

Ribosomal RNA genes in eukaryotic microorganisms: witnesses of phylogeny?

Ana Lilia Torres-Machorro, Roberto Hernández, Ana María Cevallos & Imelda López-Villaseñor

Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México D.F., Mexico

Correspondence: Imelda López-Villaseñor, Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Apartado postal 70-228, 04510 México D.F., Mexico. Tel.: +52 55 5622 8952; fax: +52 55 5622 9212; e-mail: imelda@biomedicas.unam.mx

Received 4 September 2009; revised 25 October 2009; accepted 27 October 2009. Final version published online 23 November 2009.

DOI:10.1111/j.1574-6976.2009.00196.x

Editor: Colin Berry

Keywords

rRNA; 5S rRNA; unicellular eukaryote; ribosomal cistron organization; extrachromosomal gene; repeated sequence.

Introduction

The most recent phylogenetic model for relationships among eukaryotes clusters them into six supergroups, probably monophyletic (Simpson & Roger, 2004; Adl *et al.*, 2005). Microbial eukaryotes are found in all six groups and have considerable morphological, ultrastructural and genetic diversity. Several unique features have been described in these organisms, such as trans-splicing and RNA editing in trypanosomatids (Madison-Antenucci *et al.*, 2002; Haile & Papadopolou, 2007) as well as DNA splicing and rearrangements in the ciliate *Tetrahymena* (Prescott, 2000). *Microsporidia* (*Encephalitozoon cuniculi*) possess genomes in the size range of bacteria (Keeling & Slamovits, 2004), while the genomes of dinoflagellates lack histones and nucleosomes (Moreno Díaz de la Espina *et al.*, 2005). Cryptomonad and chlorarachniophyte unicellular algae conserve a relict miniaturized nucleus of a formerly independent alga (nucleomorph) (Cavalier-Smith, 2002) and specialized infection organelles (rhoptries and micronemes) are present in apicomplexans such as

Abstract

The study of genomic organization and regulatory elements of rRNA genes in metazoan paradigmatic organisms has led to the most accepted model of rRNA gene organization in eukaryotes. Nevertheless, the rRNA genes of microbial eukaryotes have also been studied in considerable detail and their atypical structures have been considered as exceptions. However, it is likely that these organisms have preserved variations in the organization of a versatile gene that may be seen as living records of evolution. Here, we review the organization of the main rRNA transcription unit (rDNA) and the 5S rRNA genes (5S rDNA). These genes are reiterated in the genome of microbial eukaryotes and may be coded alone, in tandem repeats, linked to each other or linked to other genes. They may be found in the chromosome or extrachromosomally in linear or circular units. rDNA coding regions may contain introns, sequence insertions, protein-coding genes or additional spacers. The 5S rDNA can be found in tandem repeats or genetically linked to genes transcribed by RNA polymerases I, II or III. Available information from about a hundred microbial eukaryotes was used to review the unexpected diversity in the genomic organization of rRNA genes.

Plasmodium (Kats *et al.*, 2006) and *Toxoplasma* parasites (Boothroyd & Dubremetz, 2008). Unusual characteristics extend to the organization of rRNA genes, which evidence the peculiarities, diversity and divergence of the genome structure in microbial eukaryotes. An overview of the biology of key microbial eukaryotes is given in Box 1.

The typical eukaryotic translation machinery, the ribosome, is composed of two subunits with four rRNA species and > 70 proteins. The large subunit (LSU) contains the 28S, the 5.8S and the 5S rRNAs. The small subunit (SSU) contains the 18S rRNA (SSU rRNA). The four rRNA mature molecules are coded in two rRNA genes transcribed by two different RNA polymerases. The 18S, 5.8S and 28S rRNAs are coded in a single transcription unit called a ribosomal cistron or the main transcription unit, transcribed by RNA polymerase I (pol I). The 5S rRNA gene is not usually linked to the ribosomal cistron and is transcribed by pol III (Mandal, 1984; Paule & White, 2000).

rRNA genes were among the first genes to be studied in detail due to their highly repetitive nature, ease of manipulation and biological importance (Miller & Beatty, 1969;

Box 1. Characteristics of some microbial eukaryotes

We used the most recent and accepted classification of eukaryotes, based on multiple gene molecular phylogenies and structural analyses. This system divides eukaryotes into six supergroups: Amoebozoa, Opisthokonta, Rhizaria, Plantae, Chromalveolata and Excavata (Simpson & Roger, 2004; Adl et al., 2005; Dacks et al., 2008). Here, we describe key microorganisms of each supergroup (Margulis & Schwartz, 2000).

1. AMOEBOZOA: Organisms that show amoeboid locomotion with pseudopodia.

Pelomyxa palustris: Giant anaerobic amoeba that contains three types of bacterial endosymbionts that replace the functions of some lacking organelles such as the mitochondria.

Acanthamoeba castellanii (Acanthamoebidae): Freelifving soil amoeba.

Entamoeba histolytica (Entamoebida): Uninucleate amitochondriate amoeba that infects the intestine of animals, causing amoebiasis.

Physarum polycephalum (Eumycetozoa, Myxogastria): Amoeboid cells that can differentiate into fungus-like reproductive structures. During its life cycle, a diploid zygote divides repeatedly to form a multinucleated cytoplasmic mass called the plasmodium. Under dry conditions, the plasmodium may mature into spore-producing organs.

Didymium iridis (Eumycetozoa, Myxogastria): Plasmodial slime mold.

Dictyostelium discoideum (Eumycetozoa, Dictyostelia): Land-dwelling cellular slime mold. Independent amoebas may aggregate into a slimy mass (slug) that eventually transforms into a reproductive body that produces spores.

2. OPISTHOKONTA: Organisms with a single posterior flagellum in at least one stage of the life cycle.

Fungi: Dominant osmotrophs that play crucial roles as decomposers and as symbionts or parasites.

Ascomycota: Hold a microscopic reproductive structure called ascus.

Pneumocystis carinii (Ascomycota, Taphrinomycotina): Causes fatal pneumonia in immunocompromised humans.

Schizosaccharomyces pombe (Ascomycota, Taphrinomycotina, Schizosaccharomycetes): Fission yeast.

Saccharomyces cerevisiae (Ascomycota, Saccharomycetes): Budding yeast that ferments sugars to ethyl alcohol.

Candida albicans (Ascomycota, Saccharomycetes): Causes infections in humans.

Microsporidia: Intracellular asexual parasites that lack mitochondria. The microsporean resting stages are the chitinous spores, which contain a polar filament and an infective body.

Encephalitozoon cuniculi: Parasites of warm-blooded vertebrates, holds one of the smallest known eukaryotic genomes (2.9 Mbp) (Biderre et al., 1997).

Nosema bombycis: causes disease in insects.

3. RHIZARIA: Organisms with pseudopodia of various types.

Foraminifera: Planktonic or benthic free-swimming organisms that have pore-studded shells. They show nuclear dimorphism and complex life cycles.

4. PLANTAE (ARCHAEPLASTIDA): Organisms that hold a photosynthetic plastid derived from a primary endosymbiosis with a cyanobacterium.

Rhodophyceae (Red algae): Mostly marine organisms that hold rhodoplasts (red plastids).

Cyanidoschyzon merolae: Unicellular organisms that inhabit sulfate-rich hot springs.

Chloroplastida: Organisms that hold green chloroplasts.

Acetabularia mediterranea (Chlorophyta, Ulvophyceae): Syncytial green algae.

Chlorella (Chlorophyta, Trebouxiophyceae).

Chlamydomonas (Chlorophyta, Chlorophyceae).

5. CHROMALVEOLATA: Organisms that contain a plastid that comes from a secondary endosymbiosis with an ancestral archaeplastid.

Bacillariophyta (Stramenopiles): Single cells or colonies covered by an elaborate, symmetrical two-part shell.

Alveolata:

Dinoflagellata (Dinozoa): Mostly unicellular marine plankton, holding two undulipodia and complex rigid walls (tests). Some species produce toxins.

Pfiesteria piscicida (Dinophyceae, Peridiniophyceae).

Ciliophora: Unicellular organisms covered with cilia (short undulipodia). They have two types of nuclei: small genetic micronuclei (MIC, containing standard chromosomes) and large transcriptionally active macronuclei (MAC, it develops from the micronuclei).

Euplotes (Intramacronucleata, Spirotrichea, Hypotrichia).

Paramecium (Intramacronucleata, Oligohymenophorea, Peniculia).

Tetrahymena (Intramacronucleata, Oligohymenophorea, Hymenostomatia).

Apicomplexa: Specialized obligate intracellular parasites named for the 'apical complex' that hold structures such as rhoptries and micronemes, the specialized machinery used for invasion (Kats et al., 2006; Boothroyd & Dubremetz, 2008).

Plasmodium (Aconoidasida, Haemosporida): The causative agent of malaria exists in association with an invertebrate host (sexual stage in the mosquito) and a vertebrate host (asexual stage). *Plasmodium falciparum* and *Plasmodium vivax* infect human red blood cells, while *Plasmodium berghei* infects rodents.

Babesia bovis (Aconoidasida, Piroplasmorida).

Theileria parva (Aconoidasida, Piroplasmorida).

Cryptosporidium (Conoidasida, Coccidiasina).

Eimeria (Conoidasida, Coccidiasina).

6. EXCAVATA: Organisms that typically have a suspension-feeding groove and flagella.

Giardia intestinalis (Fornicata, Eopharyngia, Diplomonadida, Giardiae): A parasite of the small intestine of vertebrates through infective cysts. It has two transcriptionally active karyomastigonts (nuclei attached to undulipodia by thin fibers), and lacks mitochondria and the Golgi apparatus.

Trichomonas vaginalis (Parabasalia, Trichomonadida): Amitochondriate parasite causative of trichomoniasis in humans. The organelles known as parabasal bodies are involved in the synthesis, storage and transport of proteins.

Trichomonas tenax (Parabasalia, Trichomonadida): Infects the human mouth.

Tritrichomonas foetus (Parabasalia, Trichomonadida): Infects the urogenital tract of cattle.

Naegleria gruberi: (Heterolobosea, Vahlkampfiidae) Soil and freshwater freelifving amoeba that transforms into undulipodiated cells.

Euglena gracilis (Euglenozoa, Euglenida, Euglenea): Unicellular organism living in stagnant water. It can be found with or without chloroplasts.

Kinetoplastea (Euglenozoa): Contain a large mitochondrion called a kinetoplast.

Trypanosoma (Metakinetoplastina, Trypanosomatida): The change of host and some differentiation steps are associated with characteristic movements of the kinetoplast along the cell. *Trypanosoma brucei* infection (transmitted to humans through the bite of infected tsetse flies) causes the sleeping sickness, while *Trypanosoma cruzi* infection (transmitted through the bite of infected reduviid bugs) leads to Chagas disease in humans.

Leishmania (Metakinetoplastina, Trypanosomatida): Parasite responsible for the leishmaniasis disease. It multiplies within the lysosomes of vertebrate macrophages and within the digestive system of sand-flies.

Bodo saltans (Metakinetoplastina, Eubodonida): Freelifving bi-undulipodiated cell.

Crithidia (Metakinetoplastina, Trypanosomatida).

Trypanoplasma (Metakinetoplastina, Parabodonida).

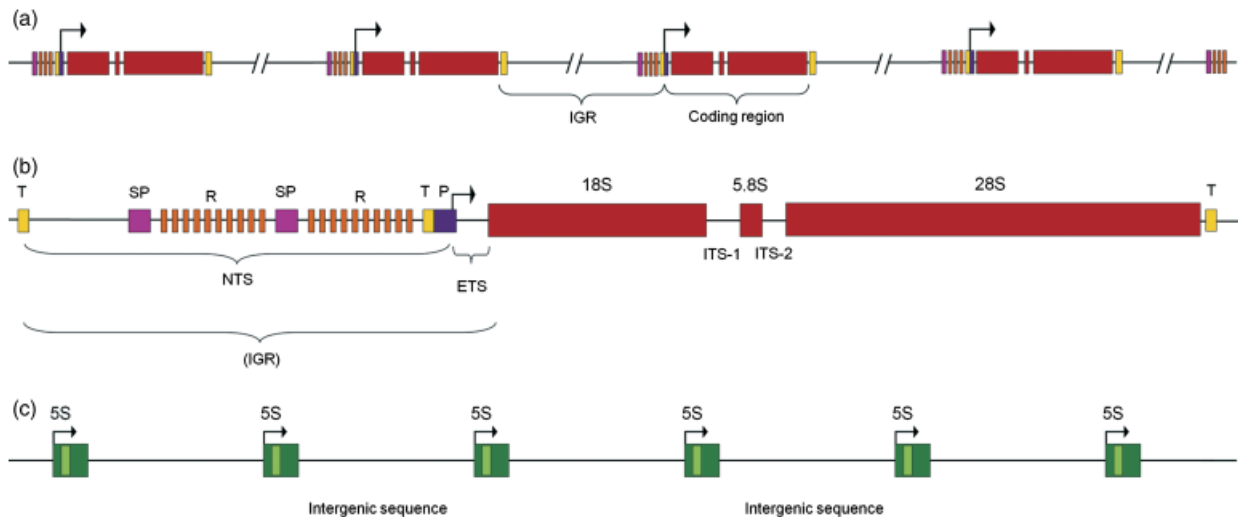


Fig. 1. General organization of the ribosomal main transcription unit (rDNA) and 5S rDNA. (a) Schematic representation of *Xenopus laevis* rDNA. About 600 U of the ribosomal cistron are encoded in the chromosome in head-to-tail tandem repeats. Each unit contains a coding region (red) and an IGR. (b) A single unit of the *X. laevis* rDNA. The 18S, 5.8S and 28S rRNA molecules are transcribed as a single RNA precursor that is post-transcriptionally processed to produce the mature rRNA molecules. Transcription regulatory elements for RNA polymerase I are found in the NTS: tandem-repeated sequences (R), spacer promoters (SP), transcription terminators (T) and the promoter (P). The IGR comprises both the NTS and the ETS. (c) Organization of somatic 5S rDNA in *X. laevis*. The 5S rDNA is organized in tandem head-to-tail repeats that include a coding region (green box) and an intergenic sequence (black line). The 5S rDNA promoter is internal to the coding region (light green box). Arrows represent the transcription start point. ETS, external transcribed spacer.

Long & Dawid, 1980; Sollner-Webb & Mougey, 1991). The thorough study and description of genomic organization and regulatory elements in the rRNA genes of *Xenopus*, *Drosophila* and mouse led to the most accepted model of rRNA gene organization in eukaryotes (Long & Dawid, 1980; Mandal, 1984; Sollner-Webb & Mougey, 1991) (Fig. 1). The rRNA genes of microbial eukaryotes have also been intensively studied, although they were considered to be the exception to the rule, as their organization differs from the general models (Long & Dawid, 1980; Mandal, 1984). Here, we focus on the rRNA gene organization of microbial eukaryotes where many examples of gene diversity can be found. This work also summarizes the variability of motifs present in the rDNA intergenic region (IGR), which may include general and species-specific elements. For simplicity, in this review, the ribosomal cistron is referred to as rDNA and the term 5S rDNA is used for the 5S rRNA gene.

Overview of the eukaryotic rRNA genes

The ribosomal cistron (rDNA)

In most species, the rDNA is present in multiple copies organized as tandem head-to-tail repeats. The rDNA unit is composed of a transcribed region and an IGR (also called the intergenic spacer) consisting of a nontranscribed spacer (NTS) 2–30 kbp long and an external transcribed spacer. The NTS contains most of the regulatory elements for

transcription, while the external transcribed spacer is part of the primary transcript (pre-rRNA, 7–14 kb long) (Sollner-Webb & Mougey, 1991) (Fig. 1).

Several regulatory elements may be found in the IGR such as enhancers, spacer promoters, a proximal terminator and the gene promoter. This region may also contain several repetitive sequences that may improve the transcription efficiency, with additive effects (Paule & White, 2000). A schematic representation of the *Xenopus laevis* rDNA is shown in Fig. 1a and b as an example of the ‘typical’ eukaryotic rDNA organization (Sollner-Webb & Mougey, 1991). The rDNA pol I core promoter and other nonrepeated rDNA regulatory elements have been described and studied in detail in some unicellular eukaryotes such as *Trypanosoma cruzi*, *Acanthamoeba castellanii* and yeast (Kownin *et al.*, 1985; Neigeborn & Warner, 1990; Wai *et al.*, 2000; Figueroa-Angulo *et al.*, 2006).

Transcription of the rDNA proceeds from the promoter through the 5′ external transcribed spacer – 18S rRNA – internal transcribed spacer-1 (ITS-1) – 5.8S rRNA – ITS-2 and 28S rRNA, until pol I comes across a transcription termination signal (Long & Dawid, 1980). In most cases, the rDNA primary transcript is post-transcriptionally processed in three rRNA mature molecules: 18S, 5.8S and 28S, resulting from the elimination of the external transcribed spacers ITS-1 and ITS-2 (Fig. 1) (Long & Dawid, 1980). Additional processing of the rRNAs into several smaller molecules has also been described. As most eukaryotic LSU rRNAs (eLSU

rRNAs) are fragmented by removal of ITS-2, the eLSU rRNA should be defined as 5.8S+28S rRNA. Because the term LSU rRNA has been used as equivalent to bacterial 23S rRNA, here we refer to the 5.8S+28S rRNA as eLSU rRNA.

The 5S rRNA gene (5S rDNA)

The 5S rDNA is reiterated in the eukaryotic genome in tandem head-to-tail arrays (Paule & White, 2000) (Fig. 1c). The 5S rDNA promoter (Internal Control Region) is found downstream of the transcription start point and within the transcribed region. Upstream regulatory elements can also be found in some 5S rDNAs that may be necessary for transcription. The Internal Control Region is sufficient for transcription of 5S rRNA in *Xenopus* (Bogenghagen *et al.*, 1980; Sakonju *et al.*, 1980) whereas in *Saccharomyces cerevisiae* two upstream regulatory elements (start site element and upstream promoter element) are necessary for its efficient transcription *in vivo* (Lee *et al.*, 1997).

Redundancy and relative ratio of the rRNA genes

rRNA genes are reiterated in almost every eukaryotic genome studied, and the gene copy number is maintained at a characteristic constant level for each organism. In most organisms, not all of the rDNA copies are transcribed (Conconi *et al.*, 1989), suggesting that the total rDNA copy number is not directly related to the synthesis of rRNA (Kobayashi *et al.*, 1998; Grummt, 2003; Raska *et al.*, 2004). It has been proposed that the rDNA may participate in roles other than transcription, such as maintenance of the nucleolar structure and rDNA stability (Nogi *et al.*, 1991; Oakes *et al.*, 1993).

A considerable variation in the rDNA and 5S rDNA gene copy number exists among eukaryotes (Table 1): the rDNA copy number can range from one and two copies in the *Ascomycota Pneumocystis carinii* and the *Apicomplexa Theileria parva* to 4800 copies in the green alga *Acetabularia mediterranea* and 9000 copies in the ciliate *Tetrahymena thermophila*. The 5S rDNA copies can range from three in the red alga *Cyanidioschyzon merolae* to about one million in the ciliate *Euplotes eurystomus*. In the slime mold *Dictyostelium discoideum* and in the yeast *S. cerevisiae*, the rDNA and 5S rDNA genes are present in equal numbers, although a strict relationship between both types of genes is not always observed. For example, *Euglena gracilis* has 800–4000 copies of rDNA and only 300 copies of the 5S rDNA; in contrast, 110 copies of the rDNA are present in the genome of the kinetoplastid *T. cruzi*, while the 5S rDNA is repeated 1600 times. The copy number of rRNA genes in some microbial eukaryotes and the rDNA/5S rDNA ratio are given in Table 1. The considerable variability in this ratio suggests that different species may have particular regulatory mechanisms

to maintain the 18S, 5.8S, 28S and 5S rRNA homeostasis for the efficient synthesis of ribosomes.

The ribosomal cistron: localization, gene linkage and IGR

The typical rDNA organization

Chromosomally localized tandem head-to-tail repeats of rDNA units containing a coding region and an IGR represent the typical rDNA organization in eukaryotes. Some microbial organisms of various phylogenetic branches share this general organization, as shown in Fig. 2 and Table 2. Tandem rDNA units may be located in a single chromosome and locus (e.g. *Kluyveromyces lactis* and *Trichomonas vaginalis*), or in various chromosomes and loci (e.g. *T. cruzi*). Atypical tail-to-tail and head-to-head rDNA repeats (interspersed with typical tandem head-to-tail repeats) are observed in *Acetabularia exigua* (Berger *et al.*, 1978; Spring *et al.*, 1978). In various yeast species such as *K. lactis* and *S. cerevisiae*, tandem rDNA units are genetically linked to the 5S rDNA. In these cases, the 5S rDNA can be coded either in a sense or an antisense orientation relative to the rDNA coding strand (Table 2, Fig. 2b and c).

Depending on the species (and isolate) studied, several tandemly repeated sequences may be found in the IGR with variations in size, sequence and number (Table 3). For example, the IGR of *Leishmania* species contain a 60–64 bp repeated element reiterated between 16 and 275 times, causing length variations in the IGR that range from 4 to 12 kbp. The size, number and sequence of these motifs are species- and isolate-specific elements in *Leishmania* spp. (Gay *et al.*, 1996; Uliana *et al.*, 1996; Yan *et al.*, 1999; Martínez-Calvillo *et al.*, 2001; Orlando *et al.*, 2002) (Table 3 and Fig. 2g).

Unlinked and heterogeneous rDNA

rDNAs with heterogeneous coding and intergenic sequences are characteristic of the *Apicomplexa* group and may be found unlinked (located in nonadjacent loci) and in low copy number (Table 4). For example, *Plasmodium* spp. may have four to eight rDNA copies per haploid genome. *Plasmodium falciparum* and *Plasmodium berghei* have two types of rDNA units (Waters *et al.*, 1989; Waters, 1994): A-type and S-type (also known as C-type in *P. berghei*) code for different SSU and eLSU rRNAs that correlate with the production of ribosomes with different GTPase activity (Rogers *et al.*, 1996; Velichutina *et al.*, 1998). The expression of rDNA genes is tightly linked to the progression of *Plasmodium* life cycle: the A-type rRNA is expressed predominantly in the vertebrate host (asexual development), whereas the S-type rRNA is expressed in the mosquito stage (sexual development) (Mercereau-Puijalon *et al.*, 2002) (Fig.

Table 1. rRNA genes: copy number, unit size and organization

Organism	rDNA copies	rDNA unit size (kbp) and organization	5S rDNA copies	5S rDNA unit size (kbp) and organization	rDNA/5S ratio	References
Excavata						
<i>Crithidia fasciculata</i>		11–12 (T)	250–300	0.23 (T)		Köck & Cornelissen (1990), Schnare <i>et al.</i> (2000)
<i>Diplonema papillatum</i>				0.68 and 1.7 (T)		Sturm <i>et al.</i> (2001)
<i>Euglena gracilis</i>	800–4000 (C), 4(chr)	11.5 (C)	330	0.6 (T)	2.42–12.12	Ravel-Chapuis (1988), Keller <i>et al.</i> (1992)
<i>Herpetomonas</i>				0.6 (T)		Aksoy (1992)
<i>Leishmania donovani</i>	166	12.5 (T)				Yan <i>et al.</i> (1999), León <i>et al.</i> (1978)
<i>Leishmania major</i>	63	14 (T)				Martínez-Calvillo <i>et al.</i> (2001), Ivens <i>et al.</i> (2005)
<i>Trypanosoma brucei</i>	56	(T)	1500	0.75	0.04	Hasan <i>et al.</i> (1984), Berriman <i>et al.</i> (2005)
<i>Trypanosoma cruzi</i>	110	30 (T)	1600	0.48 (T)	0.07	Castro <i>et al.</i> (1981), Hernández-Rivas <i>et al.</i> (1992), Hernández <i>et al.</i> (1993)
<i>Trypanosoma rangeli</i>			330	0.9		Aksoy <i>et al.</i> (1992)
<i>Trypanosoma vivax</i>				0.72		Roditi (1992)
<i>Naegleria gruberi</i>	3000–5000 (C)	14 (C)				Clark & Cross (1988)
<i>Giardia intestinalis</i>	60 (H), 300	5.6 (T)				Le Blancq <i>et al.</i> (1991)
<i>Trichomonas tenax</i>		(T)	0.01% *	0.307 and 0.316 (T)		Torres-Machorro <i>et al.</i> (2009)
<i>Trichomonas vaginalis</i>		6 (T)	0.1% *	0.334 and 0.335 (T)		López-Villaseñor <i>et al.</i> (2004), Torres-Machorro <i>et al.</i> (2006)
<i>Tritrichomonas foetus</i>	12	6 (T)	0.04%*	0.86 and ~1.3 (T)		Chakrabarti <i>et al.</i> (1992), Torres-Machorro <i>et al.</i> (2009)
<i>Trypanoplasma borreli</i>				0.59		Maslov <i>et al.</i> (1993)
Chromalveolates						
<i>Babesia bigemina</i>	3	10.65, 10.8 and 13.35 (U)				Reddy <i>et al.</i> (1991)
<i>Babesia bovis</i>	3	7				Dalrymple (1990)
<i>Babesia canis</i>	4					Dalrymple <i>et al.</i> (1992)
<i>Cryptosporidium parvum</i>	5 (H)	6.5 (T)	6	0.55 and 0.79, 3 (T), 3(U)	0.83	Taghi-Kilani <i>et al.</i> (1994), Le Blancq <i>et al.</i> (1997)
<i>Eimeria tenella</i>	140	(T)	500	0.73 (T)	0.28	Stucki <i>et al.</i> (1993), Shirley (2000)
<i>Plasmodium berghei</i>	4	(U)	3	(T)	1.33	Dame & McCutchan (1983), Waters (1994)
<i>Plasmodium falciparum</i>	5–8	(U)	3		1.67–2.67	Shippen-Lentz & Vezza (1988), Gardner <i>et al.</i> (2002)
<i>Plasmodium lophurae</i>	6	(U)				Unnasch & Wirth (1983)
<i>Plasmodium vivax</i>	7	(U)				Li <i>et al.</i> (1997)
<i>Theileria parva</i>	2	(U)	3		0.67	Kibe <i>et al.</i> (1994), Gardner <i>et al.</i> (2005)
<i>Toxoplasma gondii</i>	110	7.5 (T)	110		1	Guay <i>et al.</i> (1992)
<i>Perkinsus andrewsi</i>		7.7–7.8				Pecher <i>et al.</i> (2004)
<i>Euplotes crassus</i>		7 (L)				Erbezniak <i>et al.</i> (1999)
<i>Euplotes eurytomus</i>			1 000 000	0.93 (L)		Roberson <i>et al.</i> (1989)
<i>Glaucoma chattoni</i>		9.3 (L)				Challoner <i>et al.</i> (1985)
<i>Oxytricha fallax</i>		7.49 (L)		0.69 (L)		Rae & Spear (1978), Swanton <i>et al.</i> (1982)
<i>Paramecium tetraurelia</i>		9 (T,L,C)				Preer <i>et al.</i> (1999)
<i>Tetrahymena pyriformis</i>	200 MAC (H), 1 MIC		350 MAC, 350 MIC (H)	0.28	0.29	Kimmel & Gorovsky (1976), Kimmel & Gorovsky (1978)
<i>Tetrahymena thermophila</i>	9000 MAC, 1MIC	21 (L,P)	150 MAC, 150 MIC (H)	0.25–0.29	30	Yao & Gall (1977), Allen <i>et al.</i> (1984), Eisen <i>et al.</i> (2006)

Table 1. Continued.

Organism	rDNA copies	rDNA unit size (kbp) and organization	5S rDNA copies	5S rDNA unit size (kbp) and organization	rDNA/5S rDNA copy ratio	References
Amoebozoa						
<i>Acanthamoeba castellanii</i>	24(H), 600	12 (T)	480	(U)	1.25	Zwick <i>et al.</i> (1991), Yang <i>et al.</i> (1994)
<i>Dictyostelium discoideum</i>	180 (L), 1 (chr)	88 (L,P)	180	88 (L,P)	1	Cockburn <i>et al.</i> (1978), Hofmann <i>et al.</i> (1993)
<i>Didymium iridis</i>		20 (L)				Johansen <i>et al.</i> (1992)
<i>Entamoeba histolytica</i> (HM-1:IMSS)	200 (C), 0 (chr)	24.5 (C)				Huber <i>et al.</i> (1989), Bagchi <i>et al.</i> (1999)
<i>Physarum polycephalum</i>	1 × 10 ¹¹	> 60 (L,P)		0.68		Campbell <i>et al.</i> (1979)
Plantae						
<i>Acetabularia mediterranea</i>	3500–4800	(T)				Spring <i>et al.</i> (1978)
<i>Cyanidioschyzon merolae</i>	3	(U)	3	(U)	1	Maruyama <i>et al.</i> (2004), Matsuzaki <i>et al.</i> (2004)
Opisthokonta						
<i>Candida albicans</i>	100 (C), 200 (chr)	11.6–12.5 (T)				Huber & Rustchenko (2001)
<i>Candida glabrata</i>	> 115	(T)	> 230		0.5	Maleszka & Clark-Walker (1993), Bergeron & Drouin (2008)
<i>Hansenula polymorpha</i>	50–60	8 (T)	50–60		1	Ramezani-Rad <i>et al.</i> (2003)
<i>Kluyveromyces lactis</i>	60	8.6 (T)	60		1	Verbeet <i>et al.</i> (1984)
<i>Pneumocystis carinii</i>	1					Tang <i>et al.</i> (1998), Fischer <i>et al.</i> (2006)
<i>Saccharomyces cerevisiae</i>	100–200	9.1 (T)	100–200		1	Rubin & Sulston (1973), Rustchenko & Sherman (1994)
<i>Schizosaccharomyces pombe</i>	100–120	10.4 (T)	30	(U)	3.33–4	Wood <i>et al.</i> (2002)
<i>Yarrowia lipolytica</i>	100	7.7 and 8.7 (T)	108	(U)	0.93	van Heerikhuizen <i>et al.</i> (1985), Acker <i>et al.</i> (2008)
<i>Encephalitozoon cuniculi</i>	22	8.9 (tel)	3	(U)	7.33	Peyretilade <i>et al.</i> (1998), Katinka <i>et al.</i> (2001)
<i>Nosema apis</i>		18 (T)				Gatehouse & Malone (1998)
<i>Nosema bombycis</i>		4.3 (T)				Huang <i>et al.</i> (2004)

All copy numbers are approximate, mainly based on quantitative hybridization analyses. If the rDNA is chromosomal, the unit size corresponds to the complete unit containing the rDNA coding region and the intergenic spacer.

If the rDNA unit is extrachromosomal, the unit size corresponds to the whole molecule size.

*In trichomonads the percentage of the genome sequence that corresponds to the 5S rRNA coding region is indicated.

C, extrachromosomal circle; chr, chromosomal; H, haploid; L, extrachromosomal linear molecule; MAC, macronucleus; MIC, micronucleus; P, palindrome; T, tandem; tel, telomeric; U, unlinked and nontandem.

3). During the transfer of *Plasmodium* from the vertebrate host to the mosquito, drastic changes in glucose concentration and temperature are involved in regulating the expression of A- and S-type rDNA genes from different promoter elements (Mack *et al.*, 1979; Fang & McCutchan, 2002; Fang *et al.*, 2004). A third type, the O-type rDNA (oocyst), has been described in the human malaria parasite *Plasmodium vivax*, whose synthesis takes place in ookinetes inside the mosquito's gut (Li *et al.*, 1997). Comprehensive reviews of rRNA genes from *Plasmodium* describe in detail the characteristic organization and function of these genes (Waters, 1994; McCutchan *et al.*, 1995). Other *Apicomplexa* species

with a similar rDNA organization are described in Table 4, as well as the non-*Apicomplexa* red alga *C. merolae*. This organism has only three unlinked rDNA units in two different chromosomes, with similar rRNA coding sequences (Matsuzaki *et al.*, 2004).

Telomeric rDNA

The microsporidian obligate intracellular parasite *E. cuniculi* has 22 rDNA units located as single copies in all telomeres of its 11 chromosomes (Brugère *et al.*, 2000). *Candida albicans* rDNA is found in two subtelomeric loci (Dujon *et al.*, 2004),

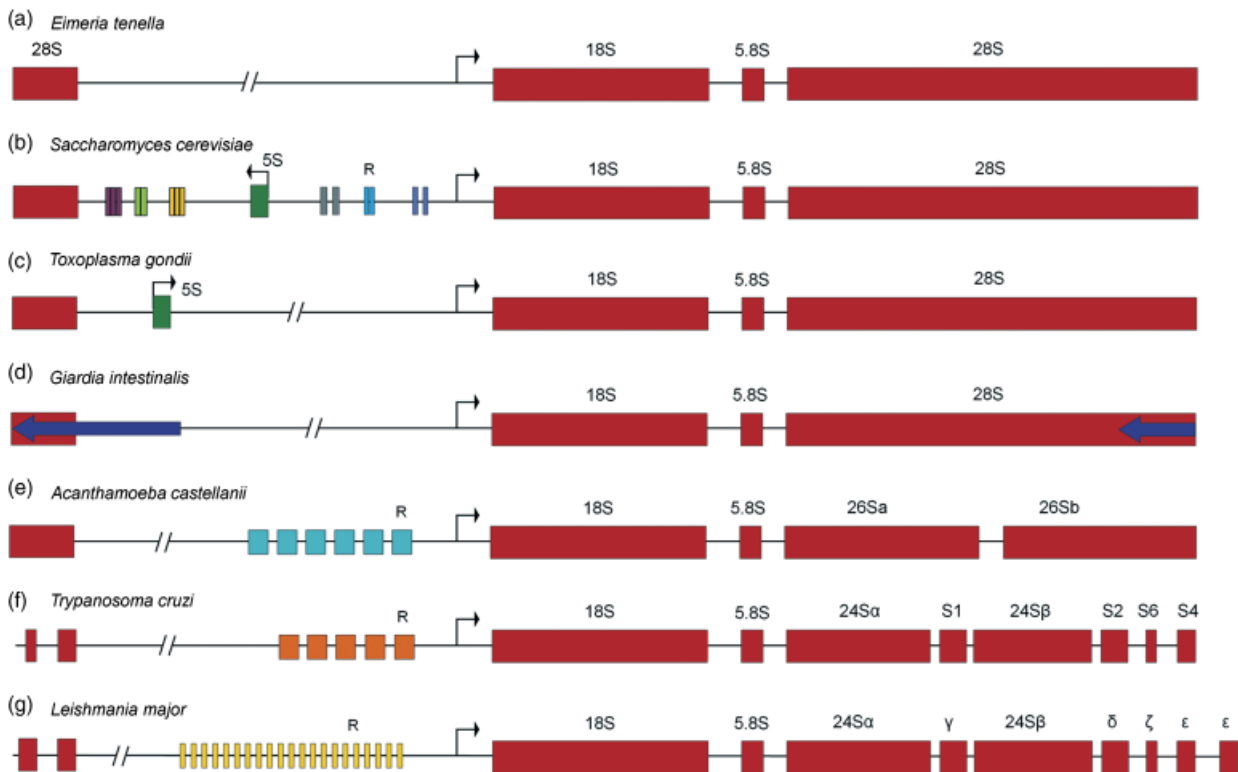


Fig. 2. Different organization of tandem head-to-tail rDNA repeats in microbial eukaryotes. (a) *Eimeria tenella* (Apicomplexa) exemplifies microbial eukaryotes with the typical rDNA organization. (b, c) rDNA copies in *Saccharomyces cerevisiae* (Ascomycota) and *Toxoplasma gondii* (Apicomplexa) are linked to the 5S rDNA (green), but in opposite polarities. Intergenic short direct repeats present in *S. cerevisiae* are shown as colored bars (see also Table 3). (d) In *Giardia intestinalis* (Diplomonadida), a 32-kDa antigenic protein (dark blue arrow) is coded in the complementary rDNA strand. (e) In *Acanthamoeba castellanii* (Acanthamoebidae), the mature eLSU rRNA is fragmented into three molecules: 5.8S, 26Sa (2.4 kb) and 26Sb (2 kb); the IGR contains six repeats of a 140-bp element (R, aqua boxes). (f, g) In *Trypanosoma cruzi* and *Leishmania major* (kinetoplastids), the eLSU rRNA is fragmented into seven molecules. The *T. cruzi* IGR contains a 172-bp repeated sequence (orange boxes). In *Leishmania* spp., the IGR is characterized by the presence of multiple repeated units (yellow). *Leishmania major* Friedlins ϵ region is duplicated once. Drawings are not to scale. The size of rDNA units is shown in Table 1. Arrows show the polarity of transcription.

while *Yarrowia lipolytica* and *Giardia intestinalis* rDNA units are positioned in seven and six subtelomeric loci, respectively (Le Blancq *et al.*, 1991; Dujon *et al.*, 2004). The *G. intestinalis* chromosome I varies 5–20% in size due to subtelomeric rearrangements including variations in rDNA copy number and size (Hou *et al.*, 1995), while some subtelomeric rDNA copies are linked to transcriptional gene units, including protein-coding genes such as ankyrin. Some of these regions may also hold incomplete rDNA sequences (Upcroft *et al.*, 2005). Interestingly, fragments of the rDNA unit are found in all chromosomal ends in *D. discoideum*. These regions encode complex repeated sequences (transposable elements–rDNA junctions) that generate novel telomeric structures (Eichinger *et al.*, 2005).

The subtelomeric localization of rDNA sequences as those found in *E. cuniculi*, *D. discoideum* and *G. intestinalis* suggests a physiological role for these elements. Telomeres have an ordered structure in the nucleus and can be clustered or associated with the nuclear matrix, at least in

some stage during the life cycle (Pryde *et al.*, 1997). Telomeres are regions of great plasticity within a heterochromatic context, with dynamics that allow for the amplification and/or variation in the number of telomeric genes and repeated sequences. It is not known whether subtelomeric rDNA is involved in the maintenance of the characteristic telomeric structure or whether the rDNA exploits this particular structure to regulate its expression and to maintain the sequence and copy number (Pryde *et al.*, 1997).

rDNA may be located extrachromosomally

Extrachromosomal rDNA has been found in ciliates, cellular and plasmodial slime molds and in yeasts (Table 5). The polyploid somatic macronucleus of the ciliate *T. thermophila* contains about 9 000 copies of a palindromic self-replicating linear minichromosome, which codes for two rDNA units (Fig. 4a and Table 5). The IGR of this palindrome contains six types of repeated sequences (Fig. 4a and

Table 2. Organisms with typical rDNA organization

Organism	Localization	IGS repeated elements	5S linkage	References
Excavata				
<i>Crithidia fasciculata</i>		Yes	X	Schnare et al. (2000)
<i>Leishmania</i> spp.	1Ch, 1L	Yes	X	Uliana et al. (1996), Yan et al. (1999), Martínez-Calvillo et al. (2001), Orlando et al. (2002), de Andrade Stempliuk & Floeter-Winter (2002)
<i>Trypanosoma brucei</i>	4Ch	X	X	Hasan et al. (1984), Melville et al. (1998)
<i>Trypanosoma cruzi</i>	≥ 2Ch	Yes	X	Hernández et al. (1993)
<i>Giardia</i> spp.	6L (intestinalis)	intestinalis and muris	X	Edlind & Chakraborty (1987), Boothroyd et al. (1987), van Keulen et al. (1992), Upcroft et al. (1994)
<i>Trichomonas tenax</i>		X	X	Torres-Machorro et al. (2009)
<i>Trichomonas vaginalis</i>	1Ch, 1L	X	X	López-Villaseñor et al. (2004), Torres-Machorro et al. (2009)
<i>Tritrichomonas foetus</i>	1Ch, 1L	X	X	Torres-Machorro et al. (2009)
Chromalveolates				
<i>Eimeria tenella</i>	1Ch, 1L	X	X	Shirley (2000)
<i>Toxoplasma gondii</i>		X	Sense	Guay et al. (1992)
Amoebozoa				
<i>Acanthamoeba castellanii</i>		Yes	X	D'Alessio et al. (1981), Yang et al. (1994)
Opisthokonta				
<i>Hansenula polymorpha</i>		X	Linked	Klabunde et al. (2002)
<i>Kluyveromyces lactis</i>	1Ch, 1L	X	Antisense	Verbeet et al. (1984)
<i>Saccharomyces cerevisiae</i>	1Ch, 1L	Yes	Antisense	Rubin & Sulston (1973), Skryabin et al. (1984), Srivastava & Schlessinger (1991), Dujon et al. (2004), Kim et al. (2006)
<i>Schizosaccharomyces pombe</i>	1Ch, 2L	X	X	Schaak et al. (1982), Wood et al. (2002)
<i>Torulopsis utilis</i>		X	Antisense	Tabata (1980)
<i>Yarrowia lipolytica</i>	7L	Yes	X	van Heerikhuizen et al. (1985)
<i>Nosema apis</i>		X	Sense	Gatehouse & Malone (1998), Iiyama et al. (2004)

IGS, intergenic spacer; X, not identified or not present; Ch, chromosome; L, locus or loci.

Table 3). The rDNA organization in *Tetrahymena pyriformis* is similar to the *T. thermophila* rDNA palindrome, with variations in the intergenic repeated motifs (Tables 3 and 5).

Dictyostelium discoideum and *Physarum polycephalum* rDNA is also encoded in palindromic extrachromosomal molecules (Fig. 4 and Table 5). Both rDNA minichromosomes contain several repeated sequence elements (Table 3). In the *D. discoideum* rDNA palindrome, two 5S rDNA copies are present near the telomeric ends, in the same polarity as the rDNA unit (Fig. 4b) (Cockburn et al., 1978). Additionally, a single rDNA palindrome is located in chromosome IV (Sucgang et al., 2003). Even though *D. discoideum* has six chromosomes, a seventh 'chromosome' can be observed in some chromosomal spreads. This additional 'chromosome' corresponds to a chromosome-sized cluster of palindromic rDNA minichromosomes (Sucgang et al., 2003), which suggests a physical interaction of the extrachromosomal rDNA. This particular organization may play a role in the expression and segregation of mitotic rDNA.

Extrachromosomal linear molecules containing one rDNA unit are found in *Didymium iridis* (Fig. 4 and Table 5). Ciliates such as *Euplotes crassus* and *Glaucoma chattoni* have extrachromosomal rDNA copies in single gene-

sized linear molecules within the macronucleus with characteristic intergenic repeated elements (Tables 3 and 5). Finally, tandem rDNA genes in *Paramecium tetraurelia* can be found both in circular and in linear extrachromosomal molecules (Table 5), which can contain > 13 rDNA copies.

rDNA units coded in extrachromosomal circular plasmids may be found in Amoebozoa (*Entamoeba histolytica*) and Excavata (*E. gracilis* and *Naegleria gruberi*) (Fig. 5 and Table 5). The most-studied *E. histolytica* isolate HM-1:IMSS lacks an rDNA chromosomal copy (Bagchi et al., 1999) but possesses about 200 copies of an extrachromosomal circular molecule, with two inverted rDNA units and repeated sequences in the IGR (Table 3 and Fig. 5a). This molecule starts replication at multiple sites; the primary replication origins are located near the pol I promoters, but other replication origins found all the way through the circle are activated under stress conditions (Ghosh et al., 2003). Interestingly, a 0.7-kb RNA of unknown function is encoded in the upstream region of one rDNA unit (Bhattacharya et al., 1998) (Fig. 5a). Depending on the isolate, variations are found in the size and organization of the rDNA circular molecules: the 200:NIH *E. histolytica* isolate has a

Table 3. Repeated sequences in the ribosomal cistron intergenic spacer

Organism	Size of repeated sequences	Number of repeated sequences	Function	Intergenic region size (kb)	References
Excavata					
<i>Crithidia fasciculata</i>	19 bp	28			Schnare <i>et al.</i> (2000)
	55 bp*	4		3.5	
<i>Euglena gracilis</i>	14 bp, imperfect repeat	6			Greenwood <i>et al.</i> (2001)
	30 bp imperfect palindromes	2			
<i>Leishmania major</i>	63 bp	16–275		4–12	Martinez-Calvillo <i>et al.</i> (2001)
<i>Leishmania amazoniensis</i>	60 bp	35–70			Uliana <i>et al.</i> (1996)
<i>Leishmania chagasi</i>	64 bp	9	Enhancer-like		Gay <i>et al.</i> (1996)
<i>Leishmania infantum</i>	63 bp	40		4–12	Requena <i>et al.</i> (1997)
	LiR3 348 bp		Transcription termination?		
<i>Leishmania donovani</i>	64 bp	39		5.8	Yan <i>et al.</i> (1999)
	63 bp	12		4	
<i>Leishmania hoogstraali</i>	63 bp	40		5.5	Orlando <i>et al.</i> (2002)
<i>Trypanosoma cruzi</i>	172 bp	2–8			Pulido <i>et al.</i> (1996)
<i>Giardia intestinalis</i>	78 bp	2 in GS strain, 3 in GK strain	Recombination?		Upcroft <i>et al.</i> (1994)
<i>Giardia muris</i>	73 bp	Varies, 6 min			van Keulen <i>et al.</i> (1992)
Chromalveolates					
<i>Euplotes crassus</i>	I. ~30 bp IV. ~17 bp				Erbeznik <i>et al.</i> (1999)
<i>Glaucoma chattoni</i>	I. 32 bp (5')	3	Spacer promoter? Enhancer?		Challoner <i>et al.</i> (1985)
	II. 14–24 bp (5')	7			
	III. 18 bp (5')	1			
	IV. 17 bp (3')	9	Termination?		
	V. 130 bp (3')	3	Gene packing?		
<i>Tetrahymena thermophila</i>	I. 32 bp (5') [†]	4	Spacer promoter?	1.9	Challoner <i>et al.</i> (1985)
	II. 20–21 bp (5')	13	Enhancer?		
	III. 20 bp (5')	7			
	430 bp (5')	4 (2 tandem in each chromosome half)	Replication, include type I and III repeats		
	IV. 17 bp (3')	5	Termination?		
	V. 130 bp (3') [‡]	4	Gene packing?		
<i>Tetrahymena pyriformis</i>	I. 33 bp (5')	3	Spacer promoter?		Challoner <i>et al.</i> (1985)
	II. 10–24 bp (5')	11	Enhancer?		
	III. 14–21 bp (5')	5			
	IV. 17 bp (3')	ND			
	V. 130 bp (3')	ND			
<i>Perkinsus andrewsi</i>	15 bp approximately	8			Coss <i>et al.</i> (2001)
<i>Pfiesteria piscicida</i>	9 bp				Saito <i>et al.</i> (2002)
Amoebozoa					
<i>Acanthamoeba castellanii</i>	106–174 bp	6	UBF binding	2.33	Yang <i>et al.</i> (1994)
<i>Dictyostelium discoideum</i>	29 bp (3') 5'	4			Sucgang <i>et al.</i> (2003)
			§		
<i>Entamoeba histolytica</i>	Dral 170 bp (3')	10	ARS-like sequences	9.2 (5')	Huber <i>et al.</i> (1989), Mittal <i>et al.</i> (1992), Bhattacharya <i>et al.</i> (1998)
	Scal 144 bp (3')	7		3.5 (3')	
	Scal 144 bp (5')	6			
	Pvul 145 bp (5')	11	Pathogenic-specific sequence		
	Hinfl 653 bp (5') [¶]	2 and 4	Recombination		
	74 bp (5')	2			
	Avall 153 bp (5')	5			
	1408 bp (5')	2			

Table 3. Continued.

Organism	Size of repeated sequences	Number of repeated sequences	Function	Intergenic region size (kb)	References
<i>Physarum polycephalum</i>	I. 130 bp (5')			≥22	Hattori <i>et al.</i> (1984)
	I'. 130 bp (3')				
	II. 50 bp (5')	16			
	II'. 52 bp (3')	6			
Opisthokonta					
<i>Saccharomyces cerevisiae</i>	I. 6 bp (5')	3		1.1 (5')	Skryabin <i>et al.</i> (1984)
	II. 16 bp (5')	2		1.25 (3')	
	III. 8 bp (5')	3			
	IV. 11 bp (3')	2			
	V. 9 bp (3')	2			
	VI. 9 bp (3')	2			
<i>Yarrowia lipolytica</i>	140–150 bp	Varies			van Heerikhuizen <i>et al.</i> (1985), Fournier <i>et al.</i> (1986)
	11 bp	14			
<i>Encephalitozoon cuniculi</i>	29 bp (5')	2			Peyretailade <i>et al.</i> (1998)
	19 bp (5')	8			
	51 bp (3')				
	43 bp (3')				

Repeated sequences may not be identical.

*The 55-bp repeat has an internal repeated inverted sequence.

†Type I and III repetitions are within a 430-bp repeated segment involved in the replication of the minichromosome.

‡Inverted repeat.

§Two blocks of highly repetitive DNA bracket the transcribed region.

¶Polymorphic locus with a 65-bp internal inverted repeated sequence.

ND, not determined; UBF, upstream binding factor; IGS, intergenic spacer.

Table 4. Organisms with unlinked rDNA units

Organism	Copy number	Chromosomal localization and rDNA types	References
Chromalveolates			
<i>Babesia bigemina</i>	3		Reddy <i>et al.</i> (1991)
<i>Babesia bovis</i>	3	Two in chr III and one in chr IV	Dalrymple (1990), Brayton <i>et al.</i> (2007)
<i>Babesia canis</i>	4	Two in one chr and two in the other chr	Dalrymple <i>et al.</i> (1992)
<i>Cryptosporidium parvum</i>	5	3 tandem, 2 alone in different chr	Le Blancq <i>et al.</i> (1997)
<i>Theileria parva</i>	2	A and B units in different chr	Kibe <i>et al.</i> (1994), Bishop <i>et al.</i> (2000)
<i>Plasmodium berghei</i>	4	Type A in chr XII and VII; Type C in chr VI and V	Waters (1994)
<i>Plasmodium falciparum</i>	5–8	Subtelomeric: Type A in chr V and VII; Type S in chr XI and XIII	Langsley <i>et al.</i> (1983), Mercereau-Puijalon <i>et al.</i> (2002), Gardner <i>et al.</i> (2002)
<i>Plasmodium lophurae</i>	6 (7–9)		Unnasch & Wirth (1983)
<i>Plasmodium vivax</i>	7	A, S and O rDNA types	van Spaendonk <i>et al.</i> (2000)
Plantae			
<i>Cyanidioschyzon merolae</i>	3	3 different loci, two in chr XVII and one in chr XVIII	Maruyama <i>et al.</i> (2004)

Chr, chromosome.

palindromic circular organization (25.9 kb), while the HK-9 (15.3 kb) and the Rahm (18.3 kb) isolates possess single rDNA units in their circular extrachromosomal molecules (Sehgal *et al.*, 1994; Bhattacharya *et al.*, 1998).

Most *E. gracilis* rDNA is found in extrachromosomal circular molecules that code for a single rDNA unit (Fig. 5b and Table 5). The whole *E. gracilis* rDNA circle is transcribed, suggesting a read-around transcription without the

need for transcriptional terminators (Greenwood *et al.*, 2001). *Naegleria gruberi* rDNA plasmid contains two ORFs: a large one downstream of the 28S rRNA (similar to a homing endonuclease gene, HE gene) and a short one that codes for a hypothetical protein (Maruyama & Nozaki, 2007) (Fig. 5c and Table 5).

The yeast *C. albicans* possesses both chromosomal and extrachromosomal rDNA. About 200 copies in tandem,

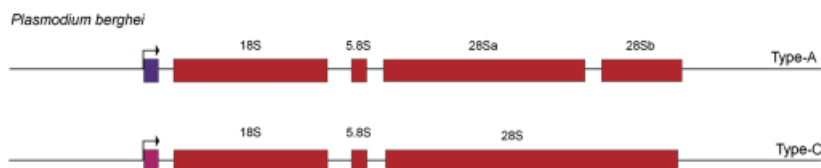


Fig. 3. rDNA organization in *Plasmodium berghei*. Four unlinked rDNA copies are coded in the telomeres of *P. berghei*. Type-A rDNAs contain a mature eLSU rRNA fragmented into three molecules (5.8S, 28Sa and 28Sb) and are expressed during the asexual stage in vertebrate hosts. Type-C rDNAs (with a nonfragmented 28S rRNA) are expressed during the sexual development in mosquitoes. The differential expression of rDNAs is regulated by specific promoter sequences (purple and pink boxes). Drawings are not to scale. The size of rDNA units is shown in Table 1.

Table 5. Extrachromosomal rDNA units

Organism	Lineal/ circular	Size	Organization	Copy number	Additional copies in chromosome	References
Excavata						
<i>Euglena gracilis</i>	C	11.5 kbp	Single	800–4000	4	Ravel-Chapuis (1988), Schnare <i>et al.</i> (1990)
<i>Naegleria gruberi</i>	C	14 kbp	Single	300–5000	None	Clark & Cross (1987)
Chromalveolates						
<i>Euplotes crassus</i>	L	7 kbp	Single gene		MIC	Erbeznik <i>et al.</i> (1999)
<i>Glaucoma chattoni</i>	L	9.3 kbp	Single gene		MIC	Katzen <i>et al.</i> (1981), Challoner <i>et al.</i> (1985)
<i>Nyctotherus ovalis</i>	L		Single gene		MIC	Ricard <i>et al.</i> (2008)
<i>Oxytricha fallax</i>	L	7.49 kbp	Single gene		MIC	Rae & Spear (1978), Swanton <i>et al.</i> (1982)
<i>Oxytricha nova</i>	L	7.49 kbp	Single gene		MIC	Swanton <i>et al.</i> (1982)
<i>Paramecium tetraurelia</i>	C and L	*	Tandem		MIC	Findly & Gall (1978)
<i>Stylonychia mytilus</i>	L		Single gene		MIC	Lipps & Steinbrück (1978)
<i>Tetrahymena pyriformis</i>	L		Palindrome	200/haploid MAC	1	Engberg <i>et al.</i> (1976), Yao & Gall (1977), Niles <i>et al.</i> (1981)
<i>Tetrahymena thermophila</i>	L	21 kbp	Palindrome	9000 MAC		Engberg (1985), Eisen <i>et al.</i> (2006)
Amoebozoa						
<i>Dictyostelium discoideum</i>	L	88 kbp	Palindrome	90	1 Palindrome	Cockburn <i>et al.</i> (1978), Hofmann <i>et al.</i> (1993), Eichinger <i>et al.</i> (2005)
<i>Didymium iridis</i>	L	20 kbp	Single			Johansen <i>et al.</i> (1992)
<i>Entamoeba histolytica</i>	C	24.5 kbp	Palindrome	200/haploid	None	Huber <i>et al.</i> (1989), Bhattacharya <i>et al.</i> (1998), Bagchi <i>et al.</i> (1999)
HM-1:IMSS						
<i>Physarum polycephalum</i>	L	> 60 kbp	Palindrome	1×10^{11}		Vogt & Braun (1976), Campbell <i>et al.</i> (1979)
Opisthokonta						
<i>Candida albicans</i>	C and L	1.2 Mbp	Tandem	100	200	Huber & Rustchenko (2001)

*The size depends on the number of rDNA repeats.

MAC, macronucleus; MIC micronucleus.

varying in size, are present in chromosome R while roughly 100 copies are found in an ~1.2 Mbp autonomously replicating circle. Some rDNA sequences are also found in 50–150-kbp linear molecules (Huber & Rustchenko, 2001).

rDNA plasmids have only been observed in old *S. cerevisiae* cell cultures. During the aging process, the tandem rDNA copies are excised from the chromosome and replicate autonomously. The accumulation of rDNA circles leads to yeast sterility and shortening of the life span. An association between rDNA locus instability and loss of

epigenetic silencing has also been observed (Sinclair & Guarente, 1997).

The ribosomal cistron: the coding region

The typical eukaryotic rDNA coding region is composed of the 18S, 5.8S and 28S rRNA coding sequences separated by ITS-1 and ITS-2. The rDNA coding sequence consists of a common core of domains that may be interspersed with a distinct set of variable regions (also called expansion

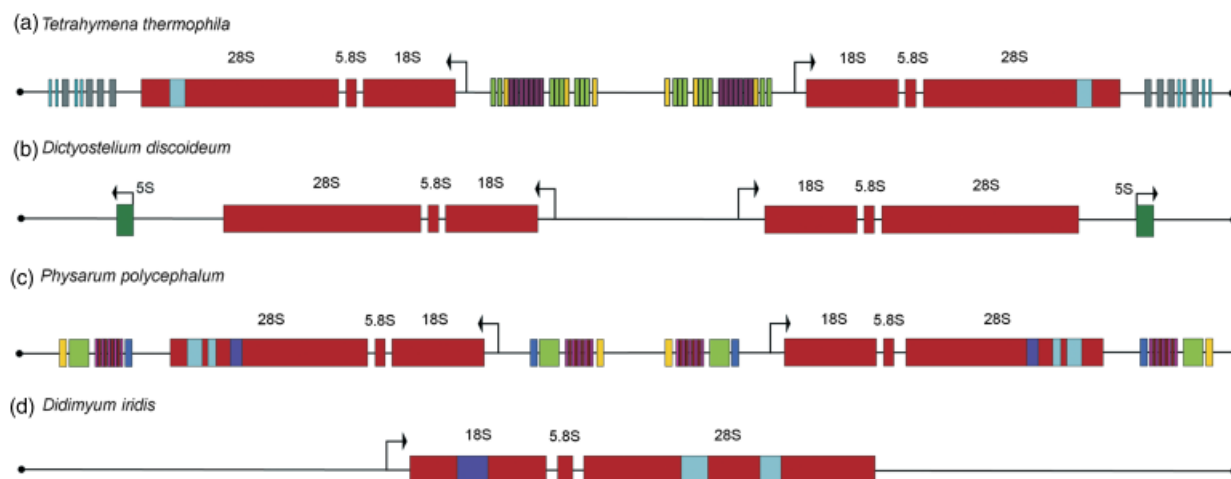


Fig. 4. Linear extrachromosomal rDNA units in microbial eukaryotes. (a) rDNA is coded in extrachromosomal linear palindromes in *Tetrahymena thermophila*. Two head-to-head rDNA units are coded in a macronuclear minichromosome that contains typical telomeric sequences (black dots). Upstream and downstream IGRs contain various types of repeated sequences (colored bars; see also Table 3). A group I intron (blue square) is found within the eLSU rRNA. (b) The rDNA and 5S rDNA in *Dictyostelium discoideum* are linked and coded in extrachromosomal linear palindromes containing telomeres (black dots). (c) rDNA in *Physarum polycephalum* is coded in palindromic head-to-head units in a minichromosome with various repeated sequences both in the 5' and 3' IGRs (see also Table 3). The eLSU rRNA is interrupted by two group I introns (blue squares), and a third group I intron that includes an HE gene (purple square). (d) In *Didymium iridis* the rDNA is coded in linear minichromosomes containing one rDNA unit. The SSU rRNA contains a twintron (purple box) and two group I introns are found within the 28S rRNA (blue boxes). Drawings are not to scale. The size of rDNA units is shown in Table 1. Arrows show the polarity of transcription.

segments; Dover, 1988). Ten and 18 variable regions have been identified in the SSU and LSU rRNAs of all organisms (Raué *et al.*, 1988). Three types of sequence insertions have been found within these variable regions: (1) expansion segments, encoding RNA sequences conserved in the mature molecule; (2) group I introns, located within highly conserved regions and removed after transcription; and (3) transcribed spacers, sequences removed from the mature rRNA, thus producing fragmented eLSU rRNA molecules (Clark *et al.*, 1984). *Babesia bovis* (Dalrymple *et al.*, 1992), *Cryptosporidium parvum* (Le Blancq *et al.*, 1997), *D. discoideum* (Frankel *et al.*, 1977), *E. histolytica* (Huber *et al.*, 1989), *G. intestinalis* (Healey *et al.*, 1990), *K. lactis* (Verbeet *et al.*, 1984), *S. cerevisiae* (Bell *et al.*, 1977), *Toxoplasma gondii* (Gagnon *et al.*, 1996) and *T. vaginalis* (López-Villaseñor *et al.*, 2004) are microbial eukaryotes with the typical rDNA organization of the coding region (Fig. 2a–e).

Insertions of expansion segments in the SSU

The average length of eukaryotic SSU rRNA is 2 kb. Unusually long SSU rRNAs have been found in *Pelobionta* (*Pelomyxa palustris*), *Foraminifera* (*Hemisphaerammina bradyi*) and *Euglenozoa* (*Distigma sennii*) (Table 6). The longest SSU rRNA known is found in the Euglenid *D. sennii*, comprising > 4.5 kb. In most cases, the insertions are found in the SSU rRNA variable regions V2, V4 and V7. The only exceptions are an extended V5 region in

A. castellanii and an expansion in a nonvariable region in *P. palustris* (Gunderson & Sogin, 1986; Milyutina *et al.*, 2001). The rRNA variable regions are located in the mature ribosome surface and their evolutionary implications are unknown (Katz & Bhattacharya, 2006).

Group I introns and twintrons

Group I introns can be found as insertions in the SSU and eLSU rRNA coding regions that are removed from the mature molecule by means of a self-splicing reaction (Einvik *et al.*, 1998), generating a completely functional rRNA molecule (Mandal, 1984). The ribozymes encoded in group I introns have conserved secondary structures of 10 base-paired segments, as well as some additional paired segments depending on the intron subclass (Michel & Westhof, 1990). The splicing reaction initiates with a nucleophilic attack of a guanosine cofactor at the 5' splice site and, after two sequential transesterification reactions, the exons are ligated and the RNA intron is removed (Einvik *et al.*, 1998).

Group I introns are widely distributed in nature and can be found in bacteria, mitochondrial and chloroplast genomes, and in the eukaryotic nucleus (Johansen *et al.*, 2007). Group I introns may interrupt the SSU rRNA coding sequence in 40 distinct conserved sites of several microbial eukaryotes (Jackson *et al.*, 2002) such as *Acanthamoeba griffini* and the green alga *Characium saccatum*. These introns may also be present in the eLSU rRNA, as is the case

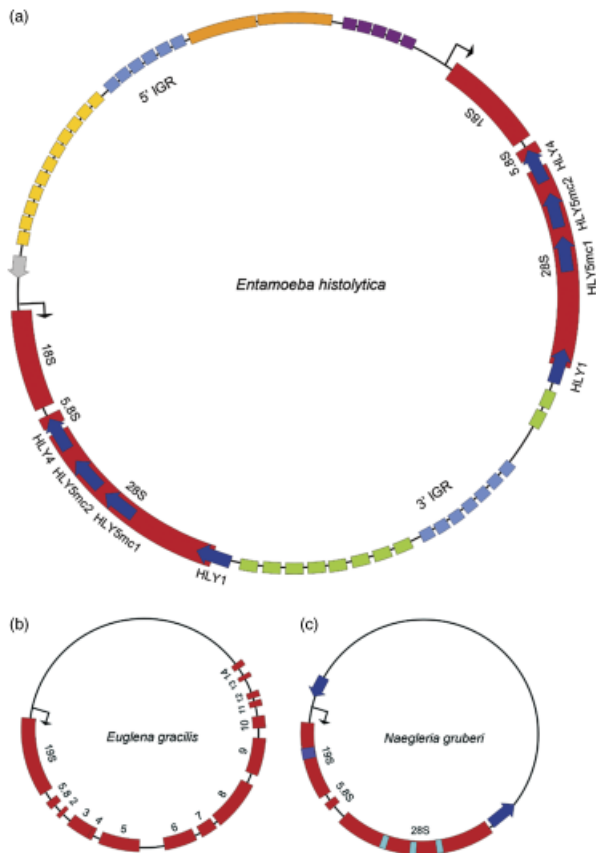


Fig. 5. Circular extrachromosomal rDNA units in microbial eukaryotes. (a) In *Entamoeba histolytica* extrachromosomal self-replicating molecules encode two palindromic rDNA units (red). The upstream and downstream IGRs contain several repeated sequences (coloured boxes, detailed in Table 3). The 5' IGR also encodes a 0.7-kb mRNA (grey arrow). In the HM1 strain, four hemolysine virulence proteins (HLYs) are coded within the rRNA coding region, in antisense orientation relative to the rRNA coding strand (navy blue arrows). (b) In *Euglena gracilis* the rDNA is coded in circular plasmids and the eLSU is fragmented in 14 segments. (c) The *Naegleria gruberi* rDNA plasmid encodes one rDNA unit containing one twintron in 18S rRNA (purple box) and three type-I introns in the eLSU rRNA (blue boxes). The IGR contains two ORFs (dark blue arrows). Black arrows show the polarity of transcription.

for *P. falciparum* A-type eLSU rRNA and the 26S rRNA of some *Tetrahymena* isolates (Fig. 4a, Table 7). It is interesting that some organisms may have both the SSU and the eLSU rRNAs interrupted by group I introns (e.g. *P. carinii*, *Chlorella ellipsoidea* and *D. iridis*, Fig. 4d). Table 7 describes some of the introns found in the rDNA of several microbial eukaryotes.

Twintrons are more complex insertions in the rDNA that consist of two group I introns (ribozymes) and an ORF encoding an HE (Einvik *et al.*, 1998; Johansen *et al.*, 2007). The *D. iridis* and *N. gruberi* SSU twintrons contain a small ribozyme (GIR1), followed by the HE ORF inserted into a second ribozyme (GIR2). Two different isolates of *D. iridis*

Table 6. Variations in the size of the SSU rRNA due to insertions

Organism	SSU size (kb)	References
Excavata		
<i>Astasia curvata</i>	2.56	Busse & Preisfeld (2002)
<i>Astasia torta</i>	2.9	Busse & Preisfeld (2002)
<i>Distigma curvatum</i>	3.4–3.7	Busse & Preisfeld (2002)
<i>Distigma elegans</i>	3.9	Busse & Preisfeld (2002)
<i>Distigma sennii</i>	4.5	Busse & Preisfeld (2002)
<i>Pleoetia costata</i>	2.4	Busse & Preisfeld (2003)
<i>Euglena gracilis</i>	2.3	Gunderson & Sogin (1986)
Amoebozoa		
<i>Acanthamoeba castellanii</i>	2.3	Gunderson & Sogin (1986)
<i>Acanthamoeba griffini</i>	2.9	Gast <i>et al.</i> (1994)
<i>Acanthamoeba lenticulata</i>	3	Schroeder-Diedrich <i>et al.</i> (1998)
<i>Pelomyxa palustris</i>	3.5	Milyutina <i>et al.</i> (2001)
<i>Phreatamoeba balamuthi</i>	2.74	Hinkle <i>et al.</i> (1994)
<i>Entamoeba histolytica</i>	2.3	Loftus <i>et al.</i> (2005)
Plantae		
<i>Ankistrodesmus stipitatus</i>	2.2	Dávila-Aponte <i>et al.</i> (1991)
Rhizaria		
<i>Foraminifera</i>	2.3–4	Katz & Bhattacharya (2006)

have two types of introns, containing an HE gene in both polarities relative to the SSU rRNA gene (Fig. 6) (Johansen *et al.*, 2007). The twintron contains the HE ORF (I-DirI) in the same polarity as the 18S rRNA coding region. GIR2 is a self-splicing ribozyme that releases the HE transcript. A second intron encoding a ribozyme (GIR1) is also found within the twintron. GIR1 modifies the 5' end of the HE transcript to form a 2'5' cap that increases its translational efficiency (Fig. 6a) (Einvik *et al.*, 1998; Johansen *et al.*, 2007). In contrast, the intron II contains an HE gene (I-DirII) in opposite polarity relative to the SSU rRNA and ribozyme-coding sequences (Johansen *et al.*, 2006). Transcription of I-DirII is established from a pol II-like promoter located immediately upstream of the HE gene (Fig. 6b) (Johansen *et al.*, 2006). Both *D. iridis* HE transcripts are processed through the nuclear spliceosomal complex to remove a 50-nt noncoding spliceosomal intron, found within the HE coding sequences, and are polyadenylated (Vader *et al.*, 1999; Johansen *et al.*, 2007).

Physarum polycephalum eLSU rDNA contains an optional group I intron holding an HE gene (Ruoff *et al.*, 1992). The full-length RNA intron can be excised or alternatively processed (immediately downstream of the HE gene) to produce a smaller transcript. Only the full-length RNA intron (lacking a 5' cap and a poly-A tail) is translated into the HE I-PpoI protein (Ruoff *et al.*, 1992). The cleavage of this transcript in the internal processing site seems to downregulate HE I-PpoI expression by decreasing the stability of the transcript in yeast transintegrated introns (Johansen *et al.*, 2007). Table 7 summarizes rDNA group I introns and HE gene insertions.

Table 7. Group I rDNA introns

Organism	Location	Number of introns	Size	HE gene	References
Excavata					
<i>Ploeotia costata</i>	SSU rRNA	1	494 bp		Busse & Preisfeld (2003)
<i>Naegleria gruberi</i>	SSU rRNA	1		I-NgrI	Wikmark <i>et al.</i> (2006)
	LSU rRNA	3			Einvik <i>et al.</i> (1998)
Chromalveolates					
<i>Plasmodium falciparum</i>	LSU rRNA	1		*	Langsley <i>et al.</i> (1983)
<i>Tetrahymena pigmentosa</i>	LSU rRNA	1	400 bp		Wild & Gall (1979)
<i>Tetrahymena thermophila</i>	LSU rRNA	1	370–410 bp		Sogin <i>et al.</i> (1986)
Amoebozoa					
<i>Acanthamoeba griffini</i>	SSU rRNA	1	519 bp		Gast <i>et al.</i> (1994)
<i>Acanthamoeba lenticulata</i>	SSU rRNA	1	656 bp		Gast <i>et al.</i> (1994)
<i>Didymium iridis</i>	SSU rRNA	1	1.43 kbp	I-DirI [†]	Johansen & Vogt (1994), Johansen <i>et al.</i> (2006), Johansen <i>et al.</i> (2007)
	LSU rRNA	2	688 and 573 bp		Johansen <i>et al.</i> (1992)
<i>Physarum polycephalum</i>	LSU rRNA	3	0.7, 0.6 and 0.94 kbp	I-Ppol [‡]	Ruoff <i>et al.</i> (1992), Johansen <i>et al.</i> (2007)
Plantae					
<i>Ankistrodesmus stipitatus</i>	SSU rRNA	1	334 bp		Dávila-Aponte <i>et al.</i> (1991)
<i>Characium saccatum</i>	SSU rRNA	1	477 bp		Wilcox <i>et al.</i> (1992)
<i>Chlorella ellipsoidea</i>	SSU rRNA	1	442 bp		Aimi <i>et al.</i> (1994)
	LSU rRNA	1	445 bp		Aimi <i>et al.</i> (1993)
<i>Dunaliella parva</i>	SSU rRNA	2	381 and 419 bp		Van Oppen <i>et al.</i> (1993), Wilcox <i>et al.</i> (1992)
<i>Dunaliella salina</i>	SSU rRNA	1 [§]	397/8 bp		Van Oppen <i>et al.</i> (1993), Wilcox <i>et al.</i> (1992)
Opisthokonta					
<i>Candida albicans</i>	LSU rRNA	1	379 bp		Miletti-González & Leibowitz (2008)
<i>Pneumocystis carinii</i>	SSU rRNA	1	390 bp		Sogin & Edman (1989)
	LSU rRNA	1			Lin <i>et al.</i> (1992), Liu <i>et al.</i> (1992)
<i>Histoplasma capsulatum</i>	SSU rRNA	1	403–425 bp		Okeke <i>et al.</i> (1998), Lasker <i>et al.</i> (1998)
<i>Nosema bombycis</i>	SSU rRNA	1			Iiyama <i>et al.</i> (2004)

*Only in one of the eight LSU rRNA copies.

[†]Depends on the isolate.

[‡]Coded in the 0.94-kbp intron.

[§]Can have two types of intron differing in sequence.

The ITS-1 and -2

The rDNA transcript is generally post-transcriptionally processed in three rRNA mature molecules: 18S, 5.8S and 28S rRNAs that result from elimination of ETS, ITS-1 and ITS-2 from the precursor transcript (Fig. 1). In microbial eukaryotes, ITS-1 ranges from 100 to 400 bp, while ITS-2 is 200–500 bp. Unusually long ITSs are found in the red alga *C. merolae* (Maruyama *et al.*, 2004), where ITS-1 and ITS-2 average sizes are 862 and 1738 bp, respectively. *Euglena gracilis* has the largest known ITS-1, 1188 bp in length (Schnare *et al.*, 1990). The dinoflagellate *Cochlodinium polykrikoides* ITS-1 contains a 101-bp sequence in six tandem repeats, resulting in an ITS-1 length of 813 bp (Ki & Han, 2007). *Yarrowia* and *Giardia* have the shortest known ITSs in microbial eukaryotes: the sum of ITS-1 and ITS-2 lengths in *Y. lipolytica* is only 150 bp (van Heerikhui-

zen *et al.*, 1985), while the *G. intestinalis* ITS-1 and ITS-2 are 37 and 52 bp in length, respectively (Boothroyd *et al.*, 1987). Some *Microsporidia* species completely lack the ITS-2 (Fig. 7) (Vossbrinck & Woese, 1986), as discussed below. The biological relevance of the ITSs' length and the presence of internal repeats are currently unknown, although their sequence has been useful in molecular phylogenetic studies of closely related species.

Additional ITSs generate fragmented eLSU rRNA

Some microbial eukaryotes process the pre-rRNA into more than three mature molecules due to the presence of additional ITSs. Well-known examples of fragmented rRNA are found among kinetoplastids, with the eLSU rRNA fragmented in seven molecules. The nomenclature of these rRNAs varies according to the organism, the size of the rRNA

molecule and the position in the coding region. The eLSU rRNA of *Leishmania* spp. is fragmented into seven elements, which are cotranscribed in the pre-rRNA and processed by exo- and endonucleolytic activities to produce the functional eLSU fragments: 5.8S, LSU α , γ , LSU β , δ , ζ and ϵ (Martínez-Calvillo *et al.*, 2001) (Fig. 2g). *Trypanosoma cruzi* and *Crithidia fasciculata* also code for an eLSU rRNA fragmented into seven elements: 5.8S, 24S α , S1, 24S β , S2, S6 and S4 in *T. cruzi* (Fig. 2f) (Hernández *et al.*, 1988), and rRNAs 5.8S, c, d, e, f, g and j in *C. fasciculata* (Spencer *et al.*, 1987).

Processing of the eLSU rRNA into several fragments has also been found in nonkinetoplastid eukaryotes (e.g. *P. berghei* and *Plasmodium chabaudi* blood stages and *A. castellanii*) (D'Alessio *et al.*, 1981; da Silveira & Mercereau-Puijalon, 1983; Johansen *et al.*, 1992) (Figs 2e and 3). *Euglena gracilis* has the most fragmented eLSU rRNA currently known with 14 mature molecules that result from

the processing of 14 ITSs (ITS-1 to ITS-14, Fig. 5b) (Schnare *et al.*, 1990).

Protein-coding regions within the rDNA coding region

A correlation has been observed between the virulence of *E. histolytica* isolates and the sequence composition of the rDNA circular molecule described above (Clark & Diamond, 1991; Zindrou *et al.*, 2001). Virulence associates with the striking presence of genes encoding hemolysins (proposed as virulence factors) within and overlapping the rRNA coding sequence, but in opposite polarity. Three hemolysins overlap with the eLSU coding region, while the fourth (HLY4) is coded in the ITS-1 between the SSU and 5.8S rRNAs (Jansson *et al.*, 1994) (Fig. 5a). In *G. intestinalis*, a gene coding for a 32-kDa flagellum antigen has been identified in the rDNA IGR that overlaps the 3' region of the 28S rRNA (Fig. 2d). The motif that directs transcription of this gene seems to be a hybrid pol II/pol III promoter (Upcroft *et al.*, 1990).

Unusual rDNA coding regions

Microsporidia are obligate intracellular eukaryotes that possess many prokaryotic characteristics in their rRNA genes (Weiss, 2001). The rDNA units are smaller than the standard eukaryotic size and lack the ITS-2; consequently, the 5.8S rRNA is fused to the 5' region of the 28S rRNA, as is found in bacteria (Vossbrinck & Woese, 1986) (Fig. 7). *Microsporidia* are the only eukaryotes known to lack an individual 5.8S rRNA molecule (e.g. *E. cuniculi* and *Vairimorpha necatrix* (Vossbrinck & Woese, 1986; Peyretailade *et al.*, 1998). The relevance of this eukaryotic 5.8S–28S rRNA fusion is unknown. In addition to these characteristics, *Nosema bombycis* and *Nosema spodopterae* have an unusual rDNA gene organization (Huang *et al.*, 2004; Iiyama *et al.*, 2004; Tsai *et al.*, 2005) because the LSU rRNA is coded and transcribed upstream to the SSU rRNA (Fig. 7b) in contrast to the almost universal order of the rRNA coding regions (Fig. 1).

Downloaded from https://academic.oup.com/femsre/article/34/1/59/590747 by guest on 18 April 2024

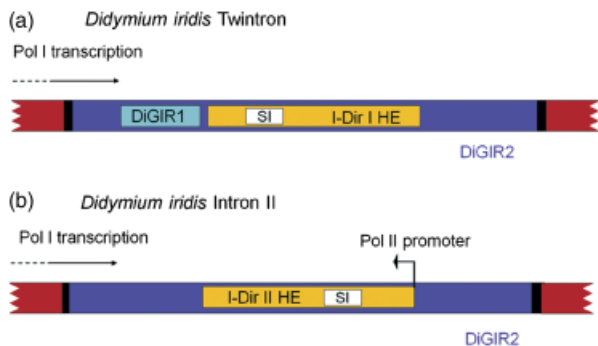


Fig. 6. Group I introns that contain an HE gene. (a) Twintron present in *Didymium iridis*: the DiGIR2 intron (purple) is encoded in the SSU rRNA and transcribed by pol I as part of the pre-rRNA; it self-splices to generate the HE pre-mRNA (splicing sites are represented as black bars). Subsequently, DiGIR1 intron (blue) self-splices and processes the I-Dir I HE pre-mRNA in the 5' side, producing a 2'5'-cap. The I-Dir I HE pre-mRNA is additionally processed by the removal of spliceosomal intron SI (white box) and polyadenylation of the 3' side to generate a functional I-Dir I HE mRNA (yellow region). (b) Intron II present in *D. iridis*: the I-Dir II HE rRNA found within the DiGIR2 intron is coded in antisense orientation and is transcribed from a pol II promoter. The HE pre-mRNA is processed by pol II-associated factors to generate a typical 5'-cap and a 3' polyadenylated tail. The spliceosomal intron SI is removed by the spliceosome machinery.

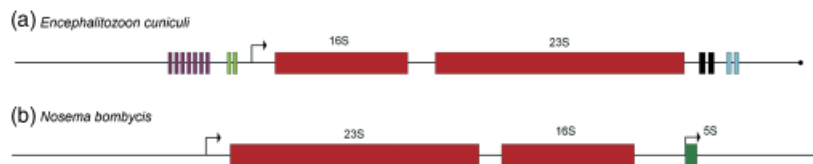


Fig. 7. Unusual rDNA organization in *Microsporidia*. *Microsporidia* lack a 5.8S rRNA mature molecule and the typical 5.8S rRNA sequence is fused to the 23S rRNA. (a) Single telomeric rDNA units are surrounded by different repeated sequences in *Encephalitozoon cuniculi* (see also Table 3) and the rDNA lacks ITS-2. (b) Some *Nosema* species have an atypical rDNA coding organization, with the LSU rRNA coded upstream of the 16S rRNA. The typical 5.8S rRNA sequence is fused to the 23S rRNA and the 5S rDNA is linked to the rDNA unit.

Table 8. Variability found in the rDNA

Organism	rDNA types*	Localization of variability	References
Chromalveolates			
<i>Babesia bigemina</i>	2	IGR and SSU coding region	Reddy <i>et al.</i> (1991), Dalrymple <i>et al.</i> (1992)
<i>Babesia bovis</i>	A, B and C	ITS and SSU coding region	Laughery <i>et al.</i> (2009)
<i>Cryptosporidium parvum</i>	At least 2	ITS and coding region	Le Blancq <i>et al.</i> (1997), Spano & Crisanti (2000)
<i>Plasmodium berghei</i>	A and C	IGR, promoter and coding regions	Waters (1994)
<i>Plasmodium falciparum</i>	A and S	IGR, promoter and coding regions	Waters (1994)
<i>Plasmodium vivax</i>	A, S and O	IGR, promoter and coding regions	Li <i>et al.</i> (1997)
<i>Theileria parva</i>	2	ITS and LSU coding region	Bishop <i>et al.</i> (2000)
<i>Toxoplasma gondii</i>		IGR and SSU coding region	Fazaeli <i>et al.</i> (2000)
<i>Oxytricha fallax</i>	2	LSU coding region	Doak <i>et al.</i> (2003)
<i>Paramecium tatraurelia</i>	6 MAC/ 4 MIC	IGR	Preer <i>et al.</i> (1999)
<i>Perkinsus andrewsi</i>	A and B	IGR and coding regions	Pecher <i>et al.</i> (2004)
Plantae			
<i>Dunaliella salina</i>	2	Group I introns	Wilcox <i>et al.</i> (1992)
Opisthokonta			
<i>Saccharomyces cerevisiae</i>		IGR	Skryabin <i>et al.</i> (1984), Jemtland <i>et al.</i> (1986)
<i>Candida albicans</i>	2	Intron-containing and intronless LSU	Miletti-González & Leibowitz (2008)
<i>Yarrowia lipolytica</i>	2–5	IGR	van Heerikhuizen <i>et al.</i> (1985), Fournier <i>et al.</i> (1986), Clare <i>et al.</i> (1986)
<i>Nosema apis</i>	At least 3	IGR	Gatehouse & Malone (1998)
<i>Nosema bombi</i>	At least 2	ITS	O'Mahony <i>et al.</i> (2007)

*The number or names of the rDNA types for each species are shown. IGR, intergenic region; MAC, macronucleus; MIC, micronucleus.

Different rDNA genes may be found within an organism

As has been mentioned, the rRNA genes within one organism are generally conserved in the coding region with an occasional sequence variation in the IGRs and with little variation in the coding sequences. Sequence variability in the IGRs may result from sequence divergence or disparity in the number of repeated sequences, involved in both up- and downregulation of rDNA transcription. Therefore, the heterogeneous composition of rDNA units may influence rDNA expression. Sequence divergence in the coding region and/or IGR within the same organism has led to a classification of rDNA units. For example, different types of rDNA may be found in *Paramecium*, *Y. lipolytica* and the *Apicomplexa* group. A detailed description of this variability is included in Table 8.

The 5S rDNA

The organization of the 5S rDNA is simpler than that of the rDNA. Most 5S rDNAs are found in tandem head-to-tail repeats consisting of a conserved ~120-bp coding region and an IGR of variable size and sequence. An internal pol III promoter is present in all 5S rDNA studied to date (Schramm & Hernandez, 2002) (Fig. 1c).

The 5S rDNAs are found as tandem head-to-tail repeats

The 5S rDNA in *T. cruzi*, *Trypanosoma brucei*, *T. vaginalis*, *Trichomonas tenax*, *C. fasciculata*, *Eimeria tenella* and *C. parvum* is typically organized in tandem head-to-tail repeats. *Tritrichomonas foetus* has two types of 5S rDNAs, while *P. falciparum* has only three 5S rDNA copies in tandem, differing in the length of the IGRs. The main characteristics of the 5S rDNA tandem head-to-tail repeats of several organisms are described in Table 9.

The 5S rDNA may be interspersed with genes transcribed by any of the three RNA polymerases

The 5S rDNA has been found linked to the rDNA (transcribed by pol I) in an alternate distribution in *T. gondii*, *Hansenula polymorpha*, *Perkinsus andrewsi* and various *Nosema* species (Guay *et al.*, 1992; Coss *et al.*, 2001; Klabunde *et al.*, 2002; Huang *et al.*, 2004; Iiyama *et al.*, 2004; Tsai *et al.*, 2005; Liu *et al.*, 2008) (Figs 2c and 7b). Two tandem 5S rDNA copies are linked to each repeated rDNA unit in *Candida glabrata* (Dujon *et al.*, 2004). In contrast, the 5S rDNA is linked to the rDNA in opposite polarity in various yeast species, such as *Torulopsis utilis*, *K. lactis* and *S. cerevisiae* (Fig. 2b, Table 2). Two copies of the 5S rDNA are coded in the extrachromosomal DNA molecule in

Table 9. Typical 5S rDNA organization

Organism	5S rDNA types*	Localization of variability	Other	References
Excavata				
<i>Trichomonas tenax</i>	A and B	IGR and coding region	Type B IGR palindrome	Torres-Machorro <i>et al.</i> (2009)
<i>Trichomonas vaginalis</i>	A and B	IGR	IGS 10-bp palindrome	Torres-Machorro <i>et al.</i> (2006)
<i>Tritrichomonas foetus</i>	A and B	IGR, repeated sequences vs. ubiquitin gene		Torres-Machorro <i>et al.</i> (2009)
<i>Trypanosoma brucei</i>				Lenardo <i>et al.</i> (1985)
<i>Trypanosoma cruzi</i>			6 Sp1 binding sites in IGR	Hernández-Rivas <i>et al.</i> (1992)
Chromalveolates				
<i>Cryptosporidium parvum</i>		IGR		Taghi-Kilani <i>et al.</i> (1994)
<i>Eimeria tenella</i>				Stucki <i>et al.</i> (1993)
<i>Plasmodium falciparum</i>	3	IGR lengths		Shippen-Lentz & Vezza (1988)
<i>Tetrahymena pyriformis</i>		IGR	Some linked to ubiquitin genes	Guerreiro <i>et al.</i> (1993)
<i>Tetrahymena thermophila</i>		IGR lengths	IGR 12 and 16-bp palindromes	Allen <i>et al.</i> (1984)

*The number or names of the 5S rDNA types for each species are shown. IGR, intergenic region.

Table 10. 5S rRNA gene linkage to pol II transcribed genes

Organism	Pol II gene	Orientation	References
Excavata			
<i>Trypanosoma vivax</i>	Spliced leader	Sense	Roditi (1992)
<i>Trypanosoma rangeli</i>	Spliced leader	Sense	Aksoy <i>et al.</i> (1992)
<i>Bodo saltans</i>	Spliced leader	Sense	Santana <i>et al.</i> (2001)
<i>Bodo caudatus</i>	Spliced leader	Sense	Campbell (1992)
<i>Diplonema papillatum</i>	Spliced leader	Sense	Sturm <i>et al.</i> (2001)
<i>Herpetomonas</i> spp.	Spliced leader	Sense	Aksoy (1992)
<i>Trypanoplasma borreli</i>	Spliced leader	Antisense	Maslov <i>et al.</i> (1993)
<i>Trypanosoma avium</i>	Spliced leader	Antisense	Santana <i>et al.</i> (2001)
<i>Euglena gracilis</i>	Spliced leader	Sense	Keller <i>et al.</i> (1992)
<i>Tritrichomonas foetus</i>	Ubiquitin	Sense	Torres-Machorro <i>et al.</i> (2009)
Chromalveolates			
<i>Tetrahymena pyriformis</i>	Ubiquitin	Sense	Guerreiro <i>et al.</i> (1993)

D. discoideum, in the same polarity as the two rDNA copies (Hofmann *et al.*, 1993) (Fig. 4b). Finally, a *S. cerevisiae* 5S rDNA variant is found in five repeats of 3.6 kbp, located next to the rDNA tandem cluster locus, in the centromere-distal side (McMahon *et al.*, 1984).

Some *Trypanosoma* species such as *Trypanosoma vivax* and *Trypanosoma rangeli* have the 5S rDNA copies linked to the spliced-leader (SL) tandem repeated genes, transcribed by pol II. SL transcripts are necessary to process the mRNAs in kinetoplasts by a trans-splicing reaction (Simpson *et al.*, 2006). A similar linkage has been found in other *Euglenozoa* such as *Diplonema papillatum* and *Bodo caudatus*. In *Trypanoplasma borreli* and *Trypanosoma avium*, the 5S rDNA is coded in opposite polarity relative to the SL gene (Table 10). Interestingly, the *T. borreli* SL can also be linked to 5S rRNA pseudogenes (with a truncated 5' end) (Maslov *et al.*, 1993). Some 5S rDNA units in *T. pyriformis* and *T. foetus* are associated with ubiquitin genes transcribed by pol II (Fig.

8b). Table 10 describes the relative polarity of the 5S rDNA linked to the genes transcribed by pol II.

Some 5S rDNA copies in *E. histolytica* and one copy in *Leishmania tarentolae* are linked to tRNA genes (Shi *et al.*, 1994; Clark *et al.*, 2006), also transcribed by pol III. Interestingly, 48 of the 108 5S rDNA copies of *Y. lipolytica* produce pol III dicistronic transcripts: tRNA–5S rRNA hybrid molecules. The synthesis of an ~200-nt transcript is driven by the tRNA pol III promoter, resulting in a transcription independent of the 5S rDNA-specific transcription factor, TFIIIA. The dicistronic transcripts, as well as a unique tricistronic transcript [Lys(CTT) tRNA–Glu(CTC) tRNA–5S rDNA] are post-transcriptionally processed to generate the typical mature RNA molecules: tRNAs and 5S rRNA (Acker *et al.*, 2008) (Fig. 8c).

Nontandem 5S rDNA copies are found dispersed throughout the genome of some microbial eukaryotes. Some examples are *A. castellanii* (Zwick *et al.*, 1991),

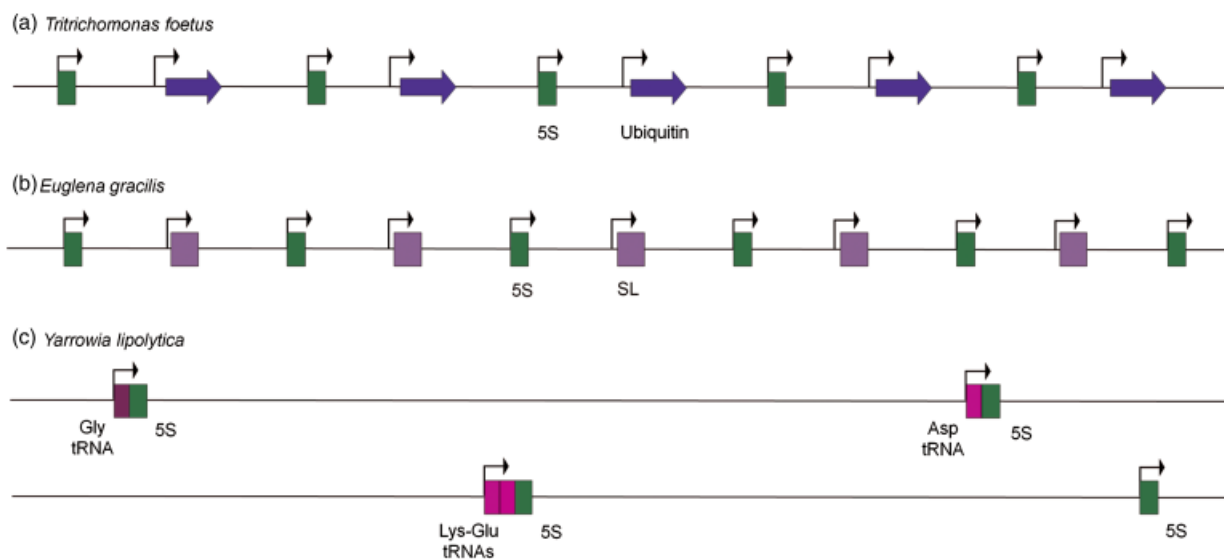


Fig. 8. The 5S rDNA may be linked to pol II or pol III transcribed genes. (a) In *Trichomonas foetus* the 5S rDNA is linked to a multigenic ubiquitin family. (b) In *Euglena gracilis* the 5S rDNA is linked to the SL gene. (c) In *Yarrowia lipolytica* dicistronic genes consisting of a tRNA gene (pink) and a 5S rDNA (green) are dispersed in the genome. One tricistronic gene: Lys(CTT) tRNA–Glu(CTC) tRNA–5S rDNA is also found. These genes are transcribed from the pol III promoter of the tRNA gene. Dispersed, single 5S rDNAs are also found (green).

Y. lipolytica and *Schizosaccharomyces pombe* (Tabata, 1981; Dujon *et al.*, 2004). The 5S rDNA may also be found in extrachromosomal DNA. Noteworthy, ciliate organisms such as *Oxytricha fallax* have single 5S rDNA copies coded in macronucleus extrachromosomal molecules (Rae & Spear, 1978; Roberson *et al.*, 1989). Moreover, about one million copies of the 5S rDNA are coded in linear minichromosomes flanked by telomeres in *E. eurystomus* (Roberson *et al.*, 1989).

Concluding remarks

Ribosomes are complex organelles that require the intricate collaboration of three types of RNAs (rRNA, mRNA and tRNA) and > 70 proteins for the synthesis of proteins. rRNAs must maintain their convoluted structural motifs in order to be functional. It is therefore not surprising that their sequence is highly conserved among related organisms and this similarity is gradually lost as organisms diverge. For this reason, sequence comparison of the SSU rRNA has been widely used in the field of molecular phylogeny (Van de Peer *et al.*, 2000).

The ‘typical’ eukaryotic rDNA genomic organization was proposed > 30 years ago, based on the analysis of the rDNA in higher eukaryotes (Long & Dawid, 1980). The tandemly repeated head-to-tail organization has been considered the standard for eukaryotic rDNA. Surprisingly, analyses of the genomic organization of ribosomal genes in microbial eukaryotes demonstrate that although some organisms do hold the typical rDNA configuration, the majority reveal

unusual characteristics. As shown in this review, the eukaryotic rDNA may be arranged in a wide variety of genomic configurations, suggesting the existence of several regulatory mechanisms (probably species-specific) within a conserved rDNA regulatory context.

Reiteration is one of the most conserved rDNA characteristics. The rDNA copy number is extremely variable and appears to be highly regulated within species. Nevertheless, the total number of rDNA repeats does not always correlate with the rate of rRNA synthesis (French *et al.*, 2003), implying that individual rDNA units may hold different epigenetic marks that result in variable transcriptional rates (Grummt, 2007). rDNA structure and transcription are also important in the establishment of the nucleolar structure, which also plays regulatory roles at the cellular level (Carmo-Fonseca *et al.*, 2000). *Dictyostelium discoideum*, *T. gondii* and some yeast species have equal numbers of rDNA and 5S rDNA copies. This organization was considered coherent, as the pool for rRNA molecules was supposed to hold equimolar amounts of 18S, 5.8S, 28S and 5S rRNA mature molecules for the efficient synthesis of ribosomes (Prokopowich *et al.*, 2003). However, rDNA and 5S rDNA are transcribed by different RNA polymerases with dissimilar transcription rates and number of transcriptionally open rDNA units. Therefore, the rDNA/5S rDNA dosage is not directly related to the stoichiometry of the total rRNA pool and expression process. The rDNA/5S rDNA dosage varies widely throughout evolution. The processes that allow the maintenance of the pool of rRNA mature molecules in appropriate stoichiometry must be a complex network including epigenetic,

transcriptional, post-transcriptional and structural mechanisms that may vary according to the rDNA/5S rDNA dosage. Additionally, the location of the rDNA and 5S rDNA in the genome may be related to its expression and physiology. The chromosomal context, together with different chromatin environments, may be involved in the maintenance of gene copy number, recombination frequency, sequence conservation and transcription regulation of rDNA.

The organization of rDNA in extrachromosomal molecules may be associated with the cellular need for quick changes in the rDNA copy number under stress conditions. Some organisms have most of their rDNA in self-replicating extrachromosomal molecules, but retain additional copies in the chromosome probably as a backup. Interestingly, some organisms hold the totality of their rDNA in extrachromosomal molecules. It has been shown that the accumulation of extrachromosomal rDNA copies in old *S. cerevisiae* cultures affects cells' health. Therefore, cells that hold most or all rDNA copies extrachromosomally may have special mechanisms to allow for the accumulation of large rDNA minichromosomes without affecting the cell fitness. However, it is possible that yeasts lack this mechanism, resulting in cell damage when episomes accumulate.

Post-transcriptional processing of ITS-1 and ITS-2 is observed in almost all eukaryotes. However, some organisms possess additional ITS sequences in the 28S rRNA that generate fragmented rRNA molecules that maintain the core rRNA active elements in the mature ribosome. It has been found that some organisms possess additional sequences in variable regions internal to the SSU rRNA coding sequence that remain in the mature molecule, thus generating unusually large SSU rRNAs. It is interesting to note that the SSU rRNA has not been found as a fragmented molecule in nuclear genomes; in contrast, the fragmented eLSU rRNA could be regarded as 28S molecules that have processed their variable regions. The structure and functionality of these rRNAs in the ribosome may help to understand the importance of variable regions and the differences/restrictions between subunits.

The rDNA coding region of several microbial eukaryotes is interrupted by group I introns. Different transcriptional and post-transcriptional mechanisms are involved in the processing of introns and HE transcripts. Some *C. albicans* strains have heterogeneous rDNA populations with both intron-containing and intron-less rDNA units. Because no function related to rDNA expression has been proposed for group I introns, *C. albicans* may provide a good model to study the role (if any) of these introns in rRNA expression, processing and stability.

The linkage of 5S rDNA to a variety of tandem repeated families may be the result of homogenizing mechanisms responsible for concerted evolution (Drouin & de Sá, 1995). The finding that the 5S rDNA can be linked to all polymer-

ase transcribed genes, coded alone in different chromosomal loci or coded in extrachromosomal molecules underscores the possibility of various mechanisms acting to regulate the expression of different types of 5S rDNAs. The simultaneous expression of pol I, pol II and pol III transcribed genes in a particular locus may alter the chromatin context as well as the availability of transcription factors in the proximity. Nevertheless, the significance of the linkage between multi-genic families and 5S rDNA has not been studied. The presence of unlinked 5S rDNA copies is also interesting because the chromosomal context for each gene may affect its regulation.

Widespread rDNA characteristics present among eukaryotic supergroups as well as particular features predominating in some eukaryotic subgroups reflect the complexity of evolution. The typical tandem head-to-tail organization of the rDNA and 5S rDNA is found in all eukaryotic supergroups (Fig. 9), suggesting that the eukaryotic common ancestor held this organization. Later on, the evolutionary process probably led to a specialization and divergence of the rDNA structure, resulting in the different variants described here. Other features, such as the group I introns, were acquired by horizontal transfer and are therefore widespread among microbial eukaryotes (Fig. 9) (Sogin *et al.*, 1986; Van Oppen *et al.*, 1993).

Some particular rDNA characteristics are conserved among related species, suggesting that the common ancestor for each group held these traits before current speciation. Examples of this can be found in the unlinked differentially expressed rDNA units in *Apicomplexa*, the extralong SSU rRNAs in Amoebozoa and *Foraminifera*, the extrachromosomal rDNA and 5S rDNA in ciliates and Amoebozoa, the 5.8S–23S rRNA fusion in *Microsporidia* and both the SL–5S rDNA linkage and the eLSU fragmentation in *Euglenozoa*. The 5S rDNA nontandem organization as well as the linkage between 5S rDNA with the ribosomal cistron can also be seen as predominant in the Fungi group (Fig. 9). Particular traits such as the 28S rRNA fragmentation could have appeared more than once and independently, leading to non-*Euglenozoa* organisms containing fragmented rRNAs such as *A. castellanii* and *Plasmodium*. Distinctive characteristics shared among non-closely related species may represent phylogenetic evidence of yet unknown linkages among eukaryotic subgroups and species. Nevertheless, a thorough and integrated comparative characterization of rRNA genes in poorly or nonstudied eukaryotes may help to understand the diversity and relationship among different forms of life.

Many unanswered questions regarding the regulation of rRNA gene expression still remain, for example, the mechanism(s) that determine which rRNA gene copies will be transcriptionally and/or epigenetically active (Lawrence *et al.*, 2004; Grummt, 2007) or the relevance of the genomic context that surrounds the rRNA genes. Finally, it should be

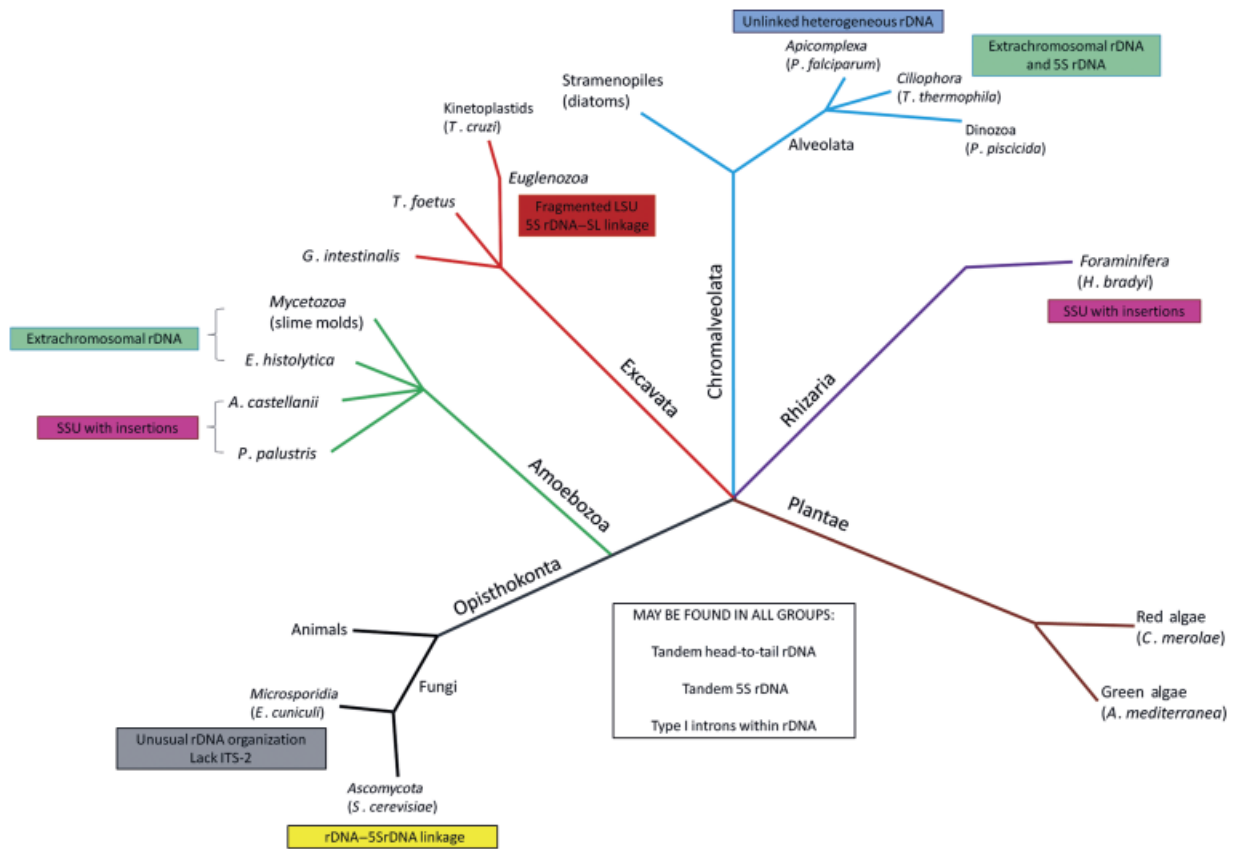


Fig. 9. Schematic phylogenetic tree describing the prevailing rDNA structures in subgroups of microbial eukaryotes. Phylogeny of eukaryotes is based on the six supergroup classification (Box 1). Conserved rDNA structures that predominate in eukaryotic subgroups are boxed in color. Some of the described characteristics may be shared by more than one subgroup.

pointed out that the rDNA organization is only one fundamental step in its regulation, because its expression is interrelated with most, if not all, of the cell's regulation levels (Paule & White, 2000; Schramm & Hernandez, 2002; Grummt, 2003).

Acknowledgements

This work was supported by grants IN214006 from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Universidad Nacional Autónoma de México (UNAM) and P45037-Q from Consejo Nacional de Ciencia y Tecnología (CONACYT), Mexico. A.L.T.-M. was supported by a scholarship from CONACYT Mexico.

References

- Acker J, Ozanne C, Kachouri-Lafond R, Gaillardin C, Neuveglise C & Marck C (2008) Diccistronic tRNA–5S rRNA genes in *Yarrowia lipolytica*: an alternative TFIIB-independent way for expression of 5S rRNA genes. *Nucleic Acids Res* **36**: 5832–5844.
- Adl SM, Simpson AG, Farmer MA *et al.* (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* **52**: 399–451.
- Aimi T, Yamada T & Murooka Y (1993) Group I self-splicing introns in both large and small subunit rRNA genes of *Chlorella*. *Nucleic Acids Symp Ser* 159–160.
- Aimi T, Yamada T & Murooka Y (1994) A group-I self-splicing intron in the nuclear small subunit rRNA-encoding gene of the green alga, *Chlorella ellipsoidea* C-87. *Gene* **139**: 65–71.
- Aksoy S (1992) Spliced leader RNA and 5S rRNA genes in *Herpetomonas* spp. are genetically linked. *Nucleic Acids Res* **20**: 913.
- Aksoy S, Shay GL, Villanueva MS, Beard CB & Richards FF (1992) Spliced leader RNA sequences of *Trypanosoma rangeli* are organized within the 5S rRNA-encoding genes. *Gene* **113**: 239–243.
- Allen SL, Ervin PR, McLaren NC & Brand RE (1984) The 5S ribosomal RNA gene clusters in *Tetrahymena thermophila*: strain differences, chromosomal localization, and loss during micronuclear ageing. *Mol Gen Genet* **197**: 244–253.

- Bagchi A, Bhattacharya A & Bhattacharya S (1999) Lack of a chromosomal copy of the circular rDNA plasmid of *Entamoeba histolytica*. *Int J Parasitol* **29**: 1775–1783.
- Bell GI, DeGennaro LJ, Gelfand DH, Bishop RJ, Valenzuela P & Rutter WJ (1977) Ribosomal RNA genes of *Saccharomyces cerevisiae*. I. Physical map of the repeating unit and location of the regions coding for 5 S, 5.8 S, 18 S, and 25 S ribosomal RNAs. *J Biol Chem* **252**: 8118–8125.
- Berger S, Zellmer DM, Kloppeck K, Richter G, Dillard WL & Schweiger HG (1978) Alternating polarity in rRNA genes. *Cell Biol Int Rep* **2**: 41–50.
- Bergeron J & Drouin G (2008) The evolution of 5S ribosomal RNA genes linked to the rDNA units of fungal species. *Curr Genet* **54**: 123–131.
- Berriman M, Ghedin E, Hertz-Fowler C *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**: 416–422.
- Bhattacharya S, Som I & Bhattacharya A (1998) The ribosomal DNA plasmids of *entamoeba*. *Parasitol Today* **14**: 181–185.
- Bidre C, Peyretailade E, Duffieux F, Peyret P, Méténier G & Vivarès C (1997) The rDNA unit of *Encephalitozoon cuniculi* (Microsporidia): complete 23S sequence and copy number. *J Eukaryot Microbiol* **44**: 76S.
- Bishop R, Gobright E, Spooner P, Allsopp B, Sohanpal B & Collins N (2000) Microsequence heterogeneity and expression of the LSU rRNA genes within the two single copy ribosomal transcription units of *Theileria parva*. *Gene* **257**: 299–305.
- Bogenhagen DE, Sakonju S & Brown DD (1980) A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region. *Cell* **19**: 27–35.
- Boothroyd JC & Dubremetz JF (2008) Kiss and spit: the dual roles of *Toxoplasma* rhoptries. *Nat Rev Microbiol* **6**: 79–88.
- Boothroyd JC, Wang A, Campbell DA & Wang CC (1987) An unusually compact ribosomal DNA repeat in the protozoan *Giardia lamblia*. *Nucleic Acids Res* **15**: 4065–4084.
- Brayton KA, Lau AO, Herndon DR *et al.* (2007) Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. *PLoS Pathog* **3**: 1401–1413.
- Brugère JF, Cornillot E, Méténier G, Bensimon A & Vivarès CP (2000) *Encephalitozoon cuniculi* (Microsporida) genome: physical map and evidence for telomere-associated rDNA units on all chromosomes. *Nucleic Acids Res* **28**: 2026–2033.
- Busse I & Preisfeld A (2002) Unusually expanded SSU ribosomal DNA of primary osmotrophic euglenids: molecular evolution and phylogenetic inference. *J Mol Evol* **55**: 757–767.
- Busse I & Preisfeld A (2003) Discovery of a group I intron in the SSU rDNA of *Ploetia costata* (Euglenozoa). *Protist* **154**: 57–69.
- Campbell DA (1992) *Bodo caudatus* medRNA and 5S rRNA genes: tandem arrangement and phylogenetic analyses. *Biochem Biophys Res Commun* **182**: 1053–1058.
- Campbell GR, Littau VC, Melera PW, Allfrey VG & Johnson EM (1979) Unique sequence arrangement of ribosomal genes in the palindromic rDNA molecule of *Physarum polycephalum*. *Nucleic Acids Res* **6**: 1433–1447.
- Carmo-Fonseca M, Mendes-Soares L & Campos I (2000) To be or not to be in the nucleolus. *Nat Cell Biol* **2**: E107–E112.
- Castro C, Hernández R & Castañeda M (1981) *Trypanosoma cruzi* ribosomal RNA: internal break in the large-molecular-mass species and number of genes. *Mol Biochem Parasit* **2**: 219–233.
- Cavalier-Smith T (2002) Nucleomorphs: enslaved algal nuclei. *Curr Opin Microbiol* **5**: 612–619.
- 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 Parasit* **52**: 75–83.
- Challoner PB, Amin AA, Pearlman RE & Blackburn EH (1985) Conserved arrangements of repeated DNA sequences in nontranscribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. *Nucleic Acids Res* **13**: 2661–2680.
- Clare JJ, Davidow LS, Gardner DC & Oliver SG (1986) Cloning and characterisation of the ribosomal RNA genes of the dimorphic yeast, *Yarrowia lipolytica*. *Curr Genet* **10**: 449–452.
- Clark CG & Cross GA (1987) rRNA genes of *Naegleria gruberi* are carried exclusively on a 14-kilobase-pair plasmid. *Mol Cell Biol* **7**: 3027–3031.
- Clark CG & Cross GA (1988) Circular ribosomal RNA genes are a general feature of schizopyrenid amoebae. *J Protozool* **35**: 326–329.
- Clark CG & Diamond LS (1991) Ribosomal RNA genes of 'pathogenic' and 'nonpathogenic' *Entamoeba histolytica* are distinct. *Mol Biochem Parasit* **49**: 297–302.
- Clark CG, Tague BW, Ware VC & Gerbi SA (1984) *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications. *Nucleic Acids Res* **12**: 6197–6220.
- Clark CG, Ali IK, Zaki M, Loftus BJ & Hall N (2006) Unique organisation of tRNA genes in *Entamoeba histolytica*. *Mol Biochem Parasit* **146**: 24–29.
- Cockburn AE, Taylor WC & Firtel RA (1978) *Dictyostelium* rDNA consists of non-chromosomal palindromic dimers containing 5S and 36S coding regions. *Chromosoma* **70**: 19–29.
- Conconi A, Widmer RM, Koller T & Sogo JM (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell* **57**: 753–761.
- Coss CA, Robledo JA, Ruiz GM & Vasta GR (2001) Description of *Perkinsus andrewni* n. sp. isolated from the Baltic clam (*Macoma balthica*) by characterization of the ribosomal RNA locus, and development of a species-specific PCR-based diagnostic assay. *J Eukaryot Microbiol* **48**: 52–61.
- Dacks JB, Walker G & Field MC (2008) Implications of the new eukaryotic systematics for parasitologists. *Parasitol Int* **57**: 97–104.
- D'Alessio JM, Harris GH, Perna PJ & Paule MR (1981) Ribosomal ribonucleic acid repeat unit of *Acanthamoeba castellanii*: cloning and restriction endonuclease map. *Biochemistry* **20**: 3822–3827.
- Dalrymple BP (1990) Cloning and characterization of the rRNA genes and flanking regions from *Babesia bovis*: use of the genes

- as strain discriminating probes. *Mol Biochem Parasit* **43**: 117–124.
- Dalrymple BP, Dimmock CM, Parrodi F & Wright IG (1992) *Babesia bovis*, *Babesia bigemina*, *Babesia canis*, *Babesia microti* and *Babesia rodhaini*: comparison of ribosomal RNA gene organization. *Int J Parasitol* **22**: 851–855.
- Dame JB & McCutchan TF (1983) The four ribosomal DNA units of the malaria parasite *Plasmodium berghei*. Identification, restriction map, and copy number analysis. *J Biol Chem* **258**: 6984–6990.
- da Silveira JF & Mercereau-Puijalón O (1983) *Plasmodium chabaudi* antigens synthesized in an mRNA-dependent cell-free translation system. *Mol Biochem Parasit* **7**: 159–172.
- Dávila-Aponte JA, Huss VA, Sogin ML & Cech TR (1991) A self-splicing group I intron in the nuclear pre-rRNA of the green alga, *Ankistrodesmus stipitatus*. *Nucleic Acids Res* **19**: 4429–4436.
- de Andrade Stempliuk V & Floeter-Winter LM (2002) Functional domains of the rDNA promoter display a differential recognition in *Leishmania*. *Int J Parasitol* **32**: 437–447.
- Doak TG, Cavalcanti AR, Stover NA, Dunn DM, Weiss R, Herrick G & Landweber LF (2003) Sequencing the *Oxytricha trifallax* macronuclear genome: a pilot project. *Trends Genet* **19**: 603–607.
- Dover GA (1988) Molecular evolution. rDNA world falling to pieces. *Nature* **336**: 623–624.
- Drouin G & de Sá MM (1995) The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. *Mol Biol Evol* **12**: 481–493.
- Dujon B, Sherman D, Fischer G et al. (2004) Genome evolution in yeasts. *Nature* **430**: 35–44.
- Edlind TD & Chakraborty PR (1987) Unusual ribosomal RNA of the intestinal parasite *Giardia lamblia*. *Nucleic Acids Res* **15**: 7889–7901.
- Eichinger L, Pachebat JA, Glöckner G et al. (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**: 43–57.
- Einvik C, Elde M & Johansen S (1998) Group I twintrons: genetic elements in myxomycete and schizopyrenid amoebiflagellate ribosomal DNAs. *J Biotechnol* **64**: 63–74.
- Eisen JA, Coyne RS, Wu M et al. (2006) Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol* **4**: e286.
- Engberg J (1985) The ribosomal RNA genes of *Tetrahymena*: structure and function. *Eur J Cell Biol* **36**: 133–151.
- Engberg J, Andersson P, Leick V & Collins J (1976) Free ribosomal DNA molecules from *Tetrahymena pyriformis* GL are giant palindromes. *J Mol Biol* **104**: 455–470.
- Erbezniak M, Yao MC & Jahn CL (1999) Characterization of the *Euplotes crassus* macronuclear rDNA and its potential as a DNA transformation vehicle. *J Eukaryot Microbiol* **46**: 206–216.
- Fang J & McCutchan TF (2002) Thermoregulation in a parasite's life cycle. *Nature* **418**: 742.
- Fang J, Sullivan M & McCutchan TF (2004) The effects of glucose concentration on the reciprocal regulation of rRNA promoters in *Plasmodium falciparum*. *J Biol Chem* **279**: 720–725.
- Fazaeli A, Carter PE & Pennington TH (2000) Intergenic spacer (IGS) polymorphism: a new genetic marker for differentiation of *Toxoplasma gondii* strains and *Neospora caninum*. *J Parasitol* **86**: 716–723.
- Figueroa-Angulo E, Cevallos AM, Zentella A, López-Villaseñor I & Hernández R (2006) Potential regulatory elements in the *Trypanosoma cruzi* rRNA gene promoter. *Biochim Biophys Acta* **1759**: 497–501.
- Findly RC & Gall JG (1978) Free ribosomal RNA genes in *Paramecium* are tandemly repeated. *P Natl Acad Sci USA* **75**: 3312–3316.
- Fischer JM, Keely SP & Stringer JR (2006) Evolutionary rate of ribosomal DNA in *Pneumocystis* species is normal despite the extraordinarily low copy-number of rDNA genes. *J Eukaryot Microbiol* **53** (suppl 1): S156–S158.
- Fournier P, Gaillardin C, Persuy MA, Klootwijk J & van Heerikhuizen H (1986) Heterogeneity in the ribosomal family of the yeast *Yarrowia lipolytica*: genomic organization and segregation studies. *Gene* **42**: 273–282.
- Frankel G, Cockburn AF, Kindle KL & Firtel RA (1977) Organization of the ribosomal RNA genes of *Dictyostelium discoideum*. Mapping of the transcribed region. *J Mol Biol* **109**: 539–558.
- French SL, Osheim YN, Cioci F, Nomura M & Beyer AL (2003) In exponentially growing *Saccharomyces cerevisiae* cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Mol Cell Biol* **23**: 1558–1568.
- Gagnon S, Bourbeau D & Levesque RC (1996) Secondary structures and features of the 18S, 5.8S and 26S ribosomal RNAs from the Apicomplexan parasite *Toxoplasma gondii*. *Gene* **173**: 129–135.
- Gardner MJ, Hall N, Fung E et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**: 498–511.
- Gardner MJ, Bishop R, Shah T et al. (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* **309**: 134–137.
- Gast RJ, Fuerst PA & Byers TJ (1994) Discovery of group I introns in the nuclear small subunit ribosomal RNA genes of *Acanthamoeba*. *Nucleic Acids Res* **22**: 592–596.
- Gatehouse HS & Malone LA (1998) The ribosomal RNA gene region of *Nosema apis* (Microspora): DNA sequence for small and large subunit rRNA genes and evidence of a large tandem repeat unit size. *J Invertebr Pathol* **71**: 97–105.
- Gay LS, Wilson ME & Donelson JE (1996) The promoter for the ribosomal RNA genes of *Leishmania chagasi*. *Mol Biochem Parasit* **77**: 193–200.
- Ghosh S, Satish S, Tyagi S, Bhattacharya A & Bhattacharya S (2003) Differential use of multiple replication origins in the ribosomal DNA episome of the protozoan parasite *Entamoeba histolytica*. *Nucleic Acids Res* **31**: 2035–2044.

- Greenwood SJ, Schnare MN, Cook JR & Gray MW (2001) Analysis of intergenic spacer transcripts suggests 'read-around' transcription of the extrachromosomal circular rDNA in *Euglena gracilis*. *Nucleic Acids Res* **29**: 2191–2198.
- Grummt I (2003) Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev* **17**: 1691–1702.
- Grummt I (2007) Different epigenetic layers engage in complex crosstalk to define the epigenetic state of mammalian rRNA genes. *Hum Mol Genet* **16** (Spec No 1): R21–R27.
- Guay JM, Huot A, Gagnon S, Tremblay A & Levesque RC (1992) Physical and genetic mapping of cloned ribosomal DNA from *Toxoplasma gondii*: primary and secondary structure of the 5S gene. *Gene* **114**: 165–171.
- 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.
- Gunderson JH & Sogin ML (1986) Length variation in eukaryotic rRNAs: small subunit rRNAs from the protists *Acanthamoeba castellanii* and *Euglena gracilis*. *Gene* **44**: 63–70.
- Haile S & Papadopoulou B (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol* **10**: 569–577.
- Hasan G, Turner MJ & Cordingley JS (1984) Ribosomal RNA genes of *Trypanosoma brucei*: mapping the regions specifying the six small ribosomal RNAs. *Gene* **27**: 75–86.
- Hattori M, Ljljana A & Sakaki Y (1984) Direct repeats surrounding the ribosomal RNA genes of *Physarum polycephalum*. *Nucleic Acids Res* **12**: 2047–2053.
- Healey A, Mitchell R, Upcroft JA, Boreham PF & Upcroft P (1990) Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from *Giardia intestinalis*. *Nucleic Acids Res* **18**: 4006.
- Hernández R, Díaz de León F & Castañeda M (1988) Molecular cloning and partial characterization of ribosomal RNA genes from *Trypanosoma cruzi*. *Mol Biochem Parasit* **27**: 275–279.
- Hernández R, Martínez-Calvillo S, Hernández-Rivas R & Gómez E (1993) *Trypanosoma cruzi* ribosomal RNA genes: a review. *Biol Res* **26**: 109–114.
- Hernández-Rivas R, Martínez-Calvillo S, Romero M & Hernández R (1992) *Trypanosoma cruzi* 5S rRNA genes: molecular cloning, structure and chromosomal organization. *FEMS Microbiol Lett* **71**: 63–67.
- Hinkle G, Leipe DD, Nerad TA & Sogin ML (1994) The unusually long small subunit ribosomal RNA of *Phreatamoeba balamuthi*. *Nucleic Acids Res* **22**: 465–469.
- Hofmann J, Winckler T, Hanenkamp A, Bukenberger M, Schumann G, Marschalek R & Dingermann T (1993) The *Dictyostelium discoideum* 5S rDNA is organized in the same transcriptional orientation as the other rDNAs. *Biochem Biophys Res Commun* **191**: 558–564.
- Hou G, Le Blancq SM, E Y, Zhu H & Lee MG (1995) Structure of a frequently rearranged rRNA-encoding chromosome in *Giardia lamblia*. *Nucleic Acids Res* **23**: 3310–3317.
- Huang WF, Tsai SJ, Lo CF, Soichi Y & Wang CH (2004) The novel organization and complete sequence of the ribosomal RNA gene of *Nosema bombycis*. *Fungal Genet Biol* **41**: 473–481.
- Huber D & Rustchenko E (2001) Large circular and linear rDNA plasmids in *Candida albicans*. *Yeast* **18**: 261–272.
- Huber M, Koller B, Gitler C, Mirelman D, Revel M, Rozenblatt S & Garfinkel L (1989) *Entamoeba histolytica* ribosomal RNA genes are carried on palindromic circular DNA molecules. *Mol Biochem Parasit* **32**: 285–296.
- Iiyama K, Chieda Y, Yasunaga-Aoki C, Hayasaka S & Shimizu S (2004) Analyses of the ribosomal DNA region in *Nosema bombycis* NIS 001. *J Eukaryot Microbiol* **51**: 598–604.
- Ivens AC, Peacock CS, Worthey EA *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**: 436–442.
- Jackson S, Cannone J, Lee J, Gutell R & Woodson S (2002) Distribution of rRNA introns in the three-dimensional structure of the ribosome. *J Mol Biol* **323**: 35–52.
- Jansson A, Gillin F, Kagardt U & Hagblom P (1994) Coding of hemolysins within the ribosomal RNA repeat on a plasmid in *Entamoeba histolytica*. *Science* **263**: 1440–1443.
- Jemtland R, Maehlum E, Gabrielsen OS & Oyen TB (1986) Regular distribution of length heterogeneities within non-transcribed spacer regions of cloned and genomic rDNA of *Saccharomyces cerevisiae*. *Nucleic Acids Res* **14**: 5145–5158.
- Johansen S & Vogt VM (1994) An intron in the nuclear ribosomal DNA of *Didymium iridis* codes for a group I ribozyme and a novel ribozyme that cooperate in self-splicing. *Cell* **76**: 725–734.
- Johansen S, Johansen T & Haugli F (1992) Extrachromosomal ribosomal DNA of *Didymium iridis*: sequence analysis of the large subunit ribosomal RNA gene and sub-telomeric region. *Curr Genet* **22**: 305–312.
- Johansen SD, Vader A, Sjøttem E & Nielsen H (2006) *In vivo* expression of a group I intron HEG from the antisense strand of *Didymium* ribosomal DNA. *RNA Biol* **3**: 157–162.
- Johansen SD, Haugen P & Nielsen H (2007) Expression of protein-coding genes embedded in ribosomal DNA. *Biol Chem* **388**: 679–686.
- Katinka MD, Duprat S, Cornillot E *et al.* (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414**: 450–453.
- Kats LM, Black CG, Proellocks NI & Coppel RL (2006) *Plasmodium* rhoptries: how things went pear-shaped. *Trends Parasitol* **22**: 269–276.
- Katz LA & Bhattacharya D (2006) *Genomics and Evolution of Microbial Eukaryotes*. 1st edn. Oxford University press, New York.
- Katzen AL, Cann GM & Blackburn EH (1981) Sequence-specific fragmentation of macronuclear DNA in a holotrichous ciliate. *Cell* **24**: 313–320.
- Keeling PJ & Slamovits CH (2004) Simplicity and complexity of microsporidian genomes. *Eukaryot Cell* **3**: 1363–1369.

- Keller M, Tessier LH, Chan RL, Weil JH & Imbault P (1992) In *Euglena*, spliced-leader RNA (SL-RNA) and 5S rRNA genes are tandemly repeated. *Nucleic Acids Res* **20**: 1711–1715.
- Ki JS & Han MS (2007) Cryptic long internal repeat sequences in the ribosomal DNA ITS1 gene of the dinoflagellate *Cochlodinium polykrikoides* (dinophyceae): a 101 nucleotide six-repeat track with a palindrome-like structure. *Genes Genet Syst* **82**: 161–166.
- Kibe MK, ole-MoiYoi OK, Nene V, Khan B, Allsopp BA, Collins NE, Morzaria SP, Gobright EI & Bishop RP (1994) Evidence for two single copy units in *Theileria parva* ribosomal RNA genes. *Mol Biochem Parasit* **66**: 249–259.
- Kim YH, Ishikawa D, Ha HP, Sugiyama M, Kaneko Y & Harashima S (2006) Chromosome XII context is important for rDNA function in yeast. *Nucleic Acids Res* **34**: 2914–2924.
- Kimmel AR & Gorovsky MA (1976) Numbers of 5S and tRNA genes in macro- and micronuclei of *Tetrahymena pyriformis*. *Chromosoma* **54**: 327–337.
- Kimmel AR & Gorovsky MA (1978) Organization of the 5S RNA genes in macro- and micronuclei of *Tetrahymena pyriformis*. *Chromosoma* **67**: 1–20.
- Klabunde J, Diesel A, Wasch D, Gellissen G, Hollenberg CP & Suckow M (2002) Single-step co-integration of multiple expressible heterologous genes into the ribosomal DNA of the methylotrophic yeast *Hansenula polymorpha*. *Appl Microbiol Biot* **58**: 797–805.
- Kobayashi T, Heck DJ, Nomura M & Horiuchi T (1998) Expansion and contraction of ribosomal DNA repeats in *Saccharomyces cerevisiae*: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. *Genes Dev* **12**: 3821–3830.
- Köck J & Cornelissen AW (1990) The 5S ribosomal RNA genes of *Crithidia fasciculata*. *Mol Biochem Parasit* **38**: 295–298.
- Kownin P, Iida CT, Brown-Shimer S & Paule MR (1985) The ribosomal RNA promoter of *Acanthamoeba castellanii* determined by transcription in a cell-free system. *Nucleic Acids Res* **13**: 6237–6248.
- Lansley G, Hyde JE, Goman M & Scaife JG (1983) Cloning and characterisation of the rRNA genes from the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res* **11**: 8703–8717.
- Lasker BA, Smith GW, Kobayashi GS, Whitney AM & Mayer LW (1998) Characterization of a single group I intron in the 18S rRNA gene of the pathogenic fungus *Histoplasma capsulatum*. *Med Mycol* **36**: 205–212.
- Laughery JM, Lau AO, White SN, Howell JM & Suarez CE (2009) *Babesia bovis*: transcriptional analysis of rRNA gene unit expression. *Exp Parasitol* **123**: 45–50.
- Lawrence RJ, Earley K, Pontes O, Silva M, Chen ZJ, Neves N, Viegas W & Pikaard CS (2004) A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol Cell* **13**: 599–609.
- Le Blancq SM, Korman SH & Van der Ploeg LH (1991) Frequent rearrangements of rRNA-encoding chromosomes in *Giardia lamblia*. *Nucleic Acids Res* **19**: 4405–4412.
- Le Blancq SM, Khramtsov NV, Zamani F, Upton SJ & Wu TW (1997) Ribosomal RNA gene organization in *Cryptosporidium parvum*. *Mol Biochem Parasit* **90**: 463–478.
- Lee Y, Wong WM, Guyer D, Erkin AM & Nazar RN (1997) *In vivo* analyses of upstream promoter sequence elements in the 5S rRNA gene from *Saccharomyces cerevisiae*. *J Mol Biol* **269**: 676–683.
- Lenardo MJ, Dorfman DM, Reddy LV & Donelson JE (1985) Characterization of the *Trypanosoma brucei* 5S ribosomal RNA gene and transcript: the 5S rRNA is a spliced-leader-independent species. *Gene* **35**: 131–141.
- León W, Fouts DL & Manning J (1978) Sequence arrangement of the 16S and 26S rRNA genes in the pathogenic haemoflagellate *Leishmania donovani*. *Nucleic Acids Res* **5**: 491–504.
- Li J, Gutell RR, Damberger SH, Wirtz RA, Kissinger JC, Rogers MJ, Sattabongkot J & McCutchan TF (1997) Regulation and trafficking of three distinct 18 S ribosomal RNAs during development of the malaria parasite. *J Mol Biol* **269**: 203–213.
- Lin H, Niu MT, Yoganathan T & Buck GA (1992) Characterization of the rRNA-encoding genes and transcripts, and a group-I self-splicing intron in *Pneumocystis carinii*. *Gene* **119**: 163–173.
- Lipps HJ & Steinbrück G (1978) Free genes for rRNAs in the macronuclear genome of the ciliate *Stylonychia mytilus*. *Chromosoma* **69**: 21–26.
- Liu H, Pan G, Song S, Xu J, Li T, Deng Y & Zhou Z (2008) Multiple rDNA units distributed on all chromosomes of *Nosema bombycis*. *J Invertebr Pathol* **99**: 235–238.
- Liu Y, Rocourt M, Pan S, Liu C & Leibowitz MJ (1992) Sequence and variability of the 5.8S and 26S rRNA genes of *Pneumocystis carinii*. *Nucleic Acids Res* **20**: 3763–3772.
- Loftus B, Anderson I, Davies R et al. (2005) The genome of the protist parasite *Entamoeba histolytica*. *Nature* **433**: 865–868.
- Long EO & Dawid IB (1980) Repeated genes in eukaryotes. *Annu Rev Biochem* **49**: 727–764.
- López-Villaseñor I, Contreras AP, López-Griego L, Álvarez-Sánchez E & Hernández R (2004) *Trichomonas vaginalis* ribosomal DNA: analysis of the intergenic region and mapping of the transcription start point. *Mol Biochem Parasit* **137**: 175–179.
- Mack SR, Samuels S & Vanderberg JP (1979) Hemolymph of *Anopheles stephensi* from noninfected and *Plasmodium berghei*-infected mosquitoes. 3. Carbohydrates. *J Parasitol* **65**: 217–221.
- Madison-Antenucci S, Grams J & Hajduk SL (2002) Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435–438.
- Maleszka R & Clark-Walker GD (1993) Yeasts have a four-fold variation in ribosomal DNA copy number. *Yeast* **9**: 53–58.
- Mandal RK (1984) The organization and transcription of eukaryotic ribosomal RNA genes. *Prog Nucleic Acid Res* **31**: 115–160.
- Margulis L & Schwartz KV (2000) *Five Kingdoms*. 3rd edn. W.H. Freeman and Company, New York.

- Martínez-Calvillo S, Sunkin SM, Yan S, Fox M, Stuart K & Myler PJ (2001) Genomic organization and functional characterization of the *Leishmania major* Friedlin ribosomal RNA gene locus. *Mol Biochem Parasit* **116**: 147–157.
- Maruyama S & Nozaki H (2007) Sequence and intranuclear location of the extrachromosomal rDNA plasmid of the amoeba-flagellate *Naegleria gruberi*. *J Eukaryot Microbiol* **54**: 333–337.
- Maruyama S, Misumi O, Ishii Y *et al.* (2004) The minimal eukaryotic ribosomal DNA units in the primitive red alga *Cyanidioschyzon merolae*. *DNA Res* **11**: 83–91.
- Maslov DA, Elgort MG, Wong S, Pecková H, Lom J, Simpson L & Campbell DA (1993) Organization of mini-exon and 5S rRNA genes in the kinetoplastid *Trypanoplasma borreli*. *Mol Biochem Parasit* **61**: 127–135.
- Matsuzaki M, Misumi O, Shin I *et al.* (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **428**: 653–657.
- McCutchan TF, Li J, McConkey GA, Rogers MJ & Waters AP (1995) The cytoplasmic ribosomal RNAs of *Plasmodium* spp. *Parasitol Today* **11**: 134–138.
- McMahon ME, Stamenkovich D & Petes TD (1984) Tandemly arranged variant 5S ribosomal RNA genes in the yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res* **12**: 8001–8016.
- Melville SE, Leech V, Gerrard CS, Tait A & Blackwell JM (1998) The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. *Mol Biochem Parasit* **94**: 155–173.
- Mercereau-Puijalon O, Barale JC & Bischoff E (2002) Three multigene families in *Plasmodium* parasites: facts and questions. *Int J Parasitol* **32**: 1323–1344.
- Michel F & Westhof E (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* **216**: 585–610.
- Miletti-González KE & Leibowitz MJ (2008) Molecular characterization of two types of rDNA units in a single strain of *Candida albicans*. *J Eukaryot Microbiol* **55**: 522–529.
- Miller OL Jr & Beatty BR (1969) Visualization of nucleolar genes. *Science* **164**: 955–957.
- Milyutina IA, Aleshin VV, Mikrjukov KA, Kedrova OS & Petrov NB (2001) The unusually long small subunit ribosomal RNA gene found in amitochondriate amoebaflagellate *Pelomyxa palustris*: its rRNA predicted secondary structure and phylogenetic implication. *Gene* **272**: 131–139.
- Mittal V, Sehgal D, Bhattacharya A & Bhattacharya S (1992) A second short repeat sequence detected downstream of rRNA genes in the *Entamoeba histolytica* rDNA episome. *Mol Biochem Parasit* **54**: 97–100.
- Moreno Díaz de la Espina S, Alverca E, Cuadrado A & Franca S (2005) Organization of the genome and gene expression in a nuclear environment lacking histones and nucleosomes: the amazing dinoflagellates. *Eur J Cell Biol* **84**: 137–149.
- Neigeborn L & Warner JR (1990) Expression of yeast 5S RNA is independent of the rDNA enhancer region. *Nucleic Acids Res* **18**: 4179–4184.
- Niles EG, Sutiphong J & Haque S (1981) Structure of the *Tetrahymena pyriformis* rRNA gene. Nucleotide sequence of the transcription initiation region. *J Biol Chem* **256**: 12849–12856.
- Nogi Y, Yano R & Nomura M (1991) Synthesis of large rRNAs by RNA polymerase II in mutants of *Saccharomyces cerevisiae* defective in RNA polymerase I. *P Natl Acad Sci USA* **88**: 3962–3966.
- Oakes M, Nogi Y, Clark MW & Nomura M (1993) Structural alterations of the nucleolus in mutants of *Saccharomyces cerevisiae* defective in RNA polymerase I. *Mol Cell Biol* **13**: 2441–2455.
- Okeke CN, Kappe R, Zakikhani S, Nolte O & Sonntag HG (1998) Ribosomal genes of *Histoplasma capsulatum* var. *duboisii* and var. *farcimosum*. *Mycoses* **41**: 355–362.
- O'Mahony EM, Tay WT & Paxton RJ (2007) Multiple rRNA variants in a single spore of the microsporidian *Nosema bombi*. *J Eukaryot Microbiol* **54**: 103–109.
- Orlando TC, Rubio MA, Sturm NR, Campbell DA & Floeter-Winter LM (2002) Intergenic and external transcribed spacers of ribosomal RNA genes in lizard-infecting *Leishmania*: molecular structure and phylogenetic relationship to mammal-infecting *Leishmania* in the subgenus *Leishmania* (*Leishmania*). *Mem I Oswaldo Cruz* **97**: 695–701.
- Paule MR & White RJ (2000) Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res* **28**: 1283–1298.
- Pecher WT, Robledo JA & Vasta GR (2004) Identification of a second rRNA gene unit in the *Perkinsus andrewsi* genome. *J Eukaryot Microbiol* **51**: 234–245.
- Peyretailade E, Biderre C, Peyret P, Duffieux F, Méténier G, Gouy M, Michot B & Vivarès CP (1998) Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. *Nucleic Acids Res* **26**: 3513–3520.
- Preer LB, Rudman B, Pollack S & Preer JR Jr (1999) Does ribosomal DNA get out of the micronuclear chromosome in *Paramecium tetraurelia* by means of a rolling circle? *Mol Cell Biol* **19**: 7792–7800.
- Prescott DM (2000) Genome gymnastics: unique modes of DNA evolution and processing in ciliates. *Nat Rev Genet* **1**: 191–198.
- Prokopowich CD, Gregory TR & Crease TJ (2003) The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**: 48–50.
- Pryde FE, Gorham HC & Louis EJ (1997) Chromosome ends: all the same under their caps. *Curr Opin Genet Dev* **7**: 822–828.
- Pulido M, Martínez-Calvillo S & Hernández R (1996) *Trypanosoma cruzi* rRNA genes: a repeated element from the non-transcribed spacer is locus specific. *Acta Trop* **62**: 163–170.
- Rae PM & Spear BB (1978) Macronuclear DNA of the hypotrichous ciliate *Oxytricha fallax*. *P Natl Acad Sci USA* **75**: 4992–4996.

- Ramezani-Rad M, Hollenberg CP, Lauber J *et al.* (2003) The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res* **4**: 207–215.
- Raska I, Koberna K, Malínský J, Fidlerová H & Masata M (2004) The nucleolus and transcription of ribosomal genes. *Biol Cell* **96**: 579–594.
- Raué HA, Klootwijk J & Musters W (1988) Evolutionary conservation of structure and function of high molecular weight ribosomal RNA. *Prog Biophys Mol Bio* **51**: 77–129.
- Ravel-Chapuis P (1988) Nuclear rDNA in *Euglena gracilis*: paucity of chromosomal units and replication of extrachromosomal units. *Nucleic Acids Res* **16**: 4801–4810.
- Reddy GR, Chakrabarti D, Yowell CA & Dame JB (1991) Sequence microheterogeneity of the three small subunit ribosomal RNA genes of *Babesia bigemina*: expression in erythrocyte culture. *Nucleic Acids Res* **19**: 3641–3645.
- Requena JM, Soto M, Quijada L, Carrillo G & Alonso C (1997) A region containing repeated elements is associated with transcriptional termination of *Leishmania infantum* ribosomal RNA genes. *Mol Biochem Parasit* **84**: 101–110.
- Ricard G, de Graaf RM, Dutilh BE *et al.* (2008) Macronuclear genome structure of the ciliate *Nyctotherus ovalis*: single-gene chromosomes and tiny introns. *BMC Genomics* **9**: 587.
- 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.
- Rogers MJ, Gutell RR, Damberger SH, Li J, McConkey GA, Waters AP & McCutchan TF (1996) Structural features of the large subunit rRNA expressed in *Plasmodium falciparum* sporozoites that distinguish it from the asexually expressed subunit rRNA. *RNA* **2**: 134–145.
- Rubin GM & Sulston JE (1973) Physical linkage of the 5 S cistrons to the 18 S and 28 S ribosomal RNA cistrons in *Saccharomyces cerevisiae*. *J Mol Biol* **79**: 521–530.
- Ruoff B, Johansen S & Vogt VM (1992) Characterization of the self-splicing products of a mobile intron from the nuclear rDNA of *Physarum polycephalum*. *Nucleic Acids Res* **20**: 5899–5906.
- Rustchenko EP & Sherman F (1994) Physical constitution of ribosomal genes in common strains of *Saccharomyces cerevisiae*. *Yeast* **10**: 1157–1171.
- Saito K, Drgon T, Robledo JA, Krupatkina DN & Vasta GR (2002) Characterization of the rRNA locus of *Pfiesteria piscicida* and development of standard and quantitative PCR-based detection assays targeted to the nontranscribed spacer. *Appl Environ Microb* **68**: 5394–5407.
- Sakonju S, Bogenhagen DF & Brown DD (1980) A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. *Cell* **19**: 13–25.
- Santana DM, Lukes J, Sturm NR & Campbell DA (2001) Two sequence classes of kinetoplastid 5S ribosomal RNA gene revealed among bodonid spliced leader RNA gene arrays. *FEMS Microbiol Lett* **204**: 233–237.
- Schaak J, Mao J & Söll D (1982) The 5.8S RNA gene sequence and the ribosomal repeat of *Schizosaccharomyces pombe*. *Nucleic Acids Res* **10**: 2851–2864.
- Schnare MN, Cook JR & Gray MW (1990) Fourteen internal transcribed spacers in the circular ribosomal DNA of *Euglena gracilis*. *J Mol Biol* **215**: 85–91.
- Schnare MN, Collings JC, Spencer DF & Gray MW (2000) The 28S–18S rDNA intergenic spacer from *Crithidia fasciculata*: repeated sequences, length heterogeneity, putative processing sites and potential interactions between U3 small nucleolar RNA and the ribosomal RNA precursor. *Nucleic Acids Res* **28**: 3452–3461.
- Schramm L & Hernandez N (2002) Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**: 2593–2620.
- Schroeder-Diedrich JM, Fuerst PA & Byers TJ (1998) Group-I introns with unusual sequences occur at three sites in nuclear 18S rRNA genes of *Acanthamoeba lenticulata*. *Curr Genet* **34**: 71–78.
- Sehgal D, Mittal V, Ramachandran S, Dhar SK, Bhattacharya A & Bhattacharya S (1994) Nucleotide sequence organisation and analysis of the nuclear ribosomal DNA circle of the protozoan parasite *Entamoeba histolytica*. *Mol Biochem Parasit* **67**: 205–214.
- Shi X, Chen DH & Suyama Y (1994) A nuclear tRNA gene cluster in the protozoan *Leishmania tarentolae* and differential distribution of nuclear-encoded tRNAs between the cytosol and mitochondria. *Mol Biochem Parasit* **65**: 23–37.
- Shippen-Lentz DE & Vezza AC (1988) The three 5S rRNA genes from the human malaria parasite *Plasmodium falciparum* are linked. *Mol Biochem Parasit* **27**: 263–273.
- Shirley MW (2000) The genome of *Eimeria* spp., with special reference to *Eimeria tenella*-a coccidium from the chicken. *Int J Parasitol* **30**: 485–493.
- Simpson AG & Roger AJ (2004) The real 'kingdoms' of eukaryotes. *Curr Biol* **14**: R693–R696.
- Simpson AG, Stevens JR & Lukes J (2006) The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol* **22**: 168–174.
- Sinclair DA & Guarente L (1997) Extrachromosomal rDNA circles – a cause of aging in yeast. *Cell* **91**: 1033–1042.
- Skryabin KG, Eldarov MA, Larionov VL, Bayev AA, Klootwijk J, de Regt VC, Veldman GM, Planta RJ, Georgiev OI & Hadjiolov AA (1984) Structure and function of the nontranscribed spacer regions of yeast rDNA. *Nucleic Acids Res* **12**: 2955–2968.
- Sogin ML & Edman JC (1989) A self-splicing intron in the small subunit rRNA gene of *Pneumocystis carinii*. *Nucleic Acids Res* **17**: 5349–5359.
- Sogin ML, Ingold A, Karlok M, Nielsen H & Engberg J (1986) Phylogenetic evidence for the acquisition of ribosomal RNA introns subsequent to the divergence of some of the major *Tetrahymena* groups. *EMBO J* **5**: 3625–3630.

- Sollner-Webb B & Mougey EB (1991) News from the nucleolus: rRNA gene expression. *Trends Biochem Sci* **16**: 58–62.
- Spano F & Crisanti A (2000) *Cryptosporidium parvum*: the many secrets of a small genome. *Int J Parasitol* **30**: 553–565.
- Spencer DF, Collings JC, Schnare MN & Gray MW (1987) Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata*. *EMBO J* **6**: 1063–1071.
- Spring H, Grierson D, Hemleben V, Stöhr M, Krohne G, Stadler J & Franke WW (1978) DNA contents and numbers of nucleoli and pre-rRNA-genes in nuclei of gametes and vegetative cells of *Acetabularia mediterranea*. *Exp Cell Res* **114**: 203–215.
- Srivastava AK & Schlessinger D (1991) Structure and organization of ribosomal DNA. *Biochimie* **73**: 631–638.
- Stucki U, Braun R & Roditi I (1993) *Eimeria tenella*: characterization of a 5S ribosomal RNA repeat unit and its use as a species-specific probe. *Exp Parasitol* **76**: 68–75.
- Sturm NR, Maslov DA, Grisard EC & Campbell DA (2001) *Diplonema* spp. possess spliced leader RNA genes similar to the Kinetoplastida. *J Eukaryot Microbiol* **48**: 325–331.
- Sucgang R, Chen G, Liu W, Lindsay R, Lu J, Muzny D, Shaulsky G, Loomis W, Gibbs R & Kuspa A (2003) Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in *Dictyostelium*. *Nucleic Acids Res* **31**: 2361–2368.
- Swanton MT, McCarroll RM & Spear BB (1982) The organization of macronuclear rDNA molecules of four hypotrichous ciliated protozoans. *Chromosoma* **85**: 1–9.
- Tabata S (1980) Structure of the 5-S ribosomal RNA gene and its adjacent regions in *Torulopsis utilis*. *Eur J Biochem* **110**: 107–114.
- Tabata S (1981) Nucleotide sequences of the 5S ribosomal RNA genes and their adjacent regions in *Schizosaccharomyces pombe*. *Nucleic Acids Res* **9**: 6429–6437.
- 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.
- Tang X, Bartlett MS, Smith JW, Lu JJ & Lee CH (1998) Determination of copy number of rRNA genes in *Pneumocystis carinii* f. sp. hominis. *J Clin Microbiol* **36**: 2491–2494.
- Torres-Machorro AL, Hernández R, Sánchez J & López-Villaseñor I (2006) The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis*. *Mol Biochem Parasit* **145**: 269–273.
- Torres-Machorro AL, Hernández R, Alderete JF & López-Villaseñor I (2009) Comparative analyses among the *Trichomonas vaginalis*, *Trichomonas tenax*, and *Tritrichomonas foetus* 5S ribosomal RNA genes. *Curr Genet* **55**: 199–210.
- Tsai SJ, Huang WF & Wang CH (2005) Complete sequence and gene organization of the *Nosema spodopterae* rRNA gene. *J Eukaryot Microbiol* **52**: 52–54.
- Uliana SR, Fischer W, Stempliuk VA & Floeter-Winter LM (1996) Structural and functional characterization of the *Leishmania amazonensis* ribosomal RNA promoter. *Mol Biochem Parasit* **76**: 245–255.
- Unnasch TR & Wirth DF (1983) The avian malaria *Plasmodium lophurae* has a small number of heterogeneous ribosomal RNA genes. *Nucleic Acids Res* **11**: 8443–8459.
- Upcroft JA, Healey A, Mitchell R, Boreham PF & Upcroft P (1990) Antigen expression from the ribosomal DNA repeat unit of *Giardia intestinalis*. *Nucleic Acids Res* **18**: 7077–7081.
- Upcroft JA, Healey A & Upcroft P (1994) A new rDNA repeat unit in human *Giardia*. *J Eukaryot Microbiol* **41**: 639–642.
- Upcroft JA, Abedinia M & Upcroft P (2005) Rearranged subtelomeric rRNA genes in *Giardia duodenalis*. *Eukaryot Cell* **4**: 484–486.
- Vader A, Nielsen H & Johansen S (1999) *In vivo* expression of the nucleolar group I intron-encoded I-dirI homing endonuclease involves the removal of a spliceosomal intron. *EMBO J* **18**: 1003–1013.
- Van de Peer Y, Baldauf SL, Doolittle WF & Meyer A (2000) An updated and comprehensive rRNA phylogeny of (crown) eukaryotes based on rate-calibrated evolutionary distances. *J Mol Evol* **51**: 565–576.
- van Heerikhuizen H, Ykema A, Klootwijk J, Gaillardin C, Ballas C & Fournier P (1985) Heterogeneity in the ribosomal RNA genes of the yeast *Yarrowia lipolytica*; cloning and analysis of two size classes of repeats. *Gene* **39**: 213–222.
- van Keulen H, Gutell RR, Campbell SR, Erlandsen SL & Jarroll EL (1992) The nucleotide sequence of the entire ribosomal DNA operon and the structure of the large subunit rRNA of *Giardia muris*. *J Mol Evol* **35**: 318–328.
- Van Oppen MJ, Olsen JL & Stam WT (1993) Evidence for independent acquisition of group I introns in green algae. *Mol Biol Evol* **10**: 1317–1326.
- van Spaendonk RM, Ramesar J, Janse CJ & Waters AP (2000) The rodent malaria parasite *Plasmodium berghei* does not contain a typical O-type small subunit ribosomal RNA gene. *Mol Biochem Parasit* **105**: 169–174.
- Velichutina IV, Rogers MJ, McCutchan TF & Liebman SW (1998) Chimeric rRNAs containing the GTPase centers of the developmentally regulated ribosomal rRNAs of *Plasmodium falciparum* are functionally distinct. *RNA* **4**: 594–602.
- Verbeet MP, van Heerikhuizen H, Klootwijk J, Fontijn RD & Planta RJ (1984) Evolution of yeast ribosomal DNA: molecular cloning of the rDNA units of *Kluyveromyces lactis* and *Hansenula wingei* and their comparison with the rDNA units of other Saccharomycetoideae. *Mol Gen Genet* **195**: 116–125.
- Vogt VM & Braun R (1976) Structure of ribosomal DNA in *Physarum polycephalum*. *J Mol Biol* **106**: 567–587.
- Vossbrinck CR & Woese CR (1986) Eukaryotic ribosomes that lack a 5.8S RNA. *Nature* **320**: 287–288.
- Wai HH, Vu L, Oakes M & Nomura M (2000) Complete deletion of yeast chromosomal rDNA repeats and integration of a new rDNA repeat: use of rDNA deletion strains for functional analysis of rDNA promoter elements *in vivo*. *Nucleic Acids Res* **28**: 3524–3534.

- Waters AP (1994) The ribosomal RNA genes of *Plasmodium*. *Adv Parasitol* **34**: 33–79.
- Waters AP, Syin C & McCutchan TF (1989) Developmental regulation of stage-specific ribosome populations in *Plasmodium*. *Nature* **342**: 438–440.
- Weiss LM (2001) Microsporidia: emerging pathogenic protists. *Acta Trop* **78**: 89–102.
- Wikmark OG, Einvik C, De Jonckheere JF & Johansen SD (2006) Short-term sequence evolution and vertical inheritance of the *Naegleria* twin-ribozyme group I intron. *BMC Evol Biol* **6**: 39.
- Wilcox LW, Lewis LA, Fuerst PA & Floyd GL (1992) Group I introns within the nuclear-encoded small-subunit rRNA gene of three green algae. *Mol Biol Evol* **9**: 1103–1118.
- Wild MA & Gall JG (1979) An intervening sequence in the gene coding for 25S ribosomal RNA of *Tetrahymena pigmentosa*. *Cell* **16**: 565–573.
- Wood V, Gwilliam R, Rajandream MA *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**: 871–880.
- Yan S, Lodes MJ, Fox M, Myler PJ & Stuart K (1999) Characterization of the *Leishmania donovani* ribosomal RNA promoter. *Mol Biochem Parasit* **103**: 197–210.
- Yang Q, Zwick MG & Paule MR (1994) Sequence organization of the *Acanthamoeba* rRNA intergenic spacer: identification of transcriptional enhancers. *Nucleic Acids Res* **22**: 4798–4805.
- Yao MC & Gall JG (1977) A single integrated gene for ribosomal RNA in a eucaryote, *Tetrahymena pyriformis*. *Cell* **12**: 121–132.
- Zindrou S, Orozco E, Linder E, Téllez A & Björkman A (2001) Specific detection of *Entamoeba histolytica* DNA by hemolysin gene targeted PCR. *Acta Trop* **78**: 117–125.
- Zwick MG, Wiggs M & Paule MR (1991) Sequence and organization of 5S RNA genes from the eukaryotic protist *Acanthamoeba castellanii*. *Gene* **101**: 153–157.