tipulidae_1	25
Gammaridae_2	8	Simuliidae_2	5	Simuliidae_1	5
Gammaridae_3	7	Simuliidae_3	6	Simuliidae_2	2.5
Gammaridae_4	7	Simuliidae_4	4	Simuliidae_3	5
Gammaridae_5	8	Simuliidae_5	3	Simuliidae_4	7
Gammaridae_6	6	Simuliidae_6	2	Simuliidae_5	6.5
Gammaridae_7	5	Simuliidae_7	4	Simuliidae_6	6
Gammaridae_8	6	Simuliidae_8	3	Simuliidae_7	5
Gammaridae_9	6	Simuliidae_9	3	Simuliidae_8	7
Gammaridae_10	6	Simuliidae_10	4	Simuliidae_9	5
Gammaridae_11	7	Simuliidae_11	5	Simuliidae_10	6
Gammaridae_12	5	Hydropsychidae_1	15	Simuliidae_11	5.5
Gammaridae_13	5	Hydropsychidae_2	10	Simuliidae_12	6
Gammaridae_14	3	Hydropsychidae_3	7	Simuliidae_13	6
Gammaridae_15	3	Hydropsychidae_4	6	Simuliidae_14	5.5
Hydropsychidae_1	11	Hydropsychidae_5	5	Simuliidae_15	5
Hydropsychidae_2	10	Hydropsychidae_6	4	Simuliidae_16	6
Hydropsychidae_3	10	Hydropsychidae_7	4.5	Simuliidae_17	5
Hydropsychidae_4	3	Hydropsychidae_8	7	Simuliidae_18	5
Hydropsychidae_5	3	Hydropsychidae_9	12	Simuliidae_19	3
Hydropsychidae_6	4	Hydropsychidae_10	4	Hydropsychidae_1	5
Hydropsychidae_7	5	Hydropsychidae_11	6	Hydropsychidae_2	6
Hydropsychidae_8	4	Hydropsychidae_12	8	<i>Leuctra_1</i>	6
Hydropsychidae_9	4	Hydropsychidae_13	4	<i>Leuctra_2</i>	6
Hydropsychidae_10	5	Hydropsychidae_14	3.5	<i>Leuctra_3</i>	6
Hydropsychidae_11	3	Hydropsychidae_15	2.5	<i>Leuctra_4</i>	6
Hydropsychidae_12	5	Hydropsychidae_16	4.5	<i>Leuctra_5</i>	6
Hydropsychidae_13	7	Hydropsychidae_17	5	<i>Leuctra_6</i>	6
Hydropsychidae_14	7	Hydropsychidae_18	7	<i>Leuctra_7</i>	6
Hydropsychidae_15	3	Hydropsychidae_19	3.5	<i>Leuctra_8</i>	6
Hydropsychidae_16	4	Hydropsychidae_20	6	<i>Leuctra_9</i>	6
Hydropsychidae_17	3	Hydropsychidae_21	3	<i>Leuctra_10</i>	5
Hydropsychidae_18	2	Hydropsychidae_22	4	<i>Leuctra_11</i>	5
Philopotamidae_1	8	Hydropsychidae_23	4	<i>Leuctra_12</i>	6
Philopotamidae_2	9	Hydropsychidae_24	4	<i>Leuctra_13</i>	8
Sialidae_1	5	Hydropsychidae_25	5	<i>Leuctra_14</i>	4
<i>Protonemura_1</i>	8	Hydropsychidae_26	5.2	<i>Leuctra_15</i>	5
<i>Protonemura_2</i>	7.5	Hydropsychidae_27	3	<i>Leuctra_16</i>	6
<i>Leuctra_1</i>	8	Hydropsychidae_28	8.5	<i>Leuctra_17</i>	6
<i>Leuctra_2</i>	8	Hydropsychidae_29	4	<i>Leuctra_18</i>	6
<i>Leuctra_3_juv</i>	2.5	Hydropsychidae_30	4	<i>Leuctra_19</i>	6
<i>Leuctra_4_juv</i>	3	Hydropsychidae_31	7	<i>Leuctra_20</i>	6
<i>Ecdyonurus_1</i>	10	Hydropsychidae_32	9	<i>Leuctra_21</i>	5

<i>Ecdyonurus_2</i>	7	Hydropsychidae_33	6	<i>Leuctra_22</i>	4.5
<i>Ecdyonurus_3</i>	5.5	Hydropsychidae_34	6	<i>Leuctra_23</i>	5
<i>Ecdyonurus_4</i>	5	Hydropsychidae_35	6	<i>Leuctra_24</i>	8
<i>Ecdyonurus_5</i>	3.5	Hydropsychidae_36	3	<i>Leuctra_25</i>	8
<i>Ecdyonurus_6</i>	5	Hydropsychidae_37	3	<i>Leuctra_26</i>	5
<i>Ecdyonurus_7</i>	4	Hydropsychidae_38	4.5	<i>Leuctra_27</i>	6
<i>Ecdyonurus_8</i>	3.5	Hydropsychidae_39	4	<i>Leuctra_28</i>	7
<i>Electrogena grandiae_1</i>	6	Hydropsychidae_40	3	<i>Leuctra_29</i>	5
<i>Electrogena grandiae_2</i>	6	Hydropsychidae_41	3	<i>Leuctra_30</i>	4
<i>Electrogena grandiae_3</i>	7	<i>Onychogomphus_1</i>	26	<i>Ecdyonurus venosus_1</i>	7
<i>Electrogena grandiae_4</i>	6.5	<i>Onychogomphus_2</i>	15	<i>Baetis rhodani gr_1</i>	5
<i>Electrogena grandiae_5</i>	6.5	<i>Onychogomphus_3</i>	11	<i>Baetis rhodani gr_2</i>	5
<i>Electrogena grandiae_6</i>	5.5	<i>Onychogomphus_4</i>	4	<i>Baetis rhodani gr_3</i>	7
<i>Electrogena grandiae_7</i>	7.5	<i>Onychogomphus_5</i>	5	<i>Baetis rhodani gr_4</i>	7
<i>Electrogena grandiae_8</i>	7	<i>Onychogomphus_6</i>	7	<i>Baetis rhodani gr_5</i>	6
<i>Electrogena grandiae_9</i>	6.2	<i>Onychogomphus_7</i>	5	<i>Baetis rhodani gr_6</i>	5
Heptageniidae juv_1	2	<i>Onychogomphus_8</i>	6	<i>Baetis rhodani gr_7</i>	4
Heptageniidae juv_2	3	<i>Onychogomphus_9</i>	8	<i>Baetis rhodani gr_8</i>	3
Heptageniidae juv_3	2.5	<i>Onychogomphus_10</i>	3.5	<i>Baetis rhodani gr_9</i>	4
Heptageniidae juv_4	2.5	<i>Onychogomphus_11</i>	3.5	<i>Baetis rhodani gr_10</i>	6
Heptageniidae juv_5	3	<i>Onychogomphus_12</i>	9	<i>Baetis rhodani gr_11</i>	5
Heptageniidae juv_6	2.5	Chironomidae_1	4.5	<i>Baetis rhodani gr_12</i>	5
Heptageniidae juv_7	3	Chironomidae_2	7	<i>Baetis rhodani gr_13</i>	5
Heptageniidae juv_8	1.8	Chironomidae_3	8	<i>Habroleptoides_1</i>	3
Heptageniidae juv_9	2.2	Chironomidae_4	6	<i>Dugesia_1</i>	4
Heptageniidae juv_10	2.5	Chironomidae_5	5	<i>Dugesia_2</i>	4.5
Heptageniidae juv_11	2	Chironomidae_6	6	<i>Dugesia_3</i>	2
Heptageniidae juv_12	2.5	<i>Leuctra_1</i>	6	<i>Dugesia_4</i>	4
Heptageniidae juv_13	2	<i>Leuctra_2</i>	6.5	<i>Dugesia_5</i>	5
Heptageniidae juv_14	1.8	<i>Leuctra_3</i>	4.5	<i>Dugesia_6</i>	5
Heptageniidae juv_15	2.5	<i>Leuctra_4</i>	5	<i>Dugesia_7</i>	5
<i>Ephemera danica_1</i>	6	Tubificidae_1	15	<i>Dugesia_8</i>	5
<i>Ephemera danica_2</i>	6	Dryopidae_1	6	<i>Baetis buceratus_1</i>	5
<i>Ephemera danica_3</i>	7	Dryopidae_2	6	<i>Baetis buceratus_2</i>	5
<i>Habrophlebia eldae_1</i>	6	<i>Dugesia_1</i>	4	<i>Baetis buceratus_3</i>	7
<i>Habroleptoides sp_1</i>	5	<i>Dugesia_2</i>	4	<i>Baetis buceratus_4</i>	5
<i>Habroleptoides sp_2</i>	3	Hydraenidae	2	<i>Baetis buceratus_5</i>	4.5
<i>Habroleptoides sp_3</i>	2.5	<i>Heptagenia longicauda_1</i>	8	<i>Baetis buceratus_6</i>	5
<i>Habroleptoides sp_4</i>	3	<i>Heptagenia longicauda_2</i>	4	<i>Baetis buceratus_7</i>	5
<i>Habroleptoides sp_5</i>	2.5	<i>Baetis muticus_1</i>	4	<i>Baetis buceratus_8</i>	4
<i>Habroleptoides sp_6</i>	2	<i>Baetis rhodani gr_1</i>	3	<i>Baetis buceratus_9</i>	4
<i>Habroleptoides sp_7</i>	2.5	<i>Baetis rhodani gr_2</i>	4.5	<i>Baetis fuscatus_1</i>	8
<i>Habroleptoides sp_8</i>	3	<i>Baetis rhodani gr_3</i>	5	<i>Baetis fuscatus_2</i>	6
<i>Baetis muticus_1</i>	5	<i>Baetis fuscatus_1</i>	3	<i>Baetis fuscatus_3</i>	6
<i>Baetis muticus_2</i>	5	<i>Baetis fuscatus_2</i>	5	<i>Baetis fuscatus_4</i>	6
<i>Baetis muticus_3</i>	4	<i>Baetis fuscatus_3</i>	3.5	<i>Baetis fuscatus_5</i>	5.5
<i>Baetis rhodani gr_1</i>	5	<i>Baetis fuscatus_4</i>	4	<i>Baetis fuscatus_6</i>	6

Università Degli Studi della Tuscia di Viterbo
Corso di Dottorato di Ricerca – Ecologia e Gestione delle Risorse Biologiche

Athericidae_1	8	<i>Baetis fuscatus_5</i>	3	<i>Baetis fuscatus_7</i>	4.5
Athericidae_2	10	<i>Baetis fuscatus_6</i>	4	<i>Baetis fuscatus_8</i>	6
Athericidae_3	7	<i>Baetis fuscatus_7</i>	5	<i>Baetis fuscatus_9</i>	6
Athericidae_4	5	<i>Baetis fuscatus_8</i>	3.5	<i>Baetis fuscatus_10</i>	6
Athericidae_5	4	<i>Baetis buceratus_1</i>	3	<i>Baetis fuscatus_11</i>	6
<i>Procloeon pennulatum_1</i>	6	<i>Baetis buceratus_2</i>	5	<i>Baetis fuscatus_12</i>	4
<i>Procloeon pennulatum_2</i>	5	<i>Baetis buceratus_3</i>	4	<i>Baetis fuscatus_13</i>	4
Helodidae_Larva_1	4	<i>Baetis buceratus_4</i>	5	<i>Baetis fuscatus_14</i>	6
Helodidae_L_2	3	<i>Baetis buceratus_5</i>	5	<i>Baetis fuscatus_15</i>	5
Helodidae_L_3	3	<i>Baetis buceratus_6</i>	5	<i>Baetis fuscatus_16</i>	3.5
Helodidae_L_4	3	<i>Baetis buceratus_7</i>	5	<i>Baetis fuscatus_17</i>	4
Helodidae_L_5	3	<i>Baetis buceratus_8</i>	6	<i>Baetis fuscatus_18</i>	5
Helodidae_L_6	3	<i>Baetis buceratus_9</i>	6	<i>Baetis fuscatus_19</i>	6
Helodidae_L_7	3	<i>Baetis buceratus_10</i>	4	<i>Baetis fuscatus_20</i>	4
Helodidae_L_8	3	<i>Baetis buceratus_11</i>	5	<i>Baetis fuscatus_21</i>	3
Helodidae_L_9	2	<i>Baetis buceratus_12</i>	4	<i>Baetis fuscatus_22</i>	4
Helodidae_L_10	2	<i>Baetis buceratus_13</i>	5	<i>Baetis fuscatus_23</i>	4
Helodidae_L_11	2	<i>Baetis buceratus_14</i>	5	<i>Baetis fuscatus_24</i>	5
Helodidae_L_12	2	<i>Baetis buceratus_15</i>	3.5	<i>Baetis fuscatus_25</i>	5
Helodidae_L_13	2	<i>Baetis buceratus_16</i>	5	<i>Baetis fuscatus_26</i>	5.5
Helodidae_L_14	2	<i>Baetis buceratus_17</i>	4	<i>Baetis fuscatus_27</i>	4
Helodidae_L_15	3	<i>Baetis buceratus_18</i>	3	<i>Baetis fuscatus_28</i>	5
Helodidae_L_16	3	<i>Baetis lutheri_1</i>	5	<i>Baetis fuscatus_29</i>	5
Helodidae_L_17	3	<i>Caenis macrura gr_1</i>	3	<i>Baetis fuscatus_30</i>	5
Helodidae_L_18	3	<i>Caenis macrura gr_2</i>	4	<i>Baetis fuscatus_31</i>	6
Helodidae_L_19	2	<i>Caenis macrura gr_3</i>	3	<i>Baetis fuscatus_32</i>	6
Helodidae_L_20	2	<i>Caenis macrura gr_4</i>	3	<i>Baetis fuscatus_33</i>	5
Helodidae_L_21	2	<i>Caenis macrura gr_5</i>	3	<i>Baetis fuscatus_34</i>	4
Dryopidae_Adulto_1	5	<i>Caenis pusilla_1</i>	4	<i>Baetis fuscatus_35</i>	4
Hydraenidae_Ad_1	5	<i>Caenis pusilla_2</i>	4	<i>Baetis fuscatus_36</i>	4
Elmidae_Ad_1	3	<i>Caenis pusilla_3</i>	4	<i>Baetis fuscatus_37</i>	5.5
Elmidae_Ad_2	3	<i>Caenis pusilla_4</i>	3.5	<i>Caenis pusilla_1</i>	4
Elmidae_Ad_3	3	<i>Caenis pusilla_5</i>	4	Tricottero_juv_1	1
Elmidae_Ad_4	3	<i>Caenis pusilla_6</i>	3	Tricottero_juv_2	3
Elmidae_Ad_5	2	<i>Caenis pusilla_7</i>	3	<i>Onychogomphus_1</i>	3
Elmidae_Ad_6	3	<i>Caenis pusilla_8</i>	3	<i>Onychogomphus_2</i>	3
<i>Dugesia_1</i>	4	<i>Caenis pusilla_9</i>	3	<i>Onychogomphus_3</i>	14
<i>Dugesia_2</i>	5	<i>Caenis pusilla_10</i>	4	<i>Onychogomphus_4</i>	8
<i>Dugesia_3</i>	4	<i>Caenis pusilla_11</i>	4	Athericidae_1	4
<i>Dugesia_4</i>	6	<i>Caenis pusilla_12</i>	4	Helodidae_1	1.5
<i>Dugesia_5</i>	2	<i>Caenis pusilla_13</i>	4	Tabanidae_1	7
Hydrobiidae_1	4	<i>Caenis pusilla_14</i>	3	Elmidae_1	2
Hydrobiidae_2	3.5	<i>Caenis pusilla_15</i>	3.5	Elmidae_2	5
Lumbricidae_1	10	<i>Caenis pusilla_16</i>	3	Elmidae_3	3
Lumbricidae_2	10	<i>Caenis pusilla_17</i>	4.5	Elmidae_4	2
		<i>Caenis pusilla_18</i>	3	Elmidae_5_Ad	2
		<i>Caenis pusilla_19</i>	3.5	Elmidae_6_Ad	2

	<i>Caenis pseudorivulorum</i> gr_1	4	Hydrobiidae_1	5
			Hydrobiidae_2	5
			Hydrobiidae_3	4.5
			Hydrobiidae_4	4.5
			Hydrobiidae_5	2.5
			Hydrobiidae_6	4
			Hydrobiidae_7	3.5
			Hydrobiidae_8	4
			Hydrobiidae_9	5
			Hydrobiidae_10	4
			Hydrobiidae_11	5
			Hydrobiidae_12	3
			Hydrobiidae_13	2
			Hydrobiidae_14	4
			Hydrobiidae_15	5
			Hydrobiidae_16	2.5
			Dryopidae_Ad_1	5.5
			Dryopidae_Ad_2	5.5
			Dryopidae_Ad_3	5
			Dryopidae_Ad_4	6
			Dryopidae_Ad_5	6
			Dryopidae_Ad_6	6
			Dryopidae_Ad_7	6
			Dryopidae_Ad_8	6
			Dryopidae_Ad_9	5
			Dryopidae_Ad_10	5.5
			Dryopidae_Ad_11	6
			Dryopidae_Ad_12	6
			Dryopidae_Ad_13	6
			Dryopidae_Ad_14	6
			Dryopidae_Ad_15	6
			Dryopidae_Ad_16	5.5
			Dryopidae_Ad_17	5
			Dryopidae_Ad_18	5
			Dryopidae_Ad_19	5
			Dryopidae_Ad_20	6
			Dryopidae_Ad_21	5
			Dryopidae_Ad_22	5
			Dryopidae_Ad_23	6
			Dryopidae_Ad_24	6
			Hydraenidae_Ad_1	2
			Hydraenidae_Ad_2	2
			Hydraenidae_Ad_3	2
			Hydraenidae_Ad_4	2
			Hydraenidae_Ad_5	2

Appendice III - (A) Lista degli aplotipi generata con le prime 206 bp della regione 5'-COX1, (B) lista degli aplotipi generata con le ultime 206 bp della regione 3' - COX1

A		B	
Specie presente	Aplotipo	Specie presente	Aplotipo
Baetis_buceratus_MO3	1	Baetis_buceratus_MO3	1
Baetis_buceratus_MO4	1	Baetis_buceratus_MO4	1
Baetis_buceratus_RIO	1	Baetis_buceratus_RIO	1
Baetis_buceratus_MO2	1	Baetis_buceratus_MO2	1
Baetis_buceratus_SO	1	Baetis_buceratus_SO	1
Baetis_cyrneus_1pvg	2	Baetis_cyrneus_1pvg	2
Baetis_cyrneus_2pvg	2	Baetis_cyrneus_2pvg	2
Baetis_cyrneus_4pvg	2	Baetis_cyrneus_4pvg	2
Baetis_cyrneus_SA	3	Baetis_cyrneus_SA	3
Baetis_lutheri_AT72	4	Baetis_lutheri_AT72	4
Baetis_lutheri_MP62	4	Baetis_lutheri_MP62	4
Baetis_fuscatus_SA	5	Baetis_fuscatus_SA	5
Baetis_fuscatus_SD	5	Baetis_fuscatus_SD	5
Baetis_alpinus_FG	6	Baetis_alpinus_FG	6
Baetis_alpinus_OB	6	Baetis_alpinus_OB	6
Baetis_alpinus_AN	7	Baetis_alpinus_AN	6
Baetis_albinatii_sa	8	Baetis_albinatii_sa	7
Baetis_vernus_TR80	9	Baetis_vernus_TR80	8
Baetis_muticus_vi	10	Baetis_muticus_vi	9
Cloeon_dipterum_AT	11	Cloeon_dipterum_AT	10
Cloeon_dipterum_MV75	11	Cloeon_dipterum_MV75	10
Cloeon_simile_SB	12	Cloeon_simile_SB	11
Centroptilum_luteolum_FG	13	Centroptilum_luteolum_FG	12
Centroptilum_luteolum_FT	13	Centroptilum_luteolum_FT	12
Centroptilum_luteolum_RA	13	Centroptilum_luteolum_RA	12
Procloeon_bifidum_MR74	14	Procloeon_bifidum_MR74	13
Procloeon_bifidum_ESV	15	Procloeon_bifidum_ESV	14
Procloeon_bifidum_SC	15	Procloeon_bifidum_SC	14
Procloeon_pulchrum_OR	16	Procloeon_pulchrum_OR	15
Procloeon_pulchrum_RC	17	Procloeon_pulchrum_RC	15
Procloeon_pulchrum_VI	17	Procloeon_pulchrum_VI	15
Brachycercus_harrisella_RC	18	Brachycercus_harrisella_RC	16
Brachycercus_harrisella_VI	18	Brachycercus_harrisella_VI	16
Caenis_beskidensis_RC	19	Caenis_beskidensis_RC	17
Caenis_beskidensis_AT73	19	Caenis_beskidensis_AT73	17
Caenis_beskidensis_VI	19	Caenis_beskidensis_VI	17
Caenis_martae_MO	20	Caenis_martae_MO	18
Caenis_martae_OR	20	Caenis_martae_OR	18
Caenis_martae_RIO	20	Caenis_martae_RIO	18
Caenis_pusilla_AT	21	Caenis_pusilla_AT	19
Caenis_pusilla_MR78	21	Caenis_pusilla_MR78	19
Caenis_sp._SA98	22	Caenis_sp._SA98	20
Olgoneuriella_rheana_RC	23	Olgoneuriella_rheana_RC	21
Olgoneuriella_rhena_MP	23	Olgoneuriella_rhena_MP	21
Paraleptophlebia_submarginata_AT	24	Paraleptophlebia_submarginata_AT	22
Paraleptophlebia_submarginata_VI	24	Paraleptophlebia_submarginata_VI	22
Paraleptophlebia_submarginata_VI36	24	Paraleptophlebia_submarginata_VI36	22
Habroleptoides_confusa_FG	25	Habroleptoides_confusa_FG	23

Habroleptoides_confusa_OR	25	Habroleptoides_confusa_OR	23
Habroleptoides_confusa_RA	25	Habroleptoides_confusa_RA	23
Habroleptoides_spSB	26	Habroleptoides_spSB	24
Habrophlebia_eldae_FA	27	Habrophlebia_eldae_RA	25
Habrophlebia_eldae_RA	27	Habrophlebia_eldae_FA	25
Habrophlebia_eldae_VI	27	Habrophlebia_eldae_VI	25
Choroterpes_picteti_MR77	28	Choroterpes_picteti_MR77	26
Heptagenia_longicauda_RC	29	Heptagenia_longicauda_RC	27
Heptagenia_longicauda_AT	29	Heptagenia_longicauda_AT	27
Heptagenia_longicauda_RIG	29	Heptagenia_longicauda_RIG	27
Ecdyonurus_venosus_OR76	30	Ecdyonurus_venosus_OR	28
Ecdyonurus_venosus_IN	30	Ecdyonurus_venosus_IN	28
Ecdyonurus_venosus_OR	30	Ecdyonurus_venosus_OR76	28
Ecdyonurus_venosus_PM	30	Ecdyonurus_venosus_PM	28
Ecdyonurus_helveticus_MC	31	Ecdyonurus_helveticus_FT	29
Ecdyonurus_helveticus_FT	31	Ecdyonurus_helveticus_MC	29
Ecdyonurus_helveticus_OR	31	Ecdyonurus_helveticus_OR	29
Ecdyonurus_corsicus_SB	32	Ecdyonurus_corsicus_SB	30
Electrogena_zebrata_SA104	33	Electrogena_zebrata_SA104	31
Electrogena_zebrata_SA105	33	Electrogena_zebrata_SA105	31
Electrogena_grandiae_IN	34	Electrogena_grandiae_IN	32
Electrogena_grandiae_RA	35	Electrogena_grandiae_RA	32
Rhithrogena_insularis_cor100	36	Rhithrogena_insularis_cor100	33
Rhithrogena_insularis_cor101	36	Rhithrogena_insularis_cor101	33
Rhithrogena_nuragica_3	37	Rhithrogena_nuragica_3	34
Rhithrogena_nuragica_SC	37	Rhithrogena_nuragica_SC	34
Rhithrogena_nuragica_2	38	Rhithrogena_nuragica_2	35
Rhithrogena_reatina_CA	39	Rhithrogena_reatina_CA	36
Rhithrogena_reatina_VMA79	39	Rhithrogena_reatina_VMA79	36
Rhithrogena_eatoni_cor103	40	Rhithrogena_eatoni_cor103	37
Rhithrogena_semicolorata_FT	41	Rhithrogena_semicolorata_FT	38
Rhithrogena_semicolorata_PM	41	Rhithrogena_semicolorata_PM	38
Rhithrogena_semicolorata_RS	41	Rhithrogena_semicolorata_RS	38
Epeorus_assimilis_OR	42	Epeorus_assimilis_OR	39
Epeorus_assimilis_AN	43	Epeorus_assimilis_AN	40
Epeorus_assimilis_OB	43	Epeorus_assimilis_OB	40
Serratella_ignita_OR	44	Serratella_ignita_RA	41
Serratella_ignita_RA	44	Serratella_ignita_OR	41
Serratella_ignita_FA	44	Serratella_ignita_FA	41
Serratella_ignita_VI	44	Serratella_ignita_VI	41
Siphonurus_lacustris_VMA	45	Siphonurus_lacustris_VI	42
Siphonurus_lacustris_VI	46	Siphonurus_lacustris_VMA	42
Ephemera_danica_VI	47	Ephemera_danica_VI	43

Appendice IV - Pubblicazione relativa al primo caso di studio (Efemerotteri).

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.

© 2015 Elsevier Ltd. All rights reserved.

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/© 2015 Elsevier Ltd. All rights reserved.

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 assignment 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 subarctic 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

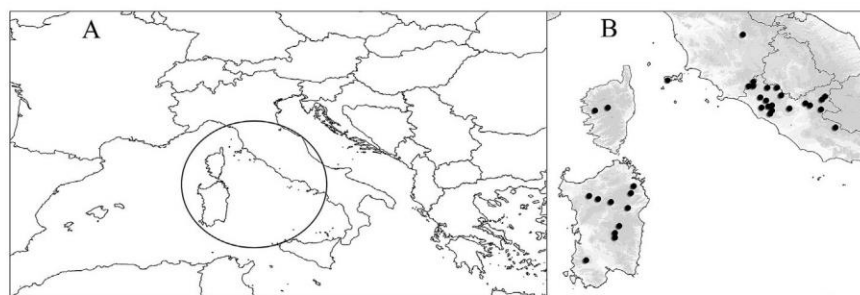


Fig. 1. Geographical area (A) and distribution map (B) of the investigated sampling sites (dotted).

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

(continued on next page)

Table 1 (continued)

Family	Genus	Species	Location	Accession number
Leptophlebiidae	<i>Habroleptoides</i>	<i>confusa</i> Sartori & Jacob	Fosso di Galantina, Latium	LN734747
			Rafanello, Latium	LN734748
			Ortolano, Latium	LN734749
	<i>Habrophlebia</i>	sp. <i>eldae</i> Jacob & Sartori	Rio Tistigliosi, Sardinia	LN734750
			Vitozza, Tuscany	LN734751
			Fosso Arlena, Latium	LN734752
			Rafanello, Latium	LN734753
	<i>Paralaptophlebia</i>	<i>submarginata</i> Stephens	Aquarella Tuscania, Latium	LN734754
			Vitozza, Tuscany, specimen 1	LN734755
	Ephemerellidae	<i>Choroterpes</i>	<i>picteti</i> Eaton	Mignone Montericcio, Latium
<i>Serratella</i>				<i>ignita</i> Poda
<i>Serratella</i>		<i>ignita</i> Poda	Vitozza, Tuscany	LN734758
			Rafanello, Latium	LN734759
			Fosso Arlena, Latium	LN734760
Oligoneuridae	<i>Oligoneuriella</i>	<i>rhenana</i> Imhoff	Ortolano, Latium	LN734761
			Meleta Pitigliano, Tuscany	LN734762
Siphonuridae	<i>Siphonurus</i>	<i>lacustris</i> Eaton	Rio Chiaro, Latium	LN734763
			Vitozza, Tuscany	LN734764
Ephemeridae	<i>Ephemera</i>	<i>danica</i> Müller	Velino AnTRODoco, Latium	LN734765
			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 1). 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 *Epeorurus 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 *Epeorurus* 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	
<i>Baetis cyrneus</i>	0.000	0.149	0.184	0.282	0.035
<i>Baetis lutheri</i>	0.007	0.007	0.217	0.300	0.211
<i>Baetis fuscatus</i>	0.000	0.000	0.200	0.296	0.200
<i>Baetis alpinus</i>	0.015	0.067	0.184	0.300	0.117
<i>Baetis buceratus</i>	0.000	0.000	0.226	0.293	0.226
<i>Procloeon pulchrum</i>	0.002	0.056	0.187	0.212	0.131
<i>Procloeon bifidum</i>	0.027	0.196	0.187	0.212	-0.009
<i>Caenis pusilla</i>	0.005	0.005	0.223	0.263	0.218
<i>Caenis beskidensis</i>	0.000	0.000	0.223	0.330	0.223
<i>Caenis martaе</i>	0.002	0.030	0.103	0.330	0.073
<i>Ecdyonurus helveticus</i>	0.020	0.028	0.212	0.225	0.184
<i>Ecdyonurus venosus</i>	0.003	0.010	0.117	0.225	0.107
<i>Electrogena grandiae</i>	0.000	0.000	0.240	0.244	0.240
<i>Electrogena zebrata</i>	0.005	0.005	0.240	0.244	0.235
<i>Rhithrogena semicolorata</i>	0.015	0.018	0.180	0.205	0.162
<i>Rhithrogena reatina</i>	0.005	0.005	0.082	0.199	0.077
<i>Rhithrogena insularis</i>	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 *Procloeon bifidum*; unexpectedly, five of the eight genera where multiple species were analyzed appeared as largely paraphyletic, with only *Ecdyonurus*, *Rhithrogena* and *Habro-leptoides* 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.

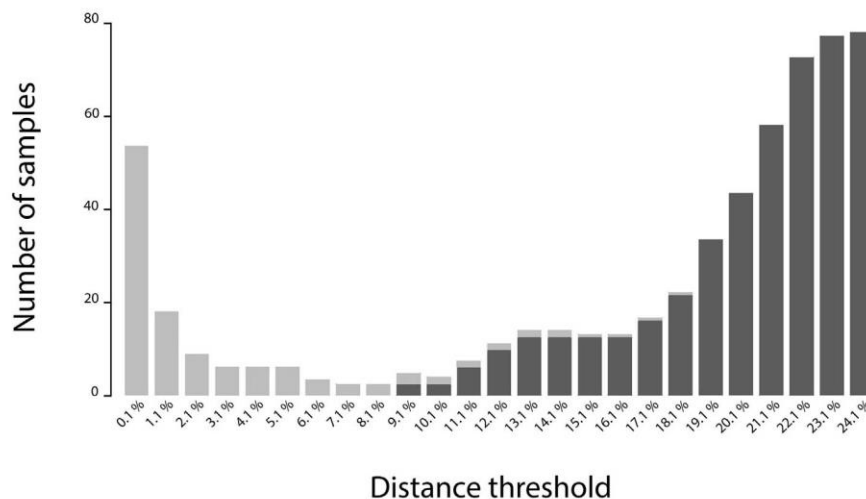


Fig. 2. Barplot showing the false positive (dark grey) and false negative (light grey) rate of mayfly identification in the present study based on COI K2P genetic distances.

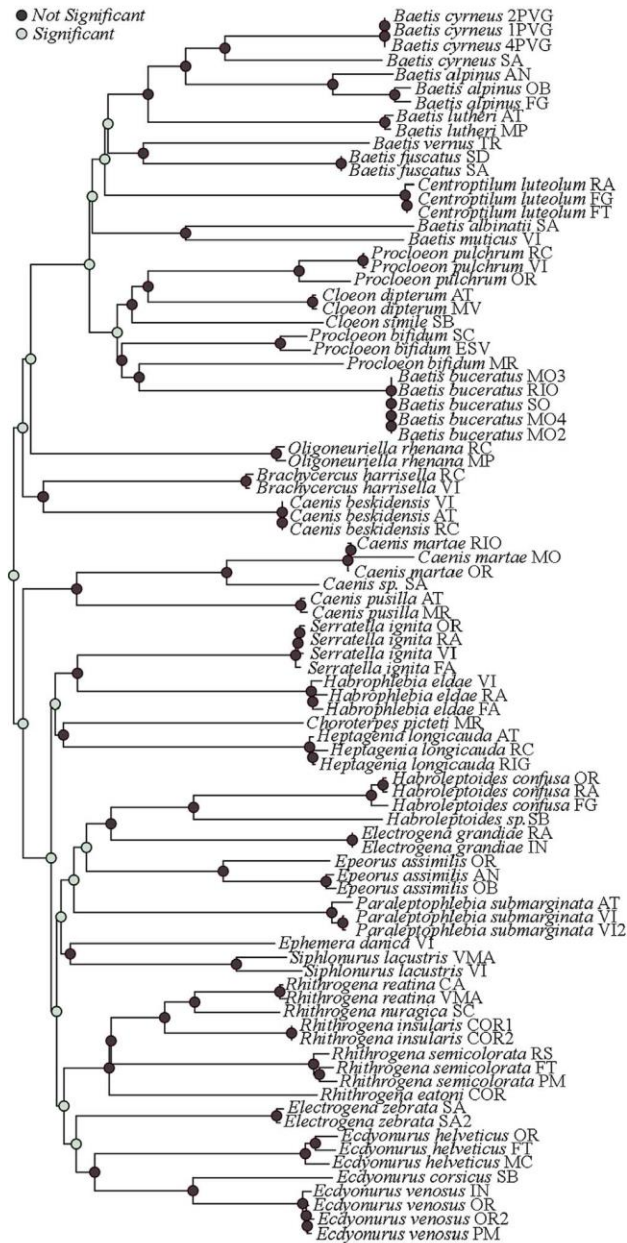


Fig. 3. Neighbor Joining dendrogram of 89 mayfly COI sequences produced with SPIDER under default options and the Rosenberg probability test of reciprocal monophyly ($\alpha = 0.005$). Light grey nodes are significant to 0.05.

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 *Ephemerella 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	
[<i>Baetis muticus_VI</i>]	Yes*	<i>Alainites cf. muticus</i>	98%	Yes*	<i>Alainites cf. muticus</i>	97.88%
[<i>Baetis vernus_TR</i>]	Yes	<i>Baetis liebenauae</i>	94%	Yes	<i>Baetis vernus</i>	96.73%
[<i>Baetis albinatii_SA</i>]	No	<i>Alainites cf. muticus</i>	99%	No	<i>Alainites cf. muticus</i>	99.35%
[<i>Baetis cyrneus_SA</i>]	Yes	<i>Baetis cf. cyrneus</i>	90%	Yes	<i>Baetis cf. cyrneus</i>	90.03%
[<i>Baetis cyrneus_4PVG, 1PVG, MO</i>]	Yes	<i>Baetis cf. cyrneus</i>	100%	Yes	<i>Baetis cf. cyrneus</i>	100%
[<i>Baetis lutheri_MP</i>]	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	94.61%
[<i>Baetis lutheri_AT</i>]	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	94.61%
[<i>Baetis fuscatus_SA, SD</i>]	Yes	<i>Baetis fuscatus clone</i>	89%	Yes	<i>Baetis fuscatus clone</i>	91.34%
[<i>Baetis alpinus_FG</i>]	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	95.09%
[<i>Baetis alpinus_AN</i>]	Yes	<i>Baetis alpinus</i>	97%	Yes	<i>Baetis alpinus</i>	96.79%
[<i>Baetis alpinus_OB</i>]	Yes	<i>Baetis alpinus</i>	95%	Yes	<i>Baetis alpinus</i>	95.60%
[<i>Baetis buceratus_MO3, RIO, SO, MO2, MO4</i>]	Yes	<i>Baetis buceratus</i>	89%	Yes	<i>Baetis buceratus</i>	90.36%
[<i>Centroptilum luteolum_RA</i>]	Yes	<i>Centroptilum luteolum</i>	89%	Yes	<i>Centroptilum luteolum</i>	89.22%
[<i>Centroptilum luteolum_FT, FG</i>]	Yes	<i>Centroptilum luteolum</i>	89%	Yes	<i>Centroptilum luteolum</i>	89.39%
[<i>Procloeon pulchrum_VI</i>]	No	<i>Procloeon mendax</i>	85%	No	<i>Cloeon dipterum</i>	84.94%
[<i>Procloeon pulchrum_OR</i>]	No	<i>Procloeon mendax</i>	86%	No	<i>Cloeon dipterum</i>	86.19%
[<i>Procloeon pulchrum_RC</i>]	No	<i>Procloeon mendax</i>	85%	No	<i>Cloeon dipterum</i>	84.77%
[<i>Procloeon bifidum_MR</i>]	No	<i>Procloeon mendax</i>	84%	Yes	<i>Procloeon bifidum</i>	86.21%
[<i>Procloeon bifidum_ES</i>]	No	<i>Procloeon mendax</i>	85%	Yes	<i>Procloeon viridiculare</i>	86.11%
[<i>Procloeon bifidum_SC</i>]	No	Unclassified Ephemeroptera	86%	Yes	<i>Procloeon viridiculare</i>	86.11%
[<i>Cloeon dipterum_MV</i>]	Yes	<i>Cloeon dipterum</i>	99%	Yes	<i>Cloeon dipterum</i>	99.82%
[<i>Cloeon dipterum_AT</i>]	Yes	<i>Cloeon dipterum</i>	99%	Yes	<i>Cloeon dipterum</i>	99.82%
[<i>Cloeon simile_SB</i>]	No	<i>Procloeon sp.</i>	90%	Yes	<i>Procloeon sp.</i>	90.78%
[<i>Caenis pusilla_MR</i>]	No	<i>Caenis anceps</i>	82%	Yes	<i>Caenis pusilla</i>	98.20%
[<i>Caenis pusilla_AT</i>]	No	<i>Caenis sp.</i>	82%	Yes	<i>Caenis pusilla</i>	98.20%
[<i>Caenis beskidensis_VI, RC, AT</i>]	No	<i>Caenis anceps</i>	83%	Yes	<i>Caenis beskidensis</i>	89.77%
[<i>Caenis martae_MO</i>]	No	<i>Leptophlebia</i>	77%	Yes	<i>Caenis macrura</i>	79.84%
[<i>Caenis martae_OR</i>]	No	<i>Caenis bajaensis</i>	80%	Yes	<i>Caenis macrura</i>	81.54%
[<i>Caenis martae_RIO</i>]	No	<i>Caenis bajaensis</i>	81%	Yes	<i>Caenis macrura</i>	81.70%
[<i>Caenis sp._SA98</i>]	No	<i>Asionurus sp.</i>	80%	No	<i>Caenis</i>	82.68%
[<i>Brachycercus harrisella_VI</i>]	Yes	<i>Brachycercus harrisella</i>	94%	Yes	<i>Brachycercus harrisella</i>	96.09%
[<i>Brachycercus harrisella_RC</i>]	Yes	<i>Brachycercus harrisella</i>	94%	Yes	<i>Brachycercus harrisella</i>	96.24%
[<i>Ecdyonurus helveticus_FT</i>]	Yes	<i>Ecdyonurus helveticus</i>	99%	Yes	<i>Ecdyonurus helveticus</i>	98.69%
[<i>Ecdyonurus helveticus_OR</i>]	Yes	<i>Ecdyonurus helveticus</i>	98%	Yes	<i>Ecdyonurus helveticus</i>	98.20%
[<i>Ecdyonurus helveticus_MC</i>]	Yes	<i>Ecdyonurus helveticus</i>	98%	Yes	<i>Ecdyonurus helveticus</i>	98.37%
[<i>Ecdyonurus corsicus_SB</i>]	Yes	<i>Ecdyonurus corsicus</i>	99%	Yes	<i>Ecdyonurus corsicus</i>	99.67%
[<i>Ecdyonurus venosus_OR</i>]	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.67%
[<i>Ecdyonurus venosus_IN</i>]	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.84%
[<i>Ecdyonurus venosus_PM</i>]	Yes	<i>Ecdyonurus venosus</i>	99%	Yes	<i>Ecdyonurus venosus</i>	99.51%
[<i>Ecdyonurus venosus_OR76</i>]	Yes	<i>Ecdyonurus venosus</i>	100%	Yes	<i>Ecdyonurus venosus</i>	100%
[<i>Epeorus assimilis_OB</i>]	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	89.70%
[<i>Epeorus assimilis_OR</i>]	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	89.87%
[<i>Epeorus assimilis_AN</i>]	No	<i>Epeorus sylvicola</i>	89%	Yes	<i>Epeorus assimilis</i>	90.70%
[<i>Electrogena grandie_IN, RA</i>]	No	<i>Electrogena sp.</i>	98%	No	<i>Electrogena sp.</i>	97.55%
[<i>Electrogena zebrata_SA104</i>]	Yes	<i>Electrogena zebrata</i>	99%	Yes	<i>Electrogena zebrata</i>	99.84%
[<i>Electrogena zebrata_SA105</i>]	Yes	<i>Electrogena zebrata</i>	99%	Yes	<i>Electrogena zebrata</i>	99.67%
[<i>Rhithrogena semicolorata_PM</i>]	Yes	<i>Rhithrogena semicolorata</i>	98%	Yes	<i>Rhithrogena semicolorata</i>	98.37%
[<i>Rhithrogena semicolorata_FT</i>]	Yes	<i>Rhithrogena semicolorata</i>	99%	Yes	<i>Rhithrogena semicolorata</i>	98.69%
[<i>Rhithrogena semicolorata_RS</i>]	Yes	<i>Rhithrogena semicolorata</i>	99%	Yes	<i>Rhithrogena semicolorata</i>	99.84%
[<i>Rhithrogena reatina_VMA79</i>]	No	<i>Rhithrogena sp.</i>	93%	No	<i>Rhithrogena sp.</i>	92.81%
[<i>Rhithrogena reatina_CA</i>]	No	<i>Rhithrogena sp.</i>	93%	No	<i>Rhithrogena sp.</i>	93.14%
[<i>Rhithrogena nuragica_SC</i>]	No	<i>Rhithrogena sp.</i>	99%	No	<i>Rhithrogena sp.</i>	99.02%
[<i>Rhithrogena insulari_COR100, COR101</i>]	No	<i>Rhithrogena sp.</i>	90%	No	<i>Rhithrogena sp.</i>	90.03%
[<i>Rhithrogena eatoni_OR103</i>]	No	<i>Rhithrogena sp.</i>	90%	No	<i>Rhithrogena sp.</i>	89.87%
[<i>Heptagenia longicauda_RIG</i>]	No	<i>Heptagenia solitaria</i>	89%	Yes	<i>Eptagenia longicauda</i>	95.42%
[<i>Heptagenia longicauda_AT</i>]	No	<i>Heptagenia solitaria</i>	89%	Yes	<i>Eptagenia longicauda</i>	95.75%
[<i>Heptagenia longicauda_RC</i>]	No	<i>Heptagenia sulfura</i>	87%	Yes	<i>Eptagenia longicauda</i>	94.77%
[<i>Habroleptoides confusa_FG</i>]	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	96.50%
[<i>Habroleptoides confusa_RA</i>]	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	96.83%
[<i>Habroleptoides confusa_OR</i>]	Yes	<i>Habroleptoides confusa</i>	84%	Yes	<i>Habroleptoides confusa</i>	97.17%
[<i>Habroleptoides sp.SB</i>]	Yes	<i>Habroleptoides modesta</i>	92%	Yes	<i>Habroleptoides modesta</i>	91.67%
[<i>Habrophlebia eldae_VI</i>]	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.35%
[<i>Habrophlebia eldae_FA</i>]	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.51%
[<i>Habrophlebia eldae_RA</i>]	Yes	<i>Habrophlebia eldae</i>	99%	Yes	<i>Habrophlebia eldae</i>	99.35%
[<i>Paraleptophlebia submarginata_AT</i>]	No	<i>Paraleptophlebia westoni</i>	82%	Yes	<i>Paraleptophlebia submarginata</i>	99.18%
[<i>Paraleptophlebia submarginata_VI</i>]	No	<i>Siphonurus sp.</i>	81%	Yes	<i>Paraleptophlebia submarginata</i>	98.69%
[<i>Paraleptophlebia submarginata_VI36</i>]	No	<i>Siphonurus sp.</i>	81%	Yes	<i>Paraleptophlebia submarginata</i>	98.85%

(continued on next page)

Table 3 (continued)

Sequence	Presence	NCBI Species match	Identity	Presence	BOLD Species match	Identity
[Choroterpes_picteti_MR77]	No	<i>Habroleptoides</i> sp.	85%	Yes	<i>Choroterpes picteti</i>	99.35%
[Serratella_ignita_VI]	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.67%
[Serratella_ignita_RA]	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.84%
[Serratella_ignita_FA]	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.51%
[Serratella_ignita_OR]	Yes	<i>Serratella ignita</i>	99%	Yes	<i>Serratella ignita</i>	99.84%
[Oligoneuriella_rhenana_MP]	No	<i>Anopheles labranchiae</i>	85%	Yes	<i>Oligoneuriella rhenana</i>	89.30%
[Oligoneuriella_rhenana_RC]	No	<i>Anopheles labranchiae</i>	84%	Yes	<i>Oligoneuriella rhenana</i>	90%
[Siphonurus_lacustris_VI]	Yes	<i>Siphonurus lacustris</i>	92%	Yes	<i>Siphonurus croaticus</i>	94.28%
[Siphonurus_lacustris_VMA]	Yes	<i>Siphonurus lacustris</i>	93%	Yes	<i>Siphonurus croaticus</i>	95.10%
[Ephemera_danica_VI]	Yes	<i>Ephemera sachalinensis</i> **	89%	Yes	<i>Ephemera danica</i>	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), *Siphonurus lacustris* (matching *Siphonurus 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](#)).

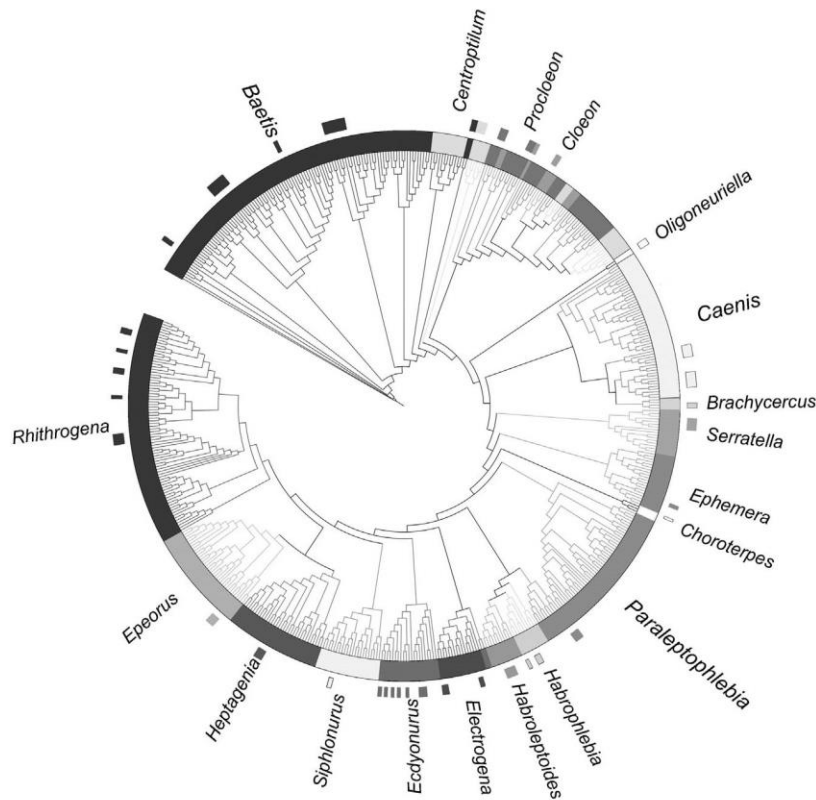


Fig. 4. Circular RaxML dendrogram of 570 mayfly COI sequences including BOLD barcodes and the investigated dataset. Bars indicate the positioning of sequences determined in this study. Different genus clusters are indicated with greyscale and the respective names.

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 Siphonuridae. 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 assignment 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 *Procleon bifidum* (separated on different macroclusters), also *Epeorus assimilis*, *Baetis alpinus*, *Procleon pulchrum*, and *Siphonurus 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 *Procleon* 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 *Procleon* 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 (*Procleon* 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 zibrata* might belong to different genera, as previously promoted by Gaino et al. (1987) and discussed by Bauerfeind and Soldan (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. zibrata* COI barcodes clustered separately from the other *Electrogena* species. Notably, egg and winged stage traits of *E. zibrata* (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., Gejjer, 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 cytochrome 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. 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-Terter, 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 Konyenburg, 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-idiiosyncratic error rates, p. e14448. *PLoS ONE* 5.
- Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K., Verdonshot, 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 macroinvertebrate 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 taxonator: 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* fig 1, 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., Hallwachs, W., Hebert, P.D.N., Hrecek, 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 palaeartic 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](http://dx.doi.org/10.1016/S0169-5347(02)00041-1).
- Theissingner, 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: Siphonuridae) 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., Ruitter, 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>.