J Mol Evol (1994) 38:593–601

Journal of Molecular Evolution

The Structure of the Small Mitochondrial DNA of *Kluyveromyces thermotolerans* Is Likely to Reflect the Ancestral Gene Order in Fungi

G.D. Clark-Walker,[1] G.F. Weiller[2]

[1] Molecular and Population Genetics Group, Research School of Biological Sciences, Australian National University, Canberra, A.C.T. 2601, Australia
[2] Centre for Molecular Structure and Function, Research School of Biological Sciences, Australian National University, Canberra, A.C.T. 2601, Australia

Received: 5 March 1993 / Revised and accepted: 31 October 1993

Abstract. Mapping the 23-kb circular mitochondrial DNA from the yeast *Kluyveromyces thermotolerans* has shown that only one change occurs in the gene order in comparison to the 19-kb mtDNA of *Candida* (*Torulopsis*) *glabrata*. Sequence analysis of the mitochondrially encoded cytochrome oxidase subunit 2 gene reveals that despite their conserved gene order, the two small genomes are more distantly related than larger mtDNA molecules with multiple rearrangements. This result supports a previous observation that larger mitochondrial genomes are more prone to rearrange than smaller forms and suggests that the architecture of the two small molecules is likely to represent the structure of an ancestor.

Key words: *Kluyveromyces thermotolerans* mtDNA — Cytochrome oxidase subunit 2 sequence — mtDNA phylogenetic tree — Genome rearrangements

Introduction

Mitochondrial DNA of yeasts is variable in length and gene order (Clark-Walker 1992). Within three separate genera (*Dekkera, Kluyveromyces,* and *Saccharomyces*) mtDNA has a three- to fourfold difference in length and several rearrangements. Even within a single species, mtDNA can vary in length and, more noteworthy, in gene order. For example, in strains of *Kluyveromyces lactis* the type A genome contains a 10-kb segment that has a different location in strain B (Skelly et al. 1991). Similarly, it has been proposed that the two species *Dekkera bruxellensis* and *Brettanomyces custersii* are strains of a single taxon (Smith et al. 1990), even though their mtDNAs of 85 and 101 kb differ in gene order (Hoeben and Clark-Walker 1986). Likewise, sequence rearrangement has been reported between the large (~70-80 kb) mtDNAs of the interfertile yeasts *S. cerevisiae* and *S. douglasii* (now classified as *S. paradoxus;* Barnett 1992), where a segment of approximately 15 kb, containing the genes for cytochrome oxidase subunit 3 (*COX3*) and the small subunit of rRNA (*SSU*), occurs at a different location in the two genomes (Tian et al. 1991). By contrast, the three smallest mtDNAs (28, 35, and 44 kb) from three species in the *Dekkera/Brettanomyces* group have identical gene topologies despite the mtDNA in these yeasts being more distantly related than larger genomes containing rearrangements (Hoeben and Clark-Walker 1986; Hoeben et al. 1993). Taken together these observations indicate that larger mtDNAs are more likely to rearrange than smaller forms (Clark-Walker 1992), and as a consequence, smaller genomes may retain gene arrangements that were present in a progenitor molecule. Following from this we reasoned that it may be possible to deduce the topology of the ancestral genome of yeasts by examining the structure of small mtDNAs from distant species. Indeed, in this report it is recorded that the 23-kb mtDNA of *K. thermotolerans* has, with the exception of one

Correspondence to: G.D. Clark-Walker

Table 1. Probes containing specific mtDNA fragments[a]

Gene	Probe name	Description	Source
A6	DS14	4.1-kb fragment of *S. cerevisiae* mtDNA	Macino and Tzagoloff 1980
A8	G5	681-bp fragment of *S. cerevisiae* mtDNA	Macreadie et al. 1983
A9	DS400/A3	1.8-kb fragment of *S. cerevisiae* mtDNA	Macino and Tzagoloff 1979
COX1 (exons 1–2)	DS6/A401	6.1-kb fragment of *S. cerevisiae* mtDNA	Bonitz et al. 1980
COX1 (exons 3–7)	DS6/A422	5.3-kb fragment of *S. cerevisiae* mtDNA	Bonitz et al. 1980
COX2	pCH14	952-bp *Fnu*DII fragment containing entire *COX2* gene from *K. lactis* mtDNA	Hardy and Clark-Walker 1990
COX3	DS31	4.5-kb fragment of *S. cerevisiae* mtDNA	Thalenfeld and Tzagoloff 1980
CYB (exon 3)	DS400M11	3.36-kb fragment of *S. cerevisiae* mtDNA	Nobrega and Tzagoloff 1980
LSU	E3	Segment of LSU from *S. cerevisiae* mtDNA (305-1,958 bp)	Sor and Fukuhara 1983
SSU	pSEM5	~1.8kb *Kpn*I fragment containing entire *SSU* gene from *S. exiguus* mtDNA	G.D. Clark-Walker (unpublished)
VAR1	pVar1	1.1-kb fragment of *S. cerevisiae* mtDNA	Hudspeth *et al.* 1982

[a] Gene abbreviations: A6, A8, and A9, subunits 6, 8, and 9 of ATPase; COX1, COX2, and COX3, subunits 1, 2, and 3 of cytochrome oxidase; CYB, apocytochrome b; LSU and SSU, large and small subunits of rRNA and VAR1, variant 1 protein

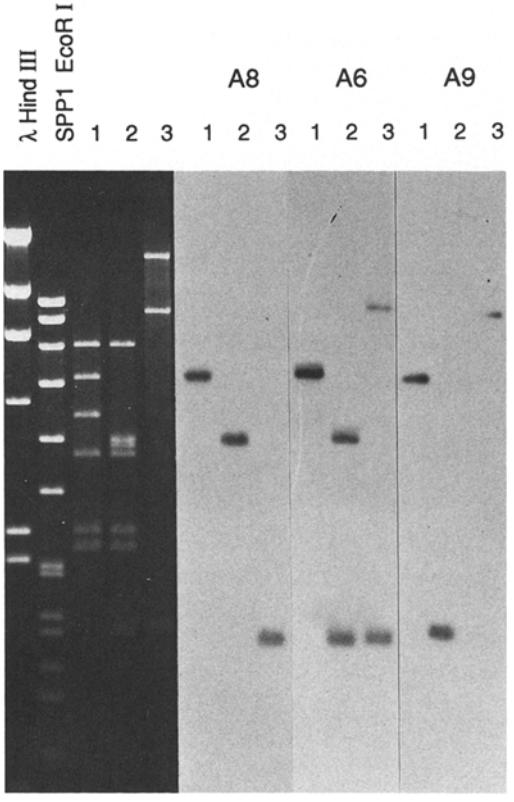

Fig. 1. Ethidium-bromide-stained gel and autoradiograms of ^{32}[P]-labeled probes hybridized to electrophoretically separated fragments of *K. thermotolerans* mtDNA. Fragments have been formed by digestion with, lane 1, *Hae*III; 2, *Hae*III and *Sca*I; 3, *Sca*I and *Eco*RI. mtDNA probes are the ATPase 8, 6, and 9 genes from *S. cerevisiae* (Table 1). Sizes of DNA markers are λ *Hind*III, 23,130, 6,557, 4,373, 2,322, 2,072, and 564 bp and SPP1 *Eco*RI; 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, and 0.36 kb.

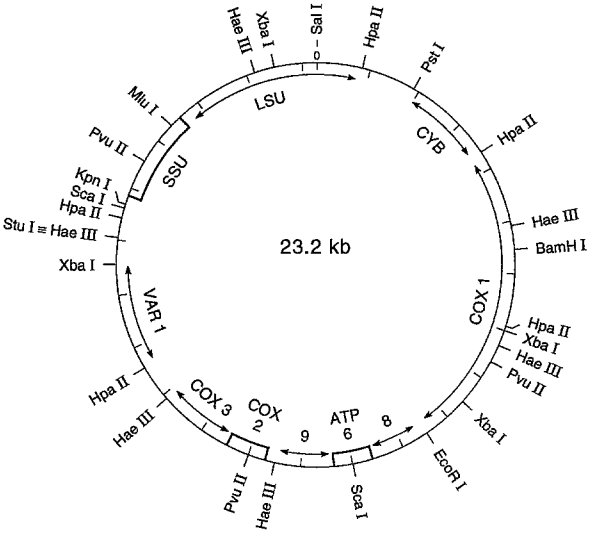

Fig. 2. Restriction endonuclease site map of *K. thermotolerans* mtDNA. Location of regions hybridizing to the various probes are shown inside the circle together with 1-kb divisions beginning with the unique *Sal*I site in the LSU gene. Regions with arrows show possible locations of genes while boxed regions for *ATP6, COX2,* and *SSU* have been positioned by sequence analysis (*COX2*) or by the occurrence of conserved restriction endonuclease sites in homologous genes from other yeast mtDNAs (*ATP6* and *SSU*).

gene, the same architecture as the 19-kb mtDNA of *Candida* (*Torulopsis*) *glabrata* (Clark-Walker and Sriprakash 1981, 1983; Clark-Walker et al. 1985). Moreover, a phylogenetic tree obtained from sequence comparisons of the mtDNA-encoded cytochrome oxidase subunit 2 (*COX2*) gene from several yeasts indicates that the shared gene order in the two small genomes is likely to represent the structure of a progenitor.

Materials and Methods

Yeast Strain and Isolation of mtDNA. K. thermotolerans CBS 2803 was obtained from the Centraal Bureau voor Schimmelcultures, Delft, The Netherlands. For isolation of mtDNA, a culture was grown in 2 L of GYP (2% glucose, 0.5% Difco yeast extract, 1.0% Difco peptone) at 28°C and spheroplasts were made by digestion with Zymolase. Spheroplasts were broken in a French press and mtDNA was

Table 2. Primers used in COX2 sequence determination

Primer	Position in *K. thermotolerans* gene	Sequence
COX2–1	367–385	5′-CAATGATATTGAAAATATG-3′
COX2–2	385–367	5′-CATATTTTCAATATCATTG-3′
COX2–3	627–644	5′-TTTAATTCAAAGAGAAGG-3′
COX 2–4[a]	592–608	5′-GATGCTGTTCCTGGTCG-3′
COX2–5[a]	245–229	5′-TCTATAAATTGTCCATG-3′
COX2–6	559–544	5′-CATGAATAACATCAGC-3′
COX2–7[a]	44–59	5′-ATGATGTACCAACACC-3′

[a] These primers differ from the sequence in *K. thermotolerans*.

Table 3. Mapping of *K. thermotolerans* mtDNA

a)	HaeIII fragment size (kb)	Gene hybridization
	6.2	LSU, CYB, COX1
	2.20	COX1
	5.20	COX1, A8, A6, A9
	2.35	COX2, COX3
	3.20	VAR1
	4.05	SSU, LSU
	23.2 kb	

b)	mtDNA fragment	Map coordinates (kb)	Gene hybridization
	*Pst*1-*Hpa*II	1.95-3.65	CYB
	*Hpa*II-*Eco*R1	3.65-9.45	COX1
	*Eco*R1-*Sca*I	9.45-11.0	A8, A6
	*Sca*I-*Hae*III	11.0-12.5	A6, A9
	*Hae*III-*Pvu*II	12.5-13.0	COX2
	*Pvu*II-*Hae*III	13.0-14.85	COX2, COX3
	*Hpa*II-*Xba*I	15.55-17.6	VAR1
	*Kpn*1-*Hae*III	18.65-22.1	SSU, LSU
	*Hae*III-*Hpa*II	22.1-0.85	LSU

isolated using bizbenzimide-CsCl buoyant density centrifugation according to a procedure described in detail (Clark-Walker et al. 1981).

Construction of a mtDNA Map. Restriction enzyme digestion of mtDNA was performed in KGB buffer (McClelland et al. 1988) and DNA fragments were separated by electrophoresis in 1% agarose gels. Transfer of DNA to nylon membrane was performed under alkaline conditions using the bidirectional transfer method as previously described (Hardy et al. 1989). Preparation of ^{32}[P]-labeled DNA probes was by the random priming method (Clark-Walker and Sriprakash 1981). Probes used in the mapping studies are listed in Table 1. Hybridization, washing of filters, and autoradiography have been described in detail (Clark-Walker and Sriprakash 1981).

Cloning and Sequencing. The 2.35-kb *Hae*III fragment containing the *COX2* gene was purified from an agarose gel using Prep-a-Gene, Biorad, Richmond, CA 94804, and cloned into the *Sma*I site of pTZ 18R. Sequence determination by the dideoxy chain termination method (Sanger et al. 1977) was performed with single-stranded DNA using oligonucleotide primers listed in Table 2.

Construction of the Phylogenetic Tree. Sources of *COX2* sequences are noted in the legend to Fig. 5. Feng-Johnson-Doolittle (FJD) distances (Feng et al. 1985) were deduced from the core peptide alignment (amino acids 16–252) using the "tree" program (Feng and Doolittle 1987) kindly provided by D.F. Feng. We used the DisCalc program (Weiller in prep.) for deriving distance matrices from the nucleotide alignment of the core fragment (nucleotides 36–756).

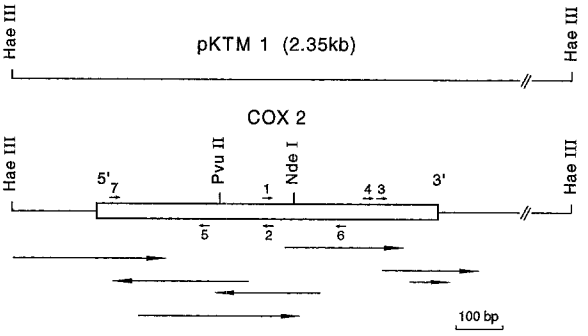

Fig. 3. Diagrams illustrating the position and orientation of the *COX2* open-reading frame relative to key restriction endonuclease sites. *Large arrows* illustrate the extent and direction of sequence determination while *small arrows* show the position and orientation of the primers listed in Table 2.

DisCalc calculates the proportion of nucleotide changes for every sequence pair in each codon position. Consequently it estimates the number of mutations per site by using either the Jukes and Cantor formula (Jukes and Cantor 1969) or the Kimura 2-, 3-, 4-, or 6-parameter corrections (Kimura 1981; Takahata and Kimura 1981). For a review on distance data correction see Gojobori et al. (1990).

Neighbor-joining trees (Saitou and Nei 1985) were reconstructed from distance data using the "njtree" program of Li Jin and J.W.H. Ferguson.

Results

Map of K. thermotolerans *mtDNA*

Mapping restriction enzymes sites and positioning genes was performed simultaneously using the probes listed in Table 1. For example, fragments of mtDNA hybridizing to probes for ATPase subunits 8, 6, and 9 are shown in Fig. 1. All three probes hybridize to the second largest *Hae*III fragment of 5.2 kb (lane 1, Fig. 1) (Table 3a). Cleavage of this *Hae*III fragment with *Sca*I leads to two bands of 3.65 and 1.5 kb that both hybridize to the ATPase subunit 6 probe (lane 2), indicating that this restriction enzyme site bisects the gene. Hybridization with probes for the other ATPase genes shows that subunit 8 is located on the 3.65-kb fragment whereas ATPase 9 is found on the 1.5-kb band. Double digestion of mtDNA with *Eco*R1, possessing a unique site and *Sca*I, positions the *Sca*I site and allows the order of the ATPase subunits to be determined (lane 3). By sim-

```
                                                    HaeIII
                                                 5'   CCTCCCTTCGGGGGTTAAGAGGATAAAGC
-158 CTTTATAGCTTAGTGGTAAAGCAGTAAACTGAAGATTTACCTACATGTAGTTCGATCTCATTAAGGGCATTAATGTATT
 -79 ATATATAAAAATAATTTTCTATAAGTAATAATAATTTTATAAATAGACAAAAAAGTCTAATAATATATAAATAAGAAAA

   1 ATG TTA AAT TTA TTA TAT AAT CAA ATT TTT AAC GTC ATT TTA AAT GAT GTA CCT ACA CCA
     Met Leu Asn Leu Leu Tyr Asn Gln Ile Phe Asn Val Ile Leu Asn Asp Val Pro Thr Pro

  61 TAT AAC CTT TAT TTT CAA GAT TCA GCT ACA CCA AAT CAA GAA GGT ATT TTA GAA TTA CAT
     Tyr Asn Thr Tyr Phe Gln Asp Ser Ala Thr Pro Asn Gln Glu Gly Ile Leu Glu Leu His

 121 GAT AAT ATT ATG TTT TAT TTA TTA GTA ATC TTA GGA TTA GTA TCA TGA TTA TTA TTT ACA
     Asp Asn Ile Met Phe Tyr Leu Leu Val Ile Leu Gly Leu Val Ser Trp Leu Leu Phe Thr

 181 ATT CTA AGA ACT TAT TCT AAA AAT CCT ATT GCA TAT AAA TAT ATT AAA CAT GGA CAA ACA
     Ile Thr Arg Thr Tyr Ser Lys Asn Pro Ile Ala Tyr Lys Tyr Ile Lys His Gly Gln Thr
                                                PvuII
 241 ATT GAA ATT ATT TGA ACA ATC TTC CCA GCT GTA ATT TTA TTA ATT ATT GCA TTC CCT TCA
     Ile Glu Ile Ile Trp Thr Ile Phe Pro Ala Val Ile Leu Leu Ile Ile Ala Phe Pro Ser

 301 TTT ATT TTA TTA TAT TTA TGT GAT GAA GTT ATT TCA CCT GCT ATG ACT ATT AAA GCT ATT
     Phe Ile Leu Leu Tyr Leu Cys Asp Glu Val Ile Ser Pro Ala Met Thr Ile Lys Ala Ile

 361 GGA TTA CAA TGA TAT TGA AAA TAC GAA TAC TCT GAT TTT ATT AAC GAT AGT GGT GAA ACA
     Gly Leu Gln Trp Tyr Trp Lys Tyr Glu Tyr Ser Asp Phe Ile Asn Asp Ser Gly Glu Thr
                      NdeI
 421 GTA GAA TTT GAA TCA TAT GTT ATT CCT GAA GAT TTA TTA GAA GAT GGT CAA TTA AGA TTA
     Val Glu Phe Glu Ser Tyr Val Ile Pro Glu Asp Leu Leu Glu Asp Gly Gln Leu Arg Leu

 481 TTA GAT ACA GAT ACA TCA GTA GTA GTA CCT GTT GAT ACT CAT ATT AGA TTT GTT GTT ACA
     Leu Asp Thr Asp Thr Ser Val Val Val Pro Val Asp Thr His Ile Arg Phe Val Val Thr

 541 GCA GCT GAT GTT ATT CAT GAT TTT GCT ATT CCT AGT TTA GGT ATT AAA GTT GAT GCA GCT
     Ala Ala Asp Val Ile His Asp Phe Ala Ile Pro Ser Leu Gly Ile Lys Val Asp Ala Ala

 601 CCT GGT AGA TTA AAT CAA GTT TCT GCT TTA ATT CAA AGA GAA GGT GTA TTC TAT GGT CAA
     Pro Gly Arg Leu Asn Gln Val Ser Ala Leu Ile Gln Arg Glu Gly Val Phe Tyr Gly Gln

 661 TGT TCA GAA TTA TGT GGT CTA GCT CAT TCT GCT ATG CCA ATC AAA ATT GAA GCA GTA TCA
     Cys Ser Glu Leu Cys Gly Thr Ala His Ser Ala Met Pro Ile Lys Ile Glu Ala Val Ser

 721 TTA CCT GCT TTC TTA GAA TGA TTA AAC GAA CAA TAA TTTATATAACTCTACGGATTTATATATATTT
     Leu Pro Ala Phe Leu Glu Trp Leu Asn Glu Gln ***

     TCGTAAATAATATCAATATTAAATAATATAGTAGTATATATCTATACATAACGGC -3'
```

Fig. 4. Nucleotide sequence of the *COX2* gene from *K. thermotolerans* mtDNA. Numbering starts at the first base of the ATG initiation codon. The sequence has been translated according to the baker's yeast mitochondrial genetic code where TGA codes for tryptophan and CTN for threonine. (See text for qualifications.) A putative transcription initiation site, identical to the consensus sequence for mtDNA promoters in yeasts, is *underlined.*

ilar experiments 10 large genes have been located on the 23.2-kb circular mtDNA (Fig. 2, Table 3). The map also illustrates the position of key restriction enzyme sites including those with unique sites: *Sal*I, *Pst*I, *Bam*HI, *Eco*RI, *Stu*I, *Kpn*I, and *Mlu*I.

From the size of the region hybridizing to the *COX1* probes it can be deduced that this gene contains introns. It is also probable that *LSU* contains the r1 or omega intron as mtDNA from a different strain of *K. thermotolerans* has this sequence (Jacquier and Dujon 1983). It is less likely that *CYB* or other genes contain introns based on the size of the regions hybridizing to the various probes. However, the most notable feature of the map is that only one rearrangement exists between this genome and that of *C. glabrata* where the *VARI* gene is located between *SSU* and *LSU* (Clark-Walker et al. 1985).

Conservation of gene order could be due to recent divergence of these yeasts or to stability of the genomes in species that are only distantly related. To decide between these alternatives we determined the sequence of the *COX2* gene of *K. thermotolerans* for use in construction of a phylogenetic tree to establish relationships between yeasts whose mtDNAs have been mapped.

Sequence of the COX2 *Gene*

From previous determinations of *COX2* sequences from other yeast mtDNAs, oligonucleotides were synthesized for conserved regions (Table 2) and used as primers to obtain the sequence of the *K. thermotolerans* gene as illustrated in Fig. 3. The gene encoding 251 amino acids, together with the upstream region extending to the flanking *Hae*III site and a short downstream sequence, is shown in Fig. 4. A conserved nonomer motif is present upstream of the coding region that is probably a promoter. Similar motifs have been found upstream of the *COX2* gene in *S. cerevisiae, C. glabrata,* and *K. lactis* mtDNAs (Coruzzi et al. 1981; Clark-Walker et al. 1985; Hardy and Clark-Walker 1990) but not in the vicinity of the *COX2* gene from *Dekkera/Brettanomyces* mtDNAs (Hoeben et al. 1993).

Table 4. Codon usage in the COX2 gene

Amino acid	Codon	Number	Amino acid	Codon	Number
Ala	GCG	0	Lys	AAG	0
	A	5		A	7
	C	0			
	T	11	Met	ATG	4
Arg	AGG	0	Phe	TTC	4
	A	5		T	9
	CGG	0			
	A	0	Pro	CCG	0
	C	0		A	4
	T	0		C	0
				T	9
Asn	AAC	4			
	T	7	Ser	TCG	0
				A	8
Asp	GAC	0		C	0
	T	14		T	4
				AGC	0
Cys	TGC	0		T	2
	T	3			
			Thr	ACG	0
Gln	CAG	0		A	9
	A	10		C	0
				T	3
Glu	GAG	0		CTG	0
	A	15		A	2
				C	0
Gly	GGG	0		T	1
	A	3			
	C	0	Trp	TGG	0
	T	8		A	5
His	CAC	0	Tyr	TAC	2
	T	5		T	11
Ile	ATA	0	Val	GTG	0
	C	3		A	10
	T	25		C	1
				T	8
Leu	TTG	0			
	A	30		Total	251

Codon Utilization

As in other yeast mtDNAs TGA is employed five times in *K. thermotolerans* to encode tryptophan (Table 4). Other triplets have been translated using the universal code except for CTN, which has been translated as threonine following the utilization of this codon in *S. cerevisiae*. However, the assignment of threonine to CTN codons is problematic because CTA only occurs twice at amino acid locations 62 and 227 and CTT once at position at 23. In *S. cerevisiae COX2* these positions are occupied by valine, threonine, and cysteine, respectively (Corruzi and Tzagoloff 1979). A second aberrant codon in *S. cerevisiae* is ATA coding for methionine but normally translated as isoleucine. This codon is not present in the *COX2* sequence of *K. thermotolerans*, but in the *LSU* r1 intron open reading frame from a different strain of this yeast the only ATA codon occurs at a location occupied by ATG (met) in *S. cerevisiae* (Jacquier and Dujon 1983). Other aspects of the data suggest that a second tRNA gene for translation of arginine is not present due to the absence of CGN codons, and G in the third position is never used apart from the methionine codon. Absence of a tRNA gene for translation of CGN has been described in *C. glabrata* mtDNA (Clark-Walker et al. 1985) and bias against the presence of codons with G at the third position is a universal property of yeast mtDNAs (Clark-Walker 1992; Hoeben et al. 1993).

Phylogenetic Analysis

Nucleotide sequences of the *COX2* gene from six yeast mtDNAs are shown in Fig. 5. The amino acid sequences were deduced from the nucleotide sequence using the translation rules outlined above. Since the 15-amino-acid NH_2-terminal *COX2* leader peptide is highly variable it was not used for tree reconstruction. No gaps had to be introduced for the alignment of the remaining core sequences. We inferred FJD distances from the amino acid alignment and nucleotide distances from the nucleotide alignment where we applied the Jukes and Cantor correction, the Kimura 2-, 3-, 4-, and 6-parameter corrections, and no correction at all. Because the six-parameter model allows for varying frequencies of different types of nucleotide changes it should suit our data best. Nevertheless all neighbor-joining trees constructed from these distance matrices have identical topology, varying only slightly in branch lengths. Consequently we show only the six parameter—and the FJD tree in Fig. 6. *Eeniella nana* was included to root the tree (Table 5).

From the figure it can be seen that the two *Kluyveromyces* species are grouped together as are the *Saccharomyces* species. Although the *C. glabrata* sequence clusters with the latter species the shallow branch responsible for this clade should be noted. However, the conserved gene order between the three small mtDNAs in the two separate clades indicates that this trait is ancestral for the five genomes.

An additional point from the nucleotide sequence comparisons (Fig. 5) is that, in general, transversions exceed transitions by 2:1. This property is also found in other yeast mtDNAs, including sequences that differ by as little as 2% (Hoeben et al. 1993).

Discussion

The circular 23.2-kb mitochondrial genome of *K. thermotolerans* is among the smallest in yeasts (Clark-Walker 1992). Furthermore, there is only one rearrangement between this genome and the 19-kb molecule in *C. glabrata*. In the former mtDNA, genes for *SSU* and *LSU* are juxtaposed, whereas in the latter these genes are separated by *VAR1*. When the gene topologies of these two molecules are viewed in conjunction with the phylogenetic tree constructed from *COX2* sequence com-

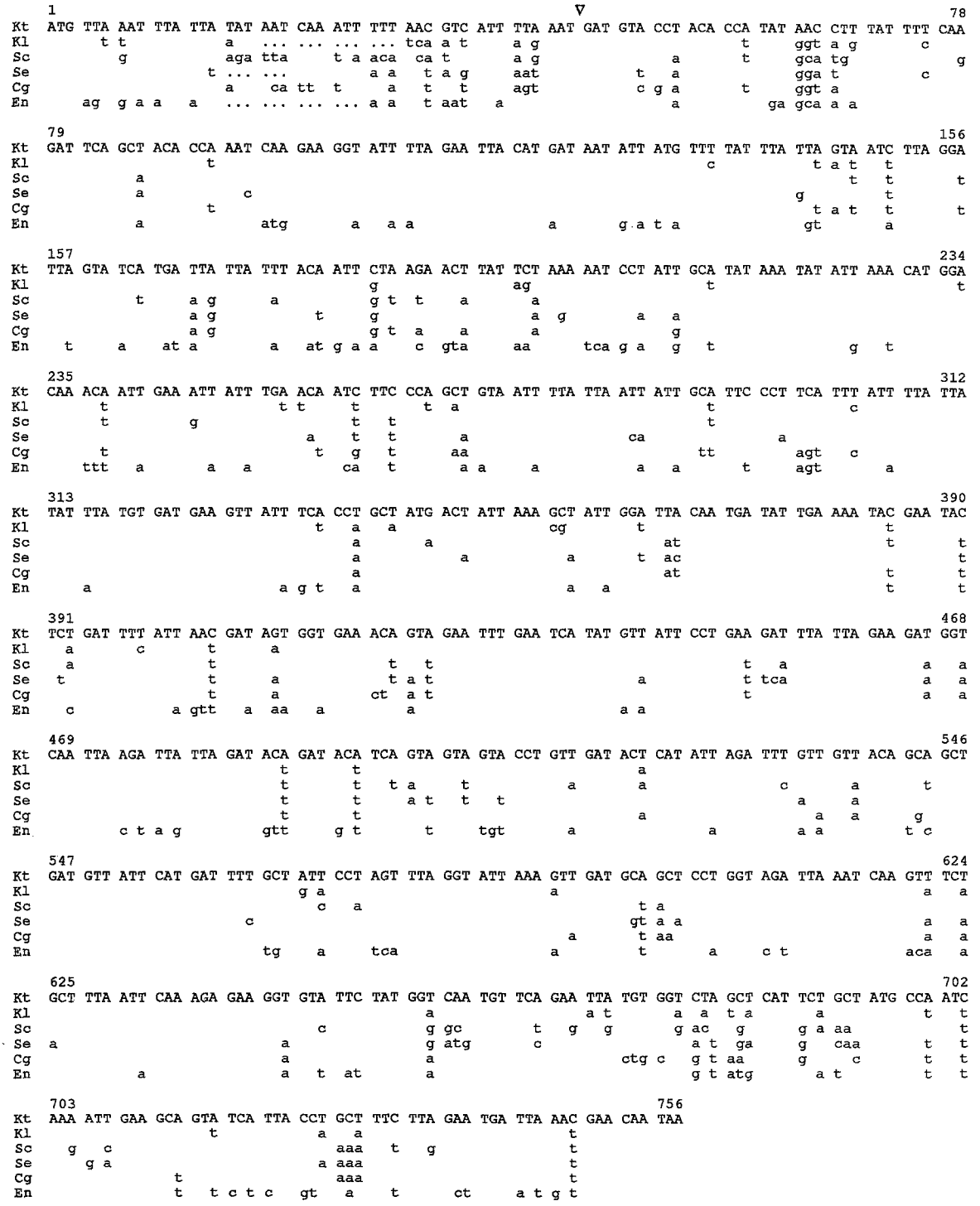

Fig. 5. Alignment of *COX2* sequences. Sources of sequence data: *K. thermotolerans* (Kt) this work, EMBL accession number X69431, *K. lactis* (K1) type A mtDNA (Hardy and Clark-Walker 1990), *S. cerevisiae* (Sc) (Coruzzi and Tzagoloff 1979), *S. exiguus* (Se) (Clark-Walker unpublished data), EMBL accession number X69429, *C. glabrata* (Cg) (Clark-Walker unpublished data), EMBL accession number X69430, *E. nana* (En) (Hoeben et al. 1993). The *arrowhead* shows the start of the core sequence corresponding to the coding region for the mature polypeptide of *S. cerevisiae COX2* (Pratje et al. 1983), which has been used for construction of the phylogenetic tree in Fig. 6.

parisons, we conclude that the structure of mtDNA in the common ancestor of the five species was similar to the two smaller genomes. By contrast, it appears that the larger *K. lactis* and *S. cerevisiae* mtDNAs, which differ quite markedly in their architecture, are the products of independent rearrangements.

In the distantly related yeasts belonging to the genus *Dekkera* or the anamorphic genus *Brettanomyces,* it has also been observed that the structure of smaller mtDNAs is more likely to represent the architecture of the ancestral molecule. In these yeasts the three largest genomes (57–101 kb) show rearrangements, yet they are

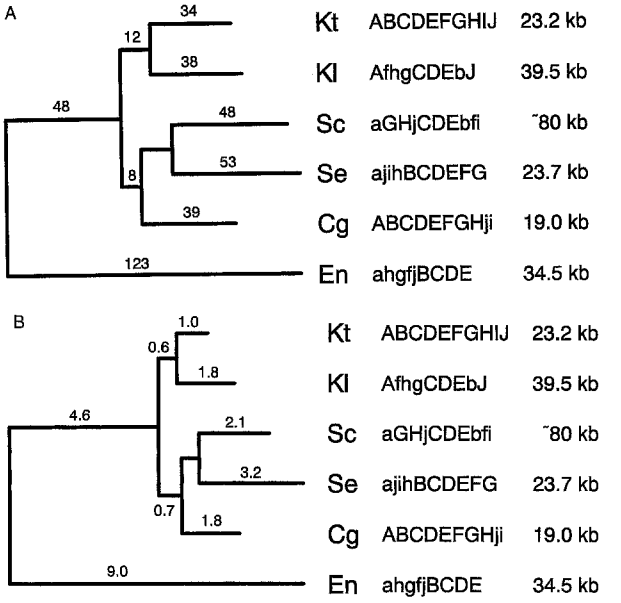

Fig. 6. Neighbor-joining trees using the nucleotide distances **(A)** and FDJ protein distances **(B)** of Table 5. Branch lengths are indicated above each branch. Denomination of yeast is as in Fig. 5. The gene order is given as follows: Large SU ribosomal RNA (A), cytochrome b (B), cytochrome oxidase SU1 (C), ATPase SU8 (D), ATPase SU6 (E), ATPase SU9 (F), cytochrome oxidase SU2 (G), cytochrome oxidase SU3 (H), VAR1 (I), small SU RNA (J), with genes in the most common order in uppercase letters. Gene orders and genome sizes are taken from: (Kt) this study, (Kl) type A mtDNA (Skelly et al 1991), (Sc) (Dujon 1981), (Se) (Clark-Walker et al. 1983), (Cg) (Clark-Walker and Sriprakash 1981), (En) (Hoeben and Clark-Walker 1986). Note that *VAR1* cannot be detected by hybridization in the En and Kl mtDNAs.

Table 5. Distance data and FDJ distances[a]

	Kl	Kt	Sc	Se	Cg	En
Kl		0.0725	0.1224	0.1328	0.0967	0.2222
Kt	2.71		0.1049	0.1143	0.1029	0.2210
Sc	5.87	3.92		0.1022	0.1001	0.2523
Se	6.99	6.19	5.31		0.10809	0.2442
Cg	4.31	4.70	4.70	5.20		0.2177
En	15.88	15.03	16.80	18.20	15.88	

[a] The upper right half matrix contains the distance data used for the tree in Fig. 6a. These distances are corrected for multiple hits using the six-parameter model. (See Materials and methods.) The lower left matrix contains the FDJ distances used for the tree in Fig. 6b. Denomination of yeasts is as in Fig. 5

more closely related than the three smaller mtDNAs (28–42 kb) with identical gene order (Hoeben et al. 1993). Extending this methodology to deduce the ancestral mitochondrial genome in a progenitor of both the *Dekkera/Brettanomyces* group and the *Saccharomyces/Kluyveromyces* yeasts shows that the only linkage preserved is *CYB-COX1-A6*. This indicates that mtDNA rearrangements have occurred in one or both lineages since divergence. Guidance in deducing which type of molecule is likely to have preserved the more ancient structure comes from two sets of observations that concern the conservation of bacterial operons and the preservation of gene order in a small plastid genome of a non-photosynthetic plant.

In the first instance, mtDNA of *K. thermotolerans* has three linkage groups that are found in bacterial genomes. Thus the arrangement of the *SSU* and *LSU* genes reflects the position of these genes in bacterial operons. Similarly, the order and orientation of genes in two other regions of mtDNA from *K. thermotolerans* and *C. glabrata* are reminiscent of gene arrangements in bacteria. In the *atp* operon of *E. coli,* analogous genes to the ATPase A6 and A9 have the same order and orientation (Downie et al. 1981; Gay and Walker 1981), while in *Paracoccus denitrificans, COX2* and *COX3* have the same order and orientation although they are separated by three open reading frames (Raitio et al. 1987). These similarities suggest that during the evolution of mtDNA in *K. thermotolerans*—as a result of overwhelming loss of genetic information from an endosymbiont—major rearrangements have not occurred.

Likewise, lack of rearrangement during formation of a smaller genome has been found by sequence analysis of the plastid genome in the parasitic plant *Epifagus*. The 70-kb molecule lacks all genes for photosynthesis that are found in the 150-kb chloroplast DNA of flowering plants (Wolfe et al. 1992). Despite the extensive reduction in genome size, the remaining genes are present in the same relative order and orientation as in tobacco plastid DNA. Hence it appears that the numerous small deletions needed to create the *Epifagus* plastid DNA are not accompanied by rearrangements. By extrapolation, it can be envisaged that deletions to an endosymbiont's genome during evolution are not a source of rearrangements.

As discussed in a recent review (Clark-Walker 1992), changes in gene order in mitochondrial genomes are viewed as more recent phenomena that sometimes accompany the chance acquisition of new sequence elements. It has been shown experimentally that short, directly repeated, G+C–rich sequences in the mtDNA of *S. cerevisiae* promote the formation of subgenomic molecules (Weiller et al. 1991). These molecules can be intermediates in the formation of rearranged mitochondrial genomes (Clark-Walker 1989). Such G+C–rich sequence tracts are not present in the mtDNA of *C. glabrata* (Clark-Walker et al. 1985). Thus it is possible that molecules which have avoided the acquisition of recombinogenic sequence elements may preserve ancient linkage groups.

Recently an attempt has been made to infer phylogenetic distances of mitochondrial genomes by a quantitative analysis of topological changes (Sankoff et al. 1992). Clearly data from the present study and our previous results (Hoeben et al. 1993) would call into question the validity of this approach when applied to yeast mitochondrial genomes.

Acknowledgments. We thank Erika Wimmer for skilled technical assistance.

References

Barnett JA (1992) The taxonomy of the genus *Saccharomyces* Meyen *ex* Reess: a short review for non taxonomists. Yeast 8:1–23

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

Clark-Walker GD, McArthur CR, Daley DJ (1981) Does mitochondrial DNA length influence the frequency of spontaneous petite mutation in yeast? Curr Genet 4:7–12

Clark-Walker GD, Sriprakash KS (1981) Sequence rearrangements between mitochondrial DNAs of *Torulopsis glabrata* and *Kloeckera africana* identified by hybridization with six polypeptide encoding regions from *Saccharomyces cerevisiae* mitochondrial DNA. J Mol Biol 151:367–387

Clark-Walker GD, McArthur CR, Sriprakash KS (1983) Order and orientation of genic sequences in circular mitochondrial DNA from *Saccharomyces exiguus:* implications for evolution of yeast mtDNAs. J Mol Evol 19:333–341

Clark-Walker GD, Sriprakash KS (1983) Analysis of a five gene cluster and unique orientation of large genic sequences in *Torulopsis glabrata* mitochondrial DNA. J Mol Evol 19:342–345

Clark-Walker GD, McArthur CR, Sriprakash KS (1985) Location of transcriptional control signals and transfer RNA sequences in *Torulopsis glabrata* mitochondrial DNA. EMBO J 4:465–473

Clark-Walker GD (1989) *In vivo* rearrangement of mitochondrial DNA in *Saccharomyces cerevisiae.* Proc Natl Acad Sci USA 86: 8847–8851

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

Coruzzi G, Tzagoloff A (1979) Assembly of the mitochondrial membrane system. DNA sequence of subunit 2 of yeast cytochrome oxidase. J Biol Chem 254:9324–9330

Coruzzi G, Bonitz SG, Thalenfeld BE, Tzagoloff A (1981) Assembly of the mitochondrial membrane system. Analysis of the nucleotide sequence and transcripts in the *OXI1* region of yeast mitochondrial DNA. J Biol Chem 256:12780–12787

Downie JA, Cox GB, Langman L, Ash G, Becker M, Gibson F (1981) Three genes coding for subunits of the membrane sector (Fo) of the *Escherichia coli* adenosine triphosphatase complex. J Bacteriol 145:200–210

Dujon B (1981) Mitochondrial genetics and functions. In: Strathern JN, Jones EW, Broach JR (eds) The molecular biology of the yeast *Saccharomyces cerevisiae*—life cycle and inheritance. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp 505–635

Feng D-F, Johnson MS, Doolittle RF (1985) Aligning amino acid sequences: comparison of commonly used methods. J Mol Evol 21: 112–125

Feng D-F, Doolittle RF (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J Mol Evol 25:351–360

Gay NJ, Walker JE (1981) The atp operon: nucleotide sequence of the promoter and genes for the membrane proteins, and the δ subunit of *Escherichia coli* ATP-synthase. Nucleic Acids Res 9:3919–3926

Gojobori T, Moriyama EN, Kimura M (1990) Statistical methods for estimating sequence divergence. Methods Enzymol 183:531–550

Hardy CM, Galeotti CL, Clark-Walker GD (1989) Deletions and rearrangements in *Kluyveromyces lactis* mitochondrial DNA. Curr Genet 16:419–428

Hardy CM, Clark-Walker GD (1990) Nucleotide sequence of the cytochrome oxidase subunit 2 and val-tRNA genes and surrounding sequences from *Kluyveromyces lactis* K8 mitochondrial DNA. Yeast 6:403–410

Hoeben P, Clark-Walker GD (1986) An approach to yeast classification by mapping DNA from *Dekkera/Brettanomyces* and *Eeniella* genera. Curr Genet 10:371–379

Hoeben P, Weiller GF, Clark-Walker GD (1993) Larger rearranged mitochondrial genomes in *Dekkera/Brettanomyces* yeasts are more closely related than smaller genomes with a conserved gene order. J Mol Evol 36:263–269

Hudspeth MES, Ainley WM, Shumard DS, Butow RA, Grossman LI (1982) Location and structure of the var1 gene on yeast mitochondrial DNA: nucleotide sequence of the 40.0 allele. Cell 30: 617–626

Jacquier A, Dujon B (1983) The intron of the mitochondrial 21 S rRNA gene: distribution in different yeast species and sequence comparison between *Kluyveromyces thermotolerans* and *Saccharomyces cerevisiae.* Mol Gen Genet 192:487–499

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munroe HN (ed) Mammalian protein metabolism II. Academic Press, New York, pp 21–132.

Kimura M (1981) Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA 78:454–458

McClelland M, Hanish J, Nelson M, Patel Y (1988) KGB: a single buffer for all restriction endonucleases. Nucleic Acids Res 16:364

Macino G, Tzagoloff A (1979) Assembly of the mitochondrial membrane system. The DNA sequence of a mitochondrial ATPase gene in *Saccharomyces cerevisiae.* J Biol Chem 254:4617–4623

Macino G, Tzagoloff A (1980) Assembly of the mitochondrial membrane system: sequence analysis of a yeast mitochondrial ATPase gene containing the *oli*2 and *oli*4 loci. Cell 20:507–517

Macreadie JG, Novitski CE, Maxwell RJ, John U, Ooi B-G, McMullen GL, Lukins HB, Linnane AW, Nagley P (1983) Biogenesis of mitochondria: the mitochondrial gene (*aap1*) coding for mitochondrial ATPase subunits in *Saccharomyces cerevisiae.* Nucleic Acids Res 11:4435–4451

Nobrega FG, Tzagoloff A (1980) Assembly of the mitochondrial membrane system. DNA sequence and organization of the cytochrome b gene in *Saccharomyces cerevisiae* D273-10B. J Biol Chem 255:9828–9837

Pratje E, Mannhaupt G, Michaelis G, Beyreuther K (1983) A nuclear mutation prevents processing of a mitochondrially encoded membrane protein in *Saccharomyces cerevisiae.* EMBO J 2:1049–1054

Raitio M, Jalli T, Saraste M (1987) Isolation and analysis of the genes for cytochrome c oxidase in *Paracoccus denitrificans.* EMBO J 6:2825–2833

Saitou N, Nei M (1987) The neighbour-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sankoff D, Guillaume L, Antoine N, Paquin B, Lang BF, Cedergren R (1992) Gene order comparisons for phylogenetic inference: evolution of the mitochondrial genome. Proc Natl Acad Sci USA 89:6575–6579

Skelly PJ, Hardy CM, Clark-Walker GD (1991) A mobile group II intron of a naturally occurring rearranged mitochondrial genome in *Kluyveromyces lactis.* Curr Genet 20:115–120

Smith M Th, Yamazaki M, Poot GA (1990) *Dekkera, Brettanomyces* and *Eeniella:* electrophoretic comparison of enzymes and DNA-DNA homology. Yeast 6:299–310

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

Takahata N, Kimura M (1981) A model of evolutionary base substitutions and its application with special reference to rapid change of pseudogenes. Genetics 98:641–657

Thalenfeld BE, Tzagoloff A (1980) Assembly of the mitochondrial membrane system. Sequence of the *oxi2* gene of yeast mitochondrial DNA. J Biol Chem 255:6173–6180

Tian G-L, Macadre C, Kruszewska C, Szczesniak B, Ragnini A, Grisanti P, Rinaldi T, Palleschi C, Frontali L, Slonimski PP, Lazowska J (1991) Incipient mitochondrial evolution in yeasts I. The physical map and gene order of *Saccharomyces douglassii* mitochondrial DNA discloses a translocation of a segment of 15,000 base-pairs and the presence of new introns in comparison with *Saccharomyces cerevisiae*. J Mol Biol 218:735–746

Weiller GF, Bruckner H, Kim SH, Pratje E, Schweyen RJ (1991) A GC-cluster repeat is a hotspot for mit$^-$ macro-deletions in yeast mitochondrial DNA. Mol Gen Genet 226:233–240

Wolfe KH, Mordeu CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a non-photosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648–10652