Curr Genet (2009) 55:199–210
DOI 10.1007/s00294-009-0237-y

RESEARCH ARTICLE

Comparative analyses among the *Trichomonas vaginalis*, *Trichomonas tenax*, and *Tritrichomonas foetus* 5S ribosomal RNA genes

Ana Lilia Torres-Machorro · Roberto Hernández · John F. Alderete · Imelda López-Villaseñor

Received: 4 December 2008 / Revised: 23 February 2009 / Accepted: 28 February 2009 / Published online: 17 March 2009

Abstract The 5S ribosomal RNA (5S rRNA) is an essential component of ribosomes. Throughout evolution, variation is found among 5S rRNA genes regarding their chromosomal localization, copy number, and intergenic regions. In this report, we describe and compare the gene sequences, motifs, genomic copy number, and chromosomal localization of the *Trichomonas vaginalis*, *Trichomonas tenax*, and *Tritrichomonas foetus* 5S rRNA genes. *T. vaginalis* and *T. foetus* have a single type of 5S rRNA-coding region, whereas two types were found in *T. tenax*. The sequence identities among the three organisms are between 94 and 97%. The intergenic regions are more divergent in sequence and size with characteristic species-specific motifs. The *T. foetus* 5S rRNA gene has larger and more complex intergenic regions, which contain either an ubiquitin gene or repeated sequences. The 5S rRNA genes were located in Trichomonads chromosomes by fluorescent in situ hybridization.

Communicated by S. Hohmann.

Nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession numbers FJ492747, FJ492748, FJ492749, FJ492750, and FJ492751.

A. L. Torres-Machorro · R. Hernández · I. López-Villaseñor (✉)
Departamento de Biología Molecular y Biotecnología,
Instituto de Investigaciones Biomédicas,
Universidad Nacional Autónoma de México,
Apartado Postal 70-228, CP 04510 Mexico D.F., Mexico
e-mail: imelda@biomedicas.unam.mx

J. F. Alderete
School of Molecular Biosciences,
Washington State University, Pullman, WA, USA

Keywords Ribosome · Intergenic region · RNA polymerase III · Ubiquitin · Repeated sequences · Chromosome

Introduction

Trichomonas vaginalis, *T. tenax*, and *T. foetus* are parasitic amitochondriate protozoa that inhabit the urogenital tract of humans, the oral cavity of humans, and the urogenital tract of cattle, respectively. Phylogenetic analyses based on the 16S ribosomal RNA (rRNA) gene place these organisms near the base of the eukaryotic tree (Sogin 1991). This characteristic, along with their amitochondriate nature and prokaryotic-like ribosomes, suggest that Trichomonads are ancestral eukaryotic organisms that conserve early regulatory mechanisms (Sogin 1991). Recently, the egalitarian model of evolution (Dacks et al. 2008) has challenged the previous proposal: it divides eukaryotes into six major super-groups, although the relationship among them and the order of branching remain unknown. This new eukaryotic phylogeny emphasizes that organisms like Trichomonads are not primitive, but instead highly evolved parasites specialized for their specific environments. Today, the controversy continues and the phylogenetic location of Trichomonads remains unclear.

Since ribosomes from Trichomonads have characteristics evocative of prokaryotes (Chakrabarti et al. 1992); the study of their rRNA genes may lead to the identification of some eukaryotic divergent and/or specific regulatory processes. The typical eukaryotic ribosome is composed of four rRNAs and multiple proteins. Three of the rRNA molecules (18S, 5.8S, and 28S) are encoded in the rRNA gene (rDNA) and are co-transcribed by RNA polymerase I (pol I). The fourth rRNA component, the 5S rRNA, is transcribed by

RNA polymerase III (pol III) and the genes are not generally linked to the rDNA unit (Paule and White 2000).

The 5S rRNA gene is generally reiterated in the genome and organized in tandem in head-to-tail repeats, which can be encoded in one or in multiple gene clusters (Long and Dawid 1980). Great variation is found in the localization and number of chromosomal sites of the 5S rRNA genes throughout evolution. Depending on the species, the gene copy number per genome can range from 3 copies in *Plasmodium falciparum* (Shippen-Lentz and Vezza 1988) to about 10^6 copies in *Euplotes carassus* (Roberson et al. 1989). These genes can be localized in a single or in multiple chromosomes or chromosome loci. In general, all the copies of the 5S rRNA gene are considered identical in a particular organism. However, some eukaryotes such as *P. falciparum* (Shippen-Lentz and Vezza 1988), *Cryptosporidium parvum* (Taghi-Kilani et al. 1994), *Neurospora crassa* (Selker et al. 1981), and mouse and human (Hallenberg et al. 1994) have sequence and length variation in the 5S rRNA intergenic region and eventually in the 5S rRNA-coding region. It has been reported that the tandem repeats of the 5S rRNA genes can be linked to different multigene families. Among these are the rDNA in *Dictyostelium discoideum* (Cockburn et al. 1978), some tRNA genes in *Entamoeba histolytica* (Clark et al. 2006), the spliced leader gene in *Trypanosoma vivax* (Roditi 1992), the histone gene cluster in crustaceans (Pelliccia et al. 2001), and the ubiquitin genes in *Tetrahymena pyriformis* (Guerreiro et al. 1993).

The 5S rRNA genes are transcribed by pol III. In contrast to pol I and pol II promoters, the pol III promoter of this gene is internal to the rRNA-coding region (internal control region, ICR) (Schramm and Hernandez 2002). In some organisms, the ICR alone is sufficient to direct pol III transcription, but in others such as yeast, the 5S rRNA gene transcription may also be regulated by upstream regulatory sequences, such as an upstream promoter element (upe) and a start site element (sse) (Lee et al. 1997).

In a previous work, we reported the sequence of the *T. vaginalis* 5S rRNA gene and identified putative pol III regulatory elements such as an internal promoter, upstream regulatory sequences, and a palindromic sequence within the intergenic region that may indicate a high-energy hairpin structure (Torres-Machorro et al. 2006). In this study, we compare the 5S rRNA gene sequence, genomic organization, and chromosome localization among three Trichomonad species. We found at least two types of 5S rRNA genes in each of the three species, although most of the sequence variation, within and between species, was found in the 5S rRNA intergenic region. *T. vaginalis* and *T. tenax* are more similar in the 5S rRNA gene sequence, organization, size, and chromosomal localization, while *T. foetus* has larger intergenic regions with particular species-specific motifs, such as an ubiquitin open reading frame (ORF) and repeated sequences.

Materials and methods

Trichomonads and culture conditions

Trichomonas vaginalis CNCD147 isolate (Alvarez-Sanchez et al. 2000) and *T. foetus* (02-97) were grown as previously described (Espinosa et al. 2002). The *T. tenax* Hs-4:NIH was grown in LYI *Entamoeba* medium supplemented with 10% fetal bovine serum as recommended by ATCC.

Trichomonads 5S rRNA gene cloning and sequencing

To clone the 5S rRNA intergenic regions of the three Trichomonads, we took advantage of the tandem organization or these genes and of their sequence conservation (Torres-Machorro et al. 2006). Two independent PCR amplifications were made with genomic DNA (prepared as described by Espinosa et al. 2002) and the 5I-F (5′-TACTGGGCT AGGAGACTT-3′) and 5I-R (5′-CGCGACTGCTTAA TTTCCA-3′) oligonucleotides (oligos), which align within the 5S rRNA gene-coding region and extend toward the intergenic region (Fig. 1a). The amplification reactions consisted of 30 cycles of 92°C denaturing for 1 min, 46°C annealing for 1 min, and 72°C extension for 3 min, with *Taq* DNA polymerase (recombinant, Invitrogen). The PCR products were cloned using the Zero Blunt® TOPO® PCR cloning kit (Invitrogen). Eighteen clones from *T. vaginalis*, 11 from *T. tenax*, and 9 from *T. foetus* were sequenced.

To clone the Type B-Tv 5S rRNA intergenic region of the *T. vaginalis* CNCD147 isolate, a genomic PCR was made using the Type B-Tv specific oligonucleotide RI335 (5′-CAAAATTAAACAAAATTTTCAGCAGAGAGTGGTGT -3′) and the 5S-R (5′-AAGCAAGCACCGCACGTTCC-3′) oligo. The annealing temperature was 54°C. The PCR product was cloned using the Zero Blunt TOPO PCR cloning kit (Invitrogen) and four clones were sequenced.

To clone the 5S rRNA gene-coding region, PCR amplifications were made using genomic DNA and the 5S-F (5′-A AGCGGCCACACCCGGCTGG-3′) and 5S-R oligos (Fig. 1a). The PCR amplification conditions used were the same as described for the intergenic region except for changing the annealing temperature to 53°C. For each of the three organisms, at least three independent clones of the coding region were sequenced. These 5S gene PCR products were also used to estimate the gene copy number. To obtain the complete 5S rRNA gene sequence, intergenic and coding region sequences were assembled as shown in

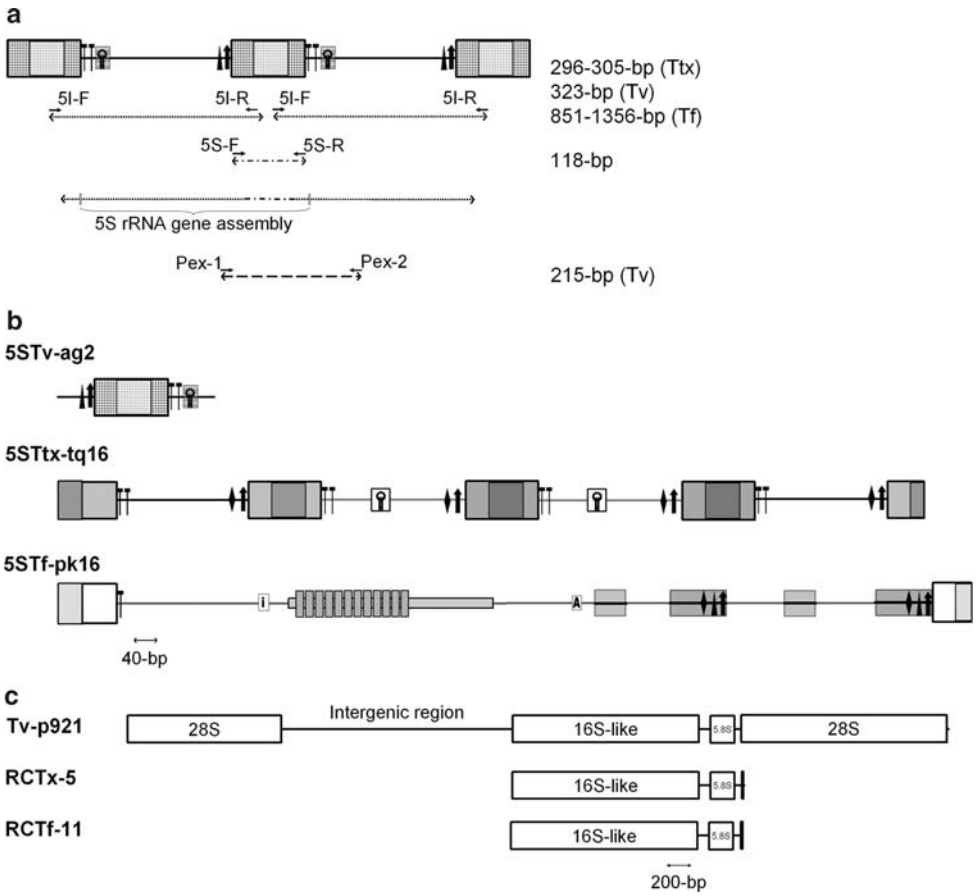

Fig. 1 Trichomonads 5S rRNA gene cloning and sequencing. **a** Schematic diagram of the tandem organization of the 5S rRNA genes. The regions where oligonucleotides anneal in the 5S rRNA genes are shown with *arrows*. The PCR amplification products are shown with *different line* styles. The assembly strategy is also shown. **b** Trichomonads 5S rRNA gene fragments (obtained from TOPO clones) used for the synthesis of probes. **c** Trichomonads rDNA fragments used as probes for the FISH assays. The *T. vaginalis* Tv-p921 clone is a genomic clone. RcTx-5 and RcTf-11 are clones obtained by PCR amplification and cloning of *T. tenax* and *T. foetus* rDNA

Fig. 1a. Two clones were obtained containing several linked 5S gene copies (see clone 5STtx-tq16, Fig. 1b), which allowed us to discriminate between the two types of coding regions in *T. tenax*.

5S rRNA gene copy number in *Trichomonas tenax* and *Tritrichomonas foetus*

The 5S rRNA gene copy number was estimated by a comparative quantitative hybridization approach essentially as described (Torres-Machorro et al. 2006; Chakrabarti et al. 1992), except for the use of a slot blot instead of a Southern blot approach. Briefly, genomic DNA from *T. tenax* and *T. foetus* was digested with *Hin*dIII, phenol extracted, ethanol precipitated, and resuspended in 0.4 M NaOH/10 mM EDTA solution. Decreasing amounts of this DNA were loaded in a slot blot apparatus (Hoefer Scientific) alongside decreasing amounts of the corresponding 5S rRNA-coding region. This 116 bp 5S rRNA gene PCR product represents one gene copy. Additionally, the same 5S gene-coding region PCR fragment was radioactively labeled (Rediprime II, GE Healthcare) and used as a probe. The conditions used for hybridization were as described (Espinosa et al. 2002).

Fluorescence in situ hybridization

Trichomonas vaginalis, *T. foetus*, and *T. tenax* cells were prepared for in situ hybridization using the hypotonic swelling technique (Xu et al. 1998) with the following modifications. Cells grown in TYM medium with 10% horse serum were not treated with colchicine. Cells were grown to a density of 2×10^6 cells/ml, diluted to 1×10^6 cells/ml in the same medium, grown further for 3 h at 37°C, and harvested. With this modification, cells with nuclei in multiple states of chromosome condensation were obtained. After centrifugation, cells were resuspended in 10 ml of 10 mM KCl and incubated at 37°C for 5 min, followed by the addition of 2 ml of fixative solution (chilled methanol acetic acid 3:1) as an important pre-fixing step. The nuclei were then centrifuged at $200 \times g$ for 10 min at room temperature, the supernatant was removed and the nuclei pellet was resuspended in 10 ml of ice-cold fixative solution. This last step was repeated two more times. Finally, nuclei were resuspended in a final volume of 500 µl. As much as 60 µl of the nuclei suspension was dropped from a 1.3 m height onto clean, grease-free slides. The smears on the slides were heat-dried. The slides were aged in a nitrogen atmosphere for 3 days. Nuclei were fixed by immersion of the

slides in 1% paraformaldehyde in 2X SSC (pH = 7.4) for 30 s, washed twice in 2X SSC (pH = 7.4), and air dried. The slides were then treated with 100 µg/ml ribonuclease A (Sigma) in 2X SSC at 37°C for 1 h. After three washes in 2X SSC, the DNA on the slides was denatured in 70% formamide/2X SSC (pH = 7.4) at 75°C for 2 min. They were immediately placed in 70% ethanol for 2 min, passed through dehydration series of 80 and 100% ethanol for 2 min each, and allowed to air dry. As much as 20 µl of the heat-denatured probe (10 ng/µl) in hybridization solution [50% formamide, 10% dextran sulfate (Sigma), 2X SSC] was added to each slide, covered with a plastic coverslip and incubated overnight in a slightly dampened hybridization box at 37°C. The slides were washed with 50% formamide/2X SSC for 5 min at 37°C, then in 2X SSC and 4X SSC at room temperature for 5 min. Spreads were stained with 0.125 µg/µl DAPI (Sigma) for 1 min, washed with PBS pH = 7.4, and mounted in Vectashield (Vector Laboratories). Samples were visualized with an epifluorescence microscope (BX5I Olympus) using objective Uplan Fl 100X/1.30 Oil ∞/0.17. Images were acquired with an Evolution VF color digital camera (Media Cybernetics) and were analyzed with the Adobe Photoshop software.

Preparation of probes for FISH assays

The *T. vaginalis* 5S gene probe corresponds to the Type A-Tv 5S rRNA gene from position −11 to +194 considering the transcription start point as +1 (Fig. 1b). It includes the whole 5S rRNA-coding region (118 bp) and 87 bp of the intergenic region. The probe was directly labeled with Cy™3-dCTP (GE Healthcare) by PCR amplification of the *T. vaginalis* 5S rRNA gene clone 5STv-ag2 (Fig. 1b) with 300 ng of each Pex1 (5′-CGAACAATCCTTTCGA AGCG-3′) and Pex2 (5′-AAGAGTTTTAGCAGGGAC AA-3′) oligos (Fig. 1a). The PCR conditions were as described for the amplification of the 5S rRNA intergenic region, replacing the dCTP with Cy3-dCTP. Salmon sperm DNA (20 µg; Invitrogen) was added to the PCR final product, which was precipitated with ethanol. After centrifugation, the pellet was dissolved in distilled water to a final concentration of ∼ 100 ng/µl. Before use, the probe was diluted (1:10) in hybridization solution. All probes were treated in a similar way before use. The *T. tenax* 5S gene probe contains two complete Type A-Ttx 5S rRNA gene-coding and intergenic regions and two complete Type B-Ttx 5S rRNA gene-coding and intergenic regions (1,235 bp insert in clone 5STtx-Tq16, Fig. 1b). The *T. foetus* 5S gene probe contains 94% of the 5S rRNA-coding region and the complete Type B-Tf 5S rRNA intergenic region (1,356 bp insert in clone 5STf-pk16, Fig. 1b). The probes were labeled with Cy3-dCTP using a Nick Translation System (Invitrogen).

The *T. vaginalis* rRNA gene probe contained the complete 6 kb rDNA unit of *T. vaginalis* (16S-like rRNA, 5.8S rRNA, 28S rRNA, and intergenic region). The Tv-p921 genomic clone insert (Fig. 1c) was labeled by nick-translation with Cy3-dCTP. The *T. tenax* and *T. foetus* rRNA gene probes correspond to a 1,894 bp fragment of the rDNA coding region, which includes the complete 16S and 5.8S rRNA-coding sequences, and 29 bp of the 28S rRNA (Fig. 1c). The inserts of clones RCTx-5 and RCTf-11 (Fig. 1c) were labeled with Cy3-dCTP using the Nick Translation System (Invitrogen).

Results

Trichomonads 5S rRNA gene cloning, sequencing, and assembly

A *T. vaginalis* 5S rRNA gene clone was previously obtained using a PCR approach (Torres-Machorro et al. 2006). The same approach was used for the amplification, cloning, and sequencing of the 5S rRNA genes of *T. tenax* and *T. foetus*, and to extend the *T. vaginalis* 5S rRNA gene analysis. For each organism, two different genomic PCRs were made: one to amplify the 5S rRNA-coding region and a second to amplify the 5S rRNA intergenic region. Both PCR products, which overlap in a 90 bp region, were cloned, sequenced, and assembled (Fig. 1a). Clones containing one or several tandem 5S rRNA genes were obtained.

The 5S rRNA gene in Trichomonads

In a previous report, only one type of 5S rRNA gene and intergenic region was experimentally found in the *T. vaginalis* CNCD147 isolate (Torres-Machorro et al. 2006). In contrast, in an in silico analysis of the *T. vaginalis* G3 strain, whose complete genome has been sequenced (Carlton et al. 2007), two types of *T. vaginalis* 5S rRNA genes could be identified. These 5S rRNA gene variants, named here by us as Type A-Tv and Type B-Tv 5S rRNA genes, have identical coding sequences and a similar gene copy number. Their intergenic regions differ in size by one nucleotide and have a sequence identity of 84%. The Type B-Tv 5S rRNA gene has a TATA-like box located in the putative upstream regulatory region, which is not present in the Type A-Tv 5S gene (Fig. 2a, b; Table 1). The G3 isolate Type A-Tv 5S rRNA gene is identical in sequence to the one previously found in the CNCD147 isolate.

Trichomonas tenax has at least two types of 5S rRNA genes, which measure 307 bp (Type A-Ttx 5S gene) and 316 bp (Type B-Ttx 5S gene) (Fig. 2c, d). The coding regions were assigned by alignment with the published

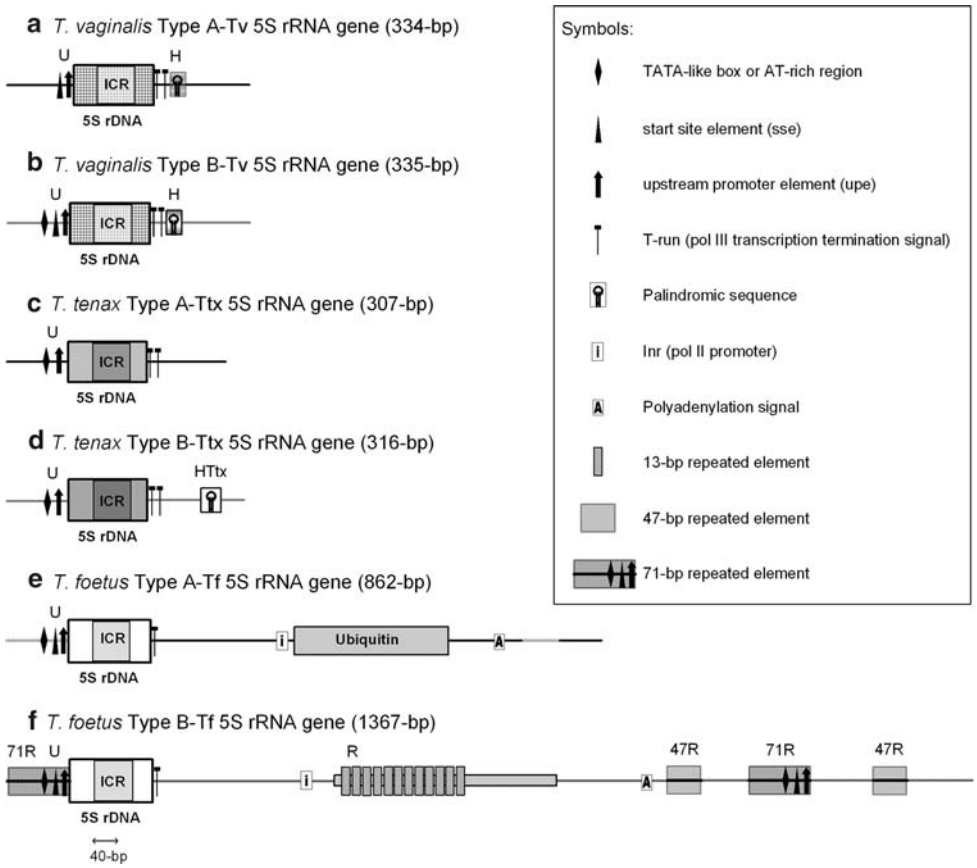

Fig. 2 The 5S rRNA gene organization in Trichomonads. Graphical representation of the two types of 5S rRNA gene organization in the three studied Trichomonads. A single copy of each variant is depicted, although the genes are found as tandem repeats in the genome. The 5S rRNA-coding regions (118 bp) are shown as *boxes*. The pol III promoter (ICR, internal control region) is internal to the coding region. The 5S rRNA upstream regulatory region (U) can have an upstream promoter element (upe, *triangle*), a start site element (sse, *arrow*), and an AT-rich region or TATA-like box (*diamond*), and its composition varies depending on the 5S rRNA gene type. The coding region is followed by runs of Ts (T) of different lengths, which act as the pol III transcription termination signal. In *T. vaginalis*, the two types of 5S rRNA genes have a palindromic sequence (H) that can have slight sequence variations. The *T. tenax* 5S rRNA gene has two types of coding and intergenic regions. The Type B-Ttx 5S gene has a unique palindromic region (HTtx) in the middle of the intergenic region. In *T. foetus*, the Type A-Tf 5S gene is characterized by the presence of an ubiquitin ORF (in the intergenic region) with elements for transcription by pol II: the promoter (Inr, i) and the polyadenylation signal (A). The Type B-Tf 5S gene is characterized by the presence of an interrupted ubiquitin ORF and repeated sequences (R) of variable lengths in the intergenic region

Table 1 Comparison of the 5S rRNA gene sequences and copy number in Trichomonads

	Trichomonas vaginalis (%)	*Trichomonas tenax*		*Tritrichomonas foetus* (%)			
		Ttx A (%)	Ttx B (%)				
Coding region identity							
Tv	100	96	97	95			
Ttx A		100	97	94			
Ttx B			100	96			
		A-Tv (%)	B-Tv (%)	A-Ttx (%)	B-Ttx (%)	A-Tf (%)	B-Tf (%)
Intergenic region identity							
A-Tv (334 bp)		100	84	49	55	42	34
B-Tv (335 bp)			100	43	43	36	43
A-Ttx (307 bp)				100	57	47	46
B-Ttx (316 bp)					100	38	35
A-Tf (862 bp)						100	87
5S gene genome percentage (%)		0.10		0.01		0.04	

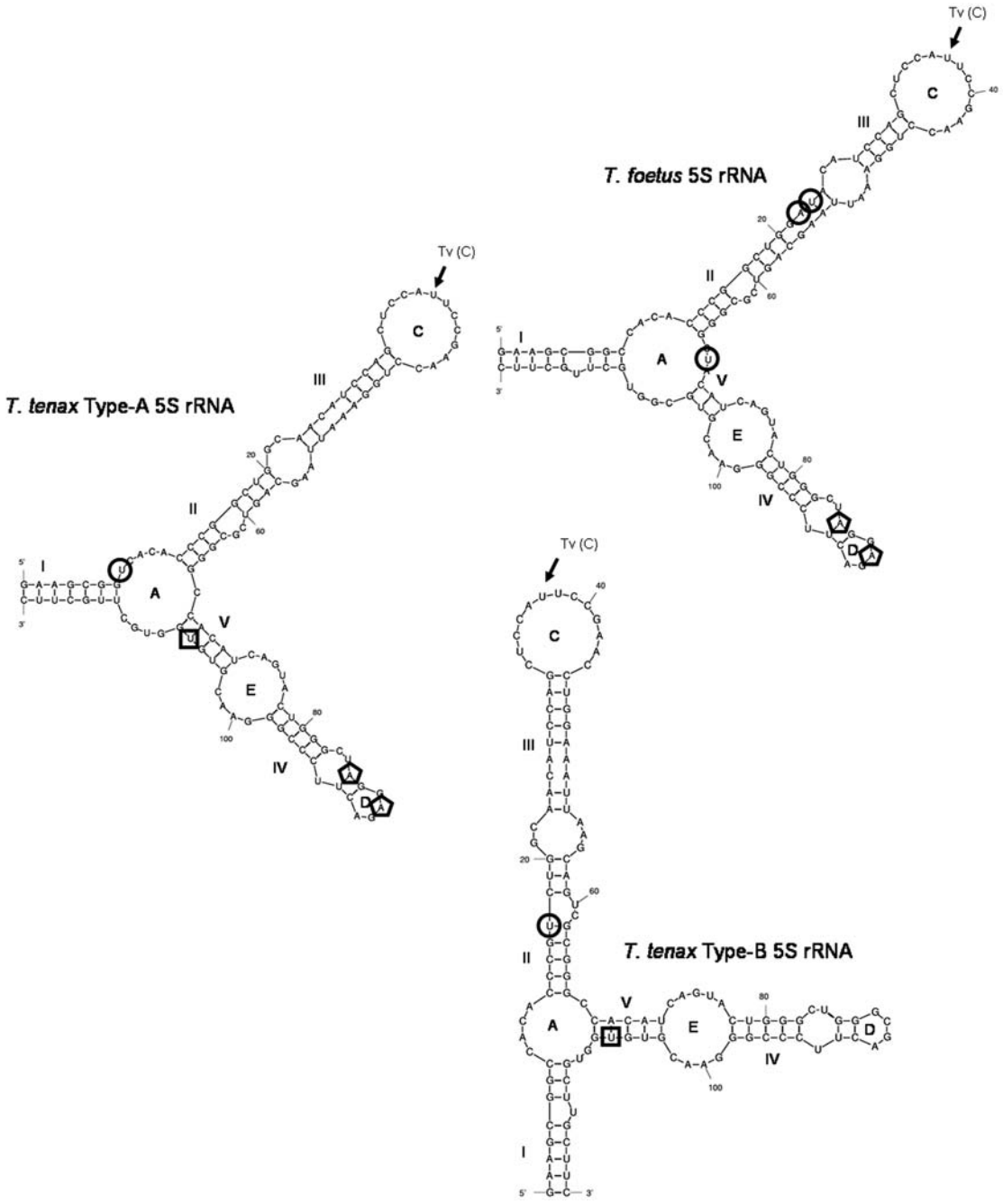

Fig. 3 Predicted secondary structure for the *T. tenax* (Type A and B) and *T. foetus* 5S rRNAs. Helixes are denoted by *roman numerals* and loops by *capital letters*. Sequences were folded with Zuker's mfold software (http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi). *Circles* depict nucleotides specific of each type of 5S rRNA species. *Squares* highlight *T. tenax* 5S rRNA specific nucleotides. *Pentagons* show *T. tenax* and *T. foetus* Type B 5S rRNA shared nucleotides. Tv(C) indicates the localization of the *T. vaginalis* 5S rRNA-specific nucleotide

T. vaginalis sequence. Both variants contain the putative internal pol III promoter. The 5S rRNA putative upstream regulatory region of both types of *T. tenax* genes is composed of a start site element, a GC-rich region, and a TATA-like box (Fig. 2c, d). The two types of 5S rRNA gene-coding regions (118 bp) are 97% identical in sequence, differing only in four nucleotides (Fig. 2; Table 1). The *T. tenax* 5S rRNA predicted secondary structures have two nucleotide changes in paired regions (stems) and two changes in the A and D loops (Fig. 3). The secondary structure of the paired regions is conserved in both types of molecules, changing only the pairing between G-C and A-U to the non-canonical pairing G-U (Fig. 3).

Fig. 4 Sequence comparison of the three Trichomonads 5S rRNA-coding region. Alignment of the 5S rRNA-coding sequences found in Trichomonads. The pol III promoter (ICR) is composed of a *box A*, an intermediate element (IE) and a *box C*. *T. tenax* has two types of 5S rRNA-coding sequences, named 5STtxA and 5STtxB, while a single coding sequence was found in *T. vaginalis* and in *T. foetus*. The species-specific sequence variations are denoted in *black fonts*

Tritrichomonas foetus has at least two types of 5S rRNA genes. These genes are identical in the coding region, but differ in the intergenic region sequence. We named these genes as Type A-Tf 5S rRNA gene (862 bp) and Type B-Tf 5S rRNA gene (~ 1,300 bp) (Fig. 2e, f). The *T. foetus* 5S rRNA-coding region (118 bp) conserves the putative pol III promoter sequence and secondary structure described in the *T. vaginalis* 5S rRNA gene (Fig. 3). Compared to the *T. vaginalis* 5S rRNA gene-coding region, it has three nucleotide variations (Fig. 4). The putative upstream regulatory region in both types of 5S rRNA gene are composed of a start site element, an upstream promoter element, and an AT-rich region (Fig. 2e, f). The typical transcription termination signal for pol III, a run of Ts, was found in the 3′-side of the 5S rRNA-coding region of the three Trichomonads (Fig. 2).

The 5S rRNA gene-coding region sequence was compared among the three Trichomonads, and a very high similarity was found among the three with a sequence identity ranging from 94 to 97% (Table 1). The Type B-Ttx and the *T. vaginalis* 5S rRNA gene-coding sequences are the most related pair with 97% sequence identity. We found that each Trichomonad has specific nucleotide variations within the 5S rRNA-coding region, and these are "species-specific" (Fig. 4) and useful for the delineation among the Trichomonad species. It is noteworthy that the expression of both types of 5S rRNA genes in *T. tenax* and *T. foetus* was experimentally confirmed by RT-PCR (data not shown).

The 5S rRNA intergenic region motifs in Trichomonads

Two different 5S rRNA intergenic regions were identified in silico in the genome of the *T. vaginalis* G3 isolate, which differ in size by one nucleotide and have a sequence identity of 84%. Since a single 5S gene sequence was previously found in the *T. vaginalis* CNCD147 isolate (Torres-Machorro et al. 2006), we experimentally searched for the second type of 5S rRNA intergenic region in this isolate. Type B-Tv 5S gene could only be identified by PCR amplification using Type B-Tv 5S gene-specific oligonucleotides. With this approach, eight sequence variants in the palindromic region were also identified (Fig. 5). Nevertheless, sequence variation in the hairpin region was also found in an in silico analysis of the 5S rRNA gene in the *T. vaginalis* G3 isolate (Fig. 5).

Trichomonas tenax has at least two types of 5S rRNA gene, with 307 and 316 bp in length (Fig. 2c, d). The *T. tenax* 5S rRNA intergenic regions are only 59% identical in sequence (Table 1) and differ by the presence of a putative high-energy hairpin (−9.6 kcal/mol) located in position −118 of Type B-Ttx 5S gene intergenic region (Fig. 2d). This hairpin shares no sequence or position similarity with the one found in the *T. vaginalis* 5S rRNA gene. Six Type A and five Type B clones were identified by sequence, suggesting a similar number of 5S rRNA genes from each type.

Trichomonas foetus has at least two types of 5S rRNA genes with 862 and 1,262–1,367 bp in length. The Type A-Tf 5S rRNA intergenic region (744 bp) contains a complete ORF for ubiquitin (Fig. 2e) and a typical pol II promoter, which is identical to the Inr previously described in *T. vaginalis* (Liston and Johnson 1999). In the 3′-end of the ubiquitin gene the mRNA polyadenylation signal for *T. vaginalis* was found (Espinosa et al. 2002). The Type B-Tf 5S gene intergenic region can vary in size from 1,144 to 1,249 bp, depending on the number of repetitions of a 13 bp repeated element. The intergenic region is characterized by the presence of two blocks of repeated sequences: the first block is composed of between 5 and 12 tandem repeated motifs (13 bp long), which were found to interrupt the ubiquitin-coding region. The second block is composed of two sequences of 47 and 71 bp, with each sequence repeated once. The 47 bp repeated sequences are located in positions −233 and −516 considering the tsp as +1 (Fig. 2f). Interestingly, the 71 bp block contains the putative 5S gene upstream regulatory region (positions −76 and −408). Remarkably, the two types of 5S rRNA intergenic regions in *T. foetus* are very similar in sequence (87% of identity, Table 1) even though they differ in size by about 450 bp. This variation in size is mainly due to the repetitive nature

Name	Sequence	dG (kcal/mol)
G3-1	GGTGCAGGGCTATTGCCCTGCACC	-13.02
G3-2	--g----------	-10.62
147-1	----------a--a----------	-13.12
147-2	----------t----	-7.96
147-3	----------a----	-7.94
147-4	----t----------	-9.49
147-5	-a----------	-10.19
147-6	---t--a----------	-3.17

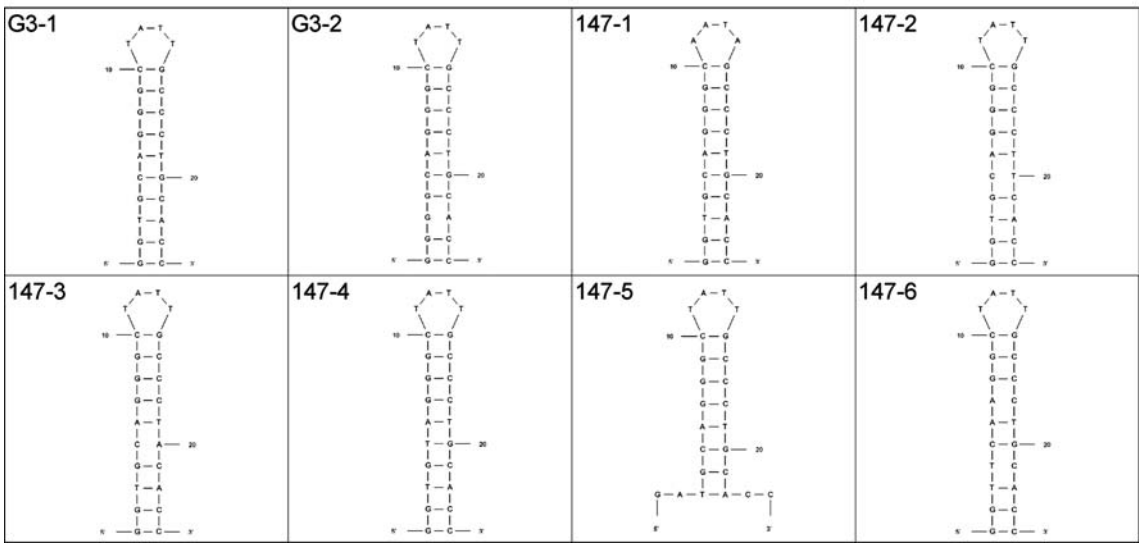

Fig. 5 Variation in the palindromic sequence found in the 5S rRNA intergenic region of the *T. vaginalis* G3 strain and CNCD147 isolate. All the hairpin sequence variations, except G3-2, correspond to sequenced palindromic region variants found in the 5S rRNA gene intergenic region of the CNCD147 isolate. The first two lines correspond to the hairpin variants found in the G3 isolate. The putative hairpin structures were obtained using the software Mfold (Zuker 2000)

of the Type B-Tf intergenic region (Fig. 2f). Three Type A-Tf and six Type B-Tf clones were obtained, suggesting the presence of more Type B-Tf gene copies in the *T. foetus* genome.

5S rRNA gene copy number and chromosomal localization in Trichomonads

To find out the approximate genomic copy number of the 5S rRNA gene in *T. tenax* and *T. foetus*, a quantitative–qualitative hybridization analysis was used (Fig. 6). This type of quantification is based on the known genome size of a given organism. The genome size of *T. vaginalis* is controversial, ranging in proposed sizes from 2.5×10^7 to 17.6×10^7 bp (Wang and Wang 1985; Carlton et al. 2007). On the other hand, the size of *T. tenax* and *T. foetus* genomes are also unclear, since their size determination was based on the *T. vaginalis* genome (Zubacova et al. 2008). Nevertheless, it is apparent that the genome size is similar among the three species (Zubacova et al. 2008). Therefore, due to the difficulty in defining a reliable 5S rRNA gene copy number, we estimated the percentage of the genome sequence that corresponds to the 5S rRNA gene-coding region in each Trichomonad genome (Table 1). The 5S rRNA genes coding region correspond to 0.10% of the genome in *T. vaginalis* and 0.04% of the genome in *T. foetus*, whereas in *T. tenax* this value is about one-tenth of the *T. vaginalis* (0.01% of the genome).

The chromosomal localization of the 5S rRNA genes in Trichomonads was determined by a fluorescent in situ hybridization (FISH) procedure. The chromosomal localization of the rDNA genes was also determined as a control. It has been reported that *T. vaginalis* and *T. tenax* have six chromosomes (Lehker and Alderete 1999; Zubacova et al. 2008), while five chromosomes have been identified in *T. foetus* (Xu et al. 1998). As shown in Fig. 7, the 5S rRNA gene is located in two chromosomes in *T. vaginalis* and in *T. tenax*, whereas this gene is found in a single chromosome in *T. foetus* (Fig. 7a).

Due to the similar size of some of the *T. vaginalis* and *T. tenax* chromosomes, it was difficult to classify and/or name them using our microscopy images. The *T. vaginalis*

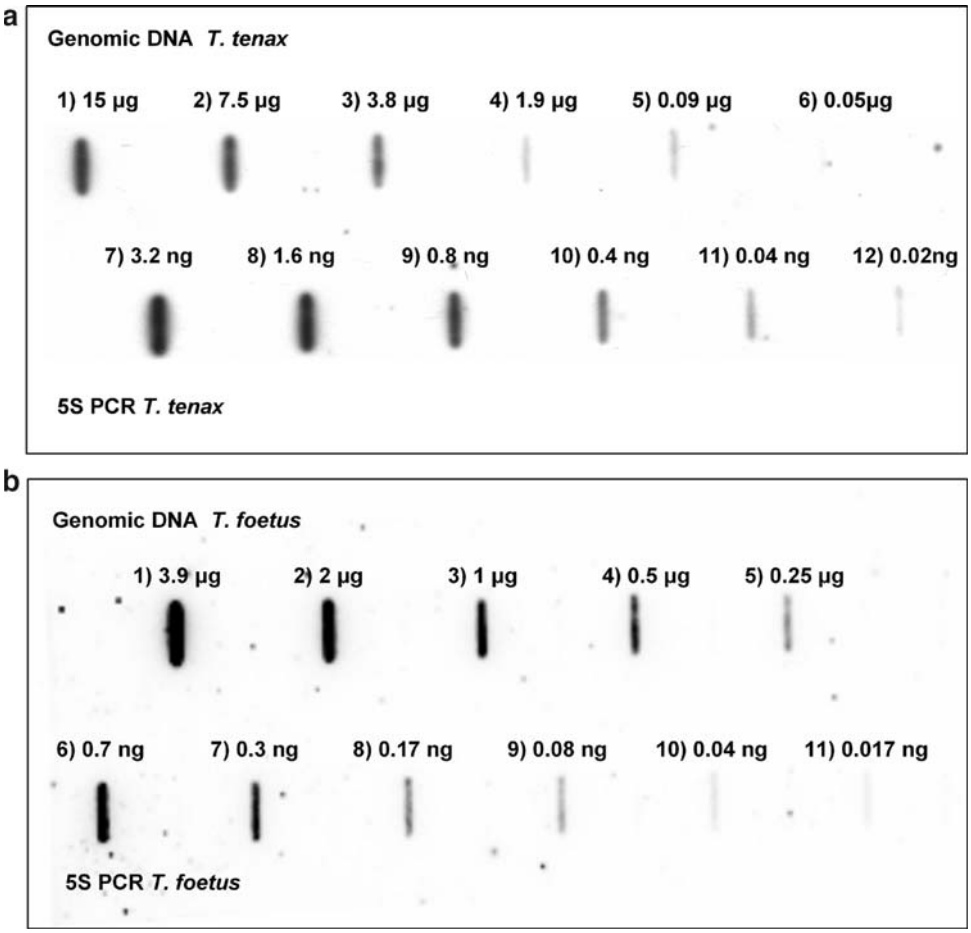

Fig. 6 5S rRNA gene copy number estimation in *T. tenax* and *T. foetus*. A comparative quantitative hybridization was performed to estimate the 5S rRNA gene copy number in *T. tenax* (**a**) and *T. foetus* (**b**). The amount of DNA loaded in each well is indicated. The probes correspond to the complete 5S rRNA-coding region of each organism, labeled with ^{32}P-dCTP by the random primer technique

chromosomes have been defined as three big, two intermediate sized, and one mini-chromosome (Lehker and Alderete 1999). For *T. tenax*, we identified two large and four small chromosomes. On the other hand, the *T. foetus* chromosomes were easily distinguished due to their clear size variation. We were able to identify all *T. foetus* chromosomes (I–V) according to their size (Fig. 7). In *T. vaginalis*, the 5S rRNA genes are localized in two of its three large chromosomes, while the rDNA gene is localized in the mini-chromosome (Fig. 7). In *T. tenax*, the 5S rRNA genes are localized in one large and one small chromosome, while the rDNA genes are localized in one of the large chromosomes (Fig. 7). In contrast, the *T. foetus* 5S rRNA genes are found in a single chromosome (II), whereas the rDNA genes are localized in the smallest chromosome (V) (Fig. 7).

Discussion

The high sequence identity in the 5S rRNA gene-coding regions of the three Trichomonad species affirms a close phylogenetic relationship among them (Felleisen 1997). Likewise, our data show a conservation of the pol III promoter. Interestingly, we found two 5S rRNA-coding sequence variants in *T. tenax*, which include variation of the promoter region. This variation within the promoter region may result in putative differential transcriptional regulation as seen in the *Plasmodium* rDNA genes (Waters 1994). On the other hand, the variation in the coding sequence may result in the synthesis of two types of 5S rRNA mature molecules worthy of future experimental analyses to understand the functional implications, if any.

The putative 5S rRNA upstream regulatory region, which includes the start site element, the upstream promoter element, and sometimes a TATA-like box or an AT-rich region is relatively well conserved among the three Trichomonad species. These kinds of elements have been confirmed as important transcriptional regulatory motifs in yeast (Lee et al. 1997). Interestingly, the 5S rRNA gene variants in the *T. vaginalis* G3 isolate have a differential TATA-like motif in the 5S rRNA upstream regulatory region whose potential role in transcription is a working hypothesis in our laboratory.

Since a very high sequence identity was found in the 5S rRNA gene-coding region among the three species, we also expected a relatively high similarity in the sequence and motifs of their intergenic regions. However, we found interesting differences in the intergenic regions of the Trichomo-

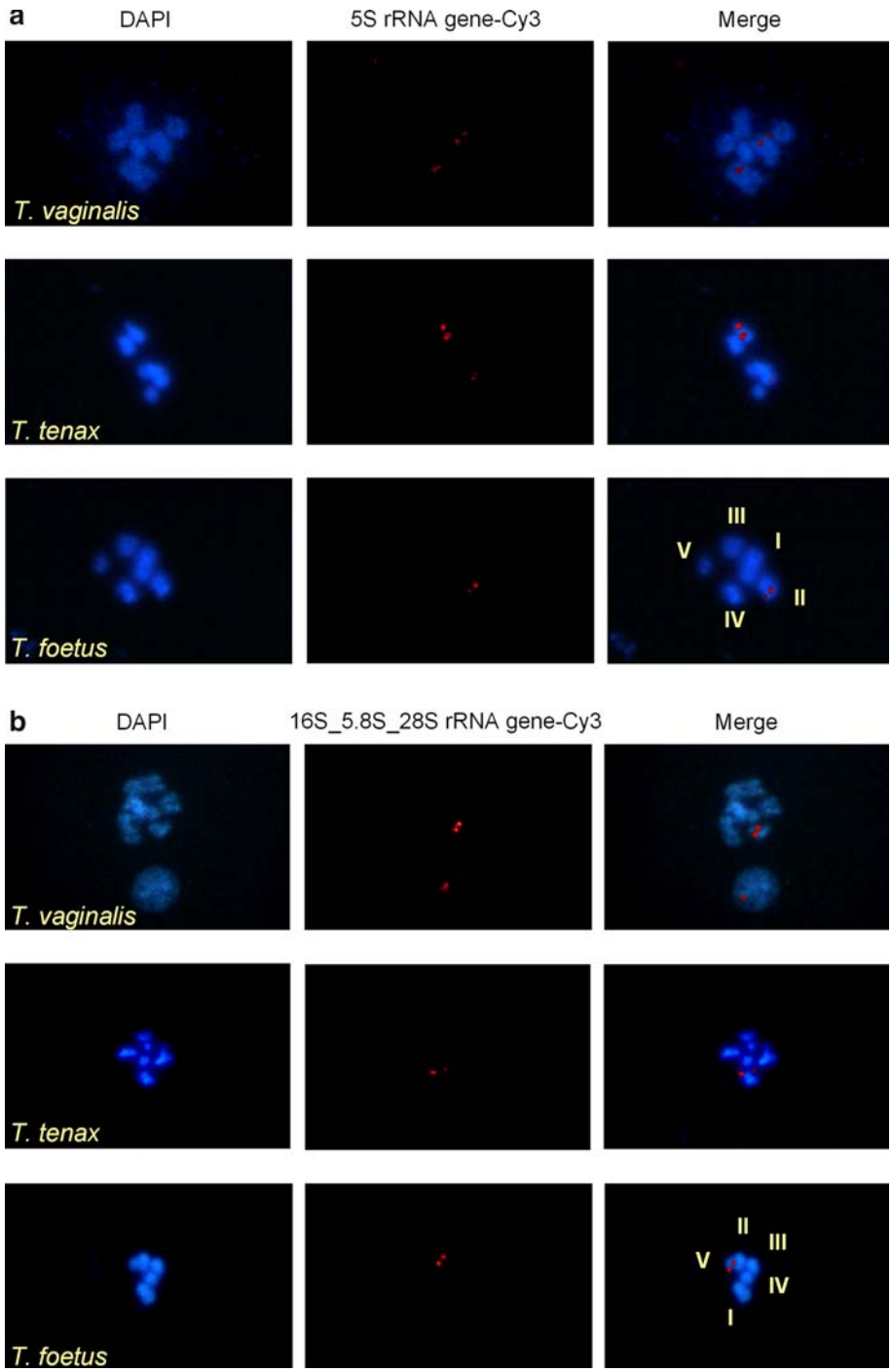

Fig. 7 Chromosomal localization of ribosomal RNA genes in Trichomonads. The 5S rRNA genes (*panel* **a**) and the rDNA genes (16S_5.8S_28S rRNA) (*panel* **b**) were localized in Trichomonads metaphases by fluorescent in situ hybridization (FISH). The numbers of the *T. foetus* chromosomes are indicated. The rRNA genes probes were labeled with Cy3 (*red*), and chromosomes were visualized with DAPI (*blue*)

nads that may reflect the fast evolutionary rate of intergenic sequences in ribosomal genes. The intergenic region sequence and motifs variation in the rRNA genes are found throughout evolution and have been studied more in the rRNA genes (rDNA). The possible involvement of the ribosomal gene sequence variants in transcriptional and epigenetic regulation of the different rDNA gene copies has not been studied in detail in most organisms. *P. falciparum* (and other Apicomplexa) are the only organisms in which the rDNA gene variants have been confirmed to be involved in a differential transcriptional regulation of the rRNA genes (Waters 1994).

Each of the Trichomonad organisms studied has characteristic species-specific 5S rRNA intergenic region motifs, such as a sequence that could structure a high-energy hairpin in *T. vaginalis*, a Type B-Ttx 5S rRNA-specific high-energy hairpin in *T. tenax* and an ubiquitin ORF and repeated sequences in *T. foetus*. It is reasonable to hypothe-

size that the role of these motifs is for the regulation of 5S rRNA expression. The ubiquitin gene linkage to the 5S rRNA gene has also been observed in an unrelated species, *Tetrahymena pyriformis* (Guerreiro et al. 1993), making interesting the study of the functional relevance of this type of multigenic families' genomic linkage.

It is interesting to note that the organization of the *T. foetus* 5S rRNA gene-repeated sequences is similar to the organization of the rDNA main transcription unit in metazoans. In the rDNA, the "spacer promoter" (a sequence similar to the ribosomal promoter, which is repeated within the intergenic region) has been proposed to have a role in recruiting the pol I transcription machinery as a mechanism to enhance the rDNA transcriptional efficiency (Paule and White 2000). Whether the repeated sequences have a role in regulating the 5S rRNA gene expression requires experimental verification in the future.

We have shown that the 5S rRNA gene copies are located in at least two unlinked clusters of two different chromosomes in *T. tenax* and *T. vaginalis*. In contrast, the 5S rRNA genes in *T. foetus* seem to be located in a single cluster in one chromosome only. The fact that *T. foetus* has one 5S rRNA cluster and one chromosome less than *T. tenax* and *T. vaginalis* makes the analysis of karyotypes in these organisms an interesting topic of research. Whether the genes from the sixth chromosome of *T. tenax* and *T. vaginalis* were reorganized, compacted or lost in the *T. foetus* genome, are questions with interesting implications in evolution.

The rDNA genes are localized in a single chromosome in the three Trichomonad species, and this may be related to the small rDNA copy number in these organisms (Chakrabarti et al. 1992). In *T. vaginalis* and *T. foetus*, the rDNA is found in the smallest chromosome, meaning that the rDNA is not linked to the 5S rRNA gene in these organisms. In contrast, in *T. tenax* the 5S rRNA is located in one of the large chromosomes. Therefore, even though *T. tenax* and *T. vaginalis* are two closely related species, their genomic organization does not seem to be completely conserved. Few karyotyping analyses have been done in Trichomonads (Lehker and Alderete 1999; Zubacova et al. 2008) and no chromosomal markers have been assigned to date. As we have shown, the 5S rRNA genes and the rDNA genes have a well-defined chromosomal localization in Trichomonads. We, therefore, propose these genes as good chromosomal markers for these organisms.

Acknowledgments This work was supported by grants P45037-Q and CB-2007-079983 from Consejo Nacional de Ciencia y Tecnología (CONACYT) México, and IN214006 from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Universidad Nacional Autónoma de México (UNAM). A scholarship from CONACYT to A.L. Torres-Machorro is acknowledged. The authors thank Karla López Pacheco for cloning the *T. tenax* and *T. foetus* rDNA fragments, and Ashwini Kucknoor for help with the *T. tenax* cultures. Useful information for the FISH assay was kindly shared by M.E. Gonsebatt's and F. Recillas-Targa's laboratories (UNAM).

Conflict of interest statement The authors declare that they have no conflict of interest.

References

Alvarez-Sanchez ME, Avila-Gonzalez L, Becerril-Garcia C, Fattel-Facenda LV, Ortega-Lopez J, Arroyo R (2000) A novel cysteine proteinase (CP65) of *Trichomonas vaginalis* involved in cytotoxicity. Microb Pathog 28:193–202

Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de PY, Miranda-Saavedra D, Barton GJ, Westrop GD, Muller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315:207–212

Chakrabarti D, Dame JB, Gutell RR, Yowell CA (1992) Characterization of the rDNA unit and sequence analysis of the small subunit rRNA and 5.8S rRNA genes from *Tritrichomonas foetus*. Mol Biochem Parasitol 52:75–83

Clark CG, Ali IK, Zaki M, Loftus BJ, Hall N (2006) Unique organisation of tRNA genes in *Entamoeba histolytica*. Mol Biochem Parasitol 146:24–29

Cockburn AF, Taylor WC, Firtel RA (1978) *Dictyostelium* rDNA consists of non-chromosomal palindromic dimers containing 5S and 36S coding regions. Chromosoma 70:19–29

Dacks JB, Walker G, Field MC (2008) Implications of the new eukaryotic systematics for parasitologists. Parasitol Int 57:97–104

Espinosa N, Hernandez R, Lopez-Griego L, Lopez-Villasenor I (2002) Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. Gene 289:81–86

Felleisen RS (1997) Comparative sequence analysis of 5.8S rRNA genes and internal transcribed spacer (ITS) regions of trichomonadid protozoa. Parasitology 115:111–119

Guerreiro P, Neves A, Rodrigues-Pousada C (1993) Clusters of 5S rRNAs in the intergenic region of ubiquitin genes in *Tetrahymena pyriformis*. Biochim Biophys Acta 1216:137–139

Hallenberg C, Nederby NJ, Frederiksen S (1994) Characterization of 5S rRNA genes from mouse. Gene 142:291–295

Lee Y, Wong WM, Guyer D, Erkine AM, Nazar RN (1997) In vivo analyses of upstream promoter sequence elements in the 5 S rRNA gene from *Saccharomyces cerevisiae*. J Mol Biol 269:676–683

Lehker MW, Alderete JF (1999) Resolution of six chromosomes of *Trichomonas vaginalis* and conservation of size and number among isolates. J Parasitol 85:976–979

Liston DR, Johnson PJ (1999) Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. Mol Cell Biol 19:2380–2388

Long EO, Dawid IB (1980) Repeated genes in eukaryotes. Annu Rev Biochem 49:727–764

Paule MR, White RJ (2000) Survey and summary: transcription by RNA polymerases I and III. Nucleic Acids Res 28:1283–1298

Pelliccia F, Barzotti R, Bucciarelli E, Rocchi A (2001) 5S ribosomal and U1 small nuclear RNA genes: a new linkage type in the genome of a crustacean that has three different tandemly repeated units containing 5S ribosomal DNA sequences. Genome 44:331–335

Roberson AE, Wolffe AP, Hauser LJ, Olins DE (1989) The 5S RNA gene minichromosome of *Euplotes*. Nucleic Acids Res 17:4699–4712

Roditi I (1992) *Trypanosoma vivax*: linkage of the mini-exon (spliced leader) and 5S ribosomal RNA genes. Nucleic Acids Res 20:1995

Schramm L, Hernandez N (2002) Recruitment of RNA polymerase III to its target promoters. Genes Dev 16:2593–2620

Selker EU, Yanofsky C, Driftmier K, Metzenberg RL, Alzner-DeWeerd B, RajBhandary UL (1981) Dispersed 5S RNA genes in *N. crassa*: structure, expression and evolution. Cell 24:819–828

Shippen-Lentz DE, Vezza AC (1988) The three 5S rRNA genes from the human malaria parasite *Plasmodium falciparum* are linked. Mol Biochem Parasitol 27:263–273

Sogin ML (1991) Early evolution and the origin of eukaryotes. Curr Opin Genet Dev 1:457–463

Taghi-Kilani R, Remacha-Moreno M, Wenman WM (1994) Three tandemly repeated 5S ribosomal RNA-encoding genes identified, cloned and characterized from *Cryptosporidium parvum*. Gene 142:253–258

Torres-Machorro AL, Hernandez R, Sanchez J, Lopez-Villasenor I (2006) The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis*. Mol Biochem Parasitol 145:269–273

Wang AL, Wang CC (1985) Isolation and characterization of DNA from *Tritrichomonas foetus* and *Trichomonas vaginalis*. Mol Biochem Parasitol 14:323–335

Waters AP (1994) The ribosomal RNA genes of *Plasmodium*. Adv Parasitol 34:33–79

Xu WD, Lun ZR, Gajadhar A (1998) Chromosome numbers of *Tritrichomonas foetus* and *Tritrichomonas suis*. Vet Parasitol 78:247–251

Zubacova Z, Cimburek Z, Tachezy J (2008) Comparative analysis of trichomonad genome sizes and karyotypes. Mol Biochem Parasitol 161:49–54

Zuker M (2000) Calculating nucleic acid secondary structure. Curr Opin Struct Biol 10:303–310