

Automated strain construction for biosynthetic pathway screening in yeast

Maria C. T. Astolfi^{1,2,3}, Sam D. Yoder^{2,3}, Marina Delfa-Lalaguna^{4,2}, Peter H. Winegar^{2,3,5}, Sara K. F. Holm^{6,2}, Mengziang Lei², Xixi Zhao^{2,3,5}, Stephen E. Tan^{2,3}, Randy Louie^{2,3}, Nathan J. Hillson^{2,3}, Graham A. Hudson^{2,3,5*} & Jay D. Keasling^{1,2,3,5,6,7*}

1- Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA

2- Joint BioEnergy Institute, Emeryville, CA 94608, USA

3- Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

4- Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

5- California Institute for Quantitative Biosciences (QB3 Institute), University of California, Berkeley, Berkeley, CA 94720, USA

6- The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark (DTU), Lyngby, Denmark

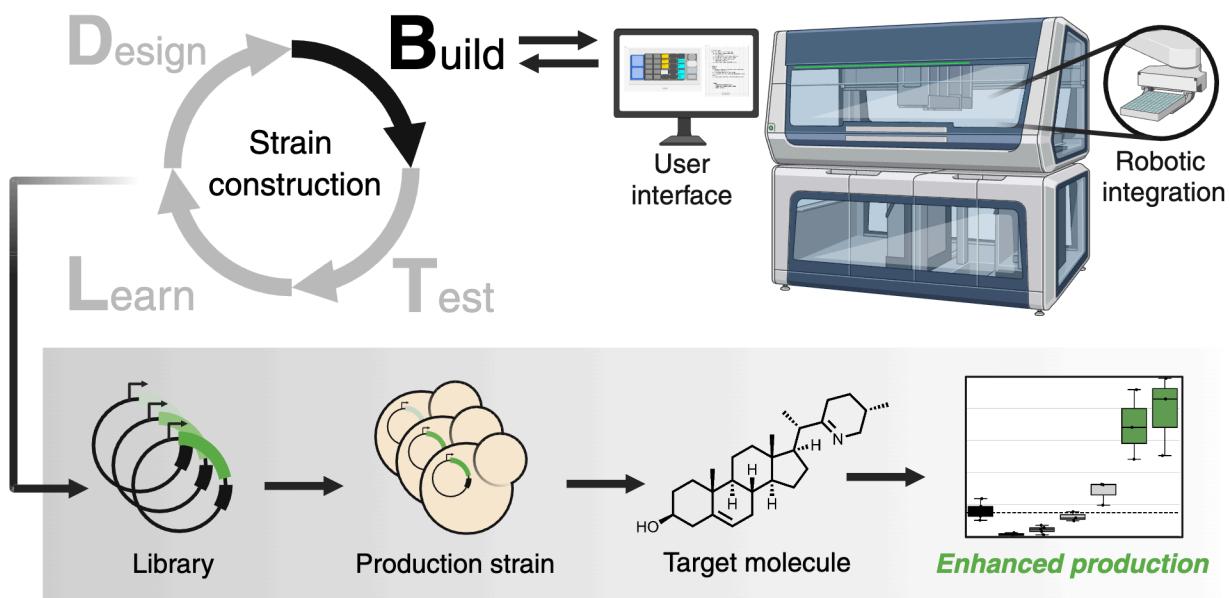
7- Department of Chemical and Biomolecular Engineering, University of California at Berkeley, Berkeley, CA 94720, USA

*Correspondence to: ghudson@berkeley.edu and keasling@berkeley.edu

Abstract *(max of 150 words; 149 words now)*

Automation accelerates the Design-Build-Test-Learn (DBTL) cycle for synthetic biology; however, most strain construction pipelines lack robotic integration. Here, we present the workflow design and source code for a modular, integrated protocol that automates the Build step in *Saccharomyces cerevisiae*. We programmed the Hamilton Microlab VANTAGE to integrate off-deck hardware via its central robotic arm, enabling automated steps that increased throughput to 2,000 transformations per week. We developed a user interface with the Hamilton VENUS software to support on-demand parameter customization. As a proof of concept, we screened a gene library in an engineered yeast strain producing verazine, a key intermediate in the biosynthesis of steroidal alkaloids. Our pipeline rapidly identified pathway bottlenecks and genes that enhanced verazine production by 2.0- to 5-fold. This technical note provides resources for synthetic biologists designing yeast workflows for biofoundries to screen libraries for pathway discovery/optimization, combinatorial biosynthesis, and protein engineering.

Graphical abstract



Keywords: automation; robotics integration; high-throughput screen; strain engineering; microbial biosynthesis; metabolic engineering.

Main text (max of 2000 words; 2063 words now)

Introduction

Biofoundries^{1–3} are facilities with broad academic and industrial interest⁴ that integrate computer-aided design, synthetic biology tools, and robotic automation. By accelerating the Design-Build-Test-Learn (DBTL) cycle for synthetic biology, biofoundries demonstrate the versatility, reproducibility, and scalability of automation, leading to their worldwide adoption⁴. As the field advances, there is a rising need to design higher-complexity workflows and deploy data infrastructures for machine learning^{5–8}. Achieving this vision will require sophisticated integration across systems.

One of the challenges of robotic workstation development is designing versatile and robust pipelines suitable for various use cases⁹. In our work, we developed a modular user interface that enables users to customize experimental parameters (**Fig. 1**). Importantly, established biofoundries play a critical role in disseminating automation worldwide. While automated yeast strain engineering has advanced significantly^{10–14}, including landmark early efforts^{15,16}, access to open and reusable automation resources remains limited. Here, we share a complete workflow, including Hamilton VENUS scripts, deck layouts, and integration design. Our pipeline is a practical resource for synthetic biologists developing yeast workflows, with robotic integration and user experience as core design principles.

Saccharomyces cerevisiae is a robust host for the biosynthesis of high-value molecules. We have demonstrated its versatility in the total or semi- biosynthesis of anti-malarials (artemisinin)¹⁷, anti-cancer drugs (vinblastine)¹⁸, cannabinoids¹⁹, and vaccine adjuvants (QS-21)²⁰. Scaling the production of such molecules relies on the rapid screening of gene libraries, a foundational strategy in metabolic engineering. Synthetic biologists routinely test novel genes, homologs, and mutants to discover, engineer, or evolve biocatalysts. However, manual and low-throughput workflows constrain these efforts to rational design, limiting the exploration of broader design spaces. In contrast, high-throughput screening accelerates the identification of pathway bottlenecks and expands the discovery of performance-enhancing genes. To address this limitation, we developed an automated pipeline at the Joint BioEnergy Institute's Robotics Lab for the iterative screening of biosynthetic pathways in yeast. As a proof of concept, we screened a gene library in an engineered *S. cerevisiae* strain to investigate the

biosynthesis of verazine, a key intermediate in the production of steroidal drug candidates such as cyclopamine²¹. Using our pipeline, we identified key pathway bottlenecks and candidate genes that enhanced verazine production by 2.0- to 5-fold (**Fig. 2**). This pipeline is broadly applicable to protein engineering, combinatorial biosynthesis, and pathway discovery/optimization for a wide range of industrial applications.

Results and Discussion

In this technical note, we report the integration of a custom robotic pipeline and user interface on the Hamilton Microlab VANTAGE for high-throughput transformation in *S. cerevisiae* (**Fig. 1**). The input, i.e., competent yeast and plasmid DNA, is converted into an output library of engineered strains compatible with automated colony picking, high-throughput culturing, and chemical extraction for downstream liquid chromatography-mass spectrometry (LC-MS) analysis. We describe the workflow integration, interface design, and its application (**Fig. 2**).

Custom integration and user interface design using Hamilton Microlab VANTAGE

The high-throughput transformation of *S. cerevisiae* is well-established²¹. We adapted the lithium acetate/ssDNA/PEG method from tubes to a 96-well format, benchmarking variables such as cell density, reagent volumes, and DNA concentration (**Fig. S1**). The optimized conditions formed the basis of a Standard Operating Procedure (SOP) described in the Methods section, which guided the programming of robot-executable tasks. The Hamilton VANTAGE platform (**Fig. 1b, S2**) was selected for its modular deck layout and capacity for hardware integration.

The workflow was programmed using Hamilton VENUS 5 (Venus on Vantage v2.2.13.4) and divided into discrete, modular steps: 1) “Transformation set up and heat shock”, 2) “Washing”, and 3) “Plating” (**Fig. 1a**). The heat-shock step, the most time-intensive, was fully automated by programming the VANTAGE robotic arm (Hamilton iSWAP) to interact with external off-deck devices, including a plate sealer, plate peeler, and thermal cycler (**Fig. 1b**). The deck layout for the method included sequences to move sample plates between the pipetting deck and off-deck devices. Integration of external equipment was achieved through instrument-specific software drivers and communication protocols using Hamilton device

Technical note

libraries, including the Inheco ODTC (96-well thermocycler), 4titude_a4S (plate sealer), and HSLBrooksAutomationXPeel (plate peeler) (**Fig. 1b**). This critical integration allowed hands-free operation after a manual deck setup. To load the deck, users are prompted to arrange labware inputs in predefined positions according to a customized deck image displayed at the start of each step. (**Fig. 1c**).

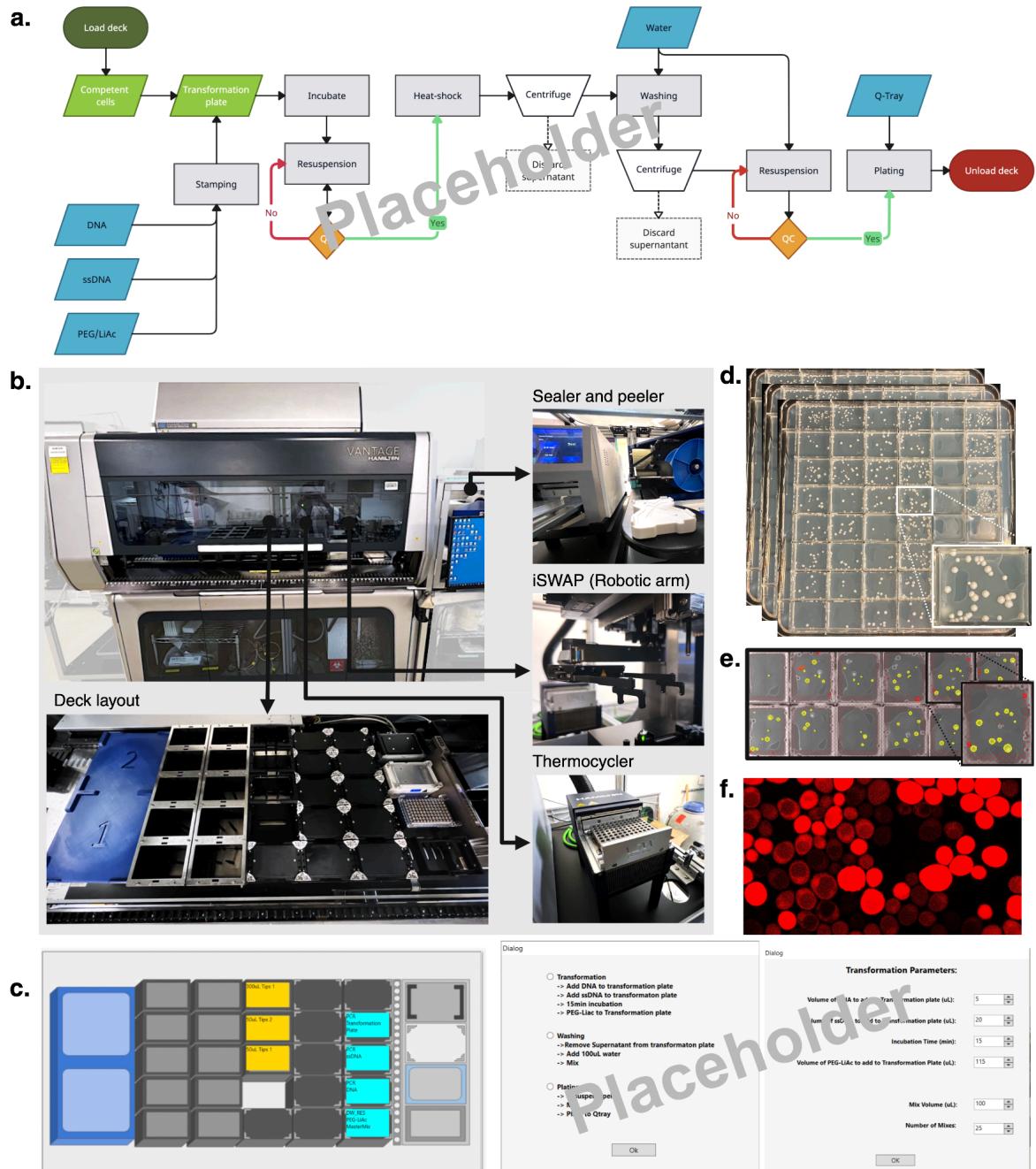


Figure 1. Automated *S. cerevisiae* transformation using the Hamilton Microlab VANTAGE. (a) Schematic of the modular and integrated yeast transformation workflow, comprising automated steps for “Transformation setup and heat shock,” “Washing,” and “Plating,” with integrated QC checkpoints and programmable decision branches. (b) Hamilton Microlab VANTAGE platform configured with a flexible on-deck layout and integrated peripherals, including off-deck devices such as the Inheco ODTC (96-well thermocycler), 4titude_a4S (plate sealer), and HSLBrooksAutomationXPeel (plate peeler) via the Hamilton iSWAP robotic arm. (c) Custom interface displaying a digital deck layout and dialog boxes for guided labware setup, modular transformation step selection, and parameter customization. (d) Output of the automated pipeline plated on a 48-well Q-Tray, displaying the growth of engineered *S. cerevisiae* harboring a high-copy 2 μ vector containing a *leu2* auxotrophic marker and an RFP expression cassette. (e) Colony detection using the robot QPix 460 automated colony picker; yellow colonies indicate selectable clones, while red ones are excluded. (f) Confocal microscopy of RFP-expressing engineered yeast, showing the heterogeneity of plasmid expression in *S. cerevisiae*.

Programming the robot involved creating and troubleshooting liquid classes for each reagent to ensure accurate pipetting. Reagents like PEG were challenging due to their viscosity, leading to unreliable transfer volumes in early trials. To improve pipetting accuracy, we adjusted aspiration and dispensing speeds, optimized air gaps, and pre- and post-dispensing parameters (**Fig. S3**). Rounds of dry runs (i.e., experiments with reagents only) were conducted during the development to verify arm movements, transfers, and interactions with external devices, reducing errors and refining protocols before wet testing.

In the design process for the user interface, each workflow step was made modular and customizable via dialog boxes (**Fig. 1c, S4**). Parameters such as DNA volume, lithium acetate/ssDNA/PEG ratios, and heat shock incubation times can be adjusted at the start of the experiment to accommodate a variety of experimental needs. Checkpoints were programmed to detect errors, such as incomplete cell resuspension, and initiate corrective loops to ensure robust performance (**Fig. S4**).

Validation of the automated pipeline was conducted by transforming competent *S. cerevisiae* with a high-copy 2 μ vector containing a *leu2* auxotrophic marker and a gene encoding red fluorescent protein (RFP) (**Fig. 1d**). Following transformation, colonies were successfully picked using the QPix 460 automated colony picker, demonstrating compatibility between the VANTAGE-generated output and downstream automation (**Fig. 1e**). Robot-picked colonies were inoculated for high-throughput culturing in 96-deep-well plates and grown in selective media. Cultures were then imaged using confocal microscopy (**Fig. 1f**).

Technical note

Our automated method was validated for 96 transformations per run, yielding a large number of colonies for each transformation (**Fig. 1d**). Each workflow required ~2 hours of robotic execution, including 1.5 hours of automated setup and hands-off heat shock. The pipeline achieved a capacity of ~400 transformations per day and up to 2,000 per week. In contrast, a human operator performing 16 transformations per day would complete 80 reactions per week, a 25-fold reduction in throughput. While manual throughput varies across laboratories, yeast transformation is broadly regarded as a labor-intensive protocol. Quantitative comparison between manual and automated workflows is provided in **Fig. S5**.

Biosynthetic pathway screening using the automated pipeline

The automated workflow was applied to overexpress a library of 32 genes in a verazine-producing *S. cerevisiae*²² (**Table S1-2**) with the goal of increasing heterologous production. Selected candidates (**Fig. 2a**) included genes encoding enzymes from the native sterol biosynthetic pathway²³ (i.e., ERG1, ERG7, ERG11, NCP1, ERG24, ERG25, ERG26, ERG27, ERG28, ERG29, ERG2, and ERG3) and the heterologous verazine biosynthetic pathway²² (i.e., StDHCR7, GgDHCR24, DzCYP90B71, AtCPR, VnCYP94N2, VcGABAT1v2, VcCYP90G1v3, and SvMSBP). Genes encoding proteins that derivatize sterols for transport/export^{24,25} (i.e., ATF1, ATF2, ARE1, and ARE2) and related to the hypothesized storage of sterols in lipid droplets^{26–31} (i.e., LDB16, SEI1, NEM1, SPO7, PAH1, DGA1, and FAS2) were also selected.

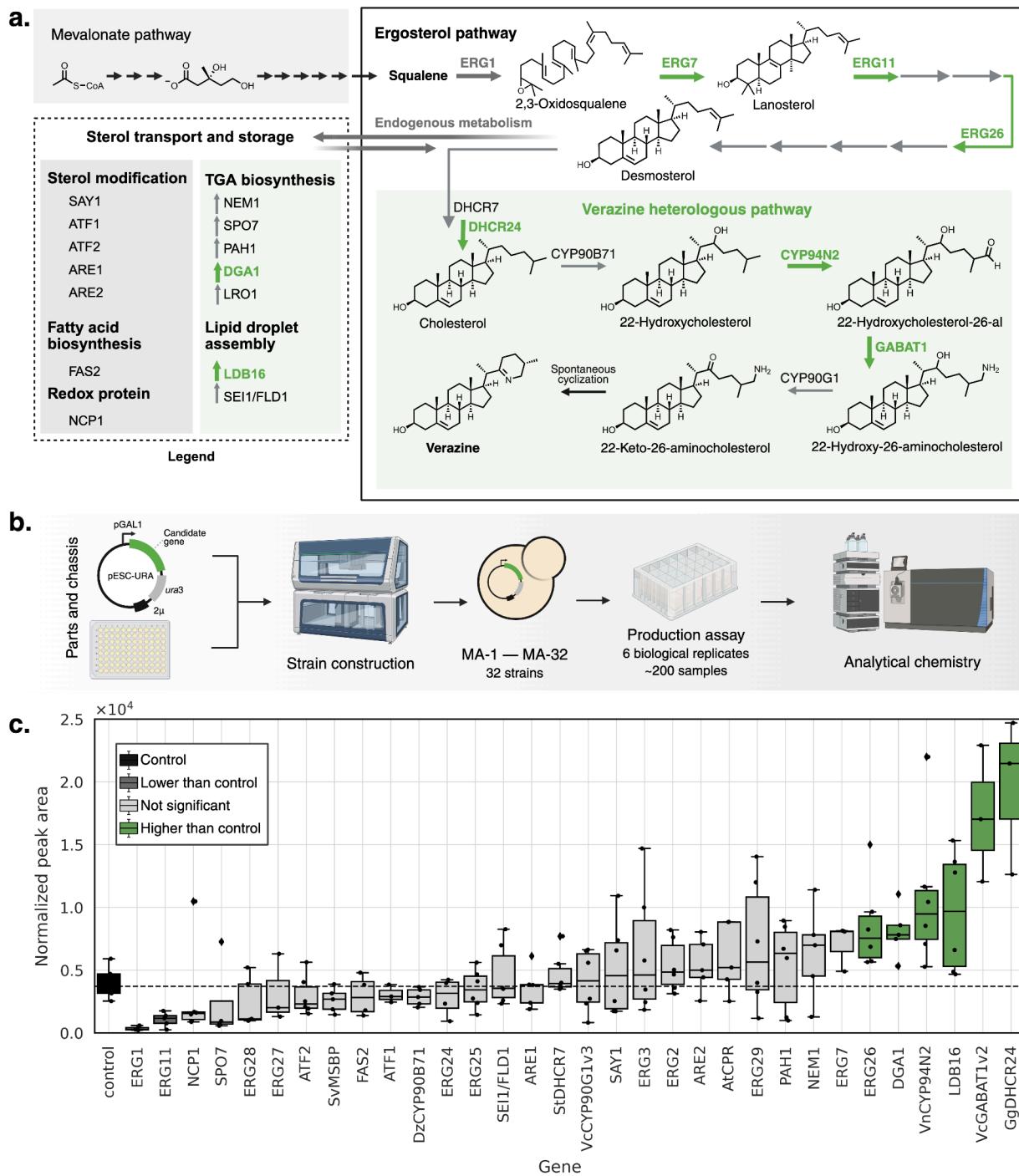


Figure 2. Automated gene library screening for verazine biosynthesis in engineered *S. cerevisiae*. (a) Genes selected for screening include the heterologous verazine biosynthetic pathway and endogenous genes involved in sterol metabolism, fatty acid biosynthesis, redox balance, TGA biosynthesis, and lipid droplet assembly. Enzymatic steps in the proposed verazine pathway are highlighted in the green box. Arrow coloring corresponds to data in panel (c): green arrows indicate genes that significantly increased verazine production; gray arrows indicate genes with no significant effect or a significant decrease. (b) Overview of the automated workflow, including transformation of a plasmid library

Technical note

into the verazine-producing *S. cerevisiae* chassis (PW-42), culture of 32 engineered strains (MA-1 to MA-32) in six biological replicates, and LC-MS analysis. **(c)** Verazine peak area normalized by optical density. Box plots represent six biological replicates per strain. Genes are colored based on statistical significance relative to the control strain in black: green, significantly higher; gray, not significant; dark gray, significantly lower (Welch's t-test, $P < 0.05$). Outliers may arise from biological variability due to heterogeneous expression of high-copy 2 μ plasmids in *S. cerevisiae*, as visualized in Figure 1f.

Each gene was cloned into a pESC-URA plasmid under the transcriptional regulation of the *GAL1* promoter (pGAL1) for inducible expression (**Fig. 2b**). The resulting plasmids were transformed into the verazine-producing *S. cerevisiae* strain PW-42²², generating a library of engineered strains (MA-1 through MA-32) (**Table S3**). Six biological replicates of each strain, including a negative control (i.e., PW-42 harboring an empty plasmid), were picked, resulting in a ~200-sample library. To enable high-throughput screening, a chemical extraction method was developed using Zymolyase-based cell lysis followed by organic solvent extraction. A rapid LC-MS protocol was adapted to reduce the verazine detection runtime from 50 minutes to 19 minutes, allowing efficient quantification of titers across the library.

Screening results revealed several genes that enhanced verazine production relative to the control (**Fig. 2c**). The top-performing strains overexpressed *erg26*, *dga1*, *cyp94n2*, *ldb16*, *gabat1v2*, or *dhcr24*, exhibiting a 2- to 5-fold increase in normalized titer. These genes span many functional categories, highlighting the role of precursor supply, sterol metabolism, and downstream processing in the verazine biosynthetic pathway. Overexpression of *erg7* and *erg26*, which encode oxidosqualene cyclase and C-3 sterol dehydrogenase²³, respectively, likely increases flux through core sterol biosynthesis, enhancing the availability of downstream intermediates. DGA1^{29,30}, a diacylglycerol acyltransferase, and LDB16²⁶, involved in lipid droplet assembly, may mitigate toxicity of verazine or its intermediates by enhancing intracellular storage capacity of yeast for steroidal products. These functions are associated with sterol storage and transport, suggesting that improved ER-to-droplet trafficking may sequester hydrophobic intermediates, buffer pathway imbalances, and reduce cellular stress.

In contrast, overexpression of *erg1*, *erg11*, and *ncp1* negatively impacted production. *erg1*, encoding squalene epoxidase (ERG1), may lead to the accumulation of reactive epoxides when dysregulated, contributing to membrane stress and cellular toxicity^{23,32}. *erg11*, which encodes lanosterol demethylase (ERG11), may disrupt sterol homeostasis or trigger feedback

Technical note

inhibition when overexpressed³². *ncp1*, the endogenous P450 reductase (NCP1), could perturb redox balance, altering NADPH availability and interfering with electron transfer to native and heterologous P450 enzymes³³. These may hinder metabolic flux and impair cell viability under pathway burden³⁴. Notably, these strains exhibited reduced growth, and verazine production was positively correlated with optical density (**Fig. S8-9**).

Interestingly, the positive effects of GgDHCR24, VnCYP94N2, and VcGABAT1v2, which catalyze three steps in the verazine pathway, highlight these enzymes as bottlenecks in the biosynthetic route²² (**Fig. 2a**). GgDHCR24, a Δ24(25)-reductase, converts 7-dehydrodesmosterol to cholesterol, establishing the sterol backbone for subsequent modifications. VnCYP94N2 catalyzes C26 oxidation of 22-hydroxycholesterol to generate 22-hydroxycholesterol-26-al, and VcGABAT1v2 mediates transamination of this aldehyde to form 22-hydroxy-26-aminocholesterol²². Their overexpression may relieve rate-limiting steps in both precursor supply and late-stage intermediate processing.

A recent study identified candidate genes for verazine biosynthesis in *Veratrum grandiflorum* using comparative transcriptomics and homolog mining, followed by heterologous expression and functional validation in *S. cerevisiae*³⁵. Their work successfully refactored the pathway and uncovered that *game4*, a gene from tomato (*Solanum lycopersicum*), catalyzes C-26 aldehyde formation, enhancing verazine titers by 2.4-fold. The verazine pathway is a striking example of how homologs from various plant species can outperform native enzymes^{22,35}. Unlocking the full genetic landscape across many species will be key to overcoming pathway bottlenecks.

These results further illustrate the value of automated platforms for screening biosynthetic pathways. In our work, we expanded the screen beyond pathway enzymes to include core sterol metabolism genes that modulate precursor flux, product transport and storage, and intracellular compartmentalization. This broader strategy uncovered previously unknown candidates. By integrating transcriptomics-guided gene discovery with systems biology tools, such as metabolic network analysis, pathway databases, and flux balance modeling, automated efforts will rapidly explore biosynthetic potential to engineer strains for robust, scalable production.

Finally, future work will extend our pipeline to automate genome engineering of biosynthetic pathways, accelerating the development of production strains. While we have reported the robotics-assisted Build step in *S. cerevisiae*, full integration of automation across the DBTL cycle remains a central objective. For example, the Design phase can be advanced by integrating pathway design tools and genome-scale metabolic models. Test throughput can be enhanced using high-throughput spectrometry platforms, such as RapidFire, enabling the acquisition of large analytical datasets. On the Learn front, robust data infrastructure will be critical to support machine learning, paving the way for self-driving and autonomous research. As the field advances, we envision a community-driven effort to establish an open repository of validated robotic protocols, liquid-handling methods, and automation scripts spanning the DBTL cycle. Such a resource would promote reproducibility, foster collaboration, and democratize access to automation in synthetic biology.

Conclusion

We present the design and source code of an integrated workflow for yeast strain construction using the Hamilton Microlab VANTAGE. The pipeline was applied to overexpress 32 candidate genes in a verazine-producing strain, enabling the screening of ~200 engineered strains. This effort identified rate-limiting steps and performance-enhancing genes across heterologous pathway and core metabolism, improving verazine titers by up to 5-fold. This technical note serves as a practical resource for synthetic biologists developing yeast-based methods for biofoundries.

Methods

Strains and chemicals

The verazine-producing engineered *S. cerevisiae* PW-42 was obtained from the Joint BioEnergy Institute's Inventory of Composable Elements^{36,37} (Part ID JBx_257708 available at <https://registry.jbei.org/entry/257708>). Authentic verazine (CAS number 14320-81-1) standard was purchased from MedChemExpress (HY-N11911). Reagents for the formulation of synthetic complete (SC) media were purchased from Sunrise Science Products. All other media

components and solvents were purchased from Sigma-Aldrich.

Plasmid construction

Coding sequences (CDSs) of endogenous genes for plasmid-based overexpression were obtained from the *Saccharomyces* Genome Database³⁸. The CDSs of heterologous genes were obtained from the literature²². Oligonucleotides for amplifying the CDSs (**Table S4**) were designed using Benchling and synthesized in a pre-mixed, 96-well plate format by Integrated DNA Technologies (IDT). CDSs were amplified from verazine-producing *S. cerevisiae* PW-42 by colony PCR (cPCR) using Q5 High-Fidelity DNA Polymerase (New England Biolabs) or Phanta Max Super-Fidelity DNA Polymerase (Vazyme). Amplified fragments were assembled into pESC-URA plasmids using Gibson Assembly (HiFi DNA Assembly Master Mix, New England Biolabs) in 96-well plates. An aliquot of each assembly reaction was transformed into *Escherichia coli* XL1-Blue competent cells (Agilent) and plated on Bioassay Q-Trays with 48-well dividers (Molecular Devices) containing LB agar supplemented with 100 µg/mL carbenicillin using the Hamilton Microlab VANTAGE platform (**Fig. S10**). The plates were incubated at 37 °C overnight. Recombinant colonies were screened by cPCR using primers that anneal to the pGAL1 promoter and tCYC1 terminator regions to verify the insert size. cPCR-verified recombinant colonies were cultured in LB medium supplemented with 100 µg/mL carbenicillin at 37 °C and 250 rpm overnight, followed by plasmid extraction using the QIAprep 96 Turbo Miniprep Kit (Qiagen). Plasmids were sequenced by Sanger sequencing (Genewiz) or next-generation sequencing (Plasmidsaurus).

Competent cells

A single, fresh *S. cerevisiae* PW-42 colony was inoculated into 5 mL of YPD medium and incubated at 30 °C with shaking at 250 rpm overnight. The following day, OD₆₀₀ was measured, and the cells were subcultured into 50 mL of fresh YPD in baffled 250 mL flasks at an initial OD₆₀₀ of 0.1. Cultures were grown for ~6 h at 30 °C with shaking at 250 rpm until OD₆₀₀ reached 0.8-1.0. Cells were harvested by centrifugation at 4,500 × g for 5 minutes. The supernatant was discarded, and the pellet was washed with 10 mL of sterile deionized water. Cells were then washed with 10 mL of 100 mM lithium acetate (LiAc), followed by a final resuspension in 100

mM LiAc. For 96 transformations, approximately 1 L of culture was resuspended in 2 mL of 100 mM LiAc. 25 µL of competent cells were aliquoted into a 96-well PCR plate and stored at -80 °C.

Automated transformation of *S. cerevisiae*

Transformation was performed using a custom workflow (**Fig. 1a**) programmed on the Hamilton Microlab VANTAGE platform (**Fig. 1b**). Competent *S. cerevisiae* cells stored at -80 °C in 96-well PCR plate (25 µL aliquots) were thawed and loaded on deck along with plates containing the library DNA (normalized to 100 ng/µL), boiled single-stranded DNA (Sigma-Aldrich), and a reservoir containing a PEG/LiAc mastermix (prepared as 100 µL of 50% w/v PEG3350 and 15 µL of 1 M lithium acetate per transformation), following the VANTAGE deck setup instructions (**Fig. 1c**).

The transformation protocol was initiated by selecting the “Transformation” step in the VENUS interface (**Fig. 1c**). Using the Hamilton CO-RE 96 Liquid Handling Head, 5 µL of library DNA and 20 µL of ssDNA were stamped onto the competent cell plate, followed by 115 µL of the PEG/LiAc mastermix. Dispensing of PEG/LiAc was carried out using custom liquid class parameters optimized for the reliable transfer of viscous solutions (**Fig. S3**). Reactions were mixed by pipetting up and down 25x. Plates were sealed using the CO-RE Gripper Arm (**Fig. 1b**), which integrated with the off-deck peripheral 4titude a4S plate sealer. Then, the sealed plate was transferred to the Inheco ODTC thermocycler for a 1 h heat shock at 42 °C. This step is fully automated after the deck setup.

Following heat shock, the “Washing” step was selected in the VENUS interface (**Fig. 1c**). Plates were centrifuged at 4,500 × g for 3 min, unsealed, and 145 µL of supernatant was carefully removed using the CO-RE 96 Liquid Handling Head. Sterile water (100 µL) was added from a reservoir and mixed 10x. The plates were resealed, centrifuged again (4,500 × g, 3 min), and the supernatant was discarded. An additional 100 µL of sterile water was added to resuspend the cells, followed by thorough mixing. Quality control dialog boxes (**Fig. S4**) were used to verify the complete resuspension of the pellet, with corrective loops triggered as needed to ensure consistent performance.

The “Plating” step was then initiated (**Fig. 1c**). Bioassay Q-Trays with 48-well dividers (Molecular Devices) containing solid SC-URA or SC-LEU medium (synthetic complete dropout medium minus uracil or minus leucine, respectively) supplemented with 2% (w/v) glucose were

loaded on deck. The Hamilton MagPip 8-Channel Liquid Handling Arm was used for plating. Dual liquid level detection (capacitance and pressure-based) enabled precise height adjustment relative to the agar surface to ensure accurate dispensing. A volume of 100 µL was dispensed into each well of the Q-Trays. Q-Trays were dried under airflow in a biosafety cabinet and then incubated at 30 °C for 3 days (**Fig. 1d**).

High-throughput culture assays

From a fresh single colony grown on Q-Trays, six biological replicates of each strain were picked using the QPix 460. Samples were inoculated into 2 mL of SC-URA medium (synthetic complete dropout medium minus uracil, Sunrise Science Products) supplemented with 2% (w/v) glucose in 24-deep-well plates using an adjustable multichannel pipette (Eppendorf). Pre-culture plates were sealed and incubated at 30 °C with 1,000 rpm shaking (Infors HT) for 48 h. To initiate the production phase, the cultures were centrifuged at 2,500 × g for 5 min, and the supernatant was discarded. Cell pellets were resuspended in SC-URA supplemented with 2% (w/v) galactose to induce gene expression. Production cultures were incubated at 30 °C with shaking at 1,000 rpm (Infors HT) for 48 h.

High-throughput chemical extraction

After 48 h of induction, cultures were transferred from 24-deep-well plates to 96-deep-well plates using an adjustable multichannel pipette (Eppendorf). Optical density at 600 nm (OD₆₀₀) was measured by diluting 5 µL of culture into 95 µL of sterile water in a clear, flat-bottom 96-well plate, followed by measurement using a plate reader (BioTek). For cell lysis, cultures were pelleted at 4,500 × g for 5 min and then resuspended in 150 µL of Zymolyase mastermix, composed of 1 µL Zymolyase 100T (Zymo Research) and 149 µL sorbitol/EDTA buffer (1.822 g sorbitol and 0.292 g Na₂EDTA in 10 mL water)³⁹. The plates were incubated at 37 °C for 1 h. Lysis was followed by extraction with 450 µL of LC-MS-grade methanol (Sigma). Plates were shaken at 1,000 rpm (Infors HT) for 1 h and centrifuged at 4,500 × g for 15 min. 100 µL aliquots were stamped into LC-MS-compatible 96-well plates (InfinityLab, round wells, V-shaped, Agilent) and sealed with silicone sealing mats (Agilent) for LC-MS analysis. All liquid transfers were performed in high-throughput using a Liquidator 96-channel pipetting system (Mettler Toledo).

Technical note

For samples with low verazine production, 600 µL of supernatant was transferred post-lysis, concentrated using a vacuum centrifuge (SpeedVac), and reconstituted in 50 µL of methanol, followed by sonication.

Detection of verazine and LC-MS data analysis

Detection of verazine was performed using a 1260 Infinity II LC/MSD XT (Agilent) or a 1260 Infinity II LC/MSD iQ (Agilent) equipped with an EC UHPLC Nucleodur C18 Htec column (1.8 µm, 100 × 2 mm; Macherey-Nagel, 760306.20). For high-throughput detection, a 5 µL injection was made from LC-MS-compatible 96-well plates (InfinityLab, round wells, V-shape; Agilent) sealed with silicone sealing mats (Agilent). For manual detection, chemically extracted samples were transferred to an LC/MS vial with an insert adapter. The LC-MS method was adapted from an established method²² to reduce runtime from 50 min to 19 min (**Table S5**), using a solvent gradient from A (LC-MS-grade water with 0.1% formic acid) to B (LC-MS-grade acetonitrile with 0.1% formic acid) as follows: 85:15 A:B held for 0.75 min, ramping to 40:60 over 9.25 min, then to 0:100 over 0.25 min, held at 100% B for 1.75 min, and re-equilibrated to 85:15 A:B over 6.75 min. Full mass spectra were generated for metabolite identification by scanning within an *m/z* range of 100-1,000 in positive-ion mode.

Data were acquired using OpenLab CDS v2.4 (Agilent). Verazine production was confirmed by co-elution with a purified standard (*m/z* = 398.3) (**Fig. S11**), and verazine levels were assessed by integrating the single-ion monitoring (SIM) peak area. Values are reported as relative peak areas. All data analysis was performed in Python using Jupyter Notebook. Approximately 10% of library samples lacked detectable peak area values due to low signal intensity, suboptimal cell growth, or technical variability during sample preparation or LC-MS injection, and were excluded from subsequent analysis (**Fig. S12**). Peak area values were normalized to optical density (OD) to account for biomass differences across samples (**Fig. S8-9**). Within each gene group, comprising 6 biological replicates, replicate-level outliers were identified using a z-score threshold of ±2.5. Group-wise statistical comparisons to the control (gene 0) were conducted using Welch's t-test, with significance defined as *p* < 0.05. Genes exhibiting significantly different production levels were classified as either higher or lower producers relative to the control.

Associated Content

Data Availability

The data underlying this study are available in the Supporting Information, which includes raw experimental data (CSV), workflow videos (MP4), and a document comprising Figures S1-S12 and Tables S1-S5 (PDF), as shown below. Source code for the “Automated Yeast Transformation” workflow is available at the [GitHub repo XXXX \(GitHub link XXXX\)](#) under the Lawrence Berkeley National Laboratory IP Office license [XXXXXXXX](#). Strains and plasmids are available through the Joint BioEnergy Institute Inventory of Composable Elements (<https://public-registry.jbei.org>, collection [XXXXXX](#)), with accession numbers provided in the Supporting Information.

Supporting Information

The Supporting Information will be available at the ACS Synthetic Biology website.

- Supporting figures showing benchmarking of yeast high-throughput transformation (Figure S1), deck layout for the Hamilton Microlab VANTAGE platform (Figure S2), workflow liquid class parameters (Figure S3), workflow user interface (Figure S4), transformation efficiency of manual and automated methods (Figure S5), optical density of engineered strains (Figure S8), scatter plots of optical density and verazine peak-area (Figure S9), high-throughput *E. coli* transformation (Figure S10), LC-MS chromatograms (Figure S11), and example of verazine production variability (Figure S12); supporting tables listing the genotype of the verazine-producing *S. cerevisiae* strain PW-42 (Table S1), library of 32 genes for overexpression (Table S2), plasmid and library strain list (MA-1 to MA-32) (Table S3), oligos list (Table S4), and adapted LC-MS method (Table S5) (PDF)
- Coding sequences (CDSs) of genes used for overexpression (CSV)
- Video files showing robotic integration and deck layout setup (Video S1), automated plating onto Q-Trays (Video S2), and the user interface with parameter customization (Video S3) (MP4)

Acknowledgements

This research was supported by the Homeworld Collective Garden Grants and the UC Berkeley Bakar Labs through the BioEnginuity Impact Grant; by the Joint BioEnergy Institute, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-AC02-05CH11231; by the National Institute of Standards and Technology (Award Number 70NANB22H017) and with the support of the Bioindustrial Manufacturing and Design Ecosystem (BioMADE); by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (Award Number 1R01AI186111-01); and by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (Award Number F32GM153046). The content is solely the responsibility of the authors and does not necessarily represent the official views of BioMADE or the NIH. The Graphical Abstract, Figures 1 and 2, were created using the software BioRender under the Open Access license XXXXXX. We thank Jacob Roberts, Allison Pearson, and Isaac Donnell for valuable discussions related to this work. We are grateful to Robin Johnston for her support in intellectual property matters and open licensing of the workflow. We also thank Madison Mennie and Olga Diaz for their essential daily administrative support.

Author Information

Author Contribution

M.C.T.A., G.A.H., and J.D.K. conceptualized the workflow. S.Y. and M.C.T.A. programmed the robot. P.H.W. designed the biosynthetic pathway study, selected the genes, and provided the verazine production strain and the authentic standard. M.D.L., S.H., M.L., and M.C.T.A. characterized the automation pipeline. M.D.L. and M.C.T.A. constructed plasmids, yeast strains, and performed assays and chemical extractions. S.Y., S.T., and R.L. contributed to robotic integration, protocol optimization, and interface development. G.A.H. and P.H.W. guided analytical chemistry and LC-MS analysis. M.C.T.A. conducted data analysis and drafted the manuscript. G.A.H., P.H.W., X.Z., N.J.H., and J.D.K. discussed, provided feedback, revisions, and final editing. All authors contributed to and approved the manuscript.

Conflict of Interest

J.D.K. has a financial interest in Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, BioMia, Cyklos Materials, Demetrix, Lygos, Napigen, ResVita Bio, and Zero Acre Farms. . N.J.H. declares financial interests in TeselaGen Biotechnologies Inc., and Ansa Biotechnologies, Inc. All other authors declare no competing interests.

Abbreviations

DBTL, Design-Build-Test-Learn cycle

SOP, standard operating procedure

PEG, polyethylene glycol

VANTAGE, Hamilton Microlab VANTAGE robotic platform

VENUS, Hamilton VENUS software

CDS, coding sequence

OD, optical density

CFU, colony-forming unit

SIM, single ion monitoring

LC-MS, liquid chromatography-mass spectrometry

Q-Tray, quad-format culture plate.

References

1. Chao, R., Yuan, Y. & Zhao, H. Building biological foundries for next-generation synthetic biology. *Science China Life Sciences* **58**, 658–665 (2015).
2. Eisenstein, M. Living factories of the future. *Nature* **531**, 401–403 (2016).
3. Chao, R., Mishra, S., Si, T. & Zhao, H. Engineering biological systems using automated biofoundries. *Metabolic Engineering* **42**, 98–108 (2017).
4. Hillson, N. *et al.* Building a global alliance of biofoundries. *Nature Communications* **10**, 1–4 (2019).
5. Carbonell, P., Radivojevic, T. & Martín, H. G. Opportunities at the Intersection of Synthetic Biology, Machine Learning, and Automation. *ACS Synthetic Biology* (2019) doi:10.1021/acssynbio.8b00540.
6. Martin, H. G. *et al.* Perspectives for self-driving labs in synthetic biology. *Current Opinion in Biotechnology* **79**, 102881 (2023).
7. Zhang, Q. *et al.* Integrating protein language models and automatic biofoundry for enhanced protein evolution. *Nature Communications* **16**, 1–16 (2025).
8. Singh, N. *et al.* A Generalized Platform for Artificial Intelligence-powered Autonomous Protein

- Engineering. *bioRxiv*.
9. Holowko, M. B., Frow, E. K., Reid, J. C., Rourke, M. & Vickers, C. E. Building a biofoundry. *Synthetic Biology* **6**, (2021).
 10. Si, T. *et al.* Automated multiplex genome-scale engineering in yeast. *Nature Communications* **8**, 1–12 (2017).
 11. Liu, G., Lanham, C., Ross Buchan, J. & Kaplan, M. E. High-throughput transformation of *Saccharomyces cerevisiae* using liquid handling robots. *PLOS ONE* **12**, e0174128 (2017).
 12. Seong, M.-J. *et al.* Automated Construction of a Yeast-Based Multigene Library via Homologous Recombination in a Biofoundry Workflow. *ACS Synthetic Biology* (2025) doi:10.1021/acssynbio.4c00812.
 13. Rajakumar, P. D. *et al.* Rapid Prototyping Platform for *Saccharomyces cerevisiae* Using Computer-Aided Genetic Design Enabled by Parallel Software and Workcell Platform Development. *SLAS Technology* **24**, 291–297 (2019).
 14. Fleming, M. S. & Gitler, A. D. High-throughput yeast plasmid overexpression screen. *J Vis Exp* (2011) doi:10.3791/2836.
 15. King, R. D. *et al.* Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252 (2004).
 16. King, R. D. *et al.* The automation of science. *Science* **324**, 85–89 (2009).
 17. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
 18. Zhang, J. *et al.* A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* **609**, 341–347 (2022).
 19. Luo, X. *et al.* Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123–126 (2019).
 20. Liu, Y. *et al.* Complete biosynthesis of QS-21 in engineered yeast. *Nature* **629**, 937–944 (2024).
 21. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**, 31–34 (2007).
 22. Winegar, P. H. *et al.* Verazine biosynthesis from simple sugars in engineered *Saccharomyces cerevisiae*. *Metabolic Engineering* **85**, 145–158 (2024).
 23. Jordá, T. & Puig, S. Regulation of Ergosterol Biosynthesis in *Saccharomyces cerevisiae*. *Genes* **11**, 795 (2020).
 24. Tiwari, R., Köffel, R. & Schneiter, R. An acetylation/deacetylation cycle controls the export of sterols and steroids from *S. cerevisiae*. *The EMBO Journal* (2007) doi:10.1038/sj.emboj.7601924.

25. Jensen-Pergakes, K., Guo, Z., Giattina, M., Sturley, S. L. & Bard, M. Transcriptional Regulation of the Two Sterol Esterification Genes in the Yeast *Saccharomyces cerevisiae*. *Journal of Bacteriology* (2001) doi:10.1128/jb.183.17.4950-4957.2001.
26. Wang, C.-W., Miao, Y.-H. & Chang, Y.-S. Control of lipid droplet size in budding yeast requires the collaboration between Fld1 and Ldb16. *J Cell Sci* **127**, 1214–1228 (2014).
27. Khondker, S., Han, G.-S. & Carman, G. M. Phosphorylation-mediated regulation of the Nem1-Spo7/Pah1 phosphatase cascade in yeast lipid synthesis. *Advances in Biological Regulation* **84**, 100889 (2022).
28. Han, G.-S. & Carman, G. M. Yeast PAH1-encoded phosphatidate phosphatase controls the expression of CHO1-encoded phosphatidylserine synthase for membrane phospholipid synthesis. *Journal of Biological Chemistry* **292**, 13230–13242 (2017).
29. Yu, Y. *et al.* Engineering *Saccharomyces cerevisiae* for high yield production of α-amyrin via synergistic remodeling of α-amyrin synthase and expanding the storage pool. *Metabolic Engineering* **62**, 72–83 (2020).
30. Sorger, D. & Daum, G. Triacylglycerol biosynthesis in yeast. *Applied Microbiology and Biotechnology* **61**, 289–299 (2003).
31. Shin, G.-H., Veen, M., Stahl, U. & Lang, C. Overexpression of genes of the fatty acid biosynthetic pathway leads to accumulation of sterols in *Saccharomyces cerevisiae*. *Yeast* **29**, 371–383 (2012).
32. Bhattacharya, S., Esquivel, B. D. & White, T. C. Overexpression or Deletion of Ergosterol Biosynthesis Genes Alters Doubling Time, Response to Stress Agents, and Drug Susceptibility in *Saccharomyces cerevisiae*. *mBio* **9**, e01291–18 (2018).
33. Otto, M., Teixeira, P. G., Vizcaino, M. I., David, F. & Siewers, V. Integration of a multi-step heterologous pathway in *Saccharomyces cerevisiae* for the production of abscisic acid. *Microbial Cell Factories* **18**, 1–17 (2019).
34. Nowrouzi, B. & Rios-Solis, L. Redox metabolism for improving whole-cell P450-catalysed terpenoid biosynthesis. *Critical Reviews in Biotechnology* 1213–1237 (2021).
35. Hong, K. *et al.* Identification of genes involved in verazine biosynthesis in *Veratrum grandiflorum* and their heterologous production in *Saccharomyces cerevisiae*. *BMC Plant Biology* **25**, 1–15 (2025).
36. Ham, T. S. *et al.* Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic acids research* **40**, (2012).
37. Plahar, H. A. *et al.* BioParts—A Biological Parts Search Portal and Updates to the ICE Parts Registry Software Platform. *ACS Synthetic Biology* (2021) doi:10.1021/acssynbio.1c00263.
38. Engel, S. R. *et al.* *Saccharomyces* Genome Database: advances in genome annotation, expanded

Technical note

- biochemical pathways, and other key enhancements. *Genetics* **229**, iyae185 (2024).
39. Chen, Y. *et al.* Automated ‘Cells-To-Peptides’ Sample Preparation Workflow for High-Throughput, Quantitative Proteomic Assays of Microbes. *Journal of Proteome Research* (2019)
doi:10.1021/acs.jproteome.9b00455.