# Comparison Of Secondary Structure Of The Ribosomal Internal Transcribed Spacer 2 (ITS2) Of Eight Lepidopteran Species From Diverse Geographical Locations

M Kuracha, B Rayavarapu, S Kumar, P Rao

## Citation

M Kuracha, B Rayavarapu, S Kumar, P Rao. *Comparison Of Secondary Structure Of The Ribosomal Internal Transcribed Spacer 2 (ITS2) Of Eight Lepidopteran Species From Diverse Geographical Locations*. The Internet Journal of Genomics and Proteomics. 2005 Volume 2 Number 1.

## Abstract

A comparative study of common core secondary structure in the ribosomal internal transcriber 2 (ITS2) of 8 Lepidopteron species selected from different geographical locations was carried out. Among the selected insects some are serious pests of agricultural crops. Ex: Helicoverpa punctigera. Multiple sequence alignment and secondary structural analysis of ITS2 was performed to elucidate the phylogenetic relationship. These studies indicated a phylogenetic relationship among the selected insects belonging to different geographical locations. Several common features of secondary structure are shared among these species, with some of them supported by compensatory changes, suggesting the significant role by ITS2 as an RNA domain during ribosome biogenesis.

## INTRODUCTION

Lepidopteran insects are basically phytophagous in nature. The larvae of these lepidopteran insects are voracious. Hence, they have become major pests of various crops [1,2,3,4]. Intra-specific conservation and variations have been reported using ITS2 region collected from diverse geographic locations that are used for phylogenetic studies [5, 6, 7, 8]. The internal transcribed spacers (ITS) are located between the repeating array of nuclear 18S, 5.8S and 28S ribosomal RNA genes, a locus that has 100–200 copies per genome. The ITS spacers are versatile genetic markers and have been used for phylogenetic analysis, evaluation of the evolutionary process, as well as for determination of taxonomic identities [9]. The lepidopteran insects selected for the present study were: Helicoverpa punctigera, Luehdorfia longicaudata, Luehdorfia chinensis, Luehdorfia puziloi puziloi, Luehdorfia puziloi yessoensis, Heliconius charithonia, Heliconius hecale, Heliconius melpomene.

Mature rRNAs are produced by the processing of a large precursor from which different transcribed spacer regions are sequentially removed through an elaborate pathway of cleavage steps [10]. In eukaryotes, transcribed spacer regions may represent a very substantial fraction of the length of the primary transcript. Although these transcribed spacer regions are obvious candidates for important roles in the control of ribosome biogenesis, elucidation of their biological function and of the molecular mechanisms involved in their accurate excision still remain a major challenge. Recent functional analyses performed on yeast ribosomal RNA genes clearly show that the structural integrity of the transcribed spacer regions is an essential prerequisite for correct processing of mature rRNA and biogenesis of active ribosomal subunits [6, 11]. The derivation of reliable secondary structure models for each transcribed spacer region would undoubtedly represent a major step towards a detailed understanding of their biological role. The comparative sequence analysis provides the most powerful tool for identifying the biologically relevant folding pattern of an RNA molecule, i.e. its native structure within the cellular context [6]. However, an essential prerequisite for its effective utilization is the availability of a collection of sequences exhibiting a substantial number of nucleotide differences while remaining similar enough for unequivocal sequence alignments. Due to the high rate of sequence variation of transcribed spacers, this may exhibit dramatic size variation and extensive sequence divergence even among moderately distant species [12, 13]. Nevertheless, the presence of phylogenetically conserved secondary structure elements in the 5' externally transcribed spacer was recently revealed by the comparative analysis of a limited set of vertebrate sequences [14].

The present study is focused on internal transcribed spacer, ITS2, which interrupts the eukaryotic large subunit rRNA molecule and has no prokaryotic equivalent [15]. The sequence and secondary structures of ITS2 of selected geographically variant Lepidopteran insect species were comprehensively investigated. Such case studies are relevant in broader phylogenetic contexts and for analyzing the function in ribosome biogenesis. Since the secondary structures of ITS region are more conserved than the nucleotide sequences their analysis helps in understanding molecular evolution and increases the number of structural characters. Thus the structure models developed in this study can be used for future phylogenetic analyses.

## MATERIALS AND METHODS
### DATA SET

ITS 2 sequences of eight Lepidopteran species belonging to diverse geographical locations (Costa Rica, China, Russia, America and Japan) that are deposited in Genbank were investigated. The accession numbers of Helicoverpa punctigera, Luehdorfia longicaudata, Luehdorfia chinensis, Luehdorfia puziloi puziloi, Luehdorfia puziloi yessoensis, Heliconius charithonia, Heliconius hecale, Heliconius melpomene are: AF047759, AB071926, AB071925, AB071923, AB071911, AF453773, AF453768, and AF453767 respectively.

### SEQUENCE ALIGNMENTS

Multiple sequence alignments were performed using CLUSTALW with a gap opening penalty of 15 and gap extension penalty of 6.66.

### SECONDARY STRUCTURE PREDICTION

The RNA secondary structures for ITS2 were predicted using RNADRAW [21]. RNADRAW predicts RNA structures by identifying suboptimal structures using the free energy optimization methodology at a default temperature of 370C. In the current study, ITS2 and 5.8S regions (the first 170 nucleotides) were used for RNA structure prediction. The minimum energy structure prediction algorithm in RNADRAW was ported from the RNAFOLD program included in the Vienna RNA package [16]. The dynamic programming algorithm employed in RNADRAW was based on the work of Zuker and Stiegler [17] and uses energy parameters taken from Freier [18] and Jaeger [19].

### RNA FOLD

The Sribo program in Sfold (statistical Folding and Rational Design of Nucleic Acids) was used to predict the probable target accessibility sites (loops) for trans-cleaving ribozymes in ITS2 [20]. The prediction of accessibility is based on a statistical sample of the Boltzmann ensemble for secondary structures. Here, we assessed the likelihood of unpaired sites for potential ribozyme target. Each mRNA exists as a population of different structures. Hence, stochastic approach to the evaluation of accessible sites was found appropriate [21]. The probability profiling approach by Ding and Lawrence [22] reveals target sites that are commonly accessible for a large number of statistically representative structures in the target RNA. This novel approach bypasses the long-standing difficulty in accessibility evaluation due to limited representation of probable structures due to high statistical confidence in predictions. The probability profile for individual bases (W = 1) is produced for the region that includes a triplet and two flanking sequences of 15 bases each in every site of the selected cleavage triplet (e.g., GUC).

## PHYLOGENETIC ANALYSIS

The phylogenetic Genebee service was used for phylogenetic tree construction [23].

## RESULTS
### SEQUENCE ANALYSIS

The length of ITS2 elements of eight selected lepidopteran species ranged in size between 644 and 699 bp. eight dispersed but unambiguously conserved sequence segments encompassing about a third of the ITS2 length have been identified. They were interspersed with variable regions and gaps where size variations accumulate. The characteristics of the sequences for each species are shown in Table 1. The length variations were observed with maximum length being 699 bp and minimum of 644 bp for Heliconius charithonia and Helicoverpa punctigera respectively. The G+C contents for the two regions of rDNA (5.8S and ITS2) of all species ranged from 54% to 69%. For ITS2 regions the sequence identities ranges, with maximum 99% similarity between Luehdorfia longicaudataand Luehdorfia chinensis; 95% between Luehdorfia longicaudata and Luehdorfia puziloi puziloi whereas the minimum being the 2% between Luehdorfia puziloi puziloi and Heliconius hecale. Alignment of ITS2 region is shown in the Figure: 1 Simple tandem repeats were present at various locations along the ITS2. The sequence similarity is more towards the 5' end and with dispersed conserved ness in the middle than towards the 3' end.

## Figure 1

Table 1: Lengths (in mers); G+C content (in %); GC, AU, GU base pairs (in numbers); No: stems and energies (in kcal) of secondary structures of the second internal transcribed spacer (ITS-2) rDNA sequences of the species.

| Species | ITS 2 | | 5.8S+ITS2 | | | No of stems | Energy |
|---|---|---|---|---|---|---|---|
| | Length | G+C | GC | AU | GU | | |
| *Helicoverpa punctigera* | 644 | 56 | 129 | 62 | 36 | 42 | -229.83 |
| *Luehdorfia longicaudata* | 681 | 54 | 140 | 66 | 38 | 48 | -251.67 |
| *Luehdorfia chinensis* | 691 | 54 | 141 | 69 | 34 | 48 | -252.52 |
| *Luehdorfia puziloi puziloi* | 668 | 55 | 140 | 64 | 35 | 48 | -253.31 |
| *Luehdorfia puziloi yessoensis* | 667 | 55 | 139 | 60 | 37 | 47 | -249.39 |
| *Heliconius charithonia* | 699 | 69 | 145 | 57 | 35 | 48 | -252.85 |
| *Heliconius hecale* | 658 | 61 | 147 | 55 | 22 | 45 | -263.21 |
| *Heliconius melpomene* | 678 | 59 | 143 | 60 | 25 | 46 | -256.14 |

## SECONDARY STRUCTURE

Secondary structural features of ITS2 regions were given in the Table: 1. The secondary structures of the mentioned Lepidopteran species were classified into three groups based on the analysis of conserved stems and loops. Class I includes L.longicaudata, L.chinensis, L.puziloi yessoensis, L.puziloi puziloi that show overall similarity in the ITS2 rDNA folding where L.longicaudata and L.chinensis have identical secondary structures. Secondary structures of remaining species are highly variant. Two common motifs, having sequences UGUCG and CUUCGGUG respectively were conserved in all classes. Apart from the common conserved motifs shared among the species that are categorized into different classes, variable regions also do exist. The observed similarities at the secondary structural level are further reflected at energy level.

## DISCUSSION

The selected 8 Lepidopteran insect species occur worldwide and they are pests of various crops. In the present investigation, the ITS2 sequences reflected the trend observed in the phylogeny. The more distantly related the less was the convergence at the ITS2 level (Figure 1 &2).

## Figure 2

Figure 1a: Alignment between species ITS2 regions

**Figure 3**



Hp: Helicoverpa punctigera, Ll: Luehdorfia longicaudata, Lch: Luehdorfia chinensis, Lpu: Luehdorfia puziloi puziloi, Lpuz: Luehdorfia puziloi yessoensis, Hchar: Heliconius charithonia, Hhec: Heliconius hecale, Hmel: Heliconius melpomene.

**Figure 4**

Figure 1b: Phylogenetic tree of 8 selected lepidatperan species.



Hp: Helicoverpa punctigera, Ll: Luehdorfia longicaudata, Lch: Luehdorfia chinensis, Lpu: Luehdorfia puziloi puziloi, Lpuz:Luehdorfia puziloi yessoensis, Hchar: Heliconius charithonia, Hhec: Heliconius hecale, Hmel: Heliconius melpomene.
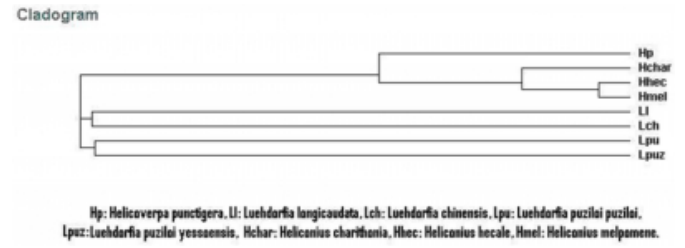
**Figure 5**
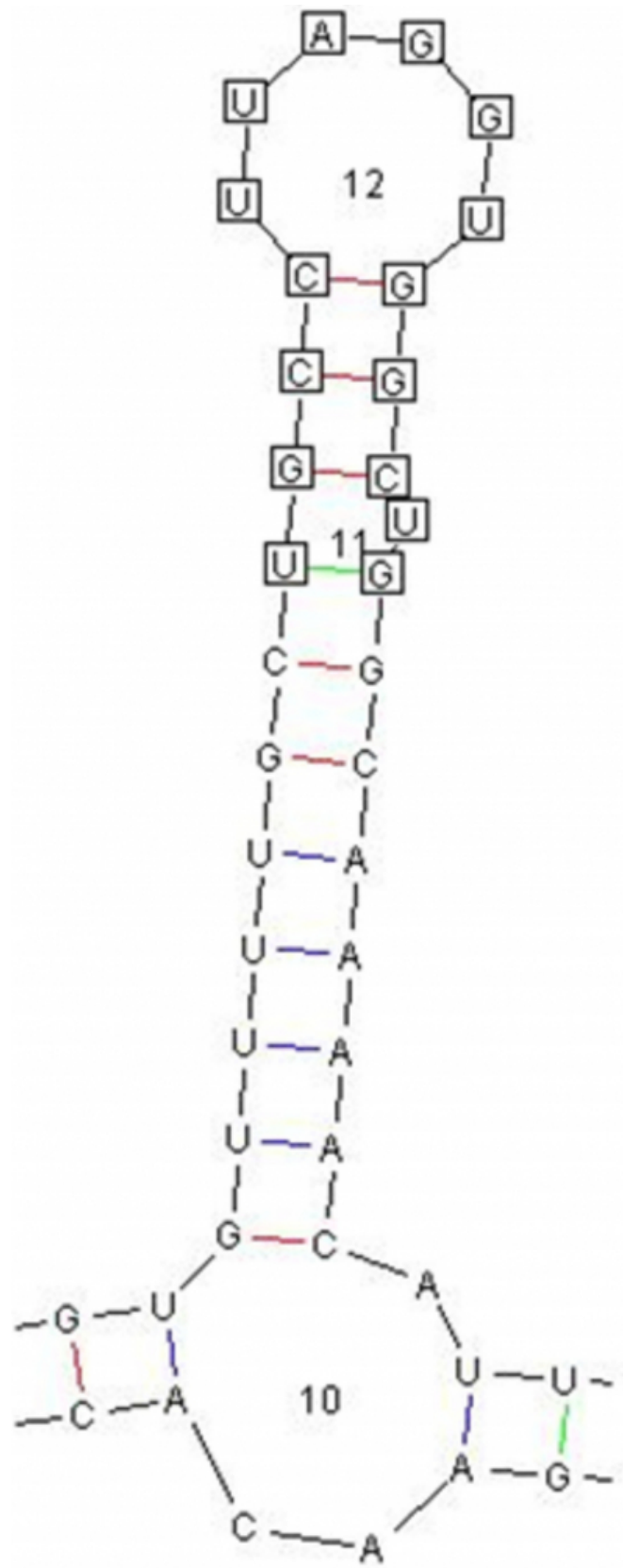
Figure 2: A secondary structural fold found in common to all 8 Lepidateran species. Identical bases are marked.



However accumulated substitutions in the ITS sequence leading to length variation also had a profound effect on the

conservedness among the structures. The length variation observed was may be due to insertions effected by many factors including genetic drift, the relative number and size of repeats, rates of unequal crossover, gene conversion, immigration and the number of the loci [24]. But still high level of sequence conservation was found between some species like Luehdorfia longicaudata and Luehdorfia chinensis i, Luehdorfia longicaudata and Luehdorfia puziloi puziloi, [25] even the simple tandem repeats were found to be conserved to a large extent among them. This conservation was further reflected at secondary structural and energy level. The predicted features of ITS2 using RNADRAW are given in Table 1. The stems (double stranded paired regions) stabilize RNA secondary structures and the number of stems present in each ITS2 is given (Table 1). ITS2 RNA structures from Heliconius hecale and Heliconius melpomene the highest negative free energy (-263.21Kcal and -256.14 Kcal) followed by Luehdorfia puziloi puziloi (-253.31), Heliconius charithonia (-252.85) Luehdorfia chinensis (-252.52) and then by Luehdorfia longicaudata (-251.67), Luehdorfia puziloi yessoensis (-249.39) and Helicoverpa punctigera (-229.83). Visual comparison shows that this is related to the trend in the cladogram given in Figure 1A.This convergence at secondary structural level among few species from different geographic isolates is may be due to the evolutionary pressure on ITS2 to maintain the RNA secondary structure involved in post-transcriptional processing of rRNA [7]. Secondary structure predictions for the ITS2 region indicate that these domains base pair to form a core region central to several stem features implying that conserved ness is more important for the proper rRNA folding pattern [26]. In contrast to the present observation with respect to Class I, Barker in his study found that [27] ITS 2 is unique in the 16 populations, of Rhipicephalus and Boophilus species, with considerable nucleotide variation among species and genera. However intraspecific variation of ITS2 sequences was also found in the populations. Analysis of ITS spacers of T. rangeli and T. cruzi allowed the distinction of two distinct groups, revealing a low-level similarity between them [28]. Figure 3 shows the distribution of different types of loops (hairpin, bulge, multi branched, interior and exterior) among different isolates. The segments of the ITS2 having score ? 50 are further probed carefully for target site to assess the likelihood of unpaired segments. Interestingly, the observed phylogenetic trend was identified with respect to the target accessibility sites for the eight Lepidopteron isolates. The order of preference is interior loop, bulge loop, multiple branched loop, hairpin loop and

exterior loops in the all the isolates.

## Figure 6

Figure 3: Distribution of different types of loops [Hairpin (H), Bulge (B), Multi branched (M), Interior (I) and Exterior (E)] among different isolates of Lepidopteron species.
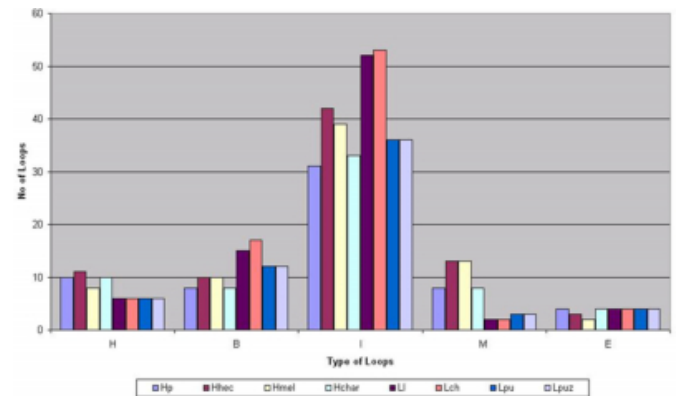


These results suggest that the differences and conserved ness observed between ITS-2 of different species are not "neutral" and are not simple accumulated random nucleotide changes, but bear a significant functional load. In the previous study of three related mosquito genera (Aedes, Psorophora, and Haemogogus), [26] it was found that intra spacer variable regions appear to co-evolve and that ITS-2 variation is constrained to some extent by its secondary structure. Further studies on yeast [29] have demonstrated that the ITS-2 is essential for the correct and efficient processing and maturation of certain ribosomal units. Furthermore, information for the efficient removal of ITS-2 from its RNA precursor is dispersed through the entire ITS-2 region, and indels that effect secondary structure differentially alters rRNA processing. Critical changes in the rRNA folding pattern brought about by sequence evolution in the ITS spacer regions may thus have an important influence on the kinetics of precursor rRNA formation, and ultimately on the efficient functioning of the rDNA cluster.

## CONCLUSION

The present study shows two contrasting aspects of ITS2 regions i.e. the general trend of variability among the species as well as the conservedness between few species. Surprisingly, the species displaying the conservedness belong to different geographical locations with diverse climatic and ecological conditions. Our study implies that the ITS2 regions though have less selective pressure than the ribosomal regions but still evolve slower than the intergenic spacers, indicating that some selective pressure do exists on them, probably from the constraint to maintain the RNA

secondary structure required for post-transcriptional processing and are more species specific than geographically influenced. Several common structural folds were shared among the selected lepidopteran insects for maintaining functional equivalents. Identifying the homologous regions and reconstructing their evolution increases the traits available for the phylogenetic analysis. Construction of an evolutionary tree using more isolates of Lepidopteron will provide an understanding for their functional selection. The present study indicates the phylogenetic relationship among the selected Lepidopteron species belonging to diverse geographical locations.

## CORRESPONDENCE TO

Prof. Poduri Nagaraja Rao, Department of Zoology, Osmania University College for Women, Koti, Hyderabad, Andhra Pradesh, India. e-mail: nagarajaraop@yahoo.com

## References

1. Metcalf, C.L. and Flint, W.P.1962. Destructive and useful insects their habits and control. Tata Megraw Hill, New Delhi.

2. Nair, M.R.G.K 1986: Insects and mites of crops in India. ICAR, New Delhi.

3. David Dent 1991: Insect pest management. C.A.B. International, U.K.

4. Srivastava, K.P 2000: A text book of Applied entomology, Kalyani publishers, Ludhiana, India

5. Van der Sande CA, Kwa M, van Nues RW, van Heerikhuizen H, Raue HA, Planta RJ.. Functional analysis of internal transcribed spacer 2 of Saccharomyces cerevisiae ribosomal DNAJ. Mol. Biol, 1992, 223: 899-910.

6. Michot B, Hassouna N, Bachellerie JP.. Secondary structure of mouse 28S rRNA and general model for the folding of the large rRNA in eukaryotes.Nucleic Acids Res, 1984, 12: 4259-4279.

7. Shinohara ML, LoBuglio KF, Rogers SO. Comparison of ribosomal DNA ITS regions among geographic isolates of Cenococcum geophilum, Current genetics, 1999, 35:527-535.

8. Crabtree MB, Savage HM, Miller BR. Development of a species-diagnostic polymerase chain reaction assay for the identification of Culex vectors of St. Louis encephalitis virus based on interspecies sequence variation in ribosomal DNAspacers.Am J Trop Med Hyg. 1995, 53, 1:105-9.

9. Atanas I. Lalev, Ross N. Nazar. Structural equivalence in the transcribed spacers of pre-rRNA transcripts in Schizosaccharomyces pombe. Nucleic Acids Research, 1999, 27, 13 3071-3078.

10. Perry RP. Processing of RNA. Annu. Rev. Biochem, 1976, 45, 605-629.

11. Musters W, Boon K, van der Sande CA, van Heerikhuizen H, Planta RJ.. Functional analysis of transcribed spacers of yeast ribosomal DNA. EMBO J, 1990, 9: 3989-3996.

12. Michot B, Bachellerie JP, Raynal F. Structure of mouse rRNA precursors. Complete sequence and potential folding of the spacer regions between 18S and 28S rRNA.Nucleic Acids Res, 1983, 11, 3375-3391.

13. Furlong JC, Maden BE. Patterns of major divergence between the internal transcribed spacers of ribosomal DNA in Xenopus borealis and Xenopus laevis, and of minimal divergence within ribosomal coding regions. EMBO J, 1983, 2, 443-448.

14. Michot B, Bachellerie JP.. Secondary structure of the 5' external transcribed spacer of vertebrate pre-rRNA. Presence of phylogenetically conserved features. Eur. J. Biochem, 1991, 195, 601-609.

15. Nusha K et al, Molecular Ecology notes. 2004, 4: 20-22.

16. Hofacker, I. L., W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker, and P. Schuster:, 'Fast Folding and Comparison of RNA Secondary Structures'. Monatsh. Chem. 1994,125, 167--188.

17. Zuker M, Stiegler P., Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucl. Acid. Res. 1981, 9: 133-148.

18. Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Nat. Acad. Sci 1986, 83, 9373-0377.

19. Jaeger JA, Turner DH, Zuker M. Improved predictions of secondary structures for RNA. Proc Natl Acad Sci U S A. 1989 Oct; 86(20):7706-10.

20. Ding Y, Chan CY, Lawrence CE. Sfold web server for statistical folding and rational design of nucleic acids. Nucleic Acids Res. 2004, 32 Web Server issue, W135-W141.

21. Christoffersen, McSwiggen & Konings 1994, J. Mol. Structure (Theochem) 311, p. 208.

22. Ding Y, Lawrence CE. Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond. .Nucleic Acids Res. 2001, 29, 1034-1046.

23. Brodsky L.I., Vasiliev A.V., Kalaidzidis Ya.L., Osipov Yu.S., Tatuzov R.L., Feranchuk S.I. GeneBee: the program package for biopolymer structure analysis, 1992, Dimacs, 8, 127-139.

24. Levinson G, Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol.bio.evolve, 1987, 4:203-221.

25. Severini C, Silvestrini F, Mancini P, La Rosa G, Marinucci M. Sequence and secondary structure of the rDNA second internal transcribed spacer in the sibling species Culex pipiens L. and Cx. quinquefasciatus Say (Diptera: Culicidae). .Insect Mol. Biology 1996, 5 (3): 181-186.

26. Wesson DM, Porter CH, Collins FH. Sequence and secondary structure comparisons of ITS rDNA in mosquitoes (Diptera: Culicidae).Molecular phylogenetics and evolution, 1992, 1: 253-2699.

27. Barker S.C., Distinguishing species and populations of rhipicephaline ticks with its 2 ribosomal RNA. J Parasitol, 1998, 84:887-892.

28. Fernandes O, Santos S, Junqueira A, Jansen A, Cupolillo E, Campbell D, Zingales B, Coura JR. Populational heterogeneity of Brazilian Trypanosoma cruzi isolates revealed by the mini-exon and ribosomal spacers. Mem. Inst. Oswaldo Cruz, 1999, 94:195-197.

29. Van der Sande CA, Kwa M, van Nues RW, van Heerikhuizen H, Raue HA, Planta RJ.. Functional analysis of internal transcribed spacer 2 of Saccharomyces cerevisiae ribosomal DNA. Journal of molecular biology, 1992, 223: 899-910.

## Author Information

**Murali R. Kuracha, M.Sc.**
Department of Zoology, Osmania University College for Women

**Bhargavi Rayavarapu, M.Sc.**
Department of Zoology, Osmania University College for Women

**Satya D.V.R. Kumar, Ph.D.**
University of Maryland

**Poduri Nagaraja Rao, Ph.D.**
Department of Zoology, Osmania University College for Women