Construction of Efficient Centromeric, Multicopy and Expression Vectors for the Yeast Kluyveromyces marxianus Using Homologous Elements and the Promoter of a Purine-Cytosine-Like Permease

Maria M. Ball¹, Alain Raynal², Michel Guérineau², and François Iborra²*

¹Present address: Laboratorio Organizacion y Expresión del Gen, Dpto. de Biología, Facultad de Ciencias, Mérida 5101 Venezuela
²Laboratoire de Biologie et Génétique Moléculaire, Institut de Génétique et Microbiologie, U.M.R. C.N.R.S. 8621, Bât 400 Université Paris 11, 91 405 Orsay Cedex France

Abstract

Efficient centromeric and multicopy vectors have been constructed for the yeast Kluyveromyces marxianus using homologous ARS and centromere sequences. A homologous promoter of a purine–cytosine permease gene called PCPL3 has been cloned, using an expression system based on GUS. Its strength has been estimated in K. marxianus by putting the homologous β-glucosidase gene under its control. This promoter is very efficient as activities higher than the ones obtained with the Saccharomyces cerevisiae PGK promoter were obtained. This promoter appears to be constitutive in various conditions tested. Its five transcription start sites have been mapped, and a derivative expression vector for K. marxianus has been constructed.

Introduction

Kluyveromyces marxianus is a yeast closely related to K. lactis. Many studies have shown that these yeasts are related to Saccharomyces cerevisiae (Lachance, 1993). In these three species, many genes have been characterized and have been shown to be often interspecifically functional and similarly organized (Webster and Dickson, 1988) and regulated by similar effectors (Breunig and Kruger, 1987). In some cases regulatory genes have been swapped from one yeast to another and proved to exert the same control (Oberye et al., 1993). Very often, similar UAS or URS are found in their promoters, and some are subject to similar regulations (Ruzzi et al., 1987; Cassart et al., 1997). Nevertheless, K. marxianus is more suitable than S. cerevisiae for biological purposes. K. marxianus is used industrially for the production of biomass by growth on whey and production of alcohol by growth on inuline rich plants (Bajpai and Bajpai, 1991; Brady et al., 1994). It is also a source of β–galactosidase and inulinase, two enzymes absent in S. cerevisiae. Inulinase can be secreted up to 1g/l, a yield by far much higher than the one obtained for the similar enzyme in S. cerevisiae. This yeast also has the property of growing at high temperatures and is consequently easier to handle in industrial fermentation.

Such “non conventional yeasts” whose secretion is very efficient are consequently used increasingly to produce metabolites and heterologous proteins difficult or impossible to obtain in S. cerevisiae. For a review see Fleer et al. (1992). As classic genetics is poorly developed in most of these yeasts, one has to develop efficient and appropriate vectors. From both evolutionary and biotechnology points of view, it was interesting to make this yeast accessible to genetic engineering. So far, the only vector available was a poorly efficient replicative vector using an ARS from the yeast K. lactis constructed with a selection for G418 (Das and Hollenberg, 1982). It was then adapted for a uracil selection (Iborra, 1993). Using this vector, ura3 mutants were isolated and a high efficiency transformation method was developed (Iborra, 1993). ARS sequences and two centromeres of K. marxianus have been cloned and described. They function in K. lactis but not in S. cerevisiae (Iborra and Ball, 1994). Using these elements, it was possible to construct homologous efficient vectors including expression vectors. In this context, it was interesting to isolate new promoters from K. marxianus and analyse their expression in this yeast assuming that homologous promoters will allow a better expression of the genes placed under their control. The cloning of these promoters was undertaken using a fusion with the β–glucuronidase (GUS) reporter gene.

In this paper, we report the construction of homologous K. marxianus multicopy and centromeric vectors as well as the cloning of a very active promoter used to construct an efficient expression vector. This promoter controls the expression of a gene, PCPL3, similar to one member of the purine-cytosine permease family of S. cerevisiae. The accession number for the described sequence is AJ011419.

Results and Discussion

Construction of Transforming Vectors for K. marxianus (Figure 1A)

Multicopy Plasmid Vector pKM2 (Escherichia coli-K. marxianus)

The 810 bp sequence of fragment B, KpnI-ClaI previously isolated from the genomic DNA of K. marxianus (Iborra and Ball, 1994) contains the origin of replication KMARS2.

Received July 14, 1999; revised September 16, 1999; accepted September 30, 1999. *For correspondence. Email iborra@igmors.u-psud.fr; Tel. +33-1-69 15 46 75; Fax. +33-1-69 15 72 96.

© 1999 Horizon Scientific Press
Further Reading

Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at caister.com

- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  [www.caister.com/malditof](http://www.caister.com/malditof)

- **Aspergillus and Penicillium in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  [www.caister.com/aspergillus2](http://www.caister.com/aspergillus2)

- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorfl, SM Roy, MA Riley (2016)
  [www.caister.com/bacteriocins](http://www.caister.com/bacteriocins)

- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  [www.caister.com/opdr](http://www.caister.com/opdr)

- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  [www.caister.com/acidophiles](http://www.caister.com/acidophiles)

- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  [www.caister.com/climate](http://www.caister.com/climate)

- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  [www.caister.com/biorem](http://www.caister.com/biorem)

- **Microalgae: Current Research and Applications**
  Edited by: MN Tsialoglou (2016)
  [www.caister.com/microalgae](http://www.caister.com/microalgae)

- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  [www.caister.com/gasplasma](http://www.caister.com/gasplasma)

- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  [www.caister.com/virus Evolution](http://www.caister.com/virus Evolution)

- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  [www.caister.com/arbo](http://www.caister.com/arbo)

- **Shigella: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  [www.caister.com/shigella](http://www.caister.com/shigella)

- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romani, H Guasch, MD Balaguier (2016)
  [www.caister.com/aquaticbiofilms](http://www.caister.com/aquaticbiofilms)

- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  [www.caister.com/alphaviruses](http://www.caister.com/alphaviruses)

- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  [www.caister.com/thermophile](http://www.caister.com/thermophile)

- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  [www.caister.com/flow](http://www.caister.com/flow)

- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  [www.caister.com/probiotics](http://www.caister.com/probiotics)

- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)

- **Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  [www.caister.com/cory2](http://www.caister.com/cory2)

- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  [www.caister.com/vaccines](http://www.caister.com/vaccines)

- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  [www.caister.com/antifungals](http://www.caister.com/antifungals)

- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  [www.caister.com/bacteria-plant](http://www.caister.com/bacteria-plant)

- **Aeromonas**
  Edited by: J Graf (2015)
  [www.caister.com/aeromonas](http://www.caister.com/aeromonas)

- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  [www.caister.com/antibiotics](http://www.caister.com/antibiotics)

- **Leishmania: Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  [www.caister.com/leish2](http://www.caister.com/leish2)

- **Acanthamoeba: Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  [www.caister.com/acanthamoeba2](http://www.caister.com/acanthamoeba2)

- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  [www.caister.com/microarrays2](http://www.caister.com/microarrays2)

- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  [www.caister.com/n2](http://www.caister.com/n2)

Order from caister.com/order
This fragment was inserted in the AatI site of pRS306 after rendering all ends blunt by treatment with the nuclease S1 (Figure 1A). This plasmid is much more stable than the ones based on the ARS of K. lactis. Its loss rate is only 12% per generation.

Centromeric Plasmid Vector pKMCA (E. coli-K. marxianus)
The 800 bp Clal-Sal sequence from fragment A, previously isolated from the genomic DNA of K. marxianus, contains the origin of replication KMARS1 and the centromere A. First, the inner HindIII site was suppressed after digestion with HindIII. Then, this fragment was inserted in the AatI site of pRS306 after rendering all ends blunt by treatment with the Klenow enzyme. This plasmid has a loss rate of 3% per generation. As the ARS and centromeres of K. marxianus function in K. lactis, but not in S. cerevisiae (Iborra and Ball, 1994), these vectors can be used in K. lactis, but not in S. cerevisiae.

Promoter Cloning with the GUS Reporter System

Construction of pYEL1/GUS Plasmid (Figure 1B)

Since K. marxianus has a strong endogenous β–galactosidase activity, another reporter gene was used: the GUS gene reporter system. It has been successfully used to detect β-galactosidase activity is as easy to detect as β–galactosidase using X-gluc and pNPG. The SalI-SalI fragment containing the modified GUS gene was inserted in the Xhol-BglII sites of plasmid pYEL1, a S. cerevisiae-E. coli shuttle plasmid. This places the GUS gene between the S. cerevisiae PGK promoter and its terminator, two well characterised and highly efficient sequences used to express heterologous proteins in S. cerevisiae.

Construction of the Promoter Cloning Vector pKM2/GUS (Figure 1C)
The 2.2 kb BamHI-SalI fragment of pYEL1/GUS containing the GUS gene and the PGK terminator was inserted in the BamHI-SalI sites of pKM2. This leads to a promoterless GUS gene followed by the PGK terminator.

Construction of the Library

2 to 5 kb Sau3A fragments resulting from a partial digestion of K. marxianus genomic DNA were cloned in the BamHI site upstream of the GUS coding sequence in the plasmid pKM2/GUS (Figure 1C). Thus, sequences having a promoter activity can be detected by the expression of the GUS gene. After transforming yeast cells, 6000 URA+ clones were examined and about 100 positive clones were detected qualitatively by their blue colour after permeabilisation in presence of X-gluc and quantitatively by the yellow coloration using pNP-glu as a substrate after lysing the cells with zymolyase. The GUS activity was spectrophotometrically measured for the most strongly coloured clones.

Cloning the Promoter

Although, 6000 clones are not expected to contain the whole genome of K. marxianus, it was sufficient to check the feasibility of the method and to find an efficient promoter. Clone 3 was chosen because it gave the highest level of GUS expression among the 10 most active clones. The insert of the clone 3 plasmid was sequenced. It revealed an ORF of 103 bp amino-acids fused in phase with the β–glucuronidase gene and preceded by a promoter like sequence.
Characterization of the PCPL3 Gene

It was of interest to know the nature of the gene under the control of the clone 3 promoter. Since the initial 103 amino-acids of the gene product were not sufficiently informative, the gene was cloned and sequenced and complete coding sequence was compared to sequence databanks.

*K. marxianus* DNA library constructed in the plasmid YEP352 was screened, using a *Spe*-Xho fragment of the promoter as a probe. One clone strongly hybridised with the probe. It contained a plasmid with a 10.5 kb insert. The promoter and the gene were localised after restriction mapping and hybridisation with the probe. DNA fragments containing the gene were then sequenced (Figure 2). Analysis of the DNA sequence revealed an ORF of 1686 bp coding for a putative 562 amino-acid protein with a predicted molecular weight of 63.17 kDa. Comparisons with sequences databanks of this predicted protein showed sequence identities with *S. cerevisiae* genes which had been classified in the purine/cytosine permease family, a subfamily of the major facilitator super family (MFS) characterised by two structural units of six transmembrane spanning α-helical segments connected by a cytoplasmic loop (Nelisen *et al.*, 1997). In *S. cerevisiae*, four genes constitute the purine/cytosine permease family: the purine/cytosine permease gene FCY2 (Weber *et al.*, 1990), two genes FCY21 and FCY22 located as tandem copies on chromosome V, and a YGL186C ORF located on chromosome VII with lower identity.

The best scores were obtained with the YGL186C ORF (Coglevina *et al.*, 1977) of unknown function coding for an hypothetical protein of 64.5 kDa. The gene product was 63% identical in a 587 amino acid overlap with YGL186C and 25% identical to FCY2 (Figure 3). Blast and Beauty analysis (Worley *et al.*, 1995) indicates three conserved domains between these three proteins. The gene was named PCPL3 for Purine-Cytosine Family n° 3.

![Figure 2. Sequence of the PCPL3 Gene and of its Promoter](image)
Figure 3. Comparison of the Deduced Protein Sequences of the *S. cerevisiae* Purine Cytosine Permease Gene Family and of the *K. marxianus* PCPL3 Gene

On first line, Sc1 is the *S. cerevisiae* FCY2 protein. On the second line, Sc2 is the *S. cerevisiae* YGL186C protein and on the third line Km is the *K. marxianus* PCPL3 protein. Black shaded amino acids are conserved in the three proteins, grey shaded amino acids are conserved between two proteins.
Analysis of the PCPL3 Promoter

Structure of the Promoter

The sequence of the PCPL3 promoter (Figure 2) contains typical promoter elements. They are very similar to the ones of S. cerevisiae genes: TATA, TATAAA and TATATA sequences, as well as a T-rich sequence found in various promoters between the TATA box and the initiation site (Heslot and Gaillardin, 1992). Five RNA initiation sites were found by the primer extension method before the putative TATA boxes: four major ones at positions -55, -96, -100, -104 and a weaker one at -109 (Figure 4).

Functional Analysis

In order to analyse the expression of this promoter in K. marxianus a XhoI site was introduced by PCR upstream of the ATG to remove the coding sequence and a BamHI site at -1049 bp to facilitate cloning (Figure 5). As a reporter gene, the endogenous β−glucosidase gene of K. marxianus was used rather than β−glucuronidase. It is well expressed in S. cerevisiae (Raynal and Guérineau, 1984) and was found to give more reliable results. The low endogenous activity in K. marxianus did not impair activity measurements. The modified promoter was then inserted either in pKM2 to get multicopy plasmids or in pKMCA to get centromeric plasmids. A SacI-SalI fragment containing the β−glucosidase ORF followed by its terminator was ligated in the XhoI site (Figure 1D).

Potential Regulatory Sequences

Three STRE consensus sequences (AGGGG) (Marchler et al., 1993) were found in this promoter (Figure 2). They are responsible upon stress conditions of the transcriptional induction of several S. cerevisiae genes. This sequence is present in the promoter of Hsp26 gene. To investigate their functionality, the yeast cells containing the recombinant gene proPCPL3-β−glucosidase-terPGK in the centromeric vector pKMCA were submitted to various stresses. Clearly in K. marxianus, these sequences by themselves are not sufficient to confer any stress induced regulation of the gene, since all the stress conditions tested (heat shock, oxidative shock, saline shock) did not induce a higher expression.

Six incomplete GCN4 consensus target sequences (TGACTN) were also found in this promoter. Although this core sequence has been proved to be essential to derepress upon amino acids starvation the synthesis of most amino acids, we examined the possibility of a general amino-acid regulation for the expression of this gene in K. marxianus, since the exact target of such regulation is not known in this yeast and could be slightly different. There is only a two fold effect upon addition of amino-acids both in K. marxianus and in S. cerevisiae, which may be explained by the more rapid growth of the cells in medium supplemented with all amino acids.

The expression of the fused gene is not sensitive to catabolic repression. Only a slight increase is observed when the cells are shifted from 2% glycerol to 2% glucose, although there is a two fold increase when cells are grown on 4% glucose. When nucleotides such as guanine, cytosine or adenine are added to the medium, no effect on the production of β−glucosidase have been noticed. Therefore, the PCPL3 gene appears to be constitutive for the various conditions tested.

To check the efficiency of the promoter PCPL3, promoters of Sacharomyces cerevisiae (GAL10, PGK, HSP26) and ADH1 of K. marxianus were modified in the

---

**Figure 4. Transcription Start Site**

Initiation starts of transcription for the promoter PCPL3 in K. marxianus determined by primer extension (-55, -96, -100, -104).

**Figure 5. PCR Modification**

Modification brought by PCR to promoters and β−glucosidase used in this study. Modifications are indicated in bold small letters.
same way (Figure 5). A XhoI site was introduced 5' of the
ATG (Figure 2), then inserted in pKM2 and a Sall-Sall
fragment containing the ORF of the β-glucosidase gene
and its terminator was ligated in this new XhoI site. As a
control, the same construction was used with the PGK
promoter of S. cerevisiae.

After transformation of K. marxianus, the transformants
were grown to early stationary phase in selective medium
and the β-glucosidase activity was determined in cell
extracts. Several clones for each construction were checked.
Figure 6 shows that the activity of the PCPL3 promoter is
about the same as the one obtained with the PGK
promoter in K. marxianus. It is also worth noticing that all S.
cerevisiae inducible promoters which were tested were also
inducible in K. marxianus but with an induction level some-
what lower (Johnston, 1987; Bossier et al., 1989). The high
activity of the PCPL3 promoter is not due to a titration ef-
fect of some repressor by a multicopy vector as high level
of expression were also obtained by a centromeric one.

Deletion Analysis

In order to further analyse this promoter two deletions
were constructed and analysed. When the promoter is short-
ened to the HindIII site at -256, there is a 25% increase
of promoter activity in K. marxianus (Figure 6). Further
reduction to -200 resulted in a loss of activity, although
the promoter still contains the TATA box and the five initia-
tion sites found by primer extension. Only an incomplete GCN4
core consensus binding sequence can be found in the -256
-200 fragment. No other known regulatory sites are present
in this fragment. This region contains two repeats of poly T
followed by poly A sequences. It is possible that they could
form a secondary structure important for the transcription
of an unknown site of regulation. The reason why a per-
mease promoter is so active in K. marxianus is unknown
and would be worth further investigation.

Construction of an Expression Vector

The high activity obtained with the -256 promoter deletion
led to the construction of a small sized expression vector
with a better expression than the one obtained with the
PGK promoter. Several restriction sites of the polylinker
are available in front of the promoter and can be used to
insert regulatory sequences which might be very useful
for specific needs. This vector can be either multicopy using
pKM2 or centromeric using pKMCA as starting vectors.
The new vectors for the yeast K. marxianus, containing
homologous ARS and centromere sequences can be
be very efficient. The replicative vector is more stable than
the only other vector available constructed with an ARS
from K. lactis. Despite the high strength of the promoter in
the expression vector, no instability of the vectors carrying
it was observed. With the cloned PCPL3 promoter, and its
-256 bp derivative, an efficient expression vector is now
available for this yeast.

Experimental procedures

Strains and Growth Media

K. marxianus Fl 47 ura3 was derived by UV mutagenesis from ATCC 12424
strain, a homothalic strain. It was grown on YPD rich medium or YNB-
glucose (Difco) minimal medium supplemented by necessary auxotrophic
requirements. It was transformed by electroporation using a Jouan square
pulse generator (Iborra, 1993), or by the Li-PEG method of Ito et al. (1983).
E. coli strain DH5α (BRL) was propagated on LB medium. It was rendered
competent and transformed by the method of Inoue et al. (1990).

Plasmids

pRS316, a Bluescript KS (Stratagene) derivative containing the S. cerevisiae
URA3 gene without a yeast origin of replication was obtained from Dr
Sikorski (Sikorski and Hieter, 1989). Plasmid pYEL1 and pYEL3 were con-
structed by Nicaud et al. (1994). Plasmid pBH101 and GUS kit were ob-
tained from Clontech. They were prepared by the method of Berghammer
and Auer (1993) or by alkaline lysis. Plasmid YEP352 was constructed by
Hill et al. (1986).

Colonies Detection and Enzymes Activities

A replica of yeast colonies was made on a Whatman filter paper W540.
The filter was frozen at -70°C, then thawed. It was then deposited on a
solution of 67 mM phosphate buffer pH7 containing zymolyase and 5-bromo-
4-chloro-3 indolyl-glucuronide (X-glc) and incubated at 30°C. The deepest
blue colonies were selected. Cells extracts were obtained after lysis by
treatment with zymolyase or disruption with glass beads followed by a 15
min centrifugation in a microfuge. Protein concentration was measured on
these extracts by the method of Bradford (1976). Then, the hydrolysis of
paranitrophenyl(pNP)Glucuronide and pNPGlucose by the cell extracts
was followed by the increase of absorbancy at 425 nm in 67 mM phosphate
buffer pH7 with an EM of 13 10⁶.

Stress Conditions Used

Cells transformed by the centromeric vector pKMCA containing the β-gluc-
osidase gene under the control of the PCPL3 promoter were grown on
YPD up to DO 0.5. They were then transferred in YPD medium containing
either NaCl 0.3 M, or Sorbitol 0.4 M, or H₂O₂ and gathered after one or two
hours of induction. Themic shock was produced by shifting the cells grown
at 28 to 50°C for K. marxianus. Such high temperatures were used for K.
marxianus since optimal growth temperature for this strain is 37°C.

General Amino Acids Repression

The cells were grown on minimal medium, then washed and grown in the
same medium with 0.2% casaminoacids.
Polymerase Chain Reaction (PCR)
PCR DNA amplification were done on a TECHNE thermocycler (OSI) using *Thermus aquaticus* DNA polymerase from Eurogentec or Appligene.

Primer Extension
Primer extension was performed with mRNA isolated with Hybond-mAP paper (Amersham) using the primer extension kit of Promega.

Double Strand DNA Sequencing
It was performed on both strands on Exonuclease III nested deletions (Pharmacia kits) or on restriction fragments using the dideoxy termination method, either manually or on AB1373 Stretch sequencer (PE Applied Biosystems) using a Dye Terminator kit. DNA sequences were compared to the EMBL library using the BLASTP + BEAUTY programs (Altschul et al., 1990; Worley et al. 1995). Protein sequences were compared using the Multialign version 5.3.3. of Corpet (1988).

Acknowledgements
The authors gratefully acknowledge Dr. Bolotin-Fukuhara and Dr. Carl Mann for critical reading of the manuscript.

References


