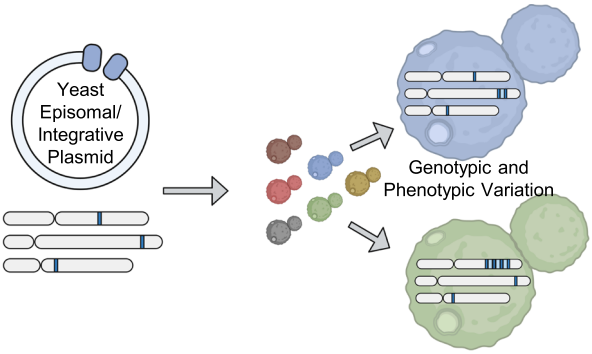
Integration of Yeast Episomal/Integrative Plasmid Causes Genotypic and Phenotypic Diversity and Improved Sesquiterpene Production in Metabolically Engineered *Saccharomyces cerevisiae*

Bingyin Peng,* Sarah J. Weintraub, Zeyu Lu, Samuel Evans, Qianyi Shen, Liam McDonnell, Manuel Plan, Thomas Collier, Li Chen Cheah, Lei Ji, Christopher B. Howard, Will Anderson, Matt Trau, Geoff Dumsday, Erin L. Bredeweg, Eric M. Young, Robert Speight, and Claudia E. Vickers

ABSTRACT: The variability in phenotypic outcomes among biological replicates in engineered microbial factories presents a captivating mystery. Establishing the association between phenotypic variability and genetic drivers is important to solve this intricate puzzle. We applied a previously developed auxin-inducible depletion of hexokinase 2 as a metabolic engineering strategy for improved nerolidol production in *Saccharomyces cerevisiae*, and biological replicates exhibit a dichotomy in nerolidol production of either 3.5 or 2.5 g L^{-1} nerolidol. Harnessing Oxford Nanopore's long-read genomic sequencing, we reveal a potential genetic cause—the chromosome integration of a 2μ sequence-based yeast episomal plasmid, encoding the expression cassettes for nerolidol synthetic enzymes. This finding was reinforced through chromosome integration revalidation, engineering nerolidol and valencene production strains, and generating a diverse pool of yeast clones, each uniquely fingerprinted by gene copy numbers, plasmid integrations, other genomic rearrangements, protein expression levels, growth rate, and target product productivities. The best clone in two strains produced 3.5 g L^{-1} nerolidol and ~0.96 g L^{-1} valencene. Comparable genotypic and phenotypic variations were also generated through the integration of a yeast integrative plasmid lacking 2μ sequences. Our work shows that multiple factors, including plasmid integration status, subchromosomal location, gene copy number, sesquiterpene synthase expression level, and genome rearrangement, together play a complicated determinant role on the productivities of sesquiterpene product. Integration of yeast episomal/integrative plasmids may be used as a versatile method for increasing the diversity and optimizing the efficiency of yeast cell factories, thereby uncovering metabolic control mechanisms.



KEYWORDS: *yeast engineering, metabolic engineering, genetic variation, nanopore sequencing, genome rearrangement*

INTRODUCTION

In metabolic engineering and synthetic biology, microbial cell factories are commonly developed through a stepwise Design-Build-Test-Learn process.[1−3] In this process, transgenes are transformed into chassis cells to construct and optimize metabolic pathways and genetic regulatory networks.[4,5] In each step of transformation, many clones of recombinant cells, known as biological replicates, are produced.[6] Biological replicates often vary in phenotypes (traits), such as productivities (titer/yield/rate) for a certain product.[3,6,7] Although the best clone can be selected for application purposes, genetic drivers for the superior trait are typically not investigated.[3,8−10] In this scenario, detecting genetic variations in biological replicates is important to establish a better knowledge about the association of genes and phenotypes.[3,6,11,12]

The budding yeast *Saccharomyces cerevisiae* is a common chassis broadly used in metabolic engineering and synthetic biology studies.[13] Engineering in *S. cerevisiae* benefits from its well-studied genetics and well-developed genetic engineering tools.[14] Many types of genetic vectors have been developed to introduce transgenes into *S. cerevisiae*, including yeast episomal plasmid (or 2μ sequence-based plasmid),[15] yeast replicative plasmid,[16] yeast integrative plasmid,[17] the multicopy integrative plasmids targeting to rDNA or δ sequences,[18] and the multicopy integrative plasmid using haploinsufficiency-driven

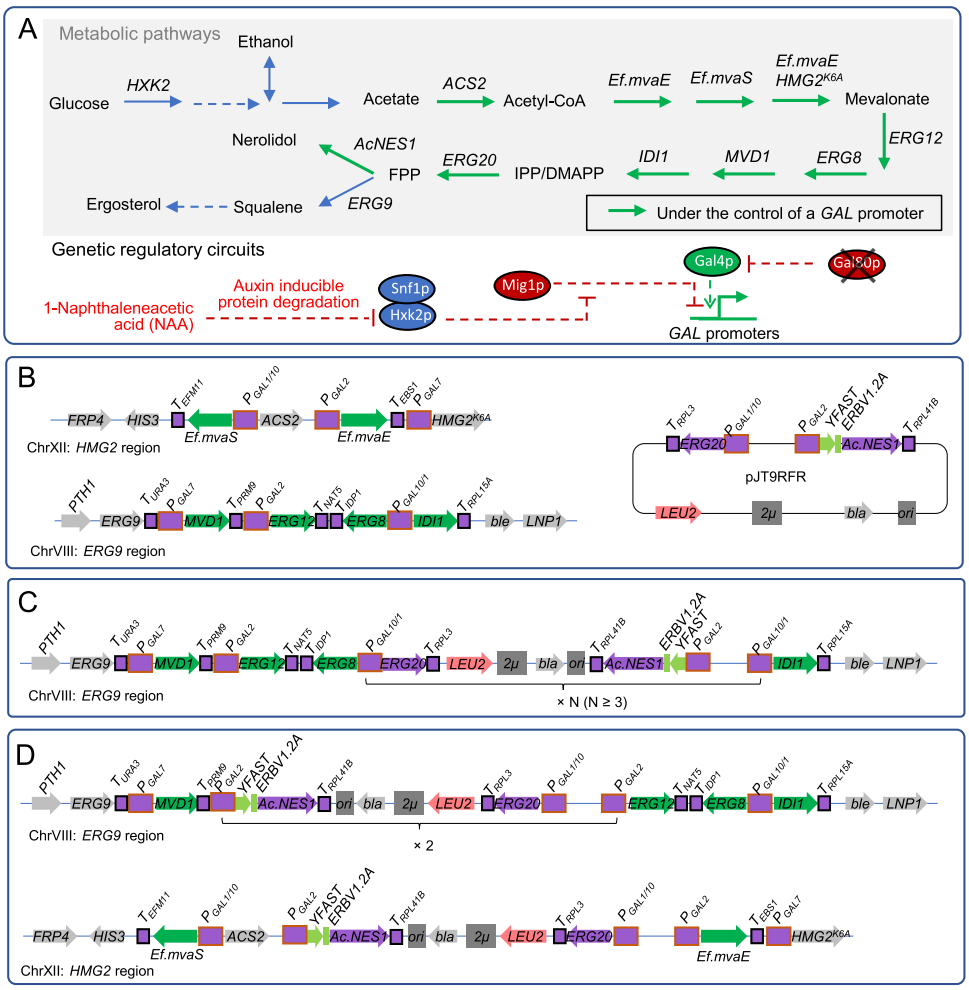


Figure 1. Genetic modifications introduced in metabolically engineered *S. cerevisiae* for nerolidol production. (A) Diagrams of metabolic pathways for nerolidol synthesis and genetic regulatory circuits for the control of *GAL* promoters. (B) Genetic features in original designs for the development of *S. cerevisiae* strain NLD128. (C) Genetic features found in ~3.5 g L^{-1} nerolidol-producing NLD128 clones NLD128H-1 and NLD128H-2. (D) Genetic features found in ~2.5 g L^{-1} nerolidol-producing NLD128 clones NLD128L-1 and NLD128L-2. Gene/enzyme: *HXK2/Hxk2p*, hexokinase 2; *ACS2*, acetyl-CoA synthase; *Ef.mvaE*, *Enterococcus faecalis* acetyl-CoA acetyltransferase/HMG-CoA reductase; *Ef.mvaS*, *E. faecalis* HMG-CoA synthase; *HMG2*^{K6A}, HMG-CoA synthase K6A mutant; *ERG12*, mevalonate kinase; *ERG8*, phosphomevalonate kinase; *MVD1*, mevalonate pyrophosphate decarboxylase; *IDI1*, isopentenyl diphosphate (IPP):dimethylallyl diphosphate (DMAPP) isomerase; *ERG20*, farnesyl pyrophosphate (FPP) synthetase; *Ac.NES1*, *Actinidia chinensis* trans-nerolidol synthase; *ERG9*, squalene synthase; Snf1p, AMP-activated S/T protein kinase; Mig1p, transcription factor involved in glucose repression; Gal4p, transcription activator for activating *GAL* gene; Gal80p, transcription repressor inhibiting Gal4p. P_{XXXN}, the promoter of defined gene; T_{XXXN}, the terminator of defined gene.

gene amplification (HapAmp).[19] Each of these plasmids has specific defined properties, including copy number and mitotic stability. These properties are used as engineering design standards and have been continuously validated in a multitude of applications. Nevertheless, significant phenotypic variations in biological replicates can still be observed.

Some specific variations have been reported and understood, for example, integration through δ sequences,[10,20] whereas other variations have not been well understood. Previously, we applied an auxin-inducible protein degradation mechanism to deplete hexokinase 2 for improved sequiterpene (nerolidol) production, which resulted in the strain with two kinds of phenotypes in biological replicates.[8] To develop this strain, the genes encoding the enzymes from the mevalonate pathway were overexpressed under the control of *GAL* promoters; and hexokinase 2 was modified with an auxin-inducible protein degradation mechanism (Figure 1A). This was followed by transformation of a 2μ sequence-based episomal plasmid that harbored *GAL* promoter-controlled expression cassettes of farnesyl pyrophosphate synthase (*ERG20*) and *trans*-nerolidol synthase (*Ac.NES1*); and the knocking out of *GAL* repressor Gal80 gene (Figure 1A). After 2μ plasmid transformation and *GAL80* disruption, two groups of biological replicates were obtained. One group of biological replicates produced ~3.5 g L^{-1} nerolidol, and another group of biological replicates produced ~2.5 g L^{-1} nerolidol. We have repeated the process of 2μ plasmid transformation and *GAL80* disruption and consistently obtained two groups of biological replicates. The genetic causes for this variation were not clear.

Whole genome sequencing is a powerful method to detect the high-resolution DNA sequence features in defined samples.[11] It builds the foundation of genomics studies to unravel the genetic cause of a biofunction, including an association of genes and traits. The advances in DNA sequencing techniques, for example, Oxford nanopore sequencing, make whole genome sequencing more acces-

sible.[1,21] With the availability of analytics platforms, software packages, and analysis pipelines, whole genome sequencing has been integrated in the Design-Build-Test-Learn strain engineering cycle to provide a better validation of the resulted genotypes.[11,22]

In this study, we sequenced the whole genomes of the biological replicates that produce either 3.5 or 2.5 g L^{-1} nerolidol through the long-read Oxford nanopore sequencing technique. Unexpectedly, the yeast episomal plasmid harboring the expression cassettes of farnesyl pyrophosphate synthase (Erg20p) and nerolidol synthase was integrated into the background strain chromosome. The integration happened at the loci of the *GAL* promoters that control the genes from the mevalonate pathway. In the replicates producing 3.5 g L^{-1} nerolidol, the integration, through the recombination between the *GAL1* promoters, formed at least three tandemly repeated copies of the whole 2μ plasmid. We validated the integration of 2μ plasmid through the *GAL1* promoter for optimized production of sesquiterpenes nerolidol and valencene, attempted to investigate the genetic causes driving the integration events, and investigated the integration using a haploinsufficiency-driven gene amplification vector and a yeast integrative plasmid that does not harbor a 2μ sequence. The genotypic and phenotypic variations in the biological replicates were examined.

RESULTS

Genomic Integration of the 2μ Sequence-Based Episomal Plasmid in Nerolidol-Producing Strain. We previously developed an *S. cerevisiae* strain (ILHA NLD128-1) that produced the sesquiterpene nerolidol at ~3.5 g L^{-1} in flask cultivation.[8] This strain was developed from CEN.PK2−1C[23] through the seven steps of sequential genetic modifications. (Step 1) A *GAL2* promoter-controlled mevalonate kinase gene (*ERG12*), a *TEF2* promoter-controlled phosphomevalonate kinase gene (*ERG8*), a *ADH2* promoter-controlled mevalonate pyrophosphate decarboxylase gene (*MVD1*), and a *TEF1* promoter-controlled isopentenyl diphosphate:dimethylallyl diphosphate isomerase gene (*IDI1*) were integrated into *pdc5* locus.[24] (Step 2) A hydroxy methylglutaryl (HMG)-CoA synthase gene (*Ef.mvaS*) and a acetyl-CoA synthase gene (*ACS2*) under the control of bidirectional *GAL1/GAL10* promoters, a bifunctional Acetyl-CoA acetyltransferase/HMG-CoA reductase gene (*Ef.mvaE*) under the control of the *GAL2* promoter, the *GAL7* promoter integrated at the upstream of HMG-CoA reductase 2 gene (*HMG2*), and a K6A mutation was introduced to *HMG2* (Figure 1a; ChrXII: *HMG2* region).[24] (Step 3) a *GAL7* promoter-controlled *MVD1*, a *GAL2* promoter-controlled *ERG12*, and the *ERG8* and *IDI1* under the control of bidirectional *GAL10/GAL1* promoters were integrated into the downstream of squalene synthase gene *ERG9* (Figure 1B; ChrVIII: *ERG9* region).[25] (Step 4) the fusion of a minimal auxin-inducible degron and a metallothionein *Cup1p* were tagged to the carboxyl-terminal of hexokinase 2 (Hxk2p).[8] (Step 5) the fusion of *SKP1* gene and auxin receptor gene (*TIR1*) under the control of the *ACS2* promoter was integrated into *ura3* locus.[8] (Step 6) A *GAL1* promoter-controlled farnesyl pyrophosphate gene (*ERG20*), and the fusion of yellow fluorescence-activating and absorption-shifting tag (YFAST), 2A sequence (ERBV1.2A), and nerolidol synthase (*Ac.NES1*) were introduced through a 2μ sequence-based episomal plasmid (Figure 1B; pJT9RFR).[8] (Step 7) the *GAL* repressor gene *GAL80* was disrupted. For the modifications at steps 1−5, only one confirmed yeast clone was stored and used in the next-step modification. For the modification at Step 6, >30 transformants were mixed and used in the next-step *GAL80* disruption. After *GAL80* disruption at step 7, two types of biological replicates were observed.[8] One type (NLD128H) was two/3-fold higher than that of another type (NLD128L) in YFAST fluorescence, and NLD128H clones produced 3.5 g L^{-1} nerolidol and NLD128L clones produced 2.5 g L^{-1} nerolidol. The repeated experiment for step 6−7 modifications generated the same types of biological replicates.

A 2- to 3-fold difference of YFAST fluorescence between NLD128H clones and NLD128L clones suggested that there be a difference in the copy numbers of the 2μ plasmid pJT9RFR (Figure 1B).[8] Aiming to unravel the genetic causes for such differences, we performed whole genome sequencing. We first considered the long-read Oxford Nanopore sequencing. We sequenced two NLD128H clones (NLD128H-1 and NLD128H-2) and two NLD128L (NLD128L-1 and NLD128L-2) clones. The sequences were *de novo* assembled and analyzed using a Prymetime pipeline.[11] The engineered sequences were mapped in the context of the entire genome assembly using BLASTN. The mapping results quite obviously showed that *YFAST*, *Ac.NES1*, and other features from plasmid pJT9RFR were localized in the region between the *ERG9* C-terminal and the *ERG9* terminator (Supplementary data 1). These features are shown as three repeats in NLD128H-1 and NLD128H-2 and a single repeat in NLD128L-2. For NLD128L-1, they were found localized in the region near *ERG9* C-terminal as two tandemly repeated copies and in the region between *Ef.mvaE* and *Ef.mvaS* as a single copy. This indicates the integration of pJT9RFR with varying copy numbers at different genome loci. The integration of the 2μ plasmid into the chromosome was not expected. To further confirm this, we reassembled the whole genome sequences using the Canu assembler and annotated all genetic features.[26,27] The results are consistent (Supplementary Figures 1−5).

In clones NLD128H-1 and NLD128H-2, pJT9RFR might be integrated through the recombination between the *GAL1* promoter in the plasmid and the *GAL1* promoter at the engineered *ERG9* locus and formed three tandemly repeated copies or more (Figure 1C). In clones NLD128L-1, pJT9RFR might be integrated through the recombination between the *GAL2* promoter in the plasmid and the *GAL2* promoter in the engineered *ERG9* locus or *HMG2* locus. However, the *HMG2* locus without pJT9RFR integration was also found in NLD128L-1 (Supplementary Figure 6). These data may indicate cross-contamination during sample preparation for NLD128L-1. In NLD128L-2, pJT9RFR was integrated through the recombination between the *GAL2* promoter in the plasmid and the *GAL2* promoter in the engineered *ERG9* locus and formed two tandemly repeated copies.

In summary, long-read Nanopore genome sequencing results show that the plasmid pJT9RFR was integrated into the yeast chromosome, and the superior nerolidol production in NLD128H clones might be related to the *GAL1* promoter-mediated multicopy chromosome integration of pJT9RFR plasmid.

Integration of the Episomal Plasmid Causes Genotypic and Phenotypic Variations in the Nerolidol-Producing Strain. Yeast episomal plasmids that contain a 2μ plasmid fragment for plasmid replication have been used in

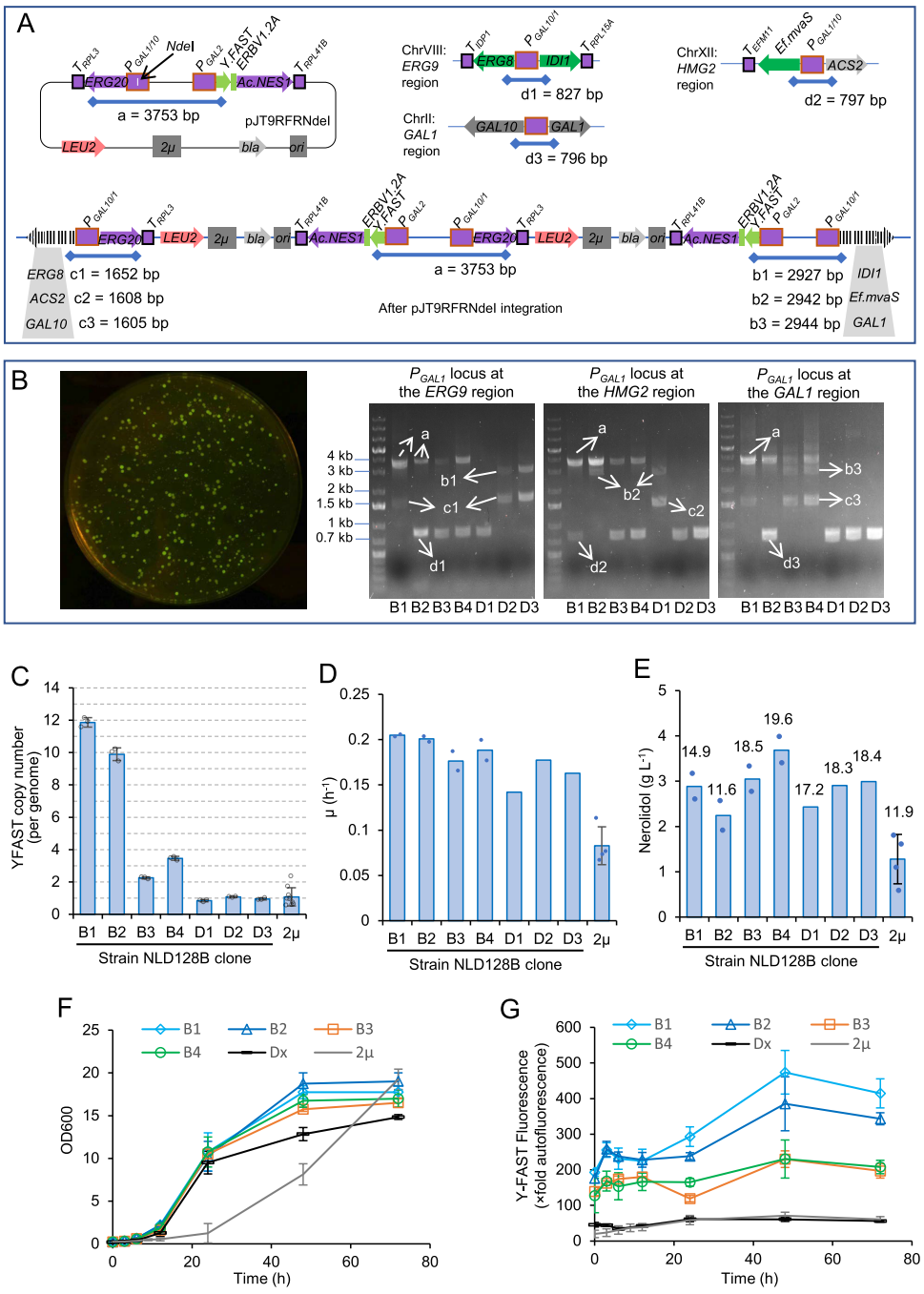


Figure 2. Genotypic and phenotypic variations in nerolidol-producing strains. (A) Genetic features in episomal plasmid pJT9RFRNdeI, three *GAL1/10* promoter loci in background strain o128RB, and *GAL1*/10 promoter loci after pJT9RFRNdeI integration. Fragments a, b1/2/3, c1/2/3, and d1/2/3 are the PCR products in PCR product length polymorphism analysis. (B) YFAST fluorescence in yeast NLD128B colonies transformed with *Nde*I-linearized pJT9RFRNdeI and the agarose gel separation of yeast colony PCR products. (C) YFAST copy numbers in the NLD128B clones and strain NLD128B2μ (2μ; transformed with circular pJT9RFRNdeI). Mean values ± standard deviation are shown (n = 4 technical replicates; and 4 biological replicates and 4 technical replicates for 2μ). (D–G) characterization of NLD128B clones and the NLD128B2μ (2μ) in flask cultivation: μ (the maximal specific exponential growth rate), nerolidol titer at 72 h, growth curve, and YFAST fluorescence. The numbers above the bars in E are specific nerolidol production rates (mg g^{-1} biomass h^{-1}). Yeast cells were grown in synthetic minimal YNB media containing 20 g L^{-1} of glucose and 100 mM MES/ammonia buffer. The numbers above the bar in (E) are the overall specific production rates (mg g^{-1} biomass h^{-1}). Mean values and mean values ± absolute error are shown for B1–B4 (n = 2 independent cultivation; D and E). For D1–D3, n = 1 independent cultivation in D and E. Mean values ± standard deviation for Dx (D1–D3) in F and G and 2μ (n = 4 biological replicates) in (D–G). YFAST fluorescence was measured using an Accuri C6 flow cytometer.

S. cerevisiae engineering for more than four decades.[15,28] To date, it has been assumed that they are episomal, i.e., remaining as a part of the eukaryotic genome without chromosome integration. The whole genome sequencing indicates the integration of the episomal plasmid pJT9RFR. It is not clear whether the integration is required for pJT9RFR maintenance or what drives pJT9RFR integration.

We attempted to revalidate the integration-phenotype phenomenon through alternative engineering designs for Steps 6 and 7. Continuing to above Step 5, there were two steps of genetic modifications. (Step 6-1) We transformed an integrative plasmid to introduce the tetracycline-mediated derepression mechanism to control *GAL80* expression to generate the strain o128RB.[5] In the absence of tetracycline, the *GAL80* transcription is repressed. This leads to the genetic effects allowing the induction of *GAL* promoters, as does *GAL80* disruption. In the presence of tetracycline, *GAL80* transcription is derepressed and causing the repression of *GAL* promoters. (Step 7-1) The plasmid pJT9RFR was transformed and the transformants were plated on the YNB 20 g L^{-1} glucose agar with 4-hydroxy-3-methylbenzylidene-rhodanine (HMBR; the fluorogen generating YFAST fluorescence)[29] and without tetracycline. However, we could not see the colonies with distinguishable fluorescence by using a Blue-Light Transilluminator. This is not consistent with the observation of the bright fluorescence from the above pJT9RFR-integrated NLD128H clones.[8] We postulated that alternative engineering designs might not result in NLD128H genotypes and pJT9RFR integration genotypes.

We hypothesized that the linearization of pJT9RFR plasmid through a restriction enzyme digestion in the middle of the *GAL1/10* promoter might facilitate chromosome integration in the background strain o128RB. Through mutating a single base, we designed a *Nde*I site in the middle of the *GAL1/10* promoter to generate a new plasmid pJT9RFRNdeI (Figure 2A). *Nde*I-digested pJT9RFRNdeI plasmid was transformed into strain o128RB, and the transformants were plated on the YNB-glucose-HMBR plate. The colonies showed a very clear fluorescence variation (Figure 2A). We postulated that pJT9RFRNdeI was successfully integrated through the recombination at the *GAL1/10* promoter. In strain o128RB, there are three *GAL1/10* promoter loci (Figure 2A). They are the endogenous *GAL1/10* promoter between gene *GAL1* and gene *GAL10* in chromosome II, the *GAL1/10* promoter used to *ERG8* and *IDI1* in the chromosome VIII *ERG9* region, and the *GAL1/10* promoter used to *Ef.mvaS* and *ACS2* in the chromosome XII *HMG2* region. We designed three sets of primers to allow the examination of the integration events at the *GAL1/10* loci through the PCR product length polymorphism (Figure 2A,B). Agarose gel electrophoresis was used to examine the PCR products in each reaction (Figure 2B). We examined four colonies showing a bright fluorescence (Figure 2B: B1−B4) and three colonies showing a dim fluorescence (Figure 2B: D1−D3). The PCR product patterns indicate that (1) the linearized pJT9RFR was integrated with these seven clones; (2) the integration could happen at any of the *GAL1/10* promoter loci (shown by the pattern of bands b1−b3, c1−c3, and d1−d3); and (3) the integration might result in a tandem repeat in clones B1−B4 but not in clones D1−D3 (shown by the existence of band a). Noted are that pJT9RFRNdeI was integrated with two *GAL1/10* loci in the *ERG9* region and the *GAL1* region in clone B1 and that the integration in the *HMG2* region in clone B2 might couple with unknown events due to the lack of band c2. We examined the copy number of the YFAST gene using quantitative real-time PCR (Figure 2C). Consistent with the indication by the yeast colony PCR results, clones D1−D3 contain a single copy of YFAST gene per copy of genome and clones B1−B4 contain 2 to 12 copies of YFAST gene per copy of genome. These data confirmed that a 2μ sequence-based plasmid can be integrated into chromosome DNA in the current background strain o128RB and that the integration causes the variations in genotype, including copy number.

We further evaluated these seven clones for nerolidol production in flask cultivation (Figure 2D−G). For reference, we also transformed the circular pJT9RFRNdeI plasmid into background strain o128RB to generate clones 2μ. Clones 2μ contain a single copy of YFAST gene per copy of genome (Figure 2C). Previously, the plasmid pJT9RFR (not having the *Nde*I site in the *GAL1* promoter) showed 14 copies per genome in the isogenic *HXK2*-wildtype background strain.[19] The decrease in the plasmid copy number might be related to the modification of auxin-inducible Hxk2p degradation, which resulted in the induction of the *GAL* promoters even in the absence of auxin.[8] But the cause remains unclear. To be consistent with previous cultivation conditions for auxin-inducible depletion of hexokinase 2 (Hxk2p), the flask cultivation was performed in synthetic minimal YNB media containing 20 g L^{-1} glucose. 2-(*N*-morpholino)ethanesulfonic acid (MES; 100 mM) was added as the buffer and the initial pH was adjusted to 6.0 using ammonia hydroxide. Auxin analogue 1-naphthalene acetic acid (1 mM) was added to induce Hxk2p depletion.

The 2μ clones transformed with the circular pJT9RFRNdeI showed a lag or slow growth in the first 24 h, and the integration clones (B1−B4 and D1−D3) grew faster than 2μ clones (Figure 2E,F). This may indicate that the circular form of plasmid pJT9RFRNdeI results in growth deficiency in the current metabolically engineered background, and chromosome integration may alleviate this deficient effect.

The YFAST fluorescence levels were similar in 2μ clones and the single-copy integration clones (D1−D3) across the 72 h cultivation. In multicopy integration clones (B1−B4), YFAST fluorescence levels and copy numbers are positively correlated. YFAST copy numbers indicate the copy numbers of the "episomal" plasmid pJT9RFRNdeI (Figure 2C). YFAST fluorescence can be used as a proxy to indicate the expression levels of the nerolidol synthase (*Ac.NES1*; Figure 2A). In this case, the YFAST fluorescence is not linearly correlated with the nerolidol titer at 72 h (Figure 2E,G). The best nerolidol productivities (titer and rate) were seen in clone B4. Clone B4, harboring three copies of integrated plasmid pJT9RFRNdeI, produced ~3.5 g L^{-1} nerolidol at 72 h. Others, harboring one or two copies or ≥10 copies with higher levels of nerolidol synthase expression, produced less. 2μ clones only produced less than half the amount of nerolidol in comparison to the single-copy integration clones, indicating the plasmid integration contributed to the improved production. Although the plasmid pJT9RFRNdeI or pJT9RFR was integrated at different loci in previous NLD128H clones and current clone B4, ~3.5 g L^{-1} nerolidol production and coupled three-copy plasmid integration were revalidated.

We further evaluated the plasmid stability in multicopy integration clones (B1, B3, and B4), single-copy integration clones (D1, D2, and D3), and 2μ clones (n = 3). After ~10-to-11-generation growth in YPD broth, only ~6% of cells for 2μ clones were recoverable on minimal mineral agar (YNB-glucose), whereas the recovery rates were ~84% for multicopy integration clones and ~93% for single-copy integration clones (Supplementary Table 1). This indicates that pJT9RFRNdeI was not mitotically stable in 2μ clones, and such instability may support the idea that pJT9RFRNdeI was episomal, not integrated into chromosomes.

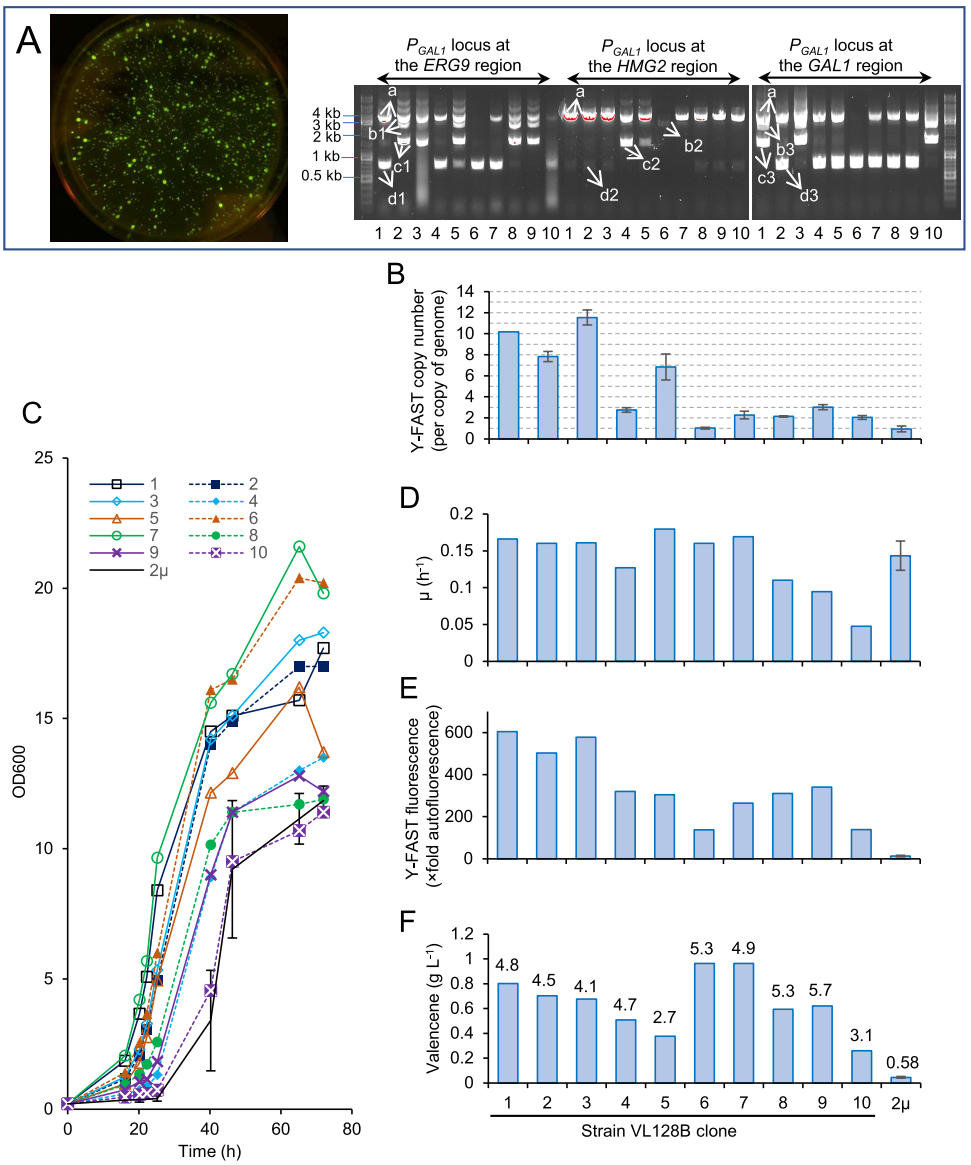


Figure 3. Genotypic and phenotypic variations in valencene-producing strains. (A) YFAST fluorescence in yeast VL128B colonies transformed with *Nde*I-linearized pJTEgVS9RRNdeI and the agarose gel separation of yeast colony PCR products. Fragments a, b1/2/3, c1/2/3, and d1/2/3 are the PCR products shown in Figure 2A. (B) YFAST copy numbers in the VL128B clones and strain VL128B2μ (2μ; transformed with circular pJTEgVS9RRNdeI). Mean values ± standard deviation are shown (n = 4 technical replicates for 2−10; and 4 biological replicates for 2μ). (C−F) characterization of VL128B clones and strain VL128B2μ (2μ) in flask cultivation: growth curve, μ (the maximal specific exponential growth rate), YFAST fluorescence at 72 h, and valencene titer at 72 h. The numbers above the bar in E are the overall specific production rate (mg g^{-1} biomass h^{-1}). Yeast cells were grown in synthetic minimal YNB media containing 20 g L^{-1} glucose and 100 mM MES/ammonia buffer. For 1−10, n = 1 independent cultivation). For 2μ, mean ± standard deviations are shown (n = 4 biological replicates). YFAST fluorescence was measured using an Accuri C6 flow cytometer.

Episomal Plasmid Integration for Genotypic and Phenotypic Variations in Valencene-Producing Strain. The integration of the nerolidol synthase-expressing episomal plasmid led to the genotypic and phenotypic variations, which showed that three-copy integration was optimal for nerolidol production. To further investigate the application of this method, we used valencene production as another example. We replaced the nerolidol synthase gene *Ac.NES1* with an *Eryngium glaciale* valencene synthase gene mutant *Eg.VS*$^{I533V\ R336K\ K325E\ R306K\ D176E\ H196R}$ (*Eg.VS**),[30] which generated the plasmid pJTEgVS9RFNdeI. *Nde*I-digested pJTEgVS9RFNdeI was transformed into the background strain o128RB, which produced colonies showing varied YFAST fluorescence on the YNB-glucose-HMBR plate (Figure 3A). Ten colonies were analyzed through yeast colony PCR. The PCR product length polymorphism (PCR products are the same as those in Figure 2A) indicates the integration of pJTEgVS9RFNdeI at the *GAL1/10* promoter loci (Figure 3A). The colonies containing 1−12 copies of integrated plasmids, shown by YFAST copy number per copy of genome (Figure 3B), were obtained. The YFAST fluorescence at 72 h positively correlated with the YFAST copy number (Figure 3B,E). Clones 6 and 7 were shown to have pJTEgVS9RFNdeI integrated at the *GAL1/10* promoter locus in the *HMG2* region, but the band c2 (Figure 3A) was missing, potentially indicating unknown genome rearrangement events. The

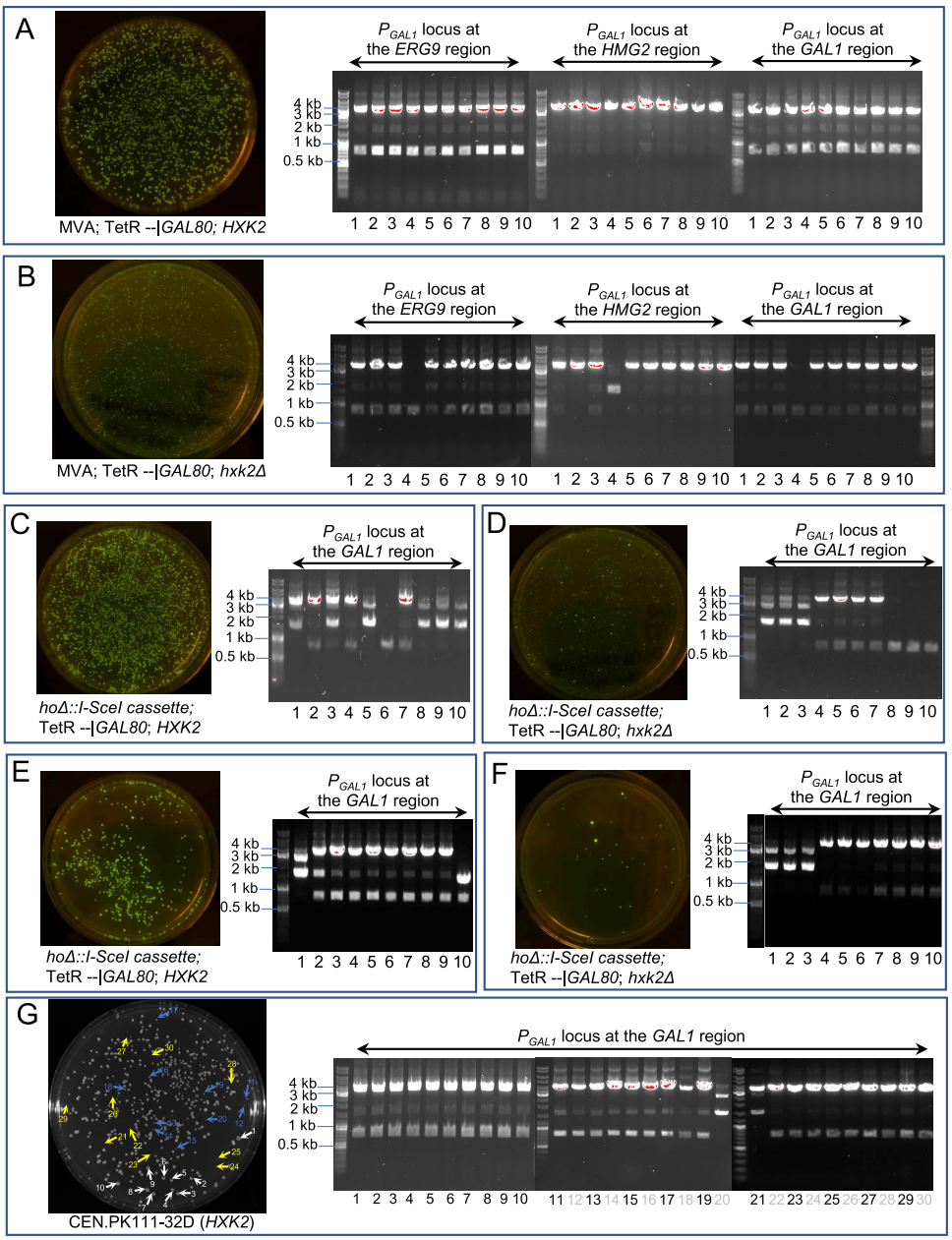


Figure 4. Transformation of linearized episomal plasmids into *HXK2*-wildtype or *hxk2*Δ background strains and agarose gel separation of yeast colony PCR products. (A–D, and G) Transformation of *Nde*I-digested pJT9RFRNdeI. (E and F) Transformation of *Nde*I-digested pJTEgVS9RNdeI. YFAST fluorescence in yeast colonies is shown in (A–F). PCR products are the same as those in Figures 1 and 2. The first three clones in D and F are the bright clones. Background strains are listed in Supplementary data 2. Genotype: MVA, augmented mevalonate pathway; TetR--|*GAL80*, *GAL80* transcription is repressed by TetR-Tup1 repressor; *hoΔ::* I-*SceI* cassette, *HO* is disrupted by a meganuclease *I-SceI* expression cassette.

circular pJTEgVS9RFNdeI was transformed into the same background strain to generate clones 2μ, which showed instability in the absence of selection pressure (Supplementary Table 1). Integration clones 1–10 and circular plasmid clones 2μ were further characterized through flask cultivation.

Similarly to nerolidol-producing 2μ clones (Figure 2F), valencene-producing 2μ clones contained a single copy of the plasmid per copy of genome and showed a lag or delayed growth in the first 24 h (Figure 3c), although the maximal growth rates were approximately 0.15 h^{-1} (Figure 3D). Different from nerolidol-producing 2μ clones (Figure 2E,G), the YFAST fluorescence and valencene titer at 72 h were quite low in valencene-producing 2μ clones (Figure 3E,F).

Among the valencene-producing clones, clones 6 and 7 showed the best productivities (titer and rate). They produced ~0.96 g L^{-1} valencene at 72 h, although they varied in YFAST copy number and fluorescence levels. Clones 6 and 7 showed higher biomass accumulation than the other eight clones. Clones 1, 2, 3, and 5, containing ≥7 copies of plasmid integration, showed medium biomass accumulation, and clones 4, 8, 9, and 10, containing ≤3 copies of plasmid integration, showed low biomass accumulation. These eight clones produced moderate levels of valencene with a titer of 0.26 to 0.8 g L^{-1}. Among these eight clones and strain 2μ, the valencene titers and YFAST fluorescence at 72 h show a positive correlation (R^2 = 0.88; Supplementary Figure 7).

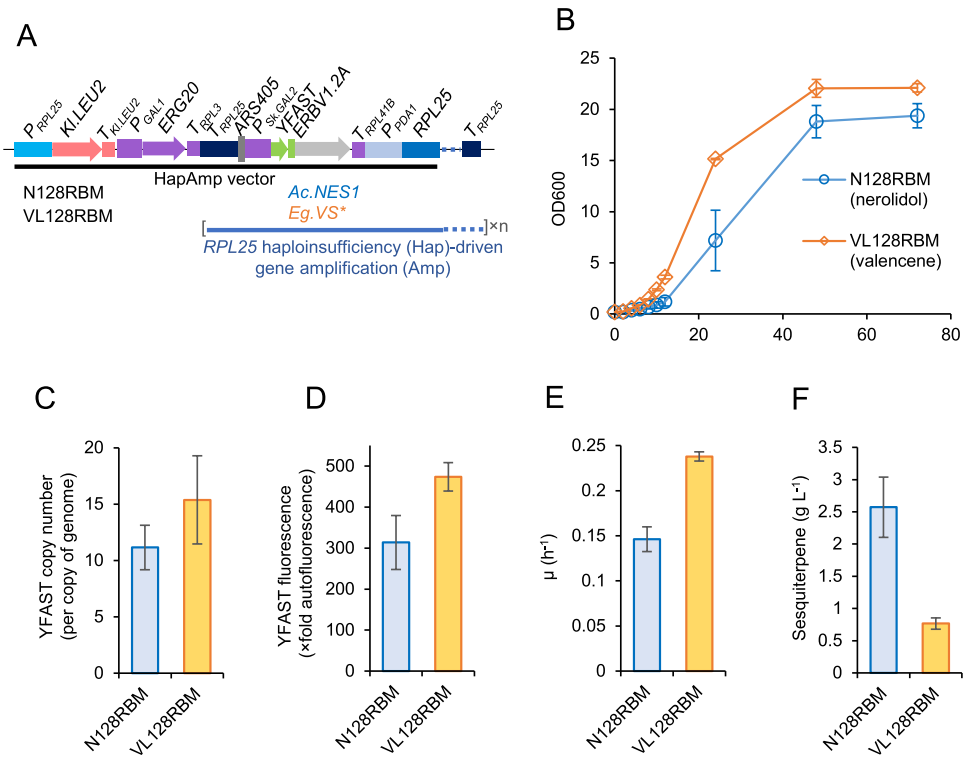


Figure 5. Characterization of nerolidol-producing strain NLD128RBM and valencene-producing strain VL128RBM in flask cultivation. (A) Genetic features for the *RPL25*-driven gene amplification of nerolidol synthase (*Ac.NES1*) or valencene synthase (*Eg.VS**) expression cassette. (B) Growth curve. (C) YFAST copy number. (D) YFAST fluorescence at 72 h. (E) The maximal specific growth rate (μ). (F) Nerolidol/valencene titer at 72 h. Yeast cells were grown in synthetic minimal YNB media containing 20 g L^{-1} glucose and 100 mM MES/ammonia buffer. YFAST fluorescence was measured using an Accuri C6 flow cytometer. Mean values ± standard deviation are shown (n = 4 biological replicates).

Missing c2 bands in yeast colony PCR indicated the unknown genome rearrangements in the best production clones (clones 6 and 7; Figure 3A). This band was designed to confirm the engineering feature around *ACS2* at the *HMG2* locus. We postulated that the missing c2 band was the result of the homologous recombination between the *GAL2* promoter controlling *Ef.mvaE* and the *GAL2* promoter controlling *YFAST-2A-Eg.VS** (Supplementary Figure 8). This postulation was confirmed by yeast colony PCR (Supplementary Figure 8) and whole genome sequencing (Supplementary Figures 9 and 10). The genome assembly shows three-copy integration of plasmid pJTEgVS9RFNdeI for clone 7, but the sequencing coverage supports two-copy integration (Supplementary Figure 10). This further suggests that missing *ACS2* under the control of the *GAL10* promoter may be beneficial for valencene production.

In summary, genotypic and phenotypic variations were characterized for the integration of valencene synthase-expressing episomal plasmid.

Hexokinase 2 Deficiency Is Not Sufficient to Cause, and Is not a Prerequisite for, Chromosome Integration of Yeast Episomal Plasmid. The integration of 2μ sequence-based episomal plasmid was validated in the background strain engineered with auxin-inducible protein degradation of hexokinase 2. It is not clear whether the integration of the episomal plasmid is caused by compromised hexokinase 2 function, which may be caused by the leaky degradation through the auxin-inducible protein degradation mechanism.[8] To investigate this hypothesis, we transformed *Nde*I-digested pJT9RFRNdeI into an isogenic background strain o401RUB that does not contain the engineered modules for auxin-inducible degradation of hexokinase 2 (Figure 4A). Yeast colony PCR results indicate that the linearized plasmid was not integrated into the *GAL1/10* promoter loci. We then transformed the linearized plasmid into the isogenetic background stain o401RUB with *HXK2* disruption. One clone, out of ten clones, contained the integrated plasmid. This indicates that the *HXK2* deficiency is not sufficient to cause the integration of the yeast episomal plasmid.

We further tested the integration of *Nde*I-digested pJT9RFRNdeI in other *HXK2*-wildtype or *hxk2*Δ background strains that were genetically engineered with the integration of a meganuclease I-*Sce*I expression cassette at the *ho*Δ locus and the tetracyline-mediated *GAL80* derepression module (Figure 4C−F). Integration of the linearized plasmids was observed in both the *HXK2*-wildtype and the *hxk2*Δ background. Consistent with this, using *Nde*I-digested pJTEgVS9RFNdeI generated similar results. The integration clones cannot be distinguished by the YFAST fluorescence in the *HXK2*-wildtype background. The integration clones were those showing a larger colony size and bright YFAST fluorescence in the *hxk2*Δ background. This shows a growth advantage caused by episomal plasmid integration, which consists of the observation of the growth of nerolidol/valencene-producing clones (Figures 2 and 3). These results confirm that *HXK2* deficiency is not sufficient to cause the integration of the yeast episomal plasmid and suggest that *HXK2* deficiency is not required for such integration.

To confirm whether the linearized yeast episomal plasmid can be integrated into *HXK2* background, we transformed *Nde*I-digested pJT9RFRNdeI into the background CEN.PK111−32D. We first tested ten colonies showing

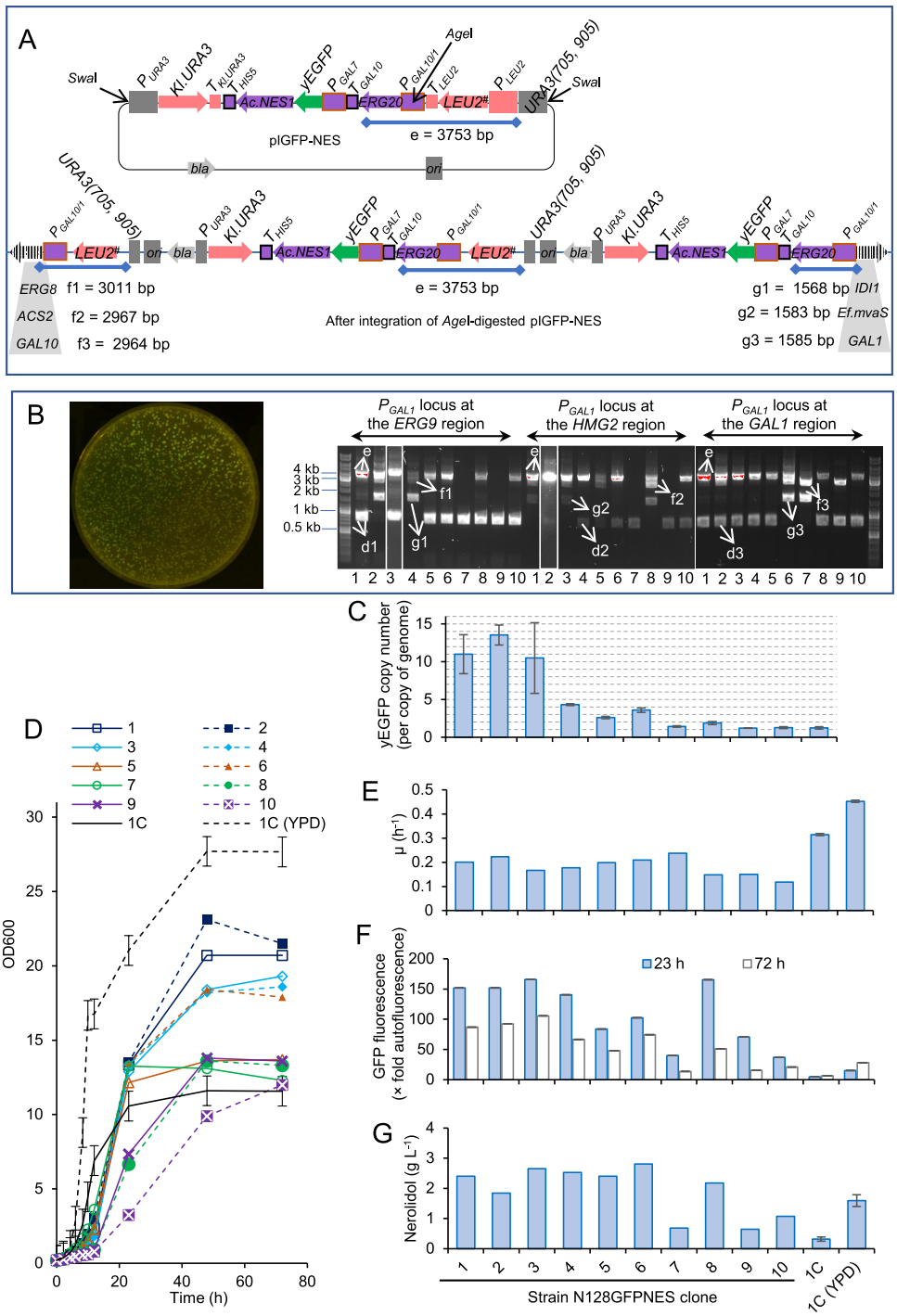


Figure 6. Genotypic and phenotypic variations in nerolidol-producing strains with the use of a yeast integrative plasmid. (A) Genetic features in the yeast integrative plasmid pIGFP-NES, and the *GAL1/10* promoter loci after pIGFP-NES integration. Fragments e, f1/2/3, and g1/2/3 are the PCR products in PCR product length polymorphism analysis. *LEU2*$^{\#}$, *LEU2* with *AgeI* restriction site removed. (B) yEGFP in yeast NLD128GFPNES colonies transformed with *AgeI*-linearized pIGFP-NES and the agarose gel separation of yeast colony PCR products. Fragments d1/2/3 are the PCR products shown in Figure 1A. (C) yEGFP copy numbers in the NLD128GFPNES clones and strain NLD128GFPNES1C (1C; transformed with *SwaI*-digested pIGFP-NES). Mean values ± standard deviation are shown (n = 4 technical replicates; and 4 four biological replicates for 1C). (C–F) characterization of NLD128GFPNES clones and strain NLD128GFPNES1C (1C) in flask cultivation: growth curve, μ (the maximal specific exponential growth rate), yEGFP fluorescence at 23 and 72 h, and nerolidol titer at 72 h. For 1–10 and 1C, yeast cells were grown in synthetic minimal YNB media containing 20 g L^{-1} glucose and 100 mM MES/ammonia buffer. For 1C (YPD), yeast cells were grown in a rich yeast extract-peptone 20 g L^{-1} glucose media. For 1–10, n = 1 independent cultivation. For 1C, mean ± standard deviation is shown (n = 3 biological replicates). yEGFP fluorescence was measured using a BD FACSCanto II flow cytometer.

normal colony size (Figure 4G: clones 1 to 10), and none of them showed the integration. We then tested another ten colonies with a smaller colony size, and clone 20 showed the integration (Figure 4G: clones 11 to 20). Clone 20 was a tiny-sized colony. We then tested another ten colonies with a tiny size, and only clone 21 showed the integration. These results

show that the linearized yeast episomal plasmid can be integrated into yeast chromosome and may confirm that *HXK2* deficiency be not required for such integration.

In summary, the integration of the episomal plasmid does not depend on hexokinase 2 function, and *hxk2*Δ may be related to the growth advantage for the clones with the integrated episomal plasmid.

Production of Nerolidol and Valencene Using the *RPL25*-Mediated HapAmp (Haploinsufficiency-Driven Gene Amplification) Vector. In the above sesquiterpene (nerolidol and valencene)-producing strains, the integration of the episomal plasmids (pJT9RFRNdeI and pJTEgVS9RFNdeI) caused genotypic variations, including in the gene copy number. The variations in the copy number of sesquiterpene synthetic genes (*ERG20*, *Ac.NES1*, and/or *Eg.VS*) may be the main genetic factors influencing the productivities. We aimed to characterize the productivities in the strains harboring a single copy of *ERG20* and multiple copies of *Ac.NES1* or *Eg.VS**. Previously, we developed a HapAmp vector to introduce the multiple copies of nerolidol synthase (*Ac.NES1*).[19] This vector can be integrated into the upstream of the open reading frame of the haploinsufficient gene *RPL25*.[31] As a result, a weak *PDA1* promoter controls the expression of *RPL25*, decreasing its expression output per gene copy. This provides an evolutionary force for the duplication of the genes between the *RPL25* terminator in the HapAmp vector and the *RPL25* terminator on the genome (Figure 5A). Here, we constructed two vectors that carry the fusion gene *YFAST-2A-Ac.NES1* or *YFAST-2A-Eg.VS** under the control of the *S. kudriavzevii GAL2* promoter (Figure 5A).[32] The *ERG20* gene in these vectors is localized in the nonamplification region (Figure 5A). We transformed these vectors into the background o128RB to generate nerolidol-producing strain NLD128RBM and valencene-producing strain VL128RBM. In these strains, the copy number of the terpene synthase module was around 11 and 15 (shown by YFAST copy number, Figure 5C), respectively. Both strains were characterized in flask cultivation under the conditions with the depletion of hexokinase 2.

The YFAST fluorescence and the YFAST copy numbers were consistently higher in valencene-producing strain VL128RBM than that in nerolidol-producing strain NLD128RBM. The YFAST fluorescence in both strains are in the range shown in the clones having ≥10 copies of episomal plasmid integration (Figures 5D, 2G, and 3E). This indicates the similar-level expression of sesquiterpene synthase in these clones. The maximum specific growth rates (μ_{max}) were higher in HapAmp valencene strains and lower in HapAmp nerolidol strains than those in plasmid integration clones, but the final biomass accumulation was similar. Sesquiterpene titers, ∼2.5 g L^{-1} nerolidol and ∼0.76 g L^{-1} valencene, were similar in these comparisons.

Multiple-Copy Integration of the Plasmid without 2μ Replicative Sequence. In *S. cerevisiae*, plasmids lacking an origin of replication can be integrated into yeast genome, which happens typically through homologous recombination.[33] This kind of plasmid is known as a yeast integrative plasmid (YIp) and is known as "being of single-copy" or 'the copy number depending on the numbers of homologous recombinant sequences on genome.' Because the integration of the episomal plasmids led to genotypic and phenotypic variations, we further raised the question as to whether the variations depended on the 2μ sequence and whether a YIp-type plasmid would cause similar variations.

We generated a new plasmid pIGFP-NES (Figure 6A), which contains the fusion of yeast-enhanced green fluorescent protein (yEGFP) gene and *Ac.NES1* under the control of the *GAL7* promoter, *ERG20* under the control of the *GAL10* promoter, selectable markers *LEU2* and *Kluyveromyces lactis URA3* (*Kl.URA3*), and recombination arms targeting to *URA3* locus.[34] The plasmid pIGFP-NES may be transformed after the linearization through *Age*I digestion in the middle of the *GAL1/10* promoter or *Swa*I digestion beside the *URA3* recombinant arms. *Age*I digestion may facilitate integration through homologous recombination at the *GAL1/10* promoter loci, and *Swa*I digestion may facilitate integration into the *URA3* locus. Noted is that the *GAL7* and *GAL10* promoters are used to control *yEGFP-Ac.NES1* and *ERG20* transcription. The *GAL7* promoter is ∼2-fold weaker than the *GAL2* promoter used in plasmid pJT9RFRNdeI.[24] The *GAL10* promoter is ∼3-fold weaker than the *GAL1* promoter.[24] This difference may allow evaluation of the effects of copy number variations when the heterologous pathway genes were controlled by weaker promoters.

We transformed *Age*I-digested pIGFP-NES into the background strain o128RB to generate strain NLD128GFPNES. The transformants showed different levels of fluorescence under blue-light illumination (Figure 6B). Colonies with bright fluorescence and colonies with dim fluorescence were analyzed through the yeast colony PCR and PCR product length polymorphism (Figure 6B). In clones 1−8, the plasmid pIGFP-NES was integrated into one of the *GAL1/10* promoter loci in *ERG9/HMG2/GAL1* regions (Figure 6B). In clone 9, plasmid pIGFP-NES was possibly integrated through the recombination with the *GAL1/10* promoter used to control *ERG8* and *IDI1* in the *ERG9* region (shown by the f3 band), but unknown genome rearrangement happened (shown by the lack of the g3 band). In clone 10, the plasmid pIGFP-NES was not integrated into any of the three *GAL1/10* loci (shown by d1, d2, and d3 bands), and the *AgeI* digestion was likely to be repaired (shown by the e band). In clones 7 and 9, the tandemly repeated integration did not happen (shown by the lack of e band). This is consistent with the quantitative PCR results showing the single copy of yEGFP per copy of genome (Figure 6C). In clones 1−6 and 8, yEGFP copy numbers ranged from 2 to 13 per copy of genome, and there was a single-copy yEGFP in clone 10. YIp and the integrated 2μ plasmids showed similar mitotic stabilities (Supplementary Table 1). These show that the genotypic variations do not depend on the 2μ sequence and that YIp can generate copy number variations at a single integration locus.

We transformed *Swa*I-digested pIGFP-NES into the background strain o128RB to generate the clones integrated with a single copy of nerolidol synthesis module at the *URA3* locus. One copy of yEGFP per copy of genome indicates single-copy integration (Figure 6C; strain 1C).

Ten NLD128GFPNES clones and the reference 1C were characterized in MES-buffered YNB-glucose media with auxin-inducible depletion of Hxk2p. The reference 1C showed an exponential growth rate of ∼0.3 h^{-1}, faster than the NLD128GFPNES clones. The NLD128GFPNES clones with the single copy or two copy of *yEGFP-Ac.NES1* module as well as the reference 1C showed slow growth or deficient growth after 24 h (Figure 6D: clones 7−10 and 1C). Other NLD128GFPNES clones, having increased copy numbers,

showed improved growth after 24 h. The yEGFP fluorescence was quite low in reference 1C in comparison with the NLD128GFPNES single-copy clones 7, 9, and 10 (Figure 6F). In NLD128GFPNES clones, clones with the single-copy integration showed the lower yEGFP fluorescence than others. The yEGFP copy number did not correlate well with the yEGFP fluorescence at 23 h, which might be reflective of different growth phases in these clones. The yEGFP copy number was positively correlated with yEGFP fluorescence at 72 h. yEGFP fluorescence decreased at 72 h in comparison with that at 23 h. This is different from the slight increase in YFAST fluorescence in NLD128RB clones (Figure 2G). The activities of the *GAL7* promoter and the half-life of the *yEGFP-Ac.NES1* fusion protein may influence the yEGFP protein content.

For nerolidol production, clone 6 showed the highest titer, ~2.8 g L^{-1}, and clones 1 and 3−5 showed the titers around ~2.5 g L^{-1}. This titer below the levels produced is in the single-copy integration clones D2 and D3 in strain NLD128B, in which a strong *GAL2* promoter is used to control nerolidol synthase expression. The single-copy clones produced ≤1 g L^{-1}. The reference 1C produced the least titer, ~0.3 g L^{-1}. We then characterized the reference 1C strain in YPD medium, which resulted in improved growth, yEGFP fluorescence, and nerolidol production (~1.6 g L^{-1}).

In summary, the *GAL1* promoter-mediated integration of a YIp plasmid may deliver genotypic and phenotypic variations. The reference 1C strain shows that the integration location has a strong effect on multiple phenotypes. Without considering clones 1−3 with ≥9 copy integration, the yEGFP fluorescence at 72 h and nerolidol titers show a strong positive correlation (Supplementary Figure 11; R^2 = 0.95). This indicates that the nerolidol synthase expression levels have a determinative effect on the nerolidol titer in a specific range.

DISCUSSION

Rational engineering and targeted selection of desired phenotypes represent two key methods to develop and improve microbial cell factories.[35] Rational engineering relies on predictable designs and precise tools to build a predefined function, whereas evolution-selection plays a random statistics "dice game" to find "odds and ends."[36] These strategies seem contradictory to each other but are often combinatorially used in metabolic engineering and synthetic biology. Our previous engineering for improved terpene production has been primarily dependent on the rational engineering strategies.[5,8,24,25,37] This study shows that the improvement that we achieved might not be aligned with the design strategy parameters and raises doubts about how precisely we understand these commonly used engineering tools.

The mechanism for improved nerolidol production in the strain with auxin-inducible protein depletion of hexokinase 2 needs to be reviewed. We previously hypothesized that depletion of hexokinase 2 (Hxk2p) may relieve glucose repression on *GAL* promoters for upregulated expression of heterologous nerolidol synthetic genes and increase respiration and biomass accumulation on glucose.[8] Proteomics analysis showed that *GAL* promoter-controlled genes were expressed at higher levels in Hxk2p-depleted cells than in *HXK2*-wildtype cells.[38] Increased biomass accumulation has been shown previously.[8] Here, we show that the production of the sesquiterpene products nerolidol and valencene is influenced by multiple factors, including (1) the chromosomal integration status vs episomal status, (2) integration location, (3) copy number, (4) the expression level of sesquiterpene synthase, and (5) other genome rearrangements.

Chromosome integration of episomal plasmids contributed to improved growth and improved nerolidol/valencene production. In Hxk2p-depleted/*hxk2*Δ background (Figures 2, 3, and 4D,F), proteomic data showed that Rep2p, a key factor for 2μ plasmid inherence,[39] was not present in Hxk2p-depleted cells. However, the activities of the *REP2* promoter (and the *TEF1* promoter as a reference) and the levels of the Rep2-yEGFP fusion under the control of the *TEF1* promoter were not dramatically impacted by *HXK2* disruption (Supplementary Figure 12). *GAL* promoters are transcription hotspots in *gal80*Δ and Hxk2p-depleted background, which may delay DNA replication at transcription initiation sites in the S phase and require DNA synthesis in the G2/M phase.[40] The replication of each copy of 2μ plasmids occurs during the S phase.[41] Uncoupled replication timing might contribute to plasmid instability and then growth deficiency. However, this is not sufficient for understanding the causes of the episomal plasmid integration and its positive effects on growth and production and requires further investigation.

The genomic location of heterologous genes has a significant effect on phenotypes. In two NLD128H clones that produced 3.5 g L^{-1} nerolidol, the episomal plasmid pJT9RFR was integrated through the *GAL1* promoter. In two NLD128L clones producing 2.5 g L^{-1} nerolidol, the integration was through the *GAL2* promoter. It seems that the integration locus results in a series of outcomes in the genotype and phenotype. Here, we only revalidated the genotypes caused by the integration through the *GAL1* promoter and left this phenomenon largely unexplained. Another genome-location effect is shown by the difference between the NLD128GFPNES clones 7 and 9 that contained a single-copy integration of pIGFP-NES at the *GAL1* locus and in strain NLD128GFPNES1C (1C) that contained the pIGFP-NES integrated at *ura3* locus (Figure 6). In NLD128GFPNES1C, yEGFP levels were much lower than those in NLD128GFPNES clones 7 and 9, suggesting the *GAL7* promoter was not fully induced at *ura3* locus. Similar location effects on heterologous gene expression have been previously observed in *S. cerevisiae* and *Komagataella phaffii*.[42,43] The epigenetic effect and enhancer-like effect from surrounding sequences may influence transgene expression.[42] The pIGFP-NES integration through the *GAL1* promoter resulted in the P_{GAL7}-*yEGFP-NES* cassette being surrounded by *GAL* promoter-controlled expression cassettes (Figure 6A). This may form a micro three-dimensional environment favoring the transcription from the *GAL7* promoter.

Integration copy numbers were positively correlated to YFAST/yEGFP fluorescence levels (Figures 2, 3, and 5). YFAST/yEGFP fluorescence levels (proxy to nerolidol/valencene synthase levels) and nerolidol/valencene production were positively correlated below a threshold (Supplementary Figures 7 and 11). Although we only characterized them once for some integration clones (n = 1; Figures 3 and 5), the coefficients of determination for the correlations were over 85%, indicating that there was a strong dependency between terpene synthase expression and sesquiterpene production. The clones with lower YFAST fluorescence showed slowed growth (Figures 3 and 6). This may indicate the metabolic imbalance between the synthesis and consumption of

sesquiterpene precursor farnesyl pyrophosphate, accumulation of which may cause cellular toxicity.[44] For nerolidol production, three-copy integration is optimal, and increasing expression dosage per gene copy is superior to increasing the copy number of the weakly expressed gene (Figures 2 and6). Previously, three-copy integration of a terpene synthase expression cassette was also reported for optimal production of sesquiterpenes, including valencene.[30,45] However, this observation is not consistent with the valencene case in this study. The best valencene-producing clones (VL128RB clones 6 and 7) have a single-copy or two-copy integration, varied in YFAST fluorescence, but produced the same amount of valencene (~0.96 g L^{-1}). Both clones share the same genome rearrangement, which resulted in the loss of the *ACS2* expression cassette. This improvement is not what we expected. In these two clones, improved expression from the valencene synthase module did not show further improvement, whereas in other clones, such expression was well correlated with production (Supplementary Figure 7). Further investigation will be necessary to understand how metabolic flux is controlled in this case.

In summary, the integration of an episomal/integrative plasmid may cause genotypic and phenotypic diversity in *S. cerevisiae*. Genotype may vary in copy number, integration location depending on the availability of homologous sequences on genome, and other genome rearrangements. A fluorescent reporter can be used to assist in the selection of copy number variants. An association between genotypes (like copy number) and phenotypes (like productivities) can be observed conditionally. Screening multiple clones may be used to select the clones with improved production of the target sesquiterpene. It suggests that plasmid-based expression should be carefully considered in microbial cell factory development, especially when heterogeneity can be a problem. Nevertheless, this work indicates that screening biological replicates, being one of the evolution-selection strategies, may further unleash the capacity of microbial cell factories that are rationally designed and engineered. Moreover, this work envisions the challenges for precise designs in engineering microbial cell factories due to the knowledge gap between genes/genomes and phenotypes/traits. To fill the gap, it will be necessary to expand the characterization at multiomics levels and comprehensively model the biological networks.

MATERIALS AND METHODS

Plasmid and Strain Construction. Primers, plasmids, plasmid construction processes, *S. cerevisiae* strains, and strain construction processes are summarized in Supplementary Data 2.

Yeast Colony PCR. Zymolyase-20T (nacalai tesque, Japan) was used to digest yeast cell wall, and PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Japan) was used in PCR amplification. Zymolase working solution was prepared by mixing phosphate-buffered saline (PBS) buffer and zymolyase stock solution (1U per μL in 1*PBS buffer containing 100 mM DTT and 50% v/v glycerol) at the ratio of 10:1. Yeast cells were resuspended in 5 μL of Zymolase working solution, digested at 37 °C for 30 min, and denatured at 95 °C for 5 min. Yeast cell digestion was mixed with 100 μL of water, and 1 μL of mixture was used as the template for PCR amplification. Each PCR reaction was prepared in 10 μL final volume and contained four primers (each 0.2 μM final concentration; Supplementary Data 2). PCR thermal-cycle conditions are (1) 98 °C for 30 s, (98 °C for 15 s, 50 °C for 10 s, 68 °C for 5 min)*3, (98 °C for 15 s, 55 °C for 10 s, 68 °C for 5 min)*27, 68 °C for 5 min; or (2) 98 °C for 30 s, (98 °C for 15 s, 55 °C for 10 s, 68 °C for 5 min)*30, 68 °C for 5 min. Former conditions resulted in a better amplification for d2 band (Figures 2 and 6).

Flask Cultivation. Dodecane-overlaid two-phase flask cultivation was performed as previously described.[8] 2-(*N*-morpholino)ethanesulfonic acid (MES)-buffered synthetic minimal medium contained 6.9 g L^{-1} YNB (FORMEDIUM#-CYN0402), 20 g L^{-1} glucose as the carbon source, 100 mM MES with the pH adjusted to 6.0 with ammonium hydroxide. YPD medium contained 20 g L^{-1} of peptone, 20 g L^{-1} of yeast extract, and 20 g L^{-1} glucose. Yeast cells from the glycerol stock were streaked on YNB agar containing 20 g L^{-1} of glucose and 125 mM tetracycline. Cells were precultured in 5 mL MES-buffered synthetic minimal medium and grown for ~ 8 h, and then grown in 15 mL MES-buffered synthetic minimal medium with addition of 1 mM NAA (500 mM stock; prepared in ethanol) overnight to exponential phase (OD_{600} between 1 and 4). Precultured cells were collected by centrifugation and resuspended in a fresh medium. Appropriate volumes of precultured cells were transferred to a fresh fermentation medium (MES-buffered synthetic medium or YPD medium) to an initial OD_{600} of 0.2 in a total volume of 23 mL medium in a 250 mL flask. Sterile dodecane (2 mL) was added after inoculation for the *in situ* extraction of the terpene products. NAA (500 mM, dissolved in ethanol) was added to a concentration of 1 mM to trigger auxin-inducible protein degradation. Cells were cultivated at 200 rpm and 30 °C. In the first 12 h of cultivation, 3 mL culture was sampled. Dodecane was sampled at 72 h and stored at −80 °C for extracellular metabolite analysis. Yeast cells were sampled and immediately used for cell density analysis and/or fluorescence-based analysis.

Flow Cytometry. YFAST fluorescence in single cells was analyzed using a BD Accuri C6 flow cytometer (BD Biosciences, USA). For analysis of YFAST fluorescence, 2 mM HMBR (100-time concentrated; dissolved in dimethyl sulfoxide) was added to the samples to a 20 μM final concentration, and the sample was mixed before analysis. FSC.H threshold was set at the value of 250,000 for exclusion of debris particles. YFAST fluorescence was excited by a 488 nm laser and monitored through a 530/20 nm bandpass filter (FL1.A), with 10,000 events recorded per sample. Mean values of FSC.A, SSC.A, and FL1.A for all detected events were extracted using BD Csampler software (BD Accuri C6 software, version 1.0.264.21). YFAST fluorescence level was expressed as the percentage of the average background autofluorescence from the exponential-phase cells of GFP-negative reference strain GH4.

yEGFP fluorescence in single cells was analyzed using a BD FACSCanto II flow cytometry (Firmware version 1.49). Yeast cells diluted with water were used directly for yEGFP analysis. yEGFP fluorescence was excited by a 488 nm laser and monitored at the FITC-A channel through a 556 long-pass dichroic mirror, a 502 long-pass dichroic mirror, and a 530/30 nm bandpass filter. Mean values of FSC-A (PMT voltage = 200; gating threshold = 3000), SSC-A (PMT voltage = 200), and FITC-A (PMT voltage = 450) for all detected events were extracted using BD FACSDiva Software (version 8.0.1; CST Version 3.0.1; PLA version 2.0). YFAST fluorescence level was expressed as the percentage of the average background

autofluorescence from the exponential-phase cells of GFP-negative reference strain CEN.PK113-7D.[23]

Metabolite Analysis. Nerolidol and valencene were analyzed using an Agilent 1200 HPLC system at the Metabolomics Australia Queensland Node. Dodecane samples (diluted with dodecane) were diluted in a 40-fold volume of ethanol. The ethanol-diluted samples (20 μL) were injected and separated using a Zorbax Extend C18 column (4.6 × 150 mm, 3.5 μm, Agilent PN: 763953-902) equipped with a guard column (SecurityGuard Gemini C18, Phenomenex PN: AJO-7597). Analytes were eluted at 35 °C and 0.9 mL/min using the mixture of solvent A (water) and solvent B (45% acetonitrile, 45% methanol, and 10% water). For nerolidol analysis, solvents A and B were mixed with a linear gradient of 5−100% solvent B from 0−24 min, then 100% from 24 to 30 min, and finally 5% from 30.1 to 35 min. For valencene analysis (Supplementary Figure 13), solvents A and B were mixed with a linear gradient of 5−100% solvent B from 0 to 24 min, then 100% from 24 to 40 min, and finally 5% from 40.1 to 45 min. *Trans*-nerolidol primary reference standard (93.7% purity; Sigma-Aldrich #04610590 purity was not used in correcting quantification) was used to prepare the standard curve. (+)-Valencene (≥80.0% purity; Sigma-Aldrich #06808-50MG; purity 80% was used in correcting quantification). A diode array detector (Agilent DAD SL, G1315C) at a 202 nm wavelength was used to monitor analytes of interest. The specific production rate for nerolidol and valencene was calculated as published previously.[24]

Quantitative Real-Time PCR. To extract genomic DNA, yeast cells were homogenized using a glass bead-phenol-chloroform homogenization method, and nucleic acids in the aqueous phase were precipitated using ethanol precipitation. Rnase I was used to remove RNA and genomic DNA was further purified using Monarch DNA Cleanup Columns (NEB). Luna Universal qPCR Master Mix (NEB) was used for quantitative real-time PCR amplification. Genomic DNA (0.25 ng) was used as the template in a 10 μL reaction. Genomic DNA from a strain (ILHA Y2A)[8] with a single-copy integration of YFAST-GFP was used as the reference genome, and *ACT1* gene was used as the reference gene. Primers are listed in Supplementary Data 2. YFAST and yEGFP copy numbers were calculated using $2^{-\Delta\Delta CT}$ method.[46]

Genome Sequencing. Genomic DNA extraction and genome sequencing were performed as described previously.[19] The procedures are summarized as below. MagAttract HMW DNA Kit (Qiangen) was used with a modified protocol. Zymolase (30 U) was used to digest yeast cells in 2 mL 1 M sorbitol solution. Yeast protoplast cells were then resuspended in 300 μL Buffer AL (MagAttract HMW DNA Kit) and then 360 buffer ATL (MagAttract HMW DNA Kit) was added. Digestion by Proteinase K and Rnase A and purification using magnetic beads were then performed. Genomic DNA was further purified with phenol:chloroform treatment and magnetic bead purification. Rapid Barcoding Kit (SQK-RBK004, Oxford Nanopore) and MinION Mk1C (Oxford Nanopore) with R9 flow cell MIN106D were used nanopore DNA sequencing. Ont-guppy-for-mk1c (version 4.2.3) installed MinION Mk1C (MinKNOW version 20.10.6) was used for base-calling. The fastQ read files have been deposited in NCBI Sequence Read Archive (SRA; Submission: SUB11091178 and SUB13887293). A Prymetime pipeline (https://github.com/sjtrauber/prymetime) was used to assemble the genome and map the engineering features in the context of genome assembly.[11] In the Prymetime pipeline, a Flye assembler was used for genome assembly.[11,47] For genome assembly through a Canu assembler (Galaxy Version 2.2+galaxy0),[26] a Galaxy server (Galaxy Australia) was used.[48] A Maker genome annotation pipeline (Galaxy Version 2.31.11+galaxy2) was used to annotate genetic features. MiniMap2 (Galaxy Version 2.24+galaxy0) was used to align the reads trimmed by Canu assembler and the contigs outputted by Canu assembler.[49] JBrowse2 (Desktop version v2.4.2) was used to display genetic features in genome assembly.[50]

■ ASSOCIATED CONTENT

Supporting Information

Supplementary Table 1: testing the stability of the strains containing the integrated/episomal plasmid; Supplementary Figures 1−6 and 9−10: genetic features shown by the whole genome sequencing; Supplementary Figures 7 and 11: the correlation between sesquiterpene production and YFAST fluorescence; Supplementary Figure 8: analysis of genome rearrangement in VL128B clone 6 and clone 7 and yeast colony PCR confirmation; Supplementary Figure 12: characterization of promoter activities in the *HXK2* or *hxk2*Δ background strains; Supplementary Figure 13: HPLC separation profiles for the valencene standard and a valencene sample from yeast fermentation (PDF)

Supplementary Data 1: localization of selected engineering features in genome assemblies in NLD128H/NLD128L clones (XLSX)

Supplementary Data 2: genetic and microbial materials (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Bingyin Peng − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia;* orcid.org/0000-0001-6918-6016; Email: bingyin.peng@qut.edu.au

Authors

Sarah J. Weintraub − *Bioinformatics and Computational Biology, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, United States of America*

Zeyu Lu − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia*

Samuel Evans − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental*

Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia

Qianyi Shen − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia; School of Chemistry and Molecular Biosciences (SCMB), The University of Queensland, Brisbane, QLD 4072, Australia*

Liam McDonnell − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia*

Manuel Plan − *Australian Institute for Bioengineering and Nanotechnology (AIBN) and Metabolomics Australia (Queensland Node), Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia*

Thomas Collier − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; School of Natural Sciences, Macquarie University, Sydney, NSW 2109, Australia;* orcid.org/0009-0006-0537-445X

Li Chen Cheah − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia*

Lei Ji − *Shandong Provincial Key Laboratory of Applied Microbiology, Ecology Institute, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250103, PR China*

Christopher B. Howard − *Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia;* orcid.org/0000-0001-9797-8686

Will Anderson − *Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia*

Matt Trau − *Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia; School of Chemistry and Molecular Biosciences (SCMB), The University of Queensland, Brisbane, QLD 4072, Australia;* orcid.org/0000-0001-5516-1280

Geoff Dumsday − *CSIRO Manufacturing, Clayton, VIC 3169, Australia*

Erin L. Bredeweg − *Functional and Systems Biology Group, Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States*

Eric M. Young − *Department of Chemical Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, United States;* orcid.org/0000-0001-5276-2873

Robert Speight − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; Advanced Engineering Biology Future Science Platform, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Black Mountain, ACT 2601, Australia;* orcid.org/0000-0003-4161-8272

Claudia E. Vickers − *ARC Centre of Excellence in Synthetic Biology, Sydney, NSW 2109, Australia; Centre of Agriculture and the Bioeconomy, School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia*

Author Contributions

B.P.: Investigation, Methodology, Conceptualization, Validation, Formal analysis, Funding acquisition, Supervision, Writing−original draft. S.J.W.: Investigation, Methodology, Formal analysis, Writing - Review & Editing. Z.L.: Investigation, Methodology, Validation, Formal analysis. S.E.: Investigation, Methodology, Validation, Formal analysis. Q.S.: Investigation, Methodology, Validation, Formal analysis. L.M.: Investigation, Methodology, Validation, Formal analysis. M.P.: Investigation, Methodology, Validation, Formal analysis. T.C.: Investigation, Writing−original draft, Writing - Review & Editing. L.C.C.: Investigation, Conceptualization, Resources. L.J.: Investigation, Conceptualization, Writing−original draft. C.B.H.: Investigation, Supervision, Resources. W.A.: Investigation, Resources. M.T.: Investigation, Supervision, Resources, Project administration. G.D.: Investigation, Supervision, Resources, Funding acquisition. E.L.B.: Investigation, Conceptualization, Resources. E.M.Y.: Investigation, Conceptualization, Methodology, Supervision, Formal analysis, Software. R.S.: Investigation, Supervision, Resources, Project administration, Funding acquisition. C.E.V.: Investigation, Conceptualization, Resources, Funding acquisition, Writing - Review & Editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported partially by the Australian Government through the Australian Research Council Centres of Excellence funding scheme (project CE200100029). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or Australian Research Council. This research and B.P. were support partially from CSIRO and the University of Queensland in the form of a Synthetic Biology Future Science Platform Fellowship. Metabolomics Australia is supported by BioPlatforms Australia through the Commonwealth Government's National Collaborative Research Infrastructure Strategy (NCRIS). This work is supported by Galaxy Australia, a service provided by the Australian Biocommons and its partners.

■ REFERENCES

(1) Malcı, K.; Watts, E.; Roberts, T. M.; Auxillos, J. Y.; Nowrouzi, B.; Boll, H. O.; Nascimento, C. Z. S. d.; Andreou, A.; Vegh, P.; Donovan, S.; et al. Standardization of Synthetic Biology Tools and Assembly Methods for Saccharomyces cerevisiae and Emerging Yeast Species. *ACS synthetic biology* **2022**, *11* (8), 2527−2547.

(2) Mante, J.; Hao, Y.; Jett, J.; Joshi, U.; Keating, K.; Lu, X.; Nakum, G.; Rodriguez, N. E.; Tang, J.; Terry, L.; et al. Synthetic Biology Knowledge System. *ACS synthetic biology* **2021**, *10* (9), 2276−2285. Liu, Z.; Wang, J.; Nielsen, J. Yeast synthetic biology advances biofuel production. *Curr. Opin. Microbiol.* **2022**, *65*, 33−39.

(3) Cardoso, J. G.; Andersen, M. R.; Herrgård, M. J.; Sonnenschein, N. Analysis of genetic variation and potential applications in genome-

scale metabolic modeling. *Frontiers in bioengineering and biotechnology* **2015**, *3*, 13.
(4) Montaño López, J.; Duran, L.; Avalos, J. L. Physiological limitations and opportunities in microbial metabolic engineering. *Nature Reviews Microbiology* **2022**, *20* (1), 35−48.
(5) Peng, B.; Bandari, N. C.; Lu, Z.; Howard, C. B.; Scott, C.; Trau, M.; Dumsday, G.; Vickers, C. E. Engineering eukaryote-like regulatory circuits to expand artificial control mechanisms for metabolic engineering in Saccharomyces cerevisiae. *Commun. Biol.* **2022**, *5* (1), 135.
(6) Wang, T.; Dunlop, M. J. Controlling and exploiting cell-to-cell variation in metabolic engineering. *Curr. Opin. Biotechnol.* **2019**, *57*, 10−16.
(7) Naseri, G.; Koffas, M. A. G. Application of combinatorial optimization strategies in synthetic biology. *Nat. Commun.* **2020**, *11* (1), 2446.
(8) Lu, Z.; Peng, B.; Ebert, B. E.; Dumsday, G.; Vickers, C. E. Auxin-mediated protein depletion for metabolic engineering in terpene-producing yeast. *Nat. Commun.* **2021**, *12* (1), 1051.
(9) Meadows, A. L.; Hawkins, K. M.; Tsegaye, Y.; Antipov, E.; Kim, Y.; Raetz, L.; Dahl, R. H.; Tai, A.; Mahatdejkul-Meadows, T.; Xu, L.; et al. Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* **2016**, *537* (7622), 694−697.
(10) Jiang, Y.; Xia, L.; Gao, S.; Li, N.; Yu, S.; Zhou, J. Engineering Saccharomyces cerevisiae for enhanced (−)-α-bisabolol production. *Synthetic and Systems Biotechnology* **2023**, *8* (2), 187−195.
(11) Collins, J. H.; Keating, K. W.; Jones, T. R.; Balaji, S.; Marsan, C. B.; Çomo, M.; Newlon, Z. J.; Mitchell, T.; Bartley, B.; Adler, A.; et al. Engineered yeast genomes accurately assembled from pure and mixed samples. *Nat. Commun.* **2021**, *12* (1), 1485.
(12) Yi, X.; Alper, H. S. Considering Strain Variation and Non-Type Strains for Yeast Metabolic Engineering Applications. *Life* **2022**, *12* (4), 510.
(13) Xu, X.; Meier, F.; Blount, B. A.; Pretorius, I. S.; Ellis, T.; Paulsen, I. T.; Williams, T. C. Trimming the genomic fat: minimising and re-functionalising genomes using synthetic biology. *Nat. Commun.* **2023**, *14* (1), 1984. Cao, X.; Yu, W.; Chen, Y.; Yang, S.; Zhao, Z. K.; Nielsen, J.; Luan, H.; Zhou, Y. J. Engineering yeast for high-level production of diterpenoid sclareol. *Metabolic engineering* **2023**, *75*, 19−28. Kwan, B. D.; Seligmann, B.; Nguyen, T.-D.; Franke, J.; Dang, T.-T. T. Leveraging synthetic biology and metabolic engineering to overcome obstacles in plant pathway elucidation. *Current opinion in plant biology* **2023**, *71*, No. 102330.
(14) He, S.; Zhang, Z.; Lu, W. Natural promoters and promoter engineering strategies for metabolic regulation in Saccharomyces cerevisiae. *Journal of Industrial Microbiology and Biotechnology* **2023**, *50*, 1. Park, J. H.; Bassalo, M. C.; Lin, G.-M.; Chen, Y.; Doosthosseini, H.; Schmitz, J.; Roubos, J. A.; Voigt, C. A. Design of Four Small-Molecule-Inducible Systems in the Yeast Chromosome, Applied to Optimize Terpene Biosynthesis. *ACS synthetic biology* **2023**, *12* (4), 1119−1132.
(15) Futcher, A.; Cox, B. Copy number and the stability of 2-micron circle-based artificial plasmids of Saccharomyces cerevisiae. *Journal of bacteriology* **1984**, *157* (1), 283−290. Guerrini, A. M.; Ascenzioni, F.; Tribioli, C.; Donini, P. Transformation of Saccharomyces cerevisiae and Schizosaccharomyces pombe with linear plasmids containing 2 micron sequences. *EMBO journal* **1985**, *4* (6), 1569−1573.
(16) Kingsman, A. J.; Clarke, L.; Mortimer, R. K.; Carbon, J. Replication in Saccharomyces cerevisiae of plasmid pBR313 carrying DNA from the yeast trpl region. *Gene* **1979**, *7* (2), 141−152.
(17) Gunge, N. Yeast DNA plasmids. *Annual review of microbiology* **1983**, *37* (1), 253−276.
(18) Lee, F. W. F.; Silva, N. A. D. Ty1-Mediated Integration of Expression Cassettes: Host Strain Effects, Stability, and Product Synthesis. *Biotechnology progress* **1996**, *12* (4), 548−554. Jacobs, E.; Dewerchin, M.; Bocke, J. D. Retrovirus-like vectors for Saccharomyces cerevisiae: integration of foreign genes controlled by efficient promoters into yeast chromosomal DNA. *Gene* **1988**, *67* (2), 259−269. Lopes, T. S.; Klootwijk, J.; Veenstra, A. E.; van der Aar, P. C.; van Heerikhuizen, H.; Raué, H. A.; Planta, R. J. High-copy-number integration into the ribosomal DNA of Saccharomyces cerevisiae: a new vector for high-level expression. *Gene* **1989**, *79* (2), 199−206.
(19) Peng, B.; Esquirol, L.; Lu, Z.; Shen, Q.; Cheah, L. C.; Howard, C. B.; Scott, C.; Trau, M.; Dumsday, G.; Vickers, C. E. An in vivo gene amplification system for high level expression in Saccharomyces cerevisiae. *Nat. Commun.* **2022**, *13* (1), 2895.
(20) Kim, M.-D.; Rhee, S.-K.; Seo, J.-H. Enhanced production of anticoagulant hirudin in recombinant Saccharomyces cerevisiae by chromosomal δ-integration. *Journal of biotechnology* **2001**, *85* (1), 41−48.
(21) Sereika, M.; Kirkegaard, R. H.; Karst, S. M.; Michaelsen, T. Y.; Sørensen, E. A.; Wollenberg, R. D.; Albertsen, M. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nat. Methods* **2022**, *19* (7), 823−826.
(22) Lv, X.; Hueso-Gil, A.; Bi, X.; Wu, Y.; Liu, Y.; Liu, L.; Ledesma-Amaro, R. New synthetic biology tools for metabolic control. *Curr. Opin. Biotechnol.* **2022**, *76*, No. 102724. Bartley, B. A.; Beal, J.; Karr, J. R.; Strychalski, E. A. Organizing genome engineering for the gigabase scale. *Nat. Commun.* **2020**, *11* (1), 689. Opgenorth, P.; Costello, Z.; Okada, T.; Goyal, G.; Chen, Y.; Gin, J.; Benites, V.; de Raad, M.; Northen, T. R.; Deng, K.; et al. Lessons from Two Design−Build−Test−Learn Cycles of Dodecanol Production in Escherichia coli Aided by Machine Learning. *ACS synthetic biology* **2019**, *8* (6), 1337−1351.
(23) Entian, K.-D.; Kötter, P. 25 Yeast Genetic Strain and Plasmid Collections. In *Methods in Microbiology*; Stansfield, I., Stark, M. J. R., Eds.; vol 36; Academic Press, 2007; pp 629−666.
(24) Peng, B.; Plan, M. R.; Carpenter, A.; Nielsen, L. K.; Vickers, C. E. Coupling gene regulatory patterns to bioprocess conditions to optimize synthetic metabolic modules for improved sesquiterpene production in yeast. *Biotechnology for biofuels* **2017**, *10*, 43.
(25) Peng, B.; Nielsen, L. K.; Kampranis, S. C.; Vickers, C. E. Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in Saccharomyces cerevisiae. *Metabolic engineering* **2018**, *47*, 83.
(26) Koren, S.; Walenz, B. P.; Berlin, K.; Miller, J. R.; Bergman, N. H.; Phillippy, A. M. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **2017**, *27* (5), 722−736.
(27) Campbell, M. S.; Holt, C.; Moore, B.; Yandell, M. Genome Annotation and Curation Using MAKER and MAKER-P. *Curr. Protoc Bioinformatics* **2014**, *48*, 48. 4 11 11−14 11 39
(28) Futcher, A. B.; Cox, B. S. Copy number and the stability of 2-micron circle-based artificial plasmids of Saccharomyces cerevisiae. *Journal of bacteriology* **1984**, *157* (1), 283−290.
(29) Plamont, M. A.; Billon-Denis, E.; Maurin, S.; Gauron, C.; Pimenta, F. M.; Specht, C. G.; Shi, J.; Querard, J.; Pan, B. Y.; Rossignol, J.; et al. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113* (3), 497−502.
(30) Ye, Z.; Huang, Y.; Shi, B.; Xiang, Z.; Tian, Z.; Huang, M.; Wu, L.; Deng, Z.; Shen, K.; Liu, T. Coupling cell growth and biochemical pathway induction in Saccharomyces cerevisiae for production of (+)-valencene and its chemical conversion to (+)-nootkatone. *Metabolic engineering* **2022**, *72*, 107−115.
(31) Deutschbauer, A. M.; Jaramillo, D. F.; Proctor, M.; Kumm, J.; Hillenmeyer, M. E.; Davis, R. W.; Nislow, C.; Giaever, G. Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **2005**, *169* (4), 1915−1925.
(32) Peng, B.; Wood, R. J.; Nielsen, L. K.; Vickers, C. E. An Expanded Heterologous GAL Promoter Collection for Diauxie-Inducible Expression in Saccharomyces cerevisiae. *ACS synthetic biology* **2018**, *7*, 748.
(33) Orr-Weaver, T. L.; Szostak, J. W.; Rothstein, R. J. Yeast transformation: a model system for the study of recombination. *Proc.*

Natl. Acad. Sci. U. S. A. **1981**, *78* (10), 6354−6358. Myers, A. M.; Tzagoloff, A.; Kinney, D. M.; Lusty, C. J. Yeast shuttle and integrative vectors with multiple cloning sites suitable for construction of lacZ fusions. *Gene* **1986**, *45* (3), 299−310.

(34) Cheah, L. C.; Liu, L.; Stark, T.; Plan, M. R.; Peng, B.; Lu, Z.; Schenk, G.; Sainsbury, F.; Vickers, C. E. Metabolic flux enhancement from the translational fusion of terpene synthases is linked to terpene synthase accumulation. *Metabolic engineering* **2023**, *77*, 143−151. Peng, B.; Williams, T.; Henry, M.; Nielsen, L.; Vickers, C. Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microbial Cell Fact.* **2015**, *14* (1), 91.

(35) Cho, J. S.; Kim, G. B.; Eun, H.; Moon, C. W.; Lee, S. Y. Designing Microbial Cell Factories for the Production of Chemicals. *JACS Au* **2022**, *2* (8), 1781−1799. Ye, M.; Gao, J.; Zhou, Y. J. Global metabolic rewiring of the nonconventional yeast Ogataea polymorpha for biosynthesis of the sesquiterpenoid *β*-elemene. *Metabolic engineering* **2023**, *76*, 225−231.

(36) de Lorenzo, V. Evolutionary tinkering vs. rational engineering in the times of synthetic biology. *Life Sciences, Society and Policy* **2018**, *14* (1), 18.

(37) Hayat, I. F.; Plan, M.; Ebert, B. E.; Dumsday, G.; Vickers, C. E.; Peng, B. Auxin-mediated induction of GAL promoters by conditional degradation of Mig1p improves sesquiterpene production in Saccharomyces cerevisiae with engineered acetyl-CoA synthesis. *Microb. Biotechnol.* **2021**, *14* (6), 2627−2642. (acccessed 2022/11/12)

(38) Lu, Z.; Shen, Q.; Liu, L.; Talbo, G.; Speight, R.; Trau, M.; Dumsday, G.; Howard, C. B.; Vickers, C. E.; Peng, B. Profiling proteomic responses to hexokinase-II depletion in terpene-producing Saccharomyces cerevisiae. *Engineering Microbiology* **2023**, *3* (3), No. 100079.

(39) Mereshchuk, A.; Johnstone, P. S.; Chew, J. S. K.; Dobson, M. J. The yeast 2-micron plasmid Rep2 protein has Rep1-independent partitioning function. *Nucleic Acids Res.* **2022**, *50* (18), 10571−10585. (acccessed 4/9/2023)

(40) Wang, J.; Rojas, P.; Mao, J.; Mustè Sadurnì, M.; Garnier, O.; Xiao, S.; Higgs, M. R.; Garcia, P.; Saponaro, M. Persistence of RNA transcription during DNA replication delays duplication of transcription start sites until G2/M. *Cell Rep.* **2021**, *34* (7), 108759.

(41) Zakian, V. A.; Brewer, B. J.; Fangman, W. L. Replication of each copy of the yeast 2 micron DNA plasmid occurs during the S phase. *Cell* **1979**, *17* (4), 923−934.

(42) Brady, J. R.; Tan, M. C.; Whittaker, C. A.; Colant, N. A.; Dalvie, N. C.; Love, K. R.; Love, J. C. Identifying Improved Sites for Heterologous Gene Integration Using ATAC-seq. *ACS synthetic biology* **2020**, *9* (9), 2515−2524.

(43) Vogl, T.; Gebbie, L.; Palfreyman Robin, W.; Speight, R. Effect of Plasmid Design and Type of Integration Event on Recombinant Protein Expression in Pichia pastoris. *Appl. Environ. Microbiol.* **2018**, *84* (6), No. e02712-17. Wu, X.-L.; Li, B.-Z.; Zhang, W.-Z.; Song, K.; Qi, H.; Dai, J.-B; Yuan, Y.-J. Genome-wide landscape of position effects on heterogeneous gene expression in Saccharomyces cerevisiae. *Biotechnology for biofuels* **2017**, *10* (1), 189. Bai Flagfeldt, D.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast (Chichester, England)* **2009**, *26* (10), 545−551.

(44) Rubat, S.; Varas, I.; Sepulveda, R.; Almonacid, D.; Gonzalez-Nilo, F.; Agosin, E. Increasing the intracellular isoprenoid pool in Saccharomyces cerevisiae by structural fine-tuning of a bifunctional farnesyl diphosphate synthase. *FEMS yeast research* **2017**, *17*, 4. Dahl, R. H.; Zhang, F.; Alonso-Gutierrez, J.; Baidoo, E.; Batth, T. S.; Redding-Johanson, A. M. Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* **2013**, *31*, 1039.

(45) Deng, X.; Shi, B.; Ye, Z.; Huang, M.; Chen, R.; Cai, Y.; Kuang, Z.; Sun, X.; Bian, G.; Deng, Z.; et al. Systematic identification of Ocimum sanctum sesquiterpenoid synthases and (−)-eremophilene overproduction in engineered yeast. *Metabolic engineering* **2022**, *69*, 122−133.

(46) Livak, K. J.; Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25* (4), 402−408.

(47) Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nature biotechnology* **2019**, *37* (5), 540−546.

(48) Community, T. G. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res.* **2022**, *50* (W1), W345−W351.

(49) Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26* (5), 589−595.

(50) Diesh, C.; Stevens, G. J.; Xie, P.; De Jesus Martinez, T.; Hershberg, E. A.; Leung, A.; Guo, E.; Dider, S.; Zhang, J.; Bridge, C.; et al. JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome biology* **2023**, *24* (1), 74.