

SCIENTIFIC REPORTS



OPEN

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

Yumiko Nambu-Nishida^{1,2}, Keiji Nishida³, Tomohisa Hasunuma³ & Akihiko Kondo^{2,3,4}

Kluyveromyces marxianus, a non-conventional thermotolerant yeast, is potentially useful for production of ethanol and other products. This species has a strong tendency to randomly integrate transforming DNA fragments, making necessary the development of more precise methods for gene targeting. In this study, we first demonstrated that *K. marxianus* NBRC1777 is cold-tolerant, and then established a highly efficient and precise technique for gene editing by introducing genes encoding deaminase-mediated targeted point mutagenesis (Target-AID) and clustered regularly interspaced short palindromic repeats (CRISPR) associated proteins (CRISPR-Cas9). We used Target-AID to introduce targeted point mutations that disrupted *Nej1* or *Dnl4*, genes that are involved in non-homologous end-joining (NHEJ). Both of the resulting mutant strains showed enhanced proportions of homology-mediated integration compared to the wild-type parent. In combination with target cleavage by CRISPR-Cas9, markerless integration was performed using short (~50 bp) flanking homologous sequences. Together, these tools render this species fully tractable for gene manipulation, permitting targeted genetic changes in the cold- and thermo-tolerant yeast *K. marxianus*.

The non-conventional yeast *Kluyveromyces marxianus* is known to be thermotolerant and Crabtree-negative. The latter term indicates that this species produces ethanol only under anaerobic conditions¹; in contrast, Crabtree-positive yeasts such as *Saccharomyces cerevisiae* obligately produce ethanol both aerobically and anaerobically. Thus, Crabtree-negative yeasts are more suitable for the generation of various products other than ethanol because these strains do not undergo aerobic alcoholic fermentation, which consumes carbon resources². The thermotolerance of *K. marxianus* also allows faster growth and production at higher temperatures, and some volatile products can be collected during synthesis more efficiently at higher temperature. Another benefit of the faster growth of *K. marxianus* at 37 °C is the ability to obtain single colonies after overnight incubation, while generally used *S. cerevisiae* strains take 2 to 3 days at 30 °C. Recently, the complete genome sequence for *K. marxianus* has been reported³⁻⁵.

Unlike *S. cerevisiae*, which demonstrates homologous recombination (HR) of transformed DNA, *K. marxianus* DMKU3-1042 tends to integrate DNA fragments randomly into its genome¹. Such random integration is attributed to the high activity of the non-homologous end-joining (NHEJ) pathway⁶. *K. marxianus* NBRC1777 reportedly needed relatively large homology arms (641 and 1475 bp) of flanking DNA for insertion by HR⁷, while *S. cerevisiae* requires only ~50-bp homology arms, sequence lengths that can be added at the 5' ends of PCR primers⁸.

¹Technology Research Association of Highly Efficient Gene Design (TRAHED), 7-1-49 Minatojimaminamimachi, Chuo-ku, Kobe, Hyogo, 650-0047, Japan. ²Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo, 657-8501, Japan. ³Graduate School of Science, Technology and Innovation, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo, 657-8501, Japan. ⁴Biomass Engineering Program, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan. Correspondence and requests for materials should be addressed to A.K. (email: akondo@kobe-u.ac.jp)

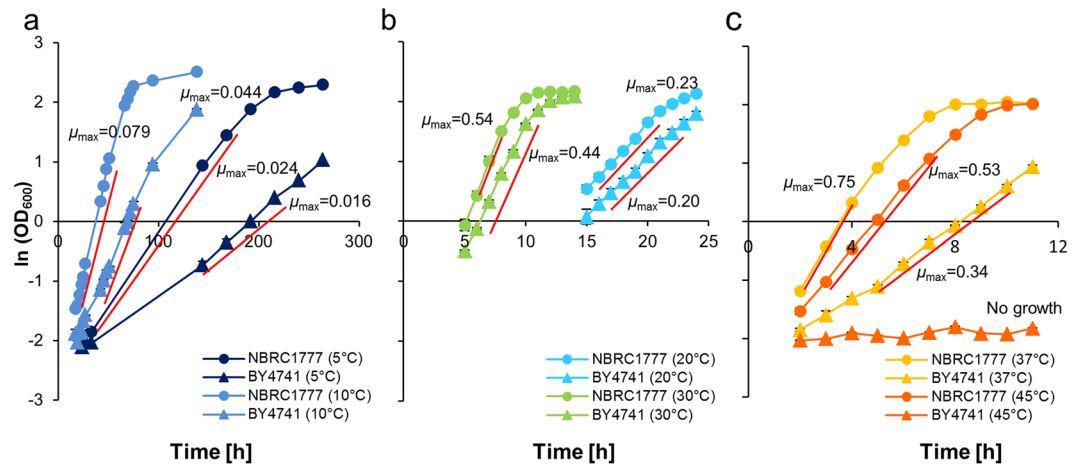


Figure 1. Cell growth of *K. marxianus* and *S. cerevisiae* at various temperatures. Cell growth was compared between *K. marxianus* NBRC1777 and *S. cerevisiae* BY4741 at 5, 10 °C (a), at 20, 30 °C (b), and at 37, 45 °C (c). The maximum specific growth rate (μ_{\max}) was determined as the slope of the mid-log phase of natural logarithm (ln) (OD₆₀₀). Timeframes for each μ_{\max} were as follows: at 5 °C from 33 h to 168 h and 144 h to 216 h, at 10 °C from 19 h to 48 h and 42 h to 75 h, at 20 °C from 16 h to 20 h and 16 h to 22 h, at 30 °C from 6 h to 8 h and from 6 h to 10 h, at 37 °C from 2 h to 4 h and 5 h to 10 h, 45 °C from 3 h to 7 h and no growth for *K. marxianus* and *S. cerevisiae*, respectively.

Disruption of the NHEJ pathway in yeasts has been shown to suppress random integration events and therefore increase the success rate of HR in these organisms^{6,9}. The core NHEJ proteins are Ku70/Ku80, Lig4/Lif1, and Nej1^{10–13}. In the first step, a stable heterodimer of Ku70/Ku80 binds to broken DNA ends and acts as a bridging complex. Lig4 and Lif1 also form a highly stable complex. Lig4 is an ATP-dependent double-strand break (DSB) repair DNA ligase and Lif1 is Lig4's stabilizing, stimulating, and targeting co-factor¹⁰. In *S. cerevisiae*, deletion of *NEJ1* reduces NHEJ by 100-fold⁹.

Clustered regularly interspaced short palindromic repeats (CRISPR) associated proteins (CRISPR-Cas9) genome-editing tools have been applied in several yeast species, facilitating NHEJ-mediated insertion/deletion formation. HR-mediated integration following CRISPR-Cas9 editing has also been demonstrated, with efficiency varying depending on the yeast species^{2,14}. More recently, deaminase-mediated targeted point mutagenesis (Target-AID) has been performed in *S. cerevisiae*; which is a hybrid system of nuclease-deficient CRISPR-Cas9 and activation-induced cytidine deaminase (AID)¹⁵. When AID ortholog PmCDA1 was fused to C-terminus of nickase Cas9(D10A), C to T/G point mutation was induced within 3–5 bases window at the 5' end of target sequence. Target-AID also appeared to be much less toxic compared to CRISPR-Cas9.

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

Results and Discussions

Cold-tolerant yeast *K. marxianus* strain. To characterize growth rates at various temperature, *K. marxianus* NBRC1777 and *S. cerevisiae* BY4741 were cultured at 5, 10, 20, 30, 37, and 45 °C (Fig. 1). The maximum specific growth rate (μ_{\max}) was higher for *K. marxianus* than for *S. cerevisiae* at all temperatures (Fig. 1), indicating that the *K. marxianus* is a fast-growing strain. *K. marxianus* showed best growth at 37 °C and was able to grow at 45 °C. The difference in the μ_{\max} was also bigger at 10 °C (0.079/0.044) compared to the difference at 20 °C (0.23/0.20). Thus, in addition to the known thermotolerance of *K. marxianus*⁷, strain NBRC1777 exhibited fast and robust growth at a wide range of temperatures, a feature that will be appealing for a variety of industrial applications.

Isolation of KmSNR52 promoter for expression of sgRNA for CRISPR-Cas9. To introduce the CRISPR system into *K. marxianus*, an efficient promoter for sgRNA expression was needed. As 5'-capped transcripts are not suitable for use as sgRNAs, PolII-driven promoters are not favoured. The *SNR52* promoter has been used to express functional sgRNAs in *S. cerevisiae*¹⁶, so we obtained the *K. marxianus* *SNR52* promoter (KmP_{SNR52}) as follows. The *K. marxianus* *SNR52* transcript sequence was first identified using the BLASTN search program with the *S. cerevisiae* *SNR52* transcript as a query. The region upstream of the putative *KmSNR52* transcript was then aligned with the ScSNR52 promoter; a 531-bp fragment was thereby defined as the putative KmP_{SNR52} (Supplementary Fig. S1).

We next constructed a vector for use as a shuttle vector by combining a *K. marxianus* autonomously replicating sequence (KmARS7)¹⁷ with a centromere sequence (KmCEN-D)¹⁸ and demonstrated (data not shown) that this plasmid is retained as an episome in *K. marxianus* NBRC1777. A human-optimized *Streptococcus pyogenes* (Sp) *cas9*¹⁶ under the control of the constitutive *S. cerevisiae* *PDC1* promoter (ScP_{PDC1}) was inserted into the shuttle vector, yielding a basal plasmid (Cas9_Base; Supplementary Fig. S2 and S3) capable of directing the expression of Cas9 in *K. marxianus*.

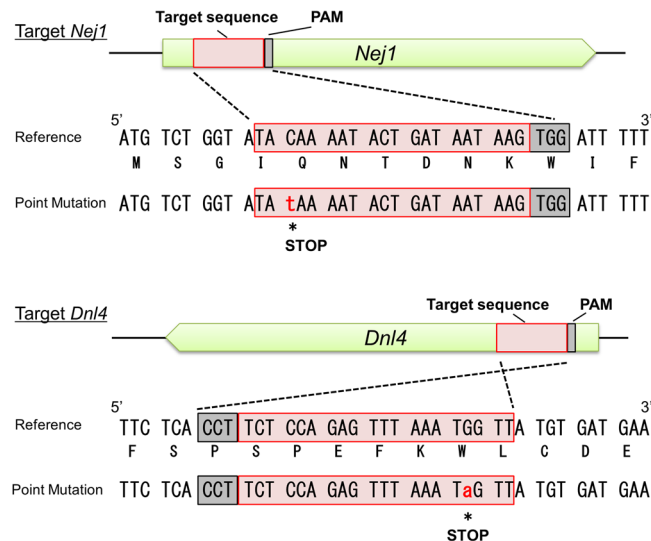


Figure 2. Induction of mutations by Target-AID. The *Nej1* and *Dnl4* target sites were selected for the introduction of stop codons by Target-AID. Red boxes indicate the target sites and grey boxes indicate the PAM sequences. Lower-case nucleotide in red font indicates the mutated base. The sequences of the induced mutations are shown in alignment with the reference sequences, with each predicted amino acid sequence below the respective nucleic acid sequence.

Inactivation of the NHEJ pathway by Target-AID point mutation of NHEJ proteins. In *S. cerevisiae*, the *NEJ1* and *DNL4* genes encode core components of the proteins mediating NHEJ. By analogy to *S. cerevisiae*, disruption of the *K. marxianus* *Nej1* and *Dnl4* homologs are expected to yield more genetically tractable *K. marxianus* hosts with higher proportions of HR. Mutation of the *K. marxianus* *Nej1* and *Dnl4* genes identified by sequence similarity to the respective budding yeast genes was performed using targeted point mutagenesis by Target-AID in *K. marxianus*. A basal plasmid (nCas9-CDA_Base) was generated by constructing a variant of the Cas9_Base vector that encoded a nickase mutant (D10A) SpCas9 protein fused to the *Petromyizon marinus* (Pm) CDA1 cytidine deaminase¹⁵ (Supplementary Fig. S2 and S4). Following a previous study in *S. cerevisiae*⁹, target sites in the *Nej1* and *Dnl4* genes of *K. marxianus* NBRC1777 were selected so as to introduce stop codons by causing C-to-T mutagenesis at 16 to 19 bp upstream of the sequences encoding protospacer adjacent motif (PAM) sequences for the target sequences. *K. marxianus* NBRC1777 cells were transformed with plasmids containing each targeting sequence (nCas9-CDA_target *Nej1* and nCas9-CDA_target *Dnl4*). G418-resistant colonies were selected; the *Nej1* and *Dnl4* genes of the transformants then were PCR amplified and subjected to sequence analysis to confirm the presence of the targeted lesions (Fig. 2). Both plasmids introduced stop codons as expected, providing null alleles of *Nej1* (*Nej1*^o) and *Dnl4* (*Dnl4*^o) with 1 of 8 (or 1 of 4) G418-resistant colonies contained the mutant allele, respectively. Thus, Target-AID was efficient enough to introduce targeted point mutation in *K. marxianus*.

Conventional HR in NHEJ null mutants. To investigate if these disruptions in genes with putative NHEJ activities increased the rate of HR in *K. marxianus*, cells of the *Nej1*^o or *Dnl4*^o mutants were transformed with DNA donor fragments with homology arms. Specifically, we employed a *GFP bleo* cassette (i.e., encoding green fluorescent protein (GFP) fused to a protein providing a selectable marker for bleomycin/zeocin resistance) flanked by *K. marxianus* *Ura3* sequences. Two separate constructs were used, one providing extended homology (1000-bp arms at both ends) and the second providing shorter homology (arms of 122 and 109 bp at the two ends). The fragments were expected to replace *Ura3* with the *GFP-bleo* cassette if HR was facilitated by mutation of genes of the NHEJ pathway. The cells were transformed and selected for zeocin-resistance; the resulting transformants were then screened for 5-fluoroorotic acid (5-FOA) resistance, which would be conferred by *Ura3* disruption. When using the construct providing homology arms of 1000 bp, 100% of the Zeo^R transformants of the *Nej1*^o or *Dnl4*^o strains (17/17 and 13/13, respectively) were 5-FOA^R; in comparison, only 26% (6/23) of the Zeo^R transformants of the wild-type (WT) strain were 5-FOA^R (Fig. 3). When using equimolar quantities of the construct providing shorter homology arms (122 bp/109 bp), 123 Zeo^R transformants were obtained from the WT strain but no 5-FOA^R transformants were obtained. The *Nej1*^o or *Dnl4*^o strains yielded one and two Zeo^R transformants (respectively), of which no and one (respectively) were 5-FOA^R (i.e., disrupted at the targeted *Ura3* locus) (Fig. 3).

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

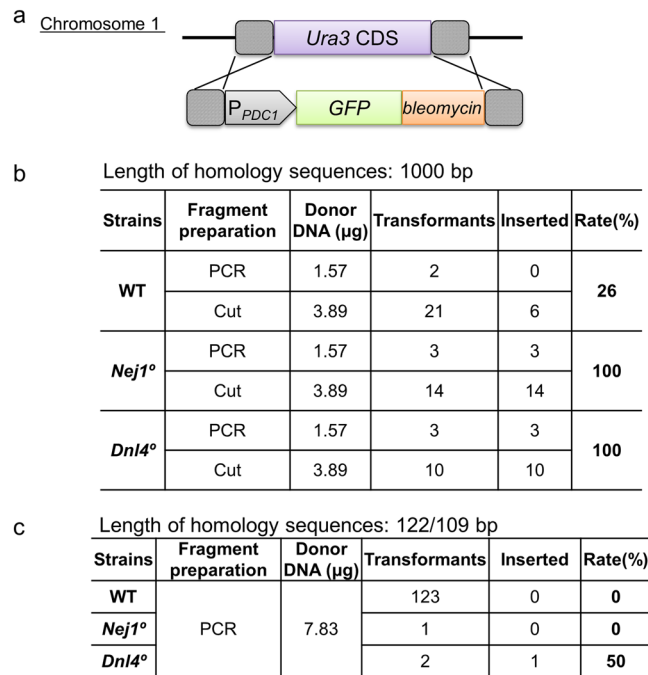


Figure 3. Effect of *Nej1*^o and *Dnl4*^o on homologous recombination. (a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus. Dashed boxes indicate homologous regions. (b), (c) Homologous recombination rates (5-FOA resistant (inserted)/zeocin resistant (transformants)). Fragments were prepared either by PCR or by restriction digestion of plasmid (Cut).

HR with CRISPR-Cas9. To boost the HR frequency, co-transformation with a construct encoding CRISPR-Cas9 was performed; the CRISPR system was expected to facilitate HR by providing cleavage of the targeted locus. A pair of target sites was selected at just inside of the homologous regions of *Ura3*. WT, *Nej1*^o, and *Dnl4*^o cells were transformed with the fragment with homology arms of 122 bp/109 bp (in combination with the Cas9 construct) and transformants were selected using either zeocin, (*Ura3*-targeting fragment), G418 (CRISPR-Cas9 vector), or both. For all three strains, a larger number of antibiotic-resistant transformants were obtained when the transformation mixes included the CRISPR-Cas9 construct. Among the antibiotic-resistant transformants, 41–79% of those derived from WT cells were 5-FOA^R (i.e., disrupted at *Ura3*), while 95–100% of those derived from the *Nej1*^o and *Dnl4*^o strains were 5-FOA^R (Fig. 4). These results indicated that target cleavage by Cas9 enhanced the frequency of HR sufficiently that even without selection by zeocin (i.e., without direct selection for the integration at *Ura3*), disruption of *Ura3* was seen in virtually all transformants.

Markerless integration with 50-bp homology arms and CRISPR-Cas9. Markerless integration was attempted by co-transformation using the combination of a donor DNA template containing a *GFP* expression cassette flanked by 50-bp homology arms along with a CRISPR-Cas9-encoding construct. Transformants were selected using G418 and then screened by both the 5-FOA test and PCR analysis. We observed successful integration (i.e., disruption of the *Ura3* target locus in G418^R transformants) at rates of 28%, 84%, and 92% in the WT, *Nej1*^o, and *Dnl4*^o backgrounds, respectively (Fig. 5). Taking advantage of the markerless integration, the *GFP*-encoding sequence was directly inserted into the coding sequence of the endogenous *Sed1* gene so as to encode a *GFP*-*Sed1* fusion protein preceded by the *Sed1* localization peptide. A single cleavage site for CRISPR-Cas9 within this target gene was designed in such a way that the site would not be retained after successful integration. Short (50-bp) homology arms flanking the target site were added to either end of the *GFP*-encoding sequence by PCR. Transformants were selected by G418 and then screened by PCR, revealing gene replacement at the *Sed1* target locus at 38%, 100%, and 100% success rates in the WT, *Nej1*^o, and *Dnl4*^o backgrounds, respectively (Fig. 6). When viewed by fluorescence microscopy, cells harboring the targeted integration exhibited peripheral localization of *GFP*-*Sed1*, consistent with the localization previously reported for the equivalent construct in *S. cerevisiae*¹⁹. Thus, this method can be applied for scar-less genome editing in *K. marxianus*.

Assembly and integration of multiple fragments without markers. *In vivo* assembly and integration of multiple fragments was performed using three DNA fragments, each containing 50 bp overlap with the adjacent fragments or target sites, such that the entire construct was designed to replace *Ura3* (Fig. 7). Compared to single PCR product donor constructs, a smaller number of transformants was obtained by G418 selection when using the multiple-fragment DNA as the donor. Nonetheless, the WT, *Nej1*^o, and *Dnl4*^o strains showed integration efficiencies of 18% (2/11), 100% (3/3), and 100% (4/4), respectively. This assembly and integration method is expected to allow combinatorial integration of gene cassettes using (for example) combinations of various promoters and coding regions.

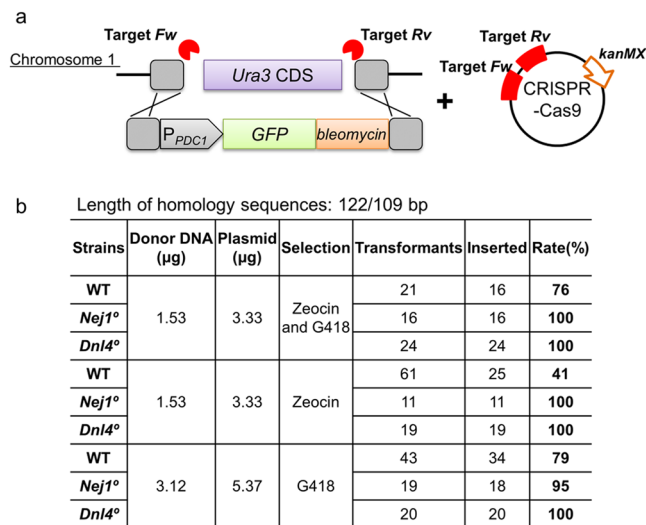


Figure 4. Homologous recombination facilitated by CRISPR-Cas9. (a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus and CRISPR-Cas9 plasmid targeting two sites (indicated by lightning bolts) flanking the *Ura3* coding sequences (CDS). Dashed boxes indicate homologous regions. Bottom panel indicates (b) Homologous recombination rate (5-FOA resistant (inserted)/zeocin and/or G418 resistant (transformants) * 100).

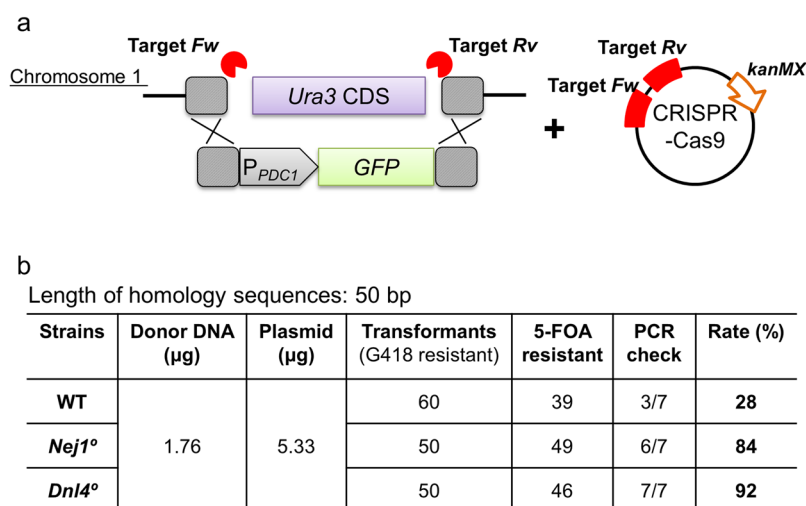


Figure 5. Markerless integration with short homologous arms. (a) Schematic of the transforming donor DNA fragment targeting the *Ura3* locus and the CRISPR-Cas9 plasmid targeting two sites (indicated by lightning bolts) flanking the *Ura3* CDS. Dashed boxes indicate homologous regions. (b) Homologous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

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

Methods

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

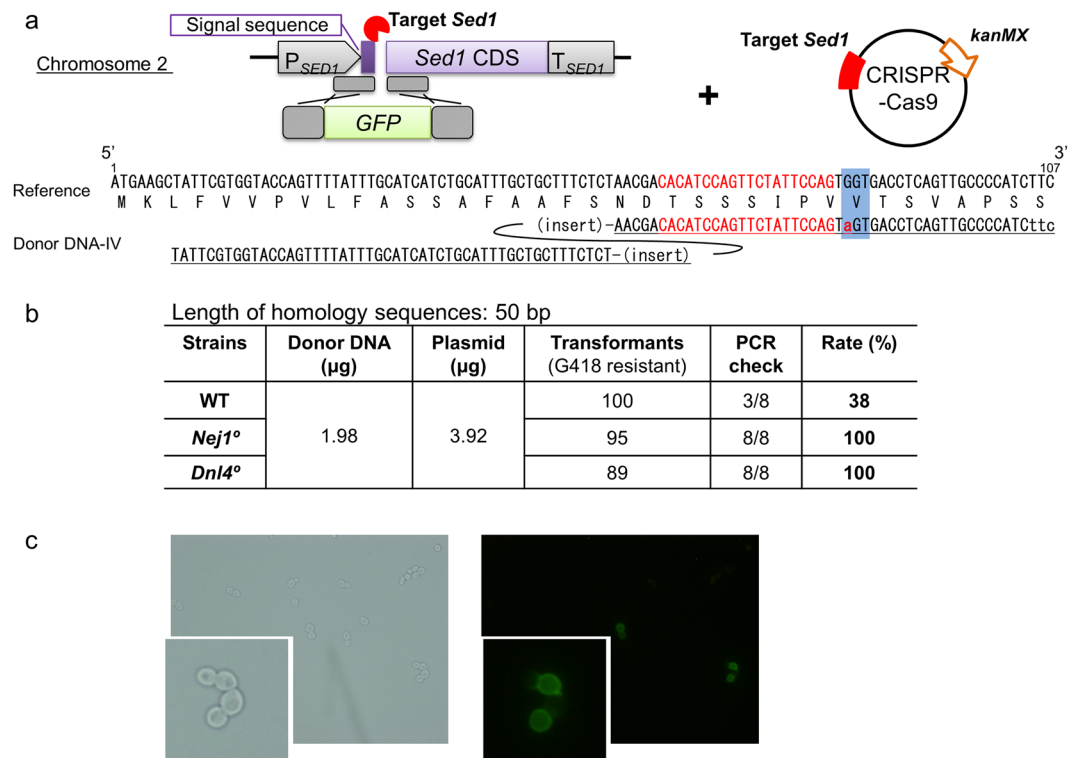


Figure 6. Direct fusion of GFP-coding sequence into a chromosomal locus. (a) Schematic showing integration of the GFP-coding fragment into the *Sed1* CDS. The CRISPR-Cas9 plasmid targeted (indicated by lightning bolts) the end of the signal peptide-coding region of *Sed1*. Dashed boxes indicate homologous regions. Left arm is homologous to the end of the signal peptide-coding sequence (bp 8 to 57 with respect to the start of the *Sed1* coding sequence). Right arm is homologous to *Sed1* sequence (bp 58 to 107 with respect to the start of the *Sed1* coding sequence) with substitution at bp 84 (G to A; indicated by lower-case letter in red font) to eliminate a PAM. (b) The successful integration rate was calculated as PCR check * 100. (c) Representative microscopic image of positive transformants (*Nej1*^o/GFP_ *Sed1*) under bright-field (left) and fluorescent (right) illumination. Insets show 3-fold magnified images.

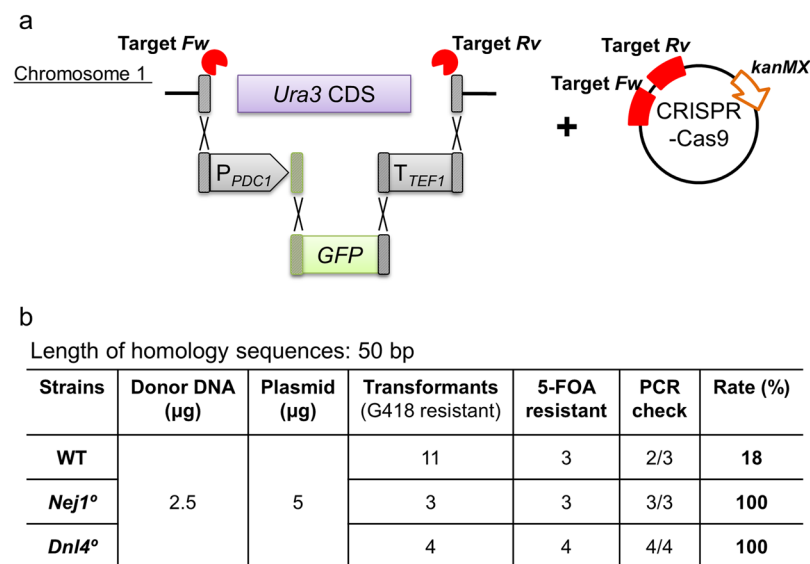


Figure 7. *In vivo* assembly and integration of three DNA fragments with 50 bp overlaps. (a) Schematic of the three transforming donor DNA fragments replacing the *Ura3* locus along with the CRISPR-Cas9 plasmid targeting two sites (indicated by lightning bolts) flanking the *Ura3* CDS. Dashed boxes indicate homologous regions. Bottom panel shows (b) Homologous recombination rate ((5-FOA resistant * PCR check)/ transformants (G418 resistant) * 100).

| Plasmids and strains | Genotype | Reference |
|---------------------------------------|--|--|
| Plasmids | | |
| nCas9-CDA_Base | ScP _{PDC1} -nCas9-T _{TDH3} , KmARS7, Km CEN D, <i>kanMX</i> , ori, and AmpR | This study |
| nCas9-CDA_target <i>Nej1</i> | KmP _{SNR52} -target- <i>Nej1</i> _sgRNA-T _{sup4} cassette in nCas9-CDA_Base | This study |
| nCas9-CDA_target <i>Dnl4</i> | KmP _{SNR52} -target- <i>Dnl4</i> _sgRNA-T _{sup4} cassette in nCas9-CDA_Base | This study |
| Cas9_Base | ScP _{PDC1} -Cas9-T _{TDH3} , KmARS7, Km CEN D, <i>kanMX</i> , ori, and AmpR | This study |
| Cas9_Base_target <i>Ura3</i> _Fw | KmP _{SNR52} -target- <i>Ura3</i> _Fw_sgRNA-T _{sup4} cassette in Cas9_Base | This study |
| Cas9_Base_target <i>Ura3</i> _Fw + Rv | KmP _{SNR52} -target- <i>Ura3</i> _Fw_sgRNA-T _{sup4} cassette and KmP _{SNR52} -target- <i>Ura3</i> _Rv_sgRNA-T _{sup4} cassette in Cas9_Base | This study |
| Cas9_Base_target <i>Sed1</i> | KmP _{SNR52} -target- <i>Sed1</i> _sgRNA-T _{sup4} cassette in Cas9_Base | This study |
| TOPO/F7 | ScP _{PDC1} -EGFP_bleomycin on both sides <i>Ura3</i> homologous sequence cassette in pCR TM 4Blunt-TOPO [®] | This study |
| Strains | | |
| <i>S. cerevisiae</i> | | |
| BY4741 | <i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</i> | ATCC |
| <i>K. marxianus</i> | | |
| NBRC1777 | Wild-type | NITE Biological Resource Center, Japan |
| <i>Nej1</i> [°] | <i>Nej1</i> disrupted by C to T point mutation at position 13 | This study |
| <i>Dnl4</i> [°] | <i>Dnl4</i> disrupted by G to A point mutation at position 44 | This study |

Table 1. Plasmids and strains used in this study.

Growth measurements. *S. cerevisiae* BY4741 and *K. marxianus* NBRC1777 strains were grown in YPD medium overnight at 30 °C by shaking in test tubes. The cells were inoculated into 5 mL of YPD medium at a starting optical density at 600 nm (OD₆₀₀) of 0.05, and cultures were cultivated at 5, 10, 20, 30, 37, or 45 °C. Cell growth was tracked by measuring OD₆₀₀ using a UV mini spectrophotometer (Shimadzu, Kyoto, Japan). The μ_{\max} was determined as the slope of the log phase of natural logarithm (ln) transformation of the linear region.

Construction of basal genome-editing plasmids. The vector sequences are provided in Supplementary Fig. S3 and S4. Vector assembly was performed by PCR using the In-fusion cloning method (Takara Bio, Shiga, Japan). The vector backbone was designed to contain KmARS7, KmCEN-D and selectable marker *kanMX*, which encodes G418 resistance. For CRISPR-Cas9 (Cas9_Base) and Target-AID (nCas9-CDA_Base) vectors, a gene encoding human-optimized nCas9-PmCDA1 or Cas9^{15, 16} (respectively) was placed between ScP_{PDC1} and the *TDH3* terminator (ScT_{TDH3}) on the vector backbone.

Construction of target sequence-containing plasmids. To generate the sgRNA expression cassettes, the *K. marxianus* *SNR52* promoter (KmP_{SNR52}) (531 bp; identified by alignment with the *S. cerevisiae* gene) was amplified by PCR using the combination of the forward primer (P_Km01-013) and each reverse primer (5'-CTAGCTCTAAAAC- reverse complement of target sequence-GATTTCGAAGTGGACGTTG-3'). The sgRNA scaffold and the *SUP4* terminator were amplified by PCR from plasmid pRS426-SNR52p-grNA.CAN1.Y-SUP4t¹⁶ using the primers P_Km01-014 and P_Km01-015 (Supplementary Table S1). The two fragments were fused by overlap extension PCR using the primers P_Km01-013 and P_Km01-015 to yield each target-containing sgRNA cassette (Supplementary Fig. S1). Each cassette was digested with restriction enzymes NotI and NheI and ligated into NotI-, SpeI-digested nCas9-CDA_Base or Cas9_Base plasmid.

Transformation. *K. marxianus* cells were transformed by the lithium acetate method²⁰. After transformation, cells were plated on YPD containing appropriate selection reagents (100 μg/mL G418 and/or 50 μg/mL zeocin) and grown for one day. For the 5-FOA test, cells were streaked onto YPD containing 3 mg/mL 5-FOA, incubated overnight, and assessed for growth. Colony PCR was performed to check fragment size and to obtain DNA for Sanger sequencing using a 3130xL Genetic Analyzer (Applied Biosystems, CA, USA).

Donor DNA templates. Each donor DNA template was constructed by PCR and cloned into the plasmid TOPO/F7 using the Zero Blunt[®] TOPO[®] PCR Cloning Kit (Invitrogen, CA, USA) (Fig. 5). Each cloned donor DNA was either excised by restriction enzymes or amplified by PCR using the respective primer pair (Supplementary Table S1), and the resulting fragment was purified by agarose gel electrophoresis before use in transformation.

References

- Nonklang, S. *et al.* High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. *Appl. Environ. Microbiol.* **74**, 7514–7521 (2008).
- Wagner, J. M. & Alper, H. S. Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. *Fungal Genet. Biol.* **89**, 126–136 (2016).

3. Jeong, H. *et al.* Genome sequence of the thermotolerant yeast *Kluyveromyces marxianus* var. *marxianus* KCTC 17555. *Eukaryot. Cell* **11**, 1584–1585 (2012).
4. Suzuki, T., Hoshino, T. & Matsushika, A. Draft genome sequence of *Kluyveromyces marxianus* strain DMB1, isolated from sugarcane bagasse hydrolysate. *Genome Announc.* **2**, e00733–14 (2014).
5. Inokuma, K., Ishii, J., Hara, K. Y., Mochizuki, M. & Hasunuma, T. Complete genome sequence of *Kluyveromyces marxianus* NBRC1777, a nonconventional thermotolerant yeast. *Yeast* **3**, e00389–15 (2015).
6. Abdel-Banat, B. M. A., Nonklang, S., Hoshida, H. & Akada, R. Random and targeted gene integrations through the control of non-homologous end joining in the yeast *Kluyveromyces marxianus*. *Yeast* **27**, 29–39 (2010).
7. Hong, J., Wang, Y., Kumagai, H. & Tamaki, H. Construction of thermotolerant yeast expressing thermostable cellulase genes. *J. Biotechnol.* **130**, 114–123 (2007).
8. Da Silva, N. a & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **12**, 197–214 (2012).
9. Valencia, M. *et al.* *NEJ1* controls non-homologous end joining in *Saccharomyces cerevisiae*. *Nature* **414**, 666–669 (2001).
10. Dudášová, Z., Dudáš, A. & Chovanec, M. Non-homologous end-joining factors of *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* **28**, 581–601 (2004).
11. Feldmann, H. & Winnacker, E. L. A putative homologue of the human autoantigen Ku from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **268**, 12895–12900 (1993).
12. Feldmann, H. *et al.* HDF2, the second subunit of the Ku homologue from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **271**, 27765–27769 (1996).
13. Moore, J. K. & Haber, J. E. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 2164–2173 (1996).
14. Jakočiūnas, T., Jensen, M. K. & Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **34**, 44–59 (2016).
15. Nishida, K. *et al.* Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729-1-aaf8729-8 (2016).
16. Dicarlo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
17. Hoshida, H., Murakami, N., Suzuki, A., Tamura, R. & Asakawa, J. Non-homologous end joining-mediated functional marker selection for DNA cloning in the yeast *Kluyveromyces marxianus*. *Yeast* **31**, 29–46 (2014).
18. Yarimizu, T., Nakamura, M., Hoshida, H. & Akada, R. Screening of accurate clones for gene synthesis in yeast. *J. Biosci. Bioeng.* **119**, 251–259 (2015).
19. Huh, W.-K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
20. Gietz, D., St Jean, A., Woods, R. A. & Schiestl, R. H. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* **20**, 1425 (1992).

Acknowledgements

This research was supported by the Project Focused on Developing Key Technology for Discovering and Manufacturing Drugs for Next-Generation Treatment and Diagnosis from the Japan Agency for Medical Research and Development (AMED), Japan.

Author Contributions

Y.N. and K.N. designed the experiment and contributed to interpretation of data. Y.N. wrote the initial draft of the manuscript. K.N. and T.H. assisted in the preparation of the manuscript. A.K. supervised the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-08356-5

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017