Regular Article

A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*

Shuobo Shi, Youyun Liang, Mingzi Zhang, Ee Lui Ang, Huimin Zhao

*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States*

**Metabolic Engineering Research Laboratory, Science and Engineering Institutes, Agency for Science, Technology and Research, Singapore, Singapore**

*Corresponding author at: Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States. Fax: +1 217 333 5052.

E-mail addresses: angel@merl.a-star.edu.sg (E.L. Ang), zhaos@illinois.edu (H. Zhao).

1 These authors contributed equally to this work.

**Abstract**

Despite recent advances in genome editing capabilities for the model organism *Saccharomyces cerevisiae*, the chromosomal integration of large biochemical pathways for stable industrial production remains challenging. In this work, we developed a simple platform for high-efficiency, single-step, markerless, multi-copy chromosomal integration of full biochemical pathways in *Saccharomyces cerevisiae*. In this Di-CRISPR (delta integration CRISPR-Cas) platform based on the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated systems (Cas), we specifically designed guide RNA sequences to target multiple delta sites in the yeast genome. The generation of double stranded breaks at the delta sites allowed simultaneous integration of multiple copies of linearized donor DNA containing large biochemical pathways. With our newly developed Di-CRISPR platform, we were able to attain highly efficient and markerless integration of large biochemical pathways and achieve an unprecedented 18-copy genomic integration of a 24 kb combined xylose utilization and (R,R)-2,3-butanediol (BDO) production pathway in a single step, thus generating a strain that was able to produce BDO directly from xylose. The simplicity and high efficiency of the Di-CRISPR platform could provide a superior alternative to high copy plasmid systems and would render this platform an invaluable tool for genome editing and metabolic engineering in *S. cerevisiae*.

© 2015 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

1. Introduction

*Saccharomyces cerevisiae* offers many advantages as a production host due to its robustness, tolerance towards harsh environment conditions, genetic tractability and long history of industrial use (Du et al., 2011; Hong and Nielsen, 2012; Hou et al., 2012; Luo et al., 2015). Moreover, this organism has the additional advantage of being considered a safe organism for the production of food and healthcare products. Thus, *S. cerevisiae* is widely used for the production of fuels, pharmaceuticals, and other value-added chemicals (Hong and Nielsen, 2012; Nevoigt, 2008). This has motivated extensive research in the development of new molecular biology and genome editing tools for metabolic engineering in *S. cerevisiae*.

The majority of metabolic engineering endeavors in *S. cerevisiae* employ non-integrative plasmids. This can be attributed to the diverse array of well-characterized plasmids available for use in *S. cerevisiae* and other advantages including ease of use, portability and high copy numbers (Karim et al., 2013). In particular, high copy 2-micron plasmids routinely used in *S. cerevisiae* can be maintained at 10–50 copies per cell (Romanos et al., 1992), providing a convenient platform for overexpression of heterologous genes. However, there are inherent problems associated with episomal plasmids. These include instability due to plasmid loss, the need to maintain selection pressure in culture, as well as variation in gene expression within the population (Da Silva and Srikrishnan, 2012; Karim et al., 2013), especially during long term and large-scale industrial cultivations in poorly defined media. Due to these limitations, the integration of pathway genes into the genome of *S. cerevisiae* is generally preferred for industrial production. However, the genomic integration of entire biochemical pathways is not trivial. Despite the highly efficient recombination machinery in *S. cerevisiae* (Oldenburg et al., 1997), traditional homologous recombination based methods, which involve site-specific integration of linear DNA fragments flanked by homologous arms, are generally plagued by low
efficiencies (Storici et al., 2003). Similarly for delta integration (Da Silva and Srikrishnan, 2012; Sakai et al., 1990; Shi et al., 2014; Yamada et al., 2010; Yuan and Ching, 2013; Yuan and Ching, 2014), which involves multi-copy integration of heterologous genes at the Ty retrotransposon delta sites in the yeast genome, integration efficiencies and number of integrated copies decline rapidly with the size of the integrated donor DNA (Lee and Silva, 1997; Yamada et al., 2010), limiting the utility of this approach for integrating large biochemical pathways. It has been found that recombination efficiencies increase dramatically when targeted double stranded breaks (DSBs) are made in the chromosomes (Storici et al., 2003; Wingler and Cornish, 2011). This finding has inspired the use of various strategies such as I-sceI homing endonucleases to generate DSBs in the chromosome of S. cerevisiae to facilitate homologous recombination (Storici et al., 2003; Wingler and Cornish, 2011). However, homing endonucleases are highly specific and can only cut their cognate DNA target sites, thus restricting the flexibility of such systems (Wingler and Cornish, 2011).

More recently, several groups have investigated the use of targetable endonucleases such as the Transcription Activator-Like Effector Nucleases (TALEN) (Christian et al., 2010; Li et al., 2011) and the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated systems (Cas) (Bao et al., 2015; DiCarlo et al., 2013; Horwitz et al., 2015; Jakočiunas et al., 2015; Mans et al., 2015) to generate specific DSBs in yeast. These endonucleases afford much greater flexibility and allow the targeting of virtually any DNA sequence. In particular, the CRISPR-Cas system trumps in terms of customizability, requiring only the expression of a guide RNA guide RNA to direct endonuclease activity (Jiang et al., 2013). Although the generation of DSBs using TALEN and CRISPR-Cas significantly increases integration efficiencies, most of these studies only investigated the generation of point mutations or integration of single genes (Bao et al., 2015; Christian et al., 2010; DiCarlo et al., 2013; Jakočiunas et al., 2015; Li et al., 2011). To our knowledge, there were only three endeavors to integrate larger DNA constructs. These involved the single copy integrations of a 6-fragment muconic acid pathway (Horwitz et al., 2015), a single-fragment pyruvate dehydrogenase complex (Mans et al., 2015) and a 3-fragment carotenoid pathway (Ronda et al., 2015). However, the simultaneous integration of multiple copies of pathway genes remains unexamined. As such, there is an urgent need for a convenient platform to achieve high efficiency multi-copy integration of large biochemical pathways.

In this study, we aim to develop a novel approach for multi-copy integration of large pathways into the yeast genome. The requirements set for this platform include multi-copy, high efficiency, one-step and ideally, markerless integration. To achieve this, we harnessed the CRISPR-Cas system to specifically generate DSBs at the delta sites in the chromosomes of S. cerevisiae to increase homologous recombination efficiency. This approach, named Di-CRISPR (delta integration CRISPR-Cas), enabled the single-step, multi-copy and markerless integration of DNA constructs ranging from 8 kb to 24 kb with high efficiencies. To demonstrate the ease with which this system could be applied to metabolic engineering applications, we integrated multiple copies of a combined xylose utilization and (RR)-2,3-butanediol (BDO) production pathway using our novel Di-CRISPR platform. Xylose is the second major constituent of lignocellulosic hydrolysate, and has become an attractive raw material due to its abundance and low cost (Shao et al., 2009; Sun et al., 2012), while BDO is an important chemical with broad industrial applications (Lian et al., 2014; Ng et al., 2012). Using our platform, we were able to obtain a strain capable of producing BDO from xylose in a single step. With the high efficiencies and simplicity of the Di-CRISPR platform, we envision that this platform will be extremely useful for metabolic engineering and synthetic biology studies in S. cerevisiae.

2. Materials and methods

2.1. Yeast transformation

Transformation of pDi-CRISPR plasmid and linearized donor DNA was carried out using the LiAC/SS carrier DNA/PEG method (Gietz and Schiestl, 2007). For all plasmids used, 300 ng of DNA were transformed. In the case of the linearized donor DNA, equimolar amounts of 8 kb (300 ng), 16 kb (600 ng) and 24 kb (900 ng) constructs were used respectively. Following transformation, the cells were washed once with water and transferred to the appropriate SC media and cultivated at 250 rpm, 30 °C. Every 2 days, 1 mL of cell culture was transferred to 1.2 mL of fresh media to ensure cell viability. On day 6, the cells were washed and prepared for flow cytometry analysis.

A proportion of cells was also diluted after transformation and plated on selective and non-selective plates. Cells were allowed to grow for three days before the numbers of fluorescent and non-fluorescent colonies were counted with the help of the SZX16 fluorescent stereo microscope (green fluorescence mode) (Olympus, Tokyo, Japan) for determination of the percentages of GFP-fluorescent colonies. Images of the colonies were obtained with the stereo microscope. The brightness of the green channel was adjusted to account for auto-fluorescence from cells not containing GFP. All images were processed using the same settings. All experiments were conducted in duplicates unless otherwise stated.

2.2. DNA manipulation and plasmid construction

All DNA manipulations were carried out in Escherichia coli DH5α as described by Sambrook and Russell (2001). Oligonucleotides used in this study are listed in Supplementary Table S1. The pDi-CRISPR plasmid was modified from the pCRCT plasmid described by Bao et al. (2015) to replace its guide RNA array with delta-specific guide RNA. For this, the oligonucleotides delta-grna1_bsaI/delta_gRNA1_bsaIR encoding the delta-specific guide RNA were mixed and incubated to form duplex DNA. The resultant duplex and pCRCT were then digested with BsaI and ligated to generate pDi-CRISPR. The pDi-CRISPR plasmid was verified by sequencing.

The pRS416 plasmid containing the three-enzyme (RR)-2,3-butanediol (BDO) biosynthetic pathway (Bacillus subtilis acetolactate synthase, R.R-butanediol dehydrogenase and Enterobacter aerogenes acetolactate decarboxylase) was constructed using the DNA assembler method that is based on homologous recombination in yeast (Shao et al., 2009). The alsS, budA and bhaA genes encoding for acetolactate synthase, acetolactate decarboxylase and R.R-butanediol dehydrogenase, respectively, were amplified from the genomic DNA by polymerase chain reaction (PCR), respectively, and assembled into helper plasmids (Du et al., 2012) each containing the respective yeast promoter and terminator using the NEB Gibson Assembly kit (New England Biolabs Inc., Massachusetts, USA). The three expression modules: ADH1p-alsS-ADH1t, PGK1p-budA-CYC1t, PYK1p-bhaA-ADH2t were individually PCR amplified with homologous arms, purified (QiAquick Gel Extraction Kit, Qiagen), and co-transformed into S. cerevisiae along with the BaniHI-digested pRS416 plasmid backbone to assemble several elements in one step for the construction of plasmid pC77.

Two DNA fragments, Delta1 and Delta2, each containing delta-sequence element were PCR amplified from S. cerevisiae HZB848 genomic DNA. The GFP expression cassette 1, TEF1p-GFP-TEF1t, was PCR amplified from plasmid pRS426-tef1p-gfp-tef1t (Sun et al., 2012). These three fragments were assembled into plasmid pRS423 linearized with CfiI and EcoRI via Gibson Assembly resulting in plasmid pS599. Plasmid pS599 was linearized by AvrII.
digestion, and adopted as donor DNA 8 kb. The BDO expression cassette was PCR amplified from plasmid pC77 and assembled into Sall/Bsp120I digested plasmid pSS99 via Gibson Assembly to generate pSS100. Plasmid pSS100 was linearized by AvrII digestion, and adopted as donor DNA 16 kb. The xylose utilization cassette was amplified from plasmid pRS426m-xylose-zeaxanthin (Shao et al., 2009) and assembled into SacI/BamHI digested plasmid pSS100 via Gibson Assembly to generate pSS101. Plasmid pSS101 was linearized by AvrII digestion, and adopted as donor DNA 24 kb.

GFP expression cassette 2 was also amplified from plasmid pRS426-tesf-gfp-tesf1t, and integrated at the LYP1 site of S. cerevisiae HZ848 to generate a control strain for qPCR analysis.

2.3. Strains and media

E. coli strain DH5α used for plasmid construction was grown in Luria-Bertani (LB) medium in the presence of ampicillin (100 mg/L) at 37 °C. Yeast strains used in this work are listed in Supplementary Table S2. Yeast strains were grown on complete synthetic (SC) medium consisting of 0.17% yeast nitrogen base, 0.5% ammonium sulfate, and the appropriate amino acid drop out mix (CSM-COMPLETE, CSM-HIS, CSM-URA, or CSM-HIS-URA, Formedium LTD, Hunstanton, England), supplemented with glucose or xylose as carbon source. Deletion of LYP1 was performed by replacement with the GFP expression cassette 2, and plates supplemented with 250 μg/ml thialysine (S-2-aminoethyl-L-cysteine) were used to select for LYP1 disrupted cells. Unless otherwise indicated, all chemicals were purchased from Sigma-Aldrich (St Louis, MO, USA).

2.4. Flow cytometry analysis

The GFP fluorescence intensities were measured by flow cytometry using the Becton Dickson LSR II flow cytometer (Franklin Lakes, NJ, USA). Briefly, cells were washed and resuspended in phosphate-buffered saline at an optical density between 0.1 and 0.2. Cells not containing GFP were used as a non-fluorescent control and the percentage of GFP-positive cells and fluorescence intensities were determined.

2.5. Copy number estimation

For genomic DNA extraction, cells were cultured overnight in YPD and extracted using the Wizard genomic DNA kit (Promega, Madison, WI, USA) according to the manufacturer’s protocol. Gene copy numbers were determined by quantitative PCR (qPCR) using the yeast total DNA extracts. The GFP gene on the integration cassette and ALG9 gene on the chromosome were chosen as the target and reference genes, respectively. Strain containing one copy of GFP and ALG9 was adopted as the template for standard curves.

The copy numbers were quantified by comparing the C_{q} values of the target and reference genes using a previously described method (Shi et al., 2014). qPCR was performed using FastStart Essential DNA Green Master (Roche, Basel, Switzerland) on a Roche Light Cycler® 96 System. Oligonucleotide primers for qPCR are listed in Supplementary Table S1. They were designed and chosen from DNA fragments utilizing the program Primer 3 Version 0.4.0 (Rozen and Skaletsky, 2000).

2.6. HPLC analysis

Extracellular metabolites, including glucose, acetoin and BDO, were detected and quantified using Shimadzu Prominence UFLC (Kyoto, Japan) equipped with a Phenomenex Rezex ROA-Organic Acid H+ (8%) column (Torrance, CA, USA) and a Shimadzu RID-10A refractive index detector. The column was maintained at 70 °C and 5 mM sulfuric acid was used as the mobile phase with a constant flow rate of 0.6 mL/min.

2.7. Functional analysis of the integrated pathways

The recombinant yeast strains for BDO production were pre-cultured in the SC-HIS-URA medium for 2–3 days until saturation, and then inoculated in fresh SC medium with a dilution ratio of 0.1. After 2 days, cells were harvested, washed, and re-suspended at OD_{600}=10 in the SC-HIS-URA medium with varied concentrations of glucose or xylose as carbon source. The fermentation for BDO production was carried out in sealed deep well plates for 5 days. Supernatants were harvested after centrifugation and injected into HPLC for analysis.

The recombinant S. cerevisiae carrying the xylose utilization pathway was streaked in SC agar plates supplemented with 2% D-xylose as the sole carbon source. The exponential growth rate on xylose or glucose was calculated by log-linear regression analysis of the biomass versus cultivation time.

3. Results

3.1. Di-CRISPR design and optimization

Our Di-CRISPR platform entailed generation of CRISPR-Cas mediated DSBs at the Ty retrotransposon delta sites in the yeast genome, coupled with homologous recombination of donor DNA at the DSBs, to achieve high-efficiency and multi-copy integration of full biochemical pathways in a single step (Fig. 1A). To develop this platform, we carried out a preliminary assay to evaluate the integration efficiencies with two different CRISPR-Cas systems developed for use in S. cerevisiae. The first system was the “regular CRISPR-Cas” described by DiCarlo et al. (2013) and the second was based on the HiCRISPR system developed by Bao et al. (2015). The HiCRISPR system comprised of an ultrahigh copy plasmid containing a truncated URA3 promoter and two point mutations in the Cas9 gene (Bao et al., 2015). In our preliminary assay, we observed improvements in integration efficiencies when the HiCRISPR system was used as compared to the “regular CRISPR-Cas” (Fig. S1). This implied that high Cas9 expression or guide RNA was central to the success of our strategy. Previous studies have demonstrated that Cas9 had low turnovers and remained tightly bound to double-stranded DNA even after inducing DSBs in vitro (Sternberg et al., 2014). This could explain the need for high Cas9 expression in order to generate the multiple cuts required in our Di-CRISPR platform. Having selected a CRISPR-Cas system, we constructed a pDi-CRISPR plasmid which comprised of a delta-targeting guide RNA and a Cas9 protein (Fig. 1B). To allow homologous recombination at the cut delta sites, we introduced 2 halves of the delta sequences at the ends of the 8 kb, 16 kb and 24 kb linear donor DNA fragments. The donor DNA comprised of different biochemical pathways and GFP (Fig. 1B) to provide an easy readout of the relative integration efficiencies and copy numbers. Integration of the donor DNA at the DSBs disrupts the guide RNA recognition site in the delta sequences.

3.2. Characterization of Di-CRISPR

Following transformation of the pDi-CRISPR plasmid together with linearized donor DNA ranging from 8 kb to 24 kb, cells were cultured in selective media to select for both the pDi-CRISPR plasmid and integration of the linear donor, and were passaged every 2 days and analyzed with the flow cytometer on day 6. The 6-day incubation period was chosen as it was found that fluorescence peaked and plateaued on that day presumably due to a delay in recovery and GFP
expression. As a comparison, we carried out conventional delta integration for the series of donor DNA without inducing DSBs. With conventional delta integration, it was possible to integrate the 8 kb fragment albeit with moderate integration efficiencies as shown by the flow cytometry analysis (Fig. 2A–D). However, very few integration events were observed with 16 kb and 24 kb fragments as shown by the low percentages of GFP-positive cells. With the use of our Di-CRISPR system, we were able to achieve dramatic improvements in all measurements (Fig. 2A). For the integration of the 8 kb, 16 kb and 24 kb donor DNA, the mean fluorescence intensities (Fig. 2B), percentages of GFP-positive cells (Fig. 2C) and fluorescence intensities corresponding to 95th percentile of the respective cell populations (Fig. 2D) were significantly higher for the Di-CRISPR processed cells. This indicated that our Di-CRISPR approach was superior in terms of integration efficiencies and copy numbers when compared to conventional delta integration. Using a fluorescent microscope, we further randomly picked 8 fluorescent colonies per condition and determined the average copy numbers of GFP by qPCR. The average copy numbers for the conventional delta integration method decreased from 2.6 to 1.7 as the size of the donor DNA increased from 8 kb to 24 kb, whereas much higher copy numbers were obtained with our Di-CRISPR approach (Fig. 2E). The highest copy numbers obtained with the 8 kb, 16 kb and 24 kb donor DNA were 11, 10 and 10 respectively. These results demonstrated the
effectiveness of our Di-CRISPR approach for the integration of large pathways up to 24 kb with high efficiency and copy numbers. As with the conventional delta integration method, we observed a decrease in fluorescence and copy numbers for larger donor DNA. The average copy number for the 8 kb donor DNA was approximately 1.2-fold that of the 24 kb donor DNA, thus implying a slight decrease in integration efficiency as the size of donor DNA increased. This was accompanied by a 2.7-fold difference in mean fluorescence.

Given the high integration efficiency achieved with the Di-CRISPR platform, we postulated that given the deleterious nature of unrepaired DSBs in S. cerevisiae genome (Oza et al., 2009), there will be an intrinsic selection for cells that have repaired Di-CRISPR-induced DSBs by homologous recombination with the available donor DNA. To test this, we used the same constructs for integration and plated the cells on non-selective plates. The number of fluorescent colonies versus non-fluorescent colonies on each plate was then quantified. The percentages of fluorescent colonies were comparable for the selective and non-selective plates (Fig. 3), confirming that the Di-CRISPR system could be employed for markerless genomic integration, whereas the percentages of fluorescent colonies in conventional delta integration dropped to almost zero.

Fig. 2. Comparison of the Di-CRISPR system with conventional delta integration. (A) Flow cytometry analysis. Conventional delta integration without DSBs and Di-CRISPR were carried out for 8 kb, 16 kb and 24 kb donor DNA. The histograms revealed that GFP fluorescence was markedly higher for Di-CRISPR populations as compared to conventional delta integration populations. (B) Mean fluorescence intensities. Di-CRISPR populations had higher mean fluorescence intensities than their delta integration counterparts. (C) Percentages of GFP-fluorescent cells. Di-CRISPR populations had higher percentages of GFP-fluorescent cells than their delta integration counterparts. (D) Fluorescence intensities corresponding to 95th percentile of the respective populations. Di-CRISPR populations had higher 95th percentile intensities than their delta integration counterparts. (E) Copy numbers. The copy numbers of 8 random fluorescent colonies were determined by qPCR and the average copy number is depicted by the blue dotted line. Di-CRISPR cells had higher copy numbers than their delta integration counterpart. In general, the copy numbers for the 8 kb donor DNA were higher than that of the 16 kb and 24 kb donor DNA. For figures B and C, the values and error bars represent means and standard deviations for 2 transformations. For figure D, the values represent maximum readings from 2 transformations. For figure E, the values and error bars represent average readings and standard deviations for 3 qPCR reactions. For B, C and D, * indicates statistically significant differences between Di-CRISPR and conventional delta integration populations (student’s t-test, p < 0.05).
As the amounts of donor DNA used in this study (900 ng of 24 kb DNA or 0.06 pmol) were low, we further investigated whether increasing the amounts of donor DNA could enhance integration efficiencies. To this end, we carried out Di-CRISPR integration for the 24 kb construct with different amounts of donor DNA, namely 1-fold (900 ng), 3-fold and 10-fold (Fig. 4A–E). Increasing the amounts of donor DNA 3-fold did not yield clear differences in all measurements. However, increasing donor DNA 10-fold resulted in ~15% more GFP-positive cells (Fig. 4C) with the mean fluorescence increasing by 30% (Fig. 4B). Additionally, we observed ~20% increase in the fluorescence intensity corresponding to the 95th percentile of the respective populations (Fig. 4D) and average copy numbers also increased by 1.5-fold (Fig. 4E). When 0.6 pmol (9 μg) of donor DNA was used in our Di-CRISPR system, we were able to integrate up to 18 copies of the 24 kb pathway with a single round of integration. These enhancements indicated that homologous recombination rather than generation of DSBs by CRISPR was likely to be limiting in our system.

3.3. Functional analysis of the integrated pathways

To demonstrate the utility and generality of the Di-CRISPR system in metabolic engineering applications, we used the system for the engineering of another S. cerevisiae strain (CEN.PK2-1C). As a proof of concept, a 16 kb donor DNA containing the BDO biosynthetic pathway was integrated into S. cerevisiae CEN.PK2-1C without the use of selection markers. To further demonstrate the capacity and versatility of Di-CRISPR facilitated integration for introduction of even longer biochemical pathways, the DNA cassette (24 kb donor DNA) containing both the BDO production pathway and the xylose utilization pathway was also integrated into CEN.PK2-1C. The metabolic pathways for BDO production and xylose utilization are shown in Fig. S2.

We randomly selected three colonies from each condition, and analyzed their ability to produce BDO using different sugar solutions at varied concentrations (Table 1). Generally, when glucose was used as the carbon source, strains containing 16 kb donor DNA could produce more BDO than strains containing 24 kb donor DNA. In particular, strain BDO2 produced up to 12.51 g/L BDO from 80 g/L glucose. As the xylose utilization genes were only present in the 24 kb donor DNA, only strains containing the 24 kb donor DNA could use xylose to produce BDO (Table 1). Table S3 showed the corresponding BDO yield of the engineered strains. Beside BDO production, the specific growth rates of the strains were also measured to evaluate the ability to use glucose or xylose (Table 1).

The copy numbers of integrated genes in these 6 transformants varied from 4.4 to 10.3 (Table 1). In general, a higher copy number gave a higher BDO production in strains containing the same donor DNA. However, strains containing 24 kb donor DNA (i.e., BDO2, BDO4 and BDO8) showed lower BDO production than strains containing 16 kb donor DNA (i.e., BDO10, BDO11 and BDO14) even when BDO10 and BDO14 had higher copy numbers than BDO4. In addition, we also studied the effect of xylose utilization pathway copy number on the ability of these 6 transformants to use xylose. The strains were streaked on agar plates with either glucose or xylose as the sole carbon source, and only strains containing 24 kb donor DNA grew on xylose (Fig. S3). As shown in Table 1, strains with higher copy numbers of 24 kb donor DNA containing the xylose utilization pathway exhibited a higher growth rate on xylose, thus suggesting that the ability to use xylose is proportional to the copy numbers. Finally, we tested the long-term genetic stability of the integration strain (BDO2). After more than 100 generations of cell division without selection, the strain still produced about 3.02 g/L BDO from 20 g/L glucose, over 90% of the amount obtained from a fresh transformant of strain BDO2 (Table 1).

4. Discussion

While delta integration was useful for multi-copy integration of single genes, integration of long biochemical pathways remains a challenge (Yamada et al., 2010; Yuan and Ching, 2013; Yuan and Ching, 2014). It was reported that integration efficiencies using homologous recombination increased dramatically when targeted DSBs were introduced in the chromosomes (Storici et al., 2003; Wingler and Cornish, 2011). This inspired our use of CRISPR-Cas to generate DSBs in the delta sites of S. cerevisiae chromosome in order to increase integration efficiencies and integrated copy numbers of large DNA fragments.

The benefits of our work were notable in terms of significantly improved efficiency and multi-copy integration of large pathways. As shown in Fig. 2, and in accordance with previous reports (Kuhlman and Cox, 2010; Shao et al., 2009), integration efficiencies for traditional delta integration ranged from modest to low and decreased sharply for larger donor DNA (Fig. 2A and C). Such low efficiencies would limit the application of delta integration in the production of many value-added compounds, whose biosynthetic pathways are usually long and involve several genes. In contrast, the Di-CRISPR approach has the potential to overcome this limitation by increasing metabolic efficiencies for the engineering of another S. cerevisiae strain. In this study, we randomly selected three colonies from each condition, and analyzed their ability to produce BDO using different sugar solutions at varied concentrations (Table 1). Generally, when glucose was used as the carbon source, strains containing 16 kb donor DNA could produce more BDO than strains containing 24 kb donor DNA. In particular, strain BDO2 produced up to 12.51 g/L BDO from 80 g/L glucose. As the xylose utilization genes were only present in the 24 kb donor DNA, only strains containing the 24 kb donor DNA could use xylose to produce BDO (Table 1). Table S3 showed the corresponding BDO yield of the engineered strains. Beside BDO production, the specific growth rates of the strains were also measured to evaluate the ability to use glucose or xylose (Table 1).

The copy numbers of integrated genes in these 6 transformants varied from 4.4 to 10.3 (Table 1). In general, a higher copy number gave a higher BDO production in strains containing the same donor DNA. However, strains containing 24 kb donor DNA (i.e., BDO2, BDO4 and BDO8) showed lower BDO production than strains containing 16 kb donor DNA (i.e., BDO10, BDO11 and BDO14) even when BDO10 and BDO14 had higher copy numbers than BDO4. In addition, we also studied the effect of xylose utilization pathway copy number on the ability of these 6 transformants to use xylose. The strains were streaked on agar plates with either glucose or xylose as the sole carbon source, and only strains containing 24 kb donor DNA grew on xylose (Fig. S3). As shown in Table 1, strains with higher copy numbers of 24 kb donor DNA containing the xylose utilization pathway exhibited a higher growth rate on xylose, thus suggesting that the ability to use xylose is proportional to the copy numbers. Finally, we tested the long-term genetic stability of the integration strain (BDO2). After more than 100 generations of cell division without selection, the strain still produced about 3.02 g/L BDO from 20 g/L glucose, over 90% of the amount obtained from a fresh transformant of strain BDO2 (Table 1).
method showed a remarkable increase in the percentage of GFP-positive cells (Fig. 2A and C). The enhancement in efficiency was more obvious in the integration of larger DNA. Although we did not explore the upper limits of the insert size, our method enabled efficient integration of 24 kb DNA, and has the potential to integrate even larger DNA at a reasonable efficiency. In addition, a 24 kb construct would allow the expression of 6 to 8 genes in yeast, and is sufficient for most reported metabolic pathways (Du et al., 2011; Hong and Nielsen, 2012; Hou et al., 2012; Luo et al., 2015). If the minimal promoter and terminator developed by the Alper group were used (Redden and Alper, 2015), even more genes could be included in the 24 kb donor DNA. Beside the percentage of GFP-positive cells, the fluorescence of the different cell populations dramatically increased when Di-CRISPR was used (Fig. 2A–D), with more obvious improvements for 16 kb and 24 kb integration events. The enhancements suggested that Di-CRISPR generated strains with more integrated

![Figure 4](image-url)
levels (Parekh et al., 1996; Plessis and Dujon, 1993). These DNA integration had been observed to increase with elevated DNA integrated into multiplex genomic loci. However, compared to these much larger DNA (up to 24 kb) at single loci.

in a single step. Second, our system could be used to integrate CRISPR enabled integration of multiple copies of pathway genes and will be a useful feature in the guide to tune the copy numbers and expression of target genes number with elevated amounts of donor DNA could be used as a fold amount of donor DNA was used (Fig. 4). The increase in copy enhancements were also observed in the present study when 10-

without the need for selection markers (Bao et al., 2015; reports have also shown that CRISPR-Cas worked ef

clearly linear relationship. This phenomenon was also reported in a previous study for 1,2-propanediol production (Lee and Da Silva, 2006).

The best BDO producer we constructed was strain BDO2, and it could produce 3.22 g/L BDO from 20 g/L glucose and 12.51 g/L BDO from 80 g/L glucose in sealed 96 deep well plates. Genetic stability of the strain was also demonstrated by long-term culture without selection pressure. Such high genetic stability is typical of integration strains (Cho et al., 1999; Kim et al., 2001) and is an important feature for industrial applications. Currently, the best S. cerevisiae BDO producer reported can produce up to 100 g/L BDO using fed-batch fermentation by introducing a BDO biosynthetic pathway into a pyruvate decarboxylase-knockout (PDC−) strain (Lian et al., 2014). However, when comparing BDO production by PDC+ yeast strains, our BDO2 strain is the best BDO producer with the highest reported titer and yield of BDO in glucose (Ng et al., 2012). The Di-CRISPR system, for the first time, was also used to construct strains (i.e. strain BDO10, BDO11, and BDO14) capable of producing BDO from xylose, with strain BDO14 producing 0.31 g/L BDO from xylose. The function of the integrated xylose utilization pathway in the 24 kb donor DNA was clearly demonstrated by growing the cells with only xylose (Table 1 and Fig. S3). While there was no relationship between copy numbers and specific growth rates on glucose, there was an obvious positive relationship between copy numbers and specific growth rates on xylose. In summary, the single step integration of the combined BDO production and xylose utilization pathway provided evidence for the usefulness of the Di-CRISPR system for multiple-copy, markerless, integration of large pathways in a single step and would likely work well for the integration of other biochemical pathways in yeast. Although we only examined the integration of single constructs at individual loci, the Di-CRISPR system should be generalizable to multi-fragment integrations to allow fine-tuning of expression levels.

5. Conclusion

Metabolic engineering requires the introduction of multi-enzyme pathways and stable expression of the associated genes.
It is advantageous but challenging to simultaneously integrate multiple copies of pathway genes. Here we demonstrated the ability of the Di-CRISPR system to achieve this purpose in a simple way. Compared to previous efforts, Di-CRISPR is the first reported method to combine delta integration and CRISPR-Cas facilitated DSBs for markerless multi-copy integration. Di-CRISPR was used to integrate large pathways up to 24 kb with high efficiency and multiple copies in a single step. It was also revealed here that the integration could be further enhanced with greater amounts of donor DNA used in the transformation. As a proof of concept, the method was used to construct strains that can produce BDO from glucose or xylose. With its ease of use, utility in markerless genomic integration and unparalleled efficiency and efficacy, we envision that Di-CRISPR would become an invaluable tool in the genome editing toolbox for S. cerevisiae and would finally provide a stable alternative to high-copy.

Acknowledgments

We thank Agency for Science, Technology, and Research, Singapore for supporting various research projects in the Metabolic Engineering Research Laboratory (MERL) through the Visiting Investigator Programme to H.Z.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2015.10.011.

References


