

# Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*

Nancy A. Da Silva & Sneha Srikrishnan

Department of Chemical Engineering and Materials Science, University of California, Irvine, CA, USA

**Correspondence:** Nancy A. Da Silva, Department of Chemical Engineering and Materials Science, University of California, Irvine, CA 92697-2575, USA. Tel.: +1 949 824 8288; fax: +1 949 824 2541; e-mail: ndasilva@uci.edu

Received 15 October 2011; revised 23 November 2011; accepted 23 November 2011; Final version published online 12 January 2012.

DOI: 10.1111/j.1567-1364.2011.00769.x

Editor: Jens Nielsen

## Keywords

yeast; plasmids; genomic integration; promoters.

## Introduction

The yeast *Saccharomyces cerevisiae* is a key laboratory and industrial microorganism and an excellent host for metabolic pathway engineering. As a eukaryote, *S. cerevisiae* can synthesize a variety of fungal and mammalian proteins that can prove problematic in bacteria. Ease of cultivation, success at the industrial level over many years, and generally recognized as safe (GRAS) status by the U. S. Food and Drug Administration contribute to the interest in this microorganism. The broad array of tools available for molecular-level manipulation of this species and knowledge of the *S. cerevisiae* metabolic, secretory, transport, signaling, and other pathways enable the successful engineering of this yeast for a diverse range of applications.

Pathway engineering requires the regulated expression of multiple foreign or native genes. In contrast to the operons that can be employed in prokaryotic cells, a series of independently transcribed genes must be introduced in yeast. Both the level and the timing of enzyme synthesis can be essential for the successful introduction of new pathways. In addition to gene number and tran-

## Abstract

Metabolic pathway engineering in the yeast *Saccharomyces cerevisiae* leads to improved production of a wide range of compounds, ranging from ethanol (from biomass) to natural products such as sesquiterpenes. The introduction of multienzyme pathways requires precise control over the level and timing of expression of the associated genes. Gene number and promoter strength/regulation are two critical control points, and multiple studies have focused on modulating these in yeast. This MiniReview focuses on methods for introducing genes and controlling their copy number and on the many promoters (both constitutive and inducible) that have been successfully employed. The advantages and disadvantages of the methods will be presented, and applications to pathway engineering will be highlighted.

scription level/timing, translational and post-translational control can be important for modulating protein levels.

This MiniReview focuses on methods developed to introduce and control the expression of genes in *S. cerevisiae*, focusing on those most useful for metabolic engineering applications. The review will consider the control of gene number, an important method for regulating overall expression. This will include the advantages and limitations of plasmid vectors, and useful new vector series. Chromosomal gene integration is efficient in yeast and offers precise control over gene copy number and stability. We will consider the important issues and methods for chromosomal gene integration. Controlling transcription level via constitutive and inducible promoters offers another critical method of gene expression regulation. Useful promoters (including new series of promoters), the range of promoter strengths available, and regulation will be summarized. Recent metabolic engineering studies in *S. cerevisiae* will be used to demonstrate the application of these methods.

While the review will focus on the critical elements of introducing, maintaining, and expressing genes, we recognize the importance of additional methods to control the

enzyme levels in yeast, including those modulating translation and post-translational processing.

## Introduction of genes and regulation of gene number

Both plasmid vectors and chromosomal integration are widely used to introduce genes and control copy number in *S. cerevisiae*. Each has an important role, and the choice depends on the overall goal (e.g. overexpression, precise control of gene number). While the plasmids available for use in yeast are much more limited than those for *Escherichia coli*, they have been successfully employed for many metabolic engineering applications. They are extremely useful for gene expression; however, plasmids offer limited control of copy number, and segregational stability can be a significant issue even in selective medium. As homologous recombination is very efficient in *S. cerevisiae*, integration of genes into the genome offers an alternate, straightforward mechanism for gene introduction. Chromosomal integration also allows the insertion of precise numbers of the same or different genes. This is particularly important for the regulated expression of metabolic pathway genes.

## Plasmid vectors

The three classes of autonomously replicating plasmids in yeast are YRp, YE<sub>p</sub>, and YC<sub>p</sub>. All are *S. cerevisiae*/*E. coli* shuttle vectors that typically carry a multiple cloning site (MCS) for the insertion of expression cassettes. YRp vectors carry a *S. cerevisiae* origin of replication (e.g. ARS sequence) with no partitioning control. These are extremely unstable (Murray & Szostak, 1983a; Da Silva & Bailey, 1991) and not generally used for metabolic engineering applications. In contrast, the widely used YC<sub>p</sub> and YE<sub>p</sub> vectors have proven successful for many applications. YC<sub>p</sub> (CEN/ARS) vectors carry both an origin of replication and a centromere sequence, have high segregational stability in selective medium, and are maintained at 1–2 copies per cell (Clarke & Carbon, 1980). YE<sub>p</sub> vectors are based on the *S. cerevisiae* native 2 $\mu$  episomal plasmid and contain either the full 2 $\mu$  sequence or, more commonly, a 2 $\mu$  sequence including both the origin and the *REP3* (*STB*) stability locus (Futcher & Cox, 1983; Kikuchi, 1983). For the full sequence, use of a *cir*<sup>0</sup> strain lacking the native plasmid is recommended to prevent the recombination between the vectors and to keep copy number of the recombinant vector high. For the partial 2 $\mu$  plasmids, a *cir*<sup>+</sup> host carrying the native 2 $\mu$  is required to provide the transacting factors (*REP1* and *REP2*) required for stability. These vectors are generally more structurally stable than the full 2 $\mu$  plasmids, but may be

maintained at lower copy number. Although the maintenance of YE<sub>p</sub> vectors at 10–40 copies (Romanos *et al.*, 1992) is generally assumed, copy number is not controlled and can vary widely with the gene product and level of expression. Expression from a strong constitutive promoter or synthesis/secretion of complex products can reduce the average copy number and plasmid stability significantly, or overload a pathway in the cell (e.g. Moore *et al.*, 1990; Ro *et al.*, 2008; Fang *et al.*, 2011). In extreme cases, use of a CEN/ARS vector may give higher product levels (Wittrup *et al.*, 1994).

The general lack of yeast plasmids that are maintained at very high copy number has led to the development of 2 $\mu$ -based vectors carrying selection markers such as *LEU2-d* and *URA3-d* (Beggs, 1978; Erhart & Hollenberg, 1983; Loison *et al.*, 1989). The defective promoters on the markers result in an increase in copy number; hundreds of copies have been reported in selective medium, although these high copy numbers are not required for viability (Lopes *et al.*, 1991). Such vectors are generally more useful for the overexpression of a product gene than for metabolic engineering applications. However, Ro *et al.* (2008) demonstrated the successful use of a plasmid carrying the *LEU2-d* marker (and three pathway genes) for the synthesis of artemisinic acid in nonselective, complex medium.

Several vector series (Table 1) have been developed carrying a series of selection markers on YE<sub>p</sub> and YC<sub>p</sub> plasmids (and YIp integrating vectors discussed below). Ma *et al.* (1987) constructed a series of YC<sub>p</sub> and YE<sub>p</sub> plasmids with *LEU2*, *HIS3*, *LYS2*, *URA3*, and *TRP1* selection markers. The YE<sub>p</sub>lac and YC<sub>p</sub>lac plasmids (Gietz & Sugino, 1988) carry a MCS and *URA3*, *TRP1*, and *LEU2* selection markers on both 2 $\mu$ - and CEN/ARS-based plasmids, respectively. The pRS series (Sikorski & Hieter, 1989; Christianson *et al.*, 1992) are similar useful cloning vectors with *URA3*, *TRP1*, *HIS3*, and *LEU2* markers on both 2 $\mu$  and CEN/ARS plasmids. Brachmann *et al.* (1998) and Taxis & Knop (2006) extended the pRS series to include the *MET15*, *ADE2*, *kanMX*, *hphNT1*, and *natNT2* selectable markers. The various vector series have been widely used for gene expression in yeast.

Groups of vectors carrying constitutive and inducible promoters have also been developed (Table 1). These include the variants of the pRS series carrying *ADH1*, *TEF1*, *GPD1*, *MET25*, *CYC1*, *GAL1*, and *GALL* or *GALS* (*GAL1* variant) promoters (Mumberg *et al.*, 1994, 1995) and the *CUP1* promoter (Labbe & Thiele, 1999). Cartwright *et al.* (1994) developed YE<sub>p</sub> and YC<sub>p</sub> expression vectors carrying a *URA3* selection marker and the *PGK*, *GAL1*, *GAL10*, *PHO5*, and *CUP1* promoters. The YE<sub>p</sub>lac and YC<sub>p</sub>lac plasmids were modified to carry the tetracycline-responsive *tet-on/off* promoters (Gari *et al.*, 1997).

**Table 1.** Vector series for gene expression in *Saccharomyces cerevisiae* (autonomous plasmids and integration vectors/templates)

Series name	Origins for YCp & Yep vectors	Selection markers	Promoters	Comments	References
YCplac, YEplac, Ylplac	ARS1/CEN4, 2 $\mu$	URA3, TRP1, LEU2		Contain pUC19 MCS*	Gietz & Sugino (1988)
pSXXX (integration vectors)		URA3	P <sub>TEF1</sub> , P <sub>GAL1</sub> , P <sub>STE12</sub>	Ylplac211 backbone	Sadowski et al. (2007)
YCp4XX, YEp4XX, YRp4XX	ARS1/CEN4, 2 $\mu$ , ARS1	URA3, TRP1, HIS3, LEU2, LYS2		pBR322 <sup>+</sup> backbone	Ma et al. (1987)
pRS series (YCp, YEp, Ylp)	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2, kanMX, ADE2, MET15, hphNT1, natNTZ		pBLUESCRIPT <sup>+</sup> -based vectors	Sikorski & Hieter (1989), Christianson et al. (1992), Brachmann et al. (1998), Taxis & Knop (2006)
p4XX prom. (pRS variant)	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>CYC1</sub> , P <sub>ADH1</sub> , P <sub>TEF1</sub> , P <sub>GPD1</sub> , P <sub>MET25</sub> , P <sub>GAL1</sub> , P <sub>GALL</sub> , P <sub>GALS</sub>		Mumberg et al. (1994, 1995)
pCu4XX prom (pRS variant)	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>CTRB</sub> , P <sub>CTRI</sub> , P <sub>CUP1</sub>		Labbe & Thiele (1999)
pGREG series (pRS variant)	ARSH4/CEN6	URA3, TRP1, HIS3, LEU2, kanMX	P <sub>GAL1</sub>	Nine N- and C-epitope tags	Jansen et al. (2005)
p4XX prom.att (Gateway <sup>TM</sup> )	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>TEF1</sub> , P <sub>GPD1</sub> , P <sub>MET25</sub> <sup>**</sup> , P <sub>GAL1</sub> <sup>**</sup>	**Four N- and C-epitope tags	Funk et al. (2002)
pV2XX (Gateway <sup>TM</sup> )	ARS1/CEN4	URA3, TRP1	P <sub>PSK1</sub> , P <sub>REG-CYC1</sub>	Four N- and C-epitope tags	Van Mullen et al. (2003)
pLGXXX (Gateway <sup>TM</sup> )	ARSH4/CEN6	URA3, TRP1, HIS3, LEU2	P <sub>GAL1</sub>	Some with V5-6xHis tag	Geiser (2005)
pAG (Gateway <sup>TM</sup> )	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>TDH3</sub> , P <sub>GAL1</sub>	Seven N- and C-epitope tags, integration vectors also available	Alberti et al. (2007)
pCMXXX series (YCplac, YEplac variant)	ARS1/CEN4, 2 $\mu$	URA3, TRP1	P <sub>REG-CYC1</sub> with tTA activator	Varied number of tetO boxes	Gari et al. (1997)
pYC, pYES	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, bsd <sup>§</sup>	P <sub>GAL1</sub>	YES <sup>TM</sup> Vector Collection	Invitrogen
YCp-SPB series, YEp-SPB series, Ylp-GAL1-SPB	ARS1/CEN4, 2 $\mu$	URA3	P <sub>PSK1</sub> , P <sub>GAL1</sub> , P <sub>GAL10</sub> , P <sub>HIS5</sub> , P <sub>CUP1</sub>		Cartwright et al. (1994)
pXP series	ARSH4/CEN6, 2 $\mu$	URA3, TRP1, HIS3, LEU2, ADE2, MET15	P <sub>PSK1</sub> , P <sub>TEF1</sub> , P <sub>HXT7-391</sub> , P <sub>ADH2</sub> , P <sub>GAL1</sub> , P <sub>CUP1</sub>	Allows marker recycling (CreA-loxP system) when used as template for integration	Fang et al. (2011), Shen et al. (unpublished data)
pBEVY, pBEVY-G	2 $\mu$	URA3, TRP1, LEU2, ADE2	P <sub>TDH3</sub> -P <sub>ADH1</sub> , P <sub>GAL1</sub> -P <sub>GAL10</sub>	Bidirectional promoter system	Miller et al. (1998)
pY2x-GAL(1/10)-GPD	2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>GAL1</sub> -P <sub>TDH3</sub> , P <sub>GAL10</sub> -P <sub>TDH3</sub>	Bidirectional promoter system	Li et al. (2008)
pESC Vectors	2 $\mu$	URA3, TRP1, HIS3, LEU2	P <sub>GAL1</sub> -P <sub>GAL10</sub>	Bidirectional promoter system	Agilent Technologies
pSP-G1, pSP-G2 (pESC-URA variant)	2 $\mu$	URA3	P <sub>TEF1</sub> -P <sub>PSK1</sub>	Bidirectional promoter system	Partow et al. (2010)

\*Multiple cloning site. †Bollivar et al. (1977). ‡Stratagene. §Blasticidin resistance gene (Takeuchi et al., 1958).

The commercially available pYES and pYC series (Invitrogen) offer expression from the *GAL1* promoter on 2 $\mu$  or CEN/ARS vectors, respectively. These vectors are available with *URA3*, *TRP1*, and blasticidin resistance selection markers. The pGREG vectors (Jansen *et al.*, 2005) are derivatives of the pRS series with five different selectable markers on CEN/ARS-based plasmids with a *GAL1* promoter and can be useful tools for plasmid construction via *in vivo* recombination and for the expression of N- and C-terminal-tagged fusion proteins (nine tags available). Vectors series such as those by Funk *et al.* (2002), Van Mullem *et al.* (2003), Geiser (2005), and Alberti *et al.* (2007) have been constructed utilizing the Gateway<sup>TM</sup> cloning technology (Invitrogen, review by Walhout *et al.*, 2000), a versatile method that allows for the insertion of ORFs into vectors by *in vitro* recombination using the bacteriophage lambda *att* sites. The vector series contain various promoters and selection markers, 2 $\mu$  or CEN/ARS sequences, and additional features such as epitope tags. Recently, Fang *et al.* (2011) constructed a series of 32 pXP shuttle vectors, utilizing three constitutive promoters, *PGK1*, *TEF1*, and *HXT7-391*, and six reusable selection markers on both 2 $\mu$  and CEN/ARS vectors. These vectors can be used as templates for gene integration as described below. M.W.Y. Shen, F. Fang, S. Sandmeyer & N.A. Da Silva (unpublished data) have extended this series to include the *GAL1*, *ADH2*, and *CUP1* promoters.

These various vector series have been widely used for introducing single genes and also multiple genes for metabolic engineering in *S. cerevisiae*. Jiang *et al.* (2005) introduced a vector built from pYES expressing three genes for the biosynthesis of naringenin. Yan *et al.* (2005) introduced four genes for this pathway onto a single plasmid built from YEplac. Despite the presence of four *GAL1* promoters on the same vector, no structural instability was observed for the 65 h of culture. However, repetitive sequences on a plasmid can be an issue as seen in the work by Verwaal *et al.* (2007) on carotenoid synthesis in *S. cerevisiae*, where significant structural instability led to the decision to integrate the genes. Leonard *et al.* (2005) introduced a 2 $\mu$  plasmid carrying four genes and a CEN/ARS vector carrying three genes (built from YEplac and YCplac, respectively) for flavone biosynthesis. Carlson & Srien (2006) used modified versions of the pRS vectors to engineer *S. cerevisiae* for the production of the biopolymer poly[(R)-3-hydroxybutyrate] (PHB). Two plasmids (one containing the bidirectional *GAL1-GAL10* promoter) were used to express three pathway genes. These are just a few examples of the widespread use of such vector series for metabolic pathway engineering.

Vectors carrying two or more promoters allow the expression of more than one gene on a single plasmid.

The bidirectional promoter plasmid series pBEVY and pBEVY-G were constructed by Miller *et al.* (1998) and have the *TDH3* and *ADH1* promoters (constitutive), or the *GAL1-GAL10* (galactose-inducible) promoters, on 2 $\mu$  vectors with four different selectable markers. Li *et al.* (2008) extended this series to include eight new 2 $\mu$ -based pY2x-GAL(1/10)-GPD plasmids; these carry an inducible *GAL1* or *GAL10* promoter and a constitutive *TDH3* promoter with four different markers. The pESC vector series is commercially available (Agilent Technologies) and has the bidirectional *GAL1-GAL10* promoter cassette on 2 $\mu$ -based vectors carrying one of the four selectable markers. The recently described pSP-G1 and pSP-G2 vectors (Partow *et al.*, 2010) carry the constitutive *TEF1* and *PGK1* promoters in two different orientations; these are based on the original pESC-URA3 plasmids and aimed at high-level expression in yeast. All of these vectors allow the expression of two genes from the same construct, avoiding the need to carry two different plasmids in the cell.

Vectors carrying bidirectional promoters are thus very useful for metabolic engineering in *S. cerevisiae*. In particular, pESC variants carrying the bidirectional *GAL1-GAL10* promoters have been used to express two to four genes on a single 2 $\mu$  plasmid. Kim *et al.* (2011) used a pESC vector to evaluate the expression of pathway genes on the biosynthesis of ceramides. Maury *et al.* (2008) introduced the bacterial isoprenoid pathway into *S. cerevisiae* with eight genes carried on two plasmids (four genes under the control of two bidirectional promoters per plasmid).

To introduce multiple genes simultaneously, yeast artificial chromosomes (YACS) may be useful. Originally described by Murray & Szostak (1983b), these vectors carry an origin of replication, a centromere sequence, and telomeres and can be constructed by both *in vivo* recombination and *in vitro* ligation of the DNA fragments (Burke *et al.*, 1987). Because of their ability to carry large fragments of DNA, such vectors have been useful in studying genomes, but can also be used to carry large pathway constructs on a single new 'chromosome'. In a recent study by Naesby *et al.* (2009), eYACs (expressible yeast artificial chromosomes) were used to randomly assemble a group of genes for a seven-step flavonoid pathway. Fifty percent of the clones produced naringenin when grown in the presence of coumaric acid. Stability was comparable to normal YACS, and flavonoid production was maintained for more than 50 generations.

### Chromosomal integration of genes

Plasmid vectors are ideal for the overexpression of genes at 'high' (YE<sub>p</sub>) or 'low' (YC<sub>p</sub>) levels and are convenient and easy to use. In combination with integrated genes,

plasmids allow a quick assessment of the degree of over-expression required in a pathway. However, two or more 2 $\mu$  and/or CEN/ARS vectors can be difficult to stably maintain simultaneously in a single cell (Futcher & Carbon, 1986; Mead *et al.*, 1986). For the introduction of multiple genes, long-term stability, and precise control of expression, integration of the genes into the chromosome holds several advantages.

The ease of homologous recombination in *S. cerevisiae* (Oldenburg *et al.*, 1997; Raymond *et al.*, 1999; Schaerer-Brodbeck & Barberis, 2004; Gibson *et al.*, 2008) makes genomic integration an attractive method for the introduction of pathway genes. A variety of vector- and PCR-based methods have been developed for single- or multicopy integration. In combination with reusable selection markers and characterized integration sites, chromosomal integration provides precise control over gene copy number and can ensure segregational stability. Integration may not be ideal for the overexpression of genes but is a key method for metabolic engineering in yeast. Important issues include the ease of gene integration and selection, stability of the inserts, and the ability for multiple simultaneous or sequential integrations (e.g. by marker recycling).

### Vector-based gene integration

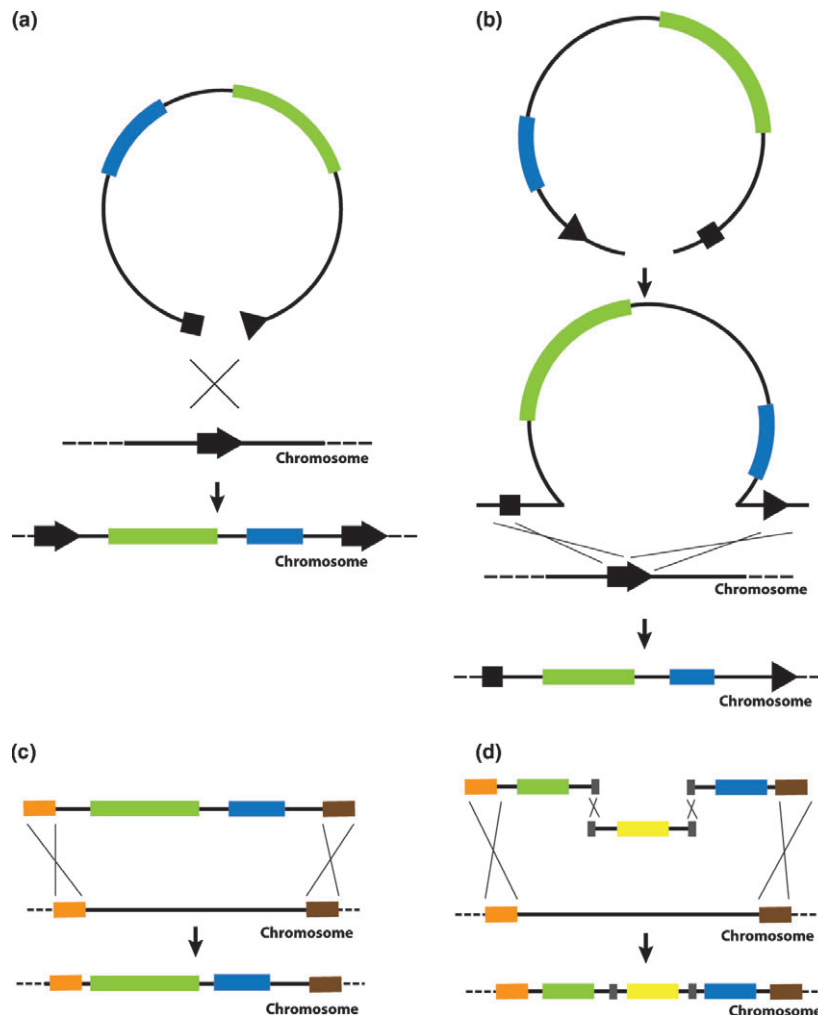
Yeast integrating plasmids (YIp) are vectors that carry a MCS, selection marker, target site(s), and no replication origin. These cannot be maintained in the cell unless integrated into the chromosomes. Depending on the position of the target sequences, integration into the genome can occur via single-crossover (Fig. 1a) or double-crossover (Fig. 1b) homologous recombination. For single-crossover, the vector is linearized within the target sequence. Integration results in duplicated sequences flanking the insert; this can result in structural instability as excision can occur via homologous recombination at these sites. For double-crossover integration (gene conversion and omega integration), the cassette to be integrated is inserted within the target sequence and the vector is linearized outside of this sequence. The genomic insertions are structurally stable as no duplicate sequences occur upon integration. The efficiency of integration via double-crossover is lower relative to single-crossover events, and construction of the vectors can be more difficult.

YIp vector series (Gietz & Sugino, 1988; Sikorski & Hieter, 1989; Cartwright *et al.*, 1994; Alberti *et al.*, 2007; Sadowski *et al.*, 2007) have been developed carrying a series of auxotrophic markers that also act as target sequences, and integration occurs by a single-crossover event. In addition to the standard YIp vectors, useful variants have been developed through the incorporation

of resistance selection markers, reusable selection markers, and repeated target sites. Use of a resistance gene, such as the bacterial *neo<sup>r</sup>* for resistance to the aminoglycoside G418 in *S. cerevisiae* (Jimenez & Davies, 1980), allows the selection of multiple simultaneous integrations through G418 selection. The success at obtaining multiple insertions is because of the weak expression of the bacterial resistance gene and increased concentrations of the G418. Selection via resistance to G418 and other compounds (e.g. hygromycin and zeocin) can be particularly effective when combined with repetitive target sequences including the Ty1 elements, the ribosomal DNA cluster, and delta elements (Kingsman *et al.*, 1985; Lopes *et al.*, 1989; Fujii *et al.*, 1990; Sakai *et al.*, 1990, 1991).

Delta ( $\delta$ ) elements are the long terminal repeats (LTRs) of the *S. cerevisiae* Ty1 and 2 retrotransposons (Boeke, 1989; Boeke & Sandmeyer, 1991). Based on the sequence of strain S288C (Goffeau *et al.*, 1996), there are several hundred delta elements dispersed in the *S. cerevisiae* chromosomes as solo  $\delta$  elements or associated with Ty elements (Dujon, 1996; Kim *et al.*, 1998; Wyrick *et al.*, 2001). The multiple possible integration locations in combination with G418 or ethionine selection have led to a high efficiency of multicopy integration in a single transformation. A range of integrations (up to 80) has been reported using this method (Shiomi *et al.*, 1995; Parekh *et al.*, 1996; Wang *et al.*, 1996; Lee & Da Silva, 1997a). Despite the large number of available sites, the majority of inserts are in long tandem repeats at one location. While the method is easy to employ, there are two major drawbacks: (1) the method can generally be used only once or twice (repeated G418 selection tends to lead to other mutations allowing resistance, not higher integrated copy number) and (2) the tandem nature of the integrations can lead to high instability, particularly following induction or for products that place a burden on the cell (Wang *et al.*, 1996; Lee & Da Silva, 1997a). However, this method has been shown to be quite effective for optimizing gene copy number when an inducible *GAL* promoter and late induction of gene expression are combined (Parekh & Wittrup, 1997; Shusta *et al.*, 1998). An example of the use of  $\delta$ -neo method for pathway engineering is the integration of cellulase and  $\beta$ -glucosidase for the conversion of cellulose to ethanol (with no instability observed) (Cho *et al.*, 1999). This method is not ideal for introducing precise numbers of a group of different genes.

$\delta$  elements have been chosen as the target site for several integration methods (discussed below) and successfully applied for pathway engineering. A recent novel approach, 'cocktail  $\delta$ -integration', was used to create strains for the surface expression of  $\beta$ -glucosidase, endoglucanase, and cellobiohydrolase (Yamada *et al.*, 2010).



**Fig. 1.** (a) Vector integration into a chromosome via single-crossover integration. The vector is linearized within the target sequence (black arrow) prior to integration. (b) Vector integration into a chromosome via double-crossover integration. The cassettes to be integrated are inserted within the target sequence on the vector, and the vector is linearized outside of the target sequence prior to integration. (c) Integration of a linear DNA fragment (e.g. PCR product) with homology at 5' and 3' ends. (d) *In vivo* recombination and integration of nested DNA fragments.

Three successive rounds of cotransformation of the three integrating vectors were performed (using a different selection marker for each round: *URA3*, *HIS3*, and *TRP1*), with different numbers of each gene integrated in each transformation. This allowed the creation of a group of strains and selection of an optimum strain for PASC (phosphoric acid-swollen cellulose) degradation.

The incorporation of reusable selection markers enables multiple sequential gene integrations owing to marker recycling. These markers include a selection gene flanked by direct repeats allowing excision by homologous recombination. Because spontaneous loss by recombination is a low-frequency event, counterselectable markers are generally used, enabling easy selection of cells that have lost the marker. In combination with a repetitive target site

(e.g. a  $\delta$  sequence), multiple integrations of the same (or different) gene are possible with the one general vector construct. Lee & Da Silva (1997b) developed an integrating vector carrying the reusable *URA3* blaster cassette (developed by Alani *et al.* (1987) for multiple gene knockouts) and the Ty1  $\delta$  sequence. Genes are integrated sequentially, and *URA3* marker excision is selected using 5-FOA (Boeke *et al.*, 1984) after each vector (or partial vector) insertion. Expression is generally linearly correlated with the number of integrations (Lee & Da Silva, 1997b, 2006). A major advantage of this method is the dispersed nature of the genes, resulting in much higher stability than with tandem inserts. Disadvantages are the length of time required for multiple integrations and knockouts and the need for Southern blots or qPCR to

determine the number of insertions (as the exact delta location is not specified). The integrations via this vector are also single-crossover events that lead to approximately 300 bp delta repeats on either side. While much more stable than tandem insertions, loss of the inserts can occur with time. A double-crossover version of the vector (Fig. 1b) was developed by placing the integration cassette within the  $\delta$  element and linearizing outside of this sequence prior to transformation (Ching, 2005). While successful for the integration of a series of different genes, integrations of multiple copies of the same gene were not possible due to the large homologies; during subsequent integrations, replacements of genes/marker at the same location as the previous integration were selected.

The combination of delta target site and reusable selection cassette has been successfully applied for the engineering of strains for the synthesis of a variety of compounds. Examples include the engineering of *S. cerevisiae* for the production of 1,2-propanediol (Lee & Da Silva, 2006) and, in combination with plasmids, for the synthesis of artemisinic acid (Ro *et al.*, 2006).

Several other reusable selection markers have been developed. Particularly powerful are those with an active recombination system for marker recycling (Prein *et al.*, 2000; Johansson & Hahn-Hagerdal, 2002; Radhakrishnan & Srivastava, 2005). In these cases, the marker need not be counterselectable. The Cre/*loxP* and FLP/FRT systems (Sauer, 1987; Sauer, 1994; Guldener *et al.*, 1996; Gueldener *et al.*, 2002; Radhakrishnan & Srivastava, 2005) have been successfully applied in yeast. In the former, the selection marker is flanked by *loxP* sequences, and the expression of the Cre recombinase allows extremely efficient marker excision. Multiple markers can be excised simultaneously, significantly decreasing the time for strain construction. The small footprint of the *loxP* sequence also reduces the chance of off-target homologous recombination during subsequent integrations.

### PCR-based gene integration

PCR-generated fragments have been widely used for gene knockouts (and gene knockins) in yeast (Langle-Rouault & Jacobs, 1995; Lorenz *et al.*, 1995; Manivasakam *et al.*, 1995; Goldstein *et al.*, 1999; Gueldener *et al.*, 2002). In recent years, integration of PCR-generated fragments has been a key method for the insertion of pathway genes into the yeast genome and the generation of libraries (Schaerer-Brodbeck & Barberis, 2004; Shao *et al.*, 2009). The ease of homologous recombination in *S. cerevisiae* and the reduction in time and effort make this a preferred approach in many instances. Use of one of the many high-fidelity polymerases allows errors during PCR to be minimized, and genomic inserts can be recovered

to confirm the correct sequence. In addition, insertions typically occur by double-crossover integration leading to the construction of stable strains (Fig. 1c).

In *S. cerevisiae*, homologous recombination requires only limited flanking homology (Manivasakam *et al.*, 1995). Efficient targeting increases with the length of the homology, and an overlap of approximately 50 bp (25 bp on each side) is sufficient to easily screen and recover integrants in specific genomic locations. Standard length primers can thus be used to amplify a desired gene cassette with flanking target regions, and the PCR product can be transformed into the yeast for insertion by double-crossover. The efficiency of recombination in yeast also allows the use of nested primers or the assembly of two or more fragments (Fig. 1d) (Erdeniz *et al.*, 1997; Hawkins & Smolke, 2008; Flagfeldt *et al.*, 2009; Shao *et al.*, 2009).

The pXP vector series mentioned earlier (Fang *et al.*, 2011; M.W.Y. Shen, F. Fang, S. Sandmeyer & N.A. Da Silva, unpublished data) was designed to allow the seamless transition from plasmid-based (2 $\mu$  or CEN/ARS) expression to PCR-based chromosomal gene integration. All selection markers on the vectors are flanked by *loxP* sequences to allow marker recycling following integration. The promoter-gene-terminator and selection cassettes can be amplified using standard priming sequences with overhangs to target specific chromosomal sites. Although less efficient than traditional vector-based integration where homologies are typically much longer, transformants can be rapidly screened via colony PCR to select the correct integration location. This system can be used for multiple sequential (and possibly simultaneous) gene integrations and multiple simultaneous marker excisions, and is thus a useful tool for pathway engineering. These vectors have now been utilized for the insertion of metabolic pathway genes in several studies currently underway in our laboratory, including the modification of central carbon metabolism, polyketide production, and arsenic uptake and sequestration (unpublished results).

Shao *et al.* (2009) made use of the efficiency of recombination in yeast to develop the DNA Assembler method for the insertion of entire pathways into the yeast genome (or onto plasmids). Expression cassettes are assembled and integrated in a single step, and large inserts of up to 19 kb have been successfully inserted demonstrating the promise of this tool for metabolic engineering. A group of different promoters and terminators were utilized to avoid repetitive sequences and subsequent instability of the inserts. The yeast Ty1 delta sequence was chosen as the insertion site, and integration occurs by single-crossover integration into one of the hundreds of potential target sites in the yeast chromosomes, contributing to the high efficiency reported. The method has been applied for

the introduction of a 3-gene xylose utilization pathway, a 5-gene zeaxanthin biosynthesis pathway, and the two pathways together (at one site) (Shao *et al.*, 2009).

The insertion of metabolic pathways via independent gene integration requires multiple target sites that provide well-characterized expression levels for the inserted genes. Traditional target integration sites have been nutritional marker genes (e.g. *URA3*, *TRP1*, and *LEU2*), although a variety of other locations have also been utilized. Prior work on single-crossover integration into the delta sequences (Lee & Da Silva, 1997b, 2006) found similar expression levels for multiple different integrations; however, the specific delta sites used were not determined.

Two recent studies have focused on the expression of reporter genes following integration (and full target replacement) at a group of well-defined sites. Using the pXP vectors as templates and primers containing targeting regions, Fang *et al.* (2011) compared the expression of the luciferase gene under the control of the strong *PGK1* promoter ( $P_{PGK1}$ -*Rluc*- $T_{CYC1}$  cassette) at 14 different locations. The expression cassette replaced *URA3*, *MET15*, *LEU2*, *TRP1*, seven full-length Ty elements (two with replacements in both orientations), and one full-length Ty3 element. Similar levels of expression were observed at the locations studied, with a maximum difference of approximately 50% in luciferase fluorescence. The values were approximately 60% of that on a CEN/ARS vector, as expected based on a plasmid copy number of 1–2. In another recent study by Flagfeldt *et al.*, 2009, expression of *lacZ* under the control of the two strong *TEF1* and *ACT1* promoters ( $P_{TEF1}$ -*lacZ*- $T_{CYC1}$  and  $P_{ACT1}$ -*lacZ*- $T_{CYC1}$  cassettes) was compared at 20 different locations: replacements of *URA3*, *SPB1/PBN1*, *PDC6*, and 17 different solo LTRs. Similar levels were observed for the three gene replacements, but up to eightfold differences were observed at the LTR locations. These results were consistent for both promoters. In combination, the two studies provide 31 characterized sites for gene integration and also demonstrate the importance of evaluating new integration sites prior to utilizing them for gene insertion and expression.

## Regulation of expression level via promoter choice

A wide variety of promoters are available for the control of transcription level in *S. cerevisiae*, including constitutive and inducible promoters of various strengths. In this review, the term ‘constitutive’ will be used broadly as several such promoters have a dependence on glucose level (e.g. transcription drops as glucose is exhausted) and growth phase. The regulated promoters will include those for which a higher level of control is possible or a clear

inducer/repressor exists. Choice of promoter in combination with gene number allows a very wide range of gene expression levels.

A useful analysis tool to facilitate a systematic overview of available promoters and regulatory elements in *S. cerevisiae* is available on SCPD (The Promoter Database of *S. cerevisiae*) by Zhu & Zhang (1999) at <http://rulai.cshl.edu/SCPD/>. The database offers information on approximately 6223 open reading frames and can be used to study the yeast transcription factor binding sites.

## Constitutive promoters and promoter series

Constitutive promoters have been widely used for controlling gene expression in *S. cerevisiae*. These promoters offer simplicity (no inducers or repressors needed) and relatively constant levels of expression. This may be desired for the introduction of new pathways in yeast, particularly if the pathway must be active during cell growth. Constitutive promoters may not be ideal for the synthesis of deleterious gene products or when separation of growth and production are desired. Plasmid stability and copy number can also be an issue, particularly when a strong constitutive promoter is used. Mumberg *et al.* (1995) found that the ratio of  $\beta$ -galactosidase expression between  $2\mu$  and CEN/ARS vectors was 30 for the *ADH1* promoter, but only three for the stronger *GPD* promoter. In a study by Fang *et al.* (2011), when *Rluc* was expressed using a *PGK1* promoter on a *URA3*-marked  $2\mu$ -based vector, copy number was 5.6 (vs. 11.6 for the empty vector).

The most widely used constitutive promoters have often been from the yeast glycolytic pathway. These glucose-dependent promoters include those for phosphoglycerate kinase  $P_{PGK1}$  (Holland & Holland, 1978; Ogden *et al.*, 1986), pyruvate decarboxylase  $P_{PDC1}$  (Kellermann *et al.*, 1986), triosephosphate isomerase  $P_{TPI1}$  (Alber & Kawasaki, 1982), alcohol dehydrogenase I  $P_{ADH1}$  (Hitzeman *et al.*, 1981; Denis *et al.*, 1983), glyceraldehyde-3-phosphate dehydrogenase  $P_{TDH3}$  (*GAP491*) or  $P_{GPD}$  (Holland & Holland, 1980; Bitter & Egan, 1984; McAllister & Holland, 1985), and pyruvate kinase  $P_{PYK1}$  (Nishizawa *et al.*, 1989). The constitutive promoter for translation elongation factor  $P_{TEF1}$  (Gatignol *et al.*, 1990) has also been widely used and has recently been modified by error-prone PCR to provide a group of promoters offering a range of expression levels (Alper *et al.*, 2006a; Nevoigt *et al.*, 2006). Other commonly used native promoters include  $P_{CYC1}$  (Guarente *et al.*, 1984),  $P_{ACT1}$  (Gallwitz & Seidel, 1980),  $P_{MFx1}$  (Brake *et al.*, 1984), and those for hexose transport, for example  $P_{HXT7}$  (Reifenberger *et al.*, 1995; Diderich *et al.*, 1999) – although this list is by no means exhaustive. Constitutive promoters have been suc-



cessfully employed for metabolic pathway engineering in yeast; examples include for ammonia assimilation (Roca *et al.*, 2003), xylose metabolism (Ho *et al.*, 1998; Shao *et al.*, 2009), arabinose metabolism (Becker & Boles, 2003; Wisselink *et al.*, 2007), and malic acid production (Zelle *et al.*, 2008).

Several variants of these promoters have been developed in order to change response to glucose or to allow inducibility. The 1500-bp promoter  $P_{ADHI}$  (Denis *et al.*, 1983) is activated during growth on glucose and is down-regulated following glucose depletion and during ethanol consumption. A short variant of the promoter ( $P_{ADH1s}$ ) with a deletion of 1100 bp in the upstream sequence shifts expression to the early ethanol growth phase with activity increasing into the late ethanol consumption phase (Ruohonen *et al.*, 1991). Restoring 300 bp of the upstream fragment resulted in a 'middle' *ADHI* promoter ( $P_{ADH1m}$ ) that is activated in early exponential growth and maintains activity into the late ethanol consumption phase (Ruohonen *et al.*, 1995). A common approach to confer regulation is to fuse the UAS region from inducible promoters upstream of constitutive promoters and therefore render the hybrid promoter responsive to temperature (Walton & Yarranton, 1989) or inducer metabolites and ions (Bitter & Egan, 1988; Hinnen *et al.*, 1989; Purvis *et al.*, 1991).

Several methods for developing 'synthetic promoter libraries' for the modulation of gene expression are in place (Jensen & Hammer, 1998; Alper *et al.*, 2006a; Hammer *et al.*, 2006; De Mey *et al.*, 2007). For example, mutagenesis of the constitutive  $P_{TEF1}$  promoter sequence (Alper *et al.*, 2006a; Nevoigt *et al.*, 2006) by error-prone PCR resulted in the selection of 11 mutant promoters with strengths ranging between 8% and 120% that of the native  $P_{TEF1}$ . The same technique has been extended to optimizing the regulatory properties of the oxygen-responsive *DANI* promoter in *S. cerevisiae* (Nevoigt *et al.*, 2007) where two mutant promoters with quicker induction and improved expression levels were isolated. Another successful method for synthetic promoter library (SPL) creation described by Jensen & Hammer (1998) is based on saturation mutagenesis of the nucleotide spacer regions flanking key promoter elements. Jeppsson *et al.* (2003) successfully employed this technique to generate a promoter library called YRP, using the RPG (Nieuwint *et al.*, 1989) and CT (Baker, 1991) regulatory elements and the *CUP1* promoter, to control the expression level of glucose-6-phosphate dehydrogenase between 0% and 179% of the wild type.

In contrast to the creation of promoter libraries, the multiple-gene-promoter-shuffling (MGPS) method developed by Lu & Jeffries (2007) can be used to achieve the optimal levels of overexpression of several genes at once,

thereby allowing synergistic control over several rate-limiting steps and metabolic flux.

### Regulated promoters

Regulated promoters enable control over the timing and level of gene expression. They are thus more suitable when expression of genes is desired at a specific stage of cell growth, or to prevent the build-up of toxic pathway intermediates. The use of inducible promoters is limited by the sensitivity of the promoter to the inducer (including the strength and time for response to the inducer or repressor), background levels of expression because of 'leaky' promoters, and the cost of metabolites for induction. In addition, strain response to the inducer may add to the complexity of expression control.

A large variety of regulated native or engineered promoters have been successfully used to control the gene expression in *S. cerevisiae*. The most tightly regulated native promoters are from the galactose-inducible *S. cerevisiae* genes *GAL1*, *GAL7*, and *GAL10* (Douglas & Hawthorne, 1964; Bassel & Mortimer, 1971). These promoters are induced approximately 1000-fold in the presence of galactose and strongly repressed in the presence of glucose (Adams, 1972). While several genes have been identified to be involved in the regulation of these promoters, the *GAL4* gene encoding a transactivator and *GAL80* gene encoding the repressor for Gal4 control the central regulation mechanism along with the *GAL* upstream activation site (UAS) (Johnston *et al.*, 1994). Modifications in the *GAL1* and *GAL10* promoters (Li *et al.*, 2008) and engineering key enzymes involved in galactose catabolism and transport (Hawkins & Smolke, 2006, 2008) have provided tunable control to galactose-driven expression under these promoters. The *GAL*-regulated promoters have been widely used in yeast metabolic pathway engineering, including for artemisinic acid synthesis (Ro *et al.*, 2006), increased acetyl-CoA synthesis for isoprenoid production (Shiba *et al.*, 2007), expression of the bacterial isoprenoid pathway in *S. cerevisiae* (Maury *et al.*, 2008), and n-butanol synthesis (Steen *et al.*, 2008).

The *S. cerevisiae* *CUP1* promoter contains four metal regulatory elements and controls the expression of copper metallothionein in yeast.  $P_{CUP1}$  can be induced around 20-fold in the presence of  $\text{Cu}^{2+}$  (Etcheverry, 1990). The activation of this promoter is independent of other culture parameters, with ionic concentration being the limiting factor based on the copper resistance of the host strain (Macreadie *et al.*, 1991; Hottiger *et al.*, 1995; Labbe & Thiele, 1999). Jeppsson *et al.* (2003) achieved a range of expression levels driven by the *CUP1* promoter using the aforementioned SPL technique. The *CUP1* promoter has also been used to express pathway genes in yeast, for

example, for the synthesis of 1,2-propanediol (Lee & Da Silva, 2006).

The *S. cerevisiae* *ADH2* promoter (573 bp) is tightly regulated by glucose repression, with over a 100-fold repression in the presence of glucose (Price *et al.*, 1990). This promoter requires no inducer, as expression begins as glucose is depleted and ethanol consumption begins. Two UASs contained in a 260-bp region upstream of the initiation site have been identified to render a fully active and regulatable  $P_{ADH2}$ . Cultivation in complex medium is required for high-level expression from  $P_{ADH2}$  (Price *et al.*, 1990; Lee & Da Silva, 2005); however, because the promoter is off until glucose levels fall, plasmid stability is usually not a major issue. The promoter has been successfully employed for several metabolic engineering studies, including the synthesis of polyketides (Kealey *et al.*, 1998; Mutka *et al.*, 2006; Lee *et al.*, 2009) and triacetic acid lactone (Xie *et al.*, 2006) in *S. cerevisiae*.

Several other inducible *S. cerevisiae* promoters have been employed in yeast including  $P_{PHO5}$ ,  $P_{MET25}$ , and  $P_{MET3}$ . The *PHO5* promoter of the acid phosphatase gene is regulated by inorganic phosphate ( $P_i$ ) in the medium with approximately 200-fold repression in the presence of phosphate. The *MET25* gene (Sangsoda *et al.*, 1985; Kerjan *et al.*, 1986) encodes *O*-acetyl homoserine sulfhydrylase, and the *MET3* gene encodes ATP sulfurylase (Cherest *et al.*, 1987). Both promoters are repressed in the presence of methionine or *S*-adenosylmethionine.

One of the most useful heterologous promoter systems in yeast utilizes the bacterial tetracycline operator (*tetO*) and hybrid transactivator, based on the expression system developed for mammalian cells (Gossen & Bujard, 1992; Gossen *et al.*, 1995) and shown to be active in yeast by Dingermann *et al.* (1992). Gari *et al.* (1997) constructed a set of expression vectors with a hybrid promoter system *tetO-CYC1* consisting of seven or fewer *tetO* boxes, the *S. cerevisiae* *CYC1* TATA region, and tTA activator derivatives (*tetR*- $\lambda$ CI-VP16), with a 1000-fold induction ratio in the absence of tetracycline. The strongest system thus developed has expression levels comparable with that of the *GAL1* promoter. Another study by Murphy *et al.* (2007) demonstrates the applicability of the *tetO*<sub>2</sub> operator in the design of combinatorial promoters. Belli *et al.* (1998) developed a *tetR* regulator-based expression system that allows for tight *tetO*-driven expression in a dual (tetracycline-repressible and inducible) manner.

### Comparisons of promoter strength

A number of studies have compared the strengths of the common constitutive and regulated promoters using gene cassettes either carried on plasmids or integrated into the genome. The latter comparisons are particularly useful as

they control for both gene stability and copy number. While the comparisons will depend on the specific gene expressed from the promoters, useful trends have been observed and can be used as a guide for fine-tuning gene expression in metabolic engineering applications. These studies (including the rank order of promoter strengths determined in each study) are summarized in Table 2.

Several promoter comparisons evaluated the expression levels from plasmid-based constructs. Mumberg *et al.* (1995) compared four constitutive promoters of varying strengths ( $P_{ADH1}$ ,  $P_{CYC1}$ ,  $P_{TDH3}$ , and  $P_{TEF1}$ ) on both *CEN/ARS*- and  $2\mu$ -based plasmids using *E. coli lacZ* as the reporter gene during growth in glucose medium. Only a 5-fold difference was observed in the expression level of the strongest promoter ( $P_{TDH3}$ ) for the  $2\mu$  plasmid as compared to the *CEN/ARS* plasmid. The truncated version of the  $P_{CYC1}$  promoter used in this study lacks the *UAS2* sequence and thereby showed weakest activity. Monfort *et al.* (1999) compared the strength of the four constitutive promoters  $P_{PGK1}$ ,  $P_{ADH1}$ ,  $P_{ADH1m}$ , and  $P_{TDH1}$  (Hadfield *et al.*, 1993) using *Geotrichum candidum* lipase2 as a reporter from  $2\mu$ -based plasmids using the constitutive  $P_{ACT1}$  as a control. The results of their studies considered distinct phases of cell growth during glucose consumption, early ethanol consumption, and late ethanol consumption (where only  $P_{ADH1s}$  remained active). Fang *et al.* (2011) did a similar comparison using three relatively strong constitutive promoters  $P_{PGK1}$ ,  $P_{TEF1}$ , and  $P_{HXT7-391}$  on the pXP vector series using Renilla luciferase as the reporter protein during exponential growth in glucose medium. Further, the authors identified an inverse relationship between the promoter strength and plasmid copy number.

Additional studies evaluated the promoters for the expression of secretory proteins. Park *et al.* (1993) compared constitutive promoters  $P_{PGK1}$  and  $P_{SUC2}$  and galactose-inducible promoter  $P_{GAL7}$  for the production of secreted  $\alpha$ -amylase in fed-batch cultures. Cartwright *et al.* (1994) compared the levels of secreted SP $\beta$ -lactamase (SP $\beta$ la) during late exponential growth on various carbon sources when five different promoters ( $P_{GAL1}$ ,  $P_{GAL10}$ ,  $P_{PGK1}$ ,  $P_{PHO5}$ , and  $P_{CUP1}$ ) were used. The difference in expression between YEp-SPB and YCp-SPB was found to be 5-fold for the *GAL* promoters, 13-fold for  $P_{PGK1}$ , 15-fold for  $P_{PHO5}$ , and 10-fold for  $P_{CUP1}$ .

A key comparison of seven constitutive promoters ( $P_{PGK1}$ ,  $P_{TEF1}$ ,  $P_{TDH3}$ ,  $P_{TPI1}$ ,  $P_{PYK1}$ ,  $P_{ADH1}$ , and  $P_{HXT7}$ ) was made by Partow *et al.* (2010) using *lacZ* as the reporter gene. In this study, the expression cassettes were integrated at a single copy into the *URA3* locus of the yeast genome ensuring both stability and constant copy number. Similar to the plasmid-based study by Monfort *et al.* (1999), the strengths of the promoters were assessed at

**Table 2.** Studies comparing the level of expression with various promoters in *Saccharomyces cerevisiae*

Promoter sources	Expression systems	Reporters	Promoter strengths	Carbon source; inducer	References
ADH1, CYC1, TDH3, TEF1	P-C/A*, P-2 $\mu$ <sup>†</sup>	<i>E. coli</i> $\beta$ -galactosidase	$P_{TDH3} > P_{TEF1} >> P_{ADH1} >>> P_{CYC1}$	Glucose	Mumberg et al. (1995)
PGK1, ADH1s, ADH1m, TDH1, ACT1	P-2 $\mu$	<i>G. candidum</i> lipase2	Glucose consumption phase: $P_{PGK1} > P_{ADH1m} \sim P_{ACT1} >> P_{ADH1s} \sim P_{TDH1}$ Early ethanol consumption phase: $P_{ACT1} > P_{PGK1} \sim P_{ADH1s} > P_{TDH3} > P_{ADH1m}$ Late ethanol consumption phase: Only $P_{ADH1s}$ active	Liquid molasses <sup>‡</sup>	Monfort et al. (1999)
PGK1, TEF1, HXT7-391	P-C/A; P-2 $\mu$	Renilla luciferase	Glucose consumption phase: P-C/A: $P_{TEF1} \sim P_{PGK1} >> P_{HXT7-391}$ P-2 $\mu$ : $P_{TEF1} \sim P_{PGK1} > P_{HXT7-391}$	2% glucose	Fang et al. (2011)
PGK1, SUC2, GAL7	P-2 $\mu$	Mouse $\alpha$ -amylase	Late exponential growth phase: $P_{PGK1} > P_{SUC2} > P_{GAL7}$	Initial carbon source (fed-batch): 1% glucose or 3% galactose	Park et al. (1993)
PGK1, GAL1, GAL10, PHO5, CUP1	P-C/A; P-2 $\mu$	<i>E. coli</i> $\beta$ -lactamase	Late exponential growth phase: $P_{CUP1(e)} > P_{GAL1} \sim P_{GAL10} (b) >> P_{GAL1} \sim P_{GAL10} (c) > P_{PGK1(a)} > P_{PHO5(d)} \sim P_{CUP1(a)}$	(a) 2% glucose (b) 2% galactose (c) 2% galactose + 2% sucrose (d) 2% glucose + 0.22 mM PI (e) 2% glucose + 0.1 mM Cu <sup>2+</sup>	Cartwright et al. (1994)
PGK1, ADH2, GAL1, CUP1	P-C/A; P-2 $\mu$	<i>E. coli</i> $\beta$ -galactosidase	Early exponential growth phase: P-C/A: $P_{GAL1(b)} > P_{PGK1(a)} \sim P_{CUP1(c)} >> P_{ADH2(a)}$ P-2 $\mu$ : $P_{GAL1(b)} > P_{PGK1(a)} > P_{CUP1(c)} >> P_{ADH2(a)}$	(a) 1% glucose (complex) (b) 1% galactose (c) 1% glucose + 1 mM Cu <sup>2+</sup>	M.W.Y. Shen, F. Fang, S. Sandmeyer & N.A. Da Silva (unpublished data)
PGK1, TEF1, TDH3, TPI, PYK1, ADH1, HXT7	Integration (1 copy; URA3 locus)	<i>E. coli</i> $\beta$ -galactosidase	Stationary growth phase: P-C/A: $P_{ADH2} > P_{CUP1} > P_{GAL1} \sim P_{PGK1}$ P-2 $\mu$ : $P_{ADH2} \sim P_{CUP1} > P_{PGK1} > P_{GAL1}$	2% glucose	Partow et al. (2010)
PGK1, ENO2, TPI, PYK1, PDC1, ADH1, HXT7	Integration (1 copy; URA3 locus)	<i>E. coli</i> $\beta$ -galactosidase	Glucose consumption phase: $P_{TEF1} \sim P_{PGK1} \sim P_{TDH3} > P_{TPI} \sim P_{PYK1} > P_{ADH1} > P_{HXT7}$ Early ethanol consumption phase: $P_{HXT7} \sim P_{TEF1} > P_{PGK1} > P_{TPI} \sim P_{TDH3} > P_{PYK1} \sim P_{ADH1}$	1% ethanol	Hauf et al. (2000)
HXT7, TEF1, GAL1, GAL10	Integration (1 copy; URA3 locus)	<i>E. coli</i> $\beta$ -galactosidase	Ethanol growth phase: $P_{HXT7} >> P_{PDC1} \sim P_{TPI} \sim P_{PGK1} > P_{ADH1} > P_{PYK1} > P_{ENO2}$	Initial carbon source (fed-batch): 2% glucose or 2% galactose	Partow et al. (2010)
ADH2, GAL1, CUP1	Integration (1 copy; $\delta$ loci)	<i>E. coli</i> $\beta$ -galactosidase	Stationary growth phase: $P_{ADH2} (a) > P_{CUP1(c)} > P_{GAL1(b)}$	(a) 1% glucose (complex) (b) 1% galactose (complex) (c) 1% glucose + 1 mM Cu <sup>2+</sup> (complex)	Lee & Da Silva (2005)
PGK1, ADH2, GAL1, CUP1	Integration (1 copy; Ty1 locus)	<i>E. coli</i> $\beta$ -galactosidase	Early exponential growth phase: $P_{GAL1(b)} >> P_{PGK1(a)} \sim P_{CUP1(c)} >> P_{ADH2(a)}$ Stationary growth phase: $P_{ADH2} > P_{GAL1} \sim P_{CUP1} \sim P_{PGK1}$	(a) 1% glucose (complex) (b) 1% galactose (complex) (c) 1% glucose + 1 mM Cu <sup>2+</sup> (complex)	M.W.Y. Shen, F. Fang, S. Sandmeyer & N.A. Da Silva (unpublished data)

\*Plasmid with CENARS origin of replication. †Plasmid with 2 $\mu$  origin of replication. ‡Randez-Gil et al. (1995).

different stages of growth (glucose or ethanol consumption). In an earlier study by Hauf *et al.* (2000), five of these promoters as well as  $P_{ENO2}$  and  $P_{PDC1}$  were compared during growth on ethanol using single-copy integrated cassettes. Partow *et al.* (2010) also compared the constitutive promoters with the  $GAL1$  and  $GAL10$  promoters in fed-batch or continuous culture. The result for the  $GAL1$ - $GAL10$  promoter comparison,  $P_{GAL10} > P_{GAL1}$  (Table 2), is in contrast to earlier studies comparing these promoters (Yocum *et al.*, 1984; West *et al.*, 1987; Da Silva & Bailey, 1991; Cartwright *et al.*, 1994). The authors note that the difference in reporter gene cloning sites used in this study may affect the translational efficiency. This observation is supported by Crook *et al.* (2011); the authors show that the distance between the promoter and gene because of the MCS can result in significant mRNA secondary structure in the 5' untranslated region, thereby affecting translational efficiency. Further differences in protein translation were observed with the length, codon optimization, or reporter gene ( $\gamma$ ECitrine, eGFP, and LacZ) used.

In another study evaluating the strength of inducible promoters, Lee & Da Silva (2005) compared the expression levels under  $P_{ADH2}$ ,  $P_{GAL1}$ , and  $P_{CUP1}$  for a single integrated copy of the *lacZ* gene into the delta sequences and also looked at induction at various culture times. Early induction was better suited to  $P_{CUP1}$  and late induction for  $P_{GAL1}$ . In recent studies in our laboratory, expression from  $P_{PGK1}$ ,  $P_{ADH2}$ ,  $P_{GAL1}$ , and  $P_{CUP1}$  was compared on both plasmids (CEN/ARS and 2 $\mu$ -based) and from single integrated copies (each replacing the same Ty1 locus in the chromosomes) using the extended pXP vector series described earlier (M.W.Y. Shen, F. Fang, S. Sandmeyer & N.A. Da Silva, unpublished data). In these studies, identical sequences separated the native promoter from the start codon of *lacZ*, and  $\beta$ -galactoside levels were compared at 12-h intervals up to 48 h.

A summary of all promoter comparison studies described is presented in Table 2. Much greater detail on the studies can be found in the listed references.

## Conclusions

Metabolic engineering requires the introduction of multi-enzyme pathways and precise control over the expression of the associated genes. This MiniReview has focused on two key elements in regulating enzyme synthesis in *S. cerevisiae*, both at the pretranslation level: (1) gene introduction/control of gene number and (2) choice of promoter (both strength and regulation). A large number of approaches are available to modulate the expression at these points, and the combination of copy number control and promoter choice allows a wide range of expression levels for pathway genes.

We have also included several examples of the use of vectors, chromosomal integration, constitutive promoters, and inducible promoters for pathway engineering in *S. cerevisiae*. This is only a small sampling of the numerous studies described in the literature and does not include many studies that combine multiple different approaches. For example, constitutive and inducible promoters have been employed together to control gene expression at specific times during the culture. In addition, combinations of multiple promoters, gene knockouts, gene integration, and plasmids have been used to introduce complex pathways. One classic example is the engineering of yeast to produce hydrocortisone from glucose (Szczepara *et al.*, 2003).

There are several other important factors influencing protein levels in *S. cerevisiae* that have not been addressed in this MiniReview, including additional factors at the transcription level and those modulating translation and post-translational processing. These range from transcription factor engineering (e.g. Alper *et al.*, 2006b) and mRNA stability (e.g. the RNA control modules described by Babiskin & Smolke, 2011) to translation efficiency and protein stability. All provide additional levels of regulation that, with gene number and promoter choice, contribute to the fine tuning of enzyme levels needed for metabolic pathway engineering.

## References

- Adams BG (1972) Induction of galactokinase in *Saccharomyces cerevisiae*- kinetics of induction and glucose effects. *J Bacteriol* **111**: 308–315.
- Alani E, Cao L & Kleckner N (1987) A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. *Genetics* **116**: 541–545.
- Alber T & Kawasaki G (1982) Nucleotide sequence of the triose phosphate isomerase gene of *Saccharomyces cerevisiae*. *J Mol Appl Genet* **1**: 419–434.
- Alberti S, Gitler AD & Lindquist S (2007) A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* **24**: 913–919.
- Alper H, Fischer C, Nevoigt E & Stephanopoulos G (2006a) Tuning genetic control through promoter engineering. *P Natl Acad Sci USA* **102**: 12678–12683.
- Alper H, Moxley J, Nevoigt E, Fink GR & Stephanopoulos G (2006b) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**: 1565–1568.
- Babiskin AH & Smolke CD (2011) A synthetic library of RNA control modules for predictable tuning of gene expression in yeast. *Mol Syst Biol* **7**: 471.
- Baker HV (1991) *GCR1* of *Saccharomyces cerevisiae* encodes a DNA-binding protein whose binding is abolished by

- mutations in the CTCC sequence motif. *P Natl Acad Sci USA* **88**: 9443–9447.
- Bassel J & Mortimer R (1971) Genetic order of the galactose structural genes in *Saccharomyces cerevisiae*. *J Bacteriol* **108**: 179–183.
- Becker J & Boles E (2003) A modified *Saccharomyces cerevisiae* strain that consumes L-arabinose and produces ethanol. *Appl Environ Microbiol* **69**: 4144–4150.
- Beggs JD (1978) Transformation of yeast by a replicating hybrid plasmid. *Nature* **275**: 104–109.
- Belli G, Gari E, Piedrafita L, Aldea N & Herrero E (1998) An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res* **26**: 942–947.
- Bitter GA & Egan KM (1984) Expression of heterologous genes in *Saccharomyces cerevisiae* from vector utilizing the glyceraldehydes-3-phosphate dehydrogenase gene promoter. *Gene* **32**: 263–274.
- Bitter GA & Egan KM (1988) Expression of interferon-gamma from hybrid yeast *GPD* promoters containing upstream regulatory sequences from the *GALI-GAL10* intergenic region. *Gene* **69**: 193–207.
- Boeke JD (1989) Transposable elements in *Saccharomyces cerevisiae*. *Mobile DNA* (Berg DE & Howe MM, eds), pp. 335–374. American Society of Microbiology, Washington, DC.
- Boeke JD & Sandmeyer SB (1991) Yeast transposable elements. *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis, and Energetics* (Broach J, Jones E & Pringle J, eds), pp. 193–261. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Boeke JD, Lacroute F & Fink GR (1984) A positive selection for mutants lacking Orotidine-5'-phosphate decarboxylase activity in yeast-5-fluoro-orotic acid resistance. *Mol Gen Genet* **197**: 345–346.
- Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH & Falkow S (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**: 25–113.
- Brachmann CB, Davies A, Cost GJ, Caputo E, Li JC, Hieter P & Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**: 115–132.
- Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P & Barr PJ (1984) Alpha-factor directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *P Natl Acad Sci USA* **81**: 4642–4646.
- Burke DT, Carle GF & Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806–812.
- Carlson R & Srien F (2006) Effects of recombinant precursor pathway variations on poly (R)-3-hydroxybutyrate synthesis in *Saccharomyces cerevisiae*. *J Biotechnol* **124**: 561–573.
- Cartwright CP, Li Y, Zhu YS, Kang YS & Tipper DJ (1994) Use of beta-lactamase as a secreted reporter of promoter function in yeast. *Yeast* **10**: 497–508.
- Cherest H, Kerjan P & Surdinkerjan Y (1987) The *Saccharomyces cerevisiae* *MET3* gene-nucleotide-sequence and relationship of the 5' noncoding region to that of *ME25*. *Mol Gen Genet* **210**: 307–313.
- Ching CG (2005) Sequential cloned gene integration: enhancements in *Saccharomyces cerevisiae*, extension to polyploid yeast strains, and application to polyketide. PhD Thesis, University of California, Irvine, Irvine, CA.
- Cho KM, Yoo YJ & Kang HS (1999) delta-Integration of endo/exo-glucanase and beta-glucosidase genes into the yeast chromosomes for direct conversion of cellulose to ethanol. *Enzyme Microb Technol* **25**: 23–30.
- Christianson TW, Sikorski RS, Dante M, Shero JH & Hieter P (1992) Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**: 119–122.
- Clarke L & Carbon J (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**: 504–509.
- Crook NC, Freeman ES & Alper HS (2011) Re-engineering multicloning sites for function and convenience. *Nucleic Acids Res* **39**: e92.
- Da Silva NA & Bailey JE (1991) Influence of plasmid origin and promoter strength in fermentations of recombinant yeast. *Biotechnol Bioeng* **37**: 318–324.
- De Mey M, Maertens J, Lequeux GJ, Soetaert WK & Vandamme EJ (2007) Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. *BMC Biotechnol* **7**: 34.
- Denis CL, Ferguson J & Young ET (1983) Messenger-RNA levels for the fermentative alcohol-dehydrogenase of *Saccharomyces cerevisiae* decrease upon growth on a nonfermentable carbon source. *J Biol Chem* **258**: 1165–1171.
- Diderich JA, Teusink B, Valkier J, Anjos J, Spencer-Martins I, van Dam K & Walsh MC (1999) Strategies to determine the extent of control exerted by glucose transport on glycolytic flux in the yeast *Saccharomyces bayanus*. *Microbiology* **145**: 3447–3454.
- Dingermann T, Frankstoll U, Werner H, Wissmann A, Hillen W, Jacquet M & Marschalek R (1992) RNA polymerase-III catalyzed transcription can be regulated in *Saccharomyces cerevisiae* by the bacterial tetracycline repressor operator system. *EMBO J* **11**: 1487–1492.
- Douglas HC & Hawthorne DC (1964) Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*. *Genetics* **49**: 837–844.
- Dujon B (1996) The yeast genome project: what did we learn? *Trends Genet* **12**: 263–270.
- Erdenez N, Mortensen UH & Rothstein R (1997) Cloning-free PCR-based allele replacement methods. *Genome Res* **7**: 1174–1183.
- Erhart E & Hollenberg CP (1983) The presence of a defective *LEU2* gene on 2 $\mu$  DNA recombinant plasmids of

- Saccharomyces cerevisiae* is responsible for curing and high copy number. *J Bacteriol* **156**: 625–635.
- Etcheverry T (1990) Induced expression of using yeast copper-metallothionein promoter. *Methods Enzymol* **185**: 319–329.
- Fang F, Salmon K, Shen MWY, Aeling KA, Ito E, Irwin B, Tran UPC, Hatfield GW, Da Silva NA & Sandmeyer S (2011) A vector set for systematic metabolic engineering in *Saccharomyces cerevisiae*. *Yeast* **28**: 123–136.
- Flagfeldt DB, Siewers V, Huang L & Nielsen J (2009) Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* **26**: 545–551.
- Fujii T, Kondo K, Shimizu F, Sone H, Tanaka J-I & Inoue T (1990) Application of a ribosomal DNA integration vector in the construction of a brewer's yeast having alpha-acetolactate decarboxylase activity. *Appl Environ Microbiol* **56**: 997–1003.
- Funk M, Niedenthal R, Mumberg D, Brinkmann K, Ronicke V & Henkel T (2002) Vector systems for heterologous expression of proteins in *Saccharomyces cerevisiae*. Guide to Yeast Genetics and Molecular and Cell Biology, Pt B *Methods Enzymol* **350**: 248–257.
- Futcher B & Carbon J (1986) Toxic effects of excess cloned centromeres. *Mol Cell Biol* **6**: 2213–2222.
- Futcher AB & Cox BS (1983) Maintenance of the 2 $\mu$  circle plasmid in populations of *Saccharomyces cerevisiae*. *J Bacteriol* **154**: 612–622.
- Gallwitz D & Seidel R (1980) Molecular cloning of the actin gene from yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res* **8**: 1043–1059.
- Gari E, Piedrafitra L, Aldea M & Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* **13**: 837–848.
- Gatignol A, Dassain M & Tiraby G (1990) Cloning of *Saccharomyces cerevisiae* using a probe vector based on Phleomycin resistance. *Gene* **91**: 35–41.
- Geiser JR (2005) Recombinational cloning vectors for regulated expression in *Saccharomyces cerevisiae*. *BioTechniques* **38**: 378–382.
- Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO & Hutchison CA III (2008) On-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *P Natl Acad Sci USA* **105**: 20404–20409.
- Gietz RD & Sugino A (1988) New yeast-*E.coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking 6-base pair restriction sites. *Gene* **74**: 527–534.
- Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274**: 546, 563–567.
- Goldstein AL, Pan X & McCusker JH (1999) Heterologous URA3MX cassettes for gene replacement in *Saccharomyces cerevisiae*. *Yeast* **15**: 507–511.
- Gossen M & Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-response promoters. *P Natl Acad Sci USA* **89**: 5547–5551.
- Gossen M, Freundlieb S, Bender G, Muller G, Hillen W & Bujard H (1995) Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**: 1766–1769.
- Guarente L, Lalonde B, Gifford P & Alani E (1984) Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* **36**: 503–511.
- Gueldener U, Heinisch J, Koehler GJ, Voss D & Hegemann JH (2002) A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* **30**: e23.
- Guldener U, Heck S, Fiedler T, Beinhauer J & Hegemann JH (1996) A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* **24**: 2519–2524.
- Hadfield C, Raina KK, Shashi-Menon K & Mount RC (1993) The expression and performance of cloned genes of yeasts. *Mycol Res* **97**: 897–944.
- Hammer K, Mijakovic I & Jensen PR (2006) Synthetic promoter libraries- tuning of gene expression. *Trends Biotechnol* **24**: 53–55.
- Hauf J, Zimmermann FK & Muller S (2000) Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast *Saccharomyces cerevisiae*. *Enzyme Microb Technol* **26**: 688–698.
- Hawkins KM & Smolke CD (2006) The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in *Saccharomyces cerevisiae*. *J Biol Chem* **281**: 13485–13492.
- Hawkins KM & Smolke CD (2008) Production of benzyloisoquinoline alkaloids in *Saccharomyces cerevisiae*. *Nat Chem Biol* **4**: 564–573.
- Hinnen A, Meyhack B & Heim J (1989) Heterologous gene expression in yeast. *Biotechnology* **13**: 193–213.
- Hitzeman RA, Hagie FE, Levine HL, Goeddel DV, Ammerer G & Hall BD (1981) Expression of a human gene for interferon in yeast. *Nature* **293**: 717–722.
- Ho NWY, Chen Z & Brainard AP (1998) Genetically engineered *Saccharomyces* yeast capable of effective cofermentation of glucose and xylose. *Appl Environ Microbiol* **64**: 1852–1859.
- Holland MJ & Holland JP (1978) Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase, and phosphoglycerate kinase. *Biochemistry* **17**: 4900–4907.
- Holland JP & Holland MJ (1980) Structural comparison of 2 non-tandemly repeated yeast glyceraldehydes-3-phosphate dehydrogenase genes. *J Biol Chem* **255**: 2596–2605.
- Hottiger T, Kuhla J, Pohlig G, Fürst P, Spielmann A, Garn M, Haemmerli S & Heim J (1995) 2 $\mu$  vectors containing the *Saccharomyces cerevisiae* metallothionein gene as a selectable marker: excellent stability in complex media, and high-level expression of a recombinant protein from a *CUP1*-promoter-controlled expression cassette in *cis*. *Yeast* **11**: 1–14.
- Jansen G, Wu CL, Schade B, Thomas DY & Whiteway M (2005) Drag&Drop cloning in yeast. *Gene* **344**: 43–51.

- Jensen PR & Hammer K (1998) Artificial promoters for metabolic optimization. *Biotechnol Bioeng* **58**: 191–195.
- Jeppsson M, Johansson B, Jensen PR, Hahn-Hagerdal B & Gorwa-Grauslund MF (2003) The level of glucose-6-phosphate dehydrogenase activity strongly influences xylose fermentation and inhibitor sensitivity in recombinant *Saccharomyces cerevisiae* strains. *Yeast* **20**: 1263–1272.
- Jiang HX, Wood KV & Morgan JA (2005) Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**: 2962–2969.
- Jimenez A & Davies J (1980) Expression of a transposable antibiotic resistance element in *Saccharomyces*. *Nature* **287**: 869–871.
- Johansson B & Hahn-Hagerdal B (2002) Overproduction of pentose phosphate pathway enzymes using a new CRE-*loxP* expression vector for repeated genomic integration in *Saccharomyces cerevisiae*. *Yeast* **19**: 225–231.
- Johnston M, Flick JS & Pexton T (1994) Multiple mechanisms provide rapid and stringent glucose repression of *GAL* gene expression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **14**: 3834–3841.
- Kealey JT, Liu L, Santi DV, Betlach MC & Barr PJ (1998) Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *P Natl Acad Sci USA* **95**: 505–509.
- Kellermann E, Seeboth PG & Hollenberg CP (1986) Analysis of the primary structure and promoter function of a pyruvate decarboxylase gene (*PDC1*) from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **14**: 8963–8977.
- Kerjan P, Cherest H & Surdinkerjan Y (1986) *Nucleotide sequence of the Saccharomyces cerevisiae MET25 gene*. *Nucleic Acids Res* **14**: 7861–7871.
- Kikuchi Y (1983) Yeast plasmid requires a cis-acting locus and two plasmid proteins for its stable maintenance. *Cell* **35**: 487–493.
- Kim JM, Vanguri S, Boeke JD, Gabriel A & Voytas DF (1998) Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res* **8**: 464–478.
- Kim SK, Noh YH, Koo JR, DaSilva NA & Yun HS (2011) Effects of expression of *lcb1/lcb2* and *lac1/lag1* genes on the biosynthesis of ceramides. *Biotechnol Bioprocess Eng* **16**: 1–6.
- Kingsman SM, Kingsman AJ, Dobson MJ, Mellor J & Roberts NA (1985) Heterologous gene expression in *Saccharomyces cerevisiae*. *Biotechnol Genet Eng Rev* **3**: 377–416.
- Labbe S & Thiele DJ (1999) Copper ion inducible and repressible promoter systems in yeast. *Expression of Recombinant Genes in Eukaryotic Systems (Methods in Enzymology)*, Vol **306** (Abelson J, Simon M, Glorioso J & Schmidt M, eds), pp. 145–153. Academic Press, New York.
- Langle-Rouault F & Jacobs E (1995) A method for performing precise alterations in the yeast genome using a recyclable selectable marker. *Nucleic Acids Res* **23**: 3079–3081.
- Lee FWF & Da Silva NA (1997a) Improved efficiency and stability of multiple cloned gene insertions at the delta sequences of *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **48**: 339–345.
- Lee FWF & Da Silva NA (1997b) Sequential delta-integration for the regulated insertion of cloned genes in *Saccharomyces cerevisiae*. *Biotechnol Prog* **13**: 368–373.
- Lee KM & Da Silva NA (2005) Evaluation of the *Saccharomyces cerevisiae ADH2* promoter for protein synthesis. *Yeast* **22**: 431–440.
- Lee W & Da Silva NA (2006) Application of sequential integration for metabolic engineering of 1,2-propanediol production in yeast. *Metab Eng* **8**: 58–65.
- Lee KKM, Da Silva NA & Kealey JT (2009) Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Anal Biochem* **394**: 75–80.
- Leonard E, Yan YJ, Lim KH & Koffas MAG (2005) Investigation of two distinct flavone synthases for plant-specific flavone biosynthesis in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**: 8241–8248.
- Li AM, Liu ZS, Li QX, Yu L, Wang DC & Deng XM (2008) Construction and characterization of bidirectional expression vectors in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **8**: 6–9.
- Louison G, Vidal A, Findeli A, Roitsch C, Balloul JM & Lemoine Y (1989) High level of expression of a protective antigen of schistosomes in *Saccharomyces cerevisiae*. *Yeast* **5**: 497–507.
- Lopes TS, Klootwijk J, Veenstra AE, Vanderaar PC, Vanheerikhuizen H, Raue HA & Planta RJ (1989) High-copy-number integration into the ribosomal DNA of *Saccharomyces cerevisiae*: a new vector for high-level expression. *Gene* **79**: 199–206.
- Lopes TS, Hakkaart G, Koerts BL, Raue HA & Planta RJ (1991) Mechanism of high-copy-number integration of pMIRY-type vectors into the ribosomal DNA of *Saccharomyces cerevisiae*. *Gene* **105**: 83–90.
- Lorenz MC, Muir RS, Lim E, McElver J, Weber SC & Heitman J (1995) Gene disruption with PCR products in *Saccharomyces cerevisiae*. *Gene* **158**: 113–117.
- Lu CF & Jeffries T (2007) Shuffling of promoters for multiple genes to optimize xylose fermentation in an engineered *Saccharomyces cerevisiae* strain. *Appl Environ Microbiol* **73**: 6072–6077.
- Ma H, Kunes S, Schatz PJ & Botstein D (1987) Plasmid construction by homologous recombination in yeast. *Gene* **58**: 201–216.
- Macreadie IG, Horaitis O, Verkuylen AJ & Savin KW (1991) Improved shuttle vectors for cloning and high-level Cu<sup>2+</sup>-mediated expression of foreign genes in yeast. *Gene* **104**: 107–111.
- Manivasakam P, Weber SC, McElver J & Schiestl RH (1995) Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **23**: 2799–2800.

- Maury J, Asadollahi MA, Moller K, Schalk M, Clark A, Formenti LR & Nielsen J (2008) Reconstruction of a bacterial isoprenoid biosynthetic pathway in *Saccharomyces cerevisiae*. *FEBS Lett* **582**: 4032–4038.
- McAllister L & Holland MJ (1985) Differential expression of the three yeast glyceraldehyde-3-phosphate dehydrogenase genes. *J Biol Chem* **260**: 15019–15027.
- Mead DJ, Gardner DCJ & Oliver SG (1986) Autoregulation of 2 $\mu$  circle gene expression provides a model for maintenance of stable plasmid copy levels. *Mol Gen Genet* **205**: 417–421.
- Miller CA, Martinat MA & Hyman LE (1998) Assessment of aryl hydrocarbon receptor complex interactions using pBEVY plasmids: expression vectors with bi-directional promoters for use in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **26**: 3577–3583.
- Monfort A, Finger S, Sanz P & Prieto JA (1999) Evaluation of different promoters for the efficient production of heterologous proteins in baker's yeast. *Biotechnol Lett* **21**: 225–229.
- Moore PA, Bettany AJE & Brown AJP (1990) Multiple copies of the pyruvate kinase gene affect yeast cell growth. *J Gen Microbiol* **136**: 2359–2366.
- Mumberg D, Muller R & Funk M (1994) Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* **22**: 5767–5768.
- Mumberg D, Muller R & Funk M (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**: 119–122.
- Murphy KF, Balazsi G & Collins JJ (2007) Combinatorial promoter design for engineering noisy gene expression. *P Natl Acad Sci USA* **104**: 12726–12731.
- Murray AW & Szostak JW (1983a) Pedigree analysis of plasmid segregation in yeast. *Cell* **34**: 961–970.
- Murray AW & Szostak JW (1983b) Construction of artificial chromosomes in yeast. *Nature* **305**: 189–193.
- Mutka SC, Bondi SM, Carney JR, Da Silva NA & Kealey JT (2006) Metabolic pathway engineering for complex polyketide biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **6**: 40–47.
- Naesby M, Nielsen SVS, Nielsen CAF *et al.* (2009) Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microb Cell Fact* **8**: 45.
- Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U & Stephanopoulos G (2006) Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **72**: 5266–5273.
- Nevoigt E, Fischer C, Mucha O, Matthaus F, Stahl U & Stephanopoulos G (2007) Engineering promoter regulation. *Biotechnol Bioeng* **96**: 550–558.
- Nieuwint RTM, Mager WH, Maurer KCT & Planta RJ (1989) Mutational analysis of the upstream activation site of yeast ribosomal protein genes. *Curr Genet* **15**: 247–251.
- Nishizawa M, Araki R & Teranishi Y (1989) Identification of an upstream activating sequence and an upstream repressible sequence of the pyruvate kinase gene of the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* **9**: 442–451.
- Ogden JE, Stanway C, Kim S, Mellor J, Kingsman AJ & Kingsman SM (1986) Efficient expression of the *Saccharomyces cerevisiae* PGK gene depends on an upstream activation sequence but does not require TATA sequences. *Mol Cell Biol* **6**: 4335–4343.
- Oldenburg KR, Vo KT, Michaelis S & Paddon C (1997) Recombination-mediated PCR-directed plasmid construction *in vivo* in yeast. *Nucleic Acids Res* **25**: 451–452.
- Parekh RN & Wittrup KD (1997) Expression level tuning for optimal heterologous protein secretion in *Saccharomyces cerevisiae*. *Biotechnol Prog* **13**: 117–122.
- Parekh RN, Shaw MR & Wittrup KD (1996) An integrating vector for tunable, high copy, stable integration into the dispersed Ty delta sites of *Saccharomyces cerevisiae*. *Biotechnol Prog* **12**: 16–21.
- Park YS, Shiba S, Iijima S, Kobayashi T & Hishinuma F (1993) Comparison of 3 different promoter systems for secretory  $\alpha$ -amylase production in fed-batch cultures of recombinant *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **41**: 854–861.
- Partow S, Siewers V, Bjorn S, Nielsen J & Maury J (2010) Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. *Yeast* **27**: 955–964.
- Prein B, Natter K & Kohlwein SD (2000) A novel strategy for constructing N-terminal chromosomal fusions to green fluorescent protein in the yeast *Saccharomyces cerevisiae*. *FEBS Lett* **485**: 29–34.
- Price VL, Taylor WE, Clevenger W, Worthington M & Young ET (1990) Expression of heterologous proteins in *Saccharomyces cerevisiae* using the ADH2 promoter. *Methods Enzymol* **185**: 308–318.
- Purvis JJ, Chotai D, Dykes CW, Lubahn DB, French FS, Wilson EM & Hobden AN (1991) An androgen-inducible expression system for *Saccharomyces cerevisiae*. *Gene* **106**: 35–42.
- Radhakrishnan P & Srivastava V (2005) Utility of the FLP-FRT recombination system for genetic manipulation of rice. *Plant Cell Rep* **23**: 721–726.
- Randez-Gil F, Prieto JA, Murcia A & Sanz P (1995) Construction of baker's yeast strains that secrete *Aspergillus oryzae*  $\alpha$ -amylase and their use in bread making. *J Cereal Sci* **21**: 185–193.
- Raymond CK, Pownder TA & Sexson SL (1999) General method for plasmid construction using homologous recombination. *BioTechniques* **26**: 134–138.
- Reifenberger E, Freidel K & Ciriacy M (1995) Identification of novel HXT genes in *Saccharomyces cerevisiae* reveals the impact of individual hexose-transporters on glycolytic flux. *Mol Microbiol* **16**: 157–167.
- Ro DK, Paradise EM, Ouellet M *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940–943.



- Ro D-K, Ouellet M, Paradise EM, Burd H, Eng D, Paddon CJ, Newman JD & Keasling JD (2008) Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol* **8**: 83.
- Roca C, Nielsen J & Olsson L (2003) Metabolic engineering of ammonium assimilation in xylose-fermenting *Saccharomyces cerevisiae* improves ethanol production. *Appl Environ Microbiol* **69**: 4732–4736.
- Romanos MA, Scorer CA & Clare JJ (1992) Foreign gene-expression in yeast – a review. *Yeast* **8**: 423–488.
- Ruohonen L, Penttilä M & Keranen S (1991) Optimization of Bacillus  $\alpha$ -amylase production by *Saccharomyces cerevisiae*. *Yeast* **7**: 337–346.
- Ruohonen L, Aalto MK & Keranen S (1995) Modifications to the *ADH1* promoter of *Saccharomyces cerevisiae* for efficient production of heterologous proteins. *J Biotechnol* **39**: 193–203.
- Sadowski I, Su T-C & Parent J (2007) Disintegrator vectors for single-copy yeast chromosomal integration. *Yeast* **24**: 447–455.
- Sakai A, Shimizu Y & Hishinuma F (1990) Integration of heterologous genes into the chromosome of *Saccharomyces cerevisiae* using a delta sequence of yeast retrotransposon Ty. *Appl Microbiol Biotechnol* **33**: 302–306.
- Sakai A, Ozawa F, Higashizaki T, Shimizu Y & Hishinuma F (1991) Enhanced secretion of human nerve growth factor from *Saccharomyces cerevisiae* using an advanced  $\delta$ -integration system. *Biotechnology* **9**: 1382–1385.
- Sangsoda S, Cherest H & Surdinkerjan Y (1985) The expression of the *MET25* gene of *Saccharomyces cerevisiae* is regulated transcriptionally. *Mol Gen Genet* **200**: 407–414.
- Sauer B (1987) Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* **7**: 2087–2096.
- Sauer B (1994) Recycling selectable markers in yeast. *BioTechniques* **16**: 1086–1088.
- Schaerer-Brodbeck C & Barberis A (2004) Coupling homologous recombination with growth selection in yeast: a tool for construction of random DNA sequence libraries. *BioTechniques* **37**: 202–206.
- Shao ZY, Zhao H & Zhao HM (2009) DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res* **37**: e16.
- Shiba Y, Paradise EM, Kirby J, Ro DK & Keasling JD (2007) Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* **9**: 160–168.
- Shiomi N, Fukuda H, Murata K & Kimura A (1995) Improvement of S-adenosylmethionine production by integration of the ethionine-resistance gene into chromosomes of the yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **42**: 730–733.
- Shusta EV, Raines RT, Pluckthun A & Wittrup KD (1998) Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. *Nat Biotechnol* **16**: 773–777.
- Sikorski RS & Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27.
- Steen EJ, Chan R, Prasad N, Myers S, Petzold CJ, Redding A, Ouellet M & Keasling JD (2008) Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microb Cell Fact* **7**: 36.
- Szcebara FM, Chandelier C, Villeret C *et al.* (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat Biotechnol* **21**: 143–149.
- Takeuchi S, Hirayama K, Ueda K, Sakai H & Yonehara H (1958) Blastidin S, a new antibiotic. *J Antibiot (Tokyo)* **11**: 1–5.
- Taxis C & Knop M (2006) System of centromeric, episomal, and integrative vectors based on drug resistance markers for *Saccharomyces cerevisiae*. *BioTechniques* **40**: 73–78.
- Van Mullem V, Wery M, De Bolle X & Vandenhoute J (2003) Construction of a set of *Saccharomyces cerevisiae* vectors designed for recombinational cloning. *Yeast* **20**: 739–746.
- Verwaal R, Wang J, Meijnen J-P, Visser H, Sandmann G, van den Berg JA & van Ooyen AJJ (2007) High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol* **73**: 4342–4350.
- Walhout AJM, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S & Vidal M (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol* **328**: 575–592.
- Walton EF & Yarranton GTC (1989) Negative regulation of gene expression by mating type. *Molecular and Cell Biology of Yeasts* (Walton EF & Yarranton GTC, eds), pp. 43–69. Blackie and son Ltd, London.
- Wang XH, Wang ZJ & Da Silva NA (1996) G418 selection and stability of cloned genes integrated at chromosomal delta sequences of *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **49**: 45–51.
- West RW, Chen SM, Putz H, Butler G & Banerjee M (1987) GAL1-GAL10 divergent promoter region of *Saccharomyces cerevisiae* contains negative control elements in addition to functionally separate and possibly overlapping upstream activating sequences. *Genes Dev* **1**: 1118–1131.
- Wisselink HW, Toirkens MJ, Berriel MDF, Winkler AA, van Dijken JP, Pronk JT & van Maris AJA (2007) Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* **73**: 4881–4891.
- Wittrup KD, Robinson AS, Parekh RN & Forrester KJ (1994) Existence of an optimum expression level for secretion of foreign proteins in yeast. *Ann N Y Acad Sci* **745**: 321–330.
- Wyrick JJ, Aparicio JG, Chen T, Barnett JD, Jennings EG, Young RA, Bell SP & Aparicio OM (2001) Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*:

- high-resolution mapping of replication origins. *Science* **294**: 2357–2360.
- Xie DM, Shao ZY, Achkar JH, Zha WJ, Frost JW & Zhao HM (2006) Microbial synthesis of triacetic acid lactone. *Biotechnol Bioeng* **93**: 727–736.
- Yamada R, Taniguchi N, Tanaka T, Ogino C, Fukuda H & Kondo A (2010) Cocktail  $\delta$ -integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains. *Microb Cell Fact* **9**: 32.
- Yan YJ, Kohli A & Koffas MAG (2005) Biosynthesis of natural flavanones in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**: 5610–5613.
- Yocum RR, Hanley S, West R & Ptashne M (1984) Use of lacZ fusions to delimit regulatory elements of the inducible divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol* **4**: 1985–1998.
- Zelle RM, de Hulster E, van Winden WA, de Waard P, Dijkema C, Winkler AA, Geertman J-MA, van Dijken JP, Pronk JT & van Maris AJA (2008) Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. *Appl Environ Microbiol* **74**: 2766–2777.
- Zhu J & Zhang MQ (1999) SCPD: a promoter database of the yeast *Saccharomyces cerevisiae*. *Bioinformatics* **15**: 607–611.