

# UC Irvine

## UC Irvine Previously Published Works

### Title

Developing a broad-range promoter set for metabolic engineering in the thermotolerant yeast *Kluyveromyces marxianus*.

### Permalink

<https://escholarship.org/uc/item/7w95r170>

### Authors

Lang, Xuye

Besada-Lombana, Pamela

Li, Mengwan

et al.

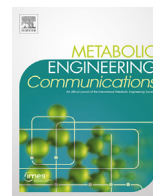
### Publication Date

2020-12-01

### DOI

10.1016/j.mec.2020.e00145

Peer reviewed



## Developing a broad-range promoter set for metabolic engineering in the thermotolerant yeast *Kluyveromyces marxianus*



Xuye Lang<sup>a,1</sup>, Pamela B. Besada-Lombana<sup>b,1</sup>, Mengwan Li<sup>a</sup>, Nancy A. Da Silva<sup>b,\*\*</sup>, Ian Wheeldon<sup>a,c,\*</sup>

<sup>a</sup> Department of Chemical and Environmental Engineering, UC Riverside, United States

<sup>b</sup> Department of Chemical and Biomolecular Engineering, UC Irvine, United States

<sup>c</sup> Center for Industrial Biotechnology, UC Riverside, United States

### ARTICLE INFO

#### Keywords:

Thermo-tolerance  
Non-conventional microbe  
Promoter  
Chemical production  
*K. marxianus*

### ABSTRACT

*Kluyveromyces marxianus* is an emerging host for metabolic engineering. This thermotolerant yeast is the fastest growing eukaryote, has high flux through the TCA cycle, and can metabolize a broad range of C5, C6, and C12 carbon sources. In comparison to the common host *Saccharomyces cerevisiae*, this non-conventional yeast suffers from a lack of metabolic engineering tools to control gene expression over a wide transcriptional range. To address this issue, we designed a library of 25 native-derived promoters from *K. marxianus* CBS6556 that spans 87-fold transcriptional strength under glucose metabolism. Six promoters from the library were further characterized in both glucose and xylose as well as across various temperatures from 30 to 45 °C. The temperature study revealed that in most cases EGFP expression decreased with elevating temperature; however, two promoters, P<sub>SSA3</sub> and P<sub>ADH1</sub>, increased expression above 40 °C in both xylose and glucose. The six-promoter set was also validated in xylose for triacetic acid lactone (TAL) production. By controlling the expression level of heterologous 2-pyrone synthase (2-PS), the specific TAL titer increased over 8-fold at 37 °C. Cultures at 41 °C exhibited a similar TAL biosynthesis capability, while at 30 °C TAL levels were lower. Taken together, these results advance the metabolic engineering tool set in *K. marxianus* and further develop this new host for chemical biosynthesis.

### 1. Introduction

Over the past few decades, the US industrial biotechnology sector has been growing at an average rate of ~10% per year and is estimated to be 2-3% of US gross domestic product (GDP) (Carlson, 2016). The global biotechnology market has been growing at a similar rate and is estimated to be upward of \$700 billion by 2025 (Kazmierczak et al., 2019). The continued expansion of this industry requires new bioprocessing technologies and new metabolic engineering hosts that can meet industrial needs. For example, microbial hosts with high native stress tolerance can benefit bioprocessing by enabling high temperature fermentations, the production of toxic solvents and compounds, and the potential use of low cost water sources including industrial wastes and sea water high in salts and other minerals (Thorwall et al., 2020).

Due to its native stress tolerance, the yeast *Kluyveromyces marxianus* is a promising eukaryotic host for bioproduction. Various strains of this

species are able to grow at 45 °C and above and in media with high salt concentration (≥5% NaCl) (Andreu and Del Olmo, 2020; Nonklang et al., 2008; Yuan et al., 2008; Lane and Morrissey, 2010). *K. marxianus* also benefits from the ability to metabolize a range of C5, C6, and C12 sugars and exhibits the fastest growth rate (~0.7 h<sup>-1</sup>) of any known eukaryote (Löbs et al., 2016, 2017; Rodrussamee et al., 2011). Like other non-conventional microbes, there has been considerably less research effort put towards developing well-characterized genetic parts for manipulating gene expression in comparison to the common laboratory host and ethanol producer *Saccharomyces cerevisiae* (Löbs et al., 2017). Despite this general lack of metabolic engineering tools, simple genetic manipulations and media optimization have been used to demonstrate the potential of *K. marxianus* for a variety of bioprocesses including protein production, high temperature ethanol fermentation, the biosynthesis of ethyl acetate (a native metabolite produced in grams per liter quantities in many *K. marxianus* strains) and heterologous products such

\* Corresponding author. Department of Chemical and Environmental Engineering, UC Riverside, United States.

\*\* Corresponding author. Department of Chemical and Biomolecular Engineering, UC Irvine, United States.

E-mail addresses: [nadasilva@uci.edu](mailto:nadasilva@uci.edu) (N.A. Da Silva), [wheeldon@ucr.edu](mailto:wheeldon@ucr.edu) (I. Wheeldon).

<sup>1</sup> Equal Contribution.

<https://doi.org/10.1016/j.mec.2020.e00145>

Received 29 April 2020; Received in revised form 17 August 2020; Accepted 31 August 2020

2214-0301/© 2020 The Authors. Published by Elsevier B.V. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY

license (<http://creativecommons.org/licenses/by/4.0/>).

as the polyketide triacetic acid lactone (TAL) (Cernak et al., 2018; Raimondi et al., 2013; Kushi et al., 2000; Lertwattanasakul et al., 2011; Yarimizu et al., 2015; Madeira and Gombert, 2018; McTaggart et al., 2019). The development of new metabolic engineering tools (e.g., promoters, gene editing systems, gene integration strategies, etc.) is needed to advance the current successes as well as to develop new strains for the high titer production of chemicals, fuels, and biologics using *K. marxianus* as a microbial host.

It has previously been demonstrated that a number of *S. cerevisiae* promoters can function in *K. marxianus*, but expression level is generally low and their ability to operate at elevated temperatures has not yet been established (Lee et al., 2013). We have leveraged this minimal functionality to adapt CRISPR-Cas9 and CRISPR interference (CRISPRi) systems for use in *K. marxianus*. Specifically, plasmid expression of Cas9 or deactivated Cas9 (dCas9) was accomplished using the heterologous *S. cerevisiae* *TEF1* promoter (Löbs et al., 2017, 2018). In addition to the development of CRISPR-based genome editing and gene regulation tools, recent efforts have identified a set of native constitutive and inducible promoters in *K. marxianus* (Rajkumar et al., 2019). While these new systems greatly enhance our ability to edit the *K. marxianus* genome and tune gene transcription, there is still a need for a set of well-characterized promoters that span a broad range of expression over a wide temperature spectrum and with different carbon sources.

We have expanded the existing promoter set by characterizing a library of 25 putative promoters comprising the region between the start codon to 700 bp upstream the start codon of selected genes in *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  (Lee et al., 2015). The resulting library enables a range of heterologous protein expression of greater than 80-fold. We also determined promoter dynamics through lag, exponential, and stationary phases, promoter strength in glucose and xylose media, and evaluated temperature effects on expression. Finally, we selected a subset of six promoters that retained a wide range of transcriptional strengths in xylose to overexpress 2-pyrone synthase (2-PS) for the synthesis of the polyketide TAL, and compared the strongest promoter with the *S. cerevisiae* *ADH2* promoter for high level production.

## 2. Materials and methods

**Media and cultivation.** *Escherichia coli* strains XL-1 Blue and TOP10 were cultivated in 5 mL of Luria Bertani (LB) medium containing 50–100  $\mu\text{g/mL}$  of ampicillin. Overnight cultures of *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  were grown in 2 mL of 2% YPD (10 g/L yeast extract (BD Difco™), 20 g/L of peptone (BD Difco™) and 20 g/L of D-glucose (Fisher Scientific) for plasmid transformation. The *HIS3* disruption was created by CRISPR-Cas9 targeting the *HIS* locus in a previously reported CBS6556  $\Delta URA3$  strain (Table 1). For expression experiments using *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$ , strains were grown in 250 mL baffled flasks using 25 mL of SD-His (20 g/L D-glucose (Fisher Scientific), 6.7 g/L yeast nitrogen base (BD Difco™), 0.77 g/L CSM-His (Sunrise)) or SX-His (20 g/L xylose (Alfa Aestar), 5 g/L ammonium sulfate (Fisher Scientific), 1.7 g/L yeast nitrogen base without amino acids (BD Difco™), and 0.77 g/L CSM-His (Sunrise)) at 250 rpm and 30, 37, 41 or 45 °C. SD stands for synthetic defined media with dextrose and SX denotes synthetic defined media with xylose. For all media, carbon sources were autoclaved separately, and yeast nitrogen base solutions were filtered. For TAL production experiments using the *K. marxianus* industrial strain KM1  $\Delta URA3$  (Table 1), cells were grown at 37 °C in 15 × 125 mm borosilicate culture tubes containing 3 mL of 1% SCA media (5 g/L caseamino acids (BD Difco™), 5 g/L ammonium sulfate (Fisher Scientific), 1.7 g/L yeast nitrogen base without amino acids (BD Difco™), 100 mg/L adenine hemisulfate (Sigma)) and 10 g/L xylose (1% SXCA) or 9.5 g/L lactose (0.95% SLCA) (providing an equimolar amount of carbon).

**Cloning and plasmid construction:** All plasmids used in this work are listed in Table 1, with primers summarized in Table S1 and promoter sequences summarized in Table S2. Primers (DNA oligos) were synthesized by Integrated DNA Technology (IDT). All promoter fragments were

amplified from extracted yeast genome of *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  by PCR using Q5® High-Fidelity DNA Polymerase purchased from New England BioLabs Inc (NEB). Expression plasmids for EGFP were created by Gibson assembly using NEBuilder® HiFi DNA Assembly Master Mix from NEB. Restriction enzymes and T4 ligase for construction of the majority of the 2-PS expression plasmids were purchased from NEB. All plasmids transformed in yeast were confirmed by Sanger sequencing from either Source BioScience or Genewiz.

For EGFP expression plasmid construction, the 700 bp long putative promoter of a given gene was identified and cloned from the genome of *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  (see Table S1 for primers). Agarose gel (1%) electrophoresis was used to confirm and purify amplified promoter fragments from PCR. Fragments with the correct length recovered from gel were then inserted into a *K. marxianus* EGFP expression vector (pIW1198) by replacing *S. cerevisiae* *P<sub>GPD1</sub>*, the original promoter upstream of EGFP. The plasmid backbone was prepared by digestion with *SacII* and *XmaI* for 1 h at 37 °C.

For construction of 2-pyrone synthase (2-PS) expression plasmids, pIW578 and pXP842-2PS (Table 1) were both digested with *SpeI* and *XhoI* and the 1233 bp fragment from pXP842-2PS was ligated to the linearized pIW578. Using *SpeI* and *NotI* restriction enzymes, *K. marxianus* *ADH1*, *HHF1*, *NC1*, *PGK1* and *SSA3* promoters and the *S. cerevisiae* *ADH2* promoter (*P<sub>ScADH2</sub>*) were then inserted upstream of 2-PS by replacing the *S. cerevisiae* *GPD1* promoter from pIW578-2PS with PCR amplicons from primer pairs FW-ADH1-2PS + RV-ADH1-2PS, FW-HHF1-2PS + RV-HHF1-2PS, FW-NC1-2PS + RV-NC1-2PS, FW-PGK-2PS + RV-PGK-2PS, FW-SSA3-2PS + RV-SSA3-2PS, and FW-ADH2Sc-2PS + RV-ADH2Sc-2PS respectively, using the corresponding EGFP-harboring vectors (*P<sub>ADH1</sub>*-EGFP, *P<sub>HHF1</sub>*-EGFP, *P<sub>NC1</sub>*-EGFP, *P<sub>PGK</sub>*-EGFP, and *P<sub>SSA3</sub>*-EGFP) and pKD-A2PS as templates. This resulted in the creation of *P<sub>ADH1</sub>*-2PS, *P<sub>HHF1</sub>*-2PS, *P<sub>NC1</sub>*-2PS, *P<sub>PGK</sub>*-2PS, *P<sub>SSA3</sub>*-2PS and *P<sub>ScADH2</sub>*-2PS. To generate *P<sub>TEF3</sub>*-2PS, the *SpeI* and *NotI*-digested pIW578-2PS plasmid was Gibson-assembled with the PCR product from primer pair FW-TEF3-2PS + RV-TEF3-2PS using *P<sub>TEF3</sub>*-EGFP as a template. This was necessary due to the presence of an internal *SpeI* restriction site in the selected *TEF3* promoter region. Finally to create pKD-N2PS, the *ADH2* promoter from *S. cerevisiae* was substituted by the *K. marxianus* *NC1* promoter in the high-copy plasmid pKD-A2PS (McTaggart et al., 2019) by introducing the PCR amplification product from primer pair FW-pKD1-NC1-2PS + RV-pKD1-NC1-2PS using *P<sub>NC1</sub>*-2PS as a template and the restriction sites *ZraI* and *SpeI*.

***K. marxianus* transformation:** CBS6556  $\Delta HIS3 \Delta URA3$  cells were transformed using the protocol described by Löbs et al., 2017. Briefly: (1) Cells were cultured in 2 mL of 2% YPD overnight at 30 °C; (2) 400  $\mu\text{L}$  of cell culture were harvested by centrifugation at 5000 rpm for 1 min; (3) harvested cells were washed with 1 mL of sterile water twice by resuspension and centrifugation at 5000 rpm for 1 min; (4) 10  $\mu\text{L}$  of carrier ssDNA (previously boiled for 5 min at 100 °C and chilled) were added to resuspend the cell pellets by gentle vortexing; (5) 500 ng of plasmid and 400  $\mu\text{L}$  of transformation buffer (40% PEG, 0.1 M LiAc, 10 mM Tris-HCl, 1 mM EDTA, 70 mM DTT pH 7.5) were added and mixed well by pipetting; (6) the mixture was incubated at room temperature for 15 min, and heat-shocked for another 15 min at 47 °C. Cell pellets were then collected by centrifugation at 5000 rpm for 1 min and resuspended in 500  $\mu\text{L}$  of selective media (SD-His). To obtain single colonies, 50  $\mu\text{L}$  of the cell resuspension was plated in SD-His plates with 2% agar. KM1  $\Delta URA3$  cells were transformed via electroporation using an adaptation of the method described by (Benatuil et al., 2010) Briefly, (1) cells were grown overnight at 30 °C in two 250 mL baffled flasks containing 25 mL of YPD 2% each; (2) cell cultures were pooled and collected by centrifugation (2400 g, 3 min, 4 °C), (3) washed twice with 25 mL of iced-cold water and (4) once with 25 mL of ice-cold Electroporation Buffer (1 M Sorbitol, 1 mM CaCl<sub>2</sub>); (5) yeast cells were conditioned for 30 min at 30 °C in 10 mL of a 0.1 M LiAc/10 mM DTT solution and (6) then washed once again with 25 mL of ice-cold Electroporation Buffer; (7) the pellet was re-suspended in 1.6 mL of electroporation buffer and 400  $\mu\text{L}$  of that

**Table 1**  
Strains and plasmids.

Name	Description	Reference
<b>Strains</b>		
<i>K. marxianus</i> CBS6556 $\Delta$ HIS3 $\Delta$ URA3	<i>his3</i> $\Delta$ ura3 $\Delta$	This study
<i>K. marxianus</i> KM1 $\Delta$ URA3	<i>ura3</i> $\Delta$	Pecota, Rajgarhia and Da Silva, 2007
<i>E. coli</i> TOP10	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galU</i> <i>galK</i> $\lambda$ <sup>-</sup> <i>rpsL</i> (Str <sup>R</sup> ) <i>endA1</i> <i>nupG</i> <i>recA1</i> <i>endA1</i> <i>gyrA</i> 96 <i>thi-1</i> <i>hsdR17</i> <i>supE</i> 44 <i>relA1</i> <i>lac</i> [F' <i>proAB</i> <i>lacIq</i> $\Delta$ M15 Tn10 (Tetr)]	Thermo Fisher
XL-1 blue		Thermo Fisher
<b>Plasmids</b>		
pIW1198	<i>K.m.</i> CEN/ARS, P <sub>GDPSc</sub> -GFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
pIW578	<i>K.m.</i> CEN/ARS, P <sub>GDPSc</sub> -T <sub>CYC1Sc</sub> , HIS3 marker	This study
pXP842-2PS	2 $\mu$ plasmid, P <sub>ADH2Sc</sub> -2-PS-T <sub>CYC1Sc</sub> , URA3 marker	Cardenas and Da Silva (2014)
P <sub>ADH1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>ADH1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>INU1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>INU1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>RPE1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>RPE1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PIR1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>PIR1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>POL4</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>POL4</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HSP26</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HSP26</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>SSA3</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>SSA3</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>ZWF</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>ZWF</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>SCL1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>SCL1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>ALD2</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>ALD2</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PST1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>PST1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>GLK1A</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>GLK1A</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>COX20</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>COX20</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>SOD1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>SOD1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>GPD1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>GPD1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>GLK1B</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>GLK1B</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HSP60</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HSP60</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>TDH3</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>TDH3</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PGK</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>PGK</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HTB2</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HTB2</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HTB1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HTB1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HHF1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HHF1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HHF2</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>HHF2</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>TEF3</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>TEF3</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>NC1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>NC1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PDC1</sub> -EGFP	<i>K.m.</i> CEN/ARS, P <sub>PDC1</sub> -EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PDC1</sub> (1400bp)- EGFP	<i>K.m.</i> CEN/ARS, P <sub>PDC1</sub> (1400bp)- EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PDC1</sub> (1700bp)- EGFP	<i>K.m.</i> CEN/ARS, P <sub>PDC1</sub> (1700bp)- EGFP-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>ADH1</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>ADH1</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>PGK</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>PGK</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>SSA3</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>SSA3</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>HHF1</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>HHF1</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>TEF3</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>TEF3</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>NC1</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>NC1</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
P <sub>ScADH2</sub> -2PS	<i>K.m.</i> CEN/ARS, P <sub>ScADH2</sub> -2PS-T <sub>CYC1Sc</sub> , HIS3 marker	This study
pKD-A2PS	<i>K. lactis</i> pKD1, P <sub>ScADH2.c</sub> -2PS-T <sub>CYC1Sc</sub> , URA3 marker	McTaggart et al. (2019)
pKD-N2PS	<i>K. lactis</i> pKD1, P <sub>NC1</sub> -2PS-T <sub>CYC1Sc</sub> , URA3 marker	This study

solution were mixed with 1  $\mu$ g of plasmid; (8) cells then electroporated in a 2 mm electrode gap cuvette (VWR) at 2.0 kV and 25  $\mu$ F; (9) and then washed with 1 mL of pre-warmed SDCA media. A 1:2500 dilution of the cell suspension was plated on SDCA to obtain single colonies.

**Cell culture and growth curve measurements.** CBS6556  $\Delta$ HIS3  $\Delta$ URA3 cells were inoculated from fresh SD-His plates into 1 mL of liquid SD-His and grown overnight at 30 °C. Cells were then transferred to 250 mL baffled flasks (initial OD<sub>600</sub> = 0.05) and grown until late stationary phase in 25 mL of SD-His. For each sample, the cell suspension was diluted 20 times and the optical density at 600 nm was measured using Nanodrop 2000c (Fisher Scientific) with a 1 cm light pathlength cuvette. For expression experiments in xylose, cells were inoculated from fresh SX-His plates (with 2% agar) into 1 mL of liquid SX-His and grown overnight at 30 °C. Cells were then transferred (initial OD<sub>600</sub> = 0.05) to fresh 1 mL of SX-His and cultivated for approximately 18 h. This culture was used as the inoculum for the experimental cultures in 25 mL of SX-His 250 mL baffled flasks (initial OD<sub>600</sub> = 0.05), which were grown until late exponential or late stationary phase. KM1  $\Delta$ URA3 strains were cultivated as described by (McTaggart et al., 2019). Briefly, cells were inoculated from fresh plates into 3 mL of liquid 1% SXCA or 0.95% SLCA and grown overnight at 37 °C. Cells were then re-inoculated into 3 mL of fresh media

at OD<sub>600</sub> = 0.1 and grown for 48 h at 37 °C in an orbital water bath shaker (Amerex Instruments, Inc., Model SK-929). For each sample, cells were diluted and optical density was measured at 600 nm in a Shimadzu UV-2450 spectrophotometer.

**EGFP thermal stability.** *K. marxianus* CBS6556  $\Delta$ HIS3  $\Delta$ URA3 harboring an EGFP expression plasmid was cultured in SD-His medium at 30 °C with 0.05 initial OD<sub>600</sub>. Cells were harvested by 10 min centrifugation at 5000g after 14 h of culture. After washing twice with 40 mL PBS, cells were resuspended with 100 mL 1XPBS. After sonication, cell lysate was collected by centrifugation at 11,000 g for 30 min at 4 °C. The cell lysate was incubated at 30, 37 and 45 °C for different time. The fluorescence signal intensity of incubated lysate was measured at 0, 1, 2, 4, 6, 8, 16, 20 h using the BioTek®Synergy™ Neo2 multi-mode microplate Reader (excitation: 488 nm, emission: 511 nm).

**Dynamic EGFP fluorescence.** One hundred microliters of the cell suspension were transferred to a 96-well Nunc black plate (Thermo Scientific) and fluorescence was measured in a plate reader (SpectraMax M2, Molecular Devices, or BioTek Synergy™ Neo2 multi-mode microplate Reader) using an excitation wavelength of 488 nm and a 511 nm emission wavelength. For measurements using the BioTek Synergy™ Neo2 multi-mode microplate Reader, the cell suspension was diluted 20 times.

For the EGFP fluorescence study, *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  with blank plasmid (pIW578, see Table 1) was employed as the background of fluorescence intensity. All relative fluorescence intensity (RFU) shown in this study are background subtracted.

**Flow cytometry.** After 14 h of culture, cells were harvested by centrifugation at 5000 rpm, and washed twice with PBS. Samples were then diluted 50-fold in water and the fluorescence signal intensity per cell was determined (excitation wavelength: 488 nm, filter range: 533/30 nm).

**Determination of triacetic acid lactone levels.** Triacetic acid lactone concentrations were determined via HPLC-UV, as described by McTaggart et al. (2019), or spectrophotometrically using a plate reader. Control experiments were used to confirm that both methods provide the same results under our experimental conditions. For the HPLC-UV assay, samples were centrifuged at 2500 $\times$ g for 5 min and 1 mL of the supernatant was collected and stored at 4 °C for further analysis. Twenty microliters of the samples were injected into an HPLC (Shimadzu, Japan) equipped with a UV-Vis detector (SPD-10 A VP; Shimadzu). A Zorbax SB-C18 reversed-phase column (2.1 Å~ 150 mm; Agilent Technologies, Santa Clara, CA) was used as a stationary phase. The mobile phase consisted of acetonitrile buffered in 1% acetic acid with a gradient of 95%–85% of water buffered in 1% acetic acid. This provided an elution time of approximately 12 min when the equipment ran at a flow rate of 0.25 mL/min and the temperature of the column was kept at 25 °C. For the spectrophotometric assay, the supernatant was diluted 20-fold and absorbance was measured in UV transparent flat bottom plate (Corning) using the plate reader at 277 nm (SpectraMax M3, Molecular Devices).

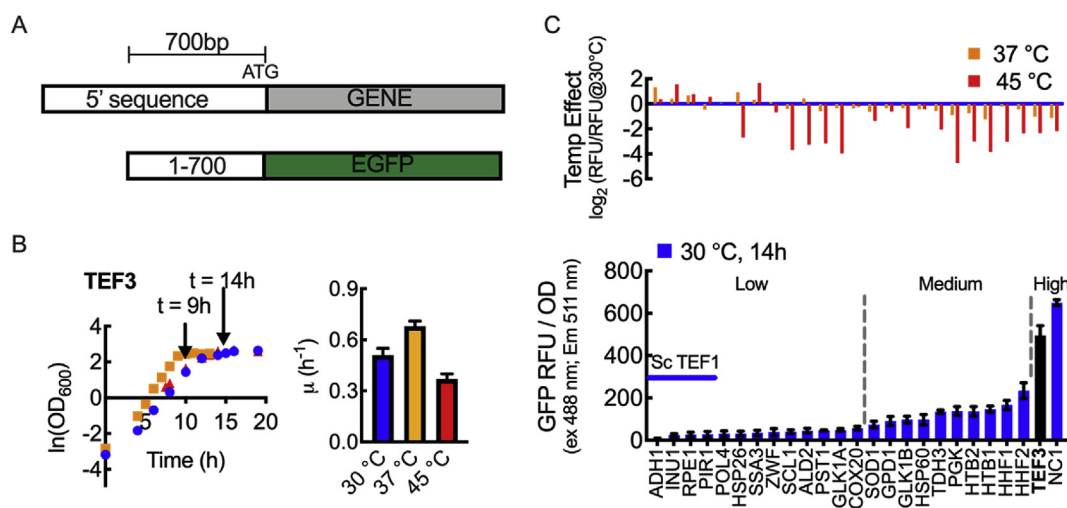
### 3. Results

All strains and plasmids used in this study are presented in Table 1. The plasmids include a set of vectors encoding EGFP with expression driven by one of 25 different promoters identified in *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$ . Promoters were selected by homology to the well-characterized *S. cerevisiae* genome and through previously published transcriptional analyses (Lee et al., 2015; Fu et al., 2019; Lertwattanasakul et al., 2015). Nine *K. marxianus* genes, *ADH1*, *HHF1*, *HHF2*,

*HTB1*, *HTB2*, *PGK*, *THD3*, *TEF3* and *NC1* (putative homolog of *ScCCW12*), were identified by translated BLAST (tblastn) using protein queries from *S. cerevisiae*. These nine promoters already showed relatively high strength in *S. cerevisiae*. We assumed that similar promoter strength might also be observed in *K. marxianus*. Sixteen additional *K. marxianus* genes were selected because they were shown to be up-regulated at 45 °C from reported transcriptomics data: *PST1*, *GLK1A*, *GLK1B*, *SCL1*, *ZWF*, *RPE1*, *GPD1*, *ALD2*, *HSP26*, *HSP60*, *PIR1*, *SSA3*, *SOD1*, *INU1*, *POL4*, and *COX20*. For each identified open reading frame (ORF), the 700 bp upstream of the start codon were cloned from the CBS6556  $\Delta HIS3 \Delta URA3$  genome and inserted upstream of EGFP to construct the promoter screening library (Fig. 1A). The stability of EGFP assay indicated that EGFP was stable at 30, 37, and 45 °C for upward of 20 h (Fig. S1).

We elected to use only 700 bp for each promoter because a similar strategy was successful in creating native-derived promoters from *S. cerevisiae* sequences and because we wanted to minimize promoter length while maintaining core promoter and upstream activating sequence (UAS) function (Lee et al., 2015). We recognize that in some cases critical upstream regions may not have been incorporated within the tested sequences, thus resulting in expression level differences from the full-length native promoters. For example, truncation of the *K. marxianus* *PDC1* promoter ( $P_{PDC1}$ ) from 1400 bp down to 700 bp dramatically reduced expression (Fig. S2). Limiting sequences to 700 bp, however, was successful in creating a series of functional native-derived promoters. The *K. marxianus* *TEF3* promoter exemplifies the strategy; expression from the 700 bp *TEF3* promoter ( $P_{TEF3}$ ) was nearly 3-fold greater than full-length  $P_{PDC1}$  at 30 °C, which was previously identified as a high expression level promoter (Rajkumar et al., 2019).

With the promoter library in-hand, we first evaluated *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  growth and EGFP expression during lag, exponential, and stationary phases with defined glucose media at 30, 37, and 45 °C. One example is illustrated in Fig. 1B, where growth curves and growth rates of CBS6556  $\Delta HIS3 \Delta URA3$  with  $P_{TEF3}$ -driven EGFP expression at various temperatures are shown. Growth rate was highest at 37 °C (0.7 h<sup>-1</sup>), with rapid growth also occurring at 30 and 45 °C (0.5 and 0.3



**Fig. 1.** Screening and characterization of *K. marxianus* promoters for heterologous protein expression. (A) Twenty-five EGFP expression cassettes were constructed by identifying and cloning the 700 base pairs upstream of a start codon of a given *K. marxianus* gene. (B) Growth and growth rate of *K. marxianus* CBS6556  $\Delta HIS3 \Delta URA3$  in 2% glucose harboring a low copy number plasmid with EGFP expression driven by  $P_{TEF3}$ . Shake flask cultures were inoculated with an initial OD<sub>600</sub> of 0.05. Triplicate cultures were grown at 30, 37, and 45 °C. Arrows indicate the early stationary phase time points used for expression characterization across the promoter library. (C) Cell density normalized EGFP fluorescence (Relative Fluorescence Unit, RFU) at 30 °C for the promoter set. The  $P_{TEF3}$  result is shown in black for ease of reference to part (B). The effect of temperature (37 and 45 °C) is indicated above as the log<sub>2</sub> fold change from fluorescence observed at 30 °C. In the lower panel, the strength of *S. cerevisiae* *TEF1* promoter is indicated by the horizontal blue line as a control. Vertical dashed lines show how promoters were grouped into low, medium and high strength sub-groups based on the EGFP RFU/OD at 30 °C. All data represent biological triplicates. Data points and bars indicate the mean; the error bars indicate the standard deviation. All RFU values are background substrate with CBS6556  $\Delta HIS3 \Delta URA3$  harboring empty plasmid pIW578 as the background. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

$\text{h}^{-1}$ , respectively). The growth curve and growth rate of CBS6556  $\Delta\text{HIS3}$   $\Delta\text{URA3}$  harboring a blank plasmid was also determined in triplicate (Fig. S3). Similar growth rates in both conditions indicate that EGFP expression did not alter *K. marxianus* growth. Fig. 1C (bottom) shows the average fluorescence per  $\text{OD}_{600}$  for each promoter as cultures reached early stationary phase ( $t = 14$  h for  $30^\circ\text{C}$ ). Figs. S4 and S5 show equivalent plots for 37 and  $45^\circ\text{C}$  cultures. Changes in promoter strength from data observed at  $30^\circ\text{C}$  are shown in Fig. 1C (top). In the majority of cases, promoter expression decreased or stayed the same at higher temperatures in glucose media; however,  $P_{\text{ADH1}}$ ,  $P_{\text{INU1}}$ ,  $P_{\text{PRE1}}$ , and  $P_{\text{SSA3}}$  resulted in higher expression as temperature increased. At  $30^\circ\text{C}$ ,  $P_{\text{NC1}}$  exhibited the highest expression level (as judged by relative EGFP fluorescence intensity), while  $P_{\text{TEF3}}$  was second highest. At the low end, 13 promoters ( $P_{\text{ADH1}}$ ,  $P_{\text{INU1}}$ ,  $P_{\text{PRE1}}$ ,  $P_{\text{PIR1}}$ ,  $P_{\text{POL4}}$ ,  $P_{\text{HSP26}}$ ,  $P_{\text{SSA3}}$ ,  $P_{\text{ZWF}}$ ,  $P_{\text{SCL1}}$ ,  $P_{\text{ALD2}}$ ,  $P_{\text{PST1}}$ ,  $P_{\text{GLK1A}}$  and  $P_{\text{COX20}}$ ) resulted in expression levels no more than 10% of  $P_{\text{NC1}}$ . The remaining 10 promoters ( $P_{\text{SOD1}}$ ,  $P_{\text{GPD1}}$ ,  $P_{\text{GLK1B}}$ ,  $P_{\text{HSP60}}$ ,  $P_{\text{TDH3}}$ ,  $P_{\text{PGK}}$ ,  $P_{\text{HTB2}}$ ,  $P_{\text{HTB1}}$ ,  $P_{\text{HHF1}}$ , and  $P_{\text{HHF2}}$ ) showed expression levels between the high and low sets. The relative groupings were found to be consistent at  $37^\circ\text{C}$ , with the exception of  $P_{\text{ADH1}}$ , which moved from the low to the medium-expression group (Fig. S4). At  $45^\circ\text{C}$ , the relative groupings changed considerably.  $P_{\text{SSA3}}$ , which had low expression at 30 and  $37^\circ\text{C}$ , became one of the strongest promoters, eight of the thirteen weak promoters (i.e.  $P_{\text{ADH1}}$ ,  $P_{\text{INU1}}$ ,  $P_{\text{PRE1}}$ ,  $P_{\text{PIR1}}$ ,  $P_{\text{POL4}}$ ,  $P_{\text{ZWF}}$ ,  $P_{\text{ALD2}}$ , and  $P_{\text{COX20}}$ ) joined the medium set, and the medium promoter  $P_{\text{HTB2}}$  moved to the low-expression set (Fig. S5). As shown in Fig. 1C, not all promoters that have previously been reported to be upregulated at  $45^\circ\text{C}$  resulted in higher EGFP expression level. In our study, at least two key differences exist: (1) This study characterizes truncations of the native promoter, thus potentially altering function; and, (2) The production of EGFP, not the native gene associated with the promoter, removes potential post-transcriptional regulation that may be present during the transcriptional studies.

From the complete list of 25 native-derived promoters, we selected the following six as a defined promoter set representing a broad range of expression levels:  $P_{\text{NC1}}$  and  $P_{\text{TEF3}}$  were classified as strong promoters,  $P_{\text{HHF1}}$  and  $P_{\text{PGK}}$  were grouped as medium, and  $P_{\text{ADH1}}$  and  $P_{\text{SSA3}}$  were defined as weak. These groupings were based on expression levels measured at  $30^\circ\text{C}$ , but the same set of promoters also provides a broad

range at both 37 and  $45^\circ\text{C}$ . The dynamics of each of the six promoters of the defined set is shown in Fig. 2 for glucose cultures at 30, 37, and  $45^\circ\text{C}$ . Overall, it is apparent that expression increased through exponential phase, reaching the maximum at early stationary phase. In most cases, EGFP fluorescence plateaued through the stationary phase until the end of the measurement period. It is important to note that expression levels are reported with background fluorescence subtracted and expression levels resulting from  $P_{\text{ADH1}}$  and  $P_{\text{SSA3}}$  at  $30^\circ\text{C}$  are low but significantly above the background. Comparison within each group over different temperatures shows that expression from  $P_{\text{NC1}}$ ,  $P_{\text{TEF3}}$ ,  $P_{\text{HHF1}}$ , and  $P_{\text{PGK}}$  decreased with increasing temperature, while  $P_{\text{SSA3}}$  and  $P_{\text{ADH1}}$  had improved performance at higher temperature. We also note here that the expression of  $P_{\text{SSA3}}$  and  $P_{\text{ADH1}}$  appeared to decrease at longer times at elevated temperature. These two promoters represent genetic parts that could be further optimized for high-level expression at the upper limit of *K. marxianus* thermotolerance.

The data thus far shows that there is a strong effect of temperature on promoter activity. It is also well-known that promoter strength can vary with carbon source (Weinhandl et al., 2014; McTaggart et al., 2019; Rajkumar et al., 2019); thus, we evaluated the six promoters in our expression set using xylose as a carbon source. Fig. 3A presents growth curves for CBS6556  $\Delta\text{HIS3}$   $\Delta\text{URA3}$  harboring control plasmid pIW578 in synthetic defined media with 2% xylose at 30, 37, 41, and  $45^\circ\text{C}$ . The additional  $41^\circ\text{C}$  data was collected due to a dramatically slower growth rate at  $45^\circ\text{C}$  with respect to that observed when cells were grown in 2% glucose ( $0.3\text{ h}^{-1}$  with glucose at  $45^\circ\text{C}$ ,  $0.14\text{ h}^{-1}$  with xylose at  $45^\circ\text{C}$ ). The growth rates on xylose at 30, 37, and  $41^\circ\text{C}$  were significantly higher than at  $45^\circ\text{C}$  ( $0.28\text{ h}^{-1}$  at  $30^\circ\text{C}$ ,  $0.34\text{ h}^{-1}$  at  $37^\circ\text{C}$ , and  $0.35\text{ h}^{-1}$  at  $41^\circ\text{C}$ ), but overall the results indicate that *K. marxianus* CBS6556  $\Delta\text{HIS3}$   $\Delta\text{URA3}$  has slower growth on xylose than on glucose. In all subsequent xylose experiments, the highest temperature investigated was  $41^\circ\text{C}$ .

With respect to promoter function, we directly compared EGFP expression with  $P_{\text{TEF3}}$  at  $30^\circ\text{C}$  in 2% glucose and 2% xylose. Flow cytometry data suggests that for  $P_{\text{TEF3}}$ , carbon source has little to no effect on protein production (Fig. S6); as such we normalized expression in glucose and xylose to the  $P_{\text{TEF3}}$  expression level (Fig. 3B). Comparison of the rank order of the six promoters in both xylose and glucose suggests that this is also true for  $P_{\text{NC1}}$ ,  $P_{\text{HHF1}}$  and  $P_{\text{SSA3}}$ . The  $\text{NC1}$  promoter resulted

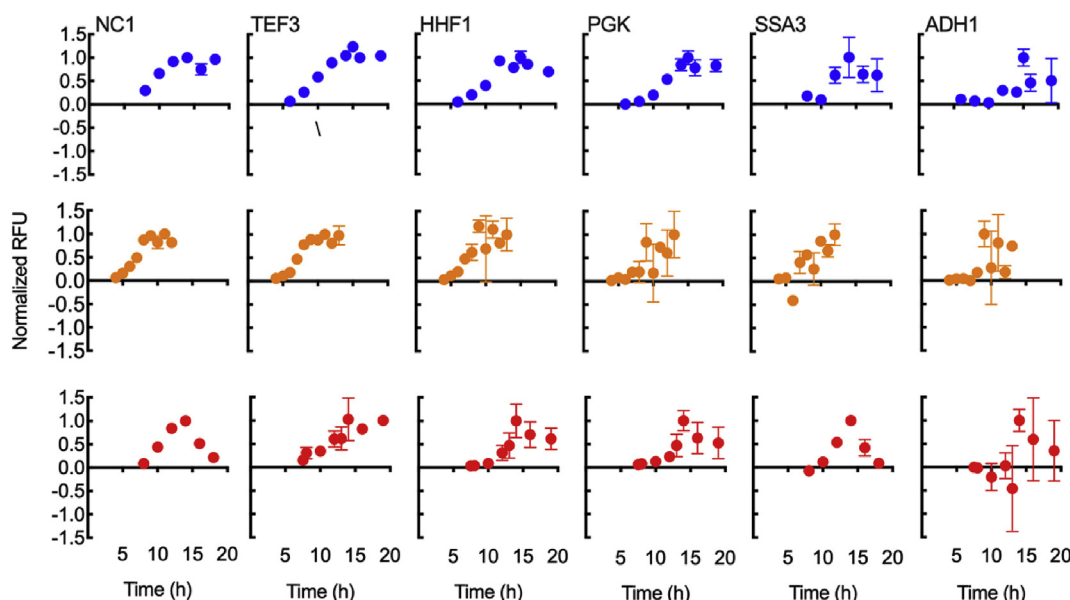
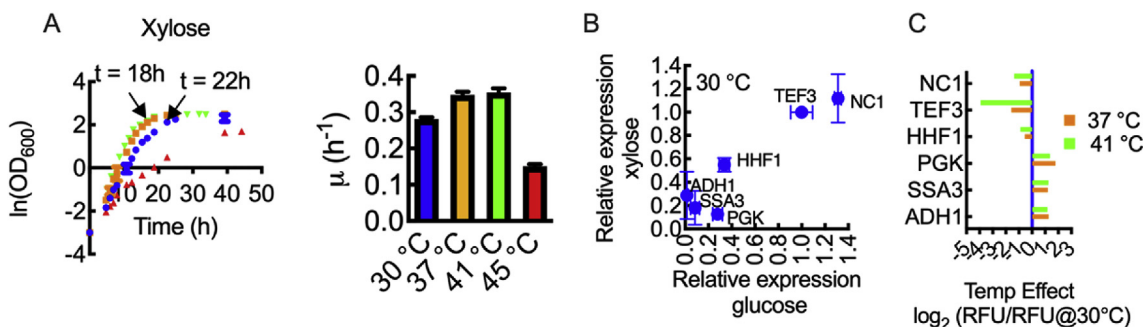


Fig. 2. *K. marxianus* promoter dynamics using glucose as the carbon source. Time course data were collected from inoculation to stationary phase for 6 different *K. marxianus* promoters. Shake flask cultures were grown at  $30^\circ\text{C}$  (blue),  $37^\circ\text{C}$  (orange), and  $45^\circ\text{C}$  (red) in synthetic defined media without histidine (SD-His) with 2% glucose. All data is normalized to the highest fluorescence signal intensity for each combination of promoter and temperature. The promoter strength upon reaching stationary phase is shown in Fig. 1. Data points represent the mean of biological triplicates, while the error bars represent the standard deviation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 3.** Characterization of six selected promoters at late exponential phase when grown on xylose. (A) Growth curves and growth rate of *K. marxianus* CBS6556  $\Delta$ HIS3  $\Delta$ URA3 harboring a blank expression vector (pIW578) in synthetic defined media with 2% xylose as a carbon source at 30, 37, 41 and 45 °C. Shake flask cultures with an initial OD<sub>600</sub> of 0.05 were inoculated with starter cultures grown on xylose. (B) Comparison of relative promoter strength with growth on xylose and glucose at 30 °C. EGFP fluorescence was normalized to  $P_{TEF3}$  for xylose and glucose separately. The absolute value of RFU/OD for each promoter is  $25 \pm 18$  (ADH1),  $47 \pm 7$  (HHF1),  $95 \pm 14$  (NC1),  $10 \pm 1$  (PGK),  $15 \pm 11$  (SSA3) and  $86 \pm 3$  (TEF3) at 30 °C. (C) The temperature effect on promoter strength with xylose as a carbon source. The effect of temperature (37 and 41 °C) is indicated as the log<sub>2</sub>-fold change from fluorescence observed at 30 °C. Time-points for late exponential phase in xylose are 22 h at 30 °C and 18 h at 37 °C and 41 °C. All data represent biological triplicates. Data points and bars indicate the mean and error bars indicate the standard deviation.

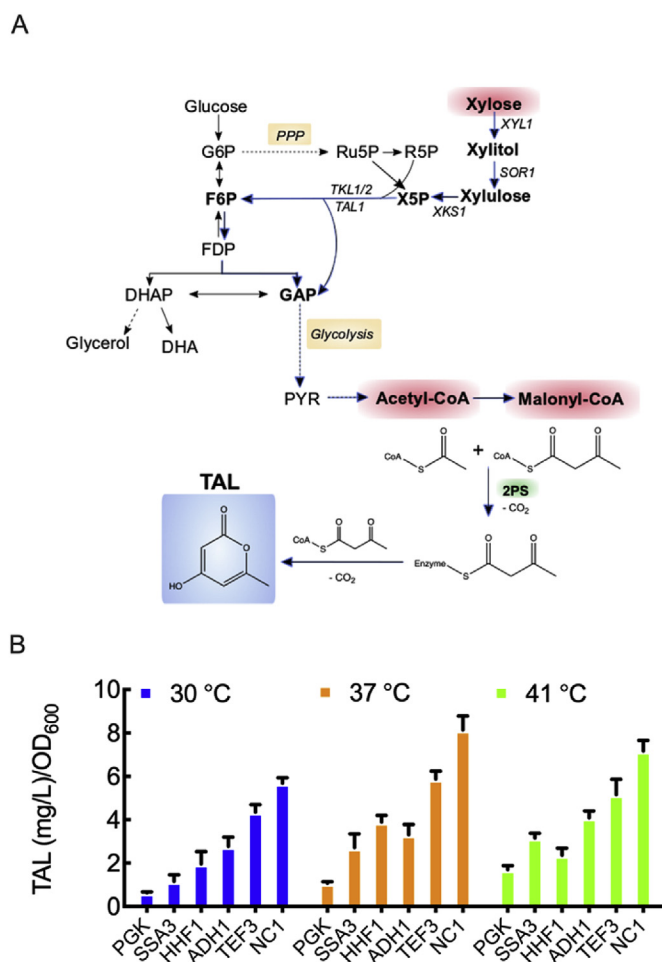
in the highest expression level, while  $P_{HHF1}$  and  $P_{SSA3}$  can be considered medium- and low-level promoters, respectively for both carbon sources at 30 °C. Variation in the rank order of promoters (and in absolute expression level) was observed in the lower range. In glucose,  $P_{PGK}$  was found to be a medium level promoter, reaching 28% of  $P_{TEF3}$ , but in xylose expression was reduced to less than 12% of  $P_{TEF3}$ . Growth in xylose had the opposite effect on  $P_{ADH1}$ , increasing expression to 28% of  $P_{TEF3}$ .

The temperature effect was also studied with 2% xylose and the corresponding data is shown in Fig. 3C. The results indicate that the three promoters with highest expression levels at 30 °C (i.e.,  $P_{NC1}$ ,  $P_{TEF3}$  and  $P_{HHF1}$ ) show reduced EGFP fluorescence at higher temperatures. In contrast, the lower range promoters (i.e.,  $P_{PGK}$ ,  $P_{ADH1}$  and  $P_{SSA3}$ ) showed between 2.7 and 3.5-fold increase in expression at 37 and 41 °C.

To demonstrate the utility of the defined promoter set, we engineered triacetic acid lactone (TAL) biosynthesis in *K. marxianus* (Fig. 4A). Our previous work demonstrated that *K. marxianus* is capable of achieving high TAL titers using low cost carbon substrates such as glycerol and xylose, without metabolic pathway engineering (McTaggart et al., 2019). TAL is a valuable platform chemical that can be derivatized into both high-value and commodity chemicals, including direct substitutes of current petrochemicals such as dienoic acid and sorbic acid, among others (Chia et al., 2012). Biological synthesis of TAL is catalyzed by the type III polyketide synthase 2-pyrone synthase (2-PS), natively found in the plant *Gerbera hybrida* (Eckermann et al., 1998). The choice of promoter controlling the synthesis of 2-PS has been shown to be an important parameter for optimal TAL production (Cardenas and Da Silva, 2014; McTaggart et al., 2019). We therefore decided to test the effect of promoter strength ( $P_{ADH1}$ ,  $P_{HHF1}$ ,  $P_{NC1}$ ,  $P_{PGK}$ ,  $P_{SSA3}$  and  $P_{TEF3}$ ) on TAL specific titers at different temperatures using xylose as the carbon source. The six promoters were placed upstream of the *G. hybrida* 2-PS gene in a low copy number plasmid and transformed into CBS6556  $\Delta$ HIS3  $\Delta$ URA3. We kept our expression system consistent to allow comparisons with the fluorescence-based measurements and for consistent plasmid copy number. The resulting strains were grown in synthetic defined medium with 2% xylose at 30, 37, and 41 °C, and TAL levels were measured at late exponential phase and late stationary phase (Fig. 4B and S7, respectively).

The six promoters tested resulted in a wide range of TAL specific titers, covering a 17.8-fold change between the highest and the lowest measured across all temperatures and promoters. Most promoters showed higher levels of TAL as temperature increased from 30 to 37 °C, except for  $P_{ADH1}$ , which did not show statistically significant changes with temperature. When temperature was increased from 37 to 41 °C, none of the six promoters resulted in a further increase in TAL production (Fig. 4B and S7). These general trends are in accordance to those reported

in our earlier work (McTaggart et al., 2019) when expressing 2-PS under the control of the full-length  $P_{PGK}$ ; higher TAL specific titers were achieved as temperature increased from 30 to 37 °C, but remained constant



**Fig. 4.** Triacetic acid lactone (TAL; 4-hydroxy-6-methyl-2-pyrone) biosynthesis in xylose medium with varying levels of 2-pyrone synthase (2-PS) expression. (A) Simplified illustration of proposed TAL biosynthesis pathway in *K. marxianus*. (B) Specific TAL production at late exponential phase at 30, 37 and 41 °C. 2-PS was overexpressed with  $PGK$ ,  $SSA3$ ,  $HHF1$ ,  $ADH1$ ,  $TEF3$ , and  $NC1$  promoters on a low copy number plasmid. All data represent biological triplicates and the error bars indicate the standard deviation.

when temperature was further increased.

For the three temperatures tested, 2-PS expression controlled by  $P_{PGK}$  and  $P_{NCI}$  resulted in the lowest and the highest TAL specific titers measured, respectively. Therefore, if TAL titers are taken as a proxy for promoter strength,  $P_{PGK}$  would be the weakest and  $P_{NCI}$  the strongest promoter. At 30 °C,  $P_{SSA3}$  showed comparable levels to  $P_{PGK}$ , but at higher temperatures it behaved as a medium-strength promoter.  $P_{ADH1}$  and  $P_{HHF1}$  performed as medium-strength promoters across temperatures and  $P_{TEF3}$  acted as a strong promoter at 37 and 41 °C.

In our previous work, the  $ADH2$  promoter from *S. cerevisiae* ( $P_{ScADH2}$ ) outperformed the native *K. marxianus*  $PGK1$  promoter when placed upstream of 2-PS (McTaggart et al., 2019). We thus decided to compare TAL production levels between our strongest promoter  $P_{NCI}$  and  $P_{ScADH2}$  using our current expression system (i.e., on a CEN/ARS plasmid in strain CBS 6556  $\Delta HIS3 \Delta URA3$ ) at 37 °C. No statistically significant difference in titers or specific titers were observed between the two promoters (Fig. S8) at late exponential or stationary phase; therefore, the *K. marxianus*  $NCI$  promoter was comparable to the strong *S. cerevisiae*  $ADH2$  promoter for TAL synthesis. TAL titers ranged from 58 mg/L to 64 mg/L at late exponential phase and from 94 to 114 mg/L in stationary phase. These low titers are expected when using a low-copy plasmid in strain CBS 6556  $\Delta HIS3 \Delta URA3$ .

McTaggart et al. reported their highest TAL titer at 48 h when 2-PS was expressed from a multi-copy plasmid under the control of the  $P_{ScADH2}$  in the industrial strain KM1  $\Delta URA3$ . To further compare the performance of  $P_{NCI}$  and  $P_{ScADH2}$ , we replaced  $P_{ScADH2}$  in pKD-A2PS with  $P_{NCI}$  to create pKD-N2PS. We then transformed both plasmids into KM1  $\Delta URA3$  and measured TAL production at 37 °C in 1% SXCA at 48 h (Fig. S9). Titers for KM1  $\Delta URA3$  pKD-A2PS and KM1  $\Delta URA3$  pKD-N2PS were 1.2 g/L and 0.82 g/L, respectively, the former being consistent with our previously reported value (McTaggart et al., 2019). While the use of  $P_{NCI}$  resulted in a 34% decrease in titer, there was no statistically significant difference between specific titers. Interestingly, when the same strains were tested using lactose as a carbon source, the use of  $P_{NCI}$  resulted in a 58% increase in titer and an 80% increase in specific titer (Fig. S9). Taken together, these results indicate that the use of  $P_{NCI}$  provides equivalent or higher TAL specific titers than  $P_{ScADH2}$ .

#### 4. Discussion

In this work, we set out to design and characterize a set of promoters for controlling gene expression in the thermotolerant yeast *K. marxianus*. From a set of 25 genes, we created a plasmid library for heterologous expression with native-derived promoters derived from *K. marxianus* sequences. These native-derived promoters were created from the 700 bp upstream of the start codon of each gene (Fig. 1A). Heterologous EGFP expression with glucose as the carbon source helped define a subset of six promoters with a broad transcriptional range (Fig. 2). We then chose a defined promoter set that includes  $P_{NCI}$ ,  $P_{TEF3}$ ,  $P_{HHF1}$ ,  $P_{PGK}$ ,  $P_{SSA3}$  and  $P_{ADH1}$  and is representative of the full 25-member library including promoters with high, medium, and low expression levels. These six promoters also maintained substantial transcriptional differences under xylose metabolism (Fig. 3B). We note that the  $ADH1$  promoter reported by Yang et al. (2015) showed a different temperature dependence to the  $ADH1$  promoter in our study. The difference was likely due to: (1) differences in  $ADH1$  promoter sequence (one sequence was from CBS 6556, while the other was from NBRC 1777); (2) differences in the studied strain (CBS 6556 vs. NBRC1777); and, (3) in our study we use EGFP as a proxy for transcriptional strength the function of which is invariant with temperature (Fig. S1), whereas Yang et al. used  $\beta$ -glucuronidase activity as a reporter, but the relationship between enzyme function and temperature was not described.

Our studies on TAL production from xylose confirmed the broad transcriptional range and utility of the library. Specifically, relative promoter strengths for 2-PS expression, a necessary heterologous reaction step in TAL biosynthesis, matched the relative EGFP expression from

the same promoter set at 30 °C (Fig. S10A). The consistency between the rank order of promoter strength judged by TAL production and separately by EGFP fluorescence was lost at 37 and 41 °C (Fig. S10). One likely reason is that TAL biosynthesis does not rely solely on the expression of 2-PS, but also on the expression of upstream enzymes that may or may not change at high temperature. This hypothesis is supported by transcriptional studies in glucose, which have shown the  $ACC1$  gene is up-regulated at higher temperatures while  $CIT1$  is down-regulated (Fu et al., 2019; Lertwattanasakul et al., 2015). Thus acetyl-CoA and malonyl-CoA pools also are likely affected by temperature and TAL titers may not necessarily correlate exclusively with levels of 2-PS protein.

A valuable trait of *K. marxianus* is its thermotolerance, the benefit of which is observable for TAL biosynthesis; specific TAL titers at 37 and 41 °C were significantly greater than those achieved at 30 °C (Fig. 4). However, the production benefits realized at elevated temperature must be balanced with increased cellular stress. High growth rates on xylose were possible at 30, 37 and 41 °C, but growth rate was reduced by 2 to 3-fold at 45 °C. Different behavior was observed when glucose was used as the carbon source as high growth rate was maintained up to 45 °C. While there have been a number of recent studies focusing on the effects of temperature on *K. marxianus* transcription (Fu et al., 2019; Lertwattanasakul et al., 2015), the global effects on metabolism and promoter function have yet to be fully understood.

#### 5. Conclusions

The results presented here help advance *K. marxianus* as a thermotolerant host for chemical biosynthesis. The primary result is the design and validation of a new set of promoters that can be used to vary gene expression by upward of 87-fold under glucose metabolism and greater than 17.8-fold with xylose as the carbon source. The defined set includes  $P_{NCI}$  and  $P_{TEF3}$  as strong promoters,  $P_{HHF1}$  and  $P_{PGK}$  as medium level promoters, and  $P_{SSA3}$  and  $P_{ADH1}$  as low-expression promoters. The temperature effect on gene expression from a subset of six promoters was also characterized, and showed that, in general, fluorescence was reduced as temperature increased from 30 °C to 37, 41, and 45 °C. Two promoters,  $P_{SSA3}$  and  $P_{ADH1}$ , were exceptions to the trend and were found to have a positive correlation with temperature in both glucose and xylose. We demonstrated the utility of the promoter set by expressing 2-PS for TAL biosynthesis from xylose and showed increased TAL specific titers at 37 and 41 °C. Finally, when the strongest promoter  $P_{NCI}$  was tested in the industrial strain KM1  $\Delta URA3$  using a multi-copy plasmid in xylose and lactose, specific TAL titers were equivalent or higher than those using the previously reported heterologous  $P_{ScADH2}$ .

#### CRedit authorship contribution statement

**Xuye Lang:** Conceptualization, Methodology, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Pamela B. Besada-Lombana:** Methodology, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Mengwan Li:** Methodology, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Nancy A. Da Silva:** Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Ian Wheeldon:** Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

#### Acknowledgements

This study was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019093 to IW and NDS, and



National Science Foundation award NSF-CBET 1706545 (IW) and 1803677 (NDS).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mec.2020.e00145>.

## References

- Andreu, C., Del Olmo, M., 2020. Whole-cell biocatalysis in seawater: new halotolerant yeast strains for the regio- and stereoselectivity reduction of 1-phenylpropane-1,2-dione in saline-rich media. *Chembiochem* 21 (11), 1621–1628.
- Benatuil, L., Perez, J.M., Belk, J., Hsieh, C.M., 2010. An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng. Des. Sel.* 23 (4), 155–159.
- Cardenas, J., Da Silva, N.A., 2014. Metabolic engineering of *Saccharomyces cerevisiae* for the production of triacetic acid lactone. *Metab. Eng.* 25, 194–203.
- Carlson, R., 2016. Estimating the biotech sector's contribution to the US economy. *Nat. Biotechnol.* 34, 247–255.
- Cernak, P., Estrela, R., Poddar, S., Skerker, J.M., Cheng, Y.F., Carlson, A.K., Chen, B., Glynn, V.M., Furlan, M., Ryan, O.W., Donnelly, M.K., Arkin, A.P., Taylor, J.W., Cate, J.H.D., 2018. Engineering *Kluyveromyces marxianus* as a robust synthetic biology platform host. *Applied and Environment Science* 9 (5), 1–16.
- Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic acid lactone as a potential biorenewable platform chemical. *Green Chem.* 14, 1850–1853.
- Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., Teeri, T., Schröder, J., 1998. New pathway to polyketides in plants. *Nature* 396, 387–390.
- Fu, X., Li, P., Zhang, L., Li, S., 2019. Understanding the stress responses of *Kluyveromyces marxianus* after an arrest during high-temperature ethanol fermentation based on integration of RNA-Seq and metabolite data. *Appl. Microbiol. Biotechnol.* 103, 2715–2729.
- Kazmierczak, M., Ritterson, R., Gardner, D., Casagrande, R., Hanemann, T., Rosen, D.H., 2019. China's Biotechnology Development- the Role of US and Other Foreign Engagement. A report prepared for the U.S.-China Economic and Security Review Commission, p. 150.
- Kushi, R.T., Monti, R., Contiero, J., 2000. Production, purification and characterization of an extracellular inulinase from *Kluyveromyces marxianus* var. *bulgaricus*. *J. Ind. Microbiol. Biotechnol.* 25, 63–69.
- Lane, M.M., Morrissey, J.P., 2010. *Kluyveromyces marxianus*: a yeast emerging from its sister's shadow. *Fungal Biology Reviews* 24, 17–26.
- Lee, K.S., Kim, J.S., Heo, P., Yang, T.J., Sung, Y.J., Cheon, Y., Koo, H.M., Yu, B.J., Seo, J.H., Jin, Y.S., Park, J.C., Kweon, D.H., 2013. Characterization of *Saccharomyces cerevisiae* promoters for heterologous gene expression in *Kluyveromyces marxianus*. *Appl. Microbiol. Biotechnol.* 97, 2029–2041.
- Lee, M.E., DeLoache, W.C., Cervantes, B., Dueber, J.E., 2015. A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol.* 4, 975–986.
- Lertwattanasakul, N., Rodrussamee, N., Suprayogi, Limtong, S., Thanonkeo, P., Kosaka, T., Yamada, M., 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).
- Lertwattanasakul, N., Kosaka, T., Hosoyama, A., Suzuki, Y., Rodrussamee, N., Matsutani, M., Murata, M., Fujimoto, N., Suprayogi, n., Tsuchikane, K., Limtong, S., Fujita, N., Yamada, M., 2015. Genetic basis of the highly efficient yeast *Kluyveromyces marxianus*: complete genome sequence and transcriptome analyses. *Biotechnol. Biofuels* 8, 47.
- Löbs, A.K., Engel, R., Schwartz, C., Flores, A., Wheeldon, I., 2017. CRISPR-Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in *Kluyveromyces marxianus*. *Biotechnol. Biofuels* 10, 164.
- Löbs, A.K., Lin, J.L., Cook, M., Wheeldon, I., 2016. High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from *Kluyveromyces marxianus* on C5, C6, and C12 carbon sources. *Biotechnol. J.* 11, 1274.
- Löbs, A.K., Schwartz, C., Wheeldon, I., 2017. Genome and metabolic engineering in non-conventional yeasts: current advances and applications. *Synth. Syst. Biotechnol.* 2, 198.
- Löbs, A.K., Schwartz, C., Thorwall, S., Wheeldon, I., 2018. Highly multiplexed CRISPRi repression of respiratory functions enhances mitochondrial localized ethyl acetate biosynthesis in *Kluyveromyces marxianus*. *ACS Synth. Biol.* 7, 2647–2655.
- Madeira, J.V., Gombert, A.K., 2018. Towards high-temperature fuel ethanol production using *Kluyveromyces marxianus*: on the search for plug-in strains for the Brazilian sugarcane-based biorefinery. *Biomass Bioenergy* 119, 217–228.
- McTaggart, T.L., Bever, D., Bassett, S., Da Silva, N.A., 2019. Synthesis of polyketides from low cost substrates by the thermotolerant yeast *Kluyveromyces marxianus*. *Biotechnol. Bioeng.* 116, 1721–1730.
- Nonklang, S., Abdel-Banat, B.M., Cha-aim, K., Moonjai, N., Hoshida, H., Limtong, S., Yamada, M., Akada, R., 2008. High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. *Appl. Environ. Microbiol.* 74, 7514–7521.
- Raimondi, S., Zanni, E., Amaretti, A., Palleschi, C., Uccelletti, D., Rossi, M., 2013. Thermal adaptability of *Kluyveromyces marxianus* in recombinant protein production. *Microb. Cell Factories* 12.
- Rajkumar, A.S., Varela, J.A., Juergens, H., Daran, J.G., Morrissey, J.P., 2019. Biological parts for *Kluyveromyces marxianus* synthetic biology. *Front. Bioeng. Biotechnol.* 7, 97.
- Rodrussamee, N., Lertwattanasakul, N., Hirata, K., Suprayogi Limtong, S., Kosaka, T., Yamada, M., 2011. Growth and ethanol fermentation ability on hexose and pentose sugars and glucose effect under various conditions in thermotolerant yeast *Kluyveromyces marxianus*. *Appl. Microbiol. Biotechnol.* 90, 1573–1586.
- Thorwall, S., Schwartz, C., Chartron, J.W., Wheeldon, I., 2020. Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. *Nat. Chem. Biol.* 16, 113–121.
- Weinhandl, K., Winkler, M., Glieder, A., Camattari, A., 2014. Carbon source dependent promoters in yeasts. *Microb. Cell Factories* 13 (5), 1–17.
- Yang, C., Hu, S., Zhou, S., Wang, D., Gao, X., Hong, J., 2015. Characterizing yeast promoters used in *Kluyveromyces marxianus*. *World J. Microbiol. Biotechnol.* 31, 1641–1646.
- Yarimizu, T., Nakamura, M., Hoshida, H., Akada, R., 2015. Synthetic signal sequences that enable efficient secretory protein production in the yeast *Kluyveromyces marxianus*. *Microb. Cell Factories* 14, 20.
- Yuan, W.J., Zhao, X.Q., Ge, X.M., Bai, F.W., 2008. Ethanol fermentation with *Kluyveromyces marxianus* from Jerusalem artichoke grown in salina and irrigated with a mixture of seawater and freshwater. *J. Appl. Microbiol.* 105, 2076–2083.