Development of yeast genome-engineering technology and its application to antibody production

Nishida, Yumiko

(Degree)
博士（工学）

(Date of Degree)
2018-03-25

(Date of Publication)
2019-03-01

(Resource Type)
doctoral thesis

(Report Number)
甲第7194号

(URL)
https://hdl.handle.net/20.500.14094/D1007194

博士論文

Development of yeast genome-engineering technology and its application to antibody production

酵母におけるゲノム編集技術の開発と抗体生産への応用

平成 30 年 1 月

神戸大学大学院工学研究科

NISHIDA Yumiko

西田　由美子

CONTENTS

General Introduction

Yeasts are one of the fundamental organisms for fermentation, research use and biotechnology. The purpose of this study is to provide basic technology and information to make most use of yeast for biomass-based production. The outline of this research is depicted in Figure 1. Knowing optimum expression level of each genes is important for both understanding and engineering of the organisms. In chapter 1, characterization of promoters, which dictate downstream gene expression is performed in various media conditions simulating biomass utilization. Providing selection of yeast species with distinct characteristics such as fast growth and thermotolerance may further be beneficial for industrial application. In chapter 2, comprehensive and highly efficient genome engineering tool is developed in unconventional thermotolerant yeast. Finally, by utilizing the genetic engineering tools and the promoter selection, valuable protein production in the unconventional yeast was performed in chapter 3.

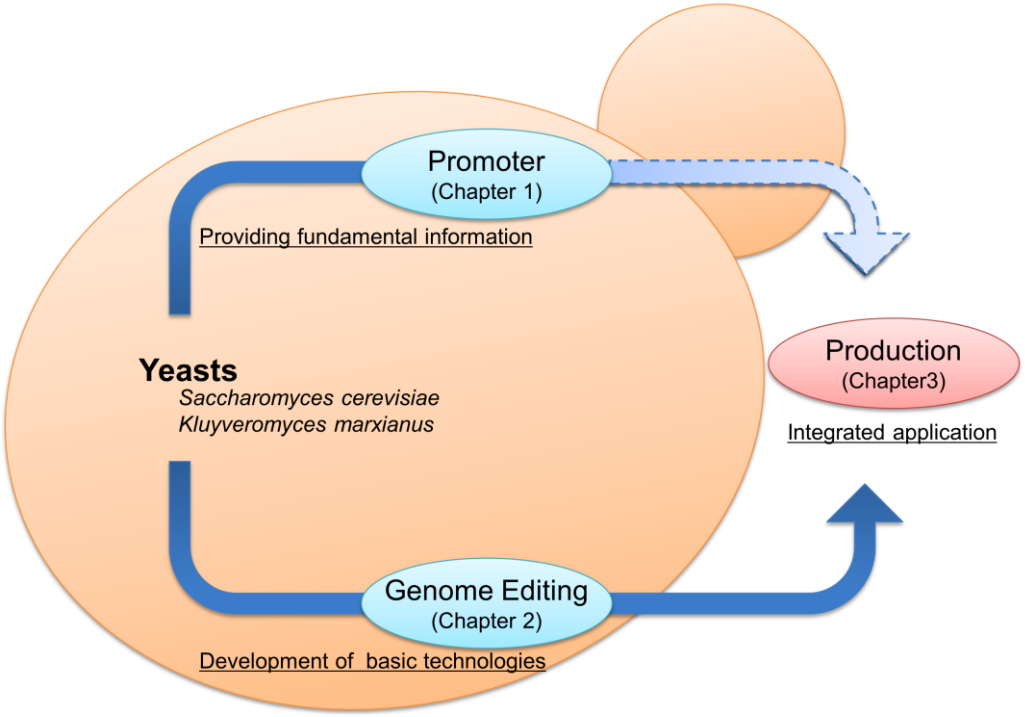


Figure 1 Outline of the research.

1.1 Yeast *Saccharomyces cerevisiae*

The budding yeast *Saccharomyces cerevisiae* has approximately 6000 genes (Lushchak 2006) and serves as eukaryotic model as it shares common cellular functions with higher organisms (Botstein et al. 1997; Schneiter 2007). Usefulness of *S. cerevisiae* is established with efficient genetic transformation tools including homologous recombination and host vector systems (Ishii et al. 2014; Gietz et al. 1992).

Therefore, *S. cerevisiae* is widely used as a model organism for both basic research and industrial bioproduction ranging from ethanol (Katahira et al. 2006; Kondo et al. 2013; Hasunuma et al. 2015) to proteins (Buckholz and Gleeson 1991; (Çelik and Çalık 2012).

Practically, biomass is the major source for bioproduction and needs pretreatment to extract cellulosic and hemicellulosic materials, which are then digested by enzymes to become simple sugars such as glucose and xylose (Hasunuma et al. 2015). As conventional *S. cerevisiae* does not assimilate xylose, genes (*XYL1* and *XYL2*) respectively encoding xylose reductase (XR) and xylitol dehydrogenase (XDH) from *Scheffersomyces stipites* have been introduced to engineer the yeast to assimilate xylose for the highest production from biomass (Eliasson et al. 2000; Johansson et al. 2001). However, the engineered strains grow on xylose much slower than on glucose (Salusjärvi et al. 2008; Matsushika et al. 2014), limiting its practical use. To enhance growth and production by genetic engineering, it is necessary to optimize gene expression network and fine-tune metabolic pathway, for which fundamental study is needed to characterize promoter activity in xylose medium. Detail of the promoter activity is further explained in Section1.3.

1.2 Yeast *Kluyveromyces marxianus*

Kluyveromyces marxianus is classified as “Crabtree negative” yeast, meaning that it does not convert sugars into ethanol under aerobic condition. *K. marxianus* is also a thermotolerant yeast and can grow fast at 37-42°C (Goshima et al. 2013) Recently, the complete genome sequence of *K. marxianus* becomes publicly available (Jeong et al. 2012; Suzuki et al. 2014; Inokuma et al. 2015). Therefore, *K. marxianus* is potentially suitable for engineering to produce valuable products such as recombinant proteins.

However, *K. marxianus* still lacks efficient genome engineering tools. Unlike *S.*

cerevisiae, transformation by homologous recombination (HR) is much less likely to occur and gene fragments tend to be inserted at random chromosomal sites (Nonklang et al. 2008). Supposedly, *K. marxianus* retains highly active non-homologous end joining (NHEJ) pathway (Dudášová et al. 2004), which repair DNA double-strand breaks (DSBs) by tethering DNA ends regardless of the sequences (Figure 2). The pathway facilitates random integration of DNA fragments and thus antagonizes with HR.

Plasmid system is also useful for heterologous gene expression. For a plasmid to be retained in a cell, autonomously replicating sequence (ARS) and centromere are indispensable. ARS initiates replication of plasmid (Newlon and Theis 1993) and centromere is required for transmission to daughter cells (Carbon 1984). Nonetheless, plasmid system is less stable than genomic integration in general. Selection marker such as drug resistance gene is also needed to select and retain plasmid-carrying cells.

In this study, comprehensive genome engineering tools are provided by introducing genome editing technology. The genome editing is described in Section 1.4.

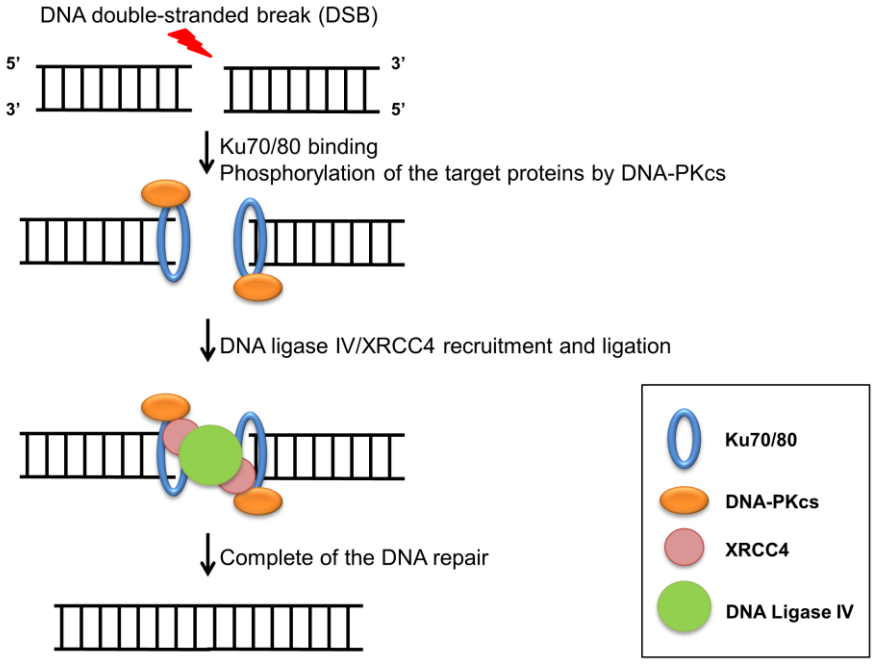


Figure 2 Mechanisms of DNA double-strand break repair by NHEJ pathway

1.3 Promoter activity and gene expression

Each gene expression level is dependent on its promoter activity, which is responsive to the cellular environment. In yeast, P_{TDH3} (promoter of *TDH3*) is assumed to be highly active in glucose media (Ishii et al. 2014). P_{PDC1}, P_{TEF1}, P_{TPI1}, and P_{PGK1} are also widely used for engineered gene expression (Sun et al. 2012; Ishii et al. 2014; Guo et al. 2016). However, promoter activity should be dependent on the environment such as carbon source and aeration, which can be changing throughout the culture. To achieve optimal set of gene expression, it is important to know the actual promoter activity under varying conditions. In chapter 1, activity of a set of promoters is quantitatively characterized in glucose or xylose medium under aerobic cultivation or microaerobic fermentation over time.

1.4 Genome editing tools

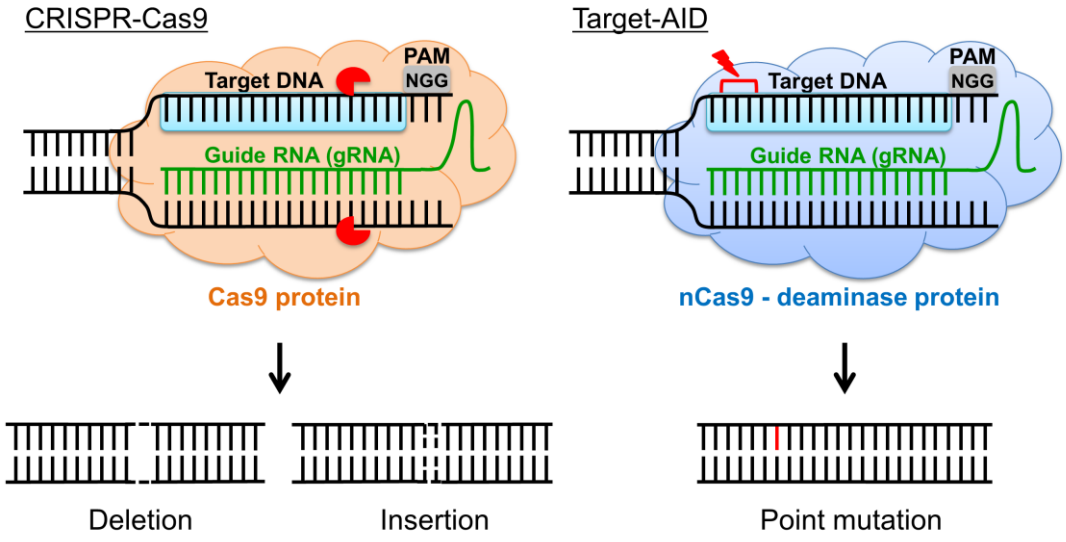


Figure 3 The principle of genome editing by CRISPR-Cas9 and Target-AID

・CRISPR-Cas9 system

In bacteria and archaea, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (CRISPR-Cas9) was discovered as an acquired immune system against viruses and phages through CRISPR RNA (crRNA)-base DNA recognition and Cas9 nuclease-mediated DNA cleavage (Gasiunas et al. 2012). CRISPR-Cas9-mediated genome editing depends on the generation of double-strand break (DSB) and subsequent cellular DNA repair process (Zhang et al. 2014). The CRISPR-Cas9 system is depicted in Figure 3. Target-recognizing sequence of crRNA or guide RNA (gRNA) in general can be customized to a specific sequence, which enables genome editing. When introduced in a cell, Cas9 and the custom gRNA binds to complementary target DNA strand in a protospacer-adjacent motif (PAM)-dependent manner, which appears as the sequence of 5'-NGG-3' directly downstream of the target sequence (Jinek et al. 2012). Cas9 nuclease induces DSB at the target site, which is then repaired by NHEJ pathway in the host cell. Error-prone nature of the NHEJ pathway may induce insertion or deletion (indel) of base pairs, resulting in gene disruption at the target locus.

・Target-AID system

Unlike nuclease-mediated genome editing, deaminase-mediated targeted mutagenesis (Target-AID) directly introduce point mutations (Nishida et al. 2016).The activation-induced cytidine deaminase (AID) is tethered to nuclease-deficient CRISPR-Cas9 to specifically induce cytosine (C) to thymine (T) mutation at -16 to -19 positions upstream of the PAM sequence.

In chapter 2, we introduced the genome editing tools CRISPR-Cas9 and Target-AID in *K. marxianus*, which enabled indel formation, highly efficient homologous recombination and direct introduction of point mutation at the target loci.

1.5 Production of biopharmaceutical protein

Biopharmaceutical products are promising targets for industrial production. There has been an increasing demand of biopharmaceutical proteins in the market (Nielsen 2013). Efficient production and safety issues such as endotoxin and virus contamination

are the major challenges for biopharmaceuticals. Yeasts are one of the preferred host cell for production of biopharmaceutical proteins (Huang et al. 2014). As a unicellular microbial organism, it is easy to culture, genetically tractable, grows fast, and adjusts well to industrial processing. Efficient secretion system of the yeasts also reduces the costs of downstream purification (Huang et al. 2014). *S. cerevisiae* and *Pichia pastoris* are the major species on which much efforts have been done to engineer for protein production (Miller et al. 2005; Tang et al. 2016). In chapter 3, we engineered the *K. marxianus* as a host for producing single-chain Fv antibody (scFv). Genetic and physiological manipulation has been applied to probe important factors for efficient production.

Reference

Botstein D, Chervitz S a., Cherry JM., 1997. Yeast as a Model Organism. 1997. Science. 277(5330), 1259-1260.

Buckholz RG, Gleeson MA., 1991. Yeast systems for the commercial production of heterologous proteins. Nat. Biotechnol. 9(11), 1067-1072.

Carbon J., 1984. Yeast centromeres: structure and function. Cell. 37(2), 351-353.

Çelik E, Çalık P., 2012. Production of recombinant proteins by yeast cells. Biotechnol. Adv. 30(5), 1108-1118.

Dudášová Z, Dudáš A, Chovanec M., 2004. Non-homologous end-joining factors of Saccharomyces cerevisiae. FEMS Microbiol. Rev. 28(5), 581-601.

Eliasson A, Christensson C, Wahlbom CF, Hahn-Hägerdal B., 2000. Anaerobic xylose fermentation by recombinant *Saccharomyces cerevisiae* carrying *XYL1*, *XYL2*, and *XKS1* in mineral medium chemostat cultures. Appl. Environ. Microbiol. 66(8), 3381–3386.

Gasiunas G, Barrangou R, Horvath P, Siksnys V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. 109(39), E2579–E2586.

Gietz D, St Jean A, Woods R a, Schiestl RH., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20(6), 1425.

Goshima T, Tsuji M, Inoue H, Yano S, Hoshino T, Matsushika A., 2013. Bioethanol production from lignocellulosic biomass by a novel *Kluyveromyces marxianus* strain. Biosci. Biotechnol. Biochem. 77(7), 1505–1510.

Guo W, Sheng J, Zhao H, Feng X., 2016. Metabolic engineering of *Saccharomyces cerevisiae* to produce 1-hexadecanol from xylose. Microb. Cell Fact. 15(1), 24.

Hasunuma T, Ishii J, Kondo A., 2015. Rational design and evolutional fine tuning of *Saccharomyces cerevisiae* for biomass breakdown. Curr. Opin. Chem. Biol. 29, 1–9.

Huang M, Bao J, Nielsen J., 2014. Biopharmaceutical protein production by *Saccharomyces cerevisiae* : current state and future prospects. Pharm. Bioprocess. 2(2), 167–182.

Inokuma K, Ishii J, Hara KY, Mochizuki M, Hasunuma T, Kondo A., 2015. Complete genome sequence of *Kluyveromyces marxianus* NBRC1777, a nonconventional thermotolerant yeast. Genome Announc. 2015 3(2), e00389-15.

Ishii J, Kondo T, Makino H, Ogura A, Matsuda F, Kondo A., 2014. Three gene expression vector sets for concurrently expressing multiple genes in *Saccharomyces cerevisiae*. FEMS Yeast Res. 14(3), 399–411.

Jeong H, Lee DH, Kim SH, Kim HJ, Lee K, Song JY, et al., 2012. Genome sequence of the thermotolerant yeast *Kluyveromyces marxianus* var. *marxianus* KCTC 17555. Eukaryot. Cell. 11(12), 1584–1585.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 337(6096), 816–821.

Johansson B, Christensson C, Hobley T, Hahn-Hägerdal B., 2001. Xylulokinase overexpression in two strains of *Saccharomyces cerevisiae* also expressing xylose reductase and xylitol dehydrogenase and its effect on fermentation of xylose and lignocellulosic hydrolysate. Appl. Environ. Microbiol. 67(9), 4249–4255.

Katahira S, Mizuike A, Fukuda H, Kondo A., 2006. Ethanol fermentation from lignocellulosic hydrolysate by a recombinant xylose- and cellooligosaccharide-assimilating yeast strain. Appl. Microbiol. Biotechnol. 72(6), 1136–1143.

Kondo A, Ishii J, Hara KY, Hasunuma T, Matsuda F., 2013. Development of microbial cell factories for bio-refinery through synthetic bioengineering. J. Biotechnol. 163(2), 204–216.

Lushchak V., 2006. Budding yeast *Saccharomyces cerevisiae* as a model to study oxidative modification of proteins in eukaryotes. Acta Biochim. Pol. 53(4), 679–684.

Matsushika A, Goshima T, Hoshino T., 2014. Transcription analysis of recombinant industrial and laboratory *Saccharomyces cerevisiae* strains reveals the molecular basis for fermentation of glucose and xylose. Microb. Cell Fact. 13(1), 16.

Miller KD, Weaver-Feldhaus J, Gray S a, Siegel RW, Feldhaus MJ., 2005. Production, purification, and characterization of human scFv antibodies expressed in *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Escherichia coli*. Protein Expr. Purif. 42(2), 255–267.

Newlon CS, Theis JF., 1993. The structure and function of yeast ARS elements. Curr. Opin. Genet. Dev. 3(5), 752–758.

Nielsen J., 2013. Production of biopharmaceutical proteins by yeast. Bioengineered. 4(4), 207–211.

Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, et al., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science. 353, aaf8729-1-aaf8729-8.

Nonklang S, Abdel-Banat BM a, Cha-aim K, Moonjai N, Hoshida H, Limtong S, et al., 2008. High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. Appl. Environ. Microbiol. 74(24), 7514–7521.

Rouwenhorst RJ, Hensing M, Verbakel J, Scheffers WA, Van Dijken JP., 1990. Structure and properties of the extracellular inulinase of *Kluyveromyces marxianus* CBS 6556. Appl. Environ. Microbiol. 56(11):3337–3345.

Salusjärvi L, Kankainen M, Soliymani R, Pitkänen J-P, Penttilä M, Ruohonen L., 2008. Regulation of xylose metabolism in recombinant *Saccharomyces cerevisiae*. Microb. Cell Fact. 7, 18.

Schneiter R., 2007. Intracellular sterol transport in eukaryotes, a connection to mitochondrial function? Biochimie. 89(2), 255–259.

Sun J, Shao Z, Zhao H, Nair N, Wen F, Xu J-H, et al., 2012. Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 109(8), 2082–2092.

Suzuki T, Hoshino T, Matsushika A., 2014. Draft genome sequence of *Kluyveromyces marxianus* strain DMB1, Isolated from sugarcane bagasse hydrolysate. Genome Announc. 24, e00733-14.

Tang H, Wang S, Wang J, Song M, Xu M, Zhang M, et al., 2016. N-hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in *Saccharomyces cerevisiae*. Sci. Rep. 6(1), 25654.

Zhang F, Wen Y, Guo X., 2014. CRISPR/Cas9 for genome editing: Progress, implications and challenges. Hum. Mol. Genet. 23(R1):40–46.

Chapter 1

Selection of yeast *Saccharomyces cerevisiae* promoters available for xylose cultivation and fermentation

Introduction

The budding yeast *Saccharomyces cerevisiae* has traditionally been used for making alcohol in the brewing industry (Lodolo et al. 2008). Because the yeast is able to predominantly produce ethanol, it has also been used for making biofuels over the last two decades (Hahn-Hägerdal et al. 2006), and recently it has been adopted as a platform for producing higher alcohols and other chemicals (Steen et al. 2008; Kondo et al. 2012; Kondo et al. 2013). Cellulosic and hemicellulosic materials, containing glucose and xylose as the major sugar components, are regarded as sustainable biomass resources, and various efforts have been made to produce bioethanol and other compounds from these materials using yeasts and other microbes (Hasunuma et al. 2013, 2015).

S. cerevisiae grows favorably on glucose, but is naturally unable to grow well on xylose. Engineered yeasts that express heterogenous *XYL1* and *XYL2* genes encoding xylose reductase (XR) and xylitol dehydrogenase (XDH) from *Scheffersomyces stipitis*, and the endogenous *XKS1* gene encoding xylulokinase (XK) (Eliasson et al. 2000; Johansson et al. 2001), can utilize xylose as a carbon source. XR reduces xylose to xylitol by an NADPH-consuming reaction, while XDH oxidizes xylitol to xylulose by an NAD^+-consuming reaction (Fernandes and Murray 2010). Then, XK phosphorylates xylulose into xylulose-5-phosphate (X5P), which is further metabolized via the pentose phosphate pathway (PPP) and joins glycolysis (Jin et al. 2003).

Even in xylose-utilizing yeast strains expressing XR, XDH, and XK, the consumption rate of xylose is much slower than that of glucose (Salusjärvi et al. 2008; Matsushika et al. 2014). To improve xylose utilization, further engineering of other metabolic pathways, such as the PPP, is required (Johansson and Hahn-Hägerdal 2002). In addition, the ability for biomass degradation and inhibitor tolerance should be conferred to obtain more sophisticated hemicellulose-utilizing yeast strains. For example, xylanase and xylosidase have been secreted outside of cells or displayed on

cell surfaces as hemicellulose degradation enzymes (La Grange et al. 2001; Guirimand et al. 2016). And formate dehydrogenase (FDH) and Haa1 transcription factor have been overexpressed as a detoxification enzyme and a global transcription activator, respectively (Hasunuma et al. 2011b; Sakihama et al. 2015). Thus, there exist varieties of gene targets to be expressed in xylose-utilizing yeast strains, and it is important to properly control their expression levels.

To date, many researchers have investigated the activities of *S. cerevisiae* promoters in the presence of glucose as a sole carbon source (Shen et al. 2012; Partow et al. 2010; Sun et al. 2012). However, there are few reports evaluating promoter activities using xylose as a carbon source. In one example, the activities of 6 promoters were measured in xylose fermentation conditions (Lu and Jeffries 2007). In addition, the shuffling of the promoter activities for gene expression, such as in the PPP, to optimize xylose utilization has been reported (Lu and Jeffries 2007; Latimer et al. 2014).

In the present study, we aimed to expand the options of promoter choice in *S. cerevisiae* for xylose cultivation and fermentation. To choose promoter candidates, comprehensive gene expressions of xylose-utilizing yeast strains were analyzed both in xylose- and glucose-fermenting conditions using DNA microarrays. Based on normalized signals during xylose fermentation, we picked 30 promoter candidates that showed strong, moderate, and weak expressions. Using the *lacZ* reporter gene, the activities of these 30 promoters in the xylose-utilizing strain were measured over time. During aerobic cultivation and microaerobic fermentation, xylose and glucose were used as respective sole carbon sources.

Materials and Methods

Media and growth conditions

The *S. cerevisiae* strains used in this study are listed in Table 1. *Escherichia coli* strain DH5α (Toyobo, Osaka, Japan) used for cloning was grown in Luria-Bertani medium (10 g/L peptone, 5 g/L yeast extract, and 5 g/L sodium chloride) supplemented with 100 mg/L ampicillin at 37°C. Yeast cells were grown in YPD medium (10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose) at 30°C. Yeast strains harboring

episomal plasmids were grown in SD (6.7 g/L yeast nitrogen base without amino acid [YNB] and 20 g/L glucose), SX (6.7 g/L YNB and 15 g/L xylose), SDX (6.7 g/L YNB, 14 g/L glucose and 6 g/L xylose), or SDX2 (6.7 g/L YNB, 35 g/L glucose and 15 g/L xylose) selective medium, which contained 20 mg/L histidine and 20 mg/L methionine.

Construction of plasmids and yeast transformation

The plasmids used in this study are listed in Table 1. The open reading frame (ORF) of *lacZ* was amplified from the genomic DNA of *E. coli* K-12 (strain MG1655) as a template by PCR using the primers lacZ_Fw and lacZ_Rv. The SAG1 terminator (T_{SAG1}) was amplified from the genomic DNA of *S. cerevisiae* BY4741 (Brachmann et al. 1998) as a template using the primers SAG1t_Fw and SAG1t_Rv. Two fragments of *lacZ* and T_{SAG1} were connected by overlap extension-PCR using the primers lacZ_SAG1t_Fw and lacZ_SAG1t_Rv to yield *lacZ*-T_{SAG1} fragment. The episomal plasmid pRS415 (Sikorski and Hieter 1989) was digested with *Not*I. The *lacZ*-T_{SAG1} fragment was inserted into *NotI*-digested pRS415 by In-fusion cloning (Takara Bio, Shiga, Japan) to yield pRS415/*lacZ*-T_{SAG1}.

Promoters used in this study are listed in Table 2. The promoter fragments were amplified from the genomic DNA of *S. cerevisiae* BY4741 by PCR using the primers listed in Table 3. Promoter fragments were inserted into the *Hind*III/*Bam*HI or *Hind*III/*Sal*I sites of pRS415/*lacZ*-T_{SAG1} by ligation or In-fusion cloning to yield plasmids pRS415/P_{xxxx}-*lacZ*-T_{SAG1} (*xxxx* indicates the gene names used as candidate promoters).

A fragment containing *S. stipitis* XR and XDH genes, the *S. cerevisiae* endogenous XK gene, and *URA3* selectable marker were amplified from plasmid pIUX1X2XKN (Ismail et al. 2013) by PCR using the primers X-I2_Fw and X-I2_Rv. The amplified fragment was integrated into the I2 genomic locus (Inokuma et al. 2015) of *S. cerevisiae* strain BY4741 with the lithium acetate method (Gietz et al. 1992). The yielded BY4741/X-I2 strain was selected on SD selective medium additionally containing 60 mg/L leucine. The episomal plasmids pRS415/*lacZ*-T_{SAG1} and pRS415/P_{xxxx}-*lacZ*-T_{SAG1} were introduced into BY4741/X-I2 with the lithium acetate method. The yielded strains BY4741/X-I2-C and BY4741/X-I2-P_{xxxx} were selected on SD selective medium.

Ethanol fermentations in YP xylose- and glucose-rich media

The strain BY4741/X-I2 was aerobically cultivated in YPD medium for 48 h at 30°C. The cells were collected by centrifugation at 3,000 rpm for 10 min at 4°C and washed twice with distilled water. Then cells were inoculated into 50 mL of fermentation medium (10 g/L yeast extract, 20 g/L peptone) containing 50 g/L xylose or glucose. Fermentation was performed at 30°C with agitation at 500 rpm in 100 mL closed bottles equipped with a bubbling CO_2 outlet under oxygen limited conditions (microaerobic conditions).

Table 1 Plasmids and *S. cerevisiae* strains used in this study

Plasmids and Strains	Genotype	Reference
Plasmids		
pIUX1X2XKN	Genomic integration plasmid with the XR-XDH-XK expression cassette (P_{TDH3}-*XYL1*, P_{CDC19}-*XYL2* and P_{TPI1}-*XKS1*) and *URA3* marker	(Ismail et al. 2013)
pRS415	Single-copy vector (*CEN/ARS* origin), *LEU2* marker	ATCC (Sikorski and Hieter 1989)
pRS415/*lacZ*-T_{SAG1}	*lacZ* and T_{SAG1} in pRS415	This study
pRS415/P_{xxxx}-*lacZ*-T_{SAG1}	Candidate promoters in pRS415/*lacZ*-T_{SAG1}	This study
Strains		
BY4741	*MAT***a** *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	ATCC (Brachmann et al. 1998)
BY4741/X-I2	BY4741 integrating pIUX1X2XKN (P_{TDH3}-*XYL1*, P_{CDC19}-*XYL2* and P_{TPI1}-*XKS1*)	This study
BY4741/X-I2-C	BY4741/X-I2 harboring pRS415/*lacZ*-T_{SAG1}	This study
BY4741/X-I2-P_{xxxx}	BY4741/X-I2 harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}	This study

Table 2 List of promoters used in this study

Gene	Description	Length of the cloned upstream region (bp)
ADH1	Alcohol dehydrogenase	700
ADH3	Mitochondrial alcohol dehydrogenase isozyme III	1021
ALD3	Cytoplasmic aldehyde dehydrogenase	574
ALD4	Mitochondrial aldehyde dehydrogenase	1000
CDC19	Pyruvate kinase	998
FBA1	Fructose 1,6-bisphosphate aldolase	551
GND1	6-phosphogluconate dehydrogenase (decarboxylating);	1000
GND2	6-phosphogluconate dehydrogenase (decarboxylating)	999
HXT2	High-affinity glucose transporter of the major facilitator superfamily	1000
HXT7	High-affinity glucose transporter	587
HXT8	Protein of unknown function with similarity to hexose transporters	1000
LAT1	Dihydrolipoamide acetyltransferase component (E2) of the PDC	1000
LPD1	Dihydrolipoamide dehydrogenase	673
NQM1	Transaldolase of unknown function	1000
PDA1	E1 alpha subunit of the pyruvate dehydrogenase (PDH) complex	190
PDC1	Major of three pyruvate decarboxylase isozymes	793
PDC6	Minor isoform of pyruvate decarboxylase	1002
PGK1	3-phosphoglycerate kinase	538
RKI1	Ribose-5-phosphate ketol-isomerase	1000
SED1a	Major stress-induced structural GPI-cell wall glycoprotein	1063
SED1b	Major stress-induced structural GPI-cell wall glycoprotein	800
SOL4	6-phosphogluconolactonase	997
STB5	Transcription factor	1000
TDH1	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), isozyme 1	1000
TDH3	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), isozyme 3	680
TEF1	Translational elongation factor EF-1 alpha	421
TEF2	Translational elongation factor EF-1 alpha	402
TKL2	Transketolase	1000
TPI1	Triose phosphate isomerase, abundant glycolytic enzyme	584
ZWF1	Glucose-6-phosphate dehydrogenase (G6PD)	1000

Table 3 Primers used in this study

Primer name	Sequence (5'-3')
lacZ_Fw	ATGACCATGATTACGGATTCACTGG
lacZ_Rv	GTACTAACTGTACAGTACCCGTTCTTATTTTTGACACCAGACCAACTGG
SAG1t_Fw	CCAGTTGGTCTGGTGTCAAAAATAAAACGGGTACTGTACAGTTAGTAC
SAG1t_Rv	TTTGATTATGTTCTTTCTATTTG
lacZ_SAG1t_Fw	AGTTCTAGAGCGGCCGCATGACCATGATTACG
lacZ_SAG1t_Rv	ACCGCGGTGGCGGCCGCTTTGATTATGTTCTTTC
pADH1_Fw	AAAAAAAAGCTTGGGTGTACAATATGGACTTCCTC
pADH1_Rv	AAAAAAGGATCCGTATATGAGATAGTTGATTGTATGC
pTDH3_Fw	AAAAAAAAGCTTCAGTTCGAGTTTATCATTATC
pTDH3_Rv	AAAAAAGGATCCTTTGTTTGTTTATGTGTGTTTATTCG
pPGK1_Fw	AAAAAAAAGCTTGCCTCTTATCGAGAAAGAAATTACC
pPGK1_Rv	AAAAAAGGATCCGTTTTATATTTGTTGTAAAAAGTAG
pHXT7_Fw	AAAAAAAAGCTTCCGTGGAAATGAGGGGTATGCAGG
pHXT7_Rv	AAAAAAGGATCCTTTTTGATTAAAATTAAAAAAAC
pTEF1_Fw	AAAAAAAAGCTTCCACACACCATAGCTTCAAAATG
pTEF1_Rv	AAAAAAGGATCCTTTGTAATTAAAACTTAGATTAGATTGC
pTEF2_Fw	AAAAAAAAGCTTATTACCCATAAGGTTGTTTGTGACG
pTEF2_Rv	AAAAAAGGATCCGTTTAGTTAATTATAGTTCGTTGACCG
pSED1(1063)_Fw	AAAAAAGTCGACgaaaaacgacaacattccacc
pSED1(1063)_Rv	AAAAAAGGATCCtcttaatagagcgaacgtattttattttgc
pTDH1_Fw	AAAAAAAAGCTTCTTCCCTTTTACAGTGCTTCGG
pTDH1_Rv	AAAAAAGGATCCTTTGTTTTGTGTGTAAATTTAG
pFBA1_Fw	AAAAAAAAGCTTTGACAGCAGGATTATCGTAATACG
pFBA1_Rv	AAAAAAGGATCCTTTGAATATGTATTACTTGGTTATG
pTPI1_Fw	AAAAAAAAGCTTGTGTTTAAAGATTACGGATATTTAAC
pTPI1_Rv	AAAAAAGGATCCTTTTAGTTTATGTATGTGTTTTTTG
pPDC1_Fw	AAAAAAAAGCTTCTGGGTGAGCATATGTTCCGCTG
pPDC1_Rv	AAAAAAGGATCCTTTGATTGATTTGACTGTGTTATTTTG
pADH3_Fw	AAAAAAAAGCTTCCACGGCATTGAATCTATTTCTCC
pADH3_Rv	AAAAAAGGATCCTGATGGTGTTAACTTATGACTATAC
pPDA1_Fw	AAAAAAAAGCTTCTTCTGATTCCTCCACCCCTTCC
pPDA1_Rv	AAAAAAGGATCCTGGCACAAATGTGGTTTCCTTTC
pLPD1_Fw	AAAAAAAAGCTTAATATCGTAGCCTCGCAAAATTCC
pLPD1_Rv	AAAAAAGGATCCTGTGAACTGTTAGTATGTTCTTTG

pLAT1_Fw	CGGTATCGATAAGCTTCTACGAGTTTGACGACTCT
pLAT1_Rv	TAGAACTAGTGGATCCAGTACGCTAATAACTATTAATTT
pALD3_Fw	AAAAAAAAGCTTAAGCACCATTATTTCAGCATGCGG
pALD3_Rv	AAAAAAGGATCCTTTTCTTTTGGCTAATTTTCTAAATG
pCDC19_Fw	AAAAAAAAGCTTAATGCTAGTATTTTGGAGATTAATC
pCDC19_Rv	AAAAAAGGATCCTGTGATGATGTTTTATTTGTTTTG
pHXT2_Fw	AAAAAAAAGCTTTTTCTACCGATGTAATACAAAAATGAATC
pHXT2_Rv	AAAAAAGGATCCTATGTTGCTTTATAAGTCTTTTTG
pHXT8_Fw	AAAAAAGTCGACTAACCCATATCTTCTTGAGGAGAGTC
pHXT8_Rv	AAAAAAGGATCCGGTTTTTGCTGGCGTTCTGAGCTAAC
pSOL4_Fw	AAAAAAAAGCTTATAGCATTATGGTATTGTCCACCAC
pSOL4_Rv	AAAAAAGGATCCTTTTCTTGAGGTGGTTGTGGATTATTTG
pGND1_Fw	AAAAAAAAGCTTGCTATGTCACTAATGATCTCGTCGG
pGND1_Rv	AAAAAAGGATCCTTTATTGTAGTTTTTGTATAGAAAGTGGC
pRKI1_Fw	AAAAAAAAGCTTGTTTGGGTGTGATAAGGTTCATTGTG
pRKI1_Rv	AAAAAAGGATCCTTCTTTCCTATGTTGTATTTATATG
pALD4_Fw	CGGTATCGATAAGCTTAGCGCTTTCCATCAAAAATAC
pALD4_Rv	TAGAACTAGTGGATCCGCTTTATTTATTTTTGTTATTGTA
pNQM1_Fw	CGGTATCGATAAGCTTGCCCTGGTTCCACAAACC
pNQM1_Rv	TAGAACTAGTGGATCCATTTAATTTCTTGTATATATGTGA
pPDC6_Fw	CGGTATCGATAAGCTTCTACCCTATTTTCTCTTACC
pPDC6_Rv	TAGAACTAGTGGATCCTTTGTTGGCAATATGTTTTTGC
pTKL2_Fw	CGGTATCGATAAGCTTTATAGTTAACTACCGGCTCG
pTKL2_Rv	TAGAACTAGTGGATCCTTTTTTGTTCTTTTTTTGATTAGT
pGND2_Fw	CGGTATCGATAAGCTTCCGTCATAACTTTGAATCCT
pGND2_Rv	TAGAACTAGTGGATCCTCTGTTCCTCGTGTTTTTTTA
pSTB5_Fw	CGGTATCGATAAGCTTGGCAAAACAATCGAGAACC
pSTB5_Rv	TAGAACTAGTGGATCCGTCCTTCTTTTACACCTTTG
pSED1(800)_Fw	CGGTATCGATAAGCTTATTGGATATAGAAAATTAACGTA
pSED1(800)_Rv	TAGAACTAGTGGATCCCTTAATAGAGCGAACGTATTT
pZWF1_Fw	CGGTATCGATAAGCTTCGCCAGCGAGCTTTCCGG
pZWF1_Rv	TAGAACTAGTGGATCCCTTGCCTTATGTGGTTTTCT
X-I2_Fw	agatctggatctgaatggcatcattttgtgatgtgtaaaagcgggaccttgagagtgcaccataccacagcttttcaatt
X-I2_Rv	cacagcgggaaactccttagtgttcataaagaaatacgatcatatgtagtaaccgcctctccccgcgcgttggccgattc

Determination of yeast cells, ethanol, glucose, glycerol, xylose, and xylitol

Cell growth was measured by a microplate reader at OD_{600}. Ethanol, glucose, glycerol, xylose, and xylitol in medium were determined by a Shimadzu HPLC system equipped with a Shim-pack SPR-Pb column (7.8 mm × 250 mm; Shimadzu, Kyoto, Japan) and an RID-10A refractive index detector (Shimadzu) after the separation of cells by centrifugation at 14,000 × *g* for 5 min. The column temperature was 80°C. Water was used at a flow rate of 0.6 mL/min as the mobile phase (Hasunuma et al. 2011a)

Microarray analyses of yeast cells in xylose and glucose fermentations

Total RNA was isolated from yeast cells cultivated for 9 h in fermentation medium containing 50g /L xylose or for 3 h in fermentation medium containing 50 g/L glucose using a Total RNA Isolation Mini Kit (Agilent Technologies, Palo Alto, CA, USA) following the manufacturer's protocol. cDNA was generated by reverse transcription, labelled with Cy3 using a Low Input Quick Amp Labeling Kit (Agilent Technologies), and hybridized to *S. cerevisiae* 4×44k microarrays. Array scans and data analyses were performed with an Agilent Single-Color DNA Microarray Scanner and Genespring GX software, respectively (Agilent Technologies). Normalized signals were calculated by the 75th percentile normalization procedure.

Glucose, xylose, or glucose/xylose cultivations under aerobic conditions

BY4741/X-I2-P_{xxxx} strains were grown aerobically in SD selective medium overnight at 30°C by shaking in 96-deep-well plates with breathable sealing film (BMBio, Tokyo, Japan). The cells were inoculated into SD or SDX selective medium ($OD_{600} \leq 0.12$), or were concentrated by centrifugation at 1,500 × *g* for 5 min and then inoculated into SX selective medium (OD_{600} = 0.1~0.2) in 96-deep-well plates with breathable sealing film. Cells were grown at 30°C by shaking at 1,200 rpm.

Glucose, xylose, or glucose/xylose fermentations under microaerobic conditions

BY4741/X-I2-P_{xxxx} were grown aerobically in 5-times concentrated SD selective medium supplemented with 2-times concentrated (40 mg/L) histidine and methionine for 72 h at 30°C with shaking in 96-deep-well plates with breathable sealing film. The cells were then concentrated by centrifugation at 1,500 × *g* for 5 min, and were inoculated into SD, SX, or SDX2 selective media in 96-deep-well plates with non-porous seals (PCR-SPS, BMBio) (OD_{600} = 0.5~0.8). Cells were grown at 30°C by shaking at 1,200 rpm.

β-galactosidase activity assays

Promoter activity was measured according to the previously described method (Partow et al. 2010) with slight modification. 0.1 mL of yeast cell culture was mixed with 0.4 mL Z buffer (45 mM Na_2HPO_4, 35 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$), 25 μL 0.1% (w/v) SDS, and 50 μL chloroform in a 96-deep-well plate. The reaction was started through the addition of 0.2 mL of 4 mg/mL *o*-nitrophenyl-β-D-galactopyranoside (ONPG) solution at 28°C and stopped through the addition of 0.5 mL of 1 M $NaCO_3$. After removal of cell debris by centrifugation, ABS_{420} and ABS_{550} were measured with a spectrophotometer (UV-mini 1240, Shimadzu). β-galactosidase activity was determined as Miller units according to the following equation:

$$\text{Miller units} = 1000\times[ABS_{420} - (1.75\times ABS_{550})] / (t\times V\times ABS_{660}) \quad (1)$$

where t represents the reaction time in minutes and V the culture volume (mL) used in the assay.

Accession number

Microarray data analyzed in this study have been deposited in the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and were assigned series record GSE97580.

Results and Discussions

Microarray analyses in xylose- and glucose-fermentation conditions

The xylose-utilizing *S. cerevisiae* strain BY4741/X-I2 was constructed by introducing the *XYL1* and *XYL2* genes from *S. stipitis* (encoding XR and XDH) and the endogenous *XKS1* gene from *S. cerevisiae* (encoding XK) into the genome of BY4741 (Table 1). Ethanol fermentations under microaerobic (oxygen-limited) conditions were tested using BY4741/X-I2 in YP rich media containing xylose and glucose as respective sole carbon sources (Figure 1). BY4741/X-I2 consumed 49 g/L xylose and produced 12 g/L ethanol and 15 g/L xylitol after a 24 h fermentation (Figure 1A), while it completely consumed 50 g/L glucose and produced 23 g/L ethanol after a 6 h fermentation (Figure 1B). To obtain microarray samples in the representative cell growth condition, the times to consume half the carbon sources were determined as approximately 9 h and 3 h in xylose- and glucose-fermentations, respectively. Microaerobic fermentation condition was selected with applications in mind.

At the above determined time points, total RNAs were extracted from the yeast cells and their gene expressions were analyzed by DNA microarray. The microarray data for the values of normalized signals (at the 75th percentile of raw signal values; $\log_2$) both during xylose and glucose fermentations have been deposited to GEO (Accession No. GSE97580).

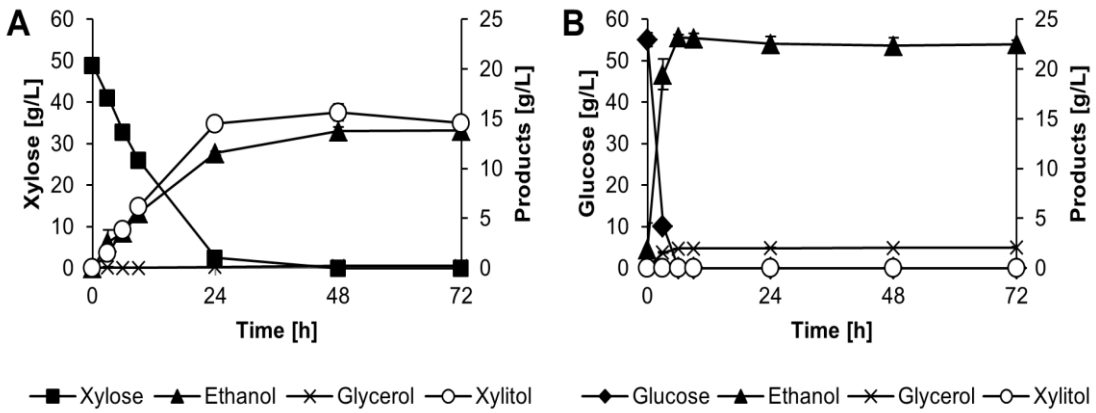


Figure 1 Time-course profiles of ethanol fermentations under microaerobic conditions using the xylose-utilizing BY4741/X-I2 yeast strain.

Fermentations were performed in YP-rich media containing (A) 50 g/L xylose and (B) 50 g/L glucose. Concentrations of each compound were measured by HPLC. Closed diamonds, glucose; closed squares, xylose; closed triangles; ethanol, open circles, xylose; crosses, glycerol. Data represent three independent experiments ± SD.

Selection of promoter candidates for xylose fermentation

Table 4 and Figure 2 show some of the normalized values of the microarray data, which are mainly focused on central metabolism. Similar to glucose fermentation, many genes serve as glycolysis factors, ethanol biosynthesis, and translational elongation factors displayed high normalized signal values during xylose fermentation (e.g. *FBA1*, *TDH1*, *PDC1*, *ADH1*, and *TEF1*) (Table 4 and Figure 2; red columns). In contrast, genes serving as PPP factors, hexose transporters, pyruvate dehydrogenases, and aldehyde dehydrogenases showed moderate (e.g. *NQM1*, *GND1*, *HXT7*, and *ALD4*) (Table 4 and Figure 2; orange columns) or weak (e.g. *RKI1*, *HXT2*, *HXT8*, and *PDA1*) (Table 4 and Figure 2; yellow columns) normalized signals during xylose fermentation. Interestingly, *NQM1* and *GND2* genes showed substantially higher expression with xylose than glucose. *HXT3*, *HXT4*, and *HXK2* genes showed much lower expression with xylose than glucose (Figure 2).

To assess whether the promoters are available for xylose cultivation and fermentation, 29 genes (30 promoters; includes one gene reported to have two different lengths as promoter regions) that showed various expression levels in xylose fermentation were chosen mainly from central metabolic pathways (Table 4 and Figure 2; red, orange, and yellow columns).

Measurements of promoter activities with xylose and glucose under aerobic and microaerobic conditions over time

Along with the *SAG1* terminator (T_{SAG1}) (Yamanishi et al. 2013), the *lacZ* gene was inserted into the multi-cloning site of a single-copy pRS415 vector, yielding

pRS415/*lacZ*-T_{SAG1} control plasmid (Table 1). The 30 selected promoters (Table 2) were then connected to the upstream of *lacZ* gene, generating pRS415/P_{xxxx}-*lacZ*-T_{SAG1} (P_{xxxx}: *xxxx* promoter) (Table 1). The BY4741/X-I2 yeast strain was transformed with plasmid pRS415/P_{xxxx}-*lacZ*-T_{SAG1}, yielding BY4741/X-I2-P_{xxxx}. A control strain, BY4741/X-I2-C, was obtained by the transformation of BY4741/X-I2 with pRS415/*lacZ*-T_{SAG1} (Table 1).

The transformants were cultured respectively in synthetic xylose and glucose media under aerobic and microaerobic conditions. Cells were inoculated into synthetic media in 96-deep-well plates at low turbidity for aerobic cultivations (for glucose: $OD_{600} \leq 0.12$; for xylose: OD_{600}=0.1~0.2) and at moderate turbidity for microaerobic fermentations (for glucose and xylose: OD_{600} = 0.5~0.8). To evaluate the activities of candidate promoters, cells were sampled over time and β-galactosidase assays using *lacZ* genes were performed. The promoter activities measured are shown in Table 5 as a heat map. Time-courses of promoter activities under each condition are displayed in Figure 3.

Because each probe of microarray has a different annealing efficiency, the normalized signal intensities obtained using the 75th percentile procedure should be generally unavailable for relative comparison. However, comparing the data shown in Table 4 (microarray data) and Table 5 (promoter activities), the order of promoter activities roughly correlated with that of the normalized signals in the gene expression analysis. This indicated that a combination of microarray data and reporter gene assays is a useful approach to choose candidate promoters that function in desired culture conditions, although the promoter activities must be carefully checked.

Table 4 Standard deviation of microarray data about pick up promoter

Xylose [log_2]
>5.0
3.0 - 5.0
Below 3.0

	Xylose					Glucose					
Gene Symbol	n=1	n=2	n=3	Average	SD	n=1	n=2	n=3	Average	SD	Probe Name
ADH1	7.07	7.62	7.41	7.37	0.28	6.71	5.61	8.48	6.93	1.45	A_06_P6278
SED1	6.85	7.08	7.02	6.98	0.12	6.21	6.38	6.86	6.48	0.34	A_06_P6281
TDH1	6.53	6.98	7.05	6.85	0.29	6.45	6.23	6.79	6.49	0.28	A_06_P6280
FBA1	6.31	6.60	6.27	6.39	0.18	6.39	6.42	6.68	6.49	0.16	A_06_P6279
CDC19	5.61	5.68	6.12	5.81	0.28	5.94	6.00	6.46	6.13	0.29	A_06_P1062
PDC1	5.53	5.72	6.04	5.77	0.25	5.95	6.46	6.67	6.36	0.37	A_06_P4856
PGK1	5.44	5.37	6.42	5.74	0.58	5.67	5.00	6.21	5.63	0.61	A_06_P1633
TPI1	5.54	5.51	5.92	5.66	0.23	6.03	5.49	6.22	5.92	0.38	A_06_P2037
TEF1	5.21	5.30	6.27	5.59	0.58	5.13	4.86	6.16	5.39	0.68	A_06_P7127
TEF2	5.07	5.09	6.41	5.52	0.77	5.10	4.68	6.70	5.49	1.06	A_06_P1370
ALD4	4.25	4.65	5.21	4.70	0.48	2.96	2.14	3.31	2.81	0.6	A_06_P6738
TDH3	4.37	4.43	4.96	4.59	0.32	4.65	4.23	5.40	4.76	0.6	A_06_P3413
GND1	4.08	4.04	4.82	4.31	0.44	4.22	2.26	4.67	3.72	1.28	A_06_P3761
HXT7	4.07	4.16	3.90	4.04	0.13	4.81	1.20	5.52	3.84	2.32	A_06_P2333
NQM1	3.34	3.46	3.90	3.56	0.29	1.89	0.83	1.87	1.53	0.61	A_06_P3263
LPD1	3.29	3.36	3.39	3.35	0.05	3.33	2.70	3.44	3.15	0.4	A_06_P2844
ADH3	2.95	2.83	3.85	3.21	0.56	3.04	2.19	3.85	3.03	0.83	A_06_P5511
GND2	2.94	2.98	3.29	3.07	0.19	2.25	1.47	2.60	2.11	0.58	A_06_P3477
LAT1	2.89	2.80	2.86	2.85	0.04	3.46	2.43	3.31	3.07	0.56	A_06_P5842
PDA1	2.36	2.40	2.25	2.34	0.08	2.69	2.34	2.43	2.49	0.18	A_06_P2810
SOL4	2.07	2.12	2.41	2.20	0.18	2.21	0.74	2.16	1.70	0.83	A_06_P3469
ALD3	1.79	1.99	1.39	1.72	0.3	1.15	-0.03	0.75	0.62	0.6	A_06_P5604
ZWF1	1.11	1.31	0.95	1.13	0.18	-0.11	1.75	-0.51	0.38	1.2	A_06_P6013
PDC6	0.55	0.63	0.60	0.59	0.04	-2.07	-0.07	-2.17	-1.44	1.18	A_06_P3307
TKL2	-0.06	0.16	0.12	0.07	0.12	-1.49	-1.97	-1.47	-1.64	0.28	A_06_P1369
HXT2	-1.06	-0.50	-0.23	-0.60	0.42	-1.94	-1.34	-1.59	-1.62	0.3	A_06_P5438
RKI1	-2.70	-2.77	-2.21	-2.56	0.31	-2.68	-3.30	-1.97	-2.65	0.67	A_06_P6457
STB5	-3.60	-3.55	-4.10	-3.75	0.3	-3.91	-2.83	-4.41	-3.72	0.81	A_06_P3756
HXT8	-4.49	-5.07	-4.26	-4.61	0.42	-5.08	-4.56	-4.53	-4.72	0.31	A_06_P4236

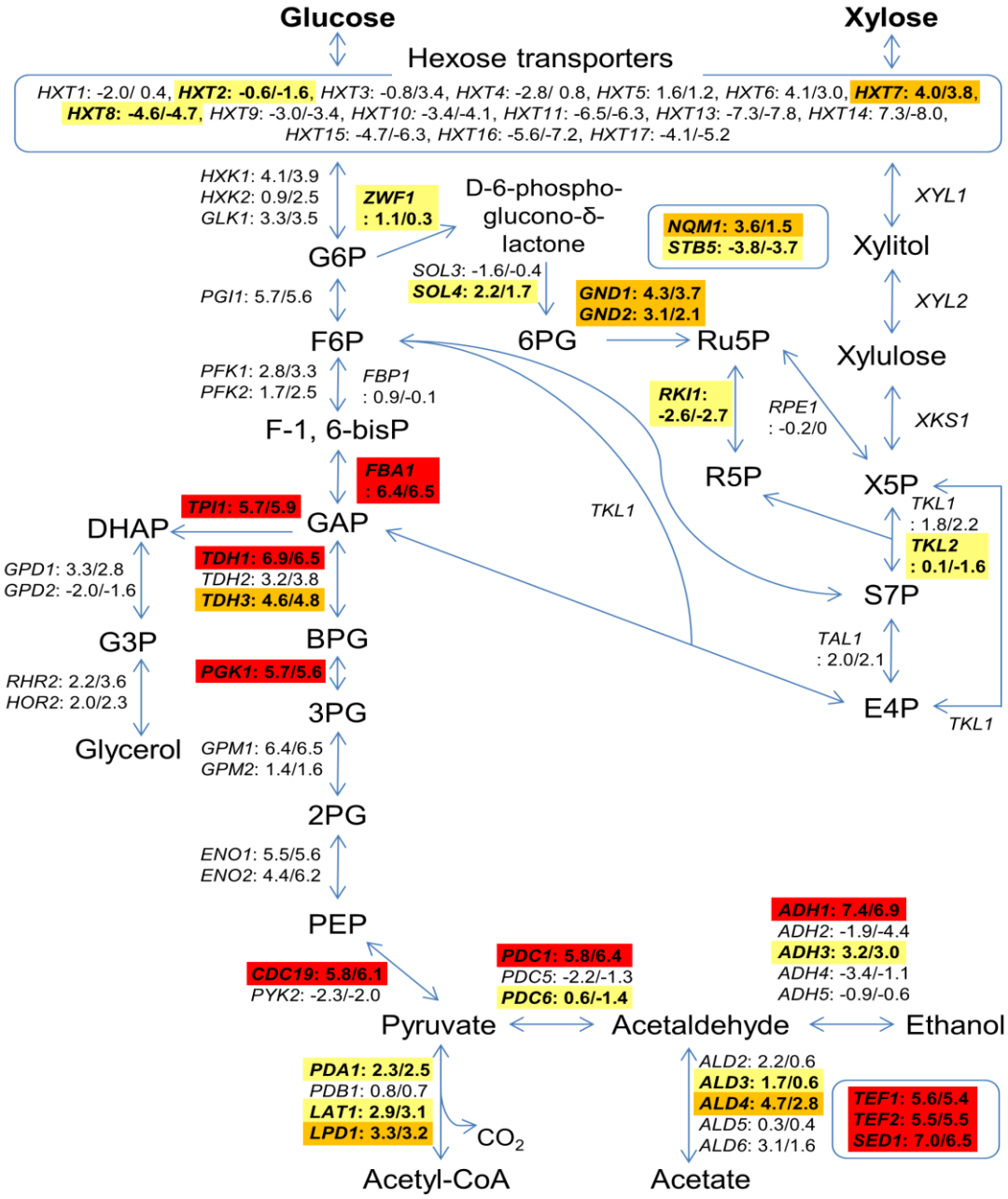
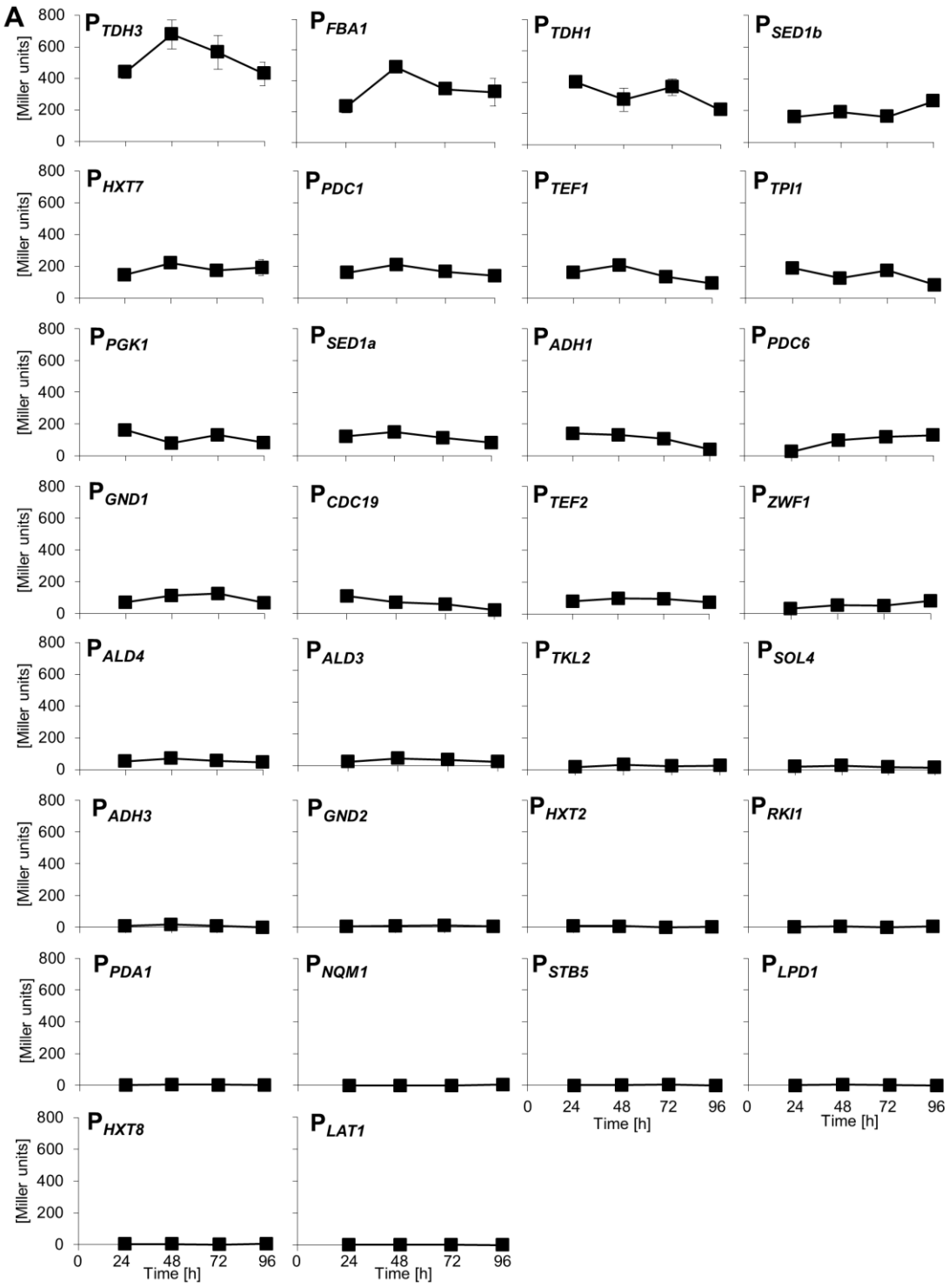
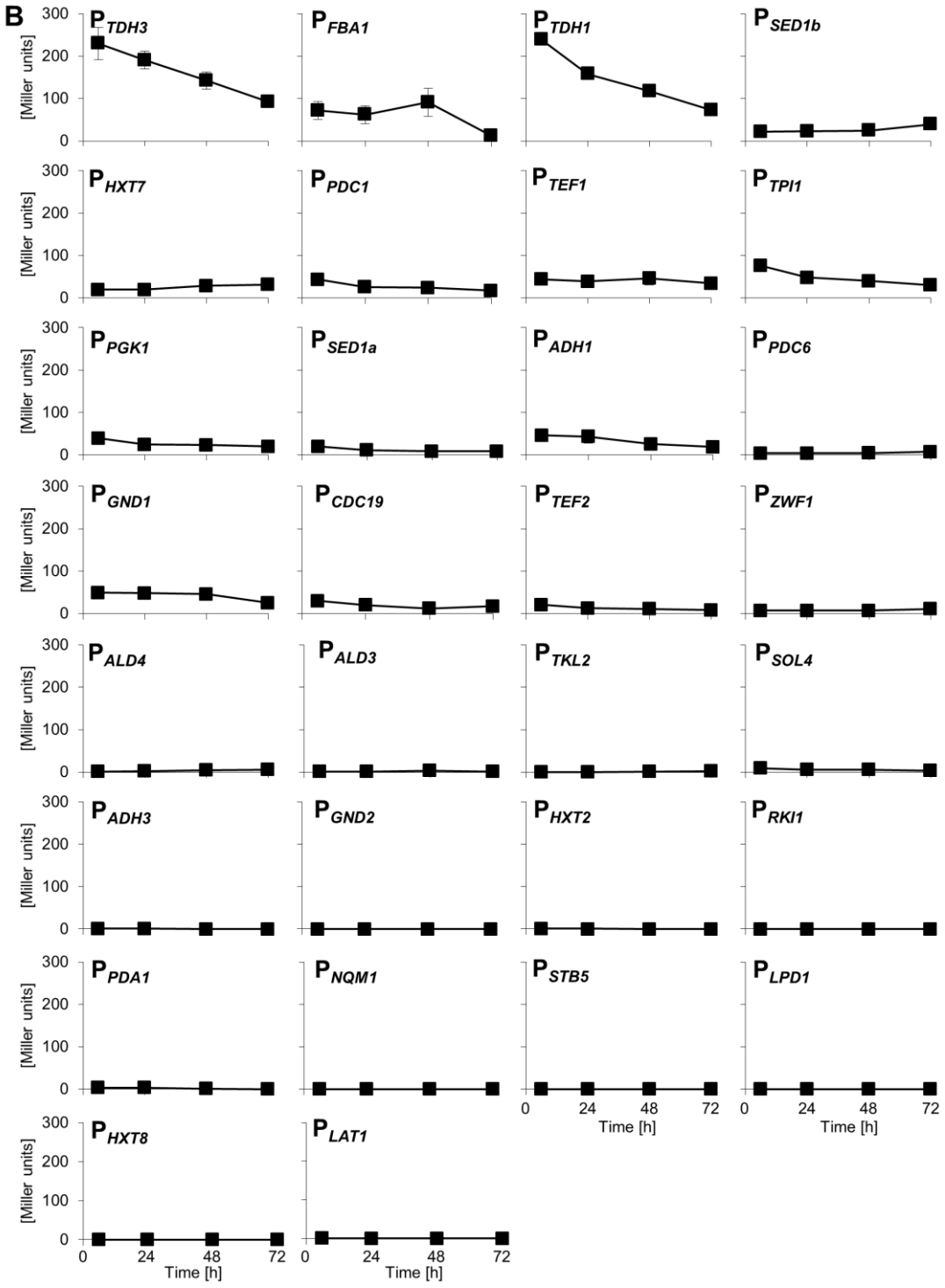
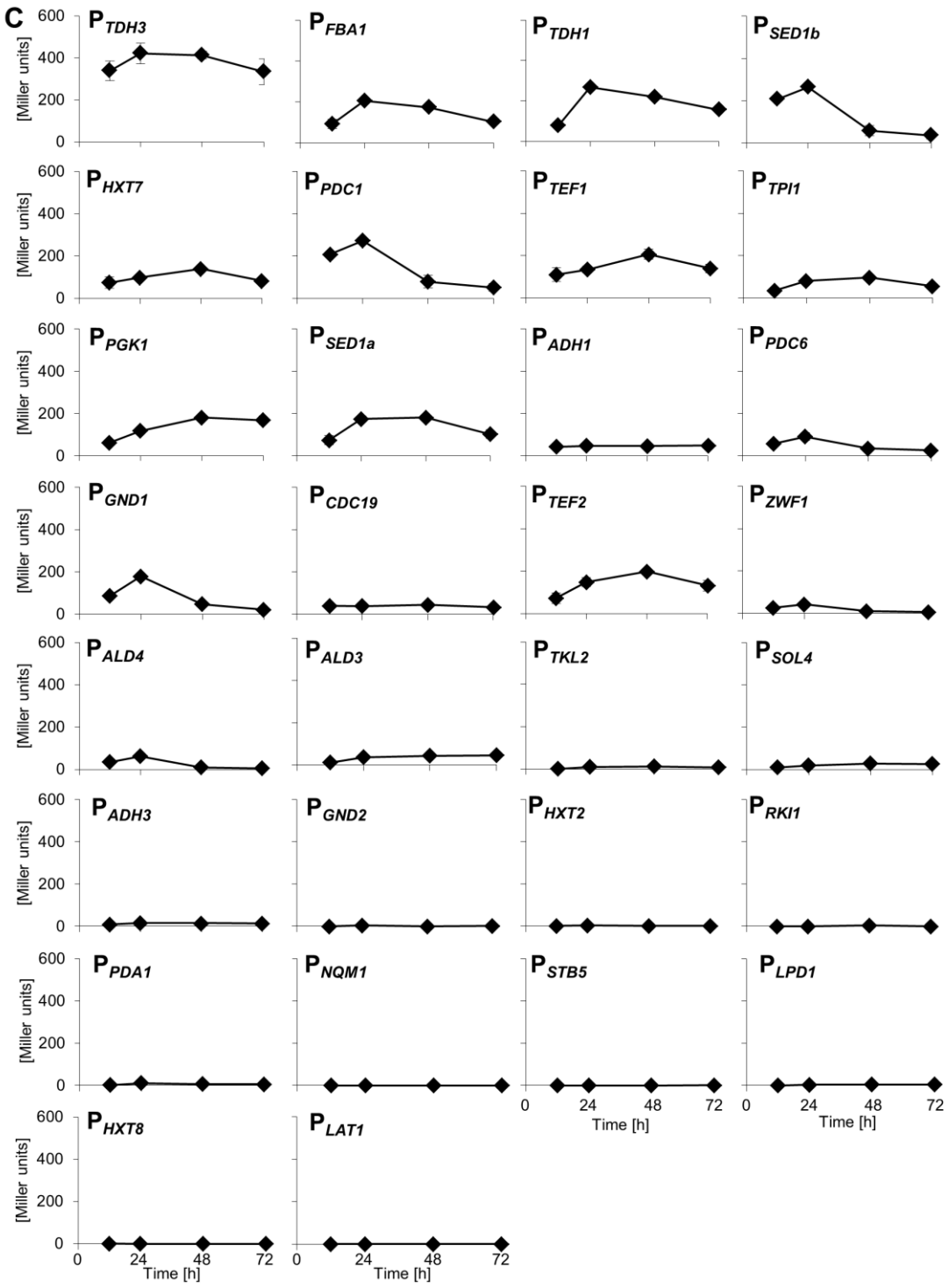


Figure 2 Normalized signals of gene expression in central carbon metabolism.

Microarray analyses were performed using the BY4741/X-I2 strain fermented in YP xylose and glucose media. Total RNAs were extracted from the yeast cells at the times at which half of the carbon source was consumed (9 h and 3 h for xylose- and glucose-fermentations, respectively). Normalized signal intensities (log_2: xylose vs. glucose) of microarray data using the 75th percentile normalization procedure are indicated. Genes marked in bold were subjected to β-galactosidase assays.


A
P TDH3
P FBA1
P TDH1
P SED1b
P HXT7
P PDC1
P TEF1
P TPI1
P PGK1
P SED1a
P ADH1
P PDC6
P GND1
P CDC19
P TEF2
P ZWF1
P ALD4
P ALD3
P TKL2
P SOL4
P ADH3
P GND2
P HXT2
P RKI1
P PDA1
P NQM1
P STB5
P LPD1
P HXT8
P LAT1
[Miller units]
0
200
400
600
800
0
24
48
72
96
Time [h]


B
P TDH3
P FBA1
P TDH1
P SED1b
P HXT7
P PDC1
P TEF1
P TPI1
P PGK1
P SED1a
P ADH1
P PDC6
P GND1
P CDC19
P TEF2
P ZWF1
P ALD4
P ALD3
P TKL2
P SOL4
P ADH3
P GND2
P HXT2
P RKI1
P PDA1
P NQM1
P STB5
P LPD1
P HXT8
P LAT1
[Miller units]
0
100
200
300
0
24
48
72
Time [h]


C
P TDH3
P FBA1
P TDH1
P SED1b
P HXT7
P PDC1
P TEF1
P TPI1
P PGK1
P SED1a
P ADH1
P PDC6
P GND1
P CDC19
P TEF2
P ZWF1
P ALD4
P ALD3
P TKL2
P SOL4
P ADH3
P GND2
P HXT2
P RKI1
P PDA1
P NQM1
P STB5
P LPD1
P HXT8
P LAT1
[Miller units]
0
200
400
600
0
24
48
72
Time [h]

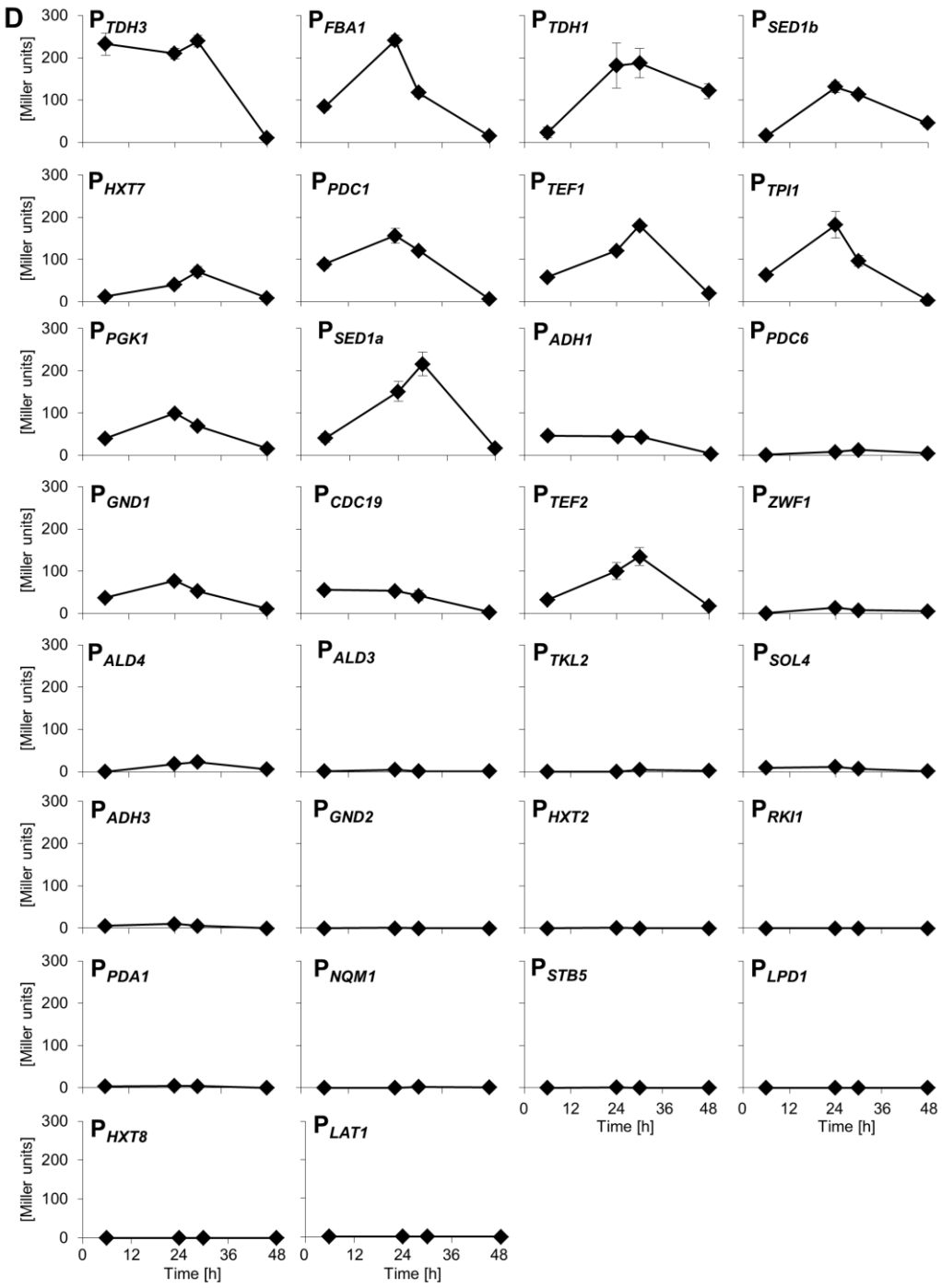


Figure 3 Time-course profiles of 30 promoter activities measured in different culture conditions.

(A) Aerobic cultivation and (B) microaerobic fermentation conditions in xylose media. (C) Aerobic cultivation and (D) microaerobic fermentation conditions in glucose media. β-galactosidase activities of BY4741/X-I2-P$_{xxxx}$ (harboring pRS415/P$_{xxxx}$-*lacZ*-T$_{SAG1}$) at each time point are indicated. Data represent three independent experiments ± SEM.

Table 5 Heatmap list of relative promoter activities during aerobic cultivations and microaerobic fermentations in xylose and glucose media

> 300 miller units
200 - 300 miller units
100 - 200 miller units
20 - 100 miller units
5 - 20 miller units
Below 5 miller units

	Xylose								Glucose							
	Aerobic cultivation				Microaerobic fermentation				Aerobic cultivation				Microaerobic fermentation			
	24 h	48 h	72 h	96 h	6 h	24 h	48 h	72 h	12 h	24 h	48 h	72 h	6 h	24 h	30 h	48 h
Control	0.00	0.19	0.01	0.00	0.06	0.00	0.00	0.00	1.95	0.27	2.48	0.64	0.00	0.00	0.01	0.02
TDH3	439.42	**679.27**	564.89	429.42	**230.34**	191.18	142.83	93.10	340.00	**423.40**	413.97	335.33	233.14	210.19	**239.56**	10.38
FBA1	232.16	**488.02**	345.99	326.68	72.46	62.44	91.39	12.86	92.71	**206.17**	174.61	104.12	84.41	**241.24**	117.14	15.07
TDH1	**398.84**	287.94	367.94	222.94	**239.54**	158.52	118.21	73.32	78.84	**268.46**	220.94	157.36	23.43	181.93	**187.87**	121.99
SED1b	158.97	190.67	161.03	**258.32**	22.02	23.10	24.60	39.89	207.29	**264.89**	58.13	36.09	15.68	**130.93**	113.05	45.50
HXT7	147.08	**220.94**	173.21	193.03	19.36	19.67	28.65	31.49	74.59	98.75	**139.37**	82.02	12.97	40.69	71.63	9.29
PDC1	160.49	**211.29**	166.53	141.35	43.98	25.79	24.47	17.14	208.23	**273.04**	79.59	52.00	88.94	**156.56**	121.06	6.36
TEF1	161.39	**207.54**	135.36	93.13	44.32	39.24	46.37	34.82	110.11	134.12	**206.17**	139.03	58.10	121.56	**181.02**	20.24
TPI1	**189.70**	126.11	173.70	83.59	76.11	47.89	40.07	30.34	34.57	80.80	95.99	55.82	63.97	**182.38**	96.78	3.88
PGK1	**161.46**	80.75	130.79	83.14	38.95	23.99	23.03	19.17	62.58	118.08	**180.44**	167.99	40.20	99.49	69.31	16.24
SED1a	122.39	**151.27**	113.69	82.14	19.63	10.91	8.32	8.03	73.50	174.81	**180.75**	102.51	40.74	150.90	**216.27**	16.79
ADH1	**140.35**	131.83	108.11	39.85	46.14	43.05	25.10	18.61	42.40	46.96	45.96	49.09	46.32	44.78	43.99	3.73
PDC6	26.11	98.67	118.86	**129.73**	3.37	3.23	4.04	6.71	57.02	90.53	35.09	25.17	1.29	8.44	13.01	4.73
GND1	71.28	113.86	**124.32**	67.24	49.47	47.62	45.53	25.49	87.54	**177.75**	47.06	20.71	37.40	77.71	53.26	11.28
CDC19	**110.53**	70.37	58.91	21.59	30.13	20.06	12.14	17.07	39.54	37.98	43.49	32.47	55.98	53.48	42.23	3.35
TEF2	78.76	97.11	92.79	74.05	20.53	13.07	11.02	8.58	72.62	148.89	**197.42**	131.35	32.62	100.87	**135.05**	18.63

ZWF1	32.06	51.80	47.82	78.96	6.47	7.06	6.60	10.90	25.19	41.00	8.60	5.15	0.89	13.96	7.68	6.01
ALD4	51.04	69.45	53.96	45.19	1.10	2.39	4.71	6.10	35.10	62.72	9.12	5.05	0.58	18.08	22.83	6.21
ALD3	23.77	45.16	36.77	24.34	1.57	0.98	3.78	0.81	10.96	35.85	42.29	45.01	0.75	5.03	0.96	0.67
TKL2	14.88	29.31	21.20	23.49	0.07	0.52	0.65	1.84	0.17	9.46	12.15	7.18	0.05	0.27	5.08	2.13
SOL4	17.81	24.39	14.26	13.66	8.87	6.04	5.82	3.42	8.79	18.61	28.52	25.74	8.66	11.38	7.29	1.49
ADH3	9.44	17.21	7.73	0.06	1.26	1.24	0.09	0.08	9.22	15.18	14.75	12.76	5.98	10.08	5.98	0.05
GND2	4.84	8.58	11.69	5.87	0.00	0.09	0.00	0.47	0.00	5.10	0.28	1.29	0.00	0.60	0.00	0.41
HXT2	8.74	7.55	0.81	3.28	1.16	0.59	0.34	0.06	2.90	4.92	2.32	1.28	0.33	1.01	0.07	0.31
RKI1	3.19	5.66	0.75	7.08	0.00	0.00	0.00	0.00	0.17	0.26	4.10	0.67	0.00	0.02	0.10	0.03
PDA1	3.33	6.90	4.28	1.78	2.94	2.94	0.70	0.00	2.07	11.14	7.47	5.50	3.51	4.23	3.88	0.31
NQM1	0.87	0.20	0.32	6.30	0.00	0.00	0.00	0.00	0.00	0.69	0.05	0.00	0.00	0.00	2.43	1.05
STB5	1.21	4.05	5.93	0.00	0.00	0.00	0.00	0.06	0.00	0.64	0.00	0.80	0.00	0.90	0.00	0.09
LPD1	1.26	4.88	2.55	0.38	0.13	0.38	0.16	0.00	0.29	3.35	3.78	4.09	0.19	0.15	0.23	0.08
HXT8	2.64	3.65	0.59	4.60	0.00	0.00	0.00	0.00	1.16	0.51	0.10	0.53	0.21	0.00	0.00	0.10
LAT1	0.30	0.15	0.47	0.00	0.27	0.16	0.00	0.00	0.12	0.26	0.14	0.39	0.40	0.86	0.24	0.02

Promoters for high expression in xylose media both during aerobic and microaerobic conditions

In xylose media, most of selected promoters showed higher activities under aerobic conditions than microaerobic (oxygen-limited) conditions (Figure 3A and 3B). A similar trend was also observed in glucose media (Figure 3C and 3D). The orders of promoter activities were largely consistent between aerobic and microaerobic conditions both in xylose (Figure 3A and 3B) and glucose media (Figure 3C and 3D).

P_{TDH3}, P_{FBA1}, and P_{TDH1} showed the highest promoter activities both in aerobic and microearobic conditions in xylose media, although P_{FBA1} in microaerobic conditions exhibited lower activity compared to the other two promoters (Figure 4). These three promoters also showed extremely high activities in glucose media (Figure 4). P_{TDH3} is well known as one of the strongest promoters and is widely used in *S. cerevisiae* (Ishii et al. 2014). Although P_{FBA1} and P_{TDH1} are less common as promoters for heterologous gene expression in *S. cerevisiae*, they would be user-friendly for use both in glucose and xylose media, similar to P_{TDH3}. When limited to aerobic conditions, the promoter activities of P_{TDH3}, P_{FBA1}, and P_{TDH1} cultured in xylose media were interestingly higher than those cultured in glucose media (Figure 4).In particular, P_{FBA1} showed much higher activity in xylose than in glucose. In a previous study, the *FBA1* promoter showed no significant difference in activity between glucose medium and glycerol/ethanol medium under aerobic cultivation (Compagno et al. 1991). The activity of P_{FBA1} is thus possibly enhanced by xylose metabolism.

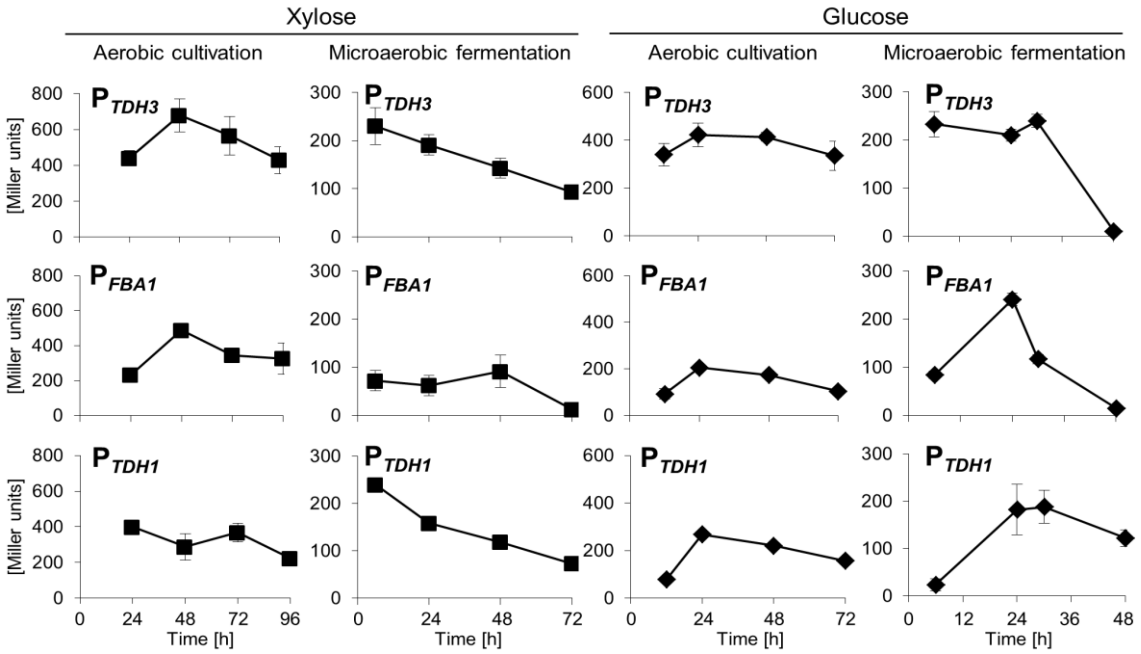


Figure 4 Time-course profiles of promoter activities (*TDH3*, *FBA1*, and *TDH1*) during aerobic cultivations and microaerobic fermentations in xylose and glucose media.

These promoters (*TDH3*, *FBA1*, and *TDH1*) showed high expressions both under aerobic and microaerobic conditions in xylose media. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

Promoters for medium-high and moderate expressions in xylose media respectively during aerobic and microaerobic conditions

P_{PDC1}, P_{TEF1}, P_{TPI1}, and P_{PGK1} serve as promoters for genes for ethanol biosynthesis, translational elongation, and glycolysis, and are also often used for engineered gene expression (Sun et al. 2012; Ishii et al. 2014; Guo et al. 2016). Hxt7p, a high-affinity glucose transporter, has been reported to be an important xylose-transporting protein among the 18 hexose-transporters (Hamacher et al. 2002). Sed1p is a major stress-induced structural glycoprotein that exists abundantly on yeast cell walls (Inokuma et al. 2014). Two different lengths of P_{SED1} have been reported as promoter

regions: one 1063 bp and the other 800 bp (designated as P_{SED1a} and P_{SED1b}, respectively) (Kotaka et al. 2010; Tanino et al. 2007). These seven promoters were available as medium-high expression promoters during aerobic cultivations both in glucose and xylose media (Figure 5). In contrast, the activities of these promoters during microaerobic fermentation in xylose media were modest and weaker than expected, even though most of them showed medium-high expression in glucose fermentation (Figure 5). P_{HXT7} exhibited comparatively lower promoter activity during glucose fermentation (Figure 5). In comparing different length *SED1* promoters, the shorter version (P_{SED1b}) showed higher maximum activity compared to the longer version (P_{SED1a}), except in microaerobic glucose fermentation (Table 5 and Figure 5).

Other promoters for moderate or weak expression in xylose media

P_{ADH1}, P_{GND1}, P_{PDC6}, and P_{CDC19} showed comparatively high promoter activities during aerobic xylose cultivation (Figure 6). Among them, P_{ADH1} and P_{GND1} showed moderate promoter activities during microaerobic xylose fermentation (Figure 6). P_{CDC19} and P_{PDC6} respectively showed weak and almost no activity in *lacZ* reporter assays during microaerobic fermentation both in xylose and glucose media (Figure 6). During both aerobic cultivation and microaerobic fermentation in glucose media, P_{GND1} displayed moderate promoter activities during glucose cultivation and fermentation, while the other promoters showed relatively weak activities (Figure 6). P_{TEF2} showed moderate promoter activity during aerobic xylose cultivation, although it exhibited weak activity during microaerobic xylose fermentation (Figure 6). In contrast, P_{TEF2} showed medium-high promoter activities both during cultivation and fermentation conditions in glucose media (Figure 6).

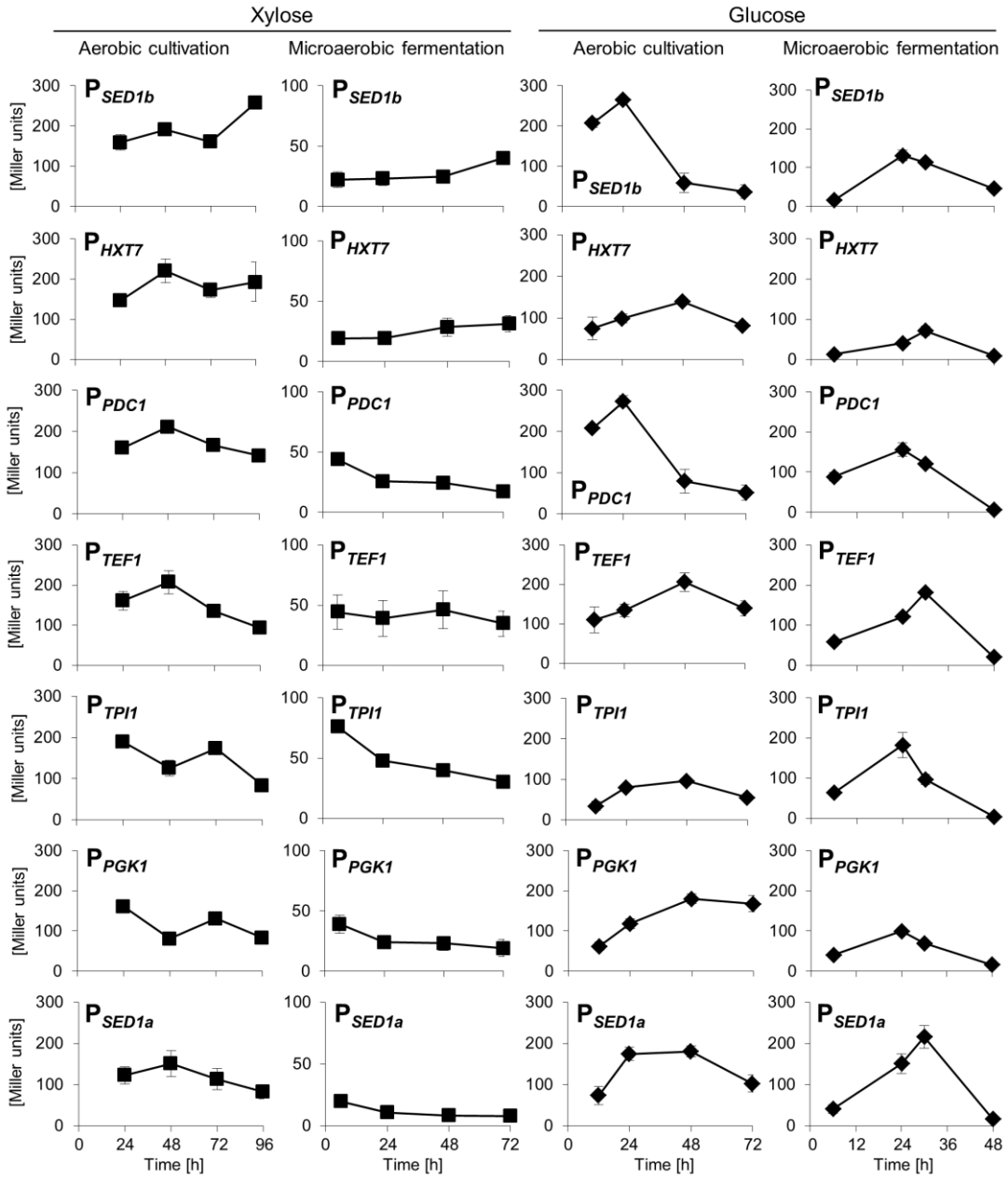


Figure 5 Time-course profiles of promoter activities (*SED1b*, *HXT7*, *PDC1*, *TEF1*, *TPI1*, *PGK1*, and *SED1a*) during aerobic cultivations and microaerobic fermentations in xylose and glucose media.

These promoters (*SED1b*, *HXT7*, *PDC1*, *TEF1*, *TPI1*, *PGK1*, and *SED1a*) showed medium-high and moderate expressions respectively under aerobic and microaerobic conditions in xylose media. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

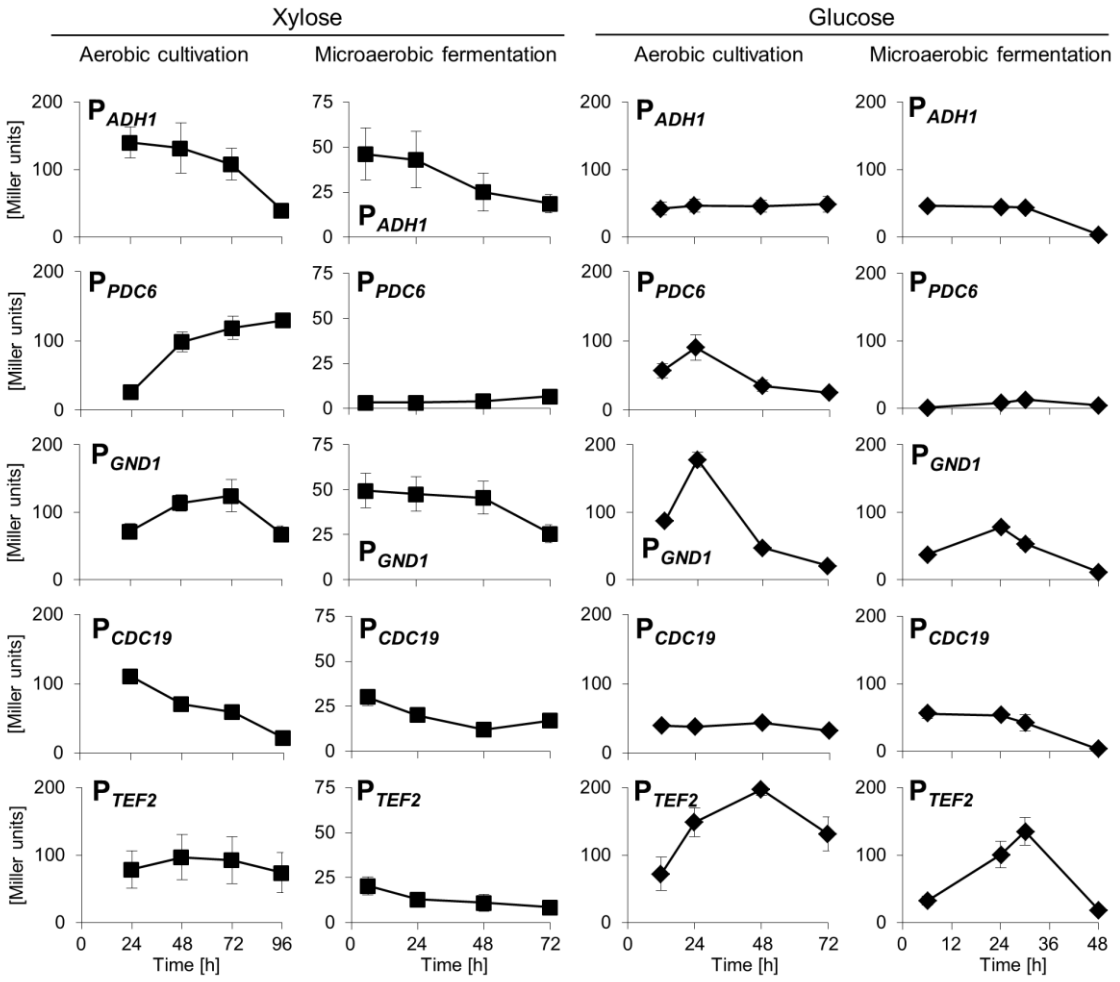


Figure 6 Time-course profiles of promoter activities (*ADH1*, *PDC6*, *GND1*, *CDC19*, and *TEF2*) during aerobic cultivations and microaerobic fermentations in xylose and glucose media.

These promoters (*ADH1*, *PDC6*, *GND1*, *CDC19*, and *TEF2*) showed moderate or weak expression under aerobic and microaerobic conditions in xylose media. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

P_{ZWF1} and P_{SOL4} displayed moderate promoter activities during aerobic cultivations, and considerably weak but definite promoter activities during microaerobic fermentations both in xylose and glucose media (Figure 7). P_{ALD4}, P_{ALD3}, and P_{TKL2} showed moderate promoter activities during aerobic cultivations in both xylose and glucose media, whereas they showed almost no activities in microaerobic fermentations

except for P_{ALD4} in glucose media (Figure 7). P_{ADH3}, P_{GND2}, P_{HXT2}, P_{RKI1}, and P_{PDA1} showed considerably weak promoter activities during aerobic cultivation both in xylose and glucose media, and they showed almost no activities in microaerobic fermentations (Figure 8). Other promoters showed no activity during either aerobic cultivations or microaerobic fermentations in either xylose or glucose media (Figure 8).

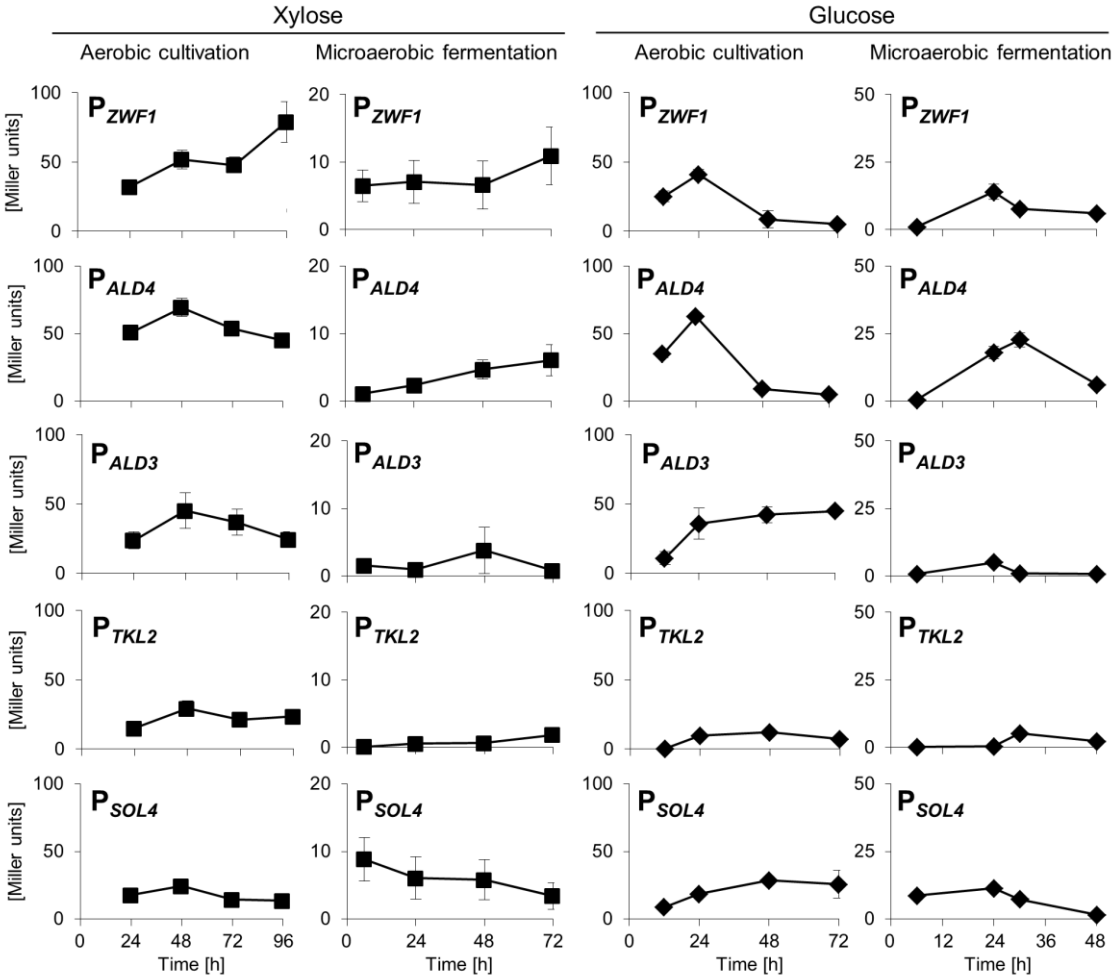


Figure 7 Time-course profiles of promoter activities (*ZWF1*, *ALD4*, *ALD3*, *TKL2*, and *SOL4*) during aerobic cultivations and microaerobic fermentations in xylose and glucose media.

These promoters (*ZWF1*, *ALD4*, *ALD3*, *TKL2*, and *SOL4*) showed moderate or considerably weak expression under aerobic and microaerobic conditions in xylose media. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

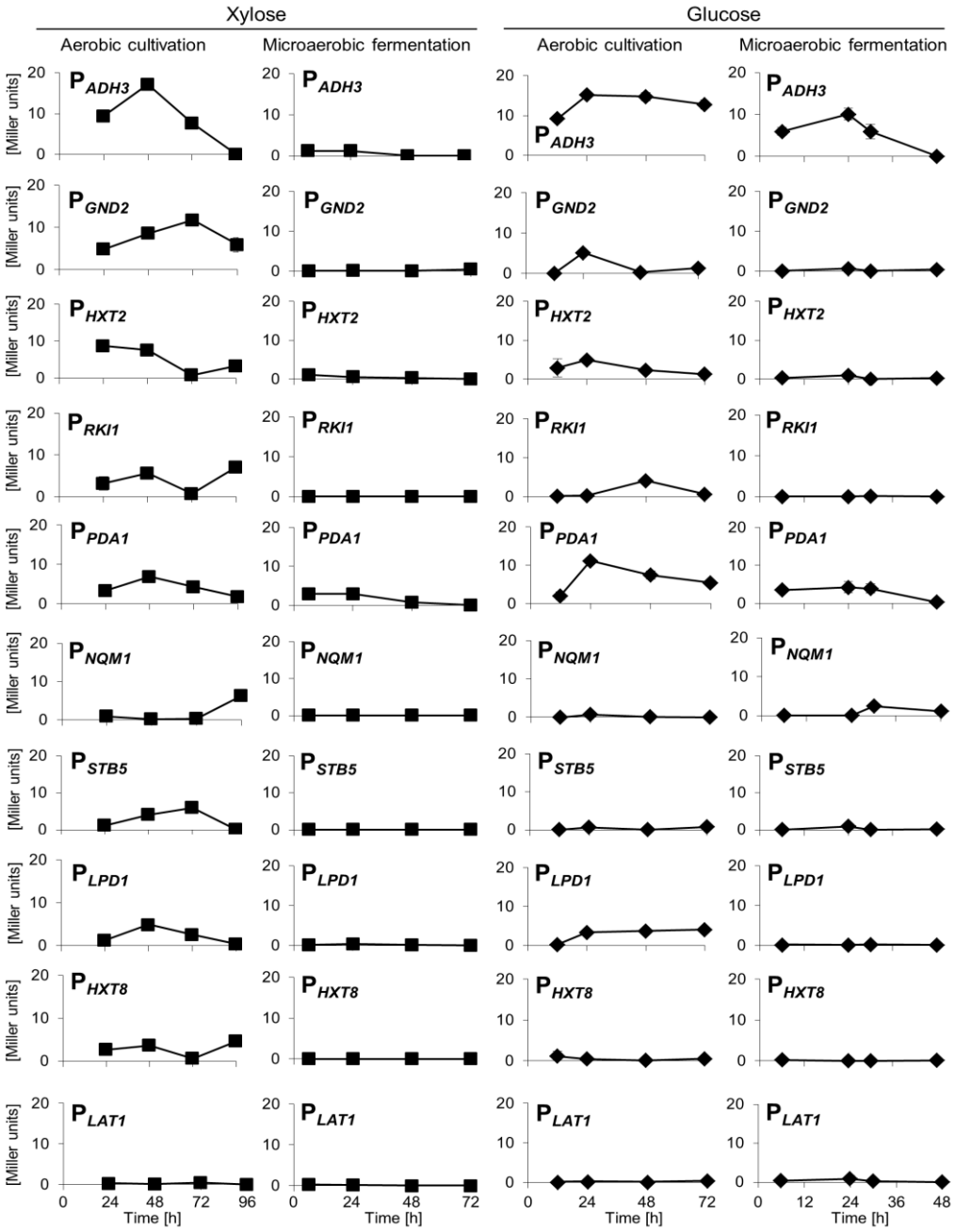


Figure 8 Time-course profiles of promoter activities (*ADH3*, *GND2*, *HXT2*, *RKI1*, *PDA1*, *NQM1*, *STB5*, *LPD1*, *HXT8*, and *LAT1*) promoters during aerobic cultivations and microaerobic fermentations in xylose and glucose media.

These promoters (*ADH3*, *GND2*, *HXT2*, *RKI1*, *PDA1*, *NQM1*, *STB5*, *LPD1*, *HXT8*, and *LAT1*) showed considerably weak or almost no expression under aerobic and microaerobic conditions in xylose media. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

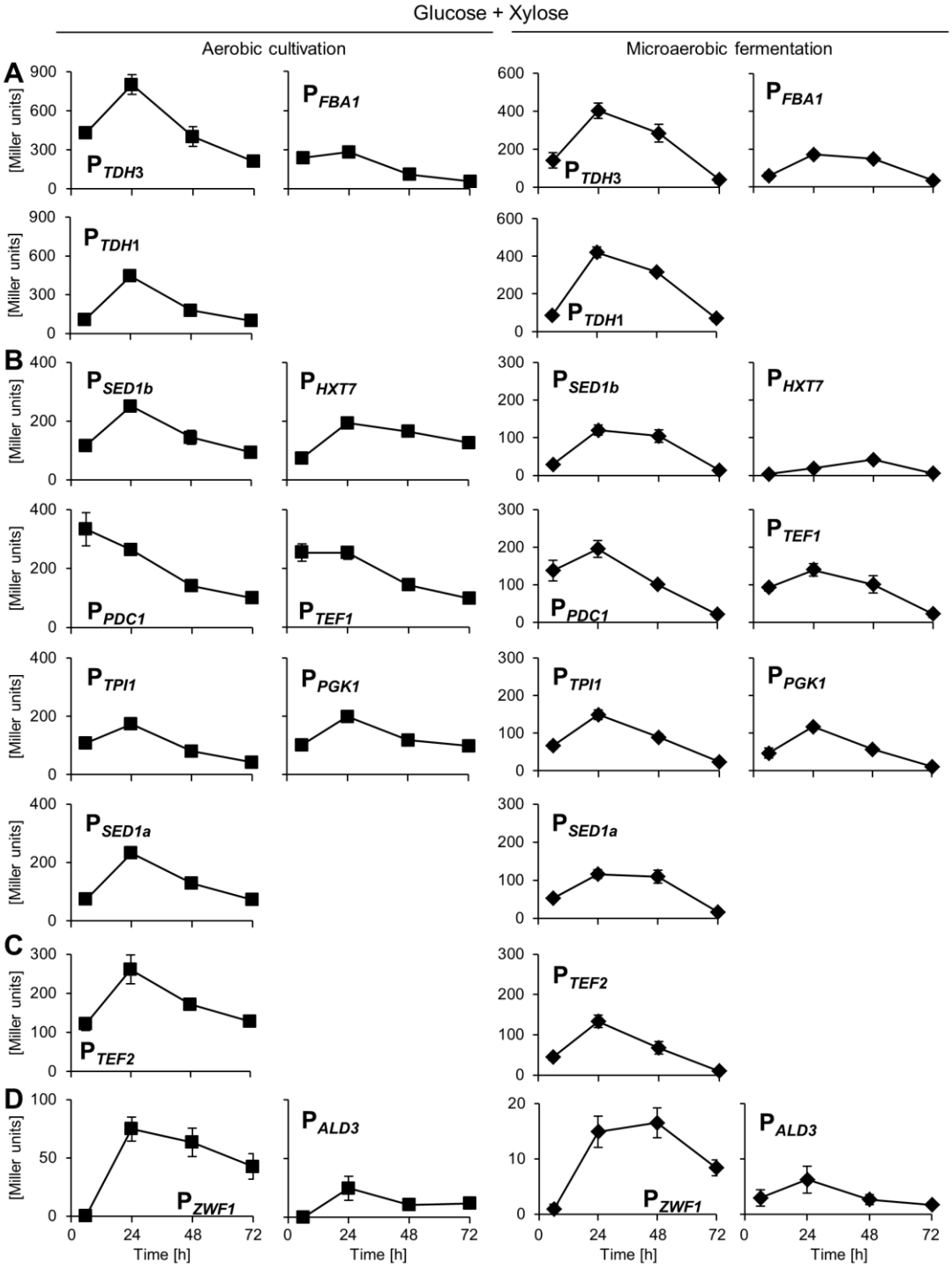

Glucose + Xylose
Aerobic cultivation
Microaerobic fermentation
A
B
C
D
[Miller units]
P TDH3
P FBA1
P TDH1
P SED1b
P HXT7
P PDC1
P TEF1
P TPI1
P PGK1
P SED1a
P TEF2
P ZWF1
P ALD3
Time [h]

Figure 9 Time-course profiles of activities of representative promoters during aerobic cultivations and microaerobic fermentations in synthetic media containing both glucose and xylose.

The promoters, which have showed (A) high expression, (B) moderate expression, (C) weak expression, and (D) considerably weak or almost no expression in xylose media under microaerobic fermentation, were selected for the assays. β-galactosidase activities of BY4741/X-I2-P_{xxxx} (harboring pRS415/P_{xxxx}-*lacZ*-T_{SAG1}) at each time point are indicated. Data represent three independent experiments ± SEM.

Measurements of activities for representative promoters in media containing both xylose and glucose during aerobic and microaerobic conditions

The activities of several representative promoters were further investigated in synthetic media containing both glucose and xylose. The transformants were cultured respectively in glucose/xylose mixed media under aerobic and microaerobic conditions, and the β-galactosidase assays were performed (Figure 9). As a consequence of the assays, most of promoters examined basically showed similar trends of the promoter activities to the cultures in xylose or glucose media (Figure 9A~D).

It has demonstrated that the use of high expression promoters for all of genes does not necessarily result in the best efficiency, and the combinatorial expression approaches are helpful to determine optimal expression profiles (Latimer et al. 2014; Ito et al. 2015). For instance, it is important to optimize the expression levels of a series of genes involved in the metabolic pathways such as xylose assimilation (XR, XDH, and XK) and PPP (e.g., RPE, RKI, TKL, and TAL) for effective xylose utilization in *S. cerevisiae*. Whereas high expression promoters were favorable for XR and TAL, medium or low expression promoters were suitable for other genes especially in micoaerobic condition (Latimer et al. 2014). As another example, tuning copy numbers for three cellulase genes using different sets of promoters and terminators enabled optimization of the degradation ability for Avicel cellulose, conducting higher ethanol production level than the previous constructed strains (Ito et al. 2015). Thus, the results of promoter activities in glucose/xylose mixed media suggested that accumulating the

knowledge of promoters would be useful to make the engineered strains for cultivation or fermentation using hemicellulosic biomass as the carbon source.

In conclusion, by reference to normalized signal values of microarray data, candidate promoters were selected and their activities were measured using *lacZ* reporter gene assays. The promoter activities were quite different depending on the culture conditions, indicating the need for a great deal of attention in promoter selection. The knowledge of promoter activity in xylose cultivation and fermentation obtained in this study will allow for the control of gene expression in engineered xylose-utilizing yeast strains used in hemicellulose fermentation.

References

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, et al., 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(2),115–132.

Compagno C, Ranzi BM, Martegani E., 1991. The promoter of *Saccharomyces cerevisiae* FBA1 gene contains a single positive upstream regulatory element. FEBS Lett. 293(1–2), 97–100.

Eliasson A, Christensson C, Wahlbom CF, Hahn-Hägerdal B., 2000. Anaerobic xylose fermentation by recombinant *Saccharomyces cerevisiae* carrying *XYL1*, *XYL2*, and *XKS1* in mineral medium chemostat cultures. Appl. Environ. Microbiol. 66(8), 3381–3386.

Fernandes S, Murray P., 2010. Metabolic engineering for improved microbial pentose fermentation. Bioeng. Bugs. 1(6), 424–428.

Gietz D, St Jean A, Woods R a, Schiestl RH., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20(6), 1425.

La Grange DC, Pretorius IS, Claeyssens M, Van Zyl WH., 2001. Degradation of Xylan to D-Xylose by Recombinant *Saccharomyces cerevisiae* coexpressing the *Aspergillus niger* β-Xylosidase (*xlnD*) and the *Trichoderma reesei* Xylanase II (*xyn2*) Genes. Appl. Environ. Microbiol. 67(12), 5512–5519.

Guirimand G, Sasaki K, Inokuma K, Bamba T, Hasunuma T, Kondo A., 2016. Cell surface engineering of *Saccharomyces cerevisiae* combined with membrane separation technology for xylitol production from rice straw hydrolysate. Appl. Microbiol. Biotechnol. 100(8), 3477–3487.

Guo W, Sheng J, Zhao H, Feng X., 2016. Metabolic engineering of *Saccharomyces cerevisiae* to produce 1-hexadecanol from xylose. Microb. Cell Fact. 15(1), 24.

Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF, Lidén G, Zacchi G., 2006. Bio-ethanol - the fuel of tomorrow from the residues of today. Trends Biotechnol. 24(12), 549–556.

Hamacher T, Becker J, Gárdonyi M, Hahn-Hägerdal B, Boles E., 2002. Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. Microbiology. 148:2783–2788.

Hasunuma T, Ishii J, Kondo A., 2015. Rational design and evolutional fine tuning of *Saccharomyces cerevisiae* for biomass breakdown. Curr. Opin. Chem. Biol. 29, 1–9.

Hasunuma T, Okazaki F, Okai N, Hara KY, Ishii J, Kondo A., 2013. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. Bioresour. Technol. 135, 513–522.

Hasunuma T, Sanda T, Yamada R, Yoshimura K, Ishii J, Kondo A., 2011a. Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*. Microb. Cell Fact. 10, 2.

Hasunuma T, Sung KM, Sanda T, Yoshimura K, Matsuda F, Kondo A., 2011b. Efficient fermentation of xylose to ethanol at high formic acid concentrations by metabolically engineered *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 90, 997–1004.

Inokuma K, Hasunuma T, Kondo A., 2014. Efficient yeast cell-surface display of exo- and endo-cellulase using the *SED1* anchoring region and its original promoter. Biotechnol. Biofuels. 7,8.

Inokuma K, Yoshida T, Ishii J, Hasunuma T, Kondo A., 2015. Efficient co-displaying and artificial ratio control of α-amylase and glucoamylase on the yeast cell surface by using combinations of different anchoring domains. Appl. Microbiol. Biotechnol. 99, 1655–1663.

Ishii J, Kondo T, Makino H, Ogura A, Matsuda F, Kondo A., 2014. Three gene expression vector sets for concurrently expressing multiple genes in *Saccharomyces cerevisia*e. FEMS Yeast Res. 14, 399–411.

Ismail KSK, Sakamoto T, Hatanaka H, Hasunuma T, Kondo A., 2013. Gene expression cross-profiling in genetically modified industrial *Saccharomyces cerevisiae* strains during high-temperature ethanol production from xylose. J. Biotechnol. 163, 50–60.

Ito Y, Yamanishi M, Ikeuchi A, Imamura C, Matsuyama T., 2015. Combinatorial screening for transgenic yeasts with high cellulase activities in combination with a tunable expression system. PLoS One. 10, 1–16.

Jin Y, Ni H, Laplaza JM, Jeffries TW., 2003. Optimal growth and ethanol production from xylose by recombinant *Saccharomyces cerevisiae* require moderate D – xylulokinase activity. Appl. Environ. Microbiol. 69, 495–503.

Johansson B, Christensson C, Hobley T, Hahn-Hägerdal B., 2001. Xylulokinase overexpression in two strains of *Saccharomyces cerevisiae* also expressing xylose reductase and xylitol dehydrogenase and its effect on fermentation of xylose and lignocellulosic hydrolysate. Appl. Environ. Microbiol. 67, 4249–4255.

Johansson B, Hahn-Hägerdal B., 2002. The non-oxidative pentose phosphate pathway controls the fermentation rate of xylulose but not of xylose in *Saccharomyces cerevisiae* TMB3001. FEMS Yeast Res. 2, 277–282.

Kondo A, Ishii J, Hara KY, Hasunuma T, Matsuda F., 2013. Development of microbial cell factories for bio-refinery through synthetic bioengineering. J. Biotechnol. 163, 204–216.

Kondo T, Tezuka H, Ishii J, Matsuda F, Ogino C, Kondo A., 2012. Genetic engineering to enhance the Ehrlich pathway and alter carbon flux for increased isobutanol production from glucose by *Saccharomyces cerevisiae*. J. Biotechnol. 159, 32–37.

Kotaka A, Sahara H, Kuroda K, Kondo A, Ueda M, Hata Y., 2010. Enhancement of β-glucosidase activity on the cell-surface of sake yeast by disruption of *SED1*. J. Biosci. Bioeng. 109, 442–446.

Latimer LN, Lee ME, Medina-Cleghorn D, Kohnz R a, Nomura DK, Dueber JE., 2014. Employing a combinatorial expression approach to characterize xylose utilization in *Saccharomyces cerevisiae*. Metab. Eng. 25,20–29.

Lodolo EJ, Kock JLF, Axcell BC, Brooks M., 2008. The yeast *Saccharomyces cerevisiae*- the main character in beer brewing. FEMS Yeast Res. 8, 1018–1036.

Lu C, 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.

Matsushika A, Goshima T, Hoshino T., 2014. Transcription analysis of recombinant industrial and laboratory *Saccharomyces cerevisiae* strains reveals the molecular basis for fermentation of glucose and xylose. Microb. Cell Fact. Microbial Cell Factories; 13, 16.

Partow S, Siewers V, Bjørn S, Nielsen J, Maury J., 2010. Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. Yeast. 27, 955–964.

Sakihama Y, Hasunuma T, Kondo A., 2015. Improved ethanol production from xylose in the presence of acetic acid by the overexpression of the *HAA1* gene in *Saccharomyces cerevisiae*. J. Biosci. Bioeng. 119, 297–302.

Salusjärvi L, Kankainen M, Soliymani R, Pitkänen J-P, Penttilä M, Ruohonen L., 2008. Regulation of xylose metabolism in recombinant *Saccharomyces cerevisiae*. Microb. Cell Fact. 7, 18.

Shen MWY, Fang F, Sandmeyer S, Da Silva NA., 2012. Development and characterization of a vector set with regulated promoters for systematic metabolic engineering in *Saccharomyces cerevisiae*. Yeast. 29, 495–503.

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., 2008. Redding A, et al. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. Microb. Cell Fact. 7, 36.

Sun J, Shao Z, Zhao H, Nair N, Wen F, Xu J-H, et al., 2012. Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 109, 2082–2092.

Tanino T, Noguchi E, Kimura S, Sahara H, Hata Y, Fukuda H, et al., 2007. Effect of cultivation conditions on cell-surface display of Flo1 fusion protein using sake yeast. Biochem. Eng. J. 33, 232–237.

Yamanishi M, Ito Y, Kintaka R, Imamura C, Katahira S, Ikeuchi A, et al., 2013. A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae provides* a ″terminatome″ toolbox. ACS Synth. Biol. 2, 337–347.

Chapter 2

Development of a comprehensive set of tools for genome engineering in a cold- and thermo-tolerant *Kluyveromyces marxianus* yeast strain

Introduction

The non-conventional yeast *Kluyveromyces marxianus* is known to be thermotolerant and Crabtree-negative. The latter term indicates that this species produces ethanol only under anaerobic conditions (Nonklang et al. 2008); in contrast, Crabtree-positive yeasts such as *Saccharomyces cerevisiae* obligately produce ethanol both aerobically and anaerobically. Thus, Crabtree-negative yeasts are more suitable for the generation of various products other than ethanol because these strains do not undergo aerobic alcoholic fermentation, which consumes carbon resources (Wagner and Alper 2016). The thermotolerance of *K. marxianus* also allows faster growth and production at higher temperatures, and some volatile products can be collected during synthesis more efficiently at higher temperature. Another benefit of the faster growth of *K. marxianus* at 37°C is the ability to obtain single colonies after overnight incubation, while generally used *S. cerevisiae* strains take 2 to 3 days at 30°C. Recently, the complete genome sequence for *K. marxianus* has been reported (Jeong et al. 2012; Suzuki et al. 2014; Inokuma et al. 2015).

Unlike *S. cerevisiae*, which demonstrates homologous recombination (HR) of transformed DNA, *K. marxianus* DMKU3-1042 tends to integrate DNA fragments randomly into its genome (Nonklang et al. 2008). Such random integration is attributed to the high activity of the non-homologous end-joining (NHEJ) pathway (Abdel-Banat et al. 2009). *K. marxianus* NBRC1777 reportedly needed relatively large homology arms (641 and 1475 bp) of flanking DNA for insertion by HR (Hong et al. 2007), while *S. cerevisiae* requires only ~50-bp homology arms, sequence lengths that can be added at the 5’ ends of PCR primers (Da Silva and Srikrishnan 2012).

Disruption of the NHEJ pathway in yeasts has been shown to suppress random integration events and therefore increase the success rate of HR in these organisms (Abdel-Banat et al. 2009; Valencia et al. 2001). The core NHEJ proteins are Ku70/Ku80, Lig4/Lif1, and Nej1 (Dudášová et al. 2004; Feldmann and Winnacker 1993; Feldmann

et al. 1996; Moore and Haber 1996). In the first step, a stable heterodimer of Ku70/Ku80 binds to broken DNA ends and acts as a bridging complex. Lig4 and Lif1 also form a highly stable complex. Lig4 is an ATP-dependent double-strand break (DSB) repair DNA ligase and Lif1 is Lig4's stabilizing, stimulating, and targeting co-factor (Dudášová et al. 2004). In *S. cerevisiae*, deletion of *NEJ1* reduces NHEJ by 100-fold (Valencia et al. 2001).

Clustered regularly interspaced short palindromic repeats (CRISPR) associated proteins (CRISPR-Cas9) genome-editing tools have been applied in several yeast species, facilitating NHEJ-mediated insertion/deletion formation. HR-mediated integration following CRISPR-Cas9 editing has also been demonstrated, with efficiency varying depending on the yeast species (Wagner and Alper 2016; Jakočiūnas et al. 2016). More recently, deaminase-mediated targeted point mutagenesis (Target-AID) has been performed in *S.cerevisiae*; which is a hybrid system of nuclease-deficient CRISPR-Cas9 and activation-induced cytidine deaminase (AID) (Nishida et al. 2016). When AID ortholog PmCDA1 was fused to C-terminus of nickase Cas9(D10A), C to T/G point mutation was induced within 3-5 bases window at the 5' end of target sequence. Target-AID also appeared to be much less toxic compared to CRISPR-Cas9.

Here, we describe the development of a comprehensive set of genome engineering tools for use with *K. marxianus* NBRC1777, a strain that we show to be both cold- and thermo-tolerant.

Materials and Methods

Strains and culture conditions

S. cerevisiae and *K. marxianus* strains used in this study are listed in Table 1. Yeast cells were grown in YPD medium (10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose) at 30°C unless otherwise noted. *Escherichia coli* strain DH5α (Toyobo, Osaka, Japan) which was used for cloning, was grown in Luria-Bertani (LB) medium (10 g/L peptone, 5 g/L yeast extract, and 5 g/L sodium chloride) supplemented with 100 mg/L ampicillin at 37°C.

Table 1 Plasmids and strains used in this study.

Plasmids and strains	Genotype	Reference
Plasmids		
nCas9-CDA_Base	ScP_{PDC1}_nCas9_T_{TDH3}, KmARS7, Km CEN D, *kanMX*, ori, and AmpR	This study
nCas9-CDA_target *Nej1*	KmP_{SNR52}_target_Nej1_sgRNA_T_{sup4} cassette in nCas9-CDA_Base	This study
nCas9-CDA_target *Dnl4*	KmP_{SNR52}_target_Dnl4_sgRNA_T_{sup4} cassette in nCas9-CDA_Base	This study
Cas9_Base	ScP_{PDC1}_Cas9_T_{TDH3}, KmARS7, Km CEN D, *kanMX*, ori, and AmpR	This study
Cas9_Base_target_*Ura3*_Fw	KmP_{SNR52}_target_Ura3_Fw_sgRNA_T_{sup4} cassette in Cas9_Base	This study
Cas9_Base_target _Ura3_Fw+Rv	KmP_{SNR52}_target_Ura3_Fw_sgRNA_T_{sup4} cassette and KmP_{SNR52}_target_Ura3_Rv_sgRNA_T_{sup4} cassette in Cas9_Base	This study
Cas9_Base_target *Sed1*	KmP_{SNR52}_target_Sed1_sgRNA_T_{sup4} cassette in Cas9_Base	This study
TOPO/F7	ScP_{PDC1}_EGFP_bleomycin on both sides Ura3 homologous sequence cassette in pCR™4Blunt-TOPO®	This study
Strains		
S. cerevisiae		
BY4741	*MAT*a *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	ATCC
K. marxianus		
NBRC1777	Wild-type	NITE Biological Resource Center, Japan
*Nej1*º	*Nej1*disrupted by C to T point mutation at position 13	This study
*Dnl4*º	*Dnl4* disrupted by G to A point mutation at position 44	This study

Growth measurements

S. cerevisiae BY4741 and *K. marxianus* NBRC1777 strains were grown in YPD medium overnight at 30°C by shaking in test tubes. The cells were inoculated into 5 mL of YPD medium at a starting optical density at 600 nm (OD_{600}) of 0.05, and cultures were cultivated at 5, 10, 20, 30, 37, or 45°C. Cell growth was tracked by measuring OD_{600} using a UV mini spectrophotometer (Shimadzu, Kyoto, Japan). The μ_{max} was determined as the slope of the log phase of natural logarithm (ln) transformation of the linear region.

Construction of basal genome-editing plasmids

The vector sequences are provided in Figure 1 and 2. Vector assembly was performed by PCR using the In-fusion cloning method (Takara Bio, Shiga, Japan). The vector backbone was designed to contain KmARS7, KmCEN-D and selectable marker *kanMX*, which encodes G418 resistance. For CRISPR-Cas9 (Cas9_Base) and Target-AID (nCas9-CDA_Base) vectors, a gene encoding human-optimized nCas9-PmCDA1 or Cas9 (Nishida et al. 2016; Dicarlo et al.2013) (respectively) was placed between ScP_{PDC1} and the *TDH3* terminator (ScT_{TDH3}) on the vector backbone.

AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGG
CTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC
CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTAT
GCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAA
GGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGTTAATTAATTGAGATAAGC
ACACTGCACCCATACCTTCCTTAAAAACGTAGCTTCCAGTTTTTGGTGGTTCCGGCTTCCTTCCCGATTCCGCCCGCTAA
ACGCATATTTTTGTTGCCTGGTGGCATTTGCAAAATGCATAACCTATGCATTTAAAAGATTATGTATGCTCTTCTGACTT

TTCGTGTGATGAGGCTCGTGGAAAAAATGAATAATTTATGAATTTGAGAACAATTTTGTGTTGTTACGGTATTTTACTAT
GGAATAATCAATCAATTGAGGATTTTATGCAAATATCGTTTGAATATTTTTCCGACCCTTTGAGTACTTTTCTTCATAAT
TGCATAATATTGTCCGCTGCCCCTTTTTCTGTTAGACGGTGTCTTGATCTACTTGCTATCGTTCAACACCACCTTATTTT
CTAACTATTTTTTTTTTAGCTCATTTGAATCAGCTTATGGTGATGGCACATTTTTGCATAAACCTAGCTGTCCTCGTTGA
ACATAGGAAAAAAAAATATATAAACAAGGCTCTTTCACTCTCCTTGCAATCAGATTTGGGTTTGTTCCCTTTATTTTCAT
ATTTCTTGTCATATTCCTTTCTCAATTATTATTTTCTACTCATAACCTCACGCAAAATAACACAGTCAAATCAATCAAAG
CTAGCGAAATGGCACCGAAGAAGAAGCGTAAAGTCGGAATCCACGGAGTTCCTGCGGCAATGGACAAGAAGTACTCCATT
GGGCTCGATATCGGCACAAACAGCGTCGGTTGGGCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGT
TCTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGAAG
CCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGT
AATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGA
GCGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCATATATCATCTGAGGAAGA
AGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATTTCGGGGACAC
TTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCA
GCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGC
GGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGG
CTGACCCCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGACACCTACGATGATGA
TCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTC
TGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAG
CACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCA
GTCTAAAAATGGCTACGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGG
AAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAAT
GGAAGCATCCCCCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTGAA
AGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCA
GATTCGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGCCTCT
GCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGCT
GTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTCCTGT
CTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGAC
TATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTA
TCACGATCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGTCC
TCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTC
ATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCAATGGGATCCGAGACAAGCA
GAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGATGACT
CTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTT
GCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGAAGGCA
TAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGA

TGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAG
AATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATCGGCTCTC
CGACTACGACGTGGATCATATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCG
ATAAAAATAGAGGGAAGAGTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTG
AACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGC
CGGCTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCATGAACA
CCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGA
AAGGACTTTCAGTTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGG
CACTGCACTTATCAAAAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAA
TGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAG
ACCGAGATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTG
GGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTAC
AGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGAC
CCCAAGAAATACGGCGGATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTC
TAAAAAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACT
TTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAA
AACGGCCGGAAACGAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTAA
TTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAAC
AACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGACGCTAACCTC
GATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATTATCCACTTGTTTAC
TCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGG
AGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGA
GACAGCAGGGCTGACCCCAAGAAGAAGAGGAAGGTGGCATGCATCCAGGCTAGCGTGAATTTACTTTAAATCTTGCATTT
AAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAATGACATTTTCGATTCATTGATTG
AAAGCTTTGTGTTTTTTCTTGATGCGCTATTGCATTGTTCTTGTCTTTTTCGCCACATGTAATATCTGTAGTAGATACCT
GATACATTGTGGATGCTGAGTGAAATTTTAGTTAATAATGGAGGCGCTCTTAATAATTTTGGGGATATTGGCTTTTTTTT
TTAAAGTTTACAAATGAATTTTTTCCGCCAGGATGCGGCCGCTCTAGACTAGTCCTAGGGGATCCGTCACCCGGCCAGCG
GAATTCAGGGATGATCTTGAGAAGTTCTTAGAGTCTTACGAGGGAACAGAAGATTTGGAACCAGCCAAAGCTGCTATAGC
GGAAGCAGATATTTTGCTAAGTAAATGATTAAATAATTAAATATGTGGAAATACATTAATCTTTTTATATTTTTGCAGTT
CGTTGTCGCTATAATTTATAGTCATCTCGTTAGTTCAAACAAGACTTCTTGAAGTGAAAACCAACTTTCAGTCTTCAAAC
TAAAAATGAAAATCAGTGGAAGAAGGTAAACGACTTCATGTTATATATGAATTGAATAGTAATGGAAATAACCAAAAACA
GCTCAACAGAAAACAAACAAAATACGTTAAGACCTGAACTCCTAGCAGAACCATAACTGCCAAATATTTATTATCTGTGG
AGATCTTATATTCTAAAACCAAAAAAAATATAACTTAAAAGTTAAAAAGAAGATGTTCTAACTGAGGTTCGAACTCAGGA
CCTTTGCCGTGTGAAGGCAACGTGATAGCCACTACACTATTAGAACTACCTTATGGGAAAAAGAAAAATAGAGTACAACT
AGAATGGTAAGATCTGTGACCTTTTCTAAACACTTAATTCCATATAGACAGTTCCCACCCACCATAAGGTCACAATTATA
ATGTCTTTAGAAGACCACTGTCGTTCATCATCTTCCTAAGCCCTCTCTCTAAAGCGGCATATTTCCGTAATTTGTTCTTC

TTTGCACAGGCACGTGAGATGACTCCGATTATTCCCACATGCATATTTAGCCTCTCTAGGGGCTCGAGCTATAGCAAGTT
CAAGGAAAGAAAACTATTATGATCTGGTCACGTGTATAAAATTTATTAATTTTAAAACTATATAATTTATTATTTTTTA
TTTTAAAGTTTAAAGTAATTTTAATAGTATTCTATATTTTAAATAAACATACTTTAAAATTTTATTTAATAATTTATTAT
TTTTAAATACAATGTTTTTATTTAAAACAAAATTATAAGTTAAAAAGTTGTTCCGAAAGTAAAATATATTTTATAGGACG
TCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTAC
ATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGG
CTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGA
GAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCT
CACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGAT
GCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCC
CGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACT
GGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCG
ATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTT
CCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAAT
CACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAA
GAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGA
CGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGA
ACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAA
TTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCA
TTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACAT
CATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGC
CTAGGCCGGCCATAGGGATAACAGGGTAATGTTAACCGGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA
CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG
GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA
AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
TGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGG
ATCTTCACCTAGATCCTTTTAAATT

Figure 1 Sequence of Cas9_Base plasmid.

Light green: ScP_{PDC1}, Red: Cas9 protein, Gray: ScT_{TDH3}, Cyan: KmARS7, Green: KmCEN-D, Yellow: *kanMX*

AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGG
CTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC
CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTAT
GCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAA
GGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGTTAATTAATTGAGATAAGC
ACACTGCACCCATACCTTCCTTAAAAACGTAGCTTCCAGTTTTTGGTGGTTCCGGCTTCCTTCCCGATTCCGCCCGCTAA
ACGCATATTTTTGTTGCCTGGTGGCATTTGCAAAATGCATAACCTATGCATTTAAAAGATTATGTATGCTCTTCTGACTT
TTCGTGTGATGAGGCTCGTGGAAAAAATGAATAATTTATGAATTTGAGAACAATTTTGTGTTGTTACGGTATTTTACTAT
GGAATAATCAATCAATTGAGGATTTTATGCAAATATCGTTTGAATATTTTTCCGACCCTTTGAGTACTTTTCTTCATAAT
TGCATAATATTGTCCGCTGCCCCTTTTTCTGTTAGACGGTGTCTTGATCTACTTGCTATCGTTCAACACCACCTTATTTT
CTAACTATTTTTTTTTTAGCTCATTTGAATCAGCTTATGGTGATGGCACATTTTTGCATAAACCTAGCTGTCCTCGTTGA
ACATAGGAAAAAAAAAATATATAAACAAGGCTCTTTCACTCTCCTTGCAATCAGATTTGGGTTTGTTCCCTTTATTTTCAT
ATTTCTTGTCATATTCCTTTCTCAATTATTATTTTCTACTCATAACCTCACGCAAAATAACACAGTCAAATCAATCAAAG
CTAGCGAAATGGCACCGAAGAAGAAGCGTAAAGTCGGAATCCACGGAGTTCCTGCGGCAATGGACAAGAAGTACTCCATT
GGGCTCGCTATCGGCACAAACAGCGTCGGTTGGGCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGT
TCTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGAAG
CCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGT
AATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGA
GCGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCATATATCATCTGAGGAAGA
AGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATTTCGGGGACAC
TTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCA
GCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGC
GGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGG
CTGACCCCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGACACCTACGATGATGA
TCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTC
TGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAG
CACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCA
GTCTAAAAATGGCTACGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGG

AAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAAT
GGAAGCATCCCCCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAA
AGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCA
GATTCGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGCCTCT
GCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGCT
GTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTCCTGT
CTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGAC
TATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTA
TCACGATCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGTCC
TCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTC
ATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCAATGGGATCCGAGACAAGCA
GAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGATGACT
CTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTT
GCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGAAGGCA
TAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGA
TGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAG
AATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATCGGCTCTC
CGACTACGACGTGGATCATATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCG
ATAAAAATAGAGGGAAGAGTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTG
AACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGC
CGGCTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCATGAACA
CCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGA
AAGGACTTTCAGTTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGG
CACTGCACTTATCAAAAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAA
TGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAG
ACCGAGATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTG
GGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTAC
AGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGAC
CCCAAGAAATACGGCGGATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTC
TAAAAAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACT
TTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAA
AACGGCCGGAAACGAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTAA
TTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAAC
AACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGACGCTAACCTC
GATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATTATCCACTTGTTTAC
TCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGG

AGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGA
GACAGCAGGGCTGACCCCAAGAAGAAGAGGAAGGTGGGTGGAGGAGGTACCGGCGGTGGAGGCTCAGCAGAATACGTACG
AGCTCTGTTTGACTTCAATGGGAATGACGAGGAGGATCTCCCCTTTAAGAAGGGCGATATTCTCCGCATCAGAGATAAGC
CCGAAGAACAATGGTGGAATGCCGAGGATAGCGAAGGGAAAAGGGGCATGATTCTGGTGCCATATGTGGAGAAATATTCC
GGTGACTACAAAGACCATGATGGGGATTACAAAGACCACGACATCGACTACAAAGACGACGACGATAAATCAGGGATGAC
AGACGCCGAGTACGTGCGCATTCATGAGAAACTGGATATTTACACCTTCAAGAAGCAGTTCTTCAACAACAAGAAATCTG
TGTCACACCGCTGCTACGTGCTGTTTGAGTTGAAGCGAAGGGGCGAAAGAAGGGCTTGCTTTTGGGGCTATGCCGTCAAC
AAGCCCCAAAGTGGCACCGAGAGAGGAATACACGCTGAGATATTCAGTATCCGAAAGGTGGAAGAGTATCTTCGGGATAA
TCCTGGGCAGTTTACGATCAACTGGTATTCCAGCTGGAGTCCTTGCGCTGATTGTGCCGAGAAAATTCTGGAATGGTATA
ATCAGGAACTTCGGGGAAACGGGCACACATTGAAAATCTGGGCCTGCAAGCTGTACTACGAGAAGAATGCCCGGAACCAG
ATAGGACTCTGGAATCTGAGGGACAATGGTGTAGGCCTGAACGTGATGGTTTCCGAGCACTATCAGTGTTGTCGGAAGAT
TTTCATCCAAAGCTCTCATAACCAGCTCAATGAAAACCGCTGGTTGGAGAAAACACTGAAACGTGCGGAGAAGCGGAGAT
CCGAGCTGAGCATCATGATCCAGGTCAAGATTCTGCATACCACTAAGTCTCCAGCCGTTGGTCCCAAGAAGAAAAGAAAA
GTCGCATGCATCCAGGCTAGCGTGAATTTACTTTAAATCTTGCATTTAAATAAATTTTCTTTTTATAGCTTTATGACTTA
GTTTCAATTTATATACTATTTTAATGACATTTTCGATTCATTGATTGAAAGCTTTGTGTTTTTTCTTGATGCGCTATTGC
ATTGTTCTTGTCTTTTTCGCCACATGTAATATCTGTAGTAGATACCTGATACATTGTGGATGCTGAGTGAAATTTTAGTT
AATAATGGAGGCGCTCTTAATAATTTTGGGGATATTGGCTTTTTTTTTTAAAGTTTACAAATGAATTTTTTCCGCCAGGA
TGCGGCCGCTCTAGACTAGTCCTAGGGGATCCGTCACCCGGCCAGCGGAATTCAGGGATGATCTTGAGAAGTTCTTAGAG
TCTTACGAGGGAACAGAAGATTTGGAACCAGCCAAAGCTGCTATAGCGGAAGCAGATATTTTGCTAAGTAAATGATTAAA
TAATTAAATATGTGGAAATACATTAATCTTTTTATATTTTTGCAGTTCGTTGTCGCTATAATTTATAGTCATCTCGTTAG
TTCAAACAAGACTTCTTGAAGTGAAAACCAACTTTCAGTCTTCAAACTAAAAATGAAAATCAGTGGAAGAAGGTAAACGA
CTTCATGTTATATATGAATTGAATAGTAATGGAAATAACCAAAAACAGCTCAACAGAAAACAAACAAAATACGTTAAGAC
CTGAACTCCTAGCAGAACCATAACTGCCAAATATTTATTATCTGTGGAGATCTTATATTCTAAAACCAAAAAAAATATAA
CTTAAAAGTTAAAAAGAAGATGTTCTAACTGAGGTTCGAACTCAGGACCTTTGCCGTGTGAAGGCAACGTGATAGCCACT
ACACTATTAGAACTACCTTATGGGAAAAAGAAAAATAGAGTACAACTAGAATGGTAAGATCTGTGACCTTTTCTAAACAC
TTAATTCCATATAGACAGTTCCCACCCACCATAAGGTCACAATTATAATGTCTTTAGAAGACCACTGTCGTTCATCATCT
TCCTAAGCCCTCTCTCTAAAGCGGCATATTTCCGTAATTTGTTCTTCTTTGCACAGGCACGTGAGATGACTCCGATTATT
CCCACATGCATATTTAGCCTCTCTAGGGGCTCGAGCTATAGCAAGTTCAAGGAAAGAAAACTATTATGATCTGGTCACGT
GTATAAAATTTATTAATTTTAAAACTATATAATTTATTATTTTTTTATTTTAAAGTTTAAAGTAATTTTAATAGTATTCT
ATATTTTAAATAAACATACTTTAAAATTTTATTTAATAATTTATTATTTTTAAATACAATGTTTTTATTTAAAACAAAAT
TATAAGTTAAAAAGTTGTTCCGAAAGTAAAATATATTTTATAGGACGTCGACATGGAGGCCCAGAATACCCTCCTTGACA
GTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTT
GCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGC
TCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGT
TCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTA
AGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGAT

AATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAA
AGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCA
AGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAA
GAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAA
TTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTG
ATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGAT
TCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGG
ACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGA
AACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTC
TAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTA
TTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAA
AGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCCTAGGCCGGCCATAGGGATAACAGGGTAATGTT
AACCGGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC
AGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGAT
ACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC
GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTG
CAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT

Figure 2 Plasmid nCas9-CDA_Base sequences.

Light green: ScP_{PDC1}, Red: nCas9 protein, Gray: ScT_{TDH3}, Cyan: KmARS7, Green: KmCEN-D, Yellow: *kanMX*

Construction of target sequence-containing plasmids

To generate the sgRNA expression casettes, the *K. marxianus SNR52* promoter (KmP_{SNR52}) (531 bp; identified by alignment with the *S. cerevisiae* gene) was amplified by PCR using the combination of the forward primer (P_Km01-013) and each reverse primer (5'-CTAGCTCTAAAAC- reverse complement of target sequence-GATTCGAACTGCGGACGTTG-3'). The sgRNA scaffold and the *SUP4* terminator were amplified by PCR from plasmid pRS426-SNR52p-gRNA.CAN1.Y-SUP4t (Dicarlo et al. 2013) using the primers

P_Km01-014 and P_Km01-015 (Table 2). The two fragments were fused by overlap extension PCR using the primers P_Km01-013 and P_Km01-015 to yield each target-containing sgRNA cassette (Figure 3). Each cassette was digested with restriction enzymes NotI and NheI and ligated into NotI-, SpeI-digested nCas9-CDA_Base or Cas9_Base plasmid.

Transformation

K. marxianus cells were transformed by the lithium acetate method (Gietz et al. 1992). After transformation, cells were plated on YPD containing appropriate selection reagents (100 μg/mL G418 and/or 50 μg/mL zeocin) and grown for one day. For the 5-FOA test, cells were streaked onto YPD containing 3 mg/mL 5-FOA, incubated overnight, and assessed for growth. Colony PCR was performed to check fragment size and to obtain DNA for Sanger sequencing using a 3130xL Genetic Analyzer (Applied Biosystems, CA, USA).

Table 2 Primers used in this study

Primer name	Sequence (5'-3')	Purpose
Universal primers for sgRNA cassette construction		
P_Km01_013(pKmSNR52_Fw)	aagctagctctagaTGCCGCAAATAGGGCAGG	amplification of $P_{KmSNR52}$ region
P_Km01_014(tracerRNA_Fw)	GTTTTAGAGCTAGAAATAGCAAG	amplification of tracrRNA_T_{sup4} fragment
P_Km01_015(tracerRNA_Rv)	cctgaattcGCGGCCGCatactagtCATAAAAAACAAAAAAAGCACCACC	amplification of tracrRNA_T_{sup4} fragment
Target-specific primers		
P_Km01_016(single1(NEJ1)_Fw)	CTAGCTCTAAAACCTTATTATCAGTATTTTGTAGATTCGAACTGCGGACGTTG	amplification of target *NEJ1*
P_Km01_017(single2(DNL4)_Fw)	CTAGCTCTAAAACTCTCCAGAGTTTAAATGGTTGATTCGAACTGCGGACGTTG	amplification of target *Dnl4*
P_Km01_018(deletion1(Ura3)_Fw)	CTAGCTCTAAAACCCTTTGATTAGTTTATGGGCGATTCGAACTGCGGACGTTG	amplification of target Ura3_Fw
P_Km01_019(deletion2(Ura3)_Rv)	CTAGCTCTAAAACAGAACAAGCAGAGGTTCGAGGATTCGAACTGCGGACGTTG	amplification of target Ura3_Rv
P_Km01_059(deletion2(Sed1)_Fw)	CTAGCTCTAAAACCTGGAATAGAACTGGATGTGGATTCGAACTGCGGACGTTG	amplification of target *Sed1*-2

Primer pairs for insert donor DNA

P_Km01_039(Ura3(5)_Fw)	ACCTTAAGTAGTCAAACAAATTGTG	amplification of GFP and bleomycin expression donor DNA containing length of 1000 bp *Ura3* homologous sequence (donor DNA-I)
P_Km01_046(Ura3(3)_Rv)	CAGAGAGTCTGCAACACCGA	
Ura3_up_100F	TGAATCTGAGCTCATCACCAAC	amplification of GFP and bleomycin expression donor DNA containing length of 122/109 bp *Ura3* homologous sequence (donor DNA-II)
Ura3_down100_R	GTATACAATGTGCGCAATGCC	
P_Km01_052(Ble_GFP_50bp_Rv)	TGGAACAGCAAGAGGAAGTATC	amplification of GFP expression donor DNA containing length of 50 bp *Ura3* homologous sequence (donor DNA-III)
P_Km01_053(GFP_50bp_Fw)	GATCAACTTGTAACTTCTGATCCGAGTACACTCGAA CCTCTGCTTGTTCTgatcccttgtacagctcgtc	
P_Km01_060(GFP_Sed1_Fw)	TATTCGTGGTACCAGTTTTATTTGCATCATCTGCATT TGCTGCTTTCTCTatggtgagcaagggcga	amplification of GFP codin region donor DNA containing length of 50 bp *Sed1* homologous sequence (donor DNA-IV)
P_Km01_061(GFP_Sed1_Rv)	gaaGATGGGGCAACTGAGGTCACtACTGGAATAGAA CTGGATGTGTCGTTacccttgtacagctcgtcca	
P_Km01_052(Ble_GFP_50bp_Rv)	TGGAACAGCAAGAGGAAGTATC	amplification of P_{PDC1} region donor DNA containing length of 50 bp homologous sequence (donor DNA-V)
P_Km01_064(pPDC1_Fw)	atgggcaccaccccggtgaa	
P_Km01_065(EGFP_Rv)	GCTAGCatggtgagcaaggg	amplification of GFP coding region donor

P_Km01_066(EGFP_Fw)	AACTATACAAATGACAAGTTCTTGAAAACAAGAAT CTTTTTATTGTCAGTgatcccttgtacagctcgtc	DNA containing length of 50 bp homologous sequence (donor DNA-VI)
P_Km01_051(Ble_GFP_50bp_Fw)	GATCAACTTGTAACTTCTGATCCG	amplification of T_{TEF1} region donor DNA containing length of 50 bp homologous sequence (donor DNA-VII)
P_Km01_067(TEF1t_Rv)	ACTGACAATAAAAAGATTCTTGT	

Red indicates each target sequence.

AAGCTAGCTCTAGATGCCGCAAATAGGGCAGGCGATATACAAGGCGGTATACAGGGCGGTATACAAGGCGTATGCCAGAC
GGTATGCCTGGAGGCATTGTCAGGAGGCATTGTCAGGAGGCATTGCCGGGCGAAATCGCCAGCTTGAGTATGGCCCCAGA
GGGAAGAAAAAACAGTTACCCAGGTGTGTGCCCCTGTATTACCCTGACTTTGAAATGTACGCGAAGTACACTTCGTGCAG
GGTGATGATGCTCGTATCAGGTGGACTTGGCGGACACGGTACGGTACGGTCCCCTTTTTTTATTTTTTCTTTTGGGAATG
AGAAGGCCACGTGACGTTACCCTGGCCCTGTGTTTCCCCCCTCGCATCTTGTACTGCCTTAGGACCCCTCGTAGTGTAGC
GGTCCGGGCCCCCAGATTGCGATGTGTTTCGAGAAAAACAACAAAAAAAATAAGTTGAAAAGCGCGGATCAGGCGGTCCG
CCCGGAGTGCGCGCGCTGTATTCTCGCCTGACAGGGCGGGATTCGCAACGTCCGCAGTTCGAATCNNNNNNNNNNNNNNN
NNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTG
GTGCTTTTTTTGTTTTTTATGACTAGTATGCGGCCGCGAATTCAGG

Figure 3 Sequence of target-containing sgRNA cassette.

"N" indicates target specific sequence. This sequence indicates as follows: Purple: KmP_{SNR52}, Gray: sgRNA, Green: Sup4 terminator.

Donor DNA templates

Each donor DNA template was constructed by PCR and cloned into the plasmid TOPO/F7 using the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen, CA, USA) (Figure 4). Each cloned donor DNA was either excised by restriction enzymes or amplified by PCR using the respective primer pair (Table 2), and the resulting fragment was purified by agarose gel electrophoresis before use in transformation.

AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGG
AAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCC
GGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCA
GAATTAACCCTCACTAAAGGGACTAGTCCTGCAGGTTTAAACGAATTCGCCCTTCAGAGAGTCTGCAACACCGATTTATC
GGTTTGCAGGTAACACAATACCGATTACTTGGTGTCTGAGGAGATCTCAATTCTCTCTTTTTCCTTCAACTTTTATTTTC
TTGGTATGCTTCTTGTGCATTCGCCTATTAATTTATATTAGCGATTCTTATGACATCGGCCATCACTGAGTATAGTGAGT
AGTAGTGGTTGATATTTGAAATCTATTATAGATAAGTTCCTTTACATATCGTGCAGCACTTGTTTATGTGAACTATTAGT
AGCTATATTGGCTATATACAGAGAGAAAGCTATAACTATTTATATCATAACTGTTTGTATAACTTAACGACCCATGCTTT
GGTTCAAAGCAGCGTTTTGCTCTTGGAAACGTTGTCCAACACGTTCACTATGTTTCAAGAACTTTTCTGAGCATTTCATT
ATGCAGCTTTGCTCCTTGGAAGTTAGCTTAGCAGAGGTAAAGTCGTTGACACAGTCACTGAAACATCTTTCGACCAAGTT
GGAGTATAGACGCATGAAGTCCTTCATTTGCTTTTGTTCCACAATTCTTTGGAACTCTTGTTGTTCCTTACCATTTAGTT
GATCCATTTTTAACTGTGTATTTGATTTTTTTGTTCTAGCAATCGACGGTCCAAGACTCAATAAAAATTGACTCTTGTAT
CTTTCTTCACACTAGTAGTAAAGCAGAAAGTCTAGCAGAAATGTGTATCTTTTATAACGGTCAATGGTCTAAGAAGCAAG
GTTTGGTTCTTCGCACTTGTATTAAACTCTTTTTCTTTGGTTGTGAAATTTCACATAATGTACTTCACCAAAAAACAAAA

AAAAAAAAAAAAAAAAAATTTTCGGTGCAAAAAACAGCTTCTAGCACGTGACTAAACTTGTAGCGTATACAATGTGCGC
AATGCCATTACACTCTAACATCTCTGTATAGTTTATATATAAACGATCAACTTGTAACTTCTGATCCGAGTACACTCGAA
CCTCTGCTTGTTCTCAGTATAGCGACCAGCATTCACATACGATTGACGCATGATATTACTTTCTGCGCACTTAACTTCGC
ATCTGGGCAGATGATGTCGAGGCGAAAAAAAATATAAATCACGCTAACATTTGATTAAAATAGAACAACTACAATATAAA
AAAACTATACAAATGACAAGTTCTTGAAAACAAGAATCTTTTTATTGTCAGTACTGATCATGAGATGCCTGCAAGCAATT
CGTTCTGTATCAGGCGCAGGAGCGTCCCGTCCGGGTCGACCAAAGCGGCCATCGTGCCTCCCCACTCCTGCAGTTCGGGG
GCATGGATGCGCGGATAGCCGCTGCTGGTTTCCTGGATGCCGACGGATTTGCACTGCCGGTAGAACTCCGCGAGGTCGTC
CAGCCTCAGGCAGCAGCTGAACCAACTCGCGAGGGGATCGAGCCCGGGGTGGGCGAAGAACTCCAGCATGAGATCCCCGC
GCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAGAAGGCGGCGGTGGGATCGAAA
TCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCGGTCATACCCATGCCGGTACCGGATCCCTTGTACAGCTCGTCCATGCC
GAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGG
ACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCG
AGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGAT
ATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCA
GCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAG
ATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCT
GAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACT
TCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCC
AGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGCTAGCTTTGATTGATTTGACTGT
GTTATTTTGCGTGAGGTTATGAGTAGAAAATAATAATTGAGAAAGGAATATGACAAGAAATATGAAAATAAAGGGAACAA
ACCCAAATCTGATTGCAAGGAGAGTGAAAGAGCCTTGTTTATATATTTTTTTTCCTATGTTCAACGAGGACAGCTAGGT
TTATGCAAAAATGTGCCATCACCATAAGCTGATTCAAATGAGCTAAAAAAAAAATAGTTAGAAAATAAGGTGGTGTTGAA
CGATAGCAAGTAGATCAAGACACCGTCTAACAGAAAAAGGGGCAGCGGACAATATTATGCAATTATGAAGAAAAGTACTC
AAAGGGTCGGAAAAATATTCAAACGATATTTGCATAAAATCCTCAATTGATTGATTATTCCATAGTAAAATACCGTAACA
ACACAAAATTGTTCTCAAATTCATAAATTATTCATTTTTTCCACGAGCCTCATCACACGAAAAGTCAGAAGAGCATACAT
AATCTTTTAAATGCATAGGTTATGCATTTTGCAAATGCCACCAGGCAACAAAAATATGCGTTTAGCGGGCGGAATCGGGA
AGGAAGCCGGAACCACCAAAAACTGGAAGCTACGTTTTTAAGGAAGGTATGGGTGCAGTGTGCTTATCTCAACCTTTGAT
TAGTTTATGGGCTAGCTTATGATACTTCCTCTTGCTGTTCCAGTTTACAAATGATTTAAACGAACTTCTACACTTTTACG
GTAGTATTTCTTGTTGGTGATGAGCTCAGATTCAATATGACTGCTTAATCAGTGAAAATTAATGGAAACAGCCAGACACA
CAGCCTTTGAAAAATGTTAAAAACGATCACATATATGTAGTCACGTGGTTAGTGCATAGAGTCAGGCGATGGTGGTTAGT
TCCCTAGCCAGAGCCATTCCCCTCTTGTTTCGAAATGCGATGACTGAATGGGGAAGAAGCTCTTAGTATTTCTGTAGTTT
TTTTTGCCGATTTCTATTGTTGCCTGCCCCGGACTTTCTTTTACCTTTTGTAGCTACAACTATTAATTATAATGCATAAT
AGATCTAGAAGCGTATTGTGATGAGCATCCTTGGAAATAAGAAAGGAAGTCACAGGTACAGTAAAAAAAGAGCCGTGACA
GGCACCTAATTTCAGTTTTGATTGAGGTATAATGAGCTCAGAAGACCAAGATGCTCTAAAAGTAAAGCAGAATGGTCTTT
CCAATCAGAATTCGTCACTTGTTGGTGCTTTAACTGCTGGTGGAAGATCTATGGTTTATCAGATGACTTCATTCTATATG
AGAACTCCACTAAAACTGTTTCGACCGGCCAGATTTGACTATACGCATTATATTCGAGTACTGTTGACTGGGAGTGATGA

TACAAATAAGGATTCTAAGGTACGACGTTCAAAATATAAGTTTTGGAGTCCCAAATATACATACTACTTCGAAAATTCCT
CCATTGGAATAGTTACTAAGGCACTAAATAAGTATGGGTGGAAAGTTATTCCTGATAGAATATTGCCTCCTTTGGTGGCA
AATTCCCTTGCTGGGGTAGTTCTTTATACCACATACTTGACTACCCTGAATAATTTCCCAGTCAGATCAGTGGAGCCAGC
TTTTCAACACAATTTGTTTGACTACTTAAGGTAAGGGCGAATTCGCGGCCGCTAAATTCAATTCGCCCTATAGTGAGTCG
TATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGC
ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTATACG
TACGGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGAC
ACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCTCCCGTGAACTTTACCCGGT
GGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAG
TGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAAATGTCAGGCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAA
TGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGC
GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAG
CCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGG
ATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCT
ATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTT
GTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGT
TCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATC
TCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCG
GCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCA
GGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCG
AGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGAC
TGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGA
ATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACG
AGTTCTTCTGAATTATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGC
ATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATAT
GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGC
GAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA
ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAA
CGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGAT
CAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAG
TTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGG

ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAG
CGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCA
TACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGACCAAAATCCCTTAACGTGAGTTTTCG
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTT
GCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTG
GCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCA
CCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGA
CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAA
CGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG
TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCC
TGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCA
GCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCT
GTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG
CGAGGAAGCGGAAG

Figure 4 Plasmid TOPO/F7 sequences.

Yellow: bleomycin gene, Cyan: GFP coding sequence, Green: ScP$_{PDC1}$

Results and Discussions

Cold-tolerant yeast *K. marxianus* strain

To characterize growth rates at various temperatures, *K. marxianus* NBRC1777 and *S. cerevisiae* BY4741 were cultured at 5, 10, 20, 30, 37 and 45°C (Figure 5). The maximum specific growth rate (μ_{max}) was higher for *K. marxianus* than for *S. cerevisiae* at all temperatures (Figure 5), indicating that the *K. marxianus* is a fast-growing strain. *K. marxianus* showed best growth at 37°C and was able to grow at 45°C. The difference in the μ_{max} was also bigger at 10°C (0.079/0.044) compared to the difference at 20°C (0.23/0.20). Thus, in addition to the known thermotolerance of *K. marxianus* (Hong et al. 2007), strain NBRC1777 exhibited fast and robust growth at a wide range of temperatures, a feature that will be appealing for a variety of industrial applications.

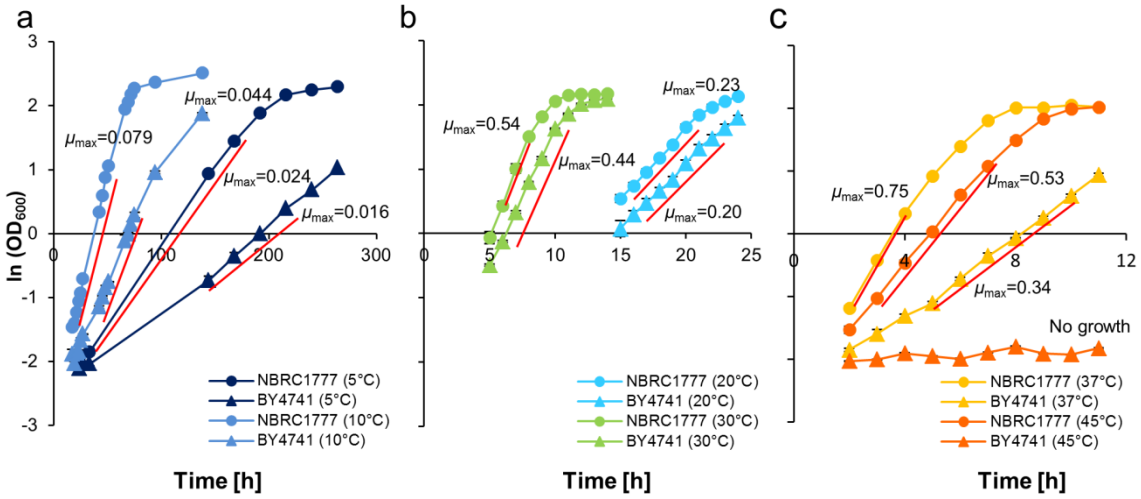


Figure 5 Cell growth of *K. marxianus* and *S. cerevisiae* at various temperatures.
Cell growth was compared between *K. marxianus* NBRC1777 and *S. cerevisiae* BY4741 at 5, 10°C (a), at 20, 30°C (b), and at 37, 45°C (c). The maximum specific growth rate (μ_{max}) was determined as the slope of the mid-log phase of natural logarithm (ln) (OD_{600}). Timeframes for each μ_{max} were as follows: at 5°C from 33 h to 168 h and 144 h to 216 h, at 10°C from 19 h to 48 h and 42 h to 75 h, at 20°C from 16 h to 20 h and 16 h to 22 h, at 30°C from 6 h to 8 h and from 6 h to 10 h, at 37°C from 2 h to 4 h and 5 h to 10 h, 45°C from 3 h to 7 h and no growth for *K. marxianus* and *S. cerevisiae,* respectively.

Isolation of KmSNR52 promoter for expression of sgRNA for CRISPR-Cas9

To introduce the CRISPR system into *K. marxianus*, an efficient promoter for sgRNA expression was needed. As 5'-capped transcripts are not suitable for use as sgRNAs, PolII-driven promoters are not favoured. The *SNR52* promoter has been used to express functional sgRNAs in *S. cerevisiae* (Dicarlo et al. 2013), so we obtained the *K. marxianus SNR52* promoter (KmP_{SNR52}) as follows. The *K. marxianus SNR52* transcript sequence was first identified using the BLASTN search program with the *S. cerevisiae SNR52* transcript as a query. The region upstream of the putative Km*SNR52*

transcript was then aligned with the Sc*SNR52* promoter; a 531-bp fragment was thereby defined as the putative KmP$_{SNR52}$ (Figure 3).

We next constructed a vector for use as a shuttle vector by combining a *K. marxianus* autonomously replicating sequence (KmARS7) (Hoshida et al. 2014) with a centromere sequence (KmCEN-D) (Yarimizu et al. 2015) and demonstrated (data not shown) that this plasmid is retained as an episome in *K. marxianus* NBRC1777. A human-optimized *Streptococcus pyogenes* (Sp) *cas9* (Dicarlo et al. 2013) under the control of the constitutive *S. cerevisiae PDC1* promoter (ScP$_{PDC1}$) was inserted into the shuttle vector, yielding a basal plasmid (Cas9_Base; Figure 1 and 6) capable of directing the expression of Cas9 in *K. marxianus*.

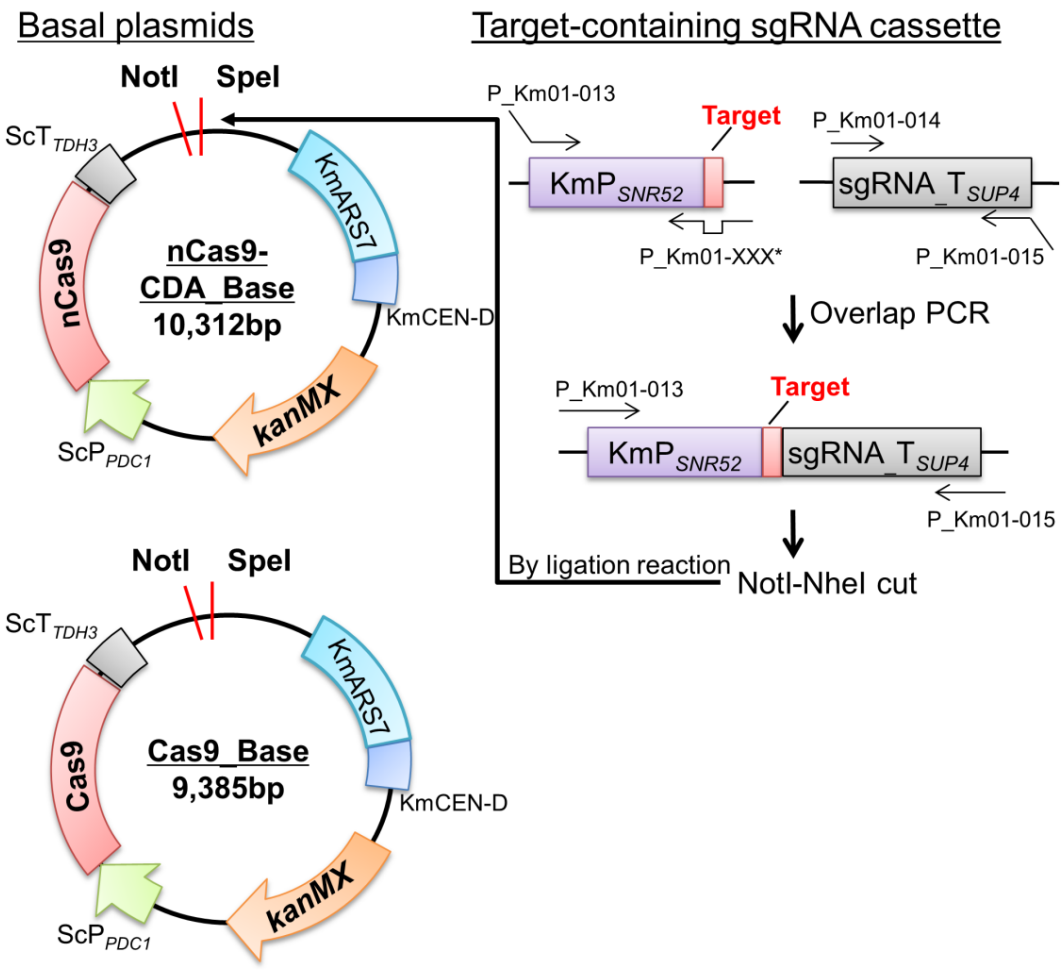


Figure 6 Construction of genome-editing plasmids.

Constructed basal plasmids nCas9-CDA_Base and Cas9_Base are shown on the left. Target-containing sgRNA cassettes were constructed by overlap PCR using target sequence-containing primers (P_Km01-XXX, where XXX was a 3-digit number as shown in the figure) and then inserted into the basal plasmids at NotI and SpeI/NheI restriction enzyme sites.

Inactivation of the NHEJ pathway by Target-AID point mutation of NHEJ proteins

In *S. cerevisiae*, the *NEJ1* and *DNL4* genes encode core components of the proteins mediating NHEJ. By analogy to *S. cerevisiae*, disruption of the *K. marxianus Nej1* and *Dnl4* homologs are expected to yield more genetically tractable *K. marxianus* hosts with higher proportions of HR. Mutation of the *K. marxianus Nej1* and *Dnl4* genes identified by sequence similarity to the respective budding yeast genes was performed using targeted point mutagenesis by Target-AID in *K. marxianus*. A basal plasmid (nCas9-CDA_Base) was generated by constructing a variant of the Cas9_Base vector that encoded a nickase mutant (D10A) SpCas9 protein fused to the *Petromyzon marinus* (Pm)CDA1 cytidine deaminase (Nishida et al. 2016) (Figure 2 and 6). Following a previous study in *S. cerevisiae* (Valencia et al. 2001), target sites in the *Nej1* and *Dnl4* genes of *K. marxianus* NBRC1777 were selected so as to introduce stop codons by causing C-to-T mutagenesis at 16 to 19 bp upstream of the sequences encoding protospacer adjacent motif (PAM) sequences for the target sequences. *K. marxianus* NBRC1777 cells were transformed with plasmids containing each targeting sequence (nCas9-CDA_target *Nej1* and nCas9-CDA_target *Dnl4*). G418-resistant colonies were selected; the *Nej1* and *Dnl4* genes of the transformants then were PCR amplified and subjected to sequence analysis to confirm the presence of the targeted lesions (Figure 7). Both plasmids introduced stop codons as expected, providing null alleles of *Nej1* (*Nej1*º) and *Dnl4* (*Dnl4*º) with 1 of 8 (or 1 of 4) G418-resistant colonies contained the mutant allele, respectively. Thus, Target-AID was efficient enough to introduce targeted point mutation in *K. marxianus*.

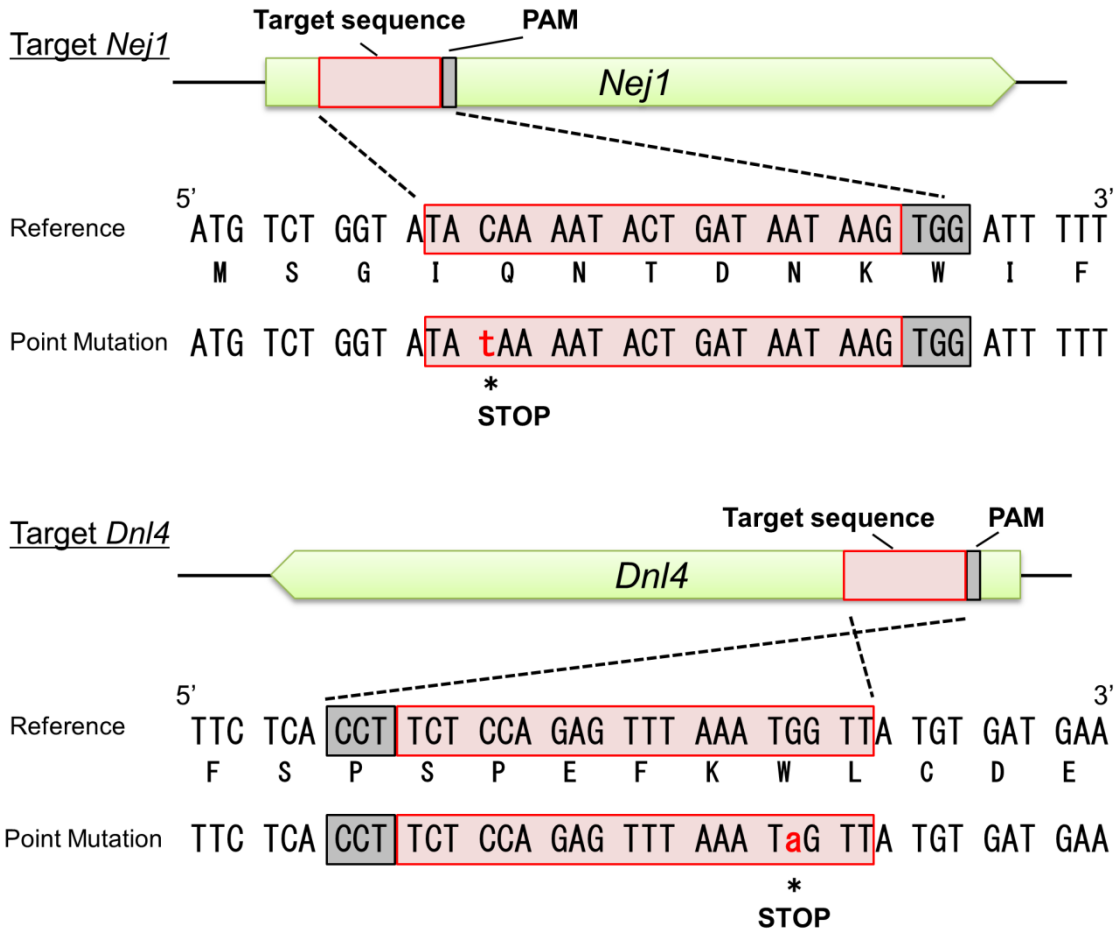


Figure 7 Induction of mutations by Target-AID.

The *Nej1* and *Dnl4* target sites were selected for the introduction of stop codons by Target-AID. Red boxes indicate the target sites and grey boxes indicate the PAM sequences. Lower-case nucleotide in red font indicates the mutated base. The sequences of the induced mutations are shown in alignment with the reference sequences, with each predicted amino acid sequence below the respective nucleic acid sequence.

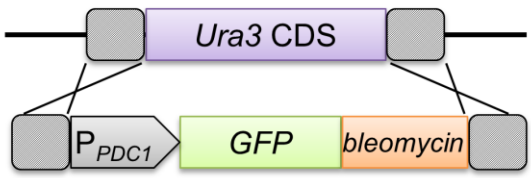
Conventional HR in NHEJ null mutants

To investigate if these disruptions in genes with putative NHEJ activities increased the rate of HR in *K. marxianus*, cells of the *Nej1*^{o} or *Dnl4*^{o} mutants were transformed with DNA donor fragments with homology arms. Specifically, we employed a *GFP bleo*

cassette (i.e., encoding green fluorescent protein (GFP) fused to a protein providing a selectable marker for bleomycin/zeocin resistance) flanked by *K. marxianus Ura3* sequences. Two separate constructs were used, one providing extended homology (1000-bp arms at both ends) and the second providing shorter homology (arms of 122 and 109 bp at the two ends). The fragments were expected to replace *Ura3* with the *GFP-bleo* cassette if HR was facilitated by mutation of genes of the NHEJ pathway. The cells were transformed and selected for zeocin-resistance; the resulting transformants were then screened for 5-fluoroorotic acid (5-FOA) resistance, which would be conferred by *Ura3* disruption. When using the construct providing homology arms of 1000 bp, 100% of the Zeo^R transformants of the *Nej1*º or *Dnl4*º strains (17/17 and 13/13, respectively) were $5\text{-}FOA^R$; in comparison, only 26% (6/23) of the Zeo^R transformants of the wild-type (WT) strain were $5\text{-}FOA^R$ (Figure 8). When using equimolar quantities of the construct providing shorter homology arms (122 bp/109 bp), 123 Zeo^R transformants were obtained from the WT strain but no 5- FOA^R transformants were obtained. The *Nej1*º or *Dnl4*º strains yielded one and two Zeo^R transformants (respectively), of which no and one (respectively) were $5\text{-}FOA^R$ (i.e., disrupted at the targeted *Ura3* locus) (Figure 8).

Apparently, less number of transformants were obtained from the mutants of genes involved in NHEJ pathway, presumably because random integration was severely suppressed, rather than by increasing the HR frequency. This decrease in NHEJ facilitates reliable HR at the targeted (by homology) locus, although the obtained efficiency may not be sufficient to perform HR with short homology arms (i.e., of the length that can be added by PCR primers).

a



b

Length of homology sequences: 1000 bp

Strains	Fragment preparation	Donor DNA (μg)	Transformants	Inserted	Rate(%)
WT	PCR	1.57	2	0	**26**
	Cut	3.89	21	6	
Nej1º	PCR	1.57	3	3	**100**
	Cut	3.89	14	14	
Dnl4º	PCR	1.57	3	3	**100**
	Cut	3.89	10	10	

c

Length of homology sequences: 122/109 bp

Strains	Fragment preparation	Donor DNA (μg)	Transformants	Inserted	Rate(%)
WT	PCR	7.83	123	0	**0**
Nej1º			1	0	**0**
Dnl4º			2	1	**50**

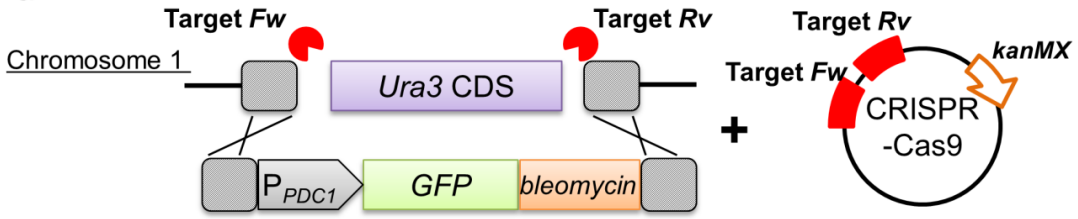
Figure 8 Effect of *Nej1º* and *Dnl4º* on homologous recombination.

(a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus. Dashed boxes indicate homologous regions. (b), (c) Homologous recombination rates (5-FOA resistant (inserted)/zeocin resistant (transformants)). Fragments were prepared either by PCR or by restriction digestion of plasmid (Cut).

HR with CRISPR-Cas9

To boost the HR frequency, co-transformation with a construct encoding CRISPR-Cas9 was performed; the CRISPR system was expected to facilitate HR by providing cleavage of the targeted locus. A pair of target sites was selected at just inside of the homologous regions of *Ura3*. WT, *Nej1*^{o}, and *Dnl4*^{o} cells were transformed with the fragment with homology arms of 122 bp/109 bp (in combination with the Cas9 construct) and transformants were selected using either zeocin, (*Ura3*-targeting fragment), G418 (CRISPR-Cas9 vector), or both. For all three strains, a larger number of antibiotic-resistant transformants were obtained when the transformation mixes included the CRISPR-Cas9 construct. Among the antibiotic-resistant transformants, 41-79% of those derived from WT cells were 5-FOA^{R} (i.e., disrupted at *Ura3*), while 95-100% of those derived from the *Nej1*^{o} and *Dnl4*^{o} strains were 5-FOA^{R} (Figure 9). These results indicated that target cleavage by Cas9 enhanced the frequency of HR sufficiently that even without selection by zeocin (i.e., without direct selection for the integration at *Ura3*), disruption of Ura3 was seen in virtually all transformants.

a



b Length of homology sequences: 122/109 bp

Strains	Donor DNA (μg)	Plasmid (μg)	Selection	Transformants	Inserted	Rate(%)
WT	1.53	3.33	Zeocin and G418	21	16	**76**
Nej1°				16	16	**100**
Dnl4°				24	24	**100**
WT	1.53	3.33	Zeocin	61	25	**41**
Nej1°				11	11	**100**
Dnl4°				19	19	**100**
WT	3.12	5.37	G418	43	34	**79**
Nej1°				19	18	**95**
Dnl4°				20	20	**100**

Figure 9 Homologous recombination facilitated by CRISPR-Cas9.

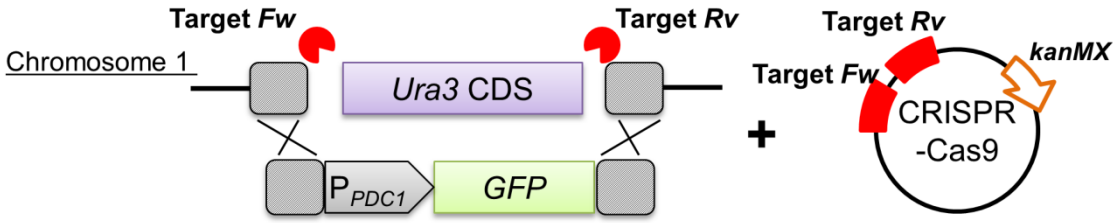
(a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus and CRISPR-Cas9 plasmid targeting two sites (indicated by red pie bolts) flanking the *Ura3* coding sequences (CDS). Dashed boxes indicate homologous regions. Bottom panel indicates (b) Homologous recombination rate (5-FOA resistant (inserted)/zeocin and/or G418 resistant (transformants) * 100).

Markerless integration with 50-bp homology arms and CRISPR-Cas9

Markerless integration was attempted by co-transformation using the combination of a donor DNA template containing a *GFP* expression cassette flanked by 50-bp homology arms along with a CRISPR-Cas9-encoding construct. Transformants were selected using G418 and then screened by both the 5-FOA test and PCR analysis. We observed successful integration (i.e., disruption of the *Ura3* target locus in $G418^R$

transformants) at rates of 28%, 84%, and 92% in the WT, *Nej1*º, and *Dnl4*º backgrounds, respectively (Figure 10). Taking advantage of the markerless integration, the GFP-encoding sequence was directly inserted into the coding sequence of the endogenous *Sed1* gene so as to encode a GFP-Sed1 fusion protein preceded by the Sed1 localization peptide. A single cleavage site for CRISPR-Cas9 within this target gene was designed in such a way that the site would not be retained after successful integration. Short (50-bp) homology arms flanking the target site were added to either end of the GFP-encoding sequence by PCR. Transformants were selected by G418 and then screened by PCR, revealing gene replacement at the *Sed1* target locus at 38%, 100%, and 100% success rates in the WT, *Nej1*º, and *Dnl4*º backgrounds, respectively (Figure 11). When viewed by fluorescence microscopy, cells harboring the targeted integration exhibited peripheral localization of GFP-Sed1, consistent with the localization previously reported for the equivalent construct in *S. cerevisiae* (Huh et al. 2003). Thus, this method can be applied for scar-less genome editing in *K. marxianus*.

a



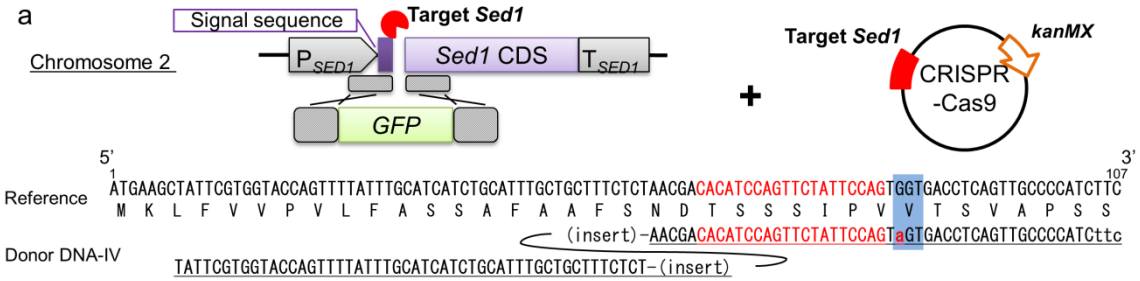
b

Length of homology sequences: 50 bp

Strains	Donor DNA (μg)	Plasmid (μg)	Transformants (G418 resistant)	5-FOA resistant	PCR check	Rate (%)
WT	1.76	5.33	60	39	3/7	**28**
Nej1º			50	49	6/7	**84**
Dnl4º			50	46	7/7	**92**

Figure 10 Markerless integration with short homologous arms.

(a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus and the CRISPR-Cas9 plasmid targeting two sites (indicated by red pie bolts) flanking the *Ura3* CDS. Dashed boxes indicate homologous regions. (b) Homologous recombination rate ((5-FOA resistant*PCR check)/ transformants (G418 resistant)*100).



b

Length of homology sequences: 50 bp

Strains	Donor DNA (μg)	Plasmid (μg)	Transformants (G418 resistant)	PCR check	Rate (%)
WT	1.98	3.92	100	3/8	**38**
Nej1°			95	8/8	**100**
Dnl4°			89	8/8	**100**

c

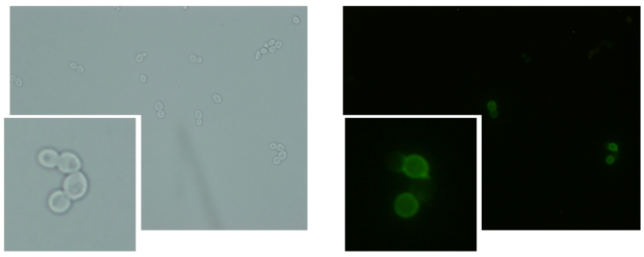


Figure 11 Direct fusion of GFP-coding sequence into a chromosomal locus.

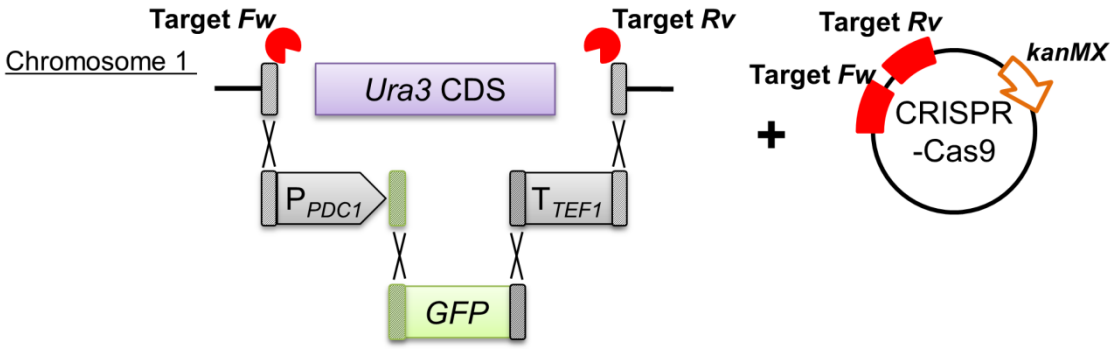
(a) Schematic showing integration of the GFP-coding fragment into the *Sed1* CDS. The CRISPR-Cas9 plasmid targeted (indicated by red pie bolts) the end of the signal peptide-coding region of *Sed1*. Dashed boxes indicate homologous regions. Left arm is homologous to the end of to the signal peptide-coding sequence (bp 8 to 57 with respect to the start of the *Sed1* coding sequence). Right arm is homologous to *Sed1* sequence (bp 58 to 107 with respect to the start of the *Sed1* coding sequence) with substitution at

bp 84 (G to A; indicated by lower-case letter in red font) to eliminate a PAM. (b) The successful integration rate was calculated as PCR check*100. (c) Representative microscopic image of positive transformants (*Nej1*°/GFP_Sed1) under bright-field (left) and fluorescent (right) illumination. Insets show 3-fold magnified images.

Assembly and integration of multiple fragments without markers

In vivo assembly and integration of multiple fragments was performed using three DNA fragments, each containing 50 bp overlap with the adjacent fragments or target sites, such that the entire construct was designed to replace *Ura3* (Figure 12). Compared to single PCR product donor constructs, a smaller number of transformants was obtained by G418 selection when using the multiple-fragment DNA as the donor. Nonetheless, the WT, *Nej1*°, and *Dnl4*° strains showed integration efficiencies of 18% (2/11), 100% (3/3), and 100% (4/4), respectively. This assembly and integration method is expected to allow combinatorial integration of gene cassettes using (for example) combinations of various promoters and coding regions.

a



b

Length of homology sequences: 50 bp

Strains	Donor DNA (μg)	Plasmid (μg)	Transformants (G418 resistant)	5-FOA resistant	PCR check	Rate (%)
WT	2.5	5	11	3	2/3	**18**
Nej1°			3	3	3/3	**100**
Dnl4°			4	4	4/4	**100**

Figure 12 In vivo assembly and integration of three DNA fragments with 50 bp overlaps.

(a) Schematic of the three transforming donor DNA fragments replacing the *Ura3* locus along with the CRISPR-Cas9 plasmid targeting two sites (indicated by red pie bolts) flanking the *Ura3* CDS. Dashed boxes indicate homologous regions. Bottom panel shows (b) Homologous recombination rate (5-FOA resistant * PCR check)/ transformants (G418 resistant)*100.

In summary, targeted point mutagenesis (by Target-AID) and markerless CRISPR-mediated integration were demonstrated in *K. marxianus* using a fast-growing strain that was shown to exhibit both thermo- and cold-tolerance. Mutation of *Nej1* or *Dnl4*, loci encoding homologs of NHEJ proteins, enhanced the proportion of HR events. These genome-engineering tools are expected to greatly facilitate industrial use of *K.* marxianus, a prototype for the class of underexploited non-conventional yeasts.

References

Abdel-Banat B. M. A., Nonklang S, Hoshida H, Akada R., 2009. Random and targeted gene integrations through the control of non-homologous end joining in the yeast *Kluyveromyces marxianus*. Yeast. 27, 19–39.

Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343.

Dudášová Z, Dudáš A, Chovanec M., 2004. Non-homologous end-joining factors of *Saccharomyces cerevisiae*. FEMS Microbiol. Rev. 28, 581–601.

Feldmann H, Driller L, Meier B, Mages G, Kellermannu J, Winnacker EL., 1996. HDF2, the second subunit of the Ku homologue from *Saccharomyces cerevisiae*. J. Biol. Chem. 271, 27765–9.

Gietz D, St Jean A, Woods R a, Schiestl RH., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20, 1425.

Hong J, Wang Y, Kumagai H, Tamaki H., 2007. Construction of thermotolerant yeast expressing thermostable cellulase genes. J. Biotechnol. 130, 114–123.

Hoshida H, Murakami N, Suzuki A, Tamura R, Asakawa J., 2014. Non-homologous end joining-mediated functional marker selection for DNA cloning in the yeast *Kluyveromyces marxianus*. 31, 29–46.

Huh W-K, Falvo J V, Gerke LC, Carroll AS, Howson RW, Weissman JS, et al., 2003. Global analysis of protein localization in budding yeast. Nature. 425, 686–691.

Inokuma K, Ishii J, Hara KY, Mochizuki M, Hasunuma T, Kondo A., 2015. Complete Genome Sequence of *Kluyveromyces marxianus* NBRC1777, a Nonconventional Thermotolerant Yeast. Genome Announc. 3, e00389-15.

Jakočiūnas T, Jensen MK, Keasling JD., 2016 CRISPR/Cas9 advances engineering of microbial cell factories. Metab. Eng. 34, 44–59.

Jeong H, Lee DH, Kim SH, Kim HJ, Lee K, Song JY, et al., 2012. Genome sequence of the thermotolerant yeast *Kluyveromyces marxianus* var. *marxianus* KCTC 17555. Eukaryot. Cell. 11, 1584–1585.

Moore JK, Haber JE., 1996. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 16, 2164–2173.

Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, et al., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science. 353, aaf8729-1-aaf8729-8.

Nonklang S, Abdel-Banat BM a, Cha-aim K, Moonjai N, Hoshida H, Limtong S, et al., 2008. High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. Appl. Environ. Microbiol. 74, 7514–7521.

Da Silva N a, Srikrishnan S., 2012. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae.* FEMS Yeast Res. 12, 197–214.

Suzuki T, Hoshino T, Matsushika A., 2014. Draft Genome Sequence of *Kluyveromyces marxianus* Strain DMB1, Isolated from Sugarcane Bagasse Hydrolysate. Genome Announc. 24, e00733-14.

Valencia M, Bentele M, Vaze MB, Herrmann G, Kraus E, Lee SE, et al., 2001. NEJ1 controls non-homologous end joining in *Saccharomyces cerevisiae*. Nature. 414, 666–669.

Wagner JM, Alper HS., 2016. Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. Fungal Genet. Biol. Elsevier Inc. 89, 126–136.

Yarimizu T, Nakamura M, Hoshida H, Akada R., 2015. Screening of accurate clones for gene synthesis in yeast. J. Biosci. Bioeng. 119, 251–259.

Chapter 3
Genetic and physiological basis for antibody production by *Kluyveromyces marxianus*

Introduction

Production of biopharmaceuticals requires the difficult choice of a host cell capable of generating the desired product in an active and safe form, devoid of unwanted modification or contamination. Additionally, some biopharmaceuticals such as antibodies have proven difficult to express at high levels. Chinese hamster ovary (CHO) cells and yeasts are the major hosts that have been engineered to produce biopharmaceutical products, including antibodies (Maccani et al. 2014). As mammalian cells, CHO cells produce mammalian-derived proteins in an active form, bearing appropriate modifications such as glycosylation. However, the development of stable cell lines takes very long times (6 to 12 months), and the cost of cell culture is very high (Lai et al. 2013). Yeast cells such as *Pichia pastoris* may provide much faster and cheaper ways of production (Çelik and Çalık 2012); while this yeast can be engineered to serve as a suitable hosts, highly complex proteins such as antibodies can be difficult to express efficiently in this system (Nielsen 2013).

Due to difficulties in expression, secretion, and post-translational modification, antibodies intended for clinical use remain a challenge to produce in a cost-effective manner (Buckholz and Gleeson 1991; Huang et al. 2014). Single-chain Fv antibody (scFv) is one of the most useful forms of antibody, consisting of a single polypeptide in which the variable regions of the heavy (V_H) and light (V_L) chain domains are fused by a short, flexible linker; the resulting product has a molecular weight of approximately 30 kDa (Damasceno et al. 2004). Unlike large immunoglobulins (IgGs), scFv proteins have demonstrated rapid tumor penetration (Yokota et al. 1992). A prototypical scFv is the anti-chicken (anti-hen) egg white lysozyme antibody (HyHEL-10), which has been used for the precise analysis of antigen-antibody interactions (Tsumoto et al. 1997).

The non-conventional yeast *Kluyveromyces marxianus* can grow on various sugars (glucose, xylose, fructose, sucrose, inulin, etc.) (Fonseca et al. 2008; Lane and Morrissey 2010; Lertwattanasakul et al. 2011). *K. marxianus* is known to secrete proteins such as inulinase into the culture medium at high levels (Rouwenhorst et al.

1990; Hu et al. 2012). Engineering of *K. marxianus* for protein production has been reported for both endogenous and heterologous enzymes (Raimondi et al. 2013; Hong et al. 2007). However, there are to date (to our knowledge) no reports on secretory antibody production in *K. marxianus*.

The *K. marxianus* NBRC1777 strain recently has been shown to exhibit rapid growth and adaptability to a wide range of temperatures (from 5°C to 45°C). Additionally, comprehensive genome engineering tools recently have been introduced for use in this strain, including a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) -associated protein (CRISPR-Cas9) system and deaminase-mediated base editing Target-AID (Nambu-Nishida et al. 2017). NBRC1777 is expected to be of use for various bio-production applications, including the secretion of high-value proteins.

In the present study, we introduced *K. marxianus* NBRC1777 as a novel host for scFv production. Several parameters were examined, including the type of secretion signal and growth conditions such as temperature, carbon source, and medium. Genetic backgrounds that affect protein production or secretion also were studied.

Materials and Methods

Strains and culture conditions

The *K. marxianus* and *S. cerevisiae* strains used in this study are listed in Table 1. *Escherichia coli* strain DH5α (Toyobo, Osaka, Japan) and yeast cells were grown as described previously (Nambu-Nishida et al. 2017).

INU1 gene disruption

The *inu1* gene-disrupted strain and homologous recombination strains were generated using the CRISPR-Cas9 system. The CRISPR-Cas9 vector plasmid (Cas9_Base) of *K. marxianus*, target sgRNA cassette construction, and methods were as described previously (Nambu-Nishida et al. 2017). The *inu1* deletion strain was generated by using a Cas9 plasmid (E02-026) containing gRNA-1 and gRNA-2 target sequences (Table 2 and Figure 1a).

Table 1 Plasmids and strains used in this study

Plasmids and Strains	Genotype	Reference
Plasmids		
E02-012	KmP_{MDH1}_KmINU1ss_scFv_T_{TDH3}, KmARS7, KmCEN D, kanMX, ori, and AmpR	This study
E02-014	KmP_{ACO1}_KmINU1ss_scFv_T_{TDH3}, KmARS7, KmCEN D, kanMX, ori, and AmpR	This study
E02-020	KmP_{MDH1}_Scα-MFss_scFv_T_{TDH3}, KmARS7, KmCEN D, kanMX, ori, and AmpR	This study
E02-022	KmP_{ACO1}_Scα-MFss_scFv_T_{TDH3}, KmARS7, KmCEN D, kanMX, ori, and AmpR	This study
Cas9_Base	ScP_{PDC1}_Cas9_T_{TDH3}, KmARS7, KmCEN D, *kanMX*, ori, and AmpR	(Nambu-Nishida et al. 2017)
E02-025	KmP_{SNR52}_target_gRNA-1_sgRNA_T_{sup4} cassette and KmP_{SNR52}_target_gRNA-3_sgRNA_T_{sup4} cassette in Cas9_Base	This study
E02-026	KmP_{SNR52}_target_gRNA-1_sgRNA_T_{sup4} cassette and KmP_{SNR52}_target_gRNA-2_sgRNA_T_{sup4} cassette in Cas9_Base	This study
Strains		
1 (NBRC1777)	Wild-type (WT)	NITE Biological Resource Center, Japan
Nej1°	*Nej1* disrupted by C to T point mutation at posision 13	(Nambu-Nishida et al. 2017)
Dnl4°	*Dnl4* disrupted by G to A point mutation at posision 44	(Nambu-Nishida et al. 2017)
2 (Km02-026)	WT/E02-012	This study
3 (Km02-032)	WT/E02-014	This study
4 (Km02-050)	WT/E02-020	This study
5 (Km02-056)	WT/E02-022	This study
Δ*inu1* (Km02-063)	*Nej1°*/Δ*inu1*	This study

6 (Km02-064)	*Nej1°*/*inu1*::P_{INU1}_scFv	This study
7 (Km02-065)	*Dnl4°*/*inu1*::P_{INU1}_scFv	This study
8 (Km02-066)	*Nej1°*/*inu1*::/E02-020	This study
BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	ATCC (Brachmann et al. 1998))

Table 2 Target sequence used in this study

Name	Sequence (5' → 3')
gRNA-1	TATAAAATGTCGCTGTGACC
gRNA-2	CAACTACAACCGGATACCTG
gRNA-3	ATGGAAGCAAGAGGGAGTAT

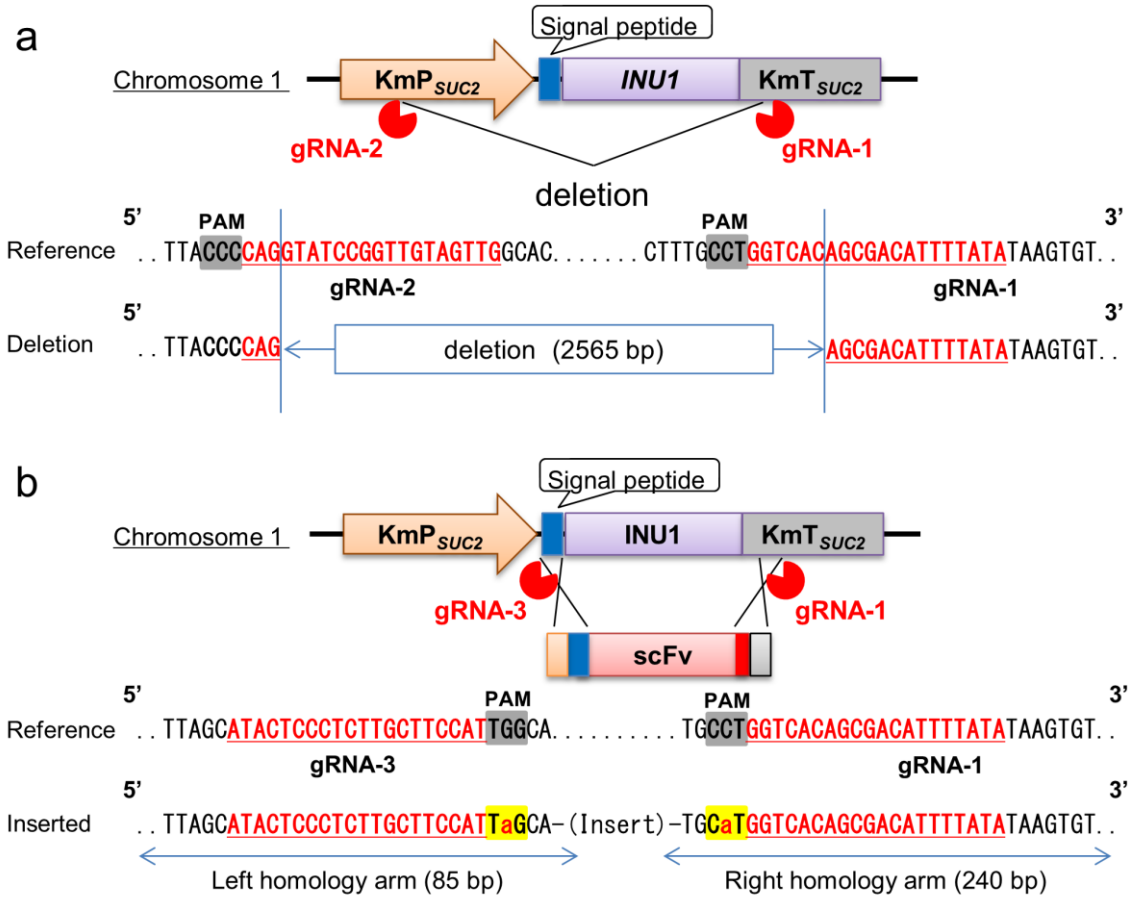


Figure 1 Design of *INU1* disruption and replacement

(a) A schematic of the *INU1* locus and its deletion is shown. Reverse complement sequences of target (red) and PAM (shadowed) are shown on the reference sequence. The lower part of the panel shows an alignment of the sequence from the deletion strain with the sequence of the wild-type locus.

(b) A schematic of the *INU1* replacement by a fragment encoding scFv is shown. The mutated PAM sequence and its reverse complement sequence are highlighted in yellow. Mutated bases are indicated as small red letters. The scFv-encoding fragment also encodes the INU1 secretion signal peptide (blue box) and a 6-histidine tag (red box) in-frame with the expressed ORF.

Replacement of *INU1* by integration of a scFv-encoding sequence

A fragment encoding HyHEL-10 scFv (scFv hereafter) (Acchione et al. 2012) was codon optimized for expression in *K. marxianus* (Figure 2). The homologous recombination strain was generated by transforming the recipient strain by the lithium acetate method (Gietz et al. 1992), using 10 μg of Cas9 plasmid (E02-025) containing gRNA-1 and gRNA-3 target sequences (Table 2 and Figure 1b) and 5 μg of the sequence-optimized fragment encoding scFv (Figure 3). The transformed cells were plated on YPD containing the appropriate selection reagent (100 μg/ml G418).

Verification of genome-edited cells

Transformants generated using the CRISPR-Cas9 system were screened by colony PCR using the primer pair P_Km01-010 + P_Km01-010-011. DNA sequence of the resulting amplicon was confirmed by sequencing using a 3130xL Genetic Analyzer (Applied Biosystems, CA, USA). The transformant cells were grown without selection reagents to isolate a clone that dropped the Cas9 plasmid.

Construction of scFv expression plasmids

The constructed plasmids are listed in Table 1. Plasmid E02-014, which includes KmARS7, KmCEN-D, the scFv-encoding fragment, KmP_{ACO1}, and the *kanMX* selectable marker (which provides G418 resistance) is shown as an example (Figure 4).

Constructs incorporated either the P_{MDH1} (KmP_{MDH1}) (Figure 5) or P_{ACO1} (KmP_{ACO1}) (Figure 6) promoters from *K. marxianus*. Constructs also incorporated sequences encoding either the secretory signal sequence from inulinase (KmINUss) from *K. marxianus* (Bergkamp et al. 1993) or that from α-MF (Scα-MFss) from *S. cerevisiae* (Melorose et al. 1986). Fragments carrying the desired promoter fragment and encoding the desired signal sequence were inserted into the NheI or SbfI/BamHI sites of the E02-014 plasmid using In-fusion cloning (Takara Bio, Shiga, Japan). The resulting scFv expression plasmids were transformed into *K. marxianus* NBRC1777 or the *inu1* deletion strain by the transformation and selection methods noted above.

DIQLQESGPSLVKPSQTLSLTCSVTGDSITSDYWSWIRKFPGNRLEYMGYVSYSGSTYYNPSLKSRISITRDTSKNQYYL
DLNSVTTEDTATYYCANWDGDYWGQGTLVTVSAGGGGSGGGGSGGGGSDIVLTQSPATLSVTPGNSVSLSCRASQSIGNN
LHWYQQKSHESPRLLIKYASQSISGIPSRFSGSGSGTDFTLSINSVETEDFGMYFCQQSNSWPYTFGGGTKLEYKHHHHH
H*

Figure 2 Amino acid sequence of scFv

Yellow: V_H, Light green: V_L, Red: 6-His-tag

TTTTTTTTTGTTAGATATGAAGTTAGCATACTCCCTCTTGCTTCCATTaGCAGGAGTCAGTGCTTCAGTGATCAATTACA
AGAGAGACATCCAATTGCAAGAATCTGGTCCATCCTTGGTTAAGCCATCCCAAACCTTGTCTTTGACTTGTTCCGTCACC
GGTGACTCTATCACTTCCGATTACTGGTCTTGGATTAGAAAGTTCCCTGGTAACAGATTGGAATACATGGGTTACGTTTC
TTACTCCGGTTCTACCTACTACAACCCATCTTTGAAGTCCAGAATCTCTATTACCAGAGACACTTCCAAGAACCAATACT
ACTTGGACTTGAACTCTGTCACCACTGAAGATACCGCTACTTACTACTGTGCCAACTGGGACGGTGACTACTGGGGTCAA
GGTACCTTGGTTACTGTCTCCGCTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGTGGTGGTGGTAGTGACATCGTTTT
GACTCAATCCCCAGCTACTTTGTCTGTTACCCCTGGTAACTCCGTCTCTTTGTCCTGTAGAGCCTCTCAATCCATCGGTA
ACAACTTGCACTGGTACCAACAAAAGTCCCATGAATCTCCAAGATTGTTGATTAAGTACGCCTCTCAATCCATCTCTGGT
ATTCCATCTAGATTCTCCGGTTCTGGTTCCGGTACCGACTTCACTTTGTCCATTAACTCTGTTGAAACCGAAGATTTCGG
TATGTACTTCTGTCAACAATCTAACTCCTGGCCATACACTTTCGGTGGTGGTACCAAGTTGGAATACAAGCACCATCACC
ATCACCATTAATAATCTGATCTGCTCTGCTTACTTTACTAACGACAAAAAAAAATCCAAAAAAAAAAAAAAATCCACTAT
TATTAAAACAAATCAGTCCTTCTCTTCTTACGATATGATATGATTAAATGATACTATGAAATCATCTTCTTCTTAACTTT
CTTAAATCTTACACGTCACTTACTCTATATACCCGTTTAGCTTTGCaTGGTCACAGCGACATTTTATATAAGTGTACGTA
TTTTCTTTTTTTTT

Figure 3 Sequence of inserted scFv fragment

Pink: Homologous regions, Yellow: V_H, Light green: V_L, Red: 6-His-tag

GTGAATTTACTTTAAATCTTGCATttaaataaattttctttttatagctttatgacttagtttcaatttatatactattt
taatgacattttcgattcattgattgaaagctttgtgtttttttcttgatgcgctattgcattgttcttgtcttttcgcc
acatgtaatatctgtagtagatacctgatacattgtggatgctgagtgaaattttagttaataatggaggcgctcttaat
aattttggggatattggctttttttttttaaagtttacaaatgaatttttccgccaggatGCGGCCGCgaattcagggat
gatcttgagaagttcttagagtcttacgagggaacagaagatttggaaccagccaaagctgctatagcggaagcagatat
tttgctaagtaaatgattaaataattaaatatgtggaaatacattaatcttttatatttttgcagttcgttgtcgctat
aatttatagtcatctcgttagttcaaacaagacttcttgaagtgaaaaccaactttcagtcttcaaactaaaaatgaaaa
tcagtggaagaaggtaaacgacttcatgttatatatgaattgaatagtaatggaaataaccaaaaacagctcaacagaaa
acaaacaaaatacgttaagacctgaactcctagcagaaccataactgccaaatatttattatctgtggagatcttatatt
ctaaaaccaaaaaaaatataacttaaaagttaaaaagaagatgttctaactgaggttcgaactcaggacctttgccgtgt

gaaggcaacgtgatagccactacactattagaactaccttatgggaaaaagaaaaatagagtacaactagaatggtaaga
tctgtgacctttctaaacacttaattccatatagacagttcccacccaccataaggtcacaattataatgtctttagaa
gaccactgtcgttcatcatcttcctaagccctctctctaaagcggcatatttccgtaatttgttcttctttgcacaggca
cgtgagatgactccgattattcccacatgcatatttagcCTCTCTAGGGGCTCGAGCTATAGCAAGTTCAAGGAAAGaaa
acTATTATGATCTGGTCACGTGTATAAAATTTATTAATTTTAAAACTATATAATTTATTATTTTTTTATTTTAAAGTTTA
AAGTAATTTTAATAGTATTCTATATTTTAAATAAACATACTTTAAAATTTTATTTAATAATTTATTATTTTTAAATACAA
TGTTTTTATTTAAAACAAAATTATAAGTTAAAAAGTTGTTCCGAAAGTAAAATATATTTTATAGgacgtcGACATGGAGG
CCcagaataccctccttgacagtcttgacgtgcgcagctcaggggcatgatgtgactgtcgcccgtacatttagcccata
catccccatgtataatcatttgcatccatacattttgatggccgcacggcgcgaagcaaaaattacggctcctcgctgca
gacctgcgagcagggaaacgctcccctcacagacgcgttgaattgtccccacgccgcgcccctgtagagaaatataaaag
gttaggatttgccactgaggttcttctttcatatacttccttttaaaatcttgctaggatacagttctcacatcacatcc
gaacataaacaaccatgggtaaggaaaagactcacgtttcgaggccgcgattaaattccaacatggatgctgatttatat
gggtataaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgattgtatgggaagcccgatgcgccaga
gttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacggaat
ttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactgcgatccccggcaaa
acagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgccggtt
gcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaata
acggtttggttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgcataag
cttttgccattctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaatt
aataggttgtattgatgttggacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtg
agttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaattgcagtttcat
ttgatgctcgatgagtttttctaatcagtactgacaataaaaagattcttgttttcaagaacttgtcatttgtatagttt
ttttatattgtagttgttctattttaatcaaatgttagcgtgatttatatttttttcgcctcgacatcatctgcccaga
tgcgaagttaagtgcgcagaaagtaatatcatgcgtcaatcgtatgtGAATGCTGGTCGCTATACTGGAGCTCCAGCTTT
TGTTCCCtttagtgagggttaattgcgcgcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgct
cacaattccacacaacataggagccggaagcataaagtgtaaagcctggggtgcctaatgagtgaggtaactcacattaa
ttgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcgggg
agaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagc
ggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaag
gccagcaaaaggccaggaaccgtaaaaggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcac
aaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccct
cgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctc
atagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcag
cccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagc
cactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctaca
ctagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggc

aaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaaga
tcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaa
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtct
gacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccc
cgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcac
cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctcc
atccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgc
tacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacat
gatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgtta
tcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagta
ctcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgc
cacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttg
agatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagc
aaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttc
aatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaata
ggggttccGCGGCGCGCCGCTAGCGCTCCGATCCTCGGAGGCAGGGTCACCCCAGCCACCTGTTACCCAACCGAAATAGG
AATAAAAAAATCCCCAAAAGAACAAACAAACTTTTCCCCCTTTGCCGAACCTCGGAAAAGACCACTTTTCAATGACAACC
CTTACTCGGCCCCGGTTTCCGGGTACGCTTCCGAATACCCTGCACATGCCCCCTCTGCATGGATGTATGGAGAGGCCCTT
CCTGCTTGCCCCATTGCAAAAAAAATATCACAAACATACCCCTACACTGCAGCCGTTTCTTAGTGCAGTCCAATCACGTC
GCGTATTTGTTCGTGATCACATAATCACAACTACAGCTCTTTCTAAGCATTGAAACCATGTCCTCTTGCATACGCTCCTA
CGGCACTAACCGCGCTGTTCTCTGTCTGGCCCTTGCGTCCCAACCATGCCATGCTGCTTCCTAGACACACACACGCACTC
AGCTCACCAGAGAACAAACAAAACTCCATTTATTCAGATCGCTCCCGTCACTTCAGAAGAGAAAATTTTTTTTCCCATTC
TTCTACTTGTGCTTCTGCTGCGCTGCCTTCGTGTTTCGAGTTGTTTTTAGGTTGGGCGTTAGCCAGTATAATGAAATTGA
GTGATCAATCGGCGGCGAATCGATCAAATTGATGACTTTTTTTTTTTTTTTTTTTGGGAGGTAAATTACGTTGAAAGC
TTGTGCTTATTGTAGTTGTATCTCTATACAAACGTAAATTTTAATATGTACAGTAGTAATTGTAGTAATAATAATAATAA
TAATAATAACACAATTAATTGTATATATAAGGGATTTCCGATTGGCTCCGTCTATATATTTTTTTGTTCTTTGGGCAAG
TATAAAAAGTTCGAATAGTTGTTGCAGTTTTTTGGGAGTCCGTATTTTTTAGTTTATTATTAGTTTATAGGCTATTAGTT
TGTTATAATTGTATACAATTCGTTTTAATCGAACAAGAACAACAAACCATATCATCTGTTGATAGCTAGCTCCGGACCTG
caggGATATGAAGTTAGCATACTCCCTCTTGCTTCCATTGGCAGGAGTCAGTGCTTCAGTGATCAATTACAAGAGAGGAT
CCGACATCCAATTGCAAGAATCTGGTCCATCCTTGGTTAAGCCATCCCAAACCTTGTCTTTGACTTGTTCCGTCACCGGT
GACTCTATCACTTCCGATTACTGGTCTTGGATTAGAAAGTTCCCTGGTAACAGATTGGAATACATGGGTTACGTTTCTTA
CTCCGGTTCTACCTACTACAACCCATCTTTGAAGTCCAGAATCTCTATTACCAGAGACACTTCCAAGAACCAATACTACT
TGGACTTGAACTCTGTCACCACTGAAGATACCGCTACTTACTACTGTGCCAACTGGGACGGTGACTACTGGGGTCAAGGT
ACCTTGGTTACTGTCTCCGCTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGTGGTGGTGGTAGTGACATCGTTTTGAC
TCAATCCCCAGCTACTTTGTCTGTTACCCCTGGTAACTCCGTCTCTTTGTCCTGTAGAGCCTCTCAATCCATCGGTAACA
ACTTGCACTGGTACCAACAAAAGTCCCATGAATCTCCAAGATTGTTGATTAAGTACGCCTCTCAATCCATCTCTGGTATT

CCATCTAGATTCTCCGGTTCTGGTTCCGGTACCGACTTCACTTTGTCCATTAACTCTGTTGAAACCGAAGATTTCGGTAT
GTACTTCTGTCAACAATCTAACTCCTGGCCATACACTTTCGGTGGTGGTACCAAGTTGGAATACAAGCACCATCACCATC
ACCATTAACCCGGGactagt

Figure 4 Sequence of E02-014 plasmid

Gray: ScT$_{TDH3}$, Cyan: KmARS7, Green: KmCEN-D, Yellow: *kanMX*, Light green: KmP$_{ACO1}$, Red: KmINU1ss, Pink: scFv

GCATGCACCATGTTACCCGGGGGGCATCCATGTCTCTCCCTTCCTCTTCCTCTTTTTTTACTAATTTCCCTGTTGTCTTC
CGTTTTTTTCTTTTTTTCCCCTATCGCCAATGCATGACCAAAGTCCTTTTTTTTTTTTCCCCATGACACAGCCCGTTGGAA
CACGCTGAAAACACGCTGAAGAAAAACAAGGGAAACACTGGAACAGTGAAACACCCGCACGAAAATGTTCGGAGTCGGCT
CCGCGTGGTCCCAATTAATAAACAACGTAGCAGCTAGCCAGCTTTAAACTGACTTAGTCCTACGGTTTTGCTGCGTGGGG
GAACCGAGGGAGGGAATCCTCTGGAAGCCGGACCTTGCCCTTTTGACTAGGGTACTACTCTGGGCAGGGTACTTACTAAA
CTGGGCCCAGACAGACAGAGGCCCAGACAGAGGCCCCAAAGAGAAGCCGCCAGGTTCCCTGCCTAGGCCTTCCTCTGGTT
GCCCCCCAGAGTACCCTCCGGAACGAAACGAAACGAAAACGCACATCGTGATGCGCGCTATTATAATTGCGTCTTGCGAA
TTCCATACGCCATGCGCTTAGTTAGCACATGAAATCCTGCCAATATTGACAATCCAACTTCCCAATTCCATTCCCAATTC
CATTCCTACCTAATTCCAGTTCCCATTCATGACCACCGCCGACTCGCCCACAGTTGGCCCACAATTGGCCCATATAATAT
CAGGATATCAGAAAAATAGTTACATATATACAAACAAATTGAAAAATATATATATTATATATTATGGATATTAATATTAT
ATATATATATATGTATATATATCATGACATAAAGCTTTGATTGATAATAATACCTCTACTTCTTGGTATGATAATGTGAT
CTCTTGCTATCTTATTCCTTTATCTAACACCCCCCCCAAGTACATACATATACTACTACACA

Figure 5 Sequence of KmP$_{MDH1}$

GCTCCGATCCTCGGAGGCAGGGTCACCCCAGCCACCTGTTACCCAACCGAAATAGGAATAAAAAAATCCCCAAAAGAACA
AACAAACTTTTCCCCCTTTGCCGAACCTCGGAAAAGACCACTTTTCAATGACAACCCTTACTCGGCCCCGGTTTCCGGGT
ACGCTTCCGAATACCCTGCACATGCCCCCTCTGCATGGATGTATGGAGAGGCCCTTCCTGCTTGCCCCATTGCAAAAAAA
ATATCACAAACATACCCCTACACTGCAGCCGTTTCTTAGTGCAGTCCAATCACGTCGCGTATTTGTTCGTGATCACATAA
TCACAACTACAGCTCTTTCTAAGCATTGAAACCATGTCCTCTTGCATACGCTCCTACGGCACTAACCGCGCTGTTCTCTG
TCTGGCCCTTGCGTCCCAACCATGCCATGCTGCTTCCTAGACACACACACGCACTCAGCTCACCAGAGAACAAACAAAAC
TCCATTTATTCAGATCGCTCCCGTCACTTCAGAAGAGAAAATTTTTTTTCCCATTCTTCTACTTGTGCTTCTGCTGCGCT
GCCTTCGTGTTTCGAGTTGTTTTTAGGTTGGGCGTTAGCCAGTATAATGAAATTGAGTGATCAATCGGCGGCGAATCGAT
CAAATTGATGACTTTTTTTTTTTTTTTTTTTTGGGAGGTAAATTACGTTGAAAGCTTGTGCTTATTGTAGTTGTATCTC
TATACAAACGTAAATTTTAATATGTACAGTAGTAATTGTAGTAATAATAATAATAATAATAACACAATTAATTGTAT
ATATAAGGGATTTCCGATTGGCTCCGTCTATATATTTTTTTTGTTCTTTGGGCAAGTATAAAAAGTTCGAATAGTTGTTG
CAGTTTTTTGGGAGTCCGTATTTTTTAGTTTATTATTAGTTTATAGGCTATTAGTTTGTTATAATTGTATACAATTCGTT
TTAATCGAACAAGAACAACAAACCATATCATCTGTTGATA

Figure 6 Sequence of KmP$_{ACO1}$

SDS-PAGE and immunoblot analysis

To analyze protein production, soluble proteins in the spent culture medium were separated on a SDS-polyacrylamide 12.5% gel (ATTO, Tokyo, Japan) and stained with Bio-Safe Coomassie Stain (Bio-Rad, Hercules, CA). MagicMark™ XP Western Protein Standard (Thermo Fisher Scientific Inc., MA, USA) was included as the molecular weight standard. For western blotting, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane (EDM Millipore, Billerica, MA, USA) by electroblotting. The membrane then was blocked by incubation for 1 h at room temperature with Blocking One (Nakarai tesque, Kyoto, Japan), followed by washing with TBST (0.1 M Tris-HCl, 0.15 M NaCl, 0.05% Tween 20). The membrane then was incubated for 1 h with the primary antibody, rabbit anti-6-His Antibody Affinity Purified (Bethyl Laboratories, TX, USA) diluted 1:5000, followed by washing with TBST and incubation for 1 h with the secondary antibody, Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch Laboratories, PA, USA) diluted 1:10000. Protein bands were detected by ImmunoStar Zeta (Wako, Osaka, Japan).

Enzyme-linked immunosorbent assay (ELISA)

Strains were cultured in YPD or YPX (10 g/L yeast extract, 20 g/L peptone, and 20 g/L xylose) supplemented with 100 mM sodium phosphate buffer, pH 6.0, and selective agent for plasmid-bearing strains in the absence or presence of 200 mM $MgSO_4$ selective medium. Culturing was performed in 96-well deep-well plates at 20°C or 30°C with shaking at 1,200 rpm.

For ELISA, a MaxiSorp plate (Thermo Fisher Scientific Inc.) was coated by distribution of 50 µL per well of lysozyme formulated at 1 µM in 1x phosphate-buffered saline (PBS (10x stock), Nacalai Tesque, Kyoto, Japan) followed by overnight incubation at 4°C. The plate then was blocked at 25°C for 1 h with the blocking solution (ImmunoBlock, DS Pharma Biomedical, Japan) diluted 1:5 in water. The plate was washed three times with PBST (1x PBS supplemented with 0.1% Tween-20 and 2% blocking solution). Supernatants (spent medium) from cultures were diluted five-fold in PBST containing 2% ImmunoBlock and distributed at 50 µL/well. All ELISAs included a blank consisting of 70 µL PBST containing 2% ImmunoBlock. Following incubation

at 25°C for 1 h, the plate was washed as above, and antibody (Anti-His-tag mAb-HRP-DirecT, MLB, Nagoya, Japan), diluted 1:8000 in PBST, was distributed at 50 μL/well. The plate was incubated at 25°C for 1 h and then washed with PBST as above. Color was developed using TMB 1-Component Microwell Peroxidase Substrate Sure Blue and TMB Stop Solution (KLP Inc, Milford, USA) according to manufacturer's instructions. Activity was then measured as absorbance at 450 nm using a SpectraMax Paradigm Multi-Mode Microplate Reader (Molecular Devices Japan, Tokyo, Japan). Relative activity of scFv was obtained by subtracting the value of blank.

Results

Disruption of *INU1* gene by CRISPR-Cas9 system

Wild-type *K. marxianus* predominantly secretes inulinase (Rouwenhorst et al. 1988; Hu et al. 2012). To facilitate the purification of the heterologous protein and re-direct cellular resources for protein production, we deleted the corresponding *INU1* gene. A CRISPR-Cas9 vector for *K. marxianus* (Nambu-Nishida et al. 2017) expressing a pair of guide RNAs (gRNA-1 and gRNA-2) flanking the *INU1* coding region was constructed and used to transform the parent strain; transformants were then screened for the *inu1* mutation (Figure 1a). For integration of the scFv-encoding sequence at the *INU1* locus, we employed strains deficient in the non-homologous end-joining (NHEJ) repair pathway (in this instance, harboring *nej1*° or *dnl4*° null mutations) (Nambu-Nishida et al. 2017) to facilitate homology-directed integration. Another vector expressing a pair of guide RNAs (gRNA-1 and gRNA-3) flanking the *INU1* coding region (Figure 1b) was designed and transformed in combination with an scFv-encoding fragment. The scFv-encoding fragment was flanked with arms (85 bp and 240 bp for the upstream and downstream sequences, respectively) with homology to the *INU1* ORF. PAM sequences of the targets in the arms were mutated to prevent re-cutting after successful integration. Vector-carrying transformant cells were PCR-amplified and subjected to agarose gel electrophoresis (Figure 7). Sequence analysis of the deletion transformant confirmed that the chromosomal *INU1* locus harbored a 2565-bp deletion between the gRNA-2 and gRNA-1 targeting sites, yielding Δ*inu1* (Km02-063) strain

(Figure 1a). Sequence analysis of the transformants from the gene-replacement experiment confirmed that the *INU1* ORF had been replaced by sequences encoding scFv; two of the resulting constructs were designated strains No. 6 (Km02-064) and No.7 (Km02-065) (Figure 1b and 8b). Note that these two strains include *nej1°* or *dnl4°*, respectively. Next, the strains were assessed for inulinase secretion.

SDS-PAGE analysis of spent culture medium recovered from the wild-type strain revealed a single major band at approximately 90 kDa (Figure 8c), consistent with the expected size of inulinase (Hong et al. 2014). Notably, this band was absent in spent medium from cultures of the *inu1* constructs, as expected (Figure 8c).

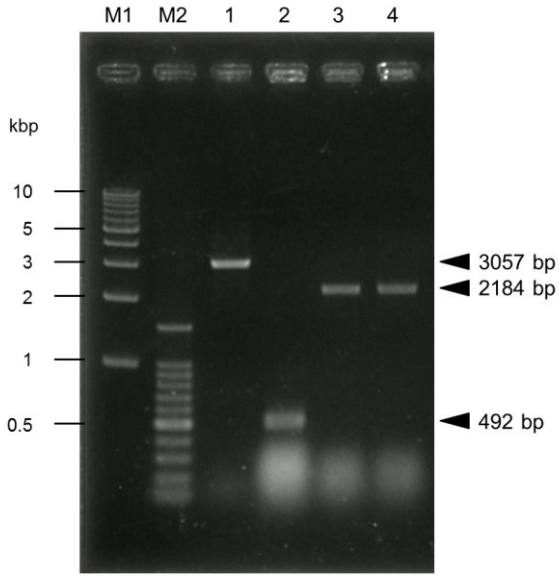
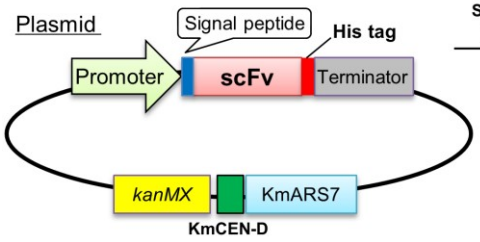


Figure 7 *INU1* deletion and replacement analyzed by agarose gel electrophoresis
Sequences from the transformed cells were PCR-amplified and analyzed by agarose gel electrophoresis. A pair of primers (P_Km01-010 + P_Km01-011) flanking the *INU1* locus were used to amplify intervening sequences. The resulting amplicons from the wild-type strain (lane 1), *inu1*-deleted strain Δ*inu1* (Km02-063) (lane 2), or scFv integration strains No. 6 (Km02-064) (lane 3) and No. 7 (Km02-065) (lane 4) matched the expected fragment sizes of 3057 bp, 492 bp, 2184 bp, and 2184 bp, respectively. DNA ladder markers (1 Kb and 100 bp) are provided as size standards (lanes M1 and M2, respectively).

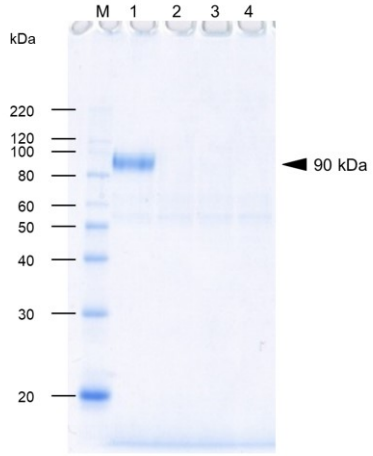
a



b

Strain No.	Host	Promoter	Signal peptide	Expression system
1	WT	-	-	-
2	WT	KmP_{MDH1}	KmINU1ss	Plasmid
3	WT	KmP_{ACO1}	KmINU1ss	Plasmid
4	WT	KmP_{MDH1}	Scα-MFss	Plasmid
5	WT	KmP_{ACO1}	Scα-MFss	Plasmid
6	*Nej1*°	KmP_{INU1}	KmINU1ss	*inu1* replacement
7	*Dnl4*°	KmP_{INU1}	KmINU1ss	*inu1* replacement
8	Δ*inu1*	KmP_{MDH1}	Scα-MFss	Plasmid

c



d

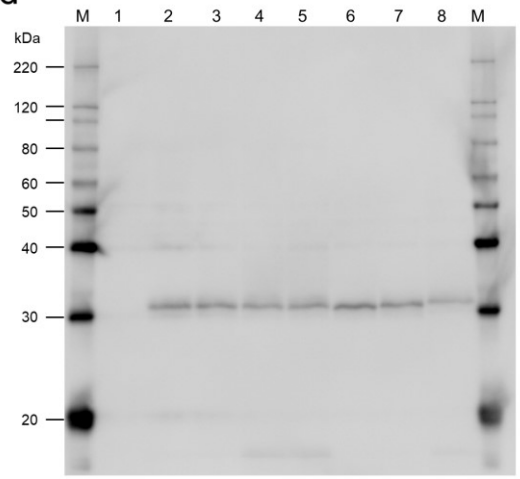


Figure 8 *INU1* deletion analyzed by SDS-PAGE and secretion of scFv analyzed by immunoblotting

(a) Schematic of the general construction of scFv-expression plasmid. *kanMX*: G418 resistance gene, KmCEN-D: centromere sequence of *K. marxianus*, KmARS7: autonomous replication sequence of *K. marxianus*.

(b) Combinations of promoters and signal peptides tested.

(c) Stains were cultivated for 48 h in YPD with or without G418, and SDS-PAGE was used to analyze the spent culture medium from strains No. 1 (wild-type) (lane 1), No. 6 (Km02-064) (lane 2), No. 7 (Km02-065) (lane 3), and No. 8 (Km02-066) (lane 4). Sizes of the molecular weight marker (lane M) are indicated on the left.

(d) Immunoblot of spent culture medium from strains cultivated for 72 h in YPD with or without G418. Lane numbers correspond to those in panel (b).

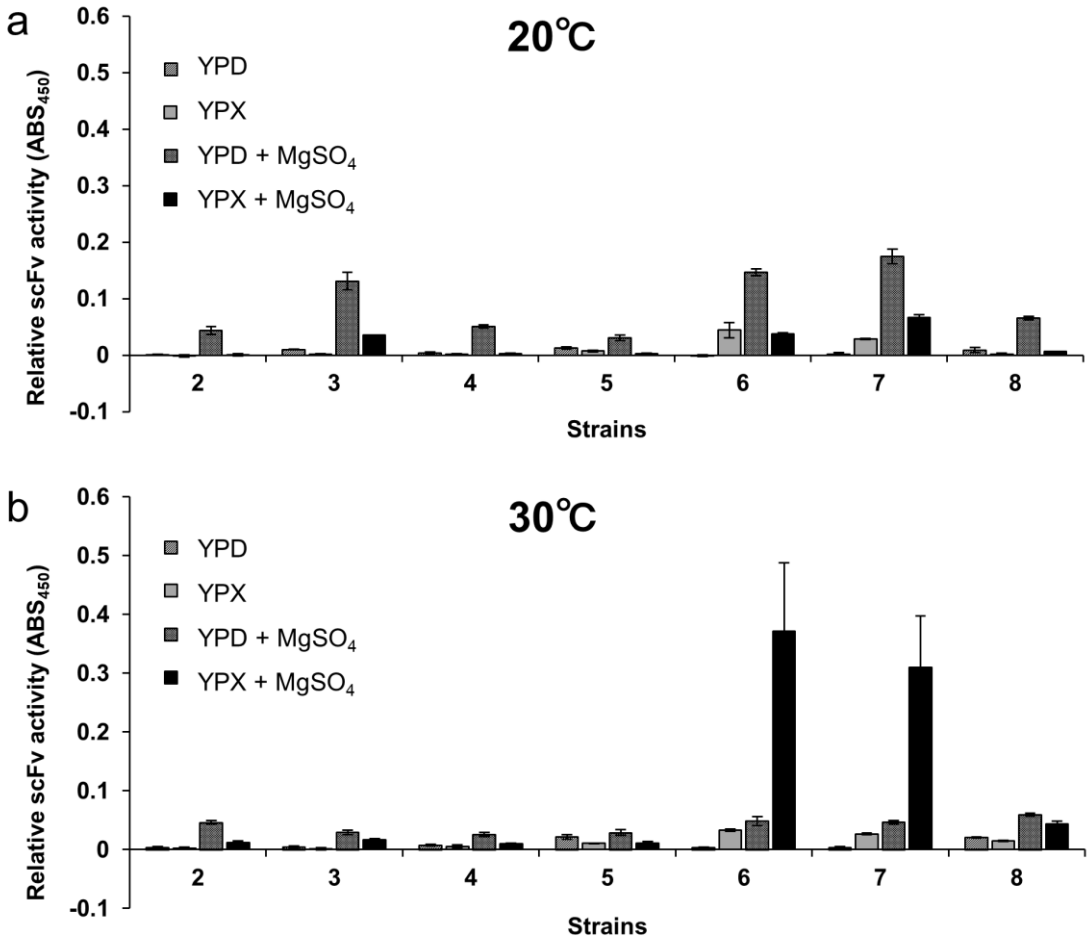


Figure 4 Immunoreactivity of secreted scFv analyzed by ELISA

Strains were cultured in YPD or YPX in the absence or presence of $MgSO_4$ at 20°C (a) or 30°C (b). Spent culture medium was harvested at 72 h and subjected to ELISA using lysozyme as an antigen for scFv. Values are presented as mean ±SEM from three independent experiments.

Expression and secretion of scFv antibody

For the expression of scFv, various expression cassettes containing combinations of promoters (P_{MDH1}, P_{ACO1}) and secretion signal (KmINUss, Scα-MFss) -sequences were introduced, either via plasmid or by genomic integration (Figure 8a and b). Protein expression and secretion was assessed by immunoblotting of spent growth medium (Figure 8d). Secretion of scFv was confirmed as the presence of an approximately 30

kDa protein (Damasceno et al. 2004) in the spent growth medium from each of the transformed strains tested (Figure 8d). These results indicated that these promoters and secretion signals functioned in a modular fashion.

Activity of scFv antibody and improved production by magnesium sulfate supplementation

We next sought to identify growth conditions, including the use of various media, that would yield enhanced expression and secretion of the intact scFv protein. Cells were grown, with shaking in 96-well deep-well plates, at temperatures of 20°C or 30°C in YPD or YPX medium in the presence or absence of various supplements and subjected to ELISA to detect the presence of intact secreted scFv. In this context, ELISA measured the immunoreactivity of scFv. $MgSO_4$ was found to have substantial impact on the antibody production in *K. marxianus* (Figure 9a and b). At 20°C, all strains exhibited increased activity when grown in YPD plus $MgSO_4$ (Figure 9a). At 30°C, more than 10-fold increased activity was observed for strains No. 6 (Km02-064) and No. 7 (Km02-065), both of which harbor constructs introduced by genomic integration, when grown in xylose medium containing $MgSO_4$ (Figure 8b and 9b). Deletion of *INU1* yielded about 4.4-fold increase in scFv activity in xylose medium containing $MgSO_4$, when comparing strain No. 4 (Km02-050) (harboring a plasmid-borne construct) and strain No. 8 (Km02-066) (*INU1* disruptant harboring a plasmid-borne construction) (Figure 8b and 9b).

Discussions

In this study, we demonstrated that *K. marxianus* NBRC1777 can be engineered to express and secrete a single-chain antibody. We showed that secretion of scFv could be changed substantially by use of various genetic constructs and by modification of the growth conditions.

The recently introduced CRISPR-Cas9 and Target-AID genome editing systems (Nambu-Nishida et al. 2017) permit genetic manipulation of organisms that had previously been underexploited because of a lack of genetic tools. In the present work, a sequence encoding scFv was integrated into the *INU1* locus without an associated selection marker. This construct allowed robust expression of the integrated gene

without a need for continued use of selection reagents. The strains (Nos. 6 and 7) carrying the integrated construct showed dramatic increases in scFv immunoreactivity, which is attributed to either increased expression and secretion or improved quality of the antibody, or both. (compared to strains expressing scFv via plasmid-borne constructs) when grown in YPX plus $MgSO_4$ at 30°C (Figure 9b). Further work will be needed to determine whether increased expression requires genomic integration in general or at the *INU1* locus specifically. It is formally possible that episomal plasmids are not well retained during outgrowth, especially if protein expression creates stress for the host cell.

In our hands, YPX medium induced increased expression from the *INU1* promoter but not from the other tested promoters (Figure 9a and b). As the *INU1* product inulinase metabolizes inuline to fructose (Rouwenhorst et al. 1988), *INU1* is downregulated in the presence of glucose, the preferred sugar (Jain et al. 2012). Moreover, *INU1* gene expression is known to be up-regulated when fructose replaces glucose as a sugar source (Schabort et al. 2016), and the *INU1* promoter has consensus binding sequences for MIG1, a known repressor of transcription in the presence of glucose (Bergkamp et al. 1993). As inulinase is the predominant protein secreted by *K. marxianus* (Rouwenhorst et al. 1990; Hu et al. 2012), deletion of the encoding locus is expected to permit re-direction of resources for expression and secretion of heterologous proteins. The present work showed that deletion of *INU1* had a positive but limited impact on scFv production in strain No. 8 (Km02-066). This limited effect may have reflected the use of the KmP_{MDH1} promoter and the Scα-MFss signal peptide. In *Kluyveromyces lactis*, the *Trichoderma reesei* CBH1 secretion signal was more efficient than that of the native α-mating factor for directing the secretion of a reporter, Enhanced Green Fluorescent Protein (EGFP) (Madhavan and Sukumaran 2014). Use of the endogenous INU1 signal peptide in *K. marxianus* may provide more efficient production by directing the heterologous protein into the secretion pathway typically used by inulinase.

In *K. marxianus*, lysine aminopeptidase activity is higher at 30°C than at 20°C (Ramírez-Zavala et al. 2004). It would be valuable to assess the *in vivo* role of various processing enzymes, for instance by suppressing the activity of endogenous proteases by using either protease inhibitors or genetic manipulations.

A positive effect of $MgSO_4$ was observed (to some extent) in all strains and conditions, indicating that $MgSO_4$ generally facilitates the production/secretion of intact scFv in *K marxianus* (Figure 9a and b). Considering the concentration of $MgSO_4$ in the defined media that typically ranges up to 10 mM or so, 200 mM of $MgSO_4$ apparently

exceeded the nutritional demands of the cell. The addition of divalent metal ions, including Mg^{+2}, has been reported to enhance bacterial ell growth and enzyme production (Venkateswarulu et al. 2017; Shahbazmohammadi and Omidinia 2017). The effect of divalent metal ions may result from changes to membrane permeability (Venkateswarulu et al. 2017). In the present study, we observed drastic increases in antibody secretion in the presence of 200 mM $MgSO_4$, a concentration that is ten times higher than that tested in bacteria. While fungal protein secretion pathways differ from those of bacteria, high concentrations of $MgSO_4$ may also affect membrane organization in eukaryotes, facilitating protein secretion and/or stimulating expression of genes that contribute to enhanced protein production and secretion. Although $MgSO_4$ is expected to be involved in a wide range of fungal cellular and biochemical processes, the mechanism of this $MgSO_4$-mediated enhancement of scFv production remains unclear. Nonetheless, our study demonstrated that there is potential for further enhancing fungal protein production by both genetic and physiological manipulations.

Reference

Acchione M, Lee YC, Desantis ME, Lipschultz CA, Wlodawer A, Li M, et al., 2012. Specific fluorine labeling of the HyHEL10 antibody affects antigen binding and dynamics. Biochemistry. 51(30), 6017–6027.

Bergkamp RJM, Bootsman TC, Toschka HY, Mooren AT a, Kox L, Verbakel JM a, et al., 1993. Expression of an a-galactosidase gene under control of the homologous inulinase promoter in *Kluyveromyces marxianus*. Appl. Microbiol. Biotechnol. 40, 309–317.

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, et al., 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(2),115–132.

Buckholz RG, Gleeson MA., 1991. Yeast systems for the commercial production of heterologous proteins. Nat Biotechnol. 9, 1067–1072.

Çelik E, Çalık P., 2012. Production of recombinant proteins by yeast cells. Biotechnol Adv. 30, 1108–1118.

Damasceno LM, Pla I, Chang HJ, Cohen L, Ritter G, Old LJ, et al., 2004. An optimized fermentation process for high-level production of a single-chain Fv antibody fragment in *Pichia pastoris*. Protein Expr. Purif. 37(1), 18–26.

Fonseca GG, Heinzle E, Wittmann C, Gombert AK., 2008. The yeast *Kluyveromyces marxianus* and its biotechnological potential. Appl. Microbiol. Biotechnol. 79(3), 339–354

Gietz D, St Jean A, Woods R a, Schiestl RH., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20(6), 1425.

Hong J, Wang Y, Kumagai H, Tamaki H. Construction of thermotolerant yeast expressing thermostable cellulase genes. 2007. J. Biotechnol. 130(2), 114–123.

Hong SJ, Kim HJ, Kim JW, Lee DH, Seo JH., 2014. Optimizing promoters and secretory signal sequences for producing ethanol from inulin by recombinant Saccharomyces cerevisiae carrying *Kluyveromyces marxianus* inulinase. Bioprocess Biosyst. Eng. 38, 263–272.

Hu N, Yuan B, Sun J, Wang SA, Li FL., 2012. Thermotolerant *Kluyveromyces marxianus* and Saccharomyces cerevisiae strains representing potentials for bioethanol production from Jerusalem artichoke by consolidated bioprocessing. Appl. Microbiol. Biotechnol. 95(5), 1359–1368.

Huang M, Bao J, Nielsen J., 2014. Biopharmaceutical protein production by *Saccharomyces cerevisiae* : current state and future prospects. Pharm Bioprocess 2(2), 167–182.

Jain SC, Jain PC, Kango N., 2012. Production of inulinase from *Kluyveromyces marxianu*s using dahlia tuber extract. Brazilian J Microbiol 43, 62–69.

Lai T, Yang Y, Ng SK., 2013. Advances in mammalian cell line development technologies for recombinant protein production. Pharmaceuticals. 6, 579–603.

Lane MM, Morrissey JP., 2010. *Kluyveromyces marxianus*: A yeast emerging from its sister's shadow. Fungal Biol. Rev. Elsevier Ltd; 24(1–2), 17–26.

Lertwattanasakul N, Rodrussamee N, Suprayogi, Limtong S, Thanonkeo P, Kosaka T, et al., 2011. Utilization capability of sucrose, raffinose and inulin and its less-sensitiveness to glucose repression in thermotolerant yeast *Kluyveromyces marxianus* DMKU 3-1042. AMB Express. 1, 20.

Maccani A, Landes N, Stadlmayr G, Maresch D, Leitner C, Maurer M,et al., 2014. *Pichia pastoris* secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins. Biotechnol J. 9, 526–537.

Madhavan A, Sukumaran RK., 2014. Promoter and signal sequence from filamentous fungus can drive recombinant protein production in the yeast *Kluyveromyces lactis*. Bioresour Technol. 165, 302–308.

Melorose J, Perroy R, Careas S. Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region., 1986. J. Biol. Chem. 261, 5858–5865.

Nambu-Nishida Y, Nishida K, Hasunuma T, Kondo A., 2017. Development of a comprehensive set of tools for genome engineering in a cold- and thermo-tolerant *Kluyveromyces marxianus* yeast strain. Sci. Rep. 7(1), 8993.

Nielsen J., 2013. Production of biopharmaceutical proteins by yeast. Bioengineered 4(4), 207–211.

Raimondi S, Zanni E, Amaretti A, Palleschi C, Uccelletti D, Rossi M., 2013. Thermal adaptability of *Kluyveromyces marxianus* in recombinant protein production. Microb. Cell Fact. 12(1), 34.

Ramírez-Zavala B, Mercado-Flores Y, Hernández-Rodríguez C, Villa-Tanaca L., 2004. Purification and characterization of a lysine aminopeptidase from *Kluyveromyces marxianus*. FEMS Microbiol Lett 235, 369–375.

Rouwenhorst RJ, Hensing M, Verbakel J, Scheffers WA, Van Dijken JP., 1990. Structure and properties of the extracellular inulinase of *Kluyveromyces marxianus* CBS 6556. Appl. Environ. Microbiol. 56(11), 3337–3345.

Rouwenhorst RJ, Visser LE, Van Der Baan a a, Scheffers W a, Van Dijken JP., 1988. Production, Distribution, and Kinetic Properties of Inulinase in Continuous Cultures of *Kluyveromyces marxianus* CBS 6556. Appl. Environ. Microbiol. 54(5), 1131–1137.

Schabort DTWP, Letebele PK, Steyn L, Kilian SG, du Preez JC., 2016. Differential RNA-seq, multi-network analysis and metabolic regulation analysis of *Kluyveromyces marxianus* reveals a compartmentalised response to xylose. PLoS One 11(6), e0156242.

Shahbazmohammadi H, Omidinia E., 2017. Medium optimization for improved production of dihydrolipohyl dehydrogenase from *Bacillus sphaericus* PAD-91 in *Escherichia coli*. Mol. Biotechnol. 59(7), 260–270

Tsumoto K, Nishimiya Y, Kasai N, Ueda H, Nagamune T, Ogasahara K, et al., 1997. Novel selection method for engineered antibodies using the mechanism of Fv fragment stabilization in the presence of antigen. Protein Eng. 10(11), 1311–1318.

Venkateswarulu TC, Prabhakar KV, Kumar RB., 2017. Optimization of nutritional components of medium by response surface methodology for enhanced production of lactase. 3 Biotech. Springer Berlin Heidelberg; 7(3), 202.

Yokota T, Milenic DE, Whitlow M, Schlom J., 1992. Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms. Cancer Res. 52(12), 3402–3408

General Conclusion

In this thesis, I aimed at providing basic information and technology for yeast genetic engineering and cost-effective culture for the sake of industrial production from biomass.

In chapter 1, gene expression profile of *S. cerevisiae* producing ethanol was studied as a model biomass-based production. Detailed time-course changes in the activity of major 30 promoters were quantitatively analyzed in various conditions including xylose utilization, which simulates biomass-derived resource. A few promoters were found to be active in microaerobic xylose fermentation, while most others showed poor activity. These will provide useful information for optimization of gene expression to improve growth and production from biomass in yeast.

In chapter 2, *K. marxianus* was selected to explore its industrial potential as it exhibited good proliferative property and thermotolerance. Comprehensive genetic engineering tools were provided by applying cutting edge genome editing technology for *K. marxianus*. Efficient gene knock-out and knock-in was achieved without use of selection marker. This approach will facilitate engineering of underexploited yeasts and other microbes for industrial applications.

Antibody production as high value-added product was performed using *K. marxianus* in chapter 3. Genetic engineering, promoter selection and optimization of media condition were together proved to be effective in the improvement of production.

The research on *S. cerevisiae* as eukaryotic model provides highly integrated and comprehensive understanding with the aid of highly accumulated previous researches. Genome editing technology now allows easy genetic manipulation of minor yeasts that have useful characteristics. This thesis demonstrated an approach in which studying model organism as reference and establishing new host strain with practical advantages coordinately facilitate industrial application for bio-based material production.

Acknowledgments

This is a thesis submitted to Kobe University for the degree of Doctor of Engineering. The studies reported here were carried out between 2014 and 2017 under the direction of Professor Akihiko Kondo in the Laboratory of Bioproduction Engineering, Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University.

I would like to express my sincere gratitude to my supervisor, Professor Akihiko Kondo for his continuous guidance and invaluable suggestions during the course for my studies. I am grateful to my second supervisor, Professor Tomohisa Hasunuma for his support and constructive comments throughout this research. I would like to express the deepest appreciation to Professor Keiji Nishida and Associate Professor Jun Ishii for their generous support and persistent assistance during the research period. I would like to extend my heartfelt gratitude to Dr. Yuri Sakihama for her help and valuable suggestions throughout the research. I would like to thank Professor Chiaki Ogino for his advice and comments on this thesis. I am also grateful to Professor Hideki Yamaji and Professor Minoru Mizuhata for providing constructive comments for this thesis.

I would like to thank Takanobu Yoshida for his technical assistance and encouragement.

I would like to thank my family, Yoshiaki, Michiyo, and Junichi Nambu and pu-ta, hana for their unfailing support and encouragement.
Last but not least, I extend my sincere appreciation to my husband, Keiji Nishida for being there for me all this while.

Yumiko Nishida
Department of Chemical Science and Engineering
Graduate School of Engineering

Publication Lists

Nambu-Nishida Y, Sakihama Y, Ishii J, Hasunuma T, Kondo A., 2017. Selection of yeast *Saccharomyces cerevisiae* promoters available for xylose cultivation and fermentation. *Journal of Bioscience and Bioengineering*, 125, 76-86.

Nambu-Nishida Y, Nishida K, Hasunuma T, Kondo A., 2017. Development of a comprehensive set of tools for genome engineering in a cold- and thermo-tolerant *Kluyveromyces marxianus* yeast strain. *Scientific Reports*, 7, 8993

Nambu-Nishida Y, Nishida K, Hasunuma T, Kondo A., Genetic and physiological basis for antibody production by *Kluyveromyces marxianus*. *– Under preparation.*

Doctor Thesis, Kobe University
“Development of yeast genome-engineering technology and its application to antibody production”, 106 pages
Submitted on January, 18th, 2018
The date of publication is printed in cover of repository version published in Kobe University Repository Kernel.