International Journal of Systematic Bacteriology (1998), 48, 295–303 Printed in Great Britain

Phylogenetic analysis of the *Saccharomyces cerevisiae* group based on polymorphisms of rDNA spacer sequences

Robert Montrocher,[1] Marie-Christine Verner,[1] Jérôme Briolay,[2] Christian Gautier[2] and Roland Marmeisse[1]

Author for correspondence: Robert Montrocher. Fax: +33 4 72 43 16 43.
e-mail: rmontroc@lercycle.univ-lyon1.fr

CNRS, UMR 5557 'Ecologie Microbienne du Sol', Bât. 405[1], and Centre d'Analyse Moléculaire de la Biodiversité[2], Université Lyon 1, 43 boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex, France

The phylogenetic relationships between species of yeasts assigned to the *Saccharomyces sensu stricto* group, which includes *Saccharomyces cerevisiae* and *Saccharomyces bayanus*, were studied together with *Saccharomyces pastorianus* and *Saccharomyces paradoxus*. The experimental approaches used were RFLP analysis of the PCR-amplified rDNA internal transcribed spacer (ITS) and intergenic spacer, and total ITS sequence analysis. Both RFLP and sequence analyses gave fairly similar results. The gene trees generated with either of the two data sets showed the distribution of the yeasts into two major, well-separated, phylogenetic clusters called '*cerevisiae*' and '*bayanus*'. The '*cerevisiae*' cluster included the *S. cerevisiae* type strain, together with most of the species (16 out of 23), whereas the '*bayanus*' cluster included the remaining seven type strains. Therefore, analysis of rDNA sequences confirmed *S. cerevisiae* and *S. bayanus* as two well-defined taxa. However, *S. pastorianus* and *S. paradoxus*, the two other usually accepted taxa of the now-defined *Saccharomyces sensu stricto* complex, could not be clearly separated from *S. bayanus* and *S. cerevisiae*, respectively. However, in both PCR-RFLP and ITS sequence analyses, *S. paradoxus* had the outermost position in the '*cerevisiae*' cluster. PCR-RFLP analysis of the ribosomal spacer sequences was also carried out on 26 *Saccharomyces* strains isolated in various wine-growing regions of France in an attempt to clarify their positions in the *Saccharomyces* phylogenetic tree. Compared to the diversity of the *Saccharomyces* type strains, less genetic diversity was detected among these yeasts and several of them exhibited identical RFLP patterns. Most of the wine yeast strains (16 out of 26) were closely related to each other and were found within the '*cerevisiae*' cluster. The remaining 10 wine yeast strains branched within the '*bayanus*' cluster. PCR-RFLP analysis of ribosomal spacer sequences thus appears to be a useful and appropriate method for the correct characterization of *Saccharomyces* yeast strains used in food processing.

Keywords: *Saccharomyces cerevisiae*, PCR-RFLP analysis of rDNA, rDNA ITS sequence analysis, molecular phylogeny

INTRODUCTION

Conventional methods used for the identification of yeast species rely extensively on the study of physiological characteristics such as the ability of isolates to assimilate and/or ferment different carbon sources. As expected, closely related species may differ from each other by a very limited number of such characteristics which, for some of them, may be controlled by single mutable genes. This is the case within the genus *Saccharomyces* Meyen *ex* Reess and more particularly for species closely related to *Saccharomyces cerevisiae*, which have been classically identified on their varying abilities to use less than 10 different carbon sources (1). The extreme proximity between species within this group has led, in the past, to many different changes in

Abbreviations: RAPD, randomly amplified polymorphic DNA; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; ITS, internal transcribed spacer; IGS, intergenic spacer; *d*, molecular distance value.

Table 1. *Saccharomyces sensu stricto* type strains studied

Species	CBS no.	Original name	Source	ITS type	IGS type	ITS sequence accession no.
S. bayanus	380^T	*S. bayanus*	Beer	I	I	Z95945
S. bayanus	395^T	*S. uvarum*	Currant juice	I	II	Z95946
S. cerevisiae	1171^T	*S. cerevisiae*	Beer	VI	VI	
S. cerevisiae	4054^T	*S. aceti*	Red wine	XII	XII	Z95931
S. cerevisiae	2247^T	*S. capensis*	Grape must	VII	XI	Z95937
S. cerevisiae	400^T	*S. chevalieri*	Palm wine	XV	XIII	Z95939
S. cerevisiae	5635^T	*S. coreanus*	Grape must	XV	XVII	Z95942
S. cerevisiae	1782^T	*S. diastaticus*	Infected wort	X	XVI	Z95935
S. cerevisiae	6006^T	*S. gaditensis*	Wine	V	X	Z95951
S. bayanus	424^T	*S. globosus*	Pear juice	IV	I	Z95947
S. bayanus	425^T	*S. heterogenicus*	Apple juice	III	I	Z95944
S. cerevisiae	4903^T	*S. hienipiensis*	'Alpechin'	XV	XIV	Z95940
S. bayanus	1546^T	*S. inusitatus*	Beer	I	I	Z95948
S. cerevisiae	459^T	*S. italicus*	Grape must	XV	VII	Z95938
S. cerevisiae	382^T	*S. logos*	Beer	XI	VIII	Z95936
S. cerevisiae	5378^T	*S. norbensis*	'Alpechin'	XV	XVII	Z95929
S. cerevisiae	3093^T	*S. oleaceus*	Olives	XIV	XVII	Z95943
S. cerevisiae	3081^T	*S. oleaginosus*	'Alpechin'	XV	XVII	Z95941
S. cerevisiae	5155^T	*S. prostoserdovi*	Wine	VIII	IX	Z95934
S. cerevisiae	423^T	*S. steineri*	Grape must	XIII	XV	Z95932
S. paradoxus	432^T	*S. paradoxus*		IX	V	Z95933
S. pastorianus	1538^T	*S. pastorianus*		II	III	Z95949
S. pastorianus	1513^T	*S. carlsbergensis*	Beer	II	IV	Z95950

their taxonomic status. The originally 21 recognized species of the *Saccharomyces sensu stricto* group (33) were grouped by Yarrow (40) in a single taxon: *S. cerevisiae*. According to Vaughan Martini & Martini (37), *S. cerevisiae sensu* Yarrow has now been split into three species: *S. cerevisiae*, *Saccharomyces bayanus* and *Saccharomyces pastorianus* which constitute, together with *Saccharomyces paradoxus* (35), the *Saccharomyces sensu stricto* complex. Furthermore, studies of the genes whose enzyme products allow the cells to utilize some of the carbon sources classically used for species identification show that the number and location of these genes in the genome can vary between strains of the same species. This is, for example, the case of the MEL genes (controlling melibiose fermentation) which can either be absent or present in one or multiple copies in the genomes of *S. cerevisiae* strains (20). This situation, which is also found for other yeast genes whose products protect the cells against toxic compounds (22), may reflect the adaptation of strains to local environmental conditions but not their taxonomic status.

The criterion originally used to separate the four above-mentioned species (*S. cerevisiae*, *S. bayanus*, *S. pastorianus* and *S. paradoxus*) was DNA relatedness, as measured in whole-cell DNA/DNA reassociation experiments (35, 36, 37). This classification was further substantiated by many other studies on electrophoretic karyotypes (2, 8, 18, 38), mitochondrial DNA (mtDNA) restriction endonuclease profiles (5), rDNA restriction analysis (12), analysis of randomly amplified polymorphic DNAs (RAPDs) (13), and by studies which showed that interspecific hybrids between these species, although viable, gave abortive ascospores (14).

Most of these studies were intended to identify the four different species and to classify different wild or industrial strains of *Saccharomyces*. Few of these studies (mtDNA and RAPD) allow a clear understanding of the phylogenetic relationships between these four different species. In this article, RFLP analysis and sequencing of nuclear rDNA spacers were carried out on the 21 original *Saccharomyces* species pooled by Yarrow (40) in one taxon. These data were used to infer phylogenetic trees. RFLP analysis was also carried out on 26 wine yeast strains to clarify their positions in the *Saccharomyces* phylogenetic tree.

In *S. cerevisiae*, as in most fungal species, the four different rDNA genes (5S, 5·8S, 18S and 26S) are grouped in repeats which are arranged head-to-tail to form a single cluster on chromosome XII. The 18S and 26S genes are separated from each other by the internal transcribed spacers (ITS) and intergenic spacers (IGS), which are non-coding sequences except for the short 5·8S and 5S genes, respectively. Both these spacer sequences can easily be amplified by PCR using oligonucleotide primers homologous to sequences present at the ends of the rDNA genes (39). A model of the secondary folding structure of the ITS sequence from *S. cerevisiae* has been obtained; this sequence has been divided into several structural domains, each

characterized by one or two loops (27). Interestingly, the sequences of ITS regions from related yeast species could be only partially aligned to the ITS sequence of *S. cerevisiae* (34), thus making this sequence a good candidate for the study of the molecular phylogeny of closely related yeast species.

METHODS

Yeast strains. The characteristics of the 49 strains studied are given in Tables 1 and 2. Twenty-three of them are the type strains of the species recognized by van der Walt (Table 1). Twenty-six are wine yeast strains collected in different regions of France (provided by C. Cuinier, ITV, Tours, France) (Table 2).

PCR-RFLP analysis. DNA was extracted from 5 ml yeast liquid cultures according to the method of Phillipsen *et al.* (25). Amplifications of the target sequences were carried out in a 50 µl volume containing 1–10 ng genomic DNA, 100 nM of each primer, 200 µM of each of the four dNTPs, 1 U *Taq* DNA polymerase and the appropriate buffer supplied by the manufacturer (Appligène). The primers used to amplify the ITS sequences were 3126T (5′-ATATGCTTAAGTTCAG-CGGGT-3′), which hybridized to the 5′ end of the 26S gene, and 2234C (5′-GTTTCCGTAGGTGAACCTGC-3′), which hybridized to the 3′ end of the 18S gene (31). The primers used to amplify the IGS sequences were 126 (5′-ACCAC-CTAGGACGGTCATCA-3′), which hybridized to the 3′ end of the 26S gene, and 127 (5′-GTAGAGTAGCCTTG-TTGTTACGATC-3′), which hybridized to the 5′ end of the 18S gene (4). For the ITS, after an initial denaturation at 95 °C for 2 min, the sequence was amplified by 35 cycles of 1 min at 95 °C, 1 min at 50 °C, and 1 min at 72 °C; this was followed by a final extension of 2 min at 72 °C. For the IGS, after an initial denaturation at 95 °C for 3 min, the sequence was amplified by 37 cycles of 2 min at 95 °C, 1 min at 55 °C, and 3 min at 72 °C; this was followed by a final extension of 7 min at 72 °C. Appropriate amounts of the amplified sequences were digested with the four-cutter enzymes *Alu*I, *Bst*UI, *Cfo*I, *Dde*I, *Hae*III, *Hpa*II, *Nde*II, *Rsa*I and *Taq*I, and *Hin*fI which has a 5 bp recognition site. Restriction fragments were separated by electrophoresis in 8% polyacrylamide gels for the ITS and in 3% agarose gels [2% NuSieve agarose (FMC Bioproducts), 1% agarose (Appligène)]. After electrophoresis, the gels were stained with ethidium bromide and photographed. For the ITS, only the restriction fragments larger than the 100 bp pBR322 *Hin*fI size marker were recorded; for the IGS, only those larger than the 75 bp pBR322 *Hin*fI size marker were recorded.

The relationships between the amplified sequences of the different *Saccharomyces* strains were evaluated by their fraction (F) of co-migrating fragments [$F = 2n_{xy}/(n_x + n_y)$, where n_{xy} is the number of co-migrating bands and n_x and n_y

Table 2. Origin and characteristics of the *Saccharomyces* wine yeast strains studied

Species*	Original name and ITV no.†	No. in Fig. 1	Origin§	ITS type	IGS type
S. bayanus	*S. uvarum* L1708	1	Sancerrois, 1986	I	XVIII
S. cerevisiae	*S. bayanus* L3336	2	Touraine, 1981	XVII	XXII
S. cerevisiae	*S. uvarum* L579	3	Beaujolais, 1986	XV	XXI
S. cerevisiae	*S. bayanus* L3571	4	Savoie, 1987	XV	XXI
S. bayanus	*S. uvarum* L19	5	Touraine, 1978	I	XVIII
S. bayanus	*S. uvarum* L99	7	Touraine, 1975	I	XVIII
S. bayanus	*S. uvarum* L490	8	Touraine, 1982	I	XVIII
S. cerevisiae	*S. uvarum* L1425	10	Beaujolais, 1986	XV	XXI
S. cerevisiae	*S. cerevisiae* L1642	11	Sancerrois, 1986	XV	XX
S. bayanus	*S. uvarum* L3583	12	Savoie, 1987	XVI	XVIII
S. cerevisiae	*S. bayanus* L206	13	Beaujolais, 1986	XVIII	XXI
S. cerevisiae	*S. cerevisiae* L1656	14	Charentes, 1986	XV	XXI
S. cerevisiae	*S. bayanus* L541	15	Beaujolais, 1985	XVIII	XXI
S. cerevisiae	*S. bayanus* L601	16	Beaujolais, 1985	XV	XXI
S. bayanus	*S. uvarum* L259	17	Touraine, 1975	I	XVIII
S. cerevisiae	*S. cerevisiae* L2835	18	Beaujolais, 1984	XV	XXI
S. cerevisiae	*S. cerevisiae* L649	19	Beaujolais, 1985	XV	XXI
S. cerevisiae	*S. cerevisiae* L3587	21	Savoie, 1987	XV	XXI
S. cerevisiae	*S. cerevisiae* L3380	22	Côtes du Rhône, 1981	XV	XIX
S. cerevisiae	*S. bayanus* L2006	24	Touraine, 1977	XV	XXI
S. bayanus	*S. uvarum* L85	25	Alsace, 1988	I	XVIII
S. cerevisiae	*S. uvarum* L2784	26	Beaujolais, 1984	XV	XX
S. bayanus	*S. pretoriensis* L3304	27	Touraine, 1981	I	XVIII
S. bayanus	*S. uvarum* L1609	28	Touraine, 1986	I	XVIII
S. cerevisiae	*S. bayanus* L1392	29	Touraine, 1985	XVII	XXII
S. bayanus	*S. cerevisiae* L926	30	Touraine, 1985	I	XVIII

* Species as defined in this study.

† ITV, Institut Technique de la Vigne et du Vin, Tours, France.

§ Wine-growing regions of France.

are the number of bands of strain x and strain y, respectively]. Molecular distances (*d*) were then estimated using the mathematical model of Nei & Li (21). Cluster analysis of these *d* values was performed using the neighbour-joining method (29). The calculations were performed using a hypercard compilation running on a Macintosh computer (Marc Neyra; ORSTOM, Dakar; Ecologie Microbienne CNRS, Lyon, France; unpublished).

Sequencing and data analysis. The ITS sequences were amplified by using the PCR conditions described above. The amplified products were purified and concentrated on Centricon columns (Amicon). The purified double-stranded PCR products were directly used for sequencing using the T7 sequencing kit from Pharmacia LKB based on the dideoxy chain-termination method (30) according to J. Briolay and others (unpublished). To sequence both strands, seven primers were used: 2234C and 3126T as described above; and AM2541 (5′-GCATCGATGAAGAACGCAGC-3′), AM2542 (5′-GCTGCGTTCTTCATCGATGC-3′), AM454 (5′-TCACTCACTACCAAACAGAA-3′), AM1701 (5′-TTTCCTTCTCAAACATTCTG-3′) and AM3393 (5′-CCTTACGGGTTTCTTTTTCAA-3′), which were deduced from the published sequence of the *S. cerevisiae* HA6 strain (GenBank accession no. U09327).

Sequences were aligned using CLUSTAL W (32) and manually refined using SEAVIEW (3). All regions corresponding to a gap in one sequence were excluded from the analyses. Matrix pairwise comparisons were corrected for multiple base substitutions by the two-parameter model of Kimura (10) and a phylogenetic tree was generated by the neighbour-joining method using PHYLO-WIN (3). Bootstrap analysis (1000 replications) was performed to assess the strength of the internal branches of the tree.

RESULTS

PCR-RFLP analysis of the rDNA spacers

The ITS and IGS sequences were amplified from the genomic DNA of all 49 studied yeast strains. In all cases, a single DNA fragment was amplified and, for each of these two DNA regions, no significant size variation could be detected between strains after agarose gel electrophoresis. The amplified ITS and IGS were approximately 780 bp and 3·2 kb long, respectively.

For the ITS, four of the 10 enzymes used (*Alu*I, *Bst*UI, *Dde*I and *Hpa*II) did not reveal any polymorphism between strains. Each of the six other enzymes gave variable numbers of restriction profiles, from four for *Nde*II and *Cfo*I to six for *Hin*fI. After the restriction profiles produced by the different enzymes had been combined, 18 different ITS-RFLP types were distinguished. Several of the ITS of different strains were identical in this analysis. This was observed with ITS type XV, which was common to six different type strains (Table 1) and 12 wine yeast strains (Table 2). Interestingly, only five different ITS profiles characterized the 26 wine strains (Table 2) while 15 different ITS patterns were recorded amongst the *Saccharomyces* type strains (Table 1). Thirteen strains were each characterized by a unique set of restriction profiles; among them was *S. cerevisiae* CBS1171 type strain which was the only one to have unique *Hae*III and *Rsa*I ITS profiles. RFLP analysis of this spacer sequence yielded 41 different restriction fragments, of which 35 (85%) were polymorphic.

The 3·2 kb IGS sequences were more polymorphic. All 10 restriction enzymes tested gave several restriction patterns, from four for *Hin*fI to 12 for *Alu*I. However, some of the strains could still not be distinguished from each other. This was particularly the case among the 26 wine yeast strains for which five sets of restriction patterns were recorded; two of these sets (IGS types XVIII and XXI) characterized 10 and 11 strains, respectively (Table 2). This is in contrast with the 23 *Saccharomyces sensu stricto* type strains which were distributed in 17 different IGS classes (Table 1). Interestingly, the *S. cerevisiae* CBS1171 type strain was the only strain which had a unique IGS restriction profile for each of the 10 enzymes used. RFLP analysis of the IGS yielded 184 different restriction fragments, of which 170 (92%) were polymorphic.

Data from RFLP analyses of the ITS and IGS sequences were pooled and used to calculate pairwise *d* values between the spacer sequences according to the model of Nei & Li (21). The matrix of *d* values was used to construct an inferred gene tree using the neighbour-joining method. This tree was characterized by two groups of phylogenetically close sequences, one of which was called the '*cerevisiae*' cluster and the other, the '*bayanus*' cluster (Fig. 1). The '*cerevisiae*' cluster included most of the studied yeasts, i.e. 15 out of the 23 type strains plus 16 out of the 26 wine yeast strains. Within this cluster, the rDNA spacers were highly similar with a maximum *d* value (d_{max}) between them of 0·033 substitutions per site. The only exception was the *S. cerevisiae* rDNA spacers, which were more distantly related to the other sequences; minimum *d* value (d_{min}) of 0·033 to *Saccharomyces aceti* and d_{max} of 0·056 to *Saccharomyces chevalieri*. The '*bayanus*' cluster included the remaining yeasts except for *S. paradoxus*, i.e. seven type strains and 10 wine yeast strains (seven of which originated from the Touraine region of France). As observed for the '*cerevisiae*' cluster, most of the sequences of the yeast type strains, including *S. bayanus*, were highly similar (d_{max} of 0·009). The sequence of the *Saccharomyces uvarum* type strain was loosely linked to the latter type strains (d_{min} of 0·063 to *Saccharomyces heterogenicus*). The rDNA spacers of the wine yeast strains included in this cluster were more related to the *S. uvarum* spacers (d_{max} of 0·034) than to the spacers of the other type strains. The two clusters '*cerevisiae*' and '*bayanus*' were clearly separated in this PCR-RFLP analysis, the d_{min} and d_{max} values separating the spacer sequences present in the two clusters were, respectively, 0·066 (*Saccharomyces oleaceus*/*S. uvarum*) and 0·101 (*S. cerevisiae*/*Saccharomyces globosus*). The rDNA spacers of *Saccharomyces paradoxus* had a somewhat intermediate position between the two clusters, being closer to the '*cerevisiae*' cluster (d_{min} of 0·057 to *S. oleaceus*) than to the '*bayanus*' cluster (d_{min} of 0·085 to *Saccharomyces carlsbergensis*). This intermediate pos-

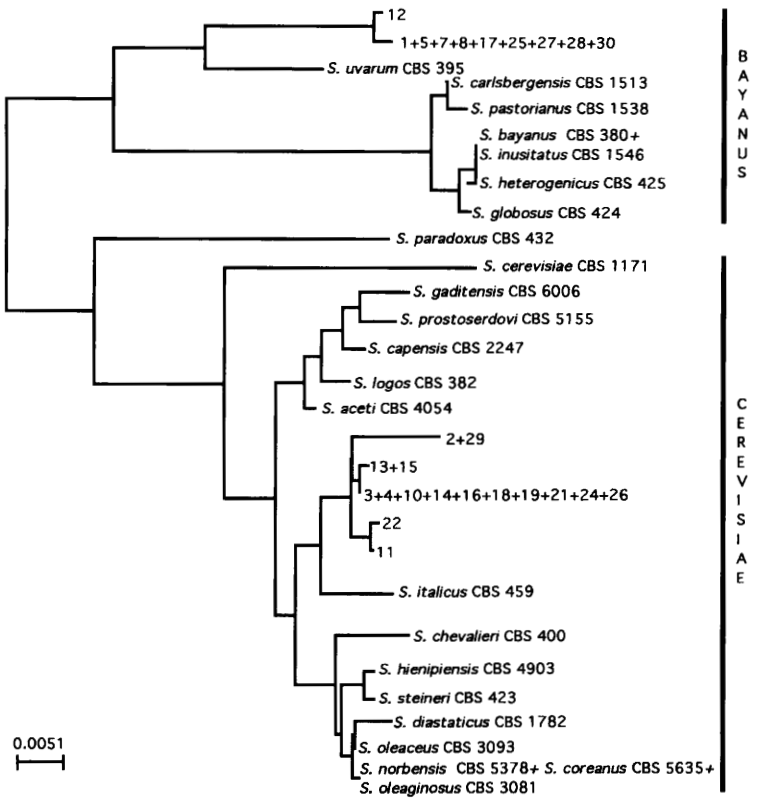

Fig. 1. Unrooted phylogenetic dendrogram showing the taxonomic positions of *Saccharomyces* type strains and wine yeast strains, based on the PCR-RFLP study of the ITS and IGS rDNA spacers. Numbers refer to strains listed in Table 2. Bar, *d* value of 0·0051.

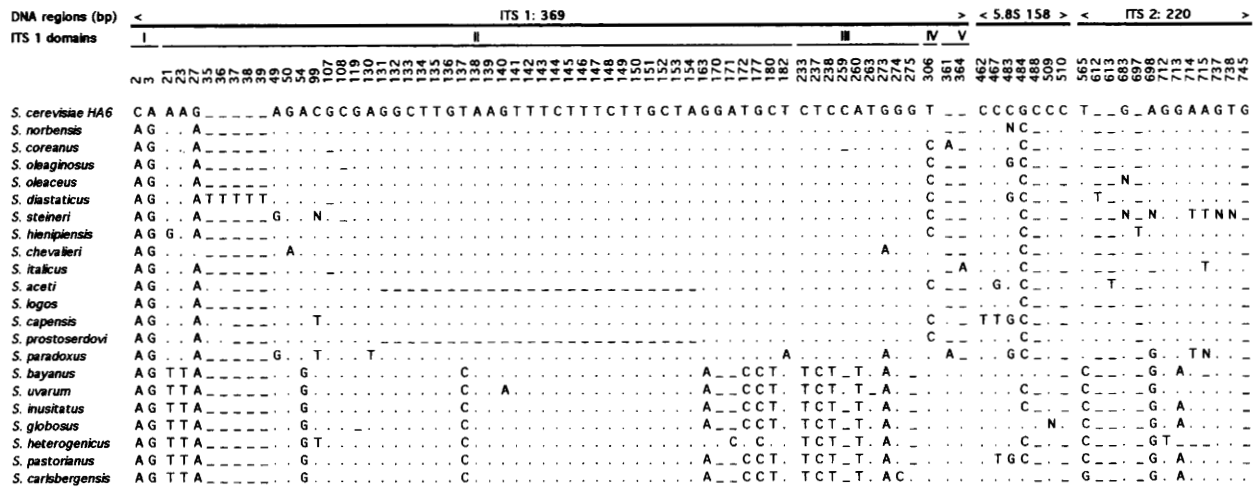

Fig. 2. Variable nucleotide positions within the ITS1-5·8S-ITS2 region of 21 *Saccharomyces* type strains (Table 1) compared to the corresponding sequence of the *S. cerevisiae* HA6 strain (GenBank accession no. U09327). Hyphens indicate deletions; N, uncertain nucleotides.

ition of *S. paradoxus* resulted only from polymorphisms in the IGS; a tree based on only the ITS included *S. paradoxus* in the '*cerevisiae*' cluster (results not shown).

Sequence analysis of the ITS regions

This was carried out on the 23 type strains. Repeated attempts to sequence the PCR-amplified ITS of *S. cerevisiae* CBS1171 type strain always gave several bands at most nucleotide positions suggesting that different sequences had been co-amplified by PCR. This strain was not studied further and the ITS sequence of *S. cerevisiae* strain HA6 (GenBank accession no. U09327) was used for comparison instead. The entire sequences of the ITS1 and 5·8S and the sequence of ITS2, excluding the last 10 nucleotides, were obtained for all other strains except *Saccharomyces gaditensis* for which the first 19 nucleotides of ITS1 and the last 150 nucleotides of ITS2 could not be

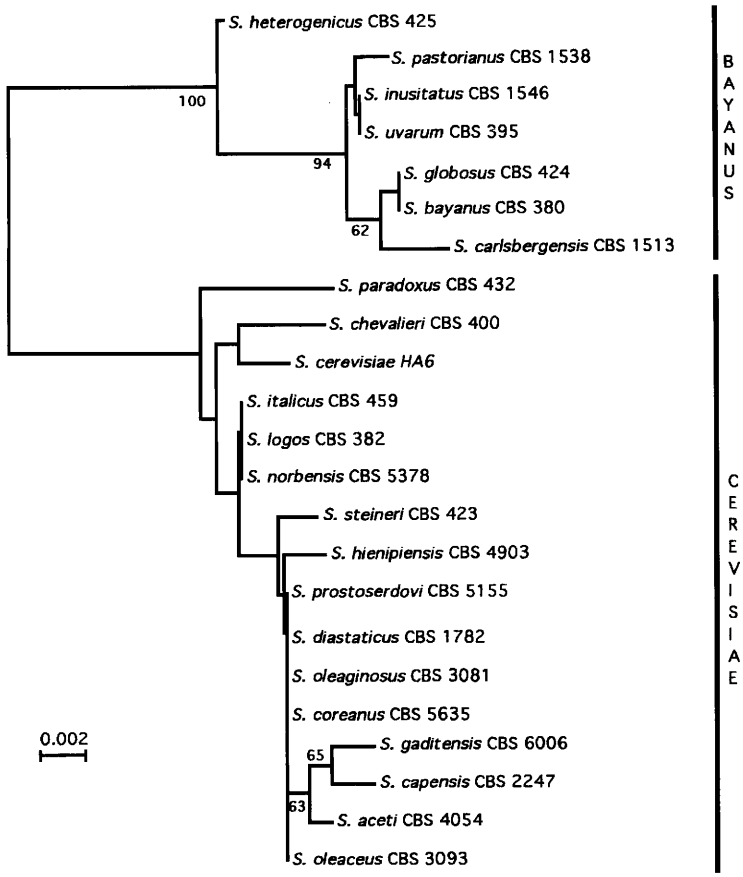

Fig. 3. Unrooted phylogenetic dendrogram showing the taxonomic positions of *Saccharomyces* type strains, based on the ITS1-5·8S-ITS2 sequences. Bootstrap supports higher than 60% are displayed. Bar, *d* value of 0·002.

read. If we exclude this latter strain, a total of 747 nucleotides including the 5·8S gene, representing the full-length sequences, could be aligned without any ambiguities. Only 8 bp were not read in the different sequences (Fig. 2). Several sequences had 1–5 bp insertions at various positions; three sequences (those of *S. aceti*, *S. gaditensis* and *Saccharomyces prostoserdovi*) shared an identical 24 bp deletion between nucleotides 131 and 154 in ITS1. Most of the variable nucleotide positions were within ITS1 (39 out of a total of about 369 nucleotides) while ITS2 was less variable (10 variable positions out of about 220 nucleotides). The 5·8S gene was the most conserved part of the sequence with six variable positions out of 157 nucleotides (Fig. 2). Within ITS1, the variable positions were not evenly distributed; most were found in domains II and III (12·6 and 13·7%, respectively, of variable positions) while domains IV and V were more conserved (2·1 and 5·5%, respectively, of variable positions).

The sequence alignment was used to construct an inferred phylogenetic tree by the neighbour-joining method (Fig. 3). For the calculation, 26 variable positions were considered. Inclusion or exclusion of the partial sequence of *S. gaditensis* in the calculation did not significantly change the topology of the tree and the bootstrap values of the nodes (results not shown). This comparative sequence analysis clearly showed the distribution of *Saccharomyces* type strains ITS sequences into two major clusters which were called '*bayanus*' and '*cerevisiae*'. Within the '*bayanus*' cluster, which included the ITS of seven type strains among which were those of *S. bayanus* and *S. pastorianus*, the different sequences were highly homologous with a d_{max} between two sequences of 0·0117 (*S. carlsbergensis*/*S. heterogenicus*). The '*cerevisiae*' cluster included the other 16 type strain sequences among which were those of *S. cerevisiae* HA6 and *S. paradoxus*. The d_{max} within this group was 0·0137 (*S. paradoxus*/*Saccharomyces capensis*).

The '*bayanus*' and '*cerevisiae*' clusters were clearly separated; the d_{min} between sequences in these two groups was 0·0196 (*S. heterogenicus*/*Saccharomyces logos*, *Saccharomyces norbensis*, *Saccharomyces italicus*). Furthermore, the two nodes between these two groups were separated by a high bootstrap value of 100% (Fig. 3). Although the *S. paradoxus* ITS sequence branched outside the '*cerevisiae*' cluster, this node was not supported by a significant bootstrap value and therefore the sequence was not excluded from the '*cerevisiae*' cluster.

DISCUSSION

Several studies on yeast strains belonging to *Saccharomyces cerevisiae sensu* Yarrow, especially those relative to nuclear DNA (nDNA)/nDNA reassociations (35, 37) have established that this species was genetically

heterogeneous and consisted of four distinct sibling taxa: *S. cerevisiae*, *S. bayanus*, *S. pastorianus* and *S. paradoxus*. In this paper, analysis of the intergenic rDNA spacers of different yeast strains belonging to the *Saccharomyces sensu stricto* group (33) has revealed sequence polymorphisms which allow inference of the phylogenetic links existing between the sequences analysed. Both combined PCR-RFLP analysis of the ITS and IGS spacers and direct sequencing of the ITS lead to a similar grouping of the studied isolates, their rDNA spacer sequences being distributed in two well-separated clusters called '*cerevisiae*' and '*bayanus*'. As for the sequences of *S. paradoxus* and *S. pastorianus*, they do not form separate clusters but instead are included in the '*cerevisiae*' and '*bayanus*' clusters, respectively.

The '*cerevisiae*' cluster includes the sequences of 16 out of the 23 studied type strains. Yeasts within this cluster are those whose genomic DNAs showed a very high level of relatedness (more than 89%) to the genomic DNA of the *S. cerevisiae* CBS1171 type strain (36). In RAPD analysis based on the use of three different 10 bp oligonucleotides, these strains appeared rigorously identical (13) whereas a certain amount of sequence polymorphism exists between their rDNA spacers.

The phylogenetic position of the *S. paradoxus* type strain relative to the other yeasts in the '*cerevisiae*' cluster cannot clearly be established from our results. This species belongs to the cluster when considering only the ITS sequences, whereas it branches outside the cluster in the RFLP analysis, which takes into account polymorphisms in both the ITS and IGS spacers. This taxonomic proximity to the '*cerevisiae*' cluster is consistent with the nDNA/nDNA reassociation values reported between the two species (in the range 40–50%), which are significantly higher than those reported with *S. bayanus* and *S. pastorianus* (34 and 23%, respectively) (35, 36, 37). The close phylogenetic relatedness between *S. paradoxus* and *S. cerevisiae* was also pointed out in two RFLP studies, one relative to the mitochondrial genome (5) and the other to the small- and large-subunit rRNA genes (11). Despite the similarities between these two species, interspecific crosses between monosporic isolates always gave hybrids producing non-viable ascospores (15, 17, 19).

The '*bayanus*' cluster includes the sequences of the seven remaining type strains. This group appears highly homogeneous as judged from the low *d* values between the sequences. The only exceptions were *S. uvarum* and *S. heterogenicus*, which branched somewhat outside the cluster in only the PCR-RFLP study for the former and in only the ITS sequence analysis for the latter. Yeasts within the '*bayanus*' cluster are those whose nDNA showed a high percentage (above 72%) of reassociation to the nDNA of the *S. bayanus* type strain (35, 36). Nevertheless, the '*bayanus*' cluster as defined in our study includes two of the recognized species of the *S. cerevisiae* complex, namely *S. bayanus* and *S. pastorianus* (35, 37). *S. pastorianus* is assumed to be a rare natural hybrid between *S. cerevisiae* and *S. bayanus* as its nDNA reassociates at a significant level to the DNAs of both these species (53 and 72%, respectively) (35). It appears that the rDNA types of *S. pastorianus* and of its synonym *S. carlsbergensis* are both extremely similar to that of *S. bayanus*, thus confirming, as already suggested (9, 24), that it has inherited of the rDNA locus of only one of its two possible parents: *S. bayanus*. A similar situation was also reported in the case of the mitochondrial genome for which only that of *S. bayanus* was shown to be present in *S. pastorianus* cells (5, 23).

A closer relatedness of *S. pastorianus* to *S. bayanus* than to *S. cerevisiae* or *S. paradoxus* was also deduced from analysis of RAPD patterns (13). *S. bayanus* and *S. pastorianus* also share at least two common physiological characteristics, i.e. an active fructose transport system and a maximum temperature for growth in the range 28–35 °C (28). Conversely, an active fructose transport system is absent in both *S. cerevisiae* and *S. paradoxus*, both of which have a higher maximum temperature for growth (in the range 35–43 °C). These data, together with our results, strongly suggest that if *S. pastorianus* is a natural hybrid between *S. cerevisiae* and *S. bayanus* it is a partial one which has inherited far more genetic material from the latter parent than from the former one. The presence of *S. cerevisiae*-specific DNA sequences in the genome of *S. pastorianus* has been established in the case of some metabolic genes such as *MET2* (6, reviewed by 9).

The study of rDNA spacers has led to the discrimination of two well-separated clusters within which some polymorphism does exist between the yeast strains. However well-separated the '*cerevisiae*' and '*bayanus*' clusters may be, they consist of very closely related yeasts when the whole genus *Saccharomyces* is considered (7, 11, 12, 13). Thus, a phylogenetic tree based on the 18S gene sequences of species belonging to *Saccharomyces* and related genera showed that the genus *Saccharomyces* is very heterogeneous, whereas the species of the *Saccharomyces sensu stricto* complex form a natural group quite separated from other *Saccharomyces* and non-*Saccharomyces* species examined (7). Another example is given by the *ori–rep–tra* mitochondrial sequence, cloned from *S. cerevisiae*, which, when used as a probe, hybridizes to the DNA of *S. cerevisiae* as well as to *S. bayanus*, *S. paradoxus* and *S. pastorianus*, but not to the DNA of any of six other species of the yeast genus *Saccharomyces*: *Saccharomyces castellii*, *Saccharomyces dairensis*, *Saccharomyces exiguus*, *Saccharomyces kluyveri*, *Saccharomyces servazzi* and *Saccharomyces unisporus* (26). This was also observed when we attempted to align the ITS sequence of *S. cerevisiae* to that of *S. kluyveri* (GenBank accession no. U09328), the only one available in the databases (R. Messner, H. Prillinger, M. Ibl & G. Himmler, unpublished). If only ITS2 is considered, the sequences of all the yeasts of the two '*cerevisiae*' and '*bayanus*' clusters could be perfectly aligned with only

10 variable positions among the 220 nucleotides. To align the same sequence of *S. kluyveri* to that of *S. cerevisiae* or *S. bayanus*, it is necessary to create in the former two large gaps of about 20 bases each and only about 65% of the remaining positions can be aligned.

We extended the PCR-RFLP study to a collection of 26 wine yeast strains assigned to the *S. cerevisiae sensu stricto* group on the basis of traditional phenotypic characters. The rDNA spacer sequences of these strains were distributed in either of the two clusters '*cerevisiae*' or '*bayanus*'. For most of the strains, this classification was in disagreement with the original identification (Table 2), thus confirming the unsuitability of the classical taxonomic criteria within the *S. cerevisiae sensu stricto* group. However, these results are in perfect agreement with those reported by Naumov *et al.* (16), who performed genetic crosses and electrophoretic karyotypes on six of the strains we studied (nos 1, 3, 5, 7, 8 and 10). Unexpectedly, a lower level of RFLP was detected in the rDNA spacers of the wine yeast strains, included in either of the two clusters, compared to the polymorphism revealed between the corresponding sequences of the type strains (Fig. 1). A possible explanation could be that a limited number of lineages have adapted to the grape must or, alternatively, that the wine yeasts studied have been isolated over a short period of time in a limited number of sites all located in France as opposed to the varied origins of the type strains.

In conclusion, the analysis of polymorphisms within the rDNA spacers demonstrates that the species of the *Saccharomyces sensu stricto* group are clearly distributed in two well-separated clusters which have recently diverged from one another as judged by the high similarities between their ITS sequences. This result also clearly shows that the study of rDNA spacer polymorphisms is a valuable tool to infer the phylogenetic position of very closely related yeasts.

ACKNOWLEDGEMENTS

We would like to thank Dr C. Cuinier (Institut Technique de la Vigne et du Vin, Tours, France) for providing the wine yeast strains. This work was supported by the Centre d'Analyse Moléculaire de la Biodiversité (Université Lyon I) and by a BQR grant of the Université Lyon I.

REFERENCES

1. **Barnett, J. A. (1992).** The taxonomy of the genus *Saccharomyces* Meyen *ex* Reess: a short review for non-taxonomists. *Yeast* **8**, 1–23.
2. **Cardinali, G. & Martini, A. (1994).** Electrophoretic karyotypes of authentic strains of the *sensu stricto* group of the genus *Saccharomyces*. *Int J Syst Bacteriol* **44**, 791–797.
3. **Galtier, N., Gouy, M. & Gautier, C. (1996).** SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* **12**, 543–548.
4. **Gryta, H., Debaud, J. C., Effosse, A., Gay, G. & Marmeisse, R. (1997).** Fine-scale structure of populations of the ectomycorrhizal fungus *Hebeloma cylindrosporum* in coastal sand dune forest ecosystems. *Mol Ecol* **6**, 353–364.
5. **Guillamon, J., Barrio, M. E., Huerta, T. & Querol, A. (1994).** Rapid characterization of four species of the *Saccharomyces sensu stricto* complex according to mitochondrial DNA patterns. *Int J Syst Bacteriol* **44**, 708–714.
6. **Hansen, J. & Kielland-Brandt, M. C. (1994).** *Saccharomyces carlsbergensis* contains two functional MET2 alleles similar to homologues from *S. cerevisiae* and *S. monacensis*. *Gene* **140**, 33–40.
7. **James, S. A., Cai, J., Roberts, I. N. & Collins, M. D. (1997).** A phylogenetic analysis of the genus *Saccharomyces* based on 18S rRNA gene sequences: description of *Saccharomyces kunashirensis* sp. nov. and *Saccharomyces martiniae* sp. nov. *Int J Syst Bacteriol* **47**, 453–460.
8. **Kaneko, Y. & Banno, I. (1991).** Reexamination of *Saccharomyces bayanus* strains by DNA/DNA hybridization and electrophoretic karyotyping. *IFO (Inst Ferment (Osaka)) Res Commun* **15**, 31–41.
9. **Kielland-Brandt, M. C., Nilsson-Tillgren, T., Gjermansen, C., Holmberg, S. & Pedersen, M. B. (1995).** Genetics of brewing yeasts. In *The Yeasts*, vol. 6, 2nd edn, pp. 223–254. Edited by A. E. Wheals, A. H. Rose & J. S. Harrison. London: Academic Press.
10. **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.
11. **Messner, R. X. & Prillinger, M. (1995).** *Saccharomyces* species assignment by long range ribotyping. *Antonie Leeuwenhoek* **67**, 363–370.
12. **Molina, F. I., Inoue, T. & Jong, S. C. (1992).** Ribosomal DNA restriction analysis reveals genetic heterogeneity in *Saccharomyces cerevisiae* Meyen ex Hansen. *Int J Syst Bacteriol* **42**, 499–502.
13. **Molnar, O., Messner, R., Prillinger, M., Stahl, U. & Slavikova, E. (1995).** Genotypic identification of *Saccharomyces* species using random amplified polymorphic DNA analysis. *Syst Appl Microbiol* **18**, 136–145.
14. **Naumov, G. I. (1986).** Genetic differentiation and ecology of yeast *Saccharomyces paradoxus* Batschinskaya. *Dokl Akad Nauk SSSR* **291**, 754–757.
15. **Naumov, G. I., Naumova, E. S., Azbukina, Z. M., Korhola, M. & Gaillardin, C. (1993).** Genetic and karyotypic identification of *Saccharomyces* yeasts from far east Asia. *Cryptogam Mycol* **14**, 85–93.
16. **Naumov, G. I., Naumova, E. S. & Gaillardin, C. (1993).** Genetic and karyotypic identification of wine *Saccharomyces bayanus* yeasts isolated in France and Italy. *Syst Appl Microbiol* **16**, 274–279.
17. **Naumov, G. I., Naumova, E. S. & Korhola, M. (1992).** Genetic identification of natural *Saccharomyces* sensu stricto yeasts from Finland, Holland and Slovakia. *Antonie Leeuwenhoek* **61**, 237–243.
18. **Naumov, G. I., Naumova, E. S., Lantto, R. A., Louis, E. J. & Korhola, M. (1992).** Genetic homology between *Saccharomyces cerevisiae* and its sibling species *S. paradoxus* and *S. bayanus*: electrophoretic karyotypes. *Yeast* **8**, 599–612.
19. **Naumov, G. I., Naumova, E. S., Sancho, E. D. & Korhola, M. (1993).** Taxogenetics of the *Saccharomyes sensu stricto* yeasts from western and south Africa. *Cryptogam Mycol* **14**, 263–270.
20. **Naumov, G. I., Naumova, E. S., Turakainen, H. & Korhola, M.**

(1996). Identification of the β-galactosidase MEL genes in some populations of *Saccharomyces cerevisiae*: a new gene MEL11. *Genet Res* **67**, 101–108.

21. **Nei, M. & Li, W. H. (1979).** Mathematical model for studying genetic variation in terms of restriction endonuclease. *Proc Natl Acad Sci USA* **76**, 5269–5273.
22. **Ness, F. & Aigle, M. (1995).** RTM1: a member of a new family of telomeric repeated genes in yeast. *Genetics* **140**, 945–956.
23. **Petersen, R. F., Groth, C., Nilsson-Tillgren, T. & Piskur, J. (1997).** Structural dynamics of the nuclear and mitochondrial genomes within the genus *Saccharomyces*. *Yeasts* **13**, S71.
24. **Petersen, J. G. L., Nilsson-Tillgren, T., Kielland-Brandt, M. C., Gjermansen, C. & Holmberg, S. (1987).** Structural heterozygosis at genes *ILV2* and *ILV5* in *Saccharomyces carlsbergensis*. *Curr Genet* **12**, 167–174.
25. **Phillipsen, P., Stotz, A. & Scherf, M. (1991).** DNA of *Saccharomyces cerevisiae*. *Methods Enzymol* **194**, 169–182.
26. **Piskur, J., Mozina, S. S., Stenderup, J. & Pedersen, M. Z. (1995).** A mitochondrial molecular marker, *ori-rep-tra* for differentiation of yeast species. *Appl Environ Microbiol* **61**, 2780–2782.
27. **Raué, H. A. & Planta, R. J. (1991).** Ribosome biogenesis in yeast. *Nucleic Acids Res Mol Biol* **41**, 89–129.
28. **Rodrigues De Sousa, M., Madeira-Lopes, A. & Spencer-Martins, I. (1995).** The significance of active fructose transport and maximum temperature for growth in the taxonomy of *Saccharomyces sensu stricto*. *Syst Appl Microbiol* **18**, 44–51.
29. **Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **44**, 406–425.
30. **Sanger, F., Nicklen, S. & Coulson, A. R. (1977).** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
31. **Sequerra, J., Marmeisse, R., Valla, G., Normand, P., Capellano, A. & Moiroud, A. (1997).** Taxonomic position and intraspecific variability of the nodule forming *Penicillium nodositatum* inferred from RFLP analysis of the ribosomal intergenic spacer and random amplified polymorphic DNA. *Mycol Res* **101**, 465–472.
32. **Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994).** CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
33. **van der Walt, J. P. (1970).** *Saccharomyces* Meyen emend. Reess. In *The Yeasts: a Taxonomic Study*, 2nd edn, pp. 555–718. Edited by E. Lodder. Elsevier: Amsterdam.
34. **Van Nues, R. W., Rientjes, J. M. J., Van Der Sande, C. A. F. M., Zerp, S. F., Sluiter, C., Venema, J., Planta, R. J. & Raué, H. A. (1994).** Separate structural elements within internal transcribed spacer 1 of *Saccharomyces cerevisiae* precursor ribosomal RNA direct the formation of 17S and 26S rRNA. *Nucleic Acids Res* **22**, 912–919.
35. **Vaughan Martini, A. (1989).** *Saccharomyces paradoxus* comb. nov., a newly separated species of the *Saccharomyces sensu stricto* complex based upon nDNA/nDNA homologies. *Syst Appl Microbiol* **12**, 179–182.
36. **Vaughan Martini, A. & Kurtzman, C. P. (1985).** Deoxyribonucleic acid relatedness among species of the genus *Saccharomyces sensu stricto*. *Int J Syst Bacteriol* **35**, 508–511.
37. **Vaughan Martini, A. & Martini, A. (1987).** Three newly delimited species of *Saccharomyces sensu stricto*. *Antonie Leeuwenhoek* **53**, 77–84.
38. **Vaughan Martini, A., Martini, A. & Cardinali, G. (1993).** Electrophoretic karyotyping as a taxonomic tool in the genus *Saccharomyces*. *Antonie Leeuwenhoek* **63**, 145–156.
39. **White, T. J., Bruns, T., Lee, S. & Taylor, J. (1990).** Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols: a Guide to Methods and Applications*, pp. 129–141. Edited by M. A. Innis, D. H. Gelfand, J. S. Sninsky & T. J. White. Academic Press: New York.
40. **Yarrow, D. (1984).** *Saccharomyces* Meyen ex. Reess. In *The Yeasts: a Taxonomic Study*, 3rd edn, pp. 379–395. Edited by N. J. W. Kreger-van Rij. Elsevier: Amsterdam.