Analysis of the chromosomal localization of the repetitive sequences (RPSs) in Candida albicans

Ariya Chindamporn, Yoshiyuki Nakagawa, Michio Homma, Hiroji Chibana, Matsuko Doi and Kenji Tanaka

Author for correspondence: Y. Nakagawa. Tel: +81 52 741 2111 ext. 2115. Fax: +81 52 731 9479. e-mail: e43187a@nucc.cc.nagoya-u.ac.jp

Laboratory of Medical Mycology, Research Institute for Disease Mechanism and Control, Nagoya University School of Medicine, Showa-ku, Nagoya 466, Japan

The location and organization of repetitive sequences, members of the RPS family, which are sequences specific to Candida albicans, were determined on each chromosome of C. albicans strain FC18. Using pulsed-field gel electrophoresis, we separated seven fractions from eight chromosomes. Each chromosome was cleaved by BamHI and XhoI to excise the RPSs, which were then detected by hybridization with an RPS probe. All chromosomes except chromosome 4 carried RPSs, and these RPSs were located within a limited region on each chromosome. From the digestion of each chromosome with Sfil and probing with the RPSs, we found that these recognition sites within the RPS region were conserved among all RPS-containing chromosomes. For further characterization of the RPSs, the locations and the boundary regions of the RPSs were examined on chromosome 6 of strain FC18 as a model chromosome. Using the restriction enzymes Sfil, Smal, XhoI, BamHI, MluI and Nrul, we constructed a semi-macro physical map of the RPSs and their boundary regions on this chromosome. We also determined which part of the RPS was adjacent to each boundary by using sub-fragments of RPS as probes. The physical configuration around the RPSs and their boundary regions are presented. The results obtained should be useful for future analysis of the function of these regions.

Keywords: repetitive sequence, RPS, physical map, Candida albicans

INTRODUCTION

The recent development of techniques, such as pulsed-field gel electrophoresis (PFGE; Schwartz & Cantor, 1984), for manipulating high-molecular-mass DNA has revealed that hyper-variability in chromosomal features is one of the most distinctive characteristics of Candida albicans (Lott et al., 1987; Magee & Magee, 1987; Iwaguchi et al., 1990; Doi et al., 1992). In some cases, changes in karyotype are known to be associated with a variation in some aspect of the phenotype, such as colony morphology or cell shape, which can be correlated with the expression of pathogenicity in this yeast (Suzuki et al., 1989; Rustchenko-Bulgac et al., 1990; Barton & Scherer, 1994). Repeated sequences in C. albicans that might be associated with chromosomal variations or translocations have been isolated in several laboratories (Scherer & Stevens, 1988; Lasker et al., 1989, 1991; Sadhu et al., 1991; Iwaguchi et al., 1992a; Thrash-Bingham & Gorman, 1993). Some of these sequences are known to be species-specific and can serve as useful probes for the epidemiological analysis of clinical isolates (Scherer & Stevens, 1988; Anderson et al., 1993). However, little is known about the precise configuration of these repetitive sequences in the genome of this yeast or about their possible functional significance.

The most variable, in terms of size, of the eight chromosomes of C. albicans was reported previously to be chromosome 2, followed by chromosome 6. A change in the number of repeating units of ribosomal DNA is responsible for the variability in length of chromosome 2 (Iwaguchi et al., 1992b). Another repetitive sequence, RPS1, was cloned from genomic DNA of C. albicans strain
NUM812 and was characterized. It hybridized to chromosome 6 of strain FC18 and also to other chromosomes (Iwaguchi et al., 1992a). Sequences homologous to the RPS1 sequence from different chromosomes of strain FC18 have been designated members of the RPS family, including RPS1 (Chibana et al., 1994). Sequence analysis revealed that members of the RPS family are composed of a few sequence repeats (172 bp) known as alt, which contain shorter, common stretches known as COM29 (29 bp; Iwaguchi et al., 1992a; Chibana et al., 1994). This structural hierarchy of the repeating sequences is very similar to the structure of sequences believed to be involved in centromeric functions in humans (Willard & Wave, 1987). Thus, characterization of the chromosomal locations of RPSs is a prerequisite for studies of the function of RPSs.

In this paper, we demonstrate that RPSs are located on every chromosome, except chromosome 4, of C. albicans and are clustered in a limited region on each chromosome. We also provide a detailed analysis of chromosome 6, which is the second most variable in size of this yeast's chromosomes. Analysis of the SfiI digestion profile revealed that RPSs are clustered in the middle region of chromosome 6. To clarify the precise position of the RPSs, a semi-macro physical map of RPSs and their boundary regions on chromosome 6 was constructed.

METHODS

Strains and plasmids. C. albicans strains FC18, NUM46, NUM55 and NUM1039 were used (Iwaguchi et al., 1990, 1992a). The homologue of chromosome 6 from each of the three NUM strains was isolated (see below). The plasmids used as sources of probe DNAs for Southern hybridization were as follows: pSI3-12 containing an RPS1 from C. albicans strain NUM812 (Iwaguchi et al., 1992a); pRPS101, pRPS116 and pRPS620 containing RPS101, RPS116 from chromosome 1, 4 and 2 and RPS620 from chromosome 6 of strain FC18, respectively (Chibana et al., 1994); pMM100 containing a telomeric sequence of Ca7 (Sadhu et al., 1991; McEachern & Hicks, 1993); and pTK2-9-1 containing the chromosome 6 marker of C. albicans (Magee et al., 1988; Iwaguchi et al., 1990, 1992b).

Preparation of chromosomal DNA for PFGE. The sample plug containing yeast chromosomal DNA for PFGE was prepared by the method described previously (Iwaguchi et al., 1990). PFGE was carried out by the CHEF method using the Pulsaphor system with a hexagonal electrode array (Pharmacia-LKB) as described previously (Iwaguchi et al., 1992a). Chromosomes of Saccharomyces cerevisiae (X2180-1A), λ phage ladder concatemeric DNA (Bio-Rad) and λ phage DNA digested by EcoT14I (Takara Shuzo) were used as molecular-size markers.

Southern hybridization. Southern hybridization was performed by using a probe prepared from an EcoRI fragment of pSI3-12 or a PstI fragment of pRPS101 for detection of RPSs and an EcoRI fragment of pTK2-9-1 for assignment of chromosome 6 (Iwaguchi et al., 1990, 1992a; Chibana et al., 1994). The probes were prepared by random primer labelling with [32P]dCTP and hybridization signals were detected by autoradiography on X-ray film or by an image analyser (BAS 2000; Fuji Photo Film).

Restriction digestion of total chromosomal DNA in sample plugs for PFGE. This was performed as described previously (Chindamporn et al., 1993).

RESULTS

Localization of RPSs on individual chromosomes

Since no recognition sites for XhoI, BamHI and SalI were found within RPSs, these enzymes were suitable for separation of RPS-containing regions from each chromosome. The pattern of digestion by XhoI or BamHI was revealed by PFGE and subsequent Southern hybridization with probe pSI3-12 (Fig. 1a, b). The digestion profiles obtained with both these enzymes from whole chromosomes (sample plugs prepared for PFGE) revealed 10–15 bands after hybridization with an RPS probe (Fig. 1a, b, lane 'Total'). This hybridization profile contained bands which corresponded to the hybridization bands derived from the different individual chromosomes (lanes 1–2 to 8). For example, in the case of digestion with BamHI (Fig. 1a), the bands of 50, 102, and 120 plus 102 kbp were derived from chromosomes 2, 3 and 6, respectively. Between one and four bands from individual chromosomes hybridized to the RPS probe. If we assume that the FC18 strain carries two homologues for every chromosome, one band would be obtained from each chromosome if the cleavage sites for a restriction enzyme that flank the RPS-containing region are conserved between the two homologues. However, when two bands are detected there are two possible locations of the RPSs: the RPSs might be located at two different loci on a chromosome, or they might be located at a single locus on a chromosome but the cleavage sites for the restriction enzymes used are not in the same position. To determine whether or not there is only one RPS(s) locus on a chromosome, we used strains in which the two homologues of chromosome 6 are separable by PFGE. After separation and isolation of chromosome 6 following gel electrophoresis, the samples were digested with XhoI and

BamHI and XhoI digestion of individual chromosomes. Each chromosome of C. albicans strain FC18 was separated from total chromosomal DNA (sample plug) on 0.7 % (w/v) gels by PFGE under conditions of 100 s at 170 V for 15 h, 300 s at 130 V for 24 h then 1000 s at 80 V for 48 h. After staining the gels with ethidium bromide and destaining with distilled water, individual bands corresponding to each chromosome were cut out from the gel under UV light (365 nm). To perform endonuclease digestion, the ethidium bromide was removed from the gel slices (about 2 mm3 in size) by three 10 min washes with ice-butanol saturated with 1 M NaCl, 10 mM Tris/HCl (pH 8.0) and 1 mM EDTA. The samples were then equilibrated in 200 μl of the appropriate restriction endonuclease buffer for 30 min, transferred into 200 μl fresh buffer with either 10 U BamHI or 10 U XhoI (Takara Shuzo) and incubated overnight at 37 °C. The digested products were separated by subsequent electrophoresis (see legend of Fig. 1) prior to hybridization analysis using RPS1 probes.
Localization of repetitive sequences in C. albicans

(a) Oxa 1 +2 3 4 5 6 7 8
kbP
145.5 -
97 -
48.5 -
19.3 -

(b) Fig. 2.
The RPS hybridization profile of XhoI fragments of chromosome 6 from various C. albicans strains. Chromosomal DNA bands corresponding to chromosome 6 of each strain used were prepared and their identity confirmed by probe pTK 2-9-1 (Iwaguchi et al., 1990; A. Chindamporn, unpublished results). The DNA in gel slices containing the individual bands was digested with XhoI (see Methods). After treatment with this enzyme, samples were subjected to PFGE with ramping conditions of 5–20 s at 200 V for 20 h (1.25%, w/v, agarose) and analysed by Southern hybridization using the 32P-labelled RPS1 as a probe. Lanes: 1 (FC18), the homologues of FC18; 2 (46-6a), the large homologue of NUM46; 3 (55-6a), the large homologue of NUM55; 4 (55-6b), the small homologue of NUM55; 5 (1039-6a), the large homologue of NUM1039. The numbers on the left side are molecular sizes calculated from mobilities of chromosomes of S. cerevisiae and λ phage concatemers. Data shown in the figure were from one set of experiments. Similar results were obtained in separate experiments.

Fig. 1. Distribution of RPSs on the chromosomes of C. albicans FC18. Each of the chromosomes separated by PFGE and the total genomic DNA (see Methods) were digested with BamHI (a) or XhoI (b), and subjected to PFGE followed by Southern hybridization with a 32P-labelled RPS1 as an RPS probe. The digested samples were separated by PFGE under the following conditions: linear ramping of 2.5–12.5 s at 200 V for 20 h on 1.2% (w/v) agarose. Molecular sizes were estimated from the mobilities of λ phage concatemeric DNA (Bio-Rad) that had been labelled with 32P. The leftmost lane (Total) indicates the digested total genomic DNA, and the other lanes indicate the digested individual chromosomes. The numbers at the top of the lanes represent the chromosome numbers. Chromosomes 1 and 2 were not clearly separated under the conditions used here. Data shown in the figure were from one set of experiments. Similar results were obtained in separate experiments.

the resulting fragments separated by PFGE and analysed by Southern hybridization. As shown in Fig. 2, only one hybridization band was detected from each separated homologue, although the mobility of the band varied from chromosome to chromosome with different strains. Similar results were obtained when MspI or NruI was used as the restriction enzyme (data not shown). In the case of chromosome 8, which yielded four bands after digestion with XhoI or BamHI (Fig. 1a, b; lane 8), a site susceptible to cleavage by these enzymes might exist within the RPS region, or the RPS-containing regions on the two homologues might be of two different lengths. However, when SalI was used instead of these two enzymes, only two RPS-positive bands were detected (data not shown). This result indicates that the RPS region on chromosome 8 might contain BamHI and XhoI recognition sites and, as a result, four bands were detected. Thus, the RPS region on this chromosome was also shown to be located in a limited region. Another interesting finding was the failure to detect a hybridization signal with the RPS probe on chromosome 4 (Fig. 1a, b, lane 4), even after both longer exposure of hybridization filters to X-ray film and low-stringency conditions of hybridization were performed (data not shown). Nevertheless, the possible presence of very short stretches homologous to the RPS cannot be excluded.
Fig. 3. Detection of small-sized fragments after SfiI treatment of each chromosome. After separation by PFGE, each chromosome in the 0.7% (w/v) agarose gel was isolated and transferred to 0.4% (w/v) agarose with subsequent treatment with SfiI (see Methods). The digested samples were fractionated by electrophoresis and subjected to Southern hybridization using a mixture of pRPS101, pRPS116 and pRPS620 (Chibana et al., 1994) as the source of the probe for RPSs. EcoT14I-treated λ phage and HindIII-treated bacteriophage ϕX174 DNA were used as molecular size markers, as shown on the left. The leftmost lane (Total) represents SfiI-digested total genomic DNA, and the respective lanes indicate digested individual chromosomes. The numbers at the top of these lanes represent the chromosome numbers (Iwaguchi et al., 1990). Chromosomes 1 and 2 were not clearly separated under the conditions used here. Data shown in the figure were from one set of experiments. Similar results were obtained in separate experiments.

From our characterization of the RPS(s), each RPS unit was shown to have several SfiI sites in all clones derived from chromosomes 1+2 and 6 (Chibana et al., 1994). In order to obtain information about whether or not the RPSs on the other chromosomes also contain such recognition sites, each chromosome was digested by this enzyme. As shown in Fig. 3, all chromosomes apart from chromosome 4 gave signals that corresponded to about 0.34, 0.68 and 1.1 kbp, namely, the sizes expected from the digestion by SfiI of the RPS, although the intensity of bands was not uniform. In addition the fragment of about 1.8 kbp in size, which was not identified from our RPS clones, was commonly detected from total genomic DNA and fractionated chromosomes except chromosome 4. The other bands observed in fractionated chromosomes were likely to be a result of partial digestion with SfiI. Moreover, a 170 bp fragment, corresponding to the size of some RPSs (Chibana et al., 1994), was detected in the analysis of chromosomes 1+2 and 8 following a longer exposure time after hybridization. This fragment was also detected in the analysis of total genomic DNA, although the signals were very weak. Thus, our results strongly suggest that all of the chromosomes except chromosome 4 include RPSs that are present within limited regions on the respective chromosomes. However, further investigation of the possible presence of RPS on chromosome 4 is necessary.

Analysis of the boundary regions of RPSs

To elucidate the biological functions of RPS, we constructed a plasmid that contained RPS620 cloned in the C. albicans vector pMK22 containing the Candida autonomously replicating sequence (CARS) and the UR A3 gene (Kurtz et al., 1987). Using this plasmid, we transformed Ura- C. albicans strains such as 1006 and SGY243 to yield Ura+ transformants. However, no transformants which contained a stable episome were obtained. Therefore, we assumed that not only single RPSs but also the boundary regions that flank the RPSs are prerequisite for the expression of some biological functions. In this way, we recognized the necessity of examining the precise structure of the boundary regions of RPS. Since the hybridization signals derived from RPSs were the most intense for chromosome 6, due to the presence of a large number of copies of the repeat in a limited region (Iwaguchi et al., 1992a) and also as only two SfiI fragments were detected (see below), we focused our attention on this chromosome. When chromosome 6 (1400 kbp) was isolated by PFGE, digested by SfiI and the resulting fragments separated by PFGE followed by Southern hybridization for detection of RPS signals, only bands of 760 and 510 kbp and a band of the intact chromosome were obtained (Fig. 4). These three RPS-hybridizing bands also
included telomeric sequences, as shown by hybridization with the pMM100 probe (Fig. 4, lane 3). This result indicated that no Sfl site was present in regions other than in the RPS-containing region, and that the RPS region was located in the middle of the chromosome or was flanked by the two fragments of 760 and 510 kbp. Furthermore, the pTK2-9-1 probe, which was used as a marker probe for chromosome 6 (Iwaguchi et al., 1990), hybridized with the 760 kbp fragment, and the snc2 and snc3 probes (B. Magee, personal communication) hybridized to the 510 kbp fragment (data not shown). These results correspond to the results of experiments with chromosome 5 of the strain 1006 reported by Chu et al. (1993), where chromosome 5 of strain 1006 and chromosome 6 of FC18 hybridized to the same chromosome marker.

Since each of the restriction enzymes Sfl, PstI, and EcoRI have only one recognition site in the RPS sequence, we were able to estimate the size of the flanking sequences of the RPS when chromosome 6 was digested with one of these enzymes. Digestion of chromosome 6 from strain FC18 with Sfl yielded three bands that hybridized with the RPS1 probe after PFGE. The fragments were estimated to be about 130, 61 and 2.1 kbp in size (Fig. 5b). Similar-sized fragments were also obtained from the homologues of chromosome 6 from other strains (data not shown). The signal of 2.1 kbp represented the RPS units themselves. The bands of 130 and 61 kbp were assumed to be fragments derived from either one of the boundary regions that flanked the RPS: we postulated that the 130 kbp band fragment contained one end of the RPS and the 61 kbp fragment contained the other end. To examine the validity of this assumption, a PstI fragment from pRPS101 (Chibana et al., 1994) was divided into three small subunits, A (PstI–Sfl), B (Sfl–ClaI), and C (ClaI–PstI), as shown schematically in Fig. 6(c), and then each was used as a probe for hybridization with the Smal fragments of chromosome 6 (Fig. 5a). All three probes hybridized to the RPS units of 2.1 kbp. Probe A hybridized to both the 130 kbp and 61 kbp fragments. Probe B hybridized to the 130 kbp but not to the 61 kbp fragment, while probe C did not hybridize to the 130 kbp fragment but did to the 61 kbp fragment. Assuming that RPS units are tandemly repeated in a limited region of each homologue of chromosome 6 (Iwaguchi et al., 1990), we deduced that the RPS region was flanked by Smal fragments of 130 kbp and 61 kbp. To obtain more details about the physical maps of the RPS boundary regions, double digestion with Sfl and XhoI, BamHI, MluI or NruI was performed (Fig. 5b). Double digestion with Sfl and XhoI gave a 66 kbp band instead of a 130 kbp band while the other signals at 61 and 2.1 kbp were unchanged. This result indicates that an XhoI site was located within the 130 kbp Smal fragment. In the same way, double digestions with Sfl and BamHI, MluI or NruI produced RPS-positive bands of 27, 62 and 78 kbp, respectively, instead of the 130 kbp band. These digestions did not affect the signals from the 61 kbp and 2.1 kbp bands, indicating that no recognition sites for these enzymes were present in the 61 kbp and 2.1 kbp fragments. From these results, a map of the entire structure of chromosome 6 was drawn, as shown in Fig. 6.
DISCUSSION

It has been estimated that the RPSs are present at a minimum of 60 copies in the diploid genome of *C. albicans* (Iwaguchi et al., 1992a). We showed here that RPSs are distributed in limited regions of all except one chromosome (Fig. 1). We could assign all the signals from the total chromosomal DNA (Fig. 1, lane Total) to those from each chromosome separately (Fig. 1, lanes 1+2 to 8). Although the numbers of copies seemed to differ among the chromosomes, as indicated by the different intensities of the various signals, RPSs were located in clusters and were not dispersed on chromosomes with multiple copies of RPS (Fig. 2). Furthermore, by digestion of each chromosome with *SfiI*, we were able to detect small-sized fragments derived from RPSs (Fig. 3). These fragments corresponded to those of the already cloned RPSs (Chibana et al., 1994). This result demonstrates the presence of conserved *SfiI* fragments in the RPS region, suggesting that *SfiI* recognition sites in the RPS region are conserved. As discussed below, if the RPS is related to some biological function, acting for example as a centromere, all the chromosomes would have to carry the RPS regardless of the location on the chromosome. However, we failed to obtain any RPS-specific signals from chromosome 4 under various experimental conditions, whereas the other chromosomes always gave significant signals. Nonetheless, it remains unknown whether chromosome 4 lacks a portion of the RPS.

Recently, Chu et al. (1993) constructed a macro-restriction map of the chromosomes of *C. albicans* by using a rare-cutting enzyme, *SfiI*, and Southern hybridization with specific gene probes for two strains, 1006 and WO-1. They found that some of the *SfiI* chromosomal fragments had been translocated to different chromosomes and they suggested that *SfiI* sites might be the break points responsible for translocation of the fragments. It is noteworthy that the RPS units contained several sites for *SfiI*. In fact, we detected small *SfiI* fragments derived from the RPS of each chromosome except for chromosome 4 (Fig. 3). Hence, chromosome 6, which contains at least 10 tandemly repeated RPS units (Iwaguchi et al., 1992a), should have more than 30 *SfiI* sites, which could be junction sites for translocation. The finding that these sites were clustered in a limited region(s) on almost all
Localization of repetitive sequences in C. albicans

Among the repeated sequences in C. albicans reported to date, some features associated with Rel-2 are very similar to those of RPS (Thrash-Bingham & Gorman, 1993). The Rel-2 sequence is about 2.7 kbp in length, it hybridizes to every chromosome, and it includes subrepeat sequences of 29 bp. RPS varies in size from 19 to 27 kbp in increments of 0.2 kbp, the RPS probe hybridizes to almost every chromosome, and it includes an inner repeated sequence, COM29, of 29 bp. Additionally, both probes hybridize not only to genomes of all strains of C. albicans tested but also to that of Candida stellatoidea (Thrash-Bingham & Gorman, 1993; A. Chindamporn, data not shown) although the intensity of the signals in the case of C. stellatoidea was weaker than that for C. albicans. However, no sequence homology was detected between the two sequences (Chibana et al., 1994). The structural similarity suggests, however, that some function might be partly shared by both sequences despite differences in nucleotide sequence.

Since the RPSs are located in limited regions of all chromosomes apart from chromosome 4 (Fig. 1), they might have some important role in chromosome function. Recently, we found that the RPS includes many repeats of the short sequence, alt, and we suggested a possible centromeric function by analogy to the structural features of human alpha-satellite DNA (Chibana et al., 1994). Most of the physically characterized centromeres include inner repetitive sequences in their higher-order units (Willard & Waye, 1987; Clarke, 1990; Takahashi et al., 1992; Vogt, 1992; Centola & Carbon, 1994). If the RPS functions as a centromere, a plasmid that includes autonomously replicating sequence (ARS) for C. albicans, a selectable marker such as Ura', and the RPS, would be expected to be maintained in Ura- cells as a stable episome. However, as mentioned in Results, no cells transformed with an RPS-carrying plasmid were stable during mitosis. Therefore, we postulated that not only the RPS sequence but also the boundaries of the RPS might be required for expression of some important role in cellular function. To examine this possibility, we physically mapped an extended region that included the RPSs on chromosome 6. Chromosome 6 could be used as a model chromosome that carries only one set of boundary regions of the RPSs since only two Sfl fragments were detected. After digestion with SmaI three fragments that hybridized with an RPS probe were obtained: that of 21 kbp represented the RPS units and the other two, of 130 kbp and 61 kbp, were inferred to be the junction fragments and to contain RPS fragments at their respective ends. The 130 kbp fragment was derived from the 510 kbp fragment generated by digestion of chromosome 6 by SfiI, and the 61 kbp fragment was derived from the 760 kbp SfiI fragment, as shown by SmaI digestion followed by hybridization with an RPS probe of the respective SfiI fragments (data not shown). To determine which part of the RPS was adjacent to the 61 kbp and the 130 kbp fragments, hybridizations using sub-RPS probes were performed (Fig. 5a). From our results, we could demonstrate a characteristic structure of the boundary region of RPS units which were tandemly arrayed on this chromosome.

ACKNOWLEDGEMENTS

The authors thank Dr B. B. Magee of the University of Minnesota for kindly providing the sncl and sncl3 probes, and Dr M. J. McEachern of the University of California, San Francisco, for the Ca7 probe DNA. The authors also thank Ms. I. Mizuguchi for expert photographic assistance. This work was supported by grants-in-aid for scientific research from the Ministry of Education, Science and Culture, Japan. A. C. is the recipient of a scholarship from the Hitachi Foundation.

REFERENCES


DNA by pulsed-field gel electrophoresis techniques. *Nucleic Acids Res* 17, 3783–3793.


Received 1 August 1994; revised 22 September 1994; accepted 19 October 1994.