Microbiology (2000), **146**, 1045–1051 Printed in Great Britain

Two distinct 18S rRNA secondary structures in *Dipodascus* (Hemiascomycetes)

Kumiko Ueda-Nishimura and Kozaburo Mikata

Author for correspondence: Kumiko Ueda-Nishimura. Tel: +81 6 6300 6555. Fax: +81 6 6300 6814. e-mail: kumiueda@mb.infoweb.ne.jp

Institute for Fermentation, Osaka, 17-85, Juso-honmachi 2-chome, Yodogawa-ku, Osaka 532-8686, Japan

The nucleotide sequences of the 18S rRNA gene from ascomycetous yeast-like fungi in the genera *Dipodascus*, *Galactomyces* and *Geotrichum* were determined and the tested strains were separated into two groups by sequence length. In group 1, the length and secondary structure of 18S rRNA corresponded to those of typical eukaryotes. In group 2, the 18S rRNA gene sequences were about 150 nt shorter than those of most other eukaryotes and the predicted secondary structure lacked helices 10 and E21-5. Many substitutions and some deletions in group 2 18S rRNA gene were not only found in variable regions, but also in regions that are highly conserved among ascomycetes. Despite the considerable differences in 18S rRNA gene sequence and secondary structure between group 2 and other fungi, including group 1, phylogenetic analysis revealed that groups 1 and 2 are closely related. These findings suggest that a number of deletions occurred in the 18S rRNA of the common ancestor of group 2 strains.

Keywords: 18S rRNA gene, 18S rRNA secondary structures, *Dipodascus*

INTRODUCTION

Dipodascus and *Galactomyces* are ascomycetous yeast-like fungi. Their anamorph genus is *Geotrichum*. All species strongly assimilate ethanol (de Hoog *et al*., 1986). *Dipodascus* is characterized by asci which have persistent walls and rupture in the apical region, and by smooth-walled ascospores with mucilaginous sheaths (de Hoog *et al*., 1986). In *Galactomyces*, on the other hand, the asci do not open at the apex and ascospores are verrucose with median rims (Redhead & Malloch, 1977). *Dipodascus* and *Galactomyces* are distinguishable by their physiological characters. Domain D1/D2 26S rRNA gene sequence analysis has shown that species of *Dipodascus* and its anamorph *Geotrichum* separate into two clades, one of which includes species of *Galactomyces* (Kurtzman & Robnett, 1995).

Since ribosomes are an indispensable component of the protein synthesis apparatus and their structures are strictly conserved, the DNA component of the small subunit ribosome has proved to be an important and useful molecular clock for quantifying evolutionary relationships between organisms (Woese *et al*., 1990). Generally, the rate of base substitutions, deletions or insertions in various regions of the rRNA gene is not uniform; some areas are highly conserved and unchanged through millions of years, some are highly variable and others are semiconserved (Johnson *et al*., 1988). 18S rRNA of most eukaryotes adopts a common secondary structure, having 48 universal helices (De Rijk *et al*., 1992). They are present in all small subunit rRNAs from Archaea, Bacteria, plastids and Eukarya except in those of the Microspora (Microsporidia) and in those of organisms classified as Parabasalia (*Trichomonas* and relatives), where some of these helices are missing (Van de Peer *et al*., 1999).

In this paper we report two distinct secondary structures of *Dipodascus* 18S rRNA, one typical of most eukaryotes and an unusual structure missing some helices, including universal helix 10.

METHODS

Isolation of genomic DNA. The strains studied are listed in Table 1. The strains were grown in YM broth at 24 or 17 °C (*Dipodascus albidus*). Genomic DNA was extracted and purified by the procedure of Holm *et al*. (1986).

Assimilation test. Assimilation tests were performed by standard methods (van der Walt & Yarrow, 1984).

Viability. Cultures were preserved at 5 °C on slanted YM agar with a cotton plug. After 6 months preservation, the cultures

The DDBJ accession numbers for the sequences reported in this paper are shown in Table 1.

Table 1. Strains studied, their 18S rDNA sequence base and accession numbers

Species	Strain		18S rRNA gene	
	IFO	CBS	DDBJ accession no.	Base no.*
Group 1				
Geotrichum klebahnii	10826^T	179.30	AB000641	1794
Dipodascus armillariae	10802^T	817.71	AB000639	1793
	10803	540.76	= AB000639	1793
	10804	458.83	= AB000639	1793
	10818	600.83	= AB000639	1793
Dipodascus macrosporus	10807^T	259.82	AB000640	1792
Dipodascus albidus	1984	766.85	AB000642	1790
Dipodascus geniculatus	10806^T	184.80	AB000644	1790
Dipodascus australiensis	10805^T	625.74	AB000643	1788
Dipodascus aggregatus	10816^T	175.53	AB000645	1784
Galactomyces reessii	10823^T	179.60	AB000646	1778
Galactomyces citri-aurantii	10821	176.89	AB000664	1777
	10822	175.89	AB000665	1777
Galactomyces geotrichum	9541^T	772.71	AB000647	1777
Geotrichum candidum	4599		AB000652	1776
	5959		= AB000652	1776
Geotrichum fermentans	1199^T	439.83	AB000651	1776
Geotrichum eriense	10584^T	694.83	AB000648	1773
Group 2				
Dipodascus ambrosiae	10801^T	749.85	AB000638	1636
Dipodascus ovetensis	1201	634.85	AB000657	1636
	10827	635.85	= AB000657	1636
Dipodascus capitatus	10819	197.35	AB000650	1635
	10820	580.82	= AB000650	1635
Dipodascus spicifer	10809^T	244.85	AB000649	1635
Geotrichum clavatum	10824^T	425.71	AB000655	1635
Dipodascus magnusii	10808	234.85	AB000653	1634
Dipodascus tetrasperma	10810^T	765.70	AB000654	1634
Geotrichum fragrans	10825^T	152.25	AB000656	1634

*Base number of 18S rRNA gene PCR products from primer set 18-F/18-R.

were transplanted into the new medium and incubated at a given temperature. Viability was evaluated by assessing growth.

18S rRNA gene sequencing. 18S rRNA gene sequences were determined with a Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech) following the supplier's protocol. PCR for amplification of the 18S rRNA gene and cycle sequencing using PCR products were performed as described by Ueda-Nishimura & Mikata (1999).

The 18S rRNA gene sequences determined in this study were deposited in DDBJ with the accession numbers listed in Table 1.

Prediction of 18S rRNA secondary structure. 18S rRNA secondary structure was predicted following the model of De Rijk *et al*. (1992). GENETYX-MAC v. 8 (Software Development) was used to help predict stem structures in the area containing many deletions in the sequence.

Phylogenetic analysis. Sequence data were manually aligned with various 18S rRNA sequences of representatives of related genera obtained from GenBank. Positions that could not be compared among all sequences, corresponding to unknown bases, deletions and insertions, and regions of ambiguity in the total alignment were removed before performing the phylogenetic analysis. A phylogenetic tree was constructed by Kimura's two-parameter method (Kimura, 1980) and the neighbour-joining method (Saitou & Nei, 1987) using CLUSTAL v (Higgins *et al*., 1992). Bootstrap values (Felsenstein, 1985) were calculated from 1000 replicates.

RESULTS

18S rRNA gene sequences and secondary structure of 18S rRNA

18S rRNA gene sequences of *Dipodascus*, *Galactomyces* and *Geotrichum* species were amplified *in vitro* by PCR and their nucleotide sequences were determined directly. Multiple strains of *Dipodascus armillariae*, *Galactomyces citri-aurantii*, *Geotrichum candidum*, *Dipodascus ovetensis* and *Dipodascus capitatus* were examined

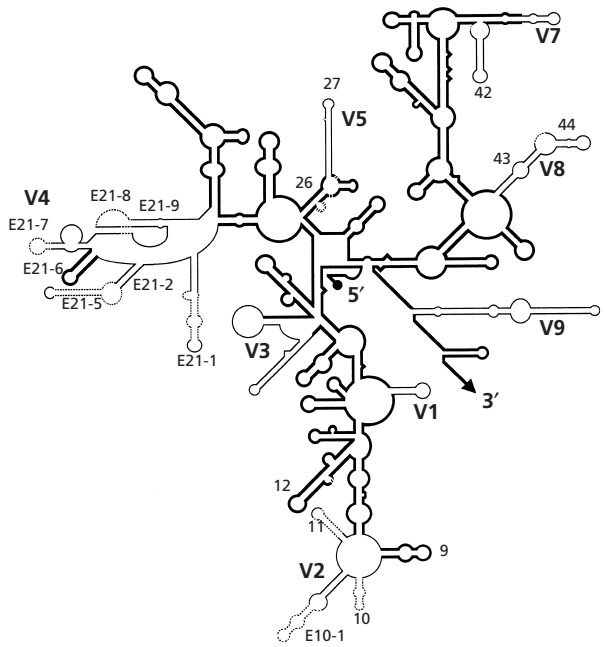

Fig. 1. Secondary structure model for eukaryotic 18S rRNAs. The 5′ terminus is symbolized by a dot, the 3′ terminus by an arrowhead. Areas of relatively conserved structure are indicated by bold lines, variable areas by thin lines. Deleted regions of group 2 strains compared with group 1 strains are indicated by dotted lines. This figure is modified from De Rijk *et al*. (1992).

(Table 1) and sequence identity was found for all species except *Gal. citri-aurantii*, in which a single base substitution between IFO 10821 and IFO 10822 was discovered at nt 1698 (nucleotide positions according to Mankin *et al*., 1986). The sequence of *Dipodascus ambrosiae* IFO 10801 was found to be identical to those of *D. ovetensis* IFO 1201 and IFO 10827. These newly determined sequences were separated by length into two groups (Table 1). Group 1 sequences are 1773–1794 nt in length, corresponding to the expected length for most eukaryotes, whereas group 2 sequences are 1634–1636 nt in length, i.e. smaller than most other eukaryotes. Group 1 consists of six *Dipodascus* species, three *Galactomyces* species and four *Geotrichum* species. Group 2 consists of six *Dipodascus* species and two *Geotrichum* species.

The 18S rRNA secondary structure of group 1 strains corresponds to that of typical eukaryotes, while that of group 2 strains differs from the general ascomycetous model (De Rijk *et al*., 1992) in that it lacks portions of the V2 and V4 variable regions (area numbering according to De Rijk *et al*., 1992) (Fig. 1). The major deletions correspond to whole helices 10 (helix numbering according to De Rijk *et al*., 1992) in the V2 region and E21-5 in the V4 region, and the ends of helices E10-1 and 11 in the V2 region and helix E21-1 in the V4 region (Fig. 2a, b). Helix 10 is deleted completely (Fig. 2a). UUG in the loop of helix 11 is conserved, but other parts of helix 11 are not conserved (Fig. 2a).

The 18S rRNA of group 2 strains has many small deletions. Most of these occur in the variable regions, but three minor deletions of 2–4 nt were found in the conserved structural region, corresponding to loops in helices (Fig. 1). Group 1 18S rRNA is more similar to *Saccharomyces cerevisiae* 18S rRNA than to that of group 2 organisms in some conserved regions around the deleted helices (Fig. 2). For example, Fig. 2(c) shows that the helix 43 sequence of *D. albidus* (group 1) is almost identical to the *S. cerevisiae* sequence, but not to the *D. ovetensis* sequence (group 2).

D-Xylose assimilation and viability

Group 1 strains assimilated D-xylose, but group 2 strains did not. D-Xylose assimilation was the only physiological character that differed between groups 1 and 2 (data not shown). Group 1 strains survived after 6 months on slanted YM agar medium at 5 °C, but again group 2 strains did not.

Phylogenetic analysis

Fig. 3. shows a phylogenetic tree calculated from published fungal 18S rRNA sequences. Basidiomycetous 18S rRNAs were chosen as the outgroup. Some regions, corresponding to positions 643–740 (V4), 1041–1070 (V5), 1336–1382 (V7) and 1680–1716 (V9), had too many gaps and substitutions for the sequences to be aligned. All alignment positions with a gap or an unknown residue and regions of ambiguous alignment were ignored for phylogenetic analysis. In this study 1389 positions were used for phylogenetic analysis. The tree shows that the tested strains form a monophyletic cluster supported at a bootstrap level of 95%. Group 1 species separate into three clusters. Species forming asci containing over 10 ascospores, i.e. *D. albidus*, *Dipodascus geniculatus* and *Dipodascus australiensis*, form subgroup 1a, supported at a bootstrap level of 100%. *Galactomyces geotrichum*, *Galactomyces reessii*, *Gal. citri-aurantii* and *Geo. candidum*, which is an anamorph of *Gal. geotrichum*, form subgroup 1b at a 100% bootstrap confidence level. The remaining species of group 1 form subgroup 1c. The *Galactomyces* strains are closely related to each other and form a monophyletic subgroup, 1b, making *Dipodascus* a paraphyletic taxon.

Despite dozens of base substitutions between *Gal. geotrichum* and *Geo. candidum*, they map to the same position on the phylogenetic tree (Fig. 3), as do *Gal. reessii* and *Gal. citri-aurantii*. However, the base substitutions are located in variable and deleted regions which were not used for the analysis because of ambiguous alignment. Accordingly, some distinctions between sequences could not be detected.

All species of group 2 are closely related to each other with a 100% bootstrap value and the stem of group 2 is very long. *D. ambrosiae* and *D. ovetensis* have identical 18S rRNA gene sequences. Only one base substitution

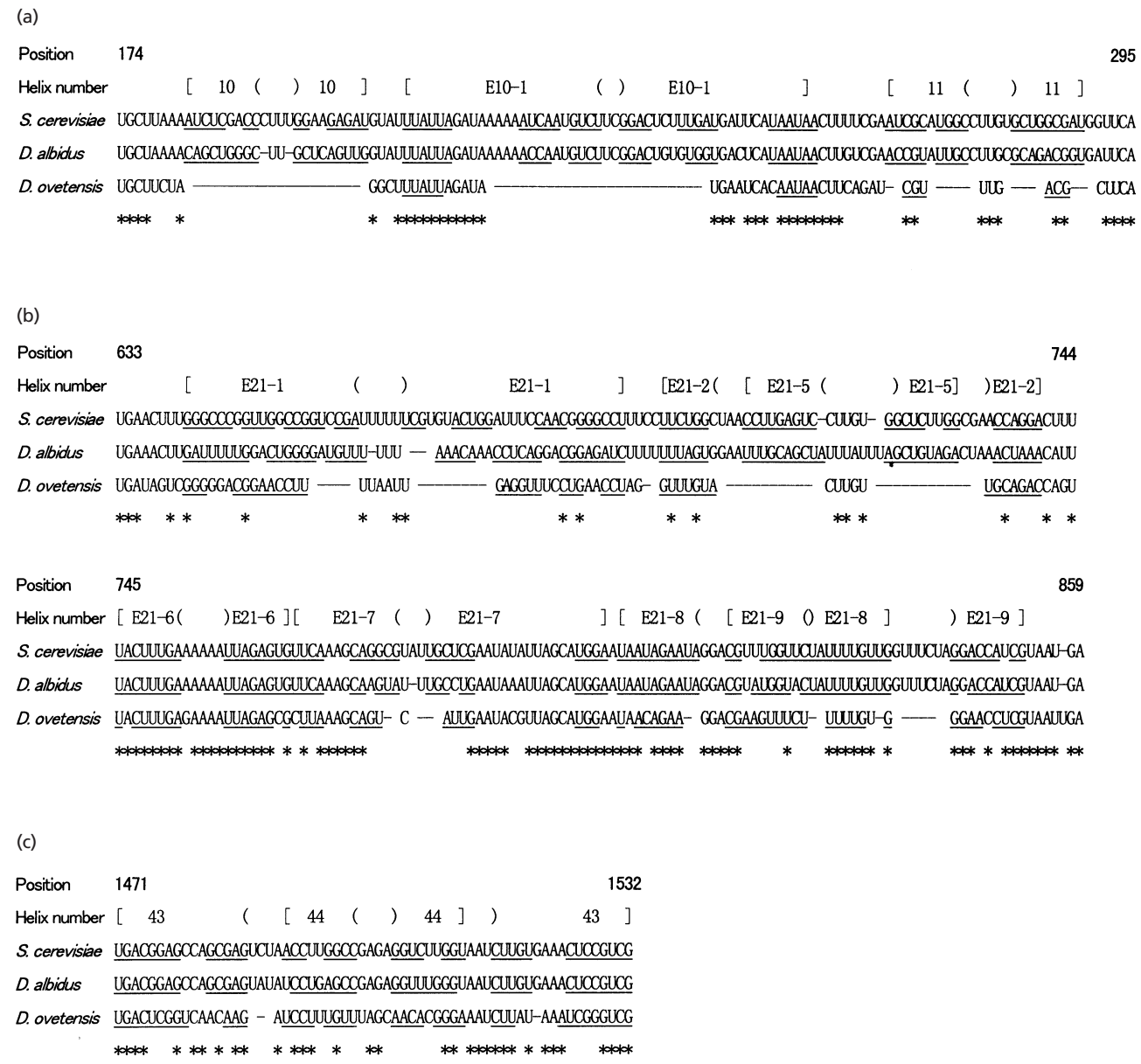

Fig. 2. Sequence deletions and substitutions in the variable regions of 18S rRNA among *S. cerevisiae*, *D. albidus* (group 1) and *D. ovetensis* (group 2): (a) V2 region, (b) V4 region, (c) V8 region. [*n* (and) *n*] indicates the upper and lower regions forming helix *n*. Helix numbering is according to De Rijk *et al.* (1992). Underlining indicates nucleotides forming hydrogen bonds with the opposite side. Asterisks represent conserved bases in the three sequences. Nucleotide positions are according to Mankin *et al.* (1986).

was found between *Dipodascus spicifer* and *Geotrichum clavatum*, and four between *Dipodascus magnusii* and *Geotrichum fragrans*. Three closely related species, *D. spicifer*, *D. capitatus* and *Geo. clavatum*, are thermotolerant, being able to grow at 40 °C (de Hoog *et al.*, 1986).

DISCUSSION

de Hoog *et al.* (1986) reported that *D. spicifer* CBS 244.85^T could assimilate D-xylose and that *D. spicifer* and *Geo. clavatum* are similar physiologically but distinguished by D-xylose assimilation. However, in this study *D. spicifer* IFO 10809^T (= CBS 244.85^T) (group 2) could not assimilate D-xylose and *D. spicifer* and *Geo. clavatum* could not be distinguished by D-xylose assimilation. *D. spicifer* and *Geo. clavatum*, whose 18S rRNA gene sequences are similar except for one substitution, have been suggested to be conspecific because of the similarity of their domain D1/D2 26S rRNA gene sequences (Kurtzman & Robnett, 1998). Therefore, it is suspected that the ability of *D. spicifer* to assimilate D-xylose as reported by de Hoog *et al.* (1986) is incorrect.

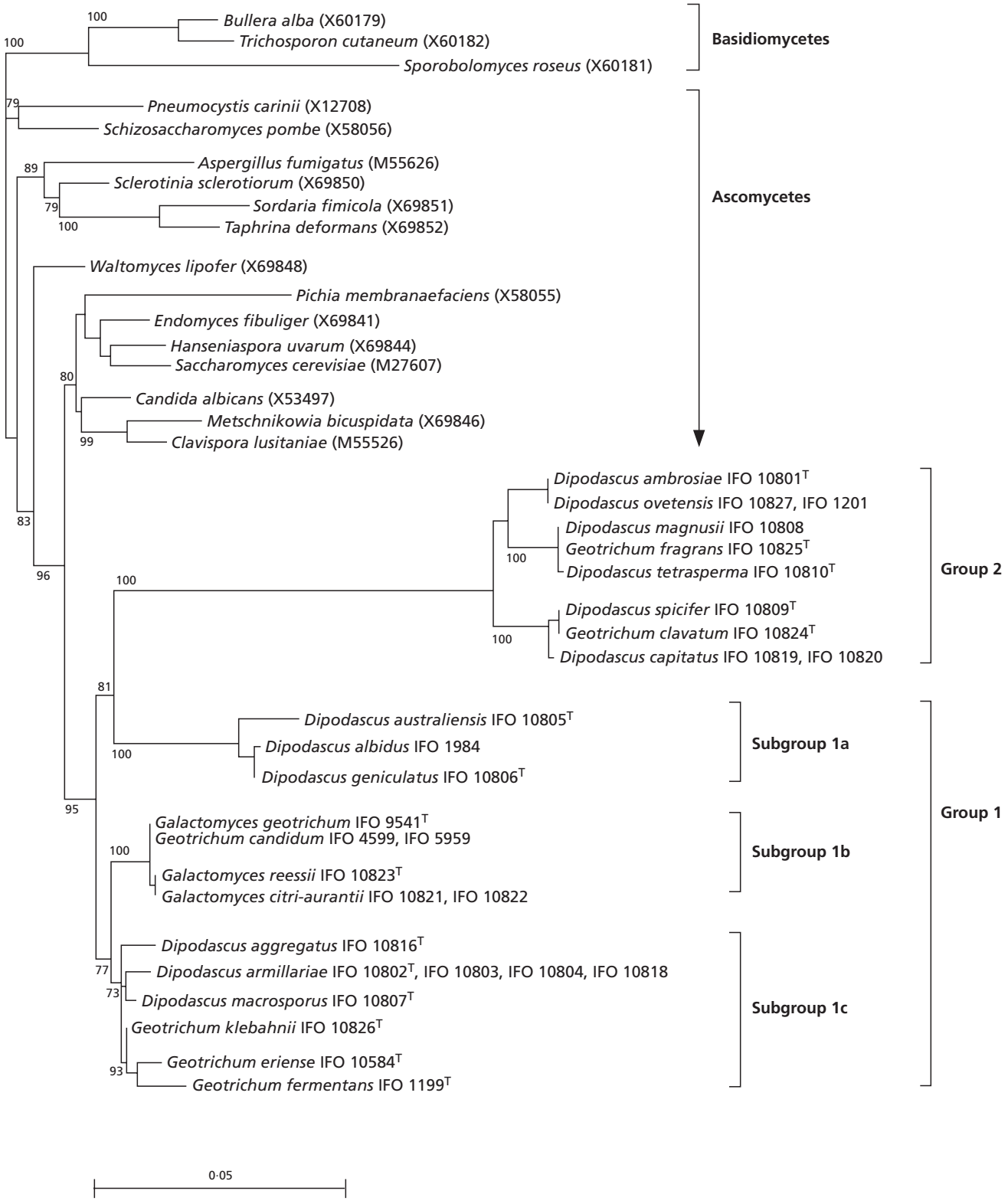

Fig. 3. Phylogenetic tree showing the relationships between species of the *Dipodascus* complex and related taxa. The tree is based on 18S rRNA gene sequence data and was constructed by using the neighbour-joining method. Bootstrap values were calculated from 1000 trees and values of $<70\%$ were omitted. GenBank, EMBL and DDBJ accession numbers are shown in parentheses. Bar, sequence dissimilarity of 5%.

It was reported that *D. ambrosiae* and *D. ovetensis* are superficially similar and distinguished by ascus size and G+C content (de Hoog *et al.*, 1986). However, they have been suggested to be conspecific because of the identity of their domain D1/D2 26S rRNA (Kurtzman & Robnett, 1998) and 18S rRNA gene sequences.

Between *Geo. candidum* (IFO 4599 and IFO 5959) and *Gal. geotrichum* (IFO 9541^T), there are 22 substitutions and one deletion in their 18S rRNA genes. Considering that *Geo. candidum* is an anamorph of *Gal. geotrichum* this difference is huge because the difference between the sequence of multiple strains of identical species, i.e. *D. armillariae*, *Gal. citri-aurantii*, *D. ovetensis* or *D. capitatus*, is usually only one substitution or no substitutions at all. Therefore, it is suggested that IFO 4599 and IFO 5959 should not be treated as *Geo. candidum*.

The predicted 18S rRNA secondary structure of group 2 species lacks helices 10 and E21-5. Helix 10 is one of the universal helices present in all hitherto known small subunit rRNAs from Archaea, Bacteria, plastids and Eukarya, but not the Microspora, which lack helices 10, 11 and 44 (De Rijk *et al.*, 1992). De Rijk *et al.* (1992) showed that 18S rRNA secondary structures of *Schizosaccharomyces pombe* and *Yarrowia lipolytica* were different from those of all other fungi. The 18S rRNAs of these species have an insert in helix E43-1, which is not found in other fungi, and *Y. lipolytica* 18S rRNA lacks helix E21-5. In this light, the lack of helices 10 and E21-5 in 18S rRNA of group 2 species shows them to be very different from other fungi.

The conserved regions are generally highly conserved among related organisms. Accordingly, it is unusual that many substitutions and some deletions in group 2 species are found not only in variable regions but also in regions that are conserved between ascomycetes. Fig. 2 shows that *D. albidus* (group 1) is more similar to *S. cerevisiae* than to *D. ovetensis* (group 2) in some conserved regions around the deleted helices. The phylogenetic analysis, using 1389 reliable positions, shows that groups 1 and 2 are related (Fig. 3). Therefore, it is suggested that the observed substitutions and deletions in the conserved regions of group 2 species happened after the branching of the *Dipodascus/Galactomyces* clade.

The grouping of species into group 1 (subgroups 1a, 1b and 1c except *Geotrichum fermentans*) and group 2 species is observed in the domain D1/D2 26S rRNA parsimony tree (Kurtzman & Robnett, 1995). Group 1 sequences of domain D1/D2 26S rRNA sequences from GenBank are 545–554 nt in length, corresponding to that of most ascomycetous yeasts (Kurtzman & Robnett, 1998), whereas group 2 sequences are 404–407 nt in length, i.e. smaller than those of most other ascomycetous yeasts. We suggest that there might be as many deletions in group 2 whole 26S rRNA as in 18S rRNA.

Group 1, as shown by Kurtzman & Robnett (1995), is monophyletic. However, subgroup 1a and group 2, as shown in Fig. 3, are joined. The branch leading to group 2 is much longer than for all other strains examined and the long branch indicates that the rate of evolutionary change is increased by substitutions in conserved regions. The branch of subgroup 1a is also relatively long. The joining of group 2 and subgroup 1a was therefore possibly due to a long-branch attraction (Felsenstein, 1978).

One possible explanation for the increased rate of substitutions is that a number of deletions may have occurred in the 18S rRNA gene of the common ancestor of group 2 strains by chromosomal rearrangement and homologous recombination. For 18S rRNA to interact with other rRNAs and ribosomal proteins for construction of the ribosome, the change in three-dimensional structure might have required many base substitutions in the general conserved regions. In addition to deletions in the 18S rRNA gene, it is expected that the length of the 26S rRNA gene of group 2 species might be shorter than that of group 1 species. Furthermore, group 2 species might have lost the ability to assimilate D-xylose and the ability of group 2 strains to survive on slanted agar medium was inferior to that of group 1. From these characteristics, it is suggested that the common ancestor of group 2 lost parts of its rDNA operon and the genes related to D-xylose assimilation and survival.

REFERENCES

De Rijk, P., Neefs, J.-M., Van de Peer, Y. & De Wachter, R. (1992). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* **20**, 2075–2089.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* **27**, 401–410.

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

Higgins, D. G., Bleasby, A. J. & Fuchs, R. (1992). CLUSTAL V: improved software for multiple sequence alignment. *Comput Appl Biosci* **8**, 189–191.

Holm, C., Meeks-Wagner, D. W., Fangman, W. L. & Botstein, D. (1986). A rapid, efficient method for isolating DNA from yeast. *Gene* **42**, 169–173.

de Hoog, G. S., Smith, M. Th. & Guého, E. (1986). A revision of the genus *Geotrichum* and its teleomorphs. *Stud Mycol* **29**, 1–131.

Johnson, A. M., Illana, S., Hakendorf, P. & Baverstock, P. R. (1988). Phylogenetic relationships of the apicomplexan protist *Sarcocystis* as determined by small subunit ribosomal RNA comparison. *J Parasitol* **74**, 847–860.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Kurtzman, C. P. & Robnett, C. J. (1995). Molecular relationships among hyphal ascomycetous yeasts and yeastlike taxa. *Can J Bot* **73** (**Suppl. 1**), S824–S830.

Kurtzman, C. P. & Robnett, C. J. (1998). Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Leeuwenhoek* **73**, 331–371.

Mankin, A. S., Skryabin, K. G. & Rubtsov, P. M. (1986). Identification of ten additional nucleotides in the primary structure of yeast 18S rRNA. *Gene* **44**, 143–145.

Redhead, S. A. & Malloch, D. W. (1977). The Endomycetaceae: new concepts, new taxa. *Can J Bot* **55**, 1701–1711.

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

Ueda-Nishimura, K. & Mikata, K. (1999). A new yeast genus, *Tetrapisispora* gen. nov.: *Tetrapisispora iriomotensis* sp. nov., *Tetrapisispora nanseiensis* sp. nov. and *Tetrapisispora arboricola* sp. nov., from the Nansei Islands, and reclassification of *Kluyveromyces phaffii* (van der Walt) van der Walt as *Tetrapisispora phaffii* comb. nov. *Int J Syst Bacteriol* **49**, 1915–1924.

Van de Peer, Y., Robbrecht, E., de Hoog, S., Cares, A., De Rijk, P. & De Wachter, R. (1999). Database on the structure of small subunit ribosomal RNA sequences. *Nucleic Acids Res* **27**, 179–183.

van der Walt, J. P. & Yarrow, D. (1984). Methods for isolation, maintenance, classification and identification of yeasts. In *The Yeasts*, *a Taxonomic Study*, 3rd edn. pp. 76–97. Edited by N. J. W. Kreger-van Rij. Amsterdam: Elsevier.

Woese, C. R., Kandler, O. & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* **87**, 4576–4579.

Received 13 September 1999; revised 29 November 1999; accepted 17 February 2000.