Development and application of the Landing Pad platform: A synthetic CRISPR/Cas9 platform for multi-copy gene integration in *Saccharomyces cerevisiae*

Leanne Bourgeois

A Thesis
in
The Department
of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Biology) at
Concordia University
Montreal, Quebec, Canada

March 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

by: Leanne Bourgeois

entitled: Development and application of the Landing Pad platform: A synthetic CRISPR/Cas9 platform for multi-copy gene integration in *Saccharomyces cerevisiae*.

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Ian Ferguson

_____ Examiner
Dr. David Walsh

_____ Examiner
Dr. Malcolm Whiteway

_____ Examiner
Dr. Paul Joyce

_____ Supervisor
Dr. Vincent Martin

Approved by _____
Chair of Department or Graduate Program Director

Date_____2018 _____
Dean of Faculty

ABSTRACT

Development and application of the Landing Pad platform: A synthetic CRISPR/Cas9 platform for multi-copy gene integration in *Saccharomyces cerevisiae*.

Leanne Bourgeois

To accelerate the construction of superior yeast strains producing high-value chemicals, we developed a modular CRISPR/Cas9 integration platform in *S. cerevisiae* that accommodates marker-less, multi-copy gene integration. Our engineering strategy introduced a series of synthetic DNA parts, called Landing Pads (LP), into the *S. cerevisiae* genome to act as modular anchors for heterologous gene integrations. The LPs have been designed to accommodate the CRISPR/Cas9 genome editing system and to facilitate multi-copy gene integration of one, two, three and four copies in a single step. First we designed ten synthetic gRNA targeting sequences and evaluated their targeting specificity, integration efficiency, and possible off-target effects. We also surveyed 16 genomic loci for landing pad integration by evaluating the integration efficiency and gene expression profiles at each site. The results gleaned from our preliminary tests informed the final configuration of the LP platform strain. To demonstrate the utility of our LP integration system, we used the platform to screen ten variants of norcoclaurine synthase (NCS), a notoriously inefficient enzyme that catalyzes the first committed step in the production of high-value benzylisoquinoline alkaloids (BIA). The platform enabled rapid integration of each NCS variant in one, two, three and four copies in parallel, yielding 40 strains total. LC/MS analysis identified two variants, *Nd*NCS and *Sc*NCS that produce higher concentrations of the BIA scaffold (*S*)-norcoclaurine by increasing copy number, suggesting that the proposed strategy may help alleviate enzyme inefficiencies in pathway engineering.

ACKNOWLEDGEMENTS

I would like to extend gratitude to my supervisor Dr. Vincent Martin for all of the opportunities he has afforded me. His guidance and expertise was invaluable and I feel very fortunate to have been a part of the Martin lab. Thank you, Vince!

Next, I would like to give a big thanks to Michael Pyne, who was a major part of this project and also the 'editor-in-chief' for my thesis. Mike is both a skilled scientist and a wordsmith and I am incredibly grateful for all of his help at the bench and throughout the writing process. Thanks, Mike!

Many thanks goes out to the members of the Martin lab, both past and present for all of their support and positivity: Audrey Morin, Lauren Narcross, Elena Fossati, Nicholas Gold, Mindy Melgar, Alex Vallerand, Meghan Davies, Mohamed Nasr, Will Cheney, Smita Amarnath. Thanks everyone!

Another thanks goes out to Marcos Di Falco for helping us with all things LC/MS related.

I would also like to thank my committee members, Dr. Malcolm Whiteway and Dr. David Walsh for taking the time to read my thesis and for providing their thoughtful feedback.

And finally, I would like to thank my Mom & Dad for their endless support ♥

AUTHOR CONTRIBUTIONS

This project was conceived by Vincent J.J. Martin. Michael Pyne designed the LP.A and LP.Z recombinogenic regions, and built the CEN.LP and CEN.LP.D strains used in this study. James Scriven selected the NCS candidates that were screened in this study. All other experiments were designed, implemented and analyzed by Leanne Bourgeois.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS & ACRONYMS

4-HPAA: 4-hydroxyphenylacetaldehyde

4-HPP: 4-hydroxyphenylpyruvate

AAA: aromatic amino acid

ARO4: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase

$Aro4^{FBR}$: ARO4 feedback resistant mutant

BIA: benzylisoquinoline alkaloids

bp: base pair

Cas: CRISPR-associated protein

Cas9: CRISPR-associated protein 9 from *Streptococcus pyogenes*

CEN.LP: LP Platform strain

CEN.LP.D: dopamine producing LP platform strain

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

CYP76AD1: tyrosine hydroxylase

DNA: deoxyribonucleic acid

DODC: dopa decarboxylase

DSB: double stranded break

FgF: Flagfeldt site

GFP: green fluorescent protein

HDR: homology direct repair

HR: homologous recombination

gRNA: guide RNA

$gRNA_X$: synthetic guide RNA (X = 1-10)

kb: kilobase

LP: landing pad

LP_X: landing pad motif (X = 1-4)

NCS: norcoclaurine synthase

nt: nucleotide

PAM: protospacer adjacent motif

RNA: ribonucleic acid

$Site_X$: genomic loci

S. cerevisiae: *Saccharomyces cerevisiae*

S. pyogenes: *Streptococcus pyogenes*

T_X: Target site (X = 1-10)

USER: uracil-specific excision reaction

USERXII-1: genomic locus based on USER cloning

YHR: yeast homologous recombination

1. INTRODUCTION

1.1. Yeast Cell Factories

Engineering microbes into cellular factories is a rapidly growing field of biotechnology that has broad applications within the energy, agriculture and human health sectors. Advances in DNA sequencing and synthesis technologies and the development of genome-editing tools have enabled us to rationally engineer microorganisms in order to produce bio-based fuels, chemicals and pharmaceuticals from renewable feedstocks [1-4]. The yeast *Saccharomyces cerevisiae* is a common industrial microbe traditionally used for ethanol fermentation and leavening bread. As such, *S. cerevisiae* is a preferred host for producing fuels and chemicals due to its GRAS (generally regarded as safe) status, high stress tolerance and general robustness for industrial fermentation [1, 2]. The availability of genetic tools and genetic tractability of *S. cerevisiae* also benefits metabolic engineering efforts towards the synthesis of high-value compounds [5-7]. In the last decade, *S. cerevisiae* has been successfully engineered for commercial production of high-value chemicals, including isobutanol and farnesene, as well as the anti-malarial drug artemisinin [4]. However, the vast majority of metabolic engineering efforts in *S. cerevisiae* have remained 'proof-of-concept', and have a long way to go before they achieve commercial viability with regard to titer, yield and productivity [8].

Engineering yeast cell factories to synthesize non-native compounds involves reconstituting heterologous pathways, as well as modifying native metabolic processes. The first step in the production of new biocompounds is to characterize the biosynthetic pathway of the target compound from the source organism. Once the target pathway has been elucidated, genes encoding the respective enzymes are either isolated directly from the host or synthesized as synthetic DNA [9, 10]. In yeast, pathway genes are assembled into individual expression cassettes and then introduced to the host on episomal plasmids or integrated directly into the chromosome. Initial proof-of-concept strains typically only produce small amounts of the target compound. This is because production hosts have evolved to prioritize native metabolic processes or the heterologous pathway does not function well in the non-native host [11]. Therefore, achieving commercially viable production strains often requires extensive pathway troubleshooting and optimization. As the design-build-test cycle for achieving industrially-relevant yeast strains is slow, labor-intensive and expensive. As such, researchers are continuously searching for new strategies that accelerate the construction of superior yeast strains.

1.2. Pathway optimization

Pathway optimization is an iterative process that typically involves many cycles through the design-build-test pipeline. After the initial proof-of-concept has been implemented, pathway bottlenecks are identified and targeted for optimization. Pathway bottlenecks are caused by enzyme inefficiencies or low substrate availability that limit flux through the pathway, resulting in low product yields [10, 12]. In order to enhance flux towards the target metabolite, a number of genetic optimization strategies have been implemented to improve enzyme activity and balance enzyme expression levels.

Enzyme bioprospecting

A common approach to overcome enzyme inefficiencies is to screen for improved catalytic activity from a library of orthologous enzyme variants. The wealth of publicly available 'omics' data has enabled us to mine transcriptome databases and select candidate genes encoding enzymes with the same putative function [9, 10, 13]. Using a plug-and-play strategy, each enzymatic conversion step in a pathway can be optimized by identifying variants with improved catalytic efficiency and substrate specificity. The reduced cost of DNA synthesis has made it easier to obtain genetic material by purchasing synthetic gene sequences rather than isolating and compiling variants from multiple sources [9, 10]. This also enables us to codon-optimize heterologous gene sequences for improved expression in *S. cerevisiae*, which further enhances flux through the target pathway [14, 15].

Enzyme expression levels

Controlling heterologous gene expression in multi-enzyme pathways is important for balancing flux and preventing the accumulation of side-products and toxic intermediates [10]. Genetic tuning involves the application of simple genetic tools to enhance or control gene expression in order to increase flux towards the target compound [16]. Regulating gene transcription is the primary method for balancing flux through a pathway and is controlled by swapping regulatory elements or titrating gene copy number. For instance, tuning gene expression can lead to six-fold differences in enzyme activity between promoters [17] and a 70-fold difference between terminators [18]. A large collection of regulatory elements has been characterized in yeast [19, 20], thus providing a diverse toolkit to optimize production of biosynthetic enzymes [6]. Strong constitutive promoters and expression-enhancing terminators are regularly chosen for pathway engineering to maximize the expression of biosynthetic genes [16].

Gene copy number

Titrating the copy number of the corresponding biosynthetic gene can further enhance enzyme levels. This strategy can be applied to an entire metabolic pathway to increase overall product yield [21], or targeted to specific enzymes to overcome pathway bottlenecks. Increasing gene copy number often enhances transcriptional levels, which can produce more enzyme and increase substrate turnover [22, 23], at least up to a certain limit [24]. However, modulating gene copy number may also be futile for some enzymes [25], and in some cases, can have detrimental effects on productivity and growth rate due to the build-up of toxic intermediates [26]. For these reasons, it is more prudent to optimize gene copy number for each enzyme individually to prevent deleterious effects and to avoid placing an unnecessary burden on the cell.

Modulating gene copy number is carried out by expressing the target gene from a multi-copy plasmid or by integrating one or more times into the host chromosome. Plasmid-borne expression is an efficient strategy to test a wide-range of expression levels in yeast by selecting one of two commonly-used expression vectors: i) low-copy CEN plasmids (1-2 copies per cell) or ii) high copy 2μ plasmids (10-50 copies per cell) [27]. However, the exact plasmid copy number for any given cell in a population is highly variable, and so plasmid-based systems offer limited control over fine-tuning gene dosage. Ryan et al. (2014) demonstrated that even for genes expressed from a low-copy plasmid, expression levels between individual cells were highly variable as well. Plasmid-based systems are also problematic for industrial use because cells must be propagated in selective media, and even then, plasmids could be lost or rearranged over long-term cultivation [2, 29]. To avoid these challenges, chromosomal integration is the preferred strategy for expressing heterologous genes, and also enables the fine-tuning of gene copy number, which could be critical for balancing metabolic pathway fluxes. Nevertheless, integrating multiple genes into the chromosome is a laborious process, and often necessitates iterative rounds of gene integration.

1.3. Genome editing in *S. cerevisiae*

Integration into the *S. cerevisiae* genome is achieved by harnessing the yeast homologous recombination (YHR) machinery [30-33]. This is accomplished by flanking the genetic construct and selection marker with site-specific homology arms, which then recombine with the homologous sequence at the genomic integration site. While homologous recombination (HR) is a relatively simple technique for integrating and assembling heterologous DNA, it suffers from low

integration efficiencies, even when selection is applied [34]. It was later demonstrated that integration efficiency using YHR dramatically improves when a double-strand break (DSB) is induced at the integration site [34, 35], which initiates homology-directed repair (HDR) using the exogenous DNA fragment. Several endonucleases have since been identified that induce targeted DSBs, including Zinc-Finger Nucleases (ZFN) [36], Transcription Activator-Like Effector Nucleases (TALEN) [37] and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) with CRISPR-associated (Cas) proteins [38, 39]. While all three systems employ programmable endonucleases to induce targeted DSBs, the CRISPR/Cas system has become the preferred genome-editing tool, offering more customizability and multiplexing capabilities.

CRISPR/Cas9-mediated genome editing

CRISPR/Cas is a bacterial defense system [38, 40, 41] that has recently been adapted into a powerful genome editing tool [38]. The type II CRISPR/Cas9 system from the bacterial species *Streptococcus pyogenes* is the most common CRISPR/Cas system that has been harnessed for targeted genome engineering in bacteria [42], yeast [39], plants [43, 44], and animals [45-47]. The CRISPR/Cas9 system has two main components, the RNA-guided Cas9 nuclease that cleaves dsDNA, and a programmable guide RNA (gRNA) that directs Cas9 to a specified target site [38]. The only prerequisite for achieving targeted Cas9-mediated DSBs is the presence of a Cas9 binding signal 5'-NGG-3', known as the protospacer adjacent motif (PAM), which can be found virtually anywhere in the genome [38, 45, 48]. Once the genomic target site is defined by the user, a 20-nucleotide site-specific ‘targeting’ sequence is programmed into the gRNA, which guides Cas9 to the complementary sequence upstream of the PAM. The Cas9-gRNA complex then binds to the target site, triggering DNA unwinding and Cas9-mediated DSB formation [49]. Following DSB induction, native HDR machinery is initiated in the presence of a DNA repair template, resulting in targeted integration of exogenous DNA without the need for selection [39].

CRISPR/Cas9 enables scar-less, marker-free genome editing with high efficiency and fidelity in *S. cerevisiae* [39]. By expressing multiple gRNAs to target different regions of the genome, the CRISPR/Cas9 system also supports multiplex engineering strategies [45, 50]. This has greatly increased the efficiency of pathway reconstruction and optimization by reducing the number of successive integration events required to integrate new constructs and improve pathway efficiency.

Multiplexed CRISPR/Cas9 genome engineering strategies in *S. cerevisiae*

The CRISPR/Cas9 genome editing system was first demonstrated in *S. cerevisiae* by DiCarlo et al. (2013), who successfully performed a single gene knock-out and knock-in using 90 bp double-stranded oligonucleotides (dsOligos) for HDR of the Cas9-mediated DSB. Since then, many different CRISPR/Cas9-based strategies have emerged that have improved the efficiency of targeted genome editing and expanded the CRISPR/Cas9 genome-editing toolbox. Ryan et al. (2014) modified the gRNA architecture by fusing the gRNA to the 3' end of a self-cleaving ribozyme, which enhanced intracellular gRNA levels and facilitated targeted integration of a three-part DNA assembly at a single locus. Bao et al. (2014) developed a multiplexed CRISPR/Cas9 strategy designed to disrupt three different genes in a single reaction by introducing frameshift mutations with a 100 bp 'heterology block'. Horwitz et al. (2015) established a multiplexed CRISPR/Cas9 strategy for integrating an entire biochemical pathway by introducing separate DNA fragments containing multiple genes into the *S. cerevisiae* genome at three separate loci. The multiplexed CRISPR/Cas9 integration strategies developed by Ronda et al. (2015) and Jakočiūnas et al. (2015) applied the EasyClone system for simultaneous CRISPR/Cas9-mediated integration into the User sites. Ronda et al. (2015) reported the highest integration efficiencies for multiplexed integration of three large 8 kb DNA constructs, while Jakočiūnas et al. (2015) achieved a 5-part *in vivo* assembly at three separate loci. Finally, by targeting delta (δ) sequences of Ty retrotransposons, Shi et al. (2016) successfully integrated up to 18 copies of a 24 kb xylose utilization and (*R,R*)-2,3-butanediol (BDO) production pathway in a single assay, though copy number and integration loci were not controlled.

1.4. Landing Pad platform

The majority of multiplexed CRISPR/Cas9 engineering strategies that have been initiated in *S. cerevisiae* have concentrated on the pathway construction process [21, 53-55] or the simultaneous disruption of multiple native genes [28, 52, 56]. To expand the multiplexed CRISPR/Cas9 toolbox, the goal of this study was to develop a novel method to advance the strain optimization process. To this end, we implemented a multiplexed CRISPR/Cas9 integration platform in *S. cerevisiae* that facilitates marker-less, multi-copy gene integration in a single step. Our strategy was designed to simplify enzyme library screens and test whether modulating gene copy number improves the activity of enzyme variants.

Project overview

The Landing Pad (LP) platform is a CRISPR/Cas9 integration system built into the *S. cerevisiae* genome that uses synthetic DNA parts, called Landing Pads (LP), to facilitate multi-copy gene integrations (Figure 1). The LP platform is comprised of four distinct LP_X constructs (LP1, LP2, LP3, LP4) that were inserted throughout the yeast genome at different copy numbers. Each LP contains a unique 5'-N_{20}NGG-3' gRNA target sequence flanked by two recombinogenic regions, and are used as anchors to accommodate CRISPR/Cas9-mediated gene integration (Figure 2). The number assigned to each LP_X represents their genomic copy number motif within the LP platform strain, i.e. the number of times the LP_X was inserted into the *S. cerevisiae* genome; LP1 was inserted at one locus, LP2 was inserted at two loci, LP3 was inserted at three loci and LP4 was inserted at four loci. Since each LP_X within the same motif contains identical target sites and recombinogenic regions, a single gRNA and donor construct are used for CRISPR/Cas9-mediated gene integration. In other words, the LP_X motifs act as modular target sites for multi-copy CRISPR/Cas9-mediated integration of genes in one, two, three or four copies in a single step.

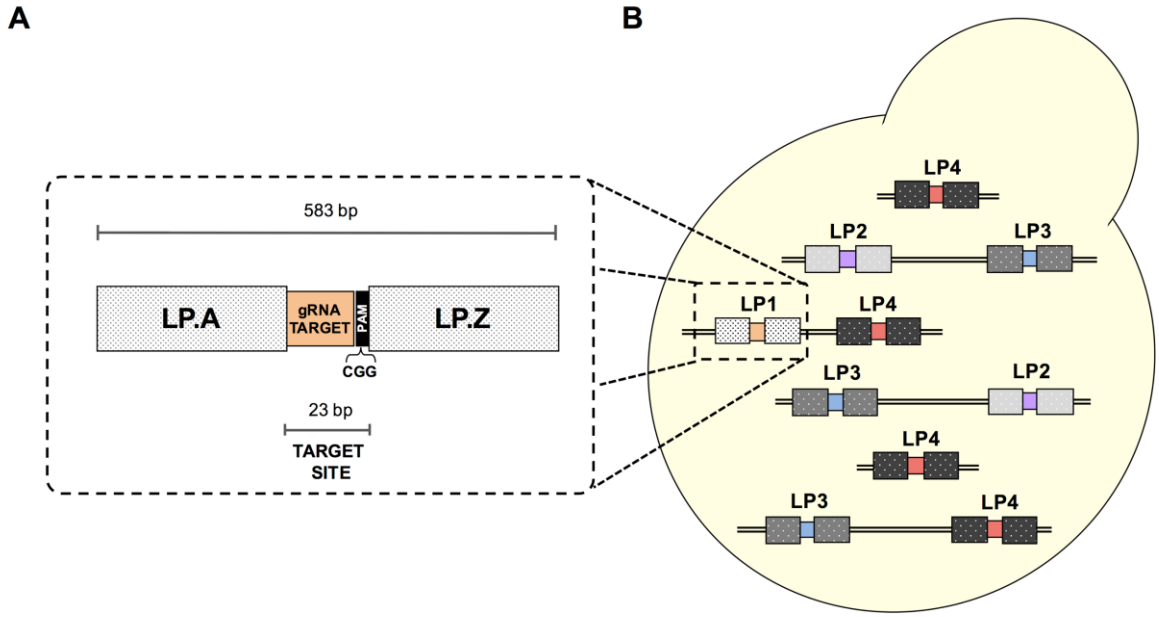


Figure 1. Schematic overview of the CRISPR/Cas9 Landing Pad platform in *S. cerevisiae*.
(A) Landing pad construct**.** A synthetic block of DNA integrated into the yeast genome that serves as a modular anchor for CRISPR/Cas9-mediated gene integration. Consists of modular parts: a target site (5'-N_{20}CGG-3') for guiding CRISPR/Cas9 endonuclease activity and LP.A and LP.Z recombinogenic regions (~250 bp) for gene integration via HDR of Cas9-induced DSBs using donor DNA flanked by A/Z homology arms. **(B)** *S. cerevisiae* landing pad platform. The platform consists of four distinct landing pads (LP1, LP2, LP3, LP4) integrated into 10 different genomic loci. Each landing pad contains a synthetic target site paired with a unique set of recombinogenic anchors. The LP_X number corresponds to the number of copies of each landing pad in the genome.

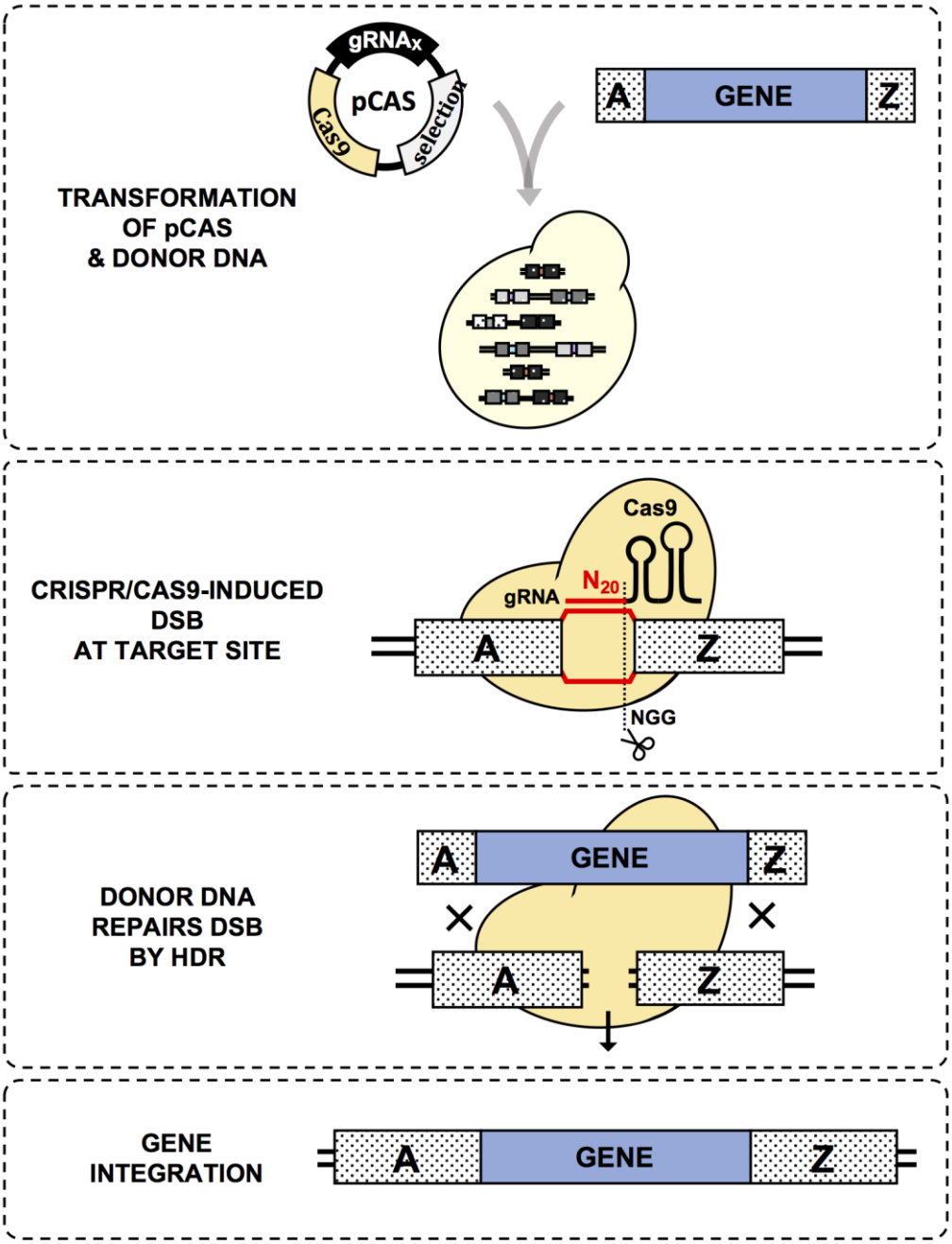


Figure 2. CRISPR/Cas9-mediated gene integration into a landing pad

Targeted integration into the landing pad platform begins by co-transforming a linear Cas9 vector and LP-gRNA cassette with a gene construct flanked by LP.A/LP.Z homology arms. Upon transformation, the LP-gRNA cassette and Cas9 vector backbone assemble into a circular expression vector via gap repair. The Cas9 and LP-gRNA form a ribonucleoprotein complex that is directed to the landing pad by the 5'-N_{20}CGG-3' targeting sequence of the LP-gRNA. Cas9 endonuclease activity is initiated upon recognition of the PAM (NGG), creating a DSB at the LP target site. The gene construct repairs the DSB through HDR, which restores the chromosome upon integration into the LP site.

Summary of the work done in this study

To construct the LP platform strain, we first designed the synthetic LP constructs and N_{20}-gRNA target sequences *in silico,* and performed a number of preliminary tests to optimize performance of the LP integration system. Previous work has shown that the N_{20}-gRNA targeting sequence is one of the most important parameters for successful CRISPR/Cas9-mediated DSB induction [39, 57]. Generally, the gRNA targeting sequence should be between 17-20 nucleotides long and contain 40-80% GC content [58]. While some additional gRNA design criteria have been reported [45, 59, 60], it is difficult to predict the functionality of a gRNA before it is evaluated *in vivo*. For this reason, we generated ten synthetic N_{20}-gRNA candidates and evaluated their fidelity towards the complementary LP target site *in vivo*. We also compared integration efficiency and gene expression profiles of 16 genomic loci that were selected from previous work [61, 62]. This was to verify that genomic sites assigned to the LP platform were i) susceptible to Cas9-mediated DSB induction, and ii) showed similar transcription levels for titrating gene expression. The results obtained from the preliminary tests informed the final configuration of the LP platform. After the platform was built into *S. cerevisiae*, we optimized the efficiency of multi-copy gene integration into each LP copy number motif, and showed that gene expression levels are proportionate to gene copy number. As proof-of-concept, we then applied the LP platform strategy to help alleviate a key pathway bottleneck towards the synthesis of benzylisoquinoline alkaloids in *S. cerevisiae*.

1.5. Benzylisoquinoline alkaloids

Benzylisoquinoline alkaloids (BIA) are a diverse class of plant secondary metabolites largely known for their important pharmaceutical properties. The most notable BIAs include the analgesics morphine and codeine, the antitussive and anti-cancer agent noscapine, the vasodilator papaverine, and the antimicrobial agents berberine and sanguinarine [63, 64]. With over ~2500 BIAs identified, only a small fraction have been clinically tested, and an even smaller fraction are commercially available [64]. Despite their therapeutic potential, many of these compounds remain understudied because they are only found in trace amounts in the source plant, which makes extraction processes too inefficient and costly for drug development [65]. Even for compounds that are currently produced at large scales, cultivation of the source plant is laborious and resource-intensive, and the supply is vulnerable to environmental degradation and climate change [10, 63, 64]. These conditions have motivated research into alternative BIA production platforms, namely microbial synthesis, in order to scale-up the production of low-yielding BIAs and secure the supply of commercially available medicines [9, 10, 64-68]. Microbial

factories also provide a more efficient platform for the discovery of new BIA molecules with novel biological activities [10, 64].

BIA biosynthesis in *S. cerevisiae*

The characterization of enzymes involved in BIA synthesis from the opium poppy (*Papaver somniferum*), as well as other BIA-producing plants, has facilitated the reconstitution of several BIA pathways in *S. cerevisiae* towards the production of (*S*)-reticuline [69], sanguinarine [66], morphine and codeine [68]. Early efforts focused on the reconstitution of partial BIA pathways as proof-of-concept, which synthesized downstream BIAs from intermediate substrates provided in the growth media. However, construction of industrially relevant strains requires that BIAs are synthesized *de novo* from simple carbon sources rather than expensive intermediates, which involves connecting BIA metabolism to primary yeast metabolism.

Entry into BIA metabolism begins with the synthesis of (*S*)-norcoclaurine, the central scaffold from which all BIAs originate [70-72]. Synthesis of (*S*)-norcoclaurine in *S. cerevisiae* requires heterologous expression of norcoclaurine synthase (NCS), which catalyzes the enantioselective Pictet-Spengler condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to produce (*S*)-norcoclaurine [71, 72]. Since dopamine and 4-HPAA are both derived from the aromatic amino acid L-tyrosine, BIA metabolism connects to yeast central metabolism via the aromatic amino acid (AAA) pathway. Unfortunately, entry into BIA metabolism in *S. cerevisiae* is very inefficient, due in part by the low catalytic activity of NCS [73, 74], as well as the low intracellular availability of substrates 4-HPAA and dopamine [64, 75]. This represents the first rate-limiting step towards the synthesis of BIAs in *S. cerevisiae* and is the major pathway bottleneck. Until we can improve the production of (*S*)-norcoclaurine, synthesis of downstream BIAs in yeast will remain proof-of-principle.

Previous efforts to improve the production of (*S*)-norcoclaurine in *S. cerevisiae* have focused on increasing substrate availability and improving NCS enzyme efficiency. *De novo* production of (*S*)-norcoclaurine in *S. cerevisiae* was first achieved by Deloache et al. (2015), who introduced a tyrosine hydroxylase and a DOPA decarboxylase into *S. cerevisiae* to generate dopamine from tyrosine via the intermediate L-DOPA. They also increased the availability of tyrosine by overexpressing a feedback-insensitive mutant of the L-tyrosine pathway enzyme *ARO4*, which more than doubled the production of dopamine. With their improved dopamine production strain, Deloache et al. evaluated three different NCS variants and identified an NCS

variant from *P. somniferum* (*Ps*NCS3) that produced the highest recorded (*S*)-norcoclaurine titer at 104.6 µg/L. Similar work done by Trenchard et al. (2015) also improved flux towards tyrosine in *S. cerevisiae* before testing different NCS variants, though they only generated ~5-10 µg/L (*S*)-norcoclaurine from *de novo* synthesis.

1.6. Application of the LP platform to improve entry into BIA metabolism

Compared to other enzymes involved in BIA metabolism, the NCS variant from *Thalictrum flavum* demonstrates low substrate affinity and a catalytic efficiency that was reportedly 100-fold lower than the average across all enzymes [73]. In order to improve (*S*)-norcoclaurine production, and by extension the production of all BIAs, it is necessary to identify a better NCS variant. It has also been shown that some NCS variants from the Papaveraceae family occur as natural fusion proteins consisting of up to four repeated domains [74]. This finding is interesting because enzymes with more repeated domains demonstrated higher catalytic efficiency than those with fewer domains [74], which suggests that the activity of single-domain NCS variants could also be improved by increasing gene copy number. Taken together, these circumstances provide an ideal opportunity to apply the LP platform to efficiently 1) screen a library of NCS variants to identify a superior enzyme variant and 2) test whether increasing the gene copy number produces higher titers of (*S*)-norcoclaurine from single domain NCS variants.

2. MATERIALS & METHODS

2.1. Strains & Media

Yeast

S. cerevisiae strains built in this study were constructed using the parent strain CEN.PK2-1D and are listed in Appendix Table A1. Yeast cultures were grown in YPD (10 g/L yeast extract, 20 g/L tryptone, 20 g/L dextrose (Thermo Fisher Scientific), YPS (10 g/L yeast extract, 20 g/L peptone 40 g/L sucrose) or synthetic complete (SC) medium made with 6.8 g/L Yeast Nitrogen Base (YNB) without amino acids, 1.92 g/L SC (-histidine), 0.76 mg/L L-histidine (Sigma Aldrich) and supplemented with 4% glucose. When appropriate, 200 µg/ml of geneticin (G418) (Sigma Aldrich) or hygromycin was added to media for selection of the pCAS plasmid.

Bacteria

Escherichia coli cultures were cultivated in Lysogeny Broth (LB) (Sigma Aldrich) supplemented with either 50 µg/ml kanamycin (Sigma Aldrich) or 100 µg/ml ampicillin (Sigma Aldrich) where applicable and grown overnight at 37°C with shaking at 200 rpm.

2.2. DNA Manipulation

Plasmids

All plasmids used in this study are listed in Appendix Table A2 and were maintained in *E. coli* DH5α. The pCAS plasmid was purchased from Addgene (Plasmid #60847). Selected constructs were cloned into plasmid pJet1.2 using the CloneJet PCR cloning kit (Thermo Fisher Scientific). The *gfp* expression cassette was amplified from pGREG503 [76]. Genes required for dopamine synthesis were amplified from the yeast integration plasmid (YIP) pWCD2249 (Genbank KR232306.1). *NCS* variants were amplified from pBOT(HIS) [10]. Plasmids were purified from *E. coli* stocks using the GeneJET plasmid mini prep kit (Thermo Fisher Scientific).

Landing Pads

Synthetic recombinogenic regions were designed using the random DNA sequence generator FaBox [77] with 50% GC content and then queried against the *S. cerevisiae* genome to ensure no sequence similarity was found with the native DNA. Four Landing Pads were synthesized as 560 bp gBlocks (Integrated DNA Technologies) (Appendix Table A3). Synthetic gRNA targeting sequences (N_{20}) were generated using >40% GC cut-off and queried against the *S. cerevisiae* genome. Ten gRNA targeting sequences (Table 1) were selected for preliminary testing and synthesized as oligonucleotides. Complementary target sites (5'-N_{20}CGG-3') were also synthesized as oligonucleotides.

Primers used in this study are listed in Appendix Tables A4-A14. DNA constructs were amplified by PCR using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific). Linearized DNA fragments were isolated using 0.8% (w/v) agarose gel electrophoresis and purified using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific).

Linearization of pCAS plasmid

The pCAS vector backbone and the pCAS gRNA expression cassette were amplified with overlapping homology regions using the primers listed in Appendix Table A4. To program novel targeting sequences, the expression cassette was amplified in two universal parts: Left (*tRNA*^{Tyr}*-3'HDV*) and Right (*scaffold-*T_{SNR52}) using primers that insert overlapping sequence containing the novel N_{20}Target sequence. The two parts were then reassembled in a second PCR to generate the full-length gRNA cassette. The (–gRNA) expression cassette was generated by fusing the left and right gRNA fragments together using overlapping primers. Primers used to program synthetic gRNAs are listed in Appendix Table A5. Primers used to program the gRNAs targeting the *S. cerevisiae* genome are listed in Appendix Table A6.

Landing Pad Constructs

To generate the LP.A and LP.Z recombinogenic regions, the LP gBlock sequences were amplified in two equal parts using primers listed in Appendix Table A7. To build the LP1 constructs harbouring target sites T1-T10, the LP1.A and LP1.Z regions were amplified using overlapping internal primers that contain the 10 x synthetic 5'-N_{20}CGG-3' gRNA target sites (Appendix Table A8). To evaluate the synthetic gRNAs, full-length LP1.T_X constructs were amplified with primers that attach ~50 bp flanking homology to the FgF20 locus. To evaluate CRISPR/Cas9-mediated gene integration at selected genomic loci, the LP1.T3-Site$_X$ constructs were amplified with primers (Appendix Table A9) to attach homologous arms complementary to each genomic integration site (Table 2). To integrate LP constructs into the genome, we amplified the ‘up’ and ‘down’ regions of each FgF locus [61] from CEN.PK2-1D genomic DNA and fused them to LP1.T3 by OE-PCR, generating 15 LP1.T3-Site$_X$ constructs. To integrate the LP1.T3 construct into USERXII-1, LP1.T3 was amplified with primers that attach ~60 bp flanking homology to the USERXII-1 locus. Primers to amplify up and down genomic regions are listed in Appendix Table A10.

LP donor DNA

LP donor DNA constructs were generated by PCR using primers listed in Appendix Table A11. For integration efficiency tests in CEN.LP the *gfp* expression cassette (*P*$_{TDH3}$*-gfp*^{S65T}*-T*$_{CYC1}$) was amplified from pGREG503 plasmid (GC975) using primer pairs to attach 60 bp flanking homology to LP1, LP2, LP3 and LP4. To generate *gfp* donors with full-length homology to the LP recombinogenic regions, LP$_X$.A and LP$_X$.Z regions were amplified from the CEN.LP genome and then fused to the LP$_X$.*gfp* donor DNA with 60 bp overlap by PCR. The full-length LP$_X$.*gfp* donor

constructs were cloned into the pJET1.2 vector and used as template for future amplifications. NCS variants (P_{TEF1}-NCS_X-T_{PGI1}) were amplified from the corresponding pBOT vector using primer pairs to attach ~60 bp flanking homology to LP1, LP2, LP3 and LP4.

Dopamine production cassette

The dopamine expression cassette (P_{TDH3}-*CYP76AD1*$^{W13L\_F309L}$-T_{TDH1}-P_{CCW12}-*DODC*-T_{ADH1}-P_{PGK1}-*ARO4*^{FBR}-T_{PGK1}) was amplified from pWCD2249 [75] in two overlapping fragments with ~60bp flanking homology to the ARO4 locus (Appendix Table A12).

2.3. Yeast Transformation

S. cerevisiae was transformed using a method modified from the Gietz PEG/LiAc protocol [78]. Strains were grown overnight in YPD medium at 30°C with shaking and then diluted to an optical density at 600 (OD_{600}) of 0.175 into fresh 2xYPD and propagated at 30°C with shaking until the culture reached OD_{600} 0.6-0.8. Cells were harvested and washed once with water followed by a second wash in 100 mM lithium acetate. The cell pellet was then suspended again in 100 mM lithium acetate (20 µl per transformation) before adding the transformation mix, which included per reaction: 100 µL 50% (w/v) PEG3350, 5.6 µL 3M lithium acetate, 4.4 µL boiled salmon sperm DNA (10 mg/ml), and 20 µl DNA + water. Transformation conditions included a 30 min incubation period at 30°C, followed by a 30 min heat shock at 42°C. Cells were recovered overnight in 500 µL YPD before plating on selective medium containing G418 or hygromycin to maintain the pCAS plasmid. For transformations requiring an extended outgrowth in selective medium, recovered cells were diluted 1:100 into fresh YPD+G418 medium, and grown for 48h at 30°C with shaking before plating on YPD+G418 medium.

2.4. CRISPR/Cas9-mediated genome integration

All genomic integrations were performed using the CRISPR/Cas9 delivery system developed by Ryan et al. (2014), which delivers *cas9* and gRNA on a single 2µ plasmid (pCAS). In pCAS the *cas9* gene from *Streptococcus pyogenes* fused to a yeast nuclear localization signal and expressed using a medium strength *RNR2* promoter (P_{RNR2}-*cas9_NLS*-T_{CYC1}). The tyrosine RNA II polymerase promoter and *SNR52* terminator are used for expressing the gRNA, which is fused to the 3' end of a self-cleaving *HDV* ribozyme (*tRNA*^{Tyr}-*3'HDV_*N_{20}*Target_scaffold*-T_{SNR52}). The pCAS plasmid encodes the *kanMX* selection marker which allows growth in media containing G418.

The pCAS plasmid is assembled *in vivo* by homologous recombination between two overlapping linear fragments: pCAS vector backbone harbouring *cas9* (250 ng) and the gRNA expression cassette (400-800 ng). Homology arms attached to any donor DNA construct were designed so that the gRNA target site would be replaced upon integration. Between 1-4 µg of donor DNA was used for CRISPR/Cas9-mediated integration.

2.5. Strain Construction

Construction of LP1.T_X strains

To test the targeting specificity and integration efficiency associated with the synthetic gRNAs, we first had to introduce the LP1 construct harbouring the complementary target sites (LP1.T_X) into the *S. cerevisiae* genome. The FgF20-LP1.T_X constructs (x10) were integrated separately into the CEN.PK2-1D strain at the FgF20 locus by co-transforming *cas9* with either FgF20-gRNA1 or FgF20-gRNA2 (Appendix Table A6). Upon transformation, a Cas9-mediated DSB was generated within the FgF20 “deletion region” (Table 2). FgF20-LP1.T_X constructs with flanking up/down homology to the FgF20 locus repaired the DSB through HDR, which replaced the deletion region upon integration. LP1.T_x integration were confirmed by colony PCR using primers that amplify the 5' and 3' genomic DNA flanking LP1 (Appendix Table A13). Positive integrant colonies were sequence verified and cultured for two days in YPD to remove the pCAS plasmid. The LP1.T_X strains were then saved in 15% glycerol and stored at -80°C. The strains were streaked onto YPD to obtain single colonies for subsequent testing of the synthetic gRNAs.

Construction of LP1.T3. $Site_X$ strains

To compare integration efficiency and gene expression levels at various genomic loci in *S. cerevisiae*, we first introduced the LP1.T3 construct at 16 genomic loci we selected based on work by Flagfeldt et al. (2009) and Mikkelsen et al. (2012). Sixteen LP1.T3 donors were generated by fusing ~500 bp homology arms for each genomic locus. LP1.T3 constructs were introduced into the CEN.PK2-1D strain at the corresponding locus by co-transforming *cas9* with site-specific gRNA(s) listed in Appendix Table A6. Upon transformation, the LP1.T3 construct replaces the “deletion region” at each genomic locus (Table 2). LP1.T3 integration at each site was confirmed by colony PCR using primers that amplify the 5' and 3' genomic regions flanking LP1 integrant (Appendix Table A13). Positive integrant colonies were sequence verified and cultured for two days in YPD to remove the pCAS plasmid. The LP1.T3.$Site_X$ strains were saved in 15% glycerol

solution and stored at -80°C. The strains were streaked onto YPD to obtain single colonies for subsequent testing of synthetic gRNAs.

Construction of CEN.LP strain

The LP platform was built into *S. cereivisiae* strain CEN.PK2-1D by Michael Pyne. Full-length $LP_X.T_X$ constructs were generated by PCR using primers to introduce target sites to the $LP_X.A$ and $LP_X.Z$ recombinogenic regions. The LP_X donor constructs were generated by PCR using primers to attach ~60 bp homology to the assigned genomic loci (Appendix Table A7). Donor constructs include: LP1.T8 with homology to site FgF20; (x2) LP2.T10 donors with homology to sites FgF18 and FgF24; (x3) LP3.T7 donors with homology to sites UserXII-1 and FgF7, FgF19; (x4) LP4.T9 donors with homology to sites FgF12, FgF16, FgF21, FgF22. The LP platform was constructed by integrating one or two LP donor constructs in successive co-transformations with the site-specific targeting gRNA(s) (Appendix Table A6) and linearized pCAS vector backbone. LP_X integrations were confirmed by colony PCR using site-specific primers that amplify the 5' and 3' genomic regions flanking LP_X integrants (Appendix Table A13). The CEN.LP strain was sequence verified and cultured for 2 days in YPD to remove the pCAS plasmid. The strain was saved in 15% glycerol solution and stored at -80°C.

Construction of CEN.LP.D

In order to test the NCS variants, a dopamine production cassette was inserted into the CEN.LP strain to provide substrate for *de novo* synthesis of (*S*)-norcoclaurine (work done by Michael Pyne). The three-gene dopamine production cassette (P_{TDH3}-*CYP76AD1*$^{W13L/F309L}$-T_{TDH1};T_{TDH1};P_{CCW12}-*DODC*-T_{ADH1};P_{PGK1}-*ARO4*^{FBR}-T_{PGK1}) was developed and provided by Deloache et al. (2015). For genomic integration into the CEN.LP genome, the cassette was amplified from the pWCD2249 plasmid (Accession number KR232306.1) in two overlapping fragments along with flanking homology to the *ARO4* locus in *S. cerevisiae* (Appendix Table A12). The four donor fragments were co-transformed with pCAS vector backbone and a gRNA targeting *ARO4* (Appendix Table A6). Overlapping fragments assembled *in vivo* and replaced the native *ARO4* sequence upon integration. Correct assembly and integration was confirmed by colony PCR using primers that amplify the 5' and 3' genomic regions adjacent to the integration (Appendix Table A13). The CEN.LP.D strain was sequence verified and cultured for two days in YPD to remove the pCAS plasmid. The strain was saved in 15% glycerol solution and stored at -80°C.

2.6. Colony PCR

Individual colonies were suspended in 15 µl sterile water. Five µl was transferred into 30 µl 20 mM NaOH and microwaved for two min. One µl was used as template for colony PCR using the Phire™ Plant Direct PCR Master Mix (Thermo Fisher Scientific). Primers used for colony PCR are site-specific and anneal to genomic DNA adjacent to the integration site (Appendix Table A13 & Table A14). Integration of donor DNA was verified by the size difference between amplicons generated for positive and negative integration events.

2.7. Preliminary tests

Synthetic gRNA efficacy

Targeting specificity of gRNAs was determined by measuring the lethality of Cas9-induced DSBs generated by cells expressing *cas9* and a targeting gRNA without repair template. We describe this relationship as the percent kill (% kill), where higher % kill is associated with lower CFU values of transformants plated on selective medium. Percent kill was measured by co-transforming linear $gRNA_X$ constructs with linearized pCAS vector backbone harbouring *cas9* into LP1.T_X strains and comparing the CFU against the same strains co-transformed with pCAS backbone and a control (–)gRNA cassette. Percent kill was calculated for each $gRNA_X$ using the formula:

$$\left(1 - \left(CFU\frac{(+)gRNA}{(-)gRNA}\right)\right) \times 100$$

We also calculated the percent kill for each $gRNA_X$ construct co-expressed with *cas9* in the CEN.PK2-1D background strain lacking an LP1.T_X target site.

Integration efficiency was measured by co-transforming linearized $gRNA_X$ and pCAS vector backbone with the LP1.*gfp* donor with full-length LP homology into the corresponding LP1.T_X strains. Integration efficiency of the *gfp* donor was calculated for each $gRNA_X$ by genotyping 12 colonies per transformation plate.

Targeted integration efficiency into selected genomic loci

Integration efficiency into LP1.T3 at each genomic locus was measured by co-transforming linearized gRNA3 and pCAS vector backbone with the LP1.*gfp* donor with full-length LP homology

into the corresponding LP1.T3.Site$_X$ strain. After selecting for assembly of the pCAS plasmid, integration efficiency of the donor DNA into each genomic loci was calculated by genotyping 12 colonies per transformation plate using the primers listed in Appendix Table A13.

2.8. Integration into the LP platform

Integration efficiency into each LP motif was measured by co-transforming linearized LP$_X$.gRNA (400-800 ng) and pCAS (250 ng) vector backbone with LP$_X$.*gfp* donor constructs (1-4 µg): LP1.*gfp* donor was co-transformed with gRNA8; LP2.*gfp* donor was co-transformed with gRNA10; LP3.*gfp* donor was co-transformed with gRNA7; and LP4.*gfp* donor was co-transformed with gRNA9. For testing the effect of homology length on transformation efficiency, we also co-transformed LP$_X$.*gfp* donor constructs harbouring full-length homology to each LP$_X$ construct. Integration efficiency was calculated by genotyping a minimum of 12 colonies per transformation using the LP primer sets listed in Appendix Table A14 to check for integration at each site within the LP motif. Strains harbouring *gfp* in each LP motif were sequence verified, saved in 15% glycerol solution and stored at -80°C.

2.9. GFP expression analysis

Fluorescence levels were measured for strains expressing GFP at different genomic loci and copy numbers by cultivating the LP1.*gfp*.Site$_X$ or LP$_X$.*gfp* strains overnight in SC with 2% glucose (w/v). Overnight cultures were then diluted 10× into fresh medium and incubated for an additional 4 h to obtain log phase cells. Fluorescence was measured from cell suspension using a microplate reader and normalized against OD_{600} for three biological replicates. Fluorescence was detected by the TECAN M200 plate reader using an excitation wavelength of 485 nm and an emission wavelength of 525 nm. Gain was adjusted for each sample. The background strain lacking GFP was used to correct for autofluorescence generated by cells and medium.

2.10. NCS enzyme candidates

The NCS enzyme library used in this study was previously curated by James Scriven by querying the PhytoMetaSyn transcriptome database (http://www.phytometasyn.net). *NCS* nucleotide sequences (Appendix Table A15) were codon-optimized for expression in *S. cerevisiae*. NCS open reading frames (ORFs) were synthesized by Gen9 and cloned into the pBOT(HIS) expression vector between the *TEF1* promoter and the GFP fusion tag. The GFP tag was subsequently removed by restriction enzyme digestion with *KasI* (NEB). NCS amino acid

sequences (Appendix Table A16) were aligned by MUSCLE (Appendix Figure A1) and phylogenetic trees were generated using the neighbour-joining method in MEGA7.0 [79] with a bootstrap value of 1000. Sequence identities between NCS variants (Appendix Table A17) were scored using BLASTp [80] multiple sequence alignment software.

NCS copy number integration into LP platform

NCS variants were integrated into each LP motif of the CEN.LP.D strain using the LP platform integration strategy outlined in section 2.8. Transformants were screened by colony PCR using primer sets listed in Appendix Table A14 to check for *NCS* integration at each site within the targeted LP motif. Forty LP.NCS variant and copy number strains were sequence verified and saved in 15% glycerol solution and stored at -80°C. Strains were streaked out on YPD to obtain single colonies for activity assays.

(*S*)-Norcoclaurine production assay

Colonies were picked in triplicate and inoculated into 200 ml of YPS in 96-well two ml deep well plates. Cultures were incubated at 30°C with shaking at 200 rpm for 16 h. Overnight cultures were diluted to an OD_{600} of 0.3 into 1.8 ml of YPS. A 180 µl aliquot was transferred to a 96-well microtiter plate and incubated at 30°C in the Sunrise® absorbance microplate reader (Tecan) with shaking at 200 rpm. Cell density was measured using an OD_{600} at 20 min intervals for 48 h to monitor growth rate (Appendix Figure A2). The remaining 900 µl of culture was grown in a 96-well two ml deep well plate and incubated at 30°C with shaking at 200 rpm. After 96 h, the cultures were centrifuged at 4,000 rpm for 5 min. Culture supernatant was collected and suspended 1:1 in 60% acetonitrile + 0.2% formic acid (final concentration 30% acetonitrile + 0.1% formic acid) for LC/MS analysis.

LC/MS detection and quantification of dopamine and (*S*)-norcoclaurine.

Production of dopamine and (*S*)-norcoclaurine was analyzed using the 1290 Infinity II LC system, (Agilent Technologies) with a Zorbax Eclipse Plus C18 50 × 4.6 mm column (Agilent Technologies). Solvent A (100% water, 0.1% formic acid) and solvent B (100% acetonitrile, 0.1% formic acid) were used in a gradient elution to separate metabolites. Samples were separated using a linear gradient: 0-5 min 98% A/ 2% B, 5-7 min 90% A/ 10% B, 7-7.1 min 15% A/ 85% B at a flow rate of 0.3 ml/min followed by a 3 min equilibration at 100% A at a flow rate of 0.4 ml/min [75]. Following liquid chromatography (LC) separation, eluent was then injected into a 6560 Ion

Mobility Q-TOF LC/MS, (Agilent Technologies). The system was operated in positive electrospray (ESI+) mode using the following parameters: capillary voltage 4000V; fragmentor voltage 400V; source temperature 325°C; nebulizer pressure 55 psig; gas flow 10 L/min. Dopamine and (*S*)-norcoclaurine standards (Toronto Research Chemicals Inc.) were used to determine retention times and to generate calibration curves. For identification and quantification of dopamine (m/z 154.086 $[M+H]^+$; R_t 1.2 min) and (*S*)-norcoclaurine (m/z 272.121 $[M+H]^+$; R_t 4.3 min), the extracted ion counts were normalized against an eight-point calibration curve ranging from 0.0078–2 mM (dopamine) and 0.078-10 µM ((*S*)-norcoclaurine) in two-fold increments. The LC/MS were analyzed using MassHunter quantitative analysis software (Agilent Technologies).

3. RESULTS

3.1. Landing Pad platform design and optimization

The LP platform was designed to accommodate marker-less, multi-copy gene integration into the *S. cerevisiae* genome in a single reaction. The platform uses a series of synthetic DNA blocks called 'landing pads' (LP_X) (Figure 1A) inserted into the *S. cerevisiae* genome at different loci, to serve as modular anchors for CRISPR-mediated gene integration. The LP platform was built by inserting four distinct LP_X constructs in one, two, three, or four copies into the *S. cerevisiae* genome at various loci (Figure 1B). LP_X motifs were then used as modular target sites for CRISPR-Cas9-mediated integration of genes in up to four copies in a single transformation.

Landing pads consist of three components to accommodate the CRISPR/Cas9 genome editing system: i) a 20-nt gRNA target sequence (N_{20}), ii) an 5'-CGG-3' PAM, and iii) two unique recombinogenic regions (LP.A/LP.Z) flanking the target site (Figure 1A). The process for CRISPR-mediated gene integration into the LP platform is illustrated in Figure 2, and begins by co-transforming a plasmid encoding *cas9* and the appropriate $gRNA_X$ along with a gene cassette repair donor possessing LP_X homology. We selected the CRISPR/Cas9 system designed by Ryan et al. (2014) to deliver *cas9* and gRNA from a single 2µ plasmid (pCAS). The gene cassette flanked by LP.A and LP.Z homology acts as template for HDR of the DSB through recombination with the LP_X recombinogenic regions. Upon integration into an LP_X, the gene cassette replaces the Cas9 target site, resulting in permanent and stable gene integration into the chromosome.

To build the LP platform, we first designed the landing pad components *in silico* using an online sequence generator to create the synthetic landing pad recombinogenic regions and gRNA

target sequences. We randomly generated four blocks of DNA sequence (560 bp) with ~50% GC to comprise the four unique LPs. The recombinogenic blocks, designated LP1, LP2, LP3 and LP4 were synthesized as gBlocks and then amplified in two halves to generate the LP.A and LP.Z recombinogenic regions (Appendix Table A3). We also generated a total of ten 20-nt DNAs with >40% GC to serve as gRNA target sequences immediately upstream of a PAM (5'-N_{20}CGG-3'). Together, the target sequence and PAM make up the LP target sites, which were synthesized as custom oligos and used to assemble the full 583 bp LP_X constructs.

We performed a number of preliminary tests to optimize the operation and efficiency of our devised LP system. First, we evaluated all 10 synthetic gRNA target sequences by measuring Cas9 endonuclease activity and targeted integration efficiency. We then compared integration efficiency and gene expression levels at different genomic loci and selected those sites that were easily modified and showed similar levels of transcription. Results gleaned from these tests informed the construction of our LP platform.

3.2. Characterization of synthetic gRNA candidates

To select four functional gRNA/target pairs to use in the LP platform, we generated 20-nt strings of DNA using a random sequence generator (http://www.faculty.ucr.edu/~mmaduro/random.htm) and selected those with >40% GC content. After including the PAM, the 5'-N_{20}CGG-3' target sites were queried against the *Saccharomyces* Genome Database (SGD) to eliminate those with potential off-target homology. A total of ten gRNA targeting sequences matching our criteria were selected for further evaluation (Table 1). To test the functionality of each synthetic gRNA *in vivo*, we first integrated LP1 constructs harbouring each of the ten gRNA target sites (LP1.T_X) into *S. cerevisiae* CEN.PK2-1D at the FgF20 locus, yielding a total of ten strains (Figure 3A & 3B). The corresponding $gRNA_X$ expression cassettes containing the selected targeting sequences were then co-transformed with *cas9* into the respective LP1.T_X strain to evaluate site-specific Cas9 nuclease activity, as well as the associated efficiency of targeted integration of donor DNA (Figure 3C).

Table 1. Synthetic gRNAx candidates

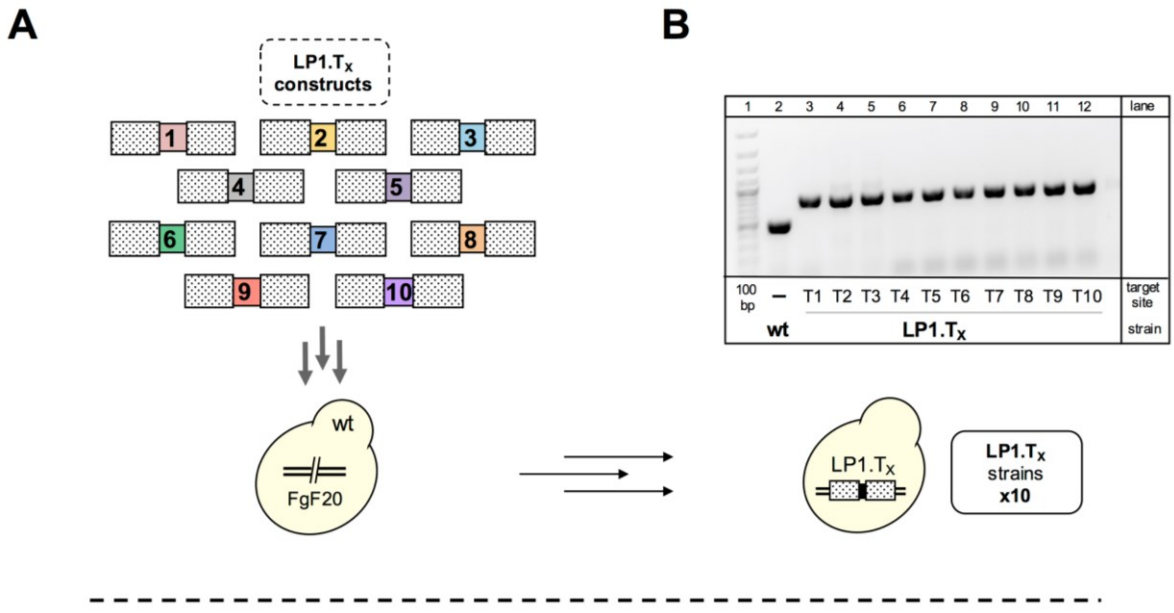
gRNA	Targeting sequence	%GC	LP target site
gRNA1	CTAGATAAGACGTGGCAGAT	45	CTAGATAAGACGTGGCAGATCGG
gRNA2	CCACCGCTCAGTAGCGGCTT	65	CCACCGCTCAGTAGCGGCTTCGG
gRNA3	GCCAGTCAGAACACTAGAGG	55	GCCAGTCAGAACACTAGAGGCGG
gRNA4	ACGCCGTTTTTCACATCTGT	45	ACGCCGTTTTTCACATCTGTCGG
gRNA5	GATACAACCCATCCGCGCTA	55	GATACAACCCATCCGCGCTACGG
gRNA6	ACGTGTCATACGAGGTATAG	45	ACGTGTCATACGAGGTATAGCGG
gRNA7	ATTGTACCCCAGCGGCGGCG	70	ATTGTACCCCAGCGGCGGCGCGG
gRNA8	TAAGTCCGCGGATAACCATT	45	TAAGTCCGCGGATAACCATTCGG
gRNA9	GTAGCCCAACAGGAGCACAT	55	GTAGCCCAACAGGAGCACATCGG
gRNA10	TCCCCAATCGTGGAGTGAAG	55	TCCCCAATCGTGGAGTGAAGCGG

gRNA$_X$ targeting specificity

To test the targeting specificity of the gRNAs, we exploited the lethal consequence of chromosomal DSBs created when Cas9 endonuclease is guided to a genomic target site by a functional gRNA. Without a repair template, successful delivery of Cas9 to the chromosomal target site results in cell death [39, 81] and is observed as a low CFU when compared to a transformation control. High % kill indicates that the gRNA successfully guides Cas9 to the genome, and induces a lethal DSB in the chromosome. Percent kill was measured by co-transforming linear gRNA$_X$ cassettes with a linearized pCAS plasmid backbone into LP1.T$_X$ strains and comparing the CFU against the same strains co-transformed with the pCAS backbone and a control (–)gRNA cassette (Figure 3C). Following a 16 h outgrowth recovery period, transformants were plated on medium selecting for assembly of the pCAS plasmid via gap repair between the gRNA cassette and the pCAS backbone. Percent kill was calculated for each gRNA$_X$ using the formula outlined in section 2.7 and ranged from 77-100% (Figure 3D). An average kill efficiency >90% was observed for all gRNA$_X$ constructs targeting the LP1.T$_X$ site. Both gRNA10 and gRNA8 produced nearly 100% kill efficiency, closely followed by gRNA5, gRNA7, gRNA9, and gRNA1 (98% kill).

To rule out possible off-target effects and ensure the % kill values reported above represent correct targeting between the synthetic gRNA$_X$ and LP1.T$_X$ landing pad, we also calculated the % kill of gRNA$_X$ and *cas9* co-expressed in the wild-type strain lacking LP1.T$_X$ target site (Figure 3C). We suspected that if gRNA$_X$ showed off-target activity, % kill would be similar between the wild-type and LP1.T$_X$ strains. The results show that % kill observed for gRNAs expressed in the parent strain were significantly less than those attained in the LP1.T$_X$ strains, suggesting highly specific targeting of gRNAs to the correct LP1.T$_X$ sites. However, potential off-target activity was observed for gRNA4 (12.2 ± 14.1%), gRNA5 (17.0 ± 25.0%) and gRNA6 (15.9 ± 13.9%) in the wild-type

strain. Still, the relatively low % kill and broad variation between replicates suggests these gRNAs are only partially complementary to off-target sites within the wild-type genome. The remaining gRNAs showed % kill values of <5% in the parent strain compared to >91% kill in LP1.T_X strains, indicating that these gRNAs show high specificity to their intended LP1.T_X target sites and are thus more suitable for use in our envisioned LP platform.



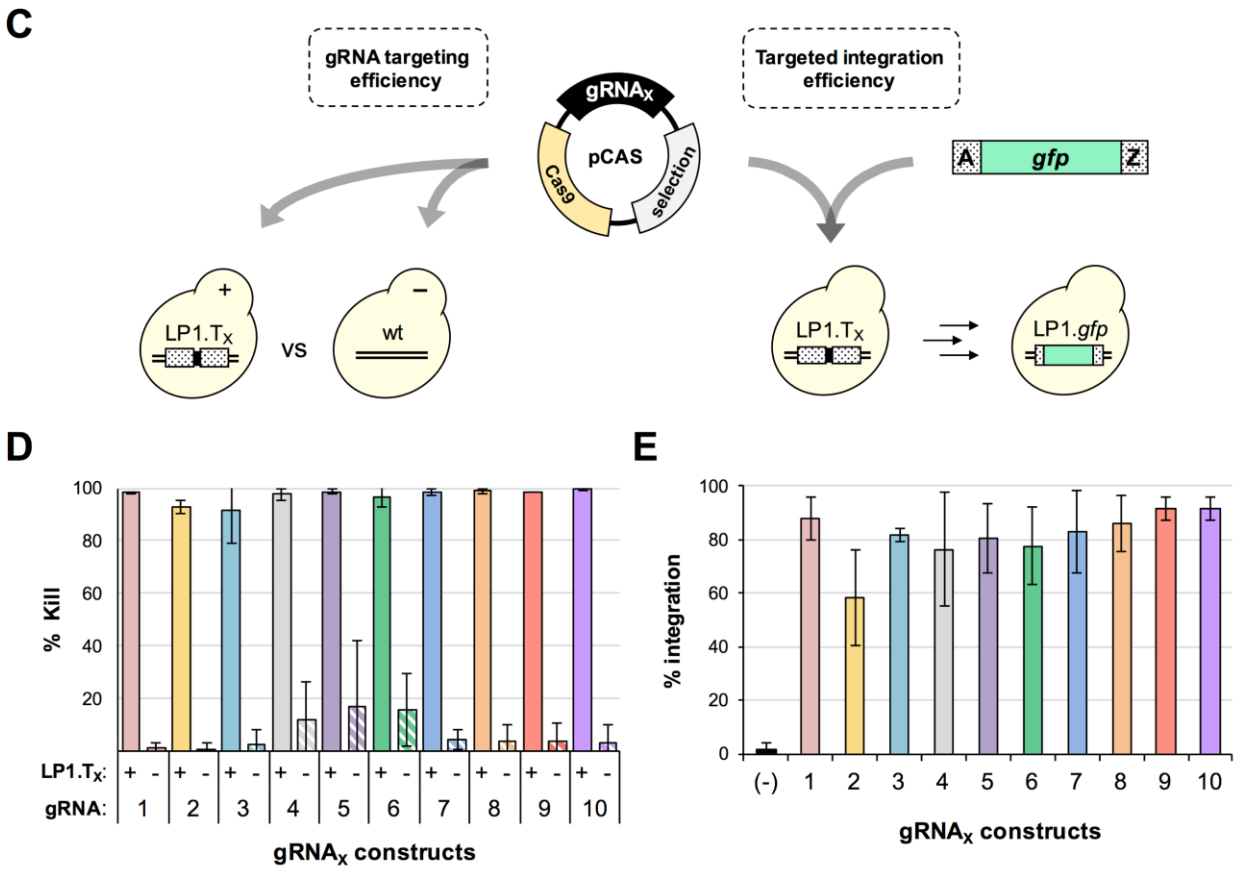


Figure 3. Targeting specificity of ten synthetic gRNAx candidates

(A) Schematic overview of LP1.T$_X$ integration into *S. cerevisiae*. Ten LP1.T$_X$ strains were generated to evaluate performance of the synthetic gRNAs. Targeting sequences complementary to the 10 synthetic gRNA candidates were assembled into LP1.T$_X$ constructs and integrated separately into locus FgF20. wt: parent strain CEN.PK2-1D. (**B**) Genotyping of LP1.T$_X$ integration into FgF20. Lane 1: 100 bp ladder, lane 2: parent strain CEN.PK2-1D (**–**) control, lanes 3 – 12: LP1.T$_X$ integrations in the following order T1, T2, T3, T4, T5, T6, T7, T8, T9, T10. **(C)** Schematic overview of synthetic gRNA efficacy tests. i) Synthetic gRNA targeting specificity tests were performed by co-transforming pCAS vector backbone with each gRNA$_X$ into the complementary LP1.T$_X$ target strain to evaluate % kill. The % kill was also calculated for gRNA$_X$ expression in the parent strain lacking a LP1.T$_X$ construct. ii) Efficiency of gRNA$_X$ targeted integration was evaluated by co-transforming the pCAS backbone and gRNA$_X$ with a linear *gfp* cassette flanked by LP1 homology into each LP1.T$_X$ strain. **(D)** Synthetic gRNA target specificity and efficiency tests. Comparison of the % kill for each synthetic gRNA co-expressed with Cas9 in wt (**–**) versus the LP1.T$_X$ strains (+). Percent kill was evaluated by comparing the number of transformants grown on medium selecting for gRNA expression against a (–)gRNA transformation control. Error bars represent the mean ± s.d. of three separate assays. **(E)** Synthetic gRNA targeted integration efficiency**.** Integration efficiency calculated for a 1.5 kb *gfp* donor cassette with LP1 homology arms into LP1.T$_X$ strains using the appropriate gRNA$_X$ in combination with pCAS. Integration efficiency was determined by genotyping selected colonies. Error bars represent the mean ± s.d. of three separate assays. wt: parent strain CEN.PK2-1D.

Targeted integration efficiency associated with synthetic gRNA$_X$ candidates

Considering that targeted gene integration is a critical component of the landing pad system, we set out to test whether the selected gRNAs facilitated efficient gene integration. To quantify the efficiency of integration into the ten LP1.T$_X$ strains, we co-transformed the linear gRNA$_X$ cassettes and pCAS backbone along with a 1.5 kb *gfp* expression cassette (P_{TDH3}-*gfp*-T_{CYC1}) flanked by 280 bp homology arms complementary to the LP1 recombinogenic regions (Figure 3C). In this experiment, *gfp* donor DNA acted as the repair template for the CRISPR/Cas9-induced DNA break at LP1.T$_X$ via homologous recombination with LP1 [28, 82]. After selecting for assembly of the pCAS plasmid, integration efficiency of the *gfp* donor was calculated for each gRNA$_X$ by genotyping 12 colonies per strain. Integration efficiencies observed for *gfp* at each LP1.T$_X$ site ranged from 58 – 92% (Figure 3E). The highest integration efficiencies were observed for gRNA9 (91.7 ± 4.2%) and gRNA10 (91.7 ± 4.2%), whereas *gfp* integration using gRNA2 was only 58.3 ± 18.1%. To eliminate the possibility that integration occurred in the absence of a Cas9-induced DSB, we transformed the *gfp* repair template without gRNA and plated on medium selecting for the pCAS (–gRNA) plasmid. Out of 72 colonies screened, only a single colony contained the *gfp* cassette integrated into LP1. Considering that native YHR-mediated integration occurred at very low frequency, the integration efficiency reported for each gRNA$_X$ is a direct result of the DSB induced by the CRISPR/Cas9 system.

Based on the integration efficiency data, we selected gRNA7, gRNA8, gRNA9 and gRNA10 to form the basis of the LP platform strain. All four gRNAs demonstrated near perfect targeting efficiency with high specificity to the LP1.T$_X$ target site, and offered among the highest targeted integration efficiencies (83-92%) of the ten synthetic gRNAs tested.

3.3. Characterization of *S. cerevisiae* genomic loci for high level gene expression

Selection of genomic loci for efficient integration and stable expression of genetic constructs is another critical component of the LP platform. To reveal a correlation between gene expression and gene dosage, it is important that the selected loci exhibit similar levels of transcription. In this context, we selected 15 loci (Table 2) based on previous work done by Flagfeldt et al. (2009) who measured gene expression levels at 20 different genomic loci using a LacZ reporter enzyme [61]. Since we were interested in the sites showing high gene expression, we excluded five sites showing low levels of expression in their study. Twelve of the sites chosen for this work contain solo long terminal repeats (LTRs), which are removed upon LP integration,

while the remaining three consist of an inactive *URA3* locus, the *PDC6* locus, and the intergenic region between housekeeping genes *SPB1* and *PBN1* [61]. We also selected the USERXII-1 site based on work done by Mikkelsen et al. (2012) to evaluate alongside the Flagfeldt (FgF) sites due its proven reliability within our laboratory. The USERXII-1 site is situated within an intergenic region on chromosome XII, and showed high gene expression levels when evaluated with a LacZ reporter enzyme [62]. We first re-evaluated the genomic loci from previous work to test the efficiency of CRISPR/Cas9-mediated integration and quantify gene expression at each site. To perform these tests, we integrated the LP1 construct harbouring target site 3 (LP1.T3) into all 16 selected loci (Figure 4A), yielding 16 LP1.T3.Site$_X$ strains for evaluation (Appendix Table A1).

Table 2. Genomic loci (Site$_X$) candidates for characterization

SiteX	Description	Chromosome	Coordinates	Deletion Region	Ref
FgF1	LTR	I	21839-23155	22221-22575	[61]
FgF5	LTR	III	290999-292320	291360-291805	[61]
FgF7	LTR	V	248671-249726	249071-249383	[61]
FgF8	LTR	XII	319986-321371	320344-320770	[61]
FgF11	LTR	XI	528976-531074	529266-530388	[61]
FgF12	LTR	IX	425603-427250	426198-426544	[61]
FgF14	LTR	XIII	480888-482601	481480-481859	[61]
FgF16	LTR	XIV	726641-728332	727268-727659	[61]
FgF18	LTR	XV	664036-665684	664793-665233	[61]
FgF19	LTR	XV	968527-970120	969238-969467	[61]
FgF20	LTR	XVI	775916-777558	776567-776868	[61]
FgF21	LTR	XVI	880684-882562	881333-881964	[61]
FgF22	URA3	V	115710-117433	116214-116932	[61]
FgF23	Intergenic region	III	33108-34547	34005-34072	[61]
FgF24	PDC6	XII	651035-653434	651504-652891	[61]
UserXII-1	Intergenic region	XII	4105-4676	4165-4616	[62]

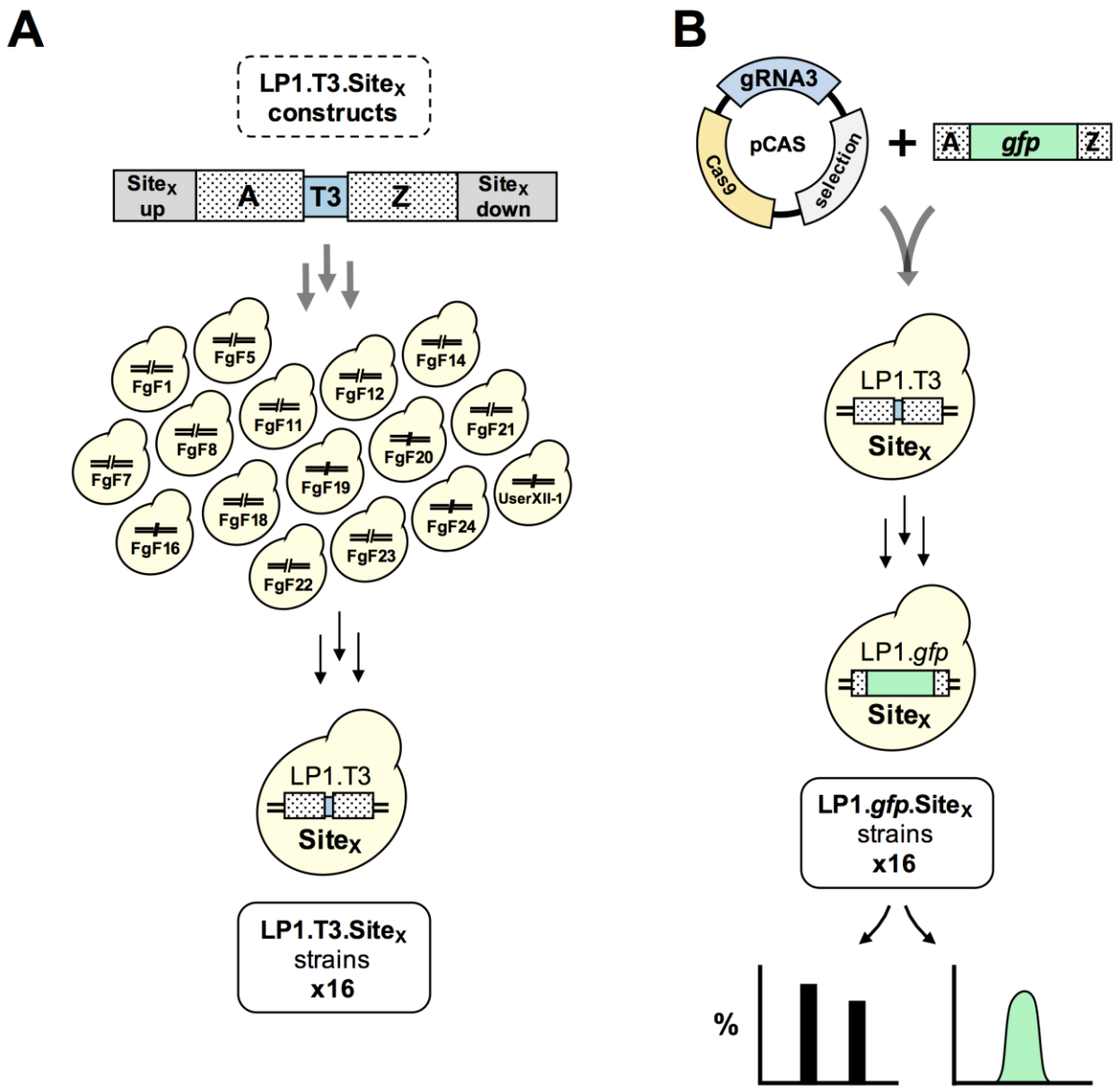
Targeted integration efficiency into candidate genomic loci (Site$_X$)

Integration efficiency at each genomic locus containing the LP1.T3 landing pad was measured by co-transforming the LP1-*gfp* donor cassette with linearized gRNA3 and pCAS plasmid into each LP1.T3.Site$_X$ strain (Figure 4B). After selecting for assembly of the pCAS plasmid, integration efficiency of the *gfp* donor was calculated for each site by genotyping 12 colonies per strain. Efficiency of targeted *gfp* integration into LP1.T3 varied from 64 – 92% between the 16 genomic loci (Figure 4C). An average integration efficiency of >80% was observed for 10 of the 16 loci (FgF7, 8, 12, 14, 16, 18, 20, 21, 24, UserXII-1), whereas sites FgF1, 5, 11,

19, 22 and 23 showed average integration efficiencies of >60%. For at least one experimental replicate, targeted integration reached 100% efficiency for sites FgF7, FgF16, FgF18 and FgF24. Since the CRISPR machinery and donor DNA remained constant between strains, it is possible that the rate of integration was influenced by the structural dynamics of chromatin at each locus, which would affect the ability of the Cas9-gRNA complex to target and bind DNA [50].

Analysis of gene expression at genomic loci

Gene expression was assayed from each of the 16 candidate sites by measuring fluorescence levels resulting from integration of *gfp* at each locus. The *gfp* cassette expresses the GFP^{S65T} mutant protein [83] controlled by regulatory elements P_{TDH3} and T_{CYC1} to achieve high levels of expression [17, 20]. Likewise, the GFP^{S65T} mutant protein (herein referred to as GFP) has improved fluorescence qualities that enables real-time quantification of gene expression in *S. cerevisiae* [84]. Fluorescence values were then normalized against OD_{600} and corrected for autofluorescence detected from the background strain and medium. The results showed a nearly seven-fold difference in fluorescence between strains expressing *gfp* from the 16 genomic sites (Figure 4D). Integration of *gfp* at sites FgF1 and FgF11 resulted in the highest levels of fluorescence, generating roughly 1.7-fold higher fluorescence than the average of all sites, whereas FgF5 showed the lowest level of fluorescence, approximately 4-fold lower than the average. These sites were not selected for use in the LP platform due to the significant deviation in gene expression compared to all other sites. The 13 remaining genomic loci exhibited similar expression profiles, showing fluorescence values within ± 1.2x from the mean fluorescence and were considered good candidates for the landing pad platform.


A
LP1.T3.Site$_X$ constructs
Site$_X$ up
A
T3
Z
Site$_X$ down
FgF1
FgF5
FgF14
FgF12
FgF11
FgF21
FgF8
FgF20
FgF7
FgF19
UserXII-1
FgF18
FgF24
FgF16
FgF23
FgF22
LP1.T3
Site$_X$
LP1.T3.Site$_X$ strains x16
B
gRNA3
pCAS
Cas9
selection
A
gfp
Z
LP1.T3
Site$_X$
LP1.gfp
Site$_X$
LP1.gfp.Site$_X$ strains x16
%

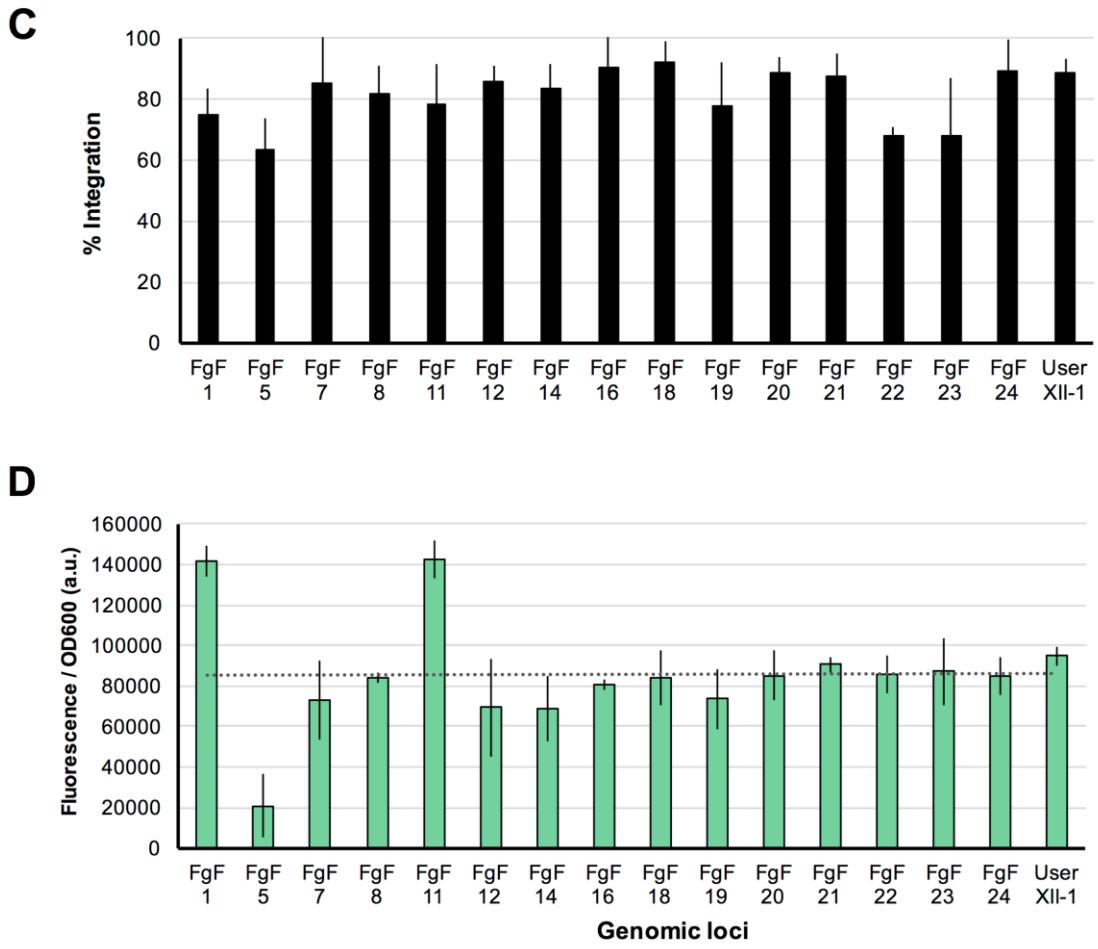


Figure 4. Evaluation of selected genomic loci in *S. cerevisiae*

(A) Construction of LP1.T3 strains for assaying integration efficiency and gene expression at each genomic locus. Sixteen LP1.T3.Site$_X$ constructs were generated by assembling and integrating LP1.T3 with ~500bp homology arms to each site. **(B)** Schematic overview of integration efficiency tests. Integration efficiency was assayed by integrating an LP1.*gfp* donor into the selected genomic sites using gRNA3 and pCAS. **(C)** Integration efficiency data at the genomic loci. Integration efficiency was determined by genotyping 12 colonies from each targeted genomic locus. Error bars represent the mean ± s.d. of three separate assays. **(D)** Gene expression levels at candidate genomic loci. Fluorescence intensity was measured for strains carrying *gfp* integrated into 16 different genomic loci. Fluorescence measurements were normalized against OD_{600} and corrected against autofluorescence of the (**–**) wt control and medium. Error bars represent the mean ± s.d. of three biological replicates. Hatched grey line represents the average fluorescence measured for all genomic loci.

3.4. Building a Landing Pad platform in *S. cerevisiae*

Overview of the LP platform

The LP platform was built into the quadruple auxotrophic *S. cerevisiae* strain CEN.PK2-1D to allow more versatility for downstream applications (work done by Michael Pyne). The composition of the final LP platform strain is illustrated in Figure 5A and includes four LP motifs: LP1, LP2, LP3 and LP4 that were integrated into the *S. cerevisiae* genome at different copy numbers reflected by their number designation (Table 3). Each of the four landing pads was assigned a single $gRNA_X$/target site: LP1 contains the gRNA8 target site (LP1.T8), LP2 contains the gRNA10 target site (LP2.T10), LP3 contains the gRNA1 target site (LP3.T1) and LP4 contains the gRNA9 target site (LP4.T9). The landing pads were then inserted into CEN.PK2-1D at the genomic loci assigned to each LP motif by integrating one or two constructs at a time using the CRISPR/Cas9 system. Previous reports have suggested that introducing multiple DSBs to a single chromosome can cause genome instability and decrease integration efficiency and cell viability [7]. Hence, to avoid an unnecessary burden on the cell, identical landing pads were integrated into sites on separate chromosomes. LP1 was integrated at FgF20, LP2 was integrated at FgF18 and FgF24, LP3 was integrated at UserXII-1, FgF7 and FgF19 and LP4 was integrated at FgF12, FgF16, FgF21 and FgF22. LP integrations were screened by colony PCR (Figure 5B) and sequence verified before the system was evaluated.

Table 3. Composition of the Landing Pad platform

LP	Targeting gRNA	LP integration loci		Copy number integration	Multiplex PCR amplicon (bp)
		Site	Chr		
LP1	gRNA8	FgF20	XVI	1	748
LP2	gRNA10	FgF18	XV	2	1181
		FgF24	XII		1419
LP3	gRNA7	UserXII-1	XII	3	645
		FgF7	V		800
		FgF19	XV		988
LP4	gRNA6	FgF12	IX	4	667
		FgF16	XIV		796
		FgF21	XVI		954
		FgF22	V		1110

After all four LP motifs had been inserted into their assigned genomic loci, we assessed the targeting specificity of the $gRNA_X$ constructs selected for the LP platform to ensure that optimal targeting specificity was maintained. The LP platform strain (CEN.LP) was transformed with pCAS plasmid harbouring one of the four $gRNA_X$ constructs assigned to LP1, LP2, LP3 and LP4 motifs (henceforth referred to as LP_X.gRNA). Targeting specificity is expressed as the % kill and was

determined by comparing the transformation efficiencies of each LP_XgRNA to the transformation efficiency of the (–)gRNA pCAS control plasmid. All four LP_X.gRNA constructs displayed high % kill values of >95% (Figure 5C), demonstrating that high targeting efficiencies of the selected LP_X.gRNAs were preserved within the LP platform. Next we set out to evaluate the usefulness of the system by i) measuring the efficiency of CRISPR/Cas9-mediated integration into LP1, LP2, LP3 and LP4 using *gfp* donor DNA, and ii) comparing the expression level of *gfp* expressed from the chromosome in 1-copy (LP1), 2-copies (LP2), 3-copies (LP3) and 4-copies (LP4). The purpose of these tests was to ensure that the LP platform supports efficient multi-copy gene integration and reliable titration of gene copy number, with the intention that it be used as a tool to help overcome pathway bottlenecks caused by enzyme inefficiencies.

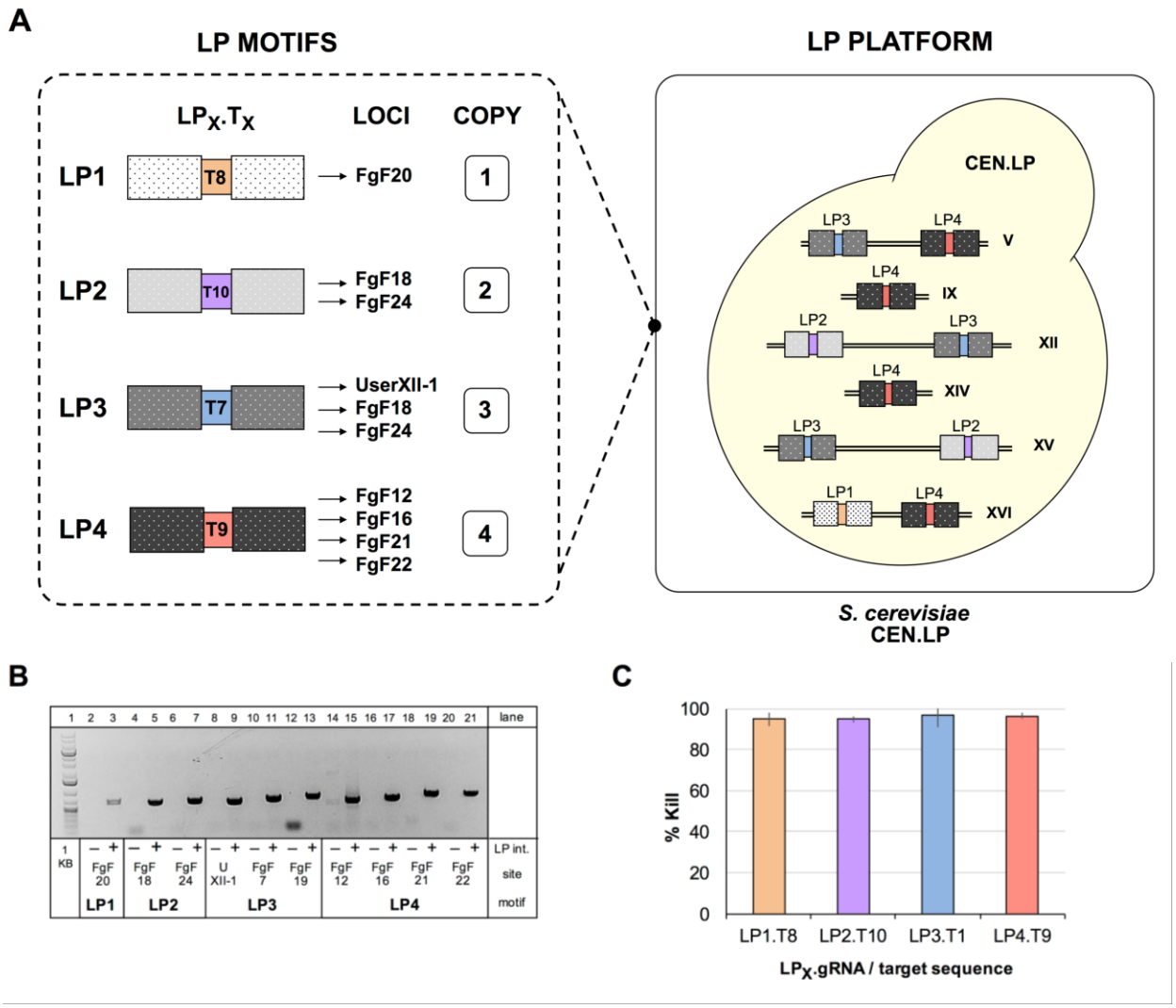


Figure 5. Overview of the LP platform in *S. cerevisiae*

(A) Left panel: Each landing pad was assigned a gRNA$_X$ target site and the resulting constructs were integrated into a single strain of *S. cerevisiae* at the selected genomic loci. **Right panel:** The LP platform strain in *S. cerevisiae* (CEN.LP) consists of four different landing pad motifs inserted into the genome at incremental frequency: LP1 = 1 locus, LP2 = 2 loci, LP3 = 3 loci, LP4 = 4 loci. Identical LPs were inserted on different chromosomes. **(B**) Genotyping of the final CEN.LP strain. LP integrations were screened using an LP-specific internal primer and a site-specific external primer. The parent strain CEN.PK2-1D (–) and the platform strain CEN.LP (+) were used as template. **(C)** Targeting specificity of selected LP.gRNAs into the LP platform. LP$_X$.gRNA targeting specificity tests were performed by co-transforming pCAS vector backbone with each of the four LP$_X$.gRNAs into the LP platform strain to evaluate the % kill associated with targeting 1-4 loci.

Overview of the LP platform integration strategy

Titration of gene copy number is performed in parallel using the devised LP platform integration strategy (Figure 6). The process begins by assembling four donor DNA constructs by PCR-amplification of a gene cassette using primers to attach 60 bp homology arms complementary to LP1, LP2, LP3 and LP4. The resulting donors are then co-transformed into CEN.LP in parallel with linear pCAS backbone and the LP_X.gRNA to target the appropriate landing pad: gRNA8 targets LP1, gRNA10 targets LP2, gRNA1 targets LP3, and gRNA9 targets LP4. Cells are plated on media selecting for expression of the pCAS plasmid, and the resulting transformants are screened by PCR to check for integration into each LP motif. Using the platform in its most basic form generates four *S. cerevisiae* strains harbouring 1-4 copies of any gene-of-interest in less than one week. The strains can then be tested in parallel to evaluate the effects of gene dosage on protein expression and activity.

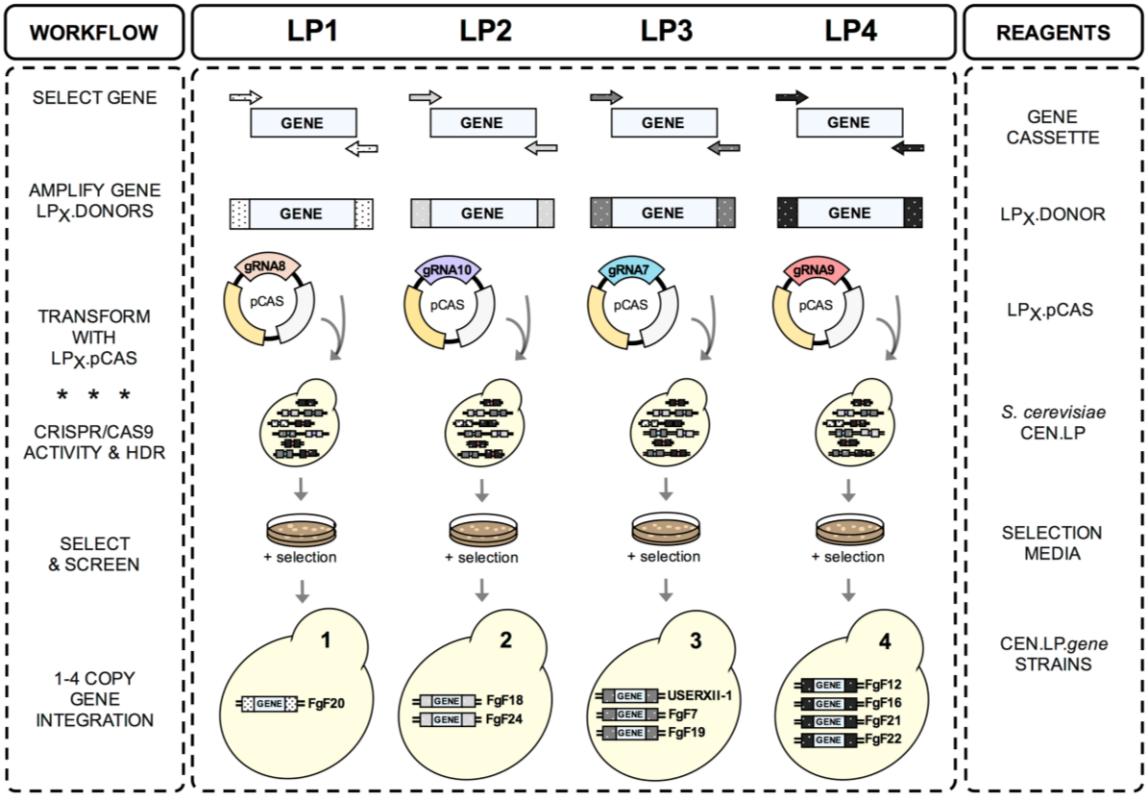


Figure 6. LP platform integration strategy

Donor DNA constructs are prepared by amplifying any gene of interest using primers to attach 60 bp homology arms. LP_X.donors are co-transformed independently with pCAS plasmid backbone and the respective LP_X.gRNA into the LP platform strain. Transformants are plated on media selecting for assembly of the pCAS plasmid. Integration of donor DNA occurs upon homology-directed repair of DSBs generated by Cas9 nuclease activity at the targeted LP.

3.5. Targeted integration efficiency into designated LP integration sites

Prior to applying our CEN.LP platform to a metabolic pathway, we first applied our devised landing pad integration strategy (Figure 6) using *gfp* as donor DNA and reporter gene. The *gfp* expression cassette (P_{TDH3}-*gfp*-T_{CYC1}) was integrated into the chromosome in 1-4 copies by targeting each of the LP motifs in four separate parallel transformations. Targeted integration into each LP motif was verified by colony PCR using locus-specific primer pairs for all loci within the LP motif. To accommodate multiplex PCR, primer pairs were originally designed so that amplification of each LP locus within the motif generated a different sized product in a single reaction (Table 3). This strategy was abandoned, however, as it often resulted in false negatives due to non-specific amplification. Instead, we used the same primer pairs to screen each LP site individually (Figure 7A). To evaluate the LP platform, we screened 12 colonies for each LP integration motif and counted the number of colonies with complete integration events. Overall, the rate of integration decreased as the number of targeted LPs increased (Figure 7B). Single-copy integrations into LP1 showed the highest efficiency of 97% compared to 81% integration into LP2, 53% integration into LP3, and 39% integration into LP4. Unexpectedly, targeting multi-copy LP2, LP3, or LP4 motifs frequently yielded partial integrant colonies in which at least one of the targeted LPs did not contain donor DNA. Efforts to increase the efficiency of the LP system and an investigation into why partial integration occurred is outlined in the next section.

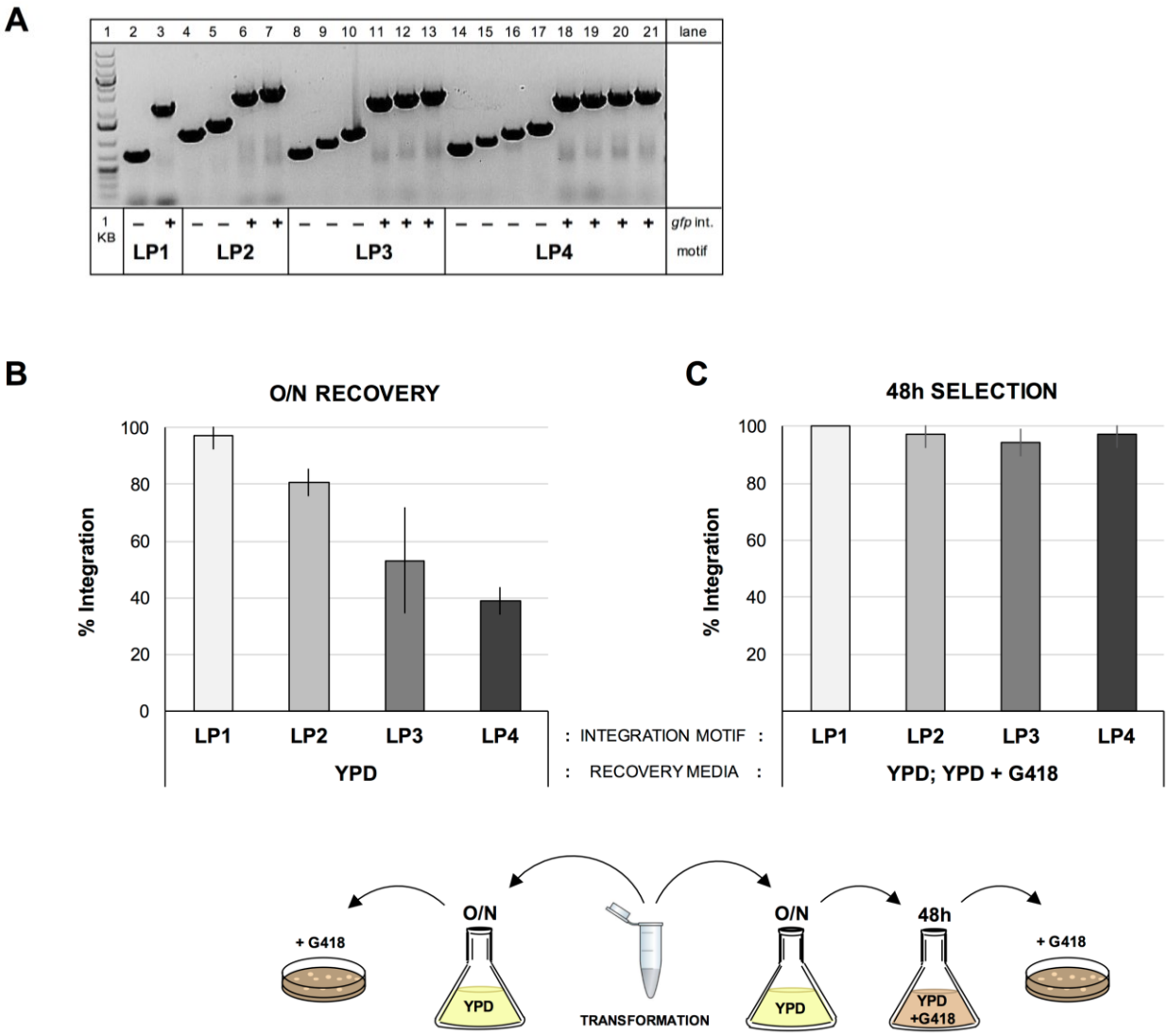


Figure 7. Evaluation of the LP platform.

(A) Genotyping of *gfp* copy number integrations into the LP platform. Lane 1: 1KB ladder; lanes 2-3: LP1.*gfp* integration; lanes 4-7: LP2.*gfp* integration; lanes 8-13: LP3.*gfp* integration; lanes 14-21: LP4.*gfp* integration. Successful integration of *gfp* at each locus (+) is shown alongside the LP platform strain (**–**). **(B)** Integration efficiency into the LP motifs. A *gfp* expression cassette flanked by LP homology arms was used as donor and co-transformed with the pCAS LP$_X$.gRNA expression vector into the LP platform strain. Transformation cultures were plated on selective medium and grown for two days. Integration efficiency was determined by genotyping 12 selected colonies for integration into each LP motif. Error bars represent the mean ± s.d. of three separate assays. **(C)** Integration efficiency into LP motifs after an additional two-day incubation in selective liquid medium. Recovered transformation cultures were transferred into rich medium containing G418 selection and incubated for two days before plating on selective medium. Integration efficiency was determined by genotyping 12 selected colonies for integration into each LP motif. Error bars represent the mean ± s.d. of three separate assays.

While integration into LP1 was highly efficient, efficiency gradually declined as the number of targeted LP integration sites increased. To optimize multi-copy gene integration using the LP platform, we performed a series of tests to uncover the rate-limiting component of our CRISPR integration system. While many variables can contribute to the success of CRISPR/Cas9-mediated gene integration, the availability of the gRNA, donor DNA and the Cas9 nuclease are the most influential and so we focused our efforts on these three components.

First, we investigated whether the efficiency of multi-loci integration events was limited by availability of the gRNA or donor DNA during transformation. For all previous assays, the gRNA expression cassettes were introduced as linearized fragments with flanking homology to the pCAS vector backbone harbouring *cas9* and the G418 selection marker. Assembly of these components is critical for: i) survival of transformants on selective medium, and ii) expression of gRNA and *cas9* nuclease. Since the transformation efficiency of the integration samples (+donor, +gRNA, +*cas9*) was consistently lower than the (–)gRNA transformation control (–donor, –gRNA, +*cas9*), we examined whether increasing the amount of the gRNA insert used for the assay would improve the rate of *in vivo* assembly of the pCAS vector and/or increase the number of transformants containing full integration of donor DNA. However, titrating the amount of gRNA from 0.4 to 1 µg did not yield significant improvements in either the transformation efficiency or the rate of full integration of donor DNA into the multi-copy LP motifs (data not shown). One possible explanation for this could be that the amount of donor DNA introduced in the previous assays was insufficient or disproportionate to the amount required to repair multiple DSBs generated upon targeting the LP2, LP3 and LP4 motifs. Any gains in the transformation efficiency or integration efficiency would be lost if the concentration of donor DNA was too low to repair all DSBs with high efficiency. To account for this, we increased the amount of donor DNA from 1 to 4 µg to accommodate repair of each additional DSB. Considering that 1 µg of donor DNA gave 97% efficiency for single-copy integrations, we reasoned that titrating the amount of donor DNA by 1 µg/LP target site may improve integration into LP2, LP3 and LP4 motifs. To further facilitate homologous recombination between the LP sites and donor DNA, we also increased the donor homology arms from 60bp to 280 bp as recommended by previous reports [54]. Implementation of these strategies, both independently and in combination, was also not effective for increasing the efficiency of multi-copy integration (data not shown). Taken together, these results indicate that neither the gRNA concentration, donor concentration, or homology length limit efficient multi-copy integration into the LP platform.

After establishing that multi-copy gene integration could not be improved by increasing the gRNA or donor concentration or increasing the donor homology length, we investigated whether our system was limited by activity of the gRNA targeting sequence and Cas9 nuclease. Evidence for this hypothesis was that partial integration of donor DNA continued to occur with the same frequency regardless of the gRNA or donor concentrations supplemented in the reaction. Furthermore, sequencing of LP3 and LP4 loci from partial integrant colonies revealed unmodified LP target sites in all cases where integration did not occur, indicating that a Cas9-mediated DSB had not been introduced to these sites. Since previous assays have already addressed the performance of the LP_X.gRNAs and the integration efficiency at selected loci, other explanations for partial integration events include: i) the occurrence of a random mutation in the gRNA targeting sequence or *cas9* coding sequence that abolishes CRISPR/Cas9 activity, or ii) expression and/or activity of the Cas9-gRNA complex is too low to generate multiple DSBs with high efficiency. To check whether gRNA or *cas9* mutations were responsible for partial integration events, the pCAS plasmid was prepped from three partial integrant colonies per LP motif and sequenced. None of the pCAS plasmids extracted from these partial integrant colonies contained mutations within the gRNA expression cassette or the *cas9* coding sequence, suggesting that poor performance of the Cas9-gRNA complex is likely a result of its abundance and activity level within the cell.

Improving the activity and abundance of the Cas9-gRNA complex would require a complete overhaul of the CRISPRm expression and delivery system adopted for this work [53], including swapping of the *cas9* or gRNA promoters to facilitate higher levels of expression, testing improved *cas9* variants, and switching to an ultra-high copy plasmid. While these modifications could be beneficial for our system, we first opted to test whether multi-copy integration into the LP platform could be improved without converting to an entirely new CRISPR/Cas9 expression strategy. To this end, we investigated whether incubating the transformation culture in liquid selection media would improve integration efficiency, as has been suggested [21, 52], by allowing more time for the Cas9-gRNA complex to generate DSBs at each of the targeted LP integration sites. Following co-transformation of pre-cloned LP_X.pCAS plasmids and LP_X.donor DNA into the LP platform, transformation cultures were first recovered overnight in non-selective medium and then transferred into fresh liquid medium containing G418 selection. These cultures were then incubated for two days in liquid selection medium to pre-select for expression of the pCAS plasmid before plating. Twelve colonies were then screened for complete integration of the LP_X.*gfp* donor into each of the LP motifs. This additional two-day growth period in liquid selection resulted in nearly 100% integration into all LP motifs (Figure 7C). Integration into LP3 and LP4 showed

remarkable improvement, achieving efficiencies of 94% for LP3 and 97% for LP4 copy number integrations. Integration efficiencies of LP1 and LP2 also improved, with LP2 reaching 97% efficiency for two-copy integration and LP1 achieving 100% efficiency for all experimental replicates. These results show that multi-copy integration efficiencies improve by growing cells in liquid selection prior to plating, which prolongs expression of the pCAS plasmid. The extended selection pressure allows more opportunity for CRISPR/Cas9 machinery to induce DSBs over multiple generations. This exposure likely enriches the culture with positive integration events by selecting for cells able to repair newly generated DSBs via chromosomal repair using the donor DNA that had previously integrated into the motif as template.

3.6. Titrating gene expression using the Landing Pad platform

After optimizing multi-loci gene integration into the LP platform to nearly 100% efficiency, next we wanted to establish whether the platform accommodates titration of gene expression by measuring fluorescence from strains expressing *gfp* from each LP copy number motif. Since *gfp* mRNA is directly proportional to the level of GFP fluorescence [85], fluorescence intensity could be used as a proxy to determine whether gene expression is proportional to copy number using the LP platform. As expected, fluorescence intensity of the four LP-GFP strains correlated with gene dosage, as higher levels of fluorescence were observed in strains carrying more copies of *gfp* (Figure 8A). A roughly 50% increase in fluorescence was observed for each additional copy of *gfp*. This result illustrates that differences in gene expression between LP motifs is modulated by gene copy number, as expression between LP sites is comparable (Figure 4C). To investigate whether multi-copy *gfp* expression levels are equivalent to the sum of expression levels from single-copy integrations at each locus, we summed fluorescence values obtained from single-copy *gfp* integrations (Figure 8B). Overall, the fluorescence generated by the sum of individual LP integrants mirrored the trend observed for strains harbouring multi-copy *gfp* integrations, though with slightly elevated fluorescence levels. Taken together, these results demonstrate that the landing pad platform can be exploited to titrate gene expression in *S. cerevisiae*.

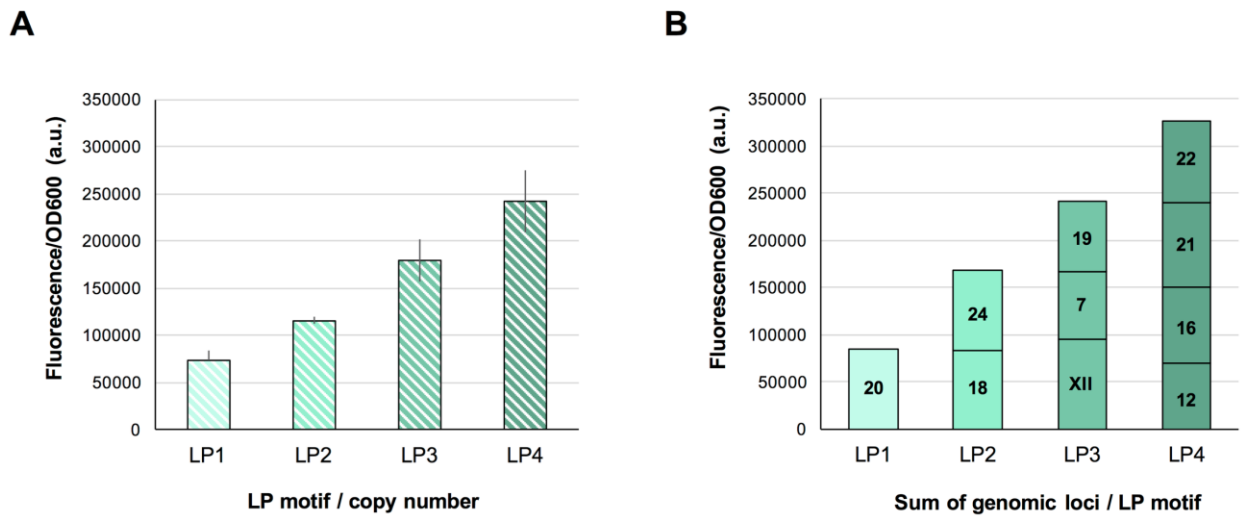


Figure 8. GFP expression analysis

(A) Fluorescence of LP platform strains expressing 1-4 copies of *gfp*. Fluorescence measurements were normalized against OD_{600} and corrected against the mean fluorescence of the wt (**–**) control. Error bars represent the mean ± s.d. of three biological replicates. **(B)** Combined fluorescence of single-copy *gfp* integrations at each LP motif locus. Fluorescence measurements were normalized against OD_{600} and corrected against the mean fluorescence of the wt (**–**) control.

3.7. Application of the Landing Pad platform

(*S*)-Norcoclaurine production in *S. cerevisiae*

Having established multi-copy integration of *gfp* using the landing pad platform, we next sought to apply our methodology to troubleshoot a key metabolic pathway bottleneck. In this regard, we targeted production of (*S*)-norcoclaurine, the central precursor to all BIAs [86]. (*S*)-Norcoclaurine is formed by the condensation of the L-tyrosine derivatives 4-hydroxyphenylacetaldehyde (4-HPAA) and dopamine, and is catalyzed by the enzyme norcoclaurine synthase (NCS) (Figure 9A). NCS catalyzes the first committed step in BIA biosynthesis and embodies one of the least efficient enzymes in the BIA biosynthetic network [64]. Conversion to (*S*)-norcoclaurine is limited by intracellular availability of 4-HPAA and dopamine substrates [75] and by the low catalytic efficiency of NCS [74], which reportedly only converts 0.25% of dopamine to (*S*)-norcoclaurine [75]. To increase titers of (*S*)-norcoclaurine in *S. cerevisiae*, we focused on improving conversion efficiency of the NCS-catalyzed reaction. Using the LP platform, we sought to: 1) identify NCS variants showing improved enzymatic properties, and 2) test whether increasing NCS gene copy number produces higher titers of (*S*)-norcoclaurine in yeast.

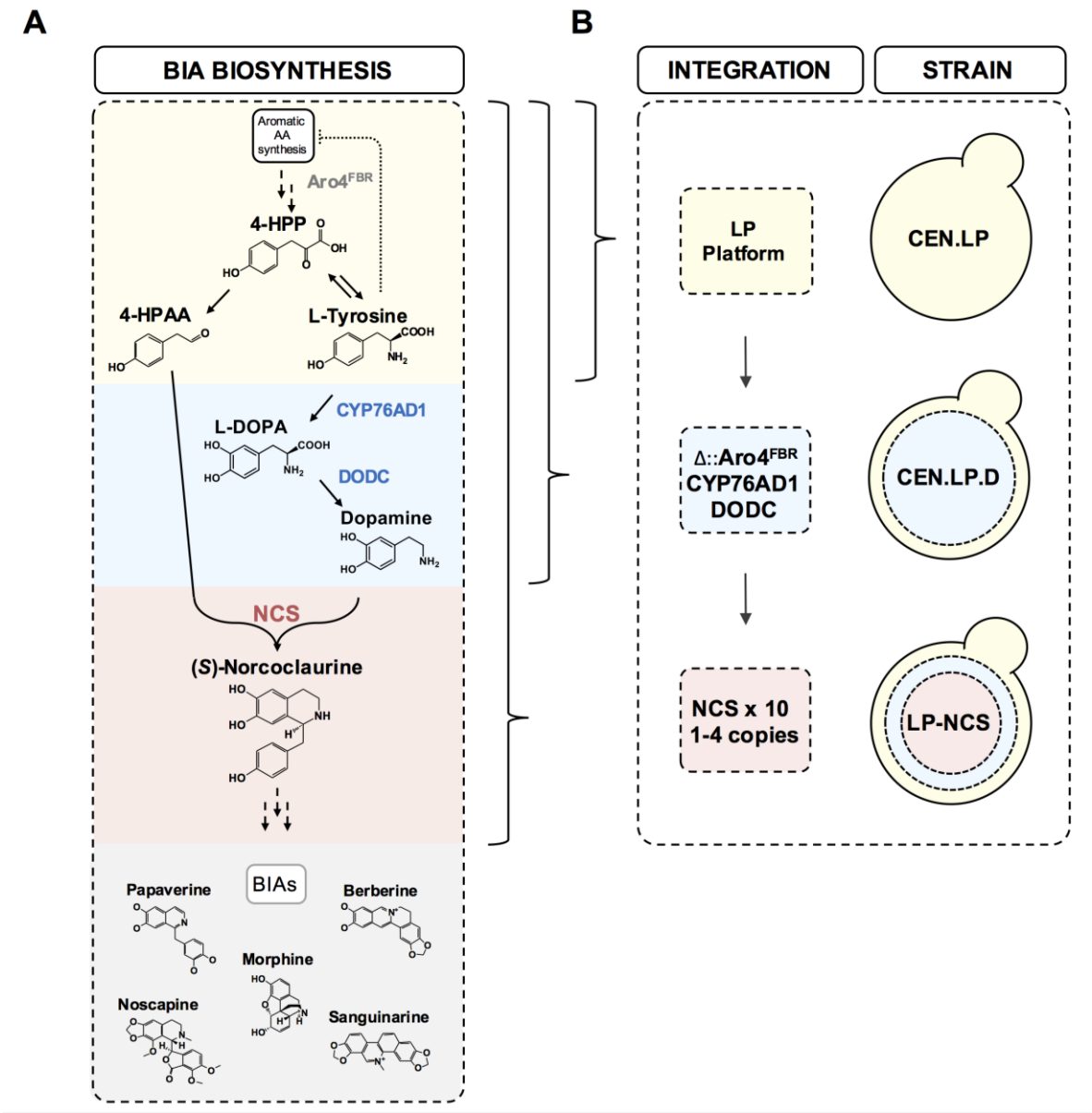


Figure 9. *De novo* synthesis of (*S*)-norcoclaurine in engineered *S. cerevisiae* strains

(A) *De novo* synthesis of (*S*)-norcoclaurine in *S. cerevisiae* towards the production of high-value BIAs. Mutant variation of *ARO4* is shown in grey text *Aro4*^{FBR} (*ARO4* feedback resistant mutant), to increase production of tyrosine heterologous enzymes are shown in coloured text. Blue text indicates the enzymes involved in the synthesis of dopamine: CYP76AD1, tyrosine hydroxylase; DODC, DOPA decarboxylase for the conversion of tyrosine to dopamine via the chemical intermediate L-DOPA. Pink text indicates the enzyme involved in the synthesis of (*S*)-norcoclaurine: NCS, norcoclaurine synthase catalyzes the Pictet-Spengler condensation of dopamine and 4-HPAA (4-hydroxyphenylacetaldehyde to generate (*S*)-norcoclaurine. Examples of high-value downstream BIA products are shown in the grey box. Dash errors indicate multiple enzymatic steps **(B)** Strain construction process towards the synthesis of (*S*)-norcoclaurine. Coloured budding yeast represent different strains that were built in this study and correspond to the enzymatic steps outlined in Figure 9A.

Engineering the CEN.LP strain to produce dopamine

(*S*)-Norcoclaurine is formed through a Pictet-Spengler condensation between the substrates 4-HPAA and dopamine. In *S. cerevisiae,* 4-HPAA is formed endogenously from 4-hydroxyphenylpyruvate (4-HPP), an intermediate in tyrosine biosynthesis [64], whereas synthesis of dopamine is dependent on introduction of heterologous enzymes into the cell (Figure 9A). Formation of dopamine from tyrosine is a two-step process involving hydroxylation and decarboxylation reactions that can occur in either order, though in yeast preference is given to the hydroxylation-first pathway, yielding the intermediate L-DOPA [64]. To impart dopamine biosynthetic capabilities to our landing pad strain and provide a background for norcoclaurine biosynthesis, we introduced a dopamine production cassette built by Deloache et al. (2015) [75] (Figure 9B). The cassette consists of three genes encoding a cytochrome P450 tyrosine hydroxylase from *Beta vulgaris* (*CYP76AD1*), a DOPA decarboxylase from *Pseudomonas putida* (*DODC*), and a feedback resistant mutant of yeast Aro4p ($Aro4^{FBR}$) to increase intracellular accumulation of tyrosine [75]. The dopamine expression cassette was amplified from pWCD2249 (Accession number KR232306.1) [75] in three overlapping fragments and integrated into the *ARO4* locus of CEN.LP, thus replacing the native *ARO4* gene with the $ARO4^{FBR}$ feedback resistant mutant. We then compared our dopamine-producing CEN.LP strain (CEN.LP.D) with yWCD745 (accession number KR232306), the dopamine production strain built by Deloache et al. (2015) [75]. After 48h growth in SC medium containing 4% glucose, millimolar concentrations of dopamine (Figure 10A) were detected from the supernatant for both strains with strain yWCD745 producing ~2.6x more dopamine than CEN.LP.D (Figure 10B). The difference in production observed between these two strains is likely an effect of the parent strains and integration loci chosen to introduce the heterologous genes. In their work, Deloache et al. used *S. cerevisiae* strain BY4741, derived from the S288c background, to integrate the dopamine cassette into the *URA3* locus, which differs from our CEN.PK host background. Differences within the shikimate pathway, which is affiliated with aromatic amino acid biosynthesis have been observed in CEN.PK and S288c backgrounds [87], and so the availability of tyrosine and other key intermediates may vary between these two strains. Likewise, the dopamine production cassette was expressed from different loci in the two strains, therefore transcriptional levels of the dopamine-producing enzymes may be disproportionate as well. While the cellular environment in the S288c background seems more favourable for dopamine production compared to CEN.PK, the amount of dopamine produced by the CEN.LP.D strain is sufficient to compare enzyme activities of the NCS variants.

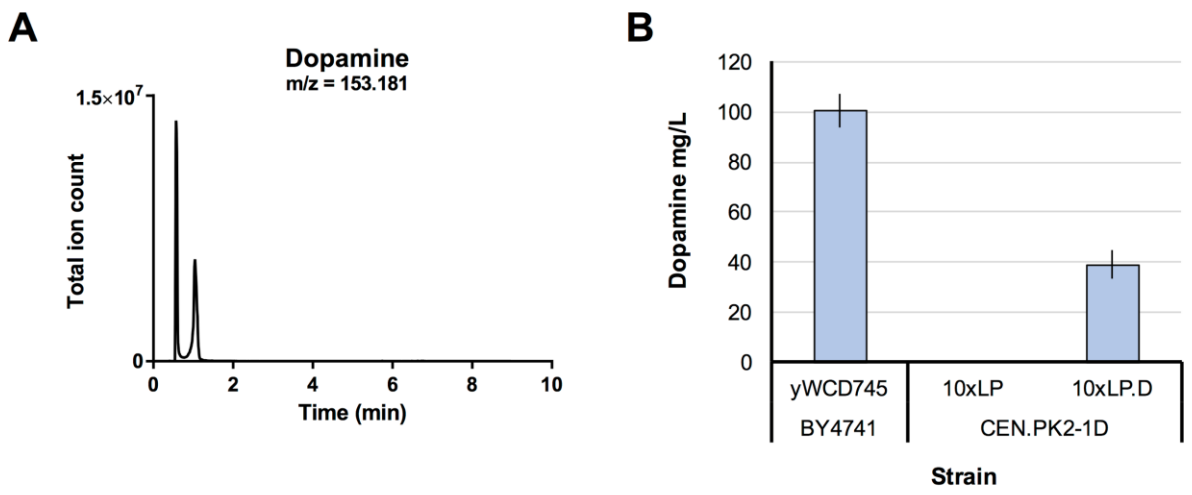


Figure 10. Analysis of dopamine production in CEN.LP platform strain

(A) LC/MS chromatogram of 1 mM dopamine standard **(B)** LC-MS analysis of dopamine in the supernatant of strain CEN.LP.D containing a dopamine production pathway (CYP76AD1; DODC; Aro4^{FBR}) after 48h growth in SC medium with 4% glucose. Error bars represent the mean ± s.d. of three biological replicates.

Compiling an NCS variant library

The selected NCS candidates were part of a library of enzymes previously identified and obtained within our laboratory by querying the PhytoMetaSyn Project databases [88]. The collection includes 10 NCS orthologs from the plant families Papaveraceae, Ranunculaceae, Berberidaceae, and Menispermaceae (Table 4). Half of the NCS candidates have been expressed in yeast, including *Ps*NCS1 [75] and *Ps*NCS3 [74, 75] from *Papaver somniferum*, *Nd*NCS (*Nandina domestica*) [74] , *Tc*NCS (*Tinospora cordifolia*) [75], and *Tf*NCS (*Thalictrum flavum*) [74, 75]. To our knowledge the remaining NCS candidates have not been tested for activity in yeast and include *Am*NCS (*Argenome mexicana*), *Ec*NCS (*Eschscholzia californica*), *Ps*NCS2 (*P. somniferum), Sc*NCS (*Sanguinaria canadensis*), and *Sd*NCS (*Stylophorum diphyllum*). BLASTp analysis revealed the NCS candidates share 35-91% sequence identity (Appendix Table A17). Multiple sequence alignment and phylogenetic analyses (Figure 11A) showed *Am*NCS from the Papaveraceae family is the most distantly related ortholog, sharing between 35-47% sequence identity with the other variants. *Sc*NCS and *Nd*NCS (*Nandina domestica*) share 91% sequence identity despite belonging to different families. Likewise, *Sd*NCS from Papaveraceae and *Tf*NCS from Ranunculaceae also share 90% sequence identity.

Screening NCS variants using the LP platform

To enhance enzyme expression levels, the NCS sequences were codon optimized for expression in yeast and cloned into the pBOT-HIS vector system [10] between the P_{TEF1} and T_{PGI1} regulatory elements. Using the integration strategy illustrated in Figure 6, 1-4 copies of each NCS variant were introduced into *S. cerevisiae* by co-transforming the LP_X-NCS donors with pCAS and the associated LP.$gRNA_X$ into the CEN.LP.D strain. Transformants were plated on selective medium and screened for complete integration into each LP motif. The assay produced a total of 40 NCS strains representing each variant integrated into S*. cerevisiae* genome in 1-4 copies. The resulting NCS variant and copy number library was assayed for *de novo* synthesis of (*S*)-norcoclaurine. Since racemic norcoclaurine has been known to form spontaneously [65], the background strain lacking NCS was included to determine the rate of spontaneous production. After 96h cultivation in rich medium and 4% sucrose (w/v), LC/MS was used to measure the concentration of (*S*)-norcoclaurine in culture supernatants (Figure 11B), as nearly all (*S*)-norcoclaurine synthesized by *S. cerevisiae* is secreted by the cell [75]. In the absence of an NCS enzyme, the CEN.LP.D strain did not produce detectable levels of norcoclaurine, suggesting that all norcoclaurine detected in cultures containing an NCS was produced enzymatically. Furthermore, enzymatic condensation of 4-HPAA and dopamine is enantioselective, therefore norcoclaurine produced by strains expressing *NCS* is expected to represent (*S*)-norcoclaurine [89].

Of the ten NCS variants, six produced detectable levels of (*S*)-norcoclaurine (*Ec*NCS, *Nd*NCS, *Ps*NCS3, *Sc*NCS, *Sd*NCS and *Tf*NCS) (Figure 11C). The best producers of (*S*)-norcoclaurine were strains containing *Sc*NCS and *Nd*NCS, which yielded 60 µg L^{-l} and 51 µg L^{-l} respectively, for single-copy integrations. Generally, increasing NCS copy number improved the production of (*S*)-norcoclaurine in strains harbouring an active NCS variant, though the degree of improvement varied between the NCS candidates and between each additional gene copy. The most pronounced improvements were observed for the top performing candidates. For example, four copies of *Sc*NCS and *Nd*NCS produced a greater than two-fold increase in (*S*)-norcoclaurine titer relative to the corresponding single copy strains, each producing approximately 130 µg L^{-l}.

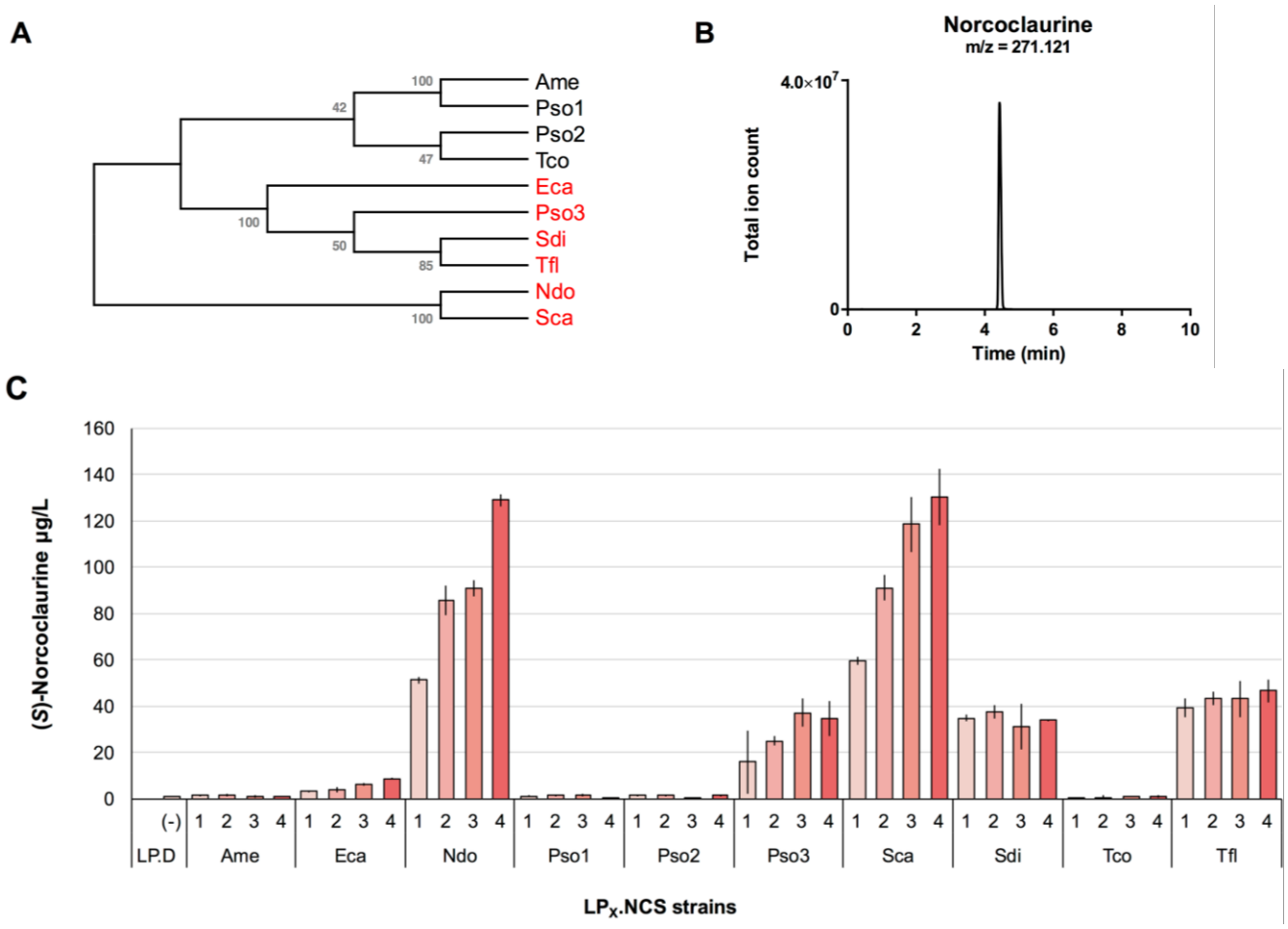


Figure 11. (*S*)-Norcoclaurine production from LP.NCSX strains

(A) Predicted evolutionary relationships among NCS candidates selected from the PhytoMetaSyn database. Candidates represent NCS variants from ten different plant species across four plant families: Papaveraceae, Ranunculaceae, Berberidaceae and Menispermaceae. NCS amino acid sequences were aligned by MUSCLE. The phylogenetic tree was generated by MEGA7 using the Neighbor-Joining method with a bootstrap value of 1000. Confidence values were generated for each branch (indicated in grey). Active NCS variants are shown in red, inactive NCS variants are shown in black **(B)** LC/MS chromatogram of 10µM (*S*)-norcoclaurine standard **(C)** LC/MS analysis of (*S*)-norcoclaurine in the supernatant of cultures expressing an NCS variant in 1 – 4 copies. Error bars represent the mean ± s.d. of three biological replicates.

Table 4. NCS candidates screened in this study

ID	Species	Family	(*S*)-norcoclaurine titer (μg L^{-1})	Ref	Accession Number
AmNCS	*Argenome mexicana*	Papaveraceae	–	This study	
EcNCS	*Eschscholzia californica*	Papaveraceae	8.7	This study	
PsNCS1	*Papaver somniferum*	Papaveraceae	` –	This study [75]	AKH61501.1
PsNCS2	*Papaver somniferum*	Papaveraceae	–	This study	
PsNCS3	*Papaver somniferum*	Papaveraceae	104.6 – 34.2	[75] [74] This study	KP262411
ScNCS	*Sanguinaria canadensis*	Papaveraceae	130.3	This study	
SdNCS	*Stylophorum diphyllum*	Papaveraceae	– 33.8	[74] This study	
NdNCS	*Nandina domestica*	Berberidaceae	– 129.0	[74] This study	
TcNCS	*Tinospora cordifolia*	Menispermaceae	N/A –	[75] This study	AKH61504.1
TfNCS	*Thalictrum flavum*	Ranunculaceae	N/A N/A 46.6	[74] [75] This study	

The LP platform enabled the simultaneous construction of 40 strains of *S. cerevisiae* expressing 10 different NCS variants from the genome in 1-4 copies. The LP platform facilitated the identification of two NCS candidates, *Sc*NCS and *Nd*NCS, producing superior (*S*)-norcoclaurine titers. Moreover, increasing gene dosage of *Sc*NCS and *Nd*NCS from one to four copies using the LP platform resulted in twice as much (*S*)-norcoclaurine produced. These results showcase the utility of the LP platform as a system to efficiently screen, test and modulate the activity levels of different enzymes variants to help improve flux towards desired products in heterologous pathways built into *S. cerevisiae*.

4. DISCUSSION

The LP platform was developed to accelerate efforts in yeast metabolic engineering towards the synthesis of high-value products. Engineering yeast to maximize production of novel compounds requires extensive pathway troubleshooting and optimization. Common genetic optimization strategies used to enhance flux through a target metabolic pathway include screening enzyme variants for improved activity [9, 65, 74, 75, 90], or tuning gene expression by testing different regulatory elements [67] and titrating gene dosage [21, 25, 90-92]. However, integrating new pathway modifications into the yeast genome remains a laborious process, and often necessitates many cycles through the design-build-test pipeline. There is a continual need to advance the pathway optimization process in order to reduce the time it takes to build superior yeast strains. The LP platform (Figure 5A) addresses this problem by providing a simple and efficient CRISPR/Cas9-based strategy for marker-less, multi-copy gene integration in *S. cerevisiae*, and was designed to facilitate enzyme library screens and test whether modulating gene copy number improves enzyme activity.

4.1. Benchmarking the Landing Pad platform

The majority of multiplexed CRISPR/Cas9 engineering strategies established in *S. cerevisiae* have focused on developing new methods for chromosomal integration of large multi-gene pathways [21, 53-55] and for simultaneous replacement of multiple native genes [28, 52, 56]. The work described here intends to expand the toolbox for multiplexed engineering in yeast by providing a well-characterized and reliable platform to facilitate multi-copy gene integrations for titrating gene dosage. The novelty of the LP platform is based on the following features: 1) synthetic target sequences and recombinogenic regions that are entirely unique within the yeast genome, 2) a single gRNA and donor DNA for integrating multiple copies of a gene in a single assay, and 3) systematic titration of gene copy number to easily test the effects of gene dosage. Unlike other multiplexed CRISPR/Cas9 engineering strategies, the LP platform is composed of synthetic parts. This allowed us to customize and scrutinize all of the components required to build the platform, including Cas9 targeting sequences (Figure 3) and recombinogenic regions (Appendix Table A3), which were then inserted into selected integration sites (Figure 4). Introducing artificial CRISPR/Cas9 ‘landing pads’ into the genome also allowed us to overcome a few limitations that can arise when targeting the same regions in the wild-type genome. For instance, customization allowed us to design a platform that supports multi-gene integration using

a single gRNA and donor DNA for gene integration into one, two, three, or four loci (Figure 6, Figure 7), whereas targeted integration at the same sites in the wild-type genome would require a gRNA and donor DNA for each individual target site. Indeed, it is possible to target multiple loci in the native genome using a single gRNA and donor DNA. By targeting the delta (δ) sequences that are distributed across the yeast genome, Shi et al. (2016) were able to integrate multiple copies of an entire biochemical pathway. Using a single targeting gRNA, they achieved integration of up to 18 copies of a 24 kb donor construct in a single reaction. Since there are hundreds of copies of the δ sequence found in the *S. cerevisiae* genome [93, 94], this approach is highly scalable. However, there is limited control over the number of copies integrated into the genome or where in the genome the integration occurs, and so targeting δ sites is not ideal for fine-tuning gene copy number and balancing pathway flux. In this regard, the LP platform provides a more systematic approach for titrating gene copy number in order to enhance flux towards desired end products.

4.2. Titrating gene expression using the Landing Pad platform

To ensure that transcription levels correlated with gene copy number in our system, we measured gene expression at various genomic loci and selected sites that showed similar expression levels (Figure 4D). This was verified by integrating *gfp* into each LP motif, which demonstrated that the configuration we chose for the LP platform facilitated titration of gene expression by copy number variation. A positive correlation between gene copy number and gene expression was observed for strains containing *gfp* integrated into the LP copy number motifs (Figure 8A). In other words, fluorescence levels increased in proportion to each additional copy of *gfp* introduced into the LP platform. This provides further confirmation that gene expression at each LP motif is proportionate to gene copy number, and is not confounded by regional differences in transcription levels between genomic loci.

While increasing copy number is an effective way to overcome enzyme inefficiencies, it can also be deleterious to the cell beyond a certain threshold. For example, Xie et al. (2015) observed increased lycopene titers from engineered lycopene-producing *S. cerevisiae* strains expressing two-copies of *CrtI* compared to strains expressing one copy of *CrtI.* However, strains expressing three-copies of *CrtI* produced lower lycopene titers and accumulate less biomass compared to the single copy strains. This study illustrates the importance of tuning gene copy number to achieve the optimal gene dosage, and also demonstrates that overexpressing genes can sometimes have detrimental effects that limit flux through the pathway [24] [91]. On the other hand, increasing gene dosage may not influence enzyme efficiency levels at all, and so expressing

additional copies of those genes would be energetically wasteful [25]. Considering these factors, the LP platform offers a simple strategy for optimizing gene dosage, and testing whether an enzyme is responsive to gene dosing or, alternatively, if other interventions are necessary to improve efficiency.

4.3. Targeted integration efficiency into the Landing Pad platform

Though different in design and purpose, the efficiency of multi-loci integration into the LP platform is comparable to other multiplexed integration strategies using CRISPR/Cas9 (Table 5). Under the initial transformation conditions, integration efficiency into the LP platform varied for each copy number motif (Figure 7B). While targeted integration of *gfp* into LP1 was highly efficient, the efficiency of multi-copy *gfp* integration declined as the number of targeted loci increased. Other groups have reported similar reductions in integration efficiency for multi-loci targets. Bao et al. (2015) reported 100% efficiency for a single gene disruption using 100 bp donor DNA, and only 27% efficiency for triple-gene disruption events. Similarly, using the CasEMBLR method, Jakočiūnas et al. (2015) achieved 97% efficiency for integration into a single locus, 58% efficiency at two loci, and 31% efficiency at three loci. At first glance, the rate of integration reported by Jakočiūnas et al. seems comparable to the initial integration efficiencies reported in this study (Figure 7B). However, their work also required the assembly of a 5-part donor for each targeted integration event (15 parts total), which likely reduced the efficiency that could be obtained using pre-assembled donor. Nevertheless, the most convincing evidence to suggest that multi-loci LP integration can be achieved at higher efficiencies came from a study by Ronda et al. (2015), who integrated three large constructs ranging from 5.1-6.6 kb into three separate loci with 84% efficiency [54]. In addition to achieving high multi-loci integration efficiency, this study was also notable because large DNA constructs are often more difficult to integrate than smaller DNA constructs. This prompted further investigation into how multi-loci integration into the LP platform could be optimized to achieve similar efficiencies.

4.4. Improving integration efficiency by propagating cells in selective media

By sequencing partial integrant colonies, we determined that the efficiency of multi-loci integration into the LP platform was ultimately limited by the rate of Cas9-mediated DSB formation. We and others have shown that generating DSBs at the target site greatly enhances integration efficiency (Figure 3E) [39, 54, 55, 95, 96], and so it is unlikely that integration will occur at any targeted integration site that remains intact. Therefore, to increase the efficiency for multi-loci

integration using CRISPR/Cas9, it was important to increase the activity of the Cas9-gRNA complex in order to generate DSBs at all targeted loci. Based on previous assays, we suspected that Cas9-gRNA expression and/or activity was too low to efficiently cleave host DNA at multiple loci. Poor performance of the pCAS expression system was likely compounded by the low turnover rate observed for Cas9, which remains bound to DNA even after cleavage has occurred [49].

To overcome the implicit limitations of the CRISPR/Cas9 system used for LP integrations, we propagated transformation outgrowth cultures in liquid selection medium over two days and achieved nearly 100% integration efficiency into all LP motifs (Figure 7C). This approach was also implemented by Bao et al. (2015) and Shi et al. (2016) to improve the efficiency of their respective multiplexed CRISPR/Cas9-based strategies. Bao et al. (2015) speculated that continued selective pressure to maintain the CRISPR/Cas9 plasmid prolongs cellular exposure to Cas9 nuclease activity, and provides more time to generate the necessary DSBs [52]. Due to the lethality of unrepaired DSBs, we also suspect that prolonged expression of the Cas9-gRNA complex enriches the culture for cells able to repair DSBs using the donor construct supplied in the initial transformation (Figure 12). Using LP4 as an example, our rationale is as follows: i) upon transformation into LP4, many of the transformants would have contained partial integrations i.e. one out of four ‘LP4’ landing pads contained integrated donor DNA, while the other three ‘LP4’ landing pads remained untouched; ii) Sequencing also showed that the gRNA and *cas9* constructs were intact and unmutated, which indicates that they are functional inside the cell; iii) therefore, prolonging the exposure of CRISPR/Cas9 machinery means that each LP4 target site that remained intact during the initial transformation would eventually be interrogated by Cas9; iv) new Cas9-induced DSBs are then repaired by HDR using the donor DNA construct that had integrated into the chromosome at an LP4 site during the initial transformation, eventually converting partial integrants to full integrants. Chromosomal repair mechanisms inherent to *S. cerevisiae* [97] likely drive HDR of newly generated DSBs using the LP.A and LP.Z recombinogenic regions of intact LPs from different loci within the same motif. Complete integration into the LP motif occurs when all of the targeted loci have been repaired by LP sites harbouring previously integrated donor DNA. Conversely, any cells lacking an integrated donor within the targeted motif will have no mechanism to replace the LP target sites to prevent Cas9-induced DSBs in future generations. Therefore, as the culture approaches saturation it becomes enriched with cells that have replaced all of the target sites with donor DNA. As well, cells unable to replace the target sites with donor DNA will be gradually eliminated from the population by Cas9 interrogation. Hence, the majority of cells that are plated after the two-day propagation in selective medium will have completed integration at

the targeted LP motif. Although this method provides a simple solution for achieving higher integration efficiencies into the LP platform, it requires more time to produce the desired results. Therefore, other solutions should be investigated in order to optimize the efficiency of the strategy.

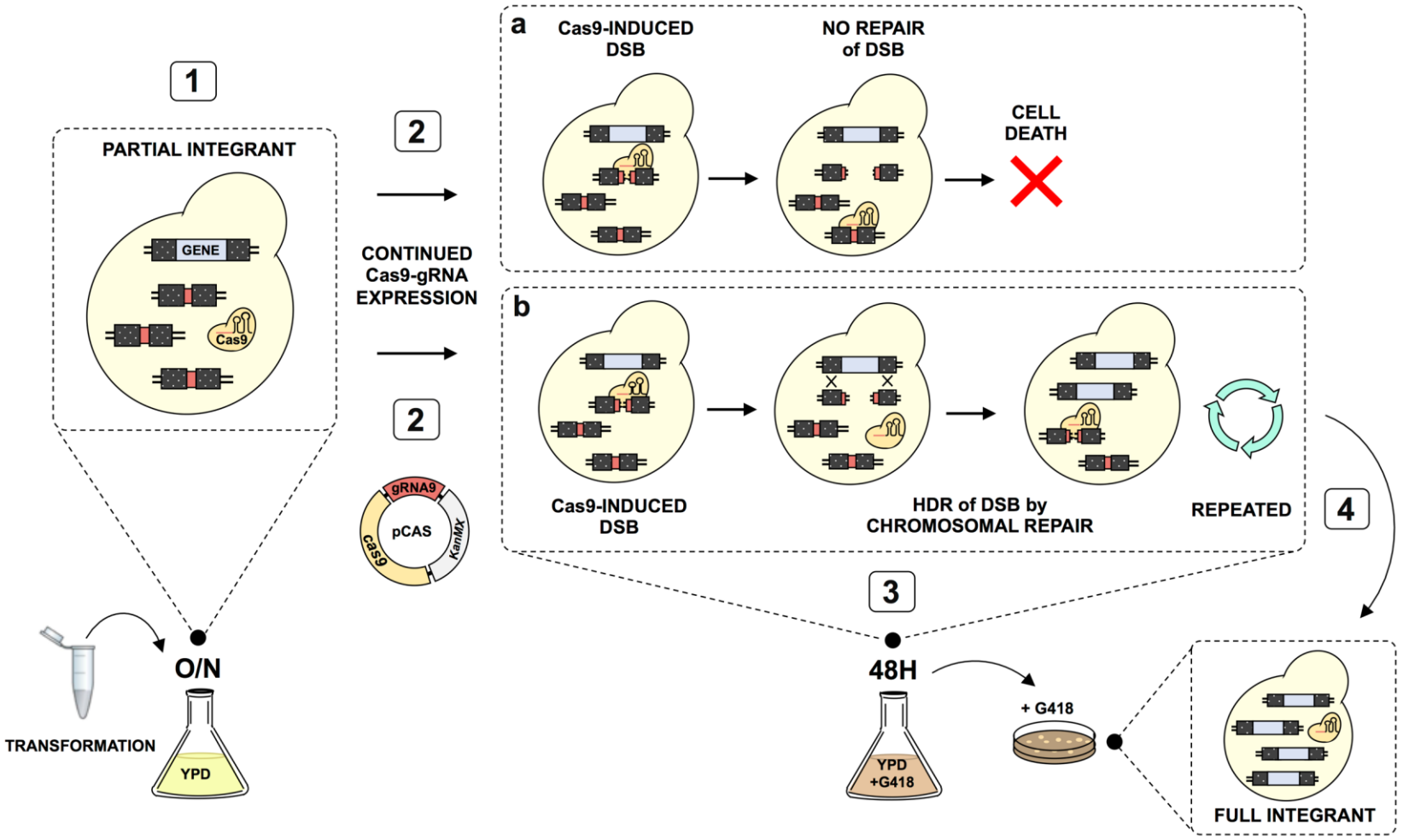


Figure 12. Outgrowth in liquid selection medium improves integration efficiency

(1) Transformation of donor DNA into LP4 results in partial integration event at a single LP4 locus. **(2)** Transformants are transferred to medium selecting for the pCAS plasmid, resulting in prolonged exposure to Cas9-gRNA for 48h. **(3)** Continued Cas9-gRNA interrogation induces DSBs at LP4 loci that remained intact in the initial transformation **(3a)** Cells unable to repair newly generated DSBs will not survive and will be gradually eliminated **(3b)** Cells able to repair new DSBs by HDR using previously integrated donor DNA will survive. **(4)** Continued LP4 interrogation by Cas9-gRNA over multiple generations will eventually convert partial integrants into full integrants. Once a cell has reached full LP4 integration, there are no target sites left for Cas9-gRNA to interrogate.

4.5. Optimizing CRISPR/Cas9 expression systems in *S. cerevisiae*

CRISPR/Cas9-mediated integration in *S. cerevisiae* is dependent on two critical processes: 1) the generation of Cas9-mediated DSBs and 2) subsequent HDR repair of the DSB with donor DNA [39]. In our effort to improve multi-loci integration efficiency into the LP platform, it became evident that our system was limited by suboptimal DSB generation, presumably caused by low-activity of the Cas9-gRNA complex. Even though we achieved high integration efficiencies after prolonged exposure of cells to the Cas9-gRNA complex, the additional two-day time commitment is not ideal for advancing the strain building process. Therefore, further optimization of Cas9-gRNA activity is needed to overcome the limitations of our current strategy. This could be accomplished by either modifying the existing CRISPR/Cas9 system presently in-use, or by exploring alternative CRISPR/Cas9 systems used by the other multiplexed engineering strategies summarized in Table 5.

Targeting gRNA

The gRNA targeting sequence is a key parameter for all CRISPR/Cas-based engineering strategies, and is often the first component that requires optimization. A poorly designed gRNA will diminish targeted Cas9 nuclease activity and drastically reduce the efficiency of targeted integration [39, 57, 98]. Therefore, multiplexed CRISPR/Cas9-based strategies that use more than one gRNA for targeting multiple loci will be only as efficient as their least effective gRNA target sequence [53-55]. Our multiplex integration strategy avoids this potential constraint by expressing a single, high-functioning synthetic gRNA for multi-loci targeting. Since the target sequences we selected for the LP platform demonstrate high targeting specificity (Figure 3D, Figure 5C) and high integration efficiency (Figure 3E), it is unlikely that the gRNA target sequences require further optimization. However, it is also important to consider the gRNA expression systems used for multiplexing, as these systems determine the intracellular availability of the gRNA for guiding Cas9 to the target sites. Multiplexed CRISPR strategies that use different gRNAs for targeting multiple loci typically express the gRNAs from individual expression cassettes cloned into the same plasmid [52, 54, 55], or assembled on separate plasmids [53]. Since each gRNA targets a single locus, these expression systems likely produce sufficient gRNA levels for guiding Cas9 to all targeted loci. For systems such as ours that use a single gRNA expression cassette for targeting more than one locus, the demands are less clear. It is possible that the intracellular levels of gRNA are too low to reliably guide Cas9 to all of the LP sites within the multi-loci motif, which would explain the low efficiency of DSB-induction within our system. However, the pCAS vector used to

express the LP gRNAs contains an enhanced gRNA architecture that greatly improves the intracellular levels of functional gRNA [28], which may help overcome the higher demands placed on gRNAs targeting multiple loci. Nevertheless, further examination into the relationship between gRNA availability and efficiency of Cas9-mediated DSB induction is needed to uncover whether the gRNA expression levels are rate-limiting in our current system.

Cas9 expression and delivery systems

Different multiplexed CRISPR/Cas-based strategies have employed various approaches for *cas9* expression and delivery, including the use of strong constitutive promoters [21, 52, 55, 82, 99], chromosomal expression of *cas9* [53, 54, 99], or using engineered Cas9 enzyme variants [21, 52]. Based on the efficiencies reported by other multiplexed CRISPR/Cas9 integration strategies, it appears our integration strategy could benefit by expressing *cas9* from a high strength promoter rather than the medium strength promoter currently employed. During the design of the pCAS plasmid, Ryan et al. (2014) showed that high expression of *cas9* reduced cellular fitness, whereas moderately expressed *cas9* maintained wild-type fitness levels [28]. However, more recent CRISPR/Cas9-based applications have shown that high *cas9* expression (–gRNA) does not adversely affect growth rates [54, 55, 82, 99]. Even including studies that used a higher functioning Cas9 variant expressed from an ultrahigh copy number plasmid (~200 per cell) [21, 52] do not report that Cas9 is detrimental to the cell when expressed without gRNA. Nevertheless, the effects of constitutive Cas9 expression over time are still largely unknown, and could vary between different strains and CRISPR/Cas9 systems. In this regard, transient expression from an episomal plasmid is preferred over chromosomal expression because *cas9*-expressing plasmids can be conveniently removed by culturing in non-selective media. Plasmid-based *cas9* expression systems are also associated with higher integration efficiencies [21, 52, 54, 55] compared to the integration efficiencies achieved when *cas9* is expressed from the chromosome [53, 54]. This might be explained by copy number differences between the two systems, where the *cas9* copy number ranges from 1-200 copies using different expression vectors, whereas chromosomal *cas9* expression is typically from a single copy. Considering the wide copy-number range offered by different expression vectors, it is also easier to enhance *cas9* expression levels with plasmid-based systems. This would be especially useful for multiplexed strategies that may require high *cas9* expression for efficient targeted DSB induction. Further improvements could be made by expressing a more active Cas9 nuclease such as the iCas9 variant used in the HI-CRISPR system [52], or probing other genome editing nucleases. Taken together, the results obtained from these strategies indicate that high constitutive expression from

an episomal plasmid is the best approach for optimizing intracellular Cas9 levels. Increasing Cas9 abundance may also necessitate further optimization of gRNA expression such that neither component limits the formation or activity of the Cas9-gRNA complex.

Table 5. Summary of multiplexed CRISPR/Cas9-mediated integration strategies

Method	Number of integration sites	Donor size	gRNA			Cas9			Efficiency	Time required	Ref
			targeting gRNAs	promoter construct	delivery	variant	promoter strength	delivery			
HI-CRISPR	3	100 bp	3 *	SNR52	Ultrahigh copy plasmid	iCas9	Strong	Ultrahigh copy plasmid	27%	1 week [O]	Bao et al. (2014)
Multiplexed CRISPR-Cas	3	24 bp ~8 kb[a] /site	3	SNR52	2µ plasmid [a]	Cas9	Moderate	Chromosome[T]	64% 4.2%	1 week	Horwitz et al. (2014)
CrEdit	3	5.1 kb 5.8 kb 6.6 kb	3	SNR52	2µ plasmid	Cas9	Strong	CEN/ARS plasmid [T]	84%	1 week	Ronda et al. (2015)
CasEMBLR	3	~1.5 kb [a] per site	3	SNR52	2µ plasmid	Cas9	Strong	CEN/ARS plasmid [T]	31%	1 week	Jakočiūnas et al. (2015)
Di-CRISPR	11 10 10	8 kb 16 kb 24 kb	1 *	SNR52	Ultrahigh copy plasmid	iCas9	Strong	Ultrahigh copy plasmid	~84% ~68% ~60%	>1 week [O]	Shi et al. (2015)
LP platform	1 2 3 4	1.5 kb	1 * 1 * 1 * 1 *	$tRNA^{Tyr}$ 3'HDV ribozyme	2µ plasmid [a]	Cas9	Moderate	2µ plasmid	100% 97% 94% 97%	1 week [O]	This study

a assembled in vivo

***** gRNA & *cas9* delivered on the same plasmid

T pre-transformed *cas9* into *S. cerevisiae*

O includes outgrowth in liquid selection medium

4.6. NCS activity and the effects of gene dosage

To demonstrate the utility of the LP platform in pathway engineering, we used the system to increase gene copy number of ten NCS enzyme orthologs in order to identify the best variant for (*S*)-norcoclaurine synthesis in yeast, and to test whether increasing gene copy number enhances production. Ten NCS variants were integrated into the LP platform in 1-4 copies, generating 40 LP.NCS strains that were then analyzed for production of (*S*)-norcoclaurine. From this assay, we identified six NCS variants that are functional in *S. cerevisiae*, and demonstrated that increasing gene copy number improves the efficiency of some variants to synthesize (*S*)-norcoclaurine (Figure 11C). The most active NCS enzymes identified in this study were *Nd*NCS and *Sc*NCS, which produced the highest (*S*)-norcoclaurine titers from single copy strains. Moreover, expressing four copies of either *Nd*NCS or *Sc*NCS generated ~130 µg/L (*S*)-norcoclaurine, which is 20 µg/L more than the highest producing strain that has been reported for *de novo* synthesis of (*S*)-norcoclaurine [75].

Generally, gene dosage effects for the majority of active NCS variants were consistent with GFP expression data. Strains expressing four copies of *EcNCS*, *NdNCS*, *PsNCS3*, or *ScNCS* produced two-fold more (*S*)-norcoclaurine compared to the respective single-copy strains. However, unlike the linear relationship observed between *gfp* copy number and GFP fluorescence (Figure 8A), the trends observed for *NCS* copy number and (*S*)-norcoclaurine production were not entirely linear. For example, varying degrees of improvement in the production levels of (*S*)-norcoclaurine were observed for each additional copy of *NdNCS* expressed in *S. cerevisiae*; i.e. a 66% increase from one to two copies, a 5% increase from two to three copies and a 42% increase from three to four copies. However, if the three-copy *NdNCS* data points are omitted, the remaining data generated for the *Nd*NCS strains closely resembles the linear trend observed between the *Sc*NCS copy number strains. Considering these variants are nearly identical, it is possible that the three-copy *NdNCS* strains are defective in some way that diminishes the level of (*S*)-norcoclaurine produced and/or detected. Similarly, the production of (*S*)-norcoclaurine was also not proportionate between strains expressing three- and four-copies of *PsNCS3*. From one- to three-copies of *PsNCS3*, the production of (*S*)-norcoclaurine increased by 60% for each additional copy expressed in yeast, while the strains expressing four copies of *PsNCS3* were less productive than the three-copy strains. While not directly comparable, a recent study done by Li et al. (2016) revealed that some NCS variants from the Papaveraceae family naturally occur as tandem fusions consisting of multiple catalytic domains, and demonstrated that enzyme efficiency increases in proportion to the number of repeated domains present. Functional expression

analysis of strains expressing *PsNCS3* variants encoding four catalytic domains showed a 20-fold improvement in catalytic efficiency compared to the single domain variant [74]. Although the catalytic activities of enzymes expressed as tandem fusions are not exactly analogous to the activity of enzymes expressed from multiple gene copies, the results imply that increasing *NCS* copy number can be an effective strategy for enhancing enzyme efficiencies.

However, not all NCS variants that are functional in *S. cerevisiae* respond to increased copy numbers. For instance, titrating the gene dosage of *SdNSC* and *TfNCS* did not affect production of (*S*)-norcoclaurine, which remained relatively stable between copy number variants (Figure 11C). This was unexpected because strains containing *Sd*NCS or *Tf*NCS variants produced more (*S*)-norcoclaurine than *Ps*NCS3 and *Ec*NCS between single-copy strains, and so we would expect that increasing gene dosage would affect the efficiency of these variants. In their study, Li et al. (2016) also showed that *Sd*NCS was more efficient than *Ps*NCS3 when expressed as a single domain protein in *S. cerevisiae*. However, functional expression of an *Sd*NCS variant encoding two catalytic domains only produced marginal improvements in enzyme efficiency compared to the multi-domain *Ps*NCS3 variant [74], further suggesting that *Sd*NCS activity is not contingent on gene copy number. Moreover, considering that the *Tf*NCS and *Sd*NCS variants are nearly identical (Appendix Table A17, Appendix Figure A1), it is conceivable that *Tf*NCS functions in a similar manner. Both *Tf*NCS and *Sd*NCS variants contain a predicted chloroplast signalling peptide at their C-termini, and so it is possible that they are localized to a membrane that becomes saturated at a single copy. Indeed, studies have shown that increasing the copy number of heterologous enzymes does not always lead to increases in activity [25, 91, 100]. Differences in translational processes or protein folding may present certain limitations on enzyme levels [91, 100], or the enzyme levels could be regulated by other mechanisms that prevent accumulation within the cell [101]. A study done in *Pichia pastoris* showed that increasing gene copy number can sometimes affect protein folding and disulfide bond formation in the endoplasmic reticulum (ER), and can prevent secretion from the ER and cause the nascent protein to accumulate inside the cell [91]. Therefore, it is possible that specific properties in the amino acid sequence shared between variants *Sd*NCS and *Tf*NCS could have affected the way these proteins are processed in the cell and limit their activity in *S. cerevisiae*. Further investigation into the protein abundance and solubility of functional NCS variants may help explain why the activity of some variants correlates with gene dosage, while other variants remain unaffected.

4.7. Future directions

This work focused on establishing the LP platform in *S. cerevisiae* as a system for multi-copy gene integration, however, the upper limits for copy-number integration into the LP platform have not been established. By targeting more than one LP motif in different combinations, the LP platform has the potential to facilitate integration of up to ten copies of a gene in a single transformation. For example, targeting LP1 and LP4 would generate a five-copy integration, LP2 and LP4 would generate a six-copy integration, and so on until all four LP motifs are targeted to generate a ten-copy integration event. This would not require any new parts, only new combinations of gRNAs and donors already used for integration into the LP platform. As well, Horwitz et al. (2015) established that it is possible to transform multiple gRNAs as linear cassettes for assembly with the same vector backbone, therefore only one selection marker is required for expressing the gRNAs and *cas9*. However, targeting more than one LP motif would likely necessitate additional outgrowth in liquid selection media to improve integration efficiency, as we observed for the multi-copy integrations targeting each LP motif independently. To further increase the efficiency of our integration strategy, we could apply other CRISPR/Cas9 delivery systems, such as the HI-CRISPR system developed by Bao et al. (2015) that co-expressed the iCas9 variant with the gRNA from ultra-high copy number plasmid. This may circumvent the need for the incubation period in liquid selection, though this is uncertain as both Bao et al. (2014) and Shi et al. (2016) included this step while using the HI-CRISPR system.

One of the drawbacks of the current LP platform is that it is limited in scale. In our current strategy, once a gene has integrated into an LP motif, that motif cannot be targeted for future integrations. This makes it difficult to build upon metabolic pathways or optimize the dosage of more than one gene in the same strain. It would be useful to fuse a new recombinogenic region and synthetic target site to either the 5' or 3' end of the gene construct, so that the LP motifs can be recycled for future use. This would extend the utility of the LP platform for building metabolic pathways and balancing flux by optimizing expression levels of multiple enzymes.

4.8. Conclusion

There is a continuous effort to advance the pathway optimization process in order to efficiently build new yeast strains capable of producing high-value bio-based chemicals. The LP platform presented in this work supports these efforts by providing a simple and efficient strategy for marker-less, multi-copy gene integration in *S. cerevisiae* using the CRISPR/Cas9 genome

editing system. We demonstrated the utility of the LP platform for titrating gene expression by integrating different copy number variations in a single step, which permits the rapid testing of gene dosage effects on enzyme efficiency for pathway optimization. As proof-of-concept, we used the LP platform to titrate gene copy number of ten NCS enzyme variants in order to improve production of (*S*)-norcoclaurine in yeast, the entry point into BIA metabolism and a major bottleneck. The LP platform facilitated the discovery of two NCS enzyme variants that exhibited higher activity levels and enhanced efficiencies by increasing gene dosage. Given the efficacy for integrating multiple gene copies and enhancing gene expression, the LP platform provides a simple tool for optimizing flux through native or heterologous metabolic pathways in *S. cerevisiae*.

REFERENCES

1. Ostergaard, S., L. Olsson, and J. Nielsen, *Metabolic engineering of Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 2000. **64**(1): p. 34-50.
2. Da Silva, N.A. and S. Srikrishnan, *Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae.* FEMS Yeast Res, 2012. **12**(2): p. 197-214.
3. Borodina, I. and J. Nielsen, *Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals.* Biotechnol J, 2014. **9**(5): p. 609-20.
4. Gustavsson, M. and S.Y. Lee, *Prospects of microbial cell factories developed through systems metabolic engineering.* Microb Biotechnol, 2016. **9**(5): p. 610-7.
5. Reider Apel, A., et al., *A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae.* Nucleic Acids Res, 2017. **45**(1): p. 496-508.
6. Lee, M.E., et al., *A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly.* ACS Synth Biol, 2015. **4**(9): p. 975-86.
7. Jessop-Fabre, M.M., et al., *EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9.* Biotechnol J, 2016. **11**(8): p. 1110-7.
8. Li, M. and I. Borodina, *Application of synthetic biology for production of chemicals in yeast Saccharomyces cerevisiae.* FEMS Yeast Res, 2015. **15**(1): p. 1-12.
9. Narcross, L., et al., *Mining Enzyme Diversity of Transcriptome Libraries through DNA Synthesis for Benzylisoquinoline Alkaloid Pathway Optimization in Yeast.* ACS Synth Biol, 2016. **5**(12): p. 1505-1518.
10. Pyne, M.E., et al., *Reconstituting Plant Secondary Metabolism in Saccharomyces cerevisiae for Production of High-Value Benzylisoquinoline Alkaloids.* Methods Enzymol, 2016. **575**: p. 195-224.
11. Bailey, J.E., *Toward a science of metabolic engineering.* Science, 1991. **252**(5013): p. 1668-75.
12. Nielsen, J. and J.D. Keasling, *Engineering Cellular Metabolism.* Cell, 2016. **164**(6): p. 1185-1197.
13. Morris, J.S., et al., *Plug-and-Play Benzylisoquinoline Alkaloid Biosynthetic Gene Discovery in Engineered Yeast.* Methods Enzymol, 2016. **575**: p. 143-78.
14. Paul M. Sharp, T.M.F.T., Krzysztof R. Mosurski, *Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed gene.* Nucleic Acids Res, 1986. **14**(13): p. 5125-5143.
15. Gustafsson, C., S. Govindarajan, and J. Minshull, *Codon bias and heterologous protein expression.* Trends Biotechnol, 2004. **22**(7): p. 346-53.
16. Redden, H., N. Morse, and H.S. Alper, *The synthetic biology toolbox for tuning gene expression in yeast.* FEMS Yeast Res, 2015. **15**(1): p. 1-10.
17. Sun, J., et al., *Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in Saccharomyces cerevisiae.* Biotechnol Bioeng, 2012. **109**(8): p. 2082-92.

18. Yamanishi, M., et al., *A genome-wide activity assessment of terminator regions in Saccharomyces cerevisiae provides a ''terminatome'' toolbox.* ACS Synth Biol, 2013. **2**(6): p. 337-47.
19. Partow, S., et al., *Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae.* Yeast, 2010. **27**(11): p. 955-64.
20. Curran, K.A., et al., *Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications.* Metab Eng, 2013. **19**: p. 88-97.
21. Shi, S., et al., *A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae.* Metab Eng, 2016. **33**: p. 19-27.
22. Clare, J.J., et al., *High-level expression of tetanus toxin fragment C in Pichia pastoris strains containing multiple tandem integrations of the gene.* Biotechnology (N Y), 1991. **9**(5): p. 455-60.
23. Oliveira, C., et al., *Development of stable flocculent Saccharomyces cerevisiae strain for continuous Aspergillus niger beta-galactosidase production.* J Biosci Bioeng, 2007. **103**(4): p. 318-24.
24. Vassileva, A., et al., *Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast Pichia pastoris.* Protein Expr Purif, 2001. **21**(1): p. 71-80.
25. Lee, W. and N.A. Dasilva, *Application of sequential integration for metabolic engineering of 1,2-propanediol production in yeast.* Metab Eng, 2006. **8**(1): p. 58-65.
26. Xie, W., et al., *Construction of lycopene-overproducing Saccharomyces cerevisiae by combining directed evolution and metabolic engineering.* Metab Eng, 2015. **30**: p. 69-78.
27. Yamada, R., et al., *Gene copy number and polyploidy on products formation in yeast.* Appl Microbiol Biotechnol, 2010. **88**(4): p. 849-57.
28. Ryan, O.W., et al., *Selection of chromosomal DNA libraries using a multiplex CRISPR system.* Elife, 2014. **3**.
29. Karim, A.S., K.A. Curran, and H.S. Alper, *Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications.* FEMS Yeast Res, 2013. **13**(1): p. 107-16.
30. Oldenburg, K.R., et al., *Recombination-mediated PCR-directed plasmid construction in vivo in yeast.* Nucleic Acids Res, 1997. **25**(2): p. 451-2.
31. Storici, F., L.K. Lewis, and M.A. Resnick, *In vivo site-directed mutagenesis using oligonucleotides.* Nat Biotechnol, 2001. **19**(8): p. 773-6.
32. Shao, Z., H. Zhao, and H. Zhao, *DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways.* Nucleic Acids Res, 2009. **37**(2): p. e16.
33. Orr-Weaver, T.L., J.W. Szostak, and R.J. Rothstein, *Yeast transformation: a model system for the study of recombination.* Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6354-8.
34. Storici, F., et al., *Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14994-9.
35. Storici, F. and M.A. Resnick, *Delitto perfetto targeted mutagenesis in yeast with oligonucleotides.* Genet Eng (N Y), 2003. **25**: p. 189-207.

36. Kim, Y.G., J. Cha, and S. Chandrasegaran, *Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1156-60.
37. Christian, M., et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
38. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
39. DiCarlo, J.E., et al., *Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems.* Nucleic Acids Res, 2013. **41**(7): p. 4336-43.
40. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.
41. Brouns, S.J., et al., *Small CRISPR RNAs guide antiviral defense in prokaryotes.* Science, 2008. **321**(5891): p. 960-4.
42. Jiang, W., et al., *RNA-guided editing of bacterial genomes using CRISPR-Cas systems.* Nat Biotechnol, 2013. **31**(3): p. 233-9.
43. Li, J.F., et al., *Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9.* Nat Biotechnol, 2013. **31**(8): p. 688-91.
44. Feng, Z., et al., *Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis.* Proc Natl Acad Sci U S A, 2014. **111**(12): p. 4632-7.
45. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
46. Jao, L.E., S.R. Wente, and W. Chen, *Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system.* Proc Natl Acad Sci U S A, 2013. **110**(34): p. 13904-9.
47. Wang, H., et al., *One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **153**(4): p. 910-8.
48. Gasiunas, G., et al., *Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. E2579-86.
49. Sternberg, S.H., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-7.
50. Mali, P., K.M. Esvelt, and G.M. Church, *Cas9 as a versatile tool for engineering biology.* Nat Methods, 2013. **10**(10): p. 957-63.
51. Ryan, O.W. and J.H. Cate, *Multiplex engineering of industrial yeast genomes using CRISPRm.* Methods Enzymol, 2014. **546**: p. 473-89.
52. Bao, Z., et al., *Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in Saccharomyces cerevisiae.* ACS Synth Biol, 2015. **4**(5): p. 585-94.
53. Horwitz, A.A., et al., *Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas.* Cell Syst, 2015. **1**(1): p. 88-96.
54. Ronda, C., et al., *CrEdit: CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae.* Microb Cell Fact, 2015. **14**: p. 97.
55. Jakociunas, T., et al., *CasEMBLR: Cas9-Facilitated Multiloci Genomic Integration of in Vivo Assembled DNA Parts in Saccharomyces cerevisiae.* ACS Synth Biol, 2015. **4**(11): p. 1226-34.

56. Jakociunas, T., et al., *Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae.* Metab Eng, 2015. **28**: p. 213-22.
57. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
58. Wang, T., et al., *Genetic screens in human cells using the CRISPR-Cas9 system.* Science, 2014. **343**(6166): p. 80-4.
59. Perez, A.R., et al., *GuideScan software for improved single and paired CRISPR guide RNA design.* Nat Biotechnol, 2017. **35**(4): p. 347-349.
60. Labuhn, M., et al., *Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications.* Nucleic Acids Res, 2018. **46**(3): p. 1375-1385.
61. Flagfeldt, D.B., et al., *Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae.* Yeast, 2009. **26**(10): p. 545-51.
62. Mikkelsen, M.D., et al., *Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform.* Metab Eng, 2012. **14**(2): p. 104-11.
63. Hagel, J.M. and P.J. Facchini, *Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world.* Plant Cell Physiol, 2013. **54**(5): p. 647-72.
64. Narcross, L., et al., *Microbial Factories for the Production of Benzylisoquinoline Alkaloids.* Trends Biotechnol, 2016. **34**(3): p. 228-41.
65. Trenchard, I.J., et al., *De novo production of the key branch point benzylisoquinoline alkaloid reticuline in yeast.* Metab Eng, 2015. **31**: p. 74-83.
66. Fossati, E., et al., *Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae.* Nat Commun, 2014. **5**: p. 3283.
67. Galanie, S., et al., *Complete biosynthesis of opioids in yeast.* Science, 2015. **349**(6252): p. 1095-100.
68. Fossati, E., et al., *Synthesis of Morphinan Alkaloids in Saccharomyces cerevisiae.* PLoS One, 2015. **10**(4): p. e0124459.
69. Hawkins, K.M. and C.D. Smolke, *Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae.* Nat Chem Biol, 2008. **4**(9): p. 564-73.
70. Samanani, N. and P.J. Facchini, *Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy.* Planta, 2001. **213**(6): p. 898-906.
71. Samanani, N. and P.J. Facchini, *Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzylisoquinoline alkaloid biosynthesis in plants.* J Biol Chem, 2002. **277**(37): p. 33878-83.
72. Samanani, N., D.K. Liscombe, and P.J. Facchini, *Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis.* Plant J, 2004. **40**(2): p. 302-13.
73. Lichman, B.R., et al., *'Dopamine-first' mechanism enables the rational engineering of the norcoclaurine synthase aldehyde activity profile.* FEBS J, 2015. **282**(6): p. 1137-51.
74. Li, J., et al., *Genes encoding norcoclaurine synthase occur as tandem fusions in the Papaveraceae.* Sci Rep, 2016. **6**: p. 39256.
75. DeLoache, W.C., et al., *An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose.* Nat Chem Biol, 2015. **11**(7): p. 465-71.

76. Jansen, G., et al., *Drag&Drop cloning in yeast.* Gene, 2005. **344**: p. 43-51.
77. Villesen, P., *FaBox: an online toolbox for fasta sequences.* Molecular Ecology Notes, 2007. **7**(6): p. 4.
78. Gietz, R.D. and R.H. Schiestl, *High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method.* Nat Protoc, 2007. **2**(1): p. 31-4.
79. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.* Mol Biol Evol, 2016. **33**(7): p. 1870-4.
80. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
81. Fisher, A.K., et al., *A review of metabolic and enzymatic engineering strategies for designing and optimizing performance of microbial cell factories.* Comput Struct Biotechnol J, 2014. **11**(18): p. 91-9.
82. Stovicek, V., I. Borodina, and J. Forster, *CRISPR–Cas system enables fast and simple genome editing of industrial Saccharomyces cerevisiae strains.* Metabolic Engineering Communications, 2015. **2**: p. 13-22.
83. Heim, R., D.C. Prasher, and R.Y. Tsien, *Wavelength mutations and posttranslational autoxidation of green fluorescent protein.* Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12501-4.
84. De Wulf, P., et al., *Real-time flow cytometric quantification of GFP expression and Gfp-fluorescence generation in Saccharomyces cerevisiae.* J Microbiol Methods, 2000. **42**(1): p. 57-64.
85. Soboleski, M.R., J. Oaks, and W.P. Halford, *Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells.* FASEB J, 2005. **19**(3): p. 440-2.
86. Stadler, R., T.M. Kutchan, and M.H. Zenk, *(S)-Norcoclaurine Is the Central Intermediate in Benzylisoquinoline Alkaloid Biosynthesis.* Phytochemistry, 1989. **28**(4): p. 1083-1086.
87. Strucko, T., O. Magdenoska, and U.H. Mortensen, *Benchmarking two commonly used Saccharomyces cerevisiae strains for heterologous vanillin-β-glucoside production.* Metabolic Engineering Communications, 2015. **2**: p. 99-108.
88. Xiao, M., et al., *Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest.* J Biotechnol, 2013. **166**(3): p. 122-34.
89. Minami, H., et al., *Functional analysis of norcoclaurine synthase in Coptis japonica.* J Biol Chem, 2007. **282**(9): p. 6274-82.
90. Galanie, S. and C.D. Smolke, *Optimization of yeast-based production of medicinal protoberberine alkaloids.* Microb Cell Fact, 2015. **14**: p. 144.
91. Hohenblum, H., et al., *Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris.* Biotechnol Bioeng, 2004. **85**(4): p. 367-75.
92. Lian, J., R. Jin, and H. Zhao, *Construction of plasmids with tunable copy numbers in Saccharomyces cerevisiae and their applications in pathway optimization and multiplex genome integration.* Biotechnol Bioeng, 2016. **113**(11): p. 2462-73.
93. Sakai, A., Y. Shimizu, and F. Hishinuma, *Integration of heterologous genes into the chromosome of Saccharomyces cerevisiae using a delta sequence of yeast retrotransposon Ty.* Appl Microbiol Biotechnol, 1990. **33**(3): p. 302-6.

94. Wang, X., Z. Wang, and N.A. Da Silva, *G418 Selection and stability of cloned genes integrated at chromosomal delta sequences of Saccharomyces cerevisiae.* Biotechnol Bioeng, 1996. **49**(1): p. 45-51.
95. Storici, F. and M.A. Resnick, *The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast.* Methods Enzymol, 2006. **409**: p. 329-45.
96. Kuijpers, N.G., et al., *One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in Saccharomyces cerevisiae.* FEMS Yeast Res, 2013. **13**(8): p. 769-81.
97. Tsaponina, O. and J.E. Haber, *Frequent Interchromosomal Template Switches during Gene Conversion in S. cerevisiae.* Mol Cell, 2014. **55**(4): p. 615-25.
98. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.
99. Mans, R., et al., *CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae.* FEMS Yeast Res, 2015. **15**(2).
100. Shi, S., et al., *Engineering of chromosomal wax ester synthase integrated Saccharomyces cerevisiae mutants for improved biosynthesis of fatty acid ethyl esters.* Biotechnol Bioeng, 2014. **111**(9): p. 1740-7.
101. Gold, N.D., et al., *Metabolic engineering of a tyrosine-overproducing yeast platform using targeted metabolomics.* Microb Cell Fact, 2015. **14**: p. 73.

APPENDIX

LIST OF APPENDIX TABLES

LIST OF APPENDIX FIGURES

STRAIN COLLECTION

Table A1. Strains used in this work

Group	Strain Name	Integration		Background	Source/ AN
		Integrant	Locus		
Parent strain	CENPK2-1D	–	–	*MATα; leu2-3,112 ; ura3-52 ; his3-Δ1; trp1-289; MAL2-8C; SUC2*	EUROSCARF
LP1.T$_X$	LP1.T1	LP1.T1	FgF20	CENPK2-1D	This study
	LP1.T2	LP1.T2	FgF20	CENPK2-1D	This study
	LP1.T3	LP1.T3	FgF20	CENPK2-1D	This study
	LP1.T4	LP1.T4	FgF20	CENPK2-1D	This study
	LP1.T5	LP1.T5	FgF20	CENPK2-1D	This study
	LP1.T6	LP1.T6	FgF20	CENPK2-1D	This study
	LP1.T7	LP1.T7	FgF20	CENPK2-1D	This study
	LP1.T8	LP1.T8	FgF20	CENPK2-1D	This study
	LP1.T9	LP1.T9	FgF20	CENPK2-1D	This study
	LP1.T10	LP1.T10	FgF20	CENPK2-1D	This study
LP1.T3. Site$_X$	LP1.T3-FgF1	LP1.T3	FgF1	CENPK2-1D	This study
	LP1.T3-FgF5	LP1.T3	FgF5	CENPK2-1D	This study
	LP1.T3-FgF7	LP1.T3	FgF7	CENPK2-1D	This study
	LP1.T3-FgF8	LP1.T3	FgF8	CENPK2-1D	This study
	LP1.T3-FgF11	LP1.T3	FgF11	CENPK2-1D	This study
	LP1.T3-FgF12	LP1.T3	FgF12	CENPK2-1D	This study
	LP1.T3-FgF14	LP1.T3	FgF14	CENPK2-1D	This study
	LP1.T3-FgF16	LP1.T3	FgF16	CENPK2-1D	This study
	LP1.T3-FgF18	LP1.T3	FgF18	CENPK2-1D	This study
	LP1.T3-FgF19	LP1.T3	FgF19	CENPK2-1D	This study
	LP1.T3-FgF20	LP1.T3	FgF20	CENPK2-1D	This study
	LP1.T3-FgF21	LP1.T3	FgF21	CENPK2-1D	This study
	LP1.T3-FgF22	LP1.T3	FgF22	CENPK2-1D	This study
	LP1.T3-FgF23	LP1.T3	FgF23	CENPK2-1D	This study
	LP1.T3-FgF24	LP1.T3	FgF24	CENPK2-1D	This study
	LP1.T3-UserXII-1	LP1.T3	UserXII-1	CENPK2-1D	This study
LP1.*gfp*. Site$_X$	LP1. *gfp* -FgF1	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF1	This study
	LP1. *gfp* -FgF5	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF5	This study
	LP1. *gfp* -FgF7	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF7	This study
	LP1. *gfp* -FgF8	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF8	This study
	LP1. *gfp* -FgF11	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF11	This study
	LP1. *gfp* -FgF12	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF12	This study
	LP1. *gfp* -FgF14	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF14	This study
	LP1. *gfp* -FgF16	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF16	This study
	LP1. *gfp* -FgF18	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF18	This study
	LP1. *gfp* -FgF19	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF19	This study
	LP1. *gfp* -FgF20	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF20	This study
	LP1. *gfp* -FgF21	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF21	This study
	LP1. *gfp* -FgF22	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF22	This study
	LP1. *gfp* -FgF23	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF23	This study
	LP1. *gfp* -FgF24	pTDH3- *gfp*^{S65T} -tCYC1	LP1.T3	LP1.T3-FgF24	This study
	LP1. *gfp* -UserXII-1	pTDH3-*gfp*^{S65T}-tCYC1	LP1.T3	LP1.T3-UserXII-1	This study
LP platform	CEN.10xLP	LP1; LP2; LP3; LP4	FgF20 FgF18; FgF24 UserXII-1; FgF7; FgF19 FgF12; FgF16; FgF21; FgF22;	CENPK2-1D	This study
***gfp* copy number**	LP1-*gfp*	pTDH3- *gfp*^{S65T} -tCYC1	LP1 [FgF20]	CEN.10xLP	This study
	LP2- *gfp*	pTDH3- *gfp*^{S65T} -tCYC1	LP2 [FgF18; FgF24]	CEN.10xLP	This study
	LP3- *gfp*	pTDH3- *gfp*^{S65T} -tCYC1	LP3 [UserXII-1; FgF7; FgF19]	CEN.10xLP	This study
	LP3- *gfp*	pTDH3- *gfp*^{S65T} -tCYC1	LP4 [FgF12; FgF16; FgF21; FgF22]	CEN.10xLP	This study
Dopamine Production	yWCD745	*pTDH3-CYP76AD1_W13L_F309L-tTDH1-pCCW12-DODC-tADH1-pPGK1-ARO4_FBR-tPGK1*	URA3	BY4741/MATα; Met- His-	[75] AN:KR232306
	CEN.10xLP.D		ARO4	CEN.10xLP	This study
LP-*NCS* Strains	LP1-*Am*NCS	*Am*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Am*NCS	*Am*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Am*NCS	*Am*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Am*NCS	*Am*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Ec*NCS	*Ec*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Ec*NCS	*Ec*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Ec*NCS	*Ec*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Ec*NCS	*Ec*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Nd*NCS	*Nd*NCS	LP1 {FgF20}	CEN.10xLP.D	This study

	LP2-*Nd*NCS	*Nd*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Nd*NCS	*Nd*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Nd*NCS	*Nd*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Ps*NCS1	*Ps*NCS1	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Ps*NCS1	*Ps*NCS1	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Ps*NCS1	*Ps*NCS1	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Ps*NCS1	*Ps*NCS1	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Ps*NCS2	*Ps*NCS2	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Ps*NCS2	*Ps*NCS2	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Ps*NCS2	*Ps*NCS2	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Ps*NCS2	*Ps*NCS2	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Ps*NCS3	*Ps*NCS3	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Ps*NCS3	*Ps*NCS3	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Ps*NCS3	*Ps*NCS3	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Ps*NCS3	*Ps*NCS3	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Sc*NCS	*Sc*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Sc*NCS	*Sc*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Sc*NCS	*Sc*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Sc*NCS	*Sc*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Sd*NCS	*Sd*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Sd*NCS	*Sd*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Sd*NCS	*Sd*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Sd*NCS	*Sd*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Tc*NCS	*Tc*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Tc*NCS	*Tc*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Tc*NCS	*Tc*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Tc*NCS	*Tc*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study
	LP1-*Tf*NCS	*Tf*NCS	LP1 {FgF20}	CEN.10xLP.D	This study
	LP2-*Tf*NCS	*Tf*NCS	LP2 {FgF18; FgF24}	CEN.10xLP.D	This study
	LP3-*Tf*NCS	*Tf*NCS	LP3 {UserXII-1; FgF7; FgF19}	CEN.10xLP.D	This study
	LP4-*Tf*NCS	*Tf*NCS	LP4 {FgF12; FgF16; FgF21; FgF22}	CEN.10xLP.D	This study

PLASMID COLLECTION

Table A2. Plasmids used in this work

Plasmid	Relevant Genotype	Source / AN
pCas-Tyr	pRNR2-Cas9_NLS-tCYC1-pUC-2µOri-tRNA_Tyr-3'HDV-gRNAx-Scaffold-tSNR52-pTEF1-KanMX-tTEF1	Ryan et al. (2014) Addgene #60847
pJet1.2_LP1	LP1.A-LP1.Z	This study
pCas-Tyr-NKgRNA	pRNR2-Cas9_NLS-tCYC1-pUC-2µOri-tRNA_Tyr-3'HDV-NKgRNA-Scaffold-tSNR52-pTEF1-KanMX-tTEF1	This study
pCas-Tyr-gRNA1	gRNA1 is targeted to T1	This study
pCas-Tyr-gRNA2	gRNA2 is targeted to T2	This study
pCas-Tyr-gRNA3	gRNA3 is targeted to T3	This study
pCas-Tyr-gRNA4	gRNA4 is targeted to T4	This study
pCas-Tyr-gRNA5	gRNA5 is targeted to T5	This study
pCas-Tyr-gRNA6	gRNA6 is targeted to T6	This study
pCas-Tyr-gRNA7	gRNA7 is targeted to T7	This study
pCas-Tyr-gRNA8	gRNA8 is targeted to T8	This study
pCas-Tyr-gRNA9	gRNA9 is targeted to T9	This study
pCas-Tyr-gRNA10	gRNA10 is targeted to T10	This study
GC975	CEN6/ARS4ori, pMB1ori, HIS3, bla, loxP-kanMX, pTDH3-GFP_S65T-TCYC1	modified from Jansen et al.
pWCD2249	pTDH3-CYP76AD1_W13L_F309L-tTDH1-pCCW12-DODC-tADH1-pPGK1-ARO4_FBR-tPGK1	Deloache et al. (2015) KR232306.1
pBOT(his3)_ *Am*NCS	pTEF1-*Am*NCS-tPGI1	This study
pBOT(his3)_ *Ec*NCS	pTEF1-*Ec*NCS-tPGI1	This study
pBOT(his3)_ *Nd*NCS	pTEF1-*Nd*NCS-tPGI1	This study
pBOT(his3)_ *Ps*NCS1	pTEF1-*Ps*NCS1-tPGI1	This study
pBOT(his3)_ *Ps*NCS2	pTEF1-*Ps*NCS2-tPGI1	This study
pBOT(his3)_ *Ps*NCS3	pTEF1-*Ps*NCS3-tPGI1	This study
pBOT(his3)_ *Sc*NCS	pTEF1-*Sc*NCS-tPGI1	This study
pBOT(his3)_ *Sd*NCS	pTEF1-*Sd*NCS-tPGI1	This study
pBOT(his3)_ *Tc*NCS	pTEF1-*Tc*NCS-tPGI1	This study
pBOT(his3)_ *Tf*NCS	pTEF1-*Tf*NCS-tPGI1	This study

LANDING PAD CONSTRUCTS

Table A3. Landing Pad sequences with assigned target sites

ID	Sequence LP.A (blue) | Target site (grey) | LP.Z (purple)
LP1.T8	CGAGAACCTCAGTTACTTTTCCTTATGGAGGTCAGGAGAGGGGAACACGGACAAAGCTACGCGCTGCC TTTCAGAAATGATCGAGTCAGCGACCTTTTGCCTTATTCACGCTAGAGGTGGGTGTAATTTTGATTCTGG GCTCACATAAAACTCTCCCCATCGCTCTGGATCTGTAAGAATCGTCACTGGGAACAGTCGTACTAGTTC GGGTAATTGGCGCGTGACTGTCAGCGCAATCCGAGGAATACTCTGAATAAAACAACTTATATAATAAAAA TGCTAAGTCCGCGGATAACCATTCGGCCGTTCTTTATATTACATCAAAATAAGAAAATAATTATAACAGAC GCTAACAAGGCGGCCTACATCTCTGTCGCCGATCATATACTGTACGTGATGTGTCCCTCTTCGCGGGTC GAGGCGATACACGGACTTTCGACGTTAATCATATCGTATAAGTAGGTAGACATTGATCATGGTGAACGG TCGAACGTGCATATCACTACAGGACTCACCATGGTATCTTCGTTCAGCTCCAGCGTATGCACATTTGACC TCCAGCCTTGAAGCCGTGTCCATAAACAA
LP2.T10	TGCTCCAAGTGTGTGACTCCTTCATCTGACAACGTGCAACCCCTATCGCCATCGATTGTTTCTGCGGAC GGTGTTGTCCTCATAGTTTGGGCATGTTTCCCTTGTAGGTGTGAAACCACTTAGCTTCGCGCCGTAGTC CTAAAGGAAAACCTATGGACTTTGTTTCGGGTAGCACCAGGAATCTGAACCATGTGAATGTGGACGTGG CGCGCGTACACCTTAATCTCCGGTTCATGCTAGGGATGTGGCTGCATGCTACGTTGACACACCTACACT GCTCTCCCCAATCGTGGAGTGAAGCGGGAAGTAGCGAGATGAAGTGTACGACCTGGCCGGAGCCGTT CCGCATCGTCACGTGTTCGTTTACTGTTAATTGGTGGCACATAAGCAATATCGTAGTCCGTCAAATTCAG CCCTGTTATCCCCGGCGTTATGTGTCAAATGGCGTAGAACTGGATTGACTGTTTGACGGTACCTGCTGA TCGGTACGGTGACCGAGAATCTGTCGGGCTATGTCACTAATACTTTCCAAACGCCCCGTATCGATGCTG AACGAATCGATGCACACTCACGTCTTTGAAGC
LP3.T7	CGCATAGACATACAAGTGGACAGATGATGGGTACGGGCCTCTAATACATCCAACACTCTACGCCCTCTT CAAGAGCTAGAAGGGCACCCTGCAGTTGGAAAGGGAATTATTTCGTAAGGCGAGCCCATACCGTCATTC ATGCGGAAGAGTTAACACGATTGGAAGTAGGAATAGTTTCGAACCACGGTTACTAATCCTAATAACGGA ACGCTGTCTGAAGGATGAGTGTCAGCGAGTGTAACTCGATGAGCTACCCAGTAGTCGTACTGGTCGAG ACAACATTGTACCCCAGCGGCGGCGCGGCCAGCTCTAATGCACTCAATCCCGAGGCCTGACGCGACAT ATCAGCTTAGACTAGGGCGGGGGTGTTGACGTTTGGGGTTGAATAAATCTATTGTACTAATCGGCTTCA ACGTGCCCCACGGGTGGCACCTCAGGAGGGGCCCACAGCGAGGAAGTAAACTGTTATTCGTCGGCGA TGGTGGTAGCTAATTATGTTCCTTGCCACTACAATAGTATCTAAGCCGTGTAATGGGAACATCCACACTT TAGTGAATCGATGTGCAGCTTCAGAATACCATTC
LP4.T9	CAATGGAGCTCGCAATACAGAGTTTACCGCATCTTGCCCTAACTGACAAACTGTGATCGACCACAAGCC AAGCCATTGCCTCTTAGACACGCCGTTACAGTGATTATGAAAACTTTGCGGGGCATGGCTACGACTTGT TCAGCCACGTCCGAGGGCAGAAACCTATCCCCATTTGTATGTTCAGCTATCTTCTACCCATCCCCGGAG GTTAAGTAGGTTGTGAGATGCGGAAGAGGCTCTCGATCATCCCGTGGGACATCAACCTTTCCCTTGATA AAGCGTAGCCCAACAGGAGCACATCGGACCCCGCTCGGGTATGGCAGAGAGAACGCCTTCTGAATTGT GCTATCCTTCGACCTTATCAAAGCTTGCTACCAATAATTAGGATTATTGCCTTGCGACAGACCTCCTACT CACACTGCCTCACATTGAGCTAGTCAGTGAGCGATTAGCTTGACCCGCTCTCTAGGGTCGCGAGTACGT GAGCTAGGGCTCCGGACTGGGCTATATAGTCGAGTCTGATCTCGCCCCGACAACTGCAAACCTCAACTT TTTTAGATAACATGGTTAGCCGAAGTTGCAC

PRIMERS USED IN THIS WORK

Table A4. Primers to linearize pCAS backbone & gRNA cassette

ID	Sequence	Description
GC2946_pCAS_F	TAGGTCTAGAGATCTGTTTAGCTTG	Amplifies pCAS vector backbone from 5' RNR2p
GC2947_pCAS_R	GCATTTAAGCATAAACACGC	Amplifies pCAS vector backbone from 3' 2μ ori
LB33_HDV_gRNA_F	CACCTATATCTGCGTGTTGC	Universal gRNA forward
GC2894_HDV_gRNA_R	AAAGTCCCATTCGCCACC	Amplifies HDV blunt left
GC2895_gRNA_scaffold_F	GTTTTAGAGCTAGAAATAGCAAGT	Amplifies scaffold blunt right
LB34_gRNA_scaffold_R	GTCAAGACTGTCAAGGAGG	Universal gRNA reverse

Table A5. Primers to program synthetic $gRNA_x$ construct into pCAS

ID	Sequence	Description
T1_gRNA1_F	CTAGATAAGACGTGGCAGATGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA1 5' overhang
T2_gRNA1_R	ATCTGCCACGTCTTATCTAGAAAGTCCCATTCGCC	HDV_R with gRNA1 3' overhang
T3_gRNA2_F	CCACCGCTCAGTAGCGGCTTGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA2 5' overhang
T4_gRNA2_R	AAGCCGCTACTGAGCGGTGGAAAGTCCCATTCGCC	HDV_R with gRNA2 3' overhang
T5_gRNA3_F	GCCAGTCAGAACACTAGAGGGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA3 5' overhang
T6_gRNA3_R	CCTCTAGTGTTCTGACTGGCAAAGTCCCATTCGCC	HDV_R with gRNA3 3' overhang
T7_gRNA4_F	ACGCCGTTTTTCACATCTGTGTTTTAGAGCTAGAAAT AGCAAGT	Scaffold_F with gRNA4 5' overhang
T8_gRNA4_R	ACAGATGTGAAAAACGGCGTAAAGTCCCATTCGCC	HDV_R with gRNA4 3' overhang
T9_gRNA5_F	GATACAACCCATCCGCGCTAGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA5 5' overhang
T10_gRNA5_R	TAGCGCGGATGGGTTGTATCAAAGTCCCATTCGCC	HDV_R with gRNA5 3' overhang
T11_gRNA6_F	ACGTGTCATACGAGGTATAGGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA6 5' overhang
T12_gRNA6_R	CTATACCTCGTATGACACGTAAAGTCCCATTCGCC	HDV_R with gRNA6 3' overhang
T13_gRNA7_F	ATTGTACCCCAGCGGCGGCGGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA7 5' overhang
T14_gRNA7_R	CGCCGCCGCTGGGGTACAATAAAGTCCCATTCGCC	HDV_R with gRNA7 3' overhang
T15_gRNA8_F	TAAGTCCGCGGATAACCATTGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA8 5' overhang
T16_gRNA8_R	AATGGTTATCCGCGGACTTAAAAGTCCCATTCGCC	HDV_R with gRNA8 3' overhang
T17_gRNA9_F	GTAGCCCAACAGGAGCACATGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA9 5' overhang
T18_gRNA9_R	ATGTGCTCCTGTTGGGCTACAAAGTCCCATTCGCC	HDV_R with gRNA9 3' overhang
T19_gRNA_10_F	TCCCCAATCGTGGAGTGAAGGTTTTAGAGCTAGAAA TAGCAAGT	Scaffold_F with gRNA10 5' overhang
T20_gRNA_10_R	CTTCACTCCACGATTGGGGAAAAGTCCCATTCGCC	HDV_R with gRNA10 3' overhang

Table A6. Primers to program the gRNAs targeting genomic loci for LP1. T3 integrations

ID	Sequence	5'-N_{20}gRNA-3'	Locus/ gRNA
FF256_FgF1.gRNA.s2_3'HDV_R	CTCCTGTTTAATATGCTGTCaaagtcccattcgcc	CTCCTGTTTAATATGCTGTC	FgF1 Sh2
FF257_FgF1.gRNA.s2_scaff_F	GACAGCATATTAAACAGGAGgttttagagctagaaatagcaagt		
FF258_FgF5.gRNA.s2_3'HDV_R	ATGAAATAAAATCCTGATATaaagtcccattcgcc	ATGAAATAAAATCCTGATAT	FgF5 Sh2
FF259_FgF5.gRNA.s2_scaff_F	ATATCAGGATTTTATTTCATgttttagagctagaaatagcaagt		
FF124_FgF7.gRNA.1_scaff_F	ACTCCTGGGAGAGAACATTCgttttagagctagaaatagcaagt	GAATGTTCTCTCCCAGGAGT	FgF7 #1
FF125_FgF7.gRNA.1_3-HDV_R	GAATGTTCTCTCCCAGGAGTaaagtcccattcgcc		
FF126_FgF7.gRNA.2_scaff_F	TCTTCTTATACCGTATATGAgttttagagctagaaatagcaagt	TCATATACGGTATAAGAAGA	FgF7 #2
FF127_FgF7.gRNA.2_3'HDV_R	TCATATACGGTATAAGAAGAaaagtcccattcgcc		
FF128_FgF8.gRNA.1_scaff_F	AGAGAAATTAGAATATTATGgttttagagctagaaatagcaagt	CATAATATTCTAATTTCTCT	FgF8 #1
FF129_FgF8.gRNA.1_3'HDV_R	CATAATATTCTAATTTCTCTaaagtcccattcgcc		
FF130_FgF8.gRNA.2_scaff_F	GTTGATAGTTAAAAGTTAACgttttagagctagaaatagcaagt	CATAATATTCTAATTTCTCT	FgF8 #2
FF131_FgF8.gRNA.2_3'HDV_R	GTTAACTTTTAACTATCAACaaagtcccattcgcc		
FF260_FgF11.gRNA.s2_3'HDV_R	GTTTAATATTCTGTCCTTAAaaagtcccattcgcc	GTTTAATATTCTGTCCTTAA	FgF11 #Sh2
FF261_FgF11.gRNA.s2_scaff_F	TTAAGGACAGAATATTAAACgttttagagctagaaatagcaagt		
FF216_FgF12.gRNA.1_3'HDV_R	CTTCTACTTTAATAAGTACAaaagtcccattcgccacc	CTTCTACTTTAATAAGTACA	FgF12 #1
FF217_FgF12.gRNA.1_scaff_F	TGTACTTATTAAAGTAGAAGgttttagagctagaaatagcaagt		
FF218_FgF12.gRNA.2_3'HDV_R	TGTTGTGGTTTCCCATACCCaaagtcccattcgccacc	TGTTGTGGTTTCCCATACCC	FgF12 #2
FF219_FgF12.gRNA.2_scaff_F	GGGTATGGGAAACCACAACAgttttagagctagaaatagcaagt		
FF140_FgF14.gRNA.1_scaff_F	ATCATTGTGATTTTATTTGTgttttagagctagaaatagcaagt	ACAAATAAAATCACAATGAT	FgF14 #1
FF141_FgF14.gRNA.1_3'HDV_R	ACAAATAAAATCACAATGATaaagtcccattcgcc		
FF240_FgF16.gRNA.s1_3'HDV_R	GTAGGATATAGGAATCTACGaaagtcccattcgcc	GTAGGATATAGGAATCTACG	FgF16 #Sh1
FF241_FgF16.gRNA.s1_scaff_F	CGTAGATTCCTATATCCTACgttttagagctagaaatagcaagt		
FF146_FgF18.gRNA.2_scaff_F	ATAGAATTACTATTGAAGAGgttttagagctagaaatagcaagt	CTCTTCAATAGTAATTCTAT	FgF18 #2
FF147_FgF18.gRNA.2_3'HDV_R	CTCTTCAATAGTAATTCTATaaagtcccattcgcc		
FF150_FgF19.gRNA.2_scaff_F	ATTCACTCTGCTAAGATTATgttttagagctagaaatagcaagt	ATAATCTTAGCAGAGTGAAT	FgF19 #2
FF151_FgF19.gRNA.2_3'HDV_R	ATAATCTTAGCAGAGTGAATaaagtcccattcgcc		
FF152_FgF20.gRNA.1_scaff_F	AAAATTCTCTCTGAGGATATgttttagagctagaaatagcaagt	ATATCCTCAGAGAGAATTTT	FgF20 #1
FF153_FgF20.gRNA.1_3'HDV_R	ATATCCTCAGAGAGAATTTTaaagtcccattcgcc		
FF154_FgF20.gRNA.2_scaff_F	GTTAGAGCTGTTACAAGTTAgttttagagctagaaatagcaagt	TAACTTGTAACAGCTCTAAC	FgF20 #2
FF155_FgF20.gRNA.2_3'HDV_R	TAACTTGTAACAGCTCTAACaaagtcccattcgcc		
FF264_FgF21.gRNA.s2_3'HDV_R	ATTGGTTCAACATTAATTATaaagtcccattcgcc	ATTGGTTCAACATTAATTAT	FgF21 #Sh2
FF265_FgF21.gRNA.sho2_scaff_F	ATAATTAATGTTGAACCAATgttttagagctagaaatagcaagt		
FF160_FgF22.gRNA.1_scaff_F	AACAAACTTGTGTGCTTCATgttttagagctagaaatagcaagt	ATGAAGCACACAAGTTTGTT	FgF22 #1
FF161_FgF22.gRNA.1_3'HDV_R	ATGAAGCACACAAGTTTGTTaaagtcccattcgcc		
FF162_FgF22.gRNA.2_scaff_F	GGTGAACGTTACAGAAAAGCgttttagagctagaaatagcaagt	GCTTTTCTGTAACGTTCACC	FgF22 #2
FF163_FgF22.gRNA.2_3'HDV_R	GCTTTTCTGTAACGTTCACCaaagtcccattcgcc		

FF228_FgF23.gRNA.1_3'HDV_R	TAAGACGTATCGCAAAGAAGaaagtcccattcgccacc	TAAGACGTATCGCAAAGAAG	FgF23 #1
FF229_FgF23.gRNA.1_scaff_F	CTTCTTTGCGATACGTCTTAgttttagagctagaaatagcaagt		
FF230_FgF23.gRNA.2_3'HDV_R	AGGTATTTGGTAACAACAAGaaagtcccattcgccacc	AGGTATTTGGTAACAACAAG	FgF23 #2
FF231_FgF23.gRNA.2_scaff_F	CTTGTTGTTACCAAATACCTgttttagagctagaaatagcaagt		
FF166_FgF24.gRNA.1_scaff_F	CCTATTGGACAAGATTTACGgttttagagctagaaatagcaagt	CGTAAATCTTGTCCAATAGG	FgF24 #1
FF167_FgF24.gRNA.1_3'HDV_R	CGTAAATCTTGTCCAATAGGaaagtcccattcgcc		
FF168_FgF24.gRNA.2_scaff_F	ATTTCGTTGTACTCTGCGTGgttttagagctagaaatagcaagt	CACGCAGAGTACAACGAAAT	FgF24 #2
FF169_HDV_FgF24.2_R	CACGCAGAGTACAACGAAATaaagtcccattcgcc		
SB78_USERXII-1.gRNA2.S	GTCTTTGCCGGTTACCCATCgttttagagctagaaatagcaagt	GTCTTTGCCGGTTACCCATC	UserXII-1 2
SB79_USERXII-1.gRNA2.AS	GATGGGTAACCGGCAAAGACaaagtcccattcgccacc		
LB679-ARO4-gRNA11.S	GCTTAGTAACACCCATGAAAgttttagagctagaaatagcaagt	GCTTAGTAACACCCATGAAA	ARO4
LB680-ARO4-gRNA11.AS	TTTCATGGGTGTTACTAAGCaaagtcccattcgcc		

Table A7. Primers to amplify Landings Pads

ID	Sequence	Description
LB1_LP1.A_F	CGAGAACCTCAGTTACTTTTC	amplifies LP1.A at 5' end
LB4_LP1.Z_R	TTGTTTATGGACACGGC	amplifies LP1.Z at 3' end
FF346_LP2.A_F	TGCTCCAAGTGTGTGACTCC	amplifies LP2.A at 5' end
FF347_LP2.Z_R	GCTTCAAAGACGTGAGTGTGC	amplifies LP2.Z at 3' end
FF348_LP3.A_F	CGCATAGACATACAAGTGGACAG	amplifies LP3.A at 5' end
FF349_LP3.Z_R	GAATGGTATTCTGAAGCTGCAC	amplifies LP3.Z at 3' end
FF350_LP4.A_F	CAATGGAGCTCGCAATACAG	amplifies LP4.A at 5' end
FF351_LP4.Z_R	GTGCAACTTCGGCTAACC	amplifies LP4.Z at 3' end
SB41-FgF18.LP2.S:	ATGTGATGTTTAGAAGAAGATTCGAACTGTTTTCAGT AGATTTGGTAACTGTGCAACCATAACTCATGCCTGC TCCAAGTGTGTGACTCC	Amplifies LP2 with homology to FgF18
SB42-FgF18.LP2.AS:	CAACTGTTGCTCTTCCTATATGCATTTAAATGTGATG AATTTTGAGAGCCCACTTTTGTTGGGGACGATTGCT TCAAAGACGTGAGTGTG	
SB43-FgF24.LP2.S:	ACCCGTAGTGGCGATCTTGTGATTTTCGTACTTTTTC GCACCAAATGCGGGCAACAGGGCGAGGTGATCCTG CTCCAAGTGTGTGACTCC	Amplifies LP2 with homology to FgF24
SB44-FgF24.LP2.AS:	AAAATACTTATTTGAAAGATTGAAGCAAGTTAATGTT AACACCATTTTTGGGCTACCAGGCGACTTCAACGCT TCAAAGACGTGAGTGTG	
SB45-FgF7.LP3.S:	CAGAAACGTAGAAAAAGAAAAACAATTAAAACATTAT ATTAAGATTATTGATTTGCCTTTTAAGGGTCCACGCA TAGACATACAAGTGGA	Amplifies LP3 with homology to FgF7
SB46-FgF7.LP3.AS	GTTACAACAATACCTCTAATATTGCTTCTGCCGAAAT CAAATTATGATAAGAACCAATCATCATCCATCGGAAT GGTATTCTGAAGCTGC	
SB47-FgF19.LP3.S:	ATTCCGCGCTTCCATCATTTAGTATAATCCATATTTT ATATAATATATAGGATAAGTAACAGCCCGCGAA CGCATAGACATACAAGTGGA	Amplifies LP3 with homology to FgF19
SB48-FgF19.LP3.AS	TTGTTTTTTATCTTGCCCATTCACAGATCTCATTTGG TAATGTCAGTAATTAGCGGATGATAGTTGGTCCGAA TGGTATTCTGAAGCTGC	
SB74.USER12-1.LP3.S	GTAGATACAATAGCACATCTCATTACCCAGTTATGAT TGACGTCATTCTGAGTTACAATGATCTTACGCATAGA CATACAAGTGGA	Amplifies LP3 with homology to USERXII-1

SB75.USER12-1.LP3.AS	GAGCATTTTGAATTTTAAGGAAATTAATTTGTCTTCTA TCTTTTGCATCCGTGCCTCCTGGATAATAACTTCAAG AATGGTATTCTGAAGCTGC	
SB51-FgF12.LP5.S	AAAATTAGCGCCATATGTTGGTATGATGTCAATAGCA GTAGAAAAATATATAGACAAGAAAGACGGGGGACAA TGGAGCTCGCAATACAG	Amplifies LP4 with homology to FgF12
SB52-FgF12.LP5.AS	TTTAGCACATAATTTTCGATGTACTTTTCTTTTTAAAA TCGAATTATCAGCGATTATTCAGCCGGCTAAGGTGC AACTTCGGCTAACCAT	
SB53-FgF16.LP5.S	GAAAGAATTAAATATTCACTAGGCTGCGATACGATA GACAAACGAAGTGATTGAAACCCGAATTAACGGACA ATGGAGCTCGCAATACAG	Amplifies LP4 with homology to FgF16
SB54-FgF16.LP5.AS:	CTTGGTTATGCGTTATTTAAATCCTCATCTGCCGCTG CTTAAAAAAAGCAGCTAAAGTGTTGCGTAGGCAGTG CAACTTCGGCTAACCAT	
SB55-FgF21.LP5.S	TGTTCTCTTATGTATTTTTAATCGTCCTTGTATGGAA GTATCAAAGGGGACGTTCTTCACCTCCTTGGAACAA TGGAGCTCGCAATACAG	Amplifies LP4 with homology to FgF21
SB56-FgF21.LP5.AS	AAGAAAGAGATTAAATAAAGAATGATTTACAATCTAG TCGCAAAAACAAGTACAGTGCTGACGTCCCATCGTG CAACTTCGGCTAACCAT	
SB57-FgF22.LP5.S	CCTGCAGGAAACGAAGATAAATCATGTCGAAAGCTA CATATAAGGAACGTGCTGCTACTCATCCTAGTCCCA ATGGAGCTCGCAATACAG	Amplifies LP4 with homology to FgF22
SB58-FgF22.LP5.AS:	AGTATACATGCATTTACTTATAATACAGTTTTTTAGTT TTGCTGGCCGCATCTTCTCAAATATGCTTCCCGTGC AACTTCGGCTAACCAT	

Table A8. Primers to build LP1.T_X constructs

ID	Sequence	Description
target1_LP1.Z_F	CTAGATAAGACGTGGCAGATCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T1 overhang
LP1.A_target1_R	CCGATCTGCCACGTCTTATCTAGgcatttttattatataagttgttttatt	Amplifies LP1.A with T1 overhang
target2_LP1.Z_F	CCACCGCTCAGTAGCGGCTTCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T2 overhang
LP1.A_target2_R	CCGAAGCCGCTACTGAGCGGTGGgcatttttattatataagttgttttatt	Amplifies LP1.A with T2 overhang
target3_LP1.Z_F	GCCAGTCAGAACACTAGAGGCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T3 overhang
LP1.A_target3_R	CCGCCTCTAGTGTTCTGACTGGCgcatttttattatataagttgttttatt	Amplifies LP1.A with T3 overhang
target4_LP1.Z_F	ACGCCGTTTTTCACATCTGTCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T4 overhang
LP1.A_target4_R	CCGACAGATGTGAAAAACGGCGTgcatttttattatataagttgttttatt	Amplifies LP1.A with T4 overhang
target5_LP1.Z_F	GATACAACCCATCCGCGCTACGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T5 overhang
LP1.A_target5_R	CCGTAGCGCGGATGGGTTGTATCgcatttttattatataagttgttttatt	Amplifies LP1.A with T5 overhang
target6_LP1.Z_F	ACGTGTCATACGAGGTATAGCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T6 overhang
LP1.A_target6_R	CCGCTATACCTCGTATGACACGTgcatttttattatataagttgttttatt	Amplifies LP1.A with T6 overhang
target7_LP1.Z_F	ATTGTACCCCAGCGGCGGCGCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T7 overhang
LP1.A_target7_R	CCGCGCCGCCGCTGGGGTACAATgcatttttattatataagttgttttatt	Amplifies LP1.A with T7 overhang
target8_LP1.Z_F	TAAGTCCGCGGATAACCATTCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T8 overhang
LP1.A_target8_R	CCGAATGGTTATCCGCGGACTTAgcatttttattatataagttgttttatt	Amplifies LP1.A with T8 overhang
target9_LP1.Z_F	GTAGCCCAACAGGAGCACATCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T9 overhang
LP1.A_target9_R	CCGATGTGCTCCTGTTGGGCTACgcatttttattatataagttgttttatt	Amplifies LP1.A with T9 overhang
target10_LP1.Z_F	TCCCCAATCGTGGAGTGAAGCGGcctctttatattacatcaaaataaga	Amplifies LP1.Z with T10 overhang
LP1.A_target10_R	CCGCTTCACTCCACGATTGGGGAgcatttttattatataagttgttttatt	Amplifies LP1.A with T10 overhang

Table A9. Primers to construct LP1.T3.Site$_X$ donors

ID	Sequence	Site	Description
FF1_FgF16_LP1.A_F	ATATTCACTAGGCTGCGATACGATAGAC AAACGAAGTGATTGAAACCCGAATTAAC GGAcgagaacctcagttacttttc	16	5' overhang to FgF site, 3' homology to LP1.A
FF2_LP1.Z_FgF16_R	ATGCGTTATTTAAATCCTCATCTGCCGCT GCTTAAAAAAAGCAGCTAAAGTGTTGCG TAGGCAttgtttatggacacggc	16	5' homology to LP1.Z, 3' overhang to FgF site
FgF1.Up_LP1.A_F	AGCACAGTGAcgagaacctcagttacttttt	1	5' overhang to FgF site, 3' homology to LP1.A
FgF1.D_LP1.Z_R	GCCATATTTCTGAttgtttatggacacggc	1	5' homology to LP1.Z, 3' overhang to FgF site
FgF5.up_LP1.A_F	TAGAAGCAcgagaacctcagttacttttt	5	5' overhang to FgF site, 3' homology to LP1.A
FgF5.D_LP1.Z_R	CTTTTTCATGACGCGTTGTTTTCttgtttatgg acacggc	5	5' homology to LP1.Z, 3' overhang to FgF site
FgF7.up_LP1.A_F	AGGGTCCAcgagaacctcagttacttttt	7	5' overhang to FgF site, 3' homology to LP1.A
FgF7.D_LP1.Z_R	TAAGAACCAATCATCATCCATCGttgtttatgg acacggc	7	5' homology to LP1.Z, 3' overhang to FgF site
FgF8.up_LP1.Z_R	GTTGCTTGATTTGCCCTGTTTGTttgtttatgg acacggc	8	5' overhang to FgF site, 3' homology to LP1.A
FgF8.D_LP1.A_F	CAAGACAGCcgagaacctcagttacttttt	8	5' homology to LP1.Z, 3' overhang to FgF site
FgF11.up_LP1.A_F	CGTAAAAGTGcgagaacctcagttacttttt	11	5' overhang to FgF site, 3' homology to LP1.A
FgF11.D_LP1.Z_R	TAATAAGTACTCATttgtttatggacacggc	11	5' homology to LP1.Z, 3' overhang to FgF site
FgF12.up_LP1.Z_R	TCAGCCGGCTAAGttgtttatggacacggc	12	5' overhang to FgF site, 3' homology to LP1.A
FgF12.D_LP1.A_F	AGACGGGGGAcgagaacctcagttacttttt	12	5' homology to LP1.Z, 3' overhang to FgF site
FgF14.up_LP1.A_F	TCTTACACCcgagaacctcagttactttc	14	5' overhang to FgF site, 3' homology to LP1.A
FgF14.D_LP1.Z_R	CGGTTGCATTTTCttgtttatggacacggc	14	5' homology to LP1.Z, 3' overhang to FgF site
FgF18.up_LP1.A_F	AACTCATGCCcgagaacctcagttacttttt	18	5' overhang to FgF site, 3' homology to LP1.A
FgF18.D_LP1.Z_R	GGGGACGATTttgtttatggacacggc	18	5' homology to LP1.Z, 3' overhang to FgF site
FgF19.up_LP1.A_F	CGCGAAcgagaacctcagttacttttt	19	5' overhang to FgF site, 3' homology to LP1.A
FgF19.D_LP1.Z_R	TGATAGTTGGTCCttgtttatggacacggc	19	5' homology to LP1.Z, 3' overhang to FgF site
FgF20.up_LP1.A_F	ATCGCAAAcgagaacctcagttacttttt	20	5' overhang to FgF site, 3' homology to LP1.A
FgF20.D_LP1.Z_R	GACCTTCCATTttgtttatggacacggc	20	5' homology to LP1.Z, 3' overhang to FgF site
FgF21.up_LP1.A_F	CCTTGGAAcgagaacctcagttacttttt	21	5' overhang to FgF site, 3' homology to LP1.A
FgF21.D_LP1.Z_R	CTGACGTCCCATCttgtttatggacacggc	21	5' homology to LP1.Z, 3' overhang to FgF site
FgF22.up_LP1.A_F	ATCCTAGTCCcgagaacctcagttacttttt	22	5' overhang to FgF site, 3' homology to LP1.A
FgF22.D_LP1.Z_R	ATATGCTTCCCttgtttatggacacggc	22	5' homology to LP1.Z, 3' overhang to FgF site
FgF23.up_LP1.A_F	CCATTCCCTTcgagaacctcagttacttttt	23	5' overhang to FgF site, 3' homology to LP1.A
FgF23.D_LP1.Z_R	GTTTTTTGATGATttgtttatggacacggc	23	5' homology to LP1.Z, 3' overhang to FgF site
FgF24.up_LP1.Z_R	GCGACTTCAACttgtttatggacacggc	24	5' homology to LP1.Z, 3' overhang to FgF site
FgF24.D_LP1.A_F	GGTGATCCcgagaacctcagttacttttt	24	5' overhang to FgF site, 3' homology to LP1.A
USER12_LP1.A_F	TTTAGTAGATACAATAGCACATCTCATTA CCCAGTTATGATTGACGTCATTCTGAGTT ACAATGATCTTAcgagaacctcagttacttttt	U XII-1	5' overhang to FgF site, 3' homology to LP1.A
USER12_LP1.Z_R	AGCATTTTGAATTTTAAGGAAATTAATTT GTCTTCTATCTTTTGCATCCGTGCCTCCT GGATAATAACTTCAAttgtttatggacacggc	U XII-1	5' homology to LP1.Z, 3' overhang to FgF site

Table A10. Primers to amplify genomic sites (up/down) with homology to LP1

ID	Sequence	Description
FF31_FgF 16 up_F	TCCGTTAATTCGGGTT	16 up
FF32_FgF 16 up_R	tccgttaattcgggtt	
FF33_FgF 16 down_F	TGCCTACGCAACACTT	16 down
FF34_FgF 16 down_R	TTGTTGGGATTCCATTG	
FF35_FgF 1 up_F	agcgaaccataggtcg	1up
FF36_FgF1up_LP1.A_R	ctgaggttctcgTCACTGTGCTTTCTGTTG	
FF37_FgF1down_LP1.Z_F	gtgtccataaacaaTCAGAAATATGGCCGG	1 down
FF38_FgF 1 down_R	aaggtttgtcacgagca	
FF39_FgF 5 up_F	aatgatttaagcgtgcg	5 up
FF40_FgF5up_LP1.A_R	ctgaggttctcgTGCTTCTACCATTGAAATC	
FF41_FgF5down_LP1.Z_F	ccttgaagccgtgtccataaacaaGAAAACAACGCGTCAT	5 down
FF42_FgF 5 down_R	GTACACGCTTTGGGG	
FF43_FgF 7 up_F	cccagttgtttgtagctg	7up
FF44_FgF7up_LP1.A_R	aactgaggttctcgTGGACCCTTAAAAGGC	
FF45_FgF7down_LP1.Z_F	cttgaagccgtgtccataaacaaCGATGGATGATGATTGG	7 down
FF46_FgF 7 down_R	TTGTTGGCATTCCATT	
FF47_FgF8up_LP1.Z_F	cttgaagccgtgtccataaacaaACAAACAGGGCAAATCA	8 up
FF48_FgF 8 up_R	AAAGAACCAGAATGGCA	
FF49_FgF 8 down_F	AGCGTTCGTTCTATGC	8 down
FF50_FgF8down_LP1.A_R	tgaggttctcgGCTGTCTTAGTTAGCTTCG	
FF51_FgF 11 up_F	ctgcaacttgtcagcc	11 up
FF52_FgF11up_LP1.A_R	ctgaggttctcgTCACTGAGCTTTCTGTTG	
FF53_FgF11down_LP1.Z_F	aaacaaATGAGTACTTATTATTAACGAGGAA	11 down
FF54_FgF 11 down_R	CACTTTTACGCCTCTGCTGA	
FF55_FgF12up_LP1.Z_F	tgtccataaacaaCTTAGCCGGCTGAATAA	12 up
FF56_FgF 12 up_R	GCCGCTCGTAAAAAC	
FF57_FgF 12 down_F	TCAACAATGTCGCTTCC	12 down
FF58_FgF12down_LP1.A_R	gaggttctcgTCCCCCGTCTTTCTT	
FF59_FgF 14 up_F	TCAATCAAAGCAACCC	14 up
FF60_FgF14up_LP1.A_R	gagtgttctcgGGTGTAAGAAAATGACATAAAGT	
FF61_FgF14down.LP1.Z_F	cataaacaaGAAAATGCAACCGAtaaa	14 down
FF62_FgF 14 down_R	GCCGTCCTCATGATG	
FF63_FgF 18 up_F	TGTGCACAAAGGCC	18 up
FF64_FgF18up_LP1.A_R	actgaggttctcgGGCATGAGTTATGGTTG	
FF65_FgF18down_LP1.Z_F	tccataaacaaAATCGTCCCCAACAAA	18 down
FF66_FgF 18 down_R	AAAGCTGGCTCCCCTTAGAC	
FF67_FgF 19 up_F	CACCGGAGCTTGG	19 up

FF68_FgF19up_LP1.A_R	gtaactgaggttctcgTTCGCGGGCTGT	
FF69_FgF19down_LP1.Z_F	tgtccataaacaaGGACCAACTATCATCCG	19 down
FF70_FgF 19 down_R	CGTGATAAACGATCGC	
FF71_FgF 20 up_F	GCCAGGCGCCT	20 up
FF72_FgF20up_LP1.A_R	aaggaaaagtaactgaggttctcgTTTGCGAaaCCCTATG	
FF73_FgF20down_LP1.Z_F	cataaacaaAATGGAAGGTCGGGA	20 down
FF74_FgF 20 down_R	ATAAAGCAGCCGCTACC	
FF75_FgF 21 up_F	AAAGGAGGTGCACGC	21 up
FF76_FgF21up_LP1.A_R	actgaggttctcgTTCCAAGGAGGTGAAGA	
FF77_FgF21down_LP1.Z_F	tccataaacaaGATGGGACGTCAGCA	21 down
FF78_FgF 21 down_R	TCAAGACACTCCGGTATTAC	
FF79_FgF 22 up_F	AAACGACGTTGAAATTGA	22 up
FF80_FgF22up_LP1.A_R	gaggttctcgGGACTAGGATGAGTAGCAGC	
FF81_FgF22down_LP1.Z_F	tccataaacaaGGGAAGCATATTTGAGAAG	22 down
FF82_FgF 22 down_R	GGAAACGCTGCCC	
FF83_FgF 23_F	AAGGCACGTATGATTTTC	23 up
FF84_FgF23up_LP1.A_R	actgaggttctcgAAGGGAATGGAAAAATAATG	
FF85_FgF23down_LP1.Z_F	taaacaaATCATCAAAAAACTTATAGGAAA	23 down
FF86_FgF 23_R	CGAGATAAGGCATGGG	
FF87_FgF24up_LP1.Z_F	gtccataaacaaGTTGAAGTCGCCTGG	24 up
FF88_FgF 24 up_R	CTTTCAAGGGTGGGG	
FF89_FgF 24 down_F	GGCTGAACAACAGTCTCT	24 down
FF90_FgF24down_LP1.A_R	aactgaggttctcgGGATCACCTCGCCC	

Table A11. Primers to amplify the LP donors

ID	Sequence	LP / donor construct	Description
FF266_LP1.A_TDH3p_F	tggcgcgtgactgtcagcgcaatccgaggaatactctgaataaa acaacttatataataaaaatgctcgagtttatcattatcaatactg	LP1.*gfp*	Attach LP1.A homology to TDH3p
FF267_LP1.Z_CYC1t_R	cgacagagatgtaggccgccttgttagcgtctgttataattattttctt attttgatgtaatataaagaacgggcaaattaaagccttcg		Attach LP1.Z homology to CYC1t
FF268_LP2.A_ TDH3p_F	ACACCTTAATCTCCGGTTCATGCTAGGGATG TGGCTGCATGCTACGTTGACACACCTACACT GCTCtcgagtttatcattatcaatactg	LP2.*gfp*	Attach LP2.A homology to TDH3p
FF269_LP2.Z_ CYC1t _R	CAATTAACAGTAAACGAACACGTGACGATGC GGAACGGCTCCGGCCAGGTCGTACACTTCA TCTCGCTACTTCgcaaattaaagccttcg		Attach LP2.Z homology to CYC1t
FF270_LP3.A_ TDH3p_F	CTGAAGGATGAGTGTCAGCGAGTGTAACTC GATGAGCTACCCAGTAGTCGTACTGGTCGA GACAACtcgagtttatcattatcaatactg	LP3.*gfp*	Attach LP3.A homology to TDH3p
FF271_LP3.Z_ CYC1t _R	AACGTCAACACCCCCGCCCTAGTCTAAGCTG ATATGTCGCGTCAGGCCTCGGGATTGAGTG CATTAGAGCTGGgcaaattaaagccttcg		Attach LP3.Z homology to CYC1t
FF272_LP4.A_ TDH3p_F	AGGTTGTGAGATGCGGAAGAGGCTCTCGAT CATCCCGTGGGACATCAACCTTTCCCTTGAT AAAGCtcgagtttatcattatcaatactg	LP4.*gfp*	Attach LP4.A homology to TDH3p

FF273_LP4.Z_ CYC1t _R	GGTAGCAAGCTTTGATAAGGTCGAAGGATAG CACAATTCAGAAGGCGTTCTCTCTGCCATAC CCGAGCGGGGTgcaaattaaagccttcg		Attach LP4.Z homology to CYC1t
TDH3p_F	TCGAGTTTATCATTATCAATACTGC	*gfp*	5' end of TDH3p blunt
CYC1t_R	GCAAATTAAAGCCTTCGAGC	*gfp*	3' end of CYC1t blunt
FF266_LP1.A_Tef1_F	ttggcgcgtgactgtcagcgcaatccgaggaatactctgaataaa acaacttatataataaaaatgcggacttttaattttcgaggaccg	LP1.*NCS*	Attach LP1.A homology to TEF1p
FF267_LP1.Z_Pgi_R	gtaggccgccttgttagcgtctgttataattattttcttattttgatgtaat ataaagaacggggtatactggaggcttcatgagttatg		Attach LP1.Z homology to PGI1t
FF268_LP2.A_Tef1_F	TACACCTTAATCTCCGGTTCATGCTAGGGAT GTGGCTGCATGCTACGTTGACACACCTACAC TGCTCggacttttaattttcgaggaccg	LP2.*NCS*	Attach LP2.A homology to TEF1p
FF269_LP2.Z_Pgi_R	TAAACGAACACGTGACGATGCGGAACGGCT CCGGCCAGGTCGTACACTTCATCTCGCTACT TCggtatactggaggcttcatgagttatg		Attach LP2.Z homology to PGI1t
FF270_LP3.A_Tef1_F	TCTGAAGGATGAGTGTCAGCGAGTGTAACTC GATGAGCTACCCAGTAGTCGTACTGGTCGA GACAACggacttttaattttcgaggaccg	LP3.*NCS*	Attach LP3.A homology to TEF1p
FF271_LP3.Z_Pgi_R	CCCCCGCCCTAGTCTAAGCTGATATGTCGCG TCAGGCCTCGGGATTGAGTGCATTAGAGCT GGggtatactggaggcttcatgagttatg		Attach LP3.Z homology to PGI1t
FF272_LP4.A_Tef1_F	AAGTAGGTTGTGAGATGCGGAAGAGGCTCT CGATCATCCCGTGGGACATCAACCTTTCCCT TGATAAAGCggacttttaattttcgaggaccg	LP4.*NCS*	Attach LP4.A homology to TEF1p
FF273_LP4.Z_Pgi_R	TTTGATAAGGTCGAAGGATAGCACAATTCAG AAGGCGTTCTCTCTGCCATACCCGAGCGGG GTggtatactggaggcttcatgagttatg		Attach LP4.Z homology to PGI1t
FF330_Tef1p_F	GGACTTTTAATTTTCGAGGAC	*NCS*	5' end of TEF1p blunt
FF331_Pgi1t_R	GGTATACTGGAGGCTTCATG	*NCS*	3' end of PGIt blunt

Table A12. Primers for building dopamine production strain

ID	Sequence	Description
LB75-delARO4up.S	GTGACACAAAAACAAAATCGAAAAC	*ARO4* UP homology
LB76-delARO4up.AS	GTGATTCTGCCATGCCTTTGC	
LB77-repARO4-NCS.S	GCAAAGGCATGGCAGAATCACCGATAATTGCAGA CGAACGCAG	Amplify half of dopamine production cassette from WCD2249 to attach to *ARO4* UP homology
LB79-ARO4.assem.AS	CCAGTCTAAGGTGGTTTCTTCCAATTCAGACAATG CTGGAGAAGATTGCCAAGATA	
LB78-ARO4-assem.S	TATCTTGGCAATCTTCTCCAGCATTGTCTGAATTG GAAGAAACCACCTTAGACTGG	Amplify half of dopamine production cassette from WCD2249 to attach to *ARO4* DOWN homology
LB80-repARO4-NCS.AS	CTTCTCTTCTTTGTCTGACAGCAGCGTAACATCTC TGTAACTGCTCAGC	
LB81-delARO4down.S	CTGCTGTCAGACAAAGAAGAGAAG	*ARO4* DOWN homology
LB82-delARO4down.AS	GTGATTTCACCAGTTGTACTTCGT	

Table A13. Colony PCR primers to screen for integration at selected genomic loci

ID	Sequence	Genomic loci
FF170-FgF1-MP.S	GAAATTCTACAGCAACAGAAAGCACAGTGA	FgF1
FF171-FgF1-MP.AS	TCAGATTAGATTCCGGCCATATTTCTGA	
FF242_FgF5cPCR_F	ACAAGCAAGTGGGGTAACTTAGACATAAGATTGAC	FgF5
FF243_FgF5cPCR_R	CACGTTCAACAAGGAACTCTACCAACTCTAGTT	

FF172-FgF7-MP.S	GCGACTTTTGGTGGAATATTATGATATGTGTTG	FgF7
FF173-FgF7-MP.AS	CAAATTATGATAAGAACCAATCATCATCCATCG	
FF244_FgF8cPCR_F	TCGGCGATACGATATGTAGAGGGGA	FgF8
FF245_FgF8cPCR_R	CTTCCTGTTGCTTGATTTGCCCTG	
FF174-FgF11-MP.S	GTACTAACGTAAGATCTAGTGTGGTTCGCTTAGG	FgF11
FF175-FgF11-MP.AS	GCAACTCATGTTACCAAGTCGAGACCTTAAAG	
FF246_FgF12cPCR_F	CGCCATATGTTGGTATGATGTCAATAGCAG	FgF12
FF247_FgF12cPCR_R	GCTCCCTCAGATTGGTGCAATCG	
FF176-FgF14-MP.S	CTCATCCACAGGCAATGAAGAACACGCTAAC	FgF14
FF177-FgF14-MP.AS	GATGGCGATGTATGGGGTTTTGGCAAAGC	
FF178-FgF16-MP.S	GAAGAGTGGGTGAGTTTTGAGATAATTGTTGG	FgF16
FF179-FgF16-MP.AS	GCTGTCATCAACTTTGGGATTACTGCATTTGC	
FF180-FgF18-MP.S	CTGTTTTCAGTAGATTTGGTAACTGTGCAACC	FgF18
FF181-FgF18-MP.AS	GAGCATTTCGTTCACTTACCAAACAATTAAGG	
FF182-FgF19-MP.S	GATTCCGCGCTTCCATCATTTAGTATAATCC	FgF19
FF183-FgF19-MP.AS	GGTAATGTCAGTAATTAGCGGATGATAGTTGG	
FF184-FgF20-MP.S	CCGTGTAGAGTTCTGTATTGTTCTTCTTAGTGC	FgF20
FF185-FgF20-MP.AS	CAGAAGTCCAAATCACGTCAAGACAAAGAAAG	
FF186-FgF21-MP.S	CTCGTAATGTTCAGGCTTATAATGTCAGTATGC	FgF21
FF187-FgF21-MP.AS	GTTGCTGAGTAATCTTCATTGCGCTTATTTATCG	
FF248_FgF22cPCR_F	CCCAGTATTCTTAACCCAACTGCACAGAAC	FgF22
FF249_FgF22cPCR_R	CTTGGTTCTGGCGAGGTATTGGATAG	
FF250_FgF23cPCR_F	CAGACAAGACAGAGAAGGACAAGGCTG	FgF23
FF251_FgF23cPCR_R	CCAAGGCCTGACACAAAGGATTATTCC	
FF188-FgF24-MP.S	CGTAGTGGCGATCTTGTGATTTTCGTAC	FgF24
FF189-FgF24-MP.AS	GCACTGTGGATTCTATGTACTTGGCAATAG	
FF298_LP3_User12_cPCR_F	GATTGACGTCATTCTGAGTTACAATGATCTTAC	UserXII-1
FF299_LP3_User12_cPCR_R	CATCCGTGCCTCCTGGATAATAACTTC	

Table A14. Multiplex colony PCR primers to screen LP platform integrations

ID	Sequence	Description
FF184-FgF20-MP.S	CCGTGTAGAGTTCTGTATTGTTCTTCTTAGTGC	LP1-FgF20
FF185-FgF20-MP.AS	CAGAAGTCCAAATCACGTCAAGACAAAGAAAG	
FF180-FgF18-MP.S	CTGTTTTCAGTAGATTTGGTAACTGTGCAACC	LP2-FgF18
FF181-FgF18-MP.AS	GAGCATTTCGTTCACTTACCAAACAATTAAGG	
FF188-FgF24-MP.S	CGTAGTGGCGATCTTGTGATTTTCGTAC	LP2-FgF24

FF189-FgF24-MP.AS	GCACTGTGGATTCTATGTACTTGGCAATAG	
FF298_LP3_User12_cPCR_F	GATTGACGTCATTCTGAGTTACAATGATCTTAC	LP3-UserXII-1
FF299_LP3_User12_cPCR_R	CATCCGTGCCTCCTGGATAATAACTTC	
FF300_LP3.FgF7.cPCR_F	GATTCGTATTCAAATTTGGTGAAGGAATAACG	LP3-FgF7
FF301_LP3.FgF7.cPCR_R	ATGACATTGGAATTTCGTTACAACAATACCTC	
FF302_LP.FgF19.cPCR_F	CTTCGTTGTATCTCAGAATGAGATCCCTC	LP3-FgF19
FF303_LP.FgF19.cPCR_R	CAAGAGTCAGCTACACGCAAATTCTG	
FF304_LP4.FgF12.cPCR_F	GCCATATGTTGGTATGATGTCAATAGCAGTAG	LP4-FgF12
FF305_LP4.FgF12.cPCR_R	CAGCGATTATTCAGCCGGCTAAG	
FF306_LP4.FgF16.cPCR_F	CTTCACTTCATTAGGTATATTCTTGGTTATGCG	LP4-FgF16
FF307_LP4.FgF16.cPCR_R	TGTTGAAAGATTCCTCTTGTGATTTTCATG	
FF308_LP4.FgF21.cPCR_F	CGAATGAAACATTCATATTTCGCATG	LP4-FgF21
FF309_LP4.FgF21.cPCR_R	CGCAGTAACCTCTTACAGAACGCC	
FF310_LP4.FgF22.cPCR_F	GGTTAATGTGGCTGTGGTTTCAGG	LP4-FgF22
FF311_LP4.FgF22.cPCR_R	CTTGGTTCTGGCGAGGTATTGGATAG	

NCS SEQUENCES

Table A15. *NCS* codon-optimized nucleotide sequences

Candidate	Sequence
AmNCS	ATGATCTTCACTGTCTACATTTTGTACTACAAGTACACCACCCACACTGCTAACTTGTTCAGAAAACACATTAGAT TAAAGATGGTTGCTTCTGTTTCTGGTGAAGTTGAGGTTAATGCTCCAGCTTCTAAGGTTTGGGAATTGTACAGAT CTTTGGAATTGTTGAGAATCACTAAAAAGGGTTTGGATCACATCGTTGATAAAATTGAAGTCTTGGAAGGTGATG GTTCCGTCGGTACCTTGTTACATTTCACTTTTCACCCTGGTGCTTTGCCTTTCTCTTCCTACAAAGAAAAATTCAC CAAAGTTGACGACGAAAACATGGTCAAGGTTGTTGAAGTTGTCGAGGGTGGTTTTTTAGAATTGGGTTTTAAATG GTACTTAGTTCGTTTGGACGTTATTGTTAAGGATGAAAAATCCTGCATTACCAAGAACACCATTGAGTACGAATTG AAAGAAGATGCTGATCCTAAGTTGGCCTCTGTCGTCTCCATTGATCCATTGATGGCCATGATGAACATTGCTGCT AACCATGTTGTTTCCGGTATTAAGGCT
EcNCS	ATGATCGGTGGTTTCTTGGATATGGGTTGTACTTTCTACATGGATAGAATCCACGTTGTTGCTAAAGGTCCAAAC TCTTGTATCATTAAATCCACCTTGATCTACGAAGTCAAGGAAGAGTATGCTGACGCTATGGCCTCTTTAATCACT GTCGAGCCATTAGCCTCCATGGCTGAAGTCGTTGCTAACTACGTTTTACACCAACAAGTTAGAGTCTTGGGTTCC GTTAAGAGAAAGGAATTGACCCATGAATTGGAAGTCGCTGCTCCAGCCGACGCTATTTGGGGTGTTTACTCTTC TCCTGATATTCCAAGATTGTTGAGAGACGTTTTGTTACCAGGTGTCTTCGAAAAGTTGGAAGTCATCCAAGGTAA CGGTGGTGTTGGTACTGTTTTAGAAATCGTTTTTCACCCAGGTGCTATTCCAAGAAGATACAAGGAAAAGTTCGT TACTATTAACCACAAAAAGCGTTTGAAGGAAGTCGTTATGATCGGTGGTTACTTGGATATGGGTTGTACTTTATAT ATGGATCGTATCCACGTCGTTTCCAAGGGTCCAAACTCTTGTGTTATCAAGTCCACCTTGATTTACGAAGTTAAG GCCGAATCCGCCGATGCTATGGCCTCCACTATTACCATCGACCCATTGGCTTCCATGGCCCAAGTTATCTCTAAT TATGTTTTGAAGAATCAAATGCAAGTCTTGGGTTCCGTCAAGAGAAGAGAATTAACTCACGAATTAGAAGTTGCT GCTTCTGCCGATGCCATCTGGGGTGTTTACGGTTCTAAGAGATACTCTAAAGCCTCCCAAGGTTGCTTCGCCTC CTGGTGTTTCCGTAAGGTTCGTTCTCAC
NdNCS	ATGAGATCTGGTATTGTCTTCTTGGTCTTGTTCTTCTTGGGTTGTGAAATTTCCCAAGGTAGACAATTGTTAGAGT CTAGATTATTCAGAAAGTCCACCATTAGAAAGGTCTTACACCACGAATTATCCGTTGCCGCTTCCGCTCAAGAAG TTTGGGATGTTTACTCCTCTCCAGAATTGCCAAAGCATTTGCCAGAGATTTTACCAGGTGCTTTCAAAAAAGTCG TCGTTACTGGTGACGGTGGTGTTGGTACTGTTATTGAGATGATTTTCCCACCAGGTGTCGTCCCACACAGATACA AAGAAAAATTCGTTTTAATTGATGACGAAAAATTTTTGAAGAAGGTTGAAATGATCGAAGGTGGTTACTTGGATAT GGGTTGTACTTTTTACATGGACACTATTCAAATCATTCCAACCGGTCCTGATTCTTGCATTATCAAGTCTTCTACT GAATACTACGTCAAGCCAGAGTTCGCTGATAAGGTTGTCCCATTAATTTCTACTGTTCCATTACAAGCTATGGCT GAAGCTATCGCTAAGATCGTTTTGGAAAACAAGGCTAAGCATAAGGGTTTTATCGAAATC
PsNCS1	ATGAAGGGTAAGATCTCTAAAGAAGTTCAAGTCCCAGTTCCAGCCTCCGATTTGTGGGCTGTTTACGGTACCTTA GAGTTAATCCGTTTAATCAAAAAGTTGTTGCCAGAGATCTTAAGAGACTTCGAAGTTGTTGTTGGTGATGGTGGT GTTGGTACTGTCTTGAAATTGACTTTCCCACCTGAATCCCCAGTTACCAATTACTCCGAAAAATTTACCAAGGTC GACAACGAAAAGAGAATCAAGGTTACTGAGGTTGTTGAAGGTGGTTATTTAGAAGTTGGTTTCTCTTTGTACAGA GTTACCTATGAAATCACTGAAAAGGGTGAACACTCTTCTGTTATTATTACTATCGAATATGAATTAGACGACGCCT TCGCTGATAATGCTTCTTTGGTTTCCATTAAGCCATTACAAGTTATCGCTAAGACTATCGGTAAGTACTTGACTGA AAAGAAGGGT
PsNCS2	ATGAAATACCAAGCCGGTTTGTCTATTTTCTTGTTATTCTTGATCGGTACCGGTGAATCCTCCAAATACACTTTGG TCAACGATTTCAACGTTGCCGCCTCCGCTGATGAGGTTTGGGCTGTTTACTCCTCTCACAATTTACCAAAGTTGA TCGTCAAATTGTTGCCAGGTATGTTTAAGAGAATTGACGTCTTGAAGGGTGATGGTGGTGTTGGTACTATCTTGA GATTAGTCTACCCTCCAGGTTCTGTTCCATTGACCTATAAGGAAAAGTTCGTTACTATTGACAACAGAAGAAGATT AAAGGAAGTTTTGCAAATCGAAGGTGGTTACTTGAAGATGGGTGTCACCTTCTACATGGATTCTTTTCAAGTCAT TAAGCGTGGTAGAGATTCCTGCATTATCAGATCTATCACTAAGTACGAAATTAGAGACGATTTGGCTGTTAAGGT CTCCCCATTAATCTCTGTTGATTCTTTAGTCACTATGGCTAGAGCTATTTCTAAGTACGTTTTGGAAAACAAAAAG AAGGCTAACTCTACCATCGTTCCA
PsNCS3	ATGAGAAAGGTTATTAAGTACGACATGGAAGTCGCCGTCTCTGCTGATTCCGTCTGGGCTGTCTACTCTTCCCC AGATATCCCAAGATTGTTGCGTGATGTCTTGTTGCCAGGTGTCTTCGAAAAGTTGGACGTTATTGAGGGTAACG GTGGTGTTGGTACCGTTTTGGATATTGTCTTTCCACCTGGTGCCGTTCCAAGATCCTACAAGGAAAAGTTCGTCA ATATCGACCGTGAAAAGAGATTGAAGGAAGTTATTATGATTGAGGGTGGTTACTTGGACATGGGTTGTACTTTCT ACTTGGATAGAATTCACGTCGTTGAGAAGACTAAGTCTTCTTGCGTCATTGAGTCCTCCATCGTTTACGATGTTA AAGAAGAGTGTGCCGACGCTATGTCCAAGTTGATTACCACCGAACCATTGAAATCCATGGCCGAAGTCATTTCTA ACTACGTTATTCAAAAGGAATTGTTCTCTGCCAGAAACATCTTATCTAAGCAATCCGTTGTCAAGAAGGAAATTCG TTACGATTTAGAAGTTCCAATTTCCGTTGATTCTATTTGGTCCGTTTACTCCTGTCCAGACATCCCACGTTTATTA
ScNCS	ATGAGATCTGGTATTGTCTTTTTGGTTTTGTTCTTCTTAGGTTGTGAAATTTCCCAAGGTAGACAATTGTTGGAAT CTAGATTGTTCAGAAAGTCTACTATTCAAAAGGTTTTGCACCACGAATTGCCAGTCGCTGCTTCTGCTCAAGAAG TTTGGGATGTCTATTCTTCCCCAGAATTGCCAAAACACTTGCCAGAAATTTTGCCAGGTGCCTTCGAAAAGGTTG TCGTCACCGGTGATGGTGGTGTTGGTACTGTTTTGGAAATGGTTTTTCCACCAGGTGAAGTTCCAAGATCCTACA AAGAAAAGTTCGTCTTAATTGACGATGAGCAATTGTTGAAGAAGGTTGAAATGATCGAAGGTGGTTACTTGGATA TGGGTTGTACTTTCTACATGGACACTATCCAAATTGTCCCAACCGGTCCAGACTCCTGTATCATCAAGTCTTCTA CTGAATACTATGTCAAGCCTGAATTCGCTGATAAGGTTGTTCCATTGATTTCTACTATCCCATTGCAAGCTATGGC CGAGGCCATCTCTAACATTGTCTTAGCTAACAAGGCCAAGAACAAGTCTATTATTATCGAAATT
SdNCS	ATGCGTAAGGAAGTCAGATACGAAATGGAAGTTCCAACCTCCGCTGATTCCATTTGGGCTGTCTACTCTTCTCAC GATATTCCAAGATTGTTGAAGGAAGTTTTGTTGCCAGGTGTCTTCGAAAAGTTGGACGTCATTGAAGGTGACGGT GGTGTTGGTACTGTTTTGGATATTGCCTTCCCACCTGGTGCTGTTCCAAGAACTTACAAGGAAAAATTCGTTACT

	ATTAACCACGAAAAGAGATTGAAAGAGGTTATTATGATCGAAGGTGGTTACTTGGATATGGGTTGTACTTTTTACA TGGATAGAATCCATGTTTTGGAAAAAGGTCCAAACTCCTGTGTTATCGAATCCGCTATTATTTACGAAGTTAAGGA AGAATTCGCTGATGTCGTCGTCCCATTGATTACCACCGAACCATTAGCTTCTATGGCTGAAGTCATTTCTAATTAC GTCTTGAAGAAGCAAATTCATGTTTTCGGTTACGTTATCAAACCAAAGTTGGGTTTATCCTTATTGTTATGTTTCAT CTTGTGTTTAGTCTTGTTAGGTGTTTTGTTAATTGGTGGTGTCCCATTG
TcNCS	ATGAGAAAGGAATTGAGACATGAATTAAAGGTCCCAGCCTCTGCTGATGATGTCTGGGAAGTTTACTCCTCCCC AGATTTGCCTAAATTGATTGTTCAATTGTTGCCATCCGTTTTTGAAAAAATTGAGGTTGTTGAAGGTGATGGTGGT GTTGGTACTGTTTTACATATTACTTTTCCACCAGGTTCTATTCCAATTTCCTACAAGGAAAAATTCGTCACTGTTGA CGATTACAAGCGTTTGAAGGAAGTTAGACAAATTGAAGGTGGTTACTTGGACATGGGTTGTACCTTCTACATGGA TTCTTTCCACATTTTAGAAGAGTGCCACGATTCCTGCGTTATCGTCTCCAAGACTGAATACGAAGTTCCACAAGA ATTGGCTGCCAACGTTGAACCATACATCTCTATCGACTCCTTGGCTGGTATGGCCACTGCTATTTCTAACTACGT TGTTGACAAGAACAAAAACAAGGAATGTGAACCAGAAGGTTGTTGTGATGACGAACGTGACCACTGTTCTGAAA AAGAAGGTAGACACGAACAATCTTCTGATGAATCTGAATCCGATTGCGAATCCGAATCCGACTGTGAC
TfNCS	ATGAGAAAGGAATTGACCCACGAAATGGAAGTTCCAGCCTCTGCTGACGCTATCTGGGCTGTCTACGGTTCCCC TGACATTCCACGTTTGTTGAAGGAGGTTTTGTTGCCAGGTGTTTTCGAAAAGTTGGATGTTATCGAAGGTGACGG TGGTGTCGGTACTGTCTTAGACATTGCCTTCCCACCAGGTGCTGTTCCACGTGCCTACAAGGAAAAGTTCATGA AAGTTAACCACGAAAAGAGATTGAAAGAAGTTGAAATGATTGAAGGTGGTTACTTGGACATGGGTTGTACCTTCT ACATGGATAGAATCCACGTTGTTGAGAAAGGTCCAAATGCTTGTGTTATTGAATCTGCCATCATTTACGAAGTCA AAGATGAATTTGCTGACGTCGTTGTTCCATTGATCACTACTGAACCTTTGGCCTCCATGGCTGAAGTCATTTCCA ACTACGTCTTGAAGAACCAATTCAGAGTCTTCGGTTATGTCATTAAGCCAAAGTTAGGTTTATCCTTGTTATTGTG TTTCATCTTGTGTTTGGTTTTGTTAGGTGGTTTGTTGATCGGTGGTGTTCCATTA

Table A16. NCS amino acid sequences

Candidate	Sequence
***Am*NCS**	MIFTVYILYYKYTTHTANLFRKHIRLKMVASVSGEVEVNAPASKVWELYRSLELLRITKKGLDHIVDKIEVLEGDGSVG TLLHFTFHPGALPFSSYKEKFTKVDDENMVKVVEVVEGGFLELGFKWYLVRLDVIVKDEKSCITKNTIEYELKEDADP KLASVVSIDPLMAMMNIAANHVVSGIKA
***Ec*NCS**	MIGGFLDMGCTFYMDRIHVVAKGPNSCIIKSTLIYEVKEEYADAMASLITVEPLASMAEVVANYVLHQQVRVLGSVK RKELTHELEVAAPADAIWGVYSSPDIPRLLRDVLLPGVFEKLEVIQGNGGVGTVLEIVFHPGAIPRRYKEKFVTINHK KRLKEVVMIGGYLDMGCTLYMDRIHVVSKGPNSCVIKSTLIYEVKAESADAMASTITIDPLASMAQVISNYVLKNQMQ VLGSVKRRELTHELEVAASADAIWGVYGSKRYSKASQGCFASWCFRKVRSH
***Nd*NCS**	MRSGIVFLVLFFLGCEISQGRQLLESRLFRKSTIRKVLHHELSVAASAQEVWDVYSSPELPKHLPEILPGAFKKVVVT GDGGVGTVIEMIFPPGVVPHRYKEKFVLIDDEKFLKKVEMIEGGYLDMGCTFYMDTIQIIPTGPDSCIIKSSTEYYVKP EFADKVVPLISTVPLQAMAEAIAKIVLENKAKHKGFIEI
***Ps*NCS1**	MKGKISKEVQVPVPASDLWAVYGTLELIRLIKKLLPEILRDFEVVVGDGGVGTVLKLTFPPESPVTNYSEKFTKVDNE KRIKVTEVVEGGYLEVGFSLYRVTYEITEKGEHSSVIITIEYELDDAFADNASLVSIKPLQVIAKTIGKYLTEKKG
***Ps*NCS2**	MKYQAGLSIFLLFLIGTGESSKYTLVNDFNVAASADEVWAVYSSHNLPKLIVKLLPGMFKRIDVLKGDGGVGTILRLV YPPGSVPLTYKEKFVTIDNRRRLKEVLQIEGGYLKMGVTFYMDSFQVIKRGRDSCIIRSITKYEIRDDLAVKVSPLISV DSLVTMARAISKYVLENKKKANSTIVP
***Ps*NCS3**	MRKVIKYDMEVAVSADSVWAVYSSPDIPRLLRDVLLPGVFEKLDVIEGNGGVGTVLDIVFPPGAVPRSYKEKFVNID REKRLKEVIMIEGGYLDMGCTFYLDRIHVVEKTKSSCVIESSIVYDVKEECADAMSKLITTEPLKSMAEVISNYVIQKE LFSARNILSKQSVVKKEIRYDLEVPISVDSIWSVYSCPDIPRLL
***Sc*NCS**	MRSGIVFLVLFFLGCEISQGRQLLESRLFRKSTIQKVLHHELPVAASAQEVWDVYSSPELPKHLPEILPGAFEKVVVT GDGGVGTVLEMVFPPGEVPRSYKEKFVLIDDEQLLKKVEMIEGGYLDMGCTFYMDTIQIVPTGPDSCIIKSSTEYYV KPEFADKVVPLISTIPLQAMAEAISNIVLANKAKNKSIIIEI
***Sd*NCS**	MRKEVRYEMEVPTSADSIWAVYSSHDIPRLLKEVLLPGVFEKLDVIEGDGGVGTVLDIAFPPGAVPRTYKEKFVTIN HEKRLKEVIMIEGGYLDMGCTFYMDRIHVLEKGPNSCVIESAIIYEVKEEFADVVVPLITTEPLASMAEVISNYVLKKQI HVFGYVIKPKLGLSLLLCFILCLVLLGVLLIGGVPL
***Tc*NCS**	MRKELRHELKVPASADDVWEVYSSPDLPKLIVQLLPSVFEKIEVVEGDGGVGTVLHITFPPGSIPISYKEKFVTVDDY KRLKEVRQIEGGYLDMGCTFYMDSFHILEECHDSCVIVSKTEYEVPQELAANVEPYISIDSLAGMATAISNYVVDKNK NKECEPEGCCDDERDHCSEKEGRHEQSSDESESDCESESDCD
***Tf*NCS**	MRKELTHEMEVPASADAIWAVYGSPDIPRLLKEVLLPGVFEKLDVIEGDGGVGTVLDIAFPPGAVPRAYKEKFMKVN HEKRLKEVEMIEGGYLDMGCTFYMDRIHVVEKGPNACVIESAIIYEVKDEFADVVVPLITTEPLASMAEVISNYVLKN QFRVFGYVIKPKLGLSLLLCFILCLVLLGGLLIGGVPL

Table A17. Sequence identity between NCS variants

	*Am*NCS	*Ec*NCS	*Nd*NCS	*Ps*NCS1	*Ps*NCS2	*Ps*NCS3	*Sc*NCS	*Sd*NCS	*Tc*NCS	*Tf*NCS
***Am*NCS**	–	37%	35%	47%	35%	36%	35%	36%	39%	37%
***Ec*NCS**	37%	–	49%	38%	47%	63%	57%	71%	60%	73%
***Nd*NCS**	35%	49%	–	39%	56%	54%	91%	56%	58%	57%
***Ps*NCS1**	47%	38%	39%	–	40%	38%	41%	40%	42%	42%
***Ps*NCS2**	35%	47%	56%	40%	–	52%	53%	54%	60%	51%
***Ps*NCS3**	36%	63%	54%	38%	52%	–	57%	78%	57%	74%
***Sc*NCS**	35%	57%	91%	41%	53%	57%	–	60%	60%	61%
***Sd*NCS**	36%	71%	56%	40%	54%	78%	60%	–	62%	90%
***Tc*NCS**	39%	60%	58%	42%	60%	57%	60%	62%	–	61%
***Tf*NCS**	37%	73%	57%	42%	51%	74%	61%	90%	61%	–

APPENDIX FIGURES

Figure A1. Amino acid sequence alignment of NCS variants

Alignment generated by MUSCLE (EMBL-EBI). Active enzymes are indicated by red font. Residues are colour-coded based on their physicochemical properties: small/hydrophobic (red), acidic (blue), basic

(magenta), hydroxyl/sulfhydryl/amine (green), and other (grey). Asterisks indicate fully conserved residues, colons and periods indicate strong and weak conservation respectively.

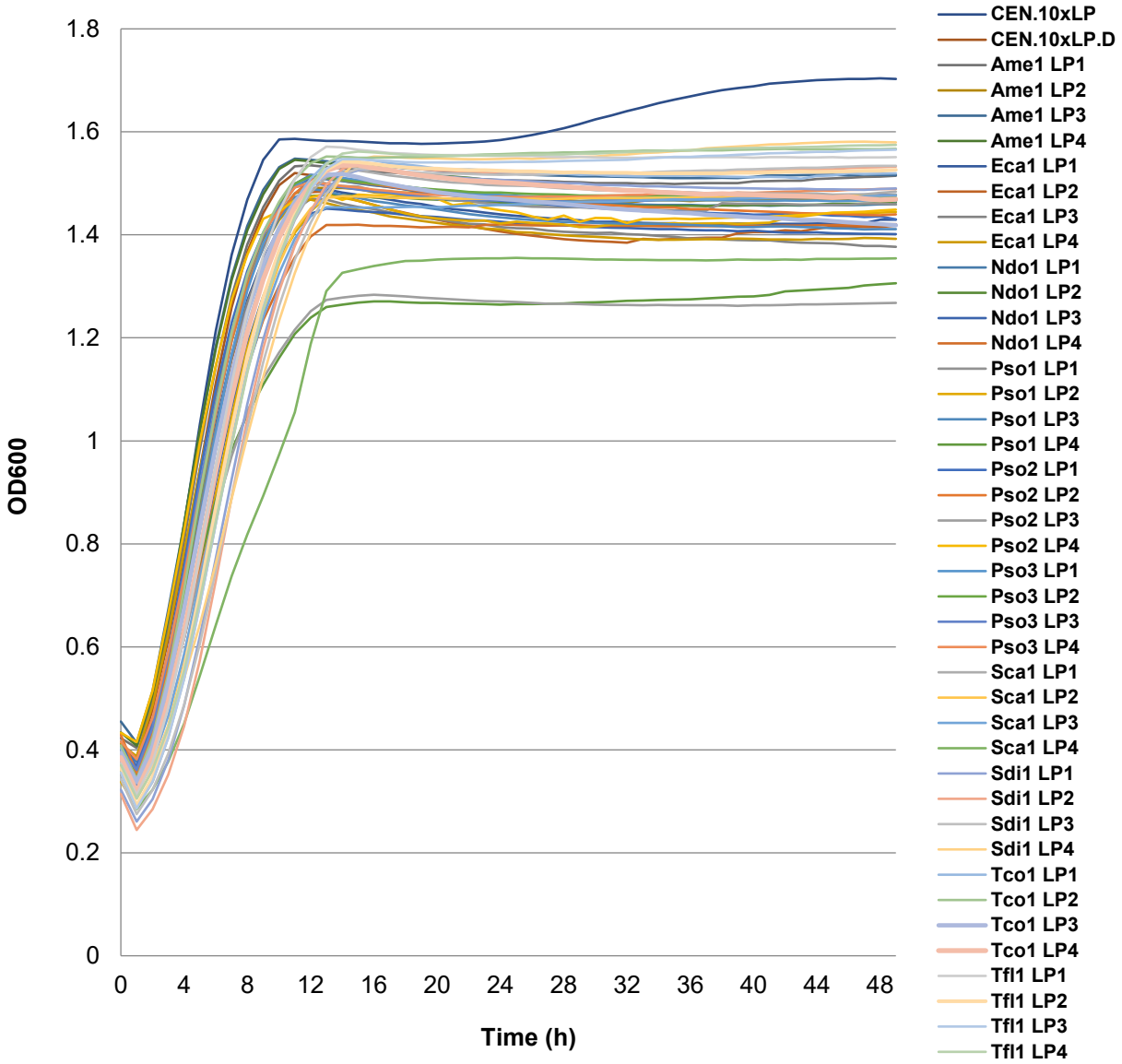


Figure A2. Growth of LP_X.NCS variant & copy number strains.

Cultures were grown in the Sunrise® absorbance microplate reader at 30°C with shaking at 200 rpm. Cell density was measured using OD_{600} in 20 min intervals for 48 h.