Applied and Environmental Microbiology, Nov. 2002, p. 5437–5444 Vol. 68, No. 11
0099-2240/02/$04.00+0 DOI: 10.1128/AEM.68.11.5437–5444.2002

SED1 Gene Length and Sequence Polymorphisms in Feral Strains of *Saccharomyces cerevisiae*

Ilaria Mannazzu,* Emanuela Simonetti, Paola Marinangeli, Emanuela Guerra,† Marilena Budroni,‡ Madan Thangavelu, and Francesca Clementi

Dipartimento di Biotecnologie Agrarie ed Ambientali, Università di Ancona, Ancona, Italy

Received 13 May 2002/Accepted 2 August 2002

The *SED1* gene (YDR077W), coding for the major cell wall glycoprotein of *Saccharomyces cerevisiae* stationary-phase cells, contains two blocks of tandem repeat units located within two distinct regions of the nucleotide sequence. A PCR survey of the *SED1* open reading frames (ORFs) of 186 previously uncharacterized grape must isolates of *S. cerevisiae* yielded 13 PCR profiles arising from different combinations of seven *SED1* length variants in individuals homozygous or heterozygous for the gene. Comparison of the nucleotide sequences of a group of representatives of each of the seven length variants with those of S288C and the type strain, CBS1171, unequivocally identified them as *SED1* alleles and provided evidence for the presence of two minisatellite-like sequences, variable in length, within the ORF of an *S. cerevisiae* gene. The segregation analyses of the *SED1* length variants and other genetic markers in 13 isolates representative of each PCR profile suggested that molecular mechanisms involved in minisatellite expansion and contraction may be responsible for *SED1* heterozygosities within a population of homothallic must isolates of *S. cerevisiae*.

SED1 (YDR077W), initially identified as a multicopy suppressor of the *ERD2* deletion (15), has more recently been shown to encode the most abundant cell wall glycoprotein of *Saccharomyces cerevisiae* stationary-phase cells (28). Sed1p is rich in serines and threonines and, like other cell wall proteins, has N- and C-terminal hydrophobic domains, multiple sites for glycosylation with both N- and O-linked sugars, and a signal sequence for the addition of a glycosylphosphatidylinositol anchor at the carboxy terminus (6, 13, 14). Sed1p is not essential for normal growth (15, 28). However, a *SED1* deletion results in decreased resistance to the action of zymolyase with respect to wild-type cells (28). This phenotype is more evident in stationary-phase cells, implicating the involvement of Sed1p in stress resistance during that growth phase. *SED1* is strongly up-regulated at the diauxic shift upon glucose depletion (9) and is highly expressed around the M phase of the cell cycle (7, 29) and in the presence of aluminum and zinc (11). The presence of both stress-responsive elements (CCCCT and AGGGG) in the promoter and upstream region of the gene (28) and of three putative PEST regions (8), which may be necessary for rapid changes in concentration of the protein in response to environmental stimuli (24, 25), has suggested a role for Sed1p in providing resistance to biotic and abiotic stresses.

Analysis of the amino acid sequence of Sed1p in *S. cerevisiae* S288C revealed the presence of repeated amino acid motifs localized within two distinct regions of the polypeptide chain (15). Based on the *SED1* nucleotide sequence of S288C, the first region (region 1) contains three repeat units of 66 bp and a truncated one of 42 bp arranged in the following order: 66, 42, 66, 66. The penultimate codon of each 66-bp unit encodes an asparagine residue that is a potential N-linked glycosylation site. The second region (region 2) contains two 153-bp repeats and the information for a total of six cysteine residues. Similar to their role in mucins, the cysteine residues may be involved in one or more functions, such as disulfide-dependent intramolecular interactions required for protein folding or inter-Sed1p interactions on the cell wall, or in interactions with other cell wall proteins, or they may be necessary for higher-order assemblies involving Sed1p.

Our objective in this study was to assess whether the observed tandem repeats could mimic a minisatellite-like behavior and cause length polymorphism in an *S. cerevisiae* gene coding for a structural protein. Thus, after a preliminary PCR survey of strain collections and previously uncharacterized *S. cerevisiae* isolates from diverse sources (I. Mannazzu and M. Thangavelu, unpublished data), we surveyed *SED1* sequence variation in a collection of previously uncharacterized grape must isolates of *S. cerevisiae* and observed abundant *SED1* polymorphisms in the population analyzed. In this paper, we report the presence of minisatellite-like sequences within the open reading frame (ORF) of the *SED1* gene and speculate on the reasons for *SED1* length variation and heterozygosis in feral strains of *S. cerevisiae*.

MATERIALS AND METHODS

Strains. *S. cerevisiae* S288C and CBS1171 were used as reference strains. CBS1171 was isolated from "brewer's top yeast," Oranjeboom Brewery, Rotterdam, The Netherlands, by A. C. van Wijk in February 1925 and was deposited at the Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands, in the same period. This strain, considered the type strain for *S. cerevisiae*, is also known as ATCC 18824, CCRC 21447, DBVPG 6173, DSM 70449, IFO 10217, IGC 4455, JCM 7255, NCYC 505, and NRRLY-12632 in the respective collections.

The wild *S. cerevisiae* strains were isolated, during the 1998 vintage, from a

* Corresponding author. Mailing address: Dipartimento di Biotecnologie Agrarie ed Ambientali, Università degli Studi di Ancona, Via Brecce Bianche 60131, Ancona, Italy. Phone: 31 071 2204782. Fax: 31 071 2204858. E-mail: ilaria@popcsi.unian.it.

† Present address: Dipartimento di Biologia Cellulare ed Oncologia, Consorzio M. Negri Sud, Chieti, Italy.

‡ Present address: Dipartimento di Scienze Ambientali Agrarie e Biotecnologie Agro-Alimentari, Università degli Studi di Sassari, Sassari, Italy.

TABLE 1. Origin of isolates and distribution of the 13 *SED1* PCR profiles within the population analyzed

Cellar	Vat no.	Grape must	No. of isolates harboring the observed *SED1* PCR profile												
			A	B	C	D	E	F	G	H	I	L	M	N	O
1	1	Pecorino		4		17	2				8				3
	2	Sangiovese				9				2	7				1
	3	Passerina		1		13	2	1			5				
	4	Montepulciano	1	5		1	7		1		2				1
2	5	Pecorino			1	16	3						1		
	6	Sangiovese				12								1	
	7	Passerina				10	5				1				
	8	Montepulciano			1	2	4			14					
3	9	Verdicchio				10	6			1	3	1			1

total of nine vats, each containing must of a single grape variety, and sampled in three separate wine cellars located in a traditional wine-producing area of the Italian region of Marche (Table 1). The musts were collected from the fermentation vats immediately after crushing of the grapes and transferred to sterile 250-ml flasks, which were taken to the laboratory within 8 h after collection and incubated statically at 25°C. The yeasts were isolated by streaking triplicate must aliquots on WL nutrient agar (Oxoid) immediately upon sampling (time zero) and at different enrichment stages (early, middle, and late) in order to collect the strains acting in sequence during must fermentation. A total of 286 colonies exhibiting *Saccharomyces* morphology on WL agar were purified by repeated streaking on YEPD, and 186 isolates, confirmed as belonging to the species *S. cerevisiae*, were utilized in the present study. They were identified according to the method of Vaughan-Martini and Martini (31), utilizing the dichotomic key proposed by Boulton et al. (4) as described by Guerra et al. (12).

Media and growth conditions. WL agar was used for yeast isolation, YEPD (2% glucose, 1% yeast extract, 2% peptone, 1.8% agar) was used for yeast cultivation, and K acetate (0.1% glucose, 0.25% yeast extract, 0.98% K acetate, 2% agar) was utilized as a sporulation medium. The media for yeast identification were as described by Guerra et al. (12). The ability to ferment galactose (GAL), melibiose (MEL), trehalose (TRE), maltose (MAL), cellobiose (CEL), sucrose (SUC), and raffinose (RAF) was tested on solid medium (1% yeast extract, 2% peptone, 2% agar) supplemented with the corresponding sugar plus the pH indicator bromothymol blue and adjusted to pH 7.5 after being autoclaved, as described by Mortimer et al. (21). The ability to assimilate nonfermentable carbon sources was assessed on GLY medium (1% yeast extract, 2% peptone, 3% glycerol, 1% ethanol [added after autoclaving], 2% agar). Copper resistance was tested as the ability to grow on CUP medium (0.7% yeast nitrogen base, 2% glucose, 2% agar, supplemented with 60 mg of copper sulfate liter^{-1}). Hydrogen sulfide production was assayed on the basis of the color of the colonies (cream or brown) on Biggy agar (Difco). Unless otherwise stated, liquid cultures were incubated at 28°C with shaking (200 rpm).

DNA extraction, PCR conditions, and restriction analyses. Total genomic DNA for PCR analysis was isolated from 24-h cultures as described by Ushinsky et al. (30). The PCR primers SED1FOR (5′-ATGAAATTATCAACTGTCCT ATTATCTGCCGG-3′; bases 1 to 32) and SED1REV (5′-TTATAAGAATAA CATAGCAACACCAGCCAAACC-3′; bases 1017 to 984) were designed on the *S. cerevisiae* S288C sequence. PCR amplification reactions were performed on a Perkin-Elmer Gene AMP PCR System 9700 in 25-μl reaction mixtures containing 2 μl of template DNA (approximately 10 ng/μl), 0.5 U of DyNAzyne II DNA polymerase (Finnzymes), 1× reaction buffer (Mg^{2+} free), 1.5 mM $MgCl_2$, 100 μM each deoxynucleoside triphosphate, and 2 pmol each of SED1FOR and SED1REV primers. The reactions were run for 35 cycles as follows: denaturation at 94°C for 1 min, annealing at 64°C for 1 min, and elongation at 72°C for 2 min. An initial denaturation step at 94°C for 3 min and a final 7-min extension at 72°C were performed. The PCR products were analyzed by electrophoresis on a 1.4% agarose gel in 1× Tris-borate-EDTA buffer. The gel images were visualized by means of a Bio-Rad Gel DOC 1000 and acquired with Multi-Analyst software (Bio-Rad). The restriction fragment length polymorphism (RFLP) analysis of the *SED1* amplicons was performed as follows: 25 μl of the PCR product was digested overnight with an excess of *Hpa*II and *Kpn*I in a final volume of 50 μl. The restriction fragments were analyzed as described above.

DNA sequencing and sequence analyses. PCR products were purified using a Microcon-PCR Centrifugal Filter Devices (Millipore) PCR template purification kit or by means of the QIAquick gel extraction kit protocol (Qiagen) and sequenced directly. Sequencing reactions were prepared by PCR by using SED1FOR and SED1REV and the internal primers S1 (5′-TTACAACCCATC TACTGACTACACC-3′) and S2 (5′-GGTGTAGTCAGTAGATGGGTTGTA A-3′) with the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer). The products of these reactions were purified on Centri-Sep spin columns (Amicon), dried, denatured by adding Template Suppression Reagent (Perkin-Elmer), and sequenced by using the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer). The primary sequence alignment was performed using the Multiple Alignment Program (http://dot.imgen.bcm.tmc.edu:9331). Minor adjustments of the output alignment were performed by hand.

Tetrad analysis. Yeasts grown on YEPD were streaked on sporulation medium and incubated at 25°C until asci were observed microscopically. Depending on spore viability, a total of 16 to 48 asci for each isolate were dissected as described by Rose et al. (26). The single-spore progeny from asci with four viable meiotic products were grown on YEPD and replica plated onto GAL, MEL, TRE, MAL, CEL, SUC, RAF, Biggy agar, CUP, and GLY as described by Mortimer et al. (21). Sporulation efficiency was determined as the sum of four- and three-spore asci divided by the number of asci dissected. Spore viability was determined as the number of viable spores divided by the total number of spores dissected (5). Homothallism was assessed by examining microscopically the abilities of the meiotic derivatives from an ascus to sporulate. Segregation of *SED1* alleles among single-spore progeny was analyzed by PCR.

Nucleotide sequence accession numbers. The sequences of the CBS1171 *SED1* gene and of *Sed1-1* to *Sed1-7* were deposited in GenBank under accession numbers AF510219, AF510220, AF510221, AF510222, AF510223, AF510224, AF510225, and AF510226, respectively.

RESULTS

***SED1* polymorphisms in grape must isolates of *S. cerevisiae*.** The indication that the *S. cerevisiae SED1* coding region is variable in length was based on PCR amplification of the *SED1* ORFs of a small and random selection of laboratory, food, and industrial strains isolated from diverse sources (Mannazzu and Thangavelu, unpublished data). These preliminary experiments also suggested that a simple PCR assay could be used to analyze *SED1* variation in a large number of isolates. We therefore decided to study *SED1* length polymorphism in a collection of wine yeasts made up of 186 *S. cerevisiae* isolates collected in 1998 from fermenting musts. These yeasts were isolated as part of a study aiming to characterize the autochthonous strains associated with traditional wine fermentation in the Marche region of Italy and, coming from musts of different grape varieties sampled from nine vats located in three different cellars, were thought to provide a good source of biodiversity.

Under relatively stringent PCR conditions, the primer pair based on the *S. cerevisiae* S288C *SED1* sequence was utilized to amplify total DNA from the two reference strains, CBS1171 and S288C, and from all the must isolates. Both of the reference strains produced a single amplicon; S288C yielded the expected 1,017-bp product, while CBS1171 yielded a product which was smaller than that of S288C (Fig. 1). The 186 *S. cerevisiae* must isolates displayed striking differences in the electrophoretic profiles of their PCR products. The major proportion of the isolates (72%) yielded a single amplification product of variable length, suggesting either a single copy of the gene per haploid genome or multiple copies of the same size in a diploid or polyploid genome. The remaining isolates contained two or three *SED1* amplicons of different lengths per genome complement. The *SED1* amplicons ranged from approximately 950 to 1,300 bp.

Based on the observed *SED1* PCR profiles, the 186 must isolates could be grouped into 13 classes designated A to O

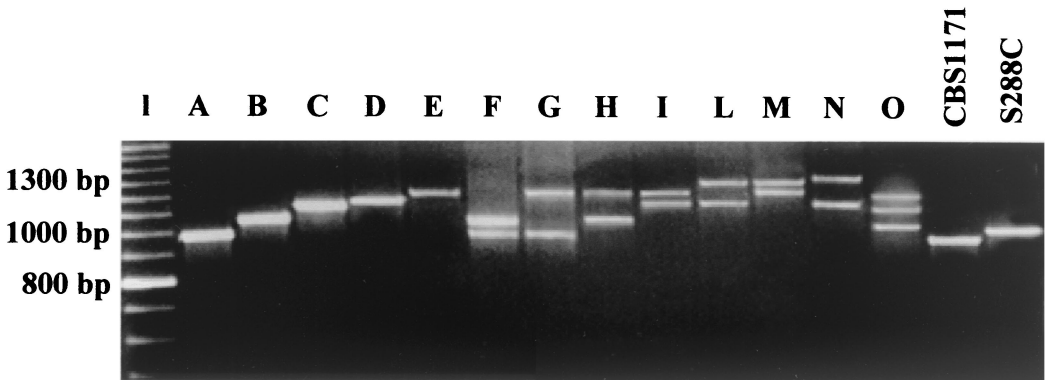

FIG. 1. *SED1* gene polymorphisms. PCR primers designed on the *SED1* sequence were utilized to amplify the *SED1* genes of the two reference strains, CBS1171 and S288C, and of the 186 must isolates of *S. cerevisiae*. Lane 1, 100-bp ladder (Amersham-Pharmacia); lanes A to O, PCR profiles observed within the population analyzed.

(Fig. 1). Interestingly, while the PCR profile and *SED1* amplicon size of the type strain, CBS1171, were similar to the class A PCR profile, the 1,017-bp PCR product of S288C did not correspond to any of the PCR profiles observed in the population. Moreover, it was interesting that the *SED1* PCR profile D occurred in all the vats sampled, while some of the others, such as A, F, G, L, M, and N, occurred very rarely and were exclusive to some of the vats (Table 1).

RFLP analysis of *SED1* amplicons. Based on the restriction map of the *SED1* ORF of S288C (http://genome-www.stanford.edu/Saccharomyces/), it was easy to observe that the *Hpa*II and *Kpn*I restriction enzymes could be used as convenient markers to localize the regions of the gene involved in the observed length variations. In fact, when digested with *Hpa*II, the *SED1* ORF of S288C yields two well-resolved fragments. The shorter one (333 bp) contains all of region 1, including the mosaic of 66- and 42-bp repeat units, while the larger fragment (654 bp) contains region 2, which includes the two 153-bp repeats. *Kpn*I cuts the *SED1* gene within region 1 and yields a band corresponding to the 42-bp block located within this region (data not shown). Comparison of the *Hpa*II and *Kpn*I restriction products of the S288C and CBS1171 *SED1* genes readily localized the observed length difference to region 1. In fact, the S288C region 1 fragment was about 40 bp longer than the corresponding fragment in CBS1171, possibly due to the lack of the 42-bp block in the latter (data not shown). Similarly, the *Hpa*II and *Kpn*I restriction profiles of the *SED1* amplification products of the isolates presenting PCR profiles A to O revealed that the observed length polymorphisms were due to length variation localized in region 1 and/or region 2. The restriction analyses of the *SED1* amplicons also showed that all the PCR profiles could invariably be considered as derived from combinations of seven *SED1* length variants labeled *Sed1-1* to *Sed1-7*, from the shortest to the longest (Fig. 2). Interestingly, the *SED1* amplicons of the must isolates never contained the short DNA fragment including the 42-bp block present in S288C and were in that respect similar to that in CBS1171 (Fig. 2).

Sequence analyses of the *SED1* length variants. In order to refine the results of the restriction analyses and obtain a detailed structure of the coding region of the gene, the *SED1* ORFs of the *S. cerevisiae* type strain, CBS1171, and a representative of each of the seven *SED1* length variants were sequenced.

Consistent with the restriction analyses, the *SED1* sequences of S288C and CBS1171 differed by a 42-nucleotide sequence in region 1. Based on the predicted amino acid sequence, this short fragment is most likely the result of an in-frame deletion of a duplication of the DNA sequence containing the first 66-bp blocks of region 1 (Fig. 3). The comparison of the nucleotide sequences of the S288C and CBS1171 *SED1* genes also revealed the existence of six single-nucleotide polymorphisms (SNPs) located in region 2 of the S288C gene (at positions 459, 510, 579, 630, 732, and 891). All these changes were at the third position of the respective codons and resulted in silent point mutations.

Interestingly, comparison of the nucleotide and amino acid sequences of the seven alleles with those of CBS1171 and S288C revealed that the CBS1171 *SED1* ORF was identical to the *Sed1-1* allele. Moreover, the 42-bp block in region 1 and the six SNPs in region 2 were unique to S288C.

The *Sed1-1* to *Sed1-7* ORFs varied in length from 972 to 1,278 bp. Multiple sequence alignment confirmed that the *SED1* length polymorphisms were due to the insertion or deletion of perfect duplicates of DNA sequences containing the 66- and 153-bp units in regions 1 and 2 and corresponding to tandem repeats of 22 and 51 amino acids, respectively. Thus, two minisatellite-like DNA sequences that exhibit the property of possessing a variable number of tandem repeats are responsible for *SED1* polymorphism in a population of wild yeast (Fig. 4).

The nucleotide sequence alignment revealed the existence of SNPs in the alleles *Sed1-2*, *Sed1-3*, *Sed1-5*, and *Sed1-6*, all positioned within region 1 (Fig. 2). Comparison of the amino acid sequences predicted on the basis of the nucleotide sequences revealed that all the single-nucleotide polymorphisms occurring at the third position of the codons are silent. On the other hand, the nucleotide substitutions observed at positions 268 in *Sed1-2* and 290 in *Sed1-2* and *Sed1-5* cause the replacement of a threonine with a serine and a leucine with an isoleucine, respectively, in the resulting amino acid sequences (Fig. 2 and 3).

The analysis of the changes in the sequences and the numbers of repeat units within regions 1 and/or 2 provided a way of

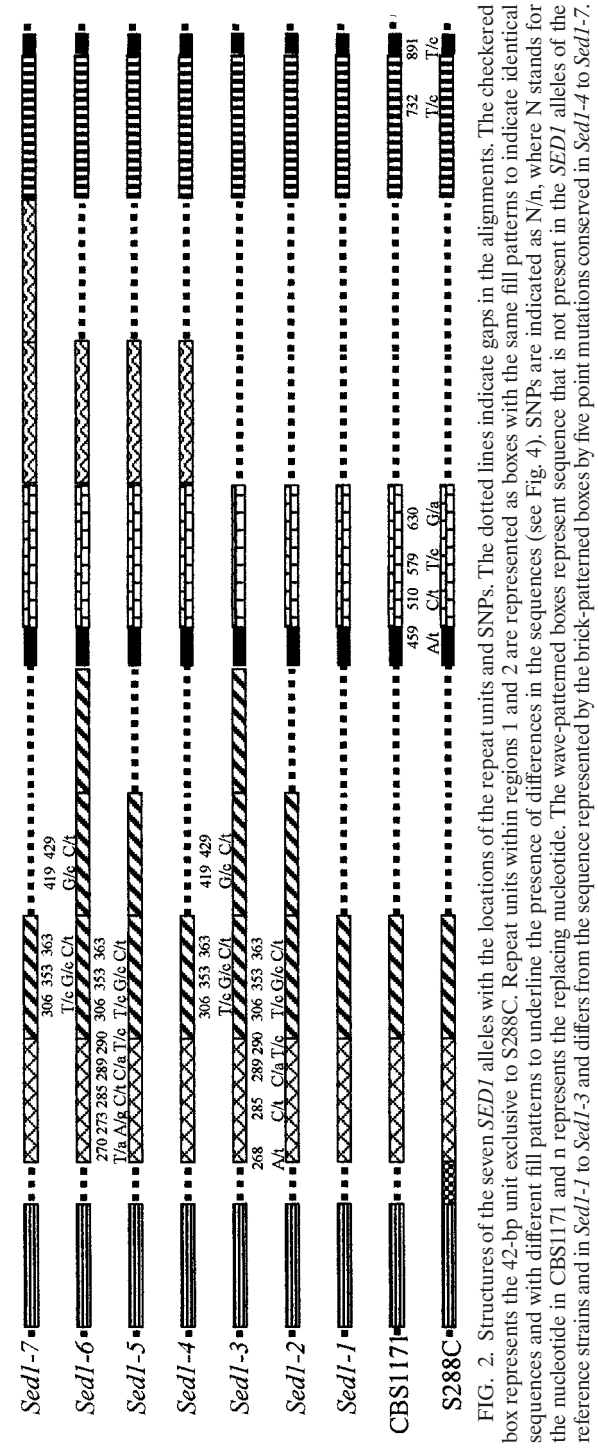

FIG. 2. Structures of the seven *SED1* alleles with the locations of the repeat units and SNPs. The dotted lines indicate gaps in the alignments. The checkered box represents the 42-bp unit exclusive to S288C. Repeat units within regions 1 and 2 are represented as boxes with the same fill patterns to indicate identical sequences and with different fill patterns to underline the presence of differences in the sequences (see Fig. 4). SNPs are indicated as N/n, where N stands for the nucleotide in CBS1171 and n represents the replacing nucleotide. The wave-patterned boxes represent sequence that is not present in the *SED1* alleles of the reference strains and in *Sed1-1* to *Sed1-3* and differs from the sequence represented by the brick-patterned boxes by five point mutations conserved in *Sed1-4* to *Sed1-7*.

exploring the routes leading to the *SED1* polymorphism observed in the must isolates and led us to hypothesize that all the alleles have arisen from a common ancestor as a consequence of tandem duplication events of the repeat units within these two regions. For example, the comparison of both the sequences and structures of *Sed1-1*, *Sed1-4*, and *Sed1-7* revealed that these three alleles differ in the number of 153-bp repeat units in region 2 (Fig. 2). Moreover, while their first and last 153-bp units are identical to the first and last such units of every other allele sequenced, the two central 153-bp repeat units of *Sed1-7* are identical to the central repeat unit of *Sed1-4* and differ by five substitutions from the first 153-bp repeat unit of region 2. These observations suggested that *Sed1-4* might have originated from *Sed1-1* subsequent to a tandem duplication of a 153-bp unit in region 2 followed by the occurrence of the five point mutations in this unit. Similarly, a plausible course of events in the history of *Sed1-7* would have involved a tandem duplication of the *Sed1-4* central 153-bp unit.

Segregation analysis of the seven *SED1* alleles and other genetic markers. Thirteen isolates representative of each PCR profile were subjected to genetic analyses as previously done by Mortimer et al. (21). As shown in Table 2, the 13 isolates were homothallic and exhibited various sporulation efficiencies and spore viabilities, as expected in wine strains of *S. cerevisiae*, due to their complex genetic constitution (2, 8). Apart from isolate Sc157, presenting PCR profile A, which was homozygous for all the characteristics considered, all the other isolates presented one or more heterozygosities, as expected in diploid strains (21).

The segregation of the *SED1* length variants in the F_1 progeny was analyzed by PCR on the monosporial cultures deriving from the four-viable-spore asci generated by each strain. The isolates presenting PCR profiles A to E showed a 4:0 segregation ratio (*SED1* amplicons of the same size in each monosporial culture from the same ascus) consistent with the presence of two copies of the same *SED1* length variant in a diploid background. Conversely, and as expected in diploid strains heterozygous for the *SED1* gene, isolates presenting PCR profiles F, G, I, L, M, and N showed a 2:2 segregation ratio for the *SED1* length variants. Isolates Sc59 and Sc93, presenting PCR profiles H and O, showed *SED1* segregation ratios which were consistent with the hypothesis of heterozygosity of this gene in a polyploid or polysomic background (Fig. 5), not unusual in wine strains (22).

DISCUSSION

By using a PCR approach, we have provided evidence for the existence of *SED1* gene polymorphism in a collection of grape must isolates of *S. cerevisiae*. The PCR screen and the PCR-RFLP procedure utilized proved to be effective tools for the analysis of this polymorphism and indicated that the 13 *SED1* PCR profiles observed within the collection of must isolates and labeled A to O were the result of different combinations of seven *SED1* length variants in individuals homozygous or heterozygous for this gene.

The sequence analysis of a representative of each of the observed *SED1* length variants unequivocally identified them as seven alleles of the *SED1* gene. However, as other alleles of the same size but with meaningful SNPs could have gone

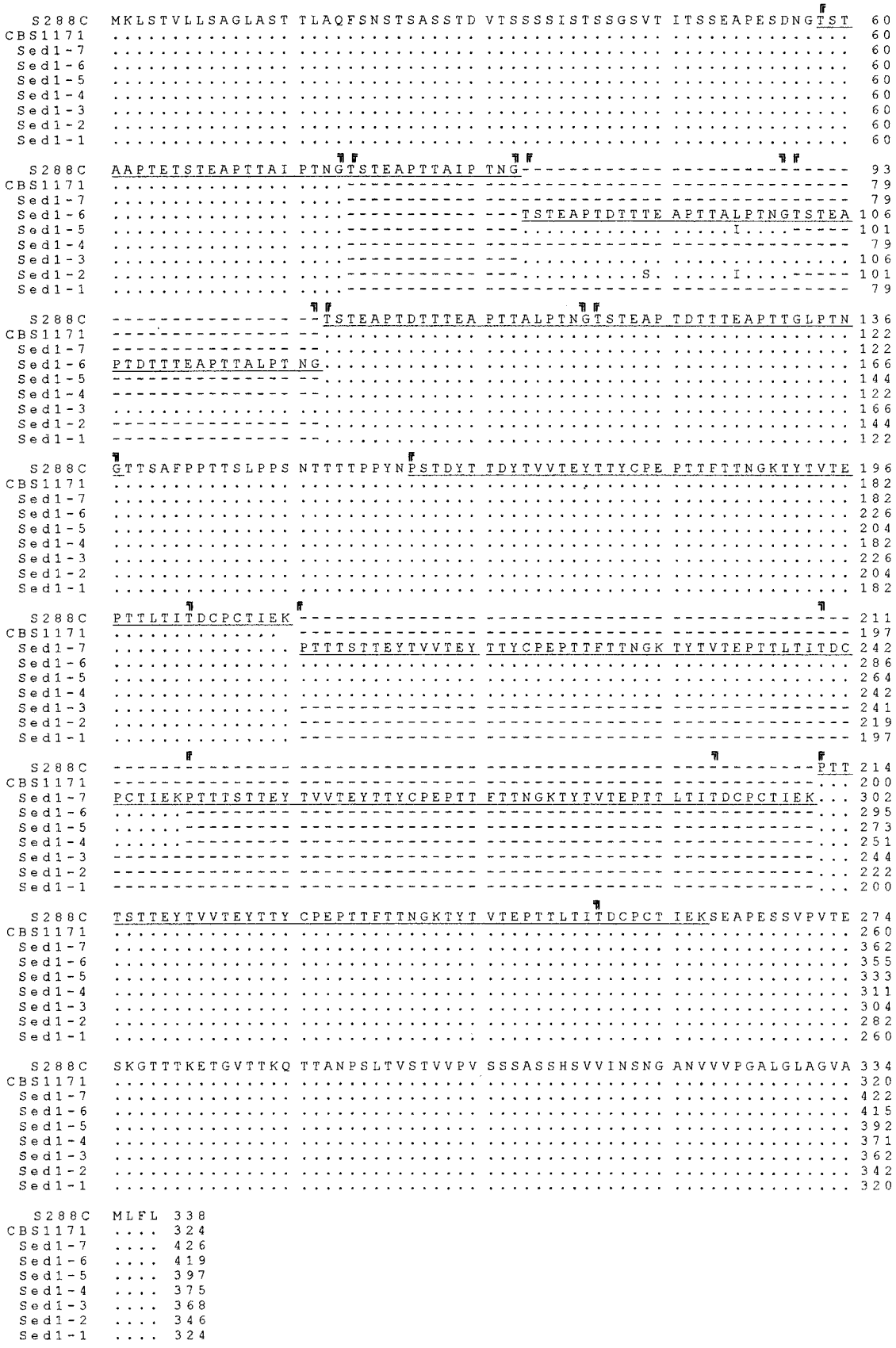

FIG. 3. Multiple alignments of the predicted amino acid sequences of the seven *SED1* alleles and the *SED1* genes of CBS1171 and S288C. The repeat units within regions 1 and 2 are underlined. The dashes indicate gaps introduced to maximize the alignment, and the dots indicate identical residues. Amino acid coordinates are provided. The beginning and end of each repeat unit is indicated by a left and right "corner," respectively, above the sequence.

```
Region 1
172                                                                  237
ACCAGCACAGCTGCACCAACTGAAACCTCAACAGAGGCTCCAACCACTGCTATCCCAACTAACGGT     Repeat 1
238                                                                  303
ACCTCTACTGAAGCTCCAACTGATACTACTACTGAAGCTCCAACCACCGCTCTTCCAACTAACGGT     Repeat 2
304                                                                  369
ACTTCTACTGAAGCTCCAACTGATACTACTACTGAAGCTCCAACCACCGGTCTTCCAACCAACGGT     Repeat 3

Region 2
610                                                                  678
ACCACTGACTACACCGTAGTCACTGAATATACTACTTACTGTCCAGAACCAACCACTTTCACCACAAAC  Repeat 1
916                                                                  984
ACCACCGAATACACTGTAGTCACTGAGTACACTACTTACTGTCCAGAACCAACCACTTTCACCACAAAC  Repeat 2
679                                                                  747
GGTAAGACTTACACTGTCACTGAACCAACCACTTTGACTATCACTGACTGTCCATGTACTATTGAAAAG  Repeat 1
985                                                                  1053
GGTAAGACTTACACTGTCACTGAACCAACCACATTGACTATCACTGACTGTCCATGCACCATTGAGAAG  Repeat 2
```

FIG. 4. Alignment between repeat units in regions 1 and 2 of CBS1171 *SED1* gene. Nucleotide coordinates are given.

undetected by the PCR-RFLP procedure, it is plausible that a higher number of alleles is present within the population analyzed.

Aiming to redefine the structure of the gene and elucidate the sequence of events that may have led to the observed length variations, we first compared the *SED1* sequences of the seven alleles with those of the two reference strains. The results obtained showed that the seven alleles analyzed have high

TABLE 2. Genetic analyses of 13 isolates representative of *SED1* PCR profiles

Characteristic[a]	Result for strain[b]:												
	Sc157	Sc205	Sc23	Sc42	Sc44	Sc200	Sc151	Sc59	Sc28	Sc112	Sc12	Sc143	Sc93
Class	A	B	C	D	E	F	G	H	I	L	M	N	O
Total no. of asci	16	16	16	16	16	16	16	32	16	16	16	16	48
Via. spore/ascus													
4	1	15	13	3	8	13	3	5	1	3	2	15	7
3	1	0	0	4	0	0	2	8	8	9	3	0	16
2	3	0	1	6	8	0	6	10	1	4	4	0	9
1	4	0	0	0	0	0	3	3	3	0	3	0	0
0	7	1	2	3	0	3	2	6	3	0	4	1	16
Spo. eff.	0.13	0.94	0.81	0.44	0.5	0.81	0.31	0.41	0.56	0.75	0.31	0.94	0.48
Via %	27	94	84	56	75	81	52	52	52	73	44	94	49
SPO	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
GAL	−/−	+/+	+/+	+/−	+/−	+/−	+/+	+/−	+/−	+/−	−/−	−/−	+/−
MEL	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−
TRE	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−
MAL	+/+	+/−	−/−	−/−	+/−	+/−	+/−	+/−	−/−	+/−	+/+	−/−	+/−
CEL	−/−	−/−	−/−	−/−	−/−	+/−	−/−	−/−	−/−	−/−	−/−	−/−	−/−
SUC	+/+	+/+	+/+	+/+	+/−	+/+	+/+	+/+	+/−	+/−	−/−	+/+	+/+
RAF	+/+	+/−	−/−	+/+	−/−	+/−	+/+	+/−	−/−	+/−	−/−	+/−	+/−
GLY	+/+	+/−	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
H_2S	+/+	+/−	+/−	+/+	+/−	+/−	+/+	+/−	+/+	+/+	+/+	+/+	+/−
Cu^{2+}	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/−	+/+	+/+
Het	0	4	1	1	4	5	1	4	2	4	1	1	4

[a] Class refers to the *SED1* PCR profile. Total no. of asci, number of four-spore asci dissected; Via. spore/ascus, number of asci with 4, 3, 2, 1, or no viable spores; Spo. eff., sporulation efficiency; Via%, spore viability; SPO, ability of individual spore clones to sporulate; GAL, MEL, TRE, MAL, CEL, SUC, and RAF, fermentation tests on the corresponding sugar; GLY, glycerol assimilation; H_2S, hydrogen sulfide production; Cu^{2+}, resistance or sensitivity to copper ions; Het, number of heterozygosities.

[b] The presence of dominant, recessive, or both alleles among spore progeny from tetrads is indicated as +/+, −/−, or +/−, respectively.

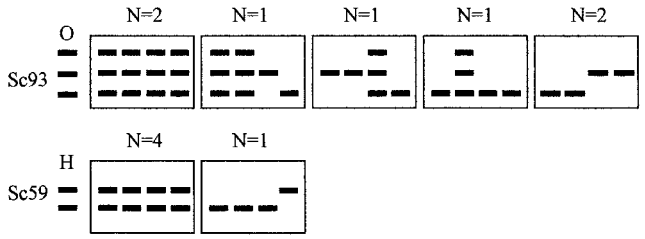

FIG. 5. Segregation of *SED1* length variants in F_1 progeny. O and H, PCR profiles of isolates Sc93 and Sc59, respectively. The boxes contain the *SED1* PCR profiles of the monosporial cultures deriving from each ascus. N, number of asci.

structural and nucleotide sequence similarities with the *SED1* allele of CBS1171, isolated as brewer's top yeast and considered the type strain of *S. cerevisiae*. On the other hand, the *SED1* allele of S288C, which has both a recent "origin" and a complex pedigree (20), differs from those of CBS1171 and all the isolates in both the presence of a 42-bp block in region 1 and six silent SNPs in region 2, which are not present in either CBS1171 or the seven alleles sequenced. Other authors have already highlighted the existence of variations between feral *S. cerevisiae* and S288C, or its isogenic strains, concerning the coding sequences of different genes (18, 19). Our data, besides adding another interesting example to this list, underline once again that this laboratory strain may not be considered fully representative of the species *S. cerevisiae.*

The annealing of the *SED1* sequences indicates that length variations are due to changes in the number of 66- and 153-bp tandem repeat units which mimic the characteristics of classic minisatellites. Thus, the observed gene length polymorphism is very likely a consequence of the different molecular mechanisms proposed for minisatellite array expansion and contraction (3, 10, 27). Hoyer et al. (16), Andersen and Nilsson-Tillgren (1), and Ingavale et al. (17) have reported the presence of repeat motifs within the ORFs of *Saccharomyces carlsbergensis*, *Candida albicans*, and *Schizosaccharomyces pombe* genes, respectively. Our results, besides suggesting that minisatellites in yeast are not as rare as initially supposed, represent the first description of polymorphic minisatellite-like sequences within the ORF of an *S. cerevisiae* gene.

Taking into account that Sed1p provides stress resistance during stationary phase, the phase in which *S. cerevisiae* carries out at least two-thirds of the alcoholic fermentation (23), one interesting question would be whether the generation and maintenance of different *SED1* alleles offer wine yeasts an adaptive advantage in response to the numerous environmental stresses that they undergo during must fermentation. Indeed, the analyses of the predicted amino acid sequences of the alleles sequenced indicate that the observed changes in the number of the 66-bp repeat units in region 1 may affect the number of potential N-glycosylation sites. This could provide a basis for a great diversity in the sugar-mediated interactions with other cell wall components. Similarly, changes in the number of 153-bp repeat units in region 2, by changing the number of cysteine residues, could have functional and biological consequences for the cell wall properties, in case cysteines are involved in molecular disulfide bonds, or for intermolecular interaction with neighboring Sed1p or other cell wall proteins.

At present, the consequences of the observed length variations for the function of the resulting Sed1p have not been explored. However, the fact that all the different *SED1* alleles are transcribed in strains homozygous or heterozygous for this gene (I. Mannazzu and P. Marinangeli, unpublished data) suggests that the observed length polymorphisms are tolerated and that any functional consequence for the resulting proteins should be compatible with the range of functional parameters seen in the population.

Evaluation of the distribution of the 13 *SED1* PCR profiles within the population under study showed that the vast majority of the isolates are homozygous for the *SED1* gene and indicated that PCR profile D predominates in the population analyzed (Table 1). Interestingly, a PCR survey of the *SED1* gene in a population of 132 *S. cerevisiae* isolates from 26 different grapes and musts sampled in six different cellars located in the Puglia region of Italy showed a similar distribution of *SED1* PCR profiles (I. Mannazzu, unpublished data), indicating that these data are probably not the consequence of a population sample bias.

The homothallic behavior, the low number of heterozygosities observed, and the 4:0 segregation ratio of the *SED1* alleles in the isolates presenting PCR profiles A to E are consistent with the "genome renewal" model, according to which diploid homothallic wine strains evolve by meiotic rearrangements which lead to fitter homozygous progeny (21). Also, isolates presenting PCR profiles F to O, heterozygous for *SED1*, are homothallic and show a low number of heterozygosities. Thus, based on this evidence, it seems quite unlikely that the coexistence of different *SED1* alleles in these isolates derives from recent conjugation events between strains carrying different alleles. On the other hand, a plausible course of events in the generation of *SED1* heterozygosity in feral strains of *S. cerevisiae* could involve the occurrence of interallelic or intra-allelic recombination in isolates originally homozygous for this gene.

ACKNOWLEDGMENTS

We are truly grateful to Alan Bakalinsky and Ann Vaughan-Martini for helpful comments on the manuscript.

This work was supported by Fondo di Ateneo, 1999 (I.M.), and by Regione Marche-REG. CEE 2081/83 ob. 5b.

REFERENCES

1. **Andersen, T. H., and T. Nilsson-Tillgren.** 1997. A fungal minisatellite. Nature **386:**771.
2. **Bakalinsky, A. T., and R. Snow.** 1990. The chromosomal constitution of wine strains of *Saccharomyces cerevisiae*. Yeast **6:**367–382.
3. **Bishop, A. J. R., E. J. Louis, and R. H. Borts.** 2000. Minisatellite variants generated in yeast meiosis involve DNA removal during gene conversion. Genetics **156:**7–20.
4. **Boulton, R. B., V. L. Singleton, L. F. Bisson, and R. E. Kunkee.** 1996. Yeast and biochemistry of ethanol fermentation, p. 102–192. *In* R. B. Boulton, V. L. Singleton, L. F. Bisson, and R. E. Kunkee (ed.), Principles and practices of winemaking. Chapman & Hall, New York, N.Y.
5. **Budroni, M., G. Giordano, G. Pinna, and G. A. Farris.** 2000. A genetic study of natural flor strains of *Saccharomyces cerevisiae* isolated during biological ageing from Sardinian wines J. Appl. Microbiol. **89:**657–662.
6. **Caro, L. H., H. Tettelin, J. H. Vossen, A. F. Ram, H. van den Hende, and F. M. Klis.** 1997. *In silicio* identification of glycosyl-phosphatidylinositol-anchored plasma-membrane and cell wall proteins of *Saccharomyces cerevisiae*. Yeast **13:**1477–1489.
7. **Caro, L. H., G. J. Smith, P. van Egmond, J. W. Chapman, and M. Klis.** 1998. Transcription of multiple cell wall protein-encoding genes in *Saccharomyces cerevisiae* is differentially regulated during the cell cycle. FEMS Microbiol. Lett. **161:**345–349.
8. **Coster, F., J. L. Jonniaux, and A. Goffeau.** 1995. Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading frames including

TPS2, *PPH3*, *RAD55*, *SED1*, *PDC2*, *AFR1*, *SSS1*, *SLU7*, and a tRNA for arginine. Yeast **11:**673–679.
9. **de Risi, J. L., V. R. Iyer, and P. O. Brown.** 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science **278:**680–686.
10. **Doege, K. J., S. N. Coulter, L. M. Meek, K. Maslen, and J. G. Wood.** 1997. A human-specific polymorphism in the coding region of the aggrecan gene. J. Biol. Chem. **272:**13974–13979.
11. **Ezaki, B., R. C. Gardner, Y. Ezaki, H. Kondo, and H. Matsumoto.** 1998. Protective roles of two aluminium (Al)-induced genes, *HSP150* and *SED1* of *Saccharomyces cerevisiae*, in Al and oxidative stresses. FEMS Microbiol. Lett. **159:**99–105.
12. **Guerra, E., G. Sordi, I. Mannazzu, F. Clementi, and F. Fatichenti.** 1999. Occurrence of wine yeasts on grapes subjected to different pesticide treatments. Ital. J. Food Sci. **11:**221–230.
13. **Hamada, K., S. Fukuchi, M. Arisawa, M. Baba, and K. Kitada.** 1998. Screening for glycosylphosphatidylinositol (GPI)-dependent cell wall proteins in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **258:**53–59.
14. **Hamada, K., H. Terashima, M. Arisawa, and N. Yabuki.** 1999. Amino acid residues in the ω-minus region participate in cellular localization of yeast glycosylphosphatidylinositol-attached proteins. J. Bacteriol. **181:**3886–3889.
15. **Hardwick, K., J. C. Boothroyd, A. D. Rudner, and H. R. Pelham.** 1992. Genes that allow yeast cells to grow in the absence of the HDEL receptor. EMBO J. **11:**4187–4195.
16. **Hoyer, L. L., S. Schere, A. R. Shatzman, and G. P. Livi.** 1995. *Candida albicans* ALS1: domains related to a *Saccharomyces cerevisiae* sexual agglutinin separated by a repeating motif. Mol. Microbiol. **15:**39–54.
17. **Ingavale, S. S., R. Kaur, P. Aggarwal, and A. Bachawat.** 1998. A minisatellite sequence within the propeptide region of the vacuolar carboxypeptidase Y gene of *Schizosaccharomyces pombe*. J. Bacteriol. **180:**3727–3729.
18. **Laizè, V., F. Tacnet, P. Ripoche, and S. Hohmann.** 2000. Polymorphism of *Saccharomyces cerevisiae* acquaporins. Yeast **16:**897–903.
19. **Liu, H., C. A. Styles, and G. R. Fink.** 1996. *Saccharomyces cerevisiae* S288C has a mutation in FLO8, a gene required for filamentous growth. Genetics **144:**967–978.
20. **Mortimer, R. K., and J. R. Johnston.** 1986. Genealogy of principal strains of the yeast genetic stock center. Genetics **113:**35–43.
21. **Mortimer, R. K., P. Romano, G. Suzzi, and M. Polsinelli.** 1994. Genome renewal: a new phenomenon revealed from a genetic study of 43 strains of *Saccharomyces cerevisiae* derived from natural fermentation of grape musts. Yeast **10:**1543–1552.
22. **Puig, S., A. Querol, E. Barrio, and J. E. Pérez-Ortìn.** 2000. Mitotic recombination and genetic changes in *Saccharomyces cerevisiae* during wine fermentation. Appl. Environ. Microbiol. **66:**2057–2061.
23. **Puig, S., and J. E. Perez-Ortìn.** 2000. Stress response and expression patterns in wine fermentations of yeast genes induced at the diauxic shift. Yeast **16:**139–148.
24. **Rechsteiner, M., and S. W. Rogers.** 1996. PEST sequences and regulation by proteolysis. Trends Biochem. Sci. **21:**267–271.
25. **Rogers, S., R. Wells, and M. Rechsteiner.** 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science **234:** 364–368.
26. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Meiotic mapping, p. 19–29. *In* M. D. Rose, F. Winston, and P. Hieter (ed.), Methods in yeast genetics—a laboratory course manual. Cold Spring Harbor Laboratory Press, New York, N.Y.
27. **Schlotterer, C., and D. Tautz.** 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Res. **20:**211–215.
28. **Shimoi, H., H. Kitagaki, H. Ohmori, Y. Iimura, and K. Ito.** 1998. Sed1p is a major cell wall protein of *Saccharomyces cerevisiae* in the stationary phase and is involved in lytic enzyme resistance. J. Bacteriol. **180:**3381–3387.
29. **Spellman, P. T., G. Sherlock, M. Q. Zhang, V. R. Iyer, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher.** 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol. Biol. Cell **12:**3273–3297.
30. **Ushinsky, S. C., H. Bussey, A. A. Ahmed, Y. Wang, J. Friesen, B. A. Williams, and R. K. Storms.** 1997. Histone H1 in *Saccharomyces cerevisiae.* Yeast **13:**151–161.
31. **Vaughan-Martini, A., and A. Martini.** 1998. *Saccharomyces* Meyen ex Rees, p. 358–371. *In* C. P. Kurtzman and J. W. Fell (ed.), The yeasts, a taxonomic study. Elsevier Press, Amsterdam, The Netherlands.