Combining 26s rDNA and the Cre-loxP system for iterative gene integration and efficient marker curation in *Yarrowia lipolytica*

Yongkun Lv [1,2], Harley Edwards[1], Jingwen Zhou [2] and Peng Xu[1,*]

[1]Department of Chemical, Biochemical and Environmental Engineering, University of Maryland Baltimore County, Baltimore, MD 21250

[2]Key Laboratory of Industrial Biotechnology, Ministry of Education and School of Biotechnology, Jiangnan University, Wuxi, Jiangsu, China

---

[*] Corresponding author Tel: +1(410)-455-2474; fax: +1(410)-455-1049.

E-mail address: pengxu@umbc.edu (Peng Xu).

Abstract: Conventional plasmid-based gene expression tends to introduce genetic instability and gene copy number variations that lead to degenerated production. The limited number of auxotrophic markers in *Yarrowia lipolytica* also restricts our ability to perform iterative genetic modifications and manipulate long gene clusters. To overcome these limitations, we combined the high recombination efficiency of the Cre-loxP system and the high integration rate of 26s rDNA, and developed a versatile framework to iteratively integrate multi-copy metabolic pathways in *Y. lipolytica*. We demonstrated the efficient genome integration of a plant-derived flavonoid pathway at random sites with multiple copies. Transient expression of Cre recombinase enabled efficient marker removal and allowed for the next round of genome integration. Investigating the recombination events demonstrated that the iterative integration is happening at sufficiently high rates (more than 80%) without disrupting the previous integration. Both the flavonoid precursor pathway and the plant-derived cytochrome *c* P450 enzymes were functionally integrated to improve flavonoid and hydroxylated flavonoid production. The engineered strains produced 71.2 mg/L naringenin, 54.2 mg/L eriodyctiol, and 48.1 mg/L taxifolin. The reported work provides a versatile platform to iteratively integrate functional gene clusters at high copy numbers. This work may streamline and expand our capability to build efficient microbial cell factories for high-value natural products and commodity chemical production in *Y. lipolytica*.

Keywords: chromosome integration, Cre-loxP, 26s rDNA, *Y. lipolytica*, flavonoids,

cytochrome *c* P450 enzymes

The oleaginous yeast *Yarrowia lipolytica* is one of the most studied non-conventional yeast. The oleogeneity is associated with its capacity to accumulate up to 20% oils under nitrogen starvation conditions, making this yeast a particularly interesting platform to produce fuels and commodity chemicals derived by fatty acids.[1] The robust cell growth and high acetyl-CoA and malonyl-CoA flux make it well suitable to be used as a chassis to produce acetyl-CoA and malonyl-CoA-derived natural products, including ω-3 polyunsaturated fatty acids (PUFAs)[2-5], flavonoids and isoprenoids *et al.*[6, 7] For instance, engineered *Y. lipolytica* strains have been patented by DuPont and Microbia as PUFA and carotenoid producers. [8, 9] *Y. lipolytica* can utilize a wide range of substrates, including low-cost volatile fatty acids (VFAs), waste glycerol, whey-containing dietary wastes and renewable sugars, like hexose and pentose sugars.[10] Recently, *Y. lipolytica* has been engineered to utilize xylose at a relatively high efficiency[11, 12], opening the possibility for bioconversion of lignocellulose-based feedstock and consolidated bioprocessing.[13, 14] Moreover, *Y. lipolytica* is also a “generally recognized as safe” (GRAS) organism and has been employed in the food/drink industry to manufacture citric acid and succinic acid.[15] These phenotype advantages present us an exciting opportunity to develop efficient biorefinery platforms that could upgrade low-cost feedstocks to high-value products.[7]

Meanwhile, a large panel of synthetic biology tools have recently been developed for targeted, precise and efficient genetic modification in *Y. lipolytica*.[7, 16] These molecular tools span a wide range of molecular scales[16], including gene overexpression with strong promoters[17, 18] and self-replicating plasmids [19], multiple copy gene integration with 26s rDNA[20-23] or Zeta DNA[24], gene knockout or integration with the Cre-loxP system[25], transposon mutagenesis[26] and genome-editing with CRISPR-Cas9[19, 27-29] *et al*. Despite of these achievements, we need more generalized tools to manipulate multi-gene pathways and maximize the potential of this yeast to produce various products. Comparing with chromosome-based gene integration, plasmid-based gene expression has been associated with genetic instability and gene copy number variations that usually lead to inconsistent strain performance.[30, 31] The other challenge is the limited number of selection markers in *Y. lipolytica*, which restricts our ability to perform multiple-round genetic modifications and manipulate long gene clusters. For example, the two auxotrophic marker (*leu2* and *ura3*) and one dominant resistance marker (hygromycin) have been routinely used in genetically modified *Y. lipolytica*. Iterative marker curation and recycle must be implemented before the next round of genetic modification. This homologous recombination based conventional approach to cure the marker is time-consuming and laborious, generally with a relatively low success rate.

To overcome these limitations, we herein combine the high recombination specificity of the Cre-loxP system and the high integration efficiency of the 26s rDNA technique, with the aim to develop an efficient approach to iteratively integrate multi-copy metabolic pathways in *Y. lipolytica*. With the plant-derived flavonoid pathway as a testbed, we demonstrated that the iterative integration and marker curation of multiple genes to expands our previous YaliBrick toolbox (**Figure 1**). Briefly, a loxP-*URA3*-loxP curation cassette was flanked with both upstream and downstream 26s rDNA homologous arms. Between the loxP-*URA3*-loxP and the downstream 26s rDNA, the heterologous pathway could be inserted following the YaliBrick gene assembly framework. This technique allows us to iteratively target genes into the genome at random sites with multiple copies. Phenotypic screening of the obtained colonies leads to a heterogenous population with a variety of strain performances. Subsequent expression of the Cre recombinase enables efficient marker removal and allows for next-round of genetic modification.

To validate our method, we probed the integration sites of a 18 kb flavonoid pathway and demonstrated that this random integration is sufficient to express multiple gene pathways. We surveyed cytosolic acetyl-CoA and malonyl-CoA pathways and improved the flavonoid production. The plant-derived cytochrome *c* P450 enzymes were also tested and functionally expressed to generate hydroxylated flavonoids. Taken together, this method provides an efficient approach to assemble metabolic pathways with varying genomic integration sites and/or gene copy numbers. By simplifying and streamlining pathway construction, the reported work facilitates the development of more efficient biocatalytic platforms to produce high-value natural products or commodity chemicals in *Y. lipolytica*.

Results and discussion

Develop random integration method

Y. lipolytica Po1f is a derivative of the wild-type W29 strain (ATCC 20460), which contains at least 200 copies of ribosome DNA (rDNA) clusters.[32, 33] These rDNA clusters have been used for random integration of heterologous pathways[21], including the yeast cell surface display of mCherry[20] and the multi-copy integration of the β-carotene pathway.[22] Because only two auxotrophic markers, Δ*leu2* and Δ*ura3,* are available in *Y. lipolytica* Po1f, a limited round of genetic modifications could be performed in this yeast. This presents a critical challenge to assemble multiple heterologous pathways or engineer the chassis cell with plasmid-based gene expression. In order to functionally express a gene cluster encoding an entire biosynthetic pathway, we first developed a 26s rDNA based random integration method. In this method, the orotidine-5'-phosphate (OMP) decarboxylase gene *URA3* was used as a selective marker. Two tandem copies of the loxP sites were inserted after the promoter and before the terminator to flank the *URA3* gene (**Figure 1**). Subsequently, the Cre-loxP recombinase system, carried by a quasi-stable plasmids, was used for marker rescue.[34] In principle, we can iterate this genetic modification cycle and repeatedly integrate genes without the constraints of the number of auxotrophic markers. The integrative plasmid prDNAloxP (**Figure 2A**) was developed on the basis of the YaliBrick plasmid pYLXP'. [19] Based on the YaliBrick gene assembly principle, multiple-gene pathways can be conveniently inserted downstream of the TEF-intron promoters. By repeated use of compatible-end ligation (*Nhe*I/*Sal*I sites compatible with *Avr*II/*Sal*I sites), functional coding sequences of entire metabolic pathways could be easily assembled.

Genomic integration of the flavonoid pathway

As a proof of concept, we sought to integrate the naringenin biosynthetic pathway RtTAL-Pc4CL-PhCHS$_{x5}$-MsCHI (T4S$_{x5}$I) to validate this method. The colony PCR results indicated that 19 out of 24 colonies contain the heterologous naringenin pathway (**Figure 2B**), and naringenin was produced by these colonies in non-selective YPD medium. All the chosen colonies produced naringenin, with a range of 5.4 – 25.7 mg/L in test tubes (**Figure 2C**). These results indicate that 26s rDNA-based platform could integrate the complete naringenin pathway into the genome with an efficiency of 79.2% (19/24). The best performing strain #7 produced 25.7 mg/L naringenin, and was named as NarPro_#7. It was essential to screen multiple transformants to select the high performing strains, because the pathway was randomly integrated into the genome. Although we did not further confirm the integration sites at the genomic loci of 26s rDNA, the 26s rDNA sequences have been frequently used as the landing pad for genomic integration in other non-conventional yeasts[35, 36], and multiple studies have confirmed the integration of heterologous genes to the 26s rDNA sites. [21, 22]

The *URA3* marker was later rescued by transforming pYLXP’-Cre into NarPro_#7, and subsequently removal of *URA3* marker was selected on

CSM-Leu+Ura+5-FOA plate. This transient expression of Cre recombinase was sufficient to remove the *URA3* marker. Colony PCR results suggested that the *URA3* auxotrophic marker was rescued with an efficiency of 95% (19/20) (**Figure 2D**). To further confirm the complete removal of *URA3* marker, these colonies were replicated onto CSM+5-FOA+Ura and CSM-Ura plates, respectively. The results showed that all the colonies except #10 grew on CSM+5-FOA+Ura plate, but did not grew on CSM-Ura plate (**Figure 2E** and **2F**). These results indicated that all *URA3* markers were removed and the cell requires uracil to survive but has developed resistance to 5-FOA. The #10 colony should be a contaminant, because it did not produce a specific band in colony PCR, nor did it grow on CSM+5-FOA+Ura or CSM-Ura plates. These results indicated that the *URA3* marker was completely removed and the naringenin-producing strain could be used for the next round of genetic modification. To test the stability of the naringenin-producing strain, we randomly picked 5 colonies and inoculated them into YPD liquid media. All the picked strains produced similar naringenin as NarPro_#7 (**Supplementary Figure 1**), indicating that the naringenin pathway was stable after removing *URA3* marker.

Finally, the Cre recombinase expressing plasmid, pYLXP'-Cre, was removed by culturing the chosen strain in YPD medium for 48 h. Removal of pYLXP'-Cre was confirmed by re-growing the cell on YPD and CSM-Leu plates, respectively (**Supplementary Figure 2**). Since Po1f is a leucine auxotrophic strain, ejection of the plasmid (pYLXP'-Cre) will leave the strain not able to survive on CSM-Leu media.

After the removal of the Cre recombinase, the best performing strain #20, producing 27.5 mg/L naringenin, was named as NarPro/Δura_#20.

Probing the genomic integration site with *ARO1* gene

In synthetic biology, it is generally required to reconstruct large numbers of heterologous and/or endogenous genes to encode a functional gene cluster for various applications. Plasmid-based gene expression is often restricted by the number of auxotrophic selection markers and the length of the metabolic pathway: plasmids have a relatively constrained carrying capacity that limits the size and number of genes that could be expressed. For instance, the available auxotrophic markers in commonly used *Y. lipolytica* Po1f are Δ*URA3* and Δ*LEU2*. This means that only two plasmids can be used in one single strain, which severely restricts the number of genes that could be expressed. Although one can assemble multiple genes into one single plasmid, a larger plasmid is generally less stable and may easily cause DNA replication burden to the host cell. Site-specific integration has been developed to insert a gene at a pre-defined genomic locus by using a docking bed, because the use of a homologous arm, this process is generally not amenable to manipulate a large number of genes. In addition, the users might need to take special cautions when choosing an integration site, which should be easy for manipulation as well as not decrease the host cells' fitness. Moreover, one has to redesign the cassette to change the homologous arms in each round of integration. Unlike the site-specific gene integration, rDNA based random integration has several advantages, including the efficient curation of the genetic marker, flexible tuning of gene copy number as well as the ease of genetic operation.

We used a *Y. lipolytica* endogenous gene, *YlARO1* (GRYC ID: YALI0F12639g, 4975 bp), to validate this design. *YlARO1* encodes for a pentafunctional AroM protein that catalyzes steps 2 through 6 in the biosynthesis of chorismate, which is the precursor to synthesize L-phenylalanine and L-tyrosine.[37] In our previous work, over-expressing *YlARO1* increased flavonoid production (unpublished results). The integration of the *YlARO1* may happen at three sites: (i) at the 26s rDNA site(s) in a random manner; (ii) at the 26s rDNA adjacent to the naringenin pathway, which will remove the previously integrated naringenin pathway; (iii) at the native *YlARO1* gene site, which will disrupt the native *YlARO1* gene. Integration at site (i) is what we want, while integrations at the other two sites are not desirable. To determine the exact integration site, we used three pairs of primers to test and probe the genomic integration events. Primers TEF_F2 and $\mathrm{ARO1}_{\mathrm{veri.}}$ R prime the *TEF* promoter and *YlARO1* gene. Using primer pair TEF_F2/$\mathrm{ARO1}_{\mathrm{veri.}}$ R, the colony containing the heterologous *YlARO1* cassette will produce a 1633-bp specific band, while the colony without the *YlARO1* cassette will not. Primers PhCHS F and PhCHS R prime the heterologous *PhCHS* gene. Using primer pair PhCHS F/PhCHS R, the colony containing naringenin pathway will produce a 1167-bp specific band. Primers ARO1up_F and ARO1down_R prime the upstream and downstream sequence flanking the native *YlARO1* gene. Using primer pair ARO1up_F/ARO1down_R, the colony with intact native *YlARO1* gene will produce a 4954-bp specific band, but the heterologous *YlARO1* pathway shouldn't produce any band. These combinations of colony PCR provide critical information to validate the exact integration site of the heterologous *YlARO1* pathway.

Colony PCR results indicate that 95.8% (23/24) of the transformants contain the heterologous *YlARO1* cassette (**Figure 3A**). All of the transformants contain the heterologous naringenin pathway gene *PhCHS*, indicating that the naringenin pathway was not replaced by the *YlARO1* cassette (**Figure 3B**). The colony PCR results also showed that all these colonies contain intact *YlARO1* genes, indicating that the native *YlARO1* genes was not disrupted by the *YlARO1* cassette (**Figure 3C**). These results showed that this 26s rDNA based random integration method can be repeatedly used to integrate genes into the *Y. lipolytica* genome. Since there are more than 200 copies of 26s rDNA, we suspect that the chance for the integration events happening at the same genomic locus is rare. We randomly screened 5 colonies using YPD medium. The results showed that the naringenin titers of these strains were between 34.7 mg/L and 44.7 mg/L (**Supplementary Figure 3**). The best performing strain #14 was used in the subsequent experiments; and the *URA3* marker and pYLXP'-Cre plasmid were subsequently removed as described in the Methods section. The resulting strain was named as NarPro/ARO1_#14. Finally, NarPro/ARO1_#14 produced 43.8 mg/L naringenin, which is 59.3% higher than that of NarPro/Δura_#20 (**Figure 3D**).

The Cre-loxP system is a powerful recombination platform for efficient genome engineering[34]. Recent development of the yeast SCRaMbLE method (synthetic chromosome rearrangement and modification by loxP-mediated evolution) presents us an exciting opportunity to massively evolve the yeast genome[38]. SCRaMbLE supports a number of important features, including combinatorial mutagenesis, capable of generating complex genotypes and a broad variety of phenotypes. This synthetic and combinatorial genomic approach has recently been applied to rearrange synthetic chromosomes [39], evolve new phenotype[39], and has been adapted to a light-controlled Cre recombinase for improved recombination events in *S. cerevisiae*.[40] *In vitro* SCRaMbLE system was also successfully implemented to optimize a β-carotene pathway. [41] SCRaMbLE-based genome evolution also demonstrated versatility for both haploid/diploid yeast[42], and heterozygous diploid and interspecies.[43]

Iterative integration of flavonoids precursor pathway

To validate the method developed herein, we harnessed the efficient gene integration and marker curation to enhance malonyl-CoA synthesis. It was well demonstrated that malonyl-CoA is a limiting precursor in the production of flavonoids.[44-46] *Y. lipolytica* acetyl-CoA synthetase 2 (*YlACS2*, YALI0F05962, 2018 bp) and acetyl-CoA carboxylase 1 (*YlACC1*, YALI0C11407, 7331 bp) were used to enhance acetyl-CoA and malonyl-CoA synthesis. The fragment YlACS2-YlACC1 containing *YlACS2* and *YlACC1* in monocistronic form was transformed into NarPro/ARO1_#14. The results showed that 79.2% (19/24) of the transformants contain YlACS2-YlACC1 (**Supplementary Figure 4**). This decreased integration efficiency is possibly linked to the relatively large fragment size for the YlACS2-YlACC1 cassette (in total more than 12 kb). We randomly screened 5 colonies, and the best performing strain #14 produced 71.2 mg/L naringenin, which is 62.6% higher than that of NarPro/ARO1_#14 (**Figure 4**). The best performing strain was named as NarPro/ACS, and was used for *URA3* rescue and pYLXP'-Cre removal, following the protocol as described in the Methods section.

While there are several cytosolic acetyl-CoA pathways[47], including the native ACL (ATP-citrate lyase), and many other candidate pathways[47] should be tested. In the current study, we primarily tested the *YlACS2* gene which can directly covert acetate to acetyl-CoA. The primary reason is that acetic acid has been used as the sole carbon source to support high lipid production (up to 115 g/L) in this yeast.[48] Since acetic acid utilization may bypass glycolytic decarboxylation (PDH or pyruvate dehydrogenase complex) steps, acetic acid should be a promising carbon source to produce natural products with relatively high yield.

Functional expression of plant cytochrome *c* P450 enzyme to produce hydroxylated flavonoids

The high flux of precursors acetyl-CoA and malonyl-CoA, along with the presence of hydrophobic lipid bodies, makes oleaginous yeast a superior host to produce various natural products (NPs) with complex structures.[19] The abundant membrane structure and subcellular compartment (i.e. ER and lipid bodies) provides the hydrophobic environment that is critical for the regioselectivity and stereoselectivity of cytochrome *c* P450 enzymes to achieve high yield and efficiency.[49, 50] For these reasons, we argued that *Y. lipolytica* might be a good host to express plant-derived cytochrome *c* P450 enzymes and produce hydroxylated flavonoids.

We further used this yeast to construct eriodictyol and taxifolin producing strains. Flavonoid 3' hydroxylase from *Gerbera hybrid* (GhF3'H), which is a cytochrome *c* P450 enzyme (CYP), was used to convert naringenin to eriodictyol.[51] The cytochrome *c* P450 reductase from *Catharanthus roseus* (CrCPR) was used to transfer electrons to GhF3'H.[52] The fragment HR_{x2} (GhF3'H-CrCPR (2 copies)) which contains *GhF3'H* and *CrCPR* in monocistronic form, was transformed into NarPro/ASC. The results showed that 14/15 of the colonies contained the heterologous fragment HR_{x2} (**Supplementary Figure 5A**). The positive transformant was named as ErioPro, and the best performing strain produced 54.2 mg/L eriodyctiol, while the control strain did not produce eriodyctiol (**Figure 5A** and **Supplementary Figure 5B**).

Flavanone 3 hydroxylase from *Solanum lycopersicum* (SlF3H) was used to convert eriodictyol to taxifolin.[53] The fragment $HR_{x2}H$ (GhF3'H-CrCPR (2 copies)-SlF3H) which contains *GhF3'H*, *CrCPR*, and *SlF3H* in monocistronic form, was transformed into NarPro/ASC. The results showed that 14/15 of the colonies contained the heterologous fragment $HR_{x2}H$ (**Supplementary Figure 5C**). The positive transformant was named as TaxiPro, and the best performing strain produced 48.1 mg/L taxifolin, while the control strain did not produce taxifolin (**Figure 5B** and

Supplementary Figure 5D). This is the first report that *Y. lipolytica* could be engineered to functionally express plant P450 enzymes and produce hydroxylated flavonoids, including both eriodyctiol and taxifolin.

Conclusions

Y. lipolytica is an oleaginous yeast that internalizes substantial portion of carbon feedstocks as lipids and fatty acids. It has been recognized as a 'generally regarded recognized as safe' (GRAS) organism for the production of organic acids and natural products in the food industry. Coupled with its low pH tolerance, its strictly aerobic nature, and its ability to degrade a wide range of substrates, including hexose/pentose, glycerol, hydrocarbons and volatile fatty acids (VFAs), making this yeast an attractive candidate for industrial applications.[10, 54] The high precursor acetyl-CoA flux along with the versatile carbon-utilization capability make this yeast a superior host to upgrade low-value carbons into high value products.

Conventional plasmid-based gene expression tends to introduce genetic instability and gene copy number variations that may lead to degenerated production. The limited number of auxotrophic markers in *Y. lipolytica* also restricts our ability to perform large-scale genetic modifications and manipulate long gene clusters. Plasmid-based gene expression is generally less efficient, time-consuming and laborious. To overcome these limitations, we take advantage of both the high recombination efficiency of the Cre-loxP system and the high integration rate of the 26s rDNA, and developed a versatile framework to iteratively integrate multi-copy metabolic pathways in *Y. lipolytica*. By expanding our previous YaliBrick toolbox, we demonstrate the efficient genome integration of a plant-derived flavonoid pathway at random sites with multiple copies. Transient expression of the Cre recombinase enables efficient marker removal and allows for iterative cycle of genetic modification. Probing the recombination events demonstrate the iterative integration is happening at sufficiently high rate (more than 80%). Flavonoid precursor pathways and plant-derived P450 enzymes were functionally integrated to improve flavonoid and hydroxylated flavonoids production. The reported work provides a versatile platform to assemble metabolic pathways with varying genomic integration sites and gene copy numbers, which may simplify and streamline pathway construction, as well as expand our capability to build efficient microbial cell factories for the production of high-value natural products and commodity chemicals.

Materials and methods

Plasmids and strains

Yarrowia lipolytica Po1f (ATCC MYA-2613, MATA ura3-302 leu2-270 xpr2-322 axp2-deltaNU49 XPR2::SUC2) was used as the host strain. *Escherichia coli* NEB5α was used for plasmid construction and proliferation. The YaliBrick plasmid pYLXP' was used as the backbone to construct other plasmids. [19]

26s rDNA and loxP pasmid prDNAloxP construction

The random integration plasmid prDNAloxP was constructed as follows. The

URA3 gene was amplified from plasmid pYLXP'2, using the primer pair YlURA3 F/YlURA3 R, and the resulting PCR product was Gibson assembled with *Sna*BI digested pYLXP' to yield pYLXP'-URA3. Synthetic single stranded oligos loxP$_{rd}$ F1 and loxP$_{rd}$ R1 were annealed to yield a 79-bp double stranded DNA, loxP1. Then the loxP1 duplex DNA was Gibson assembled with the *Xba*I digested pYLXP'-URA3 to yield pYLXP'-URA3-loxP1. Synthetic single stranded oligos loxP$_{rd}$ F2 and loxP$_{rd}$ R2 were annealed to yield an 80-bp double stranded DNA, loxP2. Then the loxP2 duplex DNA was Gibson assembled with the *Spe*I digested pYLXP'-URA3-loxP1 to yield pYLXP'-URA3-loxP. The 862-bp 26s rDNA 2s fragment was amplified from the genomic DNA of *Y. lipolytica* Po1f using primer pair 26srDNA2s F/26srDNA2s R. Then the 26s rDNA 2s PCR product was Gibson assembled with the *Avr*II digested pYLXP'-URA3-loxP to yield prDNAloxP-2s. The 623-bp 26s rDNA 1s fragment was amplified from the genomic DNA of *Y. lipolytica* Po1f using primer pair 26srDNA1s F/26srDNA1s R. Then the 26s rDNA 1s PCR product was Gibson assembled with the *Not*I digested prDNAloxP-2s to yield prDNAloxP. The resulting prDNAloxP is a YaliBrick plasmid, which can utilize the isocaudarners *Avr*II, *Xba*I, *Spe*I, and *Nhe*I for fast pathway assembly. [19, 55]

Flavonoid pathway construction

Plasmids for pathway integration were constructed as follows. Flavonoid biosynthetic pathways were constructed from Gibson assembly of synthetic genes. Specifically, we have constructed a number of flavonoid pathways containing RtTAL (tyrosine ammonia lyase), Pc4CL (4-coumaroyl-CoA ligase), PhCHS (chalcone synthase), MsCHI (chalcone isomerase), GhF3’H (flavonoid 3’ hydroxylase) CrCPR (cytochrome P450 reductase), and SlF3H (flavanone 3-hydroxylase) (Data unpublished), combed with native ACS2 (acetyl-CoA synthetase), and ACC1 (acetyl-CoA carboxylase). We followed the YaliBrick guidelines and assembled plasmids pYLXP’-RtTAL-Pc4CL-PhCHS(5 copies)–MsCHI (pYLXP’-T4S$_{\text{x5}}$I), pYLXP’2-YlACS2-YlACC1, pYLXP’2-GhF3’H-CrCPR(2 copies) (pYLXP’2-HR$_{\text{x2}}$), and pYLXP’2-GhF3’H-CrCPR(2 copies)-SlF3H (pYLXP’2-HR$_{\text{x2}}$H).[19] Plasmids pYLXP’-T4S$_{\text{x5}}$I, pYLXP’2-YlACS2-YlACC1, pYLXP’2-HR$_{\text{x2}}$, and pYLXP’2-HR$_{\text{x2}}$H were digested with *Avr*II/*Sal*I. The fragments containing pathways were inserted into prDNAloxP at *Nhe*I/*Sal*I sites, resulting in prDNAloxP-T4S$_{\text{x5}}$I, prDNAloxP-YlACS2-YlACC1, prDNAloxP-HR$_{\text{x2}}$, and prDNAloxP-HR$_{\text{x2}}$H, respectively. Due to the presence of an *Avr*II site in *YlARO1*, Gibson assembly (not YaliBrick subcloning) was used to insert *YlARO1* into prDNAloxP. Specifically, the primer pair YlARO1$_{\text{GA}}$_F/YlARO1$_{\text{GA}}$_R was used to amplify *YlARO1* from pYLXP’2-YlARO1 and the PCR product was Gibson assembled with *Nhe*I/*Sal*I digested prDNAloxP to yield prDNAloxP-YlARO1. prDNAloxP-T4S$_{\text{x5}}$I, prDNAloxP-YlACS2-YlACC1, and prDNAloxP-HR$_{\text{x2}}$H were linearized using *Avr*II/*Not*I digestion before transforming *Y. lipolytica*, while prDNAloxP-YlARO1 was linearized using *Hpa*I/*Not*I digestion before transformation. *YlARO1* contains internal *Avr*II site, which is why *Hpa*I and *Not*I were chosen as the linearization enzyme pair, in order to keep the exact length of the integrated pathway. The Cre recombinase gene was codon-optimized and synthesized by Genewiz (Frederick, MD).[56] pYLXP'-Cre was constructed by Gibson assembling the Cre synthetic gene into pYLXP' at *Sna*BI site. Plasmids used in this study were listed in **Table 2**.

26s rDNA random integration and marker curation

prDNAloxP-T4S$_{x5}$I was linearized with *Avr*II/*Not*I, and the fragment containing the naringenin pathway, T4S$_{x5}$I (*RtTAL*, *Pc4CL*, 5 copies of *PhCHS* and *MsCHI* in monocistronic form), was transformed into *Y. lipolytica* Po1f. Transformants were selected on uracil dropout CSM-Ura plates. Colonies were screened by colony PCR using primer pair PhCHS F/PhCHS R. Positive colonies will yield a 1167-bp specific band, while negative colonies will not. Naringenin producing colonies were screened in YPD medium, and the best performing colony was used in the subsequent experiments. Plasmid pYLXP'-Cre was transformed into the best performing strain to rescue the *URA3* marker. Transformants were selected on CSM-Leu+Ura+5-FOA agar plates, which is complete synthetic medium without leucine but with 50 mg/L uracil and 1 g/L 5-fluoroorotic acid (5-FOA). Colonies were screened by colony PCR using primer pair 26srDNA2s F/XPR2_Seq. Positive colonies, in which the *URA3* marker was removed, will yield a 1533-bp band, while negative colonies will yield a 2432-bp band. Naringenin producing colonies were screened in YPD medium. Plasmid pYLXP'-Cre was removed by incubating the naringenin producing strain in YPD medium at 30°C for 48 h. Marker curation was confirmed by replicating the colonies onto CSM+5-FOA+Ura and CSM-Ura agar plates. Successful removal of *URA3* marker will lead to 5-FOA resistant and uracil deficient phenotype and the cells should grow on CSM+5-FOA+Ura agar plate but not on CSM-Ura agar plate.

Iterative chromosomal integration and marker curation

The best performing strain from Section 2.4 was used for *YlARO1* integration. The integration of *YlARO1* was the same as $T4S_{x5}I$, except the primers used for colony PCR screening. Primers TEF_F2 and $ARO1_{veri.}$ R, which prime the *TEF* promoter and *YlARO1* gene respectively, were used to screening positive colonies. The positive colony will yield a 1633-bp specific band, which is composed of the *TEF* promoter and part of the *YlARO1* gene. The best performing strain was used for *URA3* marker rescue and pYLXP'-Cre removal.

The same method was used for the integration of fragments YlACS2-YlACC1, HR_{x2} (GhF3'H-CrCPR (2 copies)), and $HR_{x2}H$ (GhF3'H-CrCPR (2 copies)-SlF3H). Primer pair TEF_F2/$YlACS2_{veri.}$ R was used for screening YlACS2-YlACC1 transformants, and the positive colony will produce a 1515-bp specific band. The primer pair GhF3'H F/GhF3'H R was used for screening HR_{x2} and $HR_{x2}H$ transformants, and the positive colony will produce a 1536-bp specific band.

Analytical method

Analysis of naringenin, eriodictyol, taxifolin, and glucose was performed on Agilent HPLC 1220. For naringenin, eriodictyol, and taxifolin analysis, a ZORBAX Eclipse Plus C18 column (4.6 × 100 mm, 3.5 μm, Agilent) was used for separation, and a VWD monitor (λ=280 nm) was used for detection. Mobile phase A was 0.1% (v/v) acetic acid in water, and mobile phase B was 0.1% (v/v) acetic acid in methanol. The elution gradient was as follows (B%): 0 min 0%, 5 min 80%, 8 min 80%, 12 min 0%, and 15 min 0%. The flow rate was 0.4 mL/min. The oven temperature was 30°C. For glucose analysis, a Supelcogel H HPLC column (25 cm × 4.6 mm, Sigma) was used for separation and a RID monitor was used for detection. The mobile phase was 10 mM H_2SO_4 in water. The flow rate was 0.4 mL/min. Both the oven temperature and the RID temperature were set as 50°C.

Acknowledgements

This work was supported by the Cellular & Biochem Engineering Program of the National Science Foundation under grant no.1805139. The authors would also like to acknowledge the Department of Chemical, Biochemical and Environmental Engineering at University of Maryland Baltimore County for funding support. YK would like to thank the China Scholarship Council for funding support.

ASSOCIATED CONTENT

*Supporting Information: The Supporting Information is available free of charge on the ACS Publications website. Figures S1−S5 (PDF)

AUTHOR INFORMATION

Corresponding Author E-mail: pengxu@umbc.edu

Author contributions

PX and YL designed the study. YL performed this study. PX and YL wrote the manuscript with input from HE.

Notes

The authors declare that they have no competing interests. A provisional patent has been filed based on the results of this study.

References

1. Beopoulos, A.; and Nicaud, J.-M. (2012) Yeast: A new oil producer? *OCL 19* (1), 22-28.
2. Xie, D.; Jackson, E. N.; and Zhu, Q. (2015) Sustainable source of omega-3 eicosapentaenoic acid from metabolically engineered *Yarrowia lipolytica*: from fundamental research to commercial production. *Appl. Microbiol. Biotechnol. 99* (4), 1599-610.
3. Xue, Z.; Sharpe, P. L.; Hong, S. P.; Yadav, N. S.; Xie, D.; Short, D. R.; Damude, H. G.; Rupert, R. A.; Seip, J. E.; Wang, J.; Pollak, D. W.; Bostick, M. W.; Bosak, M. D.; Macool, D. J.; Hollerbach, D. H.; Zhang, H.; Arcilla, D. M.; Bledsoe, S. A.; Croker, K.; McCord, E. F.; Tyreus, B. D.; Jackson, E. N.; and Zhu, Q. (2013) Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. *Nat. Biotechnol. 31* (8), 734-40.
4. Cordova, L. T.; and Alper, H. S. (2018) Production of α-linolenic acid in *Yarrowia lipolytica* using low-temperature fermentation. *Appl. Microbiol. Biotechnol. 102* (20), 8809-8816.
5. Zhang, B. X.; Rong, C. C.; Chen, H. Q.; Song, Y. D.; Zhang, H.; and Chen, W. (2012) *De novo* synthesis of trans-10, cis-12 conjugated linoleic acid in oleaginous yeast *Yarrowia Lipolytica*. *Microb. Cell. Fact. 11, 51.*
6. Christen, S.; and Sauer, U. (2011) Intracellular characterization of aerobic glucose metabolism in seven yeast species by ^{13}C flux analysis and metabolomics. *FEMS Yeast Res. 11* (3), 263-72.
7. Abdel-Mawgoud, A. M.; Markham, K. A.; Palmer, C. M.; Liu, N.; Stephanopoulos, G.; and Alper, H. S. (2018) Metabolic engineering in the host *Yarrowia lipolytica*. *Metab. Eng. 50*, 192-208.
8. Bailey, R.; Madden, K. T.; and Trueheart, J. (2010) Production of carotenoids in oleaginous yeast and fungi. US 7,851,199.
9. Sharpe, P. L.; Rick, W. Y.; and Zhu, Q. Q. (2014) Carotenoid production in a recombinant oleaginous yeast. US 8,846,374.
10. Ledesma-Amaro, R.; and Nicaud, J. M. (2016) Metabolic engineering for expanding

the substrate range of *Yarrowia lipolytica*. *Trends Biotechnol. 34* (10), 798-809.

11. Ledesma-Amaro, R.; Lazar, Z.; Rakicka, M.; Guo, Z.; Fouchard, F.; Coq, A.-M. C.-L.; and Nicaud, J.-M. (2016) Metabolic engineering of *Yarrowia lipolytica* to produce chemicals and fuels from xylose. *Metab. Eng. 38*, 115-124.
12. Rodriguez, G. M.; Hussain, M. S.; Gambill, L.; Gao, D.; Yaguchi, A.; and Blenner, M. (2016) Engineering xylose utilization in *Yarrowia lipolytica* by understanding its cryptic xylose pathway. *Biotechnol. Biofuels 9* (1), 149.
13. Guo, Z.-p.; Duquesne, S.; Bozonnet, S.; Cioci, G.; Nicaud, J.-M.; Marty, A.; and O'Donohue, M. J. (2017) Conferring cellulose-degrading ability to Yarrowia lipolytica to facilitate a consolidated bioprocessing approach. *Biotechnol. Biofuels 10* (1), 132.
14. Guo, Z.-p.; Robin, J.; Duquesne, S.; O'Donohue, M. J.; Marty, A.; and Bordes, F. (2018) Developing cellulolytic *Yarrowia lipolytica* as a platform for the production of valuable products in consolidated bioprocessing of cellulose. *Biotechnol. Biofuels 11* (1), 141.
15. Groenewald, M.; Boekhout, T.; Neuveglise, C.; Gaillardin, C.; van Dijck, P. W.; and Wyss, M. (2014) *Yarrowia lipolytica*: safety assessment of an oleaginous yeast with a great industrial potential. *Crit. Rev. Microbiol. 40* (3), 187-206.
16. Larroude, M.; Rossignol, T.; Nicaud, J. M.; and Ledesma-Amaro, R. (2018) Synthetic biology tools for engineering *Yarrowia lipolytica*. *Biotechnol. Adv. 36* (8), 2150-2164.
17. Hong, S.-P.; Seip, J.; Walters-Pollak, D.; Rupert, R.; Jackson, R.; Xue, Z.; and Zhu, Q. (2012) Engineering *Yarrowia lipolytica* to express secretory invertase with strong FBA1IN promoter. *Yeast 29* (2), 59-72.
18. Blazeck, J.; Liu, L.; Redden, H.; and Alper, H. (2011) Tuning Gene Expression in *Yarrowia lipolytica* by a Hybrid Promoter Approach. *Appl. Environ. Microbiol. 77* (22), 7905-7914.
19. Wong, L.; Engel, J.; Jin, E.; Holdridge, B.; and Xu, P. (2017) YaliBricks, a versatile genetic toolkit for streamlined and rapid pathway engineering in *Yarrowia lipolytica*. *Metab. Eng. Commun. 5*, 68-77.
20. Bulani, S. I.; Moleleki, L.; Albertyn, J.; and Moleleki, N. (2012) Development of a novel rDNA based plasmid for enhanced cell surface display on *Yarrowia lipolytica*. *AMB Express 2* (1), 27.
21. Le Dall, M.-T.; Nicaud, J.-M.; and Gaillardin, C. (1994) Multiple-copy integration in the yeast *Yarrowia lipolytica*. *Curr. Genet. 26* (1), 38-44.
22. Gao, S.; Han, L.; Zhu, L.; Ge, M.; Yang, S.; Jiang, Y.; and Chen, D. (2014) One-step integration of multiple genes into the oleaginous yeast *Yarrowia lipolytica*. *Biotechnol. Lett. 36* (12), 2523-2528.
23. Juretzek, T.; Le Dall, M.-T.; Mauersberger, S.; Gaillardin, C.; Barth, G.; and Nicaud, J.-M. (2001) Vectors for gene expression and amplification in the yeast *Yarrowia lipolytica*. *Yeast 18* (2), 97-113.

24. Bordes, F.; Fudalej, F.; Dossat, V.; Nicaud, J. M.; and Marty, A. (2007) A new recombinant protein expression system for high-throughput screening in the yeast *Yarrowia lipolytica*. *J. Microbiol. Methods 70* (3), 493-502.
25. Fickers, P.; Le Dall, M. T.; Gaillardin, C.; Thonart, P.; and Nicaud, J. M. (2003) New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. *J. Microbiol. Methods 55* (3), 727-737.
26. Wagner, J. M.; Williams, E. V.; and Alper, H. S. (2018) Developing a piggyBac Transposon System and Compatible Selection Markers for Insertional Mutagenesis and Genome Engineering in *Yarrowia lipolytica*. *Biotechnol. J. 13* (5), 1800022.
27. Schwartz, C. M.; Hussain, M. S.; Blenner, M.; and Wheeldon, I. (2016) Synthetic RNA polymerase III promoters facilitate high-efficiency CRISPR-Cas9-mediated genome editing in *Yarrowialipolytica*. *ACS Synth. Biol. 5, 356-359.*
28. Holkenbrink, C.; Dam, M. I.; Kildegaard, K. R.; Beder, J.; Dahlin, J.; Doménech Belda, D.; and Borodina, I. (2018) EasyCloneYALI: CRISPR/Cas9-Based Synthetic Toolbox for Engineering of the Yeast *Yarrowia lipolytica*. *Biotechnol J. 13* (9), 1700543.
29. Morse, N. J.; Wagner, J. M.; Reed, K. B.; Gopal, M. R.; Lauffer, L. H.; and Alper, H. S. (2018) T7 Polymerase Expression of Guide RNAs in vivo Allows Exportable CRISPR-Cas9 Editing in Multiple Yeast Hosts. *ACS Synth. Biol. 7* (4), 1075-1084.
30. Englaender, J. A.; Jones, J. A.; Cress, B. F.; Kuhlman, T. E.; Linhardt, R. J.; and Koffas, M. A. G. (2017) Effect of genomic integration location on heterologous protein expression and metabolic engineering in *E. coli*. *ACS Synth. Biol. 6* (4), 710-720.
31. Jensen, N. B.; Strucko, T.; Kildegaard, K. R.; David, F.; Maury, J.; Mortensen, U. H.; Forster, J.; Nielsen, J.; and Borodina, I. (2014) EasyClone: method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res. 14* (2), 238-248.
32. Magnan, C.; Yu, J.; Chang, I.; Jahn, E.; Kanomata, Y.; Wu, J.; Zeller, M.; Oakes, M.; Baldi, P.; and Sandmeyer, S. (2016) Sequence Assembly of *Yarrowia lipolytica* Strain W29/CLIB89 Shows Transposable Element Diversity. *PLOS ONE 11* (9), e0162363.
33. Madzak, C.; Tréton, B.; and Blanchin-Roland, S. (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mo.l Microbiol. Biotechnol. 2,* 207-216.
34. Guldener, U.; Heck, S.; Fiedler, T.; Beinhauer, J.; and Hegemann, J. H. (1996) A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res. 24* (13), 2519-2524.
35. Vanheerikhuizen, H.; Ykema, A.; Klootwijk, J.; Gaillardin, C.; Ballas, C.; and Fournier, P. (1985) Heterogeneity in the ribosomal RNA genes of the yeast *Yarrowia*

Lipolytica: Cloning and analysis of two size classes of repeats. *Gene 39* (2-3), 213-222.

36. Zhang, N.; Li, J.; Li, F.; and Wang, S. A. (2018) Selectable marker recycling in the nonconventional yeast *Xanthophyllomyces dendrorhous* by transient expression of Cre on a genetically unstable vector. *Appl. Microbiol. Biotechnol.* doi: 10.1007/s00253-018-9496-1.
37. Duncan, K.; Edwards, R. M.; and Coggins, J. R. (1987) The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. *Biochem J 246* (2), 375-386.
38. Dymond, J. S.; Richardson, S. M.; Coombes, C. E.; Babatz, T.; Muller, H.; Annaluru, N.; Blake, W. J.; Schwerzmann, J. W.; Dai, J.; Lindstrom, D. L.; Boeke, A. C.; Gottschling, D. E.; Chandrasegaran, S.; Bader, J. S.; and Boeke, J. D. (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature 477*, 471-476.
39. Luo, Z.; Wang, L.; Wang, Y.; Zhang, W.; Guo, Y.; Shen, Y.; Jiang, L.; Wu, Q.; Zhang, C.; Cai, Y.; and Dai, J. (2018) Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. *Nat. Commun. 9* (1), 1930.
40. Hochrein, L.; Mitchell, L. A.; Schulz, K.; Messerschmidt, K.; and Mueller-Roeber, B. (2018) L-SCRaMbLE as a tool for light-controlled Cre-mediated recombination in yeast. *Nat. Commun. 9* (1), 1931.
41. Wu, Y.; Zhu, R.-Y.; Mitchell, L. A.; Ma, L.; Liu, R.; Zhao, M.; Jia, B.; Xu, H.; Li, Y.-X.; Yang, Z.-M.; Ma, Y.; Li, X.; Liu, H.; Liu, D.; Xiao, W.-H.; Zhou, X.; Li, B.-Z.; Yuan, Y.-J.; and Boeke, J. D. (2018) *In vitro* DNA SCRaMbLE. *Nat. Commun. 9* (1), 1935.
42. Jia, B.; Wu, Y.; Li, B.-Z.; Mitchell, L. A.; Liu, H.; Pan, S.; Wang, J.; Zhang, H.-R.; Jia, N.; Li, B.; Shen, M.; Xie, Z.-X.; Liu, D.; Cao, Y.-X.; Li, X.; Zhou, X.; Qi, H.; Boeke, J. D.; and Yuan, Y.-J. (2018) Precise control of SCRaMbLE in synthetic haploid and diploid yeast. *Nat. Commun. 9* (1), 1933.
43. Shen, M. J.; Wu, Y.; Yang, K.; Li, Y.; Xu, H.; Zhang, H.; Li, B.-Z.; Li, X.; Xiao, W.-H.; Zhou, X.; Mitchell, L. A.; Bader, J. S.; Yuan, Y.; and Boeke, J. D. (2018) Heterozygous diploid and interspecies SCRaMbLEing. *Nat. Commun. 9* (1), 1934.
44. Fowler, Z. L.; Gikandi, W. W.; and Koffas, M. A. (2009) Increased malonyl coenzyme A biosynthesis by tuning the *Escherichia coli* metabolic network and its application to flavanone production. *Appl. Environ. Microbiol. 75* (18), 5831-9.
45. Xu, P.; Ranganathan, S.; Fowler, Z. L.; Maranas, C. D.; and Koffas, M. A. G. (2011) Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metab. Eng. 13* (5), 578-587.
46. Leonard, E.; Yan, Y.; Fowler, Z. L.; Li, Z.; Lim, C. G.; Lim, K. H.; and Koffas, M. A. (2008) Strain improvement of recombinant *Escherichia coli* for

efficient production of plant flavonoids. *Mol. Pharm. 5* (2), 257-265.

47. Xu, P.; Qiao, K.; Ahn, W. S.; and Stephanopoulos, G. (2016) Engineering *Yarrowia lipolytica* as a platform for synthesis of drop-in transportation fuels and oleochemicals. *Proc. Nat. Aca. Sci. 113* (39), 10848-10853.
48. Xu, J.; Liu, N.; Qiao, K.; Vogg, S.; and Stephanopoulos, G. (2017) Application of metabolic controls for the maximization of lipid production in semicontinuous fermentation. *Proc. Nat. Aca. Sci. 114* (27), E5308-E5316.
49. Munro, A. W.; McLean, K. J.; Grant, J. L.; and Makris, T. M. (2018) Structure and function of the cytochrome P450 peroxygenase enzymes. *Biochem. Soc. Trans. 46(1),* 183-196.
50. Nthangeni, M. B.; Urban, P.; Pompon, D.; Smit, M. S.; and Nicaud, J. M. (2004) The use of Yarrowia lipolytica for the expression of human cytochrome P450 CYP1A1. *Yeast 21* (7), 583-92.
51. Zhu, S.; Wu, J.; Du, G.; Zhou, J.; and Chen, J. (2014) Efficient synthesis of eriodictyol from L-tyrosine in *Escherichia coli*. *Appl. Environ. Microbiol. 80* (10), 3072-3080.
52. Meijer, A. H.; Lopes Cardoso, M. I.; Voskuilen, J. T.; de Waal, A.; Verpoorte, R.; and Hoge, J. H. (1993) Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH:cytochrome P-450 reductase, an enzyme essential for reactions catalysed by cytochrome P-450 mono-oxygenases in plants. *Plant J. 4* (1), 47-60.
53. Aoki, K.; Yano, K.; Suzuki, A.; Kawamura, S.; Sakurai, N.; Suda, K.; Kurabayashi, A.; Suzuki, T.; Tsugane, T.; Watanabe, M.; Ooga, K.; Torii, M.; Narita, T.; Shin, I. T.; Kohara, Y.; Yamamoto, N.; Takahashi, H.; Watanabe, Y.; Egusa, M.; Kodama, M.; Ichinose, Y.; Kikuchi, M.; Fukushima, S.; Okabe, A.; Arie, T.; Sato, Y.; Yazawa, K.; Satoh, S.; Omura, T.; Ezura, H.; and Shibata, D. (2010) Large-scale analysis of full-length cDNAs from the tomato (*Solanum lycopersicum*) cultivar Micro-Tom, a reference system for the Solanaceae genomics. *BMC genomics 11*, 210.
54. Liu, H.-H.; Ji, X.-J.; and Huang, H. (2015) Biotechnological applications of *Yarrowia lipolytica*: Past, present and future. *Biotechnol. Adv. 33* (8), 1522-1546.
55. Xu, P.; Vansiri, A.; Bhan, N.; and Koffas, M. (2012) ePathBrick: A Synthetic Biology Platform for Engineering Metabolic Pathways in E-coli. *ACS Synth. Biol. 1* (7), 256-266.
56. Lambert, J. M.; Bongers, R. S.; and Kleerebezem, M. (2007) Cre-lox-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. *Appl. Environ. Microbiol. 73* (4), 1126-35.

Tables

Table 1 Primers used in this study

Primers	Sequence (5’-3’)
YlURA3 F	CCAGCACTTTTTGCAGTACTAACCGCAGCCCTCCTACGAAGCTCGAGC
YlURA3 R	CATAGCACGCGTGTAGATACCTAACAGTTAATCTTCTGGTAAGCC
loxP$_{rd}$ F1	AGTATAAGAATCATTCAAATCTAGAATAACTTCGTATAGCATACATTATACGAAGTTAT
loxP$_{rd}$ R1	CTGCCTCTGAAACTCACCATATAACTTCGTATAATGTATGCTATACGAAGTTAT
loxP$_{rd}$ F2	GATTAACTGTTAGGGTACCGATAACTTCGTATAGCATACATTATACGAAGTTAT
loxP$_{rd}$ R2	ACGTGGGGACAGGCCATGGAACTAGTATAACTTCGTATAATGTATGCTATACGAAGTTAT
26srDNA2s F	ATCCCTAAATTTGATGAAAGCCTAGGCAGACACTGCGTCGCTCCGTCC
26srDNA2s R	CCAACCCGGTCTCTGTCGTCTGCTTCGGTATGATAGGAAGAGC
26srDNA1s F	AGCTTTACCGCAGCAGATCCAGATCTTGGTGGTAGTAGC

26srDNA1s R	ATCCACTATTGGCCTATGCGGCCGCCCGCGGGTCCGGCTGCCAGTTGCCC
YlARO1$_{GA}$_F	CCGTTATCAAATCTAGTTAGCTAGGGACGACAGAGACCGG
YlARO1$_{GA}$_R	CTTTTATCAGACATAGTCGACTCCTCCGTTATTGTCTCGCTAGC
PhCHS F	GACCAGCACTTTTTGCAGTACTAACCGCAGGTTACGGTGGAAGAATACCGC
PhCHS R	GGGACAGGCCATGGAACTAGTCGTTAGGTAGCCACACTATGCAGAACC
GhF3’H F	ACGCCTCTTACACTCCTGATTGG
GhF3’H R	TTAGACCTTGGTCGTTTCATATACG
TEF_F2	GGGTATAAAAGACCACCGTCCCC
XPR2_Seq	GGTGTTGGACTCAGTAATAAGAGCC
ARO1$_{veri.}$ R	CACTCGAACGATGCGCCACCAAGC
YlACS2$_{veri.}$ R	AGCCAGTCTCGGTCTGCCAGTAGG
ARO1up_F	ATCTCCAACTACACACCACC
ARO1down_R	CAATCATTTTGATGTCTCGG

Table 2 Plasmids used in this study

Plasmid	Annotation	Reference
pYLXP'	A YaliBrick plasmid, used as backbone for integration plasmid.	[19]
pYLXP'2	A YaliBrick plasmid, used for amplifying *YlURA3* gene.	[19]
prDNAloxP	26s rDNA based integration plasmid.	This study
pYLXP'-T4S$_{x5}$I	pYLXP' containing *RtTAL*, *Pc4CL*, *PhCHS* (5 copies), and *MsCHI* in monocistronic form. "$_{x5}$" refers to 5 copies.	Unpublished data
pYLXP'2-YlACS2-YlACC1	pYLXP'2 containing *YlACS2* and *YlACC1* in monocistronic form.	Unpublished data
pYLXP'2-YlARO1	pYLXP'2 containing *YlARO1*	Unpublished data
pYLXP'2-HR$_{x2}$	pYLXP'2 containing *GhF3'H* and *CrCPR* (2 copies) in monocistronic form. "$_{x2}$" refers to 2 copies.	Unpublished data
pYLXP'2-HR$_{x2}$H	pYLXP'2 containing *GhF3'H*, *CrCPR* (2 copies), and *SlF3H* in monocistronic form. "$_{x2}$" refers to 2 copies.	Unpublished data

prDNAloxP-T4S$_{x5}$I	prDNAloxP containing T4S$_{x5}$I.	This study
prDNAloxP-YlACS2-YlACC1	prDNAloxP containing YlACS2-YlACC1.	This study
pYLXP'2-HR$_{x2}$	pYLXP'2 containing HR$_{x2}$.	This study
prDNAloxP-HR$_{x2}$H	prDNAloxP containing HR$_{x2}$H.	This study
prDNAloxP-YlARO1	prDNAloxP containing *YlARO1*.	This study
pYLXP'-Cre	pYLXP' containing Cre gene.	This study

Figures

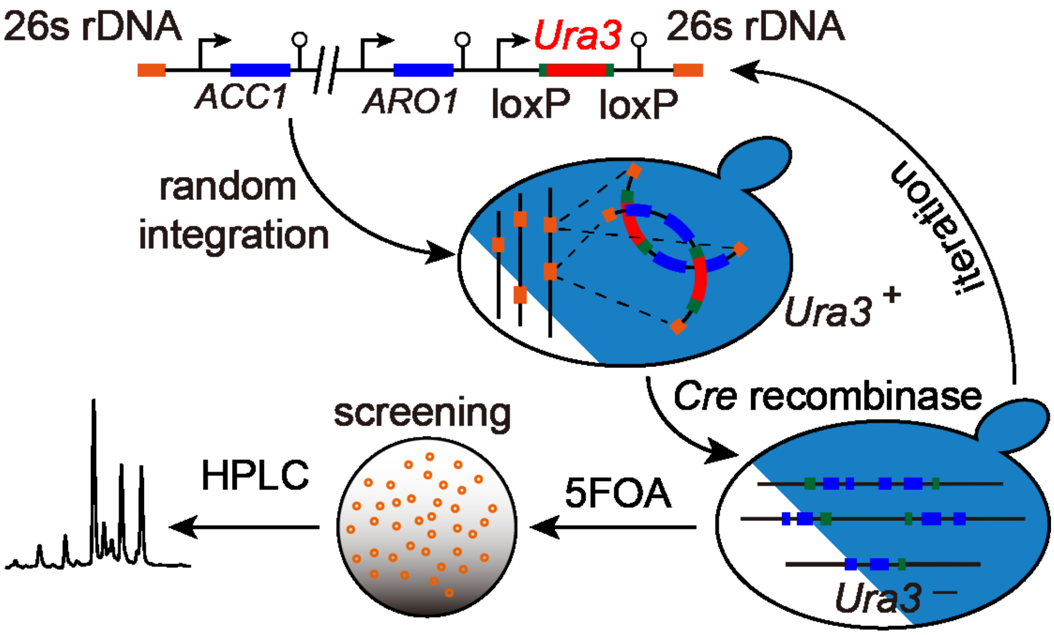


Figure 1. Demonstration of the integration method. The 26s rDNA will target gene(s) into the genome at random site(s). URA3 was used as the selective marker, and the positive transformant will grow on the uracil dropout plate. The URA3 marker with flanking loxP sites will be rescued by transforming a Cre recombinase expressing plasmid pYLXP'-Cre. The colonies can be screened by colony PCR, and the high-performing strains can be obtained by screening their performance.

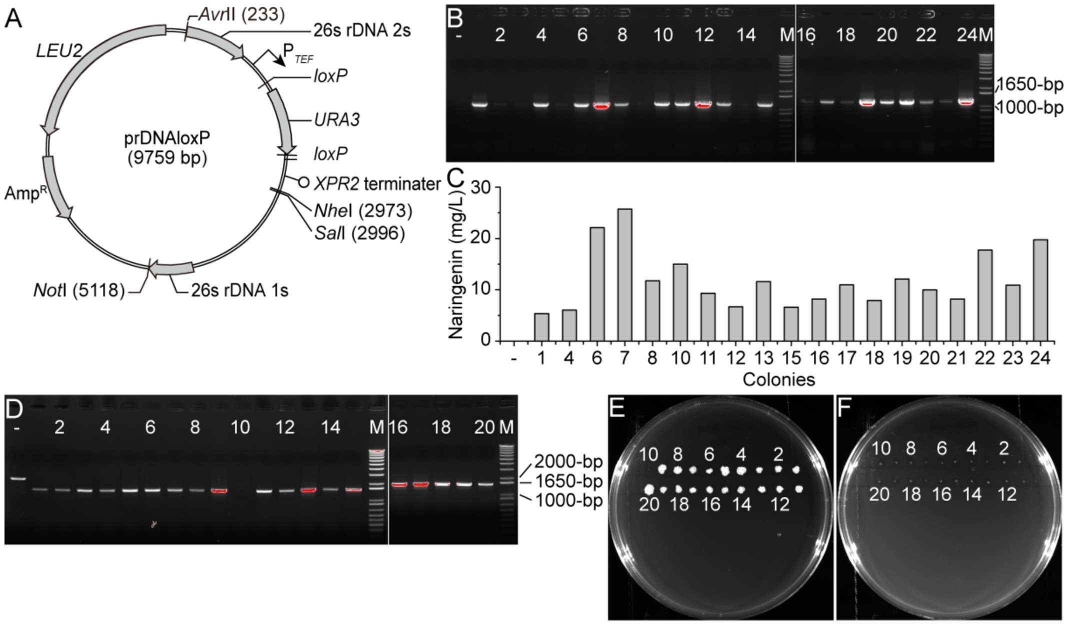


Figure 2. Analysis of random integration.

A. The map of the random integration plasmid prDNAloxP. **B.** Colony PCR analysis of naringenin pathway integration. The heterologous gene *PhCHS* was used to validate the integration. Positive colony will produce a 1167-bp band, while negative colony will not produce any band. "–" refers to the negative control, which is a wild-type *Y. lipolytica* Po1f colony. **C.** Screening of naringenin producing strains in YPD medium. "-" refers to the negative control, which is a wild-type *Y. lipolytica* Po1f strain. **D.** Colony PCR analysis of *URA3* marker rescue. Positive colony will produce 1533-bp band, while negative colony will produce 2432-bp band. "-" refers to the negative control, which is a colony without pYLXP'-Cre. **E.** Analysis of *URA3* marker recuration on SCM+5-FOA+Ura. **F.** Analysis of *URA3* marker recuration on SCM-Ura.

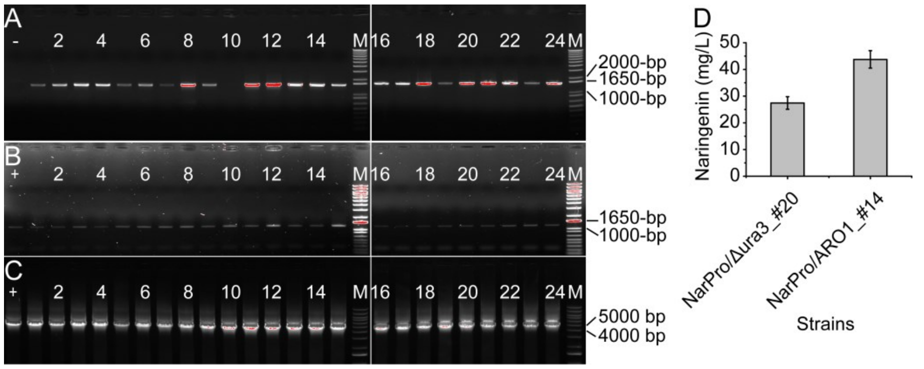


Figure 3. Analysis of repeated using of the random integration method.

A. Colony PCR analysis of ARO1 cassette integration. Colonies containing the heterologous ARO1 cassette will produce a 1633-bp specific band. "-" refers to negative control, which is a NarPro_#8 colony without transforming ARO1 cassette. **B.** Colony PCR analysis of whether the naringenin pathway was removed. Primer pair PhCHS F/PhCHS R was used. The colony containing naringenin pathway will produce a 1167-bp specific band, indicating the naringenin pathway was not removed. "+" refers to the positive control, which is a NarPro_#8 colony without transforming ARO1 cassette. **C.** Colony PCR analysis of whether the native *ARO1* gene was disrupted. The colony with intact native *ARO1* gene will produce a 4954-bp specific band, indicating that the native *ARO1* gene was not disrupted. "+" refers to the positive control, which is a NarPro_#8 colony without transforming ARO1 cassette. **D.** Analysis of the production stability of NarPro/ARO1_#14 using 3 independent experiments. NarPro/ARO1_#14 is the beat performent strain after screening.

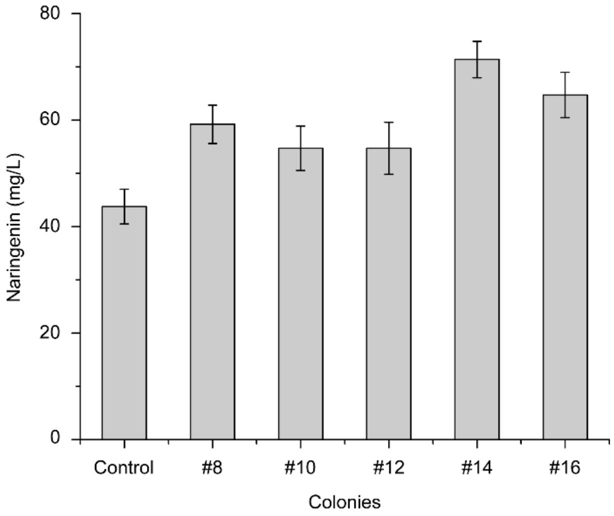


Figure 4. Screening of YlACS2-YlACC1 transformants in YPD medium. Five colonies were picked into YPD medium randomly. Fermentation was carried out in 250-mL shaking flasks for 144 h. NarPro/ARO1_#14 was used as control.

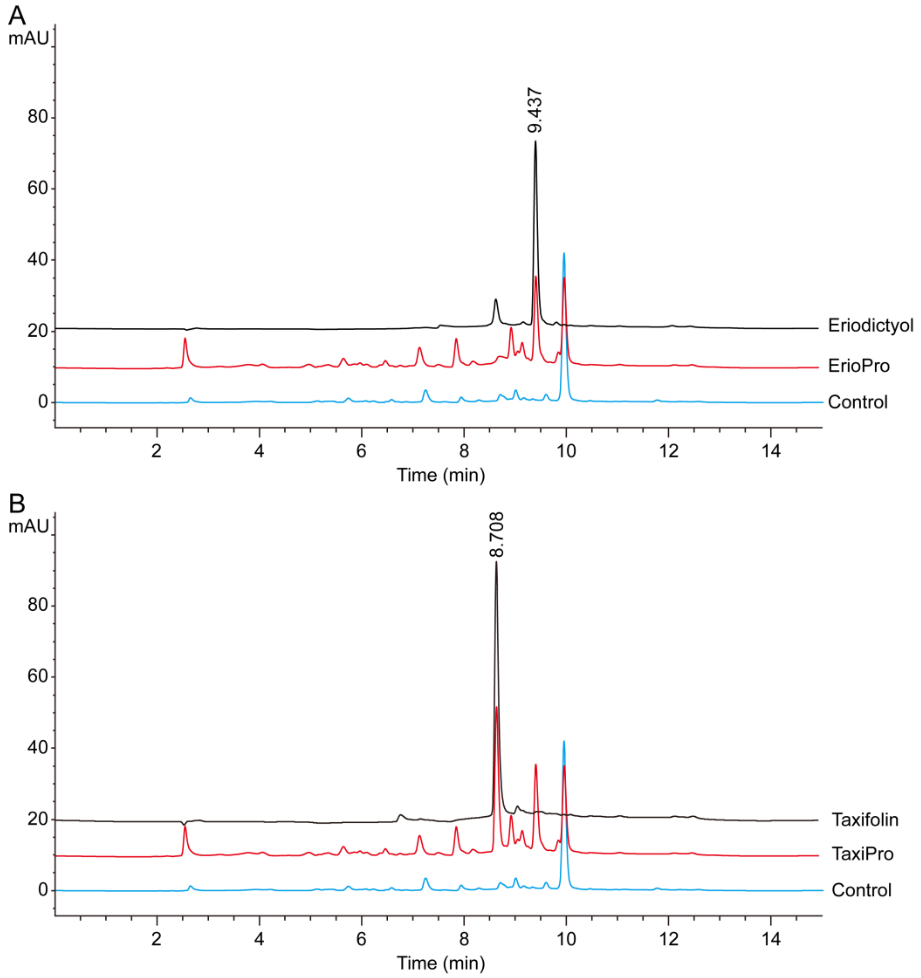
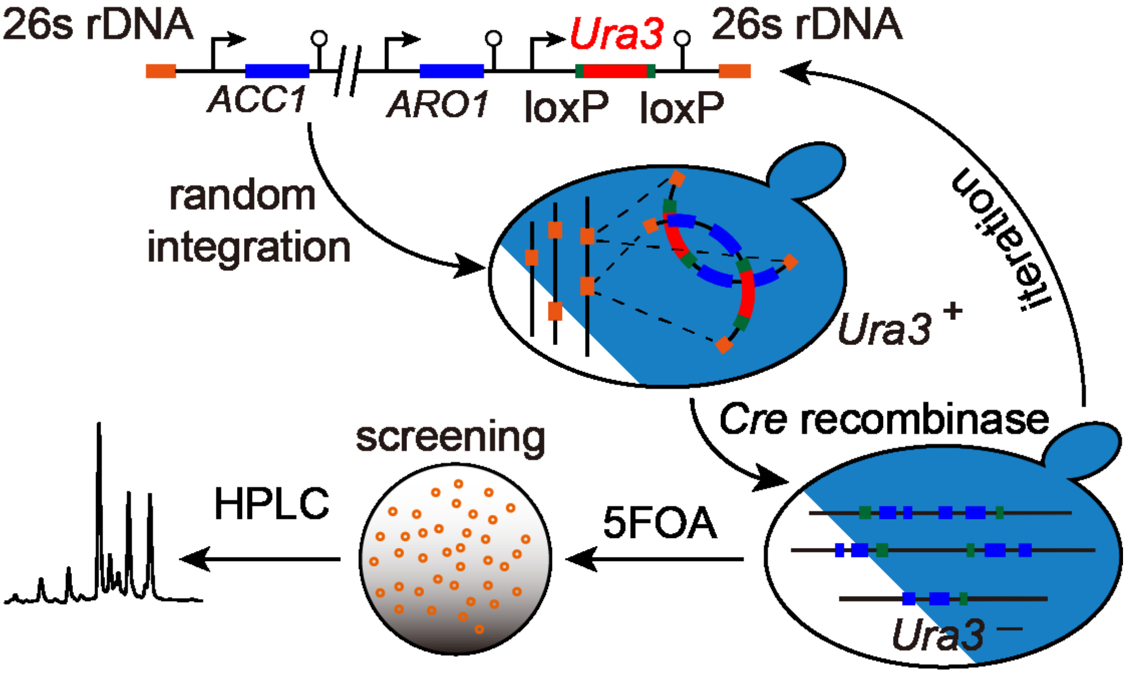


Figure 5. HPLC analysis of eriodictyol and taxifolin production. **A.** HPLC analysis of eriodictyol production. The retaintion time of eriodictyol was 9.437 min. **B.** HPLC analysis of taxifolin production. The retaintion time of taxifolin was 8.708 min. For both analysis, NarPro/ARO1_#14 was used as control.



Graphic Table of Content