RNA polymerase II-driven CRISPR-Cas9 system for efficient non-growth-biased metabolic engineering of *Kluyveromyces marxianus*

Danielle Bever\(^a\), Ian Wheeldon\(^b\), Nancy Da Silva\(^a\*,\)

\(^a\) Department of Chemical and Biomolecular Engineering, University of California, Irvine, CA, 92697, USA
\(^b\) Department of Chemical and Environmental Engineering, University of California, Riverside, CA, 92521, USA

**Article Info**

**Keywords:**
- *Kluyveromyces marxianus*
- CRISPR-Cas9
- Polyketides
- Metabolic engineering
- Thermotolerant yeast

**Abstract**

The thermotolerant yeast *Kluyveromyces marxianus* has gained significant attention in recent years as a promising microbial candidate for industrial biomanufacturing. Despite several contributions to the expanding molecular toolbox for gene expression and metabolic engineering of *K. marxianus*, there remains a need for a more efficient and versatile genome editing platform. To address this, we developed a CRISPR-based editing system that enables high efficiency marker-less gene disruptions and integrations using only 40 bp homology arms in NHEJ functional and non-functional *K. marxianus* strains. The use of a strong RNA polymerase II promoter allows efficient expression of gRNAs flanked by the self-cleaving RNA structures, tRNA and HDV ribozyme, from a single plasmid co-expressing a codon optimized Cas9. Implementing this system resulted in nearly 100% efficiency of gene disruptions in both NHEJ-functional and NHEJ-deficient *K. marxianus* strains, with donor integration efficiencies reaching 50% and 100% in the two strains, respectively. The high gRNA targeting performance also proved instrumental for selection of engineered strains with lower growth rate but improved polyketide biosynthesis by avoiding an extended outgrowth period, a common method used to enrich for edited cells but that fails to recover advantageous mutants with even slightly impaired fitness. Finally, we provide the first demonstration of simultaneous, markerless integrations at multiple loci in *K. marxianus* using a 2.6 kb and a 7.6 kb donor, achieving a dual integration efficiency of 25.5% in a NHEJ-deficient strain. These results highlight both the ease of use and general robustness of this system for rapid and flexible metabolic engineering in this non-conventional yeast.

1. Introduction

The recent advent of sophisticated sequencing and CRISPR-Cas9 technologies has enabled the engineering of a wide range of non-conventional microorganisms with diverse metabolisms and growth characteristics. *Kluyveromyces marxianus* is a non-conventional yeast that has quickly emerged as a robust heterologous host with excellent characteristics for industrial applications. *K. marxianus* is the fastest growing eukaryote with a doubling time as low as 45 min (Groeneveld et al., 2009). This species grows on a variety of carbon sources including pentose sugars and disaccharides (McTaggart et al., 2019; Nonklang et al., 2008; Unit et al., 2012), is thermotolerant up to 52 °C (Banat et al., 1992) and acid-tolerant to pH 2.3 (Amrange and Prigent, 1998), and has Generally Regarded as Safe (GRAS) and Qualified Presumption of Safety (QPS) status in the United States and European Union, respectively (Lane and Morrissey, 2010). Classified as a Crabtree-negative yeast, *K. marxianus* has also been shown to minimally accumulate common metabolic byproducts such as ethanol, glycerol, and organic acids (Fonseca et al., 2007, 2013), supporting enhanced product yield and separation. These numerous advantageous characteristics have driven the study and successful use of *K. marxianus* for the high-level production of a variety of native and non-native products including ethanol (Charoensopharat et al., 2015; Leonel et al., 2021), 2-phenylethanol (Hillman et al., 2021; Martinez et al., 2018), D-allulose (Yang et al., 2018), and triacetic acid lactone (Lang et al., 2020; McTaggart et al., 2019) from substrates other than purified glucose.

Recent contributions to the toolbox of genetic parts for tunable gene expression in *K. marxianus* include the characterization of native and non-native promoters, terminators, selection markers, and low- and high-copy origins of replication (Kumar et al., 2021; Lang et al., 2020; McTaggart et al., 2019; Rajkumar et al., 2019; Yang et al., 2015). In addition, CRISPR-Cas9 systems have been published by five groups and...
successfully applied for the synthesis of several compounds, including pectin (Cernak et al., 2018), 2-phenylethanol (Li et al., 2021; Rajkumar and Morrissey, 2020), ethyl acetate (Löbs et al., 2017, 2018), human N-glycan structures (Lee et al., 2020), and a single-chain antibody (Nambu-Nishida et al., 2018). A detailed summary of the current K. marxianus CRISPR systems can be found in Table S1. Despite the proven utility of these systems, suboptimal features include donors with long homologies (several hundreds of base pairs), donors with selection markers, a growth-biasing selection process, multiple gRNAs to target the same loci, or a high DNA input. A platform with higher efficiency and greater ease of use is thus needed for more robust metabolic engineering of K. marxianus. In the design of our system, we aimed for (1) high rates of gene knockout (KO) and gene integration using short 40 base pair (bp) homology arms and modest DNA input during transformation, (2) marker-less integrations, (3) elimination of growth bias by avoiding an extended outgrowth period, (4) reasonably efficient integrations even in a non-homologous end joining (NHEJ)-deficient strain, and (5) simultaneous gene integrations at more than one locus.

Most established CRISPR-Cas9 systems in K. marxianus use RNA polymerase (RNAP) III promoters for gRNA expression (Cernak et al., 2018; Lee et al., 2018, 2020; Löbs et al., 2017; Nambu-Nishida et al., 2017) to obtain RNA transcripts with minimal end modifications (Marko et al., 2006), a favorable quality to ensure high active expression of gRNA. However, a catalog of well-characterized RNAP III promoters does not exist (Gao and Zhao, 2014; Ma et al., 2014) and the rate of RNA expression from such promoters is typically low (Ryan et al., 2014; Turowski and Tollervey, 2016). With weak gRNA expression, many generations are generally required to accumulate sufficient gRNA for efficient complexing with Cas9 and for subsequent genome editing (Ng and Dean, 2017). RNAP II promoters are better characterized and tend to offer higher expression rates (Gao and Zhao, 2014; Ng and Dean, 2017); however, the 5’ cap and 3’ polyA tail end modifications produced by these promoters can hinder gRNA function (Gao and Zhao, 2014; Ng and Dean, 2017). Studies in various yeasts have limited the effects of these end modifications by flanking the gRNA with cleavable RNA structures, such as ribozymes and tRNAs, that remove the transcript end modifications after self-cleavage or RNase activity, respectively, leaving behind the desired, mature gRNA (Gao and Zhao, 2014; Gorter de Vries et al., 2017; Ng and Dean, 2017). It has also been suggested that the addition of these transient structures at the ends of the pre-gRNA may provide protection from degradation by exonucleases (Ng and Dean, 2017; Ryan and Cate, 2014). The hammerhead (HH) and hepatitis delta virus (HDV) ribozymes are common choices for these constructs as self-cleavage does not leave behind extraneous nucleotides (Ke et al., 2007; Pley et al., 1994). However, achieving this clean self-cleavage for the HH ribozyme requires that its 5’ end matches the first six nucleotides of the gRNA targeting sequence, complicating cloning processes (Gao and Zhao, 2014; Juergens et al., 2018). It has also been suggested that self-cleaving activity of the HH ribozyme may be a limiting factor to gRNA function. In contrast, utilization of the HDV ribozyme and tRNAs may offer higher efficiency maturation of gRNAs (Ng and Dean, 2017).

To achieve our goals, we developed a streamlined and highly efficient single plasmid CRISPR system that expresses codon-optimized Cas9 (Löbs et al., 2017) and a gRNA cassette comprised of the strong S. cerevisiae TDH3 RNAP II promoter to express a tRNA-gRNA-HDV ribozyme gene (TGR). We first evaluated the efficiency of the system for gene KO and gene integration (up to a 7.6 kb insert) in wild-type and non-homologous end joining (NHEJ)-deficient K. marxianus strains. We show that high engineering efficiencies can be achieved by this RNAP II TGR system without the use of growth-biasing outgrowth methods. As a demonstration of the utility of this system for metabolic engineering, we knocked out GPD1 and ZWF1, fitness-improving modifications, and integrated the large ACC1 gene in K. marxianus strain CBS 712 to elevate biosynthesis of the polyketide triacetic acid lactone. Finally, leveraging the high efficiency and multiplexing-compatible gRNA expression cassette design, we demonstrated, for the first time, simultaneous, markerless integrations at more than one loci in K. marxianus.

Table 1
List of strains and plasmids.

| Strain          | Features                                                                 | Source                     |
|-----------------|--------------------------------------------------------------------------|----------------------------|
| CBS 712ΔU       | ΔURA3                                                                   | McTaggart et al. (2019)    |
| CBS 712ΔAΔK     | ΔURA3ΔKU70                                                               | This work                  |
| CBS 712ΔU-ACC1  | HIS3-ACC1                                                               | This work                  |
| CBS 712ΔAΔK-ACC1| ΔURA3ΔKU70 HIS3-ACC1                                                    | This work                  |
| CBS 712ΔAΔK-ZWF1| ΔURA3ΔKU70ΔZWF1                                                         | This work                  |
| CBS 712ΔAΔK-GPD1| ΔURA3ΔKU70ΔGPD1                                                         | This work                  |

| Plasmid Name    | Features                                                                 | Source                     |
|-----------------|--------------------------------------------------------------------------|----------------------------|
| pKU70-UBlaster   | 5’ KU70Δura (432 bp) 3’-3’ KU70Δura (440 bp)                              | McTaggart (2020)           |
| pKU70-UBV2      | 5’ KU70ura (763 bp) 3’-3’ KU70ura (777 bp)                                | This work                  |
| pW601           | RPRPmRNA-tracrRNA-SUP4; TEFPmRNA-Cas9kappa; CEN/ARS                      | Lee et al. (2015)          |
| pYTK009         | TDH3pU, part plasmid                                                     | This work                  |
| pYTK050         | sgRNA dropout part plasmid                                               | This work                  |
| pDBtrg-Cas9     | TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pDBtrg-HIS3     | TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pDBtrg-LEU2     | TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| RNAP III TGR-HIS| pW601 with HIS3 gRNA and modified tracrRNA                               | This work                  |
| RNAP III TGR    | RPRPmRNA-tracrRNA-HDV-SUP4; TEFPmRNA-Cas9kappa; CEN/ARS                   | This work                  |
| RNAP III TGR-HIS3| RPRPmRNA-tracrRNA-HDV-SUP4; TEFPmRNA-Cas9kappa; CEN/ARS                   | This work                  |
| pDBtrg-Cas9     | TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pDBtrg-HIS3-LEU2| TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pDBtrg-Cas9-ZWF1| TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pDBtrg-GPD1     | TDH3pU, gRNAtracrRNA-HDV-CYC1; TEFPmRNA-Cas9kappa; CEN/ARS               | This work                  |
| pCA-A           | ADH2pU, CEN/ARS                                                         | McTaggart et al. (2019)    |
| pCA-A-ACC1      | ADH2pU, ACC1kappa, CEN/ARS                                               | This work                  |
| pKD-P2PS        | pgkpKappa, 2PSHT-CYC1; pKD4                                              | McTaggart et al. (2019)    |
2. Materials and methods

2.1. Strains and media

Escherichia coli strain DH5α (Invitrogen, Carlsbad, CA) was used for plasmid maintenance and amplification. E. coli cultures were conducted in Luria-Bertani (LB) media supplemented with 150 mg/L ampicillin. K. marxianus strain CBS 712 (ATCC, 200963; ATCC®, Manassas, VA) was previously modified for uracil auxotrophy by McTaggart et al. (2019) resulting in base strain CBS 712ΔU. K. marxianus strains were cultivated in complex YPD medium (10 g/L Bacto yeast extract, 20 g/L Bacto peptone, and 20 g/L dextrose) or selective SC medium (1.7 g/L Bacto yeast nitrogen base without ammonium sulfate, 5 g/L ammonium sulfate, 5 g/L casamino acids, and 50 mg/L adenine) supplemented with various carbon sources: SDCA (20 g/L dextrose), SXCA (10 g/L xylose), or SLCA (9.5 g/L lactose). SXCA and SLCA have equimolar concentrations of carbon. All strains utilized in this study can be found in Table 1.

2.2. CRISPR plasmid construction and donor DNA synthesis

All plasmids and primers used in this work are listed in Tables 1 and S2, respectively. Primers were purchased from IDT DNA (San Diego, CA) as single stranded DNA oligonucleotides. All PCR reactions for cloning and donor DNA synthesis used Q5 High-Fidelity DNA Polymerase (New England BioLabs, Ipswich, MA). Restriction enzymes and NEBuilder HiFi DNA Assembly Master Mix were purchased from New England BioLabs. Sanger sequencing services by Genewiz (Acenta Life Sciences; La Jolla, CA) were utilized to confirm all DNA sequences amplified by PCR for cloning.

To create the RNA II-driven gRNA expression cassette, five parts were used: (i) the S. cerevisiae TDH3 promoter isolated from pTYK009 (Lee et al., 2015) using primers pWl601-5’ScTDH3p and 5’KmtRNAgly-3’ScTDH3p; (ii) the K. marxianus tRNAgly isolated from pWl601 (Löbs et al., 2017) using primers 5’KmtRNAgly and 3’KmtRNAgly; (iii) the scaffold tracrRNA isolated from pTYK050 (Lee et al., 2015) using primers 3’KmtRNAgly-Nhel-5’’ and 5’HDV-3’’; (iv) 60 of the 68 base pairs constituting the hepatitis delta virus (HDV) ribozyme sequence (Gao and Zhao, 2014; Gorter de Vries et al., 2017) ordered as a single stranded oligo, HDV oligo; and (v) the S. cerevisiae CYC1 terminator isolated from pCA-A (McTaggart et al., 2019) using primers 3’HDV-AvrII-5’ScCYC1t and 3’ScCYC1t-loxp. A series of splicing by overlap extension PCR combinations these five parts into a single fragment. The RNAP II-driven gRNA expression cassette was then cloned into pWl601 digested with AvrII and SacII using NEBuilder HiFi DNA Assembly, replacing the original RNAP III-driven gRNA expression cassette, and completing construction of plasmid pDBtgr-Cas9.

gRNA targeting sequences were selected based on high On-Target Score (Doench et al., 2016) determined by the Benchling (2021) CRISPR tool, and were analyzed by BLAST (Altschul et al., 1990) to ensure a lack of sequence similarity between the last 12 nucleotides of the targeting sequence and the PAM sequence with the rest of the respective target site. Each pair is comprised of 40 base pairs homologous to a region upstream of the respective target site followed by up to 20 base pairs complementary to the 3’ end of the reverse oligo. The reverse oligo is comprised only of 40 base pairs homologous to a region downstream of the respective target site. Donors used for integration of the PGKlpKm-2PS-CYCItc cassette at the HIS3 and LEU2 loci were amplified from pKD-P2PS (McTaggart et al., 2019) using primers 2PS int KmHIS3 F and 2PS int KmLEU2 F or 2PS int KmHIS3 R and KmPGKlp int LEU2 F and CYCIt int LEU2 R, respectively. To synthesize donor DNA comprised of the K. marxianus ACC1 coding sequence under the control of a characterized promoter and terminator pair, the ACC1 coding sequence was first amplified from CBS 712 genomic DNA using primers pCV842 ACC1 F and pCV842 ACC1 R. NEBuilder HiFi DNA Assembly was used to insert the amplified ACC1 sequence into pCA-A (McTaggart et al., 2019) digested with Spel and Xhol, yielding pCA-A-ACC1. The ADH2psA-ACC1-KmCYCItc cassette was then amplified for integration into the HIS3 and LEU2 loci using primers ScADH2p int HIS3 F and 2PS int KmHIS3 R or ScADH2p int LEU2 F and CYCIt int LEU2 R, respectively. A small sample of each donor DNA PCR reaction was verified for synthesis of the correctly sized product by gel electrophoresis, after which the remaining volume of the PCR reaction was purified using the DNA Clean & Concentrator Kit (Zymo Research, Irvine, CA).

2.3. Transformation of K. marxianus for genomic modification

K. marxianus strains were streaked from frozen glycerol stocks onto YPD and incubated at 30 °C for 48–72 h. Single colonies were then inoculated into 3 mL YPD, grown overnight at 30 °C and 250 rpm in an air incubator shaker (New Brunswick Scientific Co. Excella E25, Edison, NJ), reincubated to OD600 0.5 (Shimadzu UV-2450 UV–Vis Spectrophotometer, Columbia, MD) in fresh 3 mL YPD media per transformation, and grown for approximately 4 h at 30 °C until OD600 2.0 was reached. The Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA) was utilized for all plasmid transformations with the following minor modifications. The volumes of Frozen-EZ Solution I, II, and III used per transformation was 250, 50, and 250 μL, respectively. For transformations involving CRISPR-related plasmids, after addition of Frozen-EZ Solution II, 11 μL of 10 mg/mL sheared salmon sperm DNA (Invitrogen, Waltham, MA) – previously boiled at 95 °C for 5 min and
immediately cooled on ice – was added to each reaction. Where applicable, donor DNA was added prior to plasmid DNA to the amount of 650 ng for KO donors or 1 μg for all others. Plasmid DNA was always added to the amount of 500 ng. After addition of Frozen-EZ Solution III, transformation reactions were incubated at 30 °C for 2 h before plating onto SDCA medium. Transformations involving CRISPR-related plasmids required about 36 h before colonies were apparent on plates, about 12 h longer than other plasmid transformations. When an outgrowth was implemented after the 2-h incubation at 30 °C, the transformation reaction was transferred to a culture tube, spun at 500 × g for 2 min in an Allegra X-22R centrifuge (Beckman Coulter, Brea, CA), and the supernatant was replaced with 5 mL SDCA. Cultures were then cultivated at 30 °C for 2 days before plating onto SDCA.

2.4. Genotype analysis of targeted loci

All targeted loci were evaluated for successful integration through colony PCR of transformants using OneTaq DNA Polymerase (New England BioLabs, Ipswich, MA). Analysis at each locus, regardless of type of donor transformed, involved two primers that bind externally to the donor homology sequences (HIS3: KmHIS3 upstream donor F and KmHIS3 downstream donor R; LEU2: KmLEU2 int check F and KmLEU2 int check R; ZWF1: ZWF1 check F and ZWF1 check R; GPD1: GPD1 check F and GPD1 check R). When analyzing colonies transformed with a 2PS or ACC1 donor, an additional primer binding internal to the donor sequence was also included (2PS 650up F or ACC1 int F2). In all cases, the selected set of primers enable simultaneous resolution of the wild-type and integrated genotypes for detection of heterogeneous transformants. Representative images of gel electrophoresis analysis for these loci for each donor evaluated, and for all possible combinations of repair by homologous recombination and non-homologous end joining using a single donor fragment, are listed in Table S3. When determining percent integration, transformants revealed to be homogeneous for the desired integration as well as those detected to be heterogeneous were counted as successful integrations.

To check for gene KOs through the formation of indels by non-homologous end joining, transformants were first restreaked on SDCA and grown at 37 °C for 36 h to eliminate heterogeneous colonies. Three daughter colonies were chosen at random per original transformant and restreaked onto an SD-HIS or SD-LEU plate for colonies transformed with a lagging NHEJ-functional strain ( Juergens et al., 2018 ), or using a 160 bp KO donor comprised solely of 959 bp homology to the gene in strain CBS 712 C overnight. A lack of growth on SD-HIS or SD-LEU paired with a lack of an integration-positive PCR result, where appropriate, was taken as sufficient evidence of a successful KO of HIS3 or LEU2 by NHEJ.

2.5. Knockout of KU70 Gene

To knockout the KU70 gene in strain CBS 712ΔU, a URA3 blaster ( Alani et al., 1987 ) integrating vector was utilized (this effort occurred in parallel to construction of our CRISPR system). This U70-URAblaster vector was built as described by McTaggart (2020), except the homology sequences were extended to 777 and 763 base pairs, creating KU70-UB2V. KU70-UB2V was digested with XhoI and XbaI to remove unnecessary sequences prior to transforming 800 ng of the isolated fragment using the Zymo Frozen-EZ Yeast Transformation II kit as above, except transformation reactions were incubated for three hours at 30 °C prior to plating. Integration of the KU70-UB2V fragment was selected for on SDCA plates. Transformants were screened by colony PCR using three primers (KU70 External F, KU70 Plas REV, and URA seq Rev) to simultaneously check for target and off-target integration. Upon identification of a correctly targeted transformant, the URA3 selection marker was excised by inoculating the transformant into 3 mL YPD for 24 h followed by two 24-h cultivations in 3 mL YPD supplemented with 1 g/L 5-FOA. Cultures were then plated on YPD for single colonies and evaluated by colony PCR for loss of URA3 ( Fig. S2 ).

2.6. TAL biosynthesis and quantification

Engineered strains were re-streaked onto SDCA after transformation to eliminate heterogenous colonies. Loss of the pDBgr-Cas9 plasmid was achieved through 5-FOA treatment as above. Strains were then transformed with pKD-P2PS and stored as frozen glycerol stocks. In preparation for the TAL studies, strains were streaked onto SDCA plates and grown for 2.5 days at 37 °C. Individual colonies were then inoculated into 3 mL SXCA or SLCA media. After 2 days, each strain was reinoculated to OD600 0.1 in 3 mL of the same medium and grown at 37 °C for 48 h in a Gyromax 929 Water Bath Shaker (Amerex Instruments, Inc., Concord, CA). OD600 was then measured, cultures were spun down at 2,600 × g for 4 min (Allegra X-22R Centrifuge, Beckman Coulter, Brea, CA), and a sample of the supernatant was collected and stored at 4 °C until measurement. Samples were prepared and measured for TAL titer using a plate reader (SpectraMax M3, Molecular Devices) as previously described by Lang et al. (2020).

3. Results and discussion

3.1. Design of a high efficiency CRISPR system for K. marxianus

For metabolic engineering of K. marxianus, a high efficiency CRISPR system using donors with short (40 bp) homology arms and no selection marker, minimal DNA input, and no extended outgrowth period would be very advantageous. The flexibility to work in wild-type NHEJ-functional strains is also desirable. To enable this, we constructed plasmid pDBgr-Cas9 by replacing the RNAI-3 driven gRNA expression cassette of pHW601 (a CEN/ARS plasmid harboring a codon-optimized Cas9 gene (Löbs et al., 2017)) with a TGR cassette ( Fig. 1 ) under the control of the strong and constitutive RNAI-2 promoter PTDH3 from S. cerevisiae. The promoter was followed by the K. marxianus glycine tRNA, an NheI restriction site for target site insertion, tracrRNA, HDV ribozyme, and CYC1 terminator from S. cerevisiae. Upon the nucleolytic activity of RNase Z to remove the 5’ tRNA (Schiffer et al., 2002) and the self-cleaving activity of the 3’ HDV ribozyme (Ferré-D’Amare and Scott, 2010), a mature gRNA transcript is produced. Design of the TGR cassette was inspired by work conducted by Ng and Dean (2017) who showed significantly better performance of a TGR cassette over a hammer head ribozyme-gRNA-HDV ribozyme (RGR) cassette in Candida albicans. Such an RGR cassette has previously been used in K. marxianus with success in creating gene KOs, demonstrating a 24% integration efficiency of a donor comprised solely of 959 bp homology to the ADE2 locus in a NHEJ-functional strain (Juergens et al., 2018), or using a 160 bp KO donor for disruption of ARO8 (Rajkumar and Morrissey, 2022).

Fig. 1. TGR gene cassette and approximate premature transcript secondary structure. Black triangles indicate sites of tRNA and HDV ribozyme cleavage.
3.2. Gene disruption and integration in wild-type and NHEJ-deficient strains

The function of our new RNAP II gRNA expression system was evaluated through targeting of the HIS3 and LEU2 open reading frames (ORF) for gene knockout (KO) and various gene integrations. Twenty base pair targeting sequences matching regions for each of these ORFs were cloned into pDBtgr-Cas9, creating pDBtgr-Cas9-HIS3 and pDBtgr-Cas9-LEU2. We chose a range of repair fragment sizes to determine the effect of insert size on efficiency and restricted the homology sequences to 40 bp per arm. Each plasmid was transformed into CBS 712ΔU alone or with one of three different linear DNA repair fragments (Fig. 2A and B) and transformants were selected on uracil-deficient plates. The first was an 80 bp fragment (KO donor) comprised of the fusion of two 40 base pair homology sequences matching regions just upstream and just downstream of the respective target site; this results in removal of approximately 240 base pairs from the ORF when integrated by homologous recombination. The second and third fragments contain the same 40 bp homology sequences flanking either a 2.6 kb expression cassette, and an ACC1 expression cassette, each containing 40 bp homology sequences matching targets of the ORF just upstream and downstream of the targeting sequence. Integration in a NHEJ-functional strain, CBS 712ΔU and E,F in a NHEJ-deficient strain, CBS 712ΔUΔK. The total height of the bars represents average total gene knockout frequency ± standard deviation and the height of the blue bars represents the proportion of KOs that were created through integration of a donor ± standard deviation. Ten transformants were analyzed for each of two transformations per condition. Unpaired t-tests showed there is no significant difference in the percent integration efficiency for the CBS 712ΔU strain when HIS3 is targeted and co-transformed with the KO, 2PS, or ACC1 donor (panel C, three blue bars), nor when LEU2 is targeted and co-transformed with the KO or 2PS donor (panel D, two blue bars). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

To improve the frequency of donor integration, we knocked out the native KU70 gene, encoding a member of the KU70/KU80 heterodimer responsible for the first step of the NHEJ pathway (Chang et al., 2017), creating strain CBS 712ΔUΔK. We first evaluated the effect of KU70 on NHEJ activity by transforming pDBtgr-Cas9-HIS3 and pDBtgr-Cas9-LEU2 in the absence of donor. For such transformations conducted in a NHEJ-deficient strain, it is expected that minimal transformants will develop since successful double-stranded DNA breaks but differential rates of homology-directed repair (HDR) that may be improved through optimization of homology sequence design.

We first evaluated editing efficiency in a K. marxianus strain with native NHEJ activity. Transformants were re-streaked onto SDCA to obtain homogeneous colonies. When these daughter colonies were streaked onto histidine- or leucine-deficient plates, 65–95% failed to grow indicating successful gene knockout through creation of an indel by NHEJ or integration of the repair fragment (Fig. 2C and D; purple bars). Colony PCR of these transformants revealed substantial levels of repair fragment integration at the HIS3 site (Fig. 2C; blue bars). For transformations targeting the LEU2 locus, integration efficiencies were less than those for the HIS3 locus despite similar levels of gene knockout. This indicates high activity of the gRNAs in producing double stranded DNA cuts but differential rates of homology-directed repair (HDR) that may be improved through optimization of homology sequence design.

It has been reported that knockout of KU70 or other genes involved in the NHEJ pathway can hinder the fitness of a strain due to their shared role in telomere maintenance (Liti and Louis, 2003; Polotnianka et al., 2012). Colony PCR of these transformants revealed substantial levels of repair fragment integration at the HIS3 site (Fig. 2C; blue bars). For transformations targeting the LEU2 locus, integration efficiencies were less than those for the HIS3 locus despite similar levels of gene knockout. This indicates high activity of the gRNAs in producing double stranded DNA cuts but differential rates of homology-directed repair (HDR) that may be improved through optimization of homology sequence design.
The deleterious effects of NHEJ gene knockouts on strain fitness has been demonstrated in some yeasts including \textit{S. cerevisiae} (Palmbos et al., 2005) and \textit{S. stipitis} (Ploessl et al., 2021). However, comparison of the growth rate of \textit{CBS 712ΔU} and \textit{CBS 712ΔΔK} grown in YPD at 37 °C revealed no significant differences between the two strains (Fig. S1). This is consistent with work completed by Choo et al. (2014) which demonstrated that knockout of \textit{KU80} in \textit{K. marxianus} KCTC 17555 did not decrease cell fitness under normal and several environmentally stressful conditions. Knockout of \textit{KU70} has the additional benefit of limiting the propensity for random integration events compared to when NHEJ is functional (Abdel-Banar et al., 2009; Rajkumar et al., 2019), facilitating more predictable engineering outcomes.

For a direct comparison with RNAP III-driven systems, we performed the same experiments (using the \textit{HIS3} locus) with two RNAP III-driven systems: \textit{RPR1}_{\text{Pkm}}-\text{rRNA}_{\text{Km}}-\text{gRNA} (RNAP III TG) and \textit{RPR1}_{\text{Pkm}}-\text{rRNA}_{\text{Km}}-\text{gRNA-HDV ribozyme} (RNAP III TGR). The RNAP II TGR system outperformed both RNAP III systems, particularly for gene integrations (Fig. S4). This data agrees with work conducted in \textit{Candida albicans} comparing the efficiency of various RNAP II and RNAP III gRNA expression systems (Ng and Dean, 2017). Overall, the high efficiency of our RNAP II-driven TGR expression cassette provides a significant improvement over the existing \textit{K. marxianus} tools summarized in Table S1. The ability to successfully use donors with only 40 bp homology sequences greatly simplifies donor synthesis since such overhangs can be added through a straight-forward, single-step PCR process. Many other demonstrations of CRISPR-Cas9-mediated gene integrations in \textit{K. marxianus} have used donors with larger homology sequences 0.7–0.9 kb in length which requires multi-step PCR processes and are sometimes cloned into vectors (Cernak et al., 2018; Li et al., 2021). Our system was also able to achieve high genomic modification efficiencies with minimal DNA input, only requiring 500 ng of plasmid and up to 1 μg of linear donor; other studies used several micrograms of each, up to 10 μg plasmid and 5 μg donor (Cernak et al., 2018; Nambu-Nishida et al., 2017, 2018). Our work also demonstrates the largest, single genomic integration (7.6 kb for \textit{ACCI}) achieved through CRISPR-Cas9 without selection in \textit{K. marxianus}, and with near 100% integration efficiency in the \textit{KU70 KO} strain. Prior studies achieved integrations of 3.5 kb or less, not including the often lengthy homology sequences (Cernak et al., 2018; Juergens et al., 2018; Li et al., 2021; Nambu-Nishida et al., 2017, 2018; Rajkumar and Morrissey, 2020). The factors above make our system significantly easier to use, while also offering greater efficiency, the flexibility to use NHEJ-functional and NHEJ-deficient \textit{K. marxianus} strains, and the ability to integrate genes of varying lengths.

### 3.3. Selection of growth-biased metabolic engineering modifications for enhanced polyketide biosynthesis

Previously characterized CRISPR-Cas9 systems for \textit{K. marxianus} (Cernak et al., 2018; Li et al., 2021) as well as other yeasts (Bao et al., 2015; Horwitz et al., 2015; Schwartz et al., 2017) often compensate for lower gRNA-targeting efficiency and low HDR rates by incorporating liquid outgrowth periods in selective media before plating and assaying for the engineered strains. While proven to increase the efficiency of integration in these studies, this method can fail to yield the desired strains if the modification significantly reduces cell growth rate. As previously demonstrated by Schwartz et al. (2017), such mutants are unlikely to be recovered after an outgrowth where the unedited, faster growing cells can vastly outnumber the slower growing mutants in just a few generations. A CRISPR-Cas9 system with universal metabolic engineering utility should allow for selection of gene modifications that potentially reduce cell fitness.

To demonstrate the ability of our new RNAP II system to select strains with reduced cell fitness without use of a liquid outgrowth period, we disrupted the genes for glucose-6-phosphate 1-dehydrogenase (\textit{ZWF1}), the enzyme catalyzing the first step of the pentose phosphate pathway, and glycerol-3-phosphate dehydrogenase (\textit{GPD1}), the enzyme that sequesters carbon away from glycolysis toward the biosynthesis of glycerol. Knockout of \textit{ZWF1} is known to cause growth defects in \textit{K. marxianus} due to impaired regeneration of NADPH (Zhang et al., 2020b); however, loss of carbon in the oxidative phase of the pentose phosphate pathway has been shown to make this gene knockout advantageous for improving the titer of triacetate acid lactone (Cardenas and Da Silva, 2014) and ethanol (Verho et al., 2003) in \textit{S. cerevisiae}.

Knockout of \textit{GPD1} has similarly been shown to cause growth defects in \textit{K. marxianus} (Zhang et al., 2020a), but has also assisted in elevated ethanol accumulation in an engineered \textit{K. marxianus} strain (Zhang et al., 2015).

Strain \textit{CBS 712ΔΔK} was transformed with pDBtgr-Cas9-ZWF1 or pDBtgr-Cas9-GPD1 along with a corresponding 80 bp KO donor and plated immediately after transformation. Screening via colony PCR showed that an average of 60% or 95% of colonies exhibited the desired \textit{ZWF1} or \textit{GPD1} KO, respectively (Fig. 3). However, when the cells were instead inoculated into selective SDCA media and cultured for two days before plating, all ten colonies from each of two biological replicates had the native \textit{ZWF1} or \textit{GPD1} genotype. The cells with targeted gene disruptions were lost during the multi-day outgrowth period due to their significantly slower growth rates relative to the cells that evaded editing (Fig. S5). For comparison, in similar transformations targeting the \textit{HIS3} locus with either the 80 bp KO donor or the \textit{ACCI} donor (modifications that do not result in a substantial change in growth rate (Fig. S5)), plating immediately after transformation or after a two-day outgrowth resulted in no significant change in integration efficiency (Fig. 3). This study demonstrates that the use of a higher efficiency CRISPR-Cas9 system such as pDBtgr-Cas9 makes outgrowth unnecessary. This is vital for recovery of engineered strains with even marginal growth defects that can be lost due to growth-bias during a liquid outgrowth period. Even though outgrowth was found to be unnecessary using pDBtgr-Cas9 plasmids, it is noted that re-streaking of transformants is often necessary to isolate homogeneous colonies before moving forward with plasmid curing and any subsequent transformations. As described previously by Ng and Dean (2017) and Schwartz et al. (2017),
heterogeneous colonies are common if plating immediately after transfection since editing may not occur until after a number of cell divisions.

To demonstrate the importance of eliminating this growth bias for metabolic engineering, we evaluated the resulting strains for improved carbon flux to a desired product in *K. marxianus*. CBS 712ΔUΔKΔZWF1, CBS 712ΔUΔKΔGPD1, and CBS 712ΔUΔKΔACC1 were transformed with pKD-P2PS (McTaggart et al., 2019) to enable production of triacetic acid lactone, a simple polyketide with promising utility as a precursor to several commodity and higher-value products (Chia et al., 2012; Shanks and Keeling, 2017). The three strains were cultured in synthetic defined media supplemented with 9.5 g/L lactose (SLCA) for 48 h at 37°C. Despite the growth impact, the *ZWF1* KO strain had a significant increase in average specific TAL titer (0.92 g/L, 0.082 g/L/OD), 48% higher than the wild-type strain (0.84 g/L, 0.056 g/L/OD) (Fig. 4). The *GPD1* KO strain exhibited similar behavior as the *ZWF1* KO, producing an increase in average specific TAL titer (0.87 g/L, 0.074 g/L/OD) over the unmodified strain. The ability of our pDBtgr-Cas9 system to facilitate selection of genome modifications, including those resulting in reduced *K. marxianus* growth rate, provides opportunities to test a broader range of metabolic engineering interventions that may result in beneficial phenotypes.

### 3.4. Expanding the TGR cassette for simultaneous integrations at multiple loci

Due to the exceptionally high efficiency of our RNAP II-driven TGR cassette for integrating one construct of varying size at a single genomic loci, we expanded the TGR cassette to enable simultaneous targeting of two loci in *K. marxianus* (Fig. 5). A Double gRNA cassette structure, approximate pre-mature transcript secondary structure, and final gRNAs after self-cleavage at the black triangles. B Percentage of transformations with simultaneous, dual integrations at the *HIS3* and *LEU2* loci in *K. marxianus* using *ACC1* and *g2ps1* gene donors, respectively. Bar represents average ± standard deviation of two experiments comprised of two and four transformations, with up to ten transformants analyzed per transformation.
site, we adapted our system for simultaneous modification of two (and potentially more) sites in the genome. pDBtDgr-Cas9 was modified to express more than one gRNA through insertion of an additional tRNA–gRNA-HDV ribozyme cassette immediately downstream of the original TGR sequence, separated by a short spacer sequence (Fig. 5A), yielding plasmid pDBtDgr-Cas9. Expression of this cassette will result in one long transcript containing two gRNAs; however, self-cleaving activity of the tRNAs and HDV ribozymes will ultimately result in two mature and independent gRNAs.

To evaluate the efficiency of this dgrRNA targeting cassette, targeting sequences for HIS3 and LEU2 were inserted into the pDBtDgr-Cas9 backbone using a single-step Gibson assembly process, forming plasmid pDBtDgr-Cas9-HIS3-LEU2. This plasmid was transformed as before into the NHEJ-deficient strain CBS 712ΔUA/K with the ACC1 donor containing 40 bp homologies to the HIS3 locus as well as the ZPS donor containing 40 bp homologies to the LEU2 locus. Across two independent experiments, the average percent of transformants with the desired dual integrations was 25.5% (Fig. 5B).

Others have been successful at integrating more than one gene into a single site in K. marxianus, but the strategies required use of either complex PCR schemes to assemble multiple genes into a single donor fragment with very large homology sequences (Li et al., 2021), or required simultaneous integration of a selection marker (Chang et al., 2012; Ha-Tran et al., 2021; Lee et al., 2020; Lin et al., 2017). This study marks the first demonstration of dual markerless integrations at multiple sites in K. marxianus. Furthermore, only 40 bp homology regions were used and no selection marker was required. By enabling simultaneous editing at two sites in the genome, the double gRNA system described here can expedite K. marxianus strain engineering by consolidating rounds of the time-consuming transformation, mutation verification, and plasmid curing processes by one half. The system also has the potential for expansion to more than two simultaneous integrations by concatenating additional tRNA–gRNA-HDV cassettes. This is an important contribution to the set of genetic tools available for efficient metabolic engineering of K. marxianus, potentially opening the door to the study of a broader set of pathways and biorenewable products in this industrially advantageous host.

4. Conclusions

The development of this highly efficient CRISPR-Cas9 system offers a streamlined platform for expedited strain engineering in K. marxianus by eliminating the need for complex donor assemblies, eliminating extended outgrowth periods, and allowing multiplexed gene editing with the flexibility for use in both NHEJ-functional and -deficient strains. This system provides a powerful addition to the synthetic biology toolbox for metabolic engineering of K. marxianus, further advancing this yeast as a promising candidate for extensive industrial biomanufacturing.

Data availability

Data will be made available on request.

Acknowledgements

This research was supported by the National Science Foundation [grant numbers CBET-1803677 (NDS) and CBET-1803630 (IW)] and the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program [grant number SC0019093 (IW and NDS)]. The authors would like to thank Tami McTaggart for her helpful advice and Shane Bassett for his general experimental support and writing assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2022.e00208.

References

Abdel-Banat, B.M.A.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, https://doi.org/10.1002/yea.1729 n/a-4.

Alani, E., Cao, L., Kleckner, N., 1987. A method for gene disruption that allows repeated use of URA3 selection in the construction of multiply disrupted yeast strains. Genetics 116, 541–545. https://doi.org/10.1534/genetics.112.541-test.

Ahleschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(03)80360-2.

Amrane, A., Prigent, Y., 1988. Effect of culture conditions of Kluyveromyces marxianus on its autolysis, and process optimization. Bioprocess Eng. 18, 383–388. https://doi.org/10.1016/S0565-0089(98)80009-6.

Banat, I.M., Nigam, P., Marchant, R., 1992. Isolation of thermotolerant, fermentative yeasts growing at 52°C and producing ethanol at 45°C and 50°C. World J. Microbiol. Biotechnol. 8, 259–263. https://doi.org/10.1007/BF02210874.

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., Shi, T., Zhao, H., 2015. Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in Saccharomyces cerevisiae. ACS Synth. Biol. 4, 585–594. https://doi.org/10.1021/sb500255b.

Benchling, 2021. Benchling [Biology Software].

Cardenas, J., Da Silva, N.A., 2014. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic lactone. Metab. Eng. 25, 194–203.

Cernak, P., Estrela, R., Poddar, S., Skoker, J.M., Cheng, Y.-F., Carlson, A.K., Chen, B., Glyn, V.M., Furlan, M., Ryan, O.W., Donnelly, M.K., Arkin, A.P., Taylor, J.W., Cate, J.H.J., 2018. Engineering Kluyveromyces marxianus as a robust synthetic biology platform host. mBio 9, https://doi.org/10.1128/mBio.01410-18.

Chang, H.H.Y., Fanunno, M.R., Adachi, N., Lieber, M.R., 2017. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat. Rev. Mol. Cell Biol. https://doi.org/10.1038/nrm.2017.46.

Chang, J.J., Ho, C.Y., Hu, F.J., Tsai, T.Y., Ke, H.M., Wang, C.H.T., Chen, H.L., Shih, M.C., Huang, C.C., Li, W.H., 2012. PGASO: a synthetic biology tool for engineering cellulosytic yeast. Biotechnol. Biofuels 5, 1–12. https://doi.org/10.1186/1754-6834-5-3/FIGURES/5.

Charoensuphan, K., Thansukho, P., Thansukho, S., Yamada, M., 2015. Ethanol production from Jerusalem artichoke artichoke tubers at high temperature by newly isolated thermotolerant inulin-utilizing yeast Kluyveromyces marxianus utilizing consolidated bioprocessing. Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. 108, 173–190. https://doi.org/10.1007/s10482-015-0476-5/TABLES/10.

Chia, M., Schwartz, T.J., Shanks, B.H., Dumesic, J.A., 2012. Triacetic lactone as a potential biorenewable platform chemical. Green Chem. 14, 1850–1853. https://doi.org/10.1039/C2GC35340A.

Choi, J.W., Da Silva, N.A., 2014. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. J. Biotechnol. 187, 56–59. https://doi.org/10.1016/j.jbiotec.2014.07.430.

Choo, J.H., Han, C., Kim, J.Y., Kang, H.A., 2014. Deletion of a KU80 homolog enhances homologous recombination in the thermotolerant yeast Kluyveromyces marxianus. Biotechnol. Lett. 36, 2059–2067. https://doi.org/10.1007/s12019-014-1576-4/TABLES/2.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343. https://doi.org/10.1093/nar/gkt135.

Doench, J.G., Fusi, N., Sullenger, M., Hegde, R., Vainberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H.W., Listgarten, J., Root, D.E., 2015. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. 34, 184–191. https://doi.org/10.1038/nbt.3437.

Eckermann, S., Schröder, G., Schmidt, J., Streek, D., Erdna, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Frohch, P., Teeri, T.I., Schröder, J.,
Yang, P., Zhu, X., Zheng, Z., Mu, D., Jiang, S., Luo, S., Wu, Y., Du, M., 2018. Cell regeneration and cyclic catalysis of engineered Kluyveromyces marxianus of a d-psicose-3-epimerase gene from Agrobacterium tumefaciens for d-allulose production. World J. Microbiol. Biotechnol. 34, 1–7. https://doi.org/10.1007/S11274-018-2451-6/TABLES/1.

Zhang, B., Ren, L., Wang, H., Xu, D., Zeng, X., Li, F., 2020a. Glycerol uptake and synthesis systems contribute to the osmotic tolerance of Kluyveromyces marxianus. Enzym. Microb. Technol. 140, 109641 https://doi.org/10.1016/J. ENZMICTEC.2020.109641.

Zhang, B., Ren, L., Zeng, S., Zhang, S., Xu, D., Zeng, X., Li, F., 2020b. Functional analysis of PGI1 and ZWF1 in thermotolerant yeast Kluyveromyces marxianus. Appl. Microbiol. Biotechnol. 104, 7991–8006. https://doi.org/10.1007/S00253-020-10808-4/FIGURES/8.

Zhang, J., Zhang, B., Wang, D., Gao, X., Sun, L., Hong, J., 2015. Rapid ethanol production at elevated temperatures by engineered thermotolerant Kluyveromyces marxianus via the NADP(H)-preferring xylose reductase-xylitol dehydrogenase pathway. Metab. Eng. 31, 140–152. https://doi.org/10.1016/J. YMBEN.2015.07.008.