

Insilico Functional Annotation for Antigenic Proteins of *Trichomonas Foetus*

Geethanjali Karli¹³, Rathnagiri Polavarapu² and Kalarani Varada^{1*}

¹Department of Biotechnology, Sri Padmavati Mahila Visvavidyalayam, Tirupati, INDIA

²2620 Braithwood Road, Atlanta, GA 30345; USA

³Department of Biotechnology, Govt. Degree College (W), Gajwel, Telangana, INDIA

Corresponding mail : vkalarani.academic@gmail.com

ABSTRACT

Introduction: Bovine Trichomonosis is considered as one of the most neglected diseases of cattle in that are known to have sexually transmission. Though cattle remain asymptomatic initially, but eventually leads to frequent abortions, and finally to complete reproductive failure. Lack of point of care diagnostics remains the major hurdle for screening for Trichomoniasis. In our earlier work, we have identified several potential immunogenic proteins that can be targeted as diagnostic markers. In this current study, we have chosen Cysteine protease 8, Surface antigen BspA-like Protein, Chlamydia polymorphic membrane protein-like, Circumsporozoite protein precursor, hypothetical protein, Immuno-dominant variable surface antigen-like, Ser/Thr protein phosphatase as frequently expressed proteins of *Trichomonas foetus* for functional annotation.

Results and discussion: Gene ontology resource for functional annotations reveal that all the proteins are involved in host pathogen interaction and establishing virulence. Except for surface antigen BspA-like Protein, all the above proteins are relatively small to medium size proteins with 300 to 800 amino acids length and are suitable for in- vitro studies. Based on sub-cellular localization predictions, Clan SB, family S8, subtilisin_like serine peptidase, Hypothetical protein, Immuno-dominant variable surface antigen-like Ser/Thr protein phosphatase, Adhesin are mostly confined to plasma membrane and out side the membrane.

Conclusions: *Trichomonas foetus* was found to have several immuogenic proteins with large outer membrane protein domains. Such proteins are most likely to have major role in host pathogen interaction. The proteins may have multiple large epitopes and exhibit efficient immune reactions. Heterologous expression of above proteins would promise to develop point of care diagnostics like Lateral flow assays and ELISA.

Keywords: Sub cellular location, Heterologous expression, functional annotation, Protein stability, Host pathogen interaction

Introduction

The animal health is the key contributor of any country`s economy as well as sustainability. But in this fast global pandemic situations, animal health and well being is paid very less attention. Animal population is under severe stress with regard to quality feed, hygiene, drastic changes in the climate conditions and various disease out breaks. If unattended at least from now, future world will have to struggle for good quality animal based food products (Standard operating procedures For Bovine breeding 2014). In view of zoonotic potential of several animal pathogens, there is a great need to develop easily available, simple and easy to perform diagnostics (OIE Reference manual 2018). Sexually transmitted diseases are major

threat to Bovine community. As per the Minimum standard Operating procedures of Bovine breeding, in recent times, there is little attention being made in developing, point of care diagnosis and treatment of some diseases like, Bovine tuberculosis (Pucken 2017). Para tuberculosis (Rudrama 2019), Brucellosis (Mallikarjuna 2017, Manasa 2019).

Bovine Trichomonosis is among the highly neglected sexually transmitted disease in cattle is responsible for huge losses in the cattle industry. (OIE Reference manual 2018). *Trichomonas foetus* (*T.foetus*) is a flagellate parasitic organism is the causative organism. It exists as trophozoite in most of its life cycle, can also take a pseudocyst form (Warton2018). Though cattle remain asymptomatic initially, but eventually leads to frequent abortions, and finally to complete reproductive failure. Presence of large number of infertile cows, requirement of multiple semen injections are the major symptoms of infestation with *T.foetus* (Schwebke 2004). *Trichomonas vaginalis* (*T.Vaginalis*) is a closely related with *T.foetus*, known to have sexual transmission in human, causing Trichomonosis. Based on several studies on surface proteome of *T.Vaginalis* (de Miguel 2010), and analysis for virulence factors (Hirt R 2013) in a comparative omics based approach it was possible to search for targets in *T.foetus*. Recently available transcriptomics and proteomics data of *T.foetus* (Huang 2013, Morin-Adeline 2014, Stroud2017, de, Andrade 2017) have further enhanced our possibility of suitable diagnostic markers and drug targets.

In our earlier studies, using comparative Omics based approach, we have identified several consistently highly proteins as possible biomarkers of Bovine Trichomonosis (Karli G 2020). As the genome sequencing of the *T.foetus* was done very recently, there is no availability of functional annotations for the proteins of this organism. In this current study, In this current study, we have chosen Cysteine protease 8, Surface antigen BspA-like Protein, Chlamydia polymorphic membrane protein-like, Circumsporozoite protein precursor, hypothetical protein, Immuno-dominant variable surface antigen-like, Ser/Thr protein phosphatase as frequently expressed proteins of *Trichomonas foetus* for further analysis. We have used several bioinformatics tools for functional annotations of proteins, prediction of physico-chemical properties of proteins, identification of most probable sub cellular localization of the protein as well as prediction of extracellular membrane proteins.. Using computational tools now it is very much possible to successfully predict various properties of proteins that aid in maximum expression of the proteins in heterologous host systems like, E.coli, Yeast or mammalian cells.

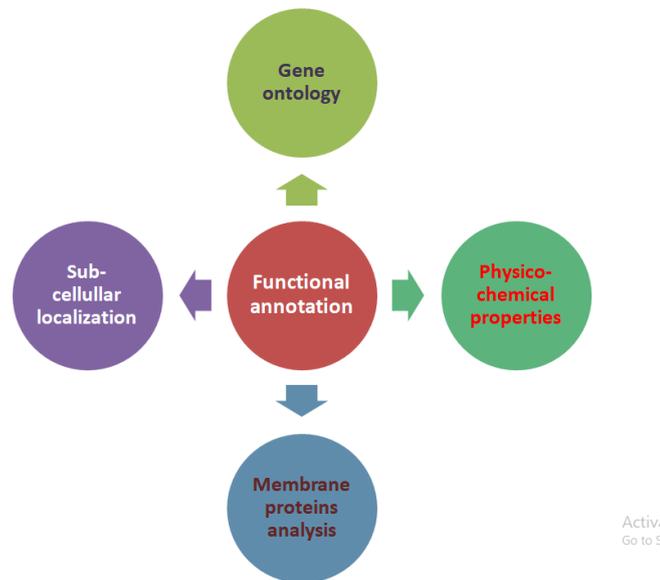


Fig.1. Over view of Functional annotation of newly identified protens. Gene ontology reveals the biological and molecular functions, sub-cellular localization for the cellular location of the protein, Physico- chemical properties for functional analysis and membrane protein prediction for possible role in pathogenicity.

Methods

Gene Ontology annotation

Protein sequences were obtained in FASTA format from NCBI protein (NCBI Resource Coordinators 2018) are used for analysis. Gene ontology annotation was used to retrieve the molecular functions and Biological processes associated with the targetable antigenic proteins of *Trichomonas foetus*. The data was obtained from Uniprot resource (https://www.uniprot.org/help/gene_ontology) (Ashburner, M 2000, UniProt Consortium 2015). FASTA sequences of the proteins are available in Supplementary file 1.0

Prediction of Physico-chemical properties of proteins

Expsy ProtParam tool was used to obtain various physico-chemical properties of the putative antigenic proteins of *Trichomonas foetus*. (<https://web.expasy.org/protparam>) (Gasteiger E 2005). These results were obtained based on the computational prediction and comparison of various chemical parameters like protein molecular weight, Isoelectric point, Composition, proportion of various amino acids as well as Physical parameters like the extinction coefficient, Half life of proteins in different cell systems, estimate of stability of proteins, Extent of aliphatic, hydrophobicity etc. that are available from Swissprot.

Sub cellular localization prediction of proteins

WoLF PSORT tool was used to predicts the intra cellular localization of proteins.(<https://wolfsort.hgc.jp/>) (Horton P 2007). This tool predicts based upon the sub cellular sorting signals as well as some amino acid sequence specific features mostly obtained through Uniprot. Depending on the predicted signals, cellular locations like, cytoplasmic, nuclear, mitochondrial, Endoplasmic reticular spaces, lysosomes as well plasma membrane proteins will be predicted.

Transmembrane protein analysis

Using Wolf PSORT majority of the proteins were predicted to be most likely present on the plasma membrane. These proteins were further analyzed by submitting the FASTA sequence to the TMMHM V.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) (Krogh A 2001) to identify the number of transmembrane domains and extent of extra cellular domain.

Results

GO Annotations

All the proteins are predicted to be highly antigenic proteins involved in both pathogen interactions. Cysteine Protease 8 is the extracellularly secreted protein involved in damaging the bovine genital tract. There are several surface antigenic proteins listed in the most expressed group may be involved in virulence. Adhesins, Surface Antigen Bspa-Like Protein, Clan SB, FamilyS8,Subtilisin_Like Serine Peptidase, Chlamydia Polymorphic Membrane Protein-Like, Immuno-Dominant Variable Surface Antigen-Like, Ser/Thr Protein Phosphatase may have a role in host pathogen interaction of the parasite to the genital epithelium. Functional annotations of the various highly expressed antigenic proteins using Gene Ontology resource (<http://geneontology.org/>) are given in the Table.1

Table 1: Functional Annotation of Putative Antigenic Proteins of *T.foetus*

Name of the protein	Functional Annotation
Cysteine Protease 8	Highly expressed extracellular protease, plays a vital role in host–parasite interactions such as virulence, adherence, nutrition acquisition and inflammation
Surface Antigen Bspa-Like Protein	Highly expressed Transmembrane protein
Clan SB, FamilyS8,Subtilisin_Like Serine Peptidase	Highly antigenic
Chlamydia Polymorphic Membrane Protein-Like	Serine –Type endo-peptidase activity
Hypothetical Protein	As auto transporter adhesins and are important in the initial phase of infection
Circumsporozoite Protein	Identified as hypothetical protein with transmembrane domain and antigenic domains in the <i>T. foetus</i> , potential diagnostic candidate gene.
	A similar Circumsporozoite protein is the immunodominant surface

Precursor	antigen on the sporozoite (the infective stage of the malaria parasite that is transmitted from the mosquito to the vertebrate host)
Immuno-Dominant Variable Surface Antigen-Like	Protein involved in effective strategy involved in Immune evasion
Ser/Thr Protein Phosphatase	A protein with hydrolase activity
Tetraspanin family protein	Novel candidate cell surface virulence factors , These are membrane proteins involved in signaling modulating adhesion, motility and tissue invasion in host systems
Adhesin AP65	Membrane associated protein involved in adhesion and pathogenesis.

Physico-chemical Properties:

ExPASy protparam analysis of FASTA sequences of these proteins are listed in Table 2.1 and 2.2 reveal that except for surface antigen BspA-like Protein, all the above proteins are relatively small to medium size proteins with 300 to 800 aa length.

Table.2.1 Chemical Properties of Potential Antigenic Proteins of *Trichomonas foetus*

NCBI ID	Name of the Protein	No of Amino acids	Molecular weight	Theoretical pI	-vely charged residues (Asp + Glu)	+vely charged residues (Arg + Lys)
OHT13704.1	Cysteine protease 8	320	35412.54	4.99	32	23
OHS99292.1	Surface antigen BspA-like Protein	3567	403762.44	4.75	448	276
OHT04136.1	Clan SB, family S8, subtilisin_like serine peptidase	866	95993.72	4.82	104	63
OHS93232.1	Chlamydia polymorphic membrane protein-like	264	29223.69	7.07	21	21
OHS95735.1	Hypothetical protein	870	96017.99	5.14	112	83
OHT08051.1	Circumsporozoite protein precursor	241	27316.93	6.24	16	11
OHT11175.1	Immuno-dominant variable surface antigen-like	270	30673.01	5.92	31	26
OHT00560.1	Ser/Thr protein phosphatase	506	58573.05	5.4	62	48
OHS99351.1	Tetraspanin family protein	208	23156.85	7.34	20	21

OHT02241.1	Adhesin AP65	376	40253.29	5.04	37	28
------------	--------------	-----	----------	------	----	----

SB-Seven stranded beta helix; Asp-Aspartic acid; Glu-Glutamic acid; Arg-Arginine;Lys-Lysine; Pi-Isoelectric pH

Table 2.2 Physical Properties Potential Antigenic Proteins of *Trichomonas foetus*

NCBI ID	Ext. coefficient	Abs Cys residues form cystines	Ext.coef ficient	All Cys reduced	Estimated half-life	Instabili ty index	Aliphati c index	(GRAVY)
OHT13704.1	81415	2.299	80790	2.281	30hours	25.67	68.34	-0.218
OHS99292.1	232190	0.575	223940	0.555	30,20,10	45.36	98.96	-0.022
OHT04136.1	75720	0.789	714720	0.778	30,20,10	40.66	85.1	-0.255
OHS93232.1	23880	0.817	23380	0.8	2.8,10,2	23.8	67.23	-0.118
OHS95735.1	87000	0.903	86750	0.903	30,20,10	32.69	95.14	-0.04
OHT08051.1	35410	1.296	35410	1.296	30,20,10	48.5	72.32	-0.509
OHT11175.1	62005	2.021	61880	2.017	30,20,10	47.76	83.04	-0.382
OHT00560.1	92640	1.582	92140	1.573	30,20,10	34.42	75.32	-0.473
OHS99351.1	27805	1.201	26930	1.163	30,20,10	31.58	126.11	0.754
OHT02241.1	44725	1.111	44350	1.102	30,28,10	27.34	86.57	-0.026

Cys-Cysteine; GRAVY- Grand average of hydropathicity

Theoretical Iso-electric point of all the a proteins is in the acidic range except for like Chlamydia polymorphic membrane protein-like and Tetraspanin family protein which fall in alkaline range. These proteins exhibit varied extinction coefficients. Except for Chlamydia polymorphic membrane protein-like, all the proteins are stable for 30 hrs in mammalian reticulocyte cells, 20 hrs in yeast cells and 10 hrs in E.coli. All the proteins are stable for in-vitro studies

Sub cellular localization:

Cellular distributions of various target proteins of *T.foetus* are consolidated with their respective probability distribution scores in WOLF PSORT tool are listed in Table.3.

Table 3. Prediction of Sub Cellular Localization of Antigenic Proteins of *Trichomonas foetus*

Name of the Protein	Output of Localization	Predominant localization	Membrane localization
Cysteine protease 8	Nucl: 13, Cyto_Nucl: 11, Cyto: 6.5, Cyto_Pero: 6.16667, Mito: 6, Extr: 4,	Nucleus	No
Surface antigen BspA-like	Nucl: 25, Plas: 3, Cyto: 2, E.R.: 1, Golg: 1	Nucleus	No
Clan SB, family S8,	Plas: 32	Plasma	Yes

subtilisin_like serine peptidase		membrane	
Chlamydia polymorphic membrane protein-like	Cyto: 10, Plas: 9, Cyto_Nucl: 8, Nucl: 4, Extr: 3, Mito: 3, Lyso: 2, Golg: 1	Cytoplasm & Plasma membrane	Yes
Hypothetical protein	Cyto: 15, Plas: 12, Nucl: 3, Cysk: 1, Golg: 1	Cytoplasm & Plasma membrane	Yes
Circumsporozoite protein precursor	Cyto_Nucl: 15, Cyto: 14.5, Nucl: 10.5, Extr: 4, Mito: 1, Pero: 1, Cysk: 1	Cytoplasm	No
Immuno-dominant variable surface antigen-like	Cyto_Nucl: 15.6667, Cyto: 15, Cyto_Plas: 9.6667, Nucl: 8, Mito: 3, Pero: 3, Extr: 2	Cytoplasm & Plasma membrane	Yes
Ser/Thr protein phosphatase	Plas: 23, Extr: 5, E.R.: 2, Pero: 1, Lyso: 1	Plasma membrane	Yes
Tetraspanin family protein	Extr: 31, Plas: 1	Extracellular	Yes
Adhesin AP65	Plas: 32	Plasma membrane	Yes

Nucl- Nucleus, Cyto- Cytoplasmic, Pero-Peroxisome, Mito-Mitochondria, Extr-Extracellular, Plas-Plasma membrane

Several proteins like Cysteine protease 8, Surface antigen BspA-like Protein, Chlamydia polymorphic membrane protein-like, Circumsporozoite protein precursor, hypothetical protein, Immuno-dominant variable surface antigen-like, Ser/Thr protein phosphatase have multiple sub cellular localizations, where as Clan SB, family S8, subtilisin_like serine peptidase, Tetraspanins and Adhesins are mostly confined to plasma membrane and out side the membrane.

Transmembrane protein analysis of T.foetus proteins

The Table 4 below describes the length of various domains of protein in the Exterior membrane region, within the membrane and in the cytoplasmic side.

Table 4. Membrane protein analysis using TMHMM analysis tool

Name of the protein	Total number of Amino acids	Outer membrane region (No of AA)	Trans membrane region (No of AA)	Cytoplasmic region (No of AA)
Clan SB, family S8, subtilisin-like serine peptidase	860	1-819	820-842	843-866
Chlamydia polymorphic membrane protein-like	726	1-640	641-663	664-726

Hypothetical protein	870	1-837	838-860	861-870
Immuno-dominant variable surface antigen	270	1-270	--	--
Ser/Thr protein phosphatase	506	1-64	65-87	88-506
Adhesin AP65	376	1-340	341-363	364-376

The graphical representations of the TMHMM V.2.0 analysis of one of the largest plasma membrane proteins, Clan SB, family S8, subtilisin-like serine peptidase are shown in Fig2. A single vertical bar represents that, it has 1 trans membrane domain predicted at the probability cut off value at 1.0 . For the above protein in the total 860 amino acids, this tools predicts it to have 819 amino acids in the extracellular domain only. Proteins with large extracellular domains tend to involve in host pathogen interaction. The figures of other trans membrane proteins are available in Supplementary file 2.

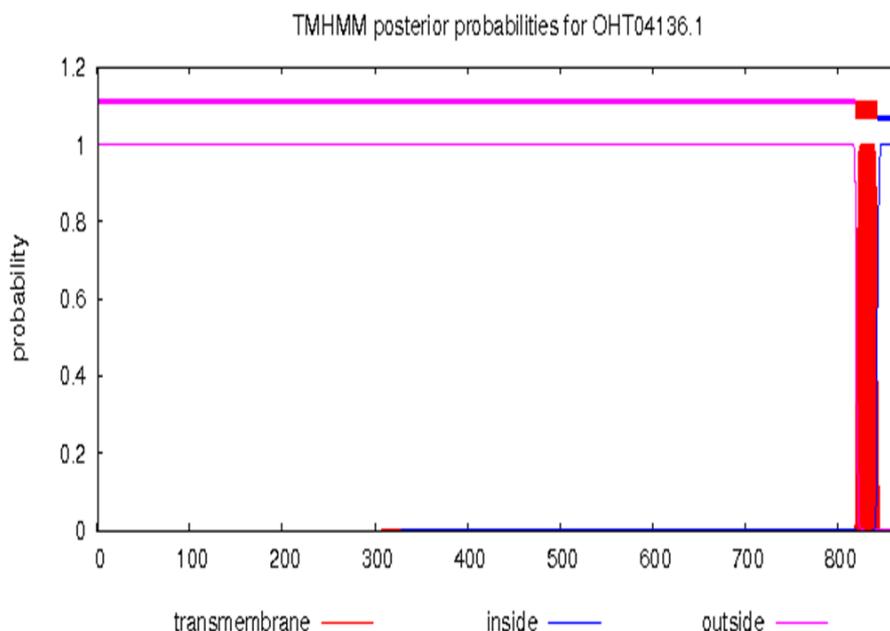


Fig.2. TMHMM analysis for Clan SB, family S8, subtilisin-like serine peptidase showing a largest extracellular domain and single trans membrane domain

Discussion

Omics based studies are helping the scientists towards a comprehensive understanding of the causative agent and possible mechanism/ mediators pathogenesis of any disease in recent times. Available data from a well studied closely related pathogen to *T.foetus* as well as few transcriptomics and proteomics have inspired us to screen the data to identify highly expressed proteins as well as consistently expressed proteins. Availability of freely accessible

Bioinformatics tools like Gene ontology resource have made us to understand the function of the target protein. The above analysis clearly indicates that, majority of amino acid sequences are falling in the exterior membrane region. It is very advantageous in the immunity/antigenicity point of view to have large outer membrane protein domains. The proteins may have multiple large epitopes and exhibit good host pathogen interaction.

In immune diagnostics point of view, it is always mandate to choose cell surface proteins as target proteins for antibodies as well as vaccine candidate. Insilico tool like, WOLF PSORT as well as TMHMM V.2.0 helped us to predict and shortlist few potential membrane proteins with moderately large amount of protein towards exterior of the cell. Extracellular portion of the proteins are more suitable to develop subunit vaccines, for developing point of care sero-diagnostics such as rapid kit LFA, ELISA as well as drug targets. Though above analysis helps in developing fast track tools, the results needs further validation with at least two more similar analytical tools for each parameter, before starting the lab experiments,

Conclusions

Trichomonas foetus is among the neglected tropical parasitic diseases. There are no diagnostics and therapeutics developed for point of care use. The above predictions would aid the scientists to perform various experiments like, cloning, expression and purification of proteins in order to develop diagnostics as well as therapeutics. Cysteine protease 8 is most abundantly expressed as well as extra cellular secretory protein of the parasite pose to induce immune response in the host. Hence using the above parameters, it can be expressed in suitable heterologous host like, E.coli, Yeast for developing serological assays. Circumsporozoite protein precursor was well exploited for sero-diagnosis in malarial infection (Zhao J 2016) which was also found to a vaccine target.

Limitations and future prospects

Insilico analysis warrants further invivo protein production and antigenicity testing in lab animals. Based on the above understanding, *T.foetus* Circumsporozoite protein need to be further exploited and characterized as therapeutic and vaccine candidate. Several proteins such as Immuno-dominant variable surface antigen-like, Tetraspanin family protein, Chlamydia polymorphic membrane protein-like, Clan SB, family S8,subtilisin_like serine peptidase, Surface antigen BspA-like Protein etc have plasma membrane localization. They can be well characterized as diagnostic targets, for vaccine development as well as drug targets.

List of Abbreviations

T.foetus (*Trichomonas foetus*); *T.vaginalis* (*Trichomonas vaginalis*); AA – Amino acids; BspA-Bacteroids Spirochetes surface antigen ; LFA –Lateral flow assay; ELISA-Enzyme linked Immunosorbent assay

Additional Information

Supplementary file 1.

Amino acid sequences of Target antigenic proteins of *Trichomonas foetus* in FASTA.

Supplementary file 2.

Graphic representations of TMHMM analysis for Chlamydia polymorphic membrane protein-like, Hypothetical protein, Immuno-dominant variable surface antigen-like, Ser/Thr protein phosphatase, Adhesin AP65.

Competing interests: Author declare No Competing interest

Funding: This Project did not receive any funding

Authors' contributions: GK has selected the list of immunogenic proteins and performed the functional annotation, and prediction of Physico-chemical properties and prepared the draft manuscript. RP has done the sub cellular localization analysis and critical analysis of results and KV has done transmembrane protein analysis and conclusions. authors have read and approved the manuscript

Acknowledgments:

We acknowledge the Team members of Genomix Molecular Diagnostics, Kukatpally, Hyderabad, INDIA and DST CURIE lab, Dept. of Biotechnology, SPMVV, Titupati, INDIA for their constant support for this current study.

References

- [1]. Compendium of Minimum standards of protocol & Standard operating procedures (22nd April 2014) For Bovine breeding Government of India Ministry of Agriculture Department of Animal Husbandry, Dairying & Fisheries. Accessed on 02.02.2019
- [2]. de, Andrade, Rosa, I., Caruso MB, de Oliveira Santos E, Gonzaga L, Zingali RB, de Vasconcelos ATR, de Souza W, Benchimol M (2017) The costa of trichomonads: A complex macromolecular cytoskeleton structure made of uncommon proteins *Biol Cell.* 109(6):238-253. doi: 10.1111/boc.201600050
- [3]. de Miguel, N., Lustig, G., Twu, O., Chattopadhyay, A., Wohlschlegel, J. A., & Johnson, P. J.(2010) Proteome analysis of the surface of *Trichomonas vaginalis* reveals novel proteins

- and strain-dependent differential expression. *Molecular & cellular proteomics*: MCP. 9(7), 1554–1566. doi:10.1074/mcp.M000022-MCP201
- [4]. Hirt R. P.(2013) Trichomonas vaginalis virulence factors: an integrative overview. *Sexually transmitted infections.*, 89(6), 439–443. doi:10.1136/sextrans-2013-051105
- [5]. Huang, K.Y., Shin, J.W., Huang, P.J., Ku,F.M., Lin, W.C., Lin, R., Hsu, W.M., Tang, P (2013) Functional profiling of the Tritrichomonas foetus transcriptome and proteome. *Mol Biochem Parasitol.*187(1):60-71. doi: 10.1016/j.molbiopara.2012.12.001
- [6]. Karli G., Polava R., Varada K.(2020) Comparative Omics Based Approach to Identify Putative Immunogenic Proteins of Trichomonas Foetus. *Learning and Analytics in Intelligent Systems*, 2020. 15. Springer, Cham. https://doi.org/10.1007/978-3-030-46939-9_51
- [7]. Mallikarjuna, R. C., Anumolu, V.K., Prudhvi, C. M., Revathi, P., Manasa, M., Rathnagiri, P., Vijayaraghavan, R. (2017) Sero Prevalence And Validation Of In-House IgM Elisa Kit For The Detection Of Brucella Antibody In Human Serum Samples *Asian Jr. of Microbiol. Biotech. Env. Sc.* 19 (4), 975-980
- [8]. Morin-Adeline, V., Lomas, R., O'Meally, D., Stack, C., Conesa, A., & Šlapeta, J.(2014) Comparative transcriptomics reveals striking similarities between the bovine and feline isolates of Tritrichomonas foetus: consequences for in silico drug-target identification. *BMC genomics.* 15(1), 955. doi:10.1186/1471-2164-15-955
- [9]. Manasa, M. , Revathi, P., Prudhvi, C. M., Maroudam, V., Navaneetha, P., Dhinakar, Raj,G., Kavi, K. PB., Rathnagiri, P. (2019) Protein-G-based lateral flow assay for rapid serodiagnosis of brucellosis in domesticated animals, *Journal of Immunoassay and Immunochemistry*, 40:2, 232-242,
- [10]. OIE Reference manual- Chapter 3.04.15- (1979).Trichomoniasis.Acessed on 10.02.2018 Warton, A., Honigberg, B.M.: Structure of trichomonads as revealed by scanning electron microscopy. *J. Protozoot.*, 26, 56–6
- [11]. OIE Reference manual (Acessed on 05.02.2018) Chapter 2.02.01.Development and optimization of Antigen detection assays. .
- [12]. Pucken V-B, Knubben-Schweizer G, Döpfer D, Groll A, Hafner-Marx A, Hörmansdorfer S, et al. (2017) Evaluating diagnostic tests for bovine tuberculosis in the southern part of Germany: A latent class analysis. *PLoS ONE* 12(6): e0179847. <https://doi.org/10.1371/journal.pone.0179847>
- [13]. Rudrama, D. P., Prudhvi, C. M., Revathi, P., Mukta, J., Soumendra, N. M., Jagdip, S. S., Sudhakar, P., Kavi, K. PB., and Rathnagiri, P. (2019) Development and Validation of Rapid, Sensitive and Inexpensive Protein G-based Point of Care Diagnostic Assay for Serodiagnosis of Paratuberculosis at Resource-Limited Areas. *Current Trends in Biotechnology and Pharmacy.* 13 (3), 232-242
- [14]. Schwebke, J. R., & Burgess, D.(2004) Trichomoniasis. *Clinical microbiology reviews* 17(4), 794–803. doi:10.1128/CMR.17.4.794-803.2004
- [15]. Stroud,L,J., Slapeta,J., PadulaM,P., Druery,D., Tsiotsioras,G., Coorsen,J,R., Stack,C,M., (2017) Comparative proteomic analysis of two pathogenic Tritrichomonas foetus genotypes: there is more to the proteome than meets the eye. *Int J Parasitol.*47(4):203-213. doi: 10.1016/j.ijpara.2016.11.004.

- [16]. NCBI Resource Coordinators (2018). Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, 46(D1), D8–D13. <https://doi.org/10.1093/nar/gkx1095>
- [17]. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1), 25–29. <https://doi.org/10.1038/75556>
- [18]. UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic acids research*, 43(Database issue), D204–D212. <https://doi.org/10.1093/nar/gku989>
- [19]. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A (2005). Protein Identification and Analysis Tools on the ExPASy Server;(In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press.pp. 571-607
- [20]. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. (2007)WoLF PSORT: protein localization predictor. *Nucleic Acids Res. Jul;35(Web Server issue):W585-7*. doi: 10.1093/nar/gkm259. Epub May 21. PMID: 17517783
- [21]. Krogh A, Larsson B, von Heijne G, Sonnhammer EL.(2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol. Jan 19;305(3):567-80*. doi: 10.1006/jmbi.2000.4315. PMID: 11152613.PMCID: PMC1933216.
- [22]. Zhao J, Bhanot P, Hu J, Wang Q A (2016) Comprehensive Analysis of Plasmodium Circumsporozoite Protein Binding to Hepatocytes. *PLoS ONE* 11(8): e0161607.