

Monosomy of a specific chromosome determines L-sorbose utilization: A novel regulatory mechanism in *Candida albicans*

(karyotypes/adaptation/gene regulation)

GUILHEM JANBON*, FRED SHERMAN, AND ELENA RUSTCHENKO†

Department of Biochemistry and Biophysics, University of Rochester Medical School, Rochester, NY 14642

Contributed by Fred Sherman, February 24, 1998

ABSTRACT We report the identification of the gene, *SOU1*, required for L-sorbose assimilation in *Candida albicans*. The level of the expression of *SOU1* is determined by the copy number of chromosome III (also denoted chromosome 5), such that monosomic strains assimilate L-sorbose, whereas disomic strains do not, in spite of the fact that *SOU1* is not on this chromosome. We suggest that *C. albicans* contains a resource of potentially beneficial genes that are activated by changes in chromosome number, and that this elaborate mechanism regulates the utilization of food supplies and possibly other important functions, thus representing a novel general means for regulating gene expression in microbes.

We previously demonstrated that *Candida albicans* spontaneously produces high frequencies of a wide range of chromosomal aberrations, and we suggested that these alterations are the means to achieve genetic variability in an organism that lacks a sexual process (1). This hypothesis was further supported by extensive findings on the relationship between changes in karyotypes and assimilation profiles (2). We also demonstrated that electrophoretic karyotypes can be specifically altered in mutants selected on alternative sugars (3). Although the relationships between altered karyotypes and changes in utilization of food supplies were established, the relationships between chromosomal rearrangements and expression of specific genes remained obscure (2).

As described in this paper, we have cloned and sequenced the *SOU1* gene, which is responsible for the utilization of L-sorbose. Most importantly, we have demonstrated that expression of *SOU1* is controlled by the number of a particular chromosome, such that strains disomic and monosomic for chromosome III (also denoted chromosome 5) are, respectively, nonutilizers and utilizers of L-sorbose. Because *SOU1* does not reside on chromosome III, we suggest that chromosome III contains a negative regulatory factor that inhibits expression of *SOU1* in two but not one copy. These results confirm our early hypothesis on the control of gene expression by chromosomal rearrangements and constitute an example of a novel general means of gene regulation based on chromosomal number.

MATERIALS AND METHODS

Nomenclature of *C. albicans* Chromosomes. *C. albicans* chromosomes are denoted by several numbering systems (4), including the nomenclature introduced by Magee *et al.* (5) and by us (1). In Magee's nomenclature, the penultimate largest to the smallest chromosome is designated by Arabic numerals 1–7, whereas the largest chromosome, containing the ribosomal DNA cluster, is designated R. The symbol R for the largest chromosome was apparently introduced because the 1–7 numerals were assigned before the largest chromosome was recognized (6). Our nomen-

clature uses roman numerals I–VIII to designate the smallest to the largest chromosomes, respectively, and “a” and “b” to designate two homologues of different sizes. The following are equivalencies for the two types of chromosome assignments:

Chromosome R	Chromosome VIII
1	VII
2	VI
3	V
4	IV
5	III
6	II
7	I

In this paper we have elected to use our nomenclature for designating chromosomes.

Designations of Genes and Strains. The symbol *SOU1* denotes the gene encoding L-sorbose dehydrogenase, which is required for L-sorbose use. *SOU2* denotes a gene that has a sequence similar to *SOU1*, but that is not required for L-sorbose use. The phenotypic symbols, *Sou*⁺ and *Sou*[−], denote the ability and inability, respectively, to grow on L-sorbose medium. Mutants derived from 3153A and CAF4–2 that affect sorbose utilization are denoted as Sor, followed by an isolation number. For example, Sor55 denotes a *Sou*⁺ mutant derived from the *Sou*[−] parental strain 3153A, whereas Sor55–1 denotes *Sou*[−], a revertant derived from the *Sou*⁺ mutant Sor55. Furthermore, Sor55–1-1 designates a *Sou*⁺ mutant subsequently derived from the *Sou*[−] revertant Sor55–1.

Strains and Plasmids. The two *Sou*[−] parental strains of *C. albicans*, CAF4–2 (obtained from W. A. Fonzi, ref. 7) and 3153A (1) were previously described. 3153A has been extensively used by us for studies of chromosomal instability, including the isolation and characterization of *Sou*⁺ strains (1–3). CAF4–2 is a *Ura*[−] strain that is conveniently used for cloning and disrupting genes and for other molecular manipulations. *Sou*⁺ mutants were obtained from independent subclones of parental strains on L-sorbose medium, as described (3), to ensure that each mutant arose from a separate event and to avoid selection of siblings (1).

CAF4–2 and 3153A, and their *Sou*⁺ derivatives, Sor17 and Sor1, respectively, were used to show that the *SOU1* sequence was identical in all four of these strains. Monosomy of chromosome III was studied in seven *Sou*⁺ mutants (Sor24, Sor25, Sor52, Sor53, Sor55, Sor56, and Sor58) from 3153A and six *Sou*⁺ mutants (Sor12, and Sor18–Sor22) from CAF4–2, as well as derivatives of these mutants. CA18, CA19, CA20, and CA21 denote, respec-

Abbreviation: YPD, yeast extract/peptone/dextrose.

Data deposition: The sequence reported in this paper has been deposited in the GeneBank database (accession no. AF002134).

*Present address: Institut Pasteur, Unité de Mycologie, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France.

†To whom reprint requests should be addressed at: Department of Biochemistry and Biophysics, Box 712, University of Rochester Medical School, Rochester, NY 14642. e-mail: elena_bulgac@urmc.rochester.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/955150-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

tively, the disrupted strains *SOU1/sou1-Δ::hisG-URA3-hisG*, *SOU1/sou1-Δ::hisG*, *sou1-Δ::hisG/sou1-Δ::hisG-URA3-hisG*, and *sou1-Δ::hisG/sou1-Δ::hisG*.

Saccharomyces cerevisiae strain B-6929 (*MAT^{ahis3-Δ1} trp1-289 ura3-52*) and *Escherichia coli* XL-1 Blue strain (8) were used for the preparation of plasmids and other manipulations.

The integrative plasmid pRC3915 and the low-copy replicative plasmid pRC2312 (9), the plasmid pCUB-6 containing the *hisG-URA3-hisG* module (7), and pUC18 (10) have been described. Other plasmids, denoted pCA and followed by a number, were constructed in this study with the above mentioned plasmids.

Media and Growth Conditions. Yeast extract/peptone/dextrose (YPD), L-sorbose, and SD media, media supplements, and conditions for growth were previously described (3, 7, 11). *C. albicans* strains were maintained by the procedure that prevented changes due to chromosomal instability (12).

Cloning of the *SOU1* Gene by Complementation of *C. albicans*.

The *SOU1* gene was cloned by complementing the *Sou*⁻ phenotype of CAF4-2 with a library prepared from Sor17 DNA. Sor17 was pregrown in L-sorbose medium supplemented with uridine and then subsequently grown for approximately seven generations in YPD medium to allow the cells to become susceptible to spheroplast formation. Spheroplasts were prepared (13) and DNA was isolated (14) as described. DNA was partially digested with *Sau3A*, and 7–10 kb fragments were separated on a 0.8% agarose gel, eluted (GeneCleave kit, Bio 101) and ligated with a low copy number vector pRC2312 (9) that had been cut with *Bam*HI and treated with calf intestinal alkaline phosphatase. *E. coli* XL-1 Blue was transformed with the ligation mixture (15), followed by the preparation of plasmid DNA from approximately 3 × 10⁴ combined colonies. More than 85% of the plasmids in the resulting library contained inserts having an average size of 7 kb. CAF4-2 was transformed with the genomic library as described (13), and the transformants were incubated for 3 days at 37°C on SD medium containing 1 M L-sorbitol. Approximately 5,000 transformants were individually transferred to SD plates and grown for 2 more days. To produce uniform and sufficient growth, each of the transformants was transferred once again on SD plates, incubated for 2 days, and subsequently replica plated on L-sorbose medium. DNA from each of the positive transformants was extracted from cells grown in SD medium and used to transform *S. cerevisiae* strain B-6929. The passage of the pertinent plasmids through *S. cerevisiae* was necessary for monomerization (16), despite the report to the contrary for pRC2312 and its derivatives (9). A total of five *S. cerevisiae* colonies from each transformant were grown in liquid SD medium containing histidine and tryptophan; DNA was prepared and used to transform *E. coli*. The plasmids, purified from the *E. coli* transformants, were subsequently used to transform *C. albicans* CAF4-2; three groups of such plasmids, originating from pCA1, pCA2, and pCA3, produced *Sou*⁺ transformants. In summary, of a total of 4,736 *Ura*⁺ transformants, 3 had the following properties: they were *Sou*⁺, as indicated by confluent growth of the spots in 2 days at 37°C; they were plasmid dependent, in that the *Ura*⁻ derivatives lacking the plasmids were *Sou*⁻; and they produced *Sou*⁺ transformants when the corresponding plasmids were reintroduced in *C. albicans*. Restriction mapping of the three plasmids, pCA1, pCA2, and pCA3, indicated that the inserted DNA contained common sequences that encompassed the *SOU1* gene. A physical map of a portion of the 7 kb insert from pCA1 is shown in Fig. 1.

DNA Sequence Analysis. The dideoxynucleotide chain termination method of Sanger *et al.* (17) was used with Sequanase enzyme (United States Biochemical) and with [α -³⁵S]dATP. DNA sequences were analyzed with the Genetics Computer Group programs (18).

Gene Disruptions, Integrations, and PCR Cloning. The *SOU1* gene was disrupted by the *URA3* blaster protocol (7), by using the plasmid pCA7 that was constructed by replacing a *Hpa*I–*Swa*I segment (Fig. 1) with the 3.8-kb *Bam*HI/*Bgl*II fragment carrying

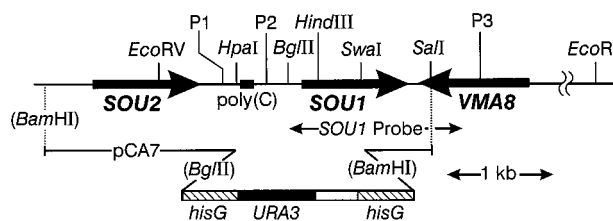


FIG. 1. Physical map of the *SOU1* region showing *SOU2* and *VMA8* flanking genes; the C-rich 40-bp poly(C) segment; P1, P2, and P3 primers used to generate PCR fragments; *SOU1* probe used in Southern and Northern blot analysis; and a portion of the pCA7 plasmid used for disrupting *SOU1*.

the *hisG-URA3-hisG* module. Integrations of *SOU1* gene into genome of parental strain CAF4-2 were performed with the plasmids pCA4 and pCA5, which were constructed with the integrative plasmid pRC3915 (9) and with inserts from, respectively, pCA1 and pCA2 containing *SOU1*; the plasmids were cut with *Sph*I and *Bgl*II so as to target the integration, respectively, to the *LEU2* and *SOU1* loci. Regions containing the *SOU1* gene in various strains were amplified by PCR by using two pairs of primers, P1/P3 and P2/P3 (Fig. 1). The PCR products were ligated with the replicative plasmid pRC2312 cut with *Sma*I.

DNA and RNA Preparations, Hybridization Probes, and Protocols. DNA and RNA were prepared, respectively, as described above and by Guo and Sherman (19). Southern and Northern blot analyses were performed as described (ref. 15; Genius System version 3.0, Boehringer Mannheim). Northern blot analyses were carried out with 10 μ g samples of RNA, as determined spectrophotometrically. Plasmids encompassing the *SOU1*, *BEN4*, and *CAG1* genes were used as probes in Southern and/or Northern blot analyses. Labeling with [α -³²P]dATP by the random priming method, as well as preparation of autoradiograms, was carried out by standard methods (15). Northern blots were calibrated by comparing the intensities of ethidium bromide stained rRNA bands on the gel with the intensities of *SOU1* probe signals from the corresponding blot, by using the FluorImager 575 (Molecular Dynamics) and PhosphorImager (Molecular Dynamics), respectively.

Pulsed-Field Gel Electrophoresis. In this work, we have designed a protocol for obtaining pure cultures of *Sou*⁺ and *Sou*⁻ cells used in the examination of electrophoretic karyotypes. Care was taken to ensure that each of the serially derived mutants was directly obtained from the previous strain of the series, as follows:



We have explored several methods to prepare serially derived electrokaryotypes. Initially, a *Sou*⁺ colony, which appeared after plating the *Sou*⁻ parental cells on a L-sorbose plate (3), was transferred to liquid L-sorbose medium and grown to a cell mass sufficient for preparing chromosomes. The same *Sou*⁺ cell suspension was streaked on a YPD plate and one large colony was transferred to liquid YPD medium and incubated overnight to enrich for *Sou*⁻ revertants. The cycle was repeated with *Sou*⁻ revertant cells plated on L-sorbose medium, thus obtaining a serially derived *Sou*⁺ mutant, which was prepared for the electrokaryotype as described above.

In another procedure, *Sou*⁺ mutants, previously obtained on L-sorbose medium (3) and deposited at -70°C on glycerol (12), were subcloned on YPD plates. After incubation of the plates at 37°C for 2–3 days, a colony, preferably a small one, was suspended in water and approximately 5,000 cells were spread on each of two YPD plates. This allowed us to avoid an accumulation of *Sou*⁻ revertants that in some strains appeared as large colonies, which could dominate the population. After the colonies grew up, the large colonies, if any, were excised from plates and the remaining small colonies were washed from the plates, inoculated in 200 ml of YPD, and allowed to

divide for several generations to condition the cells for spheroplasting. Subsequently, chromosomes were prepared from an appropriate number of cells. The same cell suspension was subcloned on YPD, one colony inoculated into 15 ml of YPD, and the culture was incubated on a rotating shaker for 3 days with daily reinoculations into fresh medium to accumulate revertants. The last passage was transferred to 200 ml of YPD for obtaining cells to prepare chromosomes. The same revertant *Sou*⁻ culture was plated on L-sorbose medium to obtain a sequential *Sou*⁺ mutant as described earlier (3). For all procedures, the phenotypes of each of the *Sou*⁺ and *Sou*⁻ cell populations were assayed by the growth or lack of growth on L-sorbose plates.

The main difference between the two procedures described above is that the second avoids prolonged growth in L-sorbose medium, a condition that can result in the accumulation of additional mutations.

In addition, the preparations of the karyotypes containing either monosomic or disomic chromosome III are conveniently obtained from certain strains having distinct small and large colonies, respectively.

The methods for preparing intact chromosomes (20) and the conditions for running orthogonal-field alternation gel electrophoresis and contour-clamped homogenous electric field system (Bio-Rad) (12, 31) have been described.

To ensure maximum resolution in separating chromosomes, three optimal conditions were used, each favoring the best separation of the shortest, middle sized, and largest bands on the gels.

RESULTS

Cloning and Sequencing of the *SOU1* Gene Responsible for L-Sorbose Utilization in *C. albicans*. The *SOU1* (SOrbose Uti-*lization*) gene was isolated with a library prepared from the *Sou*⁺ mutant Sor17 by complementing the parental strain CAF4-2 inability to grow on L-sorbose as a sole source of carbon. Each of the three *Sou*⁺ transformants yielded plasmids with either 7 kb (pCA1) or 5 kb (pCA2 and pCA3) inserts. Restriction maps of cloned fragments, subcloning, and complementation experiments established that the gene of interest encompasses an *Hind*III site, as shown in Fig. 1. Sequencing the regions on both sides of the *Hind*III site revealed an ORF of 843 bp, which encoded a potential protein of 30 kDa. This sequence shares homologies with sequences of proteins from the short-chain alcohol dehydrogenase family. A total of 23 of 25 highly conserved residues in this family, identified by Persson *et al.* (22), are present in this sequence, denoted *Sou1p* (Fig. 2). In particular, the conserved residues Gly43, Gly47, and Gly49 apparently correspond to the highly conserved NAD-binding site of many dehydrogenases (23). Comparisons of *Sou1p* revealed 57% similarities and 39.5% identities with *C. albicans* arabinol dehydrogenase (24), and 54.5% similarities and 37.4% identities with *Candida tropicalis* arabinol dehydrogenase (25). Thus, we suggest that *Sou1p* corresponds to L-sorbose dehydrogenase. Despite the failure to express *SOU1* in *S. cerevisiae* (data not shown), its sequence does not contain any CUG codons, which determine serine and not leucine residues in *C. albicans* (26).

Sequencing of Flanking Genes, *SOU2*, and a Homologue of *S. cerevisiae* *VMA8*. Sequencing upstream of *SOU1* revealed an ORF of 840 bp, encoding a putative 30 kDa protein, which is 83.5% similar and 72.5% identical to *Sou1p*, and that was denoted *Sou2p*. The similarities between *SOU1* and *SOU2* are distributed along the translated region, with no obvious similarity in the 5' and 3' untranslated regions. The 806-bp-long intergenic region contains a 40-bp-long sequence highly rich in C, with a contiguous stretch of 10 C nucleotides (Fig. 2). The similarities and the organizations of *SOU1* and *SOU2* are reminiscent of regions containing the two related aspartyl proteases genes, *SAP1* and *SAP4*, which are also separated by a similar highly C-rich region (27). The tandem positioning of two similar genes rarely occurs in yeast, although it is common in mammals (see refs. 28 and 29).

```

-560 caattaatcaacagaattataatgattacaaatcttaactccttttggaatcctctaag
-500 ttaaccattgaaaggatcaactaaagaaatagaccaactccgagagctggattaaact
-440 atttaccggagaatacaatcagagaatattccagtagcaaggattcttgatcctcttttcc
-380 tccctcttttcccccacaaaagatgcccccccacatttaagattactataacccttaa
-320 aatggccaataaagaactttttttgtgtaataatggagcatatacaaatagattttaag
-280 ccaatttgctactattatttttggatgtggtgtagcttgggtttatctaatatttttt
-220 ccaattctgcaatccttcgagtcacactcccttaaccctatctcctccgagatctcatt
-180 ttactgttagatttaactcgaatcaattgttttaatttttttaagtttgctatcaaaa
-120 aaaaaatacataatcatcaaataccoccccaacattcttcttccctttatttct
-60 tttttataaattcaatattatcaaaaaattttcaaaccaattcaaatcaaatcatt
1 ATGAGTGAAGAAATCATTTCATTCACAAATCCTGCTTTAGGTCCTACCAACAAAAGT
1 M S E E I I S F T N P A L G P L P T K A
. . . K . T . T . Y . . A K . . . . .
61 CCACAATTACCATCAAAATGTTCTTGATTTGTTTCTTTAAAGGTAAGTCGCTCCGCTG
21 P Q L P S N V L D L F S L K G K V A S V
A T I . D . I . . A . . . . .
121 ACGGGATCATCTGGAGGAATGGTTGGGCTGCGCGAAGCATTGTCTCAAGCTGGTGCT
41 T G S S G G I G W A V A E A F A Q A G A
. . . . . G . . . . .
181 GATGTTGCCATCTGGTATAATTGAAACCAGCAGATGCCAAAGCTGAATTTTAACTGAA
61 D V A I W Y N S K P A D A K A E Y L T E
. . . . . H . D . . . . . K
241 AAATATGGGTCAAAGCCAAAGCTTATAAATGTAATGTAAGTGAATGATCTAATGATGTTCT
81 K Y G V K A K A Y K C N V T D P N D V S
T . . . . S . . . . . F Q . . . E
301 AAAGTGATTAATGAAATGAAAAAGATTCGGTACTATTGATATATTGTTGCTAATGCT
101 K V I N E I E K D F G T I D I F V A N A
. . . V . Q . . . L . . . . .
361 GGAGTGCATGGACTGATGGACCAGAAATGATGTTCAAGGCTATGATCAATGGAAGAAAG
121 G V A W T D G P E I D V Q G Y D Q W K K
. . . . . E . . . K . . . K . V . K . N .
421 ATCGTTGATTGTGATTAAATGGAGTTTATTATTGCTCATACCGTGGGACAAATCTTT
141 I V D C D L N G V Y Y C A H T V V G Q I F
V . . V . . . . S . . . . . V . . P . . R
481 AAAAAAGATAAATCTGGTTCAATTAATTAATCTCATCAATGTCAGGGACAATGTTAAT
161 K K N K S G S L I I T S S M S G T I V N
. . G . . . . F . F . A . . A S . . .
541 ATCCCTCAATAGCAAGCTCCTTATAATGCTGCTAAAGCTGCACTGACTCATTTAGCCAAA
181 I P Q L Q A P Y N A A K A A G T C H L A K
V . . . . P . . . . G . . . . S .
601 TCATVAGTGTGGAATGGGCTAGTTTGGTCTAGAGTAAATTCAAATTTCCAGGGTAT
201 S L S V E W A S F G A R V N S I S P G Y
. F T . . . . P . . . . V . . . .
661 ATATTGACTGATTTGCTGATTGCTGATCCAGAAAGAAAAGAAATGGTGGCAATTG
221 I L T D I A D F A D P E M K K K W Q L
. A . H L S E . . D . . V . S . . L . .
721 ACACCTTGGGAAGAGAAGGATACCACAAGAATTAGTGGGGCATATTTACTTGGCC
241 T P L G R E G L P Q E L V G A Y L Y L A
. . . . . A K . R . . . . .
781 TCAATGCATCAACTTATACTACTGGTCAAAATATGCTGTTGATGGGGTTATACATGT
261 S N A S T Y T T G S N I A V D G G Y T C
. D . A S . . . . A D L . . . . V
841 CCAtaataaaataagggtgtgtgctttaaagaagaaaaaatcttatagaataatcgc
281 P End
I
901 actgtaaaacataaacatcatctttttggttatgtattaataaaaaataagttacta
961 ataaacatatactgtaaaaaaacacaataaccacat

```

FIG. 2. Nucleotide sequence of the *SOU1* gene and deduced amino acid sequence of the corresponding *Sou1p*. The residues conserved in the short-chain alcohol dehydrogenase family are underlined. The C-rich domain and the putative TATA elements are underlined in the 5' untranslated region. The *Sou2p* sequence is shown below the *Sou1p* sequence, where identical residues are depicted by dots.

Two similar adjacent genes are believed to arise by duplication of a region, followed by divergent evolution. It is possible that the C-rich region plays a role in the duplication of genes in *C. albicans*. Although *SOU1* and *SOU2* are similar, multicopy *SOU2* plasmids did not provide growth of CAF4-2 on L-sorbose medium.

Sequencing of the region downstream of *SOU1* revealed an 801-bp ORF that corresponded to a protein 85.2% similar and 73.8% identical to subunit 8 of the vacuolar H⁺-ATPase *Vma8p* of *S. cerevisiae* (30) (Fig. 1).

Disruption and Reintegration of the *SOU1* Gene. The function of *SOU1* was confirmed by disruption of the two *SOU1/SOU1* alleles, and by subsequently introducing *SOU1* plasmids in the resulting *sou1-Δ/sou1-Δ* strain. Because Sor17 had a high rate of the reversion to *Sou*⁻ (see below), which can confuse the formation of deleted and reverted strains, the construction was carried out with the *Sou*⁻ parental strain CAF4-2 and testing for the ability to produce *Sou*⁺ mutants on L-sorbose medium. Both *SOU1* alleles were sequentially disrupted by using the *URA3* blaster procedure (Fig. 1), and the *SOU1/SOU1*, *SOU1/sou1-Δ* and *sou1-Δ/sou1-Δ* strains were analyzed by Southern blotting and for the ability to form *Sou*⁺ mutants on L-sorbose medium (Fig. 3). Hybridization signals were, as expected, corroborating the sequential elimination of the two *SOU1* alleles. The single deleted derivative, *SOU1/sou1-Δ*, gave rise to *Sou*⁺ mutants, although their growth rates were diminished (Fig. 3). The double

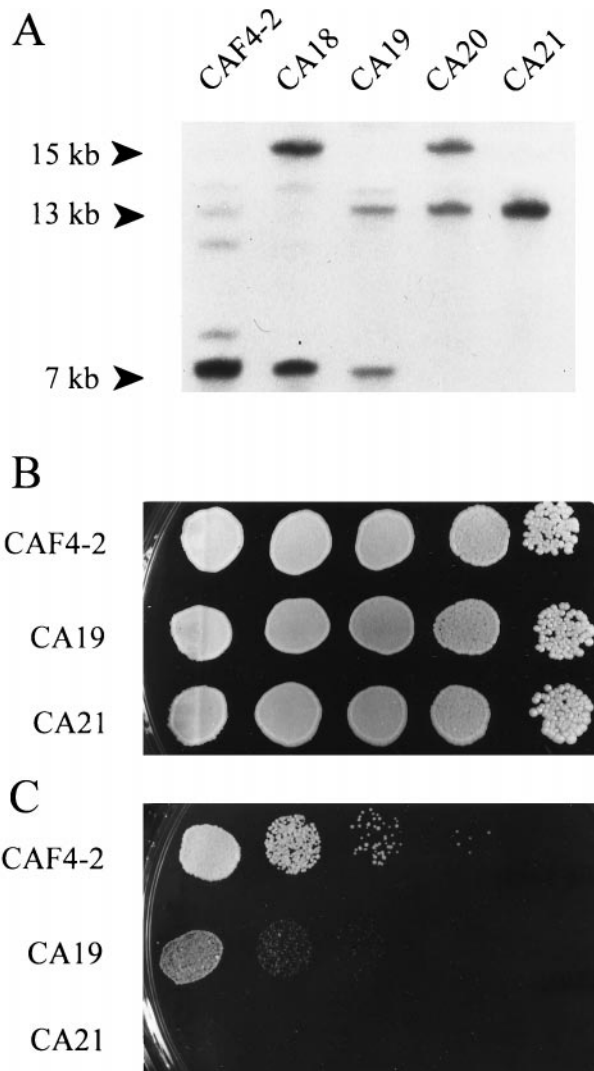


FIG. 3. (A) Southern blot analysis of DNA digestion with *Bgl*II and hybridization with the *SOU1* probe of the following strains: CAF4-2, the normal *SOU1/SOU1* strain; CA18, the singly disrupted *SOU1/sou1-Δ::hisG-URA3-hisG* strain; CA19, the singly disrupted *SOU1/sou1-Δ::hisG* strain; CA20, the doubly disrupted *sou1-Δ::hisG/sou1-Δ::hisG-URA3-hisG* strain; and CA21, the doubly disrupted *sou1-Δ::hisG/sou1-Δ::hisG* strain. (B) Comparative growth of CAF4-2 and CA19 and CA21 on glucose medium (SD plus uridine) for 6 days at 37°C and (C) on L-sorbose medium (plus uridine) for 14 days at 37°C. Each strain was spotted as serial 1/10 dilutions, from left to right, starting from an initial concentration of 10⁶ cells per ml.

deleted derivative, *sou1-Δ/sou1-Δ*, did not produce any *Sou*⁺ mutants. However, *Sou*⁺ mutants were produced after the reintegration of *SOU1* at the *LEU2* locus in the *sou1-Δ/sou1-Δ* strain (data not presented).

Chromosomal Assignment of *SOU1*. The chromosomal assignment of the *SOU1* gene was performed for both parental strains, 3153A and CAF4-2, as well as for their *Sou*⁺ mutants, Sor1 and Sor17, respectively. No differences in the position of *SOU1* in the parental strains and their derivatives were observed, indicating that the gene was retained on the same chromosome.

As shown schematically in Fig. 4A, Southern blot analysis of separated chromosomes revealed that one allele of *SOU1* in 3153A and its derivatives is located on chromosome Ib and another on chromosome IVb, whereas *SOU1* in CAF4-2 and its derivatives is located on both homologues of chromosome IV (data not presented). We suggest that the difference between the two strains may be due to a translocation.

Comparison of the *SOU1* Genes from *Sou*⁺ and *Sou*⁻. *SOU1* was amplified by PCR from the *Sou*⁻ strain CAF4-2 and its *Sou*⁺ derivative, Sor17, by using the two pairs of primers, P1/P3 and P2/P3 (Fig. 1). The PCR fragments from both CAF4-2 and Sor17, and with and without the 40 bp poly(C) region were then inserted in the low-copy replicative plasmid pRC2312 and used to transform CAF4-2. All of the transformants grew confluent on L-sorbose medium, independently of the origin of the fragments and of the presence of the C-rich region upstream of *SOU1*. Moreover, the sequence of *SOU1* from CAF4-2 was determined and shown to be identical to the one from Sor17. In addition, no differences were observed by Southern blot analysis of restriction fragments (data not shown). The same results were obtained with another pair, the *Sou*⁻ parental 3153A strain and the *Sou*⁺ mutant Sor1.

Electrophoretic Karyotypes of *Sou*⁺ and *Sou*⁻ Strains. In this work, we analyzed 33 electrophoretic karyotypes of sequential *Sou*⁺ and *Sou*⁻ strains, which were obtained as described. An examination of three complete series, i.e., *Sou*⁻ (parental) → *Sou*⁺ → *Sou*⁻ → *Sou*⁺, from 3153A, denoted Sor25, Sor52, and Sor53, as well as one complete series, Sor19, from CAF4-2, established that *Sou*⁻ and *Sou*⁺ strains were, respectively, disomic and monosomic for chromosome III and that this relationship could be continuously perpetuated with each of the mutants derived from the opposite type, as illustrated in Fig. 4. This result was supported by four additional series of 3153A (Sor24, Sor55, Sor56, and Sor58) and five of CAF4-2 (Sor12, Sor18, and Sor20–Sor22), for which only the *Sou*⁻ (parental) → *Sou*⁺ → *Sou*⁻ members were examined, but not the second *Sou*⁺ mutants. Southern blot analyses with the *CAG1* probe confirmed that the monosomic and disomic conversions were displayed by chromosome III (data not shown), consistent with our previous results (3). Because homologues IIIa and IIIb have different lengths in strain 3153A and can be electrophoretically separated, we have demonstrated that either homologue can be lost in *Sou*⁺ strains, and that the preserved homologue is duplicated in the subsequent *Sou*⁻ derivative, as schematically presented in Fig. 4D (see also ref. 3). Our precise separations, which cover the entire range of chromosomal sizes, permitted a reliable comparison of electrophoretic karyotypes. On the whole, only chromosome III was effected in each of 33 karyotypes, except for an additional change in chromosome II that appeared in Sor25 and the subsequent members of this series. The alteration of chromosome II probably reflects the general high chromosomal instability in *C. albicans* and is unrelated to the *Sou* phenotypes. We also do not consider changes in chromosome VIII to be relevant to L-sorbose utilization. Such changes were not related to the *Sou* phenotypes, and occurred in approximately one-third of the mutants analyzed in this and previous studies (3), a result that is consistent with the high variation in the number of ribosomal DNA units in this chromosome (31).

Of a total of 28 *Sou*⁺ mutants independently derived from 3153A and analyzed in this and the previous study (3), 21 (75%), had either homologue IIIa or IIIb missing and had the preserved homologue duplicated in the subsequent *Sou*⁻ derivative. Also, we wish to point out that in our earlier study we did not appreciate the selection of *Sou*⁻ cells in YPD medium, and chromosomal DNA was prepared from cells grown overnight in YPD medium, a condition that resulted in the accumulation of the *Sou*⁻ cells in the *Sou*⁺ culture and in karyotypes containing predominantly duplication of either homologue IIIa or IIIb (3), which nevertheless demonstrated the specific alteration of chromosome III.

Expression of *SOU1*. The determination of the levels of *SOU1* mRNA in the *Sou*⁻ parental strain 3153A and its *Sou*⁺ derivatives, Sor52 and Sor53, and their *Sou*⁻ revertants Sor52-1 and Sor53-1, revealed that *SOU1* is more highly expressed in *Sou*⁺ strains, compared with the *Sou*⁻ strains, with an approximate increase of 5–7 times (Fig. 5).

Relationship Between the *SOU1* Gene Copy Number and the *Sou*⁺ Phenotype. As described above, the normal *SOU1/SOU1*

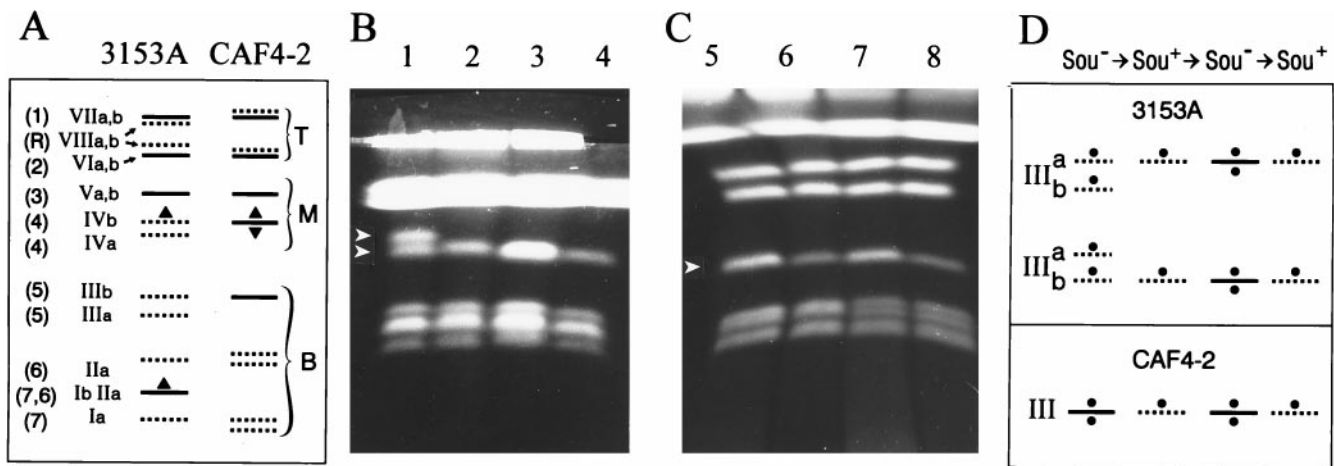


FIG. 4. (A) Schematic representation of the electrophoretic karyotypes from 3153A and CAF4-2 showing the positions of the *SOU1* gene (\blacktriangle and \blacktriangledown). Orthogonal-field-alternation gel electrophoresis separation of chromosomes of two typical representative sequential series of strains derived from the parental strains (B) 3153A and (C) CAF4-2. Orthogonal-field-alternation gel electrophoresis. Lanes: 1, 3153A (*Sou*⁻); 2, Sor55 (*Sou*⁺); 3, Sor55-1 (*Sou*⁻); 4, Sor55-1-1 (*Sou*⁺); 5, CAF4-2 (*Sou*⁻); 6, Sor19 (*Sou*⁺); 7, Sor19-1 (*Sou*⁻); 8, Sor19-1-1 (*Sou*⁺). Conditions of separation in these two representative gels were selected for the best resolution of either the smallest or the smallest and middle-sized groups of chromosomes (21). The gels showing precise separations of the other chromosomes of these strains are not presented. (D) Schematic representation of chromosomes III from the *Sou*⁺ and *Sou*⁻ strains serially derived from 3153A and CAF4-2 showing the hypothetical negative regulator *CSU1* gene (\bullet).

CAF4-2 parental strain became *Sou*⁺ when transformed with replicating plasmids containing various overlapping inserts with *SOU1*. Because the cloning vector, pRC2312, had a copy number of 2-3 (9), we conclude that transformants became *Sou*⁺ having a total of 4-5 copies of *SOU1*, even being disomic for chromosome III (see above). In addition, we have examined the normal *SOU1/SOU1* strain having the *SOU1* gene integrated at either the *LEU2* or *SOU1* loci, thus allowing the generation of strains with

a minimum number of three copies of this gene (*SOU1/SOU1 LEU2/LEU2::SOU1* or *SOU1/SOU1::SOU1*). The strains were *Sou*⁻, and gave rise to *Sou*⁺ mutants at approximately the same frequency as the normal strain with two copies. However, the *Sou*⁺ mutants arising from the strains with three copies grew faster than the *Sou*⁺ mutants derived from the *SOU1/SOU1* parental control strain that contained an integrated plasmid with no insert (data not shown). Furthermore, integration of multiple copies of *SOU1* produced confluent growth on L-sorbose medium (data not presented). Although we did not rigorously establish the minimum number of *SOU1* genes required for the *Sou*⁺ phenotype, it is clear that it is copy number dependent, as illustrated by both the method used to originally clone the gene, as well as the integration experiment.

In summary, two copies of the *SOU1* gene, along with a single copy of the hypothetical *CSU1* gene, which acts as negative regulator (Fig. 4D) result in growth on L-sorbose medium (see above). Furthermore, three copies of *SOU1* are insufficient for growth with two copies of *CSU1*. If the same ratio of 2:1 is required, than four copies of *SOU1*, along with two copies of *CSU1*, may result in the *Sou*⁺ phenotype. It should be noted that, for example, a strain with three copies of *SOU1* (i.e., *SOU1/SOU1 LEU2/LEU2::SOU1*) can readily give rise to a strain with four copies by mitotic crossing over (*SOU1/SOU1 LEU2::SOU1/LEU2::SOU1*).

Revertability of the *Sou*⁺ Strains. When plated on YPD, approximately 66% of *Sou*⁺ mutants derived from 3153A, but not from CAF4-2, showed two distinct sizes of colonies, large and small. Typically, large colonies were approximately the size of colonies from the original strain 3153A. The large colonies were shown to be *Sou*⁻ and disomic for chromosome III, and the small were shown to be *Sou*⁺ and monosomic for chromosome III. After 4 days of incubation at 37°C on YPD, all of the small colonies revealed sectoring, as a result of faster growing revertant *Sou*⁻ cells (data not shown). When a small colony was picked and an appropriate dilution was replated on YPD, from 0.1-1% of the colonies were large. This range of frequencies of *Sou*⁻ colonies among *Sou*⁺ colonies was observed with four sequential platings of small *Sou*⁺ colonies. An approximate rate of reversion of the *Sou*⁺ \rightarrow *Sou*⁻ was calculated to be 3×10^{-4} reversion/cell/generation, by using the frequencies of large colonies and the equation of Jiang and Stillman (32). Although the rate of reversion was not corrected for the differences in the rate of

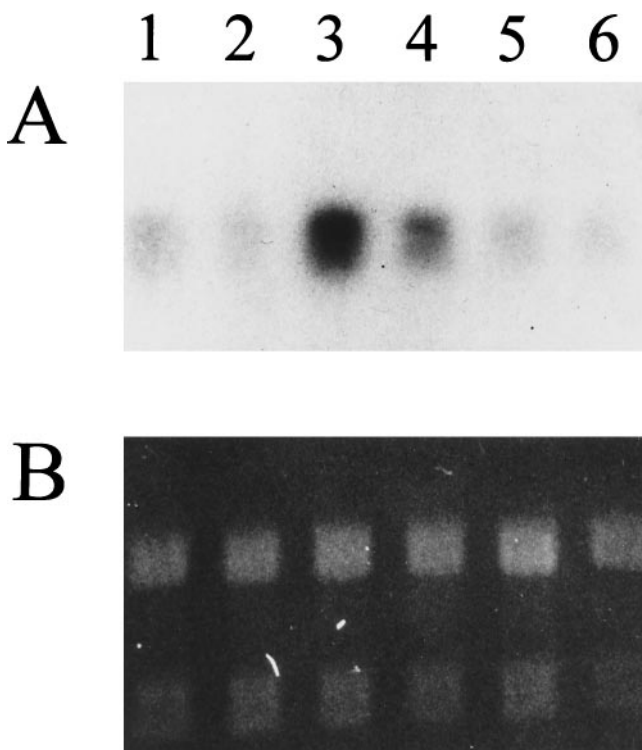


FIG. 5. (A) Expression of *SOU1* as determined by Northern blot analysis with a *SOU1* probe and with the following strains. Lanes: 1 and 2, parental strain 3153A (*Sou*⁻); 3, Sor52 (*Sou*⁺); 4, Sor53 (*Sou*⁺); 5, Sor52-1 (*Sou*⁻); and 6, Sor53-1 (*Sou*⁻). (B) Relative *SOU1* mRNA levels were estimated by comparing the hybridization signals with the ethidium bromide fluorescence intensities of the rRNA bands.

growth of the Sou^- and Sou^+ strains on YPD medium, the experiments and the corresponding calculations were carried out with newly formed small colonies to minimize this error.

Among six Sou^+ mutants derived from CAF4-2 with altered chromosome III, none showed two distinct sizes of the colonies. Instead, some mutants had uniformly looking colonies, whereas others showed different colony sizes, but all were Sou^+ (data not presented). It is important to note that populations of Sou^+ derivatives of CAF4-2 did not appear to revert during cultivation in YPD medium as rapidly as the above mentioned mutants of strain 3153A. No sectoring, characteristic of revertants, was observed for up to 10 days of incubation of the YPD plates at 37°C. The difference between 3153A and CAF4-2 could be due to differences in reversion rate or to differential growth rates of the Sou^+ and Sou^- cells in the populations.

DISCUSSION

An important feature of all microbes is their ability to profit from food opportunities. Normally, microbes do not have all of the enzymes required to metabolize diverse substrates. However, they evolved means to express certain genes in response to the changing environment. By far the most common means for a microorganism to assimilate a newly encountered carbon source is induction of enzymes at the transcriptional level (33). A major feature of the induction is phenotypic mass conversion of cells in the population. Other less frequent examples are related to the extraordinary capacity of the microbial genomes of both prokaryotes and lower eukaryotes to form beneficial genes. A well understood mechanism includes mutational activation of cryptic genes and operons as exemplified by activation of various biosynthetic pathways in *Lactobacillus* and cryptic operons for β -glucosidases in *E. coli* and *Salmonella* (34). We have provided here, for the first time, to the best of our knowledge, evidence that chromosomal number controls expression of a particular gene, establishing a novel regulatory principle. This discovery finalizes a long-term effort to document and understand the relationship between chromosomal instability and phenotypic diversity in *C. albicans* (see, for example, ref. 2 for background information).

In this study, we have demonstrated that normal parental *SOU1/SOU1* strains are Sou^- and cannot grow on L-sorbose medium because of insufficient residual expression of the *SOU1* genes, which apparently encode L-sorbose dehydrogenase (Fig. 5). However, the *SOU1/SOU1* strains spontaneously give rise to high frequencies of Sou^+ mutants (3), which acquire the ability to use L-sorbose, despite the fact that the *SOU1* genes in the Sou^- and Sou^+ strains have the same ORF and upstream-region sequences. In contrast to the normal *SOU1/SOU1* strains, the doubly deleted *sou1-Δ/sou1-Δ* strains are unable to produce Sou^+ mutants (see Fig. 3).

We have established that the phenotypic $Sou^- \rightarrow Sou^+$ change is due to the loss of either the a or b homologue of chromosome III, whereas subsequent Sou^- mutants unable to grow on L-sorbose medium are formed by duplication of the remaining homologue, as illustrated in Fig. 4. Northern blot analysis corroborated this finding on the expression of the *SOU1* gene and established that these phenotypes are determined at the transcriptional level (Fig. 5). In addition, normal *SOU1/SOU1* strains become Sou^+ by artificially introducing multiple copies of *SOU1* plasmids, which enabled us to clone the gene. We suggest that the level of transcription depends on the ratio of gene copies of *SOU1* and a hypothetical negative regulator gene *CSUI* (Control of Sorbose Utilization), which resides on chromosome III, as schematically presented in Fig. 4D. The same ORF and upstream-region sequences of the *SOU1* gene in both Sou^- and Sou^+ strains are consistent with this view and exclude such mechanisms involving *cis*-acting mutations that are found, for example, in the activation of cryptic genes of prokaryotes.

Approximately 75% of the analyzed Sou^+ mutants arose by the formation of chromosome III monosomy. Although this was a

major mechanism, the remaining mutants had different types of chromosomal alterations, possibly representing alternative pathways for regulating *SOU1*.

Another example of acquiring the ability to assimilate a particular sugar by changes in chromosome number are the previously reported D-arabinose positive mutants (3). Although these Ara^+ mutants were not rigorously analyzed, the specificity of the chromosomal alterations suggests that these mutants may be arising by a means similar to that of the Sou^+ mutants. A recent study revealed that mutants selected on fluconazole plates similarly had specific alterations in chromosome number (E.R., personal communication). In addition, a fluconazole resistant strain of a related species, *Candida glabrata*, isolated from a patient, had a correlation between the level of transcription of a gene conferring resistance and the amplification of the chromosome carrying this gene (35).

Considering the regulation of sugar assimilation by chromosomal changes discussed above, as well as fluconazole resistance, we suggest that chromosome number may be a common means to control important functions in *C. albicans*, thus establishing our early hypothesis (1) that chromosomal instability is a means for generating genetic variability.

We thank William A. Fonzi (Georgetown University), Richard D. Cannon (University of Otago), and Joachim F. Ernst (Heinrich-Heine Universität) for generously providing *C. albicans* strains and plasmids. This work was supported by U.S. Public Health Science Research Grant AI22963 from the National Institutes of Health.

- Rustchenko-Bulgac, E. P., Sherman, F. & Hicks, J. B. (1990) *J. Bacteriol.* **172**, 1276–1283.
- Rustchenko, E. P., Howard, D. H. & Sherman, F. (1997) *Microbiology* **143**, 1765–1778.
- Rustchenko-Bulgac, E. P., Howard, D. H. & Sherman, F. (1994) *J. Bacteriol.* **176**, 3221–3241.
- Pla, J., Gil, C., Monteoliva, L., Navarro-Garcia, F., Sánchez, M. & Nombella, C. (1996) *Yeast* **12**, 1677–1702.
- Magee, B. B., Koltin, Y., Gorman, J. A. & Magee, P. T. (1988) *Mol. Cell. Biol.* **8**, 4721–4726.
- Wicks, B., Staudinger, J., Magee, B. B., Kwon-Chung, K.-J., Magee, P. T. & Scherer, S. (1991) *Infect. Immun.* **59**, 2480–2484.
- Fonzi, W. A. & Irwin, M. Y. (1993) *Genetics* **134**, 717–728.
- Gough, J. & Murray, N. (1983) *J. Mol. Biol.* **166**, 1–19.
- Cannon, R. D., Jenkinson, H. F. & Shepherd, M. G. (1992) *Mol. Gen. Genet.* **235**, 453–457.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–119.
- Sherman, F. (1991) *Methods Enzymol.* **184**, 3–21.
- Rustchenko-Bulgac, E. P. (1991) *J. Bacteriol.* **173**, 6586–6596.
- Goshorn, A. K., Grindle, S. M. & Sherer, S. (1992) *Infect. Immun.* **60**, 876–884.
- Janbon, G., Magnet, R., Bigey, F., Arnaud, A. & Galzy, P. (1995) *Curr. Genet.* **28**, 150–154.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Kurtz, M. B., Cortelyou, M. W. & Kirsch, D. R. (1987) *Mol. Cell. Biol.* **7**, 209–217.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
- Guo, Z. & Sherman, F. (1996) *Mol. Cell. Biol.* **16**, 2772–2776.
- Carle, G. F. & Olson, M. V. (1984) *Nucleic Acids Res.* **12**, 5647–5664.
- Rustchenko-Bulgac, E. P. & Howard, E. P. (1993) *J. Gen. Microbiol.* **139**, 1195–1207.
- Persson, B., Krook, M. & Jörnvall, H. (1991) *Eur. J. Biochem.* **200**, 537–543.
- Wierenga, R. K., Maeyer, M. C. H. & Hol, W. G. J. (1985) *Biochemistry* **24**, 1346–1357.
- Wong, B., Murray, J. S., Castellanos, M. & Croen, K. D. (1993) *J. Bacteriol.* **175**, 6314–6320.
- Murray, J. S., Wong, M. L., Miyada, C. G., Switchenko, A. C., Goodman, T. C. & Wong, B. (1995) *Gene* **155**, 123–128.
- Santos, M. A. & Tuite, M. F. (1995) *Nucleic Acids Res.* **23**, 1481–1486.
- Miyasaki, S. H., White, T. C. & Agabian, N. (1994) *J. Bacteriol.* **176**, 1702–1710.
- Otha, T. (1993) *Genetics* **134**, 1271–1276.
- Otha, T. (1994) *Genetics* **138**, 1331–1337.
- Nelson, H., Mandiyan, S. & Nelson, N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 497–501.
- Rustchenko, E. P., Curran, T. M. & Sherman, F. (1993) *J. Bacteriol.* **175**, 7189–7199.
- Jiang, Y. W. & Stillman, D. J. (1996) *Genes Dev.* **10**, 604–619.
- Johnston, M. & Carlson, M. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, eds Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 193–281.
- Hall, B. G. & Betts, P. W. (1987) *Genetics* **115**, 431–439.
- Marichal, P., Vanden Bossche, H., Odds, F. C., Nobels, G., Warnock, D. W., Timmerman, V., Van Broeckhoven, C., Fay, S. & Mose-Larsen, P. (1997) *Antimicrob. Agents Chemother.* **41**, 2229–2237.