ANALYSIS OF CYTOCHROME OXYDASE I SEQUENCES OF ELACHISTINAE (LEPIDOPTERA: ELACHISTIDAE) SPECIES

BRIGITA PAULAVIČIŪTĖ[1,2], ALGIMANTAS PAULAUSKAS[1], VIRGINIJUS SRUOGA[3]

[1]Vytautas Magnus University, Kaunas, Lithuania.
E-mail: b.paulaviciute@gmf.vdu.lt; a.paulauskas@gmf.vdu.lt
[2]Kaunas T. Ivanauskas Zoological Museum, Kaunas, Lithuania
[3]Vilnius Pedagogical University, Vilnius, Lithuania. E-mail: virginijus.sruoga@vpu.lt

Abstract. The aim of this study was to analyse mtDNA sequences of the mtDNA cytochrome c oxidase I (COI) gene as a tool for investigating the genetic polymorphism of Elachistinae. The polymorphism of the sequenced COI gene was assessed in 31 specimens of 11 species. PCR amplification of genomic DNA with the COI primer in each of the samples yielded a specific fragment corresponding to position 2239–2944 in the sequence. The tree was constructed using the Neighbor-joining (NJ) method with the Kimura 2-parameter model. Analysis of the COI sequenced 31 samples of 11 Elachistinae species produced 640bp sequence alignment. The nucleotide diversity was obtained for Elachistinae moths. The sequences contrasted with 229 polymorphic nucleotides. The maximum parsimony analysis revealed 182 parsimony informative characters.

Key words: *Elachista*, COI gene, sequencing, polymorphism

Introduction

Molecular tools are a standard part of many conservation studies and can be informative at many different levels of analysis, although there are inherent limitations and strengths of different genes or parts of genes to find answers to specific questions. Insect DNA barcodes, 600- to 800-base-pair segments of the mitochondrial gene cytochrome c oxidase I (COI), have been proposed as a means to quantify global biodiversity. Mitochondrial (mt) DNA has a long history of use at the species level. Recent analyses suggested that the use of a single gene, particularly mitochondrial, is sufficient in taxonomic scope to recognize many species lineages. A mitochondrial genome can result in very different assessments of biodiversity (Rubinoff, 2006). The study of mitochondrial DNA sequences has become the method for a wide range of taxonomic, population and evolutionary investigations in Lepidoptera (Lunt *et al*., 1996). We used this method for investigating of *Elachistinae* moths. There are still very few publications on the investigations of mtDNA of this subfamily (Kaila & Ståhls, 2006; Sruoga *et al*., 2009; Paulavičiūtė & Paulauskas, 2010).

Elachistinae is a rather small subfamily in comparison to many other groups of Lepidoptera. Moths are small with the wingspan of 6 to 14 mm. The forewing pattern mainly consists either of a white fascia and spots on a dark background, or dark marks on a light background, or moths are unicolorous (white, yellowish or creamy). Larvae of

Elachistinae are typical leaf-miners, trophically connected mainly with Monocotyledonous grasses.

The aim of this study was to analyze the mtDNA sequences of the COI gene as a tool for investigating the genetic polymorphism of *Elachistinae*.

Material and Methods

Study area

Elachistinae moths were collected in eight places of Lithuania: Neringa and Palanga municipalities, Jurbarkas, Kaišiadorys, Kaunas, Panevėžys, Šakiai, Tauragė, Trakai, and Vilnius administrative districts (Fig. 1).

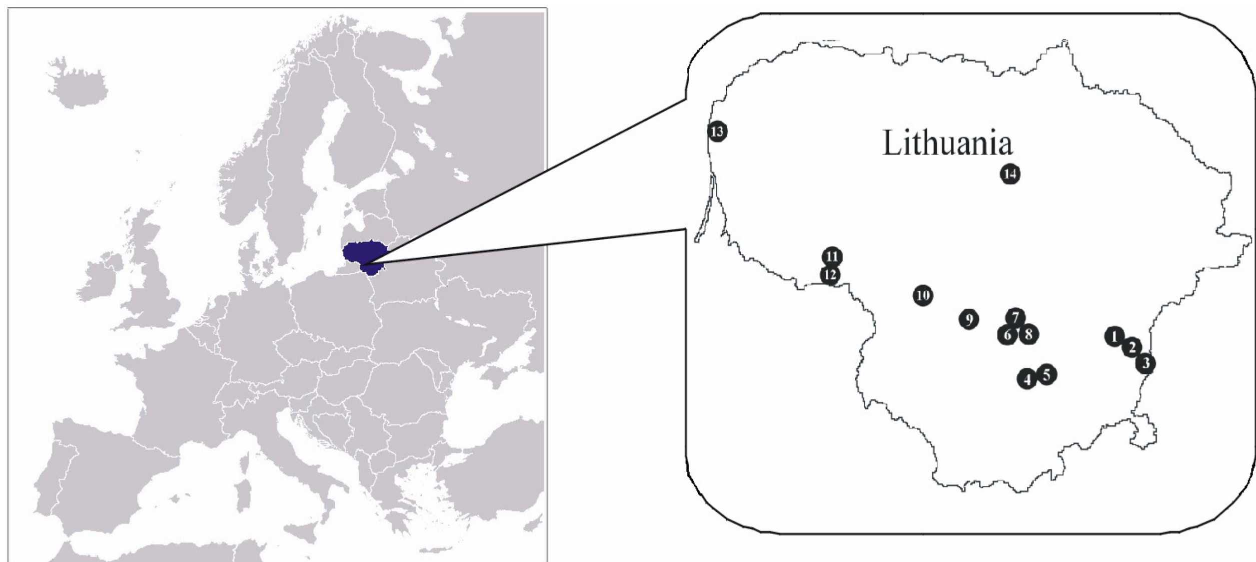

Fig. 1. Sampling sites of *Elachistinae* specimens in Lithuania: 1 – Paneriai, Vilnius env. (54°38'08,0"N, 25°11'30,61"E); 2 – Mickūnai env., Vilnius district (54°42'56,0"N, 25°32'42,1"E); 3 – Kalveliai, Vilnius district (54°38'6,4"N, 25°41'06,5"E); 4 – Aukštadvaris, Trakai district (54°34'48,9"N, 24°31'39,4"E); 5 – Čižiūnai, Trakai district (54°35'51,1"N, 24°33'50,4"E); 6 – Rumšiškės, Kaišiadorys district (54°52'38,1"N, 24°10'44,7"E); 7 – Baniškės, Kaišiadorys district (54°52'31,5"N, 24°14'03,3"E); 8 – Strėvininkų Miškas f., Kaišiadorys district (54°48'42,6"N, 24°21'37,2"E); 9 – Jiesia landscape reserve, Kaunas district (54°49'00,2"N, 23°55'02,4"E); 10 – Juškinė forest, Šakiai district (55°01'26,3"N, 23°26'54,0"E); 11 – Eičiai, Tauragė district (55°09'39,6"N, 22°28'40,7"E); 12 – Viešvilė, Jurbarkas district (55°05'10,5"N, 22°24'20,8"E); 13 – Palanga (55°55'17,5"N, 21°03'47,9"E); 14 – Kupstai, Panevėžys district (55°49'31,0"N, 24°16'04,8"E)

Sampling methods

Material was sampled from early spring to late autumn in 2004–2008 using an entomological net and during light trapping at night (160W DRL type bulb lamp was used).

Species were identified by the external appearance and genitalia of moths, using a Motic SMZ 168 stereomicroscope.

DNA extraction

We used pinned specimens and specimens which were stored in 96% ethanol. DNA was extracted from head or thorax using the Nucleospin Tissue Kit (Machery-Nagel, Düren, Germany) according to manufacturer's protocols.

DNA amplification

Primers used to amplify COI fragments were: LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') (Folmer *et al*., 1994; Herbert *et al*., 2003).

PCR reactions were carried out in 25 µl reaction mixtures containing 2 µl of DNA extract, 2 µl of each primer (at 10 pmol/µl) (MBI Fermentas, Lithuania), 0.5 µl of Amplitaq DNA polymerase (5 U/µl), 2.5 µl 25 mM of $MgCl_2$, 2.5 µl of 10X buffer (Fermentas) and 1 µl 10 mM of dNTP (Fermentas), and water. Termocycler conditions were the initial denaturing at 94°C for 2 min, 35 cycles of 30 s denaturing at 94°C, 45 s annealing at 50°C, 1 min extension at 72°C, followed by a final extension of 4 min at 72°C.

After amplification, PCR products were separated by electrophoresis in 1.5% agarose gel. Agarose gel was stained with ethidium bromide and photographed under UV light (*EASY Win32, Herolab*, Germany). DNA fragment sizes were assessed by comparison with GeneRuler™ 100bp DNA Lader Plus (MBI Fermentas, Lithuania) (Fig. 2).

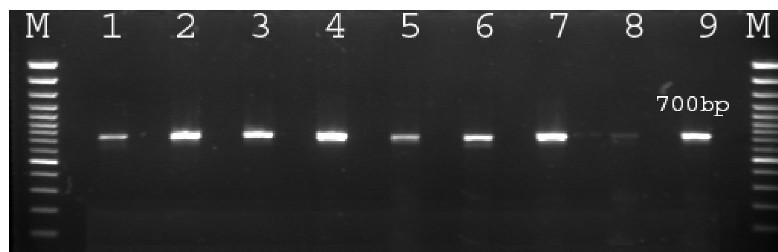

Fig. 2. PCR-amplified products of COI gene of *Elachista* moths. Lane M: 100 bp DNA ladder. Lanes 1–9: amplified products (about 700 bp)

Sequencing

PCR products were purified with the Nucleospin Extract Kit (Machery-Nagel, Düren, Germany). Cycle sequencing was done using the ABI PRISM BigDye terminator version 3.1 cycle sequencing kit. One to eight reactions were used to produce 20 µl of a cycle-sequencing product, using 8 µl of the ABI reaction mix. Products were separated and visualized using an ABI PRISM 310 Genetic analyzer. All fragments were sequenced in both directions.

Data analysis

The editing of DNA sequences, counting assembly, and the alignment of consensus sequences was performed using Bioedit 5.0.9 software. Phylogenetic analysis was performed using MEGA 4 (Tamura *et al*., 2007).

Results and discussion

Cytochrome c oxidase I partial sequences of 31 specimens were amplified and sequenced. Sequences of 640 bp (correspond to position 2239–2944) were aligned. GenBank accession numbers are presented in Table 1.

Table 1. List of specimens used for DNA sequencing and GenBank accession numbers

Examined specimens	GenBank accession number
Perittia herrichiella(1)	HM034446
Perittia herrichiella(2)	HM034447
Elachista albifrontella(1)	GU248251
Elachista albifrontella(2)	GU248252
Elachista alpinella(1)	GU248254
Elachista alpinella(2)	GU248255
Elachista alpinella(3)	GU248256
Elachista alpinella(4)	GU248257
Elachista alpinella(5)	GU248258
Elachista alpinella(6)	GU248259
Elachista argentella(1)	DQ666137
Elachista argentella(2)	DQ666138
Elachista argentella(3)	HM034449
Elachista argentella(4)	HM034448
Elachista canapennella	GU248260
Elachista consortella	GU248261
Elachista humilis(1)	GU248253
Elachista humilis(2)	HM034450
Elachista maculicerusella(1)	DQ666141
Elachista maculicerusella(2)	DQ666142
Elachista maculicerusella(3)	DQ666143
Elachista pollinariella(1)	DQ666140
Elachista pollinariella(2)	DQ666139
Elachista pollinariella(3)	DQ666144
Elachista pollinariella(4)	GU248246
Elachista pollinariella(5)	GU248247
Elachista pullicomella(1)	GU248248
Elachista pullicomella(2)	GU248249
Elachista pullicomella(3)	GU248250
Elachista utonella(1)	HM034451
Elachista utonella(2)	HM034452

Analysis of partial COI gene sequences demonstrated a different distributional rate of nucleotides (Table 2); 38.7% of T, 16.7% of C, 30.4% of A and 14.1% of G nucleotides were obtained in the sequences. The *Elachista canapennella* sequence has 41.5 % of T nucleotides. The *E. alpinella*(4) sequence has the maximum number of C nucleotides – 18.6%. The *E. maculicerusella*(3) sequence has the largest number of A nucleotides – 31.4%. The largest number of G nucleotides – 14.8 % has been obtained in the *E. humilis*(2) sequence.

Table 2. *Elachistinae* specimens and the rate of nucleotides of mtDNR COI gene fragments

Specimens	T (%)	C (%)	A (%)	G (%)	Total number of nucleotides
Perittia herrichiella(1)	38.7	17.5	29.7	14.1	576
Perittia herrichiella(2)	38.5	17.5	29.9	14.1	576
Elachista albifrontella(1)	38.0	16.7	30.9	14.4	576
Elachista albifrontella(2)	38.0	16.7	30.9	14.4	576
Elachista alpinella(1)	36.8	18.4	30.7	14.1	576
Elachista alpinella(2)	37.3	18.4	30.2	14.1	576
Elachista alpinella(3)	36.8	18.4	30.7	14.1	576
Elachista alpinella(4)	37.0	18.6	30.4	14.1	576
Elachista alpinella(5)	36.8	18.4	30.7	14.1	576
Elachista alpinella(6)	36.8	18.4	30.7	14.1	576
Elachista argentella(1)	39.0	16.7	30.3	14.1	575
Elachista argentella(2)	39.0	16.7	30.3	14.1	575
Elachista argentella(3)	38.8	16.7	30.0	14.5	640
Elachista argentella(4)	38.8	17.0	29.8	14.4	631
Elachista canapennella	41.5	14.6	30.0	13.9	576
Elachista consortella(1)	39.4	15.6	30.6	14.4	576
Elachista humilis(1)	39.4	15.6	30.6	14.4	576
Elachista humilis(2)	39.8	15.4	30.0	14.8	623
Elachista maculicerusella(1)	39.1	16.3	30.4	14.2	576
Elachista maculicerusella(2)	39.1	16.3	30.4	14.2	576
Elachista maculicerusella(3)	39.1	15.6	31.4	13.9	576
Elachista pollinariella(1)	39.5	16.0	30.8	13.7	575
Elachista pollinariella(2)	39.4	16.1	30.7	13.7	576
Elachista pollinariella(3)	39.5	16.0	30.8	13.7	575
Elachista pollinariella(4)	39.8	15.4	31.1	13.7	570
Elachista pollinariella(5)	39.1	16.5	30.7	13.7	576
Elachista pullicomella(1)	39.9	15.5	30.6	14.1	576
Elachista pullicomella(2)	40.2	15.8	30.3	13.7	575
Elachista pullicomella(3)	39.9	15.5	30.7	13.9	574
Elachista utonella(1)	37.2	18.1	30.3	14.3	623
Elachista utonella(2)	37.6	18.4	29.5	14.4	630
Average	38.7	16.7	30.4	14.1	584

COI sequences of *Elachistinae* moths have 63 codons (Table 3). The most common codon in mtDNA COI sequences of this subfamily is AUU(I) which amount to 17.3%. Codons GUC(V), GUG(V), GCA(A), GAU(D) and GGA(G) are most rarely detected in the sequences; they accounted for as little as 0.1 % only. The CUA(L) codon was not detected in COI sequences of *Elachistinae* moths.

The analysis of COI sequences of 11 *Elachistinae* species and 31 samples produced a 640 bp sequence alignment. The total nucleotide diversity and genetic divergence obtained for the *Elachistinae* moths contrasted with 229 polymorphic nucleotides. Molecular analysis results were summarized and the neighbor-joining (NJ) tree was constructed (Fig. 3).

Table 3. *Elachistinae* specimen codons and their number

Codon	Number (%)	Codon	Number (%)	Codon	Number (%)	Codon	Number (%)
UUU(F)	15.8(1.50)	UCU(S)	2.6(0.65)	UAU(Y)	16.9(1.56)	UGU(C)	4.6(1.22)
UUC(F)	5.2(0.50)	UCC(S)	1.7(0.44)	UAC(Y)	4.8(0.44)	UGC(C)	2.9(0.78)
UUA(L)	1.2(0.99)	UCA(S)	3.3(0.84)	UAA(*)	8.5(1.65)	UGA(*)	5.6(1.09)
UUG(L)	2.3(1.88)	UCG(S)	2.0(0.51)	UAG(*)	1.4(0.26)	UGG(W)	7.7(1.00)
CUU(L)	2.0(1.62)	CCU(P)	1.3(0.59)	CAU(H)	1.5(1.12)	CGU(R)	0.6(0.30)
CUC(L)	1.1(0.91)	CCC(P)	5.8(2.56)	CAC(H)	1.2(0.88)	CGC(R)	0.8(0.42)
CUA(L)	0.0(0.00)	CCA(P)	1.1(0.49)	CAA(Q)	0.3(1.23)	CGA(R)	0.4(0.20)
CUG(L)	0.7(0.60)	CCG(P)	0.8(0.37)	CAG(Q)	0.2(0.77)	CGG(R)	0.6(0.30)
AUU(I)	17.3(2.26)	ACU(T)	3.8(1.54)	AAU(N)	14.1(1.45)	AGU(S)	4.3(1.08)
AUC(I)	5.5(0.71)	ACC(T)	4.4(1.79)	AAC(N)	5.4(0.55)	AGC(S)	9.7(2.47)
AUA(I)	0.2(0.03)	ACA(T)	0.9(0.36)	AAA(K)	4.5(1.10)	AGA(R)	2.1(1.08)
AUG(M)	1.1(1.00)	ACG(T)	0.8(0.32)	AAG(K)	3.7(0.90)	AGG(R)	7.2(3.71)
GUU(V)	0.3(1.03)	GCU(A)	0.2(1.00)	GAU(D)	0.1(0.44)	GGU(G)	0.2(0.42)
GUC(V)	0.1(0.31)	GCC(A)	0.2(1.17)	GAC(D)	0.2(1.56)	GGC(G)	0.8(2.17)
GUA(V)	0.7(2.36)	GCA(A)	0.1(0.33)	GAA(E)	0.3(0.95)	GGA(G)	0.1(0.25)
GUG(V)	0.1(0.31)	GCG(A)	0.3(1.50)	GAG(E)	0.4(1.05)	GGG(G)	0.5(1.17)

The NJ tree contained two big clusters: A and B. Cluster A has two subclusters (A1, A2). The A1 subcluster includes genetically similar species: *Elachista alpinella, E. humilis, E. maculicerusella, E. albifrontella* and *E. consortella.* A total of 122 polymorphic nucleotides were detected in COI gene fragments in these species. The maximum parsimony analysis revealed 109 parsimony-informative characters. Genetically most similar were *E. humilis* and *E. consortella* species. In total 530 conservative sites were detected in their COI gene fragments. The most genetically different species in subcluster A1 are *E. alpinella* and *E. maculicerusella*: 73 polymorphic sites were detected in the sequences of these species. The A2 subcluster contains genetically similar species: *Elachista argentella*, *E. canapennella*, *E. pullicomella* and *E. pollinariella.* Genetically most similar are *E. argentella* and *E. canapennella* species. Their sequences have 578 conservative sites and 53 polymorphic nucleotides. Genetically most different species in subcluster A2 are *E. argentella* and *E. canapennella, E. argentella* and *E. pullicomella*. COI gene fragments of these moths have 60 polymorphic nucleotides from 640.

Cluster B includes *Elachistinae* species from two genera: *Perittia herrichiella* and *Elachista utonella*. COI gene fragments of this species differ in 102 polymorphic nucleotides. The largest intraspecific differences were found in the *Elachista utonella* species. Two specimens of this species have 39 variable nucleotides in COI sequences.

Conclusions

1. The investigated fragment of citochrome c oxidase I gene of *Elachistinae* moths has 229 polymorphic sites out of 640.

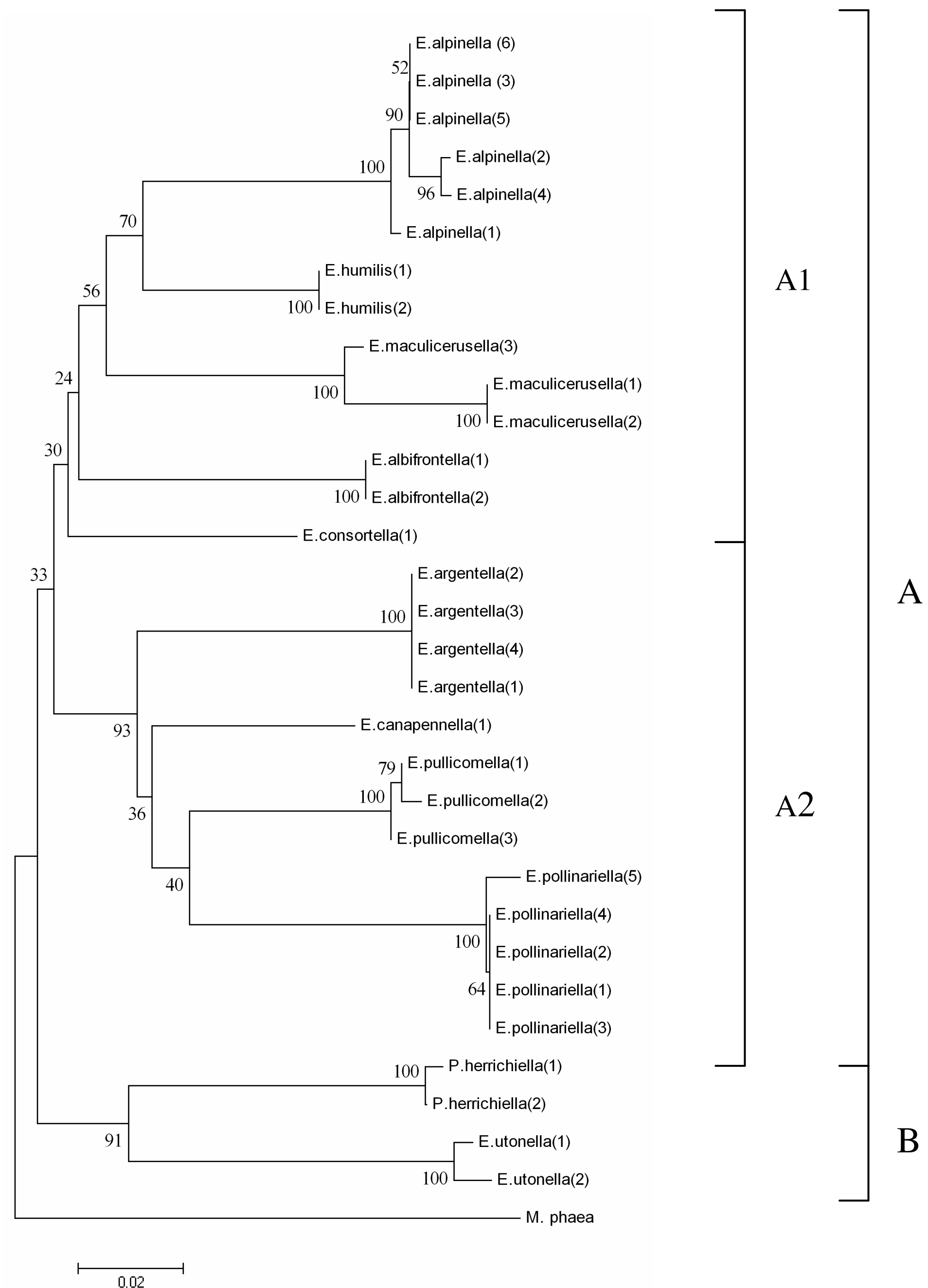

Fig. 3. The neighbor-joining tree of *Elachistinae* species. The tree was constructed by Kimura2-parameter model (bootstrap replications = 10000, complete deletion), based on analysis of 640 sites of COI gene fragment. Bootstrap values are shown above branches. The sequence of *Mythimna phaea* (GQ353295.1) was used as the outgroup.

2. COI gene fragments of *Elachistinae* sequences demonstrate a different distributional rate of nucleotides: 38.7% of T, 16.7% of C, 30.4% of A and 14.1% of G nucleotides.
3. COI sequences of *Elachistinae* moths have 63 codons, the most frequent codon being UAU(Y) (16.9%). Codon CUA(L) was not detected in the sequences of *Elachistinae* moths.
4. The greatest intraspecific differences were observed in *Elachista utonella*: two specimens of this species have 39 polymorphic nucleotides, whereas no

intraspecific differences were detected in *E. argentella, E. albifrontella* and *E. humilis* sequences.

References

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: 294–297.

Hebert P.D.N., Cywinska A., Ball S.L., De Waard J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270: 313–321.

Kaila L., Ståhls G. 2006. DNA barcodes: Evaluating the potential of COI to diffentiate closely related species of *Elachista* (Lepidoptera: Gelechioidea: Elachistidae) from Australia. *Zootaxa* 1170: 1–26.

Lunt D.H., Zhang D.X., Szymura J.M., Hewitt G.M. 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molecular Biology* 5:153–165.

Paulavičiūtė B., Paulauskas A. 2010. DNA diagnostics to identify *Elachista* (Lepidoptera: Elachistidae: Elachistinae) species. *Ecology & Safety* 4(1): 47–52.

Rubinoff D. 2006. Utility of mitochondrial DNA barcodes in species conservation. *Conservation Biology* 20(4): 1026–1033.

Sruoga V., Stunžėnas V., Paulavičiūtė B. 2009. COI gene as a molecular marker of *Elachista* species (Lepidoptera: Elachistidae: Elachistinae) from different Lithuanian populations. *Proceedings of the Latvian Academy of Sciences*, *Section B* 63(1/2): 21–24.

Tamura K., Dudley J., Masatoshi M., Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24(8): 1596–1599.

Elachistinae (Lepidoptera: Elachistidae) drugių rūšių citochromo oksidazės I geno sekų analizė

B. PAULAVIČIŪTĖ, A. PAULAUSKAS, V. SRUOGA

Santrauka

Molekulinių tyrimų metu buvo ištirta 11 *Elachistinae* pošeimio drugių rūšių. Buvo amplifikuoti 31 šios šeimos individo Citochromo c oksidazės I (COI) geno fragmentai. Gauti 640 bazių porų fragmentai, esantys 2239–2944 nukleotidų pozicijos atkarpoje. *Elachistinae* pošeimio drugių COI geno fragmentų sekų analizė parodė skirtingą nukleotidų pasiskirstymo dažnį. Užregistruoti 63 nukleotidų kodonai. Gauti rezultatai buvo apibendrinti ir artimiausių kaimynų jungimo metodu, naudojant Kimura 2 parametrų modelį, sudarytas *Elachistinae* pošeimio drugių mtDNR COI geno medis.

Received: October 15, 2010