MOLECULAR BIOLOGY & GENETIC ENGINEERING OF YEASTS

Molecular Biology and Genetic Engineering of Yeasts

Authors

Henri Heslot, D.Sc.
Emeritus Professor
Department of Genetics
Institut National Agronomique
Paris-Grignon, France

Claude Gaillardin, Ph.D.
Professor
Department of Molecular and Cellular Genetics
Institut National Agronomique
Paris-Grignon, France

Library of Congress Cataloging-in-Publication Data

Heslot, H.
Molecular biology and genetic engineering of yeasts/authors,
Henri Heslot, Claude Gaillardin.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-5645-8
1. Yeast fungi—Biotechnology. 2. Genetic engineering.
3. Molecular biology. I. Gaillardin, Claude. II. Title.
TP248.27.Y43H47 1991
660'.65—dc20 91-33676
CIP

Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., N.W., Boca Raton, Florida 33431.

International Standard Book Number 0-8493-5645-8

Library of Congress Card Number 91-33676
Printed in the United States 1 2 3 4 5 6 7 8 9 0

THE AUTHORS

Henri Heslot, D.Sc., is Emeritus Professor of Genetics at the Institut National Agronomique Paris-Grignon (France). He is Vice-Chairman of the French Society of Microbiology and member of the "Comité des Applications de l'Académie des Sciences".

He is an author of over 100 articles on yeast genetics and molecular biology and has presented his research findings at numerous national and international meetings.

Claude Gaillardin, Ph.D., is Professor of Molecular and Cellular Genetics at the Institut National Agronomique Paris-Grignon (France), Director of the CNRS-INRA associated Laboratory of Genetics of Industrial Microorganisms.

After entering Ecole Normale Supérieure in 1966, he received his degrees from the Université de Paris and completed a Ph.D. in 1976 on yeast genetics at the Institut National Agronomique. He entered the INRA in 1976 and after a postdoctoral stay at MIT (Cambridge, U.S.) held successive appointments at INRA as Research Fellow, Research Director, and currently Professor at the Institut National Agronomique.

His major interests are on the molecular genetics of nonconventional yeasts with particular emphasis on their biotechnological applications. In addition to his research activities, he is active as a teacher and as a member of several scientific evaluation or advisory committees.

TABLE OF CONTENTS

Chapter 1

YEAST TRANSFORMATION

I. YEAST TRANSFORMATION

Early efforts leading to yeast transformation[1,2] have led to the development of a sophisticated array of vectors and systems for manipulating the genome of *Saccharomyces cerevisiae*.[3] Most of these techniques and approaches now seem to be readily extendable to other yeast species: transformation of, e.g., *Schizosaccharomyces pombe*[4] or *Kluyveromyces lactis*[5] was reported quite early. Other nonconventional yeasts received attention more recently (for review, see Reference 6). The same general principles seem to apply to the ten or more species so far tested; for example, integrative vectors do generally integrate by homologous recombination, and chromosomal sequences (*ARS*) can confer extrachromosomal replication. This makes yeasts quite different from other eukaryotic cells including fungi, where nonhomologous recombination events are frequent or predominant and where *ARS* sequences could not be convincingly demonstrated up to now.

A. PROCEDURES FOR YEAST TRANSFORMATION

Early work on yeast transformation used spheroplasts stabilized in an isotonic medium.[1,4,5] In the presence of calcium ions and polyethylene glycol (PEG), DNA is taken up by a poorly understood process.[7] The efficiency of transformation (transformed cells per surviving protoplast) depends on the genetic background of the recipient[8] and on the type of selection applied: under nonselective conditions, several percent of transformed cells can be obtained.[9] Fusion of yeast spheroplasts with DNA-containing liposomes[10] or with *Escherichia coli* minicells harboring yeast *E. coli* shuttle plasmids[11] can be considered as variations of the above procedure and result in a higher transformation efficiency (up to 10% with minicells). Many strains, however, do not exhibit such a high transformation efficiency, and selective markers have to be used in order to select transformants.

The spheroplast method, although still widely used, is time consuming, and its efficiency can be severely limited by the regeneration step where protoplasts are allowed to regenerate into whole cells. Several methods using whole cells have been developed which are, in some cases, as efficient as the spheroplast method. Cells of *Saccharomyces cerevisiae* treated with alkali cations (lithium, for example)[12] or with 2-mercaptoethanol[13] were transformed at high efficiency in the presence of PEG and after a heat shock. The specificity for cation requirement (i.e., monovalent vs. divalent) was found to be strain specific, some strains being transformable only in the presence Ca^{2+}, others only with Li^{+}.[14] Although PEG is generally needed to obtain a significant transformation frequency, probably by triggering irreversible adsorption of plasmid DNA,[14] some strains can be transformed without PEG.[15] Whole cell transformation using the LiCl method have been adapted to other yeast species such as *Schizosaccharomyces pombe*,[16] *K. lactis*,[17] or *Yarrowia lipolytica*.[18]

Although the mechanisms involved in spheroplast and whole cell transformation are equally poorly understood, the two procedures probably rely on quite different mechanisms. Cotransformation seems to be the rule when spheroplasts are used, indicating that many plasmids enter the cells. Initial experiments with two integrative plasmids carrying different marker genes gave 25 to 33% cotransformation in *Saccharomyces cerevisiae*,[19] On the contrary, few plasmids seem to enter cells transformed by the LiCl method.[20] On the other hand, cells made competent by the LiCl method seem to take up linear DNA more readily than circular DNA, whereas no such difference is observed when the spheroplasting method is used: linearized *ARS* plasmids transform spheroplasts of *S. cerevisiae* at the same frequency

as circular plasmids, but they transform 80-fold more efficiently than uncut plasmids of LiCl-treated cells.[21] A similar increase of transformation efficiency of linearized *ARS* plasmids was observed on LiCl-treated *Y. lipolytica* cells.[22]

Transformation by electroporation has been reported on spheroplasts[23] and on whole cells.[24] It avoids using PEG which may be detrimental to cell viability. Since this method does not rely on a "natural competence" stage, it may be widely used for strains displaying otherwise poor transformability. The method has been recently optimized for transforming whole cells of *S. cerevisiae, Schizosaccharomyces pombe, K. lactis,* and *Y. lipolytica* yielding transformation frequencies above 10^5 transformant per microgram of input plasmid.[25]

All above methods were effective for transforming the nucleus, but failed in transforming mitochondria. Organelle transformation has been achieved recently in several systems including yeast mitochondria[26] by bombarding cells with tungsten microprojectiles coated with DNA. A strain bearing a deletion in the mit DNA *oxi3* gene was bombarded with 1-μm tungsten microprojectiles coated with a mixture of two uncut plasmids, YEp 352 (carrying a nuclear marker [URA3] and a 2-μm sequence) and pQA oxi3 (carrying a functional *oxi3* gene): about 0.1% of the nuclear transformants (Ura^+) were respiratory competent (able to grow on glycerol). All mitochondrial transformants resulted from homologous recombination between plasmid and mit DNA. Transformation of mitochondria completely devoid of mit DNA (rho^o) was reported using the *oxil* gene cloned into pBR322:[27] in this case the plasmid was amplified into concatemers of the size expected for wild-type mit DNA.

Transfer of plasmids by "transgeneric sex" may provide an alternative method to transformation for transferring plasmids between bacteria and yeasts.[28] In the experiments reported, the plasmids to be transferred carried the cis-acting sequence oriT (either derived from an F plasmid or from an Inc P1 R-plasmid) in top of a yeast marker and origin of replication. *mob* functions (F or R specific) were provided in trans by F or R plasmids. When *E. coli* strains carrying these plasmids were mixed with exponentially growing yeast cells, yeast "transconjugants" appeared at a frequency 10^{-7} to 10^{-5} per recipient cell and up to 10^{-2} per donor cell, depending on the ratio of donor to recipient cells in the assay. Transfer was 10- to 100-fold below frequency of bacterial conjugation but similarly required viable donor cells, was DNAse insensitive, required cell contact, and depended on matched *oriT* and *mob* functions. *tra* functions encoded by the R and F plasmids appeared (perhaps surprisingly) essential for transfer. "Matings" occurred on plates but not in liquid, and the size dependence was high: a complete F′ plasmid (100 kb) was transferred 10^4 times less efficiently than a 10-kb plasmid.

B. FATE OF TRANSFORMING DNA

Both double-stranded and single-stranded DNA have been used to transform yeast with high efficiency. Single-stranded circular vectors were reported to transform 10- to 30-fold more efficiently than double-stranded vectors.[29] If they carried a yeast origin of replication, they were faithfully converted to double-stranded plasmids in yeast. Single-stranded plasmids carrying no origin of replication recombined efficiently with homologous chromosomal sequences without detectable intermediate free duplex formation.[30]

Yeast cells were shown to relegate at high frequency linearized DNA, suggesting that they contain a very efficient ligase activity, or a low exonuclease activity, or both, as compared with *E. coli*. Plasmids carrying a yeast origin of replication were linearized or gapped before transformation.[31] A very limited drop in transformation efficiency was observed even when the gaps were made in plasmid sequences having no homology with the yeast genome. Recircularization occurred whether the ends were flush or protruding with compatible or uncompatible overhangs. Fragments entering the same cell were colligated efficiently: a mixture of DNA fragments carrying selectable markers (*LEU2* and a 2 μm origin of replication on a *Hind*III fragment, and *URA3* on an *Eco*RI fragment) transformed

at a frequency 100-fold lower only than the circular plasmid carrying both markers,[31] and circular plasmids were recovered in *E. coli*.

Recombination between DNA fragments entering the same yeast cell has been used to create new molecules *in vivo*. An extensive series of plasmid derivatives was built using gapped replicative plasmids cotransformed with a DNA fragment presenting homology to each side of the gap.[21] No selection was necessary for the additional sequences present on the fragment between the two homology blocks, and most plasmids recovered resulted from faithful recombination, thus bypassing tedious *in vitro* procedures. Insertion of any DNA cloned in a pBR322 derivative into a yeast plasmid by recombination *in vivo* has been documented in various yeast species: copolymerization occurred at frequencies of 10% at least in *S. cerevisiae*[32] and of more than 50% in *Schizosaccharomyces pombe*[33] or *Y. lipolytica*.[34]

The possibility to repair a gapped plasmid on a cotransformed linear fragment has been elegantly used by Pompon and Nicolas[35] to generate a library of mosaic proteins by recombining *in vivo* rabbit and mouse cytochrome P-450 (P-450) genes which are only 69 to 73% sequence related. Plasmids carrying the rabbit P-450 gene gapped at various positions were cotransformed with a cDNA of mouse P-450. Of the plasmids recovered, 90% were recombinants. There was no evidence of sequence scrambling, and recombination took place within segments of high local similarity as short as 17 bp. In the absence of any functional selection, most of the recombinants did produce an active protein, indicating that few frame shifts and stop codons were created. Since mouse and rabbit P-450 have markedly different substrate specificity, assessing the substrate specificity of chimeric proteins permitted the identification of segments critical for substrate specificity. This approach of the structure-function relationships of related proteins should be readily extendable to other proteins.

Transformation of yeast with synthetic oligonucleotides 20 to 60 bp long has been reported.[36] A strain carrying a stop codon after the third codon of *CYC1* was reverted to Cyc^+ (growth on nonfermentable carbon source) after transformation of 3.10^7 cells with 100 μg of oligonucleotide; the frequency of transformation was 10^3 times lower than with a control ars plasmid. Both single-strand and double-strand oligonucleotides were effective. Site-specific mutagenesis on single-stranded vectors has been performed using yeast transformation:[37,38] cotransformation of the single-stranded template (carrying an *ARS* sequence, a selectable marker, and a gene inactivated by a nonsense mutation) with an oligonucleotide (reverting the stop mutation) in a 30:1 ratio resulted in up to 43% of the plasmids having corrected the mutation. This method offers, thus, very high efficiency without requiring any enzymatic step *in vitro*.

Most of the former results as well as those related below (see Section II) amply demonstrate that yeasts process incoming DNA faithfully by homologous recombination and reparation, a rather unique situation among eukaryotes. Occasional deviations observed during the transformation process may reflect recombination pathways that have not been fully characterized yet. Spontaneous deletions occur on transformation of circular replicating plasmids.[39] They are markedly stimulated if the plasmid is linearized prior to transformation,[40] suggesting that spontaneous linearization during the transformation process may be a step in the generation of deletions in circular molecules, when the gap cannot be corrected into a homologous region. Recircularization in the absence of a homology region by "random" joining of free ends seems not to be, however, the major mechanism of plasmid reparation: up to 82% of the yeast transformants recovered on transformation of a linearized *ARS* plasmid contained head-to-head dimers of the plasmid,[41] suggesting that yeast may possess a double-strand break processing activity whose product is a symmetrical joint. On the other hand, high frequency deletion (up to 80%) of sequences flanked by direct repeats as short as 23 bp was observed on transformation of *Y. lipolytica* by circular plasmids:[21] deletions occurred during the transformation process and not during subsequent growth.

C. MARKER GENES

Marker genes used for *Saccharomyces cerevisiae* transformation are generally biosynthetic genes which complement the corresponding auxotrophy in the recipient. Several resistance genes have been developed and are described in Chapter 10. *S. cerevisiae* genes have been widely used for complementing auxotrophs of other yeast species (reviewed in Reference 6), although homologous genes are preferred whenever possible, e.g., *ura4* in *Schizosaccharomyces pombe*,[16] *LEU2* in *Y. lipolytica*,[18] *HIS4* in *Pichia pastoris*,[42] *ADE2* in *Candida albicans*.[43] Whereas homologous genes of defined function are relatively easy to isolate by complementation of known *Saccharomyces cerevisiae* or *E. coli* mutants, isolation of suitable auxotrophs is generally less straightforward in nonconventional yeasts and may require prior obtention of a rather large collection of mutants. These must then be screened for the loss of the chosen enzymatic activity. Alternatively, suitable recipients may be identified by complementation after transformation: this permitted identification of *trp1* mutants of *K. lactis*,[5] of *his4* mutants in *P. pastoris*,[42] and of *leu2* and *lys2* mutants in *C. maltosa*.[44,45] Genes conferring resistance to two antibiotics have also been used in nonconventional yeasts, e.g., G418^R in *K. lactis*,[5] phleomycin^R in *Y. lipolytica*.[46] Heterologous genes permitting use of normally nonassimilated carbon sources have been reported for both *Y. lipolytica* and *P. pastoris*.[47,48]

II. GENE MANIPULATION WITH INTEGRATIVE VECTORS

A. BASIC FEATURES OF INTEGRATIVE TRANSFORMATION

Plasmids carrying a DNA sequence homologous to the chromosome and devoid of yeast origin of replication preferentially integrate by homologous recombination via a Campbell-like mechanism in 80% or more of the transformants tested in *S. cerevisiae*.[19] This generates a duplication of the target sequence on both sides of the plasmid (see Figure 1). Several plasmids may cointegrate in a tandem array.[49] Cointegrates have been observed at a high frequency (up to 50% at certain loci, e.g., *LEU2*).[50,51] Integrated plasmids are mitotically quite stable and are faithfully segregated through meiosis.[19] Occasional reexcision of the plasmid by homologous recombination between the flanking sequences[52] or by sister chromatid exchange[49,53] occurs at low frequency (around 10^{-4}), but this efficiency may be site dependent. Sister chromatid exchange between tandemly repeated copies generates one chromatid with a single copy (excision of the vector) and one chromatid with three copies. Such amplifications do occur spontaneously at the same rate as excisions[49] and can be selected using markers showing a strong dosage effect like *CUP1*.[56] Multiple tandem integration in *S. cerevisiae*,[54] in *Y. lipolytica*,[34] or in *Hansenula polymorpha*[55] tends to be stable in mitosis but is lost at high frequency at meiosis.

Conversion events (i.e., transfer of genetic information from the plasmid to the chromosome without integration of the plasmid) occur in 10 to 20% of the cases in *S. cerevisiae*[19] but may represent the major outcome in *Schizosaccharomyces pombe* when the selected marker is a homologous gene.[57] Gene conversion events can be increased up to 80% in *Saccharomyces cerevisiae* when the plasmid is cut outside the homology region.[39]

Out of site integrations are observed at a very low rate in yeasts and may actually represent integration into short or imperfect regions of homology fortuitously reiterated through the genome.[19] This may account for the observation that with vectors which do not show homology to the genome, transformants do occur by "random" integration in the genome: such events are rare in *S. cerevisiae*, but occur at quite a high rate in *Schizosaccharomyces pombe*.[57-59]

The frequency of integration of circular plasmids is low, typically yielding, in all yeast systems studied, 1 to 10 transformants per microgram of input DNA. Cutting the plasmid

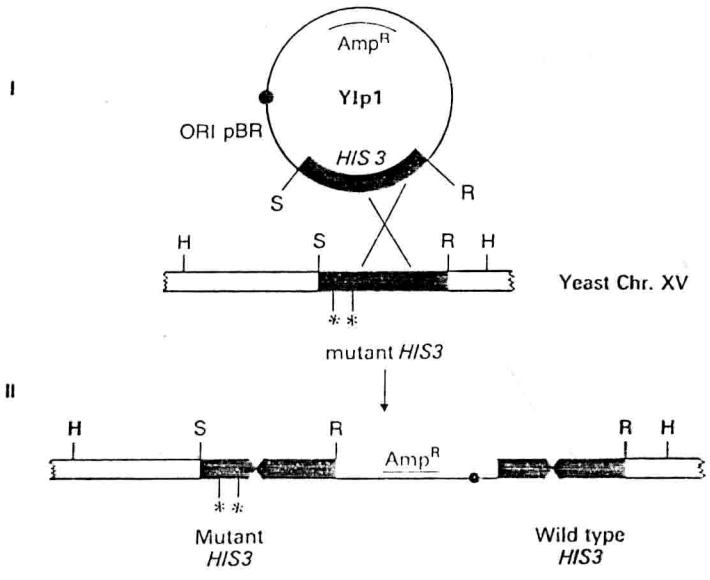

FIGURE 1. Schematic representation of homologous recombination of an integrative vector into its chromosomal target. The asterisks indicate mutations, the thin line is pBR322, and the thick line is yeast DNA. (From Kingsman, S. M. and Kingsman, J. A., in *Genetic Engineering*, Blackwell Scientific Publications, Oxford, 1988, 64. With permission.)

within the homology region increases transformation frequencies by a factor of 100 to several thousand in *Saccharomyces cerevisiae*,[50] *P. pastoris*,[60] *Y. lipolytica*,[18] and *Schizosaccharomyces pombe*.[57] Occasionally, no dramatic increase in the transformation frequency is observed.[57] This may reflect proximity of the mutated site in the chromosome and of the cut site in the plasmid. Plasmid integration in (and repair on) the chromosome involves widening of the gap on the plasmid and formation of heteroduplex tracts adjacent to the break region over hundreds of nucleotides[61] in *Saccharomyces cerevisiae*: mutations present in the chromosome can then be transferred to the plasmid and reduce the (apparent) number of transformants obtained (this has been used to retrieve mutated alleles, see below). The cut must furthermore leave sufficient homology on its sides, so that gap widening will not extend past the homology border on the plasmid. Site-directed integration has been obtained with cut plasmids leaving 170 bp of homology on one side, but not with 41 bp.[50] Almost complete loss of plasmid information occurs within the first 10 bp flanking the break, and half maximal rescue is observed with 300 bp.[62]

Complex plasmids carrying more than one region of homology with the chromosome can be targeted efficiently to only one of the possible targets by cutting the plasmid into the corresponding sequence on the plasmid as shown initially by Orr-Weaver et al.[64] with a vector carrying the *sup3*$^+$ and *HIS3* genes of this yeast.

Targeted integration has been used to retrieve mutant genes. In its simplest version, a plasmid carrying a wild-type gene is integrated at the mutated chromosomal locus using an enzyme which cuts close to one extremity of the gene. Transformants carrying an integrated plasmid at the locus are identified by Southern analysis, and their genomic DNA is then digested by an enzyme which cuts at the other extremity of the gene. This excises the plasmid and the flanking chromosomal copy which is recircularized *in vitro* and rescued in *E. coli*. In a different version (see Figure 2) the plasmid is integrated via a sequence flanking the gene of interest, and the genome is digested with an enzyme which cuts outside the gene