International Journal of Current Research and Review
DOI: https://doi.org/10.31782/IJCRR.2022.141501

Research Article

IJCRR
Section: Healthcare

Copyright@IJCRR

A Subtractive Proteomics Approach to Identify Putative Drug Targets Against Parasitic Species

Sehgal Parakh

Associate Professor, Biotechnology Department, Sanjay Rungta Group of Institution, R2, RCST, Bhilai Chhattisgarh, India.

ABSTRACT

Introduction: Parasitic diseases affecting humans continue to be the leading cause of morbidity as well as mortality, particularly in the tropical and subtropical countries. With the rise of multi-drug resistant forms of many diseases, it has become increasingly important to develop new strategies to identify alternative drug targets. One such strategy for analysis of protein sequences of pathogen that can provide useful insight and markers for drug target identification is Subtractive Proteomics Approach (SPA). In this dataset between the host and pathogen proteome is subtracted and analysed which provides information pertaining to a set of proteins that are likely to be essential to the pathogen but absent in the host. Subtractive Proteomics Approach (SPA) is one of the most efficient methods which includes the use of various precise software databases.

Aim: SPA analysis was to predict putative drug targets in parasitic species Leishmania *major, Plasmodium falciparum, Trypanosoma brucei and Trypanosoma cruzi*.

Methods: Common proteins of these species were predicted using Transporter Database 2.0. Common protein was analysed by BlastP tool that predicted proteins which are non-homologous to *homo sapiens*. Database of Essential Genes (DEG) database was used to predict essential proteins for parasites and results were subjected to CELLO tool to identify the subcellular localization of important protein targets. Prosite database identified important protein families and patterns which were present in drug targets.

Results: Drug targets identified in parasitic species *Leishmania major, Plasmodium falciparum, Trypanosoma brucei and Trypanosoma cruzi* are mitochondrial Carrier family, Major Facilitator Superfamily, Sulfate Permease Family, P-type ATPase Superfamily and Major Intrinsic Protein Family respectively.

Conclusion: Screening of drug targets using SPA approach has identified major drug targets in parasitic species and which are not present in human and can be investigated further by computational drug designing and systems biology approach.

Key Words: Subtractive proteomics approach (SPA), BlastP, DEG, CELLO, Prosite, Superfamily.

INTRODUCTION

Parasites namely *Leishmania Major*, *Plasmodium Falciparum, Trypanosome Brucei and Trypanosoma Cruzi* proving to be a threat all over the world and have approximately resulted in 48.4 million cases annually.[1,2] The diseases outbreak is due to negligence and fewer precautions and actions taken against these parasitic species. There are very fewer effective drugs and vaccines available against them. [3,4]

Diseases like Leishmaniasis (Kala-azar) is caused by *Leishmania Major* is a single cell protozoan, which is transmitted by the bite of the female phlebotomine sand fly infected by the parasite.[5] *Plasmodium falciparum* is a protozoan parasite when carried by vector female *Anopheles* mosquito to host causes an infectious disease malaria. *P.falciparum* is the most severe strain of the malaria causing species correlated with almost every malarial death. [6,7] *Trypanosoma Brucei Rhodesiense* and *Trypanosoma Brucei Gambiense are* cause of nearly all human African trypanosomiasis (Sleeping sickness). These extracellular parasites are transmitted by the bite of Glossina tsetse fly.[8,9] *Trypanosoma Cruzi*, a protozoan parasite, is the causative agent of chagas disease. The insect vector is a triatomine bug that carries the parasite *Trypanosoma Cruzi* which causes the disease. [10]

The major challenges for control of parasitic diseases in the 21st century is both exciting and daunting. There is a huge market demand to find novel drugs using various computational approaches and testing them with *in-vitro* methods

Corresponding Author:
Dr. Parakh Sehgal, Sehgal Nikunj, Milan Chowk , New Khursipar Bhilai District Durg, Chattisgarh - 490012, India.
Mobile: 7506423010; Email: parakh.sehgal09@gmail.com

ISSN: 2231-2196 (Print) **ISSN:** 0975-5241 (Online)

Received: 18.05.2022 **Revised:** 12.06.2022 **Accepted:** 05.07.2022 **Published:** 05.08.2022

to find optimum sustainable strategies for controlling the spread of parasites which is one of the major concerns across the globe and one of the major reasons of human death. The new predicted potential drugs predicted from computational screening and they can be translated to successful, sustainable strategies for regulating and extirpate the diseases caused from different forms of parasites which ultimately is a global concern and poses major human health threats. Biological target identification and validation are among the most important steps in developing a new antiparasitic drugs. *In-silico* based proteomics field has great promised multi-step drug-development and in the identification of protein targets and biochemical pathways involved in disease processes. The Subtractive Proteomic Approach (SPA) helps in the identification of essential proteins of a pathogen which must not be homologous to the host genome, as homologous protein targets may create problems of self-binding. These non-homologous essential genes ensure the survival of the pathogen and therefore can for drug development. Drugs against such potential targets are assumed to hold fewer side effects and chances of drug resistance are also expected to be low in humans. [11,12]

SPA uses high throughput screening of complete proteome using Transport DB 2.0, gives access to the large volumes of data generated by the automated transporter annotation pipeline. It provides comparative study of membrane transporters and outer membrane channels in different organisms with fully sequenced genomes.[13] BlastP program is used to detect sequence similarity of query sequences against database. UniProt database provide comprehensive resource on protein sequences and its annotated data. [14,15] Database of Essential Genes (DEG) which contains important genes of pathogen which can provide key to identify essential proteins that are indispensable for the survival. CELLO works on classification system that predicts localization of subcellular proteins. PROSITE predicts annotated collection of motif descriptors that identifies protein families and domains in sequence. [16,17,18] The present analysis is based on proteome information to predict drug targets in parasitic species. The data emphasize the challenge of identifying essential proteins required for parasite survival and are absent from its host. [19]

MATERIALS AND METHODS

Transporter DB 2.0 (www.membranetransport.org)which utilizes an augmented pipeline with completely automated transporter annotation, and hence is able to better keep pace with the rate of data acquisition.[20] The proteome of all four species were compared and analysed for common protein sequences. Each common protein was selected and its sequence was retrieved from UniProt (www.uniprot.org) which is a secondary database a combination of SWISS-PROT, TrEMBL, and PIR resources therefore has larger coverage, low redundancy, cross-references and a high quality of annotation database.[14] Further BlastP (www.uniprot.org/blast/) analysis was carried out against these FASTA sequences to find protein sequences which are non-homologous to *homo sapiens*. Sequences which were non-homologous to *homosapiens* were accepted and the ones which were not was excluded (less threshold then 10).[21,22]

The database of essential genes (DEG) contains all the essential genes that are currently available for any species. Users can BLAST the query sequences against DEG database to find non-homologous gene to query sequences.[23] Therefore BlastP was carried out in the DEG with default parameters for analysing the essential proteins required for survival of the pathogen (www.essentialgene.org/). Proteins which showed bit score more than 100 were included and the remaining were excluded. These are considered as essential proteins comprise excellent targets for anti-parasitic drugs. In this way, essential protein targets were identified. Protein sub-cellular localization prediction basically includes the computational prediction of where a protein resides in a cell. CELLO (//cello.life.nctu.edu.tw/) identifies the subcellular localization of the protein sequences and predicted function and annotation. Further Prosite database (www.prosite.expasy.org) determined the important patterns and regions of the target protein sequences. [17,18]

RESULTS

In this analysis few chosen parasitic species namely *Leishmania major, Plasmodium falciparum, Trypanosome brucei* **and** *Trypanosoma cruzi proteome* was analysed using Subtractive approach to identify putative drug targets for which **data shown in Table 1A.** [24] The species proteome were compared using Transporter DB 2.0 and resulted in twenty-two common protein families shown in **Table 1B** which shows basic gist about the number of proteins and its size, Genome size, total number of transporter proteins and the number of transporters per Mb. The proteins which showed coloured columns they specified that the protein was common for all four species which were taken for next step analysis and the blank columns were neglected. Here the colors indicate the number of transporter proteins present in each family. Further UniProt database was used to retrieve the FASTA sequences of the common families resulted in table 1(B) and were run against BlastP tool to identify the proteins sequences that are non-homologous to *Homo sapiens* **Table 2.** BlastP tool predicted the common transporter proteins this resulted in identification of non-*homosapiens* homologous essential proteins. These proteins were manually checked after obtaining the result. If the list consisted of *homosapiens* as an organism that protein family was neglected for further analysis and vice-versa. *L.Major* resulted with 17 essential transporter proteins, *P.Falciparum* with 19 essential

transporter proteins, *T.Brucei and T Cruzi* resulted 16 each essential transporter proteins which are non-homologous to humans and are used in the further analysis.

After finding the essential proteins which are non- homologous to *Homo sapiens* through the process of BlastP, database of essential genes (DEG) was used for the analysis of potential drug targets. Resulted final sequences from previous analysis of BlastP were uploaded to DEG. This step achieved list of proteins which showed bit score more than 100 as they were considered essential for survival of parasite shown in **Table 3**. Sequences with significant alignments were all manually counted. For these sequences subcellular localization of were predicted using CELLO tool shown in **Table 4** that resulted in 17 plasma membrane proteins, 3 were mitochondrial proteins, 1 was inner membrane and 2 were nuclear membrane. The resulted non- homologous proteins domains were predicted using Prosite database which identified Mitochondrial Protein Translocase (MPT) , Mitochondrial Carrier (MC) , Major Facilitator Superfamily (MFS), Sulfate Permease (SulP) and P-type ATPase (P-ATPase) Superfamily and Major Intrinsic Protein (MIP) families as potential drug targets shown in **Table 5**.

DISCUSSION

Parasites and predators possess serious threats to humans throughout our evolutionary history. Although the impact of predators in modern world is lower than previously, parasites still infl -uence morbidity and mortality of contemporary humans. Parasites are very common in the lives of their hosts, affecting everything from their physiology to their behaviour and even their life histories. This has a significant impact on the structure of ecosystems, as parasites can control the populations of their hosts. Nevertheless, the parasite has found a unique way to overcome defenses and abuse the host. SPA provides a powerful approach for identification of disease-specific and novel drug targets against L. Major, P. Falciparum, T. Brucei and T. Cruzi respectively. [25] Twenty two essential common proteins among all 4 were predicted. The sequences were screened for similarity from Blastp tool and homologous proteins were avoided in further analysis as their homology with the host (Homo sapiens, in this case) proteins can lead to unwanted toxicity and cross-reactivity inhuman if they are selected as a drug target against parasitic species. Subcellular localization were shown by CELLO and predicted 17 plasma membrane , 3 mitochondrial, 1 inner membrane and 2 nuclear membrane proteins.

Prosite shown putative functional domains namely Mitochondrial Protein Translocase (MPT), Mitochondrial Carrier (MC) ,Major Facilitator Superfamily (MFS), Sulfate Permease (SulP) and P-type ATPase (P-ATPase) Superfamily and Major Intrinsic Protein (MIP)respectively . These domains are involved in important biological processes like cell division, cell cycling, and signal transduction, ribosome assembly and protein synthesis. These protein families are components of respiratory chain complex, can be found in the C-terminal cytoplasmic part of anion transporters , cation transport ATPases. These families can help in formation of specific vaccines and drugs precisely. (Sarangi *et* al 2000) reported similar analysis with SPA using similar approach and identified several proteins that can be targeted for effective vaccine development against N. meningitides serogroup B which causes Meningococcal disease, the number of essential genes in the metabolic pathways of N. meningitides serogroup B, identified in the present study, is 33 these can be further characterized and their role in the survival of the bacteria can be verified.[26] (Hosen et al. 2014) carried out studies on the Application of a Subtractive Genomics Approach for *in-silico* identification and characterization of novel drug targets in mycobacterium tuberculosis F1 in which similar approach and tools were used , 15 proteins that are potential target against tuberculosis were identified and these proteins need to be characterized for their drug target properties *in vitro* in future to achieve newer generation anti-tuberculosis drugs.[27]

CONCLUSION

Humans are exposed to a large variety of parasites and in turn they are exposed to infectious diseases that lead to affect their survival.[28] Sometimes the disease can be fatal. Parasites in the human body can cause significant global morbidity and mortality, as they can lead to recurrent liver failure and cancer. Therefore there is a demand to find precise drugs and vaccines. Essential proteins can open a key to get better advance drugs against these parasitic infections.

The present study proved to be a powerful approach for identification of several proteins that can be targeted for effective drug development specifically against parasites which can be less or nontoxic to the host. Homology modeling of these target proteins will addon to identify the best possible sites that can be targeted for drug design by simulation modeling. In future virtual screening against these novel targets might be useful in the discovery of novel therapeutic compounds.

ACKNOWLEDGMENT

Authors acknowledge the immense help received from the scholars whose articles are cited and included in references of this manuscript. The authors are also grateful to authors / editors / publishers of all those articles, journals and books from where the literature for this article has been reviewed and discussed."

Conflict of interest

Nil

Source of funding

Nil

REFERENCES

1. Torgerson PR. World Health Organization estimates of the global and regional disease burden of 11 foodborne parasitic diseases, 2010: a data synthesis. PLoS medicine. 2015 Dec 3;12(12):e1001920.
2. Ashton PD, Curwen RS, Wilson RA. Linking proteome and genome: how to identify parasite proteins. Trends Parasitol. 2001 Apr 1;17(4):198-202.
3. Trouiller P, Olliaro P, Torreele E, Orbinski J, Laing R, Ford N. Drug development for neglected diseases: a deficient market and a public-health policy failure. Global Health. 2017 May 15:267-73.
4. Nordqvist C, Biggers A. What's to know about parasites. Retrieved on 14th February. 2019.pp1-15. Parasites_-Types__in_humans__worms__and_ectoparasites20200402-125238-1ey-f5rc-with-cover-page-v2.pdf (d1wqtxts1xzle7.cloudfront.net)
5. Roberts LJ, Handman E, Foote SJ. Science, medicine, and the future: Leishmaniasis. Bmj. 2000 Sep 30;321(7264):801-4.
6. Dabadghao VS, Singh VB, Sharma D, Meena BL. A Study Of The Clinical Profile Of Malaria And Its Complications. IJCRR. 2016 Jan 1;8(1):1-25.
7. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature. 2005 Mar;434(7030):214-7.
8. Aksoy S. Human African trypanosomiasis research gets a boost: unraveling the tsetse genome. PLoS neglected tropical diseases. 2014 Apr 24;8(4):e2624.
9. Matthews KR. 25 years of African trypanosome research: From description to molecular dissection and new drug discovery. Mol. Biochem. Parasitol. 2015 Mar 1;200(1-2):30-40.
10. Miles MA. The molecular epidemiology and phylogeography of Trypanosoma cruzi and parallel research on Leishmania: looking back and to the future. Parasitol. Res. 2009 Oct;136(12):1509-28.
11. Katsila T, Spyroulias GA, Patrinos GP, Matsoukas MT. Computational approaches in target identification and drug discovery. Comput. Struct. Biotechnol. J. 2016 Jan 1;14:177-84.
12. Ahmad S, Raza S, Uddin R, Azam SS. Comparative subtractive proteomics based ranking for antibiotic targets against the dirtiest superbug: Acinetobacter baumannii. J. Mol. Graph. Model. 2018 Jun 1;82:74-92.
13. Ren Q, Chen K, Paulsen IT. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res. 2007 Jan 1;35(suppl_1):D274-9.
14. UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 2015 Jan 28;43(D1):D204-12.
15. Tatusova TA, Madden TL. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS microbiology letters. 1999 May 1;174(2):247-50.
16. Zhang R, Lin Y. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. Nucleic Acids Res. 2009 Jan 1;37(suppl_1):D455-8.
17. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins: Structure, Function, and Bioinformatics. 2006 Aug 15;64(3):643-51.
18. Sigrist CJ. PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res .2010 Jan 1;38(suppl_1):D161-6.
19. Mabey D, Peeling RW, Ustianowski A, Perkins MD. Diagnostics for the developing world. Nat. Rev. Microbiol. 2004 Mar;2(3):231-40.
20. Elbourne LD, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic acids Res. 2017 Jan 4;45(D1):D320-4.
21. Ogbe RJ, Ochalefu DO, Olaniru OB. Bioinformatics advances in genomics-A review. IJCRR. 2016 May 15;8(10):5.
22. Madden T. The BLAST sequence analysis tool. The NCBI handbook [internet]. NCBI (US). 2013 Oct.
23. Luo H. DEG 15, an update of the Database of Essential Genes that includes built-in analysis tools. Nucleic acids research. 2021 Jan 8;49(D1):D677-86.
24. Kappy MS, Barton LL, Berkowitz CD, Carver J, Ziegler MM. Advances in Pediatrics 2016, E-Book. Elsevier Health Sciences;vol (63) 2016 Jul 26.
25. Verma M, Wright Jr GL, Hanash SM, Gopal-Srivastava RA, Srivastava S. Proteomic approaches within the NCI early detection research network for the discovery and identification of cancer biomarkers. Ann. N. Y. Acad. Sci 2001 Sep;945(1):103-15.
26. Sarangi AN, Aggarwal R, Rahman Q, Trivedi N. Subtractive genomics approach for in silico identification and characterization of novel drug targets in Neisseria Meningitides Serogroup B. J. Comput. Syst. Biol. 2009;2(5):255-8.
27. Hosen M. Application of a subtractive genomics approach for in silico identification and characterization of novel drug targets in Mycobacterium tuberculosis F11. Interdisciplinary Sciences: CLS. 2014 Mar;6(1):48-56.
28. Altizer S. Social organization and parasite risk in mammals: integrating theory and empirical studies. Annu Rev Ecol Evol Syst. 2003 Jan 1:517-47.

Table 1A: Result Shown for Parasitic Species Using Transporter Database 2.0

Different organism Compared	Leishmania Major	Plasmodium Falciparum	Trypanosoma Brucei	Trypanosoma Cruzi
Genome Size (Mb)	3280	2300	2600	3400
Total Transporter Protein	246	100	305	373
Number of Transporters Per Mb Genome	0.08	0.04	0.12	0.11

Table 1B: Detailed Chart for Four Parasitic Species Using Transporter Database 2.0

Protein families	Leishmania Major	Plasmodium Falciparum	Trypanosoma Brucei	Trypanosoma Cruzi
ATP-binding Cassette Superfamily	41	14	28	47
H+- or Na+-translocating F-type, V-type and A-type ATPase Superfamily	2	2	2	2
H+- or Na+-translocating pyrophosphate family	1	2	5	2
Type II (General) Secretory Pathway Family	3	4	5	4
Mitochondrial Protein Translocase Family	2	10	1	3
P-type ATPase Superfamily	16	11	21	31
Major Intrinsic Protein Family	5	1	4	7
CorA Metal Ion Transporter Family	2	2	8	9
Voltage-gated Ion Channel Superfamily	7	1	3	9
Amino Acid/Auxin Permease Family	28	1	82	36
Cation Diffusion Facilitator Family	4	1	7	9
Monovalent Cation:Proton Antiporter-1 Family	2	1	1	4
Drug/Metabolite Transporter Superfamily	12	6	14	21
Equilibrative Nucleoside Transporter Family	5	2	21	8
Folate-Biopterin Transporter Family	14	2	10	5
Folate-Biopterin Transporter Family	38	10	35	50
Major Facilitator Superfamily	21	15	20	39
Cytochrome Oxidase Biogenesis Family	2	1	2	3
Inorganic Phosphate Transporter Family	4	1	2	6
Resistance-Nodulation-Cell Division Superfamily	1	1	2	4
Sulfate Permease Family	1	1	2	2
Zinc ,Iron , Permease Family	4	1	5	12

Colour Schema													
Transporter Count	0	1~5	6~10	11~20	21~30	31~40	41~50	51~60	61~80	81~100	100~120	120~150	>150

Key: To read (Table 1 b) the colors indicate the number of transporter proteins present in each family.

Table 2: Protein Similarity Result from Blast P Tool for Parasitic Species

Protein families	*L. Major*	*P. Falciparum*	*T. brucei*	*T. cruzi*
ATP-binding Cassette Superfamily	%	%	%	%
H+- or Na+-translocating F-type, V-type and A-type atpase Superfamily	%	-	%	%
H+-translocating Pyrophosphatase Family	%	%	%	%
Type ii secretory pathway family	-	%	-	-

Table 2: *(Continued)*

Protein families	*L. Major*	*P. Falciparum*	*T. brucei*	*T. cruzi*
Mitochondrial protein translocase family	%	%	%	%
P-type atpase Superfamily	%	%	%	%
Major intrinsic protein family	%	-	%	-
Cora Metal Ion Transporter Family	%	%	%	%
Voltage-gated Ion Channel Superfamily	-	%	-	-
Amino acid/auxin permease family	%	%	%	%
Cation diffusion facilitator family	-	%	%	%
Monovalent cation:proton antiporter-1 family	%	%	-	%
Drug/metabolite transporter superfamily	-	%	-	-
Equilibrative nucleoside transporter family	%	%	%	%
Folate-biopterin transporter family	%	%	%	%
Mitochondrial carrier family	%	%	%	%
Major facilitator superfamily	%	%	%	%
Cytochrome oxidase biogenesis family	%	%	-	-
Inorganic phosphate transporter family	%	%	%	%
Sulfate permease family	%	%	%	%
Zinc -iron permease family	%	%	%	%

Key for Table 2: % shows that protein are nonhomologous,-no result.

Table 3: Results From Database of Essential Genes(Deg) for Paasitic Species

Protein families	*L. Major*	*P.Falciparum*	*T. brucei*	*T. cruzi*
ATP-binding Cassette Superfamily	39	46	66	All less than 100
H+- or Na+-translocating F-type, V-type and A-type ATPase Superfamily	#	-	#	#
H+-translocating Pyrophosphatase Family	All less than 100	All less than 100	Less than 100	All less than 100
Type II Secretory Pathway Family	-	#	-	-
Mitochondrial Protein Translocase Family	7	15	8	8
P-type ATPase Superfamily	#	10	13	#
Major Intrinsic Protein Family	1	-	1	-
CorA Metal Ion Transporter Family	#	#	Less than 100	#
Voltage-gated Ion Channel Superfamily	-	#	-	-
Amino Acid/Auxin Permease Family	#	#	#	#
Cation Diffusion Facilitator Family	-	1	All less than 100	1
Monovalent Cation:Proton Antiporter-1 Family	#	#	-	#
Drug/Metabolite Transporter Superfamily	-	#	-	-
Equilibrative Nucleoside Transporter Family	#	#	#	#
Folate-Biopterin Transporter Family	#	#	#	#
Mitochondrial Carrier Family	#	4	1	#

Table 3: *(Continued)*

Protein families	*L. Major*	*P.Falciparum*	*T. brucei*	*T. cruzi*
Major Facilitator Superfamily	#	#	#	#
Cytochrome Oxidase Biogenesis Family	#	All less than 100	-	-
Inorganic Phosphate Transporter Family	1	1	1	1
Sulfate Permease Family	All less than 100	1	1	1
Zinc-Iron Permease Family	#	#	#	#

Key to table 3: # indicate No hit predicted from database.

Table 4: Results of Cello Analysis for Parasitic Species

Protein families	*L. Major*	*P .Falciparum*	*T. brucei*	*T. cruzi*
ATP-binding Cassette Superfamily	$ 3.187	$ 3.153	@ 3.559	Less than 100
H+-translocating Pyrophosphatase Family	Less than 100	Less than 100	Less than 100	Less than 100
Mitochondrial Protein Translocase Family	& 2.604	@ 3.460	& 1.379	& 1.664
P-type ATPase Superfamily	#	$ 4.101	$ 4.575	#
Major Intrinsic Protein Family.	^ 4.585	-	$ 3.862	-
CorA Metal Ion Transporter Family	#	#	LESS	#
Amino Acid/Auxin Permease Family	#	#	#	# 4.960
Cation Diffusion Facilitator Family	-	$ 2.385	LESS	$ 4.923
Monovalent Cation:Proton Antiporter-1 Family	#	#	-	#
Mitochondrial Carrier Family	#	$ 4.964	$ 4.544	#
Major Facilitator Superfamily	#	#	$ 4.924	#
Cytochrome Oxidase Biogenesis Family	#	Less than 100	-	-
Inorganic Phosphate Transporter Family	$ 4.668	$ 3.565	$ 4.947	$ 4.905
Sulfate Permease Family	Less than 100	$ 4.863	$ 4.930	$ 4.961

Key for table 4: # = NO HITS ($)= Plasma, (@) =Nuclear, (&) =Mitochondrial,(^) =Inner Membrane

Table 5: Prosite Database Result for Parasitic Species

Protein families	*L.Major*	*P.Falciparum*	*T. brucei*	*T. cruzi*
The Mitochondrial Protein Translocase Family	*1.PS00301 G_TR_1 α 2.PS51722 G_TR_2 α*	*1.PS51722 G_TR_2α*	*1.PS51257 PROKAR_ LIPOPROTEIN 2.PS51722 G_TR_2 α 3.PS00301 G_TR_1α*	*1.PS51722 G_TR_2 α 2.PS00301 G_TR_1 α*
The Mitochondrial Carrier Family	-	*1 PS51003CYTB_CTER Ø*	*1.PS51002 CYTB_NTER Ø 2.PS51003 CYTB_CTER Ø*	-
The Major Facilitator Superfamily	-	-	*1.PS50850 MFS* υ	-
The Sulfate Permease Family	-	-	*2.PS50801 STAS ḍ*	*1.PS50801 STAS ḍ*
The P-type ATPase Superfamily	-	*1.PS00154 ATPA SES_E1-E2 μ*	*1.PS00154 ATPA SES_E1-E2 μ*	-
Major Intrinsic ProteinFamily	*PS00221MIP +*	-	*PS00221MIP +*	-

Key for table 5: μ =E1-E2*ATPases phosphorylation site,* α = *Translational (tr)-type guanine nucleotide-binding (G) domain signature,*ß = *Prokaryotic membrane lipoprotein lipid attachment site profile, Ø= Cytochrome b/b6 C-terminal region profile,* ḍ = *STAS domain profile,* υ = *Major facilitator superfamily profile, += MIP family signature,* PS = Prosite ID