eJBio

Electronic Journal of Biology, 2010, Vol. 6(1): 13-18

Molecular Cytogenetics of some *Anopheles* Mosquitoes (Culicidae: Diptera)

Monika Sharma*, S. Chaudhry

Mosquito Cytogenetics Unit, Department of Zoology, Panjab University,Chandigarh 160 014, India.

***** Corresponding author. Tel: 091 9876359624; E-mail: monikacyto@yahoo.com

Abstract

Experimental work pertaining to the molecular cytogenetics of malaria vector species of mosquitoes by the application of PCR technique has been carried out. The main objectives of the study included the sequence characterization of nuclear rDNA internal transcribed spacers 1 and 2 (ITS1, ITS2) and mitochondrial DNA COII gene as potential molecular markers for studying genetic relatedness and phylogenetic kinship among six important species of genus *Anopheles.* The present studies involved the extraction of genomic DNA from a single female mosquito followed by its amplification with specific primers. The total length of each DNA band with respect to the number of nucleotides was calculated along with GC:AT content, ratio of substitutions due to transitions and transversions (ts/tv), insertions/ deletions and identification of tandem and nontandem repeat sequences. Out of the three studied molecular markers, ITS1 and ITS2 were GC rich while COII gene was AT rich. As for the incidence of insertions/deletions (indels) of bases is concerned, it was found maximum in ITS1 and ITS2 and minimum in conserved COII gene sequence.

From the present results it was evident that except for *An. culicifacies* the ITS2 sequence of the remaining five species is under rapid evolutionary changes due to less conserved nature of the sequence. To the contrary ITS1 sequence was found to have highly variable length ranging from 300-900bp. It was found that *An. stephensi* + *An. culicifacies* and *An. annularis* + *An. splendidus* share a close genetic homology while *An. subpictus* and *An. maculatus* have hypervariable non-homologous genomic qualities. The phylogenetic dendrograms revealed that *An. stephensi* and *An. culicifacies* grouped together only when analysed by MP in case of ITS1 gene only while they drifted apart in case of ITS2 and COII gene. This is attributed to their host and habitat feeding preferences which have given them the dubious status of urban and rural vectors respectively.

Keywords: Phylogeny; *Anopheles;* ITS1; ITS2; COII.

1. Introduction

In integrated mosquito control programmes, taxonomic and phylogenetic studies had been quite useful in understanding the genetics of vectorial capacity and insecticide resistance. In the recent years the developments in DNA based molecular cytogenetics have offered promising possibilities for extension and future applications of genetic engineering based vector control programmes. Some of the molecular level genomic studies involving *in vitro* amplification of DNA by using the technique of polymerase chain reaction (PCR) have revealed that eukaryotic organisms have considerable nuclear and mitochondrial DNA polymorphism which provides virtually unlimited opportunities for establishing the exact taxonomic status and phylogenetics of species [1-4].

One of the most widely used regions of the genome to infer genetic variations and phylogenetic relationships is the rDNA gene cluster of tandemly repeated multigene family. This multigene family is known to evolve cohesively within species through concerted evolution, a mechanism that tends to homologize sequences within species with simultaneous variations amongst them [5-8]. In between the gene sequences coding for 18S, 5.8S and 28S rRNA are the non-coding internal transcribed spacers 1 and 2 whose sequences are being used for detecting micro and macro geographic genomic variations between species.

In addition to the rDNA domain, the mitochondrial DNA is also being exploited for comparative genomics. In fact, analysis of the parts of mt. DNA and direct sequencing of its specific regions are currently the methods of choice for majority of the population level studies [9-11]. The small size of mitochondrial genome, its single copy number, lack of introns and maternal inheritance, are some of its features for which it is preferred for DNA diagnostics.

2. Materials and Methods

Motivated by the advances made in molecular cytogenetics of mosquitoes, the present topic of research on molecular cytogenetics of some *Anopheles* mosquitoes (Diptera: Culicidae) was undertaken to carry out the sequence based phylogenetic inferences of six epidemiologically important species of subgenus *Cellia* of genus *Anopheles viz*: *An*. *stephensi*, *An. culicifacies*, *An. maculatus*, *An. subpictus*, *An. annularis* and *An. splendidus*. Out of them, *An. stephensi, An. splendidus, An. maculatus* and *An. annularis* belong to Neocellia series, *An*. *subpictus* belongs to Pyretophorus series while *An. culicifacies* belongs to Myzomyia series.

The main objective of the study was the sequence characterization of ITS1, ITS2 and COII gene as potential molecular markers. *Cx. quinquefasciatus* was used as an outgroup, so as to validate the results. The field collection of these species was carried out from villages Beladhayani near the township of Nangal, Punjab (105 kms North-west of Chandigarh) and Nadasahib, Panchkula (Haryana), 20 kms South-east of Chandigarh (30º44"N, 76º53"E). All these species are vectors of malaria in different capacities, out of which *An. stephensi* and *An. culicifacies* are rated as chief urban and rural vectors respectively. The identification of each species was carried out from their morphotaxonomic characters and the species-specific banding pattern of the salivary polytene X-chromosome. Each gravid female was handled in a test tube where it was allowed to lay eggs on a strip of wet filter paper.

In order to provide the optimal conditions of rearing, the eggs were transferred to various bowls and kept in BOD incubator. The freshly emerged unfed adults were stored at -20^0C before using them for DNA extraction. The extraction of DNA from a single female mosquito at a time was carried out by phenol-chloroform extraction method. The PCR was programmed for 35 cycles of denaturation, annealing and extension by using specific forward and reverse primers meant for ITS1, 2 and COII *viz*: ITS1- FP- 5'-CCTTTGTACACACCGCCCGT-3', RP- 5'- GTTCATGTGTCCTGCAGTTCAC - 3'; ITS2- FP - 5' – TGTGAACTGCAGGACACAT - 3', RP - 5' – TATGCTTAAATTCAGGGGGT - 3'; COII – FP - 5' – TCTAATATGGCAGATTAGTGC - 3', RP: 5' – GATCATTACTTGCTTTCAG - 3'. The amplified products were then subjected to 2% agarose gel electrophoresis along with a standard DNA ladder (Gene ruler) of 80-1031bp. The DNA of each band generated from ITS1, ITS2 and COII were sequenced and the sequence alignment data was recorded for total bp composition, GC: AT content, ratio of substitutions (ts/tv) content, insertions/ deletions (indels), incidence of dimers, trimers, tetramers and polymers, and the presence of tandem and non-tandem repeats.

Phylogenetic relationships among the six species and the dendrograms of genetic relatedness were generated by the application of maximum parsimony (MP) method which is based on the assumption that mutation is rare and the best explanation of evolutionary history is the one that requires the least mutation and Kimura-2 parameter.

3. Results

3.1 Sequence analysis of ITS1

The ITS1 region lying between 18S and 5.8S rRNA coding sequences in the rDNA domain of nuclear DNA yielded G:C rich DNA fragments ranging in length from 300-900 bp as they consisted of 639, 512, 346, 778 and 816 bp in *An. stephensi*, *An. culicifacies*, *An. maculatus, An. subpictus* and *An. annularis* respectively (Figure1).

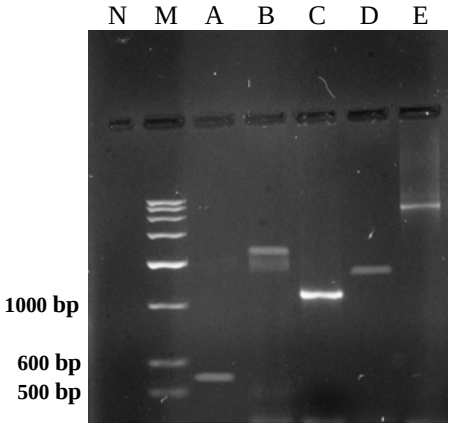

Figure 1. PCR amplified DNA bands from ITS1. Lane N: Negative control, lane M: Gene ruler, lane A: *Anopheles maculatus*, lane B: A*nopheles stephensi*, lane C: *Anopheles culicifacies*, lane D: *Anopheles subpictus*, lane E*: Anopheles annularis.*

There were a total of 393 substitutions, out of which 174 were transitions while 199 were transversions. The sequence alignment analysis revealed that ts/tv ratio ranged from 0.55 to 1.2 with 0.55 in *An. maculatus*, 0.83 in *An. culicifacies*, 0.91 in *An. annularis*, 0.93 in *An. subpictus* and 1.2 in *An. stephensi*. Indels were observed at several places along the sequence in which minimum incidence of loss or gain of base was found in *An. annularis* (16) and maximum in *An. maculatus* (51) out of the total of 191 places where insertions or deletions of bases had taken place. So far, the sequence analysis of ITS1 has been carried out only in a few species of

genus *Anopheles* where considerable variation has been found in its length. For example, the ITS1 of *An. aconitus* has an average length of 503 bp while *An. farauti* has as many as 979 bp. Outside the genus *Anopheles*, in *Aedes aegypti* its length was found to be 418 bp while in the related dipteran *Drosophila arizonae* this region varied from 500-600 bp. When all the species were compared, it was found to be longest in *An. farauti* and smallest in *An. maculatus*, a condition that demands the study of intron based evolutionary patterns in the genus.

3.2 ITS1 sequence based phylogenetics

With the application of Kimura-2 programme the present comparative study of nucleotide sequences revealed that maximum genetic divergence was between *An. subpictus* and *An. annularis* in which it gave a value of 1.384916 while it was minimum between *An. culicifacies* and *An. stephensi* with a value of 0.008244. The same was reflected in the phylogenetic trees generated by maximum parsimony (MP) method which showed that *An. annularis* and *An. subpictus* formed a single clade with a bootstrap value of 92.5 where *An. stephensi* also clustered with them with a bootstrap value of 99. Similarly, *An. culicifacies* with a bootstrap value of 100 also clustered with these species. When compared with the outgroup, *An. maculatus* closely paired with *Cx*. *quinquefasciatus* (Figure 2).

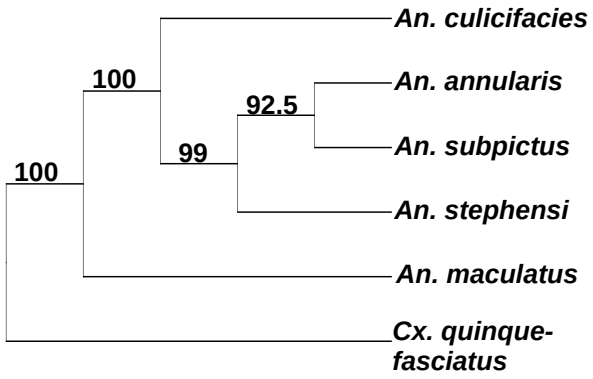

Figure 2. Phylogenetic tree based on rDNA ITS1 sequence data as generated by maximum parsimony (MP) method conducted in DNAPARS program (PHYLIP). The tree was rooted with *Cx. quinquefasciatus*. Numbers on branches are bootstrap values using 10,000 replicates.

3.3 Sequence analysis of ITS2

The PCR amplification of this spacer region produced a single DNA band from each species with a base pair length ranging from 458-506 bp. The sequence of *An. stephensi* had a minimum number of 458 bp followed by *An. maculatus*- 459 bp, *An. annularis*- 478 bp, *An. splendidus*- 488 bp, *An. subpictus*- 491, and *An. culicifacies*- 506 bp (Figure 3).

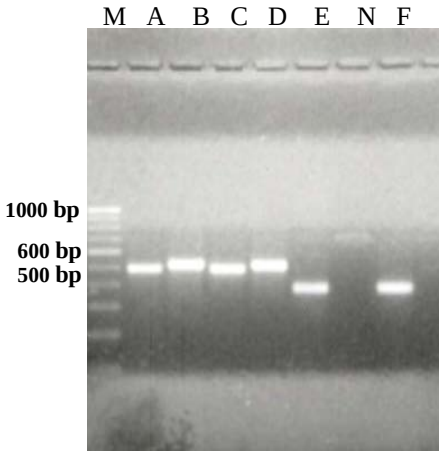

Figure 3. PCR amplified DNA bands from ITS2. Lane M: Gene ruler, lane A: *Anopheles annularis,* lane B: *Anopheles culicifacies*, lane C: *Anopheles subpictus*, lane D: *Anopheles splendidus,* lane E: *Anopheles stephensi*, lane N: Negative control, lane F: *Anopheles maculatus*.

In majority of the anopheline species studied so far, the size of ITS2 region varies in length from 300-500 bp. For example, in the subgenus *Cellia*, *An. gambiae* has an ITS2 comprising of 426 bp, *An. dirus* complex has 710-716 bp while *An. punctulatus* has 549-563 bp. Within the subgenus *Anopheles* of the genus *Anopheles*, *An. maculipennis* has an average length of 305 bp, *An. lesteri* 438 bp, *An. sinensis* 459 bp, *An. messeae* 489 bp while *An. quadrimaculatus* complex has 305-310 bp. Similarly, within the subgenus *Nyssorhynchus*, *An. nuneztovari* has a sequence ranging from 363-369 bp while *An. albitarsis* has 490 bp.

As a result of these sequence comparisons, the average G:C percentage of this spacer region in the present species was higher than A:T as it was 55.34% in *An. maculatus*, 51.26% in *An. annularis*, 52.75% in *An. subpictus*, 55.90% in *An. stephensi*, 54.10% in *An. splendidus* and 60.28% in *An. culicifacies*. When compared with other anopheline species worked out so far, the G:C content was found to vary from 45.3% to 55.7% which falls well within the range of species under study. This increase in G:C has been attributed to evolutionary genetic drift of adaptive significance. With respect to substitutions, *An. culicifacies* had maximum number of 87 single nucleotide changes out of which there were 72 transversions and 15 transitions. The most frequent transversions were between adenine and thymine and, as compared to the incidence of substitution, the rate of transversions was found to be 66.78% which was more than the transitions

which accounted for only 29.68%. As for the indels, there were a total of 274 loci with either the deletion or insertion of nucleotides. The maximum number of 59 indels was in *An. annularis* as compared to 54 in *An. stephensi* and *An. maculatus*, 50 in *An. subpictus* and *An. splendidus* and only 7 in *An. culicifacies.* With this, the transition to transversion ratio varied from 0.21 in *An. culicifacies* to 1 in *An. splendidus*. These values were considered significant, as substitutions are the main sources of evolution and speciation [12-14]. From the present results it is evident that, except for *An. culicifacies,* the ITS2 sequence of the remaining species is under rapid evolutionary changes due to its less conserved nature.

3.4 ITS2 sequence based phylogenetics

As per Kimura-2 parameter, there was maximum similarity between *An. maculatus* and *An. stephensi* which shared a value of 0.027620 while *An. splendidus* was found to be closest to *An. subpictus* with a value of 0.1127. The tree generated by MP method clearly indicated the close genetic homology and phylogenetic relationship of *An. stephensi* with *An. maculatus* due to their inclusion in a single clade with bootstrap values of 99 and 100 respectively (Figure 4).

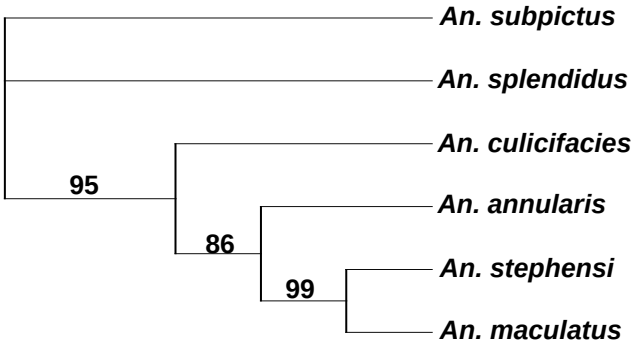

Figure 4. rDNA ITS2 sequence based phylogenetic tree for six species of genus *Anopheles* using maximum parsimony (MP) method. Numbers on branches are bootstrap values using 10,000 replicates.

For the remaining species, three different clades were generated in which one consisted of *An. stephensi* while the other had *An. culicifacies* included in it. In the same way, another clade included *An. subpictus* and *An. splendidus*. *An. subpictus* of Pyretophorus series showed a close relationship with *An. splendidus* of the series Neocellia. Since *An. culicifacies* is a member of Myzomyia series, it formed a separate clade with a bootstrap support value of 56%. In the overall assessment, it was found that maximum closeness was present between *An. stephensi* and *An. maculatus* while maximum divergence was between *An. stephensi* and *An. subpictus*.

3.5 Sequence analysis of COII gene

The amplified part of COII gene generated a band of 708-718 bp. The total sequence has ATG for initiation at 5' end and only T at the 3' end that codes for the entire set of termination sequence with 708 bp in *An. maculatus*, 710 bp in *An. culicifacies*, 711 bp each in *An. stephensi* and *An. annularis*, 713 in *An. splendidus* and 718 in *An. subpictus* (Figure 5).

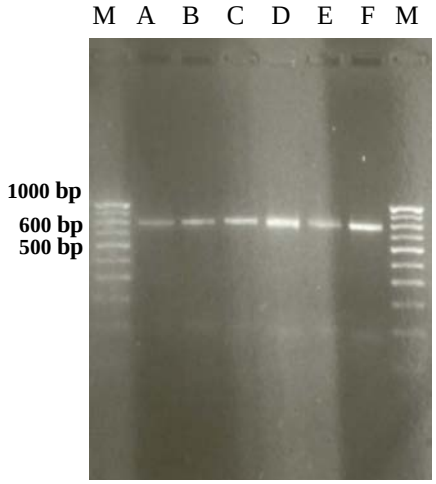

Figure 5. PCR amplified DNA bands from COII gene. Lane M: Gene ruler, lane A: *Anopheles splendidus*, lane B: A*nopheles annularis*, lane C: *Anopheles subpictus*, lane D: *Anopheles culicifacies*, lane E*: Anopheles stephensi*, lane F: *Anopheles maculatus.*

Instead of G:C rich sequences of ITS1 and ITS2, the sequence of this gene was found to be A:T rich with maximum percentage of 75.07% in *An. culicifacies* followed by 73.87% in *An. maculatus*, 73.70% in *An. stephensi*, 72.70% in *An. subpictus*, 72.51% in *An. splendidus* and 71.45% in *An*. *annularis*. The A:T% of the sequences shows that out of all the species, *An. culicifacies* was closest to *An. maculatus* with a percentage values of 75.07% and 73.87% respectively. The sequence with 718bp was the longest in *An. subpictus* followed by 713bp in *An. splendidus*, 711bp in *An. stephensi* and *An. annularis*, 710 in *An. culicifacies* and smallest with 708bp in *An. maculatus*. The incidence of 35 transitions was maximum in *An. maculatus* followed by 22 in *An. splendidus,* 21 in *An. annularis*, 19 in *An. stephensi* and 17 each in *An. subpictus* and *An. culicifacies.* Similarly, transversions were also maximum with 44 in *An. maculatus* followed by 36 in *An. splendidus*, 26 in *An. culicifacies*, 25 in *An. annularis*, 24 in *An. stephensi*, and a minimum of 21 in *An. subpictus*. With this, the transition/ transversion ratio was found to range from 0.6 to 0.8.

In addition to these variations, indels were also observed at several places along the sequence. Out of a total of 66 loci where insertions or deletions had taken place, the minimum incidence of loss or gain of bases was found in *An. subpictus* while maximum of 14 in *An. maculatus*.

3.6 COII gene sequence based phylogenetics

The phylogenetic trees showed the closeness of *An. annularis* and *An. splendidus* with bootstrap values of 100. According to the tree based on maximum parsimony (MP), *An. stephensi* formed a clade with *An. annularis* and *An. splendidus* with bootstrap value of 83 while *An. culicifacies* and *An. subpictus* formed another clade with a bootstrap value of 42 (Figure 6).

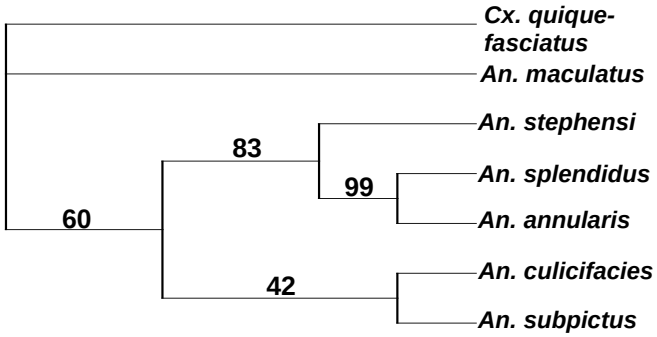

Figure 6. Phylogenetic tree based on COII gene sequence data as generated by maximum-parsimony (MP) method conducted in DNAPARS program (PHYLIP). The tree was rooted with *Cx. quinquefasciatus*. Numbers on branches are bootstrap values using 10,000 replicates.

ITS1 and ITS2 are the noncoding sequences in the rDNA domain whereas COII gene sequence is a functional subunit of the mt.DNA. Therefore, on the basis of the studies carried out so far on the coding characteristics of this sequence in mosquitoes, it has been found that the whole sequence codes for 228-229 amino acids. For example, based on the coding characteristics *An. stephensi* and *An. culicifacies* once again made a distinct group. This is attributed to their anthropophilic and zoophilic feeding preferences which have given them the dubious status of urban and rural vectors respectively.

4. Discussions and Conclusions

In order to study the phylogeny of the six most important species of the genus *Anopheles* three molecular markers were used. Out of which ITS1 and ITS2 are highly variable regions of nuclear rDNA spacer sequences, while COII gene sequence is a highly conserved sequence of mitochondrial DNA. In addition to the ITS1, ITS2 and COII based assessment of genomic novelties in the present six species, the application of SpectraL Repeat Finder (SRF) programme also revealed valuable data on different types of repeats in the sequences. For example, polymers TGACCGA and CCTCGGC were remarkably similar in their copy number in *An. stephensi* and *An. culicifacies*. Documentation of such repeats may also be helpful in the selection of matching restriction enzymes in RFLP-PCR based studies of target species.

The genetic information about them is mainly based on polytene chromosome characteristics. With the advent of molecular parameters of study, a number of molecular markers are currently being used to characterize species with doubtful taxonomic status and phylogenetic relatedness. These molecular markers include noncoding sequences like ITS1 and 2 (noncoding) and coding sequences such as cytochrome b (cyt. b), cytochrome oxidase I and II (CO I, II), NADH dehydrogenase I, II, III, IV, V and VI. In order to augment the existing data, it would be appropriate to include more number of coding and noncoding sequences in the nuclear DNA and all the 37 mitochondrial genes.

The sequencing survey and phylogenetic analysis indicate greater diversity in the ITS1 sequence among the species covered in the present research programme. The efforts at exploiting the sequence variations of ITS2 region for species discrimination and phylogenetic analysis have proved quite useful in those cases where earlier parameters of study such as banding pattern of the polytene chromosomes and isozyme/allozyme variations were inconclusive[15,16].

The true function of these spacers remains vague, seemingly based on hydrogen-bonded secondary structure of RNA which, when modified slightly in conserved regions or modified considerably in variable regions, hinder maturation of the mRNA products [17]. Although, not as widely used as ITS2 in mosquitoes, ITS1 has similar properties to ITS2 and has been used at the population level in some insect groups [18]. In the Indian subcontinent genus *Anopheles* is represented by about 59 species out of which nearly 15 are confirmed vectors of malaria. The combined data sets of protein coding versus ribosomal genes will give a better understating of the different levels of systematic hierarchy and phylogenetic relationships.

Acknowledgements

The authors are thankful to Chairperson, Department of Zoology, Panjab University, Chandigarh for providing the necessary facilities under Centre of

Advance Studies Programme of U.G.C, New Delhi, India to carry out the present research work.

References

[1] Marinucci M., Romi R., Mancini P., et al. (1999) Phylogenetic relationships of seven palearctic members of the *maculipennis* complex inferred from ITS2 sequence analysis. *Insect Mol. Biol*, **8**(4): 469-480.

[2] Proft J., Maier W.A., Kampen H. (1999) Identification of six sibling species of the *Anopheles maculipennis* complex (Diptera: Culicidae) by a polymerase chain reaction assay. *Parasitol Res*, **85**(10): 837-843**.**

[3] Wilkerson R.C., Parsons T.J., Klein T.A., et al. (1995) Diagnosis by random amplified polymorphic DNA polymerase chain reaction of four cryptic species related to *Anopheles* (*Nyssorhynchus*) *albitarsis* (Diptera: Culicidae) from Paraguay, Argentina, and Brazil. *J Med Entomol,* **32**(5): 697-704.

[4] Chaudhry S., Neetu., Dhanda R.S., et al. (2004) Random amplified polymorphic DNA-polymerase chain reaction based differentiation of some species of the genus *Anopheles* (Culicidae: Diptera). *J Cytol Genet,* **5**(NS): 173-183.

[5] Morgen J.A.T., Blair D. (1998) Relative merits of nuclear ribosomal internal transcribed spacers and mitochondrial COI and ND1 genes for distinguishing among *Echinostoma* species, *Parasitology*, **116**: 289-297.

[6] Navajas M., Lagnel J., Gutierrez J., et al. (1998) Species-wise homogeneity of nuclear ribosomal ITS2 sequences in the spider mite *Tetranchysis urticae* contrasts with extensive mitochondrial COI polymorphism. *Heredity*, **80**: 742-752.

[7] Marrelli M.T., Honorio N.A., Flores-Mendoza C., et al. (1999) Comparative susceptibility of two members of the *Anopheles oswaldoi* complex, *An. oswaldoi* and *An. konderi*, to infection by *Plasmodium vivax. Trans R Soc Med Hyg,* **93**(4): 381-384.

[8] Hackett L.W., Missiroli A. (1935) The varieties of *Anopheles maculipennis* and their relation to the distribution of malaria in Europe. *Riv Malariol*, **14**: 45-109.

[9] Hwang U-W., Kim W. (1999) General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *Kor J Parasitol*, **37**(4): 215-228.

[10] Chaudhry S., Sharma M., Gupta S., et al. (2006) Multiple technique based species discrimination in the taxon *Anopheles* (*Cellia*) *stephensi* (Culicidae: Diptera). In: *Vector Biology* (eds. Sharma, V.P. and Kirti, J.S.). The National Academy of Sciences, India, pp. 105-123.

[11] Chaudhry S., Kohli, R. (2007) Sequence analysis of mitochondrial 16S ribosomal RNA gene fragment in two populations of *Culex quinquefasciatus* Say (Culicidae: Diptera). *Natl Acad Sci Lett,* **30**(1&2): 55-60.

[12] Hillis D.M., Dixon M.T. (1991) Ribosomal DNA-molecular evolution and phylogenetic inference. *Quarterly Rev Biol,* **66**: 410-453.

[13] Manonmani A.M., Townson H., Adeniran T., et al. (2001) rDNA-ITS2 polymerase chain reaction assay for the sibling species of *Anopheles fluviatilis. Acta Trop,* **78**(1): 3-9.

[14] Porter C. H., Collins F. H. (1991) Species diagnostic differences in rDNA internal transcribed spacer from the sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). *Am J Trop Med Hyg,* **45**(2): 271-279.

[15] Severini C., Silverstrini F., Mancini P., et al. (1996) Sequence and secondary structure of the rDNA second internal transcribed spacer in the sibling species *Culex pipiens* L. and *Cx. quiquefasciatus* Say (Diptera: Culicidae). *Insect Mol Biol*, **5**(3): 181-186.

[16] Marrelli M.T., Floeter-Winter L.M., Malafronte R.S., et al. (2005) Analysis of the second internal transcribed spacers (ITS2) of ribosomal DNA sequences of the anopheline species from Amazon region, South America. *Med Vet Entomol,* **19**: 208-217.

[17] van der Sande C.A.F.M., Kwa M., Van Nues R.W., et al. (1992) Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. *J Mol Biol*, **223**: 899-910.

[18] McLain D.K., Wesson D.M., Oliver J.H. Jr., et al. (1995) Evolution of the rDNA spacer, ITS2, in the ticks *Ixodes scapularis* and *I. pacificus* (Acari: Ixodi dae). *Heredity*, **75**: 303-319.