TRICHOMONAS VAGINALIS GENOME ANALYSIS USING BIOINFORMATICS APPROACHES

Singh Satendra[1*], Singh Gurmit[2], Gautam Budhayash[1], Varadwaj Pritish[3], Farmer Rohit[1]
[1]Department of Computational Biology & Bioinformatics, JSBBE, SHIATS, Allahabad-211007, India.
[2]Department of Computer science & Information Technology, SHIATS, Allahabad- 211007, India.
[3]Department of Bioinformatics, IIIT-Allahabad, India.
***Email:** satendralike@gmail.com

ABSTRACT

The draft genome sequence of the *Trichomonas vaginalis*, a sexually transmitted human pathogen has been published. But its genome presents many unusual genomic and biochemical features like, exceptionally large genome size, the presence of hydrogenosome, gene duplication, lateral gene transfer mechanism, presence of miRNAs, drug resistance etc. The purpose of this article is to convey current understanding of *Trichomonas vaginalis* genome as it has emerged over the past decades. It also identifies some the of significant research areas and proposes how Bioinformatics or computational methods can be used for the genome analysis of the *T. vaginalis*. It is expected that this paper will help in the designing of inhibitors for the putative drug targets identified based on the genome analysis.

Keywords: *T. vaginalis*, Pseudogenes, Lateral gene transfer, Trichomoniasis, Drug resistance.

INTRODUCTION

Trichomonas vaginalis was first described by Donne in 1836[1,2]. *T. vaginalis* is a unicellular, anaerobic, protozoan with pear or oval shape. Infection with *T. vaginalis* cause of trichomoniasis, number one nonviral and second most sexually transmitted disease (STD)[2-4]. *T. vaginalis* transmitted mostly by sexual contact[5], *T. vaginalis* is a cause of urethritis and prostate cancer in men[6] and both urethritis and vaginitis in women. *T. vaginalis* infection is associated with low birth weight and preterm delivery[7,8] it also predisposes carriers to HIV/AIDS[9,10]. Infection is treated and cured with metronidazole or tinidazole, and is prescribed to any sexual partner(s) as well because they may be asymptomatic carriers[11]. At present, metronidazole-resistant and cross-resistance is a challenging problem with no universally successful treatment[12-14].

The genome sequencing of *T. vaginalis* was carried out by The Institute of Genomic Research (TIGR) [15]. But genome sequencing is only the first step in order to study about any organism. There is an urgent need to carry out genome analysis and genome annotation of *T. vaginalis* so that we can understand the various biological mechanisms involved in genome expansion, pathogenesis, drug resistance, metabolic pathways etc. The present work concentrates on extensive reviews on *T. vaginalis* and proposes the areas, which can be investigated by using Bioinformatics and computational approaches.

In-Silico Genome Analysis

The genome sequence draft of *T. vaginalis* was published by The Institute of Genomic Research (TIGR) reveals an abnormally large genome size of 160 Mb[15]. Around two-thirds of the *T. vaginalis* sequence consists of repetitive and transposable elements, this reflects a massive, evolutionarily expansion of the genome. The total number of predicted protein-coding genes is ~98,000, which includes ~38,000 'repeat' genes. Approximately 26,000 of the protein-coding genes have been classified base on predicted functions but rest of the protein remains functionally uncharacterized. These extraordinary genome statistics are likely to change after carrying out more detained study of the *T. vaginalis* genome analysis. But it appears that the number of genes of the single-celled parasite *T. vaginalis* is, much more than of its host *H. sapiens.*[16] The genome also gives the platform to construct and analyze some important signal, secretary and metabolic pathway to identify and validate novel targets, which can be harvested to designed new drug molecules. Sequence similarity search methods provide some insights into putative functions for most gene products[17].

Gene family expressions and functional characterization

Pseudogenes are DNA sequences that were derived from a functional copy of a gene but which have acquired mutations that are deleterious to function. This duplicated copy of original functional gene gets incorporated into a new chromosomal location may leading to expansion of the existing gene family. Huge number of pseudogenes was thought to be present in T. *vaginalis* due to massive gene duplication. Human has about 30,000 genes with 38% of duplicated genes (pseudogenes), and of which around 12,000 pseudogenes have been identified[19]. In case of *T. vaginalis* TIGR predicted that there are about 60,000 genes in *T. vaginalis* but did not mention pseudogenes. It was speculated that a significant portion of the 60,000 genes might be pseudogenes[16].

Transmembrane cyclase is one of the important gene families; cyclases are critical in eukaryotic signal transduction and have a unique structure, which has been studied for the gene expansion or duplication. It was found that transmembrane cyclase gene family of *T. vaginalis* has about 3000 pseudogenes. The information about number of pseudogenes in other large gene families of *T. vaginalis* is not available. But it was proposed that large number of pseudogenes are present in the family of

ankyrin proteins, hypothetical protein, conserved hypothetical protein, adenylate cyclase, vsaA, surface antigen BspA, ANK-repeat protein, CG1651-PDrelated, Dentin sialophosphoprotein precursor, ABC transporter protein, kinases, major facilitator superfamily protein, leucine rich repeat family protein, and Transmembrane amino acid transporter protein. Many of those families are involved in secretary pathway and signal transduction system, which play important role in pathogenesis. It is expected that after a larger survey on above mentioned duplicated protein families and having more experimental data on the pseudogenes, we could shed light on why *T. vaginalis* possesses such a huge genome, how genes are duplicated, the quantity of its pseudogenes, and their evolution histories and also the functional significance of genes and pseudogenes which are expressing.

Genome evolves mechanism

Lateral transfer is the process by which genetic information is passed from one genome to an unrelated genome, where it is stably integrated and maintained. There is growing evidence from whole-genome analyses that this process is a very important mechanism in genome evolution, particularly among prokaryotes[19]. The evolution of a parasitic lifestyle requires adaptation to specialized characteristics for instance, pathogenicity islands are thought to be derived from common ancestor's genomes. Examples of common adaptive traits include host interaction systems, metabolic pathways that allow the acquisition of nutrients from the host, infection-related factors and mechanisms to evade host defenses. Lateral transfer could allow a previously harmless organism to rapidly colonize a new environment by acquiring highly specific biochemical functionality by gradual adaptation.

During genome annotation of *T. vaginalis* 152 cases of possible prokaryote-to-eukaryote lateral gene transfer (LGT) were identified. The putative functions of these genes are diverse, affecting various metabolic pathways and strongly influencing the evolution of the *T. vaginalis* metabolome. One significant example of lateral gene transfer is studied in sialic acid metabolism where *N*-acetylneuraminate lyases is transferred though lateral gene transfer. The recent transfer of this enzyme from one epithelial parasite to another suggests that it may well have the same role, but confirmation of this awaits functional characterization in *T. vaginalis.* The origin of other proteins involved in sialic acid metabolism has also not been investigated in trichomonads, but it would be worthwhile to determine if the entire pathway was acquired through lateral transfer or if the bacterial acetylneuraminate lyase was integrated into an existing pathway. In addition, it would be interesting to investigate whether the neuraminate lyase gene is also present in other trichomonads[15]. It will be interesting to predict other examples of lateral gene transfer and investigate whether these example are involve in pathogenesis or not. Such analysis will be more fruitful in case of *T. vaginalis* because this parasite has undergone transition to a urogenital environment from enteric environment.

Metabolic pathway analysis

Unlike most eukaryotes, *T. vaginalis* lacks mitochondria, some necessary enzymes, cytochromes, instead uses the hydrogenosome to accomplish fermentative carbohydrate metabolism[20]. The hydrogenosome appears to have a common ancestry with mitochondria based on similarities in protein import. Nutrients are taken up by transport through the cell membrane and by phagocytosis. The organism is able to maintain energy requirements by the use of a small amount of enzymes to provide energy via glycolysis of glucose to glycerol and succinate in the cytoplasm, followed by further conversion of pyruvate and malate to hydrogen and acetate in hydrogenosome with the generation of ATP[21,22]. Huge number of metabolic enzymes in *T. vaginalis* makes it relevant to the study of protein function.

T. vaginalis is a mucosal parasite of the urogenital vaginal tract thus for pathogenesis it need to adhere to the cervicovaginal epithelium and get colonization. Other important factors involved in cytoadherence are vaginal epithelial cells (VECs) pH, time and temperature. Proteins on the surface of live trichomonads have been implicated as mediators of host cytoadherence[23,24]. These proteins are AP65, AP51, AP33, and AP23[24] that mediate the interaction of the parasite to the receptor molecules on VECs[25]. Cysteine proteinase is another virulence factor as it helps in the proper interaction of pathogen and adhesins.

The polyamines putrescine, spermine, cysteine proteinases, and spermidine are ubiquitous small cations found in almost all living species and are essential for growth and function of normal cells. Polyamines have multiple functions that include stabilization of DNA, rRNA, and tRNA association with acidic phospholipids and regulation of membrane-associated enzymes. Polyamines have been shown to also have free radical scavenger properties, thus attributing an antioxidant function to these molecules. It has been suggested that polyamines prevent glycation of proteins[26] and are linked to host cell adherence and cytotoxicity. These repertoires of *T. vaginalis* proteins (involve in a variety of metabolic pathways like secretory pathway and signal transduction, cytoadherence, energy production) and their biologic properties represent important areas for further investigation[27]. Such research work targeting metabolism of *T. vaginalis* will increase our knowledge of trichomonal virulence and pathogenesis and help in the development of novel chemotherapeutics.

Drug resistance, multidrug resistance and cross-resistance

T. vaginalis was discovered in 1836 and has been known to cause vaginitis since 1916, it was not until 1957 that an effective cure, metronidazole, was discovered. But soon after drug resistance was first reported in 1962[28]. Cross-resistance between different nitroimidazoles has been reported and is consistent with earlier studies[29].

Metronidazol is a prodrug, which need to be activated by enzymes before drug act on the desired target. The metabolic pathways and enzymes involved in activation of drugs and subsequent resistance. The mechanism of development of anaerobic resistance to metronidazole also

is controlled by hydrogenosomes, in that metronidazole competes for H as an electron acceptor. In metronidazole-resistant *T. vaginalis*, the expression levels of the hydrogenosomal enzymes pyruvateferredoxin oxidoreductase, ferridoxin, malic enzyme, and hydrogenase are reduced dramatically, which probably eliminates the ability of the parasite to activate metronidazole. A strong correlation between drug resistance and altered regulation of ferredoxin gene transcription was established. A reduction in gene transcription results in decreased intracellular levels of ferredoxin and this may play a role in metronidazole resistance by decreasing the ability of the cell to activate the drug[30]. Genetic mutation is believed to be an important factor leading to increasing drug resistance. Understanding the mutation status will help to design accurate strategies of therapy against mutant strains of *T. vaginalis.* Bioinformatics analysis has been reported to determine the positions that tend to comply peptide motifs in the amino acid sequence of ferredoxin[31].

MicoRNA expressions analysis

MicroRNAs (miRNAs) are a class of small non-protein-coding RNAs that have important regulatory roles in multicellular organisms including humans, flies, nematodes, plants, and viruses. miRNA study has be carried out by using Bioinformatics analysis pipeline to identify putative *Trichomonas vaginalis* miRNA candidates from expression sequences tag (EST) database, putative open reading frame (ORF) database, and genomic scaffold database. Till now 20 candidate miRNAs, which has significant sequence and structure homology with known miRNAs in other species, are identified[32]. The presence of putative miRNAs in highly expressed ESTs indicted that *T. vaginalis* may have a different miRNA-regulating network compared with multicellular organisms. Such analysis will be used as a basis to design *T. vaginalis* specific miRNA chip. It is therefore interesting to investigate more miRNAs in *T. vaginalis* by using Bioinformatics analysis pipeline.

The presence of a *T. vaginalis* Dicer-like gene, two Argonaute genes suggests the existence of RNA interference (RNAi) pathway. Identification of these components raises the possibility of using RNAi technology to manipulate *T. vaginalis* gene expression[15].

Identification and validation of putative novel drug target

Drug resistance, multidrug resistance and cross-resistance have already been reported with the present drugs Metronidazole, nitroimidazoles, tinidazole that are used in the treatment of Trichomoniasis. Therefore there is need to identify other novel drug targets. *T. vaginalis* is an anaerobic protozoan parasite of humans that is rely heavily on cysteine as a major redox buffer, because it lacks glutathione. This has been reported that for synthesis of cysteine from sulfide, *T. vaginalis* relies upon cysteine synthase. Humans lack cysteine synthase; therefore, this parasite enzyme could be an exploitable drug targe[33]. The thioredoxin system is one of the importance defense mechanisms in trichomonads as if offers major antioxidant activity in response to environmental changes. Increase in the levels of thioredoxin and thioredoxin peroxidase has been reported. Sequence data indicate that the thioredoxin reductase of trichomonad s differs fundamentally in structure from that of its human host and thus may represent a useful drug target[34].

Bioinformatics Approaches

The various aspects of *T. vaginalis* genome that are discussed in this article can be analysis with the help of Bioinformatics and computational methods. Numerous databases are now available which contain both sequence and functionality information. Most of these are accessible over the Internet through convenient Web browser interfaces. Many also permit downloading of sequence information for use on local servers. Sequence databases now contain the nucleotide and predicted amino acid sequences of virtually every gene in the model microbes.

The genome sequence data of *T. vaginalis* can be downloaded from JCVI. Major Biological database, which can be used, are EcoCyc, KEGG: Kyoto Encyclopedia of Genes and Genomes, GeneCensus Genome Comparisons, NCBI, DDBJ, EMBL, SwissProt, PDB. Also there are various online as well as offline Softwares like COGs, NCGR, BLAST, MAGPIE, GenobaseMicro, ANMR, Pallen, CDC, Tripos, Orthologous gene alignments at JCVI, SEQUEST for identification of proteins, Motif, Pedant, GenTHREADER, Mummer, MAGPIE: Multipurpose Automated Genome, ClustalW, Microbial genome databases, Expassy tools, MUSCLE, PIMA, PAUP, PHYLIP etc.

TrichDB.org is a free, public genomic data repository and retrieval service devoted to genome-scale trichomonad data which was started in 2007. The site currently contains all of the *T. vaginalis* sequence project data, several EST libraries, and tools for data mining[9]. YEBIS (Yet another Environment for the analysis of BIopolymer Sequences), GeneHacker : Gene Structure Prediction in Microbial Genomes, Motif Extraction from DNA Sequence Data, Motif Search within DNA Sequence Data, Homology Search Service: BLAST, PSI-BLAST, megablast, SSEARCH, BLAST-SNP, BLAT search against human genome, Genome Shovel : Web-based Dot Matrix DNA Sequence Comparison.

SUMMARY

The availability of genomic sequence information of *T. vaginalis* along with modern Bioinformatics and Computational methods provides the opportunity to analyze its genome and carry genome annotation. Such analysis will enable us to go insights into the mechanism which leads to genome expansion on such a large scale, role of different metabolic pathways, pathogenic mechanisms, and once we have sufficient data and knowledge in these research areas we can harvest the same for designing new methods of treatment which includes, identification of putative drug target for drug designing and development of novel methods for diagnosis of *T. vaginalis.*

REFERENCES

1. Donne' A., (1836) Comptes rendus hebdomadaires des se'ances de l'Academie des Sciences Paris; 3:385
2. Petrin, D., K. Delgaty, R. Bhatt, and G. Garber**.** (1998). Clinical and microbiological aspects of *Trichomonas vaginalis*. Clin. Microbiol. Rev. 11**:**300– 317.
3. Rein, M. F., and M. Mu¨ller. (1990). *Trichomonas vaginalis* and Sexually transmitted diseases 481–492. McGraw-Hill, New York, N.Y
4. World Health Organization. (1995). An overview of selected curable sexually transmitted diseases, p. 2–27. *In* Global program on AIDS, World Health Organization, Geneva, Switzerland.
5. Honigberg B. M., In B. M. Honigberg(ed), (1990) Trichomonads parasitic in humans. Springer-Verlag, New York, N. Y.
6. JR Stark, G Judson, JF Alderete, V Mundodi, AS Kucknoor, EL Giovannucci, EA Platz, S Sutcliffe, K Fall, T Kurth, J Ma, MJ Stampfer, LA Mucci (2009). "Prospective Study of Trichomonas vaginalis Infection and Prostate Cancer Incidence and Mortality: Physicians' Health Study". *Journal of the National Cancer Institute*
7. Cotch, M. F., J. G. Pastorek II, R. P. Nugent, S. L. Hillier, R. S. Gibbs, D. H. Martin, D. A. Eschenbach, R. Edelman, J. C. Carey, J. A. Regan, M. A. Krohn, M. A. Klebanoff, A. V. Rao, G. G. Rhoads, and the Vaginal Infections and Prematurity Study Group. (1997). *Trichomonas vaginalis* associated with low birth weight and preterm delivery. Sex. Transm. Dis. 24**:**353–360.
8. Saurina, G. R., and W. M. McCormack. (1997). Trichomoniasis in pregnancy. Sex. Transm. Dis. 24**:**361–362.
9. Alderete, J. F., M. W. Lehker, and R. Arroyo. (1995). The mechanisms and molecules involved in cytoadherence and pathogenesis of *Trichomonas vaginalis*. Parasitol. Today 11**:**70–74.
10. Meysick, K., and G. E. Garber. (1995). *Trichomonas vaginalis*. Curr. Opin. Infect. Dis. 8**:**22–25
11. Cudmore SL, Delgaty KL, Hayward-McClelland SF, Petrin DP, Garber GE (2004) Treatment of infections caused by metronidazole-resistant Trichomonas vaginalis. *Clinical microbiology reviews* 17 (4): 783–93.
12. Lossick, J. G. (1990). Therapy of urogenital trichomoniasis, p. 324–341. *In* B. M. Honigberg (ed.), Trichomonads parasitic in humans. Springer-Verlag, New York, N.Y.
13. Meingassner, J. G., and H. Mieth. (1976). Cross-resistance of trichomonads to 5-nitroimidazole-derivatives. Experientia 32**:**183–184.
14. Narcisi, E. M., and W. E. Secor**.** (1996). In vitro effect of tinidazole and furazolidone on metronidazole-resistant *Trichomonas vaginalis*. Antimicrob. Agents Chemother. 40**:**1121–1125.
15. Carlton JM, Hirt RP, Silva JC, *et al.* (2007) Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. *Science (New York, N.Y.)* 315 (5809): 207–12
16. Jike Cui Temple F. Smith, (2004) Gene expansion in *Trichomonas vaginalis*: a case study on transmembrane cyclases, Bioinformatics Program, Boston University, 44 Cummington St., Boston, MA, 02215
17. Donald T. Moir, Karen J. Shaw, Roberta S. Hare, and Gerald F. Vovis , (1999) Genomics and Antimicrobial Drug Discovery, Pathogen Genetics Department, Genome Therapeutics Corporation, Waltham,Massachusetts 02453-8443,1.
18. Upcroft P, Upcroft JA (2001). "Drug targets and mechanisms of resistance in the anaerobic protozoa". *Clinical microbiology reviews* 14 (1): 150–64.
19. Lawrence, J. G**.** (1999). Gene transfer, speciation, and the evolution of bacterial genomes. Curr. Opin. Microbiol. 2:519–523.
20. Lindmark, D., and M. Muller. (1973). Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. J. Biol. Chem. 248**:**7724–7728.
21. Mu¨ller, M**.** (1993). The hydrogenosome. J. Gen. Microbiol. 139**:**2879–2889.
22. Mu¨ller, M**.** (1998). Enzymes and compartmentation of core energy metabolism of anaerobic protists—a special case in eukaryotic evolution?, p. 108– 132. *In* G. H. Coombs, K. Vickerman, M. A. Sleigh, and A. Warren (ed.), Evolutionary relationships among protozoa. Chapman and Hall, London, U.K.
23. Adegbaju, A. and Morenikeji, O. A., (2008) Cytoadherence and pathogenesis of *Trichomonas vaginalis.*, Parasitology Unit, Department of Zoology, University of Ibadan, Nigeria
24. Alderete, J. F., and G. E. Garza. **(**1985) Identification and properties of *Trichomonas vaginalis* proteins involved in cytadherence. Infect. Immun. 56: 28–33.
25. Aurrecoechea C, Brestelli J, Brunk BP, *et al.* (2008) GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis. Nucleic Acids Research*
26. AndrewJ. Roger, C.graham Clark, and W. Ford Doolittle , (1996) A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*, Program in Evolutionary Biology, Canadian Institute for Advanced Research.
27. Li, W.H., Gu, Z., Wang, H., and Nekrutenko, A., (2001) Evolutionary analyses of the human genome, *Nature*, 409(6822):847-849

28. Land, K., M. G. Delgadillo-Correa, J. Tachezy, S. Vanacova, C. L. Hsieh, R. Sutak, and P. J. Johnson. (2004). Targeted gene replacement of a ferredoxin gene in *Trichomonas vaginalis* does not lead to metronidazole resistance. Mol. Microbiol. 51**:**115–122.

29. Quon, D. V., C. E. d'Oliveira, and P. J. Johnson. (1992). Reduced transcription of the ferredoxin gene in metronidazole-resistant *Trichomonas vaginalis*. Proc. Natl. Acad. Sci. USA 89**:**4402–4406.

30. Schauer, R., S. Kelm, G. Reuter, P. Roggentin, and L. Shaw. (1995). Biochemistry and role of sialic acids. Pp. 7– 67 *in* A. Rosenberg, ed. Biology of the sialic acids. Plenum Press, New York

31. Sogin, M. L. Cur.Opinion Genet. Dev. (1991); 1: 457-63

32. Zhang, Z., Harrison, P.M., Liu, Y., and Gerstein, M., (2003) Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome, *Genome Res.*, 13(12):2541-2558

33. Wei-Chen Lin, Sung-Chou Li, Wen-Chang Lin, and Petrus Tang., (2006), Bioinformatics Analysis of *Trichomonas vaginalis* microRNAome., 1Molecular Regulation & Bioinformatics Laboratory, Chang Gung University.

34. Hook, E., (1999)Trichomonas vaginalis--no longer a minor STD, *Sex. Transm. Dis.*, 26(7):388-389.

35. Cu-Uvin, S., H. Ko, J. W. Jamieson, Hogan, P. Schuman, J. Anderson, R. S. Klein, and the HIV Epidemiology Research Study (HERS) Group**.** (2002) Prevalence, incidence, and persistence or recurrence of trichomoniasis among human immunodeficiency virus (HIV)-positive women and among HIV-negative women at high risk for HIV infection. Clin. Infect. Dis. 34**:**1406–1411.

****