Curr Genet (1994) 26: 38–44

Current Genetics

Multiple-copy integration in the yeast Yarrowia lipolytica

Marie-Thérèse Le Dall, Jean-Marc Nicaud, Claude Gaillardin

Institut National Agronomique, Centre de Biotechnologie Agro-Industrielle, Laboratoire de Génétique Moléculaire et Cellulaire INRA-CNRS, F-78850 Thiverval-Grignon, France

Received: 2 November 1993

Abstract. Using an *Eco*RI-*Bgl*II fragment of the G unit of the rDNA of *Y. lipolytica* and a set of 11 deletions in the *URA3* promoter, we have constructed several plasmids to test gene amplification in the rDNA. These plasmids contain the rDNA fragment for integration, defective versions of the *URA3* gene, the *XPR2* gene encoding alkaline extracellular protease (AEP) as a reporter gene, and part of the pBR322 plasmid for selection and replication in *E. coli*. Among these plasmids, one corresponds to a deletion which allows multiple integration into the rDNA (plasmid pINA773). Two other plasmids (pINA767 and pINA772) give multiple integration only with a mutated *URA3* gene. Transformants carrying these three plasmids were tested for copy number, stability, chromosomal localization and AEP secretion. Transformants containing plasmids pINA767, 772 and 773 displayed an average copy number of 5, 12 and 25–60 copies respectively of the plasmid, as estimated by PCR and DNA hybridization. Integrations occurred in only one chromosome except for transformants containing 60 copies where copies were observed at least in two different chromosomes. Multiple integrations were found both as tandem repeats and as dispersed copies. Plasmid copy number was stable, in both minimum and rich media, for strains containing less than ten copies per cells. However, for higher copy number, multiple integrations were stable only when AEP synthesis was not induced, while in inducing medium stability of the multiple integrations was dramatically affected.

Key words: *Yarrowia lipolytica* – Gene amplification – rDNA – Secretion

Introduction

Yarrowia lipolytica is one of several yeasts currently considered as alternative hosts for heterologous gene expression (Romanos et al. 1992). Interest in it was stimulated by its capacity to naturally secrete several enzymes such as lipase, RNase, acid and alkaline proteases (for review see Heslot 1990). Indeed, under optimal conditions, this yeast is able to secrete 1–2 g/l of an alkaline extracellular protease (AEP). The *XPR2* gene encoding this protease has been cloned and analyzed (Davidow et al. 1987; Matoba et al. 1988; Nicaud et al. 1989). *XPR2* expression is tightly regulated by pH, as well as carbon and nitrogen sources (Ogrydziak and Mortimer 1977; Blanchin et al., unpublished). Expression is optimal in rich media, like YPD, and starts in late exponential phase (Nicaud et al. 1989). Its promoter (Blanchin et al., unpublished) and secretion signals (Fabre et al. 1991, 1992) have been studied and used to direct production and secretion of heterologous proteins such as bovine prochymosin (Franke et al. 1988 ; Nicaud et al. 1991), porcine α interferon (Nicaud et al. 1991), and human blood coagulation factor XIIIa (Tharaud et al. 1992).

These foreign genes were usually cloned into non-replicative vectors which integrate by homologous recombination into the genome (Davidow et al. 1985; Gaillardin et al. 1985; Gaillardin and Ribet 1987). Copy number was accordingly low, although limited amplification (5–10 copies) has been reported either through successive integration of plasmids carrying different markers (Nicaud et al. 1989), or through the use of a poorly-expressed heterologous gene (Gaillardin et al. 1985). These procedures are cumbersome and/or relatively inefficient, and no natural plasmid has yet been isolated in this species. Recently, Fournier et al. (1991) succeeded in isolating *ARS* vectors. These are present at a low copy number per cell, since they need the association of a centromeric sequence in order to be maintained (Fournier et al. 1993), and thus allow only for limited amplification of gene products (Nicaud et al. 1991). A major limitation of all these transformation systems is thus that no procedure is currently available to significantly increase the copy number of cloned genes.

Several approaches have been used in other yeasts to increase the copy number of genes inserted into chromosomes. For example, the δ-integration system in *S. cerevisiae*, coupled with the crossing of independent integrants, has been used to generate a 20-copy strain for the secre-

Correspondence to: J.-M. Nicaud

tion of nerve growth factor (Sakai et al. 1991). An apparently more general strategy was to target integration into the ribosomal DNA (rDNA) cluster and to select for multiple integrants using a defective marker. Using a *LEU2d* marker (Erhardt and Hollenberg 1983), the production of *S. cerevisiae* transformants containing up to 200 integrated copies of the transforming vector has been reported (Lopes et al. 1989). Amplification was also achieved with other defective genes such as *TRP1d* or *URA3d* (Lopes et al. 1991). This same strategy could also be extended to at least one other yeast, *Kluyveromyces lactis* (Bergkamp et al. 1992). The amplification levels achieved in both yeasts are comparable to those reported with 2μ-based vectors (Kingsman et al. 1987; Chen et al. 1989).

Our aim was to develop a similar multi-copy integration system for use in *Y. lipolytica*. In this paper we report on the construction of a set of integrative plasmids (pINA764 to pINA773). These plasmids contain a portion of the rDNA of *Y. lipolytica* as well as derivatives of the *Y. lipolytica URA3* gene as selection markers. These derivatives contain various promoter deletions either coupled, or not coupled, to a mutation in the coding region. In addition, these plasmids contained the *XPR2* gene used as a model for gene expression and protein secretion.

Materials and methods

Strains and culture conditions. E. coli strain TG1 was used for plasmid DNA propagation and single-strand DNA preparation (Sambrook et al. 1989). *Y. lipolytica* strain PO1d (*MatA, ura3-302, leu2-270, xpr2-322*) derives from the wild-type strain W29, ATCC24060 (Tharaud et al. 1992). Strain E129 (*MatA, ura3-302, leu2-270, xpr2-322*) derives from an inbred line (Fabre et al. 1991). Both strains carry non-reverting in-vitro-generated deletions of the *URA3, LEU2 and XPR2* genes. Yeast strains were grown in minimal medium YNB (Sherman et al. 1986) or in rich medium YPDm (Nicaud et al. 1991). *Y. lipolytica* was transformed by the lithium-acetate procedure as described in Xuan et al. (1988).

Construction of vectors. Plasmids pINA25, pINA154 and pINA156 were described previously (van Heerikhuizen et al. 1985; Nicaud et al. 1989). Plasmid pINA530 is a derivative of pINA156 where an 0.3-kb *Pst*I fragment was deleted to generate a *URA3* allele (*ura3d1*) containing only 41 bp upstream of the initiation codon. Plasmid pINA764 was constructed by a three-way ligation by inserting at the *Eco*RI-*Hin*dIII site of plasmid pINA154, a 1.3-kb *Eco*RI-*Bgl*II fragment carrying part of the rDNA from plasmid pINA760, together with a *Hin*dIII-*Bam*HI 1.3-kb fragment carrying the deleted *URA3* allele from plasmid pINA530 (see Fig. 1). Derivatives of this plasmid (pINA765 to pINA773) were obtained by exchanging its *Sma*I-*Hin*dIII fragment for the *Sma*I-*Hin*dIII fragment of the deletions (see below; the *Hin*dIII site came from the Bluescript™ polylinker). In the course of this ligation, a mutation destroying the *Sma*I site present in the coding region of *URA3* (*URA3ΔS*) was fortuitously generated in some clones (pINA767 and 772).

Construction of deletions. PCR reactions (ten cycles; 95 °C, 20 s.; 45 °C, 1 min.; 72 °C, 1 min.) were performed with oligonucleotide pairs URA3-1/URA3-3 and URA3-2/URA3-3 using plasmid pINA530 as a matrix. DNA sequences of URA3-1, URA3-2 and URA3-3 were respectively: CAGTCTCCTCTTCACCA, CTCTTCACCACCNAAATGCCCTCC and CTTGTCGTCAAGACCCACCCC (where N represents any nucleotide). URA3-1 and -2 hybridize upstream of, or else span, the initiation codon of *URA3*, whereas URA3-3 ends in the *Sma*I site of *URA3*. URA3-3 was gel-purified, whereas crude oligonucleotides were used in the case of URA3-1 and URA3-2. The amplified fragments (about 730 bp) were purified by Gene Clean (Stratagene) after agarose-gel electrophoresis, blunt-ended with the Klenow polymerase and treated with S1 nuclease (in order to recover incompletely-elongated products), and finally digested with *Xma*I so as to ensure a fixed 3′ end within the *URA3* coding sequence. The resulting fragments were cloned into the *Eco*RV-*Xma*I sites of the Bluescript vector KS⁻ (Stratagene). Sequence analysis was performed by the dideoxynucleotide chain-termination procedure (Sanger et al. 1977), using the T3 primer and a T7 polymerase sequencing kit (Pharmacia).

Copy-number determination. A rapid estimation of copy number was obtained by comparing the amount of amplified products derived from the deleted *XPR2* gene present in the chromosome and from the wild-type gene inserted on the amplified plasmid. A PCR reaction was performed directly on single colonies of transformants (Gussow and Clackson 1989) using two oligonucleotides (TAATGAGGGCATCGTCTTG and GATTCTAGACAGGCCCCATTG) corresponding to the DNA sequence flanking the *Apa*I deletion present in the chromosome of the recipient strains. Amplification of the *xpr2-322* allele gave a 526-bp fragment, while amplification of the wild-type allele gave a 675-bp fragment. Serial dilutions of the PCR reactions were run on an agarose gel. We estimated the copy number from the dilution factor required to lower the signal derived from the wild-type plasmid band to the level given by the smaller deleted chromosomal gene. For more accurate estimation, several independently-selected transformants were grown and chromosomal DNA was prepared. Serial dilutions of DNA were transferred to nylon membranes (Hybond, Amersham France) and hybridized separately with pBR322 and PGK probes (a fragment corresponding to the coding region of the *Y. lipolytica* phosphoglycerate kinase; B. Treton, unpublished). Plasmid copy numbers were determined by counting radioactivity for each dot hybridized with the pBR322 and PGK probes. Results were standardized using strain 20-12, which contains a single copy of the plasmid inserted at the *XPR2* locus, as a control.

Stability measurement. Stability of the multicopy integrants was assessed as follows: transformants obtained with a given plasmid were grown on non-selective medium for 12–16 generations. This was tested in both YPDm and YNB media, where AEP expression is, or is not, induced, respectively. Samples of each culture were taken at different times, plated for single colonies on YPD and checked for the loss of the plasmid-associated marker by replica-plating. Plasmid copy number was estimated by PCR on the whole culture at the starting point and after 12 generations, and on several Ura⁻ subclones obtained after plating on YPD.

Assay of AEP activity. AEP activity was assayed using a chromogenic substrate, azocoll (Nicaud et al. 1989), or by the agar diffusion method (Ogrydziak and Scharf 1982).

DNA preparation and chromosome analysis. Chromosomal DNA of *Y. lipolytica* was prepared according to Treton et al. (1992). DNA fragments were separated by electrophoresis in agarose gels and purified using a GeneClean Kit (Ozyme). Chromosomal plugs were prepared essentially as described by Naumova et al. (1993). Chromosomes were separated on an 0.8% agarose gel (Appligene) in 0.5× TAE using a CHEF DRII™ apparatus (BRL, Richmond, USA) and the following regimen: 47.7 h/40 V/3 300 s; 70 h/43 V/3 000 s; 48 h/50 V/2 400 s. (duration, voltage, switching time in seconds) at 12 °C.

Results

Design of the vectors

In order to obtain multiple-copy integrants, we have designed vectors carrying: (1) a fragment for selection and