International Journal of Innovation Sciences and Research
Vol.4, No, 10, pp.494-500, October- 2015
Available online at
http://www.ijisr.com

RESEARCH ARTICLE

COX-1 STUDIES IN EVALUATION AND ASSESSMENT OF MOLECULAR DIVERSITY AMONG GYRODACTYLIDAE, DIPLECTENIDAE, DIPLOZOIDAE AND DICTILOPHORIDAE FAMILIES (CLASS: MONOGENEA)

[1]Fozail Ahmad, [1,*]Priya Vrat Arya and [2]Singh, H.S.

[1]Department of Zoology, Dyal Singh College, University of Delhi, New Delhi-110003, India
[2]Department of Zoology, Ch. C. S. University, Meerut-250004, India

Accepted 16th September, 2015; Published Online 30th October, 2015

ABSTRACT

Proteins being mirror to molecular signature of an organism, their potential in assessment of molecular diversity may be useful. Comparison among the organisms or on a larger scale of family may further be taken to give insight on the molecular journey of the organisms. Present paper deals with the study of COX-1 in four families *viz*., Gyrodactylidae, Diplectenidae, Diplozoidae and Dictilophoridae (Class : Monogenea) using structural and other significant parameters. In all 16 species have been extensively studied across four families. Results reflecting peculiar diversity on molecular level suggesting divergence based evolution in the form of molecular molding.

Key Words: Cytochrome C Oxidase, Monogenea, Secondary Structure, Evolution.

INTRODUCTION

Proteins are more conserved than nucleic acids during evolution, providing strong platform to study conserved aspects of their structure as well as function (Butland *et al*., 2005; Socolich *et al*., 2005; Sicheritz-Ponten, 2001). Among them Cytochrome C Oxidase is one of most conserved protein and oldest one on the earth (Sicheritz-Ponté *et al*., 1998; Castresana *et al*., 1994). Cytochrome oxidase reduces oxygen to water making it essential enzyme for aerobic metabolism (Collman *et al*., 2007; Ekici *et al*., 2014). It creates a proton gradient as an intermediate step in the conversion of redox energy to ATP (Rottenberg, 1998). The enzyme complex of the electron transport chain with 13 subunits is of mixed genetic origin (Li *et al*., 2006). The three largest subunits (I-III) are encoded by mitochondrial genomes (Breek *et al*., 1997) and carry out known catalytic functions of the enzyme and show homology between eukaryotes and prokaryotes (Steffens *et al*., 1987; Smits *et al*., 2007). Other 10 subunits encoded by nuclear genome (Lenka *et al*., 1998; Wolz *et al*., 1997). The mixed origins of COX give challenge of study the evolutionary relatedness of two distinct genetic systems (Wu *et al*., 2000). COX-I, the largest subunit of the holoenzyme is important in enzyme function and only subunit conserved in all heme–copper oxidases from prokaryotes to eukaryotes (Soto *et al*., 2012). It is incorporated into the mitochondrial inner membrane, containing 12 transmembrane helices and three redox centers, heme-a, heme-a3, and CuB (Clemente *et al*., 2013). Evolution in terms of classification and placing monogeneans help integrate the large group to identify proper position in taxonomic class. A study was initiated to study the evolution of COX-I in Monogeneans and examined its protein sequences from 16 species for four families, Gyrodactylidae, Diplectenidae, Diplozoidae and Dictilophoridae. The finding may furnish a space to enumerate ancestral lineage and evolutionary pattern among selected families.

Corresponding author: Dr. P. V. Arya,
Department of Zoology, Dyal Singh College, University of Delhi, New Delhi-110003, India.

MATERIALS AND METHODS

Selection of Protein Sequences: Cytochrome C oxidase-1 was selected for 16 species from four different families based upon the availability of particular type of protein sequences for sufficient number of species in a particular family, in order to carry out analytical studies. All sequences had varying length, differ by one or two amino acids with no phylogenetic issue at all. The Gyrodactylidae, Diplozoidae, Diplectanidae and Dictylophoridae had 5, 2, 6 and 3 selected species respectively.

Phylogenetic Analysis: Sequences were subjected to alignment using ClustalW (inbuilt in MEGA 6) for multiple sequence alignment (Thompson *et al.,* 1994) with the default gap and extension penalties. The phylogenetic tree generated using neighbor joining (NJ) method in MEGA 6. The average pathway method to calculate the branch length all over the sequences. Most parsimonious tree was chosen by the close-neighbor-interchange algorithm.

Pair-wise Sequence Alignment: Pair-wise alignment was done for 3-D structure. One protein sequence from each family was taken and executed into NCBI-PBLAST (Protein-Basic Local Alignment Search Tool). The sequence with the highest score was chosen for structural modeling. The number of mutation over amino acids and comparative evaluation among 04 sequences from families was also done.

Protein Structure Prediction: Homolog protein sequences were processed in SWISS-MODEL for structure prediction and identifying its quality predicted from features of the target-template alignment for model building based on the target-template alignment using Promod-2. Insertions and deletions were remodeled using a fragment library. Side chains were then remodeled followed by regularization by force field (Guex, *et al*., 1997).

The global and per-residue model quality were assessed using the Q-Mean scoring function (Benkert, *et al*., 2011).

RESULTS

Multiple Sequence Alignment: In multiple sequence alignment every sequence was approximately contented with 145 amino acids except sequences from *Gyrodactylus* species. Initially, sequence alignment was performed with full amino acid length and later it was trimmed for being highly dissimilar and mutative (fig. 1). After removing non-matching sequences a total of 77 amino acids with conserved sites were obtained. First block of MSA reveled that rate of mutation was slow among species of the genus *Gyrodactylus*. Out of five sequences, over MSA, one block is observed for mutation or mismatching, imparting the protein Cytochrome C Oxidase-1 with high conserved occurrence in the genus.

Overall Cytochrome C Oxidase-1 has been carrying random mutation events in some genus while others with less or no mutation throughout speciation and diversification. Globally, with 77 residues, only 20 sites found conserved, indicating cytochrome c oxidase-1, a significant conservative protein for phylogenetic analysis. Residues in larger red block could be omitted without any considerations (fig. 1), though, as per the individual genus sequences are concerned, they have negligible mismatches and significant conservation sites. Overall divergence among sequences covered a broad range of mutation (fig. 2). Each group/taxon with a particular range of divergence as in case of the family Gyrodactylidae, 50-55% of mean divergence was observed. Those of Diplectenidae, Diplozoidae and Dictilophoridae had 36-70%, 22-28% and 7-9% of mean divergence respectively. Range of mean divergence characterizes to the rate of change, larger the range, faster the rate of change in amino acid composition. Diplectenidae exhibited great variability in terms of protein conservation.

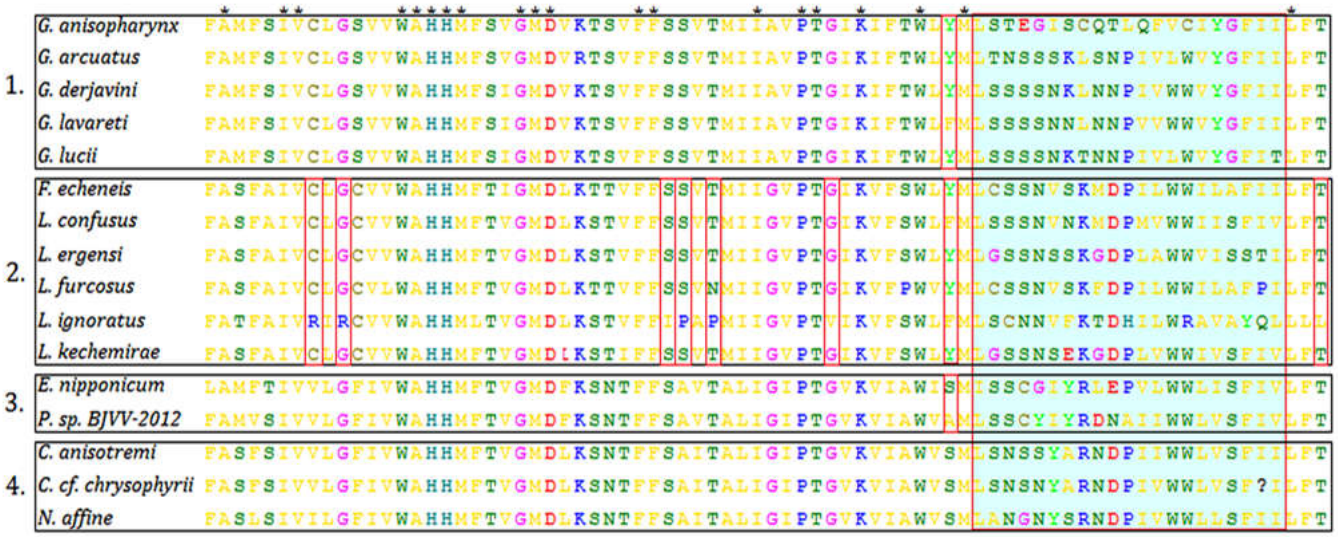

Figure 1.Multiple sequence alignment of 16 protein sequences. Number as 1-4; families Gyrodactylidae, Diplectenidae, Diplozoidae & Dictilophoridae respectively. *conserved amino acid residues; red box nonconserved regions in the sequences. Similar amino acids are given the same color; red block dissimilarities of residue in the particular genus

Contrast observed by the species of the genus *Lamellodiscus*. Six sequences of the genus reflected great diversity within the species. Other two genus *Diplozoidae* and *Dictilophoridae* with 2 and 3 species sequences had one and no mismatches respectively.

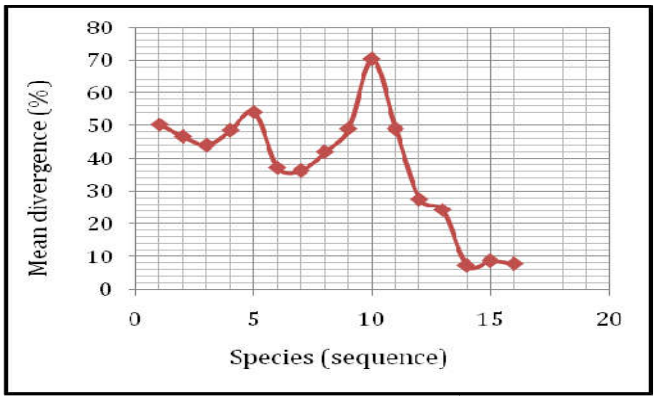

Figure 2. Plot of mean divergence between 16 sequences. Mean calculated in the form of a matrix using MEGA6 and then plotted

Pair-wise Sequence Alignment: Using NCB-PBLAST selected proteins from each genus were run and homolog were retrieved from the result with higher similarity percentage in order to assume 3-D structure of the Cytochrome C oxidase-1 (fig. 3). All of the sequences had similarity score above 85-95% that made them easy for protein homology modeling.

The query sequences were selected on the basis of their conservatory behavior in multiple sequence alignment. The phylogenetic trees from all methods produced the similar taxon group except UPGMA, projected out the out-group which was the only difference among all (fig. 4). In the process of sequence manipulation and tree construction, no sequence was given as out-group.

The first group, Gyrodactylidae, in the MSA showing just one mutation coincided with its clade about large branch of *Gyrodactylus anisopharynx* in all phylogenetic trees. So, this species may or may not be regarded as out-group or root of the tree, depending upon the kind of analysis being performed. Clade also shown monophyletic mode of species divergence, confirming small mutation among protein sequence.

Figure 3. Pair-wise sequence alignment of 4 sequences selected from each family. b *Gyrodactylus anisopharynx;* b. *Lamellodiscus furcosus;* c. *Eudiplozoon nipponicum* & d. *Neoheterobothrium affine.* Blocks- conserved/matching amino acid residues

Other group Diplectenidae in MSA depicted observable variation at various sites as it had 7 mismatches in residues and have been the group to have highest variability in the sequences. The variations in the sequences led to the greater rate of speciation than others under investigation. Other two groups Diplozoidae and Dictilophoridae were amazingly found to share a common origin (Fig. 4), though not as monophyletic but paraphyletic evolutionary pattern. As individual clade, Diplozoidae and Dictilophoridae both separately showing monophyletic pattern of evolution, suggesting origin from two different ancestors.

Protein Structure Prediction

For structural variations protein were modeled using Swiss-Model server and structure of four sequences was predicted (Fig. 5). Despite of sequence variability in cytochrome c oxidase-1 significant similarity was observed for every protein with reference to structure and function as well. In order to compare the four proteins we intended to set parameters those of Swiss-Model generated itself that include local quality estimation, development of α-helix and β-sheet structure etc. The local quality estimation (LQE) in (Fig. 6). LQE can be manipulated for justifying the variable amino acid numbers. On one axis graph showing number of residues and predicted local similarity to target on the other axis. *Gyrodactylus anisopharynx* had 250 residues and drawn a good similarity score of 0.5-0.8, a considerable range to target sequence. Other 03 sequence of *Lamellodiscus furcosus, Neoheterobothrium affine and Eudiplozoon nipponicum* are having number of amino acids just half of *Gyrodactylus anisopharynx* and hence showing lesser similarity score, 0.3-0.8 to their target sequences.

More precisely only *Lamellodiscus furcosus* had poor score to its target. This may be related with the above result of MSA in which high mutation had occurred throughout the alignment. Over local similarity, being homologous to each other, sequence coincided in structure and functions.

Fig. 6 representing a less dissimilarity in all monomeric structure as the complete protein is made up of 13 polypeptides. So, it is confirmed that structurally Cytochrome C Oxidase-1 remain conserved even if one compares individual polypeptides from different family. There might have mutation by environmental or ecological factors and great speciation event would have led the conserved protein to keep unique amino acids composition conserved, tending no change in structure. Feasibility of differences in Cytochrome comes from residues participating in core formation of protein whose removal or deletion would not affect the structural topology and so the function. Structurally all proteins are monomeric as a key enzyme in aerobic metabolism by functions. Proton pumping heme-copper oxidases represent the terminal, energy-transfer enzymes of respiratory chains in prokaryotes and eukaryotes.

Evaluation & comparison of Secondary Structure

Cytochrome C Oxidase-1 protein secondary structure was further elaborated and then compared so as to establish a clear distinction among them to identify the probable function of a protein from 3-D structure using a series of method (Fig. 8).

In secondary structure (Fig. 8) α-helix and β-sheet are very common to occur, depending upon the intrinsic propensity of amino acid sequence in a protein.

Phylogenetic Analysis

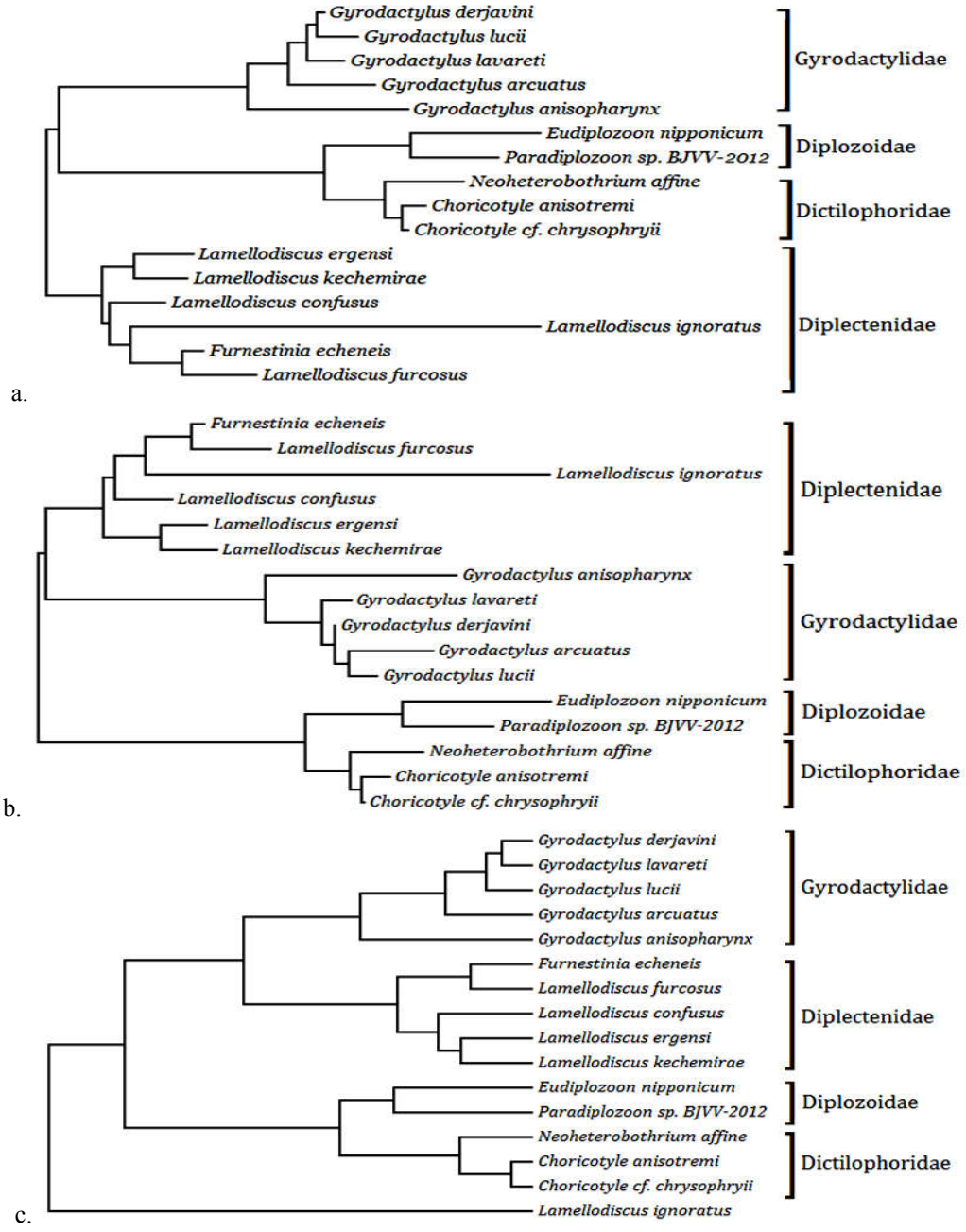

Figure 4. Phylogenetic trees a. Neighbor Joining; b. Maximum Parsimony & c. UPGMA.

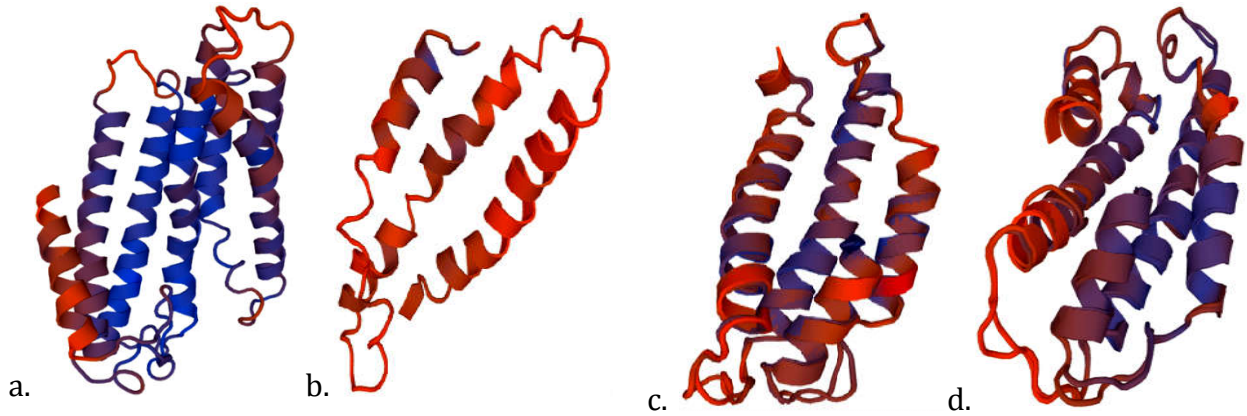

Figure 5. Predicted protein structure of a. *Gyrodactylus anisopharynx;* b. *Lamellodiscus furcosus;* c. *Neoheterobothrium affine* & d. *Eudiplozoon nipponicum*

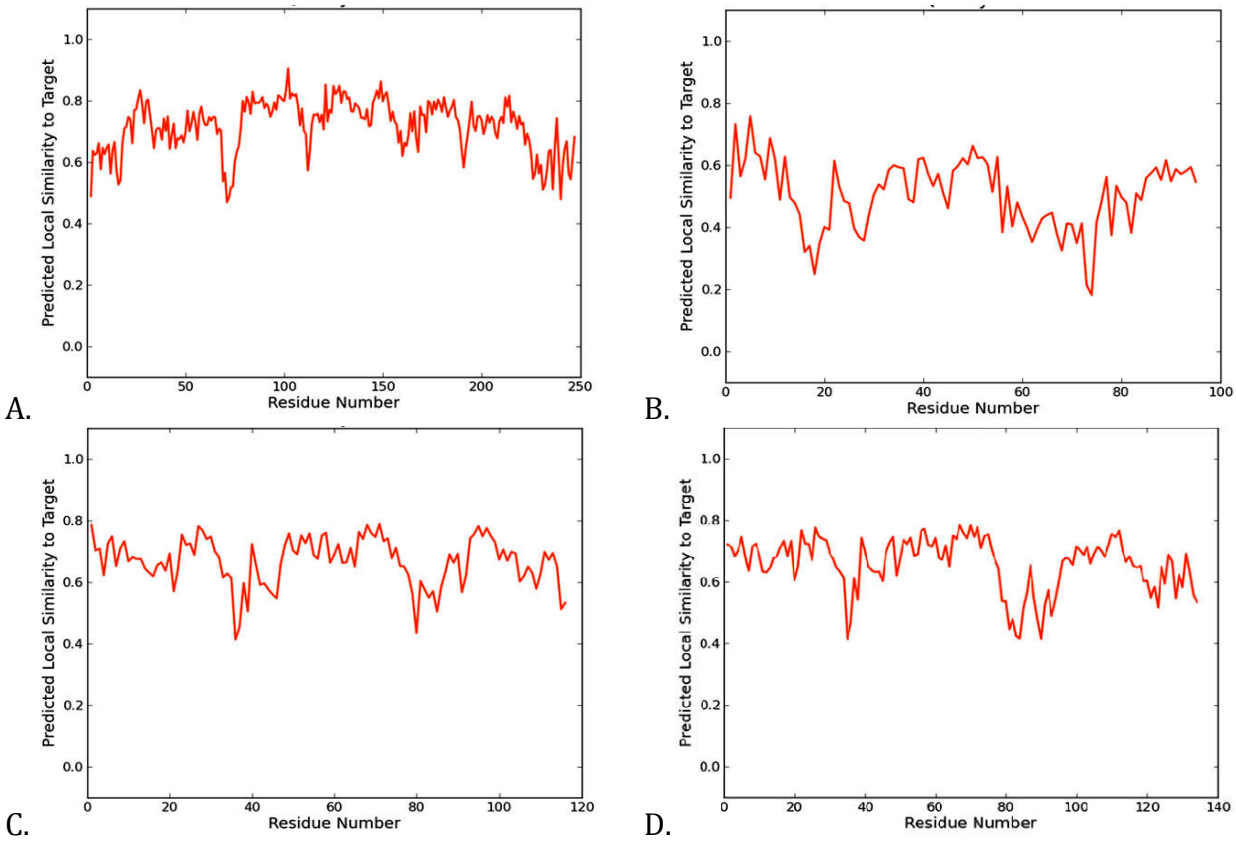

Figure 6. Local estimation of side chains to target sequences. A. *Gyrodactylus anisopharynx;* B. *Lamellodiscus furcosus;* C. *Neoheterobothrium affine,* & D. *Eudiplozoon nipponicum*

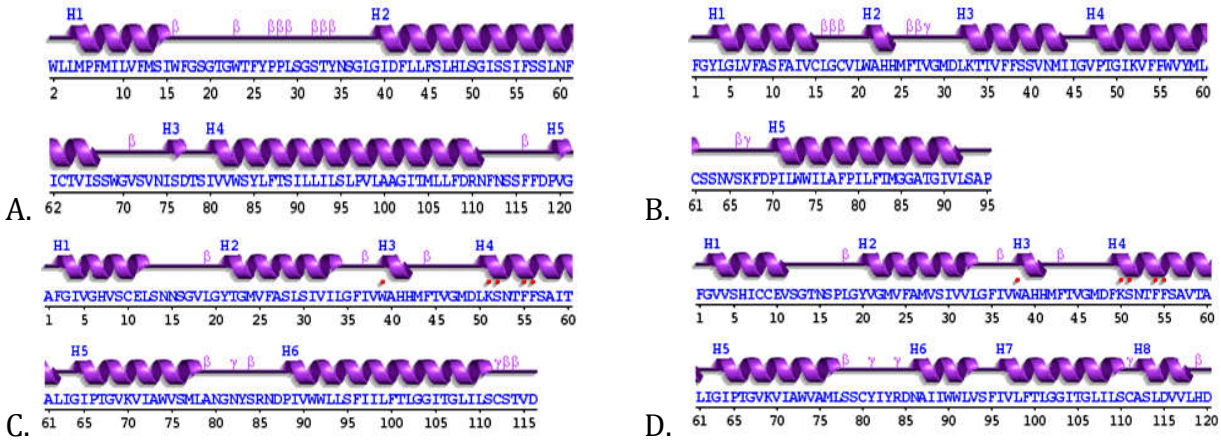

Figure 7. Open form secondary structure of Cytochrome C Oxidae-1. A. *anisopharynx;* B. *furcosus;* C. *affine* & D. *nipponicum*

Fig. 7A which belongs to *anisopharynx,* although, had 240 amino acid residues but only 120 residues were shown in the structure in order to coincide with other groups where *furcosus* had the least number of residues. Helix and sheet were indicated as H and β respectively.

For *anisopharynx* a total of 6 helices, out of which only 3 larger and rest of the sequence tend to develop β sheets. In comparison to *furcosus*, it contain 4 larger α helices and a smaller one. Likewise *affine* 5 larger α helices and 2 smaller ones. Highly significant number of α helices was developed into *nipponicum* as 5 larger and 3 smaller ones with just 120 amino acid residues. From the reference of stability, β sheet are more stable than α-helix and tend to show lesser mutation in the course of evolution.

As per results more α-helix greater the mutation or lesser the β sheet lesser the mutation. Resultantly least number of α-helix in *anisopharynx* has made it more stable than remaining three. And therefore, the evolution in that particular protein will be more than others. This result is consistent with the evolution of proteins under adverse conditions. In an order of stability to fix the relative evolution of four groups of Cytochrome C Oxidase-1, it can be represented as; *anisopharynx* > *furcosus* > *affine* > *nipponicum.* Accordingly their rate of evolution was understood with the order of relative stability and so the pattern of evolution. *G. anisopharynx* was evolved at slowest rate with highest stability and in contrast *Eudiplozoon nipponicum* evolved at fastest rate with least stability for protein.

DISCUSSION AND CONCLUSION

Measurement of sequence parameters including MSA, PSA and local quality estimation for inferring out the 3-D structure, reveals a number of facts over evolution of COX-I among monogeneans. Problems faced in carrying the study was the lack of availability of complete sequence of COX-I for monogeneans, therefore, analysis over the gene duplication and gene divergence could not be performed that would have certainly strengthened our finding for evolutionary aspect in different monogeneans and would have provided strong clues for their relatedness from across the globe. MSA provide an initial and comparative understanding of protein variability (Blackburne and Whelan, 2013). Separately, all 04 groups support intra-genus relationship by having specific mutation sites. The first group, Gyrodactylidae, with five species had highly conserved pattern in protein sequences (Fig. 1).

Only one site is found mutated with a mean divergence of 45-50% (fig. 2) with other sequences. The unique feature about the family is the monophyletic evolution of the species as shown in phylogenetic tree (Fig. 3) showing a linear evolutionary pattern from a common ancestor, withdrawing our attention towards possible relationship among species by fast but sensitive mutation. Knowingly, Gyrodactylidae represent the most diverse species for maximum number of geographical distribution. Family is rich in both number of species and adaptation to various ecological conditions. Most importantly, Gyrodactylidae show most stable form by having developed least α-helix structure (Fig. 8). Evolution in terms of gene duplication events have not been considered for COX-I in monogeneans because it fails to provide enough cue on mutation events that would be sufficient enough to create a new path of genus/group. Evolution has taken place in the protein of the family but it does not necessarily mean to have a new species or group in return.

Family Diplectenidae exhibited eight point mutations even after MSA and sequence editing had done, indicating a higher rate of mutation in the family with higher level of speciation and divergence. The tendency was supported by the molecular phylogeny of the group in the phylogenetic tree that they follow dual route (monophyletic and paraphyletic) of evolution. Among them is *Lamellodiscus ignoratus* exhibited longest branch (Fig. 4), an indication of maximum mutation in gene besides other family member. A significant variation in the mean divergence (Fig. 2) for Diplectenidae further strengthens higher species variability among the members of the group. The observation can be further rationalized with ecological attributes and geographical distribution for a clear scenario over the entire family.

In earlier studies of zoogeographical distribution and molecular phylogeny on *Lamellodiscus*, the family had found not confined in to a particular geographical zone rather it had been dispersed across the globe with significant phylogenetic anomaly (Fozail Ahmad *et al*., 2015). Members of the group were found in almost each geographical region, providing a strong support to our current study. The third group, Diplozoidae shows a single point mutation in MSA with 22-27% of mean divergence that may have either increased or decreased if more sequence had incorporated. Surprisingly, this group represents phylogenetic relationship (Fig. 4) with Dictilophoridae and forms a separate taxon.

The feasibility of monophyletic evolution or more precisely, co-evolution of COX-I in both of the group may have taken place and close relatedness among members can be inferred. As an individual group of Monogeneans, mean divergence (7-10%) of Dictilophoridae is least from others with no point mutation in MSA. Structurally, both of them are very similar in terms of having number of α-helix is 6 and 7 for Diplozoidae and Dictilophoridae respectively (fig. 8). These finding are supported by local quality estimation of protein sequences while modeling their three dimensional structures. Both show almost equal range of similarity for their target sequence.

Overall, four groups in the study provides a generalized evolutionary distinction of COX-I protein of Monogenean families in terms of sequence and structure. The four groups are highly diverging members of parasitic class, representing variability in conserved protein. Monogeneans can be evaluated on the basis of such analysis for their origin and evolution. Further studies can be performed with more families/group in order to justify the ancestral lineage. This finding just gives an idea of evolutionary relatedness in all families/genus in term of COX-I protein changing over the period or may provide the beginning of evolution of class Monogenea.

Acknowledgement

We are thankful to the authorities of UGC for financial support (F. no. 41-34/2012 (SR)) and head of institution for providing necessary facilities. Acknowledgement is also due to all the contributors whose sequences are retrieved from databases for the present study.

REFERENCES

Blackburne, B. P. and Whelan, S. Mar. 2013. "Class of Multiple Sequence Alignment Algorithm Affects Genomic Analysis," *Mol. Biol. Evol.*, vol. 30, no. 3, pp. 642–653.

Breek, C. K. Speijer, D., Dekker, H., Muijsers, A. O. and Benne, R. Aug. 1997. "Further evidence for the presence of mitochondrially encoded subunits in cytochrome c oxidase of the trypanosomatid Crithidia fasciculata," *Biol. Chem.*, vol. 378, no. 8, pp. 837–841.

Butland, G., Peregrín-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J. and Emili, A. Feb. 2005. "Interaction network containing conserved and essential protein complexes in Escherichia coli," *Nature*, vol. 433, no. 7025, pp. 531–537.

Castresana, J., Lübben, M., Saraste, M. and Higgins, D. G. Jun. 1994. "Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen," *EMBO J.*, vol. 13, no. 11, pp. 2516–2525.

Clemente, P., Peralta, S., Cruz-Bermudez, A., Echevarría, L., Fontanesi, F., Barrientos, A., Fernandez-Moreno, M. A. and Garesse, R. Mar. 2013. "hCOA3 stabilizes cytochrome c oxidase 1 (COX1) and promotes cytochrome c oxidase assembly in human mitochondria," *J. Biol. Chem.*, vol. 288, no. 12, pp. 8321–8331.

Collman, J. P., Devaraj, N. K., Decreau, R. A., Yang, Y., Yan, Y.L., Ebina, W. T., Eberspacher, A. and Chidsey, C. E. D. Mar. 2007. "A Cytochrome c Oxidase Model Catalyzes Oxygen to Water Reduction Under Rate-Limiting Electron Flux," *Science*, vol. 315, no. 5818, pp. 1565–1568.

Ekici, S., Turkarslan, S., Pawlik, G., Dancis, A., Baliga, N. S., Koch, H.G. and Daldal, F. 2014. "Intracytoplasmic copper homeostasis controls cytochrome c oxidase production," *mBio*, vol. 5, no. 1, pp. e01055–01013.

Fozail Ahmad, D. Singh and Arya, P.V. Jun. 2015 "*In-silico* phylogenetic studies on some members of parasitic genus *Gyrodactylus* (Monogenea: Gyrodactylidae) for assessment of evolutionary relatedness inferred from 28s rRNA and geomapping the sample," *Int. J. Recent Sci. Res.*, vol. 6.

Fozail Ahmad, Singh, D. and Arya, P. V. Jun. 2015. "Comparative evaluation of speciation and zoogeographical distribution for *Lamellodiscus* (Monogenea: Diplectanidae) using 18S rRNA," *Int. J. Innov. Sci. Res.*, vol. 4, no. 6, pp. 235–241.

Lenka, N., Vijayasarathy, C., Mullick, J. and Avadhani, N. G. 1998. "Structural organization and transcription regulation of nuclear genes encoding the mammalian cytochrome c oxidase complex," *Prog. Nucleic Acid Res. Mol. Biol.*, vol. 61, pp. 309–344.

Li, Y., Park, J.S., Deng, J.H. and Bai, Y. Dec. 2006. "Cytochrome c oxidase subunit IV is essential for assembly and respiratory function of the enzyme complex," *J. Bioenerg. Biomembr.*, vol. 38, no. 5–6, pp. 283–291.

Rottenberg, H. Apr. 1998. "The generation of proton electrochemical potential gradient by cytochrome c oxidase," *Biochim. Biophys. Acta BBA - Bioenerg.*, vol. 1364, no. 1, pp. 1–16.

Sicheritz-Ponten, T. Jan. 2001. "A phylogenomic approach to microbial evolution," *Nucleic Acids Res.*, vol. 29, no. 2, pp. 545–552.

Sicheritz-Pontén, T., Kurland, C. G. and Andersson, S. G. E. Jul. 1998. "A phylogenetic analysis of the cytochrome b and cytochrome c oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae1The nucleotide sequences in this paper have been deposited in the EMBL sequence database under accession Nos. Y13854 (cob) and Y13855 (cox1).1," *Biochim. Biophys. Acta BBA - Bioenerg.*, vol. 1365, no. 3, pp. 545–551.

Smits, P., Smeitink, J. A. M., van den Heuvel, L. P. Huynen, M. A. and Ettema, T. J. G. Jun. 2007. "Reconstructing the evolution of the mitochondrial ribosomal proteome," *Nucleic Acids Res.*, vol. 35, no. 14, pp. 4686–4703.

Socolich, M., Lockless, S. W., Russ, W. P., Lee, H., Gardner, K. H. and Ranganathan, R. Sep. 2005. "Evolutionary information for specifying a protein fold," *Nature*, vol. 437, no. 7058, pp. 512–518.

Soto, I. C., Fontanesi, F., Liu, J. and Barrientos, A. Jun. 2012. "Biogenesis and assembly of eukaryotic cytochrome c oxidase catalytic core," *Biochim. Biophys. Acta BBA - Bioenerg.*, vol. 1817, no. 6, pp. 883–897.

Steffens, G. C., Biewald, R. and Buse, G. Apr. 1987. "Cytochrome c oxidase is a three-copper, two-heme-A protein," *Eur. J. Biochem. FEBS*, vol. 164, no. 2, pp. 295–300.

Wolz, W., Kress, W. and Mueller, C. R. Oct. 1997. "Genomic sequence and organization of the human gene for cytochrome c oxidase subunit (COX7A1) VIIa-M," *Genomics*, vol. 45, no. 2, pp. 438–442.

Wu, W., Schmidt, T. R., Goodman, M. and Grossman, L. I. Nov. 2000. "Molecular Evolution of Cytochrome c Oxidase Subunit I in Primates: Is There Coevolution between Mitochondrial and Nuclear Genomes?," *Mol. Phylogenet. Evol.*, vol. 17, no. 2, pp. 294–304.

****