

OPOLE SCIENTIFIC SOCIETY

NATURE JOURNAL

No 47 – 2014: 68–81

EVALUATION OF THREE OPEN-SOURCE SOFTWARE FOR ANALYSIS OF DNA POLYMORPHISM PATTERNS: A CASE STUDY ON NUCLEOTIDE DIVERSITY IN THE MITOCHONDRIAL 16S rDNA AND CYTOCHROME OXIDASE 1 GENES IN THE SUBORDER HETEROPTERA (INSECTA: HEMIPTERA)

PAULINA GRADOWSKA¹, JERZY A. LIS²

Center for Biodiversity Studies, Department of Biosystematics, Opole University, Oleska 22,
45-052 Opole, Poland

¹gradowska.p@gmail.com, ²cydnus@uni.opole.pl

ABSTRACT: Three software packages utilized for genetic diversity analysis, including DnaSP, ProSeq and MEGA were compared to determine if the results generated by these programmes are comparable. The analyses of mitochondrial 16S ribosomal DNA (105 sequences) and cytochrome oxidase 1 (460 sequences) of 24 species representing the hemipteran suborder Heteroptera (true bugs) included the following parameters: position of SNP and SNP alleles, number of polymorphic sites, average number of nucleotide differences, number of haplotypes, haplotype diversity variances, standard deviation of haplotype diversity, and neutrality tests. The results suggest that DnaSP appeared to be the most comprehensive software for such analyses.

KEY WORDS: Hemiptera, Heteroptera, true-bugs, *Appasus japonicus* (Vuillefroy 1864), molecular ecology, single nucleotide polymorphism, haplotype, neutrality test, DnaSP, ProSeq, MEGA

Introduction

DNA polymorphism analyses are widely used in molecular studies among and within populations in different ecological approaches. Moreover, the results of analyses interpreting DNA polymorphism data are now very often utilized in other fields, including conservation genetics, molecular taxonomy, phylogeography, medicine, pharmacogenomics, forensic biology, and others.

There are many software packages designed for such analyses; however, in spite of the interest in this subject, no data on the comparative evaluation of different software packages are available in the literature.

In the present study, we have compared three open-source software packages, most frequently used in studies on the population genetic diversity, including DnaSP, ProSeq, and MEGA, to determine if the results generated by these programs correlate with each other and adequately characterize the population DNA polymorphism factors.

Materials and methods

Material. Two subunits of mitochondrial DNA, i.e., 16S rDNA, and cytochrome oxidase I (COI) were analyzed. Sequences for all analysed true-bug species (Insecta: Hemiptera: Heteroptera) were downloaded from GenBank [Table 1; sequence accession numbers for *Appasus japonicus* (Vuillefroy, 1864) are provided in Table 2; for all other species in the Supplementary Material]. A total of 565 sequences (105 sequences of 16S rDNA and 460 sequences of COI) of 24 species were included in the comparative studies.

Computer software. Genetic diversity calculations were performed using three open-source software packages, namely, DnaSP v. 5.10.01 (Librado and Rozas 2009), Proseq v. 4.3 (Filatov 2002) and MEGA v. 6.0 (Tamura et al. 2013).

The DnaSP (DNA Sequence polymorphism) programme can estimate several measures of DNA to sequence variation within and between populations as well as to gene flow, linkage disequilibrium and many other parameters. It computes most widely used quantitative statistics of DNA polymorphism levels, including the number of segregating sites, an average number of differences in nucleotides, a number of haplotypes and their diversity. The software consists of exhaustive molecular analysis of populations based on the coalescence theory (Librado and Rozas 2009; for the software application see, e.g., Nagaraja et al. 2004; Ciampi et al. 2005; Barques et al. 2008; Castro et al. 2010; Lv et al. 2011; Taylor et al. 2011; Wang et al. 2013).

The ProSeq (Processor sequences) software package is usually used to analyze multiple data files in the collection of DNA polymorphisms; it can handle huge data sets composed even of thousands of genes, because the sequence length and sequence number utilized for analysis is unlimited (Filatov 2002; for the software application see Mahani et al. 2008; Fan et al. 2009).

MEGA (Molecular Evolutionary Genetics Analysis) performs a large number of analyses usually related to molecular phylogeny of taxa, but can also perform the comparative analysis of sequences for population ecology studies, e.g., haplotype and nucleotide diversity, a number of polymorphic sites and alleles, and neutrality tests (Tamura et al. 2013; for the software application see Nagaraja et al. 2004; Barques et al. 2008; Castro et al. 2010; Taylor et al. 2011; Zhao and Wu 2011; Wang et al. 2013).

Analyses. All sequences originally downloaded from GenBank in FASTA format were aligned using the default parameters of Clustal W (Larkin et al. 2007); additionally, a manual matching was performed in JalView v. 2.8 (Waterhouse et al. 2009). The analysis of DNA polymorphic sites included the position of SNP and SNP alleles, the number of polymorphic sites, the nucleotide diversity per site, the average number of nucleotide differences, the number of haplotypes, the haplotype diversity variances and a standard deviation of haplotype diversity, and the G+C content. Evolutionary distances among populations were also calculated using the Tajima D' neutrality test (Tajima 1989) in all three software programmes. Moreover, the Fu's test (Fu 1997), and the Fu and Li's D and F tests were calculated (Fu and Li 1993) (Tables 3 & 5).

Results and conclusions

The comparative results for 16S rDNA and COI are presented for *Appasus japonicus* (Vuillefroy, 1864) (Heteroptera: Belostomatidae: Belostomatinae), as the example justifying our conclusions (Tables 2 and 3 for 16S rDNA, Tables 4 and 5 for COI); the same data for all remaining species were placed in the Supplementary Material (Appendix; data presented in red indicate significant differences in obtained results; data presented in blue indicated insignificant differences in obtained results).

Results regarding the number of polymorphic sites, the position of SNP and SNP alleles, the G+C content, nucleotide diversity per site, average number of nucleotide

differences, number of haplotypes and haplotype diversity, as well as the Tajima's D neutrality test are the same regardless of the software used for the analyses.

However, a variance of haplotype diversity and its standard deviation, as well as two neutrality tests (i.e., Fu and Li's F, Fu's Fs) can be calculated only with DnaSP.

These statistical tests for non-neutral evolution are, besides the classic Tajima's D neutrality test, routinely used in different types of phylogeographic and population studies (see Michel et al. 2005; Depaquit et al. 2008; Szalanski and Magnus 2010; Luzhang et al. 2010; Zhang et al. 2010, 2012, Twerdochlib et al. 2012; Weckworth et al. 2012; Malaney et al. 2013; de Volo et al. 2013; Xu et al. 2014).

Therefore, although all three compared software packages can analyze the DNA polymorphism data, DnaSP appeared to be the most comprehensive software for such analyses.

We hope our results will help at least a little in deciding which software to choose for the DNA polymorphism pattern analyses, especially for researchers starting their scientific efforts in population molecular studies.

Bibliography

Barques M. D., Klisiowicz D. R., Gonzales-Candelas F., Ramsey J. M., Monroy C., Ponce C., Salazar-Schettino P. M., Panzera F., Abad-Franch F., Sousa O. E., Schofield C. J., Dujardin J. P., Guhl F., Mas-Coma S. 2008. Phylogeography and genetic variation of *Triatoma dimidiata*, the Main Chagas Disease Vector in central America, and its position within the genus *Triatoma*. PLOS Neglected Tropical Diseases 2(5): e233.

Castro J. A., Barrio E., Gonzalez A., Picornell A., Ramon M. M., Moja A. 2010. Nucleotide diversity of a ND5 fragment confirms that population expansion is the most suitable explanation for the mtDNA haplotype polymorphism of *Drosophila subobscura*. Genetica 138: 819-829.

Ciampi M. B., Kuramae E. E., Fenille R. C., Meyer M. C., Sauza N. L., Ceresini P. C. 2005. Intraspecific evolution of *Rhizoctonia solani* AG-1 IA associated with soybean and rice in Brazil based on polymorphisms at the ITS-5.8S rDNA operon. European Journal of Plant Pathology 113: 183-196.

Depaquit J., Lienard E., Verzeaux-Griffon A., Ferté H., Bounamous A., Gantier J.-Ch., Hanafi A. H., Jacobson R. L., Maroli M., Moin-Vaziri V., Müller F., Özbel

- Y., Svobodova M., Volf P., Léger N. 2008. Molecular homogeneity in diverse geographical populations of *Phlebotomus papatasi* (Diptera, Psychodidae) inferred from ND4 mtDNA and ITS2 rDNA. Epidemiological consequences. *Infection, Genetics and Evolution* 8: 159-170.
- Fan X., Sha L.N., Yang R. W., Zhang H. Q., Kang H. Y., Ding C. B., Zhang L., Zheng Y. L., Zhou Y. H. 2009. Phylogeny and evolutionary history of *Leymus* (Triticeae: Poaceae) based on a single-copy nuclear gene encoding plastid acetyl-CoA carboxylase. *BMC Evolutionary Biology* 9(1): e247.
- Filatov D. A. 2002. ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. *Molecular Ecology Notes* 2: 621-624.
- Fu Y. X. 1997. Statistical tests of neutrality of mutation against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.
- Fu Y. X., Li W. H. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.
- Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan P. A., McWilliam H., Valentin F., Wallace I. M., Wilm A., Lopez R., Thompson J. D., Gibson T. J., Higgins D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
- Librado P., Rozas J. 2009. DNASP v.5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Luzhang R., Bei A., Backstrom N., Huaxing L., Longying W., Lixun Z., Naifa L. 2010. Phylogeographic structure and gene flow of Himalayan snowcock (*Tetraogallus himalayensis*). *Animal Biology* 60: 449-465.
- Lv S., Zhang Y., Liu H. X., Hu L., Liu W., Liu Q., Li S. Z., Hu W., Utzinger J., Zhou X.N. 2011. Population genetics of *Pomacea* spp. in mainland of China. *Chinese Journal of Schistosomiasis Control* 23: 178-182.
- Mahani M. K., Inomata N., Saboori A., Tabatabaei B. E. S., Ishiyama H., Ariama A., Schmiat A. E. 2008. Genetic variation in populations of *Allothrombium pulvirum* (Acari: Trombididae) from Northern Iran revealed by mitochondrial COXI and nuclear rDNA ITS sequences. *Experimental and Applied Acarology* 48: 273-289.
- Malaney J. L., Conroy Ch. J., Moffitt L. A., Spoonhunter H. D., Patton J. L., Cook J. A. 2013. Phylogeography of the western jumping mouse (*Zapus princeps*) detects deep and persistent allopatry with expansion. *Journal of Mammalogy* 94: 1016-1029.

- Nagaraja, Nagaraju J., Ranganath H. A. 2004. Molecular phylogeny of the *nasuta* subgroup of *Drosophila* based on 12S rRNA, 16S rRNA and COI mitochondrial genes, RAPD and ISSR polymorphisms. *Genes & Genetic Systems* 79: 293-299.
- Szalanski A. L., Magnus R. M. 2010. Mitochondrial DNA characterization of Africanized honey bee (*Apis mellifera* L.) populations from the USA. *Journal of Apicultural Research and Bee World* 49: 177-185.
- Tamura K., Stecher G., Peterson D., FilipSKI A., Kumar S. 2011. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30(12): 2725-2729.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.
- Taylor S. J., Downie D. A., Paterson I. D. 2011. Genetic diversity of introduced populations of the water hyacinth biological control agent *Eccritotarsus caterinensis* (Hemiptera: Miridae). *Biological Control* 58: 330-336.
- Twerdochlib A. L., Dalla Bona A. C., Leite S. S., Chitolina R. F., Westphal B., Navarro-Silva M. A. 2012. Genetic variability of a population of *Aedes aegypti* from Paraná, Brazil, using the mitochondrial ND4 gene. *Revista Brasileira de Entomologia* 56: 249-256.
- Volo de S. B., Reynolds R. T., Sonsthagen S. A., Talbot S. L., Antolin M. F. 2013. Phylogeography, postglacial gene flow, and population history of North American northern goshawks (*Accipiter gentilis*). *The Auk* 130: 342-354.
- Wang Y., Huang X. L., Qiao G. X. 2013 Comparative analysis of mitochondrial genomes of five aphid species (Hemiptera: Aphididae) and Phylogenetic Implications. *PlosOne* 8(10): e77511.
- Waterhouse A. M., Procter J. B., Martin D. M. A., Clamp M., Barton G. J. 2009. JalView version 2: A Multiple Sequence Alignment and Analysis Workbench. *Bioinformatics* 25: 1189-1191.
- Weckworth B. V., Musiani M., McDevitt A. D., Hebblewhite M., Mariani S. 2012. Reconstruction of caribou evolutionary history in Western North America and its implications for conservation. *Molecular Ecology* 21: 3610-3624.
- Xu J., Fonseca D. M., Hamilton G. C., Hoelmer K. A., Nielsen A. L. 2014. Tracing the origin of US brown marmorated stink bugs, *Halyomorpha halys*. *Biological Invasions* 16: 153-166.

- Zhang M., Rao D., Yang J., Yu G., Wilkinson J. A. 2010. Molecular phylogeography and population structure of a mid-elevation montane frog *Leptobrachium ailaonicum* in a fragmented habitat of southwest China. *Molecular Phylogenetics and Evolution* 54: 47-58.
- Zhang B., Edwards O. R., Kang L., Fuller S. J. 2012. Russian wheat aphids (*Diuraphis noxia*) in China: native range expansion or recent introduction? *Molecular Ecology* 29: 2130-2144.
- Zhao Y. E., Wu L. P. 2012. Phylogenetic relationships in *Demodex mites* (Acari: Demodicidae) based on mitochondrial 16S rDNA partial sequences. *Parasitology Research* 13: 1-9.

Streszczenie

**OCENA TRZECH OGÓLNODOSTĘPNYCH PROGRAMÓW KOMPUTEROWYCH
UMOŻLIWIAJĄCYCH ANALIZĘ MODELI POLIMORFIZMU DNA NA PRZYKŁADZIE BADAŃ
ZRÓŻNICOWANIA NUKLEOTYDÓW PODJEDNOSTEK 16S rDNA I OKSYDAZY
CYTOCHROMOWEJ I MITOCHONDRIALNEGO DNA W PODRZĘDZIE HETEROPTERA (INSECTA:
HEMIPTERA)**

Możliwości trzech ogólnodostępnych programów komputerowych do analiz zróżnicowania genetycznego (DnaSP, ProSeq i MEGA) zostały zestawione w celu stwierdzenia, czy wyniki generowane przez te programy są porównywalne. Analizy podjednostki 16S rDNA (105 sekwencji) oraz COI (460 sekwencji) 24 gatunków reprezentujących podrząd Heteroptera rzędu Hemiptera, uwzględniały następujące parametry: liczbę i pozycje miejsc polimorficznych, średnią liczbę różnic nukleotydowych, liczbę haplotypów i ich zmienność, standardowe odchylenie w zróżnicowaniu haplotypów, oraz testy neutralności. Rezultaty tych analiz sugerują, że tylko program DnaSP jest najbardziej odpowiednim programem do tego typu analiz.

Table 1. A list of true-bug species with data on the mtDNA subunits, numbers and length of sequences studied.

No.	SPECIES	SUBUNIT	NUMBER OF SEQUENCES	LENGTH OF SEQUENCES	SYSTEMATICS
1.	<i>Appasus japonicas</i> (Vuillefroy 1864)	16S rDNA	31	435 bp	Heteroptera – Belostomatidae
2.	<i>Appasus major</i> (Esaki 1934)	16S rDNA	27	436 bp	Heteroptera – Belostomatidae
3.	<i>Austrovelia caledonica</i> Malipatil & Monteith 1983	16S rDNA	10	426 bp	Heteroptera – Mesovellidae
4.	<i>Nepa hoffmanni</i> Esaki 1925	16S rDNA	18	509 bp	Heteroptera – Nepidae
5.	<i>Triatoma pallidipennis</i> (Stål 1872)	16S rDNA	19	709 bp	Heteroptera – Reduviidae
6.	<i>Agriosphodrus dohrni</i> (Signoret 1862)	COI	75	1218 bp	Heteroptera – Reduviidae
7.	<i>Appasus japonicas</i> (Vuillefroy 1864)	COI	31	658 bp	Heteroptera – Belostomatidae
8.	<i>Appasus major</i> (Esaki,1934)	COI	27	658 bp	Heteroptera – Belostomatidae
9.	<i>Aquarius najas</i> (deGeer 1773)	COI	61	777 bp	Heteroptera – Gerridae
10.	<i>Carpocoris fuscispinus</i> (Boheman 1851)	COI	10	653 bp	Heteroptera – Pentatomidae
11.	<i>Carpocoris mediterraneus</i> Tamanini 1958	COI	14	601 bp	Heteroptera – Pentatomidae
12.	<i>Carpocoris pudicus</i> (Poda 1761)	COI	12	589 bp	Heteroptera – Pentatomidae
13.	<i>Carpocoris purpureipennis</i> (deGeer 1773)	COI	25	602 bp	Heteroptera – Pentatomidae
14.	<i>Cimex hemipterus</i> (Fabricius 1903)	COI	12	533 bp	Heteroptera – Cimicidae
15.	<i>Corythucha ciliate</i> (Say 1832)	COI	10	320 bp	Heteroptera – Tingidae
16.	<i>Gerris costae</i> (Linnaeus 1758)	COI	13	749 bp	Heteroptera – Gerridae
17.	<i>Gerris gibbifer</i> Schummel 1832	COI	14	817 bp	Heteroptera – Gerridae
18.	<i>Gerris lacustris</i> (Linnaeus 1758)	COI	29	810 bp	Heteroptera – Gerridae
19.	<i>Gerris thoracicus</i> Schummel 1832	COI	19	723 bp	Heteroptera – Gerridae
20.	<i>Lygocoris pabulinus</i> (Linnaeus 1761)	COI	20	633 bp	Heteroptera – Miridae
21.	<i>Orius sauteri</i> (Poppius 1909)	COI	10	600 bp	Heteroptera – Anthocoridae
22.	<i>Psallivius piceicola</i> (Knight 1923)	COI	12	629 bp	Heteroptera – Miridae
23.	<i>Pseudacysta perseae</i> (Heidemann 1908)	COI	34	515 bp	Heteroptera – Tingidae
24.	<i>Sigara potamius</i> Young 1962	COI	32	738 bp	Heteroptera – Corixidae

Table 2. Sequences of 16S rDNA for *Apassus japonicus*, with GenBank accession numbers and single nucleotide polymorphism data.

No.	GenBank accession numbers	SNP position			SNP allele		
		DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
1	512380271	135	135	135	A/G	A/G	A/G
2	512380270	145	145	145	G/A	G/A	G/A
3	512380279 512380278 512380274 512380273	227	227	227	A/G	A/G	A/G
4	512380270	233	233	233	A/G	A/G	A/G
5	512380275 512380262 512380263 512380276 512380277 512380288 512380271 512380286 512380285 512380284 512380272 512380269 512380270 512380279 512380278 512380274 512380273 512380287	237	237	237	C/T	C/T	C/T
6	512380276 512380288 512380271 512380286 512380285 512380284 512380272 512380269 512380270 512380279 512380278 512380274 512380273	238	238	238	C/T	C/T	C/T
7	512380263 512380279 512380278 512380274 512380273	239	239	239	G/A	G/A	G/A
8	512380288 512380271 512380286 512380285 512380284 512380272 512380269 512380270 512380279 512380278 512380274 512380273 512380287	264	264	264	T/C	T/C	T/C
9	512380265 512380260 512380264	275	275	275	A/G	A/G	A/G
10	512380281 512380277 512380288 512380271 512380286 512380285 512380284 512380272 512380269 512380270 512380279 512380278 512380274 512380273 512380287	320	320	320	C/T	C/T	C/T
11	512380288 512380271 512380286 512380285 512380284 512380272 512380269 512380270 512380279 512380278 512380274 512380273 512380287	322	322	322	A/G	A/G	A/G
12	512380260	400	400	400	C/T	C/T	C/T
13	512380265	422	422	422	A/T	A/T	A/T

Table 3. Polymorphism analyses data of 16S rDNA for *Apassus japonicus*.

POLYMORPHIC SITES	
number of variable sites „S”	13
total number of mutations „ETA”	13
selected region	1-435
total number of sites	435
number of sequences	31

Table 4. Sequences of COI mtDNA for *Apassus japonicus*, with GenBank accession numbers and single nucleotide polymorphism data.

No.	GenBank accession numbers	SNP position			SNP allele		
		DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
1	512380180 512380170 512380166	22	22	22	G/A	G/A	G/A
2	512380172	38	38	38	G/A	G/A	G/A
3	512380200 512380198 512380196 512380194 512380192 512380176 512380168 512380166 512380164 512380162	40	40	40	C/T	C/T	C/T
4	512380180 512380170	52	52	52	A/G	A/G	A/G
5	512380200 512380198 512380196 512380194 512380192 512380186 512380182 512380180 512380178 512380176 512380174 512380172 512380170 512380168 512380166 512380164 512380162 512380150 512380148	58	58	58	G/A	G/A	G/A
6	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	67	67	67	C/T	C/T	C/T
7	512380200 512380198 512380196 512380194 512380192 512380168 512380166 512380164 512380162	79	79	79	T/C	T/C	T/C
8	512380200 512380198 512380196 512380194 512380192 512380168 512380166 512380164 512380162 512380150	88	88	88	G/A	G/A	G/A
9	512380200 512380198 512380196 512380194 512380192 512380182 512380176 512380168 512380166 512380164 512380162 512380148	103	103	103	C/T	C/T	C/T

No.	GenBank accession numbers		SNP position			SNP allele		
			DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
10	512380182 512380148	512380172	112	112	112	C/T	C/T	C/T
11	512380166		133	133	133	T/C	T/C	T/C
12	512380200 512380196 512380192 512380180 512380170 512380166 512380162	512380198 512380194 512380182 512380172 512380168 512380164	163	163	163	A/T	A/T	A/T
13	512380178 512380150	512380176	166	166	166	A/G	A/G	A/G
14	512380152		169	169	169	A/T	A/T	A/T
15	512380200 512380196 512380192 512380180 512380170 512380166 512380162 512380156 512380144	512380198 512380194 512380182 512380172 512380168 512380164	181	181	181	A/G	A/G	A/G
16	512380200 512380196 512380192 512380168 512380164	512380198 512380194 512380174 512380166 512380162	184	184	184	G/A	G/A	G/A
17	512380150		196	196	196	A/G	A/G	A/G
18	512380152		205	205	205	A/G	A/G	A/G
19	512380148		214	214	214	T/C	T/C	T/C
20	512380200 512380196 512380192 512380180 512380170 512380166 512380162	512380198 512380194 512380182 512380172 512380168 512380164	217	217	217	C/G ; C/A	C/G ; C/A	C/G ; C/A
21	512380200 512380196 512380192 512380180 512380170 512380166 512380162	512380198 512380194 512380182 512380172 512380168 512380164	238	238	238	G/A	G/A	G/A
22	512380200 512380196 512380166	512380198 512380168 512380162	262	262	262	G/A	G/A	G/A
23	512380200 512380196 512380192 512380180 512380172 512380168 512380164	512380198 512380194 512380182 512380176 512380170 512380166 512380162	268	268	268	G/A	G/A	G/A
24	512380198	512380148	271	271	271	T/C	T/C	T/C
25	512380182	512380172	274	274	274	A/G	A/G	A/G
26	512380200 512380196 512380192 512380166	512380198 512380194 512380168 512380162	280	280	280	A/G	A/G	A/G
27	512380158		290	290	290	T/C	T/C	T/C
28	512380194	512380168	310	310	310	A/G	A/G	A/G

No.	GenBank accession numbers		SNP position			SNP allele		
			DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
29	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	316	316	316	A/G	A/G	A/G	
30	512380178 512380176	319	319	319	A/G	A/G	A/G	
31	512380160	325	325	325	T/C	T/C	T/C	
32	512380164	331	331	331	A/G	A/G	A/G	
33	512380150	340	340	340	C/T	C/T	C/T	
34	512380200 512380198 512380196 512380194 512380192 512380180 512380172 512380170 512380168 512380166 512380164 512380162 512380148	343	343	343	A/G	A/G	A/G	
35	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	346	346	346	T/C	T/C	T/C	
36	512380190 512380188 512380184	373	373	373	A/G	A/G	A/G	
37	512380182 512380180 512380172 512380170	386	386	386	C/T	C/T	C/T	
38	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	388	388	388	G/A	G/A	G/A	
39	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	391	391	391	C/T	C/T	C/T	
40	512380200 512380198 512380196 512380194 512380192 512380168 512380166 512380164 512380162	400	400	400	A/G	A/G	A/G	
41	512380178 512380176 512380150	406	406	406	C/T	C/T	C/T	
42	512380194 512380182 512380180 512380172 512380170	409	409	409	G/A	G/A	G/A	
43	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	418	418	418	T/C	T/C	T/C	
44	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	430	430	430	A/G	A/G	A/G	

No.	GenBank accession numbers		SNP position			SNP allele		
			DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
45	512380200	512380198 512380196 512380194 512380192 512380190 512380188 512380186 512380184 512380182 512380180 512380178 512380176 512380174 512380172 512380170 512380168 512380166 512380164 512380162 512380160 512380158 512380156 512380154 512380152 512380150 512380148 512380146 512380144	466	466	466	G/A	G/A	G/A
46	512380200	512380198 512380196 512380168 512380166 512380162	481	481	481	A/G	A/G	A/G
47	512380200	512380198 512380196 512380168 512380166 512380164 512380162	484	484	484	C/T	C/T	C/T
48	512380174		487	487	487	C/T	C/T	C/T
49	512380156		490	490	490	C/T	C/T	C/T
50	512380200	512380198 512380196 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	496	496	496	A/G	A/G	A/G
51	512380182		499	499	499	G/A	G/A	G/A
52	512380178		500	500	500	C/T	C/T	C/T
53	512380182	512380180 512380172 512380170 512380150	502	502	502	G/A	G/A	G/A
54	512380200	512380198 512380196 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	508	508	508	G/A	G/A	G/A
55	512380168		517	517	517	A/G	A/G	A/G
56	512380146		525	525	525	C/A	C/A	C/A
57	512380192	512380164	529	529	529	C/T	C/T	C/T
58	512380174	512380164	544	544	544	A/G	A/G	A/G
59	512380200	512380198 512380196 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162 512380154 512380152 512380144	553	553	553	T/C	T/C	T/C
60	512380194	512380162	556	556	556	G/A	G/A	G/A
61	512380200	512380198 512380196 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	557	557	557	T/C	T/C	T/C

No.	GenBank accession numbers	SNP position			SNP allele		
		DnaSP	MEGA	ProSeq	DnaSP	MEGA	ProSeq
62	512380200 512380182 512380172	577	577	577	A/G	A/G	A/G
63	512380182 512380180 512380172 512380170	578	578	578	T/C	T/C	T/C
64	512380174	589	589	589	C/T	C/T	C/T
65	512380198	601	601	601	C/T	C/T	C/T
66	512380166	519	519	519	T/G	T/G	T/G
67	512380182 512380180 512380172 512380170	622	622	622	C/T	C/T	C/T
68	512380200 512380198 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	628	628	628	G/A	G/A	G/A
69	512380178 512380176 512380174 512380158	631	631	631	T/A ; T/C	T/A ; T/C	T/A ; T/C
70	512380200 512380198 512380196 512380194 512380192 512380182 512380180 512380172 512380170 512380168 512380166 512380164 512380162	634	634	634	T/C	T/C	T/C
71	512380186	649	649	649	G/A	G/A	G/A

Table 5. Polymorphism analyses data of COI mtDNA for *Apassus japonicus*.

POLYMORPHIC SITES	
number of variable sites „S”	71
total number of mutations „ETA”	74
selected region	1-658
total number of sites	658
number of sequences	31