FEMS Yeast Research, 2025, **25**, foaf032
DOI: 10.1093/femsyr/foaf032
Advance access publication date: 18 June 2025
Protocol – Protocols, Tools and Methods

Genomically integrated cassettes swapping: bringing modularity to the strain level in *Saccharomyces cerevisiae*

Pietro Butti, Francesco Bellusci, Elisa Brambilla, 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. E-mail: paola.branduardi@unimib.it
Editor: [Zongbao ZHAO]

Abstract

A large variety of synthetic biology toolkits for the introduction of multiple expression cassettes is available for *Saccharomyces cerevisiae*. Unfortunately, none of these tools is designed to allow the modification — exchange or removal — of the cassettes already integrated into the genome in a standardized way. The application of the modularity principle therefore ends to the steps preceding the final host engineering, making microbial cell factories construction stiff and strictly sequential. In this work, we describe a system that easily allows CRISPR-mediated swapping or removal of previously integrated cassettes, thus bringing the modularity to the strain level, enhancing the possibility of modifying existing strains with a reduced number of steps. In the system, each cassette is tagged with specific barcodes, which can be used as targets for CRISPR nucleases (Cas9 and Cas12a), allowing the excision of the construct from the genome and its substitution with another expression cassette or the restoration of the wild type locus in one single standardized step. The system has been applied to the previously developed Easy-MISE toolkit and tested by swapping fluorescent protein expression cassettes with an efficiency of ∼90% quantified by PCR and flow cytometry.

Keywords: CRISPR; genome editing; *Saccharomyces cerevisiae*; modularity; cassette swapping; Easy-MISE Toolkit

Introduction

Synthetic biology consists of the integration of biology with engineering principles with the aim of building complex artificial biological systems with predictable behaviors, which can be applied to address urgent issues pertaining to sustainability and health (Chen et al. 2018). For instance, synthetic biology accelerates the construction of microbial cell factories able to produce heterologous molecules difficult to obtain through other processes (Ko et al. 2020, Patra et al. 2021). While synthetic biology toolkits are becoming available for a wide range of both prokaryotic and eukaryotic microorganisms (Martins-Santana et al. 2018, Patra et al. 2021, Hwang et al. 2023, Volke et al. 2023), *Saccharomyces cerevisiae* remains the preferred chassis for developing novel tools, thanks to the ease of manipulation and the wide collection of already existing toolkits.

For instance, many toolkits for the expression of heterologous pathways are available for this host, covering a wide range of the *complexity–flexibility–development speed* balance, with both the possibility to build multilevel, highly complex assemblies (Lee et al. 2015), or simpler and faster single level constructs (Jessop-Fabre et al. 2016, Maestroni et al. 2023a). Some of them have already been expanded to add functionalities, increasing possibilities in the strain design (Otto et al. 2021, Shaw et al. 2023). These tools leverage the modularity and standardization principles of synthetic biology by using interchangeable standardized components that can be combined in various ways and at different levels to create complex constructs. In particular, collections of standard parts allow the construction of devices, which are then introduced in the yeast genome to form complex systems.

Nevertheless, the modularity is limited to the processes that precede the introduction in the expression host: as a matter of fact, none of the currently available and commonly used yeast synthetic biology toolkits includes by design the possibility to retarget the modifications after their introduction into the genome. In particular, once a heterologous expression cassette is integrated, there is no systematic way to remove or exchange it with another one without custom intervention or without restarting from the last (parental) strain not containing the undesired modification. This limits the modularity to the constructs, which are introduced in nonmodular strains that are then difficult to rearrange. As examples, the repurposing of a free fatty acids chassis strain for the production of sclareol required extensive deletions of previously overexpressed genes (Yu et al. 2018, Cao et al. 2023) or the substitution of the first enzyme of glucobrassicin biosynthetic pathway with a more efficient one implied the reintroduction of all the other modifications in the wildtype strain (Maestroni et al. 2023a). Strain construction is indeed strictly sequential and stiff: this approach hampers the flexibility in complex strain engineering strategies in which multiple enzymes combinations need to be tested and limits the possibility of "strain upgrading" when new more efficient enzymatic variants are discovered, even after long time from the initial strain design and construction (Fig. 1).

Received 11 October 2024; **revised** 27 May 2025; **accepted** 16 June 2025
© The Author(s) 2025. 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 commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site-for further information please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

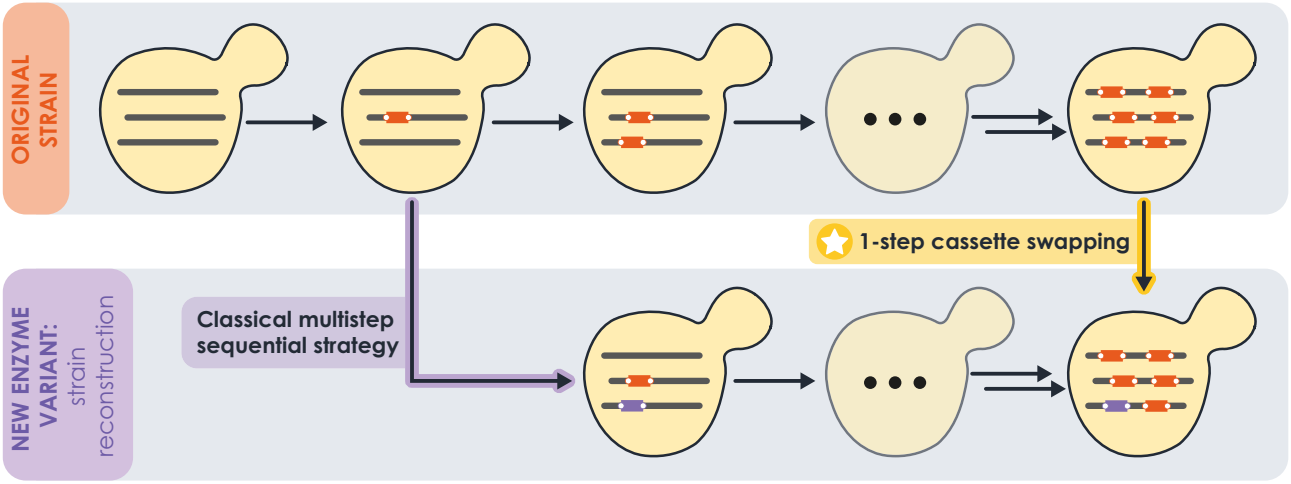

Figure 1. Classical and cassette swapping-based strategies for strain modification to express a new enzymatic variant. In order to substitute an expression cassette to test possible more performant enzymatic variants, the classical multistep sequential strategy requires starting from the last strain without the cassette of interest and repeating the following integrations. The cassette swapping strategy allows substituting any barcode-tagged cassette with one single step, at any stage of the strain development.

A possible solution to the sequentiality of the introduction of expression cassettes is represented by the simultaneous integration of many cassettes containing multiple expression units for complete or large parts of the pathway of interest, as addressed by an expansion of the Yeast Toolkit (Shaw et al. 2023). This approach can be successful but does not tackle the problem directly, since it just decreases the steps involving the nonmodular strain, delegating the need for modularity to more complex and powerful cloning steps. This solution is not effective when other metabolic engineering interventions (i.e. deletions or promoter changes) are alternated to heterologous expression of genes or when already existing strains need to be upgraded a long time after their design.

An alternative solution is to increase the modularity of the final strain allowing to target in a standardized way the introduced modifications and thus swapping heterologous expression cassettes. The exchange of expression cassettes is a practice well established in basic research but never applied systematically to strain engineering in industrial biotechnology. For instance, recombinase-mediated cassette exchange exploits exogenous recombinases to exchange a placeholder cassette previously introduced in the genome with the newly desired one (Turan et al. 2011). This technique is commonly exploited in metazoan engineering to standardize heterologous cassette integrations, overcoming random genomic integration by the nondisruptive insertion of the cassettes at a precharacterized genomic locus with appropriate landing sequences. The more recent CRISPR-mediated cassette exchange (CriMCE) technique follows the same principle with the advantage of a scar-less swapping of the placeholder cassette with the desired one, thanks to the introduction of a couple of DNA double-strand breaks (DSB) mediated by CRISPR-Cas9 before and after the locus of interest in *Anopheles gambiae* (Morianou et al. 2022).

Approaches similar to CriMCE involving a couple of DSB surrounding the genomic region of interest have been used to substitute native alleles with foreign variants both in metazoan (Kelton et al. 2017) and in yeast (Trindade de Carvalho et al. 2017). Moreover, a similar technique for a double-step allele swapping procedure in which the native allele is firstly substituted by an antibiotic resistance cassette which is then used as a Cas9 target to trigger the marker-free integration of the foreign allele has been described multiple times in yeast (see as examples Lee et al. 2019, Lutz et al. 2019).

None of the described techniques have been thought for extended, standardized, and easily maintainable swapping of multiple cassettes and has not been applied to industrial strains engineering. In this work, we have designed a strategy that exploits the principle of CRISPR-mediated cassette swapping to bring modularity to the strain level and demonstrated its effectiveness applying it to the Easy-MISE Toolkit (EMT) framework (Maestroni et al. 2023a) for the construction and integration of swappable expression cassettes.

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*; and *SUC2* — Dr P. Kötter, Institute of Microbiology, Johann Wolfgang Goethe-University, Frankfurt, Germany) (Entian and Kötter 1998, van Dijken et al. 2000). Other strains constructed and described in this work are reported in Table S1. *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 medium (10 g l^{-1} NaCl, 10 g l^{-1} peptone, and 5 g l^{-1} yeast extract) or terrific broth media (20 g l^{-1} peptone, 24 g l^{-1} yeast extract, 4 ml l^{-1} glycerol, 0.17 M KH_2PO_4, and 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, and 10 g l^{-1} yeast extract). When needed, the medium was supplemented with the antibiotic G418 (200 mg l^{-1}) or clonNAT (100 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

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

at least three times, unless otherwise indicated. 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.

Barcodes selection

The artificial barcode BRC-A1 contains a 20-bp sequence naturally absent in the *S. cerevisiae* CEN.PK113-7D genome, selected using a Python script developed on purpose (Script S1). In brief, random 20 bp sequences with %GC between 40% and 60%—for improved on-target cut efficiency (Konstantakos et al. 2022)—are generated and aligned with the target strain genome; sequences with at least four mismatches or gaps are considered as candidate artificial barcodes (Table S4). A sequence with 50% GC was then selected and flanked by Cas12a PAM (TTTA, in 5′) and Cas9 PAM (CGG, in 3′) to obtain BRC-A1. Other artificial barcodes can be obtained using the sequences reported in Table S4 or using Script S1. The accession number of the target genome sequences, as well as of already used barcodes should be indicated before use. The recycled barcodes share the same structure, where the internal 20-mer is constituted by sequences naturally present in *S. cerevisiae* genome, selected among the CRISPR-Cas9 targets used in Jessop-Fabre et al. (2016) (Table S4).

Plasmids construction

All primers and plasmids used in this work are listed in Tables S2 and S3, respectively. Golden Gate Assembly reactions were performed as described in Maestroni et al. (2023a). OneTaq® or Q5® High-Fidelity DNA Polymerases from New England Biolabs (NEB) were used on a ProFlex PCR System (Life Technologies) following NEB manuals. All enzymes used were purchased from NEB.

The EMT (Maestroni et al. 2023a) (Addgene Kit number 1000000230) was exploited as a base to apply the principle of cassette swapping, by constructing modified versions of both left (T01L) and right (T02R) terminators, as well as BF (A02L) and FL (A02R) adaptors. The new pEM parts plasmids bearing the barcode sequences were constructed according to a modified version of the protocol described in Maestroni et al. (2023a), by inserting the barcode sequence in the desired position and orientation in the primers used to amplify the inserts before the Esp3I Golden Gate Assembly in the pGA-blue acceptor vector. The 27-bp barcode was inserted between the B protruding and the insert sequences for parts of the left transcriptional unit (TU) or between the insert and the L protruding sequences for the right TU, always with Cas9 PAM toward the insert sequence and the Cas12a PAM next to the BsaI protruding (Fig. S1). Already existing canonical EMT parts were used as templates. For instance, in order to obtain pEM.T01L-A1 part, the barcode BRC-A1 was inserted between the B protruding and the sequence of left *ADH1t* terminator (T01L) by amplifying the terminator sequence from pEM.T01L with suitable primers carrying the barcode and cloning the amplicon in the pGA-blue plasmid.

Moreover, green fluorescent protein (GFP) and mCherry coding sequences were cloned in pGA-blue with G and H BsaI protruding sequences, obtaining respectively pEM.GFP.R and pEM.mCh.R according to the standard EMT protocol.

Plasmids carrying expression cassettes for GFP and mCherry were then assembled following the EMT workflow using library parts pEM.H01L (X3 upstream homology region), pEM.P01R (*ScTDH3* promoter), pEM.A01R (HI adaptor), and pEM.H01R (X3 downstream homology region), in addition to the previously described pEM.GFP.R, pEM.mCh.R, and modified versions of pEM.A02L (BF adaptor) and pEM.T02R (*ScCYC1* terminator) (Fig. S2).

CRISPR-Cas9/gRNA plasmids targeting barcode A1 and the loci described in Mikkelsen et al. (2012) were constructed by cloning suitable dsDNA oligonucleotides containing the 20-mer sequences used in Jessop-Fabre et al. (2016) (Tables S3 and S4) in the pCEC-red acceptor vector (Addgene Plasmid number 196040), as described in Maestroni et al. (2023b).

*Lb*Cpf1 BRC-A1-targeting crRNA expression plasmid was built starting from plasmid pCfB2903 (a gift from Irina Borodina, Addgene plasmid number 73275; http://n2t.net/addgene:73275; RRID:Addgene_73275). The whole backbone, except for the original gRNA sequence, was amplified with primers crCpf1-A1-b_F—annealing in *SUP4* terminator and bearing A1 spacer in 5′—and crLbCpf1_dr21_R—annealing in *SNR52* promoter and bearing the 21-bp direct repeat described in Verwaal et al. (2018) (Table S2). After digestion with DpnI to avoid template carry over, the amplicon was purified, phosphorylated and self-ligated to obtain the final plasmid pRNA-Lb21-A1. Correct clones were identified by Sanger sequencing.

*Lb*Cpf1 expression plasmid pCSN067 was a gift from Rene Verwaal (Addgene plasmid number 101748; http://n2t.net/addgene:101748; RRID:Addgene_101748).

Yeast transformation

Saccharomyces cerevisiae was transformed according to a modified version of the protocol described in Gietz and Schiestl (2007). In particular, the transformation mixture contained the DNA mix as detailed below, 240 μl of 50% PEG 3350, 36 μl of 1 M lithium acetate, 10 μl of 10 mg ml^{-1} salmon sperm ssDNA, and sterile water up to 360 μl. For Cas9-mediated genome editing, the DNA mix contained 100 ng (17.24 fmol) of the needed specific pCEC-gRNA plasmid together with a 10-fold molar quantity (172.4 fmol) of the repair fragment of interest. For *Lb*Cpf1 mediated editing, two sequential transformations were carried out. First, Sc.EMT-S1 strain was transformed with 120 ng (17.2 fmol) of pCSN067 plasmid, then Sc.EMT-S1-*Lb*Cpf1 was transformed with a DNA mix containing 554 ng of pRNA-Lb21-A1 (172.4 fmol) and a 2-fold molar quantity (334.8 fmol) of the repair fragment of interest.

S1_GFP-A1 and S2_mCh-X3 repair fragments were obtained by NheI digestion followed by gel purification of the plasmids pS1_GFP-A1 and pS2_mCh-X3, respectively; wt X3 locus fragment was amplified from wt genomic DNA using X3_UP_AB_fw and X3_DW_LM_rv primers. In addition, for each transformation, a negative control in which neither pCEC-gRNA plasmid nor the repair fragment were provided was performed. Cells were then shocked at 42°C for 40 min and then recovered for 2.5 h in YPD medium. For classical transformations (i.e. screened by PCR on genomic DNA), the cell suspension was then plated on YPD + G418 (for pCEC and pCSN067 plasmids) or YPD + G418 + clonNAT (for pRNA-Lb21-A1 plasmid) solid media and incubated for 3 days at 30°C before screening. For flow cytometry experiments, 100 μl of the cell suspension were inoculated in 20 ml of YPD + G418 and grown for up to 48 h.

Cassette swapping assessment: PCR

For each transformation, at least eight colonies from two independent transformation reactions were screened by PCR. Primers X3_UP_ctr_int, GFP_internal_fw or mCherry_internal_fw were coupled with primer X3_DW_ctr_int for Sc.EMT-S0, Sc.EMT-S1, and Sc.EMT-S2, respectively. For detailed analysis of negative clones of Sc.EMT-S2 transformations, primer couples

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

BRC-X3_fw + X3_DW_ctr_int, X3_UP_ctr_int + BRC-X3_fw, and X3_UP_ctr_int + mCherry_internal_fw were used. For nucleotide-level sequence verification of correct and incorrect clones, X3 genomic locus was amplified using X3_UP_ctr_int and X3_DW_ctr_int primers and Sanger sequenced using suitable primers.

Cassette swapping assessment: flow cytometry

Sc.EMT-S2 transformants obtained by cassette swapping in Sc.EMT-S1 were analysed by flow cytometry to assess GFP and mCherry expression in the cell population. Three independent transformations were analysed in parallel, in addition to three negative controls (without pCEC-gRNA and repair fragment) and pure cultures of wt, Sc.EMT-S1, and Sc.EMT-S2 strains as positive controls. Liquid cultures were sampled regularly until 48 h, when the percentage of mCherry positive clones reached a plateau (i.e. there was no statistically significant difference with the previous time point when a *t*-test was performed; Table S5). For each sample, ~5 × 10^6 cells were washed twice with phosphate-buffered saline, sonicated, and analysed using a flow cytometer (CytoflexS, Beckman), with excitation at 488 and 525 nm emission for GFP and 610 nm emission for mCherry. A total of 20 000 events were acquired for each sample and data were processed using FlowKit v1.1.2 Python library (White et al. 2021). Events coordinates were transformed with a Logicle transformation with the following parameters: T = 10^7, W = 0.5, M = 7, and A = 0, defined as in the GatingML 2.0 specification (Parks et al. 2006, Spidlen et al. 2015). A polygonal gate including events present only in the Sc.EMT-S2 control sample, but not in Sc.EMT-S1, or in wt and negative samples was used to select events representing mCherry-positive/GFP-negative and thus successful cassette swapping. Similarly, a gate for GFP-positive/mChery-negative clones (unsuccessful cassette swapping) was set based on the Sc.EMT-S1 control sample. Custom gates were set to compensate for slightly different population coordinates at different time points.

Results and discussion

General cassette swapping systems description

The cassette swapping system described in this work allows both to substitute or remove an expression cassette already introduced in the genome at any moment in a standardized, nonsequential, and CRISPR-mediated way. This approach brings modularity to the strain level, going beyond strictly sequential classical engineering: the substitution or removal of a feature introduced in the middle of the strategy does not require restarting from the last strain not carrying the undesired sequence, but can be performed in a single transformation step (Fig. 1).

The strategy is based on the introduction of couples of unique barcode sequences flaking the expression cassettes (Fig. 2A). Every single barcode is univocally associated with an expression cassette and is designed to be targeted by Cas9 — or other, e.g. Cas12a — CRISPR nucleases. By targeting CRISPR-mediated DSBs to a specific couple of barcodes, the desired cassette can be excised from the genome and the two genomic DNA ends left by the DSBs can be repaired by providing a suitable repair fragment. If a different integration cassette with distinct barcodes is provided, the original cassette is substituted with the new one (Fig. 2A); if the original wt locus sequence is used as a repair fragment, the original cassette is removed and the native locus is reconstructed (Fig. 2B). The cassette originally present in the genome cannot be erroneously reintroduced, since after the excision it lacks the locus-specific homology regions. By introducing the couple of barcodes in a divergent orientation, the risk of undesired cassette excision by spontaneous homologous recombination between the two identical sequences is eliminated.

For reasons inherent in the way barcodes were designed, when a cassette carrying the barcodes is integrated into *S. cerevisiae*, the barcode sequences must be absent in the genome of the parental strain, otherwise the CRISPR-Cas9 system would be targeted against these other sequences as well. For this reason, a sequence that is normally absent from the *S. cerevisiae* genome needs to be chosen as a barcode to perform the genome editing event. The system is designed in a way that only a few artificial unique barcodes need to be used (ideally a number equal to the quantity of the cassettes that are simultaneously introduced in the first modification step). Indeed, when the first CRISPR-mediated insertion editing occurs at a certain genomic locus, the used 20-mer protospacer can be eliminated by designing the homology regions a few bases upstream and downstream to it. Thus, whereas for the first editing we are forced to use an artificial barcode, for the second editing we can use a barcode corresponding to the 20-mer sequence that has been lost in the previous integration; now, since the 20-mer sequence used as a target for the second editing has been lost, a new barcode containing this sequence can be used for the third editing, and further modifications can be performed in this way, recycling the barcode sequences freed in the previous steps (Fig. 2C). Therefore, thanks to barcode recycling, with just one artificial barcode the system allows to integrate a number of cassettes, which is not limited by artificial barcodes availability, but depends only on the available target loci. Moreover, by recycling endogenous barcodes, the number of CRISPR plasmids is minimized, as the same plasmids used to target wt genomic loci for cassette integration, can be reused to target recycled barcodes.

EMT cassette swapping

The described system is, in principle, toolkit-agnostic and can be combined with different existing synthetic biology tools. Nevertheless, the benefits of the strain-modularity given by cassette swapping and the possibility of having simpler design processes can be better exploited in toolkits with a lighter design in which fewer genes are introduced in different loci and can be exchanged or removed in small groups in a more dynamic way. That is the reason we decided to implement the cassette swapping strategy in the EMT, a single-level Golden Gate Assembly framework (Maestroni et al. 2023a), in which a maximum of two divergent transcription units (TUs) are introduced in each locus (with the possibility of simultaneous multiple integrations). The EMT is available on Addgene (Addgene Kit number 1000000230), together with a detailed protocol for its use.

In brief, the EMT design is composed of two symmetrical divergent TUs, which can be built from a 10 parts Golden Gate Assembly starting from an existing parts library (Fig. 3A). Parts include, from the center to outside: a promoter, an open-reading frame (ORF), an in-frame C-terminus tag, a terminator and a homology region. A single TU can be skipped thanks to an adaptor part. The system was redesigned to include the barcode sequences while minimizing the request for new parts for the cassette swapping "add-on". Since the barcode sequences must be included between the homology regions and the terminators/adaptors these latter locus-independent parts were modified (Fig. 3B).

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

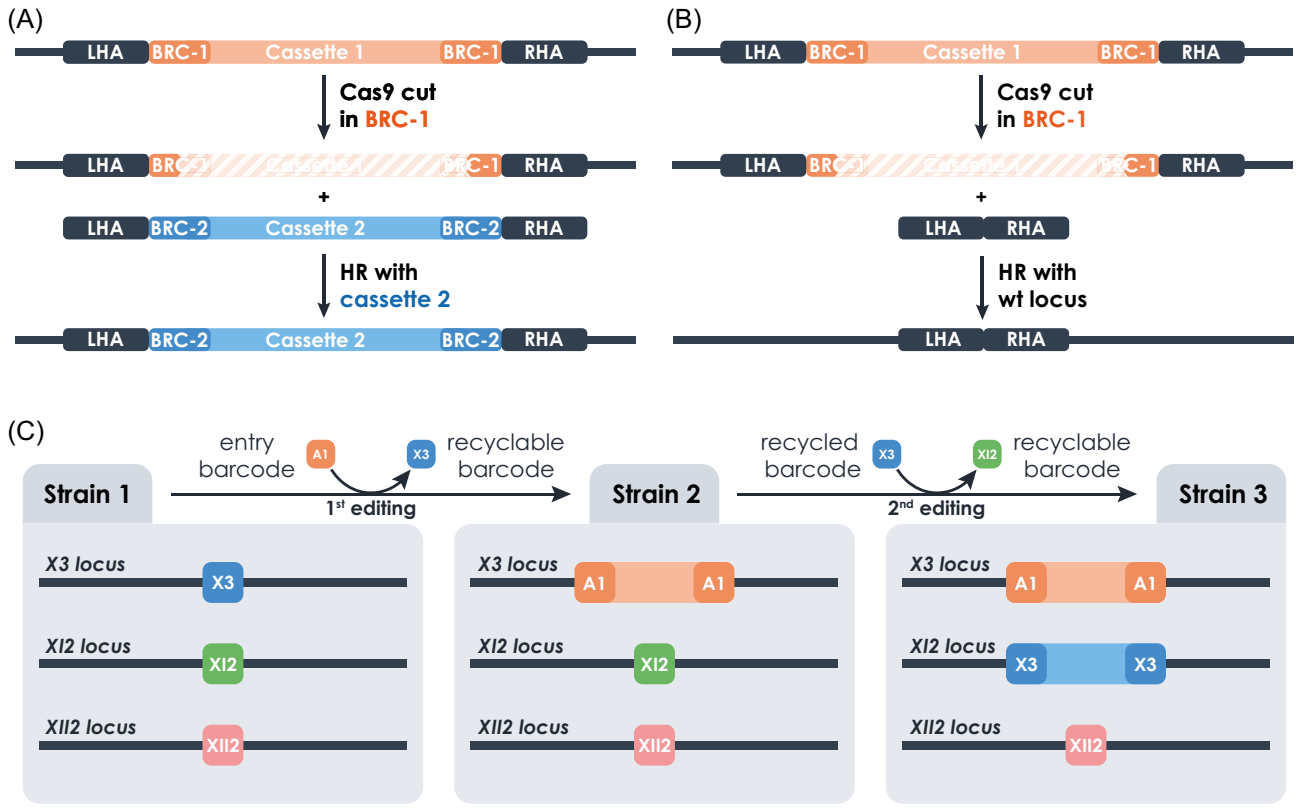

Figure 2. General principle of cassette swapping. (A) An already integrated cassette tagged with barcode BRC-1 can be substituted with a new one with a different barcode (i.e. BRC-2) by targeting Cas9 cut to BRC-1 and providing the second cassette as a repair fragment. (B) A BRC-1-tagged cassette can be replaced with the original locus in a similar fashion by providing the wt sequence as a repair fragment. (C) Endogenous sequences used to target Cas9 at the integration sites can be recycled as barcodes as they are removed by cassette integration: the integration of a first artificial barcode A1 in locus X3 frees the X3 locus-specific target sequence, which can be then reused as a barcode for the next integration, which releases a new locus-specific sequence and so on. LHA: left homology arm; RHA: right homology arm; and BRC: Barcode.

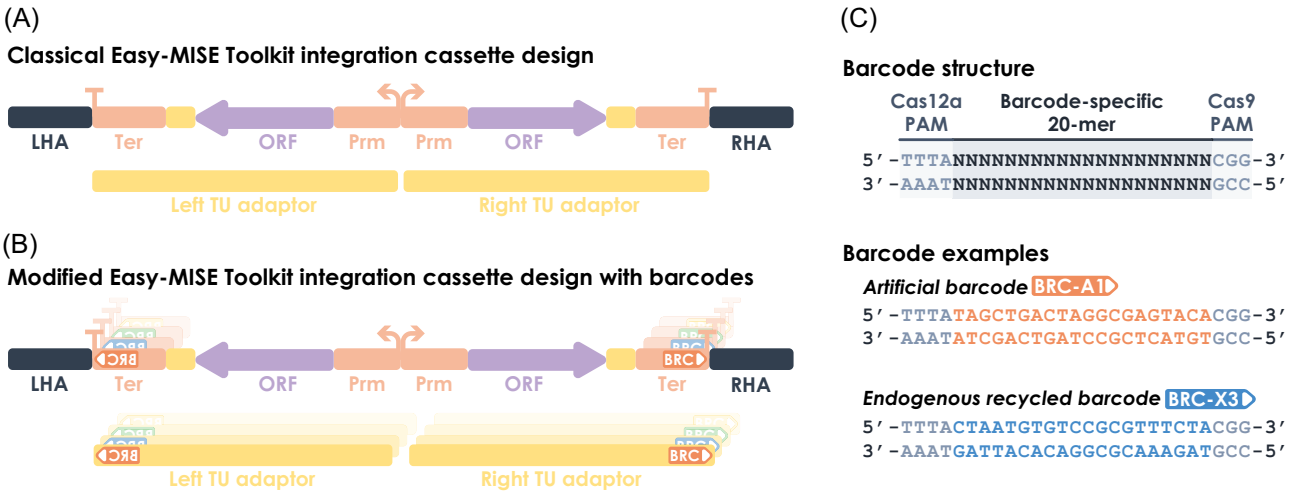

Figure 3. Cassette swapping implementation in the EMT. (A) Classical EMT cassette design, with two symmetrical TU, comprising a promoter (Prm), an ORF, an in-frame fusion module, and a terminator (Ter), flanked by a couple of homology arms (LHA and RHA). Single TUs can be skipped thanks to an adaptor part. (B) Modified cassette design in which barcodes (BRC) have been inserted in the terminators and the TU adaptors. Barcodes have a divergent orientation to avoid excision by spontaneous homologous recombination. (C) Barcode structure: each barcode, artificial (here as example BRC-A1, barcode Artificial-1) or recycled (here as example BRC-X3, barcode from locus X3) comprises a Cas12a PAM, a barcode-specific 20-mer sequence, and a Cas9 PAM.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

The barcodes sequences were designed to be targettable both by Cas9 and Cas12a CRISPR nucleases and are thus constituted by three elements, in the 5′–3′ order: a common Cas12a PAM (TTTA), a barcode-specific 20-mer sequence, a Cas9 PAM (CGG). As described, the 20-mer specific feature can be either an artificial or an endogenous and recycled sequence (Fig. 3C). For system demonstration an artificial barcode, called A1 (standing for "Artificial 1") was designed be absent in *S. cerevisiae* genome, in addition to 20-mer sequences used to target the EMT loci (Jessop-Fabre et al. 2016, Maestroni et al. 2023a). To make the system even more flexible, the EasyMISE toolkit was further modified with a transition from the original Cas9 and gRNA double plasmid system to a simpler single plasmid system based on the pCEC-red vector (Maestroni et al. 2023b), expressing both the nuclease and the gRNA from the same plasmid. This change gives multiple advantages in the context of the toolkit: (i) no needs of a preliminary transformation with Cas9 vector; (ii) the expression of Cas9 is restricted to the editing procedure, limiting possible off-target mutations due to continuous expression; and (iii) no need of double antibiotic selection, reducing cellular stress, and costs. This new set of plasmids can be used to target both the native loci or the recycled barcodes.

A set of new parts was constructed including the artificial barcode BRC-A1 and the recyclable barcodes BRC-X3, BRC-X4, BRC-XI2, BRC-XI3, BRC-XII2, and BRC-XII5, corresponding to the homonymous loci exploited in the EMT. The matching pCEC plasmids were constructed as well.

Easy-MISE cassette swapping functionality assessment

The functionality of the described system was assessed with a proof of concept experiment in which cassettes for the strong constitutive expression of GFP or mCherry fluorescent proteins were swapped one with the other.

Strain Sc.EMT-S1 was obtained by the introduction of cassette S1_GFP-A1 tagged with barcode A1 into the genomic locus X3. The cassette was then exchanged with S2_mCh-X3 by cotransformation with pCEC-A1 targeting the artificial barcode. The two swapped expression cassettes share the same parts and are identical except for the specific barcodes and for the ORF, thus simulating the possibility of exchanging very similar expression cassettes (with 85.59% of identity), for instance for the swap of two closely related genes (Fig. S2).

A preliminary screening was assessed by PCR followed by sequencing confirmation: the test included 20 clones randomly picked from two independent transformations. It demonstrated that 90% of the clones correctly swapped the entire cassette (i.e. the barcodes and the ORF), while in the remaining 10% only the barcodes were exchanged, confirming the possibility of successfully performing cassette swapping with high efficiency (Fig. 4A). The partial swapping events can be explained by an undesired quadruple crossing-over in which not only the upstream and downstream homology regions recombined, but also the shared promoter and terminator, leading to DSB repair without ORF swap. This undesired event is found in all the negative clones (2 out of 20 screened). Interestingly, locus sequencing revealed the presence of two to four base-change point mutations evenly distributed across the internal sequences shared between the exchanged cassettes—and object of the undesired recombination events—at both junctions of all the negative clones, with mutations in some cases shared even in clones coming from independent transformations (Fig. S3): we speculated that this can be due to spatial interference between the molecular machineries independently repairing the two closely positioned DSBs, as observed for other types of spatially close DNA lesions (Ma et al. 2009). The occurrence of these events suggests that the verification of the swapping should be performed using primers specific for the sequence being exchanged and, possibly, sequencing the swapped region. Moreover, the use of genetic elements differing as much as possible from those already present at the locus should minimize quadruple crossing-over events in the regions flanking the coding sequence. When this is not possible, given the low occurrence of partial swapping events (10%), the screening of a higher number of clones should be sufficient to identify correct swapping events.

To better quantify the efficiency of Easy-MISE cassette swap system on a larger number of clones, the expression of GFP and mCherry in negative and positive transformants, respectively, was tested using flow cytometry. After transformation for cassette swapping, cells were recovered for 2.5 h and inoculated in liquid selective media and analysed at multiple time points until 48 h (Fig. S4), when the percentage of mCherry-positive and GFP-positive clones populations stabilized within 3% and there was no statistically significant difference with the previous time point when a *t*-test was performed (Table S5), meaning that the population of successfully transformed cells (with or without the desired cassette swapping event) overgrew the nontransformed cells selected by the antibiotic and stabilized.

At 48 h, the population of mCherry-positive/GFP-negative events is 89.28% ± 3.22 while only 8.30% ± 2.55 of the events are GFP-positive/mCherry-negative (Fig. 4B). The remaining events can be reconducted to dead cells negatively selected by the antibiotic and cellular debris. The same phenomenon generates the prevailing population in the negative control of the transformation, where no plasmid and cassette were provided and therefore growth is inhibited by the antibiotic (Fig. 4C). This flow cytometric analysis, because of the number of the considered events, provides a more statistically robust confirmation of the effectiveness of the cassette swapping implementation and of the reliability of the efficiency estimation by PCR. The measured efficiency is in line with the 87.5% value reported for the double-step and selection-marker-based native allele swapping system described by Lutz et al. (2019), with our system being more standardized and applicable to multiple heterologous genes expression cassettes.

To demonstrate the possibility of removing a cassette and restoring the original wt locus, a second set of transformations was performed, in which S1_GFP-A1 cassette integrated in locus X3 in Sc.EMT-S1 was targeted for swapping and a ∼1 kb amplicon containing the 500-bp upstream and downstream to the native X3 locus was provided as repair fragment, to transform Sc.EMT-S1 into Sc.EMT-S0. In this strain, the native form of locus X3 is completely restored. Screening by PCR and locus sequencing confirmation demonstrated a 100% efficiency of the restoration of the native locus (Fig. 4A).

Finally, the barcode structure compatibility with Cas12a nucleases was assessed by testing the same swapping and locus restoration procedure using the *Lb*Cpf1 two-plasmids system described in Verwaal et al. (2018) instead of the pCEC system. Similarly to what obtained with Cas9, efficiencies of 90% (n = 20, from two independent transformations) and 100% (n = 20, from two independent transformations) were obtained for GFP-mCherry cassettes swapping in locus X3 using barcode A1 and for X3 native locus restoration, respectively, demonstrating the effectiveness of the double-PAMs barcode design (Fig. 4A). Nev-

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

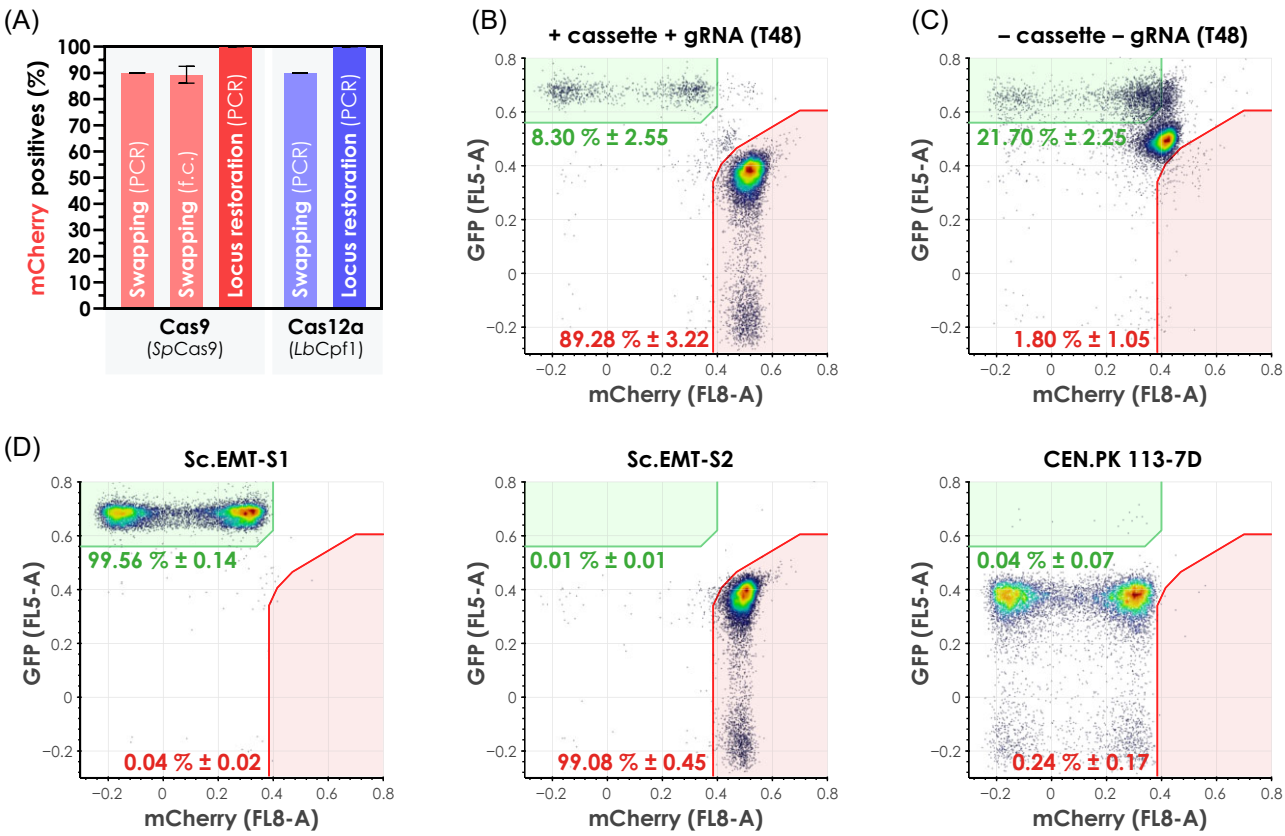

Figure 4. Assessment of cassette swapping efficiency on transformants. (A) Percentages of positive clones for GFP—mCherry cassettes swapping (indicated as "Swapping") or wt locus restoration (indicated as "Locus restoration") quantified by PCR on genomic DNA (PCR) or by flow cytometry (f.c.). The nucleases used in each experiment are indicated below the bar plot. (B) Scatter plot from flow cytometry analysis of samples taken after 48 h from the transformation of Sc.EMT-S1 with S2_mCh-X3 integration cassette and Cas9-gRNA plasmid targeting barcode A1. (C) Negative control of transformation in which integration cassette and Cas9-gRNA plasmid are omitted. (D) Scatter plots of flow cytometry analysis of control strains Sc.EMT-S1 (mCherry-negative and GFP-positive control), Sc.EMT-S2 (mCherry-positive and GFP-negative control), and CEN.PK 113–7D (mCherry-negative, GFP-negative control, and no fluorescence), grown at comparable OD as transformation samples. Coordinates of the flow cytometry data have undergone logicle transformation. The GFP-positive/mCherry-negative gate is depicted in green, while the red area indicates the mCherry-positive/GFP-negative gate. The percentages of events in each gate are reported in the same color. All flow cytometry experiments have been conducted in triplicate collecting 2 $\times 10^4$ events.

ertheless, differently from Cas9, less information is available on the sequence preferences for the protospacer, thus careful test of the available barcodes should be done to ensure complete compatibility.

In this work, the system's principle was tested by applying it to the exchange of two easily detectable expression cassettes for fluorescent proteins. Beyond this, the scope for which the tool was designed is the application to industrial-scale strain development in the context of metabolic engineering. The system has been designed for allowing the simultaneous swapping of more than one cassette, by exploiting an appropriate number of artificial barcodes (Table S4) and plasmids expressing multiple gRNAs. A detailed protocol for the application of the EMT cassette swapping "add-on" to *S. cerevisiae* strain engineering is available as supplementary material. Despite the system being an implementation of toolkits and mechanisms widely and successfully used in many different organisms, the yeast *S. cerevisiae* is amenable to being genetically engineered. Therefore, it would be interesting and relevant to test the system in different yeasts, also in hybrid ones, which are industrially relevant but at the same time challenging in terms of precision editing (Bennis et al. 2023, Gorter de Vries et al. 2017, Jayaprakash et al. 2023).

Conclusions

In the present work, we describe a system that allows to target and manipulate genomically integrated heterologous expression cassettes, allowing for their exchange or removal. The system successfully exploits CRISPR-Cas9 or CRISPR-Cas12a to allow the excision from *S. cerevisiae* genome of a previously integrated cassette and its substitution with another one (or the wild-type locus), thanks to the use of unique cassette-specific barcodes. This approach brings the modularity principle to the strain level, going beyond its limitation to the phases preceding the integration of the expression cassettes into the genome typical of the synthetic biology toolkits described up to now for yeasts (Lee et al. 2015, Jessop-Fabre et al. 2016, Maestroni et al. 2023a). This system exploits the same principle of allele swapping systems already used for genotype–phenotype association studies both in yeasts (Lee et al. 2019, Lutz et al. 2019) and in more complex eukaryotic organisms (Kelton et al. 2017), with the advantages of making it single-step, independent from the integration of selection marker cassettes, and applicable repeatedly for an indefinite number of times.

The principle of cassette swapping has been integrated with an existing synthetic biology tool, the EMT (Maestroni et al. 2023a),

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

leveraging the modularity of its expression cassettes and can be seen as an optional "add-on." Moreover, the original double-plasmid CRISPR-Cas9/gRNA system of the EMT was successfully substituted with the pCEC-red single plasmid system (Maestroni et al. 2023b), making the strain construction even more flexible. The compatibility of the barcode structure with Cas12a nucleases was assessed using an existing *Lb*Cpf1 two-plasmid system (Verwaal et al. 2018), further enlarging the applicability scope and flexibility of the described tool. In addition, given that the general idea described is applicable independently from a specific toolkit, in the future the standardized barcodes-based cassettes swapping principle can be integrated with other more complex toolkits (e.g. the Yeast Toolkit), in order to satisfy the needs of the yeast community and the compliance to other well established engineering standards. Moreover, by selecting suitable barcodes (e.g. using the Python script here proposed), the principle can be applied to other microorganisms for which efficient CRISPR-Cas9 editing techniques and mutants with improved homologous recombination rates are available.

The use of the described system from the very beginning of the design phases of a host cell should allow the construction of highly flexible strains and greatly improve the maintainability of industrial cell factories, making it possible to upgrade them as novel variants of enzymes or pathways are discovered. According to the vision of synthetic biology as the application of engineering principles to biology, this system makes strain engineering comparable to software and electronic engineering, where code-bases and machines are maintained and updated during time by upgrading the single components without the need to reassemble everything from scratch.

Acknowledgments

The authors would like to thank Dr Stefania Citterio for technical assistance with flow cytometry analysis.

Author contributions

P.B.: Conceptualization, Methodology, Validation, Investigation, Data curation, Software, Writing – original draft, Writing – review & editing. F.B.: Methodology, Validation, Investigation, Writing – review & editing. E.B.: Methodology, Validation, Investigation, Writing – original draft. P.B.: Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Supplementary data

Supplementary data is available at *FEMSYR Journal* online.

Conflict of interest: The authors declare that they have no known competing 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 FA (Fondo di Ateneo) to P.B. The PhD fellowship of P.B. was cofinanced by MUR PON Azione IV.5 and ALBINI GROUP. P.B. acknowledges the National Center 5 "National Biodiversity Future Center" (award number: Project code CN_00000033, concession decree number 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP H43C22000530001, Project title "National Biodiversity Future Center—NBFC". Funder: Project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4–Call for tender number 3138 of 16 December 2021, rectified by decree number 3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union—NextGenerationEU).

References

Bennis NX, Kostanjšek M, van den Broek M *et al.* Improving CRISPR-Cas9 mediated genome integration in interspecific hybrid yeasts. *New Biotechnology* 2023;**76**:49–62. https://doi.org/10.1016/j.nbt.2023.04.001

Cao X, Yu W, Chen Y *et al.* Engineering yeast for high-level production of diterpenoid sclareol. *Metab Eng* 2023;**75**:19–28. https://doi.org/10.1016/j.ymben.2022.11.002.

Chen B, Lee HL, Heng YC. *et al.* Synthetic biology toolkits and applications in *Saccharomyces cerevisiae*. *Biotechnol Adv* 2018;**36**:1870–81. https://doi.org/10.1016/j.biotechadv.2018.07.005.

Entian K-D, Kötter P. 23 yeast mutant and plasmid collections. In: Brown AJP, Tuite M (eds), *Methods in Microbiology, Yeast Gene Analysis*. Cambridge, MA: Academic Press, 1998, 431–49. https://doi.org/10.1016/S0580-9517(08)70344-1.

Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2007;**2**:31–34. https://doi.org/10.1038/nprot.2007.13.

Gorter de Vries AR, de Groot PA, van den Broek M *et al.* CRISPR-Cas9 mediated gene deletions in lager yeast Saccharomyces pastorianus. *Microbial Cell Factories* 2017;**16**. https://doi.org/10.1186/s12934-017-0835-1

Hwang S, Joung C, Kim W *et al.* Recent advances in non-model bacterial chassis construction. *Curr Opin Syst Biol* 2023;**36**:100471. https://doi.org/10.1016/j.coisb.2023.100471.

Jayaprakash P, Barroso L, Vajente M *et al.* CRISPR-Cas9 engineering in the hybrid yeast Zygosaccharomyces parabailii can lead to loss of heterozygosity in target chromosomes. *FEMS Yeast Research* 2023;**23**. https://doi.org/10.1093/femsyr/foad036

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. https://doi.org/10.1002/biot.201600147.

Kelton W, Waindok AC, Pesch T *et al.* Reprogramming MHC specificity by CRISPR-Cas9-assisted cassette exchange. *Sci Rep* 2017;**7**:45775. https://doi.org/10.1038/srep45775.

Ko Y-S, Kim JW, Lee JA *et al.* Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chem Soc Rev* 2020;**49**:4615–36. https://doi.org/10.1039/D0CS00155D.

Konstantakos V, Nentidis A, Krithara A *et al.* CRISPR–Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. *Nucleic Acids Res* 2022;**50**:3616–37. https://doi.org/10.1093/nar/gkac192.

Lee JT, Coradini ALV, Shen A *et al.* Layers of cryptic genetic variation underlie a yeast complex trait. *Genetics* 2019;**211**:1469–82. https://doi.org/10.1534/genetics.119.301907.

Lee ME, DeLoache WC, Cervantes B *et al.* A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth Biol* 2015;**4**:975–86. https://doi.org/10.1021/sb500366v.

Lutz S, Brion C, Kliebhan M *et al.* DNA variants affecting the expression of numerous genes in trans have diverse mechanisms of action and evolutionary histories. *PLoS Genet* 2019;**15**:e1008375. https://doi.org/10.1371/journal.pgen.1008375.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

Ma W, Panduri V, Sterling JF *et al.* The transition of closely opposed lesions to double-strand breaks during long-patch base excision repair is prevented by the coordinated action of DNA polymerase δ and Rad27/Fen1. *Mol Cell Biol* 2009;**29**:1212–21. https://doi.org/10.1128/MCB.01499-08.

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 Synth Biol* 2023a;**12**:1508–19. https://doi.org/10.1021/acssynbio.3c00015.

Maestroni L, Butti P, Senatore VG. *et al.* pCEC-red: a new vector for easier and faster CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2023b;**23**:foad002. https://doi.org/10.1093/femsyr/foad002.

Martins-Santana L, Nora LC, Sanches-Medeiros A *et al.* Systems and synthetic biology approaches to engineer fungi for fine chemical production. *Front Bioeng Biotechnol* 2018;**6**. https://doi.org/10.3389/fbioe.2018.00117.

Mikkelsen MD, Buron LD, Salomonsen B *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab Eng* 2012;**14**:104–11. https://doi.org/10.1016/j.ymben.2012.01.006.

Morianou I, Crisanti A, Nolan T *et al.* CRISPR-mediated cassette exchange (CriMCE): a method to introduce and isolate precise marker-less edits. *CRISPR J* 2022;**5**:868–76. https://doi.org/10.1089/crispr.2022.0026.

Otto M, Skrekas C, Gossing M *et al.* Expansion of the yeast modular cloning toolkit for CRISPR-based applications, genomic integrations and combinatorial libraries. *ACS Synth Biol* 2021;**10**:3461–74. https://doi.org/10.1021/acssynbio.1c00408.

Parks DR, Roederer M, Moore WA. A new "logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. *Cytometry A* 2006;**69A**:541–51. https://doi.org/10.1002/cyto.a.20258.

Patra P, Das M, Kundu P *et al.* Recent advances in systems and synthetic biology approaches for developing novel cell-factories in non-conventional yeasts. *Biotechnol Adv* 2021;**47**:107695. https://doi.org/10.1016/j.biotechadv.2021.107695.

Shaw WM, Khalil AS, Ellis T. A multiplex MoClo toolkit for extensive and flexible engineering of *Saccharomyces cerevisiae*. *ACS Synth Biol* 2023;**12**:3393–405. https://doi.org/10.1021/acssynbio.3c00423.

Spidlen J, Moore W, Brinkman RR. ISAC's Gating-ML 2.0 data exchange standard for gating description. *Cytom Part J Int Soc Anal Cytol* 2015;**87**:683–7. https://doi.org/10.1002/cyto.a.22690.

Trindade de Carvalho B, Holt S, Souffriau B *et al.* Identification of novel alleles conferring superior production of rose flavor phenylethyl acetate using polygenic analysis in yeast. *mBio* 2017;**8**. https://doi.org/10.1128/mbio.01173-17.

Turan S, Galla M, Ernst E *et al.* Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges. *J Mol Biol* 2011;**407**:193–221. https://doi.org/10.1016/j.jmb.2011.01.004.

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. https://doi.org/10.1016/S0141-0229(00)00162-9.

Verwaal R, Buiting-Wiessenhaan N, Dalhuijsen S *et al.* CRISPR/Cpf1 enables fast and simple genome editing of. *Yeast* 2018;**35**:201–11. https://doi.org/10.1002/yea.3278.

Volke DC, Orsi E, Nikel PI. Emergent CRISPR–Cas-based technologies for engineering non-model bacteria. *Curr Opin Microbiol* 2023;**75**:102353. https://doi.org/10.1016/j.mib.2023.102353.

White S, Quinn J, Enzor J *et al.* FlowKit: a Python toolkit for integrated manual and automated cytometry analysis workflows. *Front Immunol* 2021;**12**. https://doi.org/10.3389/fimmu.2021.768541.

Yu T, Zhou YJ, Huang M *et al.* Reprogramming yeast metabolism from alcoholic fermentation to lipogenesis. *Cell* 2018;**174**:1549–1558.e14. https://doi.org/10.1016/j.cell.2018.07.013.

Downloaded from https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/foaf032/8168196 by University of Milan-Bicocca user on 18 July 2025

Received 11 October 2024; **revised** 27 May 2025; **accepted** 16 June 2025

© The Author(s) 2025. 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 commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site-for further information please contact journals.permissions@oup.com