Cumhuriyet Sci. J., 44(1) (2023) 28-35
DOI: https://doi.org/10.17776/csj.1196804

Cumhuriyet Science Journal

| csj.cumhuriyet.edu.tr | Founded: 2002 ISSN: 2587-2680 e-ISSN: 2587-246X Publisher: Sivas Cumhuriyet University

DNA Barcoding of Commercial Cockroaches in Turkey

Şeyda Berk [1,2,a,*], Ayşe Nur Pektaş [2,b]

[1] *Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, Sivas, Türkiye.*
[2] *Advanced Technology Research Center (CUTAM), Sivas Cumhuriyet University, Sivas, Türkiye.*
**Corresponding author*

Research Article

History
Received: 31/10/2022
Accepted: 14/03/2023

ABSTRACT

Accurate species identification has become a precondition for accomplished biodiversity administration and further genetic research. Species acquaintance technics require molecular tools such as DNA barcoding as well as morphological identification for accurate identification. Particularly, the application of subunit I of the mitochondrial *cytochrome c oxidase (COI)* gene for DNA barcoding for insects has approved to be very useful in species acquaintance. The main aim of this study is to generate the first reference library of DNA barcode for cockroaches in Turkey using previously published data. As a result of the literature research, it has been observed that no study has been carried out on the DNA barcode of Turkish cockroaches. Therefore, in this study, we evaluated the advantage of DNA barcoding applied to two cockroach samples from Turkey for the first time. Our working samples implicated 10 DNA barcodes grounded on sequences created from our present study and 109 other DNA barcodes from BOLD. Various molecular analyzes including genetic distance-origin assessment (NeighborJoining and Maximum Likelihood trees) has been applied to accurately identify and describe species. In addition, *Blaptica dubia (B. dubia)* (Serville, 1838) and *Nauphoeta cinerea (N. cinerea)* (Olivier, 1789) have been reported as the first country records. It has been observed that reference libraries like BOLD are not yet sufficiently populated with *COI* sequences of Turkish cockroach species. In order for Turkish cockroach bio-assessment and biodiversity studies to benefit from the advantages of DNA barcoding, it is of great importance that cockroach inventories and taxonomic studies include DNA barcodes.

Keywords: *Blaptica dubia, Nauphoeta cinerea*, Cockroaches, DNA barcoding.

[a] *sberk@cumhuriyet.edu.tr* *https://orcid.org/0000-0003-4687-0223*
[b] *aysenurpektas@cumhuriyet.edu.tr* *https://orcid.org/ 0000-0001-5621-2844*

Introduction

The concept of using DNA in taxonomic identification of species was first proposed by Hebert and Gregory [1]. The concept of DNA barcoding means "small standard DNA sequence" that can be used to distinguish species from one another. It is a technique based on PCR amplification of selected specific regions in the genome and used for species identification. Since the first barcoding study, over 6000 barcoding articles have been published [2]. The idea of a barcoding zone where species could distinguish life forms from one another was quickly adopted. Later, barcoding studies expanded with the idea that other organelle regions, markers and associated primer sets could also be used for barcoding [3].

For instance *MatK* for plants [4] and *rbcL* [5]; *ITS* for fungi [6] gene regions were used as barcode gene regions. Zoological DNA barcoding continues to grow as a popular way to identify animal specimens by similarity comparison of *cytochrome c oxidase subunit I (COI)* sequences of the mitochondrial genome [7-10]. The reasons why the *COI* region is a good barcode indicator; it does not contain insertions and deletions, is easy to isolate, has great differences between species although conserved, has a high copy number, relatively few differences within the species, does not contain introns, and the range of mutation rates in different regions of the molecule [11].

Essentially, barcoding includes two key elements, each as indirectly vital as the other in allowing precise identification through molecular characterization. These are the query and the target. The query is normally represented by a partial sequence of COIs of unknown origin (approximately 650 base pairs), while the target is a sequence of COIs located in a predetermined (typically by morphology and preferably species level) database or other repository. The purpose of DNA barcoding is to help recognize the diversity of species in our ecosystem, and species-level identification is crucial to achieving this goal [12].

Recently, those who do DNA barcoding have tried to create large-scale, carefully selected storage databases such as the Barcode Of Life Data System (BOLD), [10] and it serves securely because barcodes in databases are created in coordination with GenBank (NCBI). In addition, DNA barcoding offers the most convenient method for detecting cryptic species [13, 14].Moreover, differences at the molecular level also help to identify (marking) new species. Besides to the use of barcoding for sample identification, the fact that it can also be used in fields such as biosecurity, conservation biology, epidemiology and the food industry is proof that these studies should increase [12].

While the use of morphological data in the diagnosis and identification of insect species is a requirement in classical taxonomy, it is a time-consuming and species-specific method. However, DNA barcoding is a uniform

and practical techniques for insect species identification. It offers the opportunity to detect insect species at all developmental stages (egg, pupa, nymph, adult). It also works well in cases where morphological differentiation cannot be achieved. In addition, DNA barcoding is important for the rapid detection of invasive insect species. Diagnostic studies using the DNA barcoding approach have provided better solutions than other molecular techniques in terms of identifying new species and monitoring existing pest species. It is believed that the DNA barcoding technique can reliably resolve ecological and evolutionary connections in insect-host-plant relationships. In addition, DNA barcoding studies with flies are important in terms of human health and agriculture, especially in determining the vector-disease relationship [15].

Barcode data (*COI* gene region) has been entered for approximately 217,000 species from insect groups to date (September 2022) in the BOLD database and GenBank (NCBI) databases. Of these, about 76000 species are in the order Lepidoptera, 39000 species are in the order Hymenoptera and about 40000 species are in the order Coleoptera. The BOLD system (http://www.barcodinglife.org) has data for 1436 species from the order Blattodea (October 2022).

Cockroaches are a highly diverse group of insects, with over 4000 species commonly found worldwide [16]. Because cockroaches are capable of living in habitats containing various amounts of toxic substances, including environmental pollutants, microbial toxins, insecticides, it has been proposed as a good experimental model to study stress responses and detoxification abilities [17]. In addition, because cockroaches often live close to humans, they constitute an important reservoir for human pathogens. Therefore, they are of medical importance due to their potential to spread bacteria and other pathogens [18, 19]. Especially *N. cinerea* (Olivier, 1789) is important in that it is used as a model for correlations between sexual selection, bacterial infections, toxicology studies and metabolic rate and adaptation studies [20, 21]. The aim of this study is to investigate the suitability of DNA barcodes of commercially purchased *B. dubia* (Serville, 1838) and *N. cinerea* (Olivier, 1789) to identify cockroaches in Turkey and to create the first reference library of DNA barcode for cockroach species in Turkey using previously published data.

Material and Methods

Sample statement

Adults of *B. dubia* and *N. cinerea* (Blattodea: Blaberidae) were commercially obtained from a producer in Antalya/Turkey in February 2020 (https://www.antalyacekirge.net/). *B. dubia* and *N. cinerea* adults were brought to the laboratory by controlled storage in the RNAlater® (Qiagen) that is RNA stabilization reagent and stored at freezer (-20°C) until used in further experiments.

Total Genomic DNA Isolation, PCR Amplification and Sequencing

Total genomic DNA isolation from insect tissues was carried out with the sing a commercial DNA extraction kit (PureLink Genomic DNA Mini Kit, Invitrogen, USA) in accordance with the manufacturer's recommended protocol. Additional blank negative controls without tissue samples were used to exclude possible contamination during DNA isolation. All DNA samples were stored at -20 °C until further analysis. Universal *COI* barcode region primers, sequenced below, were used for amplification of the *COI* gene region [22].

LCO1490: 5'-GGTCAACAAATCATAAAGATATTGG-3'
HC02198: 5'-TAAACTTCAGGGTGACCAAAAAATCA-3'

With a final volume of 25 µL, the PCR mix contents are as follows: 2.5 µl 10X reaction buffer (KCL buffer), 1.25 µl (1.5 mM/µl) $MgCl_2$, 10pmol of each of the primers, 0.5 µl (0.2 mM) dNTPs, 5 U Taq DNA polymerase, and 1 µl (50 ng/µl) template DNA. PCR conditions for cockroaches' species included initial denaturation at 94 °C for 2 min; followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 51.3 °C for 1 min, extension at 72 °C for 2 min; and a final elongation at 72 °C for 10 min. PCR reactions were performed in a thermal cycler (T100, BioRad). Negative controls without a DNA template were included in all PCR runs. To ensure reproducibility and accuracy, PCR products were visualized by 1% agarose gel electrophoresis to assess success in amplification. By observing under UV transilluminator (Syngene), the presence of the expected amplicons was assessed by comparing with a standard DNA marker (Fermantas, EU).

After PCR anaysis, DNA sequencing was conducted by Macrogen through BMLabosis company. Sequence operations were performed bidirectionally with the primers in the PCR process. The files with the .ab1 extension obtained after this process were checked by reciprocating the forward and reverse sequences in the Geneious Prime program. After the Contig files were created, false peaks were corrected and BLAST scan was performed. These unique gene sequences have been uploaded to NCBI and Barcode of Life Data System (BOLD) databases (Supplementary data). Genbank accession numbers and BOLD accession numbers are shown in Table 1.

Bioinformatics and Phylogenetic Data Analysis

COI barcode region sequences were downloaded from 102 Blattodea orders from the NCBI gene bank to be used in phylogenetic analyzes and from 7 Mantodea orders to be used as outgroups. The accession numbers of the sequences are shown on the trees. A data set consisting of 119 sequences was obtained by combining all sequences. Sequence alignments were made in the MAFFT program. FASTA format was converted to NEXUS and PYHLIP formats by using ALTER alignment (http://www.sing-group.org/ALTER/) program so that the data can be used in different formats in phylogenetic analysis.

In order to obtain a tree with the maximum likelihood approach, BIC (Bayesian Information Criterion) and AICc (Akaike Information Criterion, corrected) values were determined by modeltesting in the MEGA 11 program, and the model was determined accordingly. Based on this model, PhyML and RaxML trees were constructed using the heuristic search method in the Geneious program. While creating the tree, Bootstrap (1000 bootstrap) method was used based on the Maximum likelihood method. The reproducibility of the measurement and its ability to give reliable and accurate results were effective in the selection of the Boostrap method. It also has applicability to all methods used for phylogenetic structuring and is able to assign probability-like repetition percentages to each possible part of the datasets in the branches of the resulting tree. The Boostrap method was preferred because it creates many new matrices with original dimensions and makes it possible to find the best tree for the analysis of each one, and the reliability of the branches in the tree is considered to be directly proportional to the frequency of branch exposure. In addition to these, giving information about how well the node to which the boostrap values belong is supported in terms of the model used in creating the phylogenetic tree has been another reason for the use of the method. Since it is aimed to establish a phylogenetic relationship in which phenetic characters are not used in the study; Maximum likelihood, which is a method that uses a clear criterion for the comparison of all possible trees and takes into account all possible trees in order to establish a relationship with the highest similarity in the trees created, to define and reveal the best one. The bootstrap discrimination power values are written on the branches. In addition, the evolutionary distance test based on the use of the Neighbor-joining evolution principle was carried out using the Kimura 2-Parameter (K2-P) model [23], which tests multiple displacements based on the characteristic that transitional nucleotide changes are higher than transversional changes in nature. All transitions and transversions are included in the nucleotide changes, and the value of "nst=6" is entered by choosing the gamma distribution for the variation between nucleotide positions.

Results and Discussion

Total DNA Quality

Figure 1. Image of PCR products produced with universal barcode primers in 1% agarose gel electrophoresis for the *COI* region. D1-D5: *B. dubia*, L1-L5*: N. cinerea*. M: marker

Mitochondrial *COI* gene barcode region was amplified by conventional PCR in five individuals of two cockroach species and PCR products were validated by running on 1% agarose gel. Band images were obtained in the correct gene region (~ 700bp) of all samples (Figure 1).

The Properties of DNA Sequence

The products obtained after PCR were sequenced and approximately 700 bp sequence was obtained, including the 658 bp barcode region. Barcode sequences of 696 bases for *B. dubia* (n=5) and 682 bases for *N. cinerea* (n=5) were created by editing the raw sequences in the Geneious Prime program. These sequences were uploaded to the NCBI GenBank system and BOLD systems and their accession numbers are given in Table 1. No deletion, insertion or stop codons were noted in the newly amplified sequences, indicating that the amplified sequences specify functional mitochondrial *COI* sequences. Overall, 99,5% sequence similarity was accomplished during the BLAST searching in the NCBI database. NJ tree-based identification via BOLD demonstrated cockroach species clustering in our study with other similar species in the database of *COI* sequences. This allowed these species to be defined precisely at the species level. The BIN analysis includes 10 BIN sets that are compatible with other barcode data in BOLD (Table 1). At the species level, BINs were determined in agreement with the morphology-based identification. Among these BINs, 5 (55%) are specified as the other 5 (18%), meaning that BIN refers to only one species (Table 1). BIN analyzes showed that the mean intraspecific distance of BIN varied between 0.18% and 0.55%, the maximum specific distance 0.62% and 0.61%, and the mean genetic distance to nearest neighbour 7.41 and 13.86 for *B. dubia* and *N. cinerea*, respectively (Table 1). For an exact resulting of *COI* genetic similarities, we merged our newly obtained *COI* sequence with *COI* sequences from other cockroach species in the BOLD system. Taken together, our dataset for phylogenetic analyzes included 102 individuals belonging to the order Blattodea. The approximately 700 bp *COI* sequence was aligned after the indeterminate bases were trimmed. Overall, GC percentages were recorded as 35% and 38% for *B. dubia* and *N. cinerea*, respectively (Table 1). All *COI* sequences were obtained heavily AT biased, with an average A+T content of 62.7%. Estimates of Evolutionary Divergence Between Sequences are shown in Table 2 and 3. The percentage of bases that are the same among *B. dubia* sequences is min. 98%, while for *N. cinerea* this ratio is 97%.

Table 1. Information of GenBank and BOLD database

Species	GenBank Accession No	BOLD ID	BIN URI	Average Distance in BIN (%)	Max Divergence in BIN (%)	Distance to NN (nearist neigbour) (%)	Marker Code	Sequence Length	GC %
Nauphoeta cinerea	MT861034	DLB006-20	BOLD:AAG9934	0.55 (p-dist)	1.61 (p-dist)	13.86 (p-dist)	COI-5P	682	37.2
Nauphoeta cinerea	MT861035	DLB007-20	BOLD:AAG9934					682	37.4
Nauphoeta cinerea	MT861036	DLB008-20	BOLD:AAG9934					682	37.2
Nauphoeta cinerea	MT861037	DLB009-20	BOLD:AAG9934					682	37.4
Nauphoeta cinerea	MT861038	DLB010-20	BOLD:AAG9934					682	37.4
Blaptica dubia	MT861039	DLB001-20	BOLD:ADC6507	0.18 (p-dist)	0.62 (p-dist)	7.41 (p-dist)		696	35.9
Blaptica dubia	MT861040	DLB002-20	BOLD:ADC6507					696	35.8
Blaptica dubia	MT861041	DLB003-20	BOLD:ADC6507					696	35.9
Blaptica dubia	MT861042	DLB004-20	BOLD:ADC6507					696	35.8
Blaptica dubia	MT861043	DLB005-20	BOLD:ADC6507					696	35.8

Table 2. Estimates of evolutionary divergence between sequences of *B. dubia*

	D1	D2	D3	D4	D5
D1		98.849	98.750	98.851	98.189
D2	98.849		98.886	98.499	98.464
D3	98.750	98.886		98.470	98.574
D4	98.851	98.499	98.470		98.082
D5	98.189	98.464	98.574	98.082	

Table 3. Estimates of evolutionary divergence between sequences of *N. cinerea*

	N1	N2	N3	N4	N5
N1		98.107	98.519	98.090	97.535
N2	98.107		98.796	98.629	98.277
N3	98.519	98.796		99.386	99.169
N4	98.090	98.629	99.386		98.886
N5	97.535	98.277	99.169	98.886	

Phylogenetic Analyses

In order to obtain trees with the maximum likelihood approach in phylogenetic analyses, first the model was determined. Accordingly, the results obtained are as in Table 4. On the basis of hierarchical likelihood-ratio tests as implemented in Modeltest 3.0, the model General Time Reversible (GTR) model + Gamma distribution + invariable sites were used (GTR + G + I, -lnL = 16,941.61, P < 0.001, AIC = 34,374.77, BIC = 36,642.61). The gamma distribution and proportion of invariant sites were set as 0.77 and 0.42 (estimated by Modeltest), respectively (Table 4).

The K2P/NJ tree, *COI* gene datasets, clustered individuals of the same species with high bootstrapping values (Figure 2). In our present study, most of the morphologically identified species formed distinct *COI* clusters that were well differentiated and supported by high bootstrapping values (Figure 2). When we look at the information given to us by the phylogenetic tree created by the NJ method, we see three main clades except the outgroups. *B. dubia* and *N. cinerea* individuals are located in different clades (Figure 2). Considering the branch values in the tree, considering the entire Blattodea order, this tree shows the accuracy of the barcoding work, although the distinction between species is not perfect within the clades. ML trees (1000bootsrap) created with the Maximum likelihood approach have patterned to support the NJ tree (Figure 3, 4). It supports the distinction between the 3 main clades subsets. The selection of outgroups in the Mantodea order increased the success of all trees (Figure 2-4).

Our study demonstrates the first cockroach DNA barcode analysis in Turkey containing mitochondrial COI gene sequences. In present study, the general achievement rate of DNA sequencing and barcode creation varies between 81.3 and 94.8%. These values represent the percentage of sequencing quality of unclipped bases in a sequence (a total of 20 forward and reverse reads for 10 samples).

The main purpose of this study is to show the compatibility of DNA barcodes for identification of cockroaches in Turkey and to create the first DNA barcode reference library for cockroaches in Turkey using previously published data. Tree-based identifying usage of ML and NJ techniques, particularly applicable to newly barcoded samples, demonstrated that both species were clearly determinable from all other species by the formation of distinct, non-overlapping *COI* clusters. When we compared species identification methods, our study showed that it is compatible with tree-based identification using ML and NJ technics. When we examine the DNA barcoding studies on cockroaches in the literature, Evangelista et al., used molecular identification using the *COI* DNA barcode gene to confirm that the new invasive insect species, *Periplaneta japonica* Karny, 1908, which does not conform to the typical morphology of the American cockroach *Periplaneta americana L* [24]. In another study by Farah Haziqah et al., they conducted research on *Blastocystis spp* infections. They collected a total of 151 cockroaches, mostly nymph and adult stages, from various residential species in the Malaysian state of Perak and Selangor, and reported that approximately half of the scanned cockroach gut contents were positive and determined that this infection was closely related to the host stage and housing types using DNA barcoding [25]. In a review by Miskelly and Paiero, it was reported that DNA barcodes are available for more than 60% of the species belonging to the order *Blattodea*, *Orthoptera, Dermaptera*, and *Phasmida* known to be found in Canada [26]. Liao et al., used *COI* DNA barcodes in Hainan Province, China to confirm sexual dimorphism occurring in the cockroach species, *Laevifaciesquadrialata gen. et sp.*

nov [27]. Von Beeren et al., analyzed mitochondrial DNA barcodes in American cockroach (*Periplaneta americana*) by collecting 284 cockroach samples from various states of America. *P. americana* barcode sequences formed a distinct monophyletic lineage from other *Periplaneta* species, and they found three distinct *P. americana* haplogroups between groups. They suggested that this genetic pattern likely reflects multiple introductions from genetically different source populations and subsequent interbreeding in the invasive range [28]. A new species of 2021 *Bundoksia lucañas* from China has been identified and Li et al. performed molecular identification by DNA barcoding with mitochondrial *COI* data to reveal relationships between *Bundoksialongissima sp. nov.* populations. They reported details of the female genitalia in addition to the known external morphology and male genitalia of this new species found [29]. Considering all these studies, DNA barcoding is important in determining the biodiversity of a country, determining the disease vector relationship in anthropophilic species, molecular identification of agriculturally important insects, determining the place of cryptic species in classical taxonomy and identifying new species. The first DNA barcoding records from cockroaches to NCBI and BOLD databases were made with this study. Molecular identification of these commercially available species was carried out in this study. These records in the database form the basis for future *COI* barcoding studies. It has been observed that reference libraries like BOLD do not yet contain information on the *COI* sequences of Turkish cockroach species. In order for the bio-assessment and biodiversity studies of Turkish cockroaches to benefit from the advantages of DNA barcoding, it is of great importance that the inventories and taxonomic studies of cockroaches include DNA barcodes.

Table 4. Maximum Likelihood fits of 24 different nucleotide substitution models

Model	#Param	BIC	AICc	lnL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G	A=>T	A=>C	A=>G	T=>A	T=>C	T=>G	C=>A	C=>T	C=>G	G=>A	G=>T	G=>C
GTR+G+I	245	36642.61192	34374.77342	-16941.6103	0.4156809	0.771571429	2.52581738	0.306124546	0.322370195	0.203821901	0.167683358	0.08	0.01	0.05	0.08	0.23	0	0.01	0.37	0.03	0.09	0.01	0.04
GTR+G	244	36801.84756	34543.25924	-17026.85953	n/a	0.298374522	2.7181121	0.306124546	0.322370195	0.203821901	0.167683358	0.07	0.01	0.04	0.07	0.24	0	0.01	0.38	0.04	0.08	0.01	0.05
TN93+G	242	36987.60595	34747.51814	-17131.00154	0.419021182	0.794629732	2.02144266	0.306124546	0.322370195	0.203821901	0.167683358	0.05	0.03	0.06	0.05	0.2	0.03	0.05	0.31	0.03	0.11	0.05	0.03
TN93+G	241	37206.64976	34975.81228	-17246.15486	n/a	0.334693965	1.8457286	0.306124546	0.322370195	0.203821901	0.167683358	0.06	0.04	0.06	0.05	0.19	0.03	0.05	0.29	0.03	0.12	0.06	0.04
HKY+G+	241	37389.86874	35159.03127	-17337.76435	0.428963016	0.816970799	2.24549667	0.306124546	0.322370195	0.203821901	0.167683358	0.05	0.03	0.12	0.05	0.14	0.02	0.05	0.23	0.02	0.22	0.05	0.03
T92+G+I	239	37564.49977	35352.16312	-17436.34269	0.429664807	0.822659706	2.21482764	0.314247371	0.314247371	0.185752629	0.185752629	0.05	0.03	0.13	0.05	0.13	0.03	0.05	0.22	0.03	0.22	0.05	0.03
HKY+G	240	37632.26109	35410.674	-17464.59194	n/a	0.354104454	2.02569462	0.306124546	0.322370195	0.203821901	0.167683358	0.05	0.03	0.11	0.05	0.14	0.03	0.05	0.22	0.03	0.21	0.05	0.03
T92+G	238	37818.01563	35614.92946	-17568.73203	n/a	0.359260404	1.96345054	0.314247371	0.314247371	0.185752629	0.185752629	0.05	0.03	0.13	0.05	0.13	0.03	0.05	0.21	0.03	0.21	0.05	0.03
GTR+I	244	38420.48621	36161.8979	-17836.17886	0.452345793	n/a	1.90651637	0.306124546	0.322370195	0.203821901	0.167683358	0.11	0.03	0.07	0.1	0.19	0	0.04	0.31	0.01	0.13	0.01	0.01
K2+G+I	238	38599.1877	36396.10153	-17959.31806	0.436419573	0.989560432	2.0874251	0.25	0.25	0.25	0.25	0.04	0.04	0.17	0.04	0.17	0.04	0.04	0.17	0.04	0.17	0.04	0.04
K2+G	237	38672.94286	36479.10723	-18001.82705	n/a	0.271724165	1.83970723	0.25	0.25	0.25	0.25	0.04	0.04	0.16	0.04	0.16	0.04	0.04	0.16	0.04	0.16	0.04	0.04
TN93+I	241	39091.9442	36861.10673	-18188.80208	0.451613299	n/a	1.7198867	0.306124546	0.322370195	0.203821901	0.167683358	0.06	0.04	0.06	0.06	0.18	0.03	0.06	0.28	0.03	0.12	0.06	0.04
HKY+I	240	39409.94038	37188.35329	-18353.43158	0.445714872	n/a	1.67024548	0.306124546	0.322370195	0.203821901	0.167683358	0.06	0.04	0.11	0.06	0.13	0.03	0.06	0.21	0.03	0.2	0.06	0.04
T92+I	238	39557.78066	37354.69449	-18438.61454	0.451371089	n/a	1.6630279	0.314247371	0.314247371	0.185752629	0.185752629	0.06	0.03	0.12	0.06	0.12	0.03	0.06	0.2	0.03	0.2	0.06	0.03
JC+G+I	237	39674.12233	37480.2867	-18502.41679	0.433418551	1.12460805	0.5	0.25	0.25	0.25	0.25	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
JC+G	236	39735.92002	37551.33499	-18538.94705	n/a	0.270083466	0.5	0.25	0.25	0.25	0.25	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
K2+I	237	40234.18256	38040.34694	-18782.44691	0.45106856	n/a	1.49767968	0.25	0.25	0.25	0.25	0.05	0.05	0.15	0.05	0.15	0.05	0.05	0.15	0.05	0.15	0.05	0.05
JC+I	236	41254.31454	39069.72951	-19298.14431	0.45264965	n/a	0.5	0.25	0.25	0.25	0.25	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
GTR	243	43187.63347	40938.29538	-20225.38389	n/a	n/a	1.47827418	0.306124546	0.322370195	0.203821901	0.167683358	0.13	0.03	0.07	0.12	0.18	0	0.05	0.28	0	0.08	0.01	0.01
TN93	240	44194.81603	41973.22894	-20745.86941	n/a	n/a	1.46373195	0.306124546	0.322370195	0.203821901	0.167683358	0.06	0.04	0.06	0.06	0.17	0.03	0.06	0.26	0.03	0.11	0.06	0.04
HKY	239	44512.24136	42299.90471	-20910.21348	n/a	n/a	1.43248236	0.306124546	0.322370195	0.203821901	0.167683358	0.06	0.04	0.1	0.06	0.12	0.03	0.06	0.19	0.03	0.18	0.06	0.04
T92	237	44647.91347	42454.07785	-20989.31236	n/a	n/a	1.43214119	0.314247371	0.314247371	0.185752629	0.185752629	0.06	0.04	0.11	0.06	0.11	0.04	0.06	0.19	0.04	0.19	0.06	0.04
K2	236	44983.71947	42799.13444	-21162.84677	n/a	n/a	1.60385748	0.25	0.25	0.25	0.25	0.05	0.05	0.15	0.05	0.15	0.05	0.05	0.15	0.05	0.15	0.05	0.05
JC	235	45938.14175	43762.80736	-21645.68932	n/a	n/a	0.5	0.25	0.25	0.25	0.25	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08

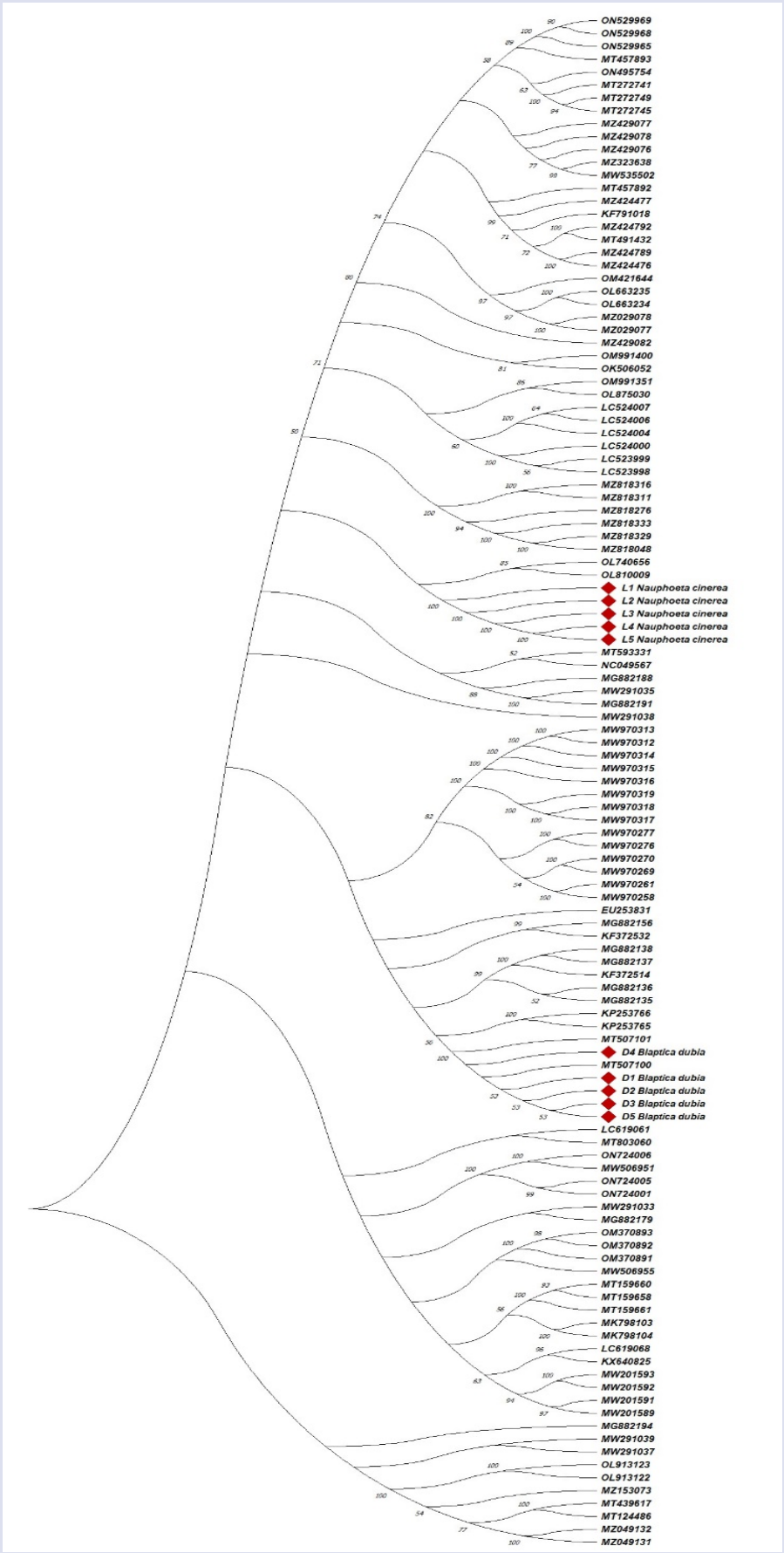

Figure 2. K2P/Neighbour-Joining (NJ) tree with bootstrap support (1000 replicates) showing clustering of species for *COI* sequences. Red nodes are the sequences obtained in this study.

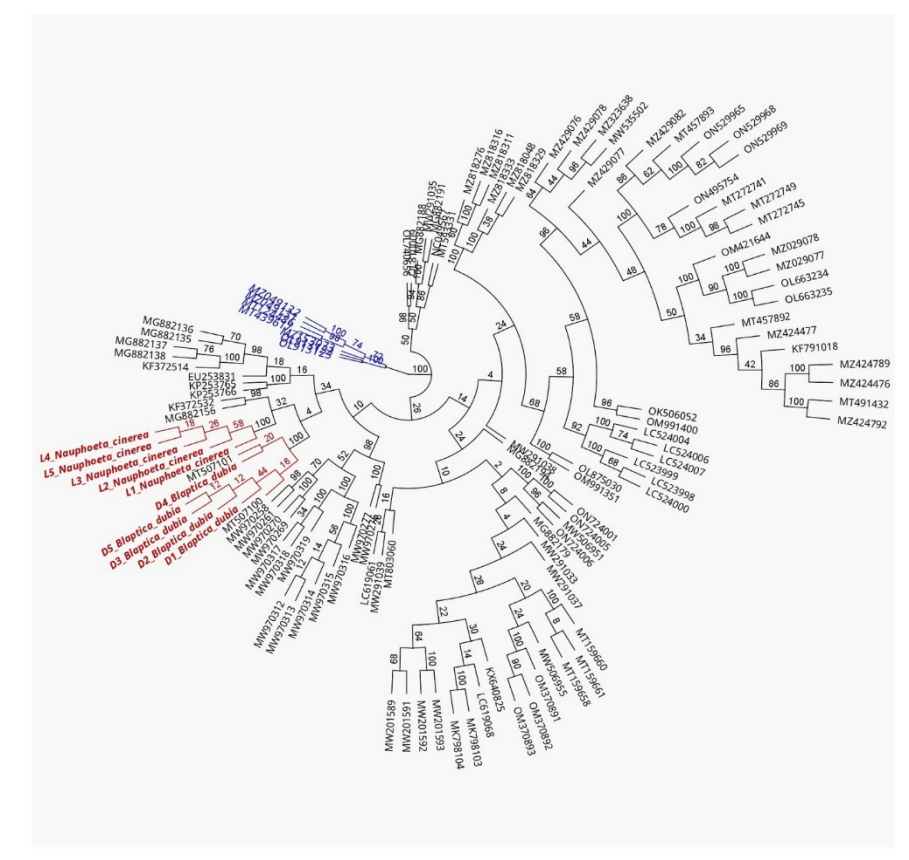

Figure 3. Mitochondrial *COI* barcode Maximum likelihood (PhyML) tree. In the tree obtained by bootstrap method, branch values are shown above. Sequences obtained in this study are shown in red and outgroup in blue.

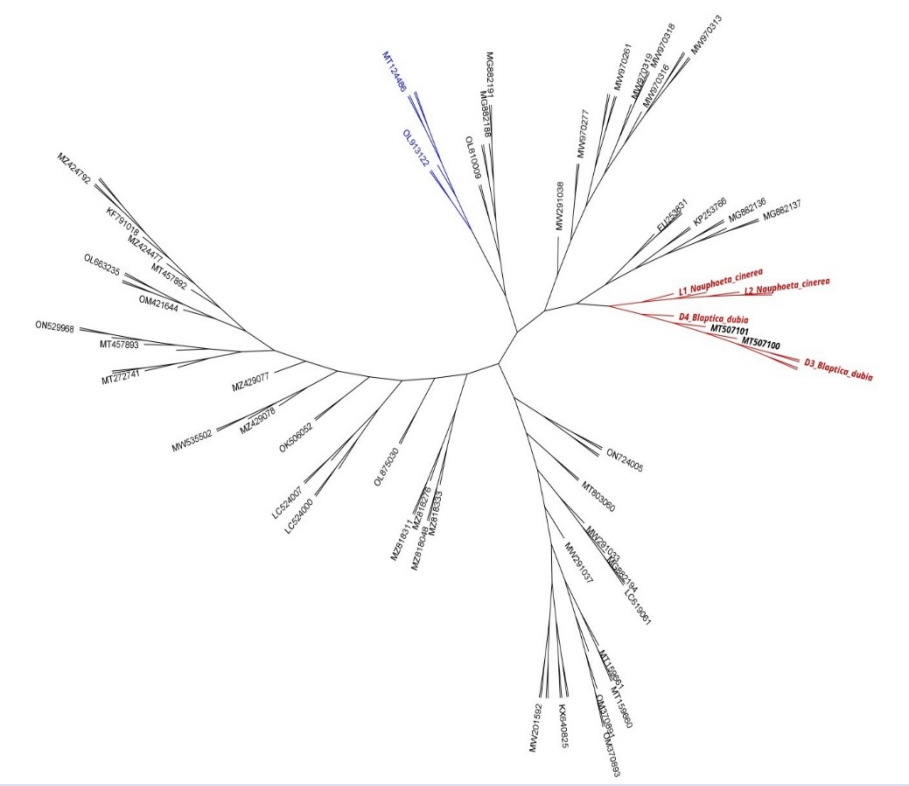

Figure 4. Mitochondrial *COI* barcode Maximum likelihood (RaxML). Sequences obtained in this study are shown in red and outgroup in blue.

Acknowledgment

All analyses in this study were performed in the Molecular Biology laboratory of the Sivas Cumhuriyet University Advanced Technology and Research Center (CUTAM).

Conflict of Interest

The authors state that did not have conflict of interests

References

[1] Hebert P.D., Gregory T.R. The promise of DNA barcoding for taxonomy, *Syst. Biol.,* 54 (2005) 852-859.

[2] Dormontt E.E., Van Dijk K.-J., Bell K.L., Biffin E., Breed M.F., Byrne M., Caddy-Retalic S., Encinas-Viso F., Nevill P.G., Shapcott A. Advancing DNA barcoding and metabarcoding applications for plants requires systematic analysis of herbarium collections—an Australian perspective, *Front. Ecol. Evol. FRONT ECOL EVOL.,* 6 (2018) 134.

[3] Bohmann K., Mirarab S., Bafna V., Gilbert M.T.P. Beyond DNA barcoding: The unrealized potential of genome skim data in sample identification, *Wiley Online Library*, 2020.

[4] Lahaye R., Van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G., Maurin O., Duthoit S., Barraclough T.G., Savolainen V. DNA barcoding the floras of biodiversity hotspots, *PNAS.*, 105 (2008) 2923-2928.

[5] Fazekas A.J., Burgess K.S., Kesanakurti P.R., Graham S.W., Newmaster S.G., Husband B.C., Percy D.M., Hajibabaei M., Barrett S.C. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well, *PloS one.*, 3 (2008) e2802.

[6] Schoch C.L., Seifert K.A., Huhndorf S., Robert V., Spouge J.L., Levesque C.A., Chen W., Consortium F.B., List F.B.C.A., Bolchacova E. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi, *PNAS.*, 109 (2012) 6241-6246.

[7] Hebert P.D., Cywinska A., Ball S.L., DeWaard J.R. Biological identifications through DNA barcodes, *Proc. R. Soc. B: Biol.*, 270 (2003) 313-321.

[8] Hebert P.D., Penton E.H., Burns J.M., Janzen D.H., Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator, *PNAS.*, 101 (2004) 14812-14817.

[9] Hajibabaei M., Janzen D.H., Burns J.M., Hallwachs W., Hebert P.D. DNA barcodes distinguish species of tropical Lepidoptera, *PNAS.*, 103 (2006) 968-971.

[10] Ratnasingham S., Hebert P.D. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org), *Mol. Ecol. Notes.*, 7 (2007) 355-364.

[11] Elyasigorji Z., Izadpanah M., Hadi F., Zare M. Mitochondrial genes as strong molecular markers for species identification, *The Nucleus*, (2022).

[12] Kvist S. Barcoding in the dark? A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge, *Mol. Phylogenet Evol.*, 69 (2013) 39-45.

[13] Witt J.D., Threloff D.L., Hebert P.D. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation, *Mol. Ecol.*, 15 (2006) 3073-3082.

[14] Pfenninger M., Schwenk K. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions, *BMC Evol. Biol.*, 7 (2007) 1-6.

[15] Jalali S., Ojha R., Venkatesan T. DNA barcoding for identification of agriculturally important insects, New horizons in insect science: Towards sustainable pest management, *Springer,* (2015) 13-23.

[16] Pomés A., Arruda L.K. Investigating cockroach allergens: aiming to improve diagnosis and treatment of cockroach allergic patients, *Methods*, 66 (2014) 75-85.

[17] Zhang J., Zhang Y., Li J., Liu M., Liu Z.,Midgut Transcriptome of the Cockroach *Periplaneta americana* and Its Microbiota: Digestion, Detoxification and Oxidative Stress Response, *PLoS One.*, 11 (2016) e0155254.

[18] Fakoorziba M.R., Eghbal F., Hassanzadeh J., Moemenbellah-Fard M.D. Cockroaches (*Periplaneta americana* and *Blattella germanica*) as potential vectors of the pathogenic bacteria found in nosocomial infections, *Ann Trop Med Parasitol.*, 104 (2010) 521-528.

[19] Atiokeng Tatang R.J., Tsila H.G., Wabo Poné J. Medically Important Parasites Carried by Cockroaches in Melong Subdivision, Littoral, Cameroon, *J Parasitol Res.*, 2017 (2017) 7967325.

[20] Bouchebti S., Durier V., Pasquaretta C., Rivault C., Lihoreau M. Subsocial Cockroaches Nauphoeta cinerea Mate Indiscriminately with Kin Despite High Costs of Inbreeding, *PLoS One.*, 11 (2016) e0162548.

[21] Schimpf N.G., Matthews P.G., White C.R.,Standard metabolic rate is associated with gestation duration, but not clutch size, in speckled cockroaches Nauphoeta cinerea, *Biol Open.*, 1 (2012) 1185-1191.

[22] Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates, *Mol Mar Biol Biotechnol.*, 3 (1994) 294-299.

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

[24] Evangelista D., Buss L., Ware J.L. Using DNA barcodes to confirm the presence of a new invasive cockroach pest in New York City, *J Econ Entomol.*, 106 (2013) 2275-2279.

[25] Farah Haziqah M.T., Nur Asyiqin M.N., Mohd Khalid M.K.N., Suresh K., Rajamanikam A., Chandrawathani P., Mohd Zain S.N. Current status of Blastocystis in cockroaches, *Trop Biomed.*, 34 (2017) 741-745.

[26] Miskelly J., Paiero S.M. Mantodea, Blattodea, Orthoptera, Dermaptera, and Phasmida of Canada, *Zookeys*, (2019) 255-269.

[27] Liao S., Wang Z., Che Y. A new genus and a new species in the subfamily *Polyzosteriinae* (Blattodea, Blattidae) from China, *Zookeys*, 852 (2019) 85-100.

[28] von Beeren C., Stoeckle M.Y., Xia J., Burke G., Kronauer D.J. Interbreeding among deeply divergent mitochondrial lineages in the American cockroach (*Periplaneta americana*), *Sci Rep.*, 5 (2015) 8297.

[29] Li Y., Luo X., Zhang J., Wang Z., Che Y. A new species of Bundoksia Lucañas, 2021 with comments on its subfamilial placement, based on morphological and molecular data, *Zookeys*, 1085 (2022) 145-163.4