International Journal of Advance Research in Science and Engineering
Vol. No.5, Issue No. 03, March 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

EVOLUTIONARY DISTANCE & CONSERVED DOMAIN ANALYSIS OF DIVERGENT PHYLOGENETIC LINEAGES OF CYTOCHROME C OXIDASE SUBUNIT III (MITOCHONDRION)

Sonu Mishra[1], **Virendra S. Gomase**[2]

[1,2]*Department of Biotechnology, MewarUniversity, Chittorgarh, (India)*

ABSTRACT

The aim of the present study is to identify the origin, evolutionary distance and conserved domain analysis of divergent phylogenetic lineage of cytochrome c oxidase subunit III (mitochondrion).We also conducted the prediction of structure and function of protein by multiple sequences analysis and observed the conserved pattern of amino acid residues and to construct the phylogenetic tree for organizing evolutionary history.

Keywords:* *cytochrome c oxidase subunit III (mitochondrion); multiple sequence alignment; conserve domain; Phylogenetic analysis

I. INTRODUCTION

Cytochrome c oxidase is an oligomeric enzymatic complex .It's occurrence has been noticed in inner membrane of the mitochondrial in eukaryotes and in plasma membrane of aerobic prokaryotes. The cytochrome c oxidase comprises of three common subunits ,I,II and III. These compose the core structure of prokaryotic and eukaryotic cytochrome c oxidase. The fourth subunit is formally seen with the combined subunits I and II in prokaryotes, whereas, in eukaryotes there are a variable number of additional small polypeptidic subunits found [Mather MW, Springer P, Hensel S, Buse G, Fee JA.,1993]. The subunit III of Cytochrome c oxidase domain is the terminal enzyme of the respiratory chain of mitochondria and many aerobic bacteria. It catalyses the transfer of electrons from reduced cytochrome c to molecular oxygen:

4 cytochrome c+2 + 4 H+ + O2 -->4 cytochrome c+3 + 2 H2O.

This reaction is coupled to the pumping of four additional protons across the mitochondrial or bacterial membrane [Michel H.,1999; Belevich I, Verkhovsky MI, Wikström M.,2006]. The bacterial respiratory system are branched, therefore they contain a number of distinct terminal oxidases, whereas, in the eukaryotic mitochondrial system contains single cytochrome oxidase .The cytochrome o oxidase does not play any role in the catalysis of the cytochrome c but in the oxidation of the quinol(ubiquinol),they belong to the same superfamily of the haem-copper oxidase as cytochrome c oxidases.Thhaem-copper oxidase superfamily member partake sequence similarities in all three core subunits(I, II,III).The more conserved subunit is subunit I and the least conversed one is subunit II [Santana M, Kunst F, Hullo MF, Rapoport G, Danchin A, Glaser

P.,1992;Chepuri V, Lemieux L, Au DC, Gennis RB.,1990;García-Horsman JA, Barquera B, Rumbley J, Ma J, Gennis RB.,1994]

II. MATERIALS AND METHODS

2.1 Sources and sequence information of cytochrome c oxidase subunit III.

cytochrome c oxidase subunit III has taken from *Dracunculusmedinensis*, in which targeted protein data were used to observe molecular resemble of related protein by phylogenetic analysis .**cytochrome c oxidase subunit III comprised of 255 amino acids residue were retrieve from the** from www.ncbi.nlm.nih.gov, UniProt databases are initially the most important [http://www.ncbi.nlm.nih.gov; Sayers, E.W.et al., 2012].

MKHNYHLLSYSGYPFMVFCSVMGLSSSLVIFLKYGVIFGVFFGVFCLFCVVMVWCKDIFMEGLSGYHN
FFVMNGFKYGMVFFIFSEFMFFFGVFWVFFDSSLVPNSELGMSWCPLGIGLINPLGVPLLNTLILLSSAV
TVTWCHNSMLCNYNSFYGLFFTCVLALFFLVFQMLEYDESGFSMSDGIYGSIFYLSTGFHGMHVFFGMI
FLFVNLFRLYMDHFNSDHHLGLEFSIVYWHFVDLIWLFLFVFVYWWSF

2.2 Multiple sequence alignment of cytochrome c oxidase subunit III

Multiple sequence alignment [MSA] is conducted by COBALT, which aligns cytochrome c oxidase subunit III protein sequences of similar protein using a combination of distance matrix and approximate parsimony methods. Numerical setting method is used to study the relative entropy threshold, in bits, that must be met for an alignment column to be displayed in red. A larger number indicates higher degree of conservation. The relative entropy is computed as: Σi fi log2 (fi / pi), where iis residue type, fi is residue frequency observed in the multiple alignment column, and pi is the background residue frequency. Identity setting used for only columns with one residue type will be colored in red [Papadopoulos JS and Agarwala R., 2007].

2.3 Construction of a phylogenetic tree cytochrome c oxidase subunit III protein.

Phylogenetic analyses were performed by fast minimum evolution algorithm and Neighbor Joining algorithms to allow the reconstruction phylogenetic tree of the molecular evolutionary history of various aligned sequences that are useful to align highly evolved gene families clearing evolutionary relationships such as multiple actin proteins [Desper R and GascuelO., 2004; Saitou N and NeiM., 1987]. Trees were obtained by the methods fast minimum evolution algorithm and Neighbor Joining algorithms. Evolutionary distance is studied by Grishin (protein) model [GrishinNV., 1995; GrishinNV., 1995]and distance between two sequences modeled as expected fraction of amino acid substitutions per site given the fraction of mismatched amino acids in the aligned region and can be computed for fraction of mismatched amino acids larger than 0.75 [Gomase VS and Tagore S.,2009; Waghmare S, Waghmare D, Bhatnagar PS., 2013; Marchler-Bauer A et al. 2011; Marchler-Bauer A. et al. , 2009].

III. RESULTS & INTERPRETATION

3.1 Evolutionary distance

This study, cytochrome c oxidase subunit III protein from *D. medinensis* is summarized to study the evolutionary distance. The identification of the origin of protein, multiple sequences analysis, observing the conserved amino acid residues and reconstruct the phylogenetic tree specify the evolutionary history, relationship of protein with different species of nematodes [**Table-1**]. Rectangle tree shows rectangular shaped rooted tree, where root is places in the longest edge. Fast minimum evolution algorithm produce un-rooted tree such as ones shown as radial or force in the tabs below. The rooted trees are created by placing a root in the middle of the longest edge. Slanted tree shows similar to rectangle, but with triangular tree shape[**Fig.1-2**]. Neighbor Joining algorithms produce un-rooted tree such as ones shown as radial or circular [**Fig.3-4**]in the tabs below. The rooted trees are created by placing a root in the middle of the longest edge.

Table 1- Sequences producing significant alignments

Description	Accession	Max score	Total score	Query cover	E value	Ident
cytochrome c oxidase subunit III [Dracunculusmedinensis]	YP_004857905.1	494	494	1	6.00E-177	1
cytochrome c oxidase subunit III [Anisakis simplex]	YP_537094.1	321	321	1	1.00E-108	0.63
cytochrome c oxidase subunit III [Toxocaramalaysiensis]	YP_001795388.1	307	307	1	5.00E-103	0.63
cytochrome c oxidase subunit III [Toxocaravitulorum]	ACM88521.1	305	305	1	2.00E-102	0.63
cytochrome c oxidase subunit III [Toxocaracanis]	YP_001905890.1	305	305	1	2.00E-102	0.63
cytochrome c oxidase subunit III [Metastrongyluspudendotectus]	YP_003433813.1	301	301	1	1.00E-100	0.63
cytochrome c oxidase subunit III [Syngamus trachea]	YP_003433939.1	300	300	0.99	3.00E-100	0.63
cytochrome c oxidase subunit III [Dictyocaulusviviparus]	YP_007183194.1	300	300	0.99	3.00E-100	0.65
cytochrome c oxidase subunit III [Steinernemacarpocapsae]	YP_026085.1	298	298	1	2.00E-99	0.66
cytochrome c oxidase subunit III [Metastrongylussalmi]	YP_003433836.1	296	296	1	1.00E-98	0.62
cytochrome c oxidase subunit III [Toxocaracati]	YP_001936462.1	295	295	1	2.00E-98	0.62
cytochrome c oxidase subunit III [Cucullanusrobustus]	YP_004927930.1	295	295	1	3.00E-98	0.65
cytochrome c oxidase subunit 3 [Angiostrongyluscostaricensis]	AKU36882.1	294	294	1	7.00E-98	0.62
cytochrome c oxidase subunit III [Anisakis simplex C Nascetti et al., 1986]	YP_009113340.1	291	291	1	8.00E-97	0.64
cytochrome c oxidase subunit III [Ascaridiagalli]	YP_008145659.1	291	291	0.97	1.00E-96	0.67
cytochrome c oxidase subunit III [Anisakis simplex]	AES86035.1	291	291	1	1.00E-96	0.64

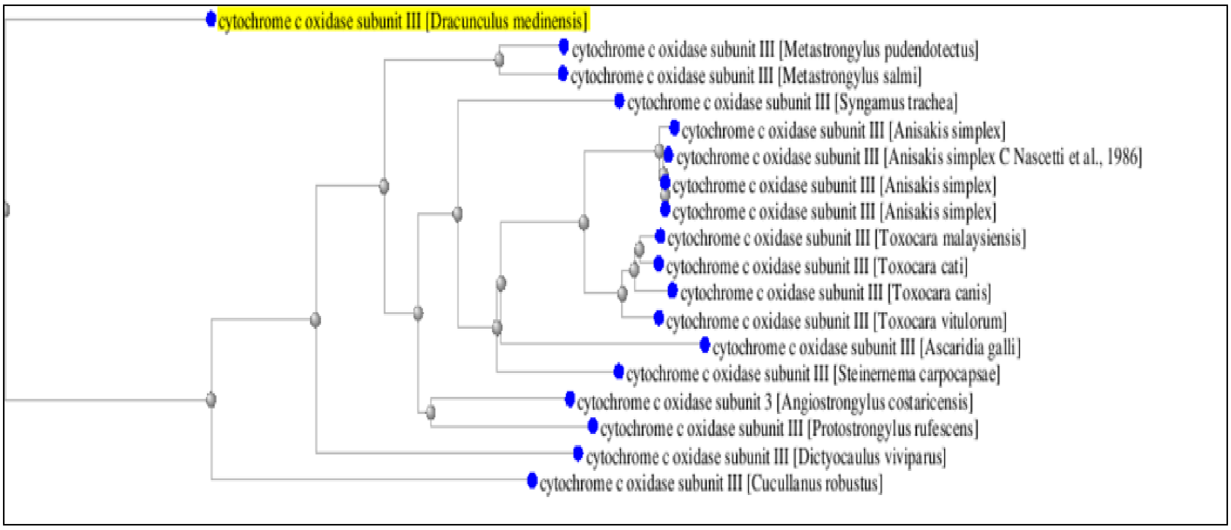

Fig. 1- Rectangle tree - Fast minimum evolution algorithm - Phylogenetic study of cytochrome c oxidase subunit III from D. mednensis with the help of rendering tree showing the evolutionary difference with other nematodes species in the rectangular shaped rooted tree, root is places in the longest edge.

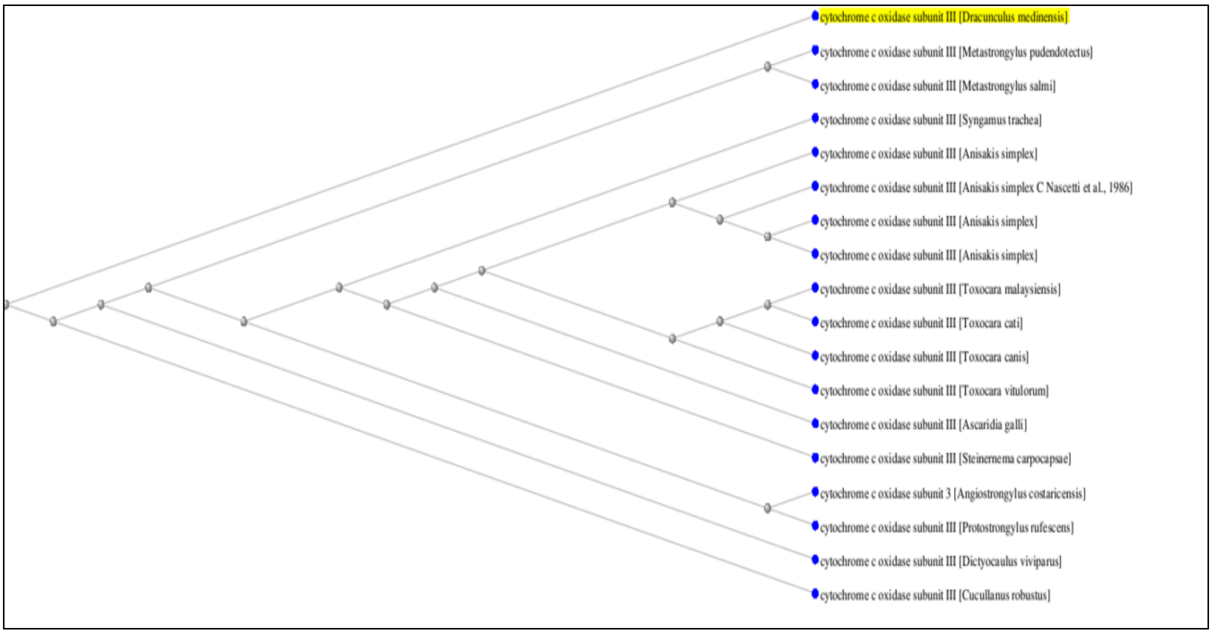

Fig. 2- Slanted tree – Fast minimum evolution algorithm - Grishin (protein) model- Phylogenetic study of cytochrome c oxidase subunit III from D. mednensis the help of rendering tree. Slanted tree is Similar to rectangle, but with triangular tree shape.

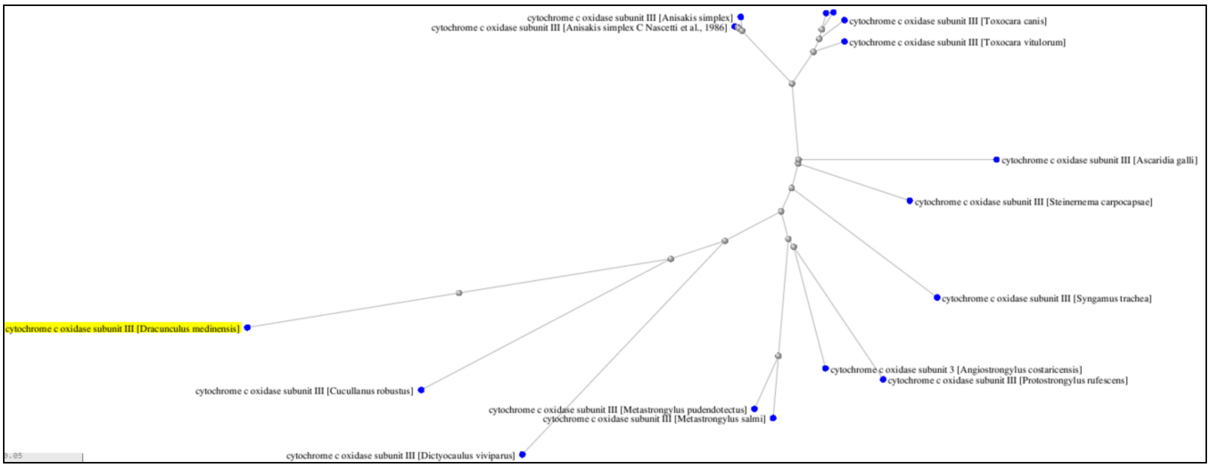

Fig. 3-Radial tree - Neighbor Joining algorithms - Grishin (protein) model- Phylogenetic study of cytochrome c oxidase subunit III from D. mednensis with the help of rendering tree. Radial tree is unrooted tree shape.

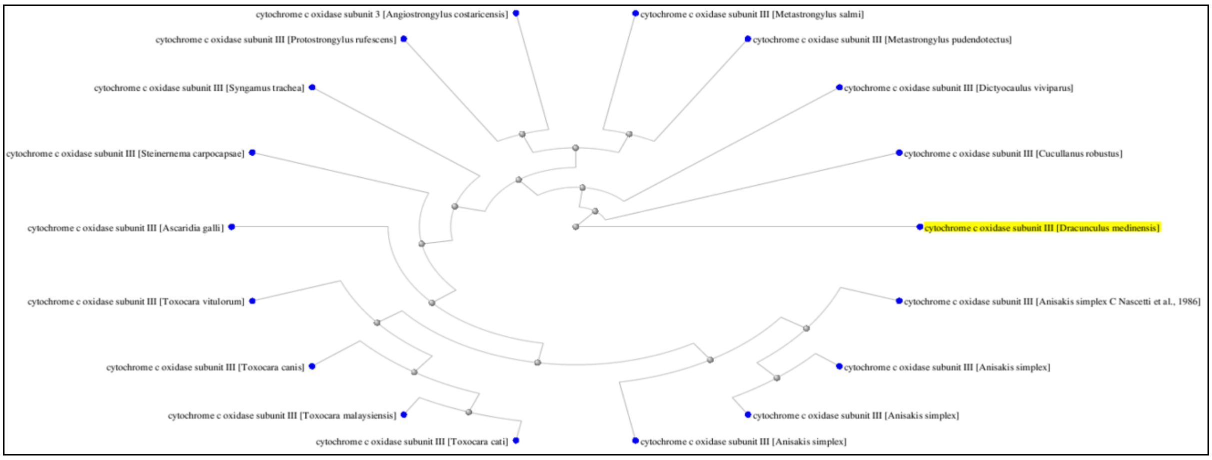

Fig. 4-Circular tree - Neighbor Joining algorithms - Grishin (protein) model- showing Phylogenetic study of cytochrome c oxidase subunit III from D. mednensis with the help of rendering tree.Radial tree is unrooted tree shape.

3.2 MSA(Multiple sequence alignment)

Multiple sequence alignment analysis shows columns with no gaps are colour in blue or red. The **red colour** indicates highly conserved regions and **blue** indicates less conserved ones. The Conservation analysis can be used to select a threshold for determining which columns are colours in red **[Fig.-5]**. Multiple sequence alignment identify conserved motifs and to predict functional role in the variable sites as well as conserved sites show the sequence divergence profile of these proteins, which demonstrate the sequence enrichment strategy of these sequences for adaptation to different physiological systems. Here we observed that from all sequences of neurotoxin proteins that Cys(c),Thr (T),Asn (N) (Hydrophilic amino acid) Phe(F),Gly(G),Ala(A),Pro(P) (hydrophobic amino acid), Lys (K),Arg(R), Positive charged, Asp(D), Negative charged which is conserved in all peptides having a common ancestor. That all of these peptides share eight highly conserved cysteines which were involved in the formation of β-strands are almost conserved. Multiple sequence alignment is carried out by COBALT of cytochrome c oxidase subunit III from GWD.

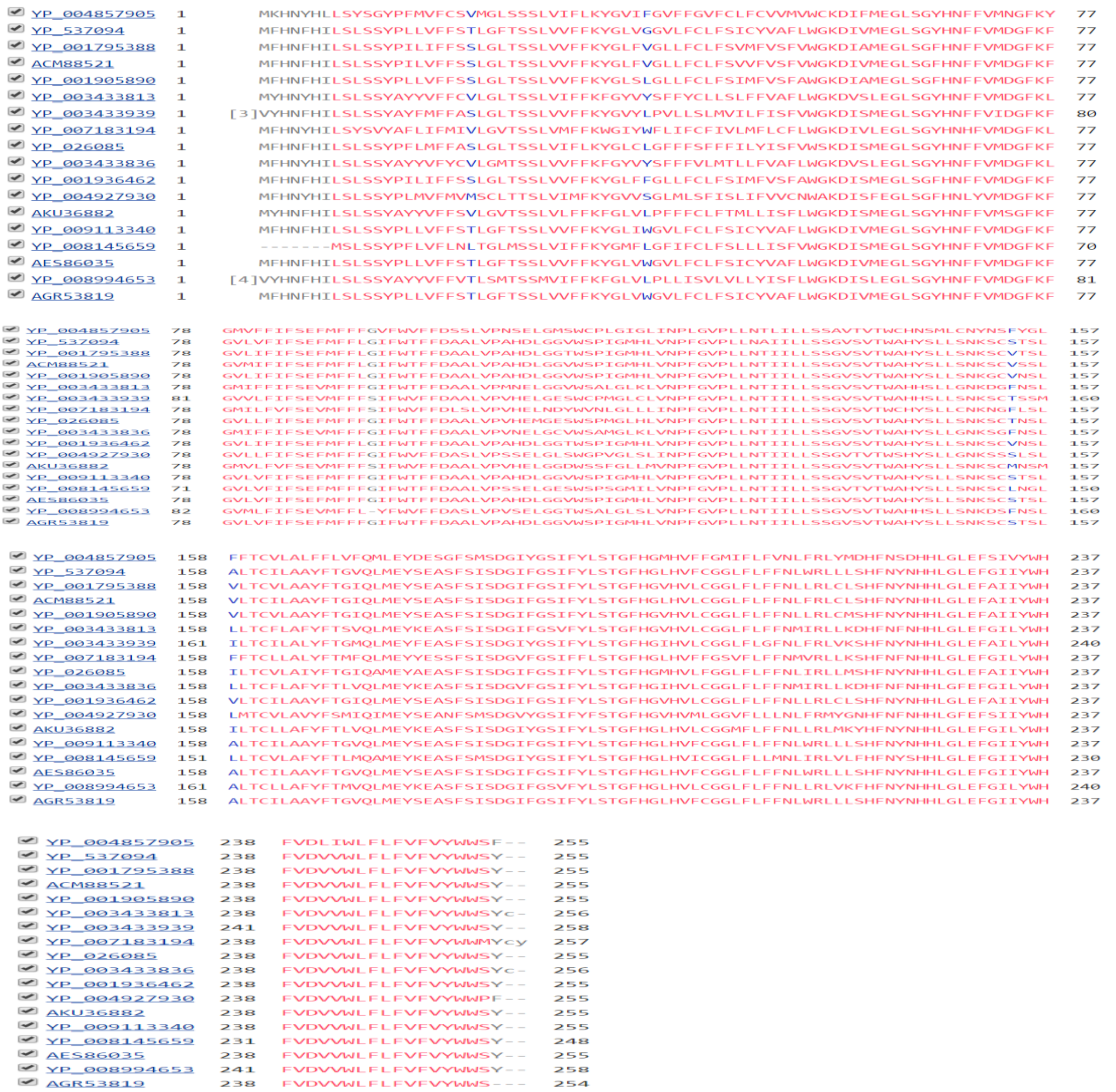

Fig 5. Multiple sequence alignment by COBALT of cytochrome c oxidase subunit III from GWD.Here columns with no gaps are colored in blue or red. The red colorCys(c),Thr (T),Asn (N) hydrophilic polar, Phe(F),Gly(G),Ala(A),Pro(P) hydrophobic nonpolar,Lys (K),Arg(R), Positive charged, Asp(D), Negative chargeindicates highly conserved columns and blue indicates less conserved ones. The Conservation Setting can be used to select a threshold for determining, which columns are coloured in red.

3.3 Conserved domain analysis

Molecular study of cytochrome c oxidase subunit III shows conserved domains and having one of protein superfamily with user query added Superfamily (**Fig.6**). Cytochrome c oxidase (CcO), the terminal oxidase in the respiratory chains of eukaryotes and most bacteria, is a multi-chain transmembrane protein located in the inner membrane of mitochondria and the cell membrane of prokaryotes. CcO catalyzes the reduction of O2 and simultaneously pumps protons across the membrane. The number of subunits varies from three to five in bacteria and up to 13 in mammalian mitochondria. Only subunits I and II are essential for function, but subunit III, which is also conserved, is believed to play a role in assembly of the multimer complex. RhodobacterCcO subunit III stabilizes the integrity of the binuclear center in subunit I. Subunits I, II, and III of mammalian CcO are encoded within the mitochondrial genome and the remaining 10 subunits are encoded within the nuclear genome. Subunit III contains bound phospholipids in several crystal structures and is proposed to contain a "lipid pool." These phospholipids are believed to intrinsic constituents similar to cofactors of the enzyme (**Table-2& Fig 7**).

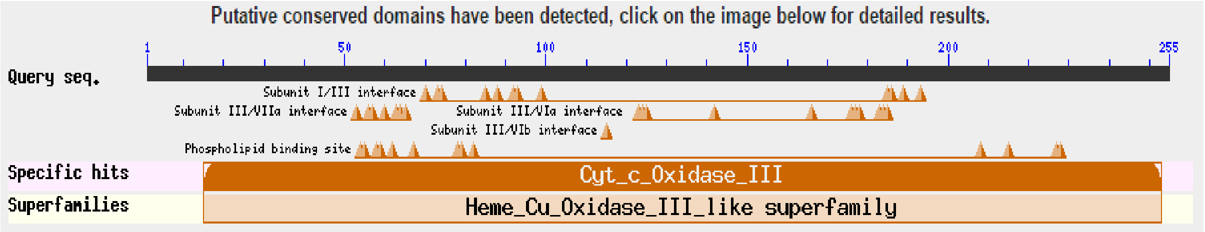
Putative conserved domains have been detected, click on the image below for detailed results.

Fig.6- Conserved domains on gi|346421508|gb|AEO27233.1|protein

Table 2-List of domain hits

Sl. No.	Description	PssmId	Multi-domain	E-value
1	**Cyt_c_Oxidase_III [cd01665]** - Cytochrome c oxidase subunit III. Cytochrome c oxidase (CcO), the terminal oxidase in the respiratory chains of eukaryotes and most bacteria, is a multi-chain transmembrane protein located in the inner membrane of mitochondria and the cell membrane of prokaryotes. CcO catalyzes the reduction of O2 and simultaneously pumps protons across the membrane. The number of subunits varies from three to five in bacteria and up to 13 in mammalian mitochondria. Only subunits I and II are essential for function, but subunit III, which is also conserved, is believed to play a role in assembly of the multimer complex. Rhodobacter CcO subunit III stabilizes the integrity of the binuclear center in subunit I. Subunits I, II, and III of mammalian CcO are encoded within the mitochondrial genome and the remaining 10 subunits are encoded within the nuclear genome. Subunit III contains bound phospholipids in several crystal structures and is proposed to contain a "lipid pool." These phospholipids are believed to intrinsic constituents similar to cofactors of the enzyme.	238834	No	4.48E-71
2	**Cytochrome c oxidase subunit III [pfam00510]**	249918	No	2.22E-83
3	cytochrome o ubiquinol oxidase, subunit III [TIGR02842] ; Cytochrome o terminal oxidase complex is the component of the aerobic respiratory chain which reacts with oxygen, reducing it to water with the concomitant transport of 4 protons across the membrane. Also known as the cytochrome bo complex, cytochrome o ubiquinol oxidase contains four subunits, two heme b cofactors and a copper atom which is believed to be the oxygen active site. This complex is structurally related to the cytochrome caa3 oxidases which utilize cytochrome c as the reductant and contain heme a cofactors, as well as the intermediate form aa3 oxidases which also react directly with quinones as the reductant. [Energy metabolism, Electron transport]	131889	No	6.18E-08
4	**cytochrome-c oxidase[PLN02194]**	177845	No	6.64e-46

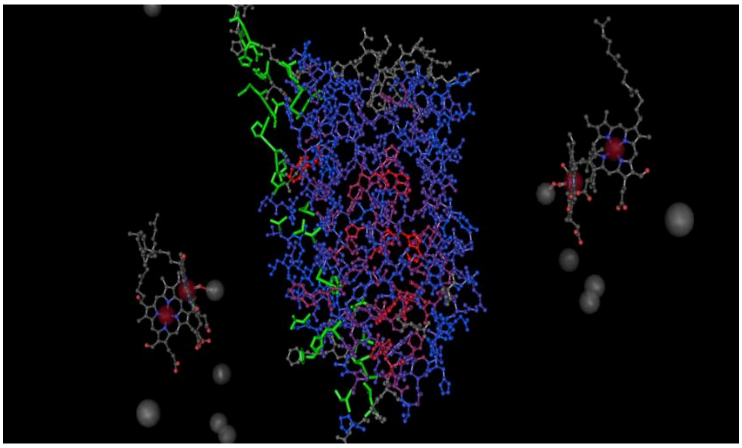

Fig.7- ***Sequence structure alignment shows similarity with 2OCC_C(interface with subunit 2OCC_A (subunit I), defined using 3.5 A contacts)*** **structure of Heme_Cu_Oxidase_III_like Superfamily(cl00211)**

IV. CONCLUSION

cytochrome c oxidase subunit III proteins of GWD are summarized the identical regions. Using multiple sequences analysis and phylogenetic tree we observe the conserved residues to specify the evolutionary history and analysing sequence structure relationship ofcytochrome c oxidase subunit III in nematodes family. Efficient utilization of Polar, nonpolar, positively and negatively charged amino acids and their distribution in protein sequence make them to be more antigenic. Comparative analyses specify that the protein demonstrates how proteins are generated within the nature's testing ground for tailor-made biologic needs. Evolutionary studies of cytochrome c oxidase subunit III sequence of *Dracunculusmedinensis*found the common ancestor of all in the nematodes family species. In future, this study will encourage to engineering to design synthetic peptide vaccine in the high positive effective role.

ABBREVIATIONS

MSA- Multiple sequence alignment
COBALT: constraint-based alignment tool for multiple protein sequences
GWD: Guinea worm disease
UniProt: The Universal Protein Resource
NCBI: National Center for Biotechnology Information
BLAST: Basic Local Alignment Search Tool

REFERENCES

[1]. Mather MW, Springer P, Hensel S, Buse G, Fee JA.,(1993). Cytochrome oxidase genes from Thermusthermophilus. Nucleotide sequence of the fused gene and analysis of the deduced primary structures for subunits I and III of cytochrome caa3. J BiolChem. ;268(8):5395-408.

[2]. Michel H.,(1999). Cytochrome c oxidase: catalytic cycle and mechanisms of proton pumping--a discussion. Biochemistry.,38(46):15129-40.

[3]. Belevich I, Verkhovsky MI, Wikström M.,(2006). Proton-coupled electron transfer drives the proton pump of cytochrome c oxidase. Nature.,440(7085):829-32.

[4]. Santana M, Kunst F, Hullo MF, Rapoport G, Danchin A, Glaser P.,(1992). Molecular cloning, sequencing, and physiological characterization of the qox operon from Bacillus subtilis encoding the aa3-600 quinol oxidase. J BiolChem. ,267(15):10225-31.

[5]. Chepuri V, Lemieux L, Au DC, Gennis RB.,(1990). The sequence of the cyo operon indicates substantial structural similarities between the cytochrome o ubiquinol oxidase of Escherichia coli and the aa3-type family of cytochrome c oxidases. J Biol Chem., 265(19):11185-92.

[6]. García-Horsman JA, Barquera B, Rumbley J, Ma J, Gennis RB.,(1994). The superfamily of heme-copper respiratory oxidases. J Bacteriol.,176(18):5587-600.

[7]. http://www.ncbi.nlm.nih.gov

[8]. Sayers EW.et al. , (2012). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 40(Database issue): D13-25.

[9]. Papadopoulos JS and Agarwala R (2007). COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics. 23:1073-79.

[10]. Desper R and Gascuel O., (2004). Fast Minimum Evolution, MolBiolEvol. 21:587-98. PMID: 14694080.

[11]. Saitou N and Nei M., (1987).Neighbor Joining, MolBiolEvol. 4:406-25.

[12]. Grishin N V.,(1995). Approximates the same model as Kimura, but can be computed for fraction of mismatched amino acids larger than 0.75, J MolEvol. 41:675-79.

[13]. Grishin N V., (1995). More general evolutionary model: substitution rates vary for both amino acids and sites, J MolEvol. 41:675-79.

[14]. Gomase VS and Tagore S., (2009).Phylogenomics: evolution and genomics intersection, Int. J. Bioinformatics Research and Applications. Vol. 5, No. 5: pp.548–563.

[15]. Waghmare S, Waghmare D, Bhatnagar PS (2013). Species Diversity of Short Horned Grasshopper (Orthoptera: Acrididae) in Selected Grasslands of Solapur District, Maharashtra, India. J Biodivers Endanger Species. 1:110.

[16]. Marchler-Bauer A et al., (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins." Nucleic Acids Res.39(D):225-9.

[17]. Marchler-Bauer A. et al., (2009). "CDD: specific functional annotation with the Conserved Domain Database." Nucleic Acids Res.37(D):205-10.