Efficient CRISPR–Cas9 mediated multiplex genome editing in yeasts

Laiyou Wang1,2, Aihua Deng1*, Yun Zhang1, Shuwen Liu1, Yong Liang1, Hua Bai1,2, Di Cui1,2, Qidi Qiu1,2, Xiuling Shang1, Zhao Yang1, Xiuping He3 and Tingyi Wen1,4*

Abstract

Background: The thermotolerant methylotrophic yeast *Ogataea polymorpha* has been regarded as an important organism for basic research and biotechnological applications. It is generally recognized as an efficient and safe cell factory in fermentative productions of chemicals, biofuels and other bio-products. However, it is difficult to genetically engineer for the deficiency of an efficient and versatile genome editing technology.

Results: In this study, we developed a CRISPR–Cas9-assisted multiplex genome editing (CMGE) approach including multiplex genes knock-outs, multi-locus (ML) and multi-copy (MC) integration methods in yeasts. Based on CMGE, various genome modifications, including gene deletion, integration, and precise point mutation, were performed in *O. polymorpha*. Using the CMGE-ML integration method, three genes TAL from *Herpetosiphon aurantiacus*, 4CL from *Arabidopsis thaliana* and STS from *Vitis vinifera* of resveratrol biosynthetic pathway were simultaneously integrated at three different loci, firstly achieving the biosynthesis of resveratrol in *O. polymorpha*. Using the CMGE-MC method, ~10 copies of the fusion expression cassette PscTEF1-TAL-PscTP1-4CL-PscTEF2-STS were integrated into the genome. Resveratrol production was increased ~20 fold compared to the one copy integrant and reached 97.23 ± 4.84 mg/L. Moreover, the biosynthesis of human serum albumin and cadaverine were achieved in *O. polymorpha* using CMGE-MC to integrate genes HSA and cadA, respectively. In addition, the CMGE-MC method was successfully developed in *Saccharomyces cerevisiae*.

Conclusions: An efficient and versatile multiplex genome editing method was developed in yeasts. The method would provide an efficient toolkit for genetic engineering and synthetic biology researches of *O. polymorpha* and other yeast species.

Keywords: CRISPR–Cas9-assisted multiplex genome editing, Markerless multi-locus integration, Markerless multicopy integration, *Ogataea polymorpha*, *Saccharomyces cerevisiae*.

Background

The thermotolerant methylotrophic yeast *Ogataea polymorpha* (*Hansenula polymorpha*), belonging to the fungal family of *Saccharomycetaceae*, subfamily *Saccharomyce toideae*, has been regarded as an attractive organism for fundamental and applied researches [1–5]. It has been an important organism for studies on methanol utilization, autophagy, peroxisome biogenesis and nitrate assimilation [6]. One attractive property of *O. polymorpha* is able to integrate up to 100 copies of target gene into the genome mediated by non-homologous end joining (NHEJ), which can be used to highly express heterologous genes and synthesize various biotechnology products [1, 4, 7, 8]. Furthermore, it can synthesize glycoproteins with human compatible oligosaccharides [9, 10]. In addition, it can grow at high temperatures up to 50 ºC that would reduce the expensive cooling cost of industrial fermentation [9]. Due to these properties, many bio-products, such as ethanol, vaccines, uricase and glutathione, have been successfully synthesized in *O. polymorpha*. Currently, several prophylactic HBV vaccines have been marketed [1, 4, 7, 8].
A versatile multiplex genome editing method is essential in construction of yeast cell factories for various bio-products [11–16]. The Cre/loxP-based site-specific recombination, the mazF-based counter-selectable and the plasmid-based CRISPR–Cas9 systems have been developed for gene modifications in *O. polymorpha* [15–17]. The Cre/loxP system was reported to leave a scar (lox72) at the target locus after gene editing, which might interfere with the subsequent genetic manipulation [17]. To overcome the problem, Song et al. used the mazF as counter-selectable marker for markerless gene deletion [16]. Recently, the CRISPR–Cas9 system has been employed to establish genome editing method [18, 25], by which homologous recombination-mediated gene replacement and NHEJ-mediated gene disruption were performed in *O. polymorpha* [18]. In addition, the pUPD system (a plasmid-based CRISPR/Cas9 system that was applicable to several yeast species) co-expressed Cas9 and gRNA in one plasmid was also developed for gene disruption as reported by Juergens et al. [25]. Although available genome editing methods can adapt for gene deletion and disruption, multiplex gene editing is not possible in *O. polymorpha*. Therefore, an efficient and markerless genome editing system that can mediate multiplex genome engineering is urgently needed in *O. polymorpha*

Recent advances in CRISPR–Cas9-assisted genome editing technologies provide an efficient approach to establish multiplex genome engineering [18–20]. In principal, a trans-activating crRNA (tracrRNA):crRNA duplex directs the Cas9 protein to bind to a target DNA sequence immediately followed by a protospacer adjacent motif (PAM). The Cas9 protein then makes a double-strand break (DSB) in three nucleotides upstream of the PAM site [18, 21–23]. The DSB can be repaired either by non-homologous end joining to result in insertions or deletions, or by homologous recombination to result in precise editing, such as gene deletion, point mutation and integration [5, 6, 17, 24].

Ribosomal DNA (rDNA) is a DNA sequence encoding ribosomal RNA. In yeasts, rDNA typically consists of high copy numbers of identical repeats that are clustered in a head-to-tail tandem array. Each rDNA repeat consists of two transcribed regions separately coding for 35S precursor rRNA and 5S rRNA, and two non-transcribed regions NTS1 and NTS2 [25]. The number and the length of rDNA repeat vary among different species [26]. In *S. cerevisiae*, the rDNA locus consists of 150–200 tandem copies of 9.1-kb rDNA repeat [25]. In *O. polymorpha*, the rDNA contains about 50–60 copies of 8-kb unit [27]. Recently, rDNA repeats have served as target loci for multi-copy integrations in some yeast species [28–30].

In this study, we developed a CRISPR–Cas9-assisted multiplex genome editing method (designated as CMGE) for various gene modifications, especially for multiplex genes knock-outs, multi-locus (ML) and multi-copy (MC) integrations of target genes in *O. polymorpha*. Using CMGE, biosynthesis of resveratrol, cadaverine and human serum albumin (HSA) were achieved in *O. polymorpha*, suggesting the practicability and effectiveness of CMGE in genetic engineering of *O. polymorpha*. In addition, the CMGE-MC method was successfully developed in the model yeast *S. cerevisiae*.

**Results**

**Scheme for CRISPR–Cas9-assisted genome editing in *O. polymorpha***

Due to unavailability of a stable episomal vector in *O. polymorpha*, the Cas9 and gRNA expression cassettes were integrated into the chromosome using recombinant plasmids pWYE3208 and pWYEN (a generic term of all gRNA delivery vector of which “N” represents the serial number. For example, “N” was 3208 in the plasmid pWYE3208), respectively (Fig. 1a). The plasmid pWYE3208 harboring the *PscTEF1*-controlled *cas9* gene and the up- and downstream homologous arms (UHA and DHA) of the *OpMET2* gene was linearized with *SpeI* and then transformed into *O. polymorpha*. Through a double-crossover homologous recombination, the *OpMET2* gene was replaced by the linearized pWYE3208 (Fig. 1b). The desired mutant OP009 (*Zeo*<sup>R</sup>, *Op001ΔOpMET2::* _PscTEF1_ *Cas9*) was verified on YPD plate containing zeocin and further confirmed by PCR.

The gRNA expression cassette was composed of the promoter *P_scsNRS2_ N<sub>20</sub>_trans::crRNA duplex and the terminator *ScSUP4t*. To conveniently screen transformants harboring gRNA expression cassette, *OpADE2* was selected as the integration site of gRNA expression cassette due to the formation of red colonies by accumulating the oxidized 5-aminomimidazole ribonucleotide.

Three unique restriction sites (*SpeI/KpnI/SphI*) between the UHA and DHA of *OpADE2* gene were used to linearize plasmid (Fig. 1a). After co-transformed into the strain OP009 with the editing template, the linearized gRNA delivery vector transcribed gRNA that can guide the Cas9 protein to recognize the PAM site. The Cas9 protein cleaved double DNA strands to generate the DSB, which was subsequently repaired by the editing template via homologous recombination. Meanwhile, the gRNA delivery vector was integrated at the gene *OpADE2* locus by homologous recombination (Fig. 1b).

The resulting transformants were screened on YPD plates containing G418. The desired mutants were obtained by two-step procedures: (i) red colonies were selected to identify the integration of gRNA expression

---

*References*

[11–16]: References to cited articles.

[15–17]: References to cited articles.

[25]: Reference to cited article.

[18]: Reference to cited article.

[19]: Reference to cited article.

[20]: Reference to cited article.

[21]: Reference to cited article.

[22]: Reference to cited article.

[23]: Reference to cited article.

[24]: Reference to cited article.

[25]: Reference to cited article.

[26]: Reference to cited article.

[27]: Reference to cited article.

[28]: Reference to cited article.

[29]: Reference to cited article.

[30]: Reference to cited article.
Fig. 1 Schematic overview of the CRISPR–Cas9-assisted genome editing in *O. polymorpha*. 

(a) The Cas9 delivery vector pWYE3208. The expression of Cas9 was driven by the promoter *P*<sub>SCTEF1</sub>. The OpMET2 5′ and OpMET2 3′ represent the ~1.5 kb of up- and downstream homologous arms of the gene OpMET2, respectively. *zeo<sup>R</sup>* represents the zeocin resistance gene. 

The plasmid pWYEN is the gRNA delivery vector. The gRNA was driven by the promoter *P*<sub>ScSNR52</sub> and ended with the terminator *SC SUP4t*. OpADE2 5′ and OpADE2 3′ represent the ~1.5 kb up- and downstream homologous arms of the gene OpADE2, respectively. *G418<sup>R</sup>* represents the G418 resistance gene.

(b) Schematic illustration of markerless genome editing. 

(i) Integration of Cas9. 

(ii) Gene modification. 

(iii) Eviction of the gRNA delivery vector and restoration of the OpADE2 gene. 

(iv) Eviction of the Cas9 delivery vector and restoration of the OpMET2 gene. PAM is the abbreviation of protospacer adjacent motif. ET represents the editing template(s).
cassette at the OpADE2 locus by PCR; (ii) The positive transformants were randomly selected from red colonies to identify the desired mutation at the target site of gRNA.

Because the integration of gRNA delivery vector was mediated by endogenous homologous recombination system (HRS), the fragment OpADE2UHA (~1.5 kb)-G418<sup>R</sup> expression cassette-OpADE2DHA (~1.5 kb) was transformed into OP009 to detect the editing efficiency mediated by the endogenous HRS. As shown in Additional file 1: Figure S1, the editing efficiency of 29.49 ± 4.44% was obtained. Given that 200~300 colonies could be obtained by transformation of 1 µg DNA, appropriately 60~90 red colonies would be obtained in the first step of screening.

After the desired mutation was verified, the gRNA delivery vector containing the marker gene G418<sup>R</sup> was evicted to restore the native OpADE2 locus by transforming with the OpADE2 editing template (Fig. 1b). The resulting transformants were selected on the SC-ADE plate and further confirmed by PCR analysis and DNA sequencing. The mutant could be used for the next round of genome editing.

When all gene manipulations were accomplished, the Cas9 expression vector containing the marker gene zeo<sup>R</sup> was evicted to restore the native OpMET2 locus by transforming with the OpMET2 editing template (Fig. 1b). The resulting transformants were verified on SC-MET plate and further confirmed by PCR and DNA sequencing. The mutant could be used for the next round of genome editing.

Gene deletion

To validate the effectiveness of CRISPR–Cas9 mediated genome editing system in markerless gene deletion, two genes OpLEU2 encoding the 3-isopropylmalate dehydrogenase and OpURA3 encoding the orotidine-5′-phosphate decarboxylase were separately selected as targets. Following the described procedure in Fig. 1b, eight colonies for each mutant were randomly selected from red-color colonies to detect editing efficiencies. Editing efficiencies for deletion of OpLEU2 and OpURA3 were 58.33 ± 7.22% and 65.28 ± 2.41%, respectively (Fig. 2a). In addition, homologous recombination frequencies without CRISPR/Cas9-induced DSB were too low to obtain the desired mutant in the control experiment by co-transforming with the vector without gRNA and the editing template into the OP009 strain (Fig. 2a; Additional file 1: Figure S2A, B). Auxotrophic phenotype analysis showed that OpURA3 or OpLEU2-deficient strains were unable to grow on SC-URA or SC-LEU medium, further confirming inactivation of the OpURA3 or OpLEU2 in O. polymorpha (Additional file 1: Figure S4A, B). Finally, the linearized gRNA delivery vector and the Cas9 protein expression cassette were successively evicted from the chromosome as described procedure in Fig. 1b (Additional file 1: Figure S3A, B).

In addition, the effect of homolog size on editing efficiency was tested. The gene OpADE2 was selected as the target gene for deletion and the gene OpURA3 was selected as the target gene for integration of the gRNA expression cassette (details were described in methods in Additional file 1). Various editing templates with different size of homologous arms (50, 100, 250, 500, 750 and 1000 bp) were used. As shown in Fig. 2b, the editing efficiency increased with the length of homologous arm, which was consistent with the effects of homologous arm sizes on gene editing efficiency in the previous report [31]. No corrected mutant was obtained when the HA was 50 bp, while the editing efficiency reached 62.18 ± 6.17% when HA was 1000 bp (Additional file 11 Figure S5). Furthermore, the editing efficiency decreased drastically when the homologous arm size was less than 500 bp. Therefore, the homologous arm size of editing template should be no less than 500 bp in the design process.

To achieve simultaneous multiple knock-outs in O. polymorpha, three genes OpLIRA3, OpHIS3 and OpLEU2 were selected as the target loci. The UHA and DHA (~1 kb) fragments of each target gene were joined by Splicing Overlapping Extension PCR as editing template. Three different editing templates for deletions of OpLIRA3, OpHIS3 and OpLEU2 were co-transformed into the strain OP009 with the linearized vector pWYE3215 harboring three gRNAs expression cassettes to separately target three genes. The vector without gRNA was used as a control. The desired mutant was obtained with the editing efficiency of 23.61 ± 6.36% by PCR analysis (Fig. 2c and Additional file 1: Figure S6). Auxotrophic phenotype analysis showed that the strain OP045 (OP001ΔOpLEU2ΔOpHIS3ΔOpURA3) was unable to grow on SC-LEU, SC-HIS3 and SC-URA plates, further confirming simultaneously inactivation of the OpLEU2, OpHIS3 and OpURA3 in O. polymorpha (Additional file 1: Figure S7).

Precise point mutation

To test the applicability of the genome editing system in precise nucleotide substitution, G73T mutation was introduced into the OpLIRA3 gene. The mutation site G73 is designed in the third nucleotide of the PAM sequence TGG to eliminate the PAM sequence upon successful mutation, thus preventing the Cas9 nuclease from breaking the mutated sequence. The substitution of nucleotide G by T resulted in the introduction of a stop codon into the OpLIRA3 gene at mutation site, which would make mutants fail to grow on SC-URA plate.
The editing template was a DNA fragment (~2 kb) containing the mutant OpURA3 gene (G73T), and its up- and downstream homologous arms. The editing template and linearized plasmid pWYE3211 (OpADE2upHA-\(P_{Sc.\, SNR52}\)-OpURA3gRNA’-OpADE2downHA) used to transcribe gRNA were co-transformed into the strain OP009 (OP001ΔOpMET2::\(P_{Sc.\, STEF1}\)-Cas9). The resulting transformants were screened on YPD plates containing G418. Desired mutants were selected from transformants by uracil auxotrophic phenotype on SC-URA plates (Additional file 1: Figure S8).

Desired mutants were generated with the editing efficiency of 31.40 ± 4.02% (Fig. 2d). The mutation site was confirmed by the nucleotide sequence of the gene OpURA3 and the uracil auxotrophic phenotype (Additional file 1: Figure S9). To detect whether there were off-target mutations, the web-based computer program CAS-OFFinder was used to predict potential off-target sites (Additional file 1: Table S3). All potential off-target sites in the mutant OP040 (OP001OpUR-A3G73T) were sequenced and the off-target mutation was not detected (Additional file 1: Figure S10). The results suggested that the CRISPR–Cas9-assisted system developed in this study could be used to precisely introduce a nucleotide mutation in the genome.

**Multi-locus gene integration**

Three genes OpLEU2, OpURA3 and OpHIS3 encoding the imidazolglycerol-phosphate dehydratase were selected as target genes to further evaluate the feasibility
of the CRISPR–Cas9 mediated genome editing system for gene integration. The *gfpmut3a* expression cassette flanked by UHA and DHA of the target gene was used as the editing template for single site. The editing efficiencies for integrations at *OpHIS3* and *OpURA3* loci were both 66.70 ± 7.22%, while that at *OpLEU2* locus was 62.50% (Fig. 3a; Additional file 1: Figure S11A–C).

To detect editing efficiencies mediated by endogenous HRS at three loci, the fragment(s) *OpLEU2* (or *OpURA3/HIS3*)UHA (~1 kb)-*G418R* expression cassette-DHA (~1 kb) was transformed into OP009. As shown in Additional file 1: Figure S12A–C, editing efficiencies mediated by endogenous HRS were only 16.03 ± 1.11% (*OpLEU2*), 17.95 ± 4.84% (*OpURA3*) and 16.67 ± 1.11% (*OpHIS3*), respectively. Therefore, CRISPR/Cas9-induced DSB significantly enhances the editing efficiency in *O. polymorpha* (Additional file 1: Figure S-15). Moreover, the expression of the *gfpmut3a* gene in the mutant was further confirmed by flow cytometry analysis (Fig. 3b).

For testing simultaneous integration of multiple genes into the genome, the non-native synthetic pathway of resveratrol was introduced in *O. polymorpha* (Fig. 3d). Three different editing templates for integration of *TAL*, *4CL* and *STS* expression cassettes at *OpURA3*, *OpHIS3* and *OpLEU2* loci were co-transformed into the strain OP009 with the linearized vector pWYE3215 harboring three gRNAs expression cassettes to separately target three loci (Fig. 3c, d). The vector without gRNA was used as a control. The mutant integrating three genes was obtained with the editing efficiency of 30.56 ± 2.40% by phenotypic screening and PCR analysis (Fig. 3e and Additional file 1: Figure S14). The maximum resveratrol production was 4.69 ± 0.17 mg/L in the shake-flask cultivation of OP021 strain (Fig. 3f, g), achieving the first biosynthesis of resveratrol in *O. polymorpha*. Therefore, the CMGE-ML was effective in simultaneous integration of multiple genes into the genome of *O. polymorpha*.

**Multi-copy integration into rDNA cluster**

To conduct multi-copy integration, the expression of Cas9 was controlled by the inducible promoter P_{Max} and the rDNA cluster consisting of 50–60 repeats were selected as integration sites (Fig. 4a). After inducible expression of Cas9, the linearized vector pWYE3220 harboring the gRNA expression cassette targeting rDNA repeats and editing template containing *gfpmut3a* expression cassette flanked by ~1 kb up- and downstream homologous arms to the target site were co-transformed into cells. DSBs at some of rDNA repeats were generated by the slicing of Cas9 and repaired by homologous recombination of donor DNA (Fig. 4b, c). Then, the expression of Cas9 was suppressed by glucose in the resuscitation medium to avoid the slicing of Cas9 at the left rDNA repeats. As a result, mutants harboring *gfpmut3a* expression cassette at rDNA repeats were obtained with the editing efficiency of 75.00 ± 12.5% (Fig. 4d and Additional file 1: Figure S15).

Eight colonies were randomly selected to determine the copy numbers of *gfpmut3a* gene in the genome. The copy numbers ranged from 2.42 ± 0.47 (colony 1) to 11.15 ± 1.10 (colony 8) (Fig. 4e). To detect the expression of *gfpmut3a*, the GFP intensities of mutants harboring different copies of *gfpmut3a* were measured. As a result, all mutants exhibited green fluorescence and the fluorescence intensity showed an increased tendency with the increase of gene copy number (Fig. 4e and Additional file 1: Figure S16).

To evaluate the stability of multi-copy gene in the chromosome, the OP025 strain harboring 11.15 ± 1.10 copies of *gfpmut3a* was cultured in YPD medium without any selective pressure. As shown in Additional file 1: Figure S17, copy numbers of the *gfpmut3a* were constant over 96-h cultivation. Furthermore, to evaluate the stability of multiplex editing in a long-term lab evolution, the strain OP025 was cultured in YPD medium without any selective pressure for 18.5 days (55 generations). As shown in Fig. 4f, copy numbers of the *gfpmut3a* gene were constant after long-term propagation. The result suggested that rDNA repeats provided a useful and stable target site for integration of multiple heterologous genes.

To demonstrate the utility and generality of CMGE-MC, three genes of *TAL*, *4CL* and *STS* were constructed into one fusion expression cassette (P_{ScTEF1} *TAL*-P_{ScPHI} *4CL*-P_{ScTEF2} *STS*), which were then integrated into rDNA repeats. Integrants harboring different copy numbers of the expression cassette were obtained. To analyze the correlation between resveratrol production and copy number of target genes, integrants harboring one to ~10 copies were selected to culture in shake flask (Fig. 5a). As shown in Fig. 5b, the resveratrol production increased...
with the increase of gene copy numbers. The OP043 strain with 9.81 ± 0.55 copies yielded the highest production of 97.23 ± 4.84 mg/L resveratrol, which increased by 20.73-fold compared to that of the single-copy integrant. The result suggested that the more copies of expression cassette were integrated into the genome, the higher level of resveratrol production was achieved. In addition, the resveratrol production increased slowly when the integrated gene reached 7-10 copies, indicating that the resveratrol production in *O. polymorpha* might be limited by other factors, such as the supply of precursor malonyl-CoA [32].

To further assess the broad applicability of CMGE-MC, the *cadA* gene from *E. coli* and the human serum albumin gene *HSA* were separately integrated into the multiple rDNA repeats. Mutants harboring different copy numbers (from 3.12 ± 0.41 to 9.54 ± 0.64) of *cadA* gene were obtained (Fig. 5c). After shake-flask cultivation of the strain OP028 harboring 9.54 ± 0.64 copies in the presence of lysine for 108 h, 2.51 ± 0.18 g/L of cadaverine was detected by HPLC (Fig. 5d, e), which was the first biosynthesis of cadaverine in *O. polymorpha*. Similarly, the copy numbers of the *HSA* gene in different integrants arranged from 2.48 ± 0.42 to 10.24 ± 1.26 (Fig. 5f). The maximum level of HSA was detected to be 97.09 ± 2.45 mg/L in the OP031 strain harboring 10.24 ± 1.26 copies of *HSA* (Fig. 5g). Taken together, the CMGE-MC was a useful and efficient tool for multi-copy integration in *O. polymorpha*.

**Multi-copy integration in *Saccharomyces cerevisiae***

To assess the effectiveness of the multi-copy integration method in other organism, *S. cerevisiae* that is one of the most intensively studied eukaryotic model organisms was selected. The gRNA delivery vector pWYE3225 and the *gfpmut3a* expression cassette flanked by up- and downstream homologous arms were co-transformed into the strain SC006 (SC001/pWYE3224) constitutively expressing Cas9 protein. As a result, the desired integrants were obtained with the editing efficiency of 45.83 ± 7.22% (Fig. 6a, b and Additional file 1: Figure S18). Eight colonies were randomly selected to detect copy numbers of *gfpmut3a* gene. As shown in Fig. 6c, the copy numbers were arranged from 1.25 ± 0.22 (colony 4) to 9.74 ± 0.79 (colony 3). Furthermore, the fluorescence intensity had an increased tendency with the increase of copy number of *gfpmut3a* gene (Fig. 6d). Therefore, the multi-copy integration method was effective in *S. cerevisiae*.

To evaluate the stability of multi-copy genes, the strain SC007 harboring 9.74 ± 0.79 copies of *gfpmut3a* was cultured in YPD medium without any selection force. The copy numbers of the *gfpmut3a* were constant over 96-h cultivation (Additional file 1: Figure S19). In addition, the strain SC007 was cultured in YPD medium without any selective pressure for 18.5 days (55 generations). As shown in Fig. 6e, copy numbers of the *gfpmut3a* were constant after long-term propagation. The results demonstrated that multi-copy genes were highly stable in *S. cerevisiae*, suggesting a possibly broad applicability of this method in various yeasts, in which there are rDNA repeats and the CRISPR–Cas9 can work efficiently.

**Discussion**

Genetic manipulation technologies, especially multiplex genome editing methods, are playing an important role in understanding gene function and developing rational design for biological engineering [22, 27, 33]. Recently, the CRISPR–Cas9 system has been employed as efficient genome editing tools in bacteria, fungi and higher eukaryotes [34–36]. In this study, we developed a CRISPR–Cas9-assisted multiplex genome editing method in yeasts.

The CRISPR–Cas9 system has enabled genome editing in different yeasts (Additional file 1: Table S6) [23, 34, 37]. Although CRISPR–Cas9-mediated genome editing methods have been developed in *O. polymorpha* by Numamoto et al. and Juergens et al. [17, 24], these methods only adapt for a single gene deletion and disruption (Additional file 1: Table S6). In this study, the CRISPR–Cas9-assisted multiplex genome editing method was developed for replacement, point mutation, multiple simultaneous knock-outs, multi-locus and multi-copy integrations of target genes in *O. polymorpha* and multi-copy gene integration in *S. cerevisiae* (Additional file 1: Table S6).

Furthermore, our study for the first time achieved CRISPR–Cas9-mediated markerless multi-copy integration in yeasts using rDNA repeats as the integration loci (Additional file 1: Table S6). Although rDNA repeats have been used extensively to integrate foreign genes in yeasts, selectable markers were required and multi-copy markers...
would be integrated into the chromosomes, which might interfere with the subsequent manipulation [28, 38]. In S. cerevisiae, Ty elements (the transposable elements) were used as the integration loci to integrate markerless multi-copy genes [33]. However, not all yeast strains have Ty elements, and the copy number and distribution on the genome of Ty elements also varied dramatically in different strains [28, 39, 40], suggesting that the available method was not applicable to the strains whose genetic background was not clear. In contrast, the copy numbers of rDNA repeats are highly constant in the same yeast species [28, 38]. Therefore, the multi-copy integration method developed in this study has broad practicability in different yeast stains.

Moreover, our method was efficient for various genetic manipulations in yeasts. To compare the editing efficiency of our method and pervious Cas9-based methods, the same gene OpADE2 was used for gene deletion in O. polymorpha. As a result, the editing efficiency for the deletion of OpADE2 was 62.18 ± 6.17% (Fig. 2b and Additional file 1: Figure S5). However, editing efficiencies of OpADE12 (OpADE2, note S2) deletion in Numamoto’s and Juergens’s methods were 47% and 9%, respectively. In addition, editing efficiencies of different gene disruptions were 17–71% in Numamoto’s method, whereas those of gene deletion, integration, precise point mutation, multi-plex genes knock-outs, multi-locus and multi-copy integration were 23.61–75.00% in our study (Additional file 1: Table S6).

Recently, more and more bio-products have been synthesized in microorganisms, which usually require the integration of multiple genes into the host’s chromosome.

**Fig. 5** The application of multi-copy integration method in O. polymorpha. a Copy numbers of the fusion expression cassette PScTEF1-TAL-PScTP1-4CL-PScTEF2-STS integrated at rDNA cluster. b Resveratrol productions of colonies harboring different copy numbers of the fusion expression cassette PScTEF1-TAL-PScTP1-4CL-PScTEF2-STS. c Copy numbers of cadA integrated at rDNA cluster. d HPLC analysis of cadaverine in O. polymorpha. e Cell growth and cadaverine productions of the wild-type strain OP001 and the mutant OP028 (OP001rDNA::cadA). f Copy numbers of gene HSA integrated at rDNA cluster in eight randomly selected colonies. g HSA productions of the mutant strain OP031 (OP001rDNA::HSA) by shake flask fermentation at different fermentation time.
Fig. 6 Multi-copy integration of gfpmut3a at rDNA cluster in S. cerevisiae. 

a Precise integration site of gfpmut3a at rDNA cluster. 

b Integration efficiencies of gfpmut3a at the rDNA repeats with and without the expression of targeting gRNA. 

c Copy numbers of gfpmut3a in eight randomly selected colonies. 

d Flow cytometry analysis of the expression of gfpmut3a in eight randomly selected colonies. 

e Stability of multi-copy integration at rDNA repeats in the mutant SC007 (SC001 rDNA:gfpmut3a)
In spite of rapid development of genetic engineering, integrations of multiple genes into the genome are still laborious and time consuming. In order to be convenient for constructing the engineering strain, it is necessary for the development of the multiplex gene integration method [37, 42]. The method using rDNA cluster as the target site can mediate simultaneous integration of multiple genes in *O. polymorpha* [27]. However, using this method, many copies of the marker gene *OpLIRA3* would be left in the genome, which might interfere with the subsequent manipulation. In addition, the editing efficiency of this method was relatively low (only 1/11). In this study, the CMGE-ML can simultaneously mediate markerless integrations of three foreign genes at three loci located in three different chromosomes with the editing efficiency of 30.56 ± 2.40%, suggesting that CMGE-ML was an effective tool in *O. polymorpha*.

High expression of gene is usually required to generate a detectable phenotype or obtain high productions of target products [43, 44]. The integration of multi-copy target gene into the genome is an important method for increasing gene expression, especially in the organism without episomal plasmids [28, 45]. In *O. polymorpha*, traditional methods for multi-copy integration include autonomously replicating plasmids-mediated random integration and recombinant plasmid-mediated site-specific integration [14, 46, 47]. Using the random integration method, up to 100 copies of target genes could be integrated at random sites, and using the site-specific integration method, 2–40 copies of target genes could be integrated at specific sites [14, 47]. However, all available multi-copy integration methods in *O. polymorpha* required selectable markers, and the whole vector containing the undesired DNA fragment would be integrated into the chromosome. Consequently, many copies of marker genes and other undesirable genetic elements (e.g., replication origin, plasmid-borne and bacterial antibiotic resistance genes) were left in the chromosome. In this study, editing templates for multi-copy integrations did not contain any antibiotic resistance cassette. Therefore, multi-copy engineered strains not containing any antibiotic resistance cassette would be stable in industrial application. Moreover, ~10 copies of the fusion expression cassette *Pstef1*-TAL-*Pstef1*-4CL-*Pstef2*-STS were integrated into the rDNA cluster via CMGE-MC by one step, achieving a ~20-fold increase of the resveratrol production, indicating that the method was a simple and powerful tool. In addition, the successful biosynthesis of cadaverine and HSA suggested that CMGE-MC had broad practicability.

The CRISPR–Cas9-assisted genome editing system in this study was established by expression of Cas9 and transcription of gRNAs in the chromosome with following advantages: (i) The expressions of Cas9 in the chromosome were more stable than the episomal plasmid which was potentially unstable in *O. polymorpha* (Note S1) and its copy numbers were variable in different cells. (ii) It was convenient for iterative genome editing without the need for construction and curing of the Cas9 expression plasmid every time. In addition, the expression of chromosome-borne Cas9 maintained at the relative low level of Cas9 protein, which might reduce the toxic impact of Cas9 protein on cell growth and metabolism [48–50]. In contrast, plasmid-based CRISPR–Cas9 methods might impose cells a fitness burden to maintain multi-copy plasmids [35].

**Conclusions**

The CMGE developed in this study is effective for markerless genome editing in yeasts. In particular, multiplex gene knock-outs, CMGE-ML and CMGE-MC were powerful tools in fundamental and applied researches. In addition, the successful application of CMGE-MC in *S. cerevisiae* suggests that it might be applicable to a wide range of yeasts.

**Methods**

**Strains, primers and genes**

The strains and plasmids used in this study are listed in Additional file 1: Table S1. *E. coli* EC135 lacking all known restriction-modification (R-M) systems and orphan DNA methyltransferases (MTases) was used as the cloning host [51]. The primer synthesis (Additional file 1: Table S2) and DNA sequencing were performed by Invitrogen company (Shanghai, China) and Beijing Genomics Institute (BGI, Beijing, China), respectively. The *cada* gene encoding lysine decarboxylase from *E. coli*, TAL gene encoding tyrosine ammonia-lyase from *Herpetosiphon aurantiacus*, 4CL gene encoding 4-coumaryl-CoA ligase from *Arabidopsis thaliana*, STS gene encoding stilbene synthase from *Vitis vinifera*, and human serum albumin gene *HSA* were all synthesized by GenScript (Nanjing, China) in codon-optimized versions for expression in *O. polymorpha*.

**Culture and growth conditions**

*Escherichia coli* cells were grown at 37 °C in Luria–Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl). *O. polymorpha* cells were grown at 37 °C in YPD medium (10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose). To induce the expression of Cas9 protein that was controlled by the inducible promoter *Ppmox* in *O. polymorpha*, YPM medium (10 g/L yeast extract, 20 g/L peptone and 5 mL/L methanol) was used [15]. *S. cerevisiae* was grown at 30 °C in YPD medium before transformation. After transformation,
cells were grown in appropriate synthetic complete (SC) medium minus the auxotrophic compound (FunGenome Company, Beijing, China) complemented by the plasmids. To induce the expression of Cas9 protein that was controlled by the inducible promoter \( P_{\text{S GAL1}} \) in \( S. \text{cerevisiae} \), cells were grown in SC medium with 2% galactose but without uracil media until the \( \text{OD}_{600} \) reached 0.5 [52]. When necessary, ampicillin (100 \( \mu \)g/mL for \( E. \text{coli} \)), kanamycin (50 \( \mu \)g/mL for \( E. \text{coli} \)), zeocin (25 \( \mu \)g/mL for \( E. \text{coli} \) or 100 \( \mu \)g/mL for \( O. \text{polymorpha} \)) and G418 (100 \( \mu \)g/mL for \( O. \text{polymorpha} \)) were added to the medium.

**Preparation of competent cell and transformation**

The preparation of electroporetant cells and DNA transformation of \( O. \text{polymorpha} \) were performed following the procedure described by Faber et al. [53]. DNA transformation of \( S. \text{cerevisiae} \) was carried out using LiAc/SS carrier DNA/PEG method [54, 55]. Approximately, 1 \( \mu \)g of plasmid DNA (or 1 \( \mu \)g of gRNA delivery plasmid and 3 \( \mu \)g of editing template) was used per transformation.

**Construction of plasmids and editing templates**

For constitutive expression of Cas9 protein in \( O. \text{polymorpha} \), the DNA fragment containing the Cas9 gene and the promoter \( P_{\text{ScTEF1}} \) was PCR amplified from the plasmid pWYE3202 (pCRCT Addgene plasmid #60621). Up- and downstream homologous arms (~1.5 kb) of the \( \text{OpMET2} \) gene (\( \text{OpMET2-}UHA-\text{DHA} \)) were amplified from the genomic DNA of \( O. \text{polymorpha} \). Three PCR fragments were then Gibson assembled into the \( Bgl II/Xba I \) site of pWYE3200 to generate the plasmid pWYE3208 (pWYE3200-\( P_{\text{ScTEF1}} \)-Cas9-\( \text{OpMET2-UHA-DHA} \)) [15]. Subsequently, the plasmid pWYE3216 (pWYE3200-\( P_{\text{OpMOX}} \)-Cas9-\( \text{OpMET2-UHA-DHA} \)) harboring the methanol-inducible promoter \( P_{\text{OpMOX}} \) from \( O. \text{polymorpha} \), the Cas9 gene and the \( \text{OpMET2-UHA-DHA} \) was constructed in a similar manner. To construct the gRNA delivery vector, the \( \text{OpADE2-UHA} \)-\( \text{DHA} \) (~1.5 kb) from \( O. \text{polymorpha} \), the promoter \( P_{\text{ScSNR52}} \) from \( S. \text{cerevisiae} \), and the synthesized crRNA, 20-bp complementary region (\( N_{20} \)) and \( ScSUP4t \) were assembled into the \( Bgl II/Bam HI \) site of pWYE3201 (pWYE3200 derivative, the zeocin resistance gene \( zeo^R \) was replaced by the G418 resistance gene \( G418^{S9} \)) to generate pWYE3201 (a generic term of all gRNA delivery vector and “N” represents the serial number).

Three gRNA expression cassettes targeting \( \text{OpLEU2} \), \( \text{OpLIRA3} \) and \( \text{OpHIS3} \) genes were constructed into the vector pWYE3215 (pWYE3201 - \( \text{OpLEU2} \)gRNA-\( \text{OpLIRA3} \)gRNA-\( \text{OpHIS3} \)gRNA-\( \text{OpADE2-UHA-DHA} \)) for multiplex genome engineering. The gRNA expression cassettes for \( \text{OpHIS3} \) and \( \text{OpLEU2} \) were PCR amplified from the vectors pWYE3213 (pWYE3201-\( \text{OpHIS3} \)gRNA-\( \text{OpADE2-UHA-DHA} \)) and pWYE3209 (pWYE3201-\( \text{OpLEU2} \)gRNA-\( \text{OpADE2-UHA-DHA} \)), respectively. The two expression cassettes were Gibson assembled into the plasmid pWYE3212 (pWYE3201-\( \text{OpLIRA3} \)gRNA-\( \text{OpADE2-UHA-DHA} \)) linearized with BamHI to generate the plasmid pWYE3215.

In \( S. \text{cerevisiae} \), the expression of Cas9 protein was controlled by the inducible promoter \( P_{\text{S GAL1}} \). The Cas9 gene fragment was Gibson assembled into the vector pWYE3222 (pYES2.0/CT, Invitrogen) linearized with BamHI and EcoRI to generate the plasmid pWYE3224 (pWYE3222-\( P_{\text{S GAL1}} \)-Cas9). The rDNA-gRNA delivery vector pWYE3225 (pWYE3223-rDNA-gRNA) was constructed by ligating the gRNA expression elements into the vector pWYE3223 (pESC-leu2, Addgene plasmid #20120).

For gene deletion or precise point mutation, the UHA and DHA (~1 kb) of the target gene (or of the cut site) were amplified from \( O. \text{polymorpha} \) genomic DNA and jointed by Splicing Overlapping Extension (SOE) PCR. For gene integration, the editing template was PCR amplified from the donor plasmid that was constructed by Gibson assembling the UHA and DHA (~1 kb) of the target gene, the promoter \( P_{\text{ScTEF1}} \) and the desired gene with a synthetic terminator into the vector pWYE3200 (see Additional file 1 for details).

**Editing efficiencies**

The mutants were identified by cell growth phenotype and/or PCR amplification. The editing efficiency was defined as the ratio of the desired mutants to the total tested colonies and calculated using the following formula: Editing efficiency = Number of desired mutants/Number of total tested colonies [58].

**Flow cytometry analysis**

The \( \text{gfpmut3a} \) gene encoding a green fluorescent protein (GFP) with enhanced intense fluorescence was used as the reporter gene. Cells integrating the \( \text{gfpmut3a} \) gene were grown in YPD plate overnight and then inoculated into YPD media at a starting \( \text{OD}_{600} \) of 0.1 for \( O. \text{polymorpha} \) (or \( \text{OD}_{600} \) of 0.1 for \( S. \text{cerevisiae} \)). Cells were harvested after reached the exponential phase, and then washed and resuspended in phosphate-buffered saline (PBS) buffer. The GFP fluorescence intensities were measured by BD FACS CaliburTM flow cytometer equipped with an argon laser (emission at 488 nm and 15 mW) and a 525-nm band-pass filter. For each sample, 30,000 events were collected at a rate of 1000–2000 events per second. Cells without integrating the \( \text{gfpmut3a} \) gene were used as a control to determine the background fluorescence.
Gene copy number estimation

Primers for qPCR are listed in Additional file 1: Table S2. Genomic DNA was extracted using the TIANamp Yeast DNA Kit (Tiangen, China) according to the manufacturer’s protocol. Gene copy numbers were determined by quantitative PCR (qPCR) as described by Kolacek et al. [56]. The OpMOX and ScALG9 genes were used as the references of O. polymorpha and S. cerevisiae, respectively. The plasmid pWYE3227 harboring partial sequences of gfpmut3a, OpMOX and ScALG9 genes was used as the template for standard curves to estimate the copy number of gfpmut3a. Similarly, to estimate the copy number of cadA, HSA or PScTEF2-TAL-PScTPH1-4CL-PScTEF2-STS expression cassette, the vector pWYE3228 harboring partial sequences of OpMOX, cadA, HSA and TAL was used for standard curves. Quantitative PCR was performed using GoTaq qPCR master mix (Promega, USA) in a 20-μL mixture with a LightCycler® 96 Real-Time PCR System (Roche, Switzerland).

Stability detection

To evaluate the stability of multi-copy genes, the integrants were cultivated in nonselective YPD medium by serial-subcultures and continuous culture, respectively. For continuous culture, transformants were cultivated in YPD medium for 96 h. 1 mL cell cultures were collected at the 24-h interval to extract the genomic DNA for estimation of the copy number.

To evaluate the stability of multi-copy integrants in a long-term lab evolution, integrants were cultivated in nonselective YPD medium for 55 generations (18.5 days). Mutants were inoculated into a 500 mL shake-flask containing 50 mL YPD medium. 2 mL culture broth was transferred to 50 mL fresh YPD medium every 8 h and this procedure was repeated fifty-four times. 2 mL cell cultures were collected every 10 generations to extract the genomic DNA for estimation of the copy number.

Potential off-target sites predict by CAS-OFFinder

In the CAS-OFFinder program, there are three key parameters affecting the prediction results: “Mismatch Number”, “DNA Bulge Size”, “RNA Bulge Size”. The “Mismatch Number” can be set as 0–9. “DNA Bulge Size” and “RNA Bulge Size” can be set as 0–2. We finally set the “Mismatch Number” as 3, the “DNA Bulge Size” and “RNA Bulge Size” both as 2 based on these reasons: (i) as the number of mismatches increases, the total number of potential off-target sites dramatically increases as well. So if set the “Mismatch Number” as 9 and sequencing all the potential off-target sites one by one is an almost impossible task. (ii) In general, two mismatches, particularly those occurring in a PAM proximal region, considerably reduced SpCas9 activity. Furthermore, three or more mismatches eliminated detectable SpCas9 cleavage in most loci [57].

Fermentation in shake flasks

*Ogataea polymorpha* cells were cultured in 500 mL shake flasks containing 50 mL YPD media at 37 °C with shaking at 200 rpm. For production of resveratrol, the media were supplemented with 5 mM tyrosine as precursor. Similarly, 50 mM lysine was added to the media to produce cadaverine. Throughout the time course of the experiment, cell cultures were collected at the 12-h interval to detect the OD660 and the concentration of resveratrol or cadaverine in the fermentation samples. The wild-type *O. polymorpha* OP001 was used as a negative control. All experiments and measurements were performed at least in triplicate.

Analysis of product concentration

The concentration of resveratrol was quantified on HPLC (Agilent) equipped with an Eclipse XDB-C18 column (4.6 × 150 mm; Agilent Technologies, USA). The eluent flow was at a constant rate of 1.0 mL/min with 70% reagent A (0.1% phosphoric acid v/v) and 30% reagent B (acetonitrile) [32]. Detection wavelength was set at 304 nm and the column was maintained at 40 °C. The concentration of cadaverine was determined by HPLC as described by Wu et al. [58]. The concentrations of resveratrol and cadaverine from the fermentation samples were calculated from the standard curves using the standards from Sigma, USA. The commercially available QuantiChrom BCG Albumin Assay Kit (DIAG-250) was used to determine the concentration of HSA in culture medium.

Additional file

Additional file 1: Table S1. Strains and plasmids used in this study. Table S2. Primers used in this study. Table S3. Potential off-target sites of CRISPR–Cas9 mediated point mutation in the gene OpURA3. Table S4. Editing efficiencies mediated by CRISPR–Cas9 in *O. polymorpha*. Table S5. Editing efficiencies mediated by endogenous HRS in *O. polymorpha*. Table S6. Editing efficiencies of CRISPR–Cas9-assisted genome engineering methods in different yeasts. Figure S1. Analysis of the editing efficiency mediated by endogenous HRS at the OpADE2 site by cell growth phenotype on YPD and SC without adenine (SC-ADE) plates. Figure S2. PCR identifications of the deletion of OpLEU2 (A) and OpURA3 (B) genes, respectively. Figure S3. Evictions of the linearized gRNA delivery vector and the linearized Cas9 protein expression vector after gene editing. Figure S4. Verification of gene deletions by auxotrophic phenotype analysis. Figure S5. Effect of HA (homologous arm) on editing efficiency mediated by CRISPR–Cas9 in *O. polymorpha*. Figure S6. PCR identifications of simultaneous deletions of genes OpLEU2, OpHIS3 and OpURA3. Figure S7. Verification of multiplex knock-outs by auxotrophic phenotype analysis. Figure S8. The identification of the point mutation by cell growth phenotype. The YPD plate and SC plate without uracil (SC-URA). Figure S9. Verifications of point mutation of the gene OpURA3.
**Figure S10.** DNA sequencing of similar genomic loci of point mutation sites in the gene OPuRA3 of the mutant OP040 (OP001 OPuRA3<sup>2398T</sup>). **Figure S11.** PCR identifications of gfpmut3a expression cassette separately integrated at OpLEU2 (A), OpHIS3 (B) and OpURA3 (C) loci. **Figure S12.** Analysis of editing efficiencies mediated by endogenous homologous recombinant system at the gene OpLEU2 (A), OpHIS3 (B) and OpURA3 (C) sites by cell-growth phenotype. **Figure S13.** Editing efficiencies at three gene sites by two different methods. **Figure S14.** PCR identification of simultaneously multi-loci genomic integration. **Figure S15.** PCR identification of gfpmut3a expression cassette by multi-copy integration at rDNA cluster in O. polymorpha. **Figure S16.** Flow cytometry analysis of the expression of gfpmut3a in eight randomly selected O. polymorpha colonies. **Figure S17.** Stability of multi-copy integration of gfpmut3a at rDNA repeats in the mutant OP025 upon continuous culture for 96 h. **Figure S18.** PCR identification of gfpmut3a expression cassette by multi-copy integration at rDNA cluster in S. cerevisiae. **Figure S19.** Stability of multi-copy integration of gfpmut3a at rDNA repeats in the mutant SC007 upon continuous culture for 96 h.

**Abbreviations**

- O. polymorpha: Ogataea polymorpha, S. cerevisiae: Saccharomyces cerevisiae, HPLC: high-performance liquid chromatography, OD<sub>600</sub>: optical density at wavelength (λ) 600 nm; OD<sub>665</sub>: optical density at wavelength (λ) 665 nm; WT: wild type; OP: Ogataea polymorpha; HA: homologous arm, DHA: downstream homologous arm; UHA: upstream homologous arm.

**Authors’ contributions**

TW, AD and LW conceived and designed the study. LW, AD, YZ, SL, YL, HB, DC, QQ, XS, ZY and XH performed the experiments. LW, AD and YZ performed all data analysis. LW, AD, YZ and TW wrote and revised the manuscript. All authors read and approved the final manuscript.

**Author details**

1. CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China.
2. University of Chinese Academy of Sciences, Beijing 100049, China.
3. CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China.
4. Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China.

**Acknowledgements**

We are grateful to Tong Zhao (Institute of Microbiology, Chinese Academy of Sciences) for excellent technical assistance with the flow cytometry analysis.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of supporting data**

Data supporting the results of the article are included within this manuscript.

**Consent for publication**

All authors have approved the manuscript for publication.

**Funding**

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA17010503), the National Natural Science Foundation of China (31870070), the National Hi-Tech Research and Development Program of China (2014AA021203), and Science and Technology Service Network Initiative (KFJ-STS-QYD-047).

**Publisher’s Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 6 April 2018 Accepted: 26 September 2018 Published online: 10 October 2018

**References**

1. Xu X, Ren S, Chen X, Ge J, Xu Z, Huang H, Sun H, Gu Y, Zhou T, Li J, et al. Generation of hepatitis B virus PreS2-S antigens in Hansenula polymorpha. Virol Sin. 2014;29(6):403–9.
2. Kurylenko OO, Ruchala J, Hyniv O, Abbas CA, Dmytryk KV, Sibirny AA. Metabolic engineering and classical selection of the methyloptrophic thermotolerant yeast Hansenula polymorpha for improvement of high-temperature xylose alcoholic fermentation. Microb Cell Fact. 2014;13:122.
3. Voronovsky AV, Rohulya OV, Abbas CA, Sibirny AA. Development of strains of the thermotolerant yeast Hansenula polymorpha capable of alcoholic fermentation of starch and xylan. Metab Eng. 2009;11(4–5):234–42.
4. Moussa M, Ibrahim M, El Ghazaly M, Rohde J, Gnoth S, Anton A, Keny F, Mueller F. Expression of recombinant staphylokinase in the methyloptrophic yeast Hansenula polymorpha. BMC Biotechnol. 2012;12:96.
5. Wagner JM, Alper HS. Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances. Fungal Genet Biol. 2016;89:126–36.
6. Saraya R, Krikken AM, Kiel JA, Baerends RJ, Veenhuis M, van der Klei J. Novel genetic tools for Hansenula polymorpha. FEMS Yeast Res. 2012;12(3):271–8.
7. Chen Z, Wang Z, He X, Guo X, Li W, Zhang B. Uracilase production by a recombinant Hansenula polymorpha strain harboring Candida utilis uracil gene. Appl Microbiol Biotechnol. 2008;79(4):545–54.
8. Lubyaviov VM, Ananim AV, Malyshay AX, Kang HA, Sibirny AA. Optimization of glutathione production in batch and fed-batch cultures by the wild-type and recombinant strains of the methyloptrophic yeast Hansenula polymorpha DL-1. BMC Biotechnol. 2011;11:8.
9. Oh D-B, Park J-S, Kim MW, Cheon SA, Kim EJ, Moon HY, Kwon O, Rhee SK, Kang HA. Glycoengineering of the methyloptrophic yeast Hansenula polymorpha for the production of glycoproteins with trimannosyl coreN-glycan by blocking core oligosaccharide assembly. Biotechnol J. 2008;3(5):659–68.
10. Gemmill TR, Trimble RB. Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochim Biophys Acta. 1999;1426(2):227–37.
11. Cheon SA, Choo J, Ubiyvovk VM, Park JN, Kim MW, Oh DB, Kwon O, Sibirny AA, Kim JY, Kang HA. New selectable host-marker systems for multiple gene manipulations based on TRP1, MET2 and ADE2 in the methyloptrophic yeast Hansenula polymorpha. Yeast. 2009;26(9):S07–21.
12. Cebollero E, Gonzalez R. Comparison of two alternative dominant selectable markers for wine yeast transformation. Appl Environ Microbiol. 2004;70(12):7018–23.
13. Merckelbach A, Godecke S, Janowicz ZA, Hollenberg CP. Cloning and sequencing of the ura3 locus of the methyloptrophic yeast Hansenula polymorpha and its use for the generation of a deletion by gene replacement. Appl Microbiol Biotechnol. 1993;40:2–3:361–4.
14. Agaphonov MO, Trusthikina PM, Sbnhr JH, Choi ES, Rhee SK, Ter-Avanesyan MD. Vectors for rapid selection of integrants with different plasmid copy numbers in the yeast Hansenula polymorpha DL1. Yeast. 1999;15(7):541–51.
15. Song P, Liu S, Guo X, Bai X, He X, Zhang B. Markerless gene deletion in methyloptrophic Hansenula polymorpha by using mazF as counter-selectable marker. Anal Biochem. 2014;468C:66–74.
16. Qian W, Song H, Liu Y, Zhang C, Niu Z, Wang H, Qiu B. Improved gene disruption method and Cre-loxP mutant system for multiple gene disruptions in Hansenula polymorpha. J Microbial Methods. 2009;79(3):253–9.
17. Numamoto M, Maekawa H, Kaneko Y. Efficient genome editing by CRISPR/Cas9 with a RNA-sgRNA fusion in the methyloptrophic yeast Ogataea polymorpha. J Biosci Biotech. 2017;124(5):487–92.
18. Ronda C, Maury J, Jakociunas T, Jacobsen SA, Nielsen AT, CrEdit. CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae. Microb Cell Fact. 2015;14:97.
19. Fuller KK, Chen S, Loros JJ, Dunlap JC. Development of the CRISPR/Cas9 system for targeted gene disruption in Saccharomyces cerevisiae. J Cell Biol. 2015;14(11):1073–80.
20. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346(6213):1258096.
21. Bao Z, Xiao H, Liang J, Zhang L, Xiong X, Sun N, Si T, Zhao H. Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in Saccharomyces cerevisiae. ACS Synth Biol. 2015;4(5):585–94.
22. Horvitz AA, Walter JM, Schubert MG, Kung SH, Hawkins K, Platt DM, Hernaday AD, Mahaddeke-Uk Eld, Szeto W, Chandran SS, et al. Efficient multiplexed integration of syngeneic alleles and metabolic pathways in yeasts via CRISPR-Cas. Cell Syst. 2015;1(1):88–96.

23. Gao S, Tong Y, Wen Z, Zhu L, Ge M, Chen D, Jiang Y, Yang S. Multiplex gene editing of the Yarrowia lipolytica genome using the CRISPR–Cas9 system. J Ind Microbiol Biotechnol. 2016;43(8):1085–93.

24. Juergens H, Varela JA, Gorter de Vries AR, Perlitz T, Gaste VIM, Gyu Verneye NV, Rajkumar AS, Mans R, Pronk JT, Morrissey JP, et al. Genome editing in Kluyveromyces and Ogataea yeasts using a broad-host-range Cas9/gRNA co-expression plasmid. FEBS Yeast Res. 2018;18(5):012.

25. Thomas DP. Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell. 1980;19(6):774–786.

26. Malesza R, Clark-Walker GD. Yeasts have a four-fold variation in ribosomal DNA copy number. Yeast. 1993;9(1):53–8.

27. Klabinde J, Diesel A, Waschik D, Gelissen G, Hollenberg CP, Suckow M. Single-step co-integration of multiple expressible heterologous genes into the ribosomal DNA of the methylotrophic yeast Hansenula polymorpha. Appl Microbiol Biotechnol. 2002;58(6):797–805.

28. Moon HY, Lee DW, Sim GH, Kwon JY, Kwon MG, Kang BK, Kim JM, Kang HA. A new set of rDNA-NTS-based multiple integrative cassettes for the development of antibiotic-marker-free recombinant yeasts. J Biotechnol. 2016;233:190–9.

29. Wu J, Fu W, Luo J, Zhang T. Expression and purification of human endostatin from Hansenula polymorpha A16. Protein Expr Purif. 2005;42(1):12–9.

30. Leite FC, Dos Anjos RS, Basilio AC, Leal GF, Simoes DA, de Morais MA. Construction of integrative plasmids suitable for genetic modification of industrial strains of Saccharomyces cerevisiae. Plasmid. 2013;69(1):114–7.

31. Gonzalez C, Perdomo G, Tejera P, Brito N, Siverio JM. One-step, PCR-mediated, gene disruption in the yeast Hansenula polymorpha. Yeast. 1999;15(13):1323–9.

32. Li M, Kildegaard KA, Chen Y, Rodriguez A, Borodina I, Nielsen J. De novo production of resveratrol from glucose or ethanol by engineered Saccharomyces cerevisiae. Metab Eng. 2015;32:1–11.

33. Shi S, Liang Y, Zhang MM, Ang EL, Zhao H. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae. Metab Eng. 2016;33:19–27.

34. Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegu M. Implementation of the CRISPR–Cas9 system in fission yeast. Nat Commun. 2014;5:3344.

35. Westbrook AW, Moc-Young M, Chou CP. Development of a CRISPR–Cas9 tool kit for comprehensive engineering of Bacillus subtilis. Appl Environ Microbiol. 2016;82(16):4876–95.

36. Kleinsteuber BP, Pattanyak V, Prew MS, Tsal SG, Nguyen TN, Zheng Z, Joung JK. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016;529(7587):490–5.

37. Schwartz C, Shabbir-Hussain M, Frogue K, Blenner M, Wheeldon I. Standardized markerless gene integration for pathway engineering in fission yeast. Nat Commun. 2014;5:5344.

38. Gidijala L, van der Kleij I, Veenhuis M, Kiel JA. Reprogramming Hansenula polymorpha for penicillin production: expression of the Penicillium chrysogenum pcp gene. FEMS Yeast Res. 2007;7(7):1160–7.

39. Liu K, Liang Y, Ang EL, Zhao H. A new era of genome-integration-cut and paste? ACS Synth Biol. 2017;6(4):601–9.

40. Sun H, Zang X, Liu Y, Cao X, Wu F, Huang X, Jiang M, Zhang X. Expression of a chimeric human/salmon calcitonin gene integrated into the Saccharomyces cerevisiae genome using rDNA sequences as recombination sites. Appl Microbiol Biotechnol. 2015;99(23):10097–106.

41. Liu J, Jin R, Zhao H. Construction of plasmids with tunable copy numbers in Saccharomyces cerevisiae and their applications in pathway optimization and multiplex genome integration. Biotechnol Bioeng. 2016;113(1):2462–73.

42. Steinborn G, Boer E, Scholtz A, Tag K, Kunze G, Geliszen G. Application of a wide-range yeast vector (CoMed) system to recombinant protein production in dimorphic Anula adenosinivorans, methylotrophic Hansenula polymorpha and other yeasts. Microb Cell Fact. 2006;5:33.

43. Sohn JH, Choi ES, Kang HA, Rhee JS, Agaphonov MO, Ter-Avanesyan MD, Rhee SK. A dominant selection system designed for copy-number-controlled gene integration in Hansenula polymorpha DL-1. Appl Microbiol Biotechnol. 1999;51(6):800–7.

44. Terentiev Y, Pico AH, Boer E, Wartmann T, Klabinde J, Breuer U, Babel W, Suckow M, Gelissen G, Kunze G. A wide-range integrative yeast expression vector system based on Anula adenosinivorans-derived elements. J Ind Microbiol Biotechnol. 2004;31(5):223–8.

45. Cho JS, Choi KR, Prabowo CPS, Shin JH, Yang D, Jang J, Lee SY. CRISPR/Cas9-coupled recombineering for metabolic engineering of Corynebacterium glutamicum. Metab Eng. 2017;42:157–67.

46. Peng F, Wang X, Sun Y, Dong G, Yang Y, Liu X, Bai E. Efficient gene editing in Corynebacterium glutamicum using the CRISPR/Cas9 system. Microb Cell Fact. 2017;16(1):201.

47. Liu J, Wang Y, Liu Y, Zheng P, Sun J, Ma Y. Development of a CRISPR/Cas9 genome editing toolbox for Corynebacterium glutamicum. Microb Cell Fact. 2017;16(1):205.

48. Zhang G, Wang W, Deng A, Sun Z, Zhang Y, Liang Y, Che Y, Wen T. A mimicking-of-DNA-methylation-patients pipeline for overcoming the restriction barrier of bacteria. PLoS Genet. 2012;8(9):e1002987.

49. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013;41(7):4336–43.

50. Faber KN, Hairma F, Harder W, Veenhuis M, Ab G. Highly-efficient electrotransformation of the yeast Hansenula polymorpha. Curr Genet. 1994;25(4):305–10.

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

52. Bai H, Deng A, Liu S, Cui D, Oiu Q, Wang L, Yang Z, Wu J, Shang X, Zhang Y, et al. A novel tool for microbial genome editing using the restriction modification system. ACS Synth Biol. 2018;7(1):98–106.

53. Kolacsek O, Pico AH, Boer E, Wartmann T, Klabunde J, Breuer U, Babel W, Suckow M, Gelissen G, Kunze G. A wide-range integrative yeast expression vector system based on Anula adenosinivorans-derived elements. J Ind Microbiol Biotechnol. 2004;31(5):223–8.

54. Terentiev Y, Pico AH, Boer E, Wartmann T, Klabunde J, Breuer U, Babel W, Suckow M, Gelissen G, Kunze G. A wide-range integrative yeast expression vector system based on Anula adenosinivorans-derived elements. J Ind Microbiol Biotechnol. 2004;31(5):223–8.

55. Sohn JH, Choi ES, Kang HA, Rhee JS, Agaphonov MO, Ter-Avanesyan MD, Rhee SK. A dominant selection system designed for copy-number-controlled gene integration in Hansenula polymorpha DL-1. Appl Microbiol Biotechnol. 1999;51(6):800–7.

56. Terentiev Y, Pico AH, Boer E, Wartmann T, Klabunde J, Breuer U, Babel W, Suckow M, Gelissen G, Kunze G. A wide-range integrative yeast expression vector system based on Anula adenosinivorans-derived elements. J Ind Microbiol Biotechnol. 2004;31(5):223–8.

57. Terentiev Y, Pico AH, Boer E, Wartmann T, Klabunde J, Breuer U, Babel W, Suckow M, Gelissen G, Kunze G. A wide-range integrative yeast expression vector system based on Anula adenosinivorans-derived elements. J Ind Microbiol Biotechnol. 2004;31(5):223–8.

58. Wu J, Deng A, Sun Q, Bai H, Sun Z, Shang X, Zhang Y, Liu Q, Liang Y, Liu S, et al. Bacterial genome editing via a designed toxin-antitoxin cassette. ACS Synth Biol. 2018;7(3):822–31.