

# ACTivE: Assembly and CRISPR-targeted *in vivo* Editing for Yeast Genome Engineering Using Minimum Reagents and Time

Koray Malcı<sup>1,2</sup>, Nestor Jonguitud-Borrego<sup>1,2</sup>, Hugo van der Straten<sup>1</sup>, Urtë Puodžiūnaitė<sup>1,2,3</sup>, Emily J. Johnston<sup>2,3</sup>, Susan J. Rosser<sup>2,3</sup>, Leonardo Rios-Solis<sup>1,2,4\*</sup>

<sup>1</sup> Institute for Bioengineering, School of Engineering, University of Edinburgh, Edinburgh EH9 3DW, UK

<sup>2</sup> Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Edinburgh EH9 3BF, UK

<sup>3</sup> School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK

<sup>4</sup> School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

\* Corresponding author: Dr. Leonardo Rios Solis (Email: [leo.rios@newcastle.ac.uk](mailto:leo.rios@newcastle.ac.uk))

## ABSTRACT

Among the numerous genetic tools developed for yeast, CRISPR/Cas system has been a widely used genome editing method thanks to its sophistication. However, CRISPR methods for yeast generally rely on pre-assembled DNAs and extra cloning steps to deliver gRNA, Cas protein, and donor DNA. These laborious steps might hinder its usefulness. Here, we propose a convenient, rapid, standardizable CRISPR method, named Assembly and CRISPR-targeted *in vivo* Editing (ACTivE), which only relies on *in vivo* assembly of linear DNA fragments for both plasmid and donor DNA construction. Thus, depending on the user's need, these parts can be easily selected and combined from a repository, serving as a toolkit for rapid genome editing without any expensive reagent. The toolkit contains verified linear DNA fragments, which are easy to store, share and transport at room temperature, drastically reducing expensive shipping costs and assembly time. After optimizing this technique, eight ARS-close loci in the yeast genome were also characterized in terms of integration and gene expression efficiencies and the impacts of the disruptions of these regions on cell fitness. The flexibility and multiplexing capacity of the ACTivE were shown by constructing  $\beta$ -carotene pathway. In only a few days, > 80% integration efficiency for single gene integration and > 50% integration efficiency for triplex integration were achieved from scratch without using *in vitro* DNA assembly methods, restriction enzymes, or extra cloning steps. This study presents a standardizable method to be readily employed to accelerate yeast genome engineering and provides well-defined genomic location alternatives for yeast synthetic biology and metabolic engineering purposes.

**Keywords:** *Saccharomyces cerevisiae*, CRISPR toolkit, genome editing, synthetic biology, standardization, locus characterization

## INTRODUCTION

Being a eukaryotic chassis, *Saccharomyces cerevisiae* has been extensively studied to produce high-value products from pharmaceuticals<sup>1–4</sup> to biofuels.<sup>5–8</sup> As a versatile and efficient genome engineering tool, the clustered regularly interspaced short palindromic repeats (CRISPR) system has been a widely used method to engineer yeast cell factories.<sup>9,10</sup>

In the last decade, a great number of CRISPR-based methods have been developed for yeast genome editing or cell factory development.<sup>11,12</sup> Mainly, delivery or expression of CRISPR-associated (Cas) protein, guide RNA (gRNA), and donor DNA differ in these yeast-specific CRISPR methods. Generally, Cas protein (generally Cas9) which is responsible for the nuclease activity on a specific genomic location<sup>13,14</sup> is expressed through a plasmid vector<sup>15,16</sup> or genomic integration.<sup>17,18</sup> The latter needs an additional transformation for genomic integration and might lead to a burden on the host. gRNA forms a complex with Cas protein and guides it towards the target sequence<sup>13,14</sup> and it is generally expressed by using plasmid vectors in the yeast. A single plasmid containing the genes of Cas protein and gRNA<sup>16,19</sup> or separate independent plasmids for each can be used with an additional selective marker.<sup>20,21</sup> Donor DNA is used as a DNA repair template for homology-directed repair (HDR)<sup>22</sup> after a double-strand break (DSB) is formed by Cas/gRNA complex. When it comes to delivery of donor DNA, single-strand oligos<sup>16</sup>, double-strand oligos<sup>23</sup>, or linear DNA fragments containing overlapping regions for *in vivo* assembly<sup>24</sup> can be used as well as linearized plasmids containing pre-assembled expression cassettes.<sup>25</sup>

Even though high efficiencies in genome editing or pathway construction were achieved with these methods, further characterization is still needed to reach a full consensus.<sup>11,12</sup> On the other hand, the main obstacle in these methods is the requirement of *in vitro* DNA assembly and additional cloning steps to construct CRISPR plasmids. Golden Gate assembly<sup>26</sup>, Gibson assembly<sup>® 27</sup>, or uracil-specific excision reaction-based cloning (USER cloning)<sup>28</sup> are the most widely used *in vitro* plasmid construction methods<sup>16,23,25,29–31</sup>. For instance, Ronda et al. (2015) developed the CrEdit method standing for CRISPR/Cas9 mediated genome editing, using the USER cloning to construct the CRISPR plasmids.<sup>25</sup> In another study, a Csy4-assisted CRISPR method was used for multiple genome editing, and the plasmid constructs were cloned using Gibson Assembly.<sup>32</sup> However, the use of these techniques needs additional steps making the CRISPR process more labor-intensive and time-consuming. Moreover, these methods require relatively expensive DNA assembly kits<sup>33,34</sup> or specific type IIS restriction enzymes that might hinder design flexibility as the DNA fragment to be cloned might contain the recognition sites. Therefore, a collection of ready-to-use DNA parts skipping these pre-assembly steps can accelerate the genome editing process. Once this collection is obtained, it can also minimize the PCR deviations caused by the use of different DNA polymerases, which can be an important problem even though there are limited reports on this issue.<sup>35</sup> Finally, this toolkit based on verified DNA parts is more stable and practical to transport at room temperature for a prolonged period than glycerol stocks or agar slants. It can drastically reduce shipping costs while improving access to lower-income communities to contribute to synthetic biology's democratization.

In addition to how CRISPR is carried out, its genome editing efficiency is another critical issue. The sequence features of CRISPR RNA (crRNA), which has a base complementarity to the target DNA<sup>36</sup> play a primary role in on-target efficiency and off-target specificity.<sup>37</sup> It has also been shown that gRNA expression might affect the CRISPR efficiency.<sup>38</sup> Apart from gRNA-dependent factors, target genomic loci might have an impact because of the chromatin structures.<sup>39</sup> Apel et al. (2017) found substantial variability in terms of integration efficiency and expression rate for 23 characterized genomic loci using green fluorescence protein (GFP) in *S. cerevisiae*.<sup>40</sup> Also, Wu et al. (2017) screened 1044 loci in the yeast genome, reporting important variations in RFP expression among the different loci tested.<sup>41</sup> Therefore, characterization of genomic regions in terms of CRISPR and gene expression efficiencies is crucial for identifying optimal target regions in the yeast genome. Especially for metabolic pathway construction involving many heterologous genes or extra copies of native genes<sup>2</sup>, the use of efficient loci is quite significant to improve production yield.

To accelerate the yeast strain development process, more convenient CRISPR methods reducing time, labor, and cost are necessary, as well as identification and characterization of optimal genomic loci for chromosomal integration of constructs. In the present study, we developed a modular, convenient and standardizable CRISPR method, named Assembly and CRISPR-targeted *in vivo* Editing (ACTivE), relying on *in vivo* assembly of co-transformed DNA modules in yeast. We used chemically synthesized gRNA cassettes and synthetic overlapping fragments as connectors for *in vivo* DNA assembly. In this way, the modules can be easily selected and combined from a part repository depending on the application. Particular parts such as a module expressing a specific type of Cas protein or a gRNA module targeting a specific genomic region can be combined along with other modules to *in vivo* construct the CRISPR plasmid. As the part repository contains only PCR-verified linear fragments, the use of plasmid purification kits to obtain the parts or extra enzymatic treatment steps such as DpnI digestion to degrade parental plasmid is not necessary. Without the use of agar stab/plate, this approach can also facilitate shipping. Also, compatible, and custom-made parts can be exchanged between different groups thanks to the connectors (synthetic fragments) at the terminals of each module. This collection can serve as a CRISPR toolkit that will be expanded with new modules and is freely available at <https://www.leorioslab.org/cost-crispr-toolkit/> so that users will be able to perform yeast genome editing by simply providing their own custom donor DNA. Therefore, ACTivE allows rapid and plasmid-free genome engineering in the yeast genome as it abolishes *in vitro* DNA assembly and cloning steps so that it does not require the use of type II restriction enzymes or DNA assembly kits which can be a considerable expense for lower-income laboratory settings. After optimizing the genome editing efficiency and reproducibility of the method, eight different genomic regions were characterized using a GFP, mNeonGreen, in terms of determining integration and gene expression efficiencies as well as integration effects on cell fitness. ~80% single gene integration and deletion efficiencies were achieved. Furthermore, the multiplexing capacity of ACTivE was tested for simultaneous integration of multi genes into multi loci in the genome. The CRISPR technique used in this study enables standardizable and rapid genome engineering in the yeast

genome, and thoroughly characterized genomic regions provide more alternatives to be used for genomic integration or pathway construction.

# **MATERIALS & METHODS**

## **Oligonucleotides, Reagents, and Plasmids**

All primers used in the study are listed in Table S1-S4. The primers were ordered from Integrated DNA Technologies (IDT) as standard DNA oligos for fragments from 20 bp to 100 bp or as DNA ultramers for fragments with ~120 bp length. Synthetic gRNA cassettes (Figure S1) were ordered from Twist Bioscience. Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific) and PrimeSTAR® GXL DNA Polymerase (TaKaRa) were used for PCR reactions while DreamTaq Green PCR Master Mix (Thermo Fisher Scientific) was used for colony PCR. FastDigest DpnI (Thermo Fisher Scientific) was used to degrade the parental plasmids. GeneJET PCR Purification Kit (Thermo Fisher Scientific) was used for PCR clean-up. p426\_Cas9\_gRNA-ARS511b (Addgene) was used to amplify uracil auxotrophic selection marker (*URA3*), yeast origin of replication (2μ ori) and bacteria storage fragment containing ampicillin resistance gene (*AmpR*), and bacterial origin of replication. pWS158 (Addgene) was used as a template to amplify *Cas9* (*Streptococcus pyogenes*) codon-optimized for expression in *Saccharomyces cerevisiae*. mNeonGreen was used as green fluorescence protein (GFP) and it was amplified from pCPS1ULA-BA6 was obtained as a gift from Matthew Dale (Rosser Lab, the University of Edinburgh). pTDH3-Re2.8-2 was gifted by Jamie Auxillos (Chris French Lab, the University of Edinburgh) and red fluorescent protein (RFP), mCherry, was amplified using this plasmid.

## **Strains and Media**

The parent strain of *S. cerevisiae*, BY4741 {MATa; *his3Δ1*; *leu2Δ0*; *met15Δ0*; *ura3Δ0*}, was used for genomic integrations and was kindly provided by Dariusz Abramczyk (Chris French Lab, the University of Edinburgh). *S. cerevisiae* CEN.PK2-1C {MATa; *his3Δ1*; *leu2-3\_112*; *ura3-52*; *trp1-289*; *MAL2-8c*; *SUC2*} from EUROSCARF Collection was used for genomic deletions. Unless otherwise stated, all chemicals were sourced from Sigma-Aldrich at the highest available purity. For cultivation of strains, YPD medium containing yeast extract (1%(w/v)), peptone (2%(w/v)), and 2% (w/v) dextrose (glucose) was used. To select positive transformants expressing *URA3* marker, synthetic defined medium containing complete supplement mixture minus uracil (CSM-Ura, MP Biomedicals™), 0.17% (w/v) yeast nitrogen base without amino acid, 0.5% (w/v) ammonium sulphate, 2% (w/v) glucose, and 2% (w/v) agar was used. For counter-selection of plasmid-free yeast cells, a synthetic defined medium supplemented with 0.1% (w/v) 5-Fluoroorotic Acid (5-FOA) (Thermo Fisher Scientific) was used. YPD media and complete synthetic defined (SD) media containing all amino acids were used for mNeonGreen expression and characterization of genomic loci.

## Yeast Heat-Shock Transformation

The chemicals were sourced from Sigma-Aldrich unless otherwise stated. All transformations were carried out according to LiAc/PEG heat-shock method<sup>42</sup> with some small modifications. After overnight cultures, fresh cultures were prepared to obtain the cells in the exponential phase. The cells were then washed once and were pelleted by centrifugation. The transformation mix containing 240  $\mu$ L PEG (50%(w/v)), 36  $\mu$ L 1.0 M lithium acetate (LiAc) and 50  $\mu$ L single-stranded carrier DNA (2.0 mg/mL) (herring sperm DNA, Promega) were added onto the cell pellet. Next, DNA fragments and water were added until the volume was made up to 360  $\mu$ L. 50 fmol equivalent molarity of each plasmid-forming DNA part, 500 ng – 1000 ng from each donor DNA part were added to the transformation mixes. As a large number of modular fragments were used for multiplexing, transformation volume was increased to 400  $\mu$ L when needed by adding more DNAs without water addition. After homogenous transformation mixes were obtained, the cells were incubated for 45 minutes at 42 °C. After plating cells to the selective media, the cells were incubated for 2-3 days at 30 °C.

## Determination of Genome Editing Efficiency

When the colonies became visible after integrating mNeonGreen into single loci, the plates were imaged on a blue-LED transilluminator (Thermo Fisher Scientific) to distinguish fluorescent mNeonGreen-expressing positive colonies from non-fluorescing negative colonies. Also, colony PCR was performed on randomly selected five positive colonies from each plate to confirm that the genes are integrated into correct locations, and 100% consistency was observed for all positive colonies controlled. Integration efficiency was determined by calculating the percentage of green light-emitting colonies. To count the colonies on the plates when numerous colonies were obtained, ImageJ<sup>43</sup>, a free distribution software, was employed with the Colony Counter plugin (Figure S4).<sup>44</sup> The plate images on the blue-LED transilluminator were first converted to 16-bit pictures. The green colonies (positive) and white colonies (negative) were distinguished using color contrast, and colonies were counted automatically. For integration efficiency of simultaneous integration of mNeonGreen and mCherry, first, the mNeonGreen expressing colonies were determined on a blue-LED transilluminator. Those colonies were then screened in CLARIOstar Plus microplate reader (BMG Labtech) to detect mCherry expression using spectral scanning with an emission wavelength ranging from 580 nm to 670 nm at 552 nm excitation wavelength. The positive colonies expressing both mNeonGreen and mCherry were also screened by colony PCR to confirm the integrations. Finally, orange colonies were counted using Colony Counter – ImageJ to determine the integration efficiency of  $\beta$ -carotene pathway. These integrations were also confirmed by employing colony PCR. The genomic deletions were first screened using colony PCR, and the deletions were confirmed by Sanger sequencing performed at GENEWIZ, Inc (Leipzig, Germany).



## Characterization of Genomic Loci with mNeonGreen Expression

Three individual colonies from each strain expressing mNeonGreen on different loci were selected after confirming the integrations. 5-FOA counter-selection was performed to select plasmid-free cells after overnight culture in YPD media. Eight mNeonGreen expressing strains and the parent strains, BY4741, were then cultured in YPD and SD media with three replicates for 72 hours to observe expression rates of mNeonGreen on each locus. Biomass of different strains was also measured to compare the integration effect on cell fitness. After overnight culture of each strain, cells were inoculated into fresh media to be grown for around six hours to obtain cells in the exponential phase. The initial OD<sub>600</sub> was adjusted to 0.1 for the growth experiments for all strains. To avoid sedimentation in low volumes, the cells were grown in 1 mL media, YPD or SD, using 24-well plates (Greiner) with shaking at 200 rpm, at 30 °C. To measure fluorescence intensity and biomass, 20 µL culture samples were taken on the 0<sup>th</sup>, 3<sup>rd</sup>, 6<sup>th</sup>, 12<sup>th</sup>, 15<sup>th</sup>, 18<sup>th</sup>, 21<sup>st</sup>, 24<sup>th</sup>, 36<sup>th</sup>, 48<sup>th</sup>, and 72<sup>nd</sup> hours and mixed with water in 200 µL total volume in a black, clear-bottom 96-well plate (Greiner). An anti-condensation solution containing 0.05% Triton X-100 (Sigma-Aldrich) in 20% ethanol <sup>45</sup> was used to cover the lids of the well-plates to prevent condensation on the lids. Fluorescence intensities and OD<sub>600</sub> measurements were taken using the CLARIOstar microplate reader (BMG Labtech). Matrix scan (2x2, 25 flashes) was used to scan the wells. To measure mNeonGreen expression, 490 nm and 525 nm wavelengths were used for excitation and emission, respectively, with 10 nm bandwidth. The emission wavelength was set to 585 nm for autofluorescence of media and yeast cells. <sup>46</sup> The gain was 1500 for both protocols.

## Data Analysis and Software

CRISPR experiments were conducted in at least three replicates. The error bars represent the standard deviations of different experiments. The one-way analysis of variance (ANOVA) ( $p$ -value<0.05) was used to determine whether there was a statistically significant difference between the experiments. The yeast genome was screened using UCSC Genome Browser <sup>47</sup> (<http://genome.ucsc.edu>) to find ARS-close intergenic regions. The potential crRNA sequences on each region were scored using CRISPOR <sup>48</sup> (<http://crispor.tefor.net>), an online gRNA selection tool giving sequence-based scores using sequence prediction algorithms. <sup>49,50</sup> The linear regression was used to determine the relationship between integration efficiencies and gRNA efficiency prediction scores, and Person's correlation coefficients (R) were calculated employing MATLAB. The fluorescence intensity and OD<sub>600</sub> data were analyzed using SciPy package <sup>51</sup> (Matplotlib, NumPy, pandas), *seaborn*, <sup>52</sup> and *omniplate* <sup>53</sup> in Python. Before analyzing growth characteristics, OD correction was performed using a standard curve for 2% (w/v) glucose-containing media as there is a non-linear relationship between biomass and OD<sub>600</sub>. To find the actual fluorescence intensity of per mNeonGreen expressing cell, autofluorescence caused by yeast cells themselves and media, YPD or SD, was corrected. <sup>46</sup> To calculate the areas under the curves, the trapezoidal rule was used. The codes and the standard curve used for plate reader data analysis can be found on <https://github.com/kmalci/plate-reader>. The illustrations were made using the BioRender. <sup>54</sup>

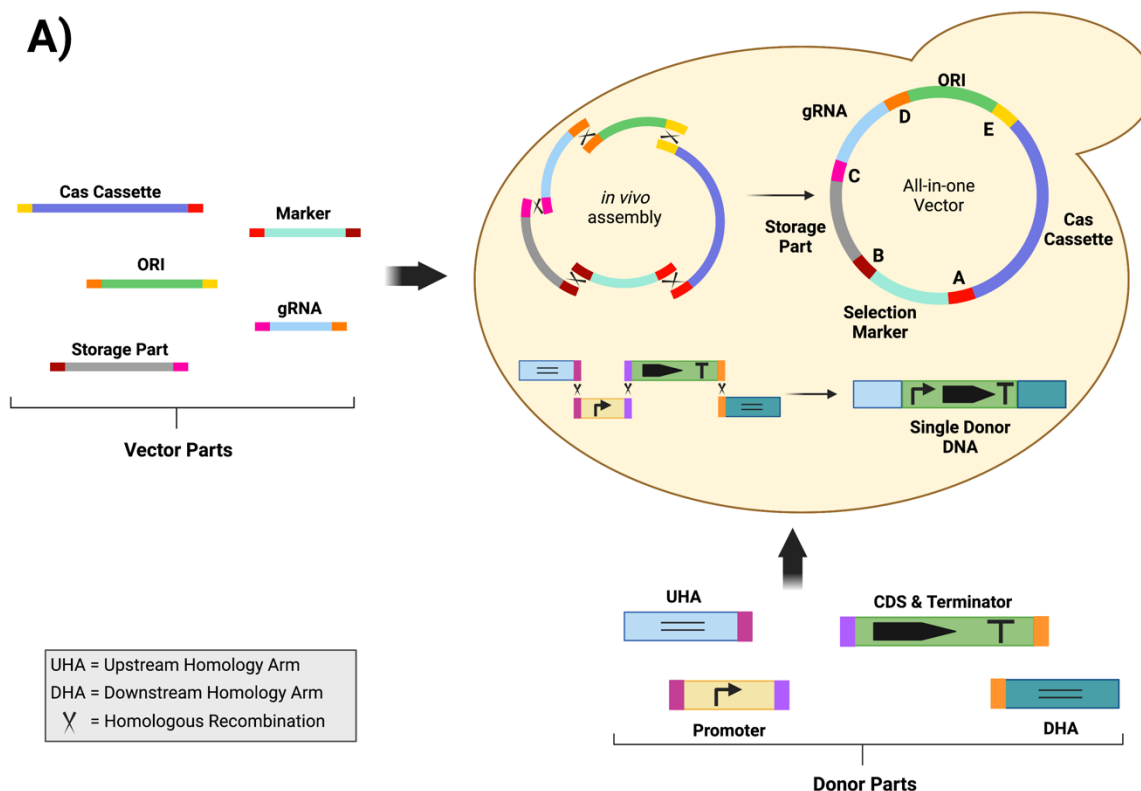
## RESULTS & DISCUSSION

### Standardizable, Rapid, Convenient Yeast CRISPR

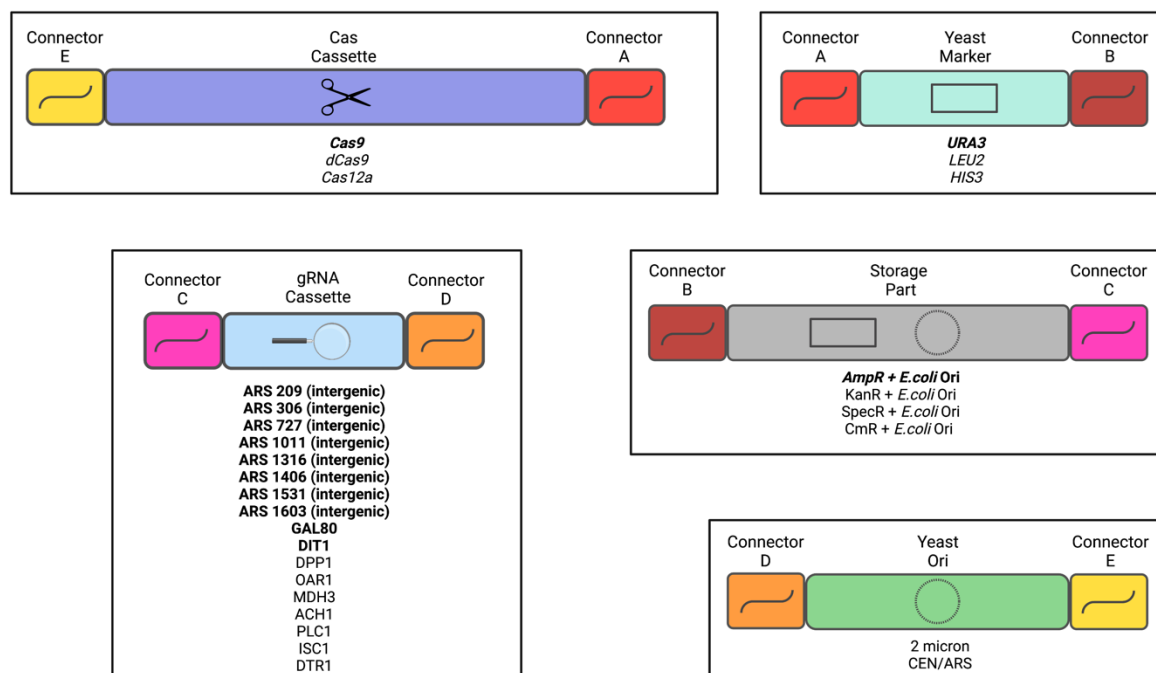
To create a CRISPR plasmid expressing the gRNA and Cas protein, *in vitro* DNA assembly procedures typically followed by a cloning step into *Escherichia coli* are the standard methods used in the yeast genome engineering processes. However, these steps retard this process, especially if sequential genome editing studies are needed to construct or design synthetic metabolic pathways. Here, we developed a more convenient, modular, and rapid CRISPR method, named ACTivE, which relies only on amplifying the functional units (donor DNA) through PCR. The modules contain small overlapping sequences assembled via *in vivo* HDR, so time-consuming and labor-demanding extra steps are omitted.

Gibson et al. (2008) succeeded in assembling a long, ~600 kb, synthetic genome consisting of 25 overlapping fragments in yeast through HDR.<sup>55</sup> After this, Kuijpers et al. (2013) reported that using synthetic overlapping sequences increases the efficiency of *in vivo* DNA assembly in yeast.<sup>56</sup> In our design, we used five modules, a *Cas9* expression cassette, a gRNA expression cassette, a selection marker, a storage part, and a yeast origin of replication (ORI), to create an all-in-one CRISPR plasmid. 60 bp synthetic fragments<sup>56</sup> were used as overlapping sequences (connectors) between each adjacent module. Among them, selection marker and ORI are essential parts for surviving and propagating the plasmid, while Cas protein and gRNA are responsible for CRISPR activity. The storage part, an optional module, contains a bacteria-specific selection marker and ORI to be used for storage of the assembled plasmid in *E. coli* if desired. As chemical DNA synthesis is a rapidly growing sector<sup>57</sup> with many alternative approaches<sup>58</sup>, synthetic production of expression cassettes is a reasonable and affordable choice, especially for short fragments, 300 bp – 500 bp, compared to traditional constructing methods. Therefore, synthetic gRNA expression cassettes were used in our method, allowing the users to choose any integration locus they desire in addition to the eight ones characterized in this work. When it comes to donor DNA to integrate heterologous genes of interest (GoI), we used four parts, ~1000 bp upstream homology arm (UHA), promoter, CDS + terminator, and ~1000 bp downstream homology arm (DHA). In this way, individual promoter parts could be easily changed as a widely used approach for fine-tuning gene expression.<sup>59</sup> Similar to plasmid assembly, each donor part had a 60 bp overlapping sequence with its adjacent part. Figure 1 illustrates the working principles and design of ACTivE and the content of the toolkit provided. Once plasmid modules are produced, they can be stored for subsequent study. Indeed, one of the benefits of this design is the CRISPR plasmid construction can be readily standardized by using the connectors (synthetic fragments) and a part collection containing different alternatives for each module. For instance, the Cas cassette has connector A and connector E at the terminals. Depending on the purpose, a particular type of Cas protein, Cas9, dead Cas9, Cas12a, etc., could be selected from the collection to be combined with other modules. This also applies to other parts, such as the selection markers, allowing more flexibility and part exchange between different research groups. Moreover, as the method relies

on only *in vivo* assembly and PCR for the donor DNA, the whole process can be finished in a single day with a good organization from scratch.



**B)**





**Figure 1:** The overview of the ACTivE method. **A)** Each plasmid module is produced by PCR and can be stored to create a part collection. The modules contain an overlapping sequence (connector) with their neighbor fragment for *in vivo* assembly via homologous recombination. Depending on the application, the plasmid modules can be selected from the part collection. In this study, the donor DNA consisted of four parts, and they were co-transformed with plasmid modules in a single transformation step. In yeast, the parts are assembled via homologous recombination, and the donor DNA is inserted into the genome thanks to its homology arms with the genome. The length of the parts may not represent the actual sizes. **B)** The parts provided in the first version of the ACTivE toolkit (the parts in bold were characterized in this study). The gRNAs of ARS-close regions target intergenic regions in the genome whereas other gRNAs target the genes. By providing their own custom repair donor DNA, the users can readily combine the plasmid modules depending on their applications for a rapid yeast genome editing study. The toolkit will be expanded with new parts and is freely available with more information on <https://www.leorioslab.org/cost-crispr-toolkit/>

## Selecting Integration Regions

To integrate the gene of interest (GoI), *mNeonGreen*, into the yeast genome, eight loci on eight different chromosomes were selected to compare the integration and gene expression efficiency and impact on cell fitness. Previous studies reported that gene expression rates of heterologous genes tend to be higher if they are located in a region close to autonomously replicating sequences (ARS).<sup>60,61</sup> Therefore, ARS-close intergenic regions were targeted for integrations, and crRNA sequences on these regions were scored using different algorithms. Moreno-Mateos et al. (2015) developed a gRNA activity prediction algorithm, CRISPRscan, using zebrafish-specific gRNAs, giving higher efficiency scores if the crRNA sequence has high guanine but low adenine content.<sup>50</sup> Whereas Doench et al. (2016) used mammalian cells to develop a gRNA efficiency algorithm.<sup>49</sup> Both algorithms were considered for selecting crRNAs as yeast-specific algorithms have not been developed yet. To minimize the off-target effects in the genome, the sequences that have the maximum MIT scores representing the highest uniqueness were selected.<sup>62</sup>

Table 1 shows the crRNA sequences used in this study and their corresponding scores. The sequences were selected considering the highest specificity and efficiency prediction scores among the other crRNA sequences on the regions.

**Table 1:** The crRNA sequences used in this study to integrate *mNeonGreen* and sequence features

Target Region	Sequence 5' – 3'	Chromosome	MIT Specificity Score <sup>62</sup>	Efficiency Prediction Scores		Off-target mismatches
				Doench <sup>49</sup>	Moreno- Mateos <sup>50</sup>	
ARS 209	ATAAAATGTGCTAGAACGTA <b>TGG</b>	II	100	51	54	0
ARS 306	GTTAGGAGAAGGAGCGGGTC <b>TGG</b>	III	100	44	66	0
ARS 727	CCGCGCCACTGCTATGACAA <b>CGG</b>	VII	100	61	67	0
ARS 1011	CTGGCGCTGCTGGAAGTGAA <b>TGG</b>	X	100	56	69	0
ARS 1316	ATAGTTACTACCTAAGAAGT <b>GGG</b>	XIII	100	65	56	0
ARS 1406	ACGGTTCTGATGGGAATGCA <b>GGG</b>	XIV	100	52	67	0
ARS 1531	GGCCTCGGCAACTTTCAGGG <b>GGG</b>	XV	100	64	63	0

ARS 1603	GCAAGAACGAAGCTAACGAG <b>AGG</b>	XVI	100	70	58	0
----------	---------------------------------	-----	-----	----	----	---

371

372 \* The last three nucleotides in **red** represent the PAM sequences

373 \*\* The scores are out of 100

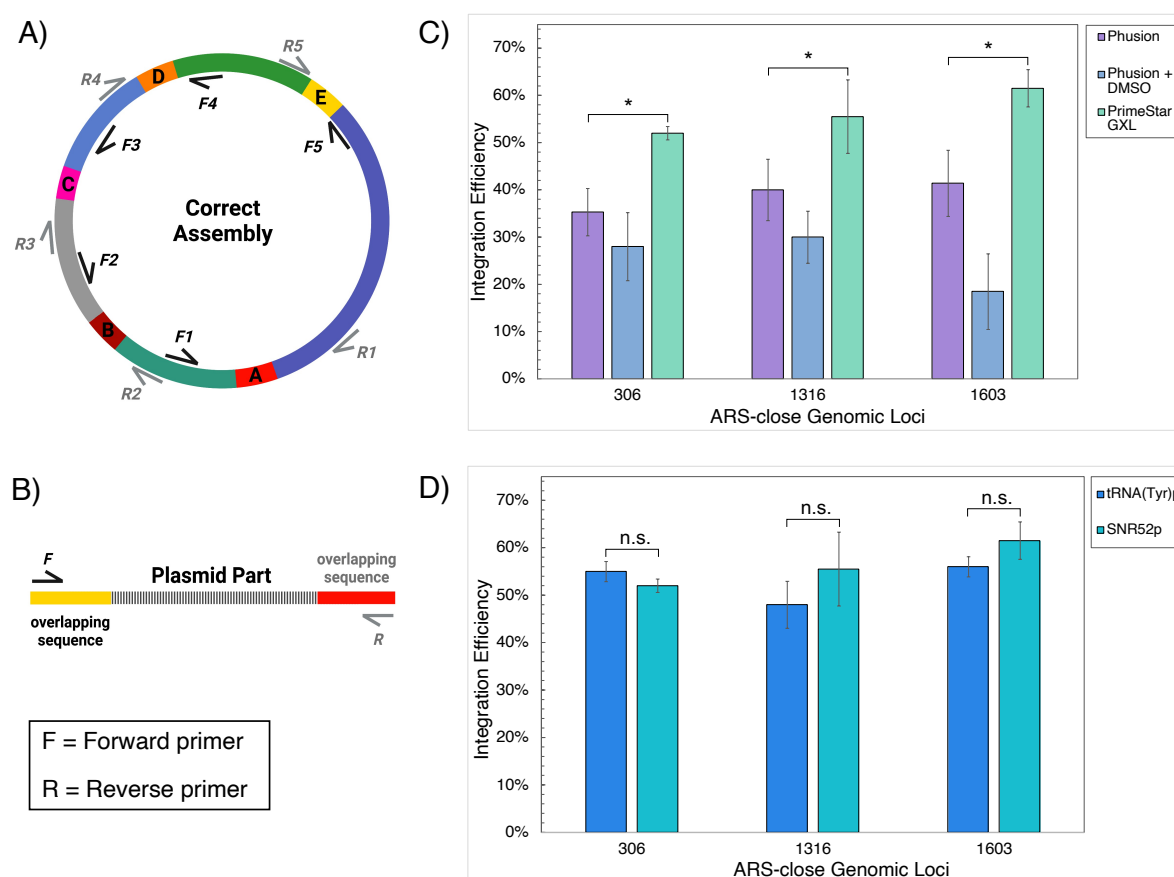
374

## 375 Optimization of the Method

376

377 The integration efficiency of ACTivE was first tested on three loci, ARS 306, ARS 1316,  
378 and ARS 1603, using high-fidelity Phusion polymerase to amplify plasmid modules  
379 and the donor parts containing a strong constitutive promoter, *TDH3p* (Figure 1). Even  
380 though the Gol was successfully integrated into all regions, the integrations  
381 efficiencies ranged from 35% to 41% (Figure 2C). Stratigopoulos et al. (2018) reported  
382 that 24 hours of DMSO feeding prior to CRISPR increased mammalian cell genome  
383 editing efficiency.<sup>63</sup> We, therefore, first tested whether DMSO feeding for 24 hours  
384 had a positive effect on CRISPR efficiency for *S. cerevisiae*. Unfortunately, DMSO  
385 feeding yielded inefficiently for integration rates with less than 30% (Figure 2C). Also,  
386 ~50% fewer colonies were obtained after transformation compared to the DMSO-free  
387 method. Following that, the false positive colonies, which were able to grow on the  
388 selective media but did not express *mNeonGreen*, were further studied to determine  
389 whether they contained correctly assembled Cas9 plasmid, using primers flanking the  
390 overlapping sequences, connectors, as indicated in Figure 2A. This revealed that none  
391 of the false-positive colonies contained connector A, and only ~10% of them had  
392 connector E in their plasmids. As both connectors were overlapping sequences  
393 flanking the *Cas9* cassette, we then tested if the *in vitro* plasmid parts could be  
394 amplified in full. The plasmid parts were used as a template, and ~20 bp primers at  
395 the end of the terminals were used for a second PCR, as demonstrated in Figure 2B.  
396 Although the four plasmid parts, selection marker, yeast ORI, storage part, gRNA,  
397 were amplified, the Cas9 gene failed. In contrast, it was amplified from a correctly  
398 assembled plasmid, proving that the terminal sequences of the Cas9 gene (connector  
399 A and E) were not completely amplified by the Phusion polymerase. Indeed, there are  
400 limited studies about inefficient, repetitive PCR problems, which have contributed to  
401 the lack of reproducibility in certain assays in the synthetic biology community.<sup>64,65</sup>  
402 Shevchuk et al. (2004) reported the “shortening phenomenon” for PCR products.<sup>35</sup>  
403 Researchers revealed that the amplicons of high-fidelity DNA polymerases might yield  
404 truncated products because of the shortening of the amplicon’s ends, especially for  
405 relatively long targets (>3-4 kb).<sup>35</sup> For this reason, we applied two alternative  
406 approaches (i) using longer primers (~120 bp ultramers, Table S1) containing non-  
407 functional extra bases at the terminal to nullify the shortening, (ii) dividing the long  
408 Cas9 gene with internal primers (Table S1) into two parts resulting ~2.5 kb fragments.  
409 However, none of these strategies resulted in higher integration rates since integration  
410 rates similar to those of Phusion polymerase are achieved. Finally, we tested another  
411 high-fidelity DNA polymerase, PrimeSTAR GXL, which has a relatively higher error  
412 rate compared to Phusion. Surprisingly, it successfully amplified the Cas9 gene in full  
413 (Figure 2B), and all plasmid parts were re-amplified with a second PCR, as explained  
414 above. Consistently, the integration efficiencies in all regions increased dramatically  
415 by approximately 1.5 fold ( $p$ -value < 0.01), as shown in Figure 2C. Thus, all the  
416 plasmid parts were produced using PrimeSTAR GXL in the subsequent experiments,  
417 while donor parts were produced using Phusion polymerase.

After improving the integration efficiency, false-positive colonies were screened to see if they contained the correctly assembled CRISPR plasmid after using the PrimeSTAR GXL. It was observed that ~95% of the false-positive colonies had incorrectly assembled plasmids without gRNA and/or Cas9 parts, and all positive mNeonGreen expressing colonies had correct plasmids. This showed that the accuracy of *in vivo* plasmid assembly was the critical factor for CRISPR success in our method. The false-positive colonies also indicated that the plasmid could be assembled with missing parts. Different combinations of plasmid parts were transformed into yeast cells to confirm this. Although transformation yield dramatically decreased with missing parts, we observed a small number of colonies as long as they contained a selection marker and yeast ORI that are essential parts for surviving, as demonstrated in Figure S5. This was probably because of assembling the plasmid parts through non-homologous end-joining (NHEJ) rather than HDR.<sup>66,67</sup> NHEJ is a well-studied pathway, and the genes involved in this process are elucidated.<sup>66,67</sup> Therefore, as a further improvement, a disruption in the NHEJ pathway could increase the integration efficiency of ACTivE or similar methods, as shown previously for a non-conventional yeast *Yarrowia lipolytica*.<sup>68</sup>



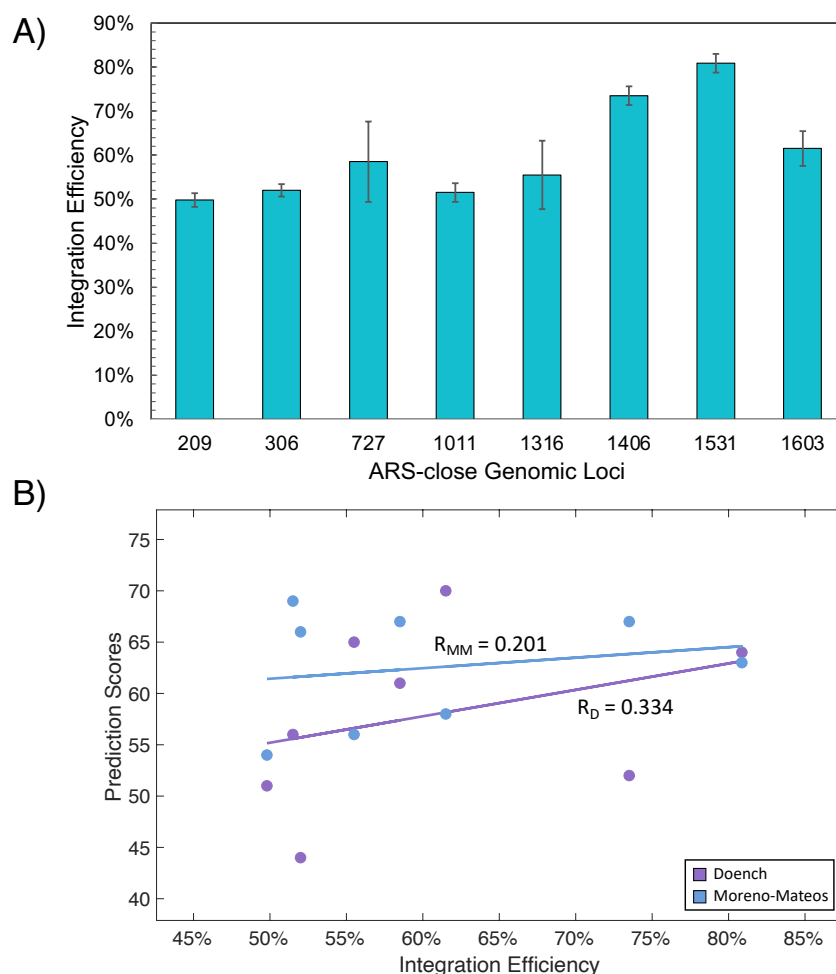
**Figure 2:** Optimization process and results of ACTivE. **A)** The all-in-one CRISPR plasmids were screened using plasmid colony PCR primers (F1-5, R1-5) targeting the connectors between adjacent plasmid modules. **B)** Each plasmid part was controlled to determine whether they were amplified in full after PCR reactions using ~20 bp primers annealing at the end of the synthetic overlapping sequences. **C)** Comparison of integration efficiencies of different conditions; Phusion polymerase-based plasmid parts, DMSO feeding for 24 hours for

Phusion polymerase-based plasmid parts, and PrimeSTAR GXL-based plasmid parts. **D)** Comparison of integration efficiencies when gRNA is expressed by two different promoters, *SNR52p* or *tRNA<sup>(Try)</sup>p*. The values are displayed as the mean of triplicate experiments, and error bars represent standard deviations. For simplification, ARS regions are shown with their corresponding numbers. The single asterisk represents a  $p$ -value  $< 0.01$  and “n.s.” stands for not statistically significant ( $p$ -value  $> 0.05$ ). The error bars show standard deviations of three replicates.

Additionally, we also compared two RNA expressing promoters to determine whether they affected gRNA expression and CRISPR activity as the differences in the promoters’ features might have an impact on the expression.<sup>37</sup> *SNR52* RNA polymerase III promoter (*SNR52p*) is one of the most widely used promoters to express gRNAs in yeast.<sup>12,23</sup> Alternatively, gRNAs can be transcribed using tRNA promoters.<sup>11,40</sup> Therefore, we tested *SNR52p* and *tRNA<sup>(Try)</sup>p* promoters as shown in Figure 2D; however, no statistically significant difference ( $p$ -value  $> 0.05$ ) was observed. Dong et al. (2020) compared *tRNA<sup>(Try)</sup>*, *tRNA<sup>(Pro)</sup>*, and *SNR52* promoters in terms of disruption yield on the *ADE2* gene and repression efficiency on a heterologous fluorescence protein.<sup>69</sup> Consistent with this study, researchers found very similar disruption and repression yields with these three gRNA promoters.<sup>69</sup> *SNR52p* consists of 269 bp, while *tRNA<sup>(Try)</sup>p* contains 118 bp (Figure S1). Therefore, promoters such as *tRNA<sup>(Try)</sup>p* could be used to minimize the cost of the synthetic gRNA cassettes.

## Integration Efficiencies on Eight ARS-close Genomic Loci

The improved A<sup>CTivE</sup> method was used to integrate *mNeonGreen* into five more ARS-close genomic loci to compare the integration efficiencies and characterize the selected eight genomic regions (Table 1). High-yield gene integrations ranging from 50% to 80% were achieved in the target regions, as shown in Figure 3A. Although 100% gene integration or deletion efficiencies have been previously reported,<sup>12</sup> genome editing efficiency of A<sup>CTivE</sup> is remarkably high considering its convenience and rapidity. In addition, previous works reported extremely low ( $< 20\%$ ) CRISPR-based heterologous gene integration efficiencies in some genomic locations.<sup>40</sup> Therefore, the regions tested in this study are suitable targets for gene integrations.



**Figure 3:** Integration efficiencies and their correlation with prediction algorithms. **A)** Integration efficiency of *mNeonGreen* on eight ARS-close genomic loci. The values are displayed as the mean of triplicate experiments and error bars represent standard deviations. For simplification, ARS regions are shown with their corresponding numbers. **B)** Linear regression models showing the correlation between integration percentages and gRNA efficiency scores of Doench's<sup>49</sup> and Moreno-Mateos's<sup>50</sup> prediction algorithms.  $R_D$  = Pearson correlation coefficient based on Doench's algorithm,  $R_{MM}$  = Pearson correlation coefficient based on Moreno-Mateos' algorithm. The error bars show standard deviations of three replicates.

The relationship between integration efficiencies and gRNA efficiency scores of the genomic regions (Table 1) was also determined using linear regression models, as shown in Figure 3B. Moderate and weak positive correlations were found on Doench<sup>49</sup> and Moreno-Mateos<sup>50</sup> algorithms with a Pearson's correlation coefficient of  $R=0.334$  and  $R=0.201$ , respectively. These findings evaluating a small dataset can be considered promising, although the best fitting gave only a moderate positive correlation. These findings suggest that computational gRNA design/scoring tools can help with selecting gRNA sequences for CRISPR-based studies in yeast, considering the algorithms' limitations. Besides this, a larger sample size is required to properly evaluate these prediction algorithms, which should be further improved by considering, for example, 3-D chromatin structures of the yeast<sup>70</sup> to develop more reliable, yeast-specific prediction tools.

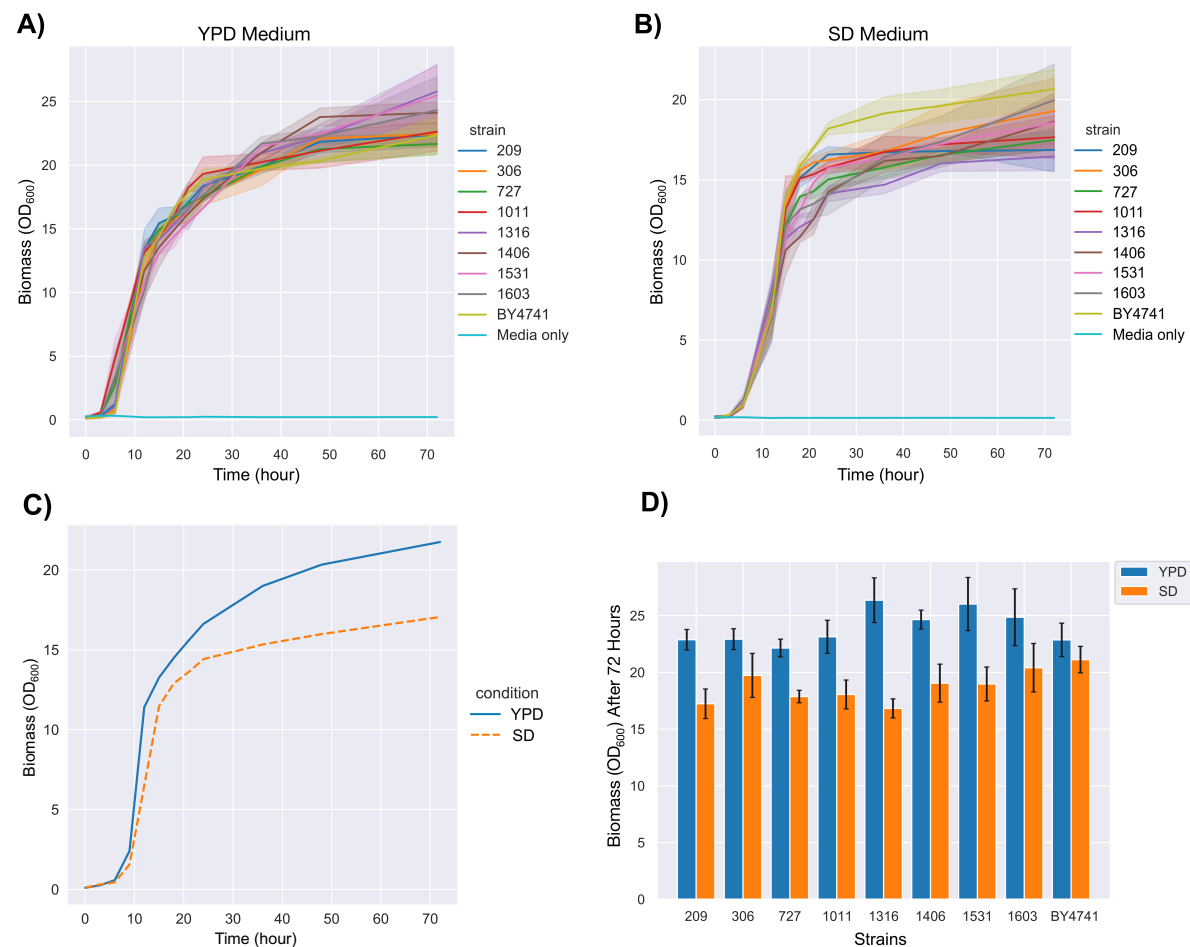


## Single-step Gene Deletion Using Only Homology Arms

Besides genomic integration, gene deletion was also tested using ACTivE. To this end, two non-essential genes, the *GAL80* gene encoding a regulator protein for galactose-related metabolic genes<sup>71</sup> and the *DIT1* encoding a sporulation-specific enzyme,<sup>72</sup> were deleted. Using only UHA and DHA flanking outside the genes (primers and crRNAs are listed in Table S5), ~75% deletion efficiency (standard deviation  $\approx$  8%) and ~93% deletion efficiency (standard deviation  $\approx$  11%) were achieved for the *GAL80* and *DIT1* genes, respectively, without the need for any heterologous gene part. In this way, the whole *GAL80* and *DIT1* genes were deleted without any scar. The deletions were confirmed using colony PCR and Sanger sequencing. This study demonstrates the flexibility of ACTivE as the same approach can be also used for scar-free gene deletion.

## Characterization of Genomic Loci

The genomic regions used as landing pads for *mNeonGreen* were characterized in terms of gene expression rate and effect on the cell fitness in two different media, YPD and SD, as described in the materials and methods section. First, plasmid-free cells were selected using 5-FOA counter-selection to eliminate plasmid burden. OD<sub>600</sub> was measured to compare biomass and growth rates of the strains containing *mNeonGreen* on different genomic loci. Figure 4 shows the biomass of the yeast strains in YPD and SD media.



**Figure 4:** The biomass comparison between *mNeonGreen* integrated strains and the parental strain **A)** The biomass over 72 hours in YPD media **B)** The biomass over 72 hours in SD media **C)** Growth trends in YPD and SD media comparing mean biomass of all strains in these two conditions over 72 hours **D)** The final OD<sub>600</sub> values represent the total biomass after 72 hours. The solid lines on the curves stand for the average values of independent colonies or different strains. The standard deviations of the three replicates are shown by shading on the curves or by error bars on the bar chart.

As seen, the growth curves presented the expected trends in both media (Figure 4A & 4B). The average biomass (Figure 4C) of nine strains, including the parental strain, BY4741, was expectedly higher in YPD media compared to SD media ( $p$ -value < 0.01) as YPD is a richer environment than SD media.<sup>73,74</sup> The biomass order in YPD media was as follows:

1316 > 1531 > 1603 > 1406 > 1011 > BY4741 > 306 > 209 > 727

Although the 1316 (OD<sub>600</sub> ≈ 26) showed the highest biomass, and the 727 (OD<sub>600</sub> ≈ 22) showed the lowest biomass at the 72<sup>nd</sup> hour (Figure 4A), there were no statistically significant differences when they were compared to the parental strain (OD<sub>600</sub> ≈ 22.5) ( $p$ -value > 0.05). This data showed that perturbations on these genomic loci are unlikely to cause a negative effect on cell fitness considering the parental strain. Apel et al. (2017) reported a similar result as they did not find a significant difference in growth rates of GFP expressing strains compared to their parental strain in YPD media.<sup>40</sup> However, a statistically significant difference ( $p$ -value < 0.05) was observed between the 1316 and the 727, showing that targeting ARS1316 likely yields better biomass than ARS727 in YPD media.

On the other hand, the average biomass order in SD media was as follows:

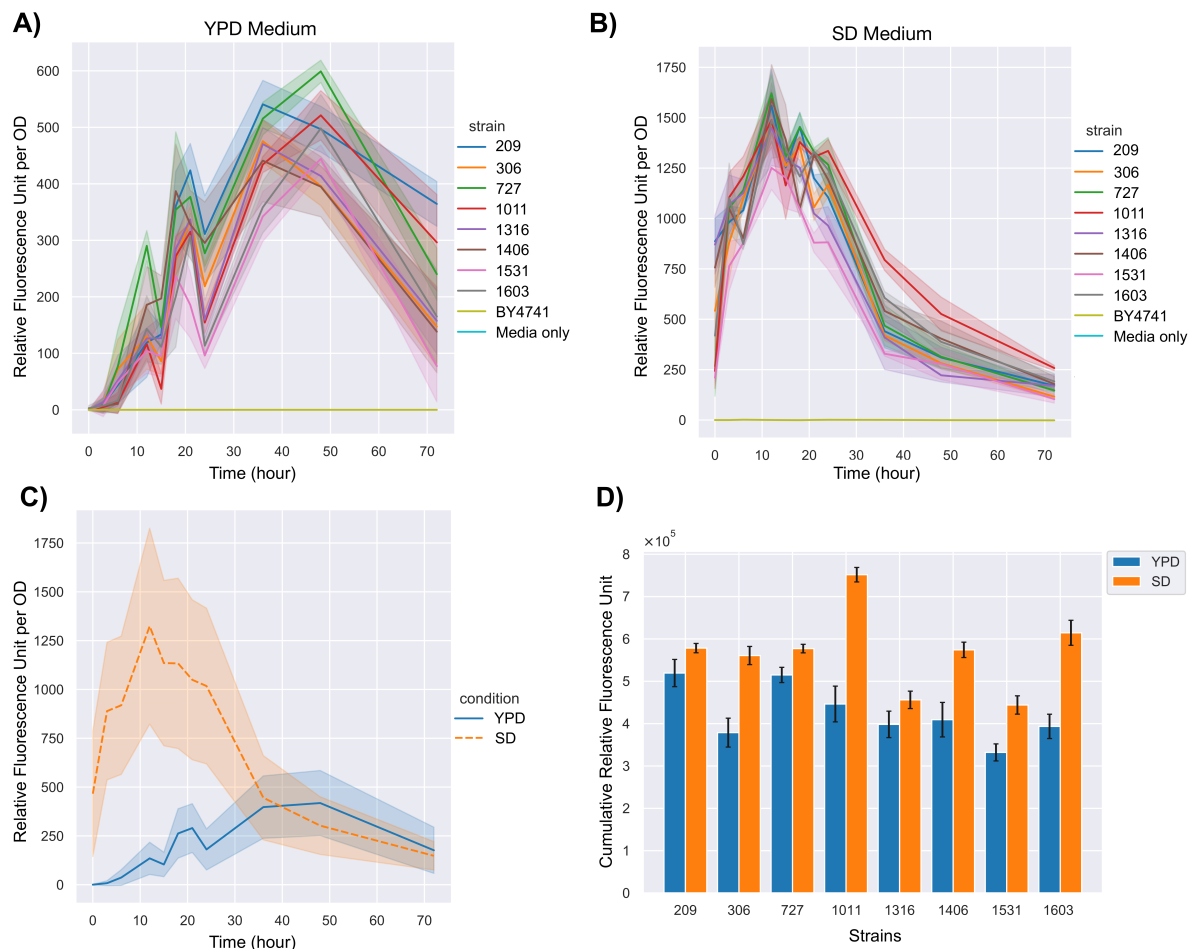
BY4741 > 1603 > 306 > 1406 > 1531 > 1011 > 727 > 209 > 1316

BY4741 showed the highest biomass amount as the parental strain in SD media (Figure 4B). Besides, a statistically significant difference ( $p$ -value < 0.05) was observed when the parental strain BY4741 was individually compared with 209, 727, 1011, and 1316, meaning that genetic perturbations on these loci might negatively affect the cell fitness in SD media. Surprisingly, 1316 resulted in the lowest biomass in SD media even though it was the best growing strain in YPD. Also, a substantially different biomass order was observed in SD media compared to YPD media (Figure 4D). These results show that genetic alterations in the corresponding locations might distinctly affect the cell fitness of yeast strains depending on the environmental conditions and the media compositions.

In addition, the growth rates of the strains were calculated using a Gaussian process-based algorithm.<sup>75</sup> As shown in Figure S6, the maximum growth rate of BY4741 was observed around the 6<sup>th</sup> hour in both YPD and SD media. Similarly, the maximum growth rates of *mNeonGreen* integrated strains were at the 6<sup>th</sup> hour in SD media (Figure S8). However, 1011, 1406, 1531, and 1603 showed the maximum growth rate at around the 3<sup>rd</sup> hour, whereas the others were around the 6<sup>th</sup> hour (Figure S7).

Probably, perturbations on ARS1011, ARS 1406, ARS1531, and ARS1603 affect the growth rate in YPD media.

To compare gene expression, fluorescence intensities produced by mNeonGreen were measured. Autofluorescence caused by the yeast cell itself and the media <sup>46</sup> was corrected using *omniplate* <sup>53</sup> to detect the fluorescence intensity per OD.



**Figure 5:** The comparison of heterologous gene expression between *mNeonGreen* integrated strains **A)** Relative fluorescence intensities (RFU) over 72 hours in YPD media **B)** RFUs over 72 hours in SD media **C)** The average RFUs of all strains in YPD and SD media separately. **D)** The cumulative RFU values represent the total gene expression or protein production by the total biomass at each time for 72 hours. The standard deviations of the three replicates are shown by shading on the curves or by error bars on the bar chart. The solid lines on the curves stand for the average values of independent colonies or different strains.

All strains showed similar expression patterns based on the selected medium, YPD or SD, with some fluctuations in the first 24 hours, as shown in Figures 5A & 5B. As mNeonGreen expression was driven by the constitutive *TDH3* promoter in all mNeonGreen expression strains, these results also showed the expression patterns of *TDH3p* in YPD and SD media (Figure 5C). In YPD media, the best mNeonGreen expressing strains depended on the hour. For instance, at 24<sup>th</sup>, 36<sup>th</sup>, and 72<sup>nd</sup> hours, the 209 was the best strain for mNeonGreen expression, whereas the 727 showed the

highest expression at the 48<sup>th</sup> hour (Figure 5A). These two strains were also the most mNeonGreen producing strains in 72 hours (Figure 5D).

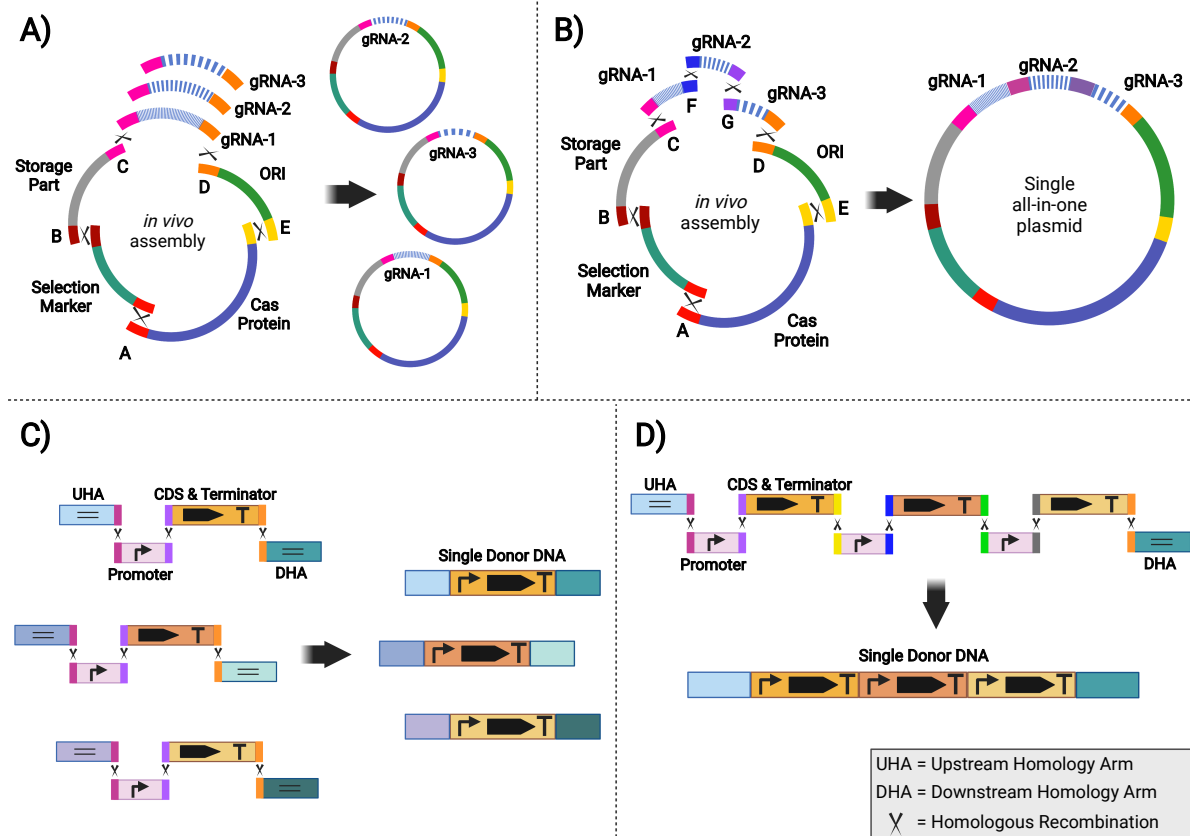
On the other hand, 1011 showed the highest expression from the 24<sup>th</sup> hour in SD media (Figure 5B). Therefore, the ARS1011 was the best integration site for the cumulative expression of mNeonGreen in SD media even though the 1011's biomass was significantly lower than the parental strain ( $p$ -value < 0.05) in SD media (Figure 4D). Moreover, Figures 4C & 5C show that more biomass was obtained in YPD media. Still, mNeonGreen expression was dramatically higher in SD media as there was a significant difference in expression rates in the first 36 hours of SD and YPD media. This shows that the expression rates were maximum in the exponential phase in SD media. However, after the 24<sup>th</sup> hour, there was a sharp decrease in the fluorescence unit. This decrease might be because of both reduction in the expression and the degradation of the mNeonGreen. Nevertheless, the cumulative expression indicating the total protein production over 72 hours by total biomass at each time can be helpful information, especially for protein secretion into the extracellular environment where the effect of intracellular degradation is minimum. Therefore, SD media was more advantageous for cumulative protein production than YPD media for these genomic loci.

The time-derivatives of mNeonGreen expressions were also found, as shown in Figures S9 & S10. The accelerations in expression reached the maximum between 10<sup>th</sup> and 20<sup>th</sup> hours in YPD media except for 1531, which showed the maximum increase in the expression rate between 30<sup>th</sup> and 40<sup>th</sup> hours (Figure S9). In contrast, the increase in the expression rates was the highest in the first 10 hours in SD media for all mNeonGreen expressing strains (Figure S10). These results were consistent with Figure 5C as the expression started with ~500 relative fluorescence units (RFU) in SD media and reached the maximum at around the 12<sup>th</sup> hour, while it began with ~0 RFU and reached the maximum at around the 48<sup>th</sup> hour in YPD media (Figure 5C).

Indeed, locus-based, *TDH3p*-driven heterologous gene expression in the yeast genome has been thoroughly studied in this work, presenting biomass, fluorescence intensity per OD, expression rate, and cumulative expression amount in 72 hours with 11 different time points in two different media. In a similar study, Apel et al. (2017) compared the GFP fluorescence levels among 23 genomic loci at the 8<sup>th</sup> and 24<sup>th</sup> hours in YPD and SD media using a *TEF1p promoter* instead of the *TDH3p* used in this work.<sup>40</sup> The researchers reported a higher average fluorescence level in YPD at the 8<sup>th</sup> hour, whereas it was clearly in favor of SD medium at the 24<sup>th</sup> hour.<sup>40</sup> In the same study, the promoters were also compared, and it was shown that the fluorescence level was the highest in SD media with *TDH3p*-driven expression at the 4<sup>th</sup> hour on a single locus, whereas it was slightly higher at the 8<sup>th</sup> and 24<sup>th</sup> hours in YPD.<sup>40</sup> Although having a higher fluorescence intensity in SD media with *TDH3p* in the early hours of the cultivation overlaps with our results; we observed a clear difference in fluorescence intensity in favor of SD media in the first 32 hours (Figure 5C) that was longer than it was reported in the previously mentioned reference.<sup>40</sup> These findings can also give insights into the expression patterns of the native *TDH3* gene that encodes an enzyme, glyceraldehyde 3-phosphate dehydrogenase, involved in glycolysis, transcriptional silencing, and rDNA recombination.<sup>76</sup>

## Multiplexing using ACTivE

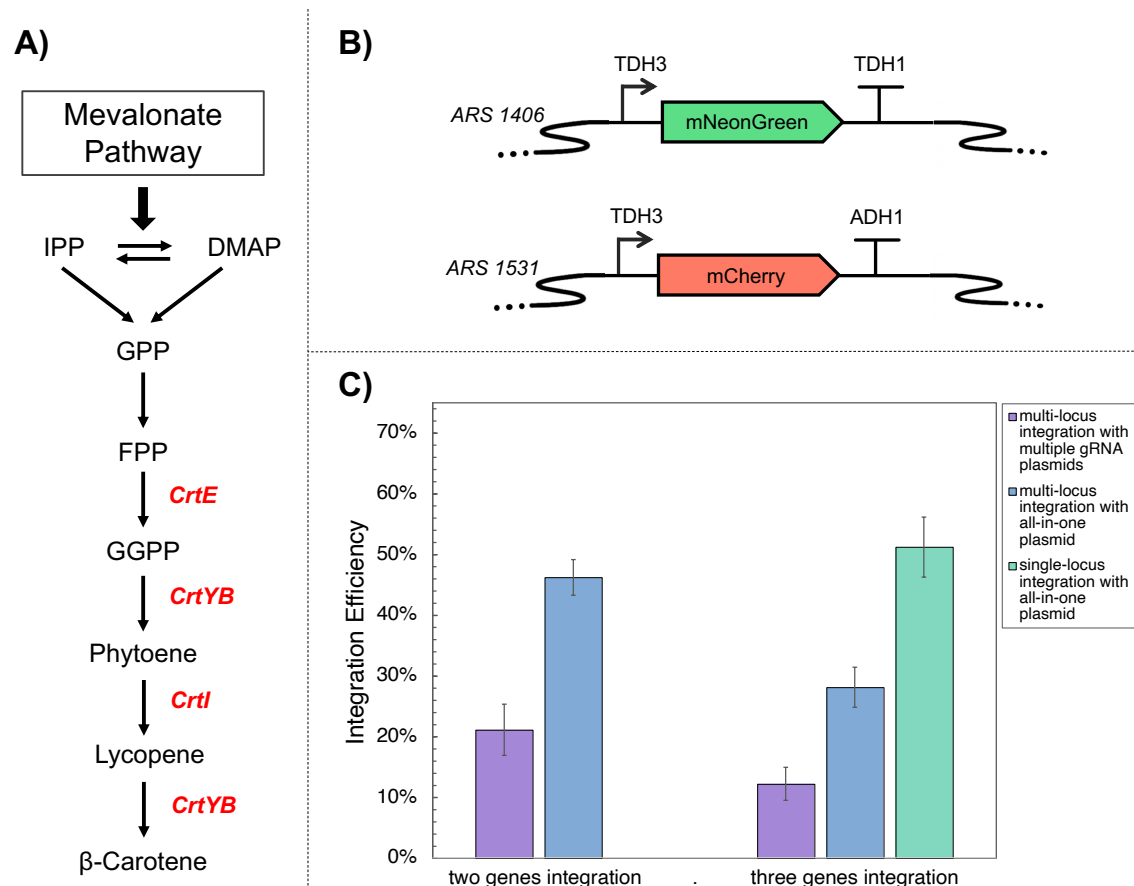
Simultaneous genome alterations in a single step can be preferential to accelerate genome editing or heterologous pathway construction. Therefore, the multiplexing capacity of ACTivE was tested for multi loci and multi-gene integrations into the yeast genome using different plasmid assembly and donor DNA delivery strategies, as illustrated in Figure 6. Initially, two fluorescent reporter proteins, mNeonGreen and mCherry, were integrated into ARS1406 and ARS1531 loci, respectively (Figure 7B), using two independent gRNA modules to be assembled into two different plasmids (Figure 6A). With this strategy, *mNeonGreen* and *mCherry* were simultaneously integrated into 21% of the colonies (Figure 7C). Alternatively, the gRNAs targeting ARS1406 and ARS1531 were *in vivo* assembled using synthetic homology sequences resulting in a single all-in-one plasmid expressing both gRNAs simultaneously (Figure 6B). The integration rate reached 46% with this approach (Figure 7C).



**Figure 6:** Alternative approaches for multiplexing **A)** Different gRNA cassettes are *in vivo* assembled, resulting in different plasmids targeting multiple loci. Each gRNA cassette contains identical overlapping sequences at their terminals so that each one individually assembles with the other plasmid parts. **B)** The gRNA cassettes are tandemly assembled, resulting in a single all-in-one plasmid that targets multiple loci. **C)** For multi-locus and multi-gene integration, each gene has its own homology arms (HA) depending on the target site. **D)** The promoters and CDSs can be tandemly assembled for single-locus multi-gene integration. The multi-gene cassette contains a single upstream homology arm (UHA) and downstream homology arm (DHA).



673 Following this approach, three heterologous genes of the  $\beta$ -carotene pathway, *CrtE*,  
674 *CrtYB*, and *CrtI* (Figure 7A), were simultaneously integrated into ARS1406, ARS1531,  
675 and ARS1603, respectively. A 12 % integration rate was achieved using the strategy  
676 of assembling independent plasmids for each gRNA (Figure 6A), but this increased to  
677 28% when three gRNAs were assembled in a single plasmid (Figure 6B). Furthermore,  
678 the genes were integrated into a single locus (ARS1531) to construct the multi-gene  
679 pathway on a single genomic location (Figure 6D), and 51% of the colonies  
680 successfully produced the orange pigment. Indeed, the differences in integration  
681 efficiencies were not surprising as higher integration rates were observed with fewer  
682 linear fragments. Even though the 12% integration rate is relatively low, the correct  
683 assembly and/or integration of 19 linear fragments (Figure 6A & 6C) shows the  
684 flexibility and capability of the A<sup>CT</sup>ivE method.



**Figure 7: A)** The heterologous  $\beta$ -carotene pathway constructed in a single-step multi-gene integration. The heterologous genes are shown in red. **B)** Genetic construct illustrations of *mNeonGreen* and *mCherry* genes used for single-step double gene integration and *CrtE*, *CrtI*, and *CrtYB* genes used for single-step triple gene integration. **C)** Integration efficiencies achieved using different strategies. The single asterisk represents a  $p$ -value < 0.01. The error bars show standard deviations of three replicates.

## CONCLUSION

The ACTivE method proposes a practical strategy to accelerate yeast genome engineering. It eliminates the use of any reagents or kits used for *in vitro* DNA assembly and bypasses extra cloning steps. Therefore, verified linear fragments have been collected to create a flexible CRISPR toolkit for many different purposes. The first version of the toolkit and a user manual are freely available at <https://www.leorioslab.org/cost-crispr-toolkit/>. The user can simply combine the modules provided in the toolkit without any plasmid purification, enzymatic treatment, *in vitro* assembly, or cloning steps. Providing custom donor DNA, the whole work for genome editing can be completed in only one day using verified linear fragments. Moreover, the modules to be used for CRISPR plasmid construction can be stored for a long time for further applications, and they can be exchanged between different groups as standard parts thanks to their synthetic ends. In this way, customized CRISPR plasmids containing specific parts such as Cas protein, selective marker, gRNA, yeast ORI or bacterial marker can be easily obtained.

To increase the genome editing efficiency and reproducibility of this method, the importance of the DNA polymerase type to be used for part amplification was underlined. In other words, the plasmid modules had to be completely amplified as their ends are critical for *in vivo* assembly. To this end, an appropriate DNA polymerase synthesizing complete amplicons containing the connectors was used to overcome the shortening problem of the DNA fragments that might be a bottleneck to amplify long amplicons.

ACTivE achieved more than 80% integration for a single gene in the ARS1531 region. For multiplexing, expressing multiple gRNAs through a single plasmid resulted in a higher integration yield for multi-loci and multi-gene integration. More than 50% integration efficiency for triple genes was reached in the ARS1531 region as a single-locus multi-gene integration. Also, more than 90% deletion efficiency was obtained on the *DIT1* gene. Considering its usefulness and pace, this method should accelerate genome editing processes as the strains of interest can be smoothly detected after a simple screening.

In addition, eight ARS-close regions in the yeast genome were thoroughly characterized using two different yeast media. The total biomass and growth rates of the strains containing heterologous genes in the corresponding loci were found. Also, RFUs were detected to characterize the gene expression rates and total protein production in these locations. All strains showed higher growth rates in YPD than in the SD medium. Considering the growth rates, ARS 1316 and ARS 1531 might be good targets for high biomass in YPD, whereas ARS 306 and ARS 1603 are preferable for higher biomass in SD media. Nonetheless, no significant difference in the final biomass was observed compared to the parental strain in YPD media. In contrast, strains 209, 727, 1011, and 1316 showed lower growth rates than the parental strain in SD media. On the other hand, gene expression rates might vary depending on the locus used. ARS 209 and ARS 727 showed better protein expressions in YPD media, while ARS 1011 was the best locus for gene expression in SD media. Besides, dynamic gene expression varied considerably mainly depending on the medium conditions, and the cumulative expressions were higher in SD media, although the biomasses were lower in this media.

*S. cerevisiae* is an important chassis organism for many applications, from metabolic engineering to disease modeling. The CRISPR/Cas system has been a versatile instrument for designing its genome. The improvements and alternative approaches presented in this paper have a great potential to accelerate the yeast genome editing process in a standardizable and easy way. The genomic loci characterized in this study provide more options for well-defined genomic landing sites, especially for yeast cell factory design.

## **Author Information**

### **Corresponding Author**

Leonardo Rios Solis – Institute for Bioengineering, School of Engineering, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BF, United Kingdom;

Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;  
School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK;  
<https://orcid.org/0000-0002-4387-984X>  
Email: leo.rios@newcastle.ac.uk

# **Authors**

Koray Malcı – Institute for Bioengineering, School of Engineering, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BF, United Kingdom;  
Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;  
<https://orcid.org/0000-0002-2942-1146>

Nestor Jonguitud-Borrego – Institute for Bioengineering, School of Engineering, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BF, United Kingdom;  
Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;

Hugo van der Straten – Institute for Bioengineering, School of Engineering, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BF, United Kingdom

Urtė Puodžiūnaitė – Institute for Bioengineering, School of Engineering, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BF, United Kingdom  
School of Biological Sciences, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3FF, United Kingdom;  
Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;

Emily J. Johnston – School of Biological Sciences, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3FF, United Kingdom;  
Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;  
<https://orcid.org/0000-0003-0592-0992>

Susan J Rosser – School of Biological Sciences, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3FF, United Kingdom;  
Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, Kings Buildings, Edinburgh, EH9 3BD, United Kingdom;  
<https://orcid.org/0000-0002-2560-6485>

## Notes

The authors declare no competing financial interest.

## Data Availability

The detailed information about the parts and modules in the toolkit and the version updates can be found at <https://www.leorioslab.org/cost-crispr-toolkit/>. The codes and the standard curve for plate reader data analysis can be found on <https://github.com/kmalci/plate-reader/>.

## Acknowledgments

This study was supported by the YLSY Program of the Ministry of National Education of the Republic of Turkey, the British Council (Grant Number: 527429894), IBCatalyst (Project No. 102297), Biotechnology and Biological Sciences Research Council (BBSRC; Project BB/S017712/1). The authors thank Dr. Dariusz Abramczyk for providing the BY4741 strain, Dr. Matthew Dale for his gift of vector pCPS1ULA-BA6, Dr. Jamie Auxillos for her gift of pTDH3-Re2.8-2 vector, Alexander Speakman for technical assistance about the microplate reader, and Prof. Peter Swain for his guide on the *omniplate* software.

## REFERENCES

- (1) Walls, L. E.; Malci, K.; Nowrouzi, B.; Li, R. A.; D’Espaux, L.; Wong, J.; Dennis, J. A.; Semião, A. J. C.; Wallace, S.; Martinez, J. L.; Keasling, J. D.; Rios-Solis, L. Optimizing the Biosynthesis of Oxygenated and Acetylated Taxol Precursors in *Saccharomyces Cerevisiae* Using Advanced Bioprocessing Strategies. *Biotechnol. Bioeng.* **2020**, *118* (1), 279–293. <https://doi.org/10.1002/BIT.27569>.
- (2) Nowrouzi, B.; Li, R.; Walls, L. E.; D’Espaux, L.; Malci, K.; Liang, L.; Jonguitud Borrego, N.; Lerma Escalera, A. I.; Morones-Ramirez, J. R.; Keasling, J. D.; Rios Solis, L. Enhanced Production of Taxadiene in *Saccharomyces Cerevisiae*. *Microb. Cell Factories* **2020**, *19*, 200. <https://doi.org/10.1186/s12934-020-01458-2>.
- (3) Madhavan, A.; Arun, K. B.; Sindhu, R.; Krishnamoorthy, J.; Reshmy, R.; Sirohi, R.; Pugazhendhi, A.; Awasthi, M. K.; Szakacs, G.; Binod, P. Customized Yeast Cell Factories for Biopharmaceuticals: From Cell Engineering to Process Scale Up. *Microb. Cell Factories* **2021**, *20* (1), 1–17. <https://doi.org/10.1186/S12934-021-01617-Z>.
- (4) Wong, J.; de Rond, T.; d’Espaux, L.; van der Horst, C.; Dev, I.; Rios-Solis, L.; Kirby, J.; Scheller, H.; Keasling, J. High-Titer Production of Lathyrane Diterpenoids from Sugar by Engineered *Saccharomyces Cerevisiae*. *Metab. Eng.* **2018**, *45*, 142–148. <https://doi.org/10.1016/J.YMBEN.2017.12.007>.
- (5) Nielsen, J.; Larsson, C.; van Maris, A.; Pronk, J. Metabolic Engineering of Yeast for Production of Fuels and Chemicals. *Curr. Opin. Biotechnol.* **2013**, *24* (3), 398–404. <https://doi.org/10.1016/J.COPBIO.2013.03.023>.



- 845 (6) Mohd Azhar, S. H.; Abdulla, R.; Jambo, S. A.; Marbawi, H.; Gansau, J. A.;  
846 Mohd Faik, A. A.; Rodrigues, K. F. Yeasts in Sustainable Bioethanol  
847 Production: A Review. *Biochem. Biophys. Reports* **2017**, *10*, 52–61.  
848 <https://doi.org/10.1016/J.BBREP.2017.03.003>.
- 849 (7) Walls, L. E.; Rios-Solis, L. Sustainable Production of Microbial Isoprenoid  
850 Derived Advanced Biojet Fuels Using Different Generation Feedstocks: A  
851 Review. *Front. Bioeng. Biotechnol.* **2020**, *8*, 1272.  
852 <https://doi.org/10.3389/FBIOE.2020.599560/BIBTEX>.
- 853 (8) Wong, J.; Rios-Solis, L.; Keasling, J. D. Microbial Production of Isoprenoids.  
854 *Consequences Microb. Interact. with Hydrocarb. Oils, Lipids Prod. Fuels*  
855 *Chem.* **2017**, 359–382. [https://doi.org/10.1007/978-3-319-50436-0\\_219](https://doi.org/10.1007/978-3-319-50436-0_219).
- 856 (9) Ding, W.; Zhang, Y.; Shi, S. Development and Application of CRISPR/Cas in  
857 Microbial Biotechnology. *Front. Bioeng. Biotechnol.* **2020**, *0*, 711.  
858 <https://doi.org/10.3389/FBIOE.2020.00711>.
- 859 (10) Meng, J.; Qiu, Y.; Shi, S. CRISPR/Cas9 Systems for the Development of  
860 *Saccharomyces Cerevisiae* Cell Factories. *Front. Bioeng. Biotechnol.* **2020**, *0*,  
861 1304. <https://doi.org/10.3389/FBIOE.2020.594347>.
- 862 (11) Stovicek, V.; Holkenbrink, C.; Borodina, I. CRISPR/Cas System for Yeast  
863 Genome Engineering: Advances and Applications. *FEMS Yeast Res.* **2017**, *17*  
864 (5), 30. <https://doi.org/10.1093/FEMSYR/FOX030>.
- 865 (12) Malcı, K.; Walls, L. E.; Rios-Solis, L. Multiplex Genome Engineering Methods  
866 for Yeast Cell Factory Development. *Front. Bioeng. Biotechnol.* **2020**, *8*, 1264.  
867 <https://doi.org/10.3389/FBIOE.2020.589468>.
- 868 (13) Hille, F.; Charpentier, E. CRISPR-Cas: Biology, Mechanisms and Relevance.  
869 *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371* (1707).  
870 <https://doi.org/10.1098/RSTB.2015.0496>.
- 871 (14) Richter, C.; Chang, J. T.; Fineran, P. C. Function and Regulation of Clustered  
872 Regularly Interspaced Short Palindromic Repeats (CRISPR) / CRISPR  
873 Associated (Cas) Systems. *Viruses* **2012**, *4* (10), 2291.  
874 <https://doi.org/10.3390/V4102291>.
- 875 (15) Deaner, M.; Alper, H. S. Systematic Testing of Enzyme Perturbation  
876 Sensitivities via Graded DCas9 Modulation in *Saccharomyces Cerevisiae*.  
877 *Metab. Eng.* **2017**, *40*, 14–22. <https://doi.org/10.1016/J.YMBEN.2017.01.012>.
- 878 (16) Generoso, W. C.; Gottardi, M.; Oreb, M.; Boles, E. Simplified CRISPR-Cas  
879 Genome Editing for *Saccharomyces Cerevisiae*. *J. Microbiol. Methods* **2016**,  
880 *127*, 203–205. <https://doi.org/10.1016/J.MIMET.2016.06.020>.
- 881 (17) Zalatan, J. G.; Lee, M. E.; Almeida, R.; Gilbert, L. A.; Whitehead, E. H.; Russa,  
882 M. La; Tsai, J. C.; Weissman, J. S.; Dueber, J. E.; Qi, L. S.; Lim, W. A.  
883 Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA  
884 Scaffolds. *Cell* **2015**, *160* (0), 339.  
885 <https://doi.org/10.1016/J.CELL.2014.11.052>.
- 886 (18) Vanegas, K. G.; Lehka, B. J.; Mortensen, U. H. SWITCH: A Dynamic CRISPR  
887 Tool for Genome Engineering and Metabolic Pathway Control for Cell Factory  
888 Construction in *Saccharomyces Cerevisiae*. *Microb. Cell Fact.* **2017**, *16* (1),  
889 25. <https://doi.org/10.1186/S12934-017-0632-X>.
- 890 (19) Laughery, M. F.; Hunter, T.; Brown, A.; Hoopes, J.; Ostbye, T.; Shumaker, T.;  
891 Wyrick, J. J. New Vectors for Simple and Streamlined CRISPR-Cas9 Genome  
892 Editing in *Saccharomyces Cerevisiae*. *Yeast* **2015**, *32* (12), 711.

- 893 <https://doi.org/10.1002/YEA.3098>.
- 894 (20) Jakoćinas, T.; Bonde, I.; Herrgård, M.; Harrison, S. J.; Kristensen, M.;  
895 Pedersen, L. E.; Jensen, M. K.; Keasling, J. D. Multiplex Metabolic Pathway  
896 Engineering Using CRISPR/Cas9 in *Saccharomyces Cerevisiae*. *Metab. Eng.*  
897 **2015**, *28*, 213–222. <https://doi.org/10.1016/J.YMBEN.2015.01.008>.
- 898 (21) Zhang, G.-C.; Kong, I. I.; Kim, H.; Liu, J.-J.; Cate, J. H. D.; Jin, Y.-S.  
899 Construction of a Quadruple Auxotrophic Mutant of an Industrial Polyploid  
900 *Saccharomyces Cerevisiae* Strain by Using RNA-Guided Cas9 Nuclease.  
901 *Appl. Environ. Microbiol.* **2014**, *80* (24), 7694.  
902 <https://doi.org/10.1128/AEM.02310-14>.
- 903 (22) Symington, L. S.; Gautier, J. Double-Strand Break End Resection and Repair  
904 Pathway Choice. <http://dx.doi.org/10.1146/annurev-genet-110410-132435>  
905 **2011**, *45*, 247–271. [https://doi.org/10.1146/ANNUREV-GENET-110410-](https://doi.org/10.1146/ANNUREV-GENET-110410-132435)  
906 [132435](https://doi.org/10.1146/ANNUREV-GENET-110410-132435).
- 907 (23) DiCarlo, J. E.; Norville, J. E.; Mali, P.; Rios, X.; Aach, J.; Church, G. M.  
908 Genome Engineering in *Saccharomyces Cerevisiae* Using CRISPR-Cas  
909 Systems. *Nucleic Acids Res.* **2013**, *41* (7), 4336.  
910 <https://doi.org/10.1093/NAR/GKT135>.
- 911 (24) Ryan, O. W.; Skerker, J. M.; Maurer, M. J.; Li, X.; Tsai, J. C.; Poddar, S.; Lee,  
912 M. E.; DeLoache, W.; Dueber, J. E.; Arkin, A. P.; Cate, J. H. Selection of  
913 Chromosomal DNA Libraries Using a Multiplex CRISPR System. *Elife* **2014**, *3*  
914 (August2014), 1–15. <https://doi.org/10.7554/ELIFE.03703>.
- 915 (25) Ronda, C.; Maury, J.; Jakoćinas, T.; Baallal Jacobsen, S. A.; Germann, S. M.;  
916 Harrison, S. J.; Borodina, I.; Keasling, J. D.; Jensen, M. K.; Nielsen, A. T.  
917 CrEdit: CRISPR Mediated Multi-Loci Gene Integration in *Saccharomyces*  
918 *Cerevisiae*. *Microb. Cell Fact.* **2015**, *14* (1), 97. [https://doi.org/10.1186/s12934-](https://doi.org/10.1186/s12934-015-0288-3)  
919 [015-0288-3](https://doi.org/10.1186/s12934-015-0288-3).
- 920 (26) Engler, C.; Kandzia, R.; Marillonnet, S. A One Pot, One Step, Precision  
921 Cloning Method with High Throughput Capability. *PLoS One* **2008**, *3* (11),  
922 e3647. <https://doi.org/10.1371/JOURNAL.PONE.0003647>.
- 923 (27) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.;  
924 Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred  
925 Kilobases. *Nat. Methods* **2009**, *6* (5), 343–345.  
926 <https://doi.org/10.1038/nmeth.1318>.
- 927 (28) Nour-Eldin, H. H.; Hansen, B. G.; Nørholm, M. H. H.; Jensen, J. K.; Halkier, B.  
928 A. Advancing Uracil-Excision Based Cloning towards an Ideal Technique for  
929 Cloning PCR Fragments. *Nucleic Acids Res.* **2006**, *34* (18), e122.  
930 <https://doi.org/10.1093/NAR/GKL635>.
- 931 (29) Akhmetov, A.; Laurent, J. M.; Gollihar, J.; Gardner, E. C.; Garge, R. K.;  
932 Ellington, A. D.; Kachroo, A. H.; Marcotte, E. M. Single-Step Precision  
933 Genome Editing in Yeast Using CRISPR-Cas9. *Bio-protocol* **2018**, *8* (6).  
934 <https://doi.org/10.21769/BIOPROTOCOL.2765>.
- 935 (30) Bao, Z.; Xiao, H.; Liang, J.; Zhang, L.; Xiong, X.; Sun, N.; Si, T.; Zhao, H.  
936 Homology-Integrated CRISPR-Cas (HI-CRISPR) System for One-Step  
937 Multigene Disruption in *Saccharomyces Cerevisiae*. *ACS Synth. Biol.* **2015**, *4*  
938 (5), 585–594. <https://doi.org/10.1021/sb500255k>.
- 939 (31) Jakoćinas, T.; Bonde, I.; Herrgård, M.; Harrison, S. J.; Kristensen, M.;  
940 Pedersen, L. E.; Jensen, M. K.; Keasling, J. D. Multiplex Metabolic Pathway

- Engineering Using CRISPR/Cas9 in *Saccharomyces Cerevisiae*. *Metab. Eng.* **2015**, *28*, 213–222. <https://doi.org/10.1016/j.ymben.2015.01.008>.
- (32) Ferreira, R.; Skrekas, C.; Nielsen, J.; David, F. Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in *Saccharomyces Cerevisiae*. *ACS Synth. Biol.* **2018**, *7* (1), 10–15. <https://doi.org/10.1021/acssynbio.7b00259>.
- (33) NEB® Golden Gate Assembly Kit (Bsal-HF®v2) | NEB [https://international.neb.com/products/e1601-neb-golden-gate-assembly-mix#Product Information](https://international.neb.com/products/e1601-neb-golden-gate-assembly-mix#Product%20Information) (accessed Aug 26, 2021).
- (34) Gibson Assembly® Cloning Kit | NEB [https://international.neb.com/products/e5510-gibson-assembly-cloning-kit#Product Information](https://international.neb.com/products/e5510-gibson-assembly-cloning-kit#Product%20Information) (accessed Aug 26, 2021).
- (35) Shevchuk, N. A.; Bryksin, A. V.; Nusinovich, Y. A.; Cabello, F. C.; Sutherland, M.; Ladisch, S. Construction of Long DNA Molecules Using Long PCR-Based Fusion of Several Fragments Simultaneously. *Nucleic Acids Res.* **2004**, *32* (2), e19. <https://doi.org/10.1093/NAR/GNH014>.
- (36) Karvelis, T.; Gasiunas, G.; Miksys, A.; Barrangou, R.; Horvath, P.; Siksnys, V. CrRNA and TracrRNA Guide Cas9-Mediated DNA Interference in *Streptococcus Thermophilus*. *RNA Biol.* **2013**, *10* (5), 841. <https://doi.org/10.4161/RNA.24203>.
- (37) Wilson, L. O. W.; O'Brien, A. R.; Bauer, D. C. The Current State and Future of CRISPR-Cas9 GRNA Design Tools. *Front. Pharmacol.* **2018**, *9* (JUN). <https://doi.org/10.3389/FPHAR.2018.00749>.
- (38) Zhang, T.; Gao, Y.; Wang, R.; Zhao, Y. Production of Guide RNAs in Vitro and in Vivo for CRISPR Using Ribozymes and RNA Polymerase II Promoters. *Bio-protocol* **2017**, *7* (4). <https://doi.org/10.21769/BIOPROTOCOL.2148>.
- (39) Horlbeck, M. A.; Witkowski, L. B.; Guglielmi, B.; Replogle, J. M.; Gilbert, L. A.; Villalta, J. E.; Torigoe, S. E.; Tjian, R.; Weissman, J. S. Nucleosomes Impede Cas9 Access to DNA in Vivo and in Vitro. *Elife* **2016**, *5* (MARCH2016). <https://doi.org/10.7554/ELIFE.12677>.
- (40) Reider Apel, A.; d'Espaux, L.; Wehrs, M.; Sachs, D.; Li, R. A.; Tong, G. J.; Garber, M.; Nnadi, O.; Zhuang, W.; Hillson, N. J.; Keasling, J. D.; Mukhopadhyay, A. A Cas9-Based Toolkit to Program Gene Expression in *Saccharomyces Cerevisiae*. *Nucleic Acids Res.* **2017**, *45* (1), 496–508. <https://doi.org/10.1093/nar/gkw1023>.
- (41) Wu, X.-L.; Li, B.-Z.; Zhang, W.-Z.; Song, K.; Qi, H.; Dai, J.; Yuan, Y.-J. Genome-Wide Landscape of Position Effects on Heterogeneous Gene Expression in *Saccharomyces Cerevisiae*. *Biotechnol. Biofuels* **2017**, *10* (1), 1–10. <https://doi.org/10.1186/S13068-017-0872-3>.
- (42) SnapGene | Software for everyday molecular biology <https://www.snapgene.com/> (accessed Aug 28, 2021).
- (43) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9* (7), 671–675. <https://doi.org/10.1038/nmeth.2089>.
- (44) Colony Counter <https://imagej.nih.gov/ij/plugins/colony-counter.html> (accessed Aug 29, 2021).
- (45) Brewster, J. D. A Simple Micro-Growth Assay for Enumerating Bacteria. *J. Microbiol. Methods* **2003**, *53* (1), 77–86. <https://doi.org/10.1016/S0167->

- 7012(02)00226-9.
- (46) Lichten, C. A.; White, R.; Clark, I. B.; Swain, P. S. Unmixing of Fluorescence Spectra to Resolve Quantitative Time-Series Measurements of Gene Expression in Plate Readers. *BMC Biotechnol.* 2014 141 **2014**, 14 (1), 1–11. <https://doi.org/10.1186/1472-6750-14-11>.
- (47) Karolchik, D.; Baertsch, R.; Diekhans, M.; Furey, T. S.; Hinrichs, A.; Lu, Y. T.; Roskin, K. M.; Schwartz, M.; Sugnet, C. W.; Thomas, D. J.; Weber, R. J.; Haussler, D.; Kent, W. J. The UCSC Genome Browser Database. *Nucleic Acids Res.* **2003**, 31 (1), 51. <https://doi.org/10.1093/NAR/GKG129>.
- (48) Concordet, J.-P.; Haeussler, M. CRISPOR: Intuitive Guide Selection for CRISPR/Cas9 Genome Editing Experiments and Screens. *Nucleic Acids Res.* **2018**, 46 (W1), W242–W245. <https://doi.org/10.1093/NAR/GKY354>.
- (49) Doench, J. G.; Fusi, N.; Sullender, M.; Hegde, M.; Vaimberg, E. W.; Donovan, K. F.; Smith, I.; Tothova, Z.; Wilen, C.; Orchard, R.; Virgin, H. W.; Listgarten, J.; Root, D. E. Optimized SgRNA Design to Maximize Activity and Minimize Off-Target Effects of CRISPR-Cas9. *Nat. Biotechnol.* 2015 342 **2016**, 34 (2), 184–191. <https://doi.org/10.1038/nbt.3437>.
- (50) Moreno-Mateos, M. A.; Vejnar, C. E.; Beaudoin, J.-D.; Fernandez, J. P.; Mis, E. K.; Khokha, M. K.; Giraldez, A. J. CRISPRscan: Designing Highly Efficient SgRNAs for CRISPR-Cas9 Targeting in Vivo. *Nat. Methods* 2015 1210 **2015**, 12 (10), 982–988. <https://doi.org/10.1038/nmeth.3543>.
- (51) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* 2020 173 **2020**, 17 (3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
- (52) Waskom, M. L. Seaborn: Statistical Data Visualization. *J. Open Source Softw.* **2021**, 6 (60), 3021. <https://doi.org/10.21105/JOSS.03021>.
- (53) omniplate <https://swainlab.bio.ed.ac.uk/software/platereader/omniplate.html> (accessed Aug 29, 2021).
- (54) BioRender <https://biorender.com/> (accessed Oct 23, 2021).
- (55) Gibson, D. G.; Benders, G. A.; Axelrod, K. C.; Zaveri, J.; Algire, M. A.; Moodie, M.; Montague, M. G.; Venter, J. C.; Smith, H. O.; Hutchison, C. A. One-Step Assembly in Yeast of 25 Overlapping DNA Fragments to Form a Complete Synthetic Mycoplasma Genitalium Genome. *Proc. Natl. Acad. Sci.* **2008**, 105 (51), 20404–20409. <https://doi.org/10.1073/PNAS.0811011106>.
- (56) Kuijpers, N. G.; Solis-Escalante, D.; Bosman, L.; van den Broek, M.; Pronk, J. T.; Daran, J.-M.; Daran-Lapujade, P. A Versatile, Efficient Strategy for Assembly of Multi-Fragment Expression Vectors in *Saccharomyces Cerevisiae* Using 60 Bp Synthetic Recombination Sequences. *Microb. Cell Factories* 2013 121 **2013**, 12 (1), 1–13. <https://doi.org/10.1186/1475-2859-12-47>.
- (57) Eisenstein, M. Enzymatic DNA Synthesis Enters New Phase. *Nat. Biotechnol.* **2020**, 38 (10), 1113–1115. <https://doi.org/10.1038/S41587-020-0695-9>.
- (58) Hughes, R. A.; Ellington, A. D. Synthetic DNA Synthesis and Assembly:



- Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **2017**, 9 (1). <https://doi.org/10.1101/CSHPERSPECT.A023812>.
- (59) Portela, R. M. C.; Vogl, T.; Kniely, C.; Fischer, J. E.; Oliveira, R.; Glieder, A. Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species. *ACS Synth. Biol.* **2016**, 6 (3), 471–484. <https://doi.org/10.1021/ACSSYNBIO.6B00178>.
- (60) Flagfeldt, D. B.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of Chromosomal Integration Sites for Heterologous Gene Expression in *Saccharomyces Cerevisiae*. *Yeast* **2009**, 26 (10), 545–551. <https://doi.org/10.1002/YEA.1705>.
- (61) Wu, X. Le; Bi, Y. H.; Gao, F.; Xie, Z. X.; Li, X.; Zhou, X.; Ma, D. J.; Li, B. Z.; Yuan, Y. J. The Effect of Autonomously Replicating Sequences on Gene Expression in *Saccharomyces Cerevisiae*. *Biochem. Eng. J.* **2019**, 149, 107250. <https://doi.org/10.1016/J.BEJ.2019.107250>.
- (62) Hsu, P. D.; Scott, D. A.; Weinstein, J. A.; Ran, F. A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E. J.; Wu, X.; Shalem, O.; Cradick, T. J.; Marraffini, L. A.; Bao, G.; Zhang, F. DNA Targeting Specificity of RNA-Guided Cas9 Nucleases. *Nat. Biotechnol.* **2013**, 31 (9), 827–832. <https://doi.org/10.1038/nbt.2647>.
- (63) Stratigopoulos, G.; Rosa, M. C. De; LeDuc, C. A.; Leibel, R. L.; Doege, C. A. DMSO Increases Efficiency of Genome Editing at Two Non-Coding Loci. *PLoS One* **2018**, 13 (6), e0198637. <https://doi.org/10.1371/JOURNAL.PONE.0198637>.
- (64) Beal, J.; Goñi-Moreno, A.; Myers, C.; Hecht, A.; Vicente, M. del C.; Parco, M.; Schmidt, M.; Timmis, K.; Baldwin, G.; Friedrichs, S.; Freemont, P.; Kiga, D.; Ordozgoiti, E.; Rennig, M.; Rios, L.; Tanner, K.; Lorenzo, V.; Porcar, M. The Long Journey towards Standards for Engineering Biosystems. *EMBO Rep.* **2020**, 21 (5). <https://doi.org/10.15252/embr.202050521>.
- (65) Why PCR product can't be re-amplified in second round? [https://www.researchgate.net/post/Why\\_PCR\\_product\\_cant\\_be\\_re-amplified\\_in\\_second\\_round](https://www.researchgate.net/post/Why_PCR_product_cant_be_re-amplified_in_second_round) (accessed Sep 1, 2021).
- (66) Gao, S.; Honey, S.; Fitcher, B.; Grollman, A. P. The Non-Homologous End-Joining Pathway of *S. Cerevisiae* Works Effectively in G1-Phase Cells, and Religates Cognate Ends Correctly and Non-Randomly. *DNA Repair (Amst)*. **2016**, 42, 1. <https://doi.org/10.1016/J.DNAREP.2016.03.013>.
- (67) Emerson, C. H.; Bertuch, A. A. Consider the Workhorse: Nonhomologous End Joining in Budding Yeast. *Biochem. Cell Biol.* **2016**, 94 (5), 396. <https://doi.org/10.1139/BCB-2016-0001>.
- (68) Jang, I.-S.; Yu, B. J.; Jang, J. Y.; Jegal, J.; Lee, J. Y. Improving the Efficiency of Homologous Recombination by Chemical and Biological Approaches in *Yarrowia Lipolytica*. *PLoS One* **2018**, 13 (3). <https://doi.org/10.1371/JOURNAL.PONE.0194954>.
- (69) Dong, C.; Jiang, L.; Xu, S.; Huang, L.; Cai, J.; Lian, J.; Xu, Z. A Single Cas9-VPR Nuclease for Simultaneous Gene Activation, Repression, and Editing in *Saccharomyces Cerevisiae*. *ACS Synth. Biol.* **2020**, 9 (9), 2252–2257. [https://doi.org/10.1021/ACSSYNBIO.0C00218/SUPPL\\_FILE/SB0C00218\\_SI\\_001.PDF](https://doi.org/10.1021/ACSSYNBIO.0C00218/SUPPL_FILE/SB0C00218_SI_001.PDF).
- (70) Buitrago, D.; Labrador, M.; Arcon, J. P.; Lema, R.; Flores, O.; Esteve-Codina,



- 1085 A.; Blanc, J.; Villegas, N.; Bellido, D.; Gut, M.; Dans, P. D.; Heath, S. C.; Gut, I.  
1086 G.; Brun Heath, I.; Orozco, M. Impact of DNA Methylation on 3D Genome  
1087 Structure. *Nat. Commun.* 2021 121 **2021**, 12 (1), 1–17.  
1088 <https://doi.org/10.1038/s41467-021-23142-8>.
- 1089 (71) Pilauri, V.; Bewley, M.; Diep, C.; Hopper, J. Gal80 Dimerization and the Yeast  
1090 GAL Gene Switch. *Genetics* **2005**, 169 (4), 1903.  
1091 <https://doi.org/10.1534/GENETICS.104.036723>.
- 1092 (72) Briza, P.; Eckerstorfer, M.; Breitenbach, M. The Sporulation-Specific Enzymes  
1093 Encoded by the DIT1 and DIT2 Genes Catalyze a Two-Step Reaction Leading  
1094 to a Soluble LL-Dityrosine-Containing Precursor of the Yeast Spore Wall. *Proc.*  
1095 *Natl. Acad. Sci. U. S. A.* **1994**, 91 (10), 4524–4528.  
1096 <https://doi.org/10.1073/PNAS.91.10.4524>.
- 1097 (73) Hahn-Hägerdal, B.; Karhumaa, K.; Larsson, C. U.; Gorwa-Grauslund, M.;  
1098 Görgens, J.; van Zyl, W. H. Role of Cultivation Media in the Development of  
1099 Yeast Strains for Large Scale Industrial Use. *Microb. Cell Fact.* **2005**, 4 (1), 1–  
1100 16. <https://doi.org/10.1186/1475-2859-4-31/TABLES/3>.
- 1101 (74) Dymond, J. S. *Saccharomyces Cerevisiae* Growth Media. *Methods Enzymol.*  
1102 **2013**, 533, 191–204. <https://doi.org/10.1016/B978-0-12-420067-8.00012-X>.
- 1103 (75) Swain, P. S.; Stevenson, K.; Leary, A.; Montano-Gutierrez, L. F.; Clark, I. B.  
1104 N.; Vogel, J.; Pilizota, T. Inferring Time Derivatives Including Cell Growth  
1105 Rates Using Gaussian Processes. *Nat. Commun.* 2016 71 **2016**, 7 (1), 1–8.  
1106 <https://doi.org/10.1038/ncomms13766>.
- 1107 (76) Ringel, A. E.; Ryznar, R.; Picariello, H.; Huang, K. lin; Lazarus, A. G.; Holmes,  
1108 S. G. Yeast Tdh3 (Glyceraldehyde 3-Phosphate Dehydrogenase) Is a Sir2-  
1109 Interacting Factor That Regulates Transcriptional Silencing and RDNA  
1110 Recombination. *PLoS Genet.* **2013**, 9 (10), 1003871.  
1111 <https://doi.org/10.1371/JOURNAL.PGEN.1003871>.
- 1112