

Analysis of sequences conferring autonomous replication in baker's yeast

Stephen Kearsey

Laboratory of Molecular Biology, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, UK

Communicated by R.A. Laskey

Received on 9 June 1983

A method is presented for rapid sequencing and mapping of elements which support autonomous replication in yeast. The strategy relies on a novel phage M13 vector which allows detection of ARS (autonomously replicating sequence) function in cloned fragments. Deletion mapping of an ARS element linked to the *HO* gene of *Saccharomyces cerevisiae* has identified a 57-bp region 3' to the gene, which is essential for autonomous replication. This region shows sequence homology to other ARS elements.

Key words: ARS elements/*HO* gene/origins of replication/phage M13 vector/*Saccharomyces cerevisiae*

Introduction

Replication of circular DNA molecules in yeast appears to need the presence of a special sequence, called an ARS (for autonomously replicating sequence, Stinchcomb *et al.*, 1979). ARS elements from the yeast genome are good candidates for chromosomal origins of replication, although critical evidence is lacking. Indirect evidence supporting this hypothesis is provided by the approximate correlation between the number of ARS sequences in the genome of *Saccharomyces cerevisiae* and the number of initiations of DNA replication during chromosomal replication (Beach *et al.*, 1980; Chan and Tye, 1980). Celniker and Campbell (1982), studying DNA replication *in vitro* using an extract made from yeast cells, have reported that an ARS increases the efficiency of replication of a DNA fragment and that replication appears to commence in the vicinity of the ARS, as judged by electron microscopic mapping and label incorporation. More recently, Fangman *et al.* (1983) have shown that different ARS elements in the genome are replicated at specific times during the yeast S phase, and that plasmid-borne ARS1 is replicated at the same time as its chromosomal counterpart.

Other formal possibilities for the function of ARS elements can be envisaged. ARS elements may be required *in cis* for the elongation or termination steps of DNA synthesis, or merely to ensure compartmentalization of plasmid DNA in the nucleus. Alternatively, ARS elements may affect segregation of replicated plasmids to daughter cells during mitosis.

The mechanism of ARS function has become particularly intriguing with results suggesting that the transcriptional control of certain yeast genes involves sites closely linked to ARS elements. Transcriptional repression of the silent mating-type cassettes is abolished by deletion of flanking regions of DNA which contain ARS elements (Abraham *et al.*, 1982; Broach *et al.*, 1982). Furthermore, the cell cycle control of histone 2A gene transcription may involve an ARS located at the 3' end of the closely linked H2B gene (Osley and Hereford, 1982). If a mechanism of transcriptional control involving ARS elements does exist, it seems likely that such a system would

regulate many genes. One good candidate for such ARS control is the *HO* gene, as suggested by Nasmyth (1983), which encodes or regulates an endonuclease responsible for the initiation of mating-type conversion, and whose expression shows a pattern of regulation which is both cell type- and cell cycle-dependent (Nasmyth, 1983). In particular, the cell cycle variation of amounts of *HO* RNA in mother cells is similar to the pattern seen with histone RNA, perhaps reflecting a similar mechanism of control.

Convincing demonstration that ARS elements are implicated in transcriptional control will clearly require precise mapping of the sequences which are functionally required for autonomous replication. Appropriate mutations can then be made to test their effects on autonomous replication, and effects on transcription due to ARS elements and those due to closely linked genetic elements can be distinguished. I describe here the deletion analysis of an ARS element closely linked to the *HO* gene. The method uses a novel M13 vector which should be generally useful for mapping and sequencing elements which confer autonomous replication.

Results

Construction of a vector for yeast transformation from phage M13

Phage M13 vectors allow rapid sequencing of cloned DNA using the dideoxy technique (Sanger *et al.*, 1980). Since relatively short sequences can allow autonomous replication, an appropriately constructed phage M13 shuttle vector would facilitate their analysis. For this purpose, the selectable *URA3* gene, which does not contain an ARS, was inserted into the *HindIII* site of phage M13mp9 RF DNA, producing a phage designated M13se102 (Figure 1). Transformation of a *ura3* yeast strain with M13se102 DNA was very inefficient (<1 transformant/ μ g DNA) and, as expected, transformants were stable for the Ura⁺ phenotype, consistent with integration of the transforming DNA. Insertion of an ARS-containing restriction fragment into the *BamHI* site of M13se102 DNA increased its transformation efficiency by over four orders of magnitude, using either single-stranded or RF DNA (see Singh *et al.*, 1982). The transformants were unstable for the Ura⁺ phenotype in the absence of selection and the original DNA could be recovered by transfecting *Escherichia coli* with total DNA made from a transformant, as expected for an autonomously replicating DNA molecule.

Mapping and sequencing of an ARS element linked to the HO gene of S. cerevisiae

The strategy used for mapping ARS elements using phage M13se102 involves fragmentation of an ARS-containing plasmid by sonication, and insertion of the randomly produced fragments into the M13 vector, using a procedure developed by Deininger (1983). DNA from the pooled hybrid phage is used to transform a *ura3* strain of yeast. Ura⁺ transformants mainly arise from clones whose yeast DNA inserts contain an ARS. These inserts can be readily sequenced after rescue of the phage clone by transfection of *E. coli* with total yeast DNA. The sequences obtained should share regions of

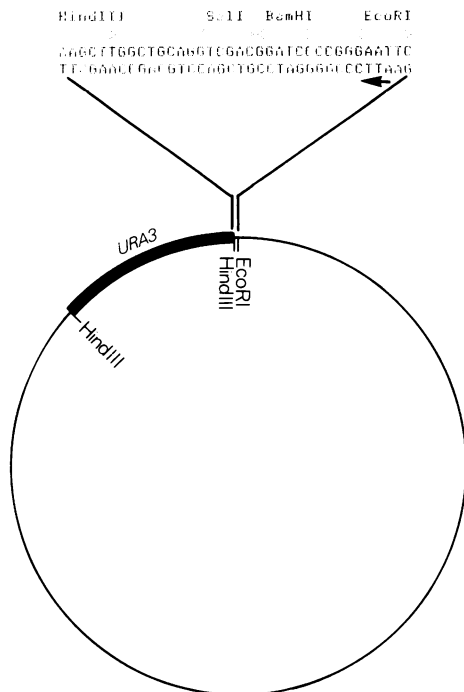


Fig. 1. Structure of M13se102. The 1.1-kb *URA3* *Hind*III restriction fragment from Y1p9-Sc2717 was inserted into the *Hind*III site of M13mp9 (Messing, 1981). The construction of the vector allows fragments inserted into the *Sac*II, *Bam*HI and *Eco*RI sites to be sequenced using the usual M13 sequencing primer (direction of sequencing is indicated), and since the vector alone does not autonomously replicate in yeast, *ARS* function in cloned inserts can be detected. The *Hinc*II, *Pst*I and *Sma*I sites in the sequence between the *Hind*III and *Eco*RI sites also occur in the *URA3* fragment, and hence cannot be used for cloning.

DNA containing *ARS* elements, and clearly the precision of the mapping is dependent on the sizes of the DNA inserts and the number compared.

The plasmid E347, consisting of a 2.5-kb *Hind*III restriction fragment of *S. cerevisiae* DNA cloned in the Y1p5 vector, contains the *HO* gene and at least one *ARS* element as judged by its ability to transform *ho* mutants and replicate as an extrachromosomal plasmid (K.A.Nasmyth and R.Jensen, personal communication). The E347 plasmid DNA was sonicated to give small fragments with sizes ranging from ~100 to 500 bp. These fragments were repaired with T4 DNA polymerase and ligated into the *Bam*HI site of M13se102 DNA. The ligated DNA was used to transfect *E. coli*, and the phage pool obtained was propagated to prepare sufficient DNA for yeast transformation. The ligation mix can be used to transform yeast directly, using carrier DNA to increase the transformation efficiency, but the amplification step in *E. coli* serves to select for hybrid phage containing small inserts, and this simplifies the subsequent sequence analysis. Transformation of a *ura3-52* yeast strain with 5 μ g single-stranded phage DNA produced 100–200 transformants, most of which lost their *Ura*⁺ phenotype during growth without selection, suggesting autonomous replication of the transforming DNA. Yeast DNA was prepared from a number of transformants and M13 clone DNA was rescued by transfection of *E. coli*. Only *ARS*-containing M13 clones can be easily rescued in this manner; M13 DNA integrated into yeast DNA cannot transfect *E. coli*.

Single-stranded phage DNA was prepared from the rescued M13 clones and partially sequenced using the dideoxy C reaction. The C reaction was chosen as it gives an easily

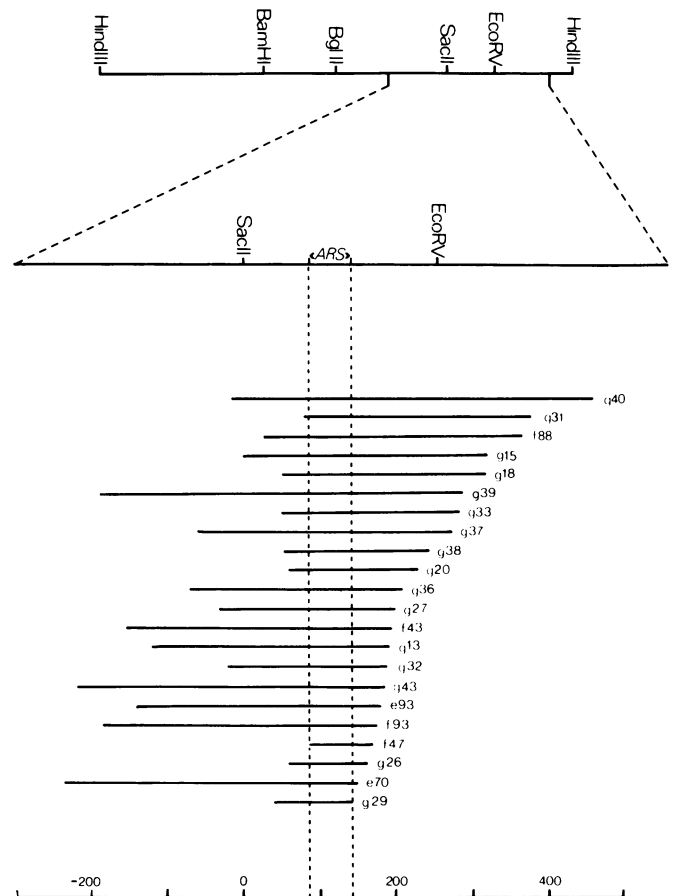


Fig. 2. Deletion mapping an *ARS* in a 2.5-kb *Hind*III fragment of *S. cerevisiae* DNA containing the *HO* gene. The restriction map of the *ARS*-containing DNA fragment is shown, with an expanded region corresponding to the part of the restriction fragment which has been sequenced (unpublished data). Sequences of *ARS*-containing fragments overlap as shown to define a small region, of 57 bp, in common with all clones, which presumably contains the *ARS* element, or at least sequences essential for *ARS* activity. 29 clones were analyzed in detail, and 22 of these are shown. Clones not spanning the *ARS*-containing region were not recovered. The *HO* gene is probably to the left-hand side of the *Hind*III restriction fragment (K.Nasmyth, unpublished results). The scale is numbered in nucleotides from the *Sac*II site.

recognizable pattern of bands for vector sequence between the *Bam*HI and *Hind*III sites, and thus facilitates sizing of the cloned inserts. Over half the clones contained fragments <300 bp, thus their sequences could be obtained from a single gel. Complete sequencing of some of these clones allowed the partial C sequences to be aligned, and the region of minimum overlap defines the maximum extent of the sequence essential for autonomous replication (Figures 2 and 3). There were no inconsistencies between the sequences obtained, indicating that any selection for mutations which improve *ARS* function is not a problem. Analysis of 29 different clones allowed the definition of only one *ARS*-containing region, all of the clones having inserts which spanned this region. Phage DNA from the f47 and g29 clones, whose inserts define the edges of the *ARS*-containing region (Figure 2), transformed yeast with a high efficiency, confirming the location of the *ARS* element.

Analysis of the sequence around the *HO* *ARS* showed it to contain sites for restriction enzymes *Sac*II and *Eco*RV (data not shown). These sites were mapped in the 2.5-kb *Hind*III

by transformation of a suitable yeast strain. Sequences of the cloned inserts overlap to define a single region, whose limits define the maximum extent of the region essential for autonomous replication. This mapping strategy should allow several *ARS* elements to be mapped simultaneously in large regions of DNA.

It should be pointed out that this method has formal limitations for defining the absence of *ARS* elements, since an *ARS* very closely linked to a region refractory to cloning in *E. coli*, such as a perfect inverted repeat (Lilley, 1981), would be difficult to detect. Also, the system inherently selects for, rather than just detects, *ARS* function. Since different *ARS* elements show different efficiencies of yeast transformation and autonomous replication (Stinchcomb *et al.*, 1980; Tschumper and Carbon, 1982), a poorly functioning *ARS* might therefore be difficult to map in a fragment which contains another stronger *ARS*. Thus, in order to unambiguously map all *ARS* elements in a given restriction fragment, those detected using the mapping strategy should be deleted from the original fragment to ensure that the remaining DNA cannot autonomously replicate.

A similar consideration applies to partially active fragments of a single *ARS*. Deletions near *ARS*I can reduce efficiency of autonomous replication, and this has been interpreted to suggest that a core element, capable of autonomous replication, is flanked by a sequence which in some way facilitates *ARS* function (Stinchcomb *et al.*, 1981). The strategy used here would tend to define boundaries of the fully active *ARS*, rather than the putative core sequence.

In addition to *ARS* mapping, the M13 vector allows rapid acquisition of short *ARS* sequences, thus facilitating analysis of the sequence requirements for autonomous replication. A strategy similar to that described here has provided a number of *ARS* sequences from total *S. cerevisiae* and *Xenopus* DNA (Figure 4, and Kearsley and Méchali, in preparation). M13 clones containing small *ARS*-containing fragments should be useful for directed deletion and site-specific mutagenesis studies.

Sequence requirements for autonomous replication in yeast

The overlap of randomly produced *ARS*-containing fragments defines a 57-bp region to the 3' side of the *HO* gene, which contains a sequence essential and probably sufficient for autonomous replication. Further deletion or mutational analyses will be necessary to show what parts of this region are important. Previous studies have identified regions as small as 75 bp or 100 bp as being essential for *ARS* function (Broach *et al.*, 1982; Tschumper and Carbon, 1982), but those results do not suggest a minimal size for the *ARS* element. In the analysis described here, the large number of deletion end points in the vicinity of the 57-bp region implies that the fully functional *ARS* is not considerably smaller than the defined region. Comparison between different *ARS* elements has so far only identified the *ARS* consensus sequence as a specifically conserved element (Stinchcomb *et al.*, 1981; Broach *et al.*, 1982). The presence of this consensus sequence in the two small *ARS* elements presented here (Figure 4) strongly supports the contention that it represents a structural requirement for autonomous replication. Apart from this consensus sequence, it is conceivable that less specific, or less obvious features of flanking sequence are also required for *ARS* function. Whether this other sequence forms an important part of the *ARS* structure remains to be seen. If a

general property, such as AT richness, is required for *ARS* function as well as a specific element, deletion analysis will not clearly distinguish between the two. Point mutations in the *HO* *ARS* region may be more useful for identifying nucleotides crucial for autonomous replication.

Materials and methods

Enzymes and chemicals

Restriction enzymes and T4 DNA polymerase were purchased from New England Biolabs and PL Biochemicals, and were used as recommended by the manufacturer. T4 DNA ligase was from D. Bentley. [α - 32 S]dATP was from Amersham.

Strains

E. coli strains used were DH1 (*recA1, endA1, gyrA96, thi-1, hsdR17* (rK^- , mK^+) *supE44*) and JM101 (Δ (*lac pro*), *thi, supE, F' traD36, proAB, lacI^f, Z Δ M15*). Yeast strain SX34-4D (*hmla, mata, hmra, mar1, ade8-10, ura3-52, leu2-3,112, trp1, his3*) was used for all transformations. YIp9-Sc2717 was provided by K. Struhl. Plasmid E347 was provided by K. Nasmyth.

Preparation of DNA

Plasmid DNA was prepared in *E. coli* using standard methods (Davis *et al.*, 1980). *E. coli* JM101 was used for the propagation of phage M13. Phage M13 RF DNA was prepared as described in Hong (1982) and single-stranded phage DNA was isolated according to Sanger *et al.* (1980).

Yeast DNA was prepared from *URA*⁺ transformants, grown on selective plates, as described in Sherman *et al.*, (1982).

Shotgun method for *ARS* mapping using M13se102

M13se102 RF DNA was cut with *Bam*HI, the staggered ends created were filled in by adding all four dNTPs (to 0.1 mM) and Klenow fragment of DNA polymerase I, and the DNA was phenol extracted and ethanol precipitated. E347 plasmid DNA (5 μ g) was sonicated in 30 μ l of 33 mM Tris-acetate (pH 7.9), 66 mM KAc, 10 mM MgAc, 0.5 mM dithiothreitol (DTT). The ends of the DNA fragment were repaired by adding all four dNTPs (to 20 μ M), and 10 units of T4 DNA polymerase, incubating overnight at 15°C. The sonicated DNA was phenol extracted and precipitated with ethanol. The sonicated fragments were ligated into the *Bam*HI-cut M13se102 in a 40 μ l mix containing 2 μ g/ml vector, 1–5 μ g/ml sonicated DNA, 70 mM Tris-HCl pH 7.9, 10 mM MgCl₂, 1 mM spermidine, 15 mM DTT and T4 DNA ligase. After 24 h the mixture was used to transfect competent *E. coli* JM101, as described in Davis *et al.* (1980), using 10 μ l ligation mix per plate. This produced almost confluent phage plaques. The plates were overlaid with 5 ml 20 mM Tris-HCl pH 7.5, 20 mM NaCl, 1 mM EDTA and left overnight at 4°C. The overlay solution was poured off, and spun at 10 000 r.p.m. for 10 min to remove debris.

Phage DNA for yeast transformation was prepared by inoculating a 40 ml culture of *E. coli* JM101 in log phase ($OD_{600} = 0.1$) with 100 μ l of the phage stock. After 5 h growth the bacterial cells were removed by centrifugation and 10 ml 20% PEG 6000, 2.5 M NaCl were added to the supernatant. After 5 min at room temperature, the phage were spun down (10 000 r.p.m., 20 min) and extracted twice with phenol and twice with chloroform. The single-stranded DNA (yield ~40 μ g) was precipitated with ethanol.

The *ura3-52* yeast strain SX34-4D was transformed with 5 μ g phage DNA using the protocol described in Sherman *et al.*, (1982), and transformants were grown on selective plates for DNA preparation. Integrative transformation by plasmids only containing the *URA3* gene is extremely inefficient with strains carrying the *ura3-52* allele (Stinchcomb *et al.*, 1980). *E. coli* JM101 was transfected with 0.1–1 μ g yeast transformant DNA, and single-stranded template was prepared from the rescued clones for sequencing. Clones were characterized using the dideoxy C sequencing reaction, and selected clones were completely sequenced (Sanger *et al.*, 1980), allowing alignment of the C sequences. Sequence data was handled using the DBUTIL program (Staden, 1980).

M13se102 does not contain an *Eco*K site, thus rK^+ *E. coli* strains such as JM101 can be used to rescue the autonomously replicating clone DNAs from yeast transformants. Ideally an rK^- strain should be used to guard against the possibility that an *ARS* element is coincident with an *Eco*K site.

Acknowledgements

I thank Ron Laskey, Marcel Méchali and Kim Nasmyth for helpful discussions. I also thank Kevin Struhl for providing a plasmid with a resectable *URA3* gene. S.K. is a Beit Memorial Research Fellow.

References

- Abraham,J., Feldman,J., Nasmyth,K.A., Strathern,J.N., Klar,A.J.S., Broach,J.R. and Hicks,J.B. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 989-998.
- Beach,D., Piper,M. and Shall,S. (1980) *Nature*, **284**, 185-187.
- Broach,J.R., Li,Y.-Y., Feldman,J., Jayaram,M., Abraham,J., Nasmyth,K.A. and Hicks,J.B. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 1165-1173.
- Celniker,S.E. and Campbell,J.L. (1982) *Cell*, **31**, 201-213.
- Chan,C.S.M. and Tye,B.-K. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 6329-6333.
- Davis,R.W., Botstein,D. and Roth,J.R. (1980), *Advanced Bacterial Genetics*, published by Cold Spring Harbor Laboratory Press, NY.
- Deininger,P.L. (1983) *Anal. Biochem.*, **129**, 216-223.
- Fangman,W.L., Hice,R.H. and Chlebowicz-Sledziewska,E. (1983) *Cell*, **32**, 831-838.
- Hong,G.F. (1982) *J. Mol. Biol.*, **158**, 539-549.
- Lilley,D.M.J. (1981) *Nature*, **292**, 380-382.
- Messing,J. (1981) in Walton,A. (ed.), *Third Cleveland Symposium on Macromolecules: Recombinant DNA*, Elsevier, Amsterdam, pp. 143-153.
- Nasmyth,K.A. (1982) *Annu. Rev. Genet.*, **16**, 439-500.
- Nasmyth,K.A. (1983) *Nature*, **302**, 670-676.
- Osley,M.A. and Hereford,L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7689-7693.
- Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
- Sherman,F., Fink,G. and Hicks,J.B. (1982) *Methods in Yeast Genetics*, published by Cold Spring Harbor Laboratory Press, NY.
- Singh,H., Bieker,J.J. and Dumas,L.B. (1982) *Gene*, **20**, 441-449.
- Staden,R. (1977) *Nucleic Acids Res.*, **4**, 4037-4051.
- Staden,R. (1980) *Nucleic Acids Res.*, **8**, 3673-3694.
- Stinchcomb,D.T., Struhl,K. and Davis,R.W. (1979) *Nature*, **282**, 39-43.
- Stinchcomb,D.T., Thomas,M., Kelly,J., Selker,E. and Davis,R.W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4559-4563.
- Stinchcomb,D.T., Mann,C., Selker,E. and Davis,R.W. (1981) *ICN UCLA Symp. Mol. Cell. Biol.*, **22**, 473-488.
- Tschumper,G. and Carbon,J. (1980) *Gene*, **10**, 157-166.
- Tschumper,G. and Carbon,J. (1982) *J. Mol. Biol.*, **156**, 293-307.

Note added in proof

In a recent publication on the regulation of *HO* gene expression by the mating-type locus, Jensen, Sprague, and Herskowitz (1983, *Proc. Natl. Acad. Sci. USA*, **80**, 3035-3039) also report that *ARS* activity is associated with the *HO* gene. They find that the larger *Bam*HI-*Hind*III subfragment of the restriction fragment shown in Figure 2 confers autonomous replication, consistent with the *ARS* mapping reported here. They also mention that the smaller *Bam*HI-*Hind*III subfragment confers high efficiency transformation on the YIp5 vector, implying that this fragment contains another *ARS*.