Published online May 9, 2006

2472–2481 *Nucleic Acids Research, 2006, Vol. 34, No. 8*
doi:10.1093/nar/gkl327

Complete DNA sequences of the mitochondrial genomes of the pathogenic yeasts *Candida orthopsilosis* and *Candida metapsilosis*: insight into the evolution of linear DNA genomes from mitochondrial telomere mutants

Peter Kosa[1], Matus Valach[1], Lubomir Tomaska[2], Kenneth H. Wolfe[3] and Jozef Nosek[1,*]

[1]Department of Biochemistry and [2]Department of Genetics, Faculty of Natural Sciences, Comenius University, Mlynska dolina CH-1 and B-1, 842 15, Bratislava, Slovak Republic and [3]Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland

Received January 13, 2006; Revised February 7, 2006; Accepted April 13, 2006

ABSTRACT

We determined complete mitochondrial DNA sequences of the two yeast species, *Candida orthopsilosis* and *Candida metapsilosis*, and compared them with the linear mitochondrial genome of their close relative, *C.parapsilosis*. Mitochondria of all the three species harbor compact genomes encoding the same set of genes arranged in the identical order. Differences in the length of these genomes result mainly from the presence/absence of introns. Multiple alterations were identified also in the sequences of the ribosomal and transfer RNAs, and proteins. However, the most striking feature of *C.orthopsilosis* and *C.metapsilosis* is the existence of strains differing in the molecular form of the mitochondrial genome (circular-mapping versus linear). Their analysis opens a unique window for understanding the role of mitochondrial telomeres in the stability and evolution of molecular architecture of the genome. Our results indicate that the circular-mapping mitochondrial genome derived from the linear form by intramolecular end-to-end fusions. Moreover, we suggest that the linear mitochondrial genome evolved from a circular-mapping form present in a common ancestor of the three species and, at the same time, the emergence of mitochondrial telomeres enabled the formation of linear monomeric DNA forms. In addition, comparison of isogenic *C.metapsilosis* strains differing in the form of the organellar genome suggests a possibility that, under some circumstances, the linearity and/or the presence of telomeres provide a competitive advantage over a circular-mapping mitochondrial genome.

INTRODUCTION

Mitochondrial genomes vary in size, gene content as well as their molecular form. The latter is represented by a number of structures such as monogenomic circular molecules typical for animals (1), concatenated networks of circles seen in kinetoplastids (2), concatemeric linear molecules heterogeneous in size displaying circular physical maps detected in most of fungal and plant species (3,4) and linear molecules with defined terminal structures observed in a range of phylogenetically diverse taxa (5). Interestingly, the form of mitochondrial genome may differ in closely related organisms or even within the same species (6). However, the biological significance of different molecular forms of the organellar genome remains obscure.

Recently, we determined the complete mitochondrial DNA (mtDNA) sequence of the pathogenic yeast *Candida parapsilosis* (7). This mitochondrial genome is represented by 30.9 kb linear double-stranded DNA molecules terminating with arrays of tandem repeats of a 738 bp unit termed mitochondrial telomeres. The very end of the molecule terminates with a single stranded 5′ overhang of ~110 nt (8,9). Our previous studies indicated that the ends of *C.parapsilosis* linear mtDNA molecules are protected by mitochondrial telomere-binding protein (mtTBP) (8–10) and telomeric loop (t-loop) structures (11). In addition to linear DNA molecules, two-dimensional gel electrophoresis and electron microscopy

*To whom correspondence should be addressed. Tel: +421 2 60296 536; Fax: +421 2 60296 452; Email: nosek@fns.uniba.sk

demonstrated that *C.parapsilosis* mitochondria harbor series of double-stranded circular DNA molecules derived from the sequence of the mitochondrial telomere thus forming integral multimers of the 738 bp unit that we named telomeric circles (t-circles) (12,13). The t-circles were shown to replicate independently of the linear mtDNA via a rolling-circle strategy thus generating arrays of telomeric repeats that may recombine with the ends of the linear mtDNA molecules to extend their termini. As mitochondria of *C.parapsilosis* represent a natural telomerase deficient system, the t-circles seem to provide a telomerase-independent mean of the telomere maintenance (14). Importantly, the t-circles were recently identified also in human cells maintaining their chromosomal termini via alternative telomere lengthening (ALT) mode pointing to their involvement in a roll-and-spread mechanism of nuclear telomere elongation (15). This indicates that exploration of the mitochondrial system may lead to results of a general significance. Importantly, the evolutionary emergence of the linear genomes in yeast mitochondria was recently proposed as a paradigm for the origin of linear chromosomes in nuclei of eukaryotes in the pre-telomerase era (16). Studies of the molecular architecture of the mtDNA and mitochondrial telomere dynamics thus may have implications for understanding the nature and evolution of alternative, telomerase-independent, pathway(s) of telomere maintenance of eukaryotic chromosomes.

With the aim of understanding how a mitochondrial genome with such a unique molecular architecture emerged in evolution we studied two yeasts, *Candida orthopsilosis* and *Candida metapsilosis*, closely related to *C.parapsilosis*. These taxa, originally classified as groups II and III of *C.parapsilosis* (17,18), were recently established as independent species (19). In this work, we determined the complete mtDNA sequences of *C.orthopsilosis* MCO456 and *C.metapsilosis* MCO448. Their comparative analysis makes it possible to study the evolution of genome organization in the three closely related organisms and provides a basis for identification of conserved elements within non-coding intergenic regions that can be implicated in the control of replication and gene expression. Moreover, we analyzed a set of clinical isolates that lack mitochondrial t-circles and harbor circular-mapping mtDNA (20). Our study sheds light on the biological role of mitochondrial telomeres and exemplifies the formation of circular-mapping mtDNA from originally a linear genome via intramolecular end-to-end fusion.

MATERIALS AND METHODS

Yeast strains

Strains of *C.orthopsilosis* (MCO456, MCO457, MCO462, MCO471, PL452) and *C.metapsilosis* (MCO448, PL448) were kindly provided by P. F. Lehmann (Medical College of Ohio, Toledo, USA) and S. A. Meyer (Georgia State University, Atlanta, USA). The *C.parapsilosis* strain SR23/CBS7157 is from the collection of the Department of Biochemistry (Comenius University, Bratislava). Strains designated as CBS were from Centraalbureau voor Schimmelcultures (Utrecht, The Netherlands). Yeasts were grown at 28°C in YPD medium [1% (w/v) yeast extract (Difco), 1% (w/v) Bacto Peptone (Difco), 2% (w/v) glucose, pH was adjusted to 6.3 with 50 mM potassium phosphate buffer].

Isolation and analysis of mtDNA

DNA was isolated from mitochondria of both species and purified by CsCl/bisbenzimide gradient centrifugation (21). Random libraries of mtDNA fragments were prepared by nebulization essentially as described by (22). The DNA sequence was determined by dideoxy-chain termination method on double-stranded plasmid templates by AGOWA (Berlin, Germany). The sequence assembly and analysis was done using the Vector NTI Advance v. 8.0 software package (Informax, Inc.). Secondary structures of ribosomal RNAs (rRNA) were inferred from models elaborated for their equivalents from *Escherichia coli* and mitochondria of *Saccharomyces cerevisiae*, using BioEdit (23) and RnaViz (24) programs. Cumulative GC skew analyses were performed using the Genome Skew software v. 1.0 (Technical University of Munich, Germany). Telomere junctions of the strains with circularized mtDNA were either cloned as HindIII restriction fragments or amplified by PCR using the primers 5′-TTAGCTGTTGTTGCTATTACT-3′ and 5′-ATTGTTGCTTTTGTTGTTGA-3′. Separation of mtDNA by pulsed-field gel electrophoresis (PFGE) was performed as described previously (20).

Phylogenetic analysis

Phylogenetical relationships among yeast species were analyzed by both the neighbor-joining (NJ) and maximum-likelihood (ML) methods. Concatenated amino acid sequences of proteins encoded by mtDNAs (i.e. *atp6-atp8-atp9-cob-cox1-cox2-cox3*) were aligned using ClustalW 1.81 (25) and the alignment was manually edited to remove gap regions, and 1689 aligned positions were used in the analysis. ML and NJ analyses were performed by PHYML (26) and MEGA 3.1 package (27), respectively. We used the Jones–Taylor–Thorton matrix for amino acid substitution and Gamma distribution model of site variation automatically estimated by maximizing the likelihood of the phylogeny. The trees were tested by 500 bootstrap replicates. The mtDNAs of following species were involved in the analysis: *Ashbya gossypii* (AE016821), *Candida albicans* (AF285261), *C.parapsilosis* (X74411), *C.orthopsilosis* (AY962590), *C.metapsilosis* (AY962591), *Candida glabrata* (AJ511533), *Hanseniaspora uvarum* (DQ058142), *Kluyveromyces lactis* (AY654900), *Kluyveromyces thermotolerans* (AJ634268), *Pichia canadensis* (D31785), *Saccharomyces castellii* (AF437291), *S.cerevisiae* (AJ011856), *Saccharomyces servazzii* (AJ430679), *Schizosaccharomyces pombe* (X54421), *Schizosaccharomyces octosporus* (AF275271), *Schizosaccharomyces japonicus* (AF547983) and *Yarrowia lipolytica* (AJ307410). The sequence of *Candida tropicalis* mtDNA was downloaded from *C.tropicalis* Sequencing Project (Broad Institute of Harvard and MIT; http://www.broad.mit.edu).

Miscellaneous

Enzymatic manipulations with DNA, cloning procedures, Southern and northern blot analyses were performed essentially as described by Sambrook and Russell (28). Following

oligonucleotides were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and used for the northern hybridizations: 5′-CTATCATAAATAGTAACACCAGGAGCA-3′ (*nad1*); 5′-ATAGAAGCAATACCACCAGTTACAAAATA-3′ (*nad2*); 5′-GATTGTTGGAAAGAAGTGAAACCACA-3′ (*nad3*); 5′-CCWGCTAATGATAATAATTCACCTATAAA-G-3′ (*nad4*); 5′-ATAGTATGACCWACTTTCCCTCTTAA-TC-3′ (*nad4L*); 5′-GCACAAACCATAGTAGCAGCATG-3′ (*nad5*); 5′-CAATAGCCCCTATAACAGAAAGAATTAAA A-3′ (*nad6*); 5′-TAACTAAACATAAACCTAATAAAGAA-CCAACAT-3′ (*cob*); 5′-CCATAGCACTAATCATACCA-TAGATAAG-3′ (*cox1*); 5′-CTAATTGCTCAAGGTGTTGG-CAC-3′ (*cox2*); 5′-GTATGATCTCCTAAATATGTACTTT-CAGC-3′ (*cox3*); 5′-AAACAACAGCAACTAATTGAGCA-GAAATA-3′ (*atp6*); 5′-GCTAAAATACCACCWGTTAA-TAAATTCATAAAG-3′ (*atp8*); and 5′-ACCGATAGCAGC - ACCACCTAAA-3′ (*atp9*).

Sequences of the D1/D2 domain of cytosolic ribosomal RNA were determined on PCR products amplified from total genomic DNA using NL-1 (5′-GCATATCAATAAGC-GGAGGAAAAG-3′) and NL-4 (5′-GGTCCGTGTTTCAA-GACGG-3′) primers (29). Open reading frames (ORFs) encoding the mitochondrial telomere-binding protein (mtTBP) homologs were amplified from genomic DNA template using the oligonucleotide primers 5′-ATGTTGC-GAGCATTCACTAGATCA-3′ and 5′-TTCTGTAGCTTCG - GCTCTATCCTCA-3′.

Nucleotide sequence accession numbers

The DNA sequences described in this article were deposited in the GenBank data library under the accession numbers: AY962590 (the mtDNA of *C.orthopsilosis* MCO456), AY962591 (the mtDNA of *C.metapsilosis* MCO448), AY986471 (right telomeric fragment of *C.orthopsilosis* MCO471), DQ026513 and DQ376035 (*cox1* region of MCO471 and CLIB214, respectively), AY986472, AY986473, AY986474, AY391853 (telomere fusion fragments between the genes *atp6* and *nad3* of *C.orthopsilosis* MCO457, MCO462, PL452 and *C.metapsilosis* PL448, respectively), DQ213055–DQ213057 (sequences of the D1/D2 domain of cytosolic 26S rRNA of *C.parapsilosis* SR23/CBS7157, *C.orthopsilosis* MCO456 and *C.metapsilosis* MCO448 strains).

RESULTS AND DISCUSSION

Genetic organization of the mitochondrial genomes of *C.orthopsilosis* and *C.metapsilosis*

We determined the complete mtDNA sequences of *C.orthopsilosis* MCO456 and *C.metapsilosis* MCO448. Their analyses confirm that, similar to the *C.parapsilosis* mtDNA (7), these genomes are highly compact and encode the same set of genes arranged in the identical order, which is conserved even at the level of individual transfer RNA (tRNA) genes (Figure 1 and Table 1). Analysis of protein coding sequences revealed that only the sequence of *nad4L* is invariant. In other cases, we identified minor changes mostly representing substitutions without any significant change in the length of the ORFs. In contrast, the *nad4* reading frame varies in length at its 3′ end.

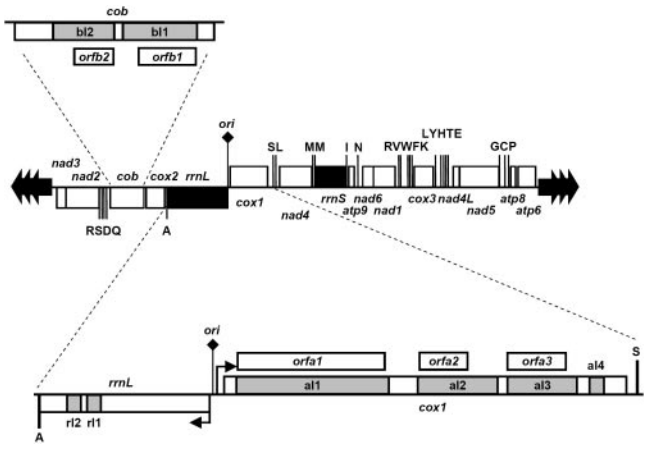

Figure 1. Linear mitochondrial genomes of the closely related yeasts *C.parapsilosis, C.orthopsilosis*, and *C.metapsilosis*. The gene order is conserved in all three species. Their mtDNAs differ in the presence/absence of introns in the *cob, cox1* and *rrnL* loci (*C.parapsilosis* strains SR23 and CLIB214 contain aI1-aI3, bI1, bI2, rI1 and rI2; *C.orthopsilosis* MCO456 has aI4 and bI2 and MCO471 has aI3, aI4 and probably also bI2; *C.metapsilosis* MCO448 mtDNA possesses solely bI1). Positions of genes for tRNAs are indicated by single letter codes of their cognate amino acids, potential origin of mtDNA replication (*ori*) was predicted by cumulative GC skew analysis (Supplementary Figure 1). Triangles at both termini represent arrays of telomeric repeats. Note the strains of *C.orthopsilosis* MCO456, MCO457, MCO462, PL452 and *C.metapsilosis* PL448 lack telomeric arrays and contain a circularized form of the mitochondrial genome (see Figure 5 for more details).

Mitochondrial rRNA sequences of the three yeast species display a number of differences. In the case of the small subunit (SSU) rRNA, the most significant alteration affects helix 46 [numbering from *E.coli* SSU rRNA; (30)]. In the large subunit (LSU) rRNA, multiple short insertions within the sequences were found. Importantly, comparison of LSU rRNA revealed two short (266 and 265 nt) group I intronic sequences within the *C.parapsilosis* gene for mitochondrial LSU rRNA that are absent in its counterparts from *C.orthopsilosis* and *C.metapsilosis*.

Each of the three mitochondrial genomes contains the same set of 24 tRNA genes. In most cases their sequences are identical in all three species, or differ only by a single nucleotide in one species. In one case, *trnS1*, the sequence differs in a single position among all three species. However, *trnN* is exceptional because there are multiple alterations within the dihydrouridine loop and the anticodon stem of the *C.parapsilosis* tRNA, which differs from both *C.orthopsilosis* and *C.metapsilosis* by nine substitutions and a single insertion of uridine.

The size of the mitochondrial genome in both *C.orthopsilosis* and *C.metapsilosis* represents about two-thirds of the *C.parapsilosis* mtDNA. To a large extent, the variation in the mtDNA length is caused by differences in intron content. *C.parapsilosis* mtDNA contains seven introns whose total length corresponds to almost one-third of the genome: two introns in *cob*, three in *cox1* (7) and two in *rrnL* gene. In contrast, the shorter mitochondrial genomes of *C.metapsilosis* strain MCO448 and *C.orthopsilosis* strain MCO456 have only one and two introns, respectively, representing in both cases only ~6% of the genome size. Interestingly, the *C.metapsilosis* mtDNA contains an

Table 1. Comparison of the mitochondrial genomes of *C.parapsilosis*, *C.orthopsilosis* and *C.metapsilosis*

Species/strain	mtDNA length	mtDNA form	Presence of t-circles[a]	Presumed telomere junction[b] (bp)	G + C contents (%)	No. of introns	No. of intronic ORFs	Introns versus total mtDNA (%)	Non-coding intergenic spacers (%)
C.parapsilosis									
SR23/CBS7157[c]	30 923 + 2nx 738 bp	Linear	nx 738 bp	—	24.3	7	5	28.9	7.6
CLIB214[d,e]	~31 000 + 2nx ~750 bp	Linear	nx 0.75 kbp	—		≥5	5		
C.orthopsilosis									
MCO456[c]	22 528 bp	Circular[f]	—	944	25.0	2	1	6.3	12.7
MCO457[d]	~25 000 bp	Circular[f]	—	820					
MCO462	~25 000 bp	Circular[f]	—	808					
MCO471	~25 000 bp + 2nx 777 bp	Linear	nx 777 bp	—		≥2	≥1		
PL452	~25 000 bp	Circular[f]	—	349					
C.metapsilosis									
MCO448[c,d,g]	23 062 + 2nx 620 bp	Linear	nx 620 bp	—	25.1	1	1	6.2	12.1
PL448[g]	~23 000 bp	Circular[f]	—	747		1	1		
PL429	~23 000 +2nx ~600 bp	Linear	nx 0.6 kb	—					
CBS2916	~23 000 +2nx ~600 bp	Linear	nx 0.6 kb	—					

[a]As determined by (12,14,20), the t-circles exist in multimeric forms.
[b]The size of non-coding region between the genes *atp6* and *nad3* in the circular-mapping mtDNAs.
[c]The complete mtDNA sequence was determined by (7) and this study.
[d]The type strain of the species.
[e]Partial sequence representing ~90% of the mitochondrial genome is available (44), it is 99.94% identical to the sequence from the strain SR23/CBS7157.
[f]Circular-mapping mtDNA as determined by restriction enzymes mapping, PFGE and DNA sequence analysis [(20), this study].
[g]PL448 is an isogenic strain of MCO448, they differ in the mtDNA form.

intronless allele of *cox1* that is rather exceptional among fungi (31). In most cases, the introns contain reading frames encoding putative RNA maturases, reverse transcriptases or DNA endonucleases.

Processing of mitochondrial transcripts involves excision of tRNA and the 3′ end signal

The coding sequences of the three mitochondrial genomes are arranged into two transcriptional units *rrnL–nad3* and *cox1–atp6*, both transcribed from a region corresponding to a putative origin of replication (*ori*) predicted using cumulative GC skew analysis between the genes *rrnL* and *cox1* (Figure 1 and Supplementary Figure 1). Results of northern blot analysis indicate that sizes of mRNA in mitochondria of all three species correspond to mono- and bi-cistronic products presumably processed from large polycistronic primary transcripts. The distribution of *trn* sequences in the mtDNAs suggests that tRNA punctuation mode of transcript processing [reviewed in (32)] may operate in the three yeast species. The occurrence of bi-cistronic mRNAs in the cases of *nad4L-nad5, nad6-nad1*, *nad2-nad3* and *atp8-atp6,* where the protein coding sequences are not separated by tRNAs, is compatible with this possibility. The only exceptions are the genes *cox2* and *cob* which are not separated by a tRNA gene but whose corresponding mRNAs have different sizes. *In silico* analysis of the mtDNA of all three yeast species revealed a conserved motif 5′-UAW-WUAUUCYUUW-3′ related to the *S.cerevisiae* dodecanucleotide consensus 5′-AAUAAYAUUCUU-3′ implicated in the 3′ end processing of mitochondrial transcripts (33,34). The motif occurs downstream of the protein coding genes *cob, cox1, cox2, cox3, atp6, atp9, nad1* and *nad3*. This indicates that maturation of mitochondrial transcripts may require excision of tRNA sequences and/or the presence of the 3′ end processing signal (Figure 2 and Supplementary Figure 2).

Mitochondrial telomeres of the three species have conserved architecture

In spite of similar genetic organization, the molecular forms of the three sequenced mitochondrial genomes are different. While mitochondria of both *C.parapsilosis* SR23 and *C.metapsilosis* MCO448 harbor linear DNA genomes, *C.orthopsilosis* MCO456 possesses a circular-mapping mtDNA. Importantly, another *C.orthopsilosis* strain MCO471, contains a linear mitochondrial genome. In order to compare mitochondrial telomeres from all three yeast species, we cloned and sequenced the telomeric region of the right arm of MCO471. Comparison of mitochondrial telomeres of all three yeast species revealed that they consist of terminal inverted repeats themselves consisting of a subterminal repeat and an array of tandem repeat units (Figure 3). The sequence corresponding to the 5′ single-stranded overhang at the termini of the *C.parapsilosis* mtDNA is conserved in both *C.orthopsilosis* MCO471 and *C.metapsilosis* MCO448 suggesting that mitochondrial telomeres in all three species have essentially the same organization. In addition, the subterminal repeat of the *C.parapsilosis* mtDNA contains a characteristic motif rich in guanine and cytosine residues named GC-box. This sequence is conserved in the subtelomeric regions of linear mtDNAs of *C.orthopsilosis* MCO471 and *C.metapsilosis* MCO448, downstream of the genes *atp6* (right arm) and *nad3* (left arm). It was also identified in the region between *atp6* and *nad3* in the circular-mapping mtDNAs of *C.orthopsilosis* MCO456. Although the biological role of this element is not known, it may play a role in transcription termination and/or be implicated in recombination-dependent mode of mitochondrial telomere maintenance (7).

Circular-mapping genomes in mitochondria of *C.orthopsilosis* and *C.metapsilosis*

To investigate differences between linear and circular-mapping genomes, we systematically investigated the mtDNA

form among strains of *C.parapsilosis*, *C.orthopsilosis* and *C.metapsilosis*. The PFGE analysis revealed a discrete band migrating at ~49 kb in all strains harboring linear mtDNA (Figure 4A). It is not clear why the size of linear molecules deduced from their electrophoretic mobility is larger than expected from its DNA sequence. However, a similar observation was made by Maleszka (35) in the *C.parapsilosis* strain SR23/CBS7157. In any case, all strains previously shown to have circular mtDNA restriction enzyme maps (20) contain, instead of the discrete band (form I), mtDNA molecules heterogeneous in size (Figure 4A and B). This electrophoretic profile of mtDNA molecules is similar to the pattern reported in baker's yeast mitochondria and termed polydisperse mtDNA (4,36). Two additional faint bands (forms II and III) (Figure 4A, lanes with MCO456 and PL448) were detected in strains containing circular-mapping mtDNAs. Since these bands were not present in all preparations from the same strain they may correspond to transient mtDNA replication intermediates of the circular-mapping genome.

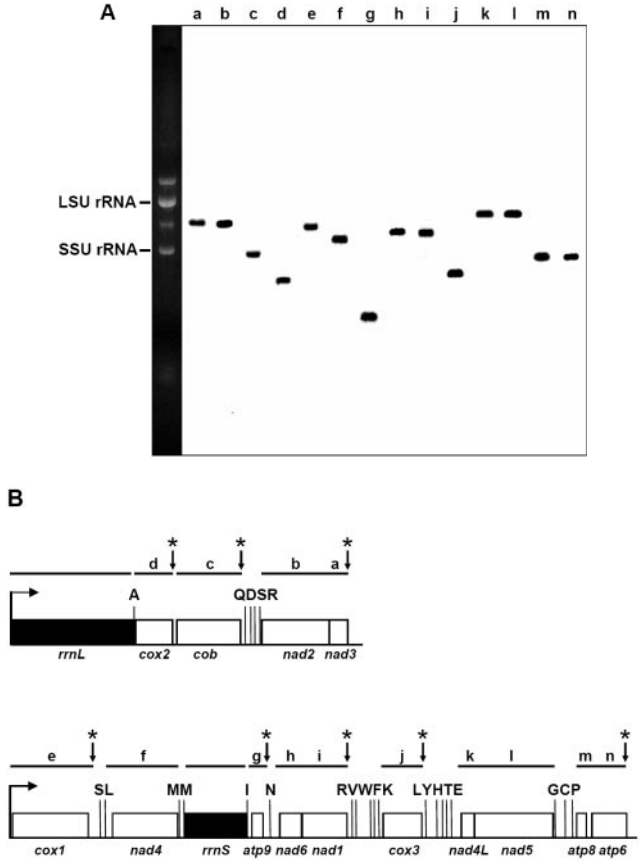

Figure 2. Mitochondrial transcripts from *C.parapsilosis, C.orthopsilosis* and *C.metapsilosis* (**A**) Northern blots with RNA isolated from mitochondria of the *C.metapsilosis* strain were probed with radiolabeled oligonucleotides (see Materials and Methods section) complementary to the corresponding sequences; a, *nad3*; b, *nad2*; c, *cob*; d, *cox2*; e, *cox1*; f, *nad4*; g, *atp9*; h, *nad6*; i, *nad1*; j, *cox3*; k, *nad4L*; l, *nad5*; m, *atp8*; n, *atp6*. The transcripts from *C.parapsilosis* SR23 and *C.orthopsilosis* MCO456 show essentially the same pattern (Supplementary Figure 2). (**B**) Summary map of the left and the right polycistronic mitochondrial transcripts, conserved in all three yeast species. Asterisks indicate the conserved position of the 3′ end processing signal. Sequences corresponding to tRNAs are labeled by single letter codes corresponding to their cognate amino acids.

Comparison of the circular-mapping mtDNA of *C.metapsilosis* PL448 with the linear genome found in mitochondria of the strain MCO448 suggests that the former lacks telomeric arrays and contains a restriction fragment corresponding to the fusion of the left and right telomere (Figure 4C). Essentially the same results were obtained by comparison of the *C.orthopsilosis* strains MCO456 (circular-mapping) and MCO471 (linear). This raises a possibility that emergence of the circular-mapping mtDNA in both species may be due to a defect in the mitochondrial telomere maintenance accompanied by the fusion of mtDNA termini.

As circular-mapping mtDNAs were identified in several isolates, we analyzed presumed telomeric junctions in the strains of *C.orthopsilosis* MCO457, MCO462 and PL452 and *C.metapsilosis* PL448 (Table 1). The corresponding region was either amplified by PCR using primers derived from the sequence of the left (*nad3*) and the right (*atp6*) subterminal gene and/or cloned as a HindIII restriction fragment. Although the junction sites differ among strains, DNA sequence analysis revealed in all cases stretches of telomeric sequence confirming that this region originated from telomere–telomere fusions accompanied by deletion of a significant portion of the telomeric sequence (Figure 5). In strains MCO456, MCO457, MCO462 and PL448, the fusion of the left and right arms of the mtDNA occurred at short identical A+T-rich stretches of sequence suggesting that the circularization may have been mediated by homologous recombination between short repetitive motifs present within the telomeric sequence. In one circular-mapping genome, *C.orthopsilosis* PL452, no repeat was apparent and instead two extra bases (TT) are present at the junction site.

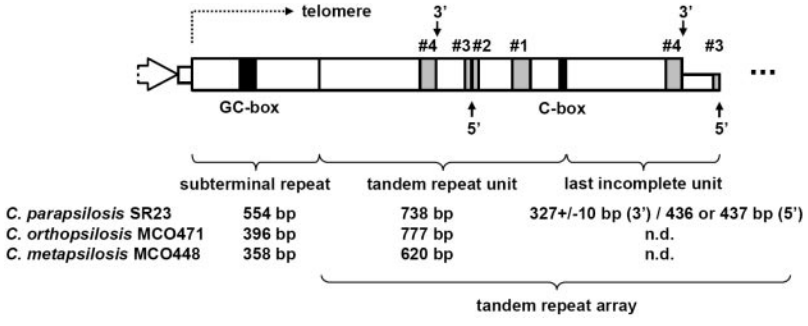

Figure 3. Conserved structure of mitochondrial telomeres. Guanine and cytosine (GC-box) and cytosine-rich sequences (C-box) are shown as indicated. Palindromic sequences (labeled as #1–4) are shaded.

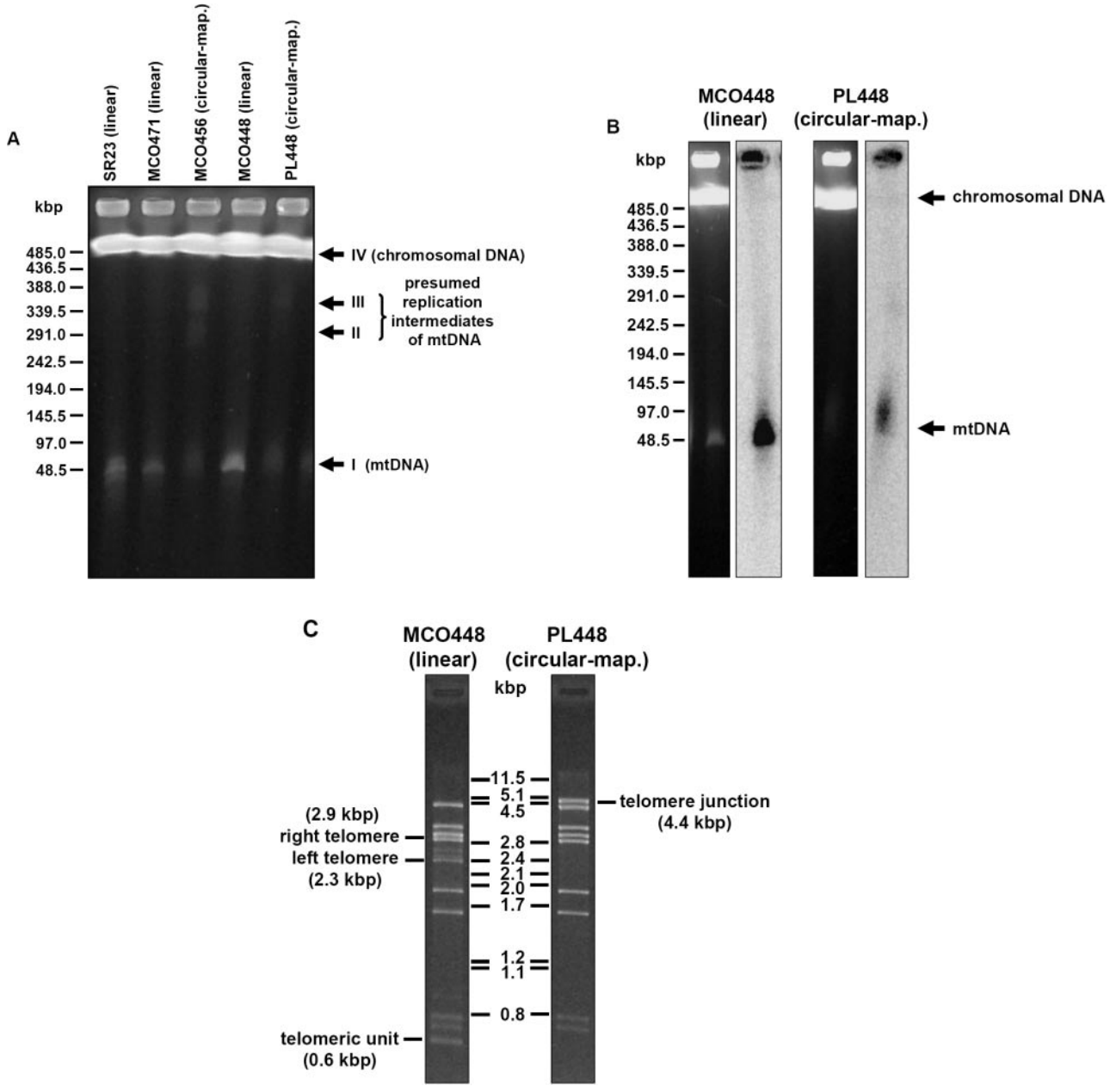

Figure 4. Comparison of strains harboring linear and circular-mapping mtDNA. (**A**) Electrophoretic profiles of DNA isolated from strains containing linear (*C.parapsilosis* SR23, *C.orthopsilosis* MCO471, *C.metapsilosis* MCO448) and circular-mapping (*C.orthopsilosis* MCO456, *C.metapsilosis* PL448) mtDNA. Arrows indicate different DNA forms: I, linear mtDNA; II and III, presumed replication intermediates present in the strains harboring circular-mapping mtDNA; IV, chromosomal DNA. (**B**) Southern blot hybridization analysis of PFGE separated mtDNA from *C.metapsilosis* strains MCO448 and PL448. Blots were probed with radioactively labeled mtDNA from the strain MCO448. (**C**) Profiles of EcoRV digested mtDNA from *C.metapsilosis* strains MCO448 and PL448.

Formation of circular derivatives of the linear mitochondrial genome might result from a defect in key telomere function(s) such as telomere capping or telomere replication. We hypothesize that a defect in the processing of the ends of the mtDNA molecules might have prevented formation of t-loops and the resulting structure may have been recognized as a double-stranded break repaired by end-to-end fusion resulting in circularization of the genome. Alternatively, alterations in mtTBP or another, yet unknown, telomere-binding factor might be responsible for changes in the molecular form of the mtDNA. To test the latter possibility we analyzed the coding sequences of mtTBP homologs from *C.orthopsilosis* (MCO456, MCO471), *C.metapsilosis* (MCO448, PL448) and *C.parapsilosis* (SR23, CBS604). We identified only minor sequence variations between the species. Importantly, strains of the same species differing in the forms of the mitochondrial genome do not differ in the sequence of this protein (Supplementary Figure 3) suggesting that the mtTBP-mediated capping function is not impaired.

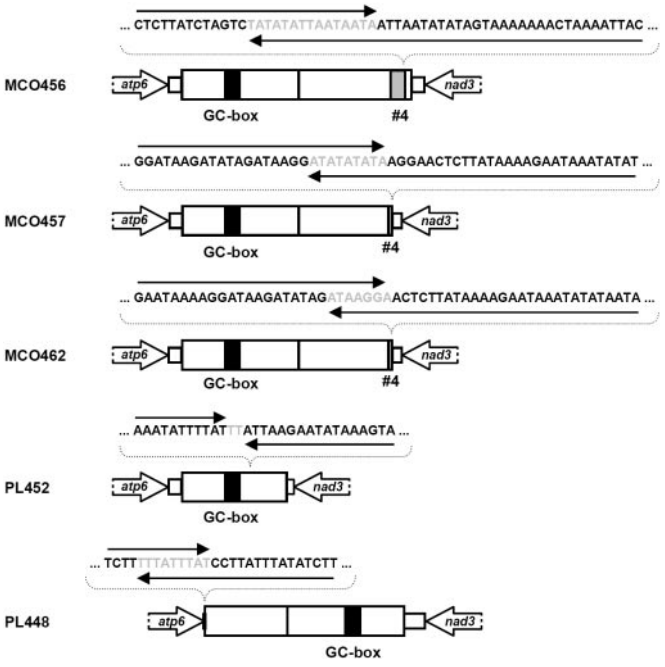

Figure 5. Region between *nad3* and *atp6* genes of the strains harboring circular-mapping mtDNAs illustrating presumed intramolecular end-to-end fusions. Arrows over and below sequence indicate the homology with the left and right telomere, respectively. GC-box and the conserved palindrome #4 (see Figure 3) are shown.

Previously we demonstrated that strains with circular-mapping mtDNA lack t-circles (20). At present we cannot formally rule out the possibility that their absence results from end-to-end fusion accompanied by deletion of tandem repeats (i.e. the generation of t-circles may be dependent on intramolecular recombination between repeat units within the terminal tandem arrays). On the other hand, the absence of t-circles may cause the genome circularization. For example, their elimination from the organelle might have an effect on the telomere maintenance pathway and result in the formation of circularized form of the genome via end-to-end fusions. A support for the scenario of DNA circularization in the absence of functional telomere maintenance comes from example described in nuclei of fission yeasts (37,38). In this case, most of cells surviving a defect in the telomerase-dependent pathway escape the problem of telomere synthesis and genome stability by circularization of all three nuclear chromosomes. Since the t-circles were proposed as important players in the telomerase-independent mode of telomere maintenance in both, mitochondrial and nuclear, systems (13), our results may be relevant for understanding the alternative telomere replication mechanisms in eukaryotic nuclei.

Phylogenetic relationships among *C.orthopsilosis, C.metapsilosis* and *C.parapsilosis*

In previous studies, Diezmann *et al.* (39) and Tavanti *et al.* (19) employing different sets of sequences (variable as well as conserved) for analysis placed *C.parapsilosis* as an outgroup to the *C.orthopsilosis–C.metapsilosis* clade. However, large datasets derived from complete mitochondrial genome sequences may produce more reliable phylogenies. Therefore,

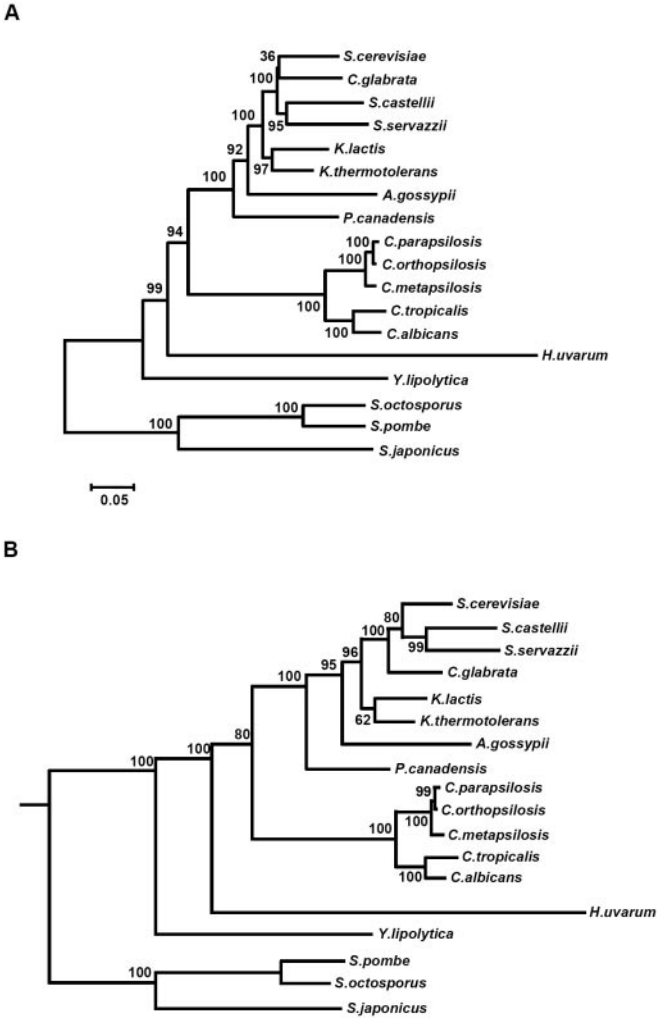

Figure 6. Phylogenetic relationship among species in the *C.parapsilosis* clade inferred from the alignments of concatenated sequences of proteins encoded by the mtDNAs (i.e. *atp6-atp8-atp9-cob-cox1-cox2-cox3*). The trees were calculated by (**A**) NJ and (**B**) ML methods. The numbers above or below branches indicate the percentage with which a given branch appeared in 500 bootstrap replicates.

we employed the concatenated amino acid sequences of mtDNA encoded proteins to analyze the relationships between the three *Candida* species. We constructed phylogenetic trees by both the NJ and ML methods (Figure 6). Our analyses clearly place the three species as a sister branch of the clade containing *C.albicans* and *C.tropicalis*. In both trees, the branching of these species exhibits the same order and indicates with high confidence (bootstrap values are 99–100%) that *C.metapsilosis* separated from a common ancestor prior the split of *C.orthopsilosis* and *C.parapsilosis*. The phylogeny based on the sequences of the D1/D2 domain of the 26S rRNA indicates the same tree topology, although in this case the statistical support is considerably lower (data not shown).

Evolution of the genome structure

Most of fungal and plant species contain circular-mapping mitochondrial genome which is represented *in vivo* mostly

by linear DNA concatemers and a minor fraction of circular DNA species (3,36). Recent studies on *S.cerevisiae* mitochondria revealed an asymmetry in segregation of mtDNA molecules during cell division. While daughter cells receive a linear monomer of the genome that circularizes and amplifies via rolling-circle mechanism, mother cells retain mostly concatemeric and branched mtDNA forms (40,41). We propose that mitochondrial telomeres may facilitate the resolution of monomeric genomes thus providing an advantage in the process of asymmetric segregation. The incorporation of mitochondrial telomeres into the replication strategy of an ancestral circular-mapping mtDNA might have caused the resolution of linear monomeric genome terminating with the telomeric arrays from a population of concatemeric mtDNA molecules. This is in agreement with the hypothesis that linear and circular-mapping mitochondrial genomes may interconvert via a relatively simple molecular mechanism (42). It also implies that the linear mitochondrial genome of *C.parapsilosis, C.orthopsilosis* and *C.metapsilosis*, and probably also *Candida salmanticensis*, *Pichia philodendri* (21) and *Hanseniaspora uvarum* (43), which possess the linear form of mtDNA terminating with structurally similar telomeres, might have evolved relatively recently from circular-mapping ancestors.

The phylogenetic analysis suggests that the common ancestor of *C.orthopsilosis, C.metapsilosis* and *C.parapsilosis* had a linear genome, and that the circular maps observed in some strains of *C.orthopsilosis* and *C.metapsilosis* are the result of recent strain-specific telomere mutations that circularized them. However, the linear structure of the genome of the common ancestor of these species can itself be inferred to be a relatively recent innovation, because the mtDNAs of their closest relative *Lodderomyces elongisporus*, as well as other species from this phylogenetic branch (i.e. *C.tropicalis, Candida maltosa, Candida viswanathii, Candida lodderae* and *C.albicans*) have circular-mapping mtDNA (P. Kosa, M. Valach, L. Tomaska, K. H. Wolfe and J. Nosek, unpublished data). This suggests that the linear form of the genome and its specific terminal structures evolved *de novo* after separation of the branch leading to the three *Candida*species.

Does the linearity of mitochondrial genome provide a selective advantage?

The strain PL448 of *C.metapsilosis* originates from the clinical isolate MCO448 and both are considered to be isogenic (S. A. Meyer, personal communication). Our results described above demonstrate they differ in the mtDNA form (Figure 4) suggesting that circular-mapping mtDNA in PL448 is a mutant derivative of the linear genome found in the strain MCO448.

The availability of these strains opens a possibility to investigate whether linearity and/or the presence of telomeric structures at the ends of mtDNA molecules offers a selective advantage over strains harboring a circular-mapping mtDNA. First, we compared the growth of *C.metapsilosis* strains MCO448 (linear) and PL448 (circular-mapping) in YPD medium. Our results indicate that cells MCO448 outgrows PL448 during exponential phase in batch cultures suggesting that the linear form of the mitochondrial genome may provide an advantage during cellular proliferation (Supplementary Figure 4). Moreover, we observed that, probably due to a difference in the growth rate, the strain MCO448 forms colonies with sizes of ~1–2 mm significantly earlier than PL448.

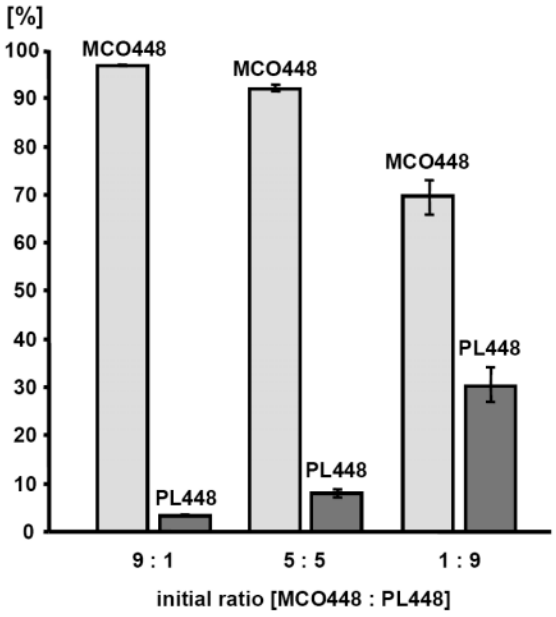

Figure 7. The *C.metapsilosis* MCO448 cells harboring the linear genome in mitochondria are able to outgrow isogenic strain PL448 with circular-mapping mtDNA. Yeast cultures MCO448 and PL448 were mixed in 9:1, 5:5 and 1:9 ratio, and cultivated in YPD medium for 24 h. Cells were then plated onto solid YPD medium and the number of MCO448 and PL448 subclones was counted. To discriminate between both strains we used the PCR approach described previously (20) as well as the difference in the colony size (mentioned in the text). Results are based on the three independent experiments. Note that after three passages of the mixed 9:1, 5:5 and 1:9 cultures, the number of PL448 cells was 0.3, 1.2 and 1.9 %, respectively.

To demonstrate that MCO448 is able to out-compete PL448 in a mixed culture, we combined the cells from late exponential phase (~5×10^8 cells/ml) of both strains in three different ratios (i.e. 9:1, 5:5 and 1:9) and cultivated them in the YPD medium. The cultures were then diluted and plated onto solid YPD media and the proportion of MCO448 and PL448 subclones was determined (Figure 7). These results demonstrate that MCO448 out-competes PL448 suggesting that the linear form of the mtDNA may provide an advantage for MCO448 cells.

The hypothesis that the linearity and/or the presence of telomeric structures at the ends of linear mtDNA provide a selective advantage needs to be addressed in more detail. Nevertheless, the results described above indicate that, under certain circumstances, linear mtDNA may increase the host's fitness. This raises a possibility that the selective advantage of a linear genome is not limited solely to yeast mitochondria but is a more general phenomenon that applies also to other linear DNA genomes including the chromosomes of eukaryotic nuclei.

SUPPLEMENTARY DATA

Supplementary data are available at *NAR* Online.

ACKNOWLEDGEMENTS

The authors wish to thank L. Kovac (Comenius University, Bratislava) for continuous support and helpful comments, S. A. Meyer (Georgia State University, Atlanta) for discussion, and members of our laboratories for discussions and/or technical assistance. This work was supported by grants from the Howard Hughes Medical Institute (55000327), the Fogarty International Research Collaboration Award (2-R03-TW005654-04A1), the Slovak grant agencies VEGA (1/2331/05, 1/0006/03), APVT (20-003902, 20-001604), the Comenius University (UK/181/2005) and Science Foundation Ireland. Funding to pay the Open Access publication charges for this article was provided by the grant from the Howard Hughes Medical Institute (55005622).

Conflict of interest statement. None declared.

REFERENCES

1. Boore,J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Res*., **27**, 1767–1780.
2. Lukes,J., Jirku,M., Avliyakulov,N. and Benada,O. (1998) Pankinetoplast DNA structure in a primitive bodonid flagellate, *Cryptobia helicis*. *EMBO J*., **17**, 838–846.
3. Bendich,A.J. (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr. Genet*., **24**, 279–290.
4. Williamson,D. (2002) The curious history of yeast mitochondrial DNA. *Nature Rev. Genet*., **3**, 475–481.
5. Nosek,J., Tomaska,L., Fukuhara,H., Suyama,Y. and Kovac,L. (1998) Linear mitochondrial genomes: 30 years down the line. *Trends Genet*., **14**, 184–188.
6. Nosek,J. and Tomaska,L. (2003) Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Curr. Genet*., **44**, 73–84.
7. Nosek,J., Novotna,M., Hlavatovicova,Z., Ussery,D.W., Fajkus,J. and Tomaska,L. (2004) Complete DNA sequence of the linear mitochondrial genome of the pathogenic yeast *Candida parapsilosis*. *Mol. Genet. Genomics*, **272**, 173–180.
8. Tomaska,L., Nosek,J. and Fukuhara,H. (1997) Identification of a putative mitochondrial telomere-binding protein of the yeast *Candida parapsilosis*. *J. Biol. Chem*., **272**, 3049–3056.
9. Nosek,J., Tomaska,L., Pagacova,B. and Fukuhara,H. (1999) Mitochondrial telomere-binding protein from *Candida parapsilosis* suggests an evolutionary adaptation of a nonspecific single-stranded DNA-binding protein. *J. Biol. Chem*., **274**, 8850–8857.
10. Tomaska,L., Makhov,A.M., Nosek,J., Kucejova,B. and Griffith,J.D. (2001) Electron microscopic analysis supports a dual role for the mitochondrial telomere-binding protein of *Candida parapsilosis*. *J. Mol. Biol*., **305**, 61–69.
11. Tomaska,L., Makhov,A.M., Griffith,J.D. and Nosek,J. (2002) t-loops in yeast mitochondria. *Mitochondrion*, **1**, 455–459.
12. Tomaska,L., Nosek,J., Makhov,A.M., Pastorakova,A. and Griffith,J.D. (2000) Extragenomic double-stranded DNA circles in yeast with linear mitochondrial genomes: potential involvement in telomere maintenance. *Nucleic Acids Res*., **28**, 4479–4487.
13. Tomaska,L., McEachern,M.J. and Nosek,J. (2004) Alternatives to telomerase: keeping linear chromosomes via telomeric circles. *FEBS Lett*., **567**, 142–146.
14. Nosek,J., Rycovska,A., Makhov,A.M., Griffith,J.D. and Tomaska,L. (2005) Amplification of telomeric arrays via rolling-circle mechanism. *J. Biol. Chem*., **280**, 10840–10845.
15. Cesare,A.J. and Griffith,J.D. (2004) Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Mol. Cell. Biol*., **24**, 9948–9957.
16. Nosek,J., Kosa,P. and Tomaska,L. (2006) On the origin of telomeres: a glimpse at the pre-telomerase world. *BioEssays*, **28**, 182–190.
17. Lin,D., Wu,L.C., Rinaldi,M.G. and Lehmann,P.F. (1995) Three distinct genotypes within *Candida parapsilosis* from clinical sources. *J. Clin. Microbiol*., **33**, 1815–1821.
18. Roy,B. and Meyer,S.A. (1998) Confirmation of the distinct genotype groups within the form species *Candida parapsilosis*. *J. Clin. Microbiol*., **36**, 216–218.
19. Tavanti,A., Davidson,A.D., Gow,N.A., Maiden,M.C. and Odds,F.C. (2005) *Candida orthopsilosis* and *Candida metapsilosis* spp. nov. to replace *Candida parapsilosis* groups II and III. *J. Clin. Microbiol*., **43**, 284–292.
20. Rycovska,A., Valach,M., Tomaska,L., Bolotin-Fukuhara,M. and Nosek,J. (2004) Linear versus circular mitochondrial genomes: intraspecies variability of mitochondrial genome architecture in *Candida parapsilosis*. *Microbiology*, **150**, 1571–1580.
21. Nosek,J., Dinouel,N., Kovac,L. and Fukuhara,H. (1995) Linear mitochondrial DNAs from yeasts: telomeres with large tandem repetitions. *Mol. Gen. Genet*., **247**, 61–72.
22. Bullerwell,C.E., Forget,L. and Lang,B.F. (2003) Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res*., **31**, 1614–1623.
23. Hall,T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser*., **41**, 95–98.
24. De Rijk,P. and De Wachter,R. (1997) RnaViz, a program for the visualisation of RNA secondary structure. *Nucleic Acids Res*., **25**, 4679–4684.
25. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*., **22**, 4673–4680.
26. Guindon,S., Lethiec,F., Duroux,P. and Gascuel,O. (2005) PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res*., **33**, W557–W559.
27. Kumar,S., Tamura,K. and Nei,M. (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform*., **5**, 150–163.
28. Sambrook,J. and Russell,D. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory, New York.
29. Kurtzman,C.P. and Robnett,C.J. (1997) Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5′ end of the large-subunit (26S) ribosomal DNA gene. *J. Clin. Microbiol*., **35**, 1216–1223.
30. Wuyts,J., Perriere,G. and Van De Peer,Y. (2004) The European ribosomal RNA database. *Nucleic Acids Res*., **32**, D101–103.
31. Clark-Walker,G.D. (1992) Evolution of mitochondrial genomes in fungi. *Int. Rev. Cytol*., **141**, 89–127.
32. Schafer,B. (2005) RNA maturation in mitochondria of *S. cerevisiae* and *S. pombe*. *Gene*, **354**, 80–85.
33. Osinga,K.A., De Vries,E., Van der Horst,G. and Tabak,H.F. (1984) Processing of yeast mitochondrial messenger RNAs at a conserved dodecamer sequence. *EMBO J*., **3**, 829–834.
34. Hofmann,T.J., Min,J. and Zassenhaus,H.P. (1993) Formation of the 3′ end of yeast mitochondrial mRNAs occurs by site-specific cleavage two bases downstream of a conserved dodecamer sequence. *Yeast*, **9**, 1319–1330.
35. Maleszka,R. (1994) The *in vivo* effects of ethidium bromide on mitochondrial and ribosomal DNA in *Candida parapsilosis*. *Yeast*, **10**, 1203–1210.
36. Maleszka,R., Skelly,P.J. and Clark-Walker,G.D. (1991) Rolling circle replication of DNA in yeast mitochondria. *EMBO J*., **10**, 3923–3929.
37. Naito,T., Matsuura,A. and Ishikawa,F. (1998) Circular chromosome formation in a fission yeast mutant defective in two ATM homologues. *Nature Genet*., **20**, 203–206.
38. Nakamura,T.M., Cooper,J.P. and Cech,T.R. (1998) Two modes of survival of fission yeast without telomerase. *Science*, **282**, 493–496.
39. Diezmann,S., Cox,C.J., Schonian,G., Vilgalys,R.J. and Mitchell,T.G. (2004) Phylogeny and evolution of medical species of *Candida* and related taxa: a multigenic analysis. *J. Clin. Microbiol*., **42**, 5624–5635.
40. Ling,F. and Shibata,T. (2002) Recombination-dependent mtDNA partitioning: *in vivo* role of Mhr1p to promote pairing of homologous DNA. *EMBO J*., **21**, 4730–4740.
41. Ling,F. and Shibata,T. (2004) Mhr1p-dependent concatemeric mitochondrial DNA formation for generating yeast mitochondrial homoplasmic cells. *Mol. Biol. Cell*., **15**, 310–322.

42. Fukuhara,H., Sor,F., Drissi,R., Dinouel,N., Miyakawa,I., Rousset,S. and Viola,A.M. (1993) Linear mitochondrial DNAs of yeasts: frequency of occurrence and general features. *Mol. Cell. Biol.*, **13**, 2309–2314.
43. Pramateftaki,P.V., Kouvelis,V.N., Lanaridis,P. and Typas,M.A. (2006) The mitochondrial genome of the wine yeast *Hanseniaspora uvarum*: a unique genome organization among yeast/fungal counterparts. *FEMS Yeast Res.*, **6**, 77–90.
44. Logue,M.E., Wong,S., Wolfe,K.H. and Butler,G. (2005) A genome sequence survey shows that the pathogenic yeast *Candida parapsilosis* has a defective MTLa1 allele at its mating type locus. *Eukaryot. Cell*, **4**, 1009–1017.