

*Research Paper*

# **Molecular Phylogeny based on nuclear ribosomal DNA Internal Transcribed Spacer 2 (rDNA ITS2) gene in the *Anopheles* species (Culcidae: Diptera)**

**Sharma Monika<sup>1</sup>, Kohli Rashmi<sup>2</sup> and Chaudhry Sudarshan<sup>3</sup>**

<sup>1</sup>Department of Biology, Ideal 21<sup>st</sup> Century Competitions Ltd., Mumbai, INDIA

<sup>2</sup>Department of Zoology, Post Graduate Government College for Girls, Sector-42 Chandigarh, INDIA

<sup>3</sup>Retd. Professor and Chairman, Department of Zoology, Panjab University, Chandigarh, INDIA

(Received 16 May, 2013, Accepted 3 July, 2013)

## **Abstract**

Phylogenetic studies mostly employ non-coding regions of large proportion of nuclear DNA markers because they are usually more variable than coding regions. DNA sequence analysis of 2<sup>nd</sup> Internal Transcribed Spacer (ITS) region was carried out among six *Anopheles* species within the subgenus *Cellia*. The results indicated high variance of spacer mutants in the genus *Anopheles*. The phylogenetic trees were generated by all the three methods i.e. MP, ML and NJ and showed that both *stephensi* and *maculatus* form a single clade. The ITS2 sequences varied between 453-506 base pairs and were slightly G-C rich. As per the incidence of substitution, the rate of transversions 66.78% was found to be more than transitions 29.68%. As a result of the present data it had been possible to examine the usefulness of the ITS 2 in the nuclear ribosomal gene array for conducting a phylogenetic analysis of selected taxa within the genus *Anopheles*. The aim of the present study was to assess the suitability of ITS2 gene for species level analysis and the current data provides useful reference for further work.

**Keywords:** *Anopheles*, Dendrogram, Phylogeny, Sequencing.

## **Introduction**

Phylogenetic studies have been remarkably productive and successful in the last decade following the invention of PCR technology and the introduction of various markers including mitochondrial DNA and nuclear DNA. Nuclear DNA analysis has been widely applicable due to the recent progress in genomic studies and molecular biology technologies. Eukaryotic organisms show wide nuclear DNA polymorphisms which provide virtually unlimited opportunities for studying the mechanisms of evolution.

One of the most widely used regions of the genome to infer genetic variation and phylogenetic relationships is the rDNA cluster, a tandemly repeated multigene family. In eukaryotic organisms, every repeated unit of rDNA consists of an intergenic spacer (IGS) followed by genes coding the 18S, 5.8S and 28S rDNA. Preceding the 18S is the ETS (External Transcribed Spacer) and surrounding the 5.8S rDNA are the ITS1 and ITS2. This multigene family evolves cohesively within species through concerted

evolution, a mechanism that tends to homogenize sequences within species while driving differentiation between species<sup>[1]</sup>. The coding regions are highly conserved, even between distantly related species, while non-coding DNA usually rapidly drifts apart, even between closely related species. Thus using primers located in the conserved rDNA regions, variable regions can be amplified from a wide range of species in absence of prior sequence information. As such the rDNA cluster has become an increasingly popular tool in molecular entomology and in particular as a means to develop diagnostic test to differentiate cryptic anopheline species<sup>[2]</sup>. For example, diagnostic PCR assays based on segregating sequence variation in the ITS and IGS are now available to identify members of the *An. gambiae* Giles, *An. quadrimaculatus* Say, *An. punctulatus* Donite, *An. maculipennis* Meigen and *An. funestus* Giles complexes<sup>[3,4,5,6,7]</sup>. It has already been suggested by Karotam *et al.* (1995)<sup>[8]</sup> that introns and 5' and 3' flanking non coding regions are already under varying degrees of functional constraint. Introns do not appear to be just non functional junk sequences. For example, it is well known that introns contain signal sequences important for splicing and even for regulation of transcription<sup>[9,10]</sup>.

The development of conserved PCR primers has played a key role in promoting phylogenetic studies at the molecular level in the past decade<sup>[11, 12]</sup>. Enormous success has been achieved with animal mt. DNA primers<sup>[13, 14, 15]</sup>. Because of this, it has been a common desire to find similarly conserved and useful nuclear DNA primers. The present study is based on the interspecific sequence differences in the ITS2 ribosomal mosquito DNA that with the help of the PCR and species specific primers, can be utilized to produce DNA fragments of species-specific length which are easily recognized and distinguished on an ethidium bromide stained agarose gel.

The present study allowed the discrimination of six *Anopheles* species within the subgenus *Cellia* from areas nearby Chandigarh among which five are major malaria vectors. The introduction of mosquitoes as a major disease vectors due to differences between habitats, changes in predators and the possible alterations in the genetic make up cannot be underestimated.

The spacer sequences have been shown to be highly variable in length and sequence, even between closely related species, making them useful tools for species identification<sup>[2]</sup>. ITS2 sequences of *Anopheles* species are preferred as targets because of its 2 important features. First, they are relatively short, less than 1kilo base pair, making the amplification of intervening ITS2 using primers from highly conserved domains of flanking coding regions relatively simple. Second, the level of intraspecies variation in them is lower than the interspecific variation.

## Materials and Methods

### DNA collection

For the present purpose of research the species of genus *Anopheles* were collected from the cattle sheds in village Beladhayani near Nangal, Punjab (105 kms Northwest of Chandigarh) which were selected as field stations for the periodic collections of the different species. Six species of genus *Anopheles* subgenus *Cellia* were selected for the present study which include *An. maculatus* Theobald, *An. splendidus* Koidzumi, *An. stephensi* Liston, *An. annularis* Vander Wulp, *An. culicifacies* Giles and *An. subpictus* Grassi. These were collected during different seasons of the year as *An. annularis*, *An. splendidus*, *An. culicifacies* were found during months of Oct – Dec, *An. maculatus* during Jan-Feb, *An. subpictus* during Sept. All are confirmed vectors of malaria out of which *An. stephensi* and *An. culicifacies* are rated as chief urban and rural vectors respectively. Out of these *stephensi*, *splendidus*, *maculatus* and *annularis* belong to Neocellia series while *subpictus* belong to Pyretophorus series and *culicifacies* belong to Myzomyia series. Four of them are major vectors of malaria, *An. annularis* being a secondary vector while *splendidus* being a non- vector. Gravid females were collected and were brought to the laboratory where they were allowed to lay eggs. The eggs were reared upto adult stages by feeding them on diet made of dog biscuits and yeast powder in the ratio of 3:1<sup>[16, 17]</sup>. The larvae in enamel bowls were maintained at 28<sup>o</sup> C in BOD incubator. The adults were stored as dry specimens at 4<sup>o</sup>C before they were used for DNA extraction.

## Extraction of Genomic DNA

Genomic DNA was extracted from adult female mosquitoes by following the modified phenol chloroform extraction method of Ausubel *et al.* (1999)<sup>[18]</sup>. Accordingly, individual mosquitoes were macerated in 100µl of lysis buffer containing 10mM Tris Cl (pH-8.0), 1 mM EDTA, 25mM NaCl and 1% SDS. The proteins were digested using sodium acetate at pH-5.2 and in the end DNA was precipitated in ethanol. Final storage of extracted DNA was done in TE buffer and kept at -20°C.

## Primers

The two primers used in this study were annealed to conserved regions of 5.8S and 28S rDNA subunits that flank the ITS2 (5'-TGTGAAGTGCAGGACACAT-3' and 5'-TATGCTTAAATTCAGGGGGT-3')<sup>[19]</sup>.

## PCR Amplification and DNA sequencing

The rDNA ITS2 region was amplified by PCR. The 50µl reaction mixture contained 5µl of 10X PCR buffer, 0.2mM dNTP's, 1 U Taq polymerase, 0.2µM each primer, 1.5mM MgCl<sub>2</sub>. The PCR temperature profile consisted of 35 cycles of amplification wherein each cycle consisted of 1 min. at 94°C, 1 min. at 56°C and 1min. at 72°C. All the products of amplification were electrophoresed using 2% agarose gel and the final products were visually analyzed under UV transilluminator. The ITS2 of each species was sequenced from M/s Bangalore Genei Pvt. Lt., Bangalore, India while the consensus sequences were obtained by direct sequencing using the same primers which were used for PCR amplification.

## Sequence analysis

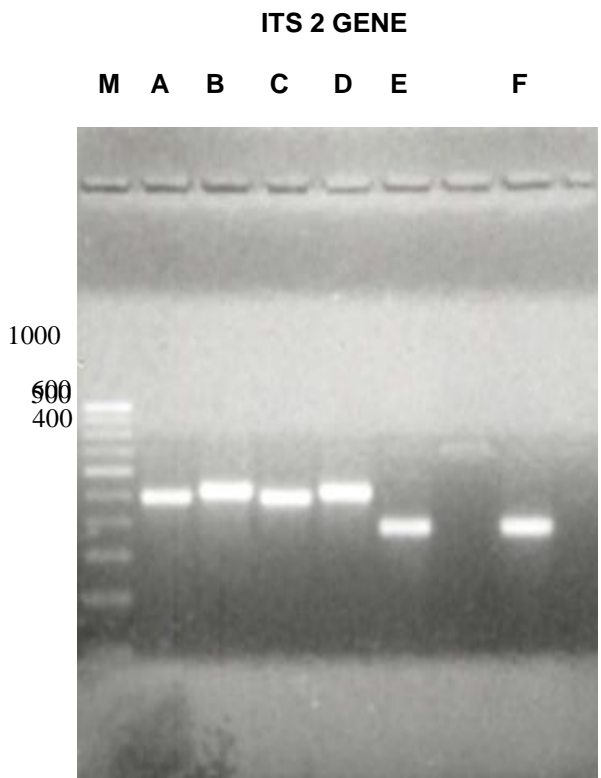
Nucleotide sequences were aligned using the default settings within the Clustal W multiple sequence alignment program<sup>[20]</sup> and the phylogenetic analysis was performed by applying several programs in PHYLIP 3.66 package<sup>[21]</sup>. The phylogenetic relationships were inferred by constructing trees using three methods viz. Maximum Parsimony (MP), Maximum Likelihood (ML) and Neighbour Joining (NJ). A distance matrix was generated using the Jukes and Cantor model (DNADIST in PHYLW) while neighbour joining method (Neighbour in PHYLIP) was used for constructing the phylogenetic tree. Two specimens from each species were sequenced one in both directions and other in forward direction only. As a part of the parsimony analysis, the robustness of the groups was analyzed by performing 1000 bootstrap replicates on the data sets. In order to test whether the data sets were evolving in a clock like mode, maximum likelihood analysis was performed with default values with and without a molecular clock hypothesis after which the likelihood values were subjected to the likelihood ratio test<sup>[21]</sup>. The transition:transversion (Ti:Tv) ratios were identified manually by comparing the sequences of closely related species from which transitions and transversions were carefully counted. The same programme mentioned above were also used for phylogeny reconstruction and seqboot and consense were used for bootstrapping. Treeview 1.40 was used to draw the phylogenetic trees presented in the following text. When two or more parsimonious trees were produced, a 50% majority rule consensus tree was constructed. ML (default model of DNAML in PHYLIP, where transversions were given double the weight of transitions and empirical base frequencies were used) and NJ (based on Kimura distance matrices) analysis of the sequences were performed with PHYLIP.

## Results and Discussion

### Sequence alignment analysis

In this study the sequence variation in the ITS2 region among six species of *Anopheles* mosquito was examined. All the six species showed high interspecies sequence variation in this region. Sequence was obtained from within the primers that annealed to the 5.8S and 28S regions flanking the ITS2 and in this sequence transition/transversion ratio with program DNAPARS was found to be 1.26. The PCR amplification of rDNA ITS2 of all the six species produced a single band (Fig.1).The ITS2 sequences varied between 453-506 base pair and were slightly G-C rich (Table 1). The sequences were submitted to

the gene bank under accession numbers EF192272 (*An. maculatus*), EF192273 (*An. annularis*), EF192274 (*An. culicifacies*), EF192275 (*An. stephensi*), EF192276 (*An. splendidus*) and EF192277 (*An. subpictus*). Figure 2 shows the multiple sequence alignment of the ITS2 sequence. The asterisks (\*) show those regions of the sequence where base pairs were identical in all the species while dashes (-) indicate the loci differing due to insertions or deletions (indels) of bases (Fig.2). The regions marked in bold indicate the repeat sequences while the arrows represent the base substitutions in *culicifacies* which were comparable with the remaining 5 species.



**Figure 1:** PCR amplified products of ITS2 gene of different anopheline species. 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 F: *Anopheles maculatus*.

**Table 1:** G-C content and length of ITS2 sequences analyzed in six *Anopheles* species.

Species	Length	Length difference	G+C %
<i>An. maculatus</i>	459/506	47	55.34
<i>An. annularis</i>	478/506	28	51.26
<i>An. subpictus</i>	491/506	15	52.75
<i>An. stephensi</i>	458/506	48	55.90
<i>An. splendidus</i>	488/506	18	54.10
<i>An. culicifacies</i>	506/506	0	60.28

maculatus -----  
stephensi -----ATGCTTT 7  
splendidus -----

*subpictus* TGTGAAGTGAAGGACACATGGTGTATGCTTAAATTCAGGGGGTTCGTCACACATCACTTG 60  
*annularis* -----  
*culicifacies* -----

↓

*maculatus* -----CGGCCTCCTTGAGAGAATGGCGAGCAGATG 30  
*stephensi* AAATTCAGGGGGTTCGTCACACATCACTTGACGCCTACTTGAGAGAATGGCGAGCAGATG 67  
*splendidus* -----CTGGAATCGATAGATTCGGG-----CATCCCGTATGTTGGTCGGA-G 41  
*subpictus* AGGCCTACTGTGCTAATATCAGATAATGAGAAG--TGCTATCCCGTATGTTGGTCGGA-G 117  
*annularis* -----GACGTAGAGGAATTTTCCACTTCCCGTGTTCCTTTGCAAGCCGATG 48  
*culicifacies* -----ACTGGAGT**GTGT**GCGGC 18  
\* \*

↓ ↓ ↓ ↓ ↓ ↓

*maculatus* GTGACGGGCGCG-TTGGGCCGCAA-**GTGT**TACTCGACC---CGAGCCAACCGGGGACCC 84  
*stephensi* GTGACGGGCGCG-TTGGGCCGCAA-**GTGT**TACTCGACC---CGAGCCAACCGGGGACCC 121  
*splendidus* ATTCCGAGCGCGGTTGTACTTCCCC**GTGT**TACTCGACC---CTAGCCAACCTAGGACCC 97  
*subpictus* ATTCCGAGCGCGGTTGTACTTCCCC**GTGT**TACTCGACC---CTAGCCAACCTAGGACCC 173  
*annularis* AAGACAGACCCG-TTGGGCCGCAA-**GTGT**TACACGACC---AGAGCCCCCAGGACCC 102  
*culicifacies* TCCTGTGTTACAATGGGTCGCTACGCGTTACAGAACTGGCCGGA**ACTTCC**CTGGACCC 78  
\* \*

*maculatus* CTTTGCTAGCGCCTGGTGTAACTGTTAGGCTGCGGGGGTTTCATCCCTAGTTGATACT 144  
*stephensi* CTTTGCTAGCGCCTGGTGTAACTGTTAGGCTGCGGGGGTTTCATCCCTAGTTGATACT 181  
*splendidus* CTTTGCTAGCGCCTGGTG-----ATAGGCTGCGGGGGTTTCATCCCGGGTTGATACT 149  
*subpictus* CTTTGCTAGCGCCTGGTG-----ATAGGCTGCGGGGGTTTCATCCCGGGTTGATACT 225  
*annularis* CTTTGCTAGCGCCCG-TG-----ATAGGCTGCGGGGGTTTATCCCGGGTTGATACT 153  
*culicifacies* GTATTGGCAAAGCCT**GTGT**TA-----GTAGGCTGCGGGGGTTCTCAGGTATCCAAAGCC 133  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* GGAGGTGCAACCGTGC-GACTTGACGCG--CGCCCGTTCATCTCTCGCCTTGCGTACCAC 201  
*stephensi* GGAGGTGCAACCGTGC-GACTTGACGCG--CGCCCGTTCATCTCTCGCCTTGCGTACCAC 238  
*splendidus* GTAAGCCGAACCGTGC-GATACG-CGCG--CGCCCGTCCATCTCTCACCATTGAACCCG 205  
*subpictus* GTAAGCCGAACCGTGC-GATACG-CGCG--CGCCCGTCCATCTCTCACCATTGAACCCG 281  
*annularis* G-AGTTCGAACCATGC-GACTTGTGCGCG--TGCTGTTTCTCAGTACCTTACCATACA 210  
*culicifacies* GAAGTTCGACCCGTGCCACTTGACGCGCGCACCCGTTTCGC-CAAACTTTTCAGCGCC 192  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* CTTGT-----TAATAGCA**CTCT**CACT---TTCAGCGCCACGGT-CCCAC-----AT 244  
*stephensi* CTTGT-----TAATAGCA**CTCT**CACT---TTCAGCGCCACGGT-CCCAC-----AT 281  
*splendidus* **TATA**TGTCTTCTATTAAC**ACTCT**CATAGTATTAAGCGCCCAAGGT-CCCGTCG-----GT 259  
*subpictus* **TATA**TGTCTTCTATTAAC**ACTCT**CATAGTATTAAGCGCCCAAGGT-CCCGTCG-----GT 335  
*annularis* AA-GTCATCTGTATTAGCTTTCAC**TCT**---TTCAGCGCCCAAGGT-CCCGTC-----AG 260  
*culicifacies* CAAGAGCCCCCTTGGCGGACTCGAGCACGCCAAGATGCACAGTGCGCCACCGTGCAGTGC 252  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* ACGGGGTC--CGAGCACGCCA---TTCTGCGACAG----CCCATCGCGGGGA----- 287  
*stephensi* ACGGGGTC--CGAGCACGCCA---TTCTGCGACAG----CCCATCGCGGGGA----- 324  
*splendidus* CGCGGGTC--CGAGCACGCCA---TGCTGCGACAG----CCCATCTCGAGG----- 301  
*subpictus* CGCGGGTC--CGAGCACGCCA---TGCTGCGACAG----CCCATCTCGAGG----- 377  
*annularis* TGCGGGTC--CGAGCACGCCA---TGATGCGACA----TCTATCCCGGAGA----- 302  
*culicifacies* TGCAA**ACCACCA**CCGCCAAGGCGTTGGGACGGGTGCTCGTCCCG**GTGT**TGCTTG 312  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* -----TGGGACAGTCA**GT**CTGGTA-GGAGTATATCGATAACTTGGTAGGCACTCAAG 338  
*stephensi* -----TGGGACAGTCA**GT**CTGGTA-GGAGTATATCGATAACTTGGTAGGCACTCAAG 375  
*splendidus* -----TGGGACAGTCA**GT**TTGATA-ATAGGATATTGACAACCTGGTAGGCACTCAAG 352  
*subpictus* -----TGGGACAGTCA**GT**TTGATA-ATAGGATATTGACAACCTGGTAGGCACTCAAG 428  
*annularis* -----TAGG-CTGTCA**GT**TTGATA-AATGCGTATAGATAACTTGGTAGGCACTCAAG 352  
*culicifacies* CCTGCACGCGCCCTGTAGT**GT**TTTATGTTGTAACACGCATTTGGTAGGCACTCAAG 372  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* GAT**GTGT**GCA-**TCGGT**CGGGTTGAAACGTCGGATGCGCAATATGCGTTCAACGTGTCGGT 397  
*stephensi* GAT**GTGT**-CA-TCGGT**CGGGT**GAAACCGAGTTTCATCCATCATGCGTTCAACGTGTCGGT 433  
*splendidus* GAT**GTGT**GCA-TCGGT**CGGGT**TGAGACGTCGGATGCGCAATATGCGTTCAACTTATCGGT 411  
*subpictus* GTATGGTGAATCCGTATGCTTAAACTCTAGGGGGTGGATCACCTCCTACACTGCACGTT 488  
*annularis* GAT**GTGT**GCA-TCGGT**CGGGT**TGAGACGTCGGATGCGCAATATGCGTTCAACTTATCAAT 411  
*culicifacies* AAT**GTGT**TACC-CCGGT**CGGGT**TGAATCGTCCCATGCCCATATGCGTTCAACGTGCCCGG 431  
\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

*maculatus* GTTCATGT---GTCTTGCAGTT-CACAACAAGTATGTGTGACGACCCCTGAAATTTTA 453

```

stephensi      GTTCATGT---GTCCTGGCAGTCACACA----- 458
splendidus    GTTCATGT---GTCCTGCAGTTCACAACACAGTAGGCCTCAAGTGATGTGTGACGACCC 468
subpictus     ACG----- 491
annularis     GTTCATGT---GTCCTGCAGTTCACAGGCCTCAAGTGATGTGTGACACCTCGAAATATAT 468
culicifacies  TTCTGTGTCCTGCACTCCCAACGCTTTATCAGTAGGCCTCAAGTGATGTGTGACAACCC 491

maculatus     AGCATA----- 459
stephensi     -----
splendidus    CCTGGAATTTAAGCATAACC 488
subpictus     -----
annularis     ATAACATATA----- 478
culicifacies  CCTGAATTAAGCATA----- 506

```

**Table 2: Transitions (Ti), transversions (Tv), Ti/Tv ratio and indels in six *Anopheles* species**

Species	Ti (%)	Tv(%)	Ti/Tv	Total number of substitutions	Indels
<i>An. maculatus</i>	8 (9.52)	14 (7.40)	0.57	22	54
<i>An. annularis</i>	24 (28.57)	28 (14.8)	0.88	52	7
<i>An. subpictus</i>	21 (25)	38 (20.1)	0.55	59	54
<i>An. stephensi</i>	12 (14.28)	23 (12.17)	0.52	35	50
<i>An. splendidus</i>	14 (16.66)	14 (7.40)	1	28	50
<i>An. culicifacies</i>	15 (17.85)	72 (38.0)	0.21	87	59
	84	189		283	274

**Figure 2: Multiple sequence alignment of ITS2 sequences of *Anopheles* mosquitoes**

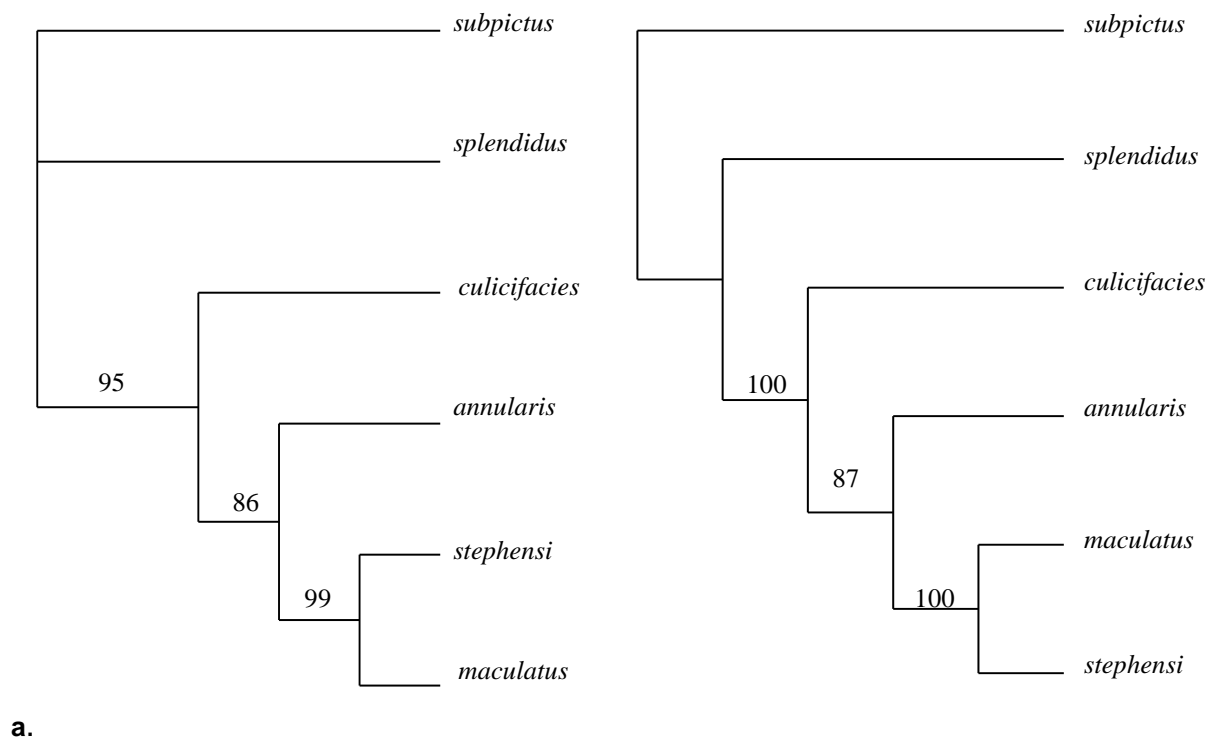
From the sequence alignment, it was observed that while comparing the 5 sequences with *culicifacies*, maximum substitutions were shown by it. Out of which transversions were more (72) than transitions (15). A careful analysis of the entire sequence reveals that this particular sequence is G-C rich as it was found to be maximum in this species wherein it was as much as 60.28% of the total bases. As per the incidence of substitution, the rate of transversions were found to be 66.78% which was more than transitions which accounted only 29.68%. The ITS2 gene was largest in *culicifacies* (506 bp) and smallest in *maculatus* (453 bp). The maximum number of 54 insertion deletion elements (indels) were present in *maculatus* as compared to only 7 in *culicifacies* (Table 2).

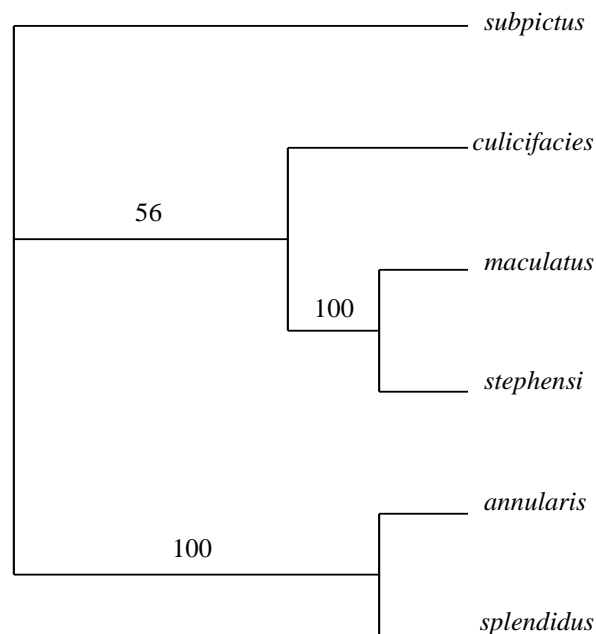
For example insertion in *culicifacies* is shown at 57-60. Indels were minimum in *culicifacies*. For example at position 150, 164 and 340 in *culicifacies*, insertion of C,G and T were found respectively. Simple tandem repeats were present at various locations in different species along ITS2. For example in *An. maculatus* there were three repeats of the sequence GAGA at position 11, two repeats of TGTG at position 341 and 430 and TCGG at position 349 (Figure 2). Triple CCA repeat was found in *culicifacies* at 260 base pair. Out of the large number of tandem repeats GT repeats were most common followed by GA which were maximum in *culicifacies*. The tandem repeats at various positions in different species indicates species-specific mutations in that particular species. The large number of base repeats in case of *culicifacies* account for most of the sequence variation observed and suggest their role as a major cause of divergence in the evolution of this spacer. According to the sequence alignment, the *splendidus* and *subpictus* were found to show maximum similarity (Figure 2). From the present results, it is evident that sequence comparisons of different species can provide an estimate of their genetic relatedness through molecular diagnostics.

## Phylogenetic analysis

A molecular phylogeny was constructed using the ITS2 sequences from six *Anopheles* species from the alignment in Figure 2. The pairwise Kimura-2 parameter distances shows the maximum similarity between *maculatus* and *stephensi* with the lowest value while among the members of Neocellia series *splendidus* was encountered as the closest taxa to *subpictus* with a value of 0.1127 (Table 3). Single base frequencies for the coding strand were as follows: A= 0.20521, C =0.27569, G =0.27326, T =0.24583. The likelihood ratio test of Felsenstein comparing the maximum likelihood tree with molecular clock to that without molecular clock rejected the molecular clock for all data sets.

The difference between probabilities of two likelihood trees was 24.48 which was larger than critical value with  $df=6$  and  $p\text{-value}=0.05$ . In maximum likelihood analysis, the likelihood ratio was found to be  $-\ln L=2509.67180$ . The gene trees were generated by all the three methods i.e. MP, ML and Distance tree. The trees are unrooted as an outgroup with an easily aligned ITS2 is not available. Both MP and ML methods of tree construction resulted in the same tree topology while the relative positions of species were altered when neighbour joining tree was used, although *stephensi* and *maculatus* still formed a single clade. Bootstrap values of 99 and 100 from all the three trees generated clearly indicate the close relationship of *stephensi* with *maculatus*. Both MP and ML tree generated formed 3 clades, one clade of *stephensi* and *maculatus* while the second clade joining *culicifacies* while third one including *subpictus* and *splendidus*. As *subpictus* falls in Pyrethrophorus series it was found to be in close relationship to *splendidus* among the Neocellia series (Figure 3a, b). *Culicifacies* being a member of Myzomyia series forms a separate clade but bootstrap support value is only 56%. The tree generated by neighbour joining method clearly indicates the closeness of *annularis* and *splendidus* which are morphologically similar while for the rest two trees it was not the same which might be due to the fact that ITS2 gene lies under highly differential constraints (Figure 3c). The maximum divergence was shown by *stephensi* and *subpictus* among the studied *Anopheles* species.





c.

**Figure 3: Dendrograms based on rDNA ITS2 sequence data for the taxa in the mosquito genus *Anopheles* using a) maximum parsimony b) maximum likelihood c) neighbour joining method. The confidence probability value is given above the branch and bootstrap confidence level from 1000 replicates.**

**Table 3: Interspecies divergence in ITS2 gene sequences in six *Anopheles* species based on Kimura-2 parameter distances**

<i>An. maculatus</i>	<i>An. stephensi</i>	<i>An. splendidus</i>	<i>An. subpictus</i>	<i>An.annularis</i>	<i>An.culicifacies</i>	
<i>An. maculatus</i>	0.000000					
<i>An. stephensi</i>	0.027620	0.000000				
<i>An. splendidus</i>	0.369819	0.319528	0.000000			
<i>An. subpictus</i>	0.350589	0.455994	0.112710	0.000000		
<i>An.annularis</i>	0.329158	0.348165	0.481074	0.500863	0.000000	
<i>An. culicifacies</i>	0.873359	0.748277	0.655745	0.952336	0.900031	0.000000

## Conclusion

Molecular variations at the intra and interspecific levels have been worked out by many workers on different members of the subgenus *Cellia*. DNA sequence divergence in *An. aconitus*, *An. varuna*, *An. minimus* A and C which are members of *minimus* group of subgenus *Cellia* has been studied<sup>[22]</sup>. PCR based diagnostic assays have been developed by many researchers due to these characteristic features of ITS2. Similarly, Hackett *et al.* (2000)<sup>[7]</sup> distinguished *An. (Cellia) funestus* Giles from *An. (Cellia) rivulorum* Leeson and detected a cryptic taxon within the *funestus* group based on divergence in the ITS2 sequence. Also Manonmani *et al.* (2001)<sup>[23]</sup> sequenced the same fragment of *An. fluviatilis* collected from Koraput and Malkangiri districts of Orissa, India. On the basis of similar studies, Naddaf *et al.* (2003)<sup>[24]</sup> also determined the composition and distribution of members of *An. fluviatilis* complex from a few locations in Iran. Numerous sequence based studies have been carried out using sequences of nuclear rDNA genes, mitochondrial genes and internal transcribed spacers<sup>[25, 26]</sup>.



The results presented here reveal high interspecific variation which is consistent with the work already done on this gene which indicates variance of spacer mutants in the genus *Anopheles*. rDNA based PCR assays have been developed based on molecular divergence based on many of the world's *Anopheles* malaria vectors<sup>[2]</sup>. As a result of the present data it had been possible to examine the usefulness of the ITS 2 in the nuclear ribosomal gene array for conducting a phylogenetic analysis of selected taxa within the genus *Anopheles*. Further study of the species by the analysis of their rDNA, mt.DNA and other genetic markers such as RFLP or microsatellites should help to study taxonomy, evolutionary systematics and population genetics. Trees recovered by the study of a single gene make it difficult to predict their utility for particular evolutionary questions. However it is indicated by earlier data that there is need to use combination of genes with better performance than a single gene to make phylogenetic inference<sup>[27, 28]</sup>. From the present results it is evident that sequence comparisons of different species can provide an estimate of their genetic relatedness through molecular diagnostics. However, more species in *Anopheles* should be analyzed to understand their systematic position.

## Acknowledgement

The authors are highly thankful to Council of Scientific and Industrial Research (CSIR) as the first author received fellowship support vide F.No. 9/135 (477)/ 2004/EMR-I.

## References

1. Arnheim N., Concerted evolution of Multigene families, Evolution of genes and Proteins (eds.). M Nei and R K Koehn, Sinauer Associates Inc., Massachusetts. pp. 38-62 (1983).
2. Collins F. M. and Paskewitz S. M., A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species., Insect Molecular Biology. 5,1-9, (1996).
3. Scott J. A., Brogdon W. G. and Collins F. H., Identification of single specimens of the *Anopheles maculipennis* complex by the polymerase chain reaction, Am. J.Trop.Med.Hyg. 49, 520-529, (1993).
4. Beebe N. W. and Saul A., Discrimination of all members of the *Anopheles punctulatus* complex by polymerase chain reaction – restriction fragment length polymorphism analysis, Am. J. Trop. Med. Hygiene. 53, 478-481, (1995).
5. Cornel A. J., Porter C. H. and Collins F. H., Polymerase chain reaction species diagnostic assay for *Anopheles quadrimaculatus* cryptic species (Diptera: Culicidae) based on ribosomal DNA ITS2 sequences, J. Med. Entomol. 33, 109- 116, (1996).
6. Proft J., Maier W. A. and Kampen H., Identification of six sibling species of the *Anopheles maculipennis* complex (Diptera: Culicidae) by a polymerase chain reaction assay, Parasitol. Res. 85(10), 837-843, (1999).
7. Hackett B. J., Gimnig J., Guelbeogo W., Costantine C., Koekemoer L. L. *et al.*, Ribosomal DNA internal transcribed spacer (ITS2) sequences differentiate *Anopheles funestus* and *An. rivulorum* and uncover a cryptic taxon, Insect Mol. Biol. 9, 369- 374, (2000).
8. Karotam J., Boyce T. M. and Oakeshott J. G., Nucleotide variation at hypervariable esterase 6 isozyme locus of *Drosophila simulans*, Mol. Bio. Evolution. 12, 113-122, (1995).
9. Mattick J. S. and Gagen M. J., The evolution of controlled multiple gene networks: the role of introns and other noncoding RNAs in the development of complex organisms, Mol. Bio. Evolution. 18, 1611-1630, (2001).
10. Zieler H. and Huynh C, Q., Intron dependent stimulation of marker gene expression in cultured insect cells, Insect Molecular Biology. 11, 87-95, (2002).

11. Zhang D. X. and Hewitt G. M., Assessment of the universality and utility of a set of conserved mitochondrial COI primers in insects, *Insect Molecular Biology*. 6, 143-150, **(1997)**.
12. Sunnucks P., Efficient genetic markers for population biology, *Trends in Ecology and Evolution* 15, 199-203, **(2000)**.
13. Kocher T. D., Thomas W. K., Meyer A. *et al.*, Dynamics of mitochondrial- DNA evolution in animals – amplification and sequencing with conserved primers, *Proceedings of the National Academy of Sciences USA*. 86 , 6196-6200, **(1989)**.
14. Simon C., Molecular systematics at species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. In: *Molecular Techniques in Taxonomy*. Hewitt, G. M., Johnston, A. W. B. and Young, J. P .W. (eds.). Springer-Verlag, Berlin. pp. 33-71 **(1991)**.
15. Simon C., Frati F., Beckenbach A., Crespi B., Liu H. and Flook P., Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers, *Ann. Ent. Soc. Amer.* 87, 651-701, **(1994)**.
16. Singh K. R. P., Patterson R. S., La Brecque G. C. and Razdan R. K., Mass rearing of *Culex pipiens fatigans* Wied, *J. Com. Dis.* 7, 31-53, **(1975)**.
17. Clements A., *The biology of mosquitoes: development, nutrition and reproduction* Chapman and Hall, London. **(1994)**.
18. Ausubel F. M., Brent R., Kingston R. E., Moore D. D., Sideman J. G., Smith J. A. and Struhl K., *Short protocols in molecular biology*, John-Wiley and Sons Inc., 4, 15.1-15.37, **(2000)**.
19. Walton C., Handley J. M., Kuvangkadilok C., Collins F. H., Harbach R. E. *et al.*, Identification of five species of the *Anopheles dirus* complex from Thailand, using allele-specific polymerase chain reaction, *Med. Vet. Entomol.* 13 (1), 24-32, **(1999)**.
20. Thompson J. D., Higgins D. G. and Gibson T. J., CLUSTAL W: improving the Sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Research*. 22(22), 4673-4680, **(1994)**.
21. Felsenstein J., PHYLIP (Phylogeny Inference Package) Version 3.5p edition Department of Genetics, University of Washington, Seattle, **(1993)**.
22. Sharpe R. G., Harbach R. E. and Butlin R. K., Molecular variation and phylogeny of members of the Minimus group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae), *Syst. Entomol.* 25, 263-272, **(2000)**.
23. Manonmani A., Townson H., Adeniren T., Jambulingam P., Sahu S. and Vijaykumar T., rDNA ITS2 polymerase chain reaction assay for the sibling species of *Anopheles fluviatilis*, *Acta tropica*. 78, 3-9, **(2001)**.
24. Naddaf S. R., Oshaghi M. A., Vatandoost H. and Assmar M., Molecular characterisation of *Anopheles fluviatilis* species complex in the Islamic Republic of Iran, *Eastern Mediterranean Health Journal*. 9(3), 257-265, **(2003)**.
25. Ballinger-Crabtree M. E., Black W. C. IV. and Miller B. R., Use of genetic polymorphisms detected by the random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) for differentiation

and identification of *Aedes aegypti* subspecies and populations, Am. J. Trop. Med. Hygiene. 47, 893-901, **(1992)**.

26. Cooper L., Cooper R. D. and Burkot T. R., The *Anopheles punctulatus* complex: DNA probes for identifying the Australian species using isotopic, chromogenic and chemiluminescence detection systems, Exp. Parasitol. 73, 27-35, **(1991)**.
27. Cummings M. P., Otto S. P. and Wakeley J., Sampling properties of DNA sequence data in phylogenetic analysis, Mol. Biol. Evol. 12, 814-822, **(1995)**.
28. Nei M., Phylogenetic analysis in molecular evolutionary genetics, Annu. Rev. Genet. 30, 371-403, **(1996)**.