 doi:10.20944/preprints202011.0294.v1

Research Article

Comparative analysis of the mitochondrial proteins reveals complex structural and functional relationships in Fasciola species

Hafiz Ishfaq Ahmad[1,2], Muhammad Bilal Bin Majeed1, Muhammad Ijaz[3], Muhammad Zulfiqar Ahmad[4], Babar Maqbool[3], Khalid Mehmood[5], Hamid Mustafa Pasha[1], Chen Jinping[2]*

[1] Department of Animal Breeding and Genetics, University of Veterinary and Animal Sciences, Lahore, Pakistan.
[2] Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Institute of Applied Biological Resources, Guangzhou, Guangdong, China.
[3] Department of Clinical Medicine, University of Veterinary and Animal Sciences, Lahore, Pakistan.
[4] Guangdong Provincial Key Laboratory of Plant Molecular Breeding, Guangdong Sub-center of National Center for Soybean Improvement, College of Agriculture, South China Agricultural University, Guangzhou 510642, Guangdong, China.
* Correspondence: chenjp@giabr.gd.cn

Abstract: Mitochondria is a cellular source of energy, playing an essential role in cellular stress induced by environmental stimuli. The genetic diversity of mitochondrial genes involved in oxidative phosphorylation affects the production of cellular energy and regional adaptation to various ecological (climatic) pressures influencing amino acid sequences (variants of protein). However, a little is known about the combined effect of protein changes on cell-level metabolic alterations in simultaneous exposure to various environmental conditions, including mitochondrial dysfunction and oxidative stress induction. Present study was designed to address this issue by analyzing the mitochondrial proteins in Fasciola species including Cytochrome C oxidase (COX1, COX2, COX3 and CYTB) and NADH dehydrogenase (ND1, ND2, ND3, ND4, ND5 and ND6). Mitochondrial proteins were used for a detailed computational investigation using available standard bioinformatics tools to explore structural and functional relationships. Our analysis shows that the mitochondrial protein family of Fasciola species are extensively diversified in all species studied, showing an extending role in various biological processes The results showed that the protein of COX1 of F. hepatica, F. gigantica and F. jacksoni consist of 510, 513 and 517 amino acids respectively. The alignment of proteins showed that these proteins are conserved in the same regions at ten positions in COX and CYTB proteins while at twelve locations in NADH. Three dimensional structure of COX, CYTB and NADH proteins were compared and the differences in additional conserved and binding sites in COX and CYTB proteins as compared to NADH were found in three Fasciola species. These results, based on the amino acid diversity pattern, were used to identify sites in the enzyme and the variations in mitochondrial proteins among Fasciola species. This study provides valuable information for future experimental studies including identification of therapeutics, diagnostics and immunoprophylactic interests with novel mitochondrial proteins.

Keywords: Cytochrome oxidase; NADH dehydrogenase; *F. hepatica*; *F. gigantica*; COX; CYTB

1. Introduction

Liver fluke (*Fasciola hepatica)* is a parasite of cattle and other ruminants, and less frequently, human. It belongs to the class Trematoda [1]. Numerous wild mammals, especially ungulates, rodents and hares, contribute to the maintenance of the natural Fascioliasis emphases in various regions [2]. Human Fascioliasis occurs mostly in South America, North Africa, Iran and Western Europe [3, 4]. Fascioliasis is an economically significant animal disease affecting more than 300

million animals (cattle, sheep, buffalo and goats) [5, 6]. The WHO recognizes it as an important neglected tropical disease with estimates ranging from 2.4 to 17 million infected people [7]. Helminth parasites *F. gigantica* and *F. hepatica*, commonly referred to as tropical and temperate liver flukes, respectively are the causative agents of this disease. *F. gigantica* is widespread all over Asia and Africa [8]. The origins of members of the Fasciolinae particularly *F. hepatica* and *F. gigantica* are more challenging to discern. *F. gigantica* is common and prevalent in Africa, Asia, and Hawaii, and *F. hepatica* is likely of Eurasian origin given its host preference for *L. truncatula* of that region, while *F. jacksoni* is a parasite of the Indian elephant [9]. It seems likely that a host switch from planorbids to lymnaeids occurred in Eurasia and that this favored the emergence of the Fasciolinae, with colonization of Africa occurring secondarily, both by *F. gigantica* and an apparent ancestor of *F. nyanzae* in hippos and *T. tragelaphi* in sitatungas, both hosts sharing common habitats [10]. Global economic losses attributed to Fasciolosis are estimated to exceed US$3 billion per annum [6]. Economic losses driven by *F. gigantica* infections alone are estimated at US$2.4 billion and US$0.84 billion in Asia and Africa, respectively [5]. Millions of people are estimated to be infected worldwide, and more than 180 million people worldwide are at risk for this disease [11]. To date, there is no available effective vaccine for Fascioliasis. Morphological characteristics such as body shape and perimeter and the length to width ratio are usually used to identify adult Fasciola worms [12]. Because of considerable variation and overlap in measurements between *Fasciola* species such phenotypic criteria, however, are unreliable for specific identification and differentiation [13]. Mitochondrial markers have been used for phylogenetic characterization and tracking the dispersal route the dispersal route of the *Fasciola* species [14].

Disease diagnosis requires different aspects of research including serological methods, imaging and clinical findings [15]. Consideration has been paid on the proteome assessment of the affected person to overcome the diagnostic complications [16]. Because of the parasite's genetic diversity its control in the endemic regions is challenging. Certain geographic areas of the world, however, have achieved degrees of success [17]. It is, thus, necessary to explore all the aspects of the disease including vital proteins, strain specifications and affinity to specific targets for organs [18, 19]. New approaches including bioinformatics analyses, which are remarkable tools to understand the pathogenicity and to identify the particular strain affinity to a target organ, should be considered to understand various aspects of genes and proteins, especially the protein models [20]. Bioinformatics is an interdisciplinary investigation approach that has been used to comprehend the biological and molecular functions of genes and proteins in living systems including biological macromolecules such as DNA, RNA and proteins [21, 22]. Use of computer-assisted methodologies help the investigators to understand the association between protein function and structure and to find precise targets for drugs and vaccines [23]. These approaches have been recognized to identify the dominant vaccine epitopes, interactions between function and structure and metabolic and putative protein roles [24]. Several proteins of the *Fasciola* species have been investigated in different studies. Regardless of other proteins in the *Fasciola* species, Cytochrome C oxidase, and NADH dehydrogenase (NADH) with oxidoreductase activity have been less investigated. This study aimed to explore the discrepancies in mitochondrial (COX, CYTB, and NADH) and associated proteins of Fasciola spp. using bioinformatics. A comprehensive study of the primary, secondary and tertiary structures of mitochondrial proteins of *Fasciola* species was performed to reveal the structural and functional configuration of the different domains present in correlation among *Fasciola species*. A thorough analysis shows that many of the basic amino acids that control the structural and functional topographies in *Fasciola* species are conserved.

2. Materials and Methods

2.1. Retrieval of sequences and analysis

The amino acid sequences of mitochondrial proteins of *F. hepatica, F. gigantica and F. jacksoni* were retrieved from Genbank and UniProt database [25]. The homologous sequences for these proteins were searched and selected using standard parameters and were envisaged using the open-source

 doi:10.20944/preprints202011.0294.v1

tool, Aliview [26]. Multiple sequence alignment of mitochondrial proteins and the selected orthologues were performed using ClustalOmega [27] from EMBL EBI Web server. The secondary structure of these proteins was generated by PSIPRED [28] and the intrinsically disordered structure was predicted using CSpritz [29].

2.2. Prediction of the protein-ligand binding site and docking analysis

Protein-ligand binding sites were analyzed to predict the number of drug binding sites in order to comprehend the interaction of proteins and the ligands by using LPIcom's webserver. The server's analytics module can recognize preferred sites in interaction and binding motif for a given ligand. The evaluation unit of the server permits multi-ligand protein-binding residues to comprehend the relationship among ligands based on their binding residues [30]. The online servers COACH [31] was used to recognize the number of binding residues and their positions using I-Tasser. COACH is a meta-server for predicting ligand binding sites by using two relative TM-SITE methods [32] and S-SITE. Furthermore, BioLiP protein function database was used to identify the ligand binding residues in mitochondrial proteins of *Fasciola* species [33].

2.3. 3D protein modeling and structural analysis

Online programs Phyre2 (http:/www.sbg.ic.ac.uk / phyre2 / html) and the Swiss model (http:/swissmodel.expasy.org) were used to generate the crystal structure of COX, CYTB and ND proteins [34]. Modeling method for homology was used to predict the structures. The I-TESSAR and Swiss modelling approaches [35] predicted the three-dimensional (3D) structures of these proteins to predict appropriate structures. Using the conjugate gradient algorithm and the Amber force field in UCSF Chimera 1.10.1 [36], the constructed target proteins have been minimized. Besides this, the ProSA webserver [37] was used to assess the stereo-chemical characteristics of the predicted mutated structures. To predict their theoretical PIs, extinction coefficients, aliphatic and instability indices, and GRAVY values, the ProtParam tool was applied [38, 39].

2.4. Prediction of Post-translation modification (PTM) sites

The post-translation modification sites were identified by using the ModPred (http:/www.modpred.org/) webserver within the mitochondrial proteins of *Fasciola* species. This determines the potential PTM residues within the target proteins (post-translational modification) based on sequences. This web server comprises of 34 sets of logistical regression models accomplished discretely with the group of 126,036 non-redundant sites which are experimentally validated for various polymorphism obtained from the public databases [40].

2.5. Analysis of KEGG enrichment

To investigate the molecular and biological significance of these proteins involved in various pathways, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis using the KOBAS 2.0 webserver and Fisher's exact statistical test [41].

3. Results

In this study, the structural and functional characteristics of the mitochondrial proteins, including cytochrome c oxidase (COX1, COX2, COX3 and CYTB) and NADH dehydrogenase (ND1, ND2, ND3, ND4, ND5 and 143 ND6) were explored. Homologs of these proteins were identified through analyzing the amino acid sequences, using a comprehensive BLASTP search approach.

3.2. Protein domain analysis

The SMART [42], an online program, was used to recognize the domains of the selected proteins. To understand the working principle of these proteins, the structures of the predicted protein domains were analyzed by data mining from the Prosite database of proteins domains and the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

families available at URL https://prosite.expasy.org. Moreover, we used DOG.2 software for concise presentation of the picture [43]. We identified the terminal domains in cytochrome c oxidase proteins (Figure 1). In contrast, in dehydrogenase proteins, NADH-dehydrogenase, proton-conducting membrane transporter and NADH-ubiquinone domain were identified (Figure 2). This domain can also be found in the subunit (CI-B8) 1.6.5.3 of mitochondrial NADH-ubiquinone oxidoreductase B8. It is not yet clear whether oxidoreductase only or they are also part of ribosomal proteins. Moreover, the positioning of active binding sites was identified on conserved domains. The crystal structures of mitochondrial proteins of F. hepatica were used as the reference sequence and binding positions were shown onto the crystal structure using UCSF Chimera tool. Several residues in the COX1 domain that contains the ligand-binding sites were identified as well.

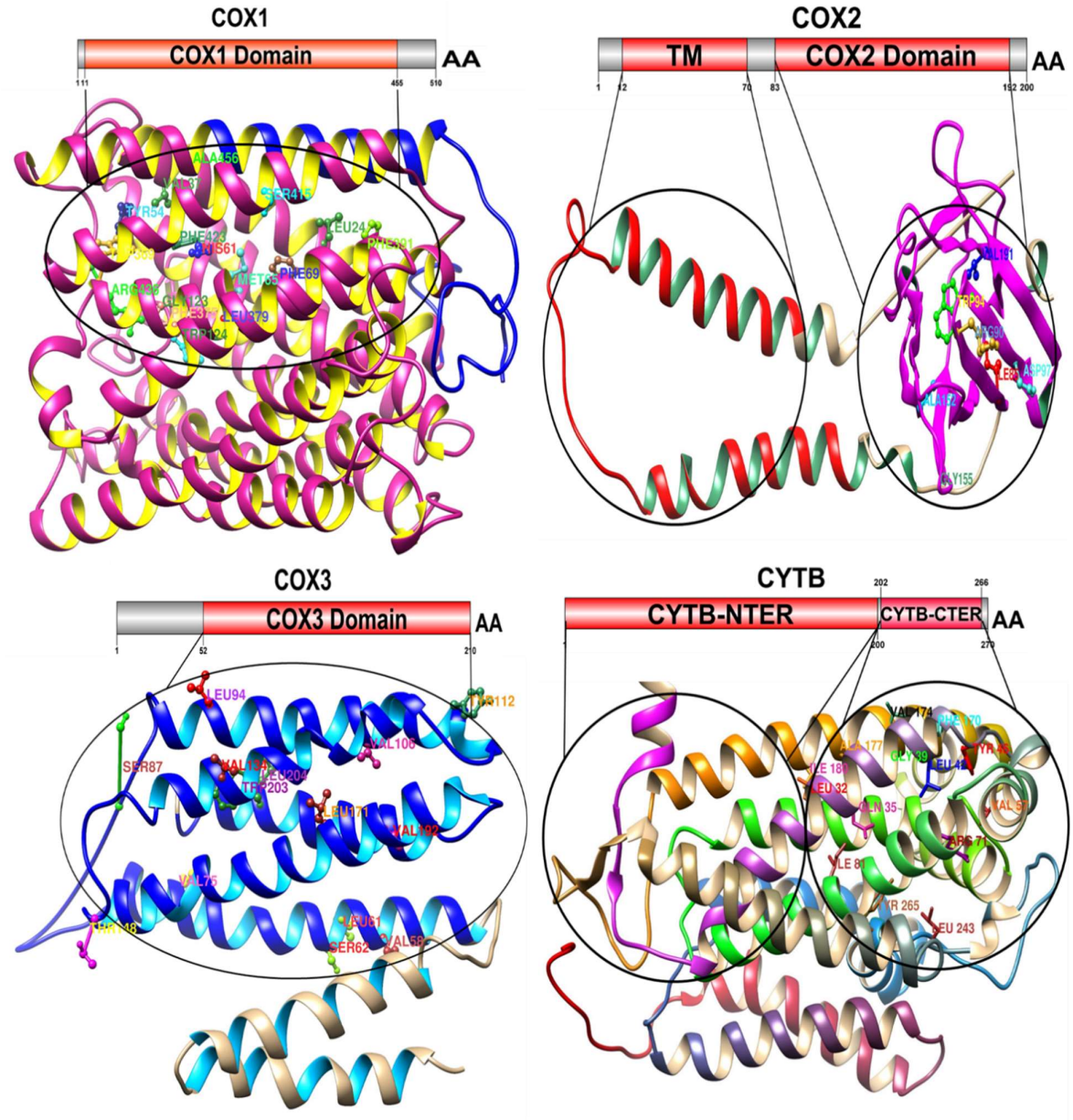

Figure 1. Domain composition of COX and CYTB proteins. The domain compositions of each protein are indicated along the amino acid sequence of each protein by specifying the start and the endpoint of the functional domain. Different colors are used to show several domains in a protein.

 doi:10.20944/preprints202011.0294.v1

3.2. 3D structure predictions and modeling

The amino acid sequences of mitochondrial proteins were downloaded from NCBI and crystal structure was generated by online web servers including Swiss modelling (http://swissmodel.expasy.org) and Phyre 2 (http://www.sbg.bio.ic.ac.uk/phyre2/html). Only top interacting and strongly bonded ligand were selected and the results of docking are shown with interacting residue showing protein-ligand sites at positions 24, 54, 56, 61, 69, 124, 379, 391, 423, and 436 that were identified in COX1 protein by ligand prediction in I-TESSAR. Six binding sites in COX2, eleven residues in COX3 and fourteen in CYTB proteins were found (Figure 1). Furthermore, we identified the binding residues in ND proteins but not in ND6 (Figure 2).

3.3. Motif prediction

The calculated sequence alignment of the most common residues of nucleotides or amino acids found at each position is called the consensus sequence. Protein sequences were aligned with MEGA6 [44] and ENDscript [45] and photographs were captured using GenDoc [46] and Modeller 9.14 [47]. A BLAST search using the PDB as a query sequence against the database of COX, CYTB, and NADH proteins was performed. The consensus alignment of these proteins showed that these are conserved in the same regions at ten positions in COX and CYTB proteins and at twelve locations in NADH. Three dimensional structure of COX, CYTB and NADH proteins were compared and the differences in additional conserved and binding sites in COX and CYTB proteins as compared to NADH are shown (Supplementary data). The sequences were compared through the multiple alignments and the similar sequence motifs were calculated. The motif distribution analysis was performed using MEME tool which identified the motifs in protein sequences. This pooled high patterns of conservation from motifs one to five (Figure 3; Table 1). All patterns were observed in protein sequences of all Fasciola species. It was found that pattern 1-4 were absent in F. hepatica, F. gigantic, and F. jacksoni in COX3, whereas motifs 2 and 3 were absent in all three species in COX1 (Figure 3). The absence of motifs in different species mean the divergence of the characteristics of the structural gene concerning exon-intron relations. These analyses showed that the variances in the sharing of motifs in these proteins of Fasciola species during adaptive evolution may have deviated from the functions of those genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

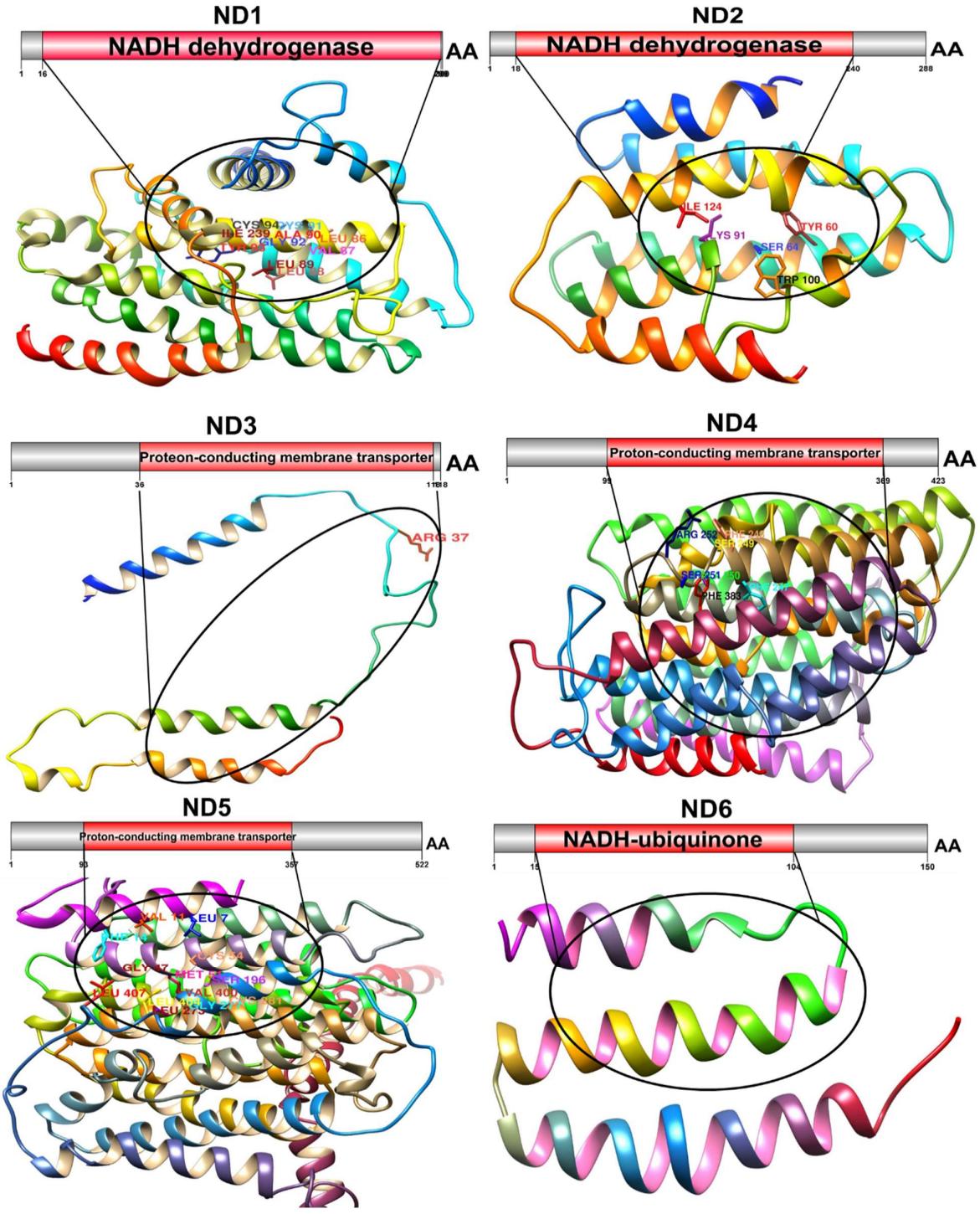

Figure 2. Domain composition of NADH proteins. The domain compositions of each protein are indicated along the amino acid sequence of each protein by specifying the start and the endpoint of the functional domain. Different colors are used to show several domains in a protein.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

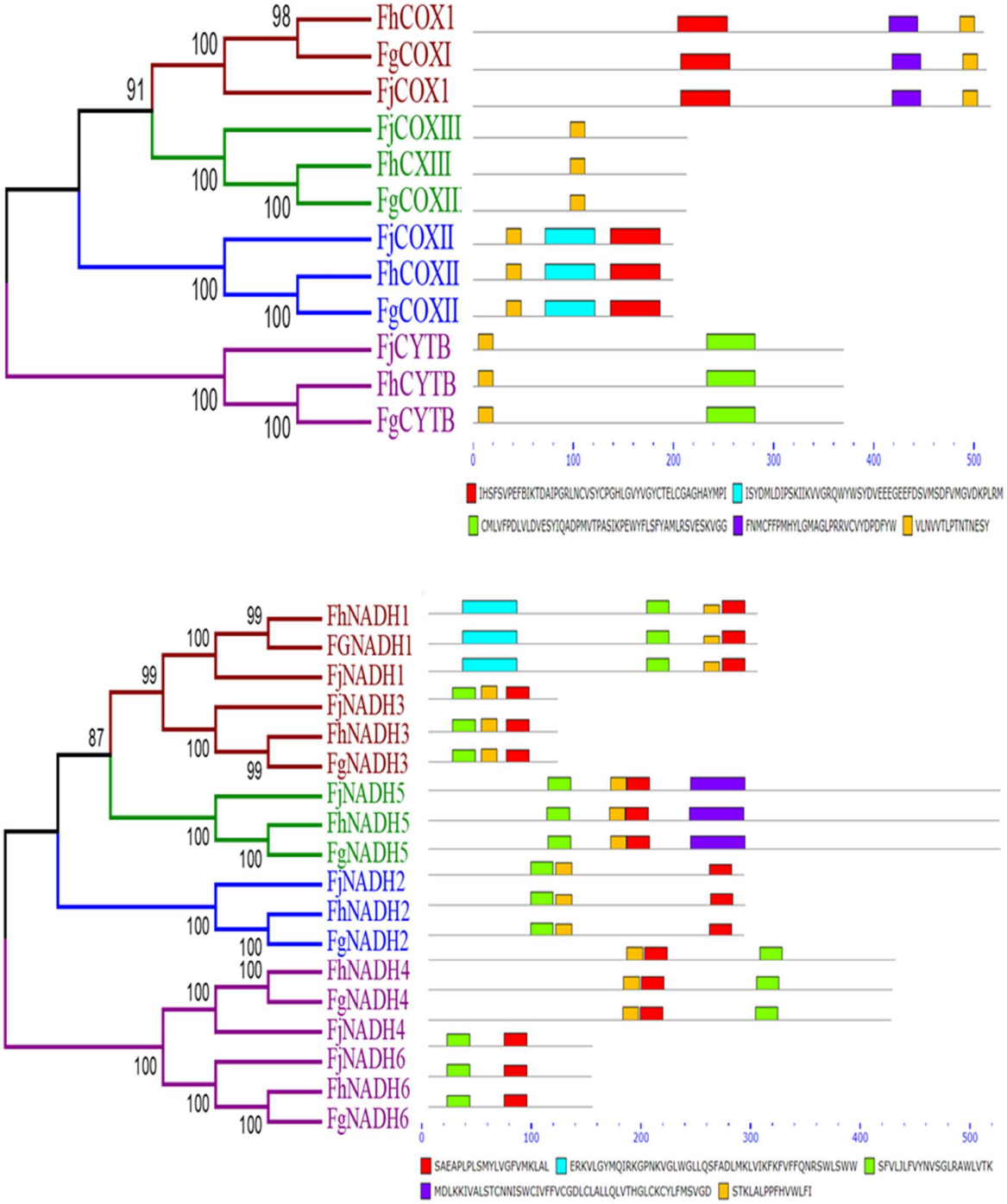

Figure 3. Phylogeny and motif distribution pattern of mitochondrial proteins within the Fasciola species. These are predicted using an amino acid sequence-based MEME suite (http:/meme.nbcr.net/meme/). Five conservative color motifs separate all sequences.

 doi:10.20944/preprints202011.0294.v1

Table 1. Details of motifs predicted using MEME suite based on amino acid sequences.

Motif detail of COX and CYTB			
No.	Sequence	Length	Description
1	IHSFSVPEFBIKTDAIPGRLNCVSYCPGHLGVYVGYCTELCGAGHAYMPI	50	Cytochrome C oxidase subunit II/periplasmic domain
2	ISYDMLDIPSKIIKVVGRQWYWSYDVEEEGEEFDSVMSDFVMGVDKPLRM	50	Cytochrome C oxidase subunit II, periplasmic domain
3	CMLVFPDLVLDVESYIQADPMVTPASIKPEWYFLSFYAMLRSVESKVGG	49	Cytochrome b(C-terminal)/b6/petD
4	FNMCFFPMHYLGMAGLPRRVCVYDPDFYW	29	Cytochrome C and Quinol oxidase polypeptide I /Bacterial signalling protein N terminal repeat
5	VLNVVTLPTNTNESY	15	No Description
Motif detail of NADH			
No.	Sequence	Length	Description
1	SAEAPLPLSMYLVGFVMKLAL	21	Proton conducting membrane transporter
2	ERKVLGYMQIRKGPNKVGLWGLLQSFADLMKLVIKFKFVFFQNRSWLSWW	50	NADH Dehydrogenase
3	SFVLJLFVYNVSGLRAWLVTK	21	Proton conducting membrane transporter
4	MDLKKIVALSTCNNISWCIVFFVCGDLCLALLQLVTHGLCKCYLFMSVGD	50	Proton-conducting membrane transporte/ TspO/MBR family
5	STKLALPPFHVWLFI	15	Proton conducting membrane transporter

3.4. Prediction of ligand binding sites and interaction of amino acids with ligands

In Fasciola species, the binding sites in mitochondrial proteins were predicted using the accessible FTSite server at (http:/ftsite.bu.edu/) to determine the epitope or binding regions in those proteins. The FTSite server executes energy-based processes and has 94 percent experimental accuracy for identifying ligand binding residues. For the target proteins the FTSite server was used to identify the ligand binding sites. For this native proteins were used to recognize ligand binding sites. It was found that the ligand binding site one has arg-99, ser-98, Val-107, leu-113, his-226, try-229, leu-445 and Val-447, ligand binding site 2 has ser-98, leu-101, cys-108, his-226 and ser-446 and ligand site 3 has ser-191, leu-378, Ile-389, Val-394 and leu-400 (Figure 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

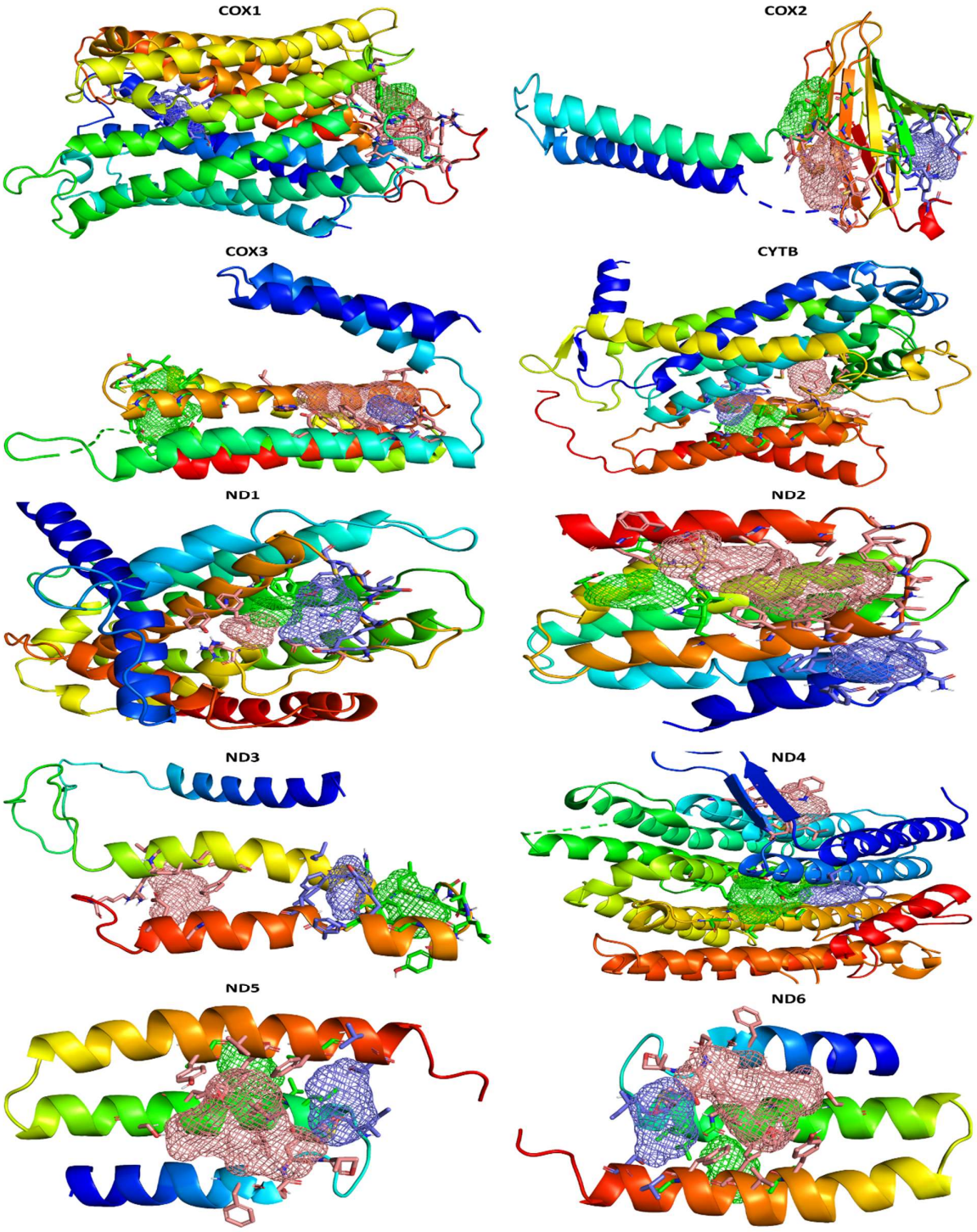

Figure 4. FT website forecast showing ligand binding sites. The first, second and third ligand binding sites are rose, green, and blue colored mesh.

To further validate the reliability of the predicted structure the protparam program was used to compute the physicochemical parameters of the predicted structures of the target proteins. By doing so, the values for physicochemical parameters including theoretical pI, estimated half-life, aliphatic index, and extinction coefficients were obtained. The principle of pK values was considered for computing the PI score of the target proteins. PI scores of 6.69, 5.9, 6.52 and 5.8 were obtained for COX1 proteins in F. hepatica, F. gigantic, and F. jacksoni, respectively. These results are comparable

 doi:10.20944/preprints202011.0294.v1

to ranges of PI values (Table 2) reported for the proteins in earlier literature [48]. This shows that the current predicted values are in a standard range (Table 2). The GRAVY values, a measure of hydropathy characteristics of the residues of the particular proteins [49], were determined. The negative values show that they are more hydrophobic and the positive value indicates that they are more hydrophilic in nature [50] as presented in Table 2.

Table 2. Physiochemical analyses of the predicted structures.

Protein	*Species*	*Amino Acids*	*Molecular weight (D)*	*ERRAT score*	*VERIFY 3D (%)*	*Theoretical pI*	*EC*	*AI*	*II*	*GRAVY*
	F. hepatica	*510*	*56463.95*	*94.12*	*76.73*	*6.69*	*83310*	*120.41*	*35.95*	*1.065*
COX1	***F. gigantica***	***513***	*56932.57*	*94.83*	*87.09*	*5.9*	*145800*	*119.94*	*32.48*	*1.008*
	F. jacksoni	*517*	*57588.28*	*95.06*	*78.39*	*5.86*	*147290*	*119.57*	*33.01*	*0.974*
	F. hepatica	*200*	*22695.37*	*96.5*	*83.73*	*4.42*	*45880*	*102.55*	*33.25*	*0.517*
COX2	***F. gigantica***	***200***	*22725.39*	*97.23*	*94.4*	*4.39*	*47370*	*101.6*	*39.55*	*0.505*
	F. jacksoni	*200*	*22757.43*	*96.48*	*89.48*	*4.46*	*45880*	*102.1*	*36.64*	*0.451*
	F. hepatica	*213*	*24264.92*	*97.56*	*93.7*	*4.73*	*69495*	*135.31*	*38.18*	*1.22*
COX3	***F. gigantica***	***213***	*24538.25*	*98.04*	*91.81*	*4.7*	*69370*	*130.19*	*36.42*	*1.271*
	F. jacksoni	*214*	*24359.19*	*98.63*	*88.3*	*5.43*	*59400*	*136.45*	*21.32*	*1.236*
	F. hepatica	*300*	*34357.25*	*89.62*	*97.58*	*8.42*	*93110*	*113.87*	*34.01*	*1.152*
ND1	***F. gigantica***	***300***	*34279.12*	*96.56*	*77.24*	*5.86*	*147290*	*119.57*	*33.01*	*0.974*
	F. jacksoni	*300*	*34388.4*	*90.06*	*91.51*	*8.32*	*98735*	*117.13*	*33.8*	*1.131*
	F. hepatica	*288*	*32365.07*	*93.59*	*90.07*	*7.95*	*94140*	*130.1*	*33.06*	*1.486*
ND2	***F. gigantica***	***288***	*32292.96*	*95.58*	*93.36*	*7.36*	*87275*	*126.04*	*36.91*	*1.481*
	F. jacksoni	*289*	*32540.24*	*94.55*	*78.15*	*6.01*	*86120*	*125.61*	*36.19*	*1.464*
	F. hepatica	*118*	*14027.74*	*88.06*	*78.39*	*5.64*	*26930*	*131.95*	*35.96*	*1.165*
ND3	***F. gigantica***	***118***	*14057.68*	*92.55*	*86.9*	*5.64*	*26930*	*119.49*	*33.96*	*1.136*
	F. jacksoni	*118*	*14035.83*	*90.24*	*84.61*	*6.03*	*27055*	*126.95*	*41.46*	*1.114*
	F. hepatica	*423*	*47367.06*	*96.37*	*94.17*	*6.18*	*117895*	*131.58*	*33.39*	*1.319*
ND4	***F. gigantica***	***422***	*47364.91*	*88.16*	*96.81*	*5.93*	*114915*	*127.94*	*31.84*	*1.306*
	F. jacksoni	*426*	*47289.93*	*97.62*	*77.15*	*5.15*	*103915*	*135.02*	*34.13*	*1.33*
	F. hepatica	*522*	*58377.86*	*85.47*	*92.64*	*5.31*	*156770*	*121.72*	*29.78*	*1.313*
ND5	***F. gigantica***	***521***	*58284.73*	*95.29*	*96.54*	*5.16*	*148280*	*122.88*	*31.22*	*1.316*
	F. jacksoni	*522*	*57672.2*	*94.9*	*96.88*	*7.34*	*142540*	*126.93*	*29.61*	*1.319*
	F. hepatica	*150*	*16366.63*	*96.5*	*81.97*	*5.3*	*17920*	*127.73*	*26.58*	*1.549*
ND6	***F. gigantica***	***150***	*16302.46*	*92.67*	*95.34*	*4.3*	*24910*	*125.13*	*33.38*	*1.511*
	F. jacksoni	*149*	*16281.78*	*95.48*	*88.16*	*5.2*	*23420*	*143.02*	*27.89*	*1.742*
	F. hepatica	*370*	*41544.67*	*88.78*	*96.35*	*7.05*	*83810*	*122.59*	*40.02*	*1.084*
CYTB	***F. gigantica***	***370***	*41701.84*	*92.45*	*93.64*	*7.05*	*93320*	*119.65*	*37.66*	*1.055*
	F. jacksoni	*370*	*41589.57*	*96.13*	*97.45*	*6.69*	*83685*	*120.41*	*35.95*	*1.065*

3.5. Consensus analysis of mitochondrial proteins

Consensus structural alignment of COX, CYTB and NADH proteins was carried out using ENDscript. The structural aliments were produced by comparing with top 20 homology structures. The ENDscript result was alternatively visualized by PyMOL 3D visualization software. The results were justified by the color displayed in the structural consensus alignment graph and the thickness of the tubes.

 doi:10.20944/preprints202011.0294.v1

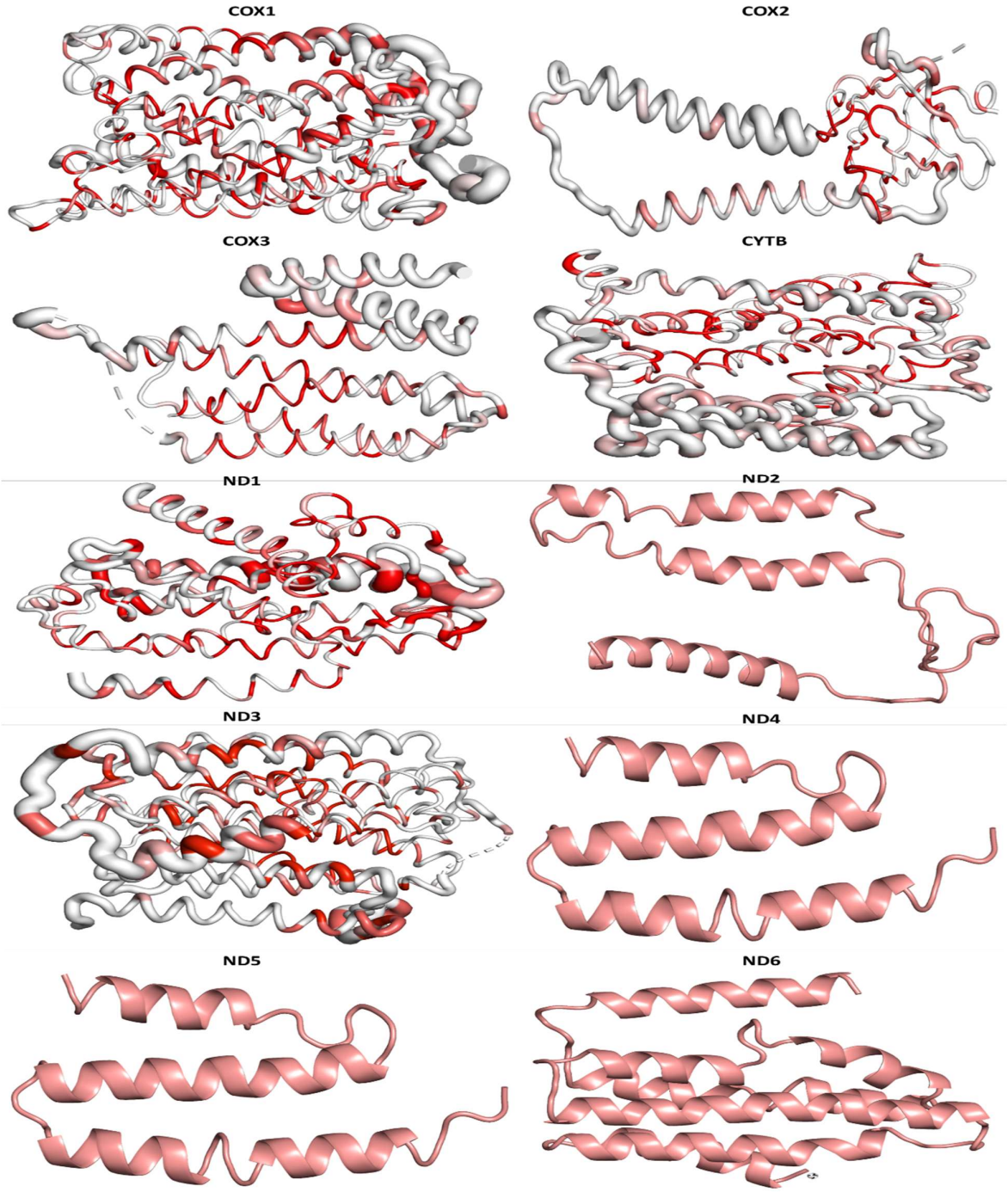

Figure 5. Predicted consensus structural alignment of COX, CYTB and NADH proteins. The color of the tubes indicates the level conservations. Red tube signifies low conservation.

3.6. Ligand consensus module

Ligands play a vital role in the expression and regulation of a protein. Three-dimensional structure of the protein is changed because of the ligand-binding interaction. The ligand binding follows the intermolecular binding forces such as ionic bonds, hydrogen bonds, hydrophobic interaction and Vander-Waals forces. These modifications in the conformational state of the protein may inhibit or activate some functions of the protein. Ligand binding interaction analysis was, thus, performed using a web server (http://crdd.osdd.net/raghava/lpicom) based on the physiochemical characteristics of amino acids. It was found that the COX1 protein has six ligands and two interacting

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

ions that interact with different ligands. The results have shown that four ligands (HEA, OXY, PEK and PDV) were interacting with critical residues such as alanine, cysteine, aspartic acid, glycine, histidine, lysine, leucine, arginine, serine, threonine, tryptophan, tyrosine and valine. The charged residues, particularly the basic amino acid residues, are more favored in HEA, OXY, PEK and PDV interaction compared to DMU and CUA interaction (Figure 6). The smaller and polar residues are correspondingly characterized in all ligands. Whereas in CYTB, five ligands were interacting with various amino acid sites, and four ligands (NAP, FMN, SEF and SF4) were showing interaction with critical residues (Figure 6).

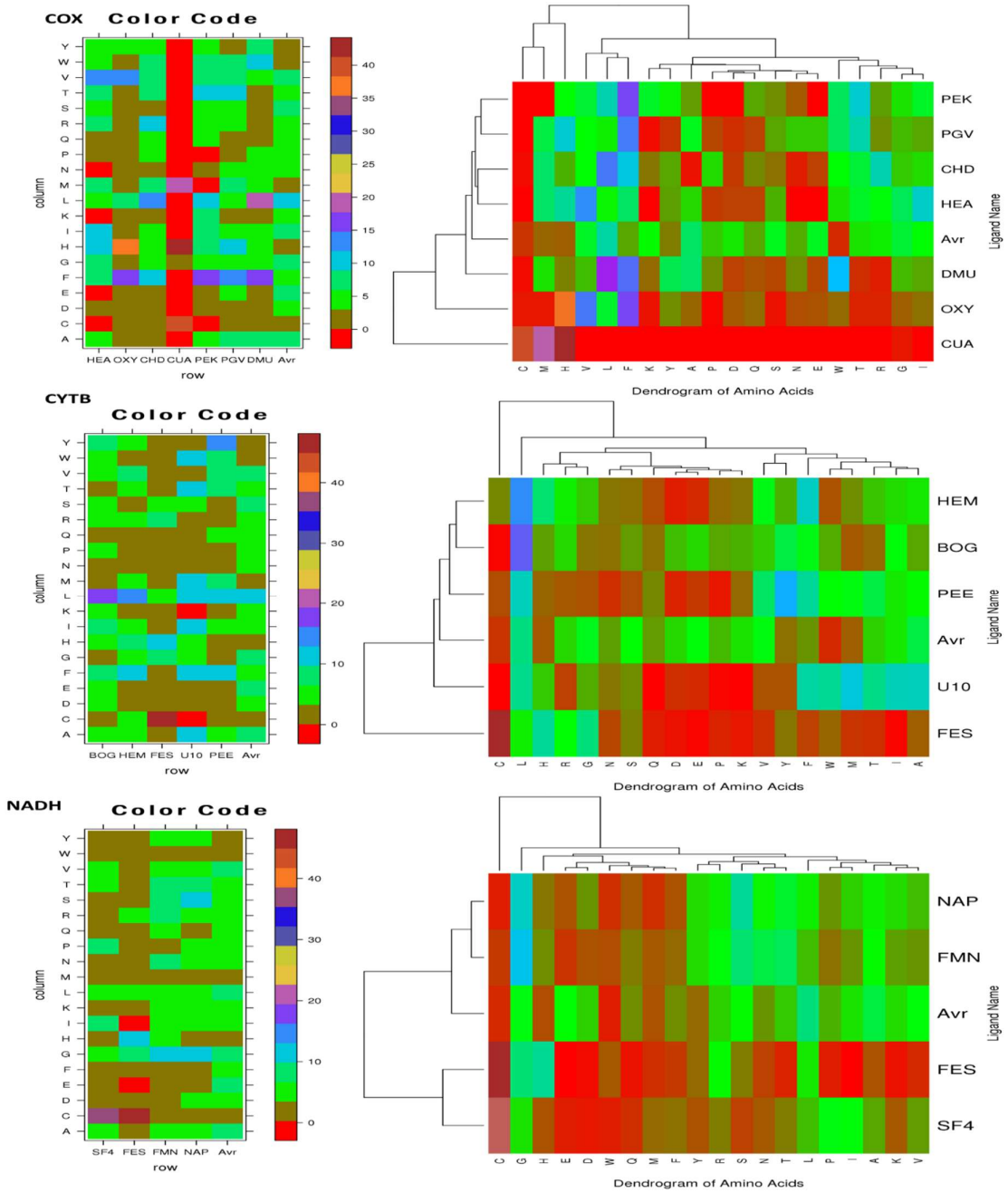

Figure 6. The clustering of interacting amino acid residues. This displays the composition of different ligand binding residues (left) and clustering of amino acids based on the physicochemical features of ligand-interacting residues.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

3.7. Predicted protein-protein interaction evidence view

Protein-protein interaction (PPI) is one of the essential molecular networks known to affect a multitude of biological functions. Protein-protein interaction analysis of all target proteins was undertaken to assess the information flow networks among these mitochondrial and other proteins. The interaction analysis was performed online STRING software (http://string91.embl.de/) [51] and visualized by Cytoscape software [52]. The interaction networks are justified by nodes and colored lines between nodes. In the PPI network COX, CYTB and NADH are interacting with the other key proteins which are co-expressed (Figure 7).

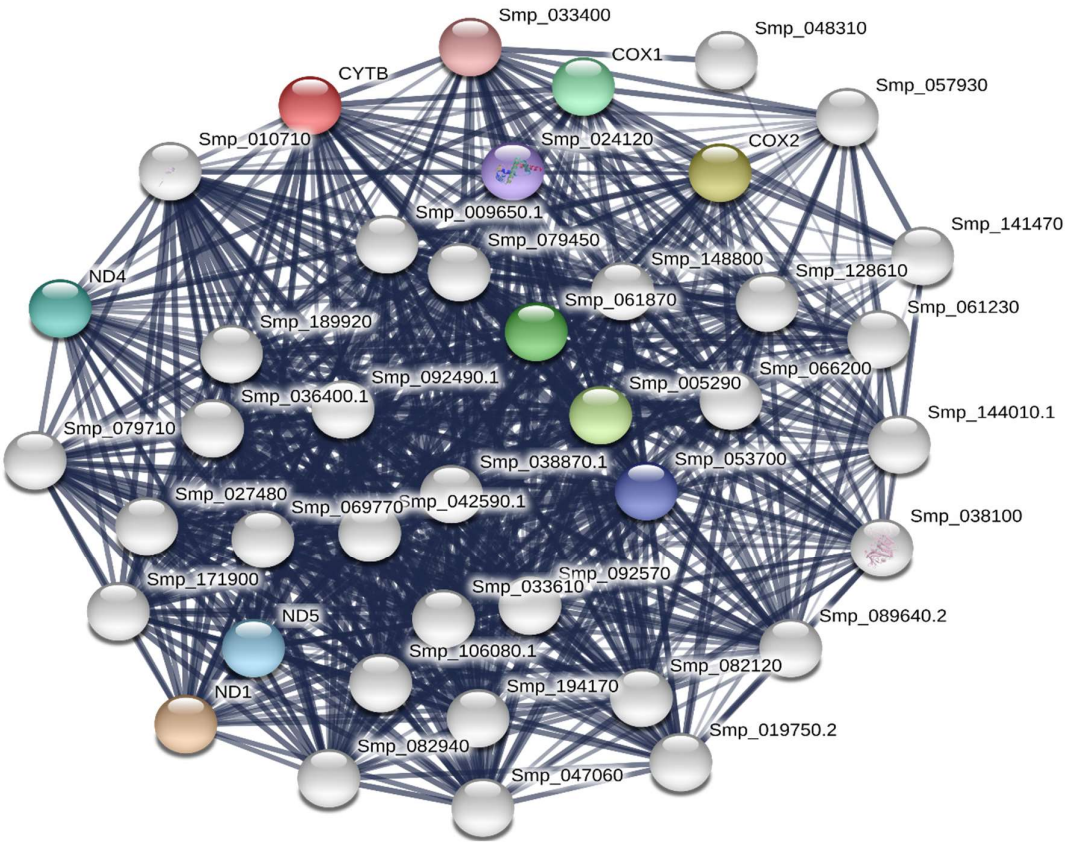

Figure 7. The protein-protein interaction (PPI) network built by STRING database. Line thickness indicates the strength of the interaction. The lengths of the connectors indicate the farness or closeness of the flow of information. Red node shows target protein; Empty nodes exhibit unknown proteins; Filled node shows known or predicted proteins; color lines between nodes indicate different types of biological interactions between proteins. .

3.8. Pathway analysis of mitochondrial proteins

The pathway enrichment analysis was performed and retrieved from KEGG pathway molecular interaction enrichment database found at http://www.genome.jp/kegg/pathway.html to understand the overall signaling pathway of oxidative phosphorylation through which the Cytochrome C oxidase and dehydrogenase enzymes are performing cellular processes. These mitochondrial proteins are enriched in complex signaling pathway and subcellular pathway of dehydrogenases and ATP synthesis. The KEGG pathway analysis revealed that mitochondrial proteins are one of the core signaling pathways regulating proteins thereby controlling several cellular processes including oxidative phosphorylation, cellular respiration and ATP synthesis (Figure 8). The subunit of cytochrome oxidase is the constituent of the respiratory chain catalyzing the reduction of oxygen into the water. Subunits 1-3 constitute the functional nucleus of the enzyme complex. Subunit-2 transfers

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 November 2020 doi:10.20944/preprints202011.0294.v1

the electrons from cytochrome C to the catalytic subunit's bimetallic center via its binuclear copper center. The complex b-c1 mediates the transfer of electrons from ubiquinol to cytochrome C (by similarity). The NADH dehydrogenase complex IV contributes to proton gradient generation throughout the mitochondrial membrane which is used for ATP synthesis. A core NADH dehydrogenase subunit (Complex I) with mitochondrial membrane respiratory chain is thought to be a part of the minimal assembly needed for catalysis.

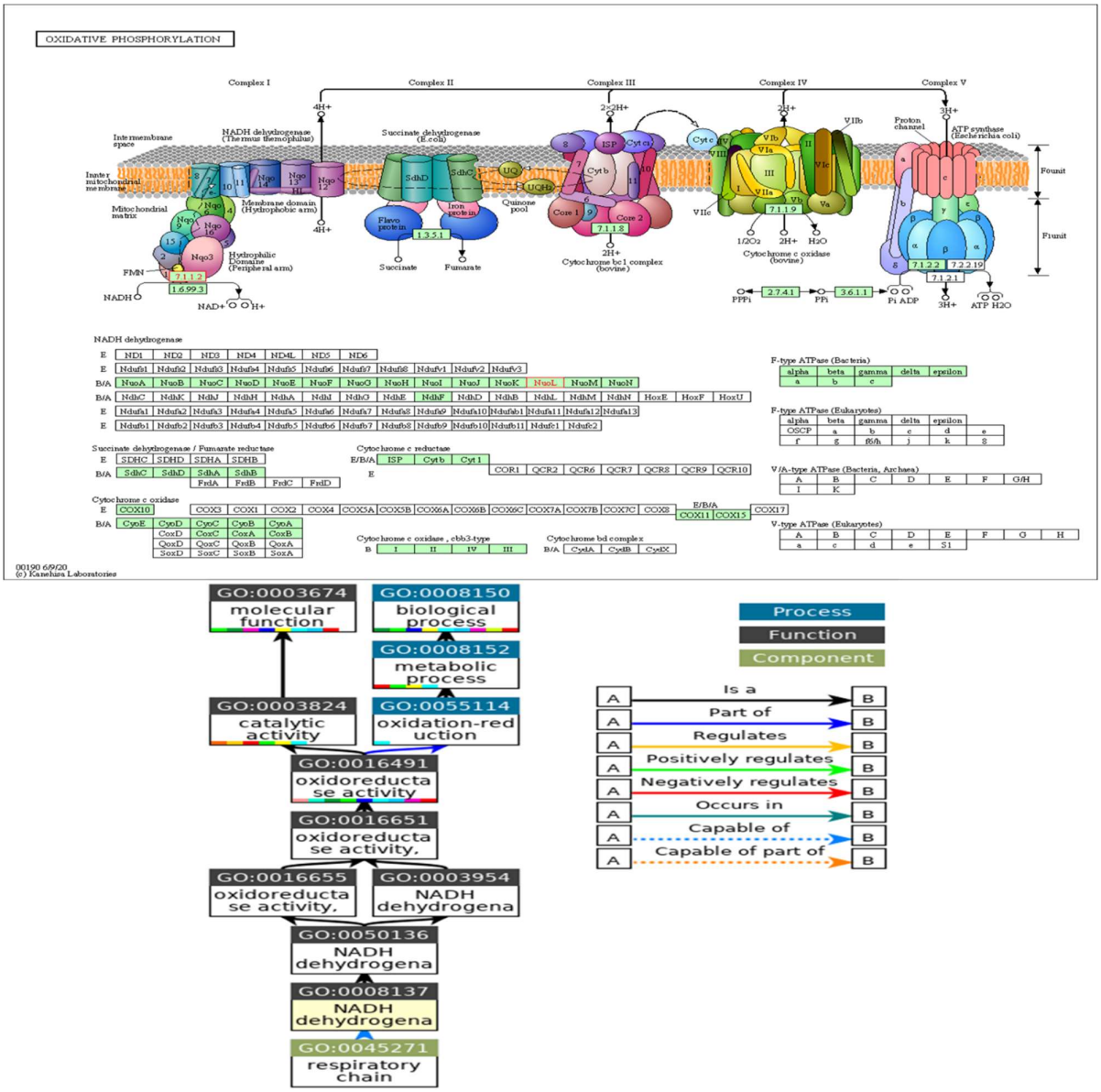

Figure 8. Signaling pathways in oxidative phosphorylation generated using KEGG pathways (Source: https://www.genome.jp/kegg/). Where the differentially expressed genes were marked. The enzymes recognized in all genes were highlighted green. The enzymes recognized in the acquiesced genes were highlighted red. The outcomes display that these genes were typically concerted in confined areas of the pathway.

 doi:10.20944/preprints202011.0294.v1

4. Discussion

Several studies have shown that *Fasciola* parasites are the causative agents of animal and human fascioliasis [53, 54]. The identification of these parasitic species, in previous studies, used to be based on the morphological and morphometric criteria. Due to morphological similarities and overlap in the values of most measurements, precise identification of Fasciola species based on the parasite speciation based on the aforementioned criteria cannot be decisive. In addition to the existence of *F. gigantica* and *F. hepatica*, an intermediate form between them has been reported from many Asian countries [55-59], which makes further difficulties in identification of Fasciola spp." Genetic methods using molecular markers are currently becoming increasingly important in parasite epidemiology especially when there is no reliable method to distinguish them morphologically [60].

The diversity of pathogens and their pathogenicity require more significant consideration in relation to various organs and hosts they affect [61]. A very few investigations have been carried out on COX and NADH proteins of other parasites than Fasciola to unravel their role in the reduction of oxygen to water in the cellular respiration and oxidoreductase [62]. In this study the differences in COX, CYTB and NAD protein sequences of *F. hepatica*, *F. gigantica*, and *F. jacksoni* have been identified. The secondary structure of COX and NAD of *F. hepatica, F. gigantica*, and *F. jacksoni* showed the differences in the construction of the coils, helix and extended enzymes. These results on the variations in alignment and secondary structures lead to differences in tertiary structures of these proteins. Tertiary protein structure of COX and NADH, functional protein forms, have been found to be identical to other NADH protein members in the ultimate function of proteins for drug susceptibility and vaccine target (Figures 1-3).

To the best of our knowledge, this is the first study on the *Fasciola* species to analyze the enzyme structure of COX and NAD. Biochemical analysis of COX and NAD in *Fasciola* species showed their index of instability to be below 40 (Table 2), which means that they are stable based on an unstable index [63]. Additionally, a comparison of the three-dimensional structures of COX proteins in *F. hepatica*, *F. gigantic*a, *F. Jacksoni* showed an extra helix of 21 carboxylate-terminal amino acids. Also, two conserved binding sites (His145 and His 174) observed in the aforesaid helix. Those COX binding sites could have different enzyme activity in three species of *Fasciola* (Figure 5). The *in silico analysis* of this study firstly performed 3D structural modeling for mitochondrial proteins of *Fasciola* species based on homology modeling principles. The predicted 3D models of COX, CYTB and ND proteins have shown different structural colors, characteristics, and conformations (Figure 1 and 2). Following 3D structural modeling, validations were carried out for all 3D modelled proteins. Moreover, the ProSA validation analysis revealed that predicted model of COX and CYTB have shown NMR and X-ray quality structures. The two properties, suggesting that the structure of the latter two proteins are not yet determined [64]. VERIFY-3D tool further confirmed the quality of the models and COX1, COX2, CYTB and ND1 had 92.75%, 94.77%, 89.87%, and 85.96% prediction accuracy, respectively. All the analyses indicate that the predicted structure can fit for further study.

In attempting to understand the ligands responsible for the functions of the mitochondrial proteins, the ligands and their binding residues were analyzed [16, 65]. Computational analysis during this study identified a total of 16 ligands in all target proteins. A total of 7, 5, and 4 ligands were identified in COX, CYTB and NADH, respectively with different residue lengths. In total, 65 amino acid residues were detected in all target proteins' ligands with 11, 28, and 26 residues in COX, CYTB and NADH respectively (Figure 6). Surprisingly, one uncharacterized ligand with unknown binding residue was detected both in COX1 and NADH indicating the complexity of additional proteins characterization. The predicted 3D models and generated ligand binding sites of the target proteins can further be used for the study of antigen-antibody interaction [66] and functional analysis binding small molecules. It, thus, can also be used as an input for protein docking software [67]. In this regard, all the target proteins in the current study can be investigated for drug binding analysis. Ligand-residue interaction analysis has shown that all identified ligands have different residue preferences. In the prediction analysis, for example, HEA and HEM ligand preferred asparagine and cysteine residue and SF4 and NAP preferred cysteine and histidine residue (Figure 6). Currently, all

the detected ligands have shown to be associated with all 20 amino acids residues showing more preference towards some amino acids than others.

The binding site of unbound structures of COX, CYTB and NADH proteins were predicted by employing FTSite tool [68] and PyMOL software (Figure 4). In all target proteins, three FT sites were detected with different amino acids residue compositions suggesting that the potential of mitochondrial proteins in small binding molecules provide evidence for therapeutic design and protein engineering [16, 68]. Further, the KEGG pathway analysis was employed to retrieve the pathway through which the mitochondrial proteins affect cellular functions. KEGG pathway analysis showed that NADH and its downstream receptors are mainly enriched in oxidative phosphorylation signaling pathway and guidance to regulate several cellular functions (Figure 8). Multiple sequence alignments were done by ENDscript software to expand our understanding of mitochondrial protein sequences and structure homology. The results of MSA alignment of these proteins showed several identical, similar, and alternate residues, however, only the similar and alternative residues were observed suggesting that NADH has less conserved residues (Figures A1 and A2). The comparative analysis of proteins of representative of Fasciola sp., *F. hepatica, F. gigantica* and using protein sequence data set exhibited that Fasciola sp. and *F. hepatica* were closely related than either was to *F. gigantica*. These findings were supported by the analysis of COX and NAD sequence data. The use of mitochondrial genomes could provide understandings in both inter-host and climatic adaptations developed by the parasites which may have been main factor in driving intergeneric change and speciation within the *Fascioloides* genus as seen has been proposed in other trematode species.

5. Conclusions

Our analysis shows that the mitochondrial protein family of Fasciola species are extensively diversified in all species studied, showing an extending role in various biological processes. Differences in the cytochrome c oxidase and NADH dehydrogenase sub-units are predicted to be fundamental to the determination of the specifics substrates that facilitate this functional diversity. These results propose that the two homologs could attain parallel catalytic and biochemical functions with the possible activity of oxidative phosphorylation and ATP synthesis. These analyses helped to understand the structural and functional homology of vital proteins. This work provides the most comprehensive analysis of these proteins, which will offer a strong background for studying these genes across model systems in other parasites as well. Fasciola species sequence analysis had differences in the mitochondrial genes (COX1, CYTB, and NAD1), and related protein models. This study, thus, provides new insights into the differences in molecular sequences among the species of *Fasciola*.

Author Contributions: Conceptualization: JC and HIA, Data curation: HIA, Formal analysis: HIA and KM, MZA. Funding acquisition: HIA and JC, Methodology: HIA, MBBM, MI and MZA, Software: HIA, KM and MBBM, Supervision: JC and MI, Validation: MI, ZA, and MBBM, writing ± original draft: HIA, KM, MBBM, Writing ± review & editing: JC, MI, MBM, HMP and MZA.

Funding: This work was supported by the GDAS project of Science and Technology Development (2019-GDASYL-0103059 and 2018GDASCX-0107).

Acknowledgments: The authors are thankful to Springer Nature Author Services for providing English language editing support and critical review of the manuscript. This work was supported by the GDAS project of Science and Technology Development (2019-GDASYL-0103059 and 2018GDASCX-0107).

Conflicts of Interest: The authors declare no conflict of interest.

 doi:10.20944/preprints202011.0294.v1

Appendix A

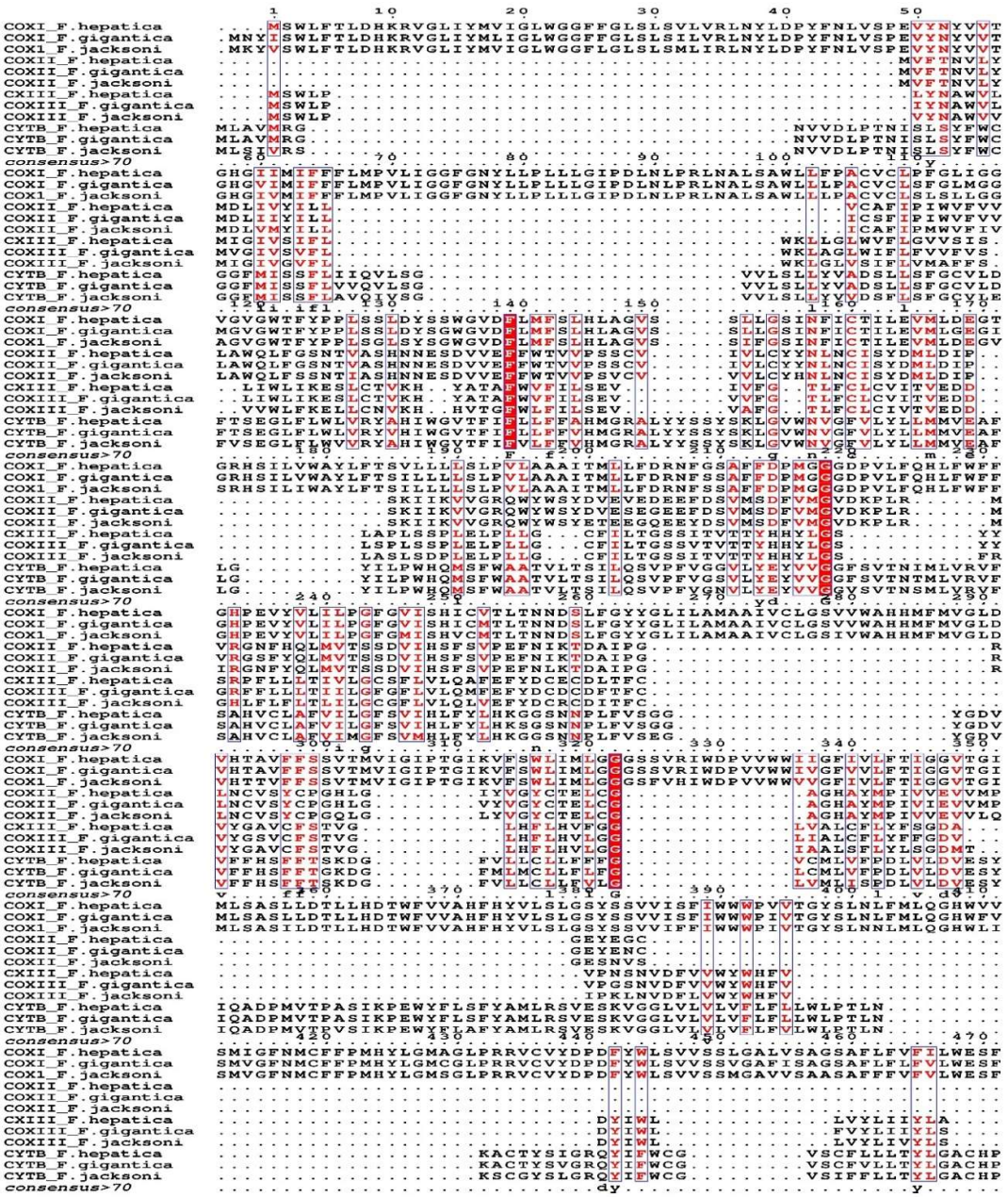

Figure 1. The result of MSA alignment of these proteins showed several identical, similar, and alternate residues however in COX and CYTB sequence alignment only similar and alternative residues were observed suggesting that had less conserved residues.

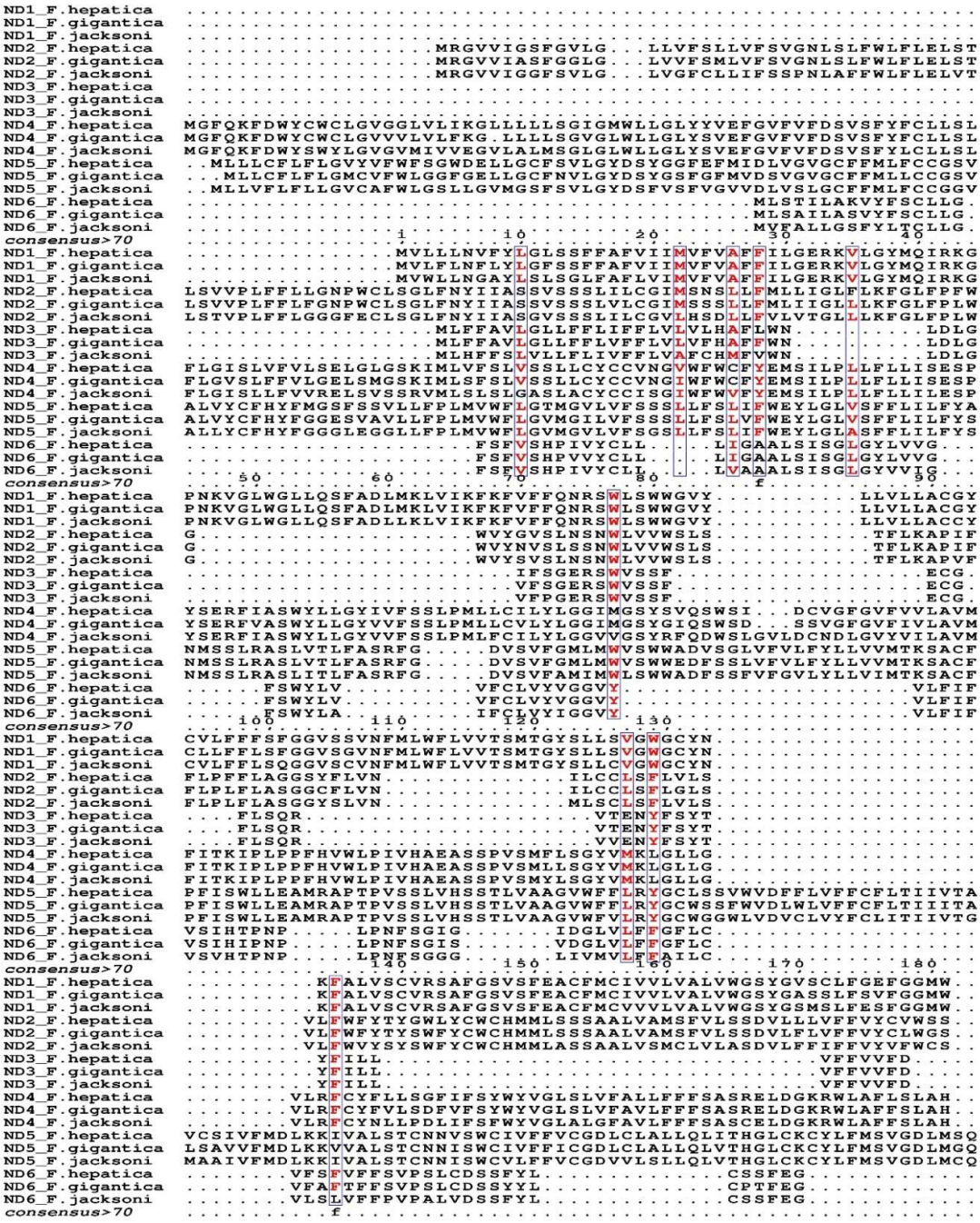

Figure A2. The result of MSA alignment of these proteins showed several identical, similar, and alternate residues however in NADH sequence alignment only similar and alternative residues were observed suggesting that had less conserved residues.

References

1. Pantelouris, E. M. *The common liver fluke: Fasciola bepatica l*. Elsevier, 2013,
2. Andrews, S. J. "The life cycle of fasciola hepatica." *Fasciolosis* 1 (1999): 20.
3. Mas-Coma, M., J. Esteban and M. Bargues. "Epidemiology of human fascioliasis: A review and proposed new classification." *Bulletin of the World Health Organization* 77 (1999): 340.
4. Shimalov, V. and V. Shimalov. "Helminth fauna of the wolf (canis lupus linnaeus, 1758) in belorussian polesie." *Parasitology research* 86 (2000): 163-64.
5. Copeman, D. and R. Copland. "Importance and potential impact of liver fluke in cattle and buffalo." *ACIAR Monograph Series* 133 (2008): 21.

6. Mehmood, K., H. Zhang, A. J. Sabir, R. Z. Abbas, M. Ijaz, A. Z. Durrani, M. H. Saleem, M. U. Rehman, M. K. Iqbal and Y. Wang. "A review on epidemiology, global prevalence and economical losses of fasciolosis in ruminants." *Microbial pathogenesis* 109 (2017): 253-62.
7. González-Miguel, J., M. Valero, M. Reguera-Gomez, C. Mas-Bargues, M. Bargues, F. Simón and S. Mas-Coma. "Numerous fasciola plasminogen-binding proteins may underlie blood-brain barrier leakage and explain neurological disorder complexity and heterogeneity in the acute and chronic phases of human fascioliasis." *Parasitology* 146 (2019): 284-98.
8. Ashrafi, K., J. Massoud, K. Holakouei, M. MAHMOUDI, M. JOUAFSHANI, M. Valero, M. Fuentes, M. Khoubbane, P. Artigas and M. Bargues. "Evidence suggesting that fasciola gigantica might be the most prevalent causal agent of fascioliasis in northern iran." (2004):
9. Malek, E. A. *Snail transmitted parasitic diseases: Volume ii*. CRC Press, 2018,
10. Lotfy, W. M., S. V. Brant, R. J. DeJong, T. H. Le, A. Demiaszkiewicz, R. J. Rajapakse, V. B. Perera, J. R. Laursen and E. S. Loker. "Evolutionary origins, diversification, and biogeography of liver flukes (digenea, fasciolidae)." *The American journal of tropical medicine and hygiene* 79 (2008): 248-55.
11. Ai, L., M.-X. Chen, S. Alasaad, H. M. Elsheikha, J. Li, H.-L. Li, R.-Q. Lin, F.-C. Zou, X.-Q. Zhu and J.-X. Chen. "Genetic characterization, species differentiation and detection of fasciola spp. By molecular approaches." *Parasites & vectors* 4 (2011): 101.
12. Periago, M., M. Valero, M. El Sayed, K. Ashrafi, A. El Wakeel, M. Mohamed, M. Desquesnes, F. Curtale and S. Mas-Coma. "First phenotypic description of fasciola hepatica/fasciola gigantica intermediate forms from the human endemic area of the nile delta, egypt." *Infection, Genetics and Evolution* 8 (2008): 51-58.
13. Ali, H., L. Ai, H. Song, S. Ali, R. Lin, B. Seyni, G. Issa and X. Zhu. "Genetic characterisation of fasciola samples from different host species and geographical localities revealed the existence of f. Hepatica and f. Gigantica in niger." *Parasitology research* 102 (2008): 1021-24.
14. Liu, G.-H., R. B. Gasser, N. D. Young, H.-Q. Song, L. Ai and X.-Q. Zhu. "Complete mitochondrial genomes of the 'intermediate form'of fasciola and fasciola gigantica, and their comparison with f. Hepatica." *Parasites & vectors* 7 (2014): 150.
15. Higuita, N. I. A., E. Brunetti and C. McCloskey. "Cystic echinococcosis." *Journal of clinical microbiology* 54 (2016): 518-23.
16. Sadjjadi, S. M., M. Ebrahimipour and F. S. Sadjjadi. "Comparison between echinococcus granulosus sensu stricto (g1) and e. Canadensis (g6) mitochondrial genes (cox1 and nad1) and their related protein models using experimental and bioinformatics analysis." *Computational biology and chemistry* 79 (2019): 103-09.
17. Deplazes, P., L. Rinaldi, C. A. Rojas, P. Torgerson, M. Harandi, T. Romig, D. Antolova, J. Schurer, S. Lahmar and G. Cringoli. "Global distribution of alveolar and cystic echinococcosis." In *Advances in parasitology*. 95. Elsevier, 2017, 315-493.
18. Rojas, C. A. A., A. R. Jex, R. B. Gasser and J.-P. Y. Scheerlinck. "Techniques for the diagnosis of fasciola infections in animals: Room for improvement." In *Advances in parasitology*. 85. Elsevier, 2014, 65-107.
19. Cwiklinski, K., S. Donnelly, O. Drysdale, H. Jewhurst, D. Smith, C. D. M. Verissimo, I. C. Pritsch, S. O'Neill, J. P. Dalton and M. W. Robinson. "The cathepsin-like cysteine peptidases of trematodes of the genus fasciola." In *Advances in parasitology*. 104. Elsevier, 2019, 113-64.
20. Zhang, X.-X., W. Cong, H. M. Elsheikha, G.-H. Liu, J.-G. Ma, W.-Y. Huang, Q. Zhao and X.-Q. Zhu. "De novo transcriptome sequencing and analysis of the juvenile and adult stages of fasciola gigantica." *Infection, Genetics and Evolution* 51 (2017): 33-40.
21. Haçarız, O. and G. P. Sayers. "Genererating a core cluster of fasciola hepatica virulence and immunomodulation-related genes using a comparative in silico approach." *Research in veterinary science* 117 (2018): 271-76.
22. Cwiklinski, K. and J. P. Dalton. "Advances in fasciola hepatica research using 'omics' technologies." *International journal for parasitology* 48 (2018): 321-31.
23. Morphew, R. M., T. J. Wilkinson, N. Mackintosh, V. Jahndel, S. Paterson, P. McVeigh, S. M. Abbas Abidi, K. Saifullah, M. Raman and G. Ravikumar. "Exploring and expanding the fatty-acid-binding protein superfamily in fasciola species." *Journal of proteome research* 15 (2016): 3308-21.
24. Martínez-Sernández, V., M. J. Perteguer, M. Mezo, M. González-Warleta, T. Garate, M. A. Valero and F. M. Ubeira. "Fasciola spp: Mapping of the mf6 epitope and antigenic analysis of the mf6p/hdm family of heme-binding proteins." *PLoS One* 12 (2017): e0188520.
25. Consortium, U. "Activities at the universal protein resource (uniprot)." *Nucleic acids research* 42 (2013): D191-D98.
26. Larsson, A. "Aliview: A fast and lightweight alignment viewer and editor for large datasets." *Bioinformatics* 30 (2014): 3276-78.
27. Sievers, F. and D. G. Higgins. "Clustal omega, accurate alignment of very large numbers of sequences." In *Multiple sequence alignment methods*. Springer, 2014, 105-16.
28. Buchan, D. W., F. Minneci, T. C. Nugent, K. Bryson and D. T. Jones. "Scalable web services for the psipred protein analysis workbench." *Nucleic acids research* 41 (2013): W349-W57.
29. Walsh, I., A. J. Martin, T. Di Domenico, A. Vullo, G. Pollastri and S. C. Tosatto. "Cspritz: Accurate prediction of protein disorder segments with annotation for homology, secondary structure and linear motifs." *Nucleic acids research* 39 (2011): W190-W96.
30. Singh, H., H. K. Srivastava and G. P. Raghava. "A web server for analysis, comparison and prediction of protein ligand binding sites." *Biology Direct* 11 (2016): 14.
31. Yang, J. and Y. Zhang. "I-tasser server: New development for protein structure and function predictions." *Nucleic acids research* 43 (2015): W174-W81.
32. Yang, J., A. Roy and Y. Zhang. "Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment." *Bioinformatics* 29 (2013): 2588-95.

33. Yang, J., A. Roy and Y. Zhang. "Biolip: A semi-manually curated database for biologically relevant ligand–protein interactions." *Nucleic acids research* 41 (2012): D1096-D103.
34. Kelley, L. A. and M. J. Sternberg. "Protein structure prediction on the web: A case study using the phyre server." *Nature protocols* 4 (2009): 363.
35. Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang. "The i-tasser suite: Protein structure and function prediction." *Nature methods* 12 (2015): 7-8.
36. Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin. "Ucsf chimera—a visualization system for exploratory research and analysis." *Journal of computational chemistry* 25 (2004): 1605-12.
37. Suhrer, S. J., M. Wiederstein and M. J. Sippl. "Qscop—scop quantified by structural relationships." *Bioinformatics* 23 (2007): 513-14.
38. Gasteiger, E., C. Hoogland, A. Gattiker, M. R. Wilkins, R. D. Appel and A. Bairoch. "Protein identification and analysis tools on the expasy server." In *The proteomics protocols handbook*. Springer, 2005, 571-607.
39. Lovell, S. C., I. W. Davis, W. B. Arendall III, P. I. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson. "Structure validation by cα geometry: ϕ, ψ and cβ deviation." *Proteins: Structure, Function, and Bioinformatics* 50 (2003): 437-50.
40. Pejaver, V., W. L. Hsu, F. Xin, A. K. Dunker, V. N. Uversky and P. Radivojac. "The structural and functional signatures of proteins that undergo multiple events of post-translational modification." *Protein Science* 23 (2014): 1077-93.
41. Xie, C., X. Mao, J. Huang, Y. Ding, J. Wu, S. Dong, L. Kong, G. Gao, C.-Y. Li and L. Wei. "Kobas 2.0: A web server for annotation and identification of enriched pathways and diseases." *Nucleic acids research* 39 (2011): W316-W22.
42. Schultz, J., R. R. Copley, T. Doerks, C. P. Ponting and P. Bork. "Smart: A web-based tool for the study of genetically mobile domains." *Nucleic acids research* 28 (2000): 231-34.
43. Ren, J., L. Wen, X. Gao, C. Jin, Y. Xue and X. Yao. "Dog 1.0: Illustrator of protein domain structures." *Cell research* 19 (2009): 271-73.
44. Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar. "Mega6: Molecular evolutionary genetics analysis version 6.0." *Molecular biology and evolution* 30 (2013): 2725-29.
45. Gouet, P. and E. Courcelle. "Endscript: A workflow to display sequence and structure information." *Bioinformatics* 18 (2002): 767-68.
46. Wenger, M. and H. Mathonet. "Gendoc: A flexible software documentation generator." Presented at Astronomical data analysis software and systems XI, 2002. 281, 462.
47. Nazmi, F., M. A. Moosavi, M. Rahmati and M. A. Hoessinpour-Feizi. "Modeling and structural analysis of human guanine nucleotide-binding protein-like 3, nucleostemin." *Bioinformation* 11 (2015): 353.
48. Bjellqvist, B., G. J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J. C. Sanchez, S. Frutiger and D. Hochstrasser. "The focusing positions of polypeptides in immobilized ph gradients can be predicted from their amino acid sequences." *Electrophoresis* 14 (1993): 1023-31.
49. Kyte, J. and R. F. Doolittle. "A simple method for displaying the hydropathic character of a protein." *Journal of molecular biology* 157 (1982): 105-32.
50. Bjellqvist, B., J. C. Sanchez, C. Pasquali, F. Ravier, N. Paquet, S. Frutiger, G. J. Hughes and D. Hochstrasser. "Micropreparative two-dimensional electrophoresis allowing the separation of samples containing milligram amounts of proteins." *Electrophoresis* 14 (1993): 1375-78.
51. Szklarczyk, D., A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguez, T. Doerks, M. Stark, J. Muller and P. Bork. "The string database in 2011: Functional interaction networks of proteins, globally integrated and scored." *Nucleic acids research* 39 (2010): D561-D68.
52. Kohl, M., S. Wiese and B. Warscheid. "Cytoscape: Software for visualization and analysis of biological networks." In *Data mining in proteomics*. Springer, 2011, 291-303.
53. Dalimi, A. and M. Jabarvand. "Fasciola hepatica in the human eye." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 99 (2005): 798-800.
54. Ashrafi, K., M. Valero, M. Panova, M. Periago, J. Massoud and S. Mas-Coma. "Phenotypic analysis of adults of fasciola hepatica, fasciola gigantica and intermediate forms from the endemic region of gilan, iran." *Parasitology International* 55 (2006): 249-60.
55. Itagaki, T., M. Kikawa, K. Sakaguchi, J. Shimo, K. Terasaki, T. Shibahara and K. Fukuda. "Genetic characterization of parthenogenic fasciola sp. In japan on the basis of the sequences of ribosomal and mitochondrial DNA." *PARASITOLOGY-CAMBRIDGE*- 131 (2005): 679.
56. Itagaki, T., K. Sakaguchi, K. Terasaki, O. Sasaki, S. Yoshihara and T. Van Dung. "Occurrence of spermic diploid and aspermic triploid forms of fasciola in vietnam and their molecular characterization based on nuclear and mitochondrial DNA." *Parasitology International* 58 (2009): 81-85.
57. Mohanta, U. K., M. Ichikawa-Seki, T. Shoriki, K. Katakura and T. Itagaki. "Characteristics and molecular phylogeny of fasciola flukes from bangladesh, determined based on spermatogenesis and nuclear and mitochondrial DNA analyses." *Parasitology research* 113 (2014): 2493-501.
58. Hayashi, K., M. Ichikawa-Seki, U. K. Mohanta, T. S. Singh, T. Shoriki, H. Sugiyama and T. Itagaki. "Molecular phylogenetic analysis of fasciola flukes from eastern india." *Parasitology International* 64 (2015): 334-38.
59. Ichikawa-Seki, M., M. Peng, K. Hayashi, T. Shoriki, U. K. Mohanta, T. Shibahara and T. Itagaki. "Nuclear and mitochondrial DNA analysis reveals that hybridization between fasciola hepatica and fasciola gigantica occurred in china." *Parasitology* 144 (2017): 206-13.
60. Rokni, M. B., H. Mirhendi, A. Mizani, M. Mohebali, M. Sharbatkhori, E. B. Kia, H. Abdoli and S. Izadi. "Identification and differentiation of fasciola hepatica and fasciola gigantica using a simple pcr-restriction enzyme method." *Experimental parasitology* 124 (2010): 209-13.

 doi:10.20944/preprints202011.0294.v1

61. Rojas, C. A. A., B. R. Ansell, R. S. Hall, R. B. Gasser, N. D. Young, A. R. Jex and J.-P. Y. Scheerlinck. "Transcriptional analysis identifies key genes involved in metabolism, fibrosis/tissue repair and the immune response against fasciola hepatica in sheep liver." *Parasites & vectors* 8 (2015): 124.
62. Umezurike, G. M. and A. O. Anya. "Carbohydrate energy metabolism in fasciola gigantica (trematoda)." *International journal for parasitology* 10 (1980): 175-80.
63. Barrett, J. "Amino acid metabolism in helminths." In *Advances in parasitology*. 30. Elsevier, 1991, 39-105.
64. Schultz, M. D. *Targeting the nad salvage pathway for the treatment of the parasitic disease schistosomiasis*. The University of Alabama at Birmingham, 2019,
65. Ohama, T., S. Osawa, K. Watanabe and T. H. Jukes. "Evolution of the mitochondrial genetic code iv. Aaa as an asparagine codon in some animal mitochondria." *Journal of molecular evolution* 30 (1990): 329-32.
66. Kulkarni-Kale, U., S. Bhosle and A. S. Kolaskar. "Cep: A conformational epitope prediction server." *Nucleic acids research* 33 (2005): W168-W71.
67. Morris, G. M., R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson. "Autodock4 and autodocktools4: Automated docking with selective receptor flexibility." *Journal of computational chemistry* 30 (2009): 2785-91.
68. Ngan, C.-H., D. R. Hall, B. Zerbe, L. E. Grove, D. Kozakov and S. Vajda. "Ftsite: High accuracy detection of ligand binding sites on unbound protein structures." *Bioinformatics* 28 (2012): 286-87.