The mitochondrial genome of the wine yeast *Hanseniaspora uvarum*: a unique genome organization among yeast/fungal counterparts

Paraskevi V. Pramateftaki[1], Vassili N. Kouvelis[2], Panagiotis Lanaridis[1] & Milton A. Typas[2]

[1]Wine Institute of Athens, National Agricultural Research Foundation, Lykovrissi, Greece; and [2]Department of Genetics and Biotechnology, Faculty of Biology, University of Athens, Athens, Greece

Correspondence: Milton A. Typas, Department of Genetics and Biotechnology, Faculty of Biology, University of Athens, Panepistimiopolis, Athens 15701, Greece. Tel.: +30 210 7274633; fax: +30 210 7274318; e-mail: matypas@biol.uoa.gr

Received 8 June 2005; revised 18 July 2005; accepted 20 July 2005.
First published online 8 December 2005.

doi:10.1111/j.1567-1364.2005.00018.x

Editor: Teun Boekhout

Keywords
Hanseniaspora uvarum; mitochondrial genome (mtDNA); linear Type 2 mtDNA; genome organization; rRNA secondary structures; phylogenetic relationships.

Abstract

The complete sequence of the apiculate wine yeast *Hanseniaspora uvarum* mtDNA has been determined and analysed. It is an extremely compact linear molecule containing the shortest functional region ever found in fungi (11 094 bp long), flanked by Type 2 telomeric inverted repeats. The latter contained a 2704-bp-long subterminal region and tandem repeats of 839-bp units. In consequence, a population of mtDNA molecules that differed at the number of their telomeric reiterations was detected. The functional region of the mitochondrial genome coded for 32 genes, which included seven subunits of respiratory complexes and ATP synthase (the genes encoding for NADH oxidoreductase subunits were absent), two rRNAs and 23 tRNA genes which recognized codons for all amino acids. A single intron interrupted the cytochrome oxidase subunit 1 gene. A number of reasons contributed towards its strikingly small size, namely: (1) the remarkable size reduction (by >40%) of the *rns* and *rnl* genes; (2) that most tRNA genes and five of the seven protein-coding genes were the shortest among known yeast homologs; and (3) that the noncoding regions were restricted to 5.1% of the genome. In addition, the genome showed multiple changes in the orientation of transcription and the gene order differed drastically from other yeasts. When all protein coding gene sequences were considered as one unit and were compared with the corresponding molecules from all other complete mtDNAs of yeasts, the phylogenetic trees constructed robustly supported its placement basal to the yeast species of the '*Saccharomyces* complex', demonstrating the advantage of this approach over single-gene or multigene approaches of unlinked genes.

Introduction

The apiculate yeast *Hanseniaspora uvarum* is the predominant species on the surface of grape berries and during the early stages of spontaneous alcoholic fermentations, the latter stages of which are invariably dominated by the alcohol-tolerant strains of *Saccharomyces cerevisiae* (Fleet, 1993). Amongst the many molecular approaches used to determine intraspecies variability in *S. cerevisiae* populations in wines, restriction analyses of mitochondrial (mt) DNAs (mtDNA-RFLPs) proved to be the most powerful tool (Vezsinhet *et al.*, 1990; Querol *et al.*, 1992). This is easily understood, considering the lengthy mt intergenic regions of *S. cerevisiae* which are well known targets for genomic rearrangements (Clark-Walker, 1992). In a recent study of yeast populations isolated from different stages of spontaneous alcoholic fermentations in Greece, *H. uvarum* was the most frequently accounted species after *S. cerevisiae* and all non-*S. cerevisiae* species exhibited high levels of genetic polymorphisms in their mt genomes (Pramateftaki *et al.*, 2000). Because non-*S. cerevisiae* yeasts are considered to contribute significantly to the final quality of wines (Lema *et al.*, 1996; Esteve-Zarzoso *et al.*, 1998), we decided to study the structure and organization of *H. uvarum* mtDNA in depth.

Today, it is well documented that the *S. cerevisiae* organelle genome consists mainly of a population of polydisperse linear DNA molecules lacking specific terminal structures, mixed with only a small fraction of circular forms (for a review, see Williamson, 2002). A similar mtDNA organization has been reported for two more yeast species, namely *Candida glabrata* and *Schizosaccharomyces*

pombe (Maleszka *et al.*, 1991; Bendich, 1996), and because this type of genomic organization results in circular mapping we might presume that it also applies to other circular mtDNAs of yeasts. True linear mtDNAs, however, represent a different situation amongst yeasts, because genome-sized linear molecules are evident and telomeres receive specific structure, upon which they are divided in two types: Type 1 telomeres, which possess inverted terminal repeats with a covalently closed single-stranded hairpin, and Type 2 telomeres, with inverted repeats that are composed of tandem arrays of large repetitive units (Fukuhara *et al.*, 1993; Nosek *et al.*, 1995).

The rest of the typical mtDNA features of yeasts are more than well-documented: extended size variation (19–85.6 kb), high AT content, two major sets of genes with either universal presence (cytochrome oxidase subunits 1–3; apocytochrome b; ATP-synthase subunits 6, 8 and 9; mt small with large ribosomal rRNAs; $\geq$23 tRNA genes) or a patchy distribution (mt small subunit protein Var1; NADH dehydrogenase subunits; RNA component of RNAseP), usage of different mt genetic codes, lengthy intergenic sequences, presence of optional introns and mostly unidirectional transcription of genes (Lang *et al.*, 1983; Sekito *et al.*, 1995; Foury *et al.*, 1998; Koszul *et al.*, 2003; Langkjaer *et al.*, 2003; Nosek *et al.*, 2004; Talla *et al.*, 2005).

In the present work, we report the complete sequence of the compact Type 2 linear mtDNA of the wine yeast *H. uvarum* and we analyze particular features of its organization and unique characteristics. This small genome is compared with the Type 2 mtDNA of the pathogenic yeast *Candida parapsilosis*, for which the complete sequence has been determined (Nosek *et al.*, 2004) and the concatenated sequences of its seven protein-coding genes are used to examine its appropriateness for studying phylogenetic relationships with other yeasts.

Materials and methods

Strain and growth conditions

The *Hanseniaspora uvarum* reference strain MUCL 31704, initially isolated from grape juice, was used as the source of mtDNA. Standard growth conditions were used (Prameteftaki *et al.*, 2000) and liquid cultures in YPD broth were grown for 16 h, with shaking (250 r.p.m.), at 28 °C.

MtDNA extraction, cloning and sequencing

MtDNA was isolated according to the method described by Defontaine *et al.* (1991) and was subsequently digested with *Hind*III and/or *Xba*I. MtDNA fragments were purified from the agarose gels and were ligated into a suitably digested pBluescriptII KS+ vector (Stratagene, La Jolla, CA), following standard methodology for all DNA manipulations (Sambrook *et al.*, 1989). The presence of mtDNA inserts in recombinant plasmids was confirmed by colony hybridization using digoxigenin (DIG)-labelled amplified mt products from *Saccharomyces cerevisiae* (DIG-labelling and detection kit, Roche Biochemicals, Basel, Switzerland). Fragments were subcloned to sizes not exceeding 700 bp and, together with overlapping fragments, were used in sequencing experiments. Primer pairs IRpr/EX1 (5′-CGAATTCCTTTAGGATATATGGTATACATTATTAAATGATTT-3′ and 5′-CGAATTCTAATGAGAATCCTAAAATAGCGTAAGGG-3′) and IRpr/EX2 (5′-CGAATTCCTTTAGGATATATGGTATACATTATTAAATGATTT-3′ and 5′-CGAATTCAATTCAATAATACCTTCCATAGCAGGAG-3′) were used for the amplification of the inverted repeats at the extremities of the genome (Fig. 1). DNA sequencing was performed on a single-dye LiCor 4200 DNA sequencer (Licor Biosciences, Lincoln, NE), using the SequiTherm EXCEL II DNA Sequencing kit (Epicentre Technologies, Madison, WI). In all cases, both strands of the mtDNA were sequenced using the M13 forward and reverse universal primers.

Sequence assembly, sequence annotation and phylogenetic analyses

Contiguous sequences were assembled using the DNASTAR software (DNASTAR Inc., Madison, WI). DNA similarity searches were performed with Basic Local Alignment Search Tool (BLAST 2.2.12; Altschul *et al.*, 1997). DNA sequence alignments were made using CLUSTALW (Thompson *et al.*, 1994) with the multiple alignment parameters set to default and then edited by visual inspection. The tRNAscan-SE search server was used to predict tRNAs (Lowe & Eddy, 1997; http://www.genetics.wustl.edu/eddy/tRNAscan-SE/). rRNA genes were initially identified by step-by-step homology comparisons with *rns* and *rnl* genes of *S. cerevisiae* (GenBank accession numbers V00704 and J01527, respectively) and subsequently by using the prokaryotic 16S and 25S models that apply to mt rRNA molecules (Neefs *et al.*, 1991; Gutell *et al.*, 1993). Secondary conformations of the more highly derived regions were obtained by using representatives from the Comparative RNA Web (CRW) site as a guide (Cannone *et al.*, 2002; http://www.rna.icmb.utexas.edu) and by analyzing sequences in sections with the m-fold server (Zuker, 2003; http://www.bioinfo.rpi.edu/applications/mfold/old/rna). The nucleotide sequence of *H. uvarum* mtDNA appears in GenBank under the accession number DQ058142.

Phylogenetic analyses of protein data sets were performed using PAUP* (Swofford, 1998) for maximum parsimony (MP) and PHYLIP (Felsenstein, 2002) for maximum likelihood (ML), as described previously (Kouvelis *et al.*, 2004). Parsimony analysis was performed using 1000 replicates

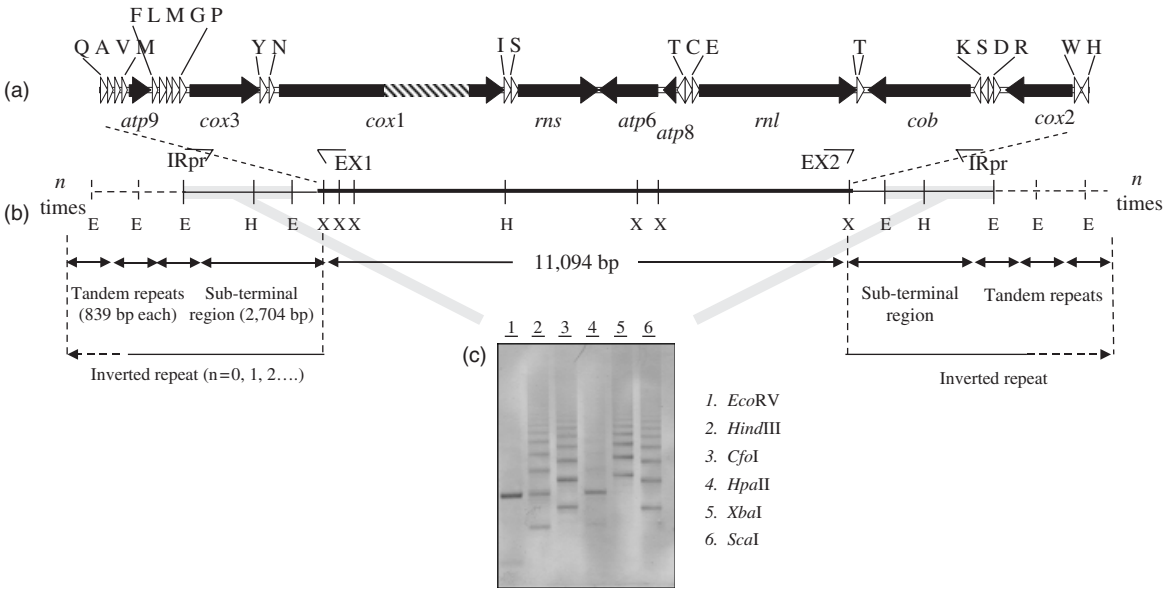

Fig. 1. (a) Genome organization of the linear Type 2 mtDNA of *Hanseniaspora uvarum*. Genes coding for proteins, *rns* and *rnl* are presented with thick black arrows, *trn* genes (named by their cognate amino acid in the one-letter code) with white arrows and the *cox*1 intron with a hatched box. Arrows indicate transcription orientation of genes. (b) A partial restriction map of the mtDNA (E, *Eco*RV; H, *Hind*III; X, *Xba*I). Hybridization of the subterminal *Eco*RV fragment (shaded gray) to mtDNA digests resulted in (c) a typical for Type 2 mtDNAs ladder pattern. Primers EX1 and EX2 were designed complementary to the *atp9* and *cox2* genes, respectively, and were paired with primer IRpr. The amplified fragments of the ladder pattern obtained by both primer pairs (IRpr/EX1 and IRpr/EX2) were sequenced and helped to determine the sequence of the tandem repeat as well as to verify the structure and identity of the two ends of the genome.

with random addition of taxa and tree-bisection reconnection branch swapping. Reliability of nodes was assessed using 1000 bootstrap iterations (Felsenstein, 1985). For the ML analysis, the Jones-Taylor-Thornton (JTT) substitution model with amino acid frequencies estimated from the data set was chosen. Aγ-distribution model of site variation ($\alpha = 0.77675$, calculated with PAML Yang, 1997) was used. In addition, phylogenetic analysis was carried out by the MRBAYES programme (Huelsenbeck & Ronquist, 2001), because MP analysis may be susceptible to long branch attraction phenomena (Felsenstein, 1978) and ML bootstrapping is time-consuming (Larget & Simon, 1999). The burn-in period was 300 000 cycles, and this was found to be clearly sufficient for the likelihood and the models parameters to reach equilibrium. After the burn-in, 3000 trees were sampled every 100 cycles during the sampling period (300 000 cycles). The amino acid sequences of the seven protein coding genes in *H. uvarum* mtDNA and the corresponding sequences from all 17 complete mt genomes of yeasts (*Saccharomycotina* and *Schizosaccharomycotina*) were used with four representatives from the subphylum of *Pezizomycotina* [*Aspergillus nidulans* (http://megasun.bch.umontreal.ca/maps/aspmap.gif), *Lecanicillium muscarium* (NC_004514), *Neurospora crassa* (http://mips.gsf.de/proj/medgen/mitop/) and *Penicillium marneffei* (NC_005256)] and one representative from each of the other phyla [namely *Cantharellus cibarius* for *Basidiomycetes* (http://megasun.bch.umontreal.ca/People/lang/FMGP/proteins.html), *Rhizopus oryzae* for *Zygomycetes* (NC_006836) and *Allomyces macrogynus* for *Chytridiomycetes* (NC_001715)] as outgroups to determine the phylogenetic relationships of *H. uvarum* (DQ058142) with the other yeasts [*Candida albicans* (NC_002653), *Candida glabrata* (NC_004691), *Candida metapsilosis* (NC_006971), *Candida orthopsilosis* (NC_006972), *Candida parapsilosis* (NC_005253), *Candida stellata* (NC_005972), *Ashbya gossypii* (NC_005789), *Kluyveromyces lactis* (NC_006077), *K. thermotolerans* (NC_006626), *Pichia canadensis* (synonym *Hansenula wingei*; NC_001762), *Saccharomyces cerevisiae* (NC_001224), *S. castelli* (NC_003920), *S. servazzii* (NC_004918), *Schizosaccharomyces pombe* (NC_001326), *S. japonicus* (NC_004332), *S. octosporus* (NC_004312), *Yarrowia lipolytica* (NC_002659)] (Fig. 4). A total of 1956 amino acid positions were included in the analysis.

Results and discussion

Gene content and genome organization

Sequence analysis of the *Hanseniaspora uvarum* mtDNA revealed a basic genome size, excluding inverted repeats, of 11 094 bp that codes for all essential yeast mt genes. The molecule is linear, terminating at identical inverted repeats that do not code for genes or ORFs. They consist of a 2704-

Table 1. Genetic content of the *Hanseniaspora uvarum* mitochondrial genome (DQ058142)

	Position			Codons		
Genetic element	From	To	Size (bp)	Start	Stop	Direction of transcription
Tandem repeat	333	1171	839			
Subterminal region	1172	3875	2704			
*trn*Q (TTG)	3887	3958	72			+
*trn*A (TGC)	3965	4034	70			+
*trn*V (TAC)	4041	4111	71			+
*trn*M (CAT)	4120	4190	71			+
*atp*9	4219	4449	231	ATG	TAA	+
*trn*F(GAA)	4467	4538	72			+
*trn*L (TAA)	4544	4626	83			+
*trn*M (CAT)	4628	4700	73			+
*trn*G (TCC)	4704	4774	71			+
*trn*P (TGG)	4781	4851	71			+
*cox*3	4891	5679	789	ATG	TAG	+
*trn*Y (GTA)	5694	5764	71			+
*trn*N (GTT)	5779	5849	71			+
*cox*1	5898–6611	7563–8423	1575	ATG	TAA	+
*cox*1intron-orf*	< 6611	7526	1629	ATG	TAG	+
*trn*I (GAT)	8436	8506	71			+
*trn*S (TGA)	8509	8591	83			+
rns	8592	9460	869			+
*atp*6	9511	10 167	657	ATG	TAA	−
atp 8	10 227	10 370	144	ATG	TAA	−
*trn*T (TAG)	10 402	10 473	72			−
*trn*C (GCA)	10 480	10 549	70			+
*trn*E (TTC)	10 558	10 628	71			+
rnl	10 629	12 419	1790			+
*trn*T (TGT)	12 420	12 490	71			+
cob	12 538	13 686	1149	ATG	TAA	−
*trn*K (TTT)	13 726	13 797	72			−
*trn*S (GCT)	13 800	13 880	81			−
*trn*D (GTC)	13 881	13 952	72			+
*trn*R (TCT)	13 954	14 024	71			+
*cox*2	14 094	14 831	738	ATG	TAA	−
*trn*W (TCA)	14 864	14 934	71			+
*trn*H (GTG)	14 942	15 012	71			−
Subterminal region	14 970	17 673	2704			
Tandem repeat	17 674	18 512	839			

trn, transfer RNA genes with anticodons in parenthesis; *rns* and *rnl*, ribosomal RNA genes for mt small and large ribosomal subunits; *atp* 6, 8 and 9, genes of ATP synthase subunits 6, 8 and 9; *cox*1, 2 and 3, genes of cytochrome oxidase subunits 1, 2 and 3; *cib*, apocytochrome *b* gene; (+), sense strand; (−), antisense strand.

*The *cox*1 gene is split by a group IB intron that contains an ORF in frame with upstream exon1 and codes for a hypothetical protein.

bp-long subterminal region and tandem repeats of an 839-bp unit (Fig. 1a, b). In consequence, when restricted mtDNA was hybridized with an *Eco*RV-labeled probe which contained the entire tandem repeat, we could detect a population of mtDNA molecules differing in the number of their telomeric reiterations from 0 to at least 10. This ladder pattern of hybridizing bands, obtained with any of the restriction endonucleases used (Fig. 1c), is typical of Type 2 telomeric structures of linear yeast genomes (Nosek *et al.*, 1995).

Undoubtedly, the *H. uvarum* mt genome size is the smallest reported yet for yeasts, and even with the addition of a complete copy of the inverted repeat (3543 bp long) on either side, it is still smaller than the smallest mtDNAs of yeasts, namely those of *Schizosaccharomyces pombe* (19.4 kb; Lang *et al.*, 1983) and *Candida glabrata* (20.1 kb; Koszul *et al.*, 2003). The exact positions and length of genes were located following nucleotide and inferred amino acid sequence comparisons with counterparts from all known yeast mtDNAs. Thus, genes coding for cytochrome oxidase subunits 1, 2 and 3 (*cox*1–3), apocytochrome b (*cob*), ATP-synthase subunits 6, 8 and 9 (*atp*6, 8 and 9), the small and large ribosomal rRNAs (*rns* and *rnl*) and 23 tRNA species (*trn*, coding for all 20 amino acids) were located on this genome (Table 1). The only exception to the compactness of this small genome was a single

951-bp-long group IB intron within its *cox*1 gene, located 714 bp downstream of the 5′ end of the gene. This position appears to be conserved in yeasts because nine out of 17 complete mtDNA genomes of ascomycetous yeasts have group I introns inserted at exactly the same position of the gene (corresponding to 720 nt of *Saccharomyces cerevisiae cox*1 homolog; data available upon request). It contains an ORF which is in frame with exon1 of *cox*1, and together they encode for a putative protein with the typical LAGLI-DADG motifs (52% similarity with the DNA site-specific endonucleases of *S. cerevisiae* and *Hansenula wingei* encoded by introns inserted after an identical nucleotide position) (Bonitz *et al*., 1980; Sekito *et al*., 1995). Therefore, the 32 genes detected in the mt genome of *H. uvarum* are the minimum set of those usually reported in yeast mtDNAs, thus verifying the accepted notion that the information content of the mt genome of yeasts is remarkably constant (Paquin *et al*., 1997).

Twenty-three *trn* genes, coding for all amino acids, were identified in the *H. uvarum* mt genome and their predicted secondary structures were drawn, having all the expected cloverleaf form and conserved anticodon sequences (data available upon request). The *trn* genes were interspersed between almost all coding sequences, in clusters of two to five genes (with the exception of *trn*T, which appeared alone), as in the case of *Sc. pombe*, *Candida parapsilosis* and in mammalian mtDNA (Boore, 1999; Schafer, 2003; Nosek *et al*., 2004). Thus, considering the overall compactness of the genome, the above organization implies that processing of primary transcripts may involve the excision of tRNAs and rRNAs, leading to the generation of mature mRNAs. The only peculiarity observed was that the left inverted repeat started 11-bp upstream of *trn*Q and its first 43 bp showed 100% identity with the *trn*H gene sequence. However, a complete and possibly functional copy of the latter gene was located at the junction of the right inverted repeat with the rest of the genome (Table 1). In a similar manner to that of all fungal mt genomes, its codon usage was biased towards the use of codons ending in A or T (80%), having as consequence the use of only 54 out of the foreseen 64 codons owing to the rarity or absence of codons with C or G in the third position. Five out of six codons of Arg are absent and in agreement with the anti-codon sequence (UCU) of the only *trn*R found. Two out of three deviations met in the *S. cerevisiae* mt genetic code were also verified here, namely that TGA was used as a Trp codon, whereas the CTN codon family coded for Thr. Assignment of ATA as a Met codon was not possible owing to its rare use and its presence in nonconserved positions. Nevertheless, because in many yeasts these deviations always occur in concert (*C. glabrata*, *Kluyveromyces thermotolerans*, *Saccharomyces castellii*, *S. cerevisiae*, *Saccharomyces servazii* (Foury *et al*., 1998; Koszul *et al*., 2003; Langkjaer *et al*., 2003; Talla *et al*., 2005)), we may assume that we have a similar situation here.

Characteristics of *Hanseniaspora uvarum* mtDNA

Size reduction of genome and secondary structures of the rRNAs

A comparison of the *H. uvarum* mt coding and noncoding sequences with corresponding sequences from other complete yeast mt genomes illustrated an extended reduction in size, almost in all regions of the mt genome (Table 2). Clearly, as a result of their conserved function, essential protein-coding genes and *trn* genes suffer smaller size reductions. Five out of the seven protein-coding genes, namely *cob*, *cox*1, *cox*3, *atp*8 and *atp*6, are smaller at their 5′ and 3′ extremities than corresponding genes of other yeasts, owing to small deletions which result in reductions of 3, 10, 7, 1 and 41 aa of their coded proteins, respectively (compared with *S. cerevisiae* mt gene products). Similarly, the length of *trn* genes was affected to reach sizes around the lower limit of corresponding molecules in yeasts, namely 70–73 bp. Intergenic regions were short (the largest being 69 bp long), though widespread (28), amounting to 571 bp, 5.1% of the complete mt genome (Table 2). However, in spite of this very compact gene organization, and although overlapping of genes is a common feature of filamentous ascomycete and animal mtDNAs (Boore, 1999; Kouvelis *et al*., 2004), no gene overlapping was observed.

Both mt rRNA genes suffer the most significant reductions in size (around 43%; see Table 2). With minimum estimated sizes of 869 and 1791 bp, *rns* and *rnl* are drastically smaller than corresponding yeast mt genes, which vary from 1300–2000 bp and from 2500–4500 bp, respectively. When putative secondary structures were drawn using the criteria of Cannone *et al*. (2002) (see also http://www.rna.icmb.utexas.edu) and were compared with corresponding fungal homologs, the elimination of specific secondary structure elements became apparent (Figs 2 and 3). More precisely, for the *rns* secondary structure, helices located in the variable periphery were either reduced (helices 8, 12, 22, 35, 47) or completely eliminated (helices 6, 9–11, 14, 15, 18, 27, 39, 41, 42, 44) leading to the loss of entire areas (Fig. 2). Interestingly enough, these helices which are stably present in most other fungal homologs (Gutell *et al*., 1985; Okamoto *et al*., 1996; Hegedus *et al*., 1998) are modified in a similar way as in reduced *rns* counterparts from other lineages such as metazoan and certain protists (Gray *et al*., 1998; Boore, 1999; Cannone *et al*., 2002). The complete absence of helix 43 is also noticeable, a vestige of which appears even in the *Caenorhabditis elegans rns*, the smallest rRNA sequence known among *Metazoa* (697 bp; Mears *et al*., 2002).

A number of characteristic helices and loops found in the secondary structures of *S. cerevisiae* and other fungal *rnl* genes were also eliminated in the corresponding secondary structure of *H. uvarum* (see boxed areas in Fig. 3). Regions

Table 2. Gene and genome sizes (in bp) of the so far known complete mtDNAs of yeasts and *Hanseniaspora uvarum*. Intergenic regions are expressed as a (%) percent of the corresponding complete mt genome

	Genome length	Protein genes†	tRNA genes†	rRNA genes	Intergenic regions (%)
Schizosaccharomyces pombe	19 431	5481	1718	4244	11.1
Candida glabrata	20 063	5418	1712	4743	14.9
*Candida metapsilosis**	21 686	5475	1689	3851	12.3
Candida orthopsilosis	22 528	5475	1689	4123	14.1
Candida stellata	23 114	5493	1711	3959	8.6
Kluyveromyces thermotolerans	23 584	5490	1714	4707	24.2
Saccharomyces castellii	25 753	5499	1717	4875	39.7
Hansenula wingei	27 694	5472	1730	4546	17.1
*Candida parapsilosi**	28 940	5475	1689	4536	7.7
Saccharomyces servazzii	30 782	5478	1733	4483	43.0
Kluyveromyces lactis‡	40 291	5430	1636	4855	55.0
Candida albicans	40 420	5478	1690	4593	36.1
Schizosaccharomyces octosporus	44 227	5499	1717	4195	49.4
Yarrowia lipolytica	47 916	5448	1700	4562	16.1
Schizosaccharomyces japonicus	80 059	5481	1730	6493	76.5
Saccharomyces cerevisiae	85 779	5487	1763	4945	62.0
Hanseniaspora uvarum*	**11 094**	**5283**	**1672**	**2660**	**5.1**

*The inverted terminal repeats of these mt genomes have been excluded.

†,‡Only proteins and tRNA genes identified in *Hansensiaspora uvarum* mtDNA are considered (except *Kluyveromyces lactis*, which has 22 tRNA genes). Intronic sequences within protein-coding genes have been excluded.

such as V6 were entirely missing and others were severely reduced. The V1 region, although significantly reduced (by 55% compared to *S. cerevisiae* corresponding region), still contained vestiges of the long-range base pairings and helices that define the 5′ end of a 5.8S-like structure. Helices defining the 3′ end were absent, but this also applies to *S. cerevisiae* and *Aspergillus nidulans* counterparts (Sor & Fukuhara, 1983; Dyson *et al*., 1989). A structure related to a 4.5S-like RNA was missing accordingly, as concluded by the much diminished size of the V15 region. Domains IV, V and the 5′ end of VI were the most structurally conserved because all significant functions are centered in these domains (Mears *et al*., 2002, and papers cited therein), hence they suffered the least reductions. They are characterized either by restricted helix reductions (V10, V11) or by the elimination of broad AU-reach insertions (V12, V13 and V14). Finally, no prominent reduction in a single-stranded region was observed, whereas all the major long-range interactions foreseen were possible (Fig. 3).

At the nucleotide level, the *rns* and *rnl* genes of *H. uvarum* (626 bp and 1578 bp, respectively; the secondary structures of which could be aligned) showed similarity levels of 58–64% when compared with corresponding genes from other yeasts. These results obviously underscore the divergence of these genes at the primary level. Yet they appear structurally conventional throughout most of the evolutionary conserved structural and functional cores. When only universally conserved residues were considered in comparisons, *rns* and *rnl* sequences were conserved at levels of 92 and 100%, respectively (for structure conservation diagrams and universally conserved residues, see Cannone *et al*. (2002); Mears *et al*. (2002). The rRNA simplification reported here is characteristic of derived mt genomes (Gray *et al*., 1999; Lang *et al*., 1999) seen in mt genomes of most animals and some protists (Gray *et al*., 1998). This is a strategy often employed when genome economization is sought, and as such it should be also considered for the *H. uvarum* mt genome (Gray *et al*., 1999; Lang *et al*., 1999). Overall, mt rRNA genes are probably the most flexible of the functional units because they can be expanded or reduced (Neefs *et al*., 1991; Gutell *et al*., 1993), split by introns (Gonzalez *et al*., 1999) or broken up into pieces and scrambled in the genome, interspersed with other genes on both strands of the mtDNA (Gray *et al*., 1998, and references cited therein).

A highly rearranged mt genome with multiple changes in transcription orientation

Recently, genome organization comparisons of all complete mtDNAs from ascomycetes have indicated several common features in their gene order (Kouvelis *et al*., 2004). However, the gene arrangement of *H. uvarum* mtDNA is totally unlike that of other yeast mt genomes, a fact that holds true even for the seemingly less rearranged 5′ half of the genome (see below; Fig. 1a). In an attempt to correlate small genome sizes with certain structural characteristics in animal mtDNAs, Noguchi *et al*. (2000) noted that smaller genomes always showed a nonconserved gene order and argued that genome size reduction may result in extensive

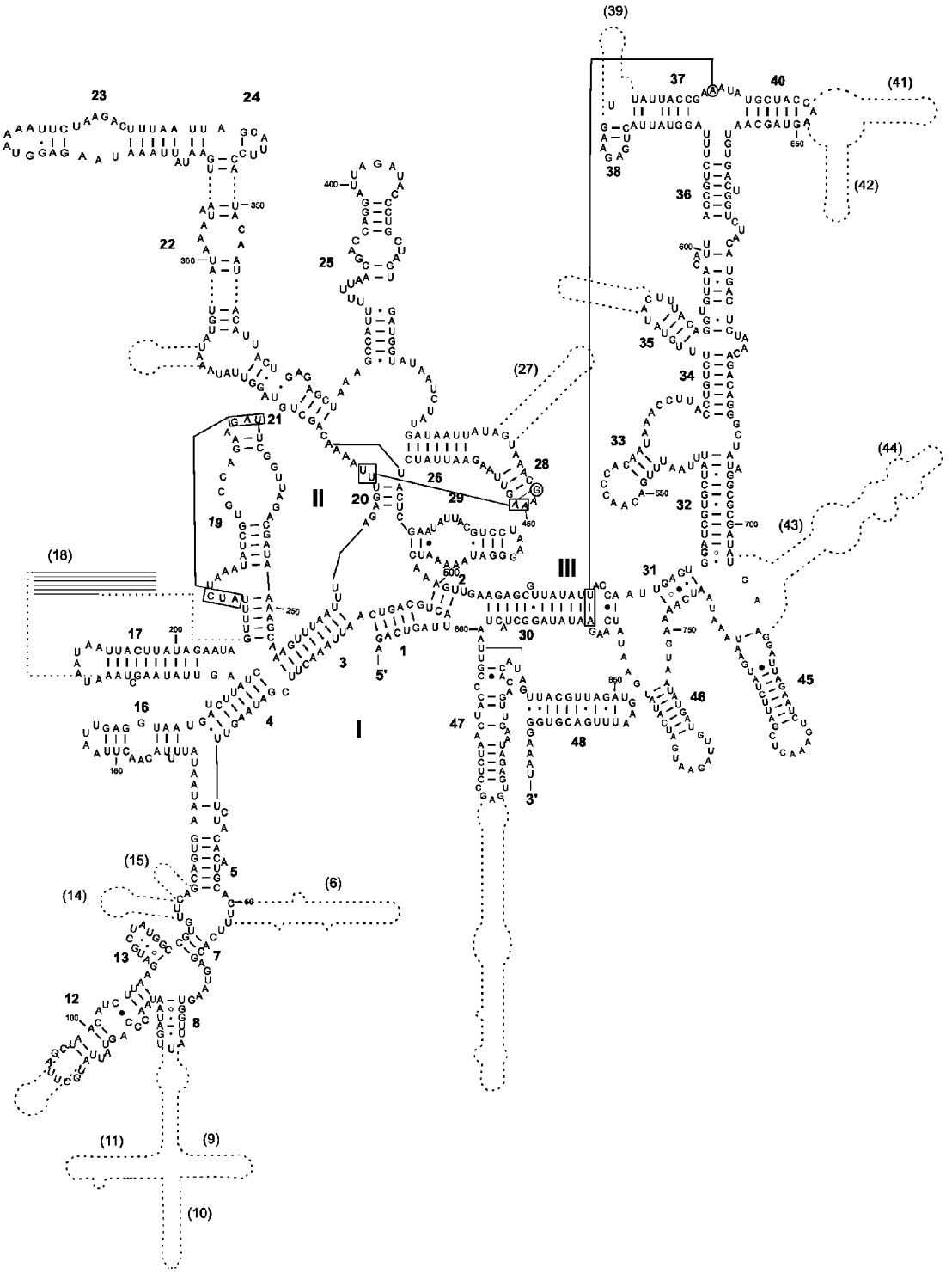

Fig. 2. Putative secondary structure of the mitochondrial *rns* gene of *Hanseniaspora uvarum*. The model is based on the *Escherichia coli* 16S according to rules set by Cannone *et al*. (2002). Helices are numbered according to Neefs *et al*. (1991) and Roman numerals denote the three major domains. Deleted helices of the structure as found by comparison to the *Saccharomyces cerevisiae* 16S rRNA (Gutell *et al*., 1985) are presented as dots and numbers in parentheses. Every 50th residue is numbered.

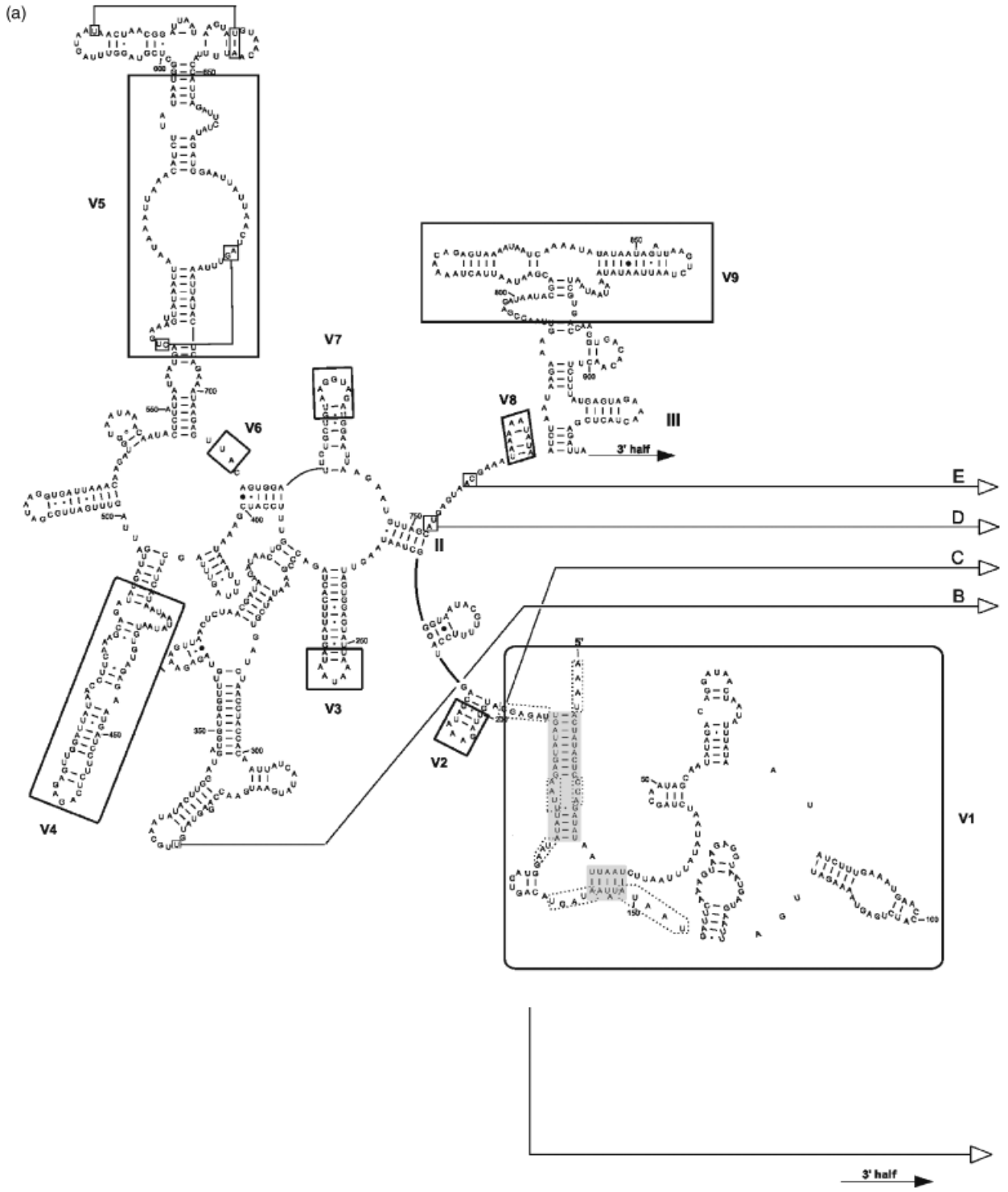

Fig. 3. Putative secondary structure of the mitochondrial *rnl* gene of *Hanseniaspora uvarum*. The model is based on the *Escherichia coli* 23S according to rules set by Cannone *et al*. (2002): (a) 5′ half; (b) 3′ half. Roman numerals denote the six major domains. Regions with extended variation from the corresponding *Saccharomyces cerevisiae* 23S rRNA model (Sor & Fukuhara, 1983) are boxed and labeled V1 to V15. Vestiges of the long-range base-pairing interactions that define a 5.8S-rRNA-like structure are shaded grey, whereas semiconservative sequence motifs flanking helical regions are enclosed by dotted lines. Every 50th residue is numbered.

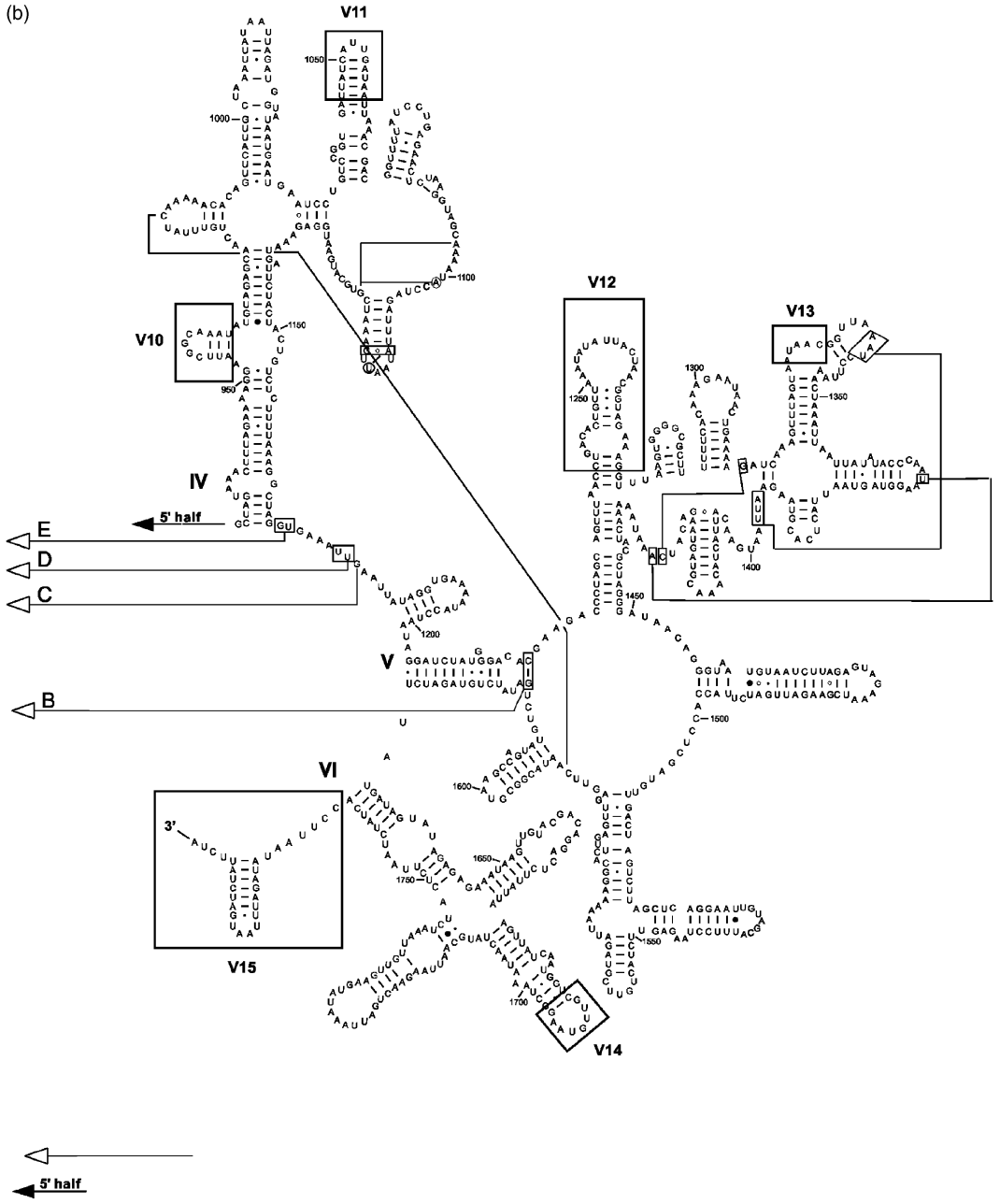

Fig. 3. Continued

reorganization of a genome. This may also be the case for *H. uvarum*, because its only conserved gene block is that of *atp8-atp6*, although transcribed in the opposite direction as in other yeast mtDNAs (the only exception being *C. albicans*; Anderson *et al.*, 2001). It is a truncated form of the *cox1-atp8-atp6* gene block found stably amongst the yeast mtDNAs of the '*Saccharomyces* complex' but only sparsely in more distant species (see Fig. 4). It is worth mentioning that *atp8-atp6* is the only pair of genes that is uninterrupted by a *trn* gene in any mtDNA from *Saccharomycotina* studied so far. The usual facilitators of mtDNA rearrangements in fungi are the intergenic regions and intronic sequences (Clark-Walker, 1992). However, the diminished size of the former and the absence of the latter from the *H. uvarum* mt genome suggest that the causative reasons of the rearrangements that shaped this genome should be sought elsewhere. Thus, taking into consideration the scattered distribution of *trn* genes in between protein-

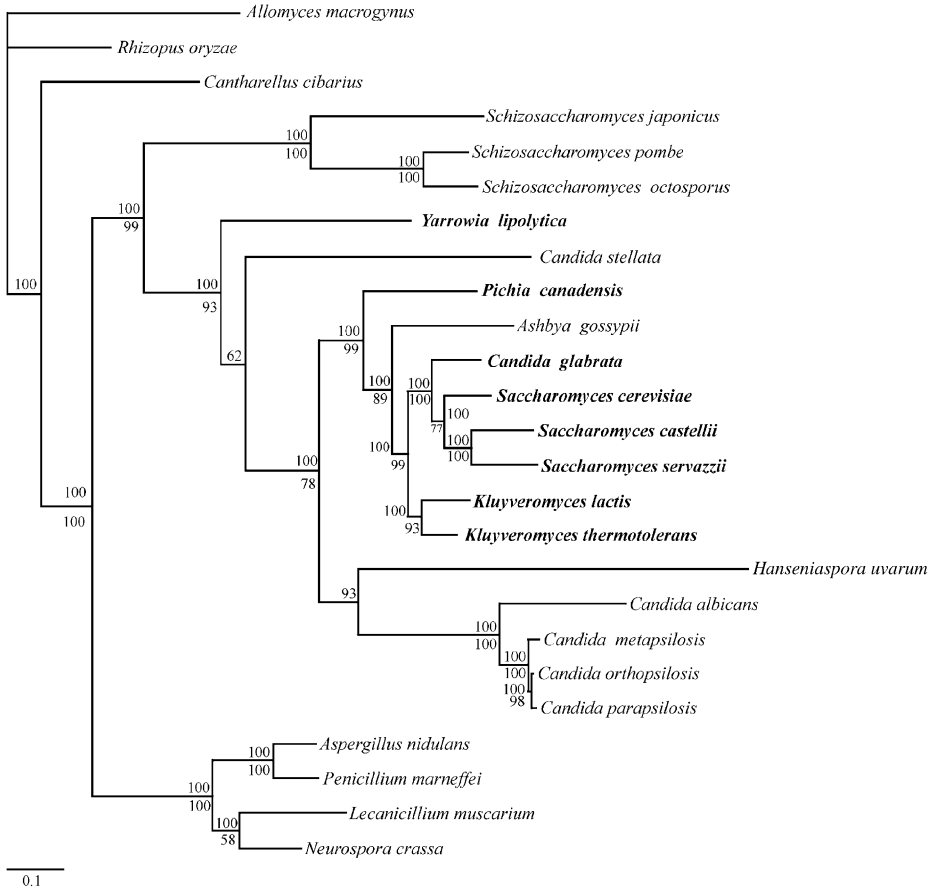

Fig. 4. The single phylogenetic tree constructed from unambiguously aligned portions of concatenated protein sequences of seven mitochondrial (mt) genes as produced by Bayesian analysis (and in accordance (100%) to the tree of maximum likelihood (ML) analysis). Clade credibility using MRBAYES (upper numbers) and parsimonial bootstrap support (lower numbers) calculated from 1000 replicates using PAUP*, are shown. The absence of lower numbers in clades reflects a different topology from maximum parsimony analysis. Yeast species with a preserved *cox*1-*atp*8-*atp*6 gene block are given in bold.

coding genes, duplication events such as that of the *trn*H in the left subterminal region and the multiple alternations of transcription orientation in the *trn* genes located at the 3′ part of the genome (see below), the involvement of *trn* genes in rearrangement events is strongly suggested. Earlier indications from *Saccharomyces douglasii* (Cardazzo *et al*., 1997), where *trn* sequences with the potential to form stable stem-and-loop structures were shown to play an important role in the rearrangements of its mtDNA, support the above hypothesis.

In contrast to what is commonly found in yeast mtDNAs, genes are located at and transcribed from both strands of the mtDNA molecule: 25 genes are transcribed rightwards and 8 leftwards (Fig. 1a; Table 1). Whereas genes at the 5′ half of the mtDNA (*trn*Q → *rns*) are transcribed unidirectionally and comprise a major cluster, genes at the 3′ half (*atp6* → *trn*H) are arranged in several small clusters showing opposite transcription orientation to each other. Transcription initiation from the inverted repeats is supported by the specific arrangement of genes adjacent to the telomeres towards the center of the molecule (the 5′ half of *trn*H is actually part of the right inverted repeat) and by the diminished intergenic regions, which do not leave the necessary space for multiple *S. cerevisiae*-like promoters (Osinga *et al*., 1984). Indeed, an *in silico* search for putative promoters similar to those proposed for yeasts failed to detect candidate sequences in the *H. uvarum* mtDNA. The above mode of transcription is unusual among the known linear mt genomes, for which transcription proceeds either from the center of the molecule towards the left and right telomeres (*C. parapsilosis*, *Polytomella parva*, *Tetrahymena pyriformis*; Edqvist *et al*., 2000; Fan & Lee, 2002; Nosek *et al*., 2004) or towards one direction (Type 1 yeast mtDNAs, (Fukuhara *et al*., 1993); *Hyaloraphidium curvatum*, Forget *et al*., 2002).

A Type 2 linear mtDNA

Hanseniaspora uvarum is the fifth yeast species reported to contain Type 2 mtDNA. Apart from the similar terminal

structures that it shares with *Candida parapsilosis*, *Candida metapsilosis* (a former subgroup of *C. parapsilosis*, now proposed to comprise a new species), *Candida salmanticensis* and *Pichia philodendra* (Nosek *et al.*, 1995; Tavanti *et al.*, 2005), no other shared features can be reported between these Type 2 mtDNAs. Even the similarly structured termini differ in detail. Inverted repeats vary in size from 978 bp (*C. metapsilosis*) to 3543 bp (*H. uvarum*), whereas tandem repeats range from 110 bp (*P. philodendra*) to 839 bp (*H. uvarum*) and show no sequence similarity between species (Nosek *et al.*, 1995; this study). In comparison with *C. parapsilosis* mtDNA, we observe different organization (in gene content, gene order and sizes), transcription pattern (two transcription units, heading in opposite directions, compared with the several in *H. uvarum*) and probably replication. [GC skew analysis in *H. uvarum* mtDNA, using the methodology of Nosek *et al.* (2004), indicated a single origin of replication located at the left inverted repeat, position 3421 nt, vs. a single bidirectional *ori* identified within the mtDNA of *C. parapsilosis* (Nosek *et al.*, 2004) (data not shown)]. Moreover, the subterminal regions of *H. uvarum* contain a *trn* gene vestige and their overall architecture implies their functional participation in transcription (and possibly in replication). These properties differentiate *H. uvarum* subterminal regions from their *C. parapsilosis* equivalents, which appear devoid of such functional traits (Nosek *et al.*, 1995, 2004). The detailed analysis of the genetic organization of *H. uvarum* mtDNA exemplifies the extended diversity that characterizes the overall organization of Type 2 mtDNAs (Nosek *et al.*, 1995) and at the same time makes it clear that none of the special features of *H. uvarum* is a common trait of linear or yeast mtDNAs in general. Because all known complete mt genomes of ascomycetous fungi are conventional, the mtDNA of *H. uvarum* must represent a derived pattern that occurred independently within the fungal lineage, maybe under similar constraints that led animal mtDNAs to genome economization. On the other hand, in spite of its differences in molecular architecture, it maintained the typical yeast features such as its gene content, presence of introns, elevated AT content, a complete set of *trn* genes, codon usage and usage of the genetic code of yeasts. We may conclude, therefore, that its existence delineates the extended structural diversity of mt genomes that does not necessarily correlate directly with the evolution of mt-encoded genes (Nosek & Tomaska, 2003).

Phylogenetic relationships to other yeast mtDNAs

Earlier, phylogenetic analysis of *Hanseniaspora* based on 26S rDNA sequences showed that this genus appears to be monophyletic and that it is not closely related to *Dekkera*, *Brettanomyces* and *Eeniella* (Boekhout *et al.*, 1994). Recent phylogenetic studies based on individual gene datasets, encompass the monophyletic genus of *Hanseniaspora* in the '*Saccharomyces* clade/complex', placing it either in the main body or in a most basal position within the clade (Cai *et al.*, 1996; Kurtzman & Robnett, 1998). In either case, its placement is only weakly supported, because single-gene phylogenies strongly resolve only the most closely related species. Even in the recent multigene phylogenetic analyses of Kurtzman & Robnett (2003), the basal placement of *Hanseniaspora* within the *Saccharomycetaceae* receives bootstrap support below 50%. Because analysis of concatenated mt protein-coding genes is considered to resolve both close and distant phylogenetic relationships better (Paquin *et al.*, 1997), the relationship of *H. uvarum* to the rest of the yeast species was examined using all protein coding mt gene sequences as one unit. Bayesian and ML analyses clearly illustrated that *H. uvarum* clustered as a sister clade to the rest of the known yeast mt genomes of the '*Saccharomyces* complex' and along with *C. metapsilosis*, *Candida orthopsilosis*, *C. parapsilosis* and *C. albicans* (Fig. 4). This placement receives both excellent clade credibility (100%) and bootstrap support (78%), demonstrating the advantages of the approach. It is, however, noted that with MP analysis *H. uvarum* branched at a basal position both in respect to the '*Saccharomyces* complex' and the four *Candida* spp. with 97% bootstrap support. Therefore, to assess the level of confidence in tree selection, statistical tests were performed employing CONSEL (Shimodaira & Hasegawa, 2001) and likelihood ratio tests were subsequently used for rigorous statistical testing of all alternative topologies of *H. uvarum* within *Saccharomycotina* and *Schizosaccharomycotina*. According to all tests included in CONSEL (data not shown), the best topology placed *H. uvarum* along with the four *Candida* species as a sister-clade of the '*Saccharomyces* complex', as was shown by the Bayesian and ML analysis (Fig. 4). This grouping of *H. uvarum* should be considered with caution because of the small number of taxa presently available for analysis and in the light of previous suggestions that linear mt genomes most probably had an evolutionary independent origin (Nosek *et al.*, 1998; Nosek & Tomaska, 2003). Undoubtedly, the continuously increasing number of complete fungal mt genomes will certainly help to resolve such questions in the near future. Data sets from complete mt genomes contain genes that are under the same evolutionary rate and therefore accurately portray or support the evolutionary history of the organisms examined (Forget *et al.*, 2002). As concatenated gene sequences were recently shown to fully resolve species trees with maximum support and without incongruences (Rokas *et al.*, 2003), our analysis of concatenated mt protein-coding genes clearly provides a better resolution over single gene or multigene approaches of unlinked genes.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipaman TJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Anderson JB, Wickens C, Khan M, Cowen LE, Federspiel N, Jones T & Kohn LM (2001) Infrequent genetic exchange and recombination in the mitochondrial genome of *Candida albicans*. *J Bacteriol* **183**: 865–872.

Bendich AJ (1996) Structural analysis of mtDNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J Mol Biol* **255**: 564–588.

Boekhout T, Kurtzman CP, O'Donell K & Smith MT (1994) Phylogeny of the yeast genera *Hanseniansporа* (anamorph *Kloeckera*), *Dekkera* (anamorph *Brettanomyces*), and *Eeniella* as inferred from partial 26S ribosomal DNA nucleotide sequences. *Int J Syst Bacteriol* **44**: 781–786.

Bonitz SG, Coruzzi G, Thalenfeld BE, Tzagoloff A & Macino G (1980) Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase. *J Biol Chem* **255**: 11927–11941.

Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Res* **15**: 1767–1780.

Cai J, Roberts IN & Collins D (1996) Phylogenetic relationships among members of the ascomycetous yeast genera *Brettanomyces*, *Debaryomyces*, *Dekkera*, and *Kluyveromyces* deduced by small subunit rRNA gene sequences. *Int J Syst Bacteriol* **46**: 542–549.

Cannone JJ, Subramanian S, Schnare MN, *et al.* (2002) The Comparative RNA Web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3, 2.

Cardazzo B, Rinaldi T, Frontali L, Carignani G & Palleschi C (1997) Evolution of mitochondrial genomes in yeast: a study of mitochondrial divergence in two closely related species, *Saccharomyces douglasii* and *Saccharomyces cerevisia*e. *Mol Biol Evol* **14**: 200–203.

Clark-Walker GD (1992) Evolution of mitochondrial genomes in fungi. *Int Rev Cyt* **141**: 89–127.

Defontaine AF, Lecocq M & Hallet JN (1991) A rapid miniprep method for the preparation of yeast mtDNA. *Nucleic Acids Res* **19**: 185.

Dyson NJ, Brown TA, Waring RB & Davies RW (1989) The mitochondrial ribosomal RNA molecules of *Aspergillus nidulans*. *Gene* **75**: 109–118.

Edqvist J, Burger G & Gray MW (2000) Expression of mitochondrial protein-coding genes in *Tetrahymena pyriformis*. *J Mol Biol* **297**: 381–393.

Esteve-Zarzoso B, Manzanares P, Ramon D & Querol A (1998) The role of non-*Saccharomyces* yeasts in industrial winemaking. *Int Microbiol* **1**: 143–148.

Fan J & Lee RW (2002) Mitochondrial genome of the colorless green alga *Polytomella parva*: two linear molecules with homologous inverted repeat termini. *Mol Biol Evol* **19**: 999–1007.

Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* **25**: 401–410.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Felsenstein J (2002) *PHYLIP (Phylogeny Inference Package) version 3.6a3. Distributed by the Author*. Department of Genome Sciences, University of Washington, Seattle, WA.

Fleet GH (1993) The microorganisms of winemaking - isolation, enumeration and identification. *Wine Microbiology and Biotechnology* (Fleet GH, ed.), pp. 1–25. Harwood Academic, Singapore.

Forget L, Ustinova J, Wang Z, Huss VAR & Lang BF (2002) *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing and an evolutionary link to lower fungi. *Mol Biol Evol* **19**: 310–319.

Foury F, Roganti T, Lecrenier N & Purnelle B (1998) The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS Lett* **440**: 325–331.

Fukuhara H, Sor F, Drissi R, Dinouël N, Miyakawa I, Rousset S & Viola A-M (1993) Linear mtDNAs of yeasts: frequency of occurrence and general features. *Mol Cell Biol* **13**: 2309–2314.

Gonzalez P, Barroso G & Labarere J (1999) Molecular gene organization and secondary structure of the mitochondrial large subunit ribosomal RNA from the cultivated Basidiomycota *Agrocybe aegerita*: a 13 kb gene possessing six unusual nucleotide extensions and eight introns. *Nucleic Acids Res* **27**: 1754–1761.

Gray MW, Burger G & Lang BF (1999) Mitochondrial evolution. *Science* **283**: 1476–1481.

Gray MW, Lang BF, Cedergren R, *et al.* (1998) Genome structure and gene content in protist mtDNAs. *Nucleic Acids Res* **26**: 865–878.

Gutell RR, Gray MW & Schnare MN (1993) A compilation of large subunit (23S and 23S-like) ribosomal RNA structures: 1993. *Nucleic Acids Res* **21**: 3055–3074.

Gutell RR, Weiser B, Woese CR & Noller HF (1985) Comparative anatomy of 16-S-like ribosomal RNA. *Prog Nucleic Acids Res Mol Biol* **32**: 155–216.

Hegedus DD, Pfeifer TA, Mulyk DS & Khachatourians GG (1998) Characterization and structure of the mitochondrial small rRNA gene of the entomopathogenic fungus *Beauveria bassiana*. *Genome* **41**: 471–476.

Huelsenbeck JP & Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Koszul R, Malpertuy A, Frangeul L, Bouchier C, Wincker P, Thierry A, Duthoy S, Ferris S, Hennequin C & Dujon B (2003) The complete mitochondrial genome sequence of the pathogenic yeast *Candida* (*Torulopsis*) *glabrata*. *FEBS Lett* **534**: 39–48.

Kouvelis VN, Ghikas DV & Typas MA (2004) The analysis of the complete mitochondrial genome of *Lecanicillium muscarium* (synonym *Verticillium lecanii*) suggests a minimum common

gene organization in mtDNAs of Sordariomycetes: phylogenetic implications. *Fung Genet Biol* **41**: 930–940.

Kurtzman CP & Robnett CJ (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek* **73**: 331–371.

Kurtzman CP & Robnett CJ (2003) Phylogenetic relationships among yeasts of the '*Saccharomyces* complex' determined from multigene sequence analyses. *FEMS Yeast Res* **3**: 417–432.

Lang BF, Ahne F, Distler S, Trinkl H, Kaudewitz F & Wolf K (1983) Sequence of the mtDNA, arrangement of genes and processing of their transcripts in *Schizosaccharomyces pombe*. *Mitochondria* (Schweyen RJ, Wolf F & Kaudewitz F, eds), pp. 313–329. Walter de Gruyter, Berlin, Germany.

Lang BF, Gray MW & Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. *Ann Rev Genet* **33**: 351–397.

Langkjaer RB, Casaregola S, Ussery DW, Gaillardin C & Piskur J (2003) Sequence analysis of three mtDNA molecules reveals interesting differences among *Saccharomyces* yeasts. *Nucleic Acids Res* **31**: 3081–3091.

Larget B & Simon D (1999) Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Mol Biol Evol* **16**: 750–759.

Lema G, Garcia-Jares C, Orriols I & Angulo L (1996) Contribution of *Saccharomyces* and non-*Saccharomyces* populations to the production of some components of Albarino wine aroma. *Am J Enol Vitic* **47**: 206–216.

Lowe TM & Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequences. *Nucleic Acids Res* **25**: 955–964.

Maleszka R, Skelly PJ & Clark-Walker GD (1991) Rolling circle replication of DNA in yeast mitochondria. *EMBO J* **10**: 3923–3929.

Mears JA, Cannone JJ, Stagg SM, Gutell RR, Agrawal RK & Harvey SC (2002) Modelling a minimal ribosome based on comparative sequence analysis. *J Mol Biol* **321**: 215–234.

Neefs J-M, Van de Peer Y, De Rijk P, Goris A & De Wachter R (1991) Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* **19**: 1987–2015.

Noguchi Y, Endo K, Tajima F & Ueshima R (2000) The mitochondrial genome of the Brachiopod *Laqueus rubellus*. *Genetics* **155**: 245–259.

Nosek J, Dinouel N, Kovac L & Fukuhara H (1995) Linear mtDNAs from yeasts: telomeres with large tandem repetitions. *Mol Gen Genet* **247**: 61–72.

Nosek J, Novotna M, Hlavatovicova Z, Ussery DW, Fajkus J & Tomaska L (2004) Complete sequence of the linear mitochondrial genome of the pathogenic yeast *Candida parapsilosis*. *Mol Gen Genet* **272**: 173–180.

Nosek J & Tomaska L (2003) Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Curr Genet* **44**: 73–84.

Nosek J, Tomaska L, Fukuhara H, Suyama Y & Kovac L (1998) Linear mitochondrial genomes: 30 years down the line. *Trends Genet* **14**: 184–188.

Okamoto K, Sekito T & Yoshida K (1996) The secondary structure and phylogenetic relationship deduced from complete nucleotide sequence of mitochondrial small subunit rRNA in yeast *Hansenula wingei*. *Genes Genet Syst* **71**: 69–74.

Osinga KA, De Vries E, Van der Horst GT & Tabak HF (1984) Initiation of transcription in yeast mitochondria: analysis of origins of replication and of genes coding for a messenger RNA and a transfer RNA. *Nucleic Acids Res* **12**: 1889–1900.

Paquin B, Laforest M-J, Forget L, Roewer I, Wang Z, Longcore J & Lang BF (1997) The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. *Curr Genet* **31**: 380–395.

Pramateftaki PV, Lanaridis P & Typas MA (2000) Molecular identification of wine yeast at species or strain level: a case study with strains from two vine-growing areas of Greece. *J Appl Microbiol* **88**: 1–14.

Querol A, Barrio E, Huerta T & Ramon D (1992) Molecular monitoring of wine fermentations conducted by active dry yeast strains. *Appl Environ Microbiol* **58**: 2948–2953.

Rokas A, Williams BL, King N & Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* **425**: 798–804.

Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schafer B (2003) Genetic conservation versus variability in mitochondria: the architecture of the mitochondrial genome in the petite-negative yeast *Schizosaccharomyces pombe*. *Curr Genet* **43**: 311–326.

Sekito T, Okamoto K & Kitano H (1995) The complete mtDNA sequence of *Hansenula wingei* reveals new characteristics of yeast mitochondrial. *Curr Genet* **28**: 39–53.

Shimodaira H & Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**: 1246–1247.

Sor F & Fukuhara H (1983) Complete DNA sequence coding for the large ribosomal RNA of yeast mitochondria. *Nucleic Acids Res* **11**: 339–348.

Swofford DL (1998) *PAUP*. Phylogenetic Analysis using Parsimony (*) and Other Methods, version 4*. Sinauer Associates, Sunderland, MA.

Talla E, Anthouard V, Bouchier C, Franguel L & Dujon B (2005) The complete mitochondrial genome of the yeast *Kluyveromyces thermotolerans*. *FEBS Lett* **579**: 30–40.

Tavanti A, Davidson AD, Gow NA, Maiden MC & Odds FC (2005) *Candida orthopsilosis* and *Candida metapsilosis spp. nov.* to replace *Candida parapsilosis groups* II and III. *J Clin Microbiol* **43**: 284–292.

Thompson JD, Higgins DG & Gibson TJ (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.

Vezinhet F, Blondin B & Hallet J-N (1990) Chromosomal DNA patterns and mtDNA polymorphism as tools for identification of enological strains of *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **32**: 568–571.

Williamson D (2002) The curious history of yeast mtDNA. *Nat Rev Genet* **3**: 475–481.

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**: 555–556.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.