

SVEU ILIŠTE U ZAGREBU
PRIRODOSLOVNO - MATEMATI KI FAKULTET
BIOLOŠKI ODSJEK

BARKODIRANJE ŽIVOTA
THE BARCODING OF LIFE

SEMINARSKI RAD

Bruno Polak

Preddiplomski studij molekularne biologije
(Undergraduate Study of Molecular Biology)

Mentor: Doc. dr. sc. Damjan Franjevi

Zagreb, 2012.

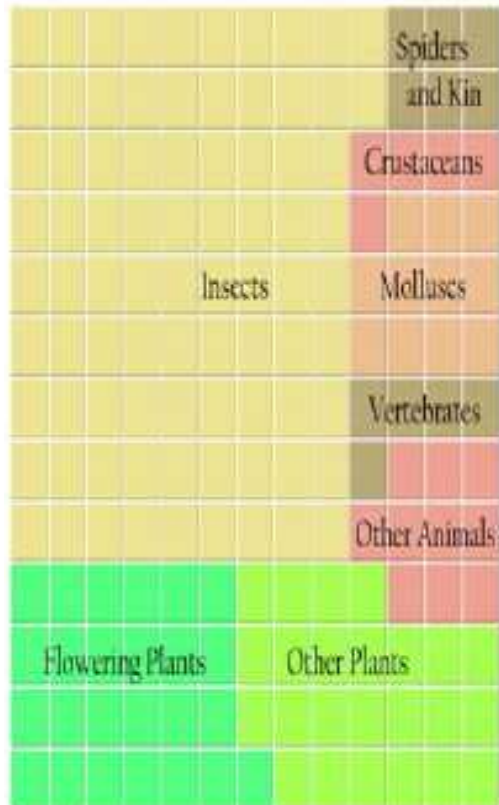
SADRŽAJ

1. UVOD	2
2. RAZVOJ, NAPREDAK I PRIMJENE DNA BARKODIRANJA	4
3. SUSTAV PODATAKA ZA BARKODIRANJE ŽIVOTA - BOLD (BARCODE OF LIFE DATA SYSTEM)	6
4. PO ECI UPOTREBE DNA BARKODOVA	8
5. NEKI PRIMJERI BARKODIRANJA	10
5.1. BARKODIRANJE DIPTERA	10
5.1.1. Chironomidae	10
5.1.2. Tachinidae	11
5.2. BARKODIRANJE EPHEMEROPTERA	13
5.2.1. <i>Baetis vernus</i>	13
5.3. BARKODIRANJE LEPIDOPTERA	13
5.3.1. <i>Hesperiidae</i>	13
6. BARKODIRANJE BILJAKA I GLJIVA	17
7. BARKODIRANJE RODA <i>HOMO</i>	18
8. NEDOSTACI BARKODIRANJA	18
9. SAŽETAK	19
10. SUMMARY	20
11. LITERATURA	21

1. UVOD

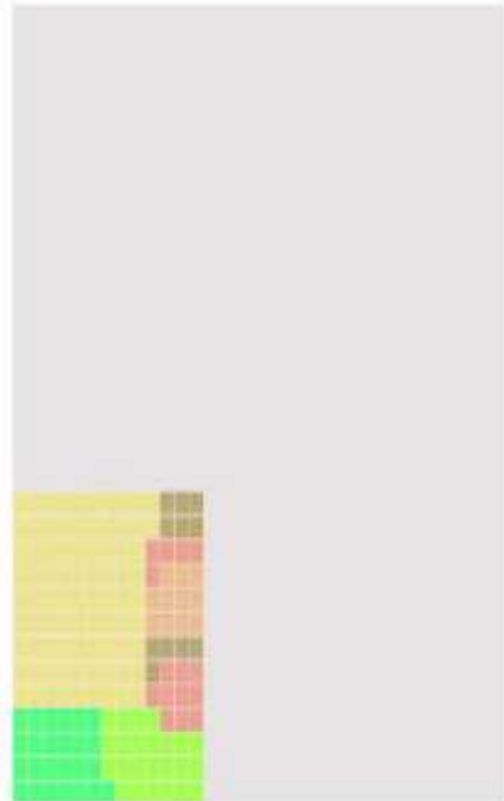
Biološka istraživanja uglavnom ovise o determinaciji vrsta te je stoga važno dodatno proširiti analize, a ne bazirati se isključivo na taksonomskim proučavanjima. Rutinsko određivanje vrsta nailazi na nekoliko nedostataka. Fenotipska plastičnost i genetička varijabilnost u karakteristikama presudnim za determinaciju vrsta mogu rezultirati pogrešnom analizom. Moguće je i previdjeti morfološki kriptične svojste prisutne u mnogim skupinama. Razumijevanje ključa za determinaciju vrsta zahtijeva visoku razinu stručnosti i iskustva te je stoga najčešća pojava u svijetu da je jedna osoba specijalizirana za određenu skupinu organizama, a mnogo je faktora koji i toj stručnoj osobi mogu predstavljati prepreke u toj taksonomskoj analizi. Morfološke karakteristike su često usmjerene na određeni stadij životnog ciklusa ili roda pa pojedine jedinice ne mogu biti klasificirane. Upravo zbog ovih nedostataka, potrebno je kreirati sustav koji bi, kad se jednom u potpunosti razvije, omogućio pouzdano, pristupačno i ekonomično rješenje za trenutne nedoumice u procesima determinacije vrsta. Takav sustav bi se trebao bazirati na DNA sekvencama kao svojstvenim barkodovima koji se nalaze na univerzalnom mjestu u genomu. Dva glavna cilja projekta su da se nepoznatim uzorcima dodijeli ime vrste te da se pospješiti otkrivanje novih vrsta i olakša identifikacija, posebno za kriptične, mikroskopske ili druge organizme komplicirane morfologije. Ustanovljeno je da se identitet vrsta može utvrditi iz malog broja polimorfnih pozicija unutar COI (citokrom *c* oksidaza 1) gena te je stoga on izabran kao univerzalni barkod. Potrebno je provesti standardizaciju barkodiranjem, koja je ključna za kreiranje razumljive i dosljedne zbirke DNA sekvenci, kojom se može služiti bilo tko, bilo gdje i bilo kad u svrhu brzog i točnog određivanja vrste iz određenog uzorka. Kada se ovaj sustav u potpunosti razvije, omogućit će analize oštećenih i nepotpunih uzoraka kao i uzoraka iz neprepoznatljivih izvora te iz zaštićenih vrsta. Također moguće biti određivanje vrsta bez obzira na stadij životnog ciklusa jedinice, a sve u svrhu razlikovanja morfološki vrlo sličnih vrsta te će pomoći u dopunjenju i doradi razvojnog stabla živog svijeta.

Known Biodiversity
(excluding microbes)
Approximately 1.7 million named
species of plants and animals.



1 square = 10,000 species

Estimated Biodiversity
(excluding microbes)
10 million species



SLIKA 1. Usporedba utvrđene bioraznolikosti u odnosu na pretpostavljenu

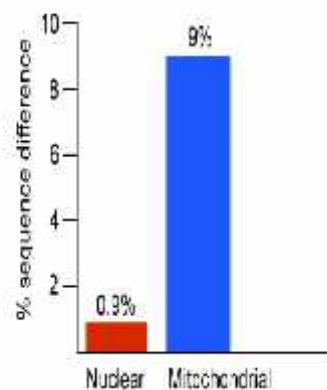
2. RAZVOJ, PRIMJENE I NAPREDAK DNA BARKODIRANJA

DNA Barkodiranje ili genska identifikacija vrsta je ideja standardiziranog molekularnog sustava identifikacije vrsta. Ova metoda ubrzano se počela razvijati tijekom 90-ih godina prošloga stoljeća, zahvaljujući pristupu identifikaciji vrsta pomoću PCR (polymerase chain reaction) metode. Lančana reakcija polimerazom je metoda kojom se određeni fragment DNA umnožava u veliki broj identičnih kopija. U metodi se provodi *in vitro* postupak sinteze DNA kroz uzastopno ponavljanje ciklusa umnažanja fragmenta uz pomoć enzima DNA polimeraze.

„The Barcode of Life“ projekt započeli su 2003. godine znanstvenici sa University of Guelph u Ontariu (Paul D. N. Hebert, Alina Cywinska, Shelley L. Ball, Jeremy R. deWaard) i prezentiran je 2004. godine međunarodnom inicijativom „Consortium for the Barcode of Life“ (CBOL). Cilj je kreirati jedinstveni inventar eukariota baziran na molekularnom pristupu. Već iste godine bio je podržan od strane 150 organizacija iz 45 zemalja uključujući prirodoslovne muzeje, zoološke i botaničke vrtove, herbare, sveučilišne odjele te privatne tvrtke i vladine organizacije. Projekt nema težnju izgraditi razvojno stablo života niti provoditi molekularnu taksonomiju, već proizvesti jednostavan dijagnostički sustav baziran na taksonomskim saznanjima, koja se poklapaju sa referentnom zbirkom DNA barkodova. Trenutna istraživanja su uglavnom usmjerena na mitohondrijski gen citokrom c oksidaza 1 (COI) kao središte globalnog sustava za identifikaciju životinja. Za carstvo Plantae ova sekvencija ne može poslužiti u procesu sekvenciranja te se druge pogodnije regije trebaju utvrditi. Kod životinja mitohondrijski genom je pogodan jer se u stanici nalazi u velikom broju kopija što osigurava veću uspješnost u radu s malim uzorcima. Prednosti mitohondrijskih barkodova su što ne sadrži introne, a razlike u sekvencama su velike od vrste do vrste što omogućava preciznu identifikaciju. Mitohondrijska DNA sadrži 13 gena koji kodiraju za proteine, 2 ribosomalne RNA i 22 gena za tRNA. Protein kodirajućih gena međusobno su više različiti nego ribosomalni te ih je stoga lakše uspoređivati kod blisko srodnih vrsta, a manje su podložni insercijama i delecijama te pokazuju manju stopu rekombinacije i haploidnog nasljeđivanja. Sam DNA barkod sadrži regiju od 648 pb od 58-705 nukleotida gledajući od 5' kraja gena za citokrom c oksidazu 1 (COI), a kao referenca je uzet mišji mitohondrijski genom. Do sada se COI gen pokazao pogodan u projektima određivanja vrsta i filogenije za puževe (Remigio et al., 2003.), skokune (Hogg et al., 2004.), leptire (Hebert et al., 2004a; Hajibabaei et al., 2006a), ptice (Hebert et al., 2004b; Kerr et al., 2007), tulare (Ball et al.,

2005), pauke (Greenstone et al., 2005), ribe (Ward et al., 2005), mrave (Smith et al., 2005), rakove (Costa et al., 2007) te nedavno za *diatomea* i protiste (Evans et al., 2007). Barkodiranje može imati svoju primjenu u prehrambenoj industriji i forenzici te u sprjeavanju ilegalne prodaje i krivolova ugroženih vrsta. Također služi i za identifikaciju organizama iz želudca anih ekstrakata u svrhu razjašnjenja prehrane određenih životinja u situacijama kad nije moguće proučavati ponašanje te nadalje služi za uvid u odnose parazita i domaćina i simbiotske odnose. S obzirom na metode koje se primjenjuju, projekt je baziran na konvencionalnim pristupnim protokolima za DNA ekstrakciju, amplifikaciju i sekvenciranje. Do 2008. godine ukupni dostupni zapisi barkodova obuhvaćali su 363 584 sekvence iz 50 039 vrsta od kojih je 136 338 sekvenci zadovoljilo kriterije DNA barkodiranja (duljina sekvence od 560 pb i više od tri jedinice po vrsti). U 5 godina prikupljeno je preko 65% svih barkodiranih uzoraka, a većina (preko 98%) su pripadnici carstva Animalia. U ovaj projekt velikih razmjera bit će uključeni znanstvenici iz 25 zemalja sa ciljem osnivanja razumljivog registra DNA barkodova za eukariote.

Average sequence differences in nuclear and mitochondrial DNA between human and chimpanzees



SLIKA 2. Prosječna razlika u jezgri i mitohondrijskoj DNA između čovjeka i šimpanze.

3. SUSTAV PODATAKA ZA BARKODIRANJE ŽIVOTA - BOLD (BARCODE OF LIFE DATA SYSTEM)

BOLD (Barcode of Life Data System) sustav podataka za barkodiranje života je službeno utemeljen 2007. godine. Ovaj projekt predstavlja stvaranje zbirke barkodova za carstvo Animalia. Ovaj sustav podataka će biti izuzetno velik i zbirka će sadržavati oko sto milijuna zapisa. Opsežnost tih podataka koji će biti prikupljeni zahtijeva kompjutorski program koji će podržati nove aspekte DNA barkodiranja, što je potaklo razvoj BOLD-a, baze podataka o barkodovima života. Ovaj sustav pruža integriranu bioinformacijsku platformu koja podupire sve faze analitičkog puta, od prikupljanja uzoraka do vrsto potvrđene zbirke barkodova. U odnosu na druge poznate zbirke sekvenci, BOLD je interaktivna alternativa, koja omogućuje da se pohranjene sekvence mogu ponovno provjeravati i taksonomski preraspodijeliti. BOLD je odabran od strane Canadian Barcode of Life Network (kanadska mreža barkodova) za projekt barkodiranja svih eukariota na tom području. Kopije svih sekvenci te ključni podaci o uzorku šalju se također i u NCBI (Nacionalni centar za biotehnoške informacije) ili u genomski repozitorij DDBJ (DNA baza podataka Japana), te i u EMBL (Europski laboratorij za molekularnu biologiju) prije nego što se rezultati javno objave. Metode u BOLD-u se zasnivaju na primjeni HMM (Hidden Markov Model) profila za COI protein te zatim slijedi pretraga zbirke, a NJ stablo se rekonstruira na osnovu podešene i ispitivane sekvence tako da se ispita veza između analizirane i susjedne referentne sekvence. NJ je statistička metoda koja služi za rekonstrukciju filogenetskih stabala i procjenu duljine ogranaka na tom stablu. Iako je ovaj sustav dostupan svakome, registracija omogućava kreiranje privatnih projekata i pristup lozinkom zaštićenim podacima. Trenutno su na BOLD-u dostupne tri funkcionalne jedinice: sustav za upravljanje i analizu, sustav za identifikaciju te eksterni sustav povezivanja. Prije nego što se vrste mogu pohraniti u BOLD, moraju biti taksonomski opisane. BOLD je spreman da prihvati opsežne podatke o vrstama životinja i da podrži njihovu upotrebu kao osnovu za automatsku identifikaciju. Budući napreci u sekvenciranju DNA te kompjutorskoj tehnologiji obave razvoju prijenosnih uređaja koji pohranjuju barkod sekvence za par minuta te koriste referentne zbirke za brzu identifikaciju.

Bats of Southeast Asia - first paper [BM]

Barcode Identifiers

Barcode ID:	BM258-04	Sample ID:	ROM 101996
Gene:	C0X1	GenBank Accession:	Record is unpublished
Last Updated:	2005-08-26	Translation Matrix:	Vertebrate Mitochondrial

Sequencing Runs

Run Date	Run Site	Direction	Trace File	PCR primers	Seq Primer	Status
2004-11-18 15:18:37	University of Guelph	Reverse	BM258-04R2_H01_ab1	VR2/VF2	VR2	high qual
2004-11-18 11:55:58	University of Guelph					

Nucleotide Sequence

Length:	619	-----XGT
Comp. A:	190	GGACAACCAGGAGCC
Comp. G:	95	TTCTTCATAGTAATA
Comp. C:	189	ATAGCATTCCCCGA
Comp. T:	185	ACAGTAGAAGCTGGAI
Updated:	2005-08-26	TCTGTAGATCTAGCAI
		ACCATCATCAATATA
		GCAGTCCTACTATTAT
		ACAACCTTCTTCGAC

Amino Acid Sequence

Length:	206	-----XGT
		HAFFPRMNMMSFWLLP
		TIINNKPPALSGYQT

Illustrative Barcode



Screenshot of the BOLD SYSTEMS interface showing sequencing run details and chromatograms. The interface includes a 'Sequencing Run' section with details like Run Site, Date, File, Status, and Direction. It also features 'Quality Scores' with a histogram and two chromatograms (Reverse and Forward) showing peaks for A, C, G, and T.

SLIKA 3. Stranica sa sekvencama za iste jedinke *Macroglossus minimus* (Chiroptera). 1. detalji o uzorku; 2. detalji o sekvenci sa uklju enim linkovima za pretragu podataka; 3. aminokiselinska sekvenca; 4. prikaz barkoda u boji; 5. prikaz traga i detaljne informacije o kvaliteti traga.

4. PO ECI UPOTREBE DNA BARKODOVA

Uzevši u obzir nedostatke rutinske determinacije vrsta, sustav mikrogenomske analize, koji se zasniva na razlikama u sekvencama DNA koje služe kao barkodovi prisutni u svakoj stanici, obećava pouzdaniji pristup u utvrđivanju bioraznolikosti. COI gen ima neke značajne prednosti u odnosu na druge gene. Jedinственe po etnicе ovog gena su vrlo kompaktne i omogućavaju uspostavljanje 5' kraja kod većine ako ne i svih životinjskih skupina (Folmer *et al.*, 1994, *et al.*, 1997). Također posjeduju veću raspon filogenetskog signala nego drugi mitohondrijski geni. Filogenetski signali omogućavaju da li srodni organizmi imaju tendenciju da nalikuju jedni drugima s obzirom na njihov genetski materijal ili fenotipske osobine. Iako i drugi mitohondrijski geni mogu poslužiti u rješavanju slučajeva nedavnih divergencija, COI gen će vjerojatno pružiti dublje filogenetske uvide (Simmons *et al.*, 2001). COI profili, dobiveni iz malih uzoraka viših taksonomskih kategorija, svrstavaju netom analizirane svojete u pripadajuće rodove. Kreiranjem jasnih COI profila moguće je klasificirati jedinke do razine vrste. Model COI profila, baziran na analizama svake jedinke od 200 blisko srodnih vrsta leptira, bio je 100% uspješan u toj noj identifikaciji prikupljenih uzoraka.

Prvo je bio kreiran COI profil za 7 najraznolikijih životinjskih koljena, baziran na 100 reprezentativnih vrsta i pokazalo se da te osnovne informacije osiguravaju svrstavanje 96% novo analiziranih jedinki u odgovarajuća koljena. Potom je analizirano potkoljeno kukaca zbog najveće bioraznolikosti na planetu. Kreiran je i profil za 8 najrazličitijih redova kukaca baziran na jednoj jedinki, predstavniku za 100 različitih porodica i svaka od 50 novo analiziranih jedinki svrstana je u odgovarajući red. Kod leptira je ispitana mogućnost toj određivanja vrste jer su razlike u sekvencama među u porodicama ovog reda vrlo male. Identifikacija 200 blisko srodnih vrsta i primjena u svrstavanju 150 novo analiziranih jedinki u odgovarajuće vrste bila je 100% uspješna. Približno ¼ COI sekvenci preuzeta je iz GenBank-a, online baze podataka u kojoj su pohranjene sve javno dostupne nukleotidne sekvence te produkti njihove translacije. Ostatak je prikupljen pripremom 30 µl ekstrakta ukupne DNA iz malog uzorka tkiva koristeći i Isoquick (Orca Research Inc, Bothell, WA, 1997) protokol. Parovi po etnicama za lananu reakciju polimerazom (primeri) LCO1490 i HCO2198 korišteni su za umnažanje fragmenta COI gena dugog 658 pb. Nakon primjene lanane reakcije polimerazom, svaki produkt je bio prošeo na gelu upotrebom Qiaex II (Qiagen) kompleta reagensa (kita) te je zatim jednosmjerno određeno primarni slijed ili

sekvencu nukleotida (sekvenciranje) na ABI 377 automatskom sekvencatoru upotrebom Big Dye kiti. Kreirana su tri COI profila u svrhu pružanja predodžbe o raznolikosti COI unutar svake skupine, a zatim su profili korišteni kao osnova za identifikaciju koljena, reda ili vrste na in da se ispituje podudarnost me u sekvencama nepoznatih jedinki i vrsta koje pripadaju odre enom profilu. Jedan profil je za 7 glavnih koljena Animalia (100 COI sekvenci dobivenih iz razli itih porodica, a predstavnici su bili iz svih dostupnih razreda), drugi profil je za 8 najve ih redova kukaca (odabrana su 4 najraznolikija reda sa više od 100 000 opisanih vrsta te dodatna 4 reda proizvoljno odabrana izme u 15 redova kukaca sa 1000 do 15 000 opisanih vrsta) i tre i profil za 200 vrsta blisko srodnih leptira (analizirane su 3 srodne superporodice: Geometroidea, Noctuoidea i Sphingoidea). Proveden je test svojti da se utvrdi mogu nost da se novo analizirane vrste svrstaju u odgovaraju u taksonomsku kategoriju. Odabrane su COI sekvence iz 55 testnih svojti koje uklju uju 5 predstavnika svakog od 5 koljena i 15 predstavnika mekušaca i lankonožaca. Od analiti kih metoda korištena je NJ (neighbour joining) metoda. U ovom slu aju služila je za ispitivanje poveznica izme u svojti u profilu te za daljnju klasifikaciju testnih svojti (Kumar et al., 2000). NJ profili za red i koljeno imaju 100 terminalnih nodija, od kojih svaki predstavlja vrste iz razli itih porodica, dok NJ profili vrsta imaju 200 nodija, od kojih svaki predstavlja zasebnu vrstu leptira. Nadalje je svaki od ova 3 NJ profila služio kao klasifikacijski pogon tako da se analiza ponavljala sa opetovanim uklju ivanjem zasebne testne svojte. Korištena je i metoda multidimenzionalnog skaliranja koja pruža grafi ki zbroj rezultata na razini vrste. Naposljetku je pokazana mogu nost kreiranja baze podataka za carstvo Animalia na temelju COI barkoda te mogu nost da razlike u aminokiselinskim sekvencama COI omogu uju pouzdano svrstavanje organizama u više taksonomske kategorije. PCR metodom je potvr eno da pseudogeni jezgre nisu utjecali na istraživanje. Vrlo je važna uspješnost COI barkoda u prepoznavanju poveznica izme u svojti jer ukazuje da karakterna konvergencija i horizontalni prijenos gena (retrovirusima) ne ometaju o ekivanu pripadnost odre enoj svojti, ve štoviše omogu avaju svrstavanje organizama na najdublje taksonomske razine. Nužno je odrediti razgrani enja koja razlikuju vrste drugih geografskih regija i taksonomskih skupina, a ista treba utvrditi i za skupine sa razli itim karakteristikama, kao što su generacijsko vrijeme ili tijek rasprostranjivanja koji su mogli utjecati na stopu evolucije ili na intenzitet odvajanja populacija. Nadalje, odre ena svojstva poput hibridizacije ili introgresije zahtijevaju dodatne analize jednog ili više gena jezgre, a kod vrsta nastalih poliploidijom treba utvrditi veli inu genoma. Hibridizacija podrazumijeva formiranje hibridnih DNA molekula iz lanaca DNA

različitih vrsta, dok introgresija opisuje infiltraciju gena jedne vrste u skup gena (zbroj svih gena u populaciji) neke druge vrste povratnim križanjem. Bez obzira na navedene nedostatke, ovaj projekt ima velike prednosti u lakšem utvrđivanju granica raznolikosti unutar vrste, omogućuje determinaciju juvenilnih jedinki a taksonomski kriteriji postat će objektivniji. Daljnji koraci i analize mogle bi biti usmjerene na vrste od akademske, medicinske ili ekonomske važnosti.

5. NEKI PRIMJERI BARKODIRANJA

5.1. Barkodiranje Diptera

5.1.1. Chironomidae

Provedeno je ispitivanje mogućnosti upotrebe COI gena kao barkoda u identifikaciji Diptera: Chironomidae. Chironomidae su najraznolikiji i najuže estaliji makrozoobentoski organizmi slatkih voda. Ličinke i ženske jedinke teško se morfološki identificiraju, a mnoge su još uvijek nepoznate za znanost. Provedeno je uzorkovanje i determinacija svojiti te potom molekularne metode ekstrakcije, amplifikacije i sekvenciranja. Analizirana je DNA iz 97 uzoraka 47 vrsta rodova *Cladotanytarsus*, *Micropsectra*, *Parapsectra*, *Paratanytarsus*, *Rheotanytarsus*, *Tanytarsus* i *Virgatanytarsus* sa naglaskom na *Micropsectra*, *Parapsectra* i *Paratanytarsus*. Dokazano je da se COI lako amplificira iz ekstrakata dobivenih iz različitih stadija životnog ciklusa upotrebom standardnih primera, uz izuzetak *Micropsectra calcifontis*. Rezultati su analizirani NJ metodom i maximum parsimony metodom (metoda najveće štedljivosti), statistička metoda koja se često koristi u kompjutorskoj filogenetici za utvrđivanje filogenije, a bazira se na principu Occamove oštrice. Rezultati su pokazali da su sve vrste taksonomski razlučive s obzirom na genetičke udaljenosti i karakterne razlike. Veliki naglasak stavljen je na analize provedene kod *Paratanytarsus grimmii*, jer je ova vrsta idealan primjer uspješne primjene DNA barkodiranja. COI sekvenca jedinke iz Norveške savršeno se podudara sa sekvencom jedinke iz Australije te će biti zanimljivo utvrditi da li je nedostatak genetičke varijabilnosti posljedica nedavnog rasprostranjenja ove vrste. Nužno je

da se barkodovi uvelike razlikuju u svrhu ispravnog uvrštavanja u skupinu ili iskljuivanja iz iste, a poželjno je da se za identifikaciju poznatih vrsta može koristiti univerzalni barkod. Ukoliko se određena vrsta već nalazi u zbirci DNA sekvenci, velika je vjerojatnost ispravne determinacije. U ovom istraživanju pronađene su i neke potencijalno kriptične vrste, a za pouzdanije razlučivanje dobivenih rezultata nužno je uključiti i markere jezgrične DNA.

5.1.2. Tachinidae

Tachinidae su porodica mušica koje parazitiraju u tisućama vrsta gusjenica Costa Rica, a mogu biti opće parazitoidne ili specijalizirane za određenu vrstu ili rod gusjenica. Istraživanje dugo 29 godina provodilo se na 400 vrsta uzgojenih u više od 390 000 u divljini ulovljenih jedinki koje su pripadnice 3500 vrsta na području Costa Rica i smatra se da je barem 90% jedinki specijalizirano za jednu ili nekoliko blisko srodnih vrsta. U svrhu potpunog definiranja parazitoidnih svojstava, provedeno je DNA barkodiranje na 2134 jedinke koje su prvotno svrstane u 16 morfološki usporedivih vrsta sa najizraženijim općim parazitoidnim karakteristikama i dobivene su 73 mitohondrijske linije sa prosječnom 4%-tnom razlikom u sekvencama. Te linije su testirane neovisnim markerima jezgre (28S i ITS1) te su trenutno smatrane odvojenim vrstama. Da bi se opće parazitoidno svojstvo u potpunosti definiralo, provele su se analize COI sekvenci kao barkoda kod 16 najizraženijih općih parazitoidnih vrsta, uzgojenih na mnogo vrsta gusjenica, a sve su uzgajane 10 do 100 puta. Tih 16 vrsta parazitiraju u provjerenim, definiranim domaćinima. Dobivene su 64 vrste specijaliziranih parazitoidea i 9 općih parazitoidnih vrsta. Ovaj rezultat pridaje još veći značaj istraživanju barkoda provedenog kod 20 morfološki različitih vrsta roda *Belvosia*, drugog pripadnika Tachinida koji obitava na istom staništu. Potvrđene su sumnje o podcijenjenoj bioraznolikosti parazitoidea te o vjerojatnosti da je na kraju života tropskih općih parazitoidnih mušica još neobičniji nego što se mislilo. Uspješno su provedene analize barkoda. Uzeta je po jedna mušica iz svakog od 2134 uzgojena legla i 14 od 16 vrsta bilo je moguće pouzdano razlikovati putem barkodova. Dvije vrste bile su međusobno komplicirane za usporediti, no ipak su bile drugačije od ostalih. Svaka od 14 vrsta pripala je u zaseban, nepreklopljivi klaster sekvenci u NJ stablu. Klasteri barkodova međusobno razlikovali su se >5%, ali je intraspecijska divergencija bila visoka, kod nekih čak preko 10%, što ukazuje na postojanje više kriptičnih vrsta. Uključene je niz ekoloških faktora kao što su gusjenice domaćini, njihova ishrana i

ekosustav te neovisni geneti ki markeri, da bi se ispitalo da li ovi podaci u cijelosti podupiru hipotezu da svaki klaster sekvenci predstavlja morfološki prikrivene vrste. Hipoteza se pokazala ispravnom i 16 vrsta je raspoređeno u 73 kriptične vrste. Samo su dvije vrste, *Hyphantrophaga virilis* i *Lespesia aletiae*, pouzdano odgovarale ve dodijeljenim imenima. Klasterima vrsta su pridodana alfanumerička privremena imena, gdje ime odražava rod u koji su svrstane u inventaru, makar ta imena nisu vrsta mjerila determinacije. činilo se i da klaster unutar jedne od 16 vrsta nije odgovarao ekološkim podacima ili je pak bilo samo malih razlika u odnosu na sekvence susjednih klastera. Zbog ovoga su analizirana i dva neovisna markera jezgre. Kao i kod *Belvosia*, analizirane su ITS1 i D2 regije 28S. Ove dodatne analize nisu provedene kod svih jedinki, za koje postoje barkodovi, jer je svrha bila identifikacija vrsta i otkrivanje prikrivenih vrsta među opće parazitoidnim vrstama. Spajanjem ovih rezultata: COI barkodovi, ekološki imbenici i jezgrine sekvence, 73 vrste je bilo moguće rasporediti u 4 uzorka:

- 1. Barkodirani opće parazitoidni organizmi ostaju opće i parazitoidi**
- 2. Barkodirana opće parazitoidna vrsta postaje dvije opće i parazitoidne vrste**
- 3. Barkodirani opće i parazitoid postaje nekoliko specijaliziranih i jedan opće i parazitoid**
- 4. Barkodirani opće i parazitoid je kompleks specijaliziranih parazitoida**

Ono što je smatrano da obuhvaća 16 vrsta ispalo je barkodiranjem 73 vrste i sve osim dvije (*B. albicauda*) mogu se identificirati putem barkoda. Barkodiranje nije efikasno samo za identifikaciju, već i za otkrivanje postojanja kriptičnih vrsta, ali je otežano kod mladih vrsta ili vrsta koje hibridiziraju pa treba uključiti druge molekularne markere što pogotovo vrijedi za skupine bogate vrstama i morfološki vrlo slične skupine kao što su parazitoidi. Takvi dodatni markeri, ovdje korišteni, bili su rRNA, zanimljivi zbog velikog obilja i relativno očuvanih boćnih regija, što olakšava kreiranje primera široke primjene.

5.2. Barkodiranje Ephemeroptera

5.2.1. *Baetis vernus*

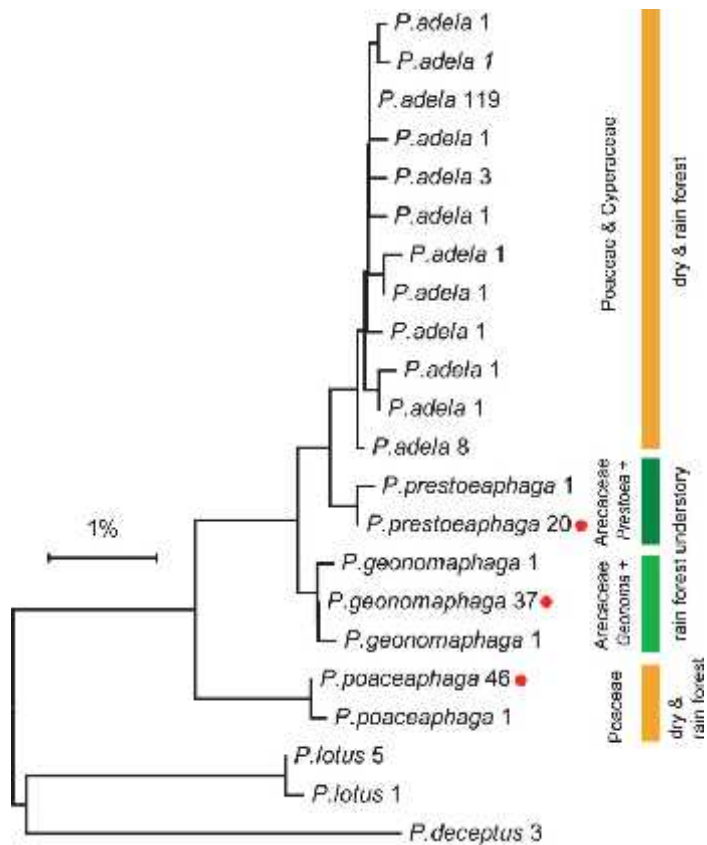
U Finskoj je poznato 19 vrsta vodencvjetova pripadnika porodice Baetidae od kojih 11 pripadaju rodu *Baetis*. Liinke *Baetis vernus* pokazuju morfološke varijacije uvjetovane okolišem (Bauernfeind et al., 2001), što ih čini vrlo kompliciranom skupinom za determinaciju bez obzira na postojanje kvalitetnog ključa Engblom, 1996 (Savolainen et al., 2007). Savolainen et al., 1981 ističu postojanje dvije varijante liinki *Baetis macani*: jedan oblik nastanjuje vode tekuće i ima manje izražene škrge sa skrivenim trahejama, dok drugi obitava u stajaćicama i ima šire škrge sa izraženim trahejama. Skupljene su jedinice svih poznatih vrsta u Finskoj, uz naglasak na svojite sa očitim taksonomskim i morfološkim nedoumicama. Analize su pokazale da su lenticke (nastanjuju stajaćice) i loticke (nastanjuju tekuće) forme *Baetis macani* međusobno reproduktivno izolirane te su svrstane u odvojene klustere. Između u ovih svojiti ustanovljene su očitne razlike, a nove svojite će naknadno biti opisane i definirane, a također će biti nužno provesti i dodatne taksonomske analize koje će moguće izmjeniti dosadašnju klasifikaciju. COI barkodovi su pokazali veliku važnost u ispravnoj determinaciji vrsta i rezultati su vrlo obećavajući.

5.3. Barkodiranje Lepidoptera

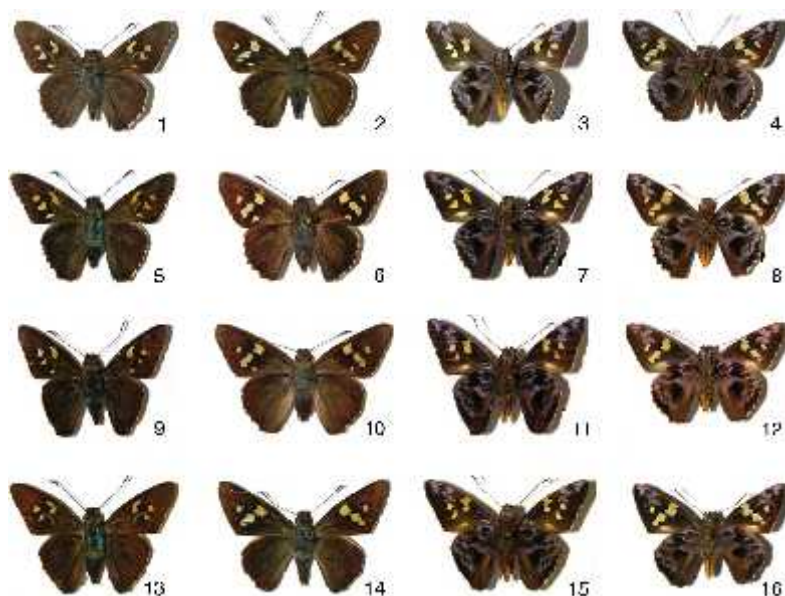
5.3.1. HesperIIDae

Leptiri HesperIIDae u usporedbi s drugim insektima taksonomski su dobro poznati. Mnogi su morfološki opisani i imenovani prije 200 godina. Kao posljedica geografskog rasprostranjivanja nastala je u estala, široko rasprostranjena, varijabilna, ali prepoznatljiva svojita s kojom su znanstvenici već toliko upoznati da rijetko potiče daljnji taksonomski interes. Idealan primjer za ovu pojavu je *Astraptes fulgerator*, veliki leptir iz porodice HesperIIDae opisan 1775. godine koji je u pogledu distribucije do sredine dvadesetog stoljeća smatran paneotropskim. Uzgojem ove vrste u velikim količinama iz gusjenica ulovljenih u divljini, utvrđeno je da je veoma polifagan (hrani se velikim brojem vrsta) što je uzrokovalo sumnju da je riječ o samo jednoj vrsti. Prvotne morfološke analize bile su neuvjerljive pa su stoga slijedile detaljne usporedbe velikog broja odraslih jedinki. Otkriveno je 6 do 7 vrsta

malih morfoloških varijacija sa razlikama u ishrani u stadiju liinke te spolu. Provedena je analiza COI gena na 466 uzgojenih odraslih jedinki. Otkriveni su klasteri koji kovariraju sa ovim već utvrđenim te 3 nova klastera, ukupno daju i 10 prikrivenih vrsta koje variraju od parapatrijskih do simpatrijskih. Druga analiza na uzgojenim leptirima iz ACG (Area de Conservacion Guanacaste) u Costa Rici ukazala je da unutar roda s jednom vrstom postoje 4 nove i vrlo slične vrste. Te 4 vrste opisuju umjerene razlike u odraslim stadijima, u reproduktivnim organima mužjaka i ženki te u uzorcima boja. Barkodi nisu uspjeli razdvojiti blisko srodne vrste unutar rodova *Polyctor*, *Cobalus* i *Neoxeniades*, iako razlike u ekološkim parametrima i morfologiji dokazuju da bi zasigurno to trebale biti različite vrste. Nemogućnost razdvajanja ovih vrsta leži u upotrebi prekratkog fragmenta u koji nisu uključene dijagnostičke regije, dok su cjeloviti barkodovi (650 pb) pouzdano razdvojili vrste u svakom paru leptira. U početku, program sistematiziranja barkodiranja svake ACG vrste macrolepidoptera uključivao je nekoliko primjeraka *Perichares philetus* koji je taksonomski bio opisan kao politipski i paneotropski. Neotekivano su prva dva uzorka pokazala jaku divergenciju unutar njihovog konospecifičnog klastera u NJ stablu, a to je potaklo intenzivnu potragu za gusjenicama *P. philetus*, više analiza barkodova odraslih jedinki te pominje ispitivanje svih stadija životnog ciklusa. Daljnje NJ statističke analize sve već ih i već ih uzoraka otkrile su 4 klastera. Taksonomsko pitanje je koji dodatni dokazi opravdavaju službeno razdvajanje prikrivenih vrsta *Perichares* na koje postojanje ukazuju DNA barkodovi. Jedna potvrda je prehrana liinke. U svakom slučaju morfološke karakteristike su varljive i teško ih je točno definirati. Trenutna klasifikacija ovih leptira još uvijek odražava temeljite analize iz sredine 20. stoljeća a u kojima su srodne vrste često klasificirane kao podvrste bez opravdana razloga. Tih mnogo tropskih vrsta za koje se smatra da se hrane različitim biljkama, zapravo obuhvaćaju setove prehrambeno specijaliziranih vrsta. Barkodiranjem 422 morfološki slične vrste utvrđeno je da su dvije ili više bioloških vrsta. Analize su još nepotpune jer je barkodiranje provedeno samo kod vrsta unutar ACG-a. Trud se ulaže u prikupljanje više ACG jedinki kad god se pojavi odstupanje u NJ stablu.



SLIKA 4. NJ stablo bazirano na Kimura dva-parametarskim udaljenostima za COI DNAbarkodove 6 vrsta *Perichares* uzgojenih u ACG; gornje 4 vrste(3 nove) pripadaju *P. philetus* kompleksu vrsta. Brojevi pokazuju koliko je jedinki sa svakim haplotipom. Trake u boji govore o u estalosti u ekosustavima te tip ishrane li inki. Crvene oznake ukazuju na haplotipove holotipova kod 3 nove vrste.



SLIKA 5. Muške jedinke (stupac 1 i 3) i ženske (stupac 2 i 4) kod 4 skrivene vrste *Perichares* sa dorzalne (lijevo) i ventralne (desno) strane. (1–4) *P. adela* (06-SRNP-45166, 06-SRNP-40498). (5–8) *P. poaceaphaga* [06-SRNP-1189 (holotype), 03-SRNP-23215]. (9–12) *P. geonomaphaga* [04-SRNP-2706 (holotype), 04-SRNP-30964]. (13–16) *P. prestoeaphaga* [01-SRNP-22227 (holotype), 06-SRNP-31113].

6. BARKODIRANJE BILJAKA I GLJIVA

DNA barkodovi mogu poslužiti i u identifikaciji biljaka i gljiva, ali uz druga iji pristup. Univerzalni COI barkod razvijen za životinje i alge nije pogodan za analizu biljaka i gljiva jer se kod ovih carstava evolucija mitohondrijskog genoma odvijala suviše sporo pa to no odre ivanje vrste nije mogu e. Stoga su odabrane druge pogodnije regije. U slu aju carstva Fungi istražuje se ITS regija ribosomalne DNA iz jezgre kao potencijalnog barkoda, a jedna od prepreka je i da su geni kod gljiva podložni duplikacijama. Za carstvo Plantae predloženo je nekoliko regija koje bi mogle poslužiti kao zajedni ki barkodovi za identifikaciju. Za istraživanje su odabrana dva atraktivna lokaliteta zna ajne bioraznolikosti biljaka: Costa Rica, kao najbogatije stanište orhideja na svijetu, te nacionalni park Kruger (KNP) u Južnoj Africi,

kao područje je velike bioraznolikosti drve a. Provedeno je sakupljanje uzoraka, analizirane su potencijalne barkod regije te su primijenjene statističke metode za usporedbu i filogenetske analize. Testirano je osam potencijalnih DNA barkodova. PCR metoda se pokazala uspješnom sa svim barkodovima, osim kod dva barkoda orhideja. Barkod koji bi bio pogodan treba biti dovoljno različit od vrste do vrste, a opet dovoljno sličan unutar pripadnika iste vrste. Izvršena je procjena svakog potencijalnog barkoda da se utvrdi je li moguće potvrditi status vrste kao monofiletske, upotrebom filogenetskih metoda te statističkih analiza poput maximum parsimony, maximum likelihood, NJ, Bayesian i bootstrap resampling. Primijenjene su analize sjedinjenja za usporedbu uzoraka razgranjenja razvojnih stabala i za identifikaciju jasnih genetskih klastera. Najveći broj nezavisnih klastera dobiven je upotrebom UPGMA (Unweighted Pair Group Method with Arithmetic Mean) sa genima *matK*, *rpoB* te *trnH-psbA*. U kombinaciji sa *matK* identificiran je 41 klaster, a od toga je 30 u potpunosti odgovaralo već determiniranim vrstama, 4 su djelomično odgovarala, a 7 klastera je svrstano u istu vrstu. Ovi rezultati ukazuju da su *matK* ili *trnH-psbA* najpogodnije regije za barkodiranje biljaka. Upotreba *matK* samostalno ili u kombinaciji sa *trnH-psbA* u monofiletskim testovima, omogućila je preko 90% ispravnu identifikaciju vrsta, a dodatkom drugih barkodova nije se postigla već uspješnost u identifikaciji vrsta. Trenutno je cilj pomoću *matK* barkoda ubrzati istraživanja i osigurati očuvanje biljne bioraznolikosti Zemlje, a usporedno se razvijaju i druge metode u iste svrhe. Još jedna primjena je i u analizi biljaka koja je prodaja regulirana sa „Convention on International Trade of Endangered Species“ (konvencijom o međunarodnoj prodaji ugroženih vrsta) kako bi se spriječila ilegalna trgovina zaštićenim vrstama. Ukoliko se tehnike barkodiranja biljaka razvijaju i postanu uspješne, nužno je da vrlo brzo postanu dostupne u siromašnim zemljama bogate bioraznolikosti.

7. BARKODIRANJE RODA HOMO

Dokazana je univerzalna pripadnost vrsti *Homo sapiens*. Usporedba COI barkoda dokazala je među pripadnicima vrste, razlike od svega 1 ili 2 pb od ukupno 648 pb.

Od šimpanzi je rod *Homo* odvojen razlikom u 60 pb, a od gorila u 70 pb.

8. NEDOSTACI BARKODIRANJA

Sve prepreke moraju biti jasno definirane i premošene u fazi izgradnje zbirke ili BOLD baza podataka nikada ne će postati značajna. Zbog majinskog nasljeđivanja mtDNA, upotreba mitohondrijskih lokusa može dovesti do zaključaka o precijenjenoj divergenciji uzoraka i potaknuti nejasne zaključke o statusu vrsta. Interspecijska hibridizacija i endosimbiontske infekcije mogu uzrokovati prijenos mitohondrijskih gena iz evolucijske skupine jedinice. NUMT su kopije mitohondrijskih DNA sekvenci koje su prenesene u genom jezgre, a mogu značajno zakomplicirati zbirku mitohondrijskih COI sekvenci. Ove sekvence mogu biti prepoznate pa se i postojanje ovakvih kopija može otkriti u trenutku provjere sekvenci, a njihova prisutnost mora biti istaknuta u BOLD-u. Visoka stopa intraspecijske divergencije može proiza i iz geografski izoliranih populacija te se i to treba uzeti u obzir prilikom kreiranja baze podataka. U slučaju ovih prepreka, potrebno je analizirati i lokuse jezgre kako bi se razriješili filogenetski odnosi. Pažnju treba posvetiti i uvođenju referentnih sekvenci, pogotovo za vrste sa već poznatim narušenim mitohondrijskim nasljeđivanjem, a takve situacije, koje mogu biti pogrešno protumačene, trebaju biti naglašene u BOLD-u. Još jedan nedostatak u barkodiranju je i nemogućnost ekstrakcije DNA iz uzoraka pohranjenih u formalinu, a većina se upravo na taj način čuva u muzejima.

9. SAŽETAK

Barkodiranje povezuje taksonomiju, molekularnu filogenetiku i populacijsku genetiku. Dokazana je mogućnost kreiranja baze podataka za životinje na temelju COI sekvence. Za druga carstva potrebno je provesti opsežnije analize u svrhu pronalaska najpogodnijeg barkoda. Daljnji razvoj i napredak ovog projekta omogućit će jasniji uvid u bioraznolikost, a analize su puno objektivnije nego kod klasičnih morfoloških analiza. Zanimljivo je i pomisliti na postojanje online enciklopedije života na Zemlji. Iako je projekt zapravo relativno nov, svakodnevno se ulaže trud da se premoste eventualne prepreke u procesima izgradnje ovakvog sustava. Metode u molekularnoj biologiji postaju jeftinije i mogu se uvesti u

postupak barkodiranja. Razvijaju se mnoge nove metode poput pirosekvenciranja koje omogućavaju brzu analizu pomiješanih uzoraka. Već postoje i razvijene tehnike nedestruktivne ekstrakcije DNA iz nedavno prikupljenih uzoraka, što pospješuje njihovu konzervaciju. Sve prepreke moraju biti jasno definirane i premošene da bi se BOLD baza podataka mogla u potpunosti uspješno razviti. Ukoliko neka organizacija poput Global Biodiversity Information Facility ili All Species Foundation preuzme ovaj projekt, to bi bio veliki korak prema dugovjernoj bazi koja nedostaje mnogim dosadašnjim bazama podataka. Zaključno, vrijeme je da se današnje društvo suoči sa mnogim biološkim problemima kao što su izbjegavanje pandemija, očuvanje bioraznolikosti, pružanje biološke sigurnosti te zaštita vrsta.

10. SUMMARY

Barcoding associates taxonomy, molecular phylogenetics and population genetics. It demonstrates the possibility of creating a database of the animals on the basis of COI sequences. For other kingdoms of life it is necessary to conduct a more comprehensive analysis in order to find the most appropriate barcode. Further development and progress of this project will provide a clearer insight into the biodiversity, and analysis are much more objective than conventional morphological analysis. It is interesting to think of the existence of the online encyclopedia of life on Earth. Although the project is actually relatively novel,

daily efforts are made to overcome any obstacles in the process of building this system. Methods in molecular biology are becoming cheaper and can be introduced into the process of barcoding. Development of many new methods such as pyrosequencing will allow rapid analysis of mixed samples. Techniques of nondestructive extraction of DNA from the recently collected samples are already developed, which promotes conservation of samples. All obstacles must be clearly defined and overridden so that BOLD database could develop fully and successfully. If an organization such as the Global Biodiversity Information Facility or All Species Foundation acquire this project, it would be a major step toward longevity lacking many current databases. Finally, it is time to confront today's society with many biological problems such as avoiding pandemics, biodiversity conservation, providing security and protection of biological species.

11. LITERATURA

Burns M., John; Janzen H., Daniel; Hajibabaei, Mehrdad; Hallwachs, Winnie; Hebert D. N., Paul, 2007., DNA barcodes and cryptic species of skipper butterflies in the genus *Perichares* in Area de Conservacion Guanacaste, Costa Rica;

Ekrem, Torbjorn; Willassen, Endre; Stur, Elisabeth, 2006., A comprehensive DNA sequence library is essential for identification with DNA barcodes;

Frezal, Lise; Leblois, Raphael, 2008., Four years of DNA barcoding: current advances and prospects;

Hebert D. N., Paul; Cywinska, Alina; Ball L., Shelley; deWard R., Jeremy, 2002.,
Biological identifications through DNA barcodes;

Lahaye, Renaud; Bank, Michelle, van der; Bogarin, Diego; Warner, Jorge; Pupulin, Franco;
Gigot, Guillaume; Maurin, Olivier; Duthoit, Sylvie; Barradough G., Timothy; Savolainen,
Vincent, 2007., DNA barcoding the floras of biodiversity hotspots;

Ratnasingham, Sujeevan; Hebert D. N., Paul, 2006., BOLD: The barcode of life data system;

Smith M. Alex; Wood D., Monty; Janzen H., Daniel; Hallwachs, Winnie; Hebert D. N., Paul;
2006., DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies
(Diptera, Tachinidae) are not all generalist;

Stahls, Gunilla; Savolainen, Eino, 2007., mtDNA COI barcodes reveal cryptic diversity in the
Baetis vernus group (Ephemeroptera, Baetidae)

Stoeckle, Mark; Waggoner E., Paul; Ausubel H., Jesse, 2005., Barcoding life, illustrated