Genetic diversity of *Trichomonas vaginalis* clinical isolates from Henan province in central China

Meng Mao, Hui Li Liu

Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Zhengzhou University, Henan Province, China

Trichomonas vaginalis is a flagellated protozoan parasite that infects the human urogenital tract, causing the most common non-viral, sexually transmitted disease worldwide. In this study, genetic variants of *T. vaginalis* were identified in Henan Province, China. Fragments of the small subunit of nuclear ribosomal RNA (18S rRNA) were amplified from 32 *T. vaginalis* isolates obtained from seven regions of Henan Province. Overall, 18 haplotypes were determined from the 18S rRNA sequences. Each sampled population and the total population displayed high haplotype diversity (Hd), accompanied by very low nucleotide diversity (Pi). In these molecular genetic variants, 91.58% genetic variation was derived from intra-regions. Phylogenetic analysis revealed no correlation between phylogeny and geographic distribution. Demographic analysis supported population expansion of *T. vaginalis* isolates from central China. Our findings showing moderate-to-high genetic variations in the 32 isolates of *T. vaginalis* provide useful knowledge for monitoring changes in parasite populations for the development of future control strategies.

Keywords: *Trichomonas vaginalis*, 18S rRNA, Genetic variation, China

Introduction

Trichomoniasis is the most prevalent non-viral sexually transmitted disease caused by the flagellated protozoan parasite, *Trichomonas vaginalis*,¹ that affects both sexes. For women primarily affected in the vulva, vagina and uterine cervix, and secondarily in the urinary tract, trichomoniasis may have major health consequences.² Additionally, *T. vaginalis* infection in pregnant women is associated with increased incidence of preterm delivery, premature rupture of membranes and low birth weight infants, and can be transmitted to neonates during passage through the birth canal.^{3,4} In men, trichomoniasis is a cause of chronic prostatitis and non-gonococcal urethritis.⁵ Trichomoniasis has also been linked to an increased risk of human immunodeficiency virus (HIV) infection and cervical cancer.^{6,7} Knowledge of the genetic characteristics of *T. vaginalis* populations is valuable for the prevention and control of trichomoniasis in humans.⁸ To date, however, limited studies have focussed on population genetics of *T. vaginalis* in China.

The lengths of specific regions in the small subunit of nuclear ribosomal RNA (SSU nrRNA, also known as 18S rRNA) are not conserved among different groups, and these differences can be significant.⁹ Thus, 18S rRNA is suitable to study genetic variations and genotypes of organisms.^{10,11} The main aim of the current study was to assess the genetic variations of *T. vaginalis* isolates collected from Henan province in central China, based on 18S rRNA data.

Methodology

Sample collection

T. vaginalis isolates were collected from infected women from seven geographical locations (Anyang, Zhengzhou, Shangqiu, Luoyang, Pingdingshan, Zhumadian and Xinyang) of Henan province in central China from May 2012 to August 2014 (See supplementary material online for this article at <http://www.maneyonline.com/doi/suppl/10.1179/2047773215Y.0000000020>; Supplementary Table S1). Parasites were obtained from urethral discharge, posterior vaginal fornix or urine sediment. *T. vaginalis* samples were cultured in modified trypticase-yeast maltose (TYM) medium supplemented with 10% heat-inactivated adult bovine serum and antibiotics (50 µg/ml gentamycin, 60 µg/ml

Correspondence to: Hui Li Liu, Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou 450014, Henan Province, China. Email: maomengaini@126.com

ciprofloxacin, 25 µg/ml clindamycin and 50 µg/ml fluconazole), and harvested at the mid-logarithmic phase (10^6 cells/ml), as described previously.¹² Axenic parasites at mid-log phase were archived as frozen isolates in 5% dimethyl sulphoxide (DMSO) at -70°C .

DNA extraction, amplification and sequencing

Total genomic DNA was extracted using the Tiangen DNeasy Blood and Tissue Kit (Tiangen, Beijing, China) following the manufacturer's protocol. The first half of the 18S rRNA gene was amplified with primers T18SF (5'-GGAAGCACACTTCGGTC-ATAG-3') and T18SRi (5'-CCTTCCTCAATTCC-TTCAA-3') and the second half using primers T18SFi (AGGGTTTCTGTGTCGATCAAGG) and T18SR (CGTTACCTTGTTACGAC TTCTCC).¹³ Polymerase chain reaction was performed in a total volume of 25 µl containing 2 mM MgCl_2 , 2.5 µM primer, 2.5 µl $10\times$ rTaq buffer, 0.5 mM each deoxyribonucleoside triphosphate (dNTP), 1.25 U of rTaq DNA polymerase (Takara, Dalian, China), and 0.25 ng DNA sample in a thermocycler under the following conditions: initial denaturation at 94°C for 5 minutes, followed by 94°C for 1 minute (denaturation), 56°C for 1 minute (annealing), 72°C for 2 minutes (extension) for 30 cycles and final extension at 72°C for 7 minutes. PCR products were purified using the High Pure PCR Product Purification Kit (Takara), and sequenced in both directions with an automated sequencer (ABI Prism 3730 XL DNA Analyzer; ABI Prism, Foster City, CA, USA) by the GENWIZ Company (Beijing, China).

Analysis of nucleotide polymorphisms

Sequences for 18S rRNA were initially aligned using default settings in Clustal X v2.0¹⁴ and adjusted in MEGA v5.0.¹⁵ Nucleotide composition, variable sites, parsimony-informative sites, singleton sites, number of haplotypes, haplotype diversity (Hd) and nucleotide diversity (Pi) were determined using the program DnaSP v5.10.01.¹⁶

Phylogenetic analysis

The phylogenetic relationships among haplotypes were estimated using maximum parsimony (MP) and Bayesian inference (BI). MP analysis was performed in PAUP*4b10 using heuristic searches with tree bisection-reconnection (TBR) branch swapping and 2000 random addition sequences. Confidence in each node was assessed by bootstrapping. Bayesian inference analyses were performed in MrBayes v3.1¹⁷ with 5 000 000 generations, sampling trees every 100 generations. Stationarity was assessed using a convergence diagnostic. An average standard deviation of split frequencies (ASDSF) <0.02 was used as the criterion of convergence between both runs. *Trichomonas gallinae* (GenBank accession

numbers: EU215373) and *Trichomonas tenax* (D49495) were employed as the outgroup to root the resulting trees. A haplotype network for 18S rRNA sequences was also inferred through the median-joining network method using the Network v.4.5 software NETWORK v4.5.0.2.¹⁸

Analyses of genetic structure

Partitions of genetic diversity within and among populations were analysed through analysis of molecular variance (AMOVA)¹⁹ using the software Arlequin v3.5.1.2²⁰ with non-parametric permutations of 1000 times (significance levels $P=0.05$). Pairwise F_{ST} ²¹ values between populations were estimated to explore the levels of genetic differentiation among the populations using Arlequin v3.5.1.2.²⁰ The significance of F_{ST} was based on 1000 random permutations. Demographic and spatial parameters were assessed using the distribution of pairwise sequence differences (mismatch distribution) of Rogers and Harpending.²² Neutrality tests using Tajima's D and Fu's F_S were also applied through Arlequin v3.5.1.2.²⁰ as an assessment of possible population expansion.^{23,24}

Results

Genetic variation

The 18S rRNA gene of all collected isolates was amplified successfully, generating fragments ranging from 1495 to 1506 bp. The final sequence alignment contained 32 sequences and 1506 positions after trimming. The average base composition of 18S rRNA was 26.2% T, 25.3% A, 27.4% G and 21.1% C, with a little AT richness in the sequences. A total of 27 polymorphic sites were found, among which 21 were parsimony-informative and six singleton-variable. Eighteen haplotypes were identified in these polymorphic sites within 32 individuals from seven localities. Each sampled population and the total population displayed high haplotype diversity, accompanied by very low nucleotide diversity (Table 1).

Phylogenetic diversity

As evident from the phylogenetic tree (Fig. 1), the 18 haplotypes of *T. vaginalis* formed a single group in both MP and BI analyses with high support values (bootstrap value 100 and Bayesian posterior probability 1.0). Although no geographical clustering was observed from both phylogenetic analyses, two subclades of haplotype clustering were revealed in the phylogenetic tree, the first containing Hap14 and Hap16 (bootstrap value 82 and posterior probability 1.0) and the other including Hap7, Hap8 and Hap9 (bootstrap value 92 and posterior probability 1.0). Additionally, in the haplotype median-joining network, no apparent genetic structure was generated in the 18 haplotypes of *T. vaginalis* within

Table 1 Sampling haplotypes with frequencies and polymorphisms of the 18S rRNA sequences.

Sample sites	SS	Haplotypes (frequencies)	Pi ± SD	Hd ± SD
AY	5	Hap1(1), Hap2(1), Hap3(1), Hap4(2)	0.00362 ± 0.00081	0.900 ± 0.161
ZZ	6	Hap2 (1), Hap5(1), Hap6(2), Hap7(1), Hap8 (1)	0.00464 ± 0.00105	0.933 ± 0.122
LY	5	Hap6 (1), Hap8(1), Hap9(1), Hap10(1), Hap11(1)	0.00496 ± 0.00101	1.000 ± 0.126
SQ	3	Hap12(2), Hap13(1)	0.00044 ± 0.00021	0.667 ± 0.314
PDS	4	Hap15(3), Hap16(1)	0.00166 ± 0.00088	0.500 ± 0.265
ZMD	4	Hap12(1), Hap14(2), Hap15(1)	0.00321 ± 0.00094	0.833 ± 0.222
XY	5	Hap13 (1), Hap14(1), Hap15(1), Hap17(1), Hap18(1)	0.00505 ± 0.00132	1.000 ± 0.126
Total	32		0.00393 ± 0.00047	0.954 ± 0.019

SS, sampling size; Pi, nucleotide diversity; Hd, haplotype diversity; SD, standard deviation.

Figure 1 Phylogenetic relationships among *Trichomonas vaginalis* isolates from Henan province inferred by maximum parsimony (MP) and Bayesian inference (BI) based on 18S rRNA sequences, using *T. gallinae* and *T. tenax* as the outgroup. Numbers along the branches indicate posterior probabilities and bootstrap values resulting from different analyses in the order MP/BI.

the dataset (See supplementary material online for this article at <http://www.maneyonline.com/doi/suppl/10.1179/2047773215Y.0000000020>; Supplementary Figure S1).

Population structure

As shown in Table 2, highest genetic variance lay within populations (91.58%), relative to that among populations (8.42%). Pairwise fixation index (F_{ST}) values between specified regions were estimated to measure population differentiation (Table 3). Apart from F_{ST} values between SQ and ZMD ($F_{ST}=0.2863$) and SQ and PDS ($F_{ST}=0.4127$), most were lower than 0.25, and no pairwise F_{ST} values estimated were statistically significant. The results of neutral test analysis of total *T. vaginalis* population revealed a significant negative value of Fu's F_S (-4.93455 , $P=0.039$) and non-significant negative value of Tajima's D (-0.44240 , $P=0.396$). Mismatch distribution analysis disclosed almost unimodal frequency distribution of pairwise differences in the *T. vaginalis* population (See supplementary material online for this article at <http://www.maneyonline.com/doi/suppl/10.1179/2047773215Y.0000000020>; Supplementary Figure S2).

Discussion

The observed AT richness of 18S rRNA sequences was consistent with earlier data.^{11,13,25} The nucleotide diversity (Pi) value is an important index to measure the degree of genetic polymorphism within

Table 2 Analysis of molecular variance (AMOVA) based on 18S rRNA sequences of the populations of *Trichomonas vaginalis* collected from Henan province of China.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation index (F_{ST})
Among populations	6	23.088	0.24967	8.42	0.08424*
Within populations	25	67.850	2.71400	91.58	
Total	31	90.938	2.96367	100	

d.f.: degrees of freedom; * P -values ≤ 0.05 .

Table 3 Genetic differentiation among populations (F_{ST}) of *Trichomonas vaginalis* collected from Henan province of China.

Population	AY	ZZ	LY	SQ	PDS	ZMD	XY
AY	0.0000						
ZZ	-0.0210*	0.0000					
LY	0.1209	-0.0959	0.0000				
SQ	0.1607	0.0499*	0.2342	0.0000			
PDS	0.1106	0.0854	0.1667	0.4127	0.0000		
ZMD	0.1545	0.1111	0.1920	0.2863	-0.0513	0.0000	
XY	0.0252*	0.0835	0.1549	0.0996	-0.1164	-0.0433*	0.0000

Significance of chi-square: * P -values ≤ 0.05 .

a population, with $P_i > 0.01$ indicating comparatively large variations in most animals.²⁶ In our study, all nucleotide diversity values were lower than 0.01, suggesting low genetic variation of the 18 haplotypes. Both monophyly of the *T. vaginalis* group and no apparent genetic structure indicated that *T. vaginalis* isolates from central China form a unique population structure, consistent with the conclusions of Conrad *et al.*⁸ who analysed several hundred isolates of *T. vaginalis* from around the globe based on microsatellite data. Conrad and co-workers identified a unique two-type structure and associated clinically relevant phenotypes. However, our phylogenetic inference and network data showed a single one-type structure. Possible reasons for the discrepancies in results between the two studies are the geographical scale of isolate sampling and that our analysis only employed one molecular marker.

The fixation index (F_{ST}) is a measure of population differentiation due to genetic structure. Generally, an F_{ST} value within the range of 0–0.05 indicates little genetic differentiation, 0.05–0.15 moderate differentiation, 0.15–0.25 great differentiation and > 0.25 very great genetic differentiation.²⁷ The majority of our F_{ST} values were lower than 0.25, suggesting moderate-to-high genetic differentiation within the different populations. Under the assumption of neutrality, population expansion produces a negative value of Tajima's D and Fu's F_S tests. Tajima's D and Fu's F_S are sensitive to bottleneck effects or population expansion, causing these values to be more significantly negative.^{23,24} Results from our neutrality tests support possible population expansion of *T. vaginalis* from central China. Meanwhile, an almost unimodal frequency distribution of pairwise differences provides evidence of population expansion.^{22,28}

Conclusions

DNA polymorphism analysis revealed low genetic diversity of *T. vaginalis* isolates. Although moderate-to-high genetic differentiation between isolates from different sites was observed, *T. vaginalis* isolates from central China should be taken as a single population. Our data support demographic expansion of the *T. vaginalis* population.

Disclaimer statements

Contributors

All named authors contributed to this article.

Funding

This work was supported by the Innovation Program of Zhengzhou University (No. 2014xjxm344).

Conflict of interest

The authors declare that they have no competing interests.

Ethics approval

Human Ethical Committees of Zheng Zhou University, and permission from Department of Department of Obstetrics and Gynecology, the Second Affiliated Hospital of Zhengzhou University.

References

- 1 World Health Organization. Global prevalence and incidence of selected curable sexually transmitted infections: overview and estimates. Geneva: WHO; 2001.
- 2 Fraga J, Rojas L, Sarrago I, Fernández-Calienes A. Genetic characterization of three Cuban *Trichomonas vaginalis* virus. Phylogeny of Totiviridae family. Infect Genet Evol. 2012; 12:113–20.
- 3 Cotch MF, Pastorek JG, Nugent RP, Hillier SL, Gibbs RS, Martin DH, *et al.* *Trichomonas vaginalis* associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group. Sex Transm Dis. 1997;24:353–60.
- 4 Smith LM, Wang M, Zangwill K, Yeh S. *Trichomonas vaginalis* infection in a premature newborn. J Perinatol. 2002;22:502–03.
- 5 Krieger JN. Consider diagnosis and treatment of trichomoniasis in men. Sex Transm Dis. 1999;27:241–42.
- 6 Laga M, Manoka A, Kivuvu M, Malele B, Tulipa M, Nzila N, *et al.* Non-ulcerative sexually transmitted diseases as risk factor for HIV-1 transmission in women: results from a cohort study. AIDS. 1993;7:95–102.
- 7 Viikki M, Pukkala E, Nieminen P, Hakama M. Gynaecological infections as risk determinants of subsequent cervical neoplasia. Acta Oncol. 2000;39:71–5.
- 8 Conrad MD, Gorman AW, Schillinger JA, Fiori PL, Arroyo R, Malla N. Extensive genetic diversity, unique population structure and evidence of genetic exchange in the sexually transmitted parasite *Trichomonas vaginalis*. PLoS Negl Trop Dis. 2012; 6:e1573.
- 9 Xie Q, Tian X, Qin Y, Bu W. Phylogenetic comparison of local length plasticity of the small subunit of nuclear rDNAs among all Hexapoda orders and the impact of hyper-length-variation on alignment. Mol Phylogenet Evol. 2009;50:310–16.
- 10 Delgado-Viscogliosi P, Viscogliosi E, Gerbod D, Kulda J, Sogin ML, Edgcomb VP. Molecular phylogeny of parabasalids based on small subunit rRNA sequences, with emphasis on the Trichomonadidae subfamily. J Eukaryot Microbio. 2000;47: 70–75.
- 11 Dimasuay KGB, Lavilla OJY, Rivera WL. New hosts of *Simplicimonas similis* and *Trichomonas batrachorum* identified by 18S ribosomal RNA gene sequences. J Parasitol Res. 2013;83: 1947.
- 12 Stiles JK, Shah PH, Xue L, Meade JC, Lushbaugh WB, Cleary JD. Molecular typing of *Trichomonas vaginalis* isolates by Hsp70 restriction fragment length polymorphism. Am J Trop Med Hyg. 2000;62:441–5.
- 13 Dimasuay KGB, Rivera WL. Molecular characterization of trichomonads isolated from animal hosts in the Philippines. Vet Parasitol. 2013;196:289–95.
- 14 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al.* Clustal W and clustal X version 2.0. Bioinformatics. 2007;23:2947–8.
- 15 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–9.
- 16 Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25: 1451–2.
- 17 Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19: 1572–4.
- 18 Polzin T, Daneshmand SV. On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett. 2003;31:12–20.
- 19 Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992;131:479–91.
- 20 Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resou. 2010;10:564–7.

- 21 Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 1995;139:457–62.
- 22 Rogers AR, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol*. 1992;9:552–69.
- 23 Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–95.
- 24 Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 1997;147:915–25.
- 25 Dimasuay KGB, Rivera WL. First report of *Trichomonas tenax* infections in the Philippines. *Parasitol Int*. 2014;63:400–02.
- 26 Neigel JE, Avise JC. Application of a random walk model to geographic distribution of animal mitochondrial DNA variation. *Genetics*. 1993;135:1209–20.
- 27 Balloux F, Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. *Mol Ecol*. 2002;11:155–65.
- 28 Zhang X, Cui J, Liu LN, Jiang P, Wang H, Qi X, *et al*. Genetic structure analysis of *Spirometra erinaceieuropaei* isolates from central and southern China. *PLoS One*. 2015;8:e78064.