International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

A Combination Study in Some Members of Monocotylidae (Monogenea) in Molecular Phylogeny Employing 28SrRNA along with Geographical Distribution

Fozail Ahmad[1], D. Singh[2], P. V. Arya[3]

Department of Zoology, Dyal Singh College (University of Delhi), Lodhi Road, New Delhi, 110003

Abstract: ***Evolution and interrelatedness among the same group of organism have been explored and debated in many ways. Traditional tools of phylogenetic investigation presented preferred scene for a considerable time. The stage was taken by the incorporation of molecular tools in later stages. Considering a combination of host specific parasite may be more informative in terms of investigating evolutionary lineage of an organism. Recently attempts have been made to incorporate secondary rRNA structure for consideration of phylogenetic studies. Present investigation is an attempt to investigate a family i.e., Monocotylidae (Class: Monogenea) for the same. Cues from geographical distribution and molecular tools have been employed in the study. The finding supports that host play substantial role in the formation of new species. Species distribution strengthened intra genus relationship, divergence and migration over period of times.***

Keywords: Zoogeographical distribution, Monocotylidae, Speciation, 28S rRNA

1. Introduction

Monocotylidae a family of monogena with more than 100 species, parasitic on chondrichthyan fishes (skin, gills, nasal cavities), the phylogeny of the family attempted on morphological and molecular basis (Chisholm et al. 1995; Chisholm et al. 2001 and Glennon et al. 2006). Finding the phylogenetic relationship (clade and cluster) among species of different genus and/or species from a particular family with their zoogeographic distribution may present evolutionary clue for diversity and speciation (Vaillant et al. 2013). Evaluation of zoogeographical distribution (Arya & Singh 2015) together with molecular clue may present evolutionary history including probable origin of the organisms (Rogers 2007; Brumfield & Edwards 2007 and Fozail et al., 2015a & b). Monogenean parasites have utilized for indirectly study of their host, zoogeographical diversity, distribution, migration and settlement over period of time(Arya & Singh 2015)(Mendlová et al. 2012)(Šimková & Morand 2008). Monocotylidae offers a broader range for evolution and zoogeographical distribution on account of multiple sites onto host (Leslie et al., 2001 and Fehlauer-Ale & Littlewood 2011). Their exposure to various sites on the same host, may be accounted for them to have special genetic compositions in order to face the different protective sites developed by chondrichthyan fishes, which is also necessary for their survival in the varying environment (Fels & Kaltz 2006). Measurement of structural parameter of 28S rRNA parameters (bond energy, geometrical features, base composition etc.) and its comparison is proved as the best methods to study molecular phylogeny and correlation with zoogeographical distribution (Tuplin et al. 2002). Phylogenetic characters of rRNA basically include bulges, loops, helices and separation of single strands as they have been conserved throughout the evolution (Lescoute 2005). Secondary structure of ribosomal RNA provides substantial information regarding evolutionary relationship that cannot be simply inferred from cladistic analyses using simple RNA sequences(Keller et al. 2010; Chaudhary & Singh 2013 and Fozail et al., 2015a & b). RNA also provides necessary information regarding the development of biomarker of individual species(Gilad et al. 2008).

Present work is an attempt of utilizing 28S rRNA, secondary structure and zoogeographical distribution reports of the parasite to investigate phylogenetic relationship along with probable pattern of speciation.

2. Materials and Methods

Genus & Species Selection-
Total 39 species from 12 genus of this family were selected based upon the availability of their 28S rRNA in NCBI, their host, distribution and environment were confirmed from literature (Table-1).

Multiple Sequence Alignment (MSA) by ClustalW- 28S rRNA sequences were aligned by ClustalW multiple alignment (Thompson *et al.* 1994) with default settings. Sequence alignment in MEGA 6 and phylogenetic tree prepared using NJ method (Figure -5)

Molecular Phylogenetic Analysis- The Kimura-2 parameter model to estimate distances for correcting the transition bias. Most parsimonious tree was secured using the close-neighbor-interchange algorithm. Bootstrap 1000 replications for every species. Subsequently, MSA were exported as part of the result (Fig: 4a-e). This was mainly exercised for analyzing the genus divergence, speciation and average similarity among species.

Inferring Secondary Structure of 28S rRNAs- Based upon the best alignment score of the sequences in each cluster, aligned using ClustalW. The inference of the secondary structure using Mfold (http://mfold.rna.albany.edu), at a fixed temperature of 37^0 C, structure was analyzed for bulges, stems, loops and negative free energy (ΔG). Every

cluster had been associated with its common rRNA averaging, evolutionary phenomenon.

Geo mapping- All the selected species (Table-1) were marked on simple world map manually (Figure -5) for the global scenario of the species relatedness and diversity. Later on joined with reference to their respective cluster for inferring molecular relatedness globally.

3. Result

Construction of Phylogenetic Tree- Phylogenetic tree forming seven clusters, consisting of two or more than two species showing evolutionary cross relationship. In the tree, Cluster A - G had 9, 2, 8, 3, 4, 2 and 9 species respectively. Cluster A with three clusters exhibited very poor bootstrap value having drastic difference among the three clusters, indicating a huge fluctuation in the event of speciation. In all, two clusters belong to the same genus and represent significant relationship and overall variations among four different genus.

Table 1: Genus with respective species (family Monocotylidae) investigated in the study

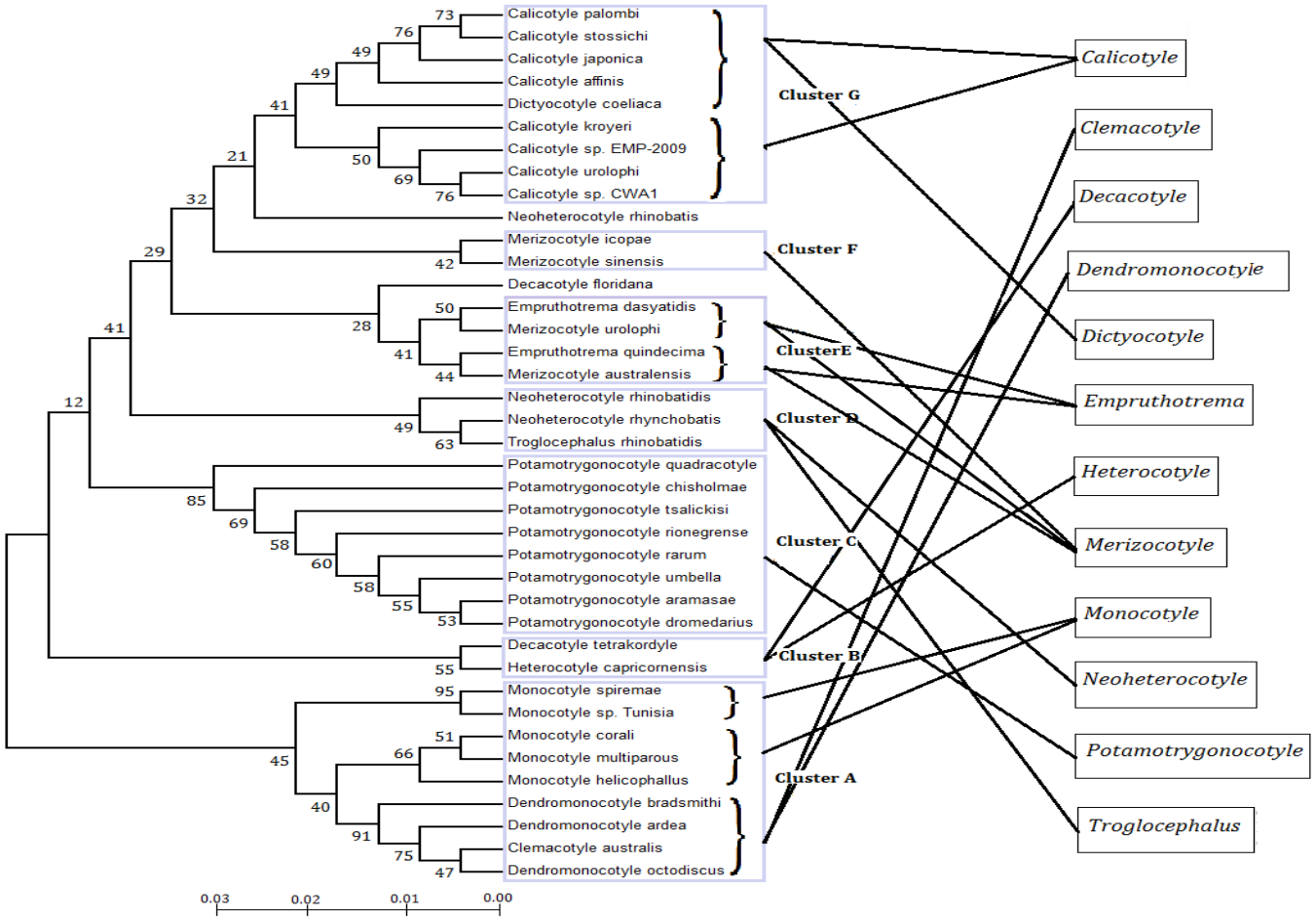
Sl.	*Genus*	*Species*	*Host*	*Envir-onment*	*Country/Area*	*Accession ID*
1	*Calicotyle*	*C. affinis* Scott, 1911	*Chimaera monstrosa*	M	N. A. Ocean	AF382061
		C. japonica Diesing, 1850	*Squalus mitsukurii*	M	Japan	AB485996
		C. kroyeri Diesing, 1850	*Anacanthobatis folirostris*	M	Mexico	AF279748
		C. palombi Euzet & William, 1960	*Mustelus mustelus*	M	N. A. Ocean	AF131709
		C. stossichi Braun, 1899	*Mustelus norrisi*	M	Mexico	AF279751
		C. urolophi Chisholm et al, 1991	*Urolophus spp.*	M	Australia	AF279752
		C. sp. CWA1 Chisholm et al, 2000	-	-	-	AF279750
		C. sp. EMP Perkins et al, 2009	-	-	-	FJ971978
2	*Clemacotyle*	*C. australis* Young, 1967	*Aetobatus narinar*	M	Australia	AF348350
3	*Decacotyle*	*D. floridana* Chisholm et al, 1998	*Aetobatus narinari*	M	Mexico	AF348357
		D. tetrakordyle Chisholm & Whittington, 1998	*Taeniura lymma*	M	Australia	AF348358
4	*Dendrocotyle*	*D. ardea* Chisholm & Whittington, 1998	*Pastinachus sephen*	M	Australia	AF348351
		D. bradsmithi Macleay, 1881	*Myliobatis australis*	F	Australia	FJ971986
		D. octodiscus Hargis, 1955	*Dasyatis americana*	M	N. A. Ocean	AF348352
5	*Dictyocotyle*	*D. coeliaca* Nybelin, 1941	*Raja naevus*	M	N. A Ocean	AY157171
6	*Empruthotrema*	*E. dasyatidis* Whittington & Kearn, 1992	*Dasyatis fluviorum*	F	Australia	AF348345
		E. quindecima Chisholm & Whittington, 1999	*Taeniura lymma*	M	Australia	AF348346
7	*Heterocotyle*	*H. capricornensis* Chisholm & Whittington, 1996	*Himantura fai*	M	Australia	AF348360
8	*Merizocotyle*	*M. australensis* Beverley-Burton and Williams, 1989	*Himantura fai*	M	Australia	AF348348
		M. icopae Beverley-Burton & Williams, 1989	*Rhinobatos typus*	M	Australia	AF348349
		M. sinensis Timofeeva, 1984		M	Taiwan	FJ514075
		M. urolophi Chisholm & Whittington, 1999	*Urolophus paucimaculatus*	M	Tasmania	AF348347
9	*Monocotyle*	*M. corali* Chisholm, 1998	*Pastinachus sephen*	M	Australia	AF348353
		M. helicophallus Beverley-Burton & Williams, 1990	-	M	Australia	AF348355
		M. multiparous Beverley-Burton & Williams, 1990	*Himantura uarnak*	M	Australia	AF348356
		M. spiremae Beverley-Burton & Williams, 1990	*Himantura uarnak*	M	Australia	AF348354
		M. sp. Tunisia Beverley-Burton & Williams, 1990	-	M	Tunisia	AF387511
10	*Neoheterocotyle*	*N. rhinobatidis* Young, 1967	*Rhinobatos typus*	M	Australia	AF026107
		N. rhinobatis Pillai & Pillai, 1976	*Rhinobatos typus*	M	Australia	AF348362
		N. rhynchobatis Tripathi, 1959	*Rhinobatos typus*	M	Australia	AF348363
11	*Potamotrygonocotyle*	*P. aramasae* Tripathi, 1959	*Paratrygon aiereba*	F	Brazil	JN379514
		P. chisholmae Mayes et al, 1981	*Potamotrygon motoro*	F	River basin (USA)	JN379519
		P. dromedarius Mayes et al, 1981	*Potamotrygon hystrix*	F	Brazil	JN379518
		P. quadracotyle Mayes et al, 1981	-	F	Brazil	FJ755807
		P. rarum Mayes et al, 1981	*Potamotrygon schroederi*	F	Brazil	FJ755809
		P. rionegrense Mayes et al, 1981	*Potamotrygon cf. motoro*	F	Brazil	FJ755810
		P. tsalickisi Mayes et al, 1981	*potamotrygonid*	F	River basin (USA)	JN379513
		P. umbella Mayes et al, 1981	*Potamotrygon*	F	Brazil	FJ755808
12	*Troglocephalus*	*T. rhinobatidis* Young, 1967	*Rhinobatos typus*	F	Australia	AF348364

Note : M= Marine; F : freshwater. All the sequences for the present study was taken from NCBI database. Acknowledgement is due to all the contributors.

Second cluster (B) having species from two different genus and represented 55% of considerable bootstrap value. Cluster C reflected a constant bootstrap value (53-59 %) for all species except *Potamotrygonocotyle quadracotyle* in the cluster, all species belong same genus. The bootstrap value of the cluster could be average to 55%. Cluster D with species from two different genus and considerable bootstrap values. Cluster E had two clusters from two different genus. Cluster F all species from the same genus not connected by considerable bootstrap values and Cluster G forming two clusters, one with species from same genus whereas, other with one species from the genus *Dictocotyle*. Both clusters with significant bootstrap (70%). The phylogenetic tree of selected species from different genus shown the level of similarities, point of deviation and time of speciation between genus that is in the same cluster. Two species from two different genus might have evolved from same ancestor regardless of any high bootstrap value. Cluster in the tree represented that a group of species might have evolved from a common ancestor, thousand years ago. This result is further supported in the subsequent sections of the result.

Secondary Structure Analysis- The inferred secondary structure of 28S rRNA by Mfold of representative species from seven cluster exhibited the evolutionary distinction between species and clusters as well (Figure -2) also provided the stability of molecules in terms of negative free energy (ΔG). Representative species were selected by multiple sequence alignment of species from each cluster, and the most conserved sequence of the species was considered based on alignment score given by ClustalW. Formation of secondary structure is characterized by the bulge loops, interior loops and hairpin loops conferred by negative free energy of molecule. Higher the negative free energy (ΔG), more stable the molecule. Negative free energy of cluster A - G (rRNA from species) were predicted to be -261 kcal/mol, -229.3 kcal/mol, -260.2 kcal/mol, -264.3 kcal/mol, -247.1 kcal/mol, -244.2 kcal/mol and 218.1 kcal/mol respectively (Figure -3). The negative free energies of cluster A, C and D fall in the range of -260 kcal/mol and discrete by -2.5 kcal/mol approximately, representing that species from both groups had followed similar pattern of evolution. Anomaly to this observation can be accounted as the varying number of different loops directly affects stability of molecule. Cluster E and cluster F had an average negative free energy of -245.5 kcal/mol (discrete by approximately ΔG =-3.9 kcal/mol), shown to be correlating each other and representing evolutionary relatedness. In case of cluster B and G, ΔG was highly discrete by -11kcal/mol, signifying a distant re

Three types of loops are formed in the secondary structure of RNA molecule (cluster/representative species) with unique pattern of occurrence (Figure -2). The formation of loops, as mentioned earlier, is almost conferred by negative free energy, resolves stability and constancy of the entire molecule.



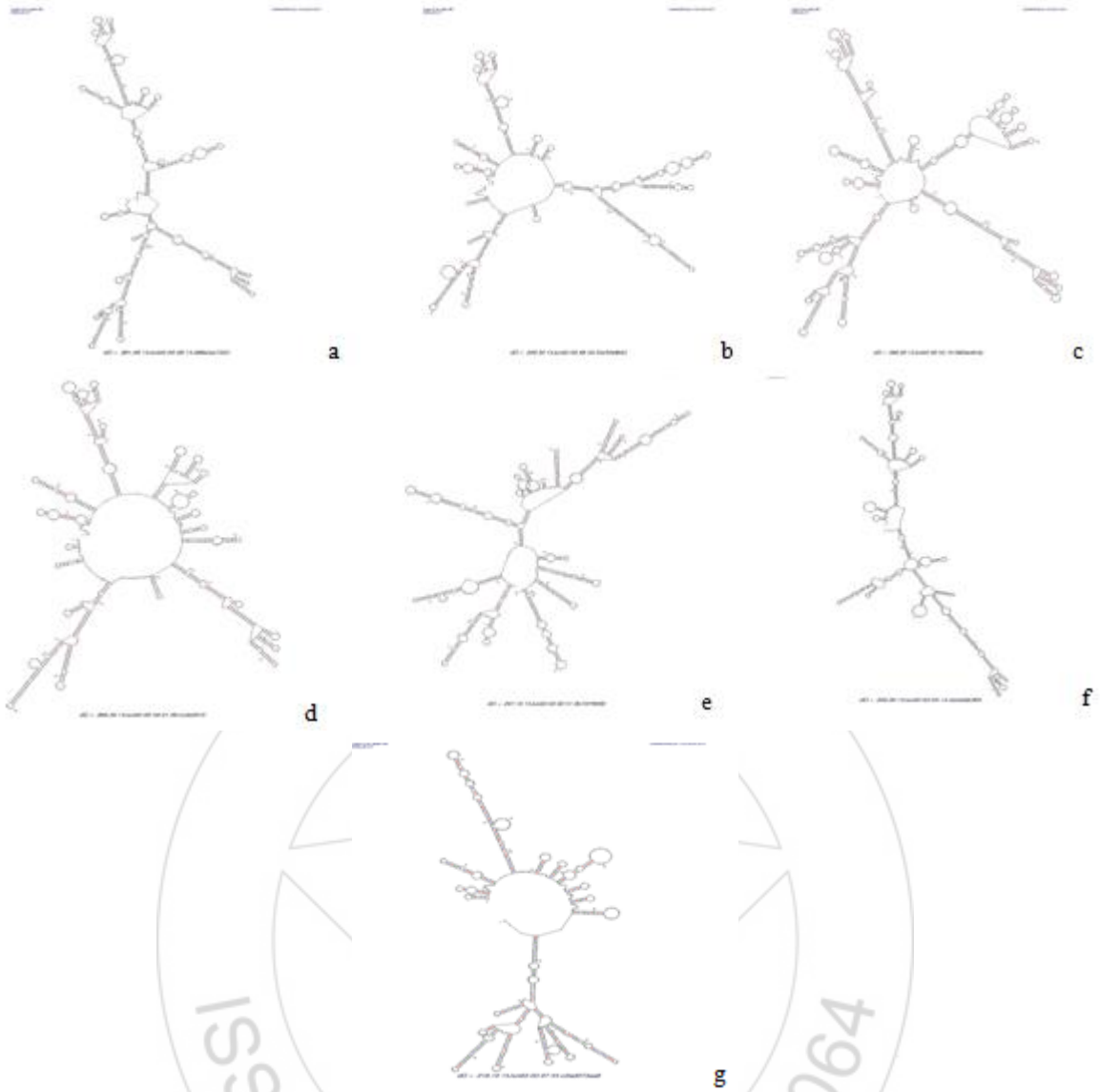


Figure 2: Secondary structure of 28S rRNA of 39 species represented by 12 genus clustered in 07 clusters family Monocotylidae

In present analysis, the sum of interior loops and hairpin loops is equal, although, their number varied for individual clusters (Figure-3). Number of bulge loops (3-6) was found to be least for all clusters. Cluster D with highest negative free energy (Figure-4) represented 40 loops, second most in number. Cluster A with the second highest **ΔG** developed a total of 35 loops that did not seem to coincide with its negative free energy (-261 kcal/mol) which should have been, thermodynamically, second most of all, mainly due to specific pattern and number of nitrogenous bases participated in forming the loops.

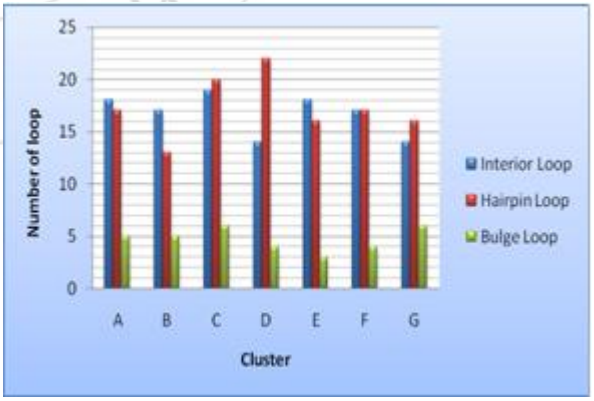


Figure 3: Number of loops from of respective RNA secondary structure for each cluster

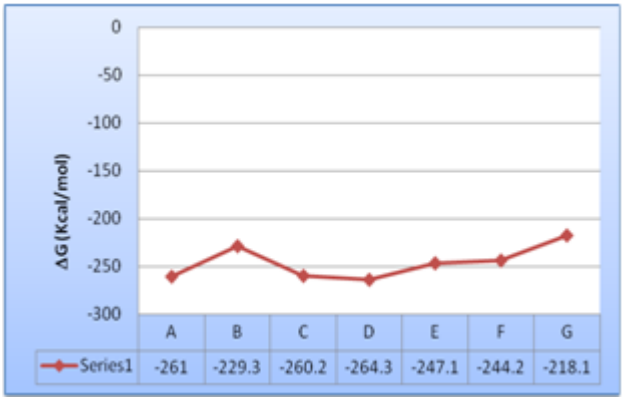


Figure 4: Negative free energy of the representative RNA structure for each cluster

Cluster C surprisingly had the greatest number (45) of loops in total. Only cluster A tend to deviate in terms of number of loop, otherwise negative free energy was peered for the three related clusters relating that species belonging to different genus had a specific pattern of evolution and later on distributed in different regions. Cluster E and F with almost equal amount of ΔG developed equal number of loops, representing strong ancestral relatedness among species. Cluster B and G were discrete by a great energy difference of -11 kcal/mol but number of loops were discrete by one loop, that could not be accounted for concern. Hence, it can be concluded that these two groups were descended from distinct ancestral lineage. Comparatively, three types of loops represented uniqueness, stability, conservation pattern, evolutionary relatedness and range of ancestral lineage.

Multiple Sequence Alignment Analysis

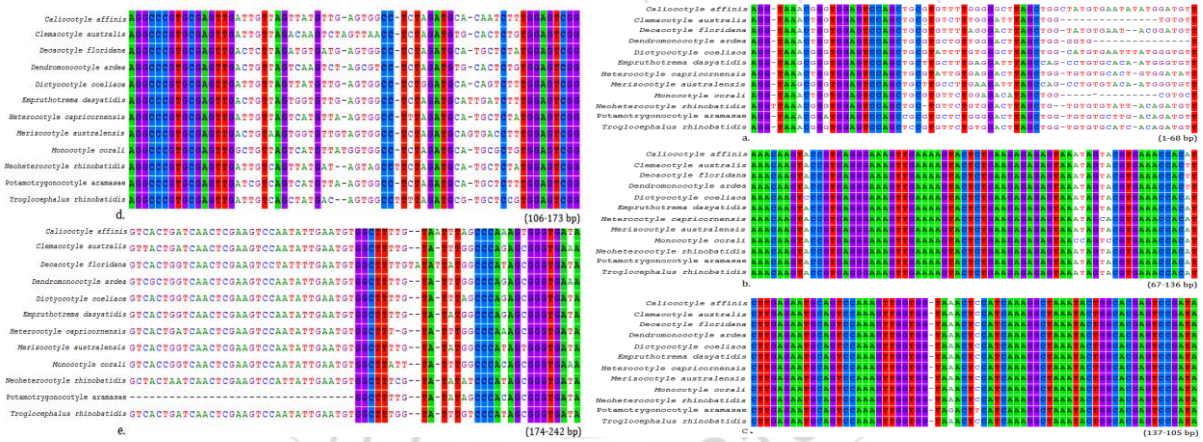
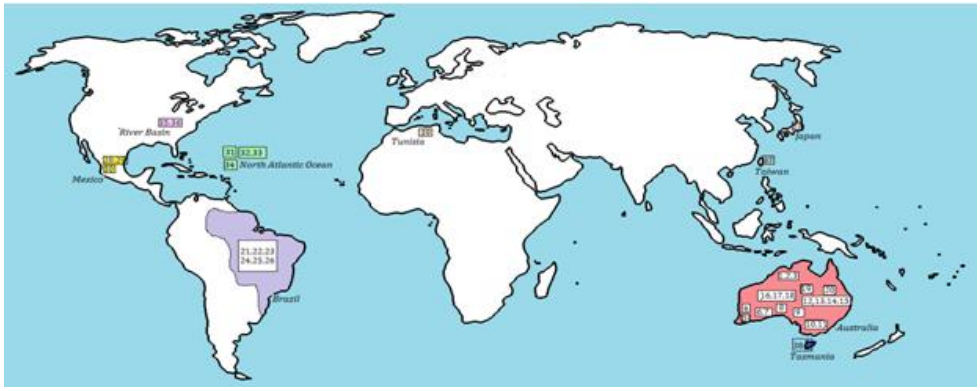


Figure 5: ClustalW alignment of 12 sequences of 28S ribosomal RNAs.

Conserved bases were shown to be highlighted with four different colors (Figure-5a- e). Out of 242 base pairs, 125 base pairs were found mismatching along with gaps (insertion & deletion). Figure-5b with bp 67-136 showed better alignment wherein six mismatches were maintained & Figure-5c with bp 137-105 had one indel and two mismatches. Figure-a, d & e represented high level of mutation over alignment. These simple alignment results provided an overview on genus variability and divergence for speciation.

Geo Mapping



4. Discussion

The molecular distinction among species from different genus provides an understanding over evolutionary process and expressing ancestral lineage to the origin of a new species. RNA in the folded form show paired and unpaired (loops) bases (Chen et al. 2005). Qualitatively, bases which are bonded tend to stabilize molecule due to higher negative free energy whereas unpaired bases tend to destabilize the molecule due to lesser negative free energy (Greatorex et al. 2002 and Geisberg et al. 2014). Quantitatively, loop that are more in number destabilize the secondary structure because they require more positive free energy (Trotta 2014). Thus, cluster A, C and D are the most stable and Cluster B, E, F and G are less stable structure, signifying that organisms belonging to the particular cluster would be of equal stability in terms of negative free energy of their RNA molecules, and therefore, could follow the same pattern of origin (Sun & Caetano-Anollés 2008). From first to seventh cluster, each organism representing its own cluster, exhibiting distinctions in term of number of neighbor/sister clade and 28S rRNA secondary structure. Although negative free energy and number of loops showed noticeable variations within all clusters with an established correlation between the two parameters. Except cluster B and G, remaining five clusters (cluster A, C & D) and (cluster E and F) represent equal stability, conservation pattern and sympatric speciation events (Figure-4). This was further supported by equal number of loops developed in the representative molecule. Cluster B and G with their respective higher and lower number of loops and negative free energies were not coinciding with other clusters in number of loops and ΔG. Because each group of organisms possess the particular pattern of evolution for ribosomal RNA. The differences among clusters about ΔG were mainly accounted due to the size of loops. Loops more in number but smaller in size are formed with less negative free energy whereas loops less in number but larger in size require more negative free energies (Katz 2003). Evidently, both, size and number of loops are accounted for estimating out the stability of a molecule (Zhang et al. 2008). The pattern of evolution and relatedness among species is reflected by the development of loops and their sizes which in turn account for the overall stability of RNA (Wongsurawat et al. 2012). Evolution, most of the time, rises the level of complexity that is strictly coincided with the necessities of situations. RNA having more complex secondary structure presenting with more loops and small size whereas molecule with lesser loops and large sizes show lower level of complexity (Gevertz 2005). Cladistic analysis corroborated that even after great speciation events, molecular information were maintained by species as two different species from two different genus represented the cladistic relationship and had fallen in the same cluster (Figure -1). Although, they were distributed in different geographical zones (Figure -6) but represented a particular group in the same family from different genus, indicating a common ancestral lineage. So their evolutionary history can be traced back to common points. A major reason is observed about species richness in the family Monocotylidae and why cladistic relationship represents inter genus similarity. The one reason behind speciation in Monocotylidae is that whenever infecting sites are changed or switched (*e.g.* from gill to inner wall of the body cavity), genetic composition automatically gets changed, determined by the extent of parasitism and resistance of host (Fels & Kaltz 2006 and Millanes et al. 2014). This creates a major molecular change followed by physiological variations, waved into conserved domain of nucleic acids (Thompson et al. 2001). Over the period of time, the developing variation is stacked and then a time reaches when the molecularly distinct species appear with novel feature and said to follow a new route for a different lineage (Nancy and Moran 1998).

The finding paves way to a hypothesis that host plays substantial role in the formation of new species especially for monogenetic parasites. Cladistic analysis giving strong clues about ancient lineage, origin and range of similarity was comprehended by secondary structure of 28S rRNA. Species distribution strengthened intra genus relationship, divergence, and migration over period of times. In the phylogenetic tree, clustering and cladistic hypothesis was supported by zoogeographical distribution of Monocotalidae in different regions of the world.

5. Acknowledgment

Authors are thankful to the authorities of UGC for financial support (F. No.: 41-34/2012 (SR)) and head of the institution for providing necessary facilities.

References

[1] Arya, P.V. & Singh, H.S., 2015. Zoogeographical distribution an evidence of co-evolution : Freshwater fish parasitic monogenean Genus Mizelleus & its host diversity. *IJSR - International Journal Of Scientific Research*, 4(6), pp.589–590.

[2] Brumfield, R.T. & Edwards, S. V., 2007. Evolution into and out of the andes: a bayesian analysis of historical diversification in *thamnophilus* antshrikes: evolution into and out of the andes. *Evolution*, 61(2), pp.346–367.

[3] Chaudhary, A. & Singh, H.S., 2013. Secondary structure and phylogenetic utility of the ribosomal large subunit (28S) in monogeneans of the genus *Thaparocleidus* and *Bifurcohaptor* (Monogenea: Dactylogyridae). *Journal of parasitic diseases : official organ of the Indian Society for Parasitology*, 37(1), pp.74–83.

[4] Chen, G., Snosko, B.M., Keneddy, S.D., Krugh, T.R. & Turner, D.H., 2005. Solution structure of an RNA internal loop with three consecutive sheared GA pairs. *Biochemistry*, 44(8), pp.2845–2856.

[5] Chisholm, L. A., Wheeler, T. A. & Beverley-Burton, M., 1995. A phylogenetic analysis and revised classification of the Monocotylidae Taschenberg, 1879 (Monogenea). *Systematic Parasitology*, 32(3), pp.159–191.

[6] Chisholm, L.A. et al., 2001. Phylogenetic analysis of the Monocotylidae (Monogenea) inferred from 28S rDNA sequences. *International journal for parasitology*, 31(11), pp.1253–1263.

[7] Fehlauer-Ale, K.H. & Littlewood, D.T.J., 2011. Molecular phylogeny of Potamotrygonocotyle (Monogenea, Monocotylidae) challenges the validity of some of its species: Phylogeny of

Potamotrygonocotyle. *Zoologica Scripta*, 40(6), pp.638–658.

[8] Fels, D. & Kaltz, O., 2006. Temperature-dependent transmission and latency of *Holospora undulata*, a micronucleus-specific parasite of the ciliate *Paramecium caudatum*. *Proceedings of the Royal Society B: Biological Sciences*, 273(1589), pp.1031–1038.

[9] Fozail, A., Singh, D. & Arya, P.V., 2015a. In silico phylogenetic studies on some members of parasitic genus *Gyrodactylus* (Monogenea: Gyrodactylidae) for assessment of evolutionary relatedness inferred from 28S ribosomal RNA and geomapping the sample. *International Journal of Recent Scientific Research;* 6(7), pp.4970-4977.

[10] Fozail, A., Singh, D. & Arya, P.V., 2015a. Comparative evaluation of speciation and zoogeographical distribution for *Lamellodiscus* (Monogenea: Diplectanidae) using 18S rRNA. *International Journal of Innovation Science and Research (IJISR)*; 4 (6), pp.235-241.

[11] Geisberg, J.V., Zarmik, M., Xiochun, F., Fatih, O. & Kevin, S., 2014. Global Analysis of mRNA Isoform Half-Lives Reveals Stabilizing and Destabilizing Elements in Yeast. *Cell*, 156(4), pp.812–824.

[12] Gevertz, J., 2005. In vitro RNA random pools are not structurally diverse: A computational analysis. *RNA*, 11(6), pp.853–863.

[13] Gilad, S., Eti, M., Yariv, V., Sima, B., Danit, L., Noga, Y., Hila, B., Michal, K., Hila, C. & Nir, M., 2008. Serum MicroRNAs Are Promising Novel Biomarkers S. Williams, ed. *PLoS ONE*, 3(9), pp.e3148.

[14] Glennon, V., Chisholm, L.A. & Whittington, I.D., 2006. Three unrelated species, 3 sites, same host--monogenean parasites of the southern fiddler ray, *Trygonorrhina fasciata*, in South Australia: egg hatching strategies and larval behaviour. *Parasitology*, 133(Pt 1), pp.55–66.

[15] Greatorex, J., Chisholm, L.A. & Whittington, I.D., 2006. Structure and stability of wild-type and mutant RNA internal loops from the SL-1 domain of the HIV-1 packaging signal. *Journal of Molecular Biology*, 322(3), pp.543–557.

[16] Katz, L., 2003. Widespread Selection for Local RNA Secondary Structure in Coding Regions of Bacterial Genes. *Genome Research*, 13(9), pp.2042–2051.

[17] Keller, A., Frank, F., Tobias, M., Thomas, D., Jorg, S. & Matthias, W., 2010. Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. *Biology Direct*, 5, pp.4-6.

[18] Lescoute, A., 2005. Recurrent structural RNA motifs, Isostericity Matrices and sequence alignments. *Nucleic Acids Research*, 33(8), pp.2395–2409.

[19] Leslie, A.C., Ian, D.W., Jess, A.T. & Robert, D.A., 2001. Phylogenetic analysis of the Monocotylidae (Monogenea) inferred from 28S rDNA sequences. , 31: 1537–1547.

[20] Nancy, A. Moran, J.D.L., 1998. Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria. *PNAS*, 95: 4458–4462.

[21] Mendlová, M., Yves, D., Kristina, C., Antony, P., Andrea, S. & Filatov, D.A., 2012. Monogeneans of West African Cichlid Fish: Evolution and Cophylogenetic Interactions D. A. Filatov, ed. *PLoS ONE*, 7(5), pp.e37268.

[22] Millanes, A.M., Camille, T., Martin W., Paul, D. & Mats, W., 2014. Host switching promotes diversity in host-specialized mycoparasitic fungi: uncoupled evolution in the *Biatoropsis-usnea* system. *Evolution; International Journal of Organic Evolution*, 68(6), pp.1576–1593.

[23] Rogers, A.D., 2007. Evolution and biodiversity of Antarctic organisms: a molecular perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1488), pp.2191–2214.

[24] Šimková, A. & Morand, S., 2008. Co-evolutionary patterns in congeneric monogeneans: a review of *Dactylogyrus* species and their cyprinid hosts. *Journal of Fish Biology*, 73(9), pp.2210–2227.

[25] Sun, F.-J. & Caetano-Anollés, G., 2008. Evolutionary Patterns in the Sequence and Structure of Transfer RNA: Early Origins of Archaea and Viruses P. E. Bourne, ed. *PLoS Computational Biology*, 4(3), p.e1000018.

[26] Thompson, J., Kim, D.F., O'Connor, M., Lieberman, K.R., Bayfield, M.A., Gregory, S.T., Green, R., Noller, H.F. & Dahlberg, A.E., 2001. Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proceedings of the National Academy of Sciences*, 98(16), pp.9002–9007

[27] Trotta, E., 2014. On the Normalization of the Minimum Free Energy of RNAs by Sequence Length D. Barash, ed. *PLoS ONE*, 9(11), pp.e113380.

[28] Tuplin, A., Jonny, W., Evans, D.J. & Peter, S., 2002. Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus. *RNA (New York, N.Y.)*, 8(6), pp.824–841.

[29] Vaillant, J.J., Bock, D.G. & Cristescu, M.E., 2013. Speciation patterns and processes in the zooplankton of the ancient lakes of Sulawesi Island, Indonesia. *Ecology and Evolution*, 3(9), pp.3083–3094.

[30] Wongsurawat, T., Jenjaroenpun, P., Kwoh, C.K. & Kuznetsov, K., 2012. Quantitative model of R-loop forming structures reveals a novel level of RNA-DNA interactome complexity. *Nucleic Acids Research*, 40(2), pp.e16–e16.

[31] Zhang, J., Ming, L., Rong, C., Wei, W. & Jie, L., 2008. Discrete state model and accurate estimation of loop entropy of RNA secondary structures. *The Journal of Chemical Physics*, 128(12), pp.125107.