OXFORD FEMS

FEMS Yeast Research, 2023, **23**, 1–9

DOI: 10.1093/femsyr/foad002
Advance access publication date: 14 January 2023
Protocol

pCEC-red: a new vector for easier and faster CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*

Letizia Maestroni†, Pietro Butti†, Vittorio Giorgio Senatore†, Paola Branduardi*

Department of Biotechnology and Biosciences, University of Milano Bicocca, Piazza della Scienza 2, 20126 Milan, Italy
***Corresponding author:** Department of Biotechnology and Biosciences, University of Milano Bicocca, Piazza della Scienza 2, 20126 Milan, Italy. Tel: +39 0264483418;E-mail: paola.branduardi@unimib.it
†L.M., P.B., and V.G.S. contributed equally to this work.
Editor: John Morrissey

Abstract

CRISPR-Cas9 technology is widely used for precise and specific editing of *Saccharomyces cerevisiae* genome to obtain marker-free engineered hosts. Targeted double-strand breaks are controlled by a guide RNA (gRNA), a chimeric RNA containing a structural segment for Cas9 binding and a 20-mer guide sequence that hybridises to the genomic DNA target. Introducing the 20-mer guide sequence into gRNA expression vectors often requires complex, time-consuming, and/or expensive cloning procedures. We present a new plasmid for CRISPR-Cas9 genome editing in *S. cerevisiae*, pCEC-red. This tool allows to (i) transform yeast with both Cas9 and gRNA expression cassettes in a single plasmid and (ii) insert the 20-mer sequence in the plasmid with high efficiency, thanks to Golden Gate Assembly and (iii) a red chromoprotein-based screening to speed up the selection of correct plasmids. We tested genome-editing efficiency of pCEC-red by targeting the *ADE2* gene. We chose three different 20-mer targets and designed two types of repair fragments to test pCEC-red for precision editing and for large DNA region replacement procedures. We obtained high efficiencies (∼90%) for both engineering procedures, suggesting that the pCEC system can be used for fast and reliable marker-free genome editing.

Keywords: CRISPR-Cas9, genome editing, *Saccharomyces cerevisiae*, new plasmid, guide RNA, *ADE2* deletion

Introduction

Synthetic biology is characterized by the development of new biological components or the manipulation of existing ones, thanks to the design and construction of core units, like parts of enzymes, genetic circuits and metabolic pathways, in a fast, scalable, and predictable way (Nielsen et al. 2022). Genome-editing technologies have become a central point in genetic manipulation strategies aimed at engineering microbial host metabolism for cell factory construction. Many genome-editing technologies have been developed during the last decades, from the Cre/LoxP system (Sauer 1987), through the homing endonuclease I-SceI (Bellaiche et al. 1999), to zinc finger nucleases (ZFNs) (Urnov et al. 2010), and transcription activator-like effector nucleases (TALENs) (Sun and Zhao 2013), there was an increased accuracy in targeting the desired modification. The last step forward in the field was made possible by the discovery of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), and the exploitation of the Cas endonucleases, together with all the modulations of the system that are nowadays available. CRISPR-Cas9 technology greatly simplified the possibility to introduce a precisely targeted double-strand break (DSB) — and consequently, the desired edit — to any genomic locus of interest, accelerating the exploitation of genome-editing technologies.

Saccharomyces cerevisiae is one of the most used cell factories in industrial biotechnology, thanks to its easy manipulation: its genetics is well established, and many genome and metabolic engineering tools are available. Thanks to its GRAS or QPS (Generally Regarded As Safe, or Qualified Presumption of Safety) status (US Food & Drug Administration — FDA, and European Food Safety Authority — EFSA), this yeast has been widely used to produce chemicals, fuels, and pharmaceuticals (Nielsen 2019, Madhavan et al. 2021, Zhang et al. 2021), and it is one of the first organisms in which CRISPR-Cas9 genome editing was successfully demonstrated (Dicarlo et al. 2013).

The standard CRISPR-Cas9 system in *S. cerevisiae* generally requires a system of two different plasmids, one carrying the Cas9 coding sequence, and the other, with the gRNA sequence to target the DSB to the desired genome locus. The use of a two-plasmid system increases the complexity of the overall system, requiring an initial double step of transformation (first with the Cas9 expression vector, then with the gRNA helper vector and the integration cassette). Moreover, the maintenance of both plasmids into transformant yeast cells requires the addition of two different antibiotics in the growth media to maintain the selective conditions, or the use of a yeast background with >1 auxotrophies, in case of auxotrophic markers. Even though there are many examples in literature of recipient plasmids, the limiting factor is always the cloning of the 20-mer guide sequence into the sgRNA expression cassette. The most common approaches in literature are either based on PCR (Dicarlo et al. 2013, Stovicek et al. 2015) or standard restriction-cloning procedures (Laughery et al. 2015, Lee et al. 2015); other approaches exploit part-assemblies, like Gibson Assembly (Apel et al. 2017), Golden Gate Assembly (Bao et al. 2015), or USER cloning (Jakočiūnas et al. 2015, Ronda et al. 2015). However, in most cases, the presented methods require labour-intensive work and/or wide use of expensive lab materials. For

Received 25 October 2022; **revised** 23 December 2022; **accepted** 13 January 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

instance, the pCRCT plasmid (Bao et al. 2015) relies on the external addition of the chromogenic compound X-Gal for convenient screening of gRNA insertion, while pTAJAK plasmids (Jakočiūnas et al. 2015, Ronda et al. 2015) require the use of expensive uracil-containing primers for USER cloning.

In this work, we describe the development of a new strategy to overcome the current limitations. We developed the pCEC-red plasmid (plasmid for CRISPR-Cas9 genome-editing in *S. cerevisiae*), which allows the expression of both Cas9 and gRNA. The advantages to this new system are three. First, the pCEC-red is a single vector harbouring information for Cas9 and gRNA expression and exploiting the sole KanR cassette for conferring resistance to two antibiotics: this simplifies the system, compared to most of the currently available systems, avoiding the use of different antibiotics or the need for multiple auxotrophies. Second, the insertion of the 20-mer guide sequence into the sgRNA expression cassette is mediated by a Golden Gate Assembly reaction for high-efficiency cloning and, lastly, a chromoprotein-based screening was introduced for easy selection of positive clones. The plasmid is available to the community as *Addgene plasmid #196040*.

As a proof-of-concept, the pCEC-red vector was tested by targeting the *ADE2* gene in *S. cerevisiae*. Indeed, the ade2Δ phenotype can be easily recognized by the red colour of the colonies, since the mutant cells, deprived of adenine, accumulate red purine precursors in the vacuole (Ugolini and Bruschi 1996). Three different targets were selected inside the *ADE2* coding sequence in order to have low off-target cut probabilities. All the targets showed comparable editing efficiencies, confirming the reliability and reproducibility granted by the pCEC system. Moreover, two different repair fragments were designed to obtain (i) precision gene editing (e.g. the final desired modification is a single-base mutation or the addition of a stop codon) and (ii) long region replacement (e.g. deletion or substitution of a genomic region of interest).

The aim of the study was to develop a novel single-plasmid system for Cas9 genome editing and gRNA expression, which allows a fast, precise, and highly efficient cloning of the 20-mer sequence in the vector backbone thanks to Golden Gate Assembly and a chromoprotein-based screening for positive clones selection.

Materials and methods

Strains

The *S. cerevisiae* parental strain used in this study was CEN.PK 113-7D (*MATa*; *HIS3*; *LEU2*; *URA3*; *TRP1*; *MAL2-8c*; *SUC2* — Dr P. Kötter, Institute of Microbiology, Johann Wolfgang Goethe-University, Frankfurt, Germany) (van Dijken et al. 2000). *Escherichia coli* strain DH5α was used to clone, propagate, and store the plasmids.

Media and growth conditions

Escherichia coli strains were stored in cryotubes at −80 °C in 50% glycerol (v v^{-1}) and grown in lysogeny broth (LB) medium (10 g L^{-1} NaCl, 10 g L^{-1} peptone, 5 g L^{-1} yeast extract) or terrific broth (TB) media (20 g L^{-1} peptone, 24 g L^{-1} yeast extract, 4 mL L^{-1} glycerol, 0.17 M KH_2PO_4, 0.72 M K_2HPO_4). When needed, the medium was supplemented with 100 μg mL^{-1} ampicillin or 50 μg mL^{-1} kanamycin.

Saccharomyces cerevisiae strains were stored in cryotubes at −80 °C in 20% glycerol (v v^{-1}) and grown on YPD medium (20 g L^{-1} glucose, 20 g L^{-1} peptone, 10 g L^{-1} yeast extract). When needed, the medium was supplemented with antibiotics G418 (200 mg L^{-1}) or nourseothricin (clonNAT) (100 mg L^{-1}), and/or adenine (60 mg L^{-1}).

Agar plates were prepared with the addition of 20 g L^{-1} agar to the liquid media. Yeast extract was provided by Biolife Italiana S.r.l., Milan, Italy. All the other reagents were provided by Sigma-Aldrich Co., St Louis, MO, USA. Each experiment was repeated at least three times. All yeast strains were grown at 30°C in an orbital shaker at 160 rpm and the ratio of tube/flask volume:medium was 5:1, while *E. coli* was grown at 37°C on an orbital shaker at 160 rpm.

pCEC-red: plasmid construction

All primers and plasmids used in this work are listed in Tables S1 and S2, respectively. The pCEC-red plasmid (Fig. 1) was generated by assembling eight different fragments obtained from eight different PCR reactions with a Golden Gate Assembly reaction, using T4 ligase and Esp3I as Type IIS restriction enzyme. Fragments one, two, three, five, six, and seven were obtained from pML104 plasmid (Laughery et al. 2015) using the primers listed in Table S1, 1–6 and 9–14. Fragments one, two, and three contain the Cas9 cassette (*THD3*p promoter, codon-optimized *Streptococcus pyogenes* Cas9, *ADH1*t); fragments were designed to domesticate and reconstitute the wt protein sequence: G171A (silent mutation) was needed to remove a BsaI site; the mutation A1836T (N612K) present in the original copy from pML104 was reverted. The RFP coding sequence (fragment four) was PCR-amplified with primers 7 and 8 from GGE114 plasmid, a gift from Macarena Larroude (Larroude et al. 2019) (*Addgene plasmid #120731*), while the kanMX coding sequence (fragment eight) was amplified with primers 15 and 16 from pGA-kanMX plasmid. The pGA-kanMX plasmid was obtained with a Golden Gate Assembly reaction with T4 ligase and Esp3I as Type IIS restriction enzyme using three PCR fragments; the first fragment was amplified from pGA-red-maxi plasmid (Fig. S1b, *Addgene plasmid #196337*) with primers 17 and 18, while the other two fragments were obtained using pZ_3 plasmid as template (Branduardi et al. 2004) and the primer couples (1) 19 and 20, and (2) 21 and 20.

Golden Gate Assembly procedures performed in this work followed the protocol optimized and described in a previous work (Maestroni et al. 2023). Q5® high-fidelity DNA Polymerase from NEB was used on a ProFlex PCR System (Life technologies) following NEB manual. All enzymes used were purchased from New England Biolabs (NEB).

gRNA cloning protocol

Synthetic DNA sequences carrying the 20-mer sequence were designed in order to contain BsaI recognition sites at both ends, the correct protruding sequences, the desired 20-mer sequence, and part of the scaffold gRNA (sgRNA) sequence (Fig. 2A). On the top strand synthetic oligo, the sequences of protruding ends are 5′-GATC-3′ and 5′-AAAT-3′, while the sgRNA sequence is 5′-GTTTTAGAGCTAG-3′. All the synthetic DNA sequences of the top-strand oligos used in this work are listed in Table S3. The bottom-strand oligos are the reverse-complementary of the previously described synthetic DNA sequence. The top and bottom strand oligos were annealed with a protocol adapted from Thermo Fisher (tools.thermofisher.com/content/sfs/brochures/TR0045-Anneal-oligos.pdf) and OpenWetWare (openwetware.org/wiki/PrbbBB:Oligo_Annealing). The annealing mix was prepared in a PCR tube adding: 5 μL of 10 μ top-strand and bottom-strand oligos, 5 μL of annealing buffer (0.1 M Tris pH 8, 10 mM EDTA pH 8, 0.5 M NaCl), and water up to 50 μL. The incubation was performed in a ProFlex PCR System (Life technologies) with the following cycles protocol: 95 °C for 5′, −1°C $cycle^{-1}$ for 1′ (until 25 °C, 70 cycles), 4 °C hold.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

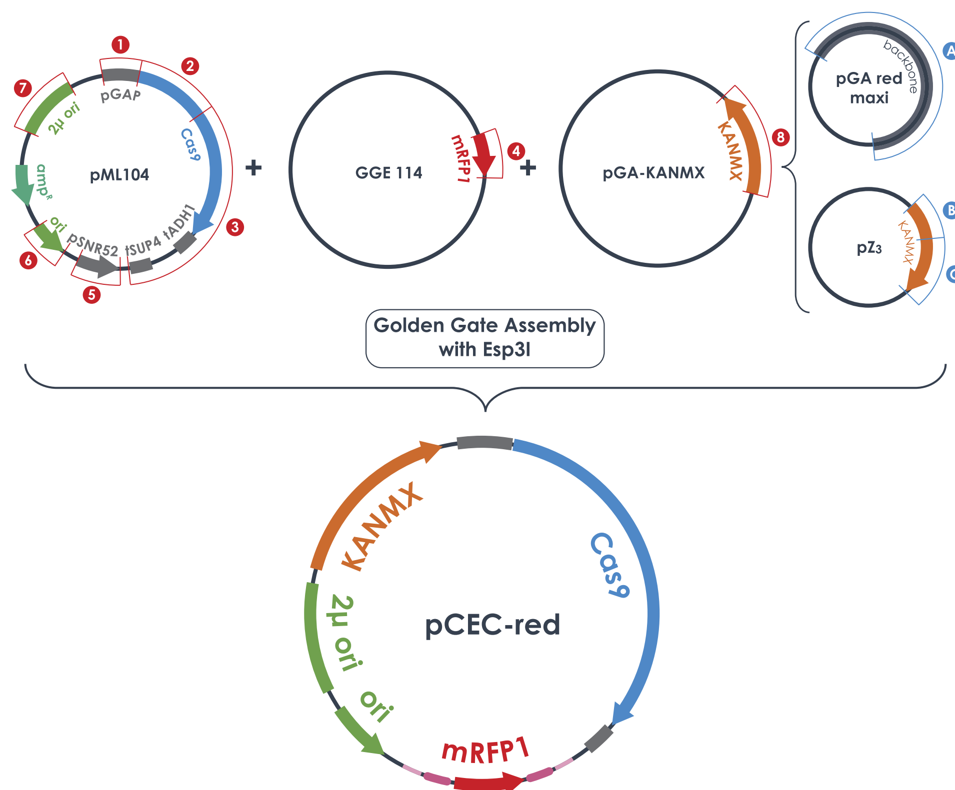

Figure 1. pCEC-red plasmid construction. pCEC-red plasmid was built starting from three different plasmids: pML104, GGE114, and pGA-kanMX. In particular, eight different fragments were obtained from these plasmids. Fragments 1 (pGAP + Cas9_1), 2 (Cas9_2), 3 (Cas9_3 + gRNA terminator), 5 (gRNA promoter), 6 (ori), and 7 (2μ ori) were amplified from pML104 (Laughery et al. 2015), while fragment 4 (mRFP1 flanked with BsaI recognition sites) from GGE 114 (*Addgene plasmid #120731*) and fragment 8 (kanMX) from pGA-kanMX (this work, see the 'Materials and methods' section for more details). All fragments were amplified in order to carry the Esp3I restriction enzyme recognition sites at their 5′ and 3′ ends. The eight parts obtained have been used as substrates of a Golden Gate reaction carried out with the addition of Esp3I enzyme and resulting in the construction of the pCEC-red plasmid. The assembly product has been transformed into DH5α *E. coli* cells for amplification. As additional control, one red colony has been checked by colony PCR and further confirmed by sequencing.

The obtained annealed oligos were cloned into pCEC-red plasmid exploiting Golden Gate Assembly reactions with T4 ligase and BsaI as Type IIS restriction enzyme. All Golden Gate Assembly reactions were performed thanks to the optimized protocol reported in a previous work (Maestroni et al. 2023). Transformants were plated in the presence of kanamycin and selected thanks to the red/white screening. Positive clones were verified by colony PCRs performed with appropriate primers (22 and 23, Table S1), then sequenced with primer number 22, Table S1.

Genome-editing efficiency evaluation by *ADE2* targeting

All 20-mer sequences used in this work and the whole gRNA sequences used to obtain the final fragments to insert into pCEC-red plasmid are listed in Tables 1 and S3, respectively. Three 20-mer sequences (B1, P1, and S1) were selected in order to have low off-target cut probabilities. B1 and S1 targets were selected from two previous works from literature (Bao et al. 2015, Stovicek et al. 2015), while P1 20-mer sequence was designed with Benchling online tool 'CRISPR Guide RNA Design' (https://www.benchling.com/crispr). These sequences were cloned in the pCEC-red plasmid following the gRNA cloning procedure, obtaining plasmids pCEC-gADE2-B1, pCEC-gADE2-P1, and pCEC-gADE2-S1, respectively.

The repair fragments for precision editing were designed and synthesized by Twist Bioscience as a single synthetic DNA sequence (Table S3). The whole synthetic sequence was cloned in the pGA-red-mini plasmid (Fig. S1a, *Addgene number #196338*) thanks to a Golden Gate Assembly reaction with BsaI restriction enzyme and T4 DNA ligase, obtaining the plasmid pRF_ScADE2_60H_BRC. The final single precision editing repair fragments were obtained by PCR using the corresponding primer couples: 24 and 25 for P1 repair fragment, 26 and 27 for S1 repair fragment, 28 and 29 for B1 repair fragment.

The repair fragment for large DNA regions replacement was built thanks to a Golden Gate Assembly between three PCR fragments and the pGA-red-maxi acceptor vector (Fig. S1b, *Addgene number #196337*). The first fragment containing about 500 bp upstream to *ADE2* was amplified from CEN.PK 113-7D genomic DNA with primers 32 and 33; the second fragment containing the NATMX cassette was amplified from pCfB3041 plasmid (Jessop-Fabre et al. 2016) with primers 34 and 35; and the third fragment containing about 500 bp downstream to *ADE2* was amplified from CEN.PK 113-7D genomic DNA with primers 36 and 37. All primers contained BsaI recognition sites and the correct protruding ends to form the final repair fragment into pGA-red-maxi acceptor vector (Fig. S1b, *Addgene number #196337*), obtaining pRF_ScADE2_LH_NAT plasmid. The final repair fragment for large DNA regions replacement was obtained by PCR using primers 32 and 37.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

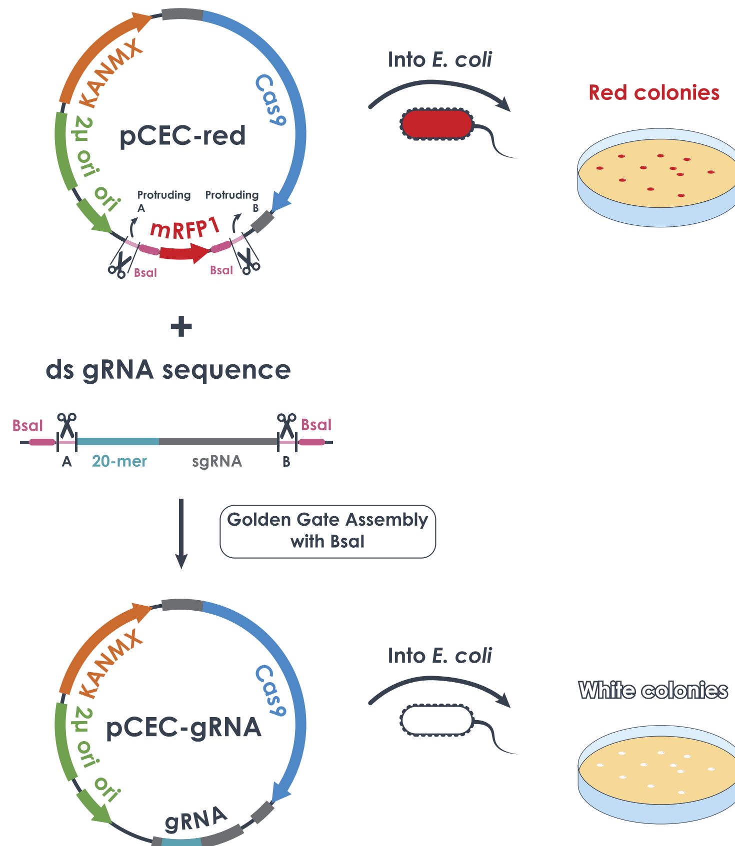

Figure 2. gRNA design and construction of pCEC-gRNA plasmids. (A) gRNA cloning sequence consists in two random sequences of 6 nt at both ends, followed by BsaI type IIS restriction enzyme recognition sites in dark pink, A and B protruding sequences in light pink, the specific 20-mer sequence in blue, and the sgRNA sequence in grey. (B) pCEC plasmids containing the gRNA sequence of interest are created by cloning the ds-gRNA sequence obtained after the oligo annealing in the pCEC-red acceptor plasmid. The plasmid carries an *E. coli* mRFP expression cassette, allowing for a red/white screening system. The cloning is obtained by digesting pCEC-red and the gRNA sequence with BsaI sites, exploiting A and B protruding sequences for ligation. As result, BsaI cutting sites and mRFP1 are replaced with the gRNA of interest (white colonies).

Table 1. 20-mer sequences used in this study.

Name	20-mer sequence on top strand oligo (5′–3′)	References
B1	GATATCAAGAGGATTGGAAA	Bao et al. (2015)
P1	AGTTACCCAAAGTGTTCCTG	This work
S1	AATTGTAGAGACTATCCACA	Stovicek et al. (2015)

Yeast transformants were obtained exploiting the constructs created in this work, while the transformation procedure was adapted from a previous work from literature (Gietz and Woods 2002). In particular, the transformation mix was prepared adding 100 ng (18 fmol) of the needed specific pCEC-gADE2 plasmid to the transformation mixture (with the correct gRNA in it), with or without a ten-fold molar quantity (180 fmol, corresponding to 13.39 ng for precision editing, or 244.5 ng for large DNA region replacement) of the repair fragment of interest. An equimolar quantity of empty pCEC-red plasmid (110 ng) was used as a positive control for transformation efficiency. The recovery time was increased to 3 hours in YPD medium with the addition of 60 mg L^{-1} of adenine. For each transformation, cells were diluted 1:50 and 49:50, plated onto two different YPD + G418 plates, and incubated for 5 days at 30 °C. After the incubation time, red and white colonies were counted to

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

establish transformation and *ADE2* disruption/deletion efficiencies.

To further confirm the result, up to 50 colonies for each transformation (40 red colonies and 10 white colonies, or less where not present) were restreaked on nonselective YPD plates. After 5 days of incubation, the red/white phenotype was confirmed and the correct integration of precision editing repair fragments into the genome was verified by colony PCR using primers 30 and 31 (Table S2); the integration of NatMX expression cassette was verified by restreaking on YPD + clonNAT plates.

Once positive clones were obtained and verified, pCEC plasmid was removed with the following curing protocols: a single colony was inoculated in 5 mL YPD at 30 °C, 160 rpm overnight. Cells were streaked on a YPD plate to obtain single colonies and incubated at 30 °C for 2 days. To verify the gRNA helper vector loss, single colonies were grown overnight in two different media: YPD with no selection and YPD with G418: cells without pCEC plasmid will not be able to grow on media with G418.

Colony PCRs

To perform colony PCRs on *E. coli*, at least five different colonies were picked for each transformation plate and dissolved (i) in 20 μL of growth media with the proper antibiotic as a colony back-up and (ii) into the PCR tube with the appropriate PCR mix. To boost cell disruption, the initial denaturation step must lasts at least 5 minutes. The positive *E. coli* clones are then inoculated starting from the 20 μL liquid cultures prepared at the beginning.

To perform colony PCRs on *S. cerevisiae* colonies, genomic DNA was extracted in 96-well plates, optimizing the LiOAc-SDS procedure of Lõoke et al. (2011). Briefly, a small amount of biomass was taken from each reastreak and resuspended in different wells filled with 50 μL of a 200 mM LiOAc, 1% SDS solution. The plate was incubated for 5 minutes at 70 °C and 150 μL of ethanol 96% were added in each well. After a centrifugation step at 3220 *g* for 5 minutes, each well was washed with 200 μL of ethanol 70% at −20 °C. Finally, the pellets obtained by a second round of centrifugation in the same conditions were resuspended in 50 μL of Tris-EDTA buffer. After a third step of centrifugation, 5 μL of the supernatant (containing genomic DNA) were used as PCR template.

Wonder Taq DNA polymerase (Euroclone) was used on a ProFlex PCR System (Life technologies) to perform colony PCR reactions.

Results and discussion

pCEC-red: a new vector for single-plasmid CRISPR–Cas9 genome editing

Cloning of the 20-mer guide sequence into its expression vector is often the low-efficiency step for CRISPR-Cas9 genome editing, requiring time-consuming and expensive cloning steps. This step is not strictly related to the CRISPR-Cas9 genome editing procedure, but it is an issue strongly connected to the design of the currently available vectors. For example, there are vectors where the cloning of the 20-mer guide sequence into a sgRNA expression cassette requires performing a single-step PCR amplification of the whole vector, including the Cas9 coding sequence (Stovicek et al. 2015). This method is susceptible to possible mutations arising from the amplification step, it is time and cost-consuming, and given the large size of the Cas9 coding sequence, the efficiency is low. Other vectors currently available exploit cloning methods. For example, there are vectors where the cloning of the 20-mer guide sequence into a sgRNA expression cassette requires performing a single-step PCR amplification of the whole vector, including the Cas9 coding sequence (Stovicek et al. 2015). This method is time and cost-consuming, and given the large size of the Cas9 coding sequence, the PCR amplification is error-prone and the efficiency is low. Other vectors currently available exploit traditional cloning methods. For example, in the works of Laugherey and colleagues, and Lee and colleagues, the restriction and ligation procedures exploit the use of type II restriction enzymes (Laughery et al. 2015, Lee et al. 2015): digestion and ligation have low efficiency, particularly when cloning a fragment of only 20 bp in a backbone of >10 kb. More advanced and efficient cloning methods can be exploited, and are used in other works present in literature (USER cloning, Jakočiūnas et al. 2015, Ronda et al. 2015; and Gibson assembly, Apel et al. 2017). However, these strategies require a complex and labour-intensive step of *in-silico* design or more expensive materials. For example, Gibson Assembly requires long synthetic homology regions for the insertion of the gRNA into the final expression plasmids; USER cloning requires long uracil-containing primers, which are generally much more expensive than regular primers.

In this scenario, we designed and built a new vector for CRISPR-Cas9 genome editing with a more efficient and easier procedure to design and insert the desired 20-mer guide sequence, based on Golden Gate Assembly and a quick red/white screening. The higher cloning efficiency allowed us to develop the pCEC single-plasmid system, based on a 10.373 bp empty Cas9 expression vector called pCEC-red ready for the insertion of a user-defined gRNA sequence (Figs 1 and 2B). The vector is called pCEC, which stands for plasmid for CRISPR-Cas9 genome editing in *S. cerevisiae*, while the word 'red' refers to the selection method to screen positive clones with the correct insertion of the 20-mer sequence. The pCEC-red plasmid is publicly available from the *Addgene* repository as *Addgene plasmid #196040*.

Our pCEC-red plasmid is an upgraded version of the previously presented vector pML104 (Laughery et al. 2015). pCEC-red presents KanMX instead of *URA3*, which can be used as a dominant marker both in *E. coli* and *S. cerevisiae*, avoiding the need of an additional bacterial resistance sequence on the plasmid and the need of an auxotrophic yeast strain, making the plasmid functional in all genotypic backgrounds. To limit the final size of the plasmid, we removed the AmpR expression cassette and other sequences with no relevant utility. The final vector carries the Cas9 expression cassette, the KanMX expression cassette, the 2μ plasmid replication origin, the origin of replication for *E. coli*, and the bacterial expression cassette of mRFP1 chromoprotein inserted between the *SNR52* promoter and *SUP4* terminator (see the 'Materials and methods' section and Fig. 1 for more details). In the final plasmid, the Cas9 coding sequence is under the control of the *TDH3* promoter, a strong constitutive promoter (Peng et al. 2015); however, it is known that strong expression of Cas9 causes toxicity (Generoso et al. 2016): here, we speculate that a not-in-frame start codon a few bases upstream of Cas9 ATG (present in the original pML104 vector as well) reduces its translation efficiencies. This probably lowers Cas9 expression levels, and thus alleviates its toxicity (see Fig. S2 for more details).

The new designed pCEC-red vector exploits Golden Gate Assembly combined with a coloured screening method leading to a time and cost-saving procedure. Indeed, the empty vector generates red *E. coli* colonies (Fig. 2B). The cloning strategy is based on the substitution of the mRFP1 *E. coli* expression cassette with the gRNA sequence by a Golden Gate Assembly reaction. This

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

is achieved by the presence of BsaI Type IIS restriction enzyme recognition sites at both ends of the mRFP1 cassette.

When considering the pML104 vector (Laughery et al. 2015), our system is more efficient, because the former is based on a traditional cloning procedure. Most importantly, the previous cloning procedure exploited BclI as restriction enzyme, which is sensitive to Dam methylation: this requires working with *dam*$^{-}$ *E. coli* strains. The present strategy exploits the Golden Gate Assembly approach with BsaI type IIS restriction enzyme to cut out the mRFP expression cassette and generate two protruding ends, A — ATTT and B — GATC (Fig. 2A), that can be used to insert any 20-mer sequence of interest in the pCEC backbone.

The desired gRNA is obtained by designing two specific oligos of ∼70 bp, one the reverse complement of the other, which prior to Golden Gate Assembly are annealed to each other to obtain a dsDNA sequence; the oligos are designed to carry the 20-mer guide sequence and the sgRNA sequence flanked by BsaI recognition sites, leaving the protruding sequences A and B. Figure 2A shows the structure of the final dsDNA sequence for Golden Gate Assembly in the pCEC-red acceptor vector. The screening procedure to select *E. coli* positive clones is based on the loss of the mRFP1 chromoprotein: red clones are considered as negative, while the positive clones will appear white (Fig. 2B).

CRISPR-Cas9 genome editing of *ADE2* using pCEC-red vector

Plasmid construction and fragments preparation

We tested the genome-editing efficiency of pCEC-red vector by targeting the yeast *ADE2* gene. For this purpose, three different 20-mer sequences were selected to target *ADE2* genomic coding sequence (Table 1): P1 20-mer sequence was designed with Benchling online tool 'CRISPR Guide RNA Design' (https://www.benchling.com/crispr); the target is positioned 352 bp from *ADE2* start codon; B1 and S1 were selected from two previous works from literature (Bao et al. 2015, Stovicek et al. 2015), and are respectively positioned at 157 and 623 bp from the start codon. The three sequences described were designed and synthesized as top and bottom stranded oligos of ∼70 bp, in a sequence containing BsaI recognition sites at both ends, two specific protruding ends called A and B, part of the sgRNA sequence and the specific 20-mer sequence. The general structure of the gRNA-bearing oligonucleotide is reported in Fig. 2A, while the specific sequence of each gRNA is reported in Table S3. The top and bottom strands are annealed thanks to a specific hybridization procedure (see the 'Materials and methods' section) and then cloned into pCEC-red plasmid thanks to a Golden Gate Assembly reaction with BsaI as type IIS restriction enzyme. This way, we obtained three different final vectors (pCEC-gADE2-B1, pCEC-gADE2-P1, and pCEC-gADE2-S1), each containing Cas9 expression cassette, and a different gRNA (B1, P1, or S1, respectively) targeting a different region of the *ADE2* coding sequence (Fig. 3).

The three pCEC-gADE2 plasmids were tested to calculate the efficiency of (a) precision gene editing (e.g. the final desired modification is a single-base mutation or the addition of a stop codon) and (b) long region replacement (e.g. deletion or substitution of a genomic region of interest). For this purpose, two different kinds of repair fragments were designed and built.

To check the efficiency of the precision gene editing, the repair fragment was designed to have short homology regions before and after the target sequence, and thus it was called 'short repair fragment', or RF_SH. The RF_SH was designed as follows: 50 bp homology region before the PAM, a 19 bp barcode (5′-TGACTGACTAGGCGAGTAC-3′), a random base, and 50 bp homology region after the PAM. The barcode sequence is an artificial sequence naturally absent in *S. cerevisiae* genome, designed to carry out a double function: simulate precision editing by the introduction of premature stop codons in *ade2* mutants, and allow easy verification of the integration by PCR. Indeed, the barcode was designed to incorporate three stop codons over the three possible frames (to ensure a translation interruption independently from the site of insertion), followed by a set of bases to allow the annealing of a specific primer. The additional bases were chosen to reduce the homology with other existing native sequences. This feature allows to easily verify the presence of the repair fragment in the predicted insertion site, while maintaining the inserted sequence as short as possible to simulate precision editing. A total length of 120 bp for RF_SH was chosen to mimic the length of a synthetic oligonucleotide sequence.

As the target sequences selected in this work were three (B1, P1, and S1), three different RF_SH were designed, each with different homology regions specific to the target: RF_ScADE2-B1_SH, RF_ScADE2-P1_SH, and RF_ScADE2-S1_SH (see Table S3 for sequence details). For easier handling, the repair fragments were clustered in a single synthetic DNA fragment (RF_ScADE2_60H_BRC) with BsaI sites on both ends to clone it in pGA-red-mini plasmid (Fig. S1b, *Addgene number* #196338), obtaining pRF_ScADE2_60H_BRC plasmid. The single repair fragments were then amplified by PCR using this plasmid as a template (see the 'Materials and methods' section for primer details).

For long region replacement, the repair fragment was designed to have a sequence of ∼500 bp upstream to *ADE2* ORF, finishing at −280 bp from ATG, a NatMX cassette, and ∼500 bp downstream of *ADE2* ORF (see Table S3 for sequence details). The three parts were PCR amplified and assembled in pGA-red-maxi (Fig. S1b, *Addgene number* #196337) with Golden Gate Assembly and BsaI as type IIS restriction enzyme, obtaining the pRF_ScADE2_LH_NAT plasmid. The region was PCR-amplified (primers number 32 and 37, Table S1) and the amplicon (2192 bp) was called 'long repair fragment', or RF_*Sc*ADE2_LH_NAT, as it has 10 times longer homology regions to promote homologous recombination (HR) far from the cutting site, and a NatMX cassette to allow the easy identification of clones in which the *ADE2* ORF was substituted with the fragment of interest.

Quali-quantitative analysis of S. cerevisiae transformants

Once the final pCEC plasmids bearing the gRNAs and the repair fragments were prepared, we proceeded with *S. cerevisiae* transformations. For each target (B1, P1, and S1), we evaluated (a) the efficiency of non-homologous end joining (NHEJ) by only adding to the transformation mix the pCEC plasmid, (b) the efficiency of precision gene editing by adding to the transformation mix the corresponding pCEC plasmid (pCEC-gADE2-B1, pCEC-gADE2-P1, or pCEC-gADE2-S1) and a short repair fragment (RF_ScADE2-B1_SH, RF_ScADE2-P1_SH, or RF_ScADE2-S1_SH, respectively), and (c) the efficiency of long region deletion/substitution by adding to the transformation mix the corresponding pCEC plasmid (pCEC-gADE2-B1, pCEC-gADE2-P1, or pCEC-gADE2-S1) and the long repair fragment (RF_*Sc*ADE2_LH_NAT). Figure 3 shows a schematic representation of the combinations of pCEC plasmids and repair fragments and the expected results from each experiment.

Plates were incubated until the appearance of small, red colonies. The observed slow growth phenotype is probably due to the combination of G418 addition and the mutation in *ADE2*. Nevertheless, red and white colonies were visible and counted after 5

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

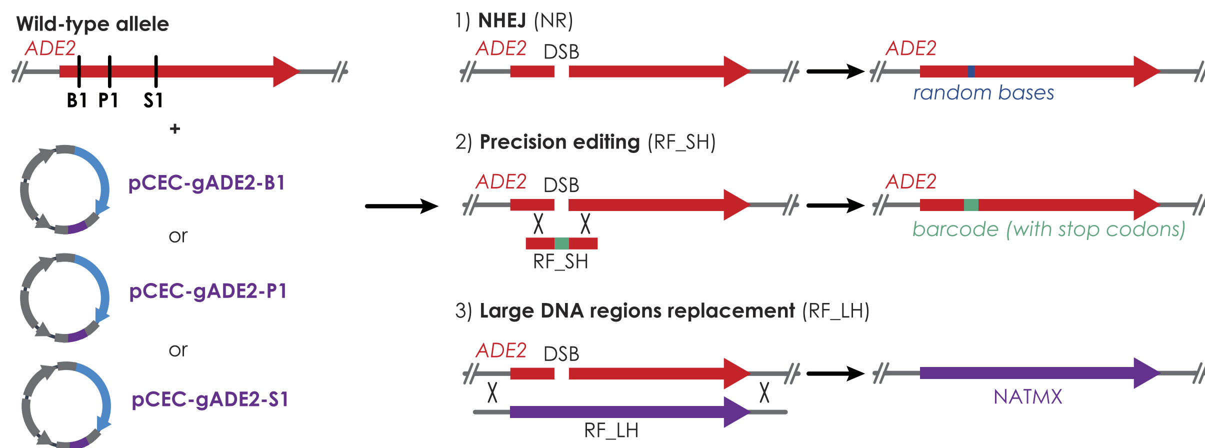

Figure 3. CRISPR-Cas9 genome editing of *ADE2* using pCEC-red vectors: summary outline. Three different 20-mer sequences were selected to target *ADE2* coding sequence, leading to the construction of three different pCEC-gRNAs plasmids: pCEC-gADE2-P1, pCEC-gADE2-S1, and pCEC-gADE2-B1 (on the top). The three plasmids were combined with two different repair fragments to test pCEC-red vector and its efficiency in (i) precision gene-editing (RF_SH) and (ii) long region replacement (RF_LH). Moreover, the different pCEC-gRNAs plasmids were also tested in the absence of a repair fragment, to verify the NHEJ efficiency in repairing DSBs. All the three pCEC gRNAs plasmids were then individually used to transform *S. cerevisiae* with (i) no repair fragment, (ii) its specific RF_SH repair fragment, comprising a specific DNA sequence called barcode, and (iii) the RF_LH repair fragment, containing NatMX expression cassette. After the transformation, white and red colonies were analysed for checking and quantifying the efficiency of the expected result.

days of incubation 30°C to calculate the transformation efficiencies.

Single colonies (up to 40 red colonies and 10 white colonies for each transformation) were restreaked on fresh YPD plates without selection to confirm the phenotype.

For the three transformations with short repair fragments (one for each of the three different targets — B1, P1, and S1), genomic DNA was extracted from 15 out of the 40 red colonies restreaked. The genomic-DNA extraction protocol from Lõoke and colleagues (Lõoke et al. 2011) was optimized to obtain a high-throughput economic and less time-consuming method of extraction in 96-well plates (see the 'Materials and methods' section for more details). Successful editing was confirmed by colony PCR exploiting a couple of primers specific for the barcode sequence and the *ADE2* ORF (outside the homology region). Therefore, the addition of the barcode made the verification procedure even easier.

For the transformations with the long repair fragment, the restreaked colonies (up to 40 red colonies and 10 white colonies) were restreaked once again on YPD + clonNAT and the correct insertion of the NatMX was confirmed by the ability of the colonies to grow in the presence of the antibiotic.

For transformations without any repair fragments, the NHEJ efficiency was calculated on the base of white/red colony rates onto transformation's plates.

Transformation results are shown in Fig. 4. Editing efficiencies were comparable between the different chosen targets (Fig. 4A), demonstrating the reliability and reproducibility granted by the pCEC system. Conversely, the ratio of red colonies over the total number of colonies largely depended on the type of repair fragment provided. The absence of an externally provided DNA fragment as a template to repair the DSB induced by Cas9 led to a very low number of colony forming units, when compared to a control where Cas9 was not guided by a gRNA (Table S4). Most of these colonies were white, confirming the toxicity of Cas9-gRNA complex and the low frequency of NHEJ as a DNA-damage repair mechanism in *S. cerevisiae*, as already described in literature (Laughery et al. 2015).

The addition of a repair fragment to the transformation mix led to a substantial increase in the number of total colonies, mainly related to an increase in the number of red-edited colonies, suggesting the activation of HR as a repair mechanism. Editing efficiencies reached overall values of 91.7% ± 5.49% for precision editing and 86.2% ± 7.95% for large DNA region replacement (Fig. 4B). PCR verification and restreaks confirmed that ~100% of red colonies integrated the provided fragment (Table S5), suggesting that the pCEC system can be used for marker-free genome editing.

Conclusions

In the present work, we demonstrated that redesigning a single expression vector might lead to advantages in terms of inserting the 20-mer sequences into the gRNA expression cassette. pCEC-red not only allows simpler, more rapid, and less expensive procedures, but also gives the possibility to have both Cas9 and gRNA expression cassettes in a single plasmid. The final plasmid with the target-specific 20-mer sequence is obtained by a Golden Gate Assembly reaction between pCEC-red plasmid and *ad hoc*-designed annealed oligos with BsaI as type IIS restriction enzyme; *E. coli* positive clones are selected thanks to a coloured red/white screening as a result of the exploitation of mRFP1 red chromoprotein.

The chosen target of genome editing was the *ADE2* gene, which allowed us to test different possible applications of pCEC-red plasmid, proving its effectiveness both in precision editing and large DNA region replacement. In addition, the pCEC-system could be combined and integrated with existing synthetic biology tools, such as the EasyClone-MarkerFree (Jessop-Fabre et al. 2016) or the MoClo-YTK and its evolutions (Lee et al. 2015; Otto et al. 2021). Thanks to the principle of modularity on which synthetic biology tools rely, the pCEC-system allows a transition to a simpler single-plasmid CRISPR/Cas9 system, when the complexity of the traditionally used two-plasmids systems is not required.

Moreover, we also demonstrated that the repair fragments can be designed *ad hoc* to simplify screening procedures for positive

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

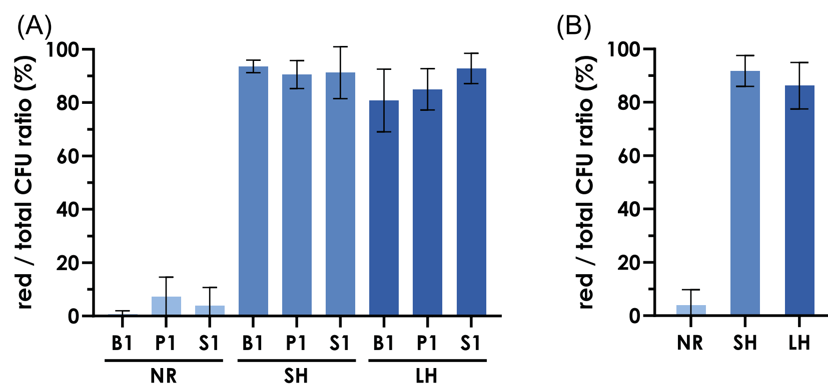

Figure 4. pCEC-gADE2 vectors transformation results. Transformation efficiencies in gene editing of the three different pCEC-gADE2 plasmids, which target three different locations on *ADE2* gene (B1, P1, and S1), are reported in panel (A). Each of the three plasmids was tested for the occurrence of NHEJ as a repair mechanism (NR pale grey columns), for precision gene editing (SH grey columns) and large DNA region replacement (LH black columns). Transformation efficiency was calculated on the basis of white/red colony rates onto transformation's plates. Data shown are representative of three independent experiments. (B) The red/white phenotype was confirmed and the correct integration of precision editing repair fragments into the genome was verified by colony PCR or by restreaks on YPD + clonNAT plates. Data are shown as mean ± SD. Error bars correspond to standard deviation of triplicate samples.

genome-edited clones' confirmation. In our case, we added a short sequence of 20 nt in the repair fragment, which allowed us to insert stop codons, while at the same time creating a barcode, a unique DNA sequence for specific annealing of primers.

The availability of this new vector should help accelerate the adoption of this technology within the yeast scientific community and its integration with other existing toolkits.

Author contributions

L.M.: conceptualization, methodology, investigation, and writing; P.B.: conceptualization, methodology, validation, investigation, and writing; V.G.S.: methodology, validation, investigation, and writing; and P.B.: funding acquisition, project administration, supervision, writing, and revision.

Acknowledgements

The authors would like to thank Luca Mastella for fruitful discussions on the project.

Supplementary data

Supplementary data are available at *FEMSYR* online.

Conflict of interest. The authors declare that they have no known conflict of financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the University of Milano–Bicocca with the FA (Fondo di Ateneo) to P.B., the PhD fellowship of the University of Milano–Bicocca to L.M. and V.G.S., and the Milano–Bicocca–ALBINI GROUP co-financed PhD fellowship to P.B. This work was also partially supported by CARDoon valorization by InteGrAted biorefiNery (CARDIGAN) and by REconnecting PLAstics life cycle to biogeochemical cycles by sustainable hydrolysis and Yeasts fermentation (REPLAY) projects, funded by the Italian Ministry of Education, University and Research (PRIN 2017KBTK93 and PRIN 2020SBNHLH) to P.B.

References

Apel AR, D'Espaux L, Wehrs M, *et al.* A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2017;**45**:496–508.

Bao Z, Xiao H, Liang J, *et al.* Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth Biol* 2015;**4**:585–94.

Bellaiche Y, Mogila V, Perrimon N. I-SceI endonuclease, a new tool for studying DNA double-strand break repair mechanisms in Drosophila. *Genetics* 1999;**152**:1037.

Branduardi P, Valli M, Brambilla L, *et al.* The yeast *Zygosaccharomyces bailii*: a new host for heterologous protein production, secretion and for metabolic engineering applications. *FEMS Yeast Res* 2004;**4**:493–504.

Dicarlo JE, Norville JE, Mali P, *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 2013;**41**:4336–43.

Generoso WC, Gottardi M, Oreb M, *et al.* Simplified CRISPR-cas genome editing for *Saccharomyces cerevisiae*. *J Microbiol Methods* 2016;**127**:203–5.

Gietz DR, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 2002;**350**:87–96.

Jakočiūnas T, Bonde I, Herrgård M, *et al.* Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab Eng* 2015;**28**:213–22.

Jessop-Fabre MM, Jakočiūnas T, Stovicek V, *et al.* EasyClone-MarkerFree: a vector toolkit for marker-less integration of genes into *Saccharomyces cerevisiae* via CRISPR-Cas9. *Biotechnol J* 2016;**11**:1110–7.

Larroude M, Park YK, Soudier P, *et al.* A modular Golden Gate toolkit for Yarrowia lipolytica synthetic biology. *Microbial biotechnology* 2019;**12(6)**:1249–59.

Laughery MF, Hunter T, Brown A, *et al.* New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast* 2015;**32**:711–20.

Lee ME, DeLoache WC, Cervantes B, *et al.* A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synthetic Biology* 2015;**4**:975–86.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

Lõoke M, Kristjuhan K, Kristjuhan A. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* 2011;**50**:325–8.

Madhavan A, Arun KB, Sindhu R, *et al.* Customized yeast cell factories for biopharmaceuticals: from cell engineering to process scale up. *Microb Cell Fact* 2021;**20**, https://doi.org/10.1186/S12934-021-01617-Z

Maestroni L, Butti P, Milanesi R, *et al.* Easy Modular Integrative fuSion-ready Expression (Easy-MISE) Toolkit for Fast Engineering of Heterologous Productions in Saccharomyces cerevisiae. *ACS Synthetic Biology Apr 14* 2023; https://doi.org/10.1021/acssynbio.3c00015

Nielsen J, Tillegreen CB, Petranovic D. Innovation trends in industrial biotechnology. *Trends Biotechnol* 2022;**40**:1160–72.

Nielsen J. Yeast systems biology: model organism and cell factory. *Biotechnol J* 2019;**14**:1800421. https://doi.org/10.1002/BIOT.201800421

Otto M., Skrekas C., Gossing , *et al.* Expansion of the Yeast Modular Cloning Toolkit for CRISPR-Based Applications, Genomic Integrations and Combinatorial Libraries. *ACS Synthetic Biology* 2021;**10(12)**: 3461–3474.

Peng B, Williams TC, Henry M, *et al.* Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microb Cell Fact* 2015;**14**: 1–11.

Ronda C, Maury J, Jakočiunas T, *et al.* CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. *Microb Cell Fact* 2015;**14**:97. https://doi.org/10.1186/S12934-015-0288-3

Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* 1987;**7**:2087.

Stovicek V, Borodina I, Forster J. CRISPR-cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Metab Eng Commun* 2015;**2**:13–22.

Sun N, Zhao H. Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnol Bioeng* 2013;**110**: 1811–21.

Ugolini S, Bruschi CV. The red/white colony color assay in the yeast *Saccharomyces cerevisiae*: epistatic growth advantage of white ade8-18, ade2 cells over red ade2 cells. *Curr Genet* 1996;**30**: 485–92.

Urnov FD, Rebar EJ, Holmes MC, *et al.* Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* 2010;**11**:636–46.

van Dijken JP, Bauer J, Brambilla L, *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 2000;**26**: 706–14.

Zhang ZX, Wang LR, Xu YS, *et al.* Recent advances in the application of multiplex genome editing in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2021;**105**: 3873–82.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foad002/6987636 by guest on 25 March 2024

Received 25 October 2022; **revised** 23 December 2022; **accepted** 13 January 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For permissions, please e-mail: journals.permissions@oup.com