

A roadmap to establish a comprehensive platform for sustainable manufacturing of natural products in yeast

Received: 1 December 2022

Gita Naseri ^{1,2} 

Accepted: 24 March 2023

Published online: 06 April 2023

 Check for updates

Secondary natural products (NPs) are a rich source for drug discovery. However, the low abundance of NPs makes their extraction from nature inefficient, while chemical synthesis is challenging and unsustainable. *Saccharomyces cerevisiae* and *Pichia pastoris* are excellent manufacturing systems for the production of NPs. This Perspective discusses a comprehensive platform for sustainable production of NPs in the two yeasts through system-associated optimization at four levels: genetics, temporal controllers, productivity screening, and scalability. Additionally, it is pointed out critical metabolic building blocks in NP bioengineering can be identified through connecting multilevel data of the optimized system using deep learning.

More than 50% of small-molecule therapeutic agents in current clinical usage are natural products (NPs)¹. Importantly, NPs have been sources of new drugs over the past four decades², and are used as prototypes to discover new drugs because of their massive scaffold diversity and structural complexity³. Plants are excellent producers of NPs with diverse biological activities^{4,5}. Currently, the principal source of most NP-derived pharmaceuticals remains direct extraction from plants³. However, the lack of continuous availability of sufficient amounts of natural sources and variations in yields have limited NP extraction⁵. Considering the pressures of food security and population ageing⁶, alternative approaches are urgently needed to enable environmentally friendly production of NP drugs.

Thanks to the rapid progress of synthetic biology, the microbial synthesis of more complex NPs has been possible. A prime example is the successful construction of a yeast microbial cell factory for strictosidine and its downstream vinblastine, positioning yeast as a scalable system to produce more than 3000 reported NPs⁷. However, only a few natural compounds produced by microbes have reached commercial scale^{8,9}. Opportunely, the increasing sizes of multi-omics datasets and the acceleration of technologies supporting structural analysis and enzyme characterization, the time to discovery of enzymatic steps for synthesis of complex NPs in plants has decreased¹⁰. To capitalize on this opportunity, it's urgent to establish a standardized chemical engineering approach. Here, I discuss yeast as a microbial cell factory for production of complex

plant NPs with pharmaceutical importance, such as alkaloids¹¹ and flavonoids¹².

Yeast for the heterologous production of NPs

Multi-purpose yeasts, including *Saccharomyces cerevisiae* and *Pichia pastoris* (syn. *Komagataella phaffii*), are the best characterized and the most widely used microorganisms in the fermentation industry¹³. They can potentially resolve issues associated with low heterologous NP product yield observed in other expression hosts, such as *E. coli* and mammalian systems¹⁴. Yeast is able to synthesize the required precursors of diverse plant NPs including alkaloids, an imperative group of NPs containing at least one nitrogen atom¹⁵. Yeast expresses endoplasmic reticulum (ER)-localized cytochrome P450 monooxygenases (P450s, also called CYPs), membrane-anchored cytochrome P450 reductases (CPRs), scaffold Pictet-Spenglerases, and berberine-bridge enzymes, which are essential enzymes for producing NPs^{16,17}. Yeast, like higher eukaryotic organisms, is equipped with subcellular organelles, such as mitochondria, that have been employed to shield the heterologous enzymes¹⁸. Compared to mitochondria, peroxisomes are more appropriate organelles for the compartmentalization of engineered NP pathways. Peroxisomes are detoxifying organelles with greater ability than other organelles to tolerate and sequester toxic molecules such as NPs¹⁹. Moreover, they are near the ER, thus facilitating efficient cytochrome P450-mediated oxidation of the intermediates required to produce NPs such as monoterpene alkaloids²⁰.

¹Max Planck Unit for the Science of Pathogens, Charitéplatz 1, 10117 Berlin, Germany. ²Institut für Biologie, Humboldt-Universität zu Berlin, Philippstrasse 13, 10115 Berlin, Germany. ✉e-mail: naseri@mpusp.mpg.de

The *S. cerevisiae* is typically used as a host for manufacturing complex NPs^{7,21}. Recently, the non-conventional methylotrophic yeast *P. pastoris* has emerged as an attractive host for NP production²². *P. pastoris* has the advantages of methanol utilization, compatibility with higher density cell growth and large-scale fermentation. Furthermore, it shows a higher expression level of heterologous protein than *S. cerevisiae*. For example, recombinant proteins expressed under the control of the yeast alcohol oxidase 1 (*AOXI*) promoter can contribute more than 30% of the total proteins in *P. pastoris*²³. A recent study has shown that a cytochrome P450 enzyme, which catalyzes the conversion of scoulerine (a product downstream of reticuline) to cheilanthifoline, exhibits a higher conversion rate in *P. pastoris* cells (70%) than *S. cerevisiae* cells (20%)²⁴. Interestingly, when *P. pastoris* is grown on methanol, its peroxisomes substantially multiply²⁵, which can be deployed to increase the production of heterologous enzymes.

The genes coding for NPs are typically randomly distributed through the genome in the native host²⁶. To enable yeast cells to produce a chemical-of-interest (Fig. 1), each individual gene encoding one enzyme of the biosynthetic pathway or biosynthesis gene clusters needs to be expressed from the individual promoter and terminator²⁷. Recently several advanced tools such as inducible synthetic transcription factors (synTFs) and promoters^{28,29}, terminators³⁰, and genome editing tools³¹ have been developed to facilitate regulation of heterologous protein expression in *S. cerevisiae*. Fortunately, these tools have been found to be functional in tuning heterologous protein expression levels in *P. pastoris*, for which a limited number of tools and methods are available for metabolic engineering. Considering the dual application of engineering tools and methods in both chassis' organisms, *P. pastoris* has attracted attention as an ideal platform for metabolic engineering of plant NPs.

Steps to establish yeast as a sustainable platform for NP production

To increase the opportunities for carbon-neutral or carbon-negative manufacturing, the core metabolism of yeast needs to shift from its dependence on sugars to one-carbon compounds. To this end, the yeast NP production platform can be potentially equipped with two optimizable inputs: (i) one-carbon (C_1) resources, e.g., carbon dioxide (CO_2); and (ii) an optogenetic inducer (Fig. 2, Step1). The fixed C_1 can provide yeast the required carbon sources for NP production. Then, light can activate the transcription of synthetic regulators that control the expression of the NP biosynthetic pathway genes in engineered

yeast (Fig. 2, Step 2). To adjust the regulation of intermediate metabolites, many diverse synthetic transcriptional units for each individual gene, consisting of synTFs and pathway genes, should be created through combinatorial optimization methods. Next, the library of synthetic transcriptional units is engineered into the genome of host cells. Compartmentalization of the precursor biosynthetic pathways in yeast organelles allows for increased production of the required substrate and its efficient channeling to downstream enzymes for NP biosynthesis³². Screening the generated cell library with biosensors allows for high-throughput quantification of NP production in yeast cells. The fermentation conditions and light induction system must be optimized in the automated bioreactors to increase the output of the entire system (Fig. 2, Step 3). The high-throughput computational sequencing analysis of the top members of the yeast combinatorial library that are labeled with barcodes is used to feed the deep learning (DL). The obtained dataset drives biosynthetic pathways navigation for increased production of diverse natural small molecules in heterologous hosts (Fig. 2, Step 4). The following sections review current and emerging technologies to conceptually advance the individual steps of the roadmap and align them to illustrate the platform described in Fig. 2.

Input optimization

Yeast tends to offer modest emissions savings over conventional agricultural processes in the production of pharmaceutical chemicals, including less consumption of water, nutrients, energy, and land, lower costs, faster and more stable production cycles, and higher quality products³³. However, the cost and land resources required to produce sugar feedstocks make manufacturing pharmaceutical products in yeast not entirely environmentally friendly. In addition, with rising concerns regarding global climate change, a transition to clean energy sources (e.g., light and hydrogen) to achieve net zero emissions of natural gases will be necessary³⁴. In the next three subsections, I discuss strategies for carbon-negative manufacturing through enabling yeast to metabolize CO_2 or pairing yeast with autotrophic microorganisms, as well as approaches to using light to activate NP production.

Carbon-negative manufacturing in yeast

To switch to carbon-negative manufacturing, a recent study has suggested a promising metabolic engineering approach by harnessing autotrophic acetogens that uses the Wood-Ljungdahl pathway³⁵.

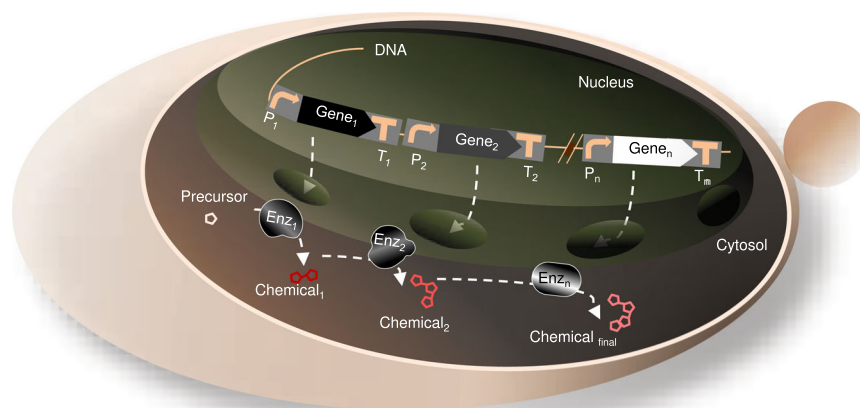


Fig. 1 | Schematic overview of the principle of heterologous chemical production in yeast. The DNA sequences of genes and regulatory elements, promoters, and terminators (to, respectively, start and terminate the transcription of the individual genes) are engineered into the DNA located inside the nucleus in the cell. Followed by translation using the host translational machinery, the expressed

enzymes convert the yeast precursor to the intermediate chemicals and, subsequently, the chemical of interest. The schematic shows a chemical biosynthetic pathway consisting of genes 'Gene₁', 'Gene₂', and 'Gene_n', requiring three promoters 'P₁', 'P₂', and 'P_n', and three terminators 'T₁', 'T₂', 'T_n'). Enz enzyme, P promoter, T terminator.

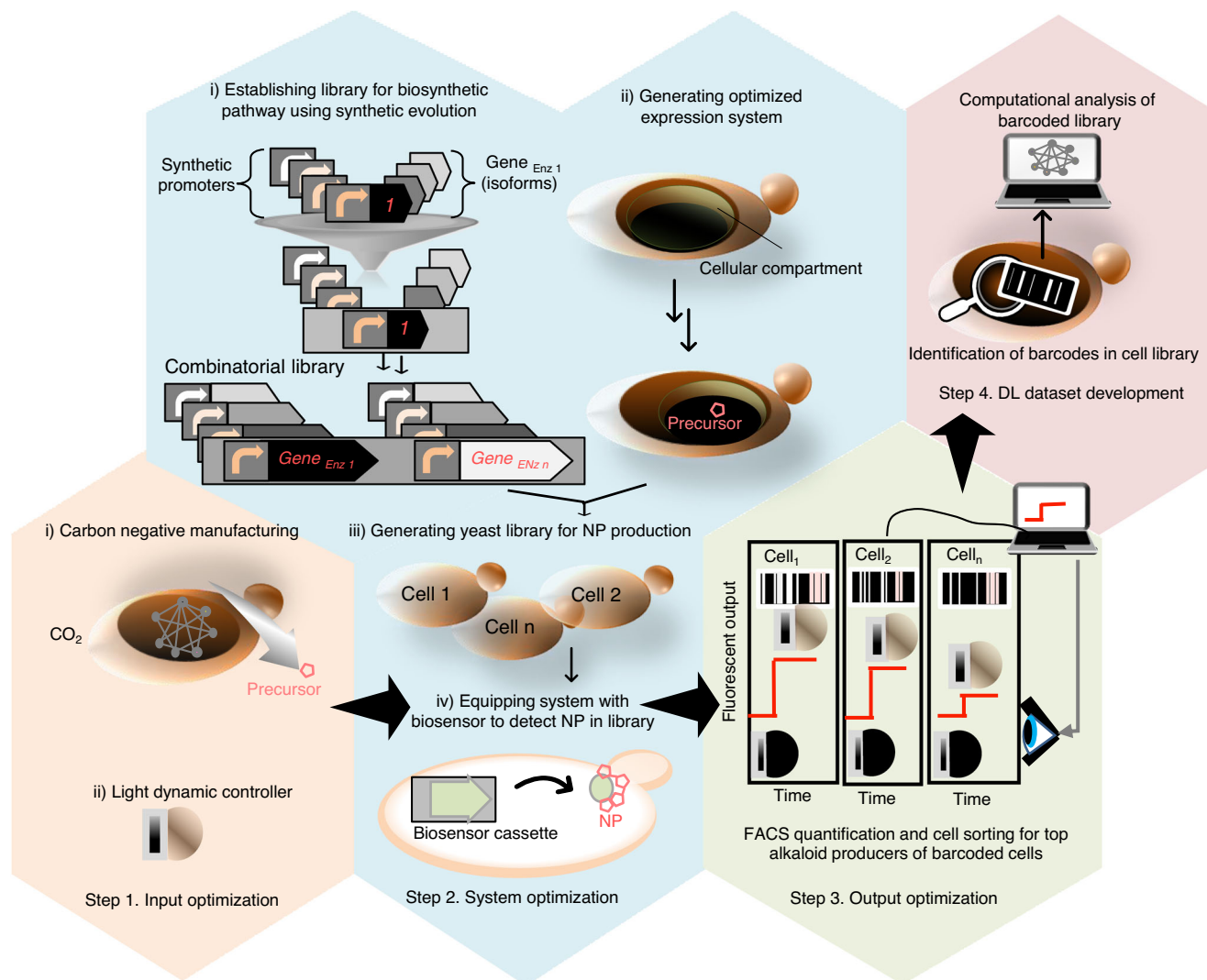


Fig. 2 | Schematic overview of the platform for sustainable production of NPs in yeast. To sustainably produce the highly required pharmaceutical NPs in an environmentally friendly manner, a four-step platform is designed. Step 1, input optimization (orange box): For carbon-negative manufacturing (i), a photosynthetic microorganism or engineered yeast fixes CO₂. The fixed carbon can provide yeast with the required sources for production NP precursors. A light-inducible system is used to dynamically control NP metabolite production in yeast (ii). Step 2, system optimization (blue box): The library of light-inducible synthetic regulators is assembled upstream of genes of enzyme isoforms by using combinatorial assembly approaches (i). Protein engineering strategies are applied to design enzyme derivatives with enhanced activity. Next, yeast cells with rewired endogenous metabolites for the enhanced production of the required precursor in peroxisomes (blue box, ii) are chromosomally engineered with the pathway library

(iii). The presence of the NP-responsive biosensor in the yeast cell enables rapid detection of NP production (iv). Step 3, output optimization (green box): Cells with a high level of fluorescence output are subjected to automated process optimization, and the top producers are sorted with FACS flow cytometry. Step 4, DL dataset development (red box): The library variants are subjected to single-cell droplet-based transcriptomics and NGS analyses to feed the DL dataset. The multilayer metadata obtained from dynamic optogenetic stimulation, the productivity of library members and growth fitness in real-time, and the genetic identity of low-to-top producers are used to fit model parameters and extract models to predict optimal conditions for heterologous production of other NPs. To simplify the figure, only three cells, Cell 1, Cell2, and Cell n, from the combinatorial library were shown. CO₂ carbon dioxide, DL deep learning, FACS fluorescence-activated cell sorting, NGS next-generation sequencing, NP natural product.

However, applying acetogenic bacteria or other natural C₁-utilizers for industrial-scale production is hindered by the limited availability of genetic manipulation tools, slow growth, and/or a narrow product spectrum³⁶. Hence, chemical engineering strategies enabling *S. cerevisiae* and *P. pastoris* to utilize CO₂, as a carbon source, rather than sugar, are increasingly being considered.

Methylotrophic *P. pastoris* can obtain all the carbon and energy needed for growth and metabolism from reduced C₁-molecule methanol³⁷. However, it leads to release of CO₂ that cannot be used by *P. pastoris*³⁷. By adding eight heterologous genes and deleting three native genes, Gassler et al. have engineered the peroxisomal methanol-assimilation pathway of *P. pastoris* into a CO₂-fixation pathway resembling the Calvin–Benson–Bassham cycle (hereafter denoted the

Calvin cycle)³⁸. The Calvin cycle is then targeted to either the cytosol or peroxisome. Next, the team have replaced the formaldehyde-assimilating dihydroxyacetone synthase enzyme of *P. pastoris* with the bacterial ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzyme. After introducing phosphoribulose kinase from spinach, the generated yeast strain directly assimilated CO₂ instead of methanol, and converted it to organic acids³⁹. Portions of the Calvin cycle have been introduced into the yeast *S. cerevisiae*⁴⁰. To enhance CO₂-fixation efficiency via the Calvin cycle in yeast, several optimization routes, for example, expressing RuBisCo derived from various origins have been evaluated⁴¹. Li et al. have increased the production of adenosine triphosphate (ATP), and relieved the nicotinamide adenine dinucleotide phosphate (NADPH) imbalance involved in the

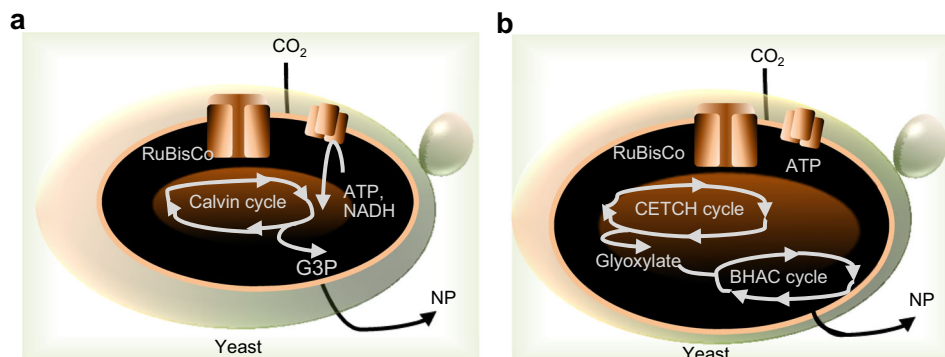


Fig. 3 | Synthetic routes for carbon-negative manufacturing to produce NP in yeast. **a** The yeast is engineered for the Calvin cycle to convert CO₂ to G3P, RuBisCo expression, and enhanced production of ATP and NADPH. **b** The yeast is engineered for the synthetic CETCH cycle to produce glyoxylate from CO₂, the β-hydroxyaspartate cycle to convert glyoxylate to NP, RuBisCo expression, and

enhanced production of ATP and NADPH. ATP, adenosine triphosphate, BHAC β-hydroxyaspartate cycle, CETCH crotonyl-coenzyme A (CoA)/ethylmalonyl-CoA/hydroxybutyryl-CoA, CO₂ carbon dioxide, G3P glyceraldehyde-3-phosphate, NADPH nicotinamide adenine dinucleotide phosphate, NP natural product.

CO₂-fixation Calvin cycle⁴². A highly active RuBisCo from the endosymbiont of the deep-sea tubeworm *Riftia pachytila*⁴³ may offer a unique application avenue for creating an efficient artificial photosynthetic pathway. Ultimately, yeast requires conversion of the Calvin cycle's final product, glyceralate 3-phosphate (G3P), to precursors for NP manufacturing (Fig. 3a). Kildegaard et al. have optimized the conversion of G3P to acetyl-CoA, required precursor for production monoterpenoid indole alkaloids (MIAs), by overexpressing native enzymes and heterologous acetyl-CoA synthase⁴⁴. Combining the above engineering efforts may enable establishment of a yeast strain that can efficiently assimilate CO₂ into acetyl-CoA-derived NPs via the Calvin cycle. As an alternative to natural CO₂ fixation, crotonyl-coenzyme A (CoA)/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) has been established for the in vitro fixation of CO₂⁴⁵. The CETCH cycle, based on crotonyl-CoA carboxylases, allows converting CO₂ into glyoxylate at a higher carbon fixation rate than the Calvin cycle. CETCH has been successfully coupled to (i) encapsulated photosynthetic membranes for light-powered CO₂ fixation to glyoxylate⁴⁶, and (ii) the β-hydroxyaspartate cycle, a glyoxylate assimilation cycle from marine bacteria, to convert CO₂ to acetyl-CoA in vitro⁴⁷. Although light-responsive CETCH cycle has not yet been implemented in vivo, it may offer a way to couple CO₂ assimilation pathways with NP-producing pathways in yeast (Fig. 3b). CETCH cycle requires two ATPs and three NADPH cofactors to convert two CO₂ molecules into organic compounds⁴⁵.

CETCH and other routes relying on C₁-assimilation require an energy source to convert thermodynamically stable CO₂. The relatively low-price methanol has been considered a promising energy source for CO₂ assimilation. To produce the required energy for CO₂ fixation, methanol is initially oxidized in the peroxisome of *P. pastoris* to hydrogen peroxide (H₂O₂), which may lead to oxidative damage to cells, and the toxic formaldehyde⁴⁸. To minimize oxidative damage, balanced expression of the peroxisomal H₂O₂-generating oxidases and catalases through combinatorial optimization is an interesting subject for investigation. To decrease the production level of formaldehyde, Gassler et al. have deleted the *AOX1* in *P. pastoris* (formaldehyde can still be synthesized by *AOX2*)³⁸. In fact, a small fraction of the formaldehyde is required to generate energy (and glyceraldehyde-3-phosphate carbon sugar)³⁸. Therefore, the balanced expression of formaldehyde synthesizing and consuming enzymes may enhance energy production in *P. pastoris* using methanol. To improve the performance of *P. pastoris* rewire for CO₂ assimilation and methanol-based energy system, Gassler et al. have enhanced growth of the autotrophic *P. pastoris* by reverse engineering and adaptive laboratory evolution^{38,49}. However, to deploy C₁'s potential for producing complex NPs in yeast, it should be stressed that a long way has to be taken

concerning an efficient energy-producing route establishment in yeast. One promising path to circumvent the problem is the electro-microbial conversion of C₁ sources into NPs using “green hydrogen” energy carriers⁵⁰. Theoretically, in a custom-designed bioreactor, yeast cells engineered to express hydrogen-oxidizing enzymes can consume hydrogen produced by solar electricity to drive CO₂ fixation. Cheng et al. have established a bacterial mixotrophic CO₂ assimilation system utilizing hydrogen, generated by hydrolysis of sodium borohydride, as an energy source⁵¹. This is an important step toward utilizing green hydrogen to support carbon-negative manufacturing in yeast. However, low solubility, poor mass transfer rate, and process safety issues of hydrogen in microbial cultures are the conceivable obstacles. Hopefully, these can possibly be addressed via a systematic design of chemical-biological interface⁵².

Artificial communities

In engineering the C₁-fixation pathway to produce NPs in yeast, we must consider that yeast needs to be engineered to express many heterologous enzymes (in addition to C₁-fixation enzymes), that may cause a metabolic burden on the yeast cell. Division of labor between cells can alleviate the metabolic burden. However, a few artificial cell-cell communication systems in yeast have been developed⁵³. Alternatively, the use of autotrophic microorganisms and yeast in synthetic microbial consortia (SMC) divides labor between cells. Recent progress in SMC designing strategies allows the production of value-added chemicals in an environmentally friendly manner. In an established SMC through co-culture of bacteria *Clostridium acetobutylicum* and *C. ljungdahlii*, CO₂ has been fixed and converted to C₂-C₄ alcohols⁵⁴. In another example, a consortium for a cooperative relationship between alginate-utilizing *Vibrio* sp. dhg and 3-hydroxypropionic acid (3-HP)-producing *Escherichia coli* strains has been established, allowing the direct conversion of alginate to 3-HP⁵⁵. For carbon-negative processing, pairwise SMC can be defined between C₁-consuming autotrophic bacteria and yeast⁵⁶. However, establishing such artificial communities fails in most cases, as one organism dominates in nutrient competition in co-cultivation condition⁵⁷. Molitor et al. have established a two-stage bioprocessing system between acetogenic bacteria *C. ljungdahlii* and *S. cerevisiae*⁵⁸. First, in a bioreactor with a pure culture of *C. ljungdahlii*, CO₂ with oxygen and nitrogen supplies (sourced by wastewater) has been converted to acetate. Next, in the other bioreactor with a pure culture of yeast, the acetate has been converted to animal feeding protein⁵⁸. In another two-stage system, CO₂ (mixed with reducing agents carbon monoxide or hydrogen) has been converted into acetic acid using acetogen *Moorella thermoacetica* in the first-stage anaerobic bioreactor⁵⁹.

Subsequently, acetic acid has been utilized for lipid production by engineered oleaginous yeast *Yarrowia lipolytica* in the second-stage aerobic bioreactor⁵⁹. Although establishing a two-stage bioprocessing system can overcome the drawbacks of co-culture schemes, the start-up of two connected bioreactors still generates high capital costs.

Alternatively, living materials combining a material scaffold with a porous structure may address the imbalanced co-culture composition issue by enabling the spatial organization of SMC species in co-cultures as that in nature (where spatial organization enables multiple species to coexist)⁶⁰. Recently a methacrylate HEMA-EDMA porous microplate living material has been designed to maximize neighboring well interactions⁶¹. This culturing system enables exchange of small molecules, but not cells, between wells⁶¹. With this porous microplate, a spatially segregated co-culture of autotrophic bacteria and yeast can potentially be established. However, extensive follow-up work must be performed on the co-culture of autotrophic bacteria and yeast⁶¹. Another technique called microbial swarmbot consortia (MSBC) has been created by encapsulating subpopulations in polymeric microcapsules⁶². The cross-linked structure of microcapsules spatially segregates the microbial cells but permits the small-size molecules and proteins to pass. In an MSBC for sustainable production of small NP drugs, photosynthetic *Synechococcus elongatus* can capture and convert inorganic CO₂ to the carbon-based organic compound G3P in the presence of blue light (Fig. 4a). In fact, blue light activates expression of chlorophyll and the key Calvin cycle enzymes. The produced G3P is then taken up by yeast to provide energy and carbon sources⁴⁴. To overcome the futile CO₂ cycling, which is due to reassimilation of CO₂ by RuBisCo⁶³, the synthetic pathway of malonyl-CoA-oxaloacetate-glyoxylate (employing the efficient phosphoenolpyruvate carboxylase to capture CO₂)⁶⁴ can be engineered into *S. elongatus* (the autotrophic partner of SMC). Alternatively, through artificial endosymbiosis (Fig. 4b), various mutants of model photosynthetic *S. elongatus* as endosymbionts have been engineered within yeast cells⁶⁵. In principle, the engineered cyanobacteria perform bioenergetic functions to support the growth of yeast cells under optimal photosynthetic growth conditions to produce NPs.

Solar-powered controller

For sustainable and reliable NP production, heterologous enzymes must be expressed after the biomass accumulation phase⁶⁶. To date, exogenous chemical inducers (e.g., L-rhamnose, IPTG, and arabinose)

have been used to control gene expression⁶⁷. However, the application of chemical inducers is limited owing to hypersensitivity and toxic effects⁶⁸. Therefore, the development of cost-effective inducers that enable modulation of protein expression levels in a fast-acting and robust manner would support stable biomanufacturing of pharmaceutical drugs⁶⁹. In the context of the proposed platform, an optogenetic system is required to regulate the expression of synTFs. A customizable synthetic DNA-binding domain (DBD), such as a transcription activator-like effector⁷⁰, to target the upstream promoter of synTF and a yeast activation domain (AD) to activate transcription are, respectively, linked with the light-using domain and its interacting partner domain (Fig. 5). The DBD binds the promoter of synTF regardless of the light condition. However, dimerization of light partner domains, allowing for activation of synTF, is light-induced (Fig. 5a). The synTF targets a synthetic promoter controlling the expression of a gene of interest. In the absence of light, the interaction of the light sensor and its partner is abolished, thus resulting in the disassociation of DBD and AD (Fig. 5b). Since blue light is required for C₁-negative manufacturing, a system based on the red light-sensitive dimerization of the photoreceptor PhyB and its interacting partner, PIF3, may hold the promise⁷¹. Benzinger and Khammash have developed a stochastic ordinary differential equation-based gene expression model to predict the output of blue light-mediated gene expression⁷². Moreover, they have demonstrated how pulsatile blue light signaling allows graded expression of multiple genes at fixed expression ratios, regardless of differences in promoter dose-response features⁷². Using the same principle, pulsatile red light-induced temporal control can be established to optimize gene expression of NP biosynthetic pathways.

System optimization

Evolution has been optimizing the expression level and the function of the enzymes to improve the structure of NPs for definite biological functions, such as cellular defence mechanisms and communication with other organisms⁵. To generate yeast cells with non-natural features, careful attention must be paid to balance the expression level of the required enzymes and to speed up their catalytic efficiency⁷³. In addition, capturing cell resources for heterologous pathway gene expression may interfere with critical cellular processes and result in energetic inefficiency in the cell⁷⁴. As a consequence, the enzyme activity and yield of the desired product are negatively influenced⁷⁴. To overcome these challenges, approaches for rewiring cellular networks

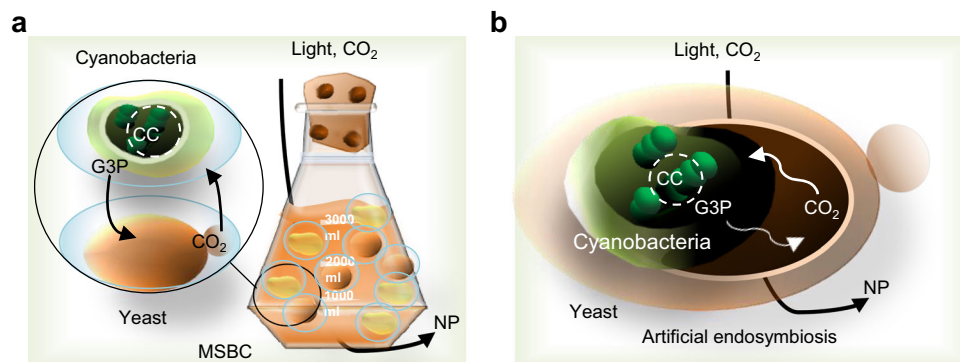


Fig. 4 | Artificial microbial communities to promote circular acetyl-CoA production in yeast. **a** An MSBC is created in polymeric microcapsules by encapsulation of subpopulations of cyanobacteria and yeast. Phototrophic cyanobacteria capture the CO₂ in the presence of light. The produced G3P is converted to the required carbon for production of metabolites, including NPs, in yeast. CO₂ produced by the yeast is reused by the autotrophic partner. **b** Artificial endosymbiosis.

Phototrophic cyanobacteria endosymbionts within yeast cells, in which the cyanobacteria perform bioenergetic functions to provide an energy source for the yeast cells to grow and produce chemicals. CC Calvin cycle, CO₂ carbon dioxide, G3P glyceraldehyde-3-phosphate, MSBC microbial swarmbot consortia, NP natural product.

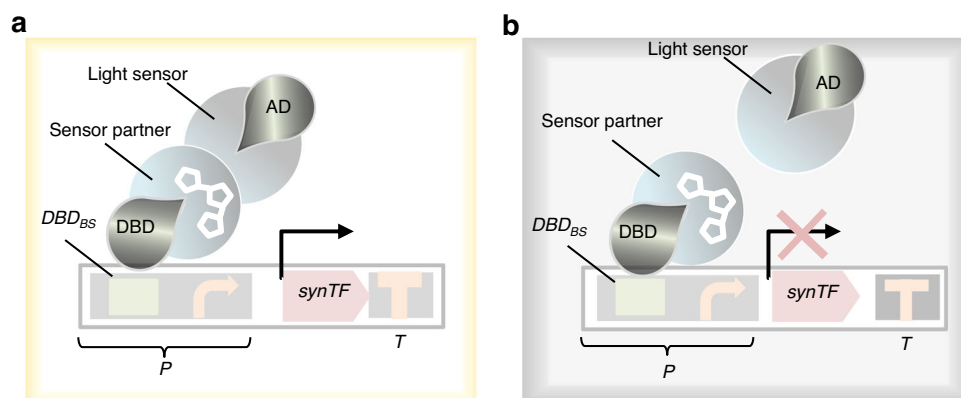


Fig. 5 | Light-dependent expression system. a Light triggers the dimerization of the light-sensing domain and its partner, thus leading to the joining of DBD and AD to form a functional synthetic activator. Next, the synthetic activator targets its BS within a promoter, thus resulting in the expression of a synTF. **b** In the absence of

light, the synthetic activator is not formed, and thus the downstream synTF is not expressed. AD activation domain, BS binding site, DBD DNA binding domain, P synthetic promoter, synTF synthetic transcription factor, T yeast terminator.

and synthetic evolution in the context of microbial production of NPs are needed.

Rewiring cellular network

A considerable hindrance in microbial bioprocess scale-up for NP production is that the accumulation of NPs within intracellular membranes decreases cell growth (and consequently the yield of the desired NP)⁷⁵. Another issue arises from the lower catalytic efficiency of plant NP enzymatic reactions in microbial cells compared to that in plants⁷³. A strategy to overcome the aforementioned challenges involves targeting membranes of cellular organelles such as mitochondria, peroxisomes, and vacuoles for compartmentalization and the storage of synthesized NP. In addition, placing the heterologous enzymes inside the cellular compartment enables enzymes such as monoterpene synthases (which have lower catalytic efficiency than the enzymes involved in the cytosolic ergosterol pathway) to access their substrates¹⁸. Increasing the ER size to cope with P450 protein folding and accumulation has yielded a concentration of 31.01 mg/L for parthenolide, a precursor for the anti-glioblastoma drug ACT001⁷⁶. In another study, to maximize the peroxisome membrane storage capacity of *S. cerevisiae*, the number and size of peroxisomes have been increased through the expression of peroxisome biogenesis-associated peroxins⁷⁷. The peroxisomal compartmentalization of acetyl-CoA, coupled with increased expression of sesquiterpenoid synthase and redirected precursor supply from the mitochondria to the cytosol, has increased production of terpene (+)-valencene 549-fold (1.2 g/L)⁷⁴. Dusseaux et al. have introduced the mevalonate pathway into the peroxisomes of *S. cerevisiae*³², which results in 125-fold greater plant terpene limonene production compared to that achieved in cytosolic production. Srinivasan and Smolke have reported an engineered yeast for increased production of hyoscyamine and scopolamine through chemical engineering of cytosol, mitochondrion, peroxisome, ER membrane, and vacuole²¹. The expression of four plant transporters, engineering cofactor regeneration mechanisms, and optimized growth conditions have resulted in an increase in de novo tropane hyoscyamine production of more than 100-fold (480 µg/L)⁷⁸. Overall, depending on the complexity of the NPs^{79,80}, different modifications should apply to the yeast strain to enhance precursor level, provide energy supply, and deliver an adequate level of NAD(P)H.

Synthetic evolution

The balance of metabolism in yeast cells has been optimized by natural evolution, wherein natural selection for desired traits creates iterative rounds of genetic diversity-selection cycles over time⁸¹. To achieve the optimal level of NP production in yeast, besides rewiring the native

global regulatory networks, the optimal expression level of heterologous enzymes must be evaluated. Therefore, many genetic diversities must be created and tested, in so-called synthetic evolution inspired by nature, which is a time-consuming task⁸². Here, I focus on combinatorial optimization methods to simultaneously assemble millions of variants, followed by high-throughput screening of optimal variants, to enable rapid synthetic evolution.

Combinatorial optimization library

Strong promoters of yeast genes are repetitively implemented for metabolic engineering purposes in *S. cerevisiae*, and *P. pastoris*, which may create a metabolic burden on the cell (thus negatively affecting growth)⁶⁶. The absent or very low transcription levels of the main biosynthetic genes for complex metabolites plays an essential role in the failure to detect compound production in native or heterologous hosts¹⁵. These issues are more severe and challenging for the long biosynthetic pathways of NPs, particularly alkaloids, which are highly toxic metabolites for the host cell¹⁹. Fast-growing collections of synTFs for *S. cerevisiae*⁶⁷ and *P. pastoris*²⁸ provide a sophisticated solution to this problem, because they enable orthogonal, tunable, and inducible expression of the target gene(s) at desired processing time. The synTFs are well suited to the platform discussed herein, especially for controlling heterologous gene expression with a light inducer^{27,83}. Using combinatorial optimization strategies, a library of synTFs can be assembled upstream of the individual genes of a biosynthetic pathway to create synthetic transcriptional units⁶⁹. Combinatorially assembled units in artificial operons are integrated into the multilocus of the yeast genome for stable metabolite production. For example, nine synTFs have been assembled with four genes (to convert farnesyl pyrophosphate (FPP) to aromatic β-ionone) in only seven combinatorial cloning reactions²⁷. The resulting 36 synthetic transcriptional unit modules have been used to generate a genomically engineered yeast library, in which various cells produce different levels of β-ionone.

In alkaloid biosynthetic pathways, master builder P450s play key roles in generating the structural variety underlying the functional diversity of produced NPs⁸⁴. Modifying substrate recognition sites of a P450 responsible for the first putative oxygenation step of the forskolin pathway has been found to increase its catalytic activity up to 6-fold in plants⁸⁵. Davies et al. have reported optimized expression ratio between P450s G8H and its redox partner increases the production of the insect-repellent nepetalactone⁸⁶. It has been shown that combined synthetic biology and protein engineering allows to identify a suitable cytochrome P450 to decorate the terpene scaffolds leading to produce 28 different noncanonical terpenes⁸⁷. Concerning several ambiguities about obtaining catalytically active heterologous enzymes

(naturally occurring in plants), yeast must be engineered with multiple gene isoforms to identify the optimal enzyme variant. Including different isoforms of various enzymes in the context of combinatorial optimization strategies enables automated optimization at the levels of transcription and translation, thus opening new routes to enhance NP production in yeast cell factories. As a starting point, Zhang et al. have optimized the production of nepetalactol in *S. cerevisiae* using screening a combinatorial library of four isoforms for iridoid synthases and cyclases⁷.

Generalized strategies for biosensor development

Detecting enzymes that can improve yield is a crucial challenge in the microbial production of chemicals⁶⁹. The problem is more pronounced when therapeutic NPs, the end product of long-biosynthetic pathways, are participants in combinatorial optimization techniques⁸⁸. Moreover, stable scale-up of manufacturing processes faces a bottleneck due to reliance on low-throughput analytical methods for assessing strain performance⁸⁹. The identification of genetically encoded transcription factor (TF) biosensors may provide alternatives to natural selection through fast screening of top-performing cells. Thus genetically encoded biosensors may potentially become part of a general platform for sustainable production of NPs. d'Oelsnitz et al. have developed a generalized approach to identify TF biosensors to detect plant NPs that are based on the TetR-family of multidrug-resistant TFs from *Salmonella typhimurium*, RamR, and its derivatives⁸⁹. RamR derivative biosensor libraries containing more than 10⁵ members have been filtered into several high-performing variants in less than one week, and used to detect several benzylisoquinoline alkaloids (BIAs)⁸⁹. TF-derived biosensors were subsequently used to translate the BIAs production level into a fluorescence signal. In addition, the TetR-family regulator CamR from *Pseudomonas putida* has been established to detect the monoterpene NPs⁹⁰. These data altogether highlight the generalized abilities of the approach to detect various alkaloid-based NPs. This strategy can also be used to screen the intermediate metabolites produced by various isoforms of, for example, P450s and their partners within a combinatorial library to identify the most efficient enzyme isoforms. Another general method to generate biosensors for user-defined molecules has been developed by Beltran et al.⁹¹. They have used the plant abscisic acid receptor PYRL, with a malleable ligand-binding pocket, to evolve 21 sensors for a range of small molecules, including structurally diverse natural and synthetic cannabinoids⁹¹. Using the systematic evolution of ligands by exponential enrichment (SELEX) strategy, we are potentially able to screen a library of synthetic riboswitches for one with high binding affinity to a NP of interest⁶⁹. Jang et al. have reported RNA aptamers for naringenin using SELEX⁹². In this study, the high-affinity naringenin-responsive riboswitches have been identified through applying in vivo selection route to the library by supplementing naringenin at different concentrations during enrichments. With the established NP biosensors, yeast cells producing high levels of NP within the combinatorial library can be detected and sorted with a flow cytometer for further optimization at bioreactor scale.

Output optimization

Bioprocess engineering aims to address the poor productivity of biomanufacturing systems arising from the manual handling of samples, dynamically changing growth environments, external sources of noise, and a lack of control over the conditions experienced by cells before and during growth⁹³. Cybergenetics involves advanced bioprocess engineering approaches that combine engineering and biology, and support the control of engineered metabolic pathways through computer interfaces and feedback controlling biological processes in real time. In this context, Chi.Bio⁹⁴ and eVOLVER⁹⁵ are two optogenetic-ready, open-source bioreactors that have been developed and made commercially available. The possibility of continuous cell culturing in

automated bioreactors enables optimizing growth conditions for the top NP producers of the combinatorial library. ReacSight, a complementary strategy to enhance bioreactor arrays for automated measurements, provides a flexible instrument control architecture⁹⁶. The combined platform of ReacSight and Chi.Bio has been used to demonstrate the real-time optogenetic control of gene expression and to achieve dynamic control of the composition of a two-strain consortium⁹⁶. Similar strategies can be applied to evaluate the target metabolite production in yeast using the light-inducible synTFs and to establish optimal microbial consortia between yeasts and bacteria for carbon-negative manufacturing. In addition, the NP production output of the system can be quantified by real-time measuring the biosensor-derived fluorescent output using flow cytometry linked to an automated bioreactor. Finally, the protocol extracted from the in situ optimization of NP-producing yeast cells in bulk culture would enable unprecedented strategies for bioreactor operation.

Dive into deep-learning data

Previous sections have discussed a platform to simplify the complex tasks of maximizing the synthesis of different NPs in microbial cells. To develop rational engineering strategies for producing rare NPs in the microbial cell, understanding the sophisticated feedback regulations controlled by the intermediate compounds is essential⁹⁷. The multi-layer optimization platform discussed herein may potentially generate a large dataset for developing DL algorithms with direct applications in biomanufacturing. Next-generation sequencing (NGS) approaches have revolutionized the identification of DNA diversity⁹⁸. NGS can unlock new ways to tackle the DNA diversity in not only top producers but also subpopulations of low- to non-producer cells within the (combinatorial) library. Nevertheless, it alone is not appropriate for identifying the combination of DNA sequence of synTFs and genes encoding enzymatic isoforms within a combinatorial library of yeast population. The high similarity of DNA sequences encoding various synTFs (identical activation domains, minimal promoters, and terminators are typically used to generate synTFs) makes it challenging to infer the relative combination of DNAs of synTF and enzymes during assembly analysis of a (sub) population⁶⁹. The droplet-based single-cell RNA sequencing allows flag-tagging of individual cells within the population⁹⁹. The detailed description of droplet-based single-cell RNA sequencing to evaluate a yeast NP-producing library is illustrated in Box 1. However, a drawback to single-cell RNA sequencing is the missing values or dropouts that often occur, mainly because of the low amount of starting material (and consequently failed amplification of the original RNAs)¹⁰⁰. Nanopore sequencing produces reads as long as 10 kb, thus resolving the sequencing issue for the repetitive region of synTFs used to modulate the production of NPs. Analytic tools that consider the characteristics of long-read data are mandatory¹⁰¹, but the development of such novel tools might take time. Large-scale single-cell genomics and transcriptomic datasets can be potentially generated using RNA sequencing for both undesirable and desirable populations. Together with induction, production, and bioprocessing optimization data, these data can feed DL algorithms to fit mathematical models, including model predictive control algorithms^{102,103}. Consequently, upon tight control of the experimental conditions and properly processing the generated data, the behavior of the biological system (e.g., dynamics of the gene expression) can be quantitatively described, and its functionality can be enhanced. Ultimately, the mechanisms that trigger loading problems are pinpointed and emergence of non-producer mutants is prevented in scaled-up fermentation.

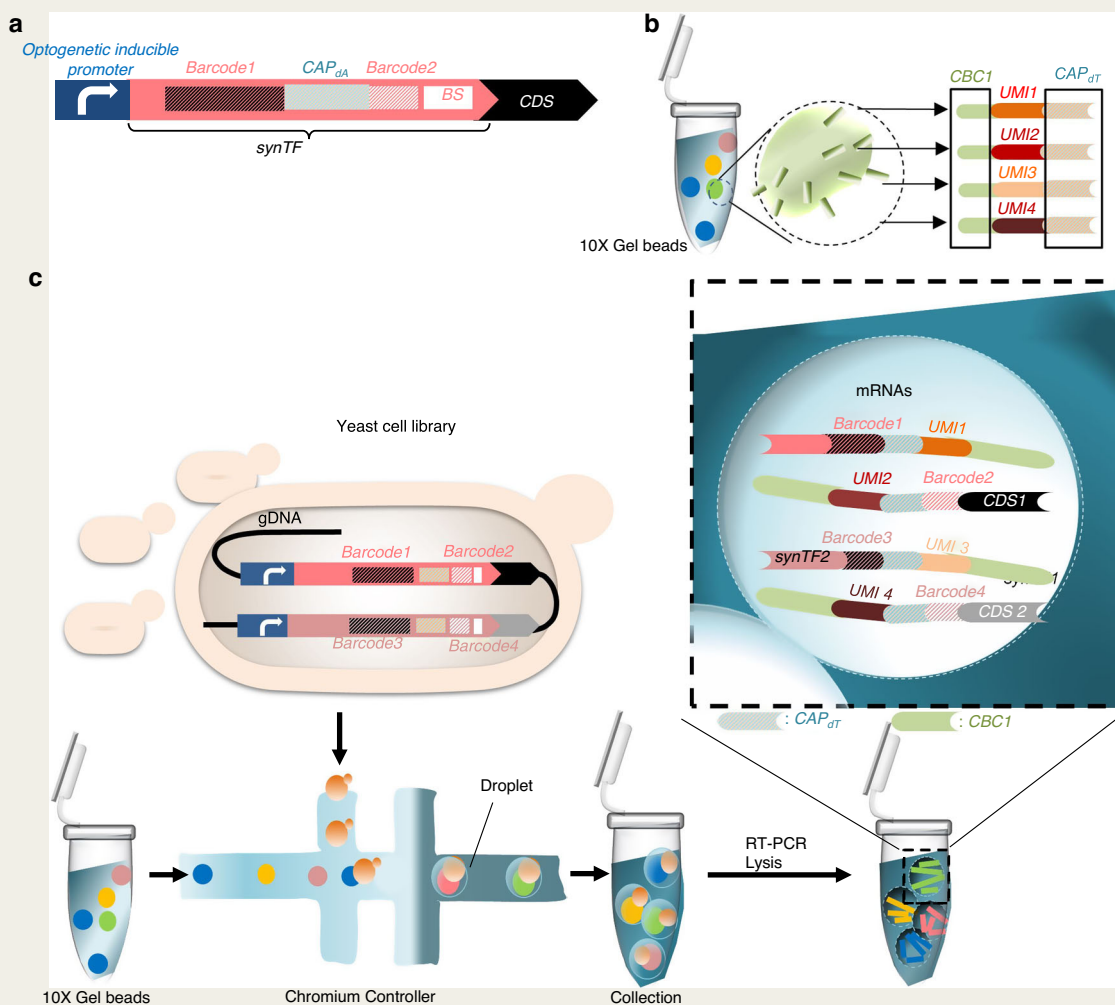
How DL algorithms may specifically contribute to extracting maximum information from an extensive combinatorial library of NP-producing variants warrants an in-depth review. To call synthetic biologists and bioinformaticians to action in this regard, promising DL programs that are applicable to substantially improving NP production

BOX 1

Schematic of gDNA capture during RNA-seq of a combinatorial library

Within the combinatorial libraries, each synthetic DNA cassette encoding *synTF*, and pathway genes in a single cell can conceivably be labeled with unique barcodes. **a**, “Barcode 1” is added to the terminator region of the *synTF*-encoding sequence, whereas a fragment containing artificially added “CAP_{dT}” and “Barcode 2” is inserted upstream of the *synTF* binding site. The 10x Genomics gel beads consist of a pool of primer sequences that are used to separately index each cell’s transcriptome. The presence of CAP_{dT} in all primer sequences enables RT of the cellular mRNAs within a library. All primers on a single gel bead contain identical CBC sequences to identify all transcriptomes of a single cell. Each primer has a unique UMI to tag each individual cDNA fragment. **b**, The Chromium Controller from 10x Genomics enables generation of tens of thousands of labeled single-cell partitions in emulsion droplets within minutes. Each droplet contains an identifying barcoded gel bead. **c**, In each droplet, the gel beads and cells are

dissolved. All fragments from the same cell share a standard 10x barcode. In the given example, “Barcode 1” and “Barcode 2” enable identifying the combination of *synTF* 1 and *CDS* 1; “Barcode 3” and “Barcode 4” enable identifying the combination of *synTF* 2 and *CDS* 2. UMIs 1, 2, 3, and 4 enable molecular counting of the captured mRNA molecules *synTF* 1, *CDS* 1, *synTF* 2, and *CDS* 2, respectively. After reverse transcription, the oil is removed, and all tagged cDNA reads are pooled to be sequenced (alongside whole single-cell transcriptomics) with NGS. The turnkey bioinformatics tools map sequencing reads back by using the identifying barcodes. Abbreviations: CAP_{dT}, capture poly dT; CBC, cell barcode capture; cDNA, coding DNA; CDS, coding DNA sequences; gDNA, genomic DNA; NGS, next-generation sequencing; RT, reverse transcription; *synTF*, synthetic transcription factor; UMI, unique molecular identifier.



are suggested here. First, Enformer can learn to predict promoter-enhancer interactions directly from the DNA sequence¹⁰⁴, thus allowing the design of more robust *synTF*s for bioengineering heterologous pathways. Second, an array of machine learning methods and amino acid property descriptors are applicable to protein redesign problems, essential tasks in NP biomanufacturing¹⁰⁵. Third, the PRISM 4 platform

can predict chemical structure¹⁰⁶. Through combinatorial optimization of NP pathway engineering in yeast, new derivatives of NPs might be produced for new drug discovery. However, the isolation of novel NP molecules is hindered by the challenge of connecting sequence information to the chemical structures of the molecules. PRISM 4 enables the development of machine-learning methods to predict the

likely biological activity of new encoded molecules. In addition to new DL algorithms, powerful computational hardware would be essential to improve the quality and reproducibility of small molecule studies. An important subject that must be addressed is the lack of accessibility of such data and their maintenance. More than 120 natural product databases and collections have been published since 2000¹⁰⁷. However, many of them are not automated facilities and do not provide open access to researchers. Generating the dataset based on the FAIR data principle (Findable, Accessible, Interoperable, and Reusable)¹⁰⁸, followed by uploading the dataset to third-party facilities such as CloudLabs (<https://cloudlabs.us/>) or biofoundries (<https://agilebiofoundry.org/>, <http://syncti.org/synthetic-biology-foundry/>), can accelerate the design-build-test-learn cycle of NP metabolic engineering in microbial cell factories¹⁰⁹.

Concluding remarks

We will face the consequences of the largest population explosion in human history in this century¹¹⁰, leading to demand for increasing cropland that competes with the need for land to produce plant-made pharmaceuticals. In a step toward sustainable manufacturing, a desirable strategy is to switch to low-carbon or carbon-neutral manufacturing. Some progresses have been made in converting C₁ to value-added chemicals at the industrial scale by acetogens. For example, a recent study has suggested a promising approach for carbon-negative manufacturing of isopropanol chemicals at a large scale by harnessing autotrophic acetogens through metabolic engineering¹¹¹. However, the conversion of C₁ gases to complex NPs of plant, like MIAs, BIAs, by engineering acetogens has still been hindered mainly by the limited availability of genetic manipulation tools, compared to a variety of advanced tools available for yeast³⁶. The slow growth and low productivity under autotrophic conditions are other barriers preventing the commercialization of acetogens on the industrial scale¹¹².

To face possible difficulties to achieve sustainable NP production, a generalized microbial biomanufacturing platform using a wide range of state-of-the-art tools and techniques, as discussed herein, is necessary. To practically use the tools and methods in various laboratories, we must consider the requirement for standardization efforts¹¹³. Encouraging chemical engineering laboratories to apply similar principles in designing their toolboxes and methods, for example, according to the BioBrick design rules, and delivering the parts to the BioBrick repository may be a large part of the solution¹¹⁴. To rapidly iterate and improve the NP production protocol, the discussed platform (Fig. 2) was equipped with lab-scale bioreactors for in situ optimization of the output of NP-producing yeast in bulk culture. The generated optimized bioprocessing protocols significantly leverage developing large-scale production strategies. In scale-up bioreactors, the applicability of optogenetic induction systems and the possibility of the mixed microbial co-culture are two specific features for consideration. Open infrastructures such as UNLOCK bioreactor platform¹¹⁵ may offer promising avenues toward scale-up production of the NPs using CO₂ and optogenetic inducer. The flexible arrangement of different types of bioreactors of UNLOCK platform allows for high-throughput cultivation of mixed cultures in MSBC and further optimization of the upstream and downstream bioprocessing protocols. Notably, safely storing the multilayer metadata produced according to the FAIR guiding principles¹⁰⁸ is of particular relevance for simultaneously ingesting, exploring, and analyzing using DL approaches.

Before transitioning NP production technology to the commercial production level, an essential consideration is developing effective regulatory compliance, besides establishing the scale-up protocol. One of the highest commercial potentials is to focus on carbon-negative biomanufacturing strategies, as they may offer promising global warming solutions while concurrently providing further economic benefits from cheap C₁ feedstocks by producing high-value

NPs¹¹⁶. Recent progresses, like establishing anaerobic biofoundries, leads to the fast growth of this path¹¹⁷. Nevertheless, analytic studies on the quality, stability, and safety of C₁-derived NPs, and the potential effect of by-products on the efficiency of processing during fermentation¹¹⁸ should be seriously considered. Furthermore, to develop cost-effective, large-scale biomanufacturing, deliberating on the process cost and technical and economic performances of the entire supply chain is necessary. For example, 30–50% of the manufacturing costs of supplies are typically allocated to downstream processing¹¹⁸. In addition, to successfully establish the discussed platform, communication issues between the disciplines must be addressed. Therefore, organizing annual interdisciplinary conferences focusing on sustainable NP production and workshops teaching various aspects of the involved disciplines are of critical importance. Providing more funding opportunities to support interdisciplinary projects can be an impetus for collaboration between different disciplines. For example, the willingness of biopharmaceutical companies to potentially commercialize C₁-derived products opens new doors to defray some of the research costs to continue the investigation on improving the potential of the C₁ technology. As a breakthrough step toward facilitating the funding of interdisciplinary fields, ‘Executive Order 14081’ on advancing biotechnology and biomanufacturing was recently approved in the USA, which aims to encourage federal investment in biotechnology and biomanufacturing. Hopefully, more countries will also take into consideration to further support research on developing the interdisciplinary field of sustainable biomanufacturing of small-molecule drugs.

Several tools and approaches to establish the discussed platform, e.g., synTFs, combinatorial optimization techniques, light inducers, bioprocessing optimization strategies, and single-cell sequencing technologies, can be implemented individually to optimize production of NP drugs in yeast. They can also be adapted to other bioengineering projects to facilitate the production of high nutritional value proteins, enzymes, and vitamins. Overall, the proposed biomanufacturing platform’s roadmap provides a unique niche for academia-industry collaboration.

References

1. Patridge, E., Gareiss, P., Kinch, M. S. & Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* **21**, 204–207 (2016).
2. Newman, D. J. & Cragg, G. M. Natural Products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
3. Atanasov, A. G., Zotchev, S. B., Dirsch, V. M. & Supuran, C. T. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov.* **20**, 200–216 (2021).
4. Pollier, J., Moses, T. & Goossens, A. Combinatorial biosynthesis in plants: a (p)review on its potential and future exploitation. *Nat. Prod. Rep.* **28**, 1897–1916 (2011).
5. Chong, K. W. et al. Biosynthetic Enantiodivergence in the Eburnane Alkaloids from *Kopsia*. *J. Nat. Prod.* **80**, 3014–3024 (2017).
6. Mobasher, A. et al. Recent advances in understanding the phenotypes of osteoarthritis. *F1000Res* **8**, 2091 (2019).
7. Zhang, J. et al. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* **609**, 341–347 (2022). **The authors introduced a very long biosynthetic pathway to *Saccharomyces cerevisiae*, including 34 heterologous genes from plants, as well as deletions, knock-downs, and overexpression of ten yeast genes, to produce vindoline.**
8. Meadows, A. L. et al. Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* **537**, 694–697 (2016).
9. Zhang, C., Zou, R., Chen, X., Stephanopoulos, G. & Too, H. P. Experimental design-aided systematic pathway optimization of glucose uptake and deoxyxylulose phosphate pathway for

- improved amorphadiene production. *Appl. Microbiol. Biotechnol.* **99**, 3825–3837 (2015).
10. Winter, J. M., Behnken, S. & Hertweck, C. Genomics-inspired discovery of natural products. *Curr. Opin. Chem. Biol.* **15**, 22–31 (2011).
11. Daley, S. K. & Cordell, G. A. Alkaloids in contemporary drug discovery to meet global disease needs. *Molecules* **26**, 3800 (2021).
12. Tanaka, Y., Sasaki, N. & Ohmiya, A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* **54**, 733–749 (2008).
13. Kastberg, L. L. B., Ard, R., Jensen, M. K. & Workman, C. T. Burden imposed by heterologous protein production in two major industrial yeast cell factories: identifying sources and mitigation strategies. *Front. Fungal Biol.* **3**, 827704 (2022).
14. Mattanovich, D., Sauer, M. & Gasser, B. Yeast biotechnology: teaching the old dog new tricks. *Micro. Cell Fact.* **13**, 34 (2014).
15. Brown, S., Clastre, M., Courdavault, V. & O'Connor, S. E. *De novo* production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl Acad. Sci. USA* **112**, e3205–e3210 (2015).
16. Roddan, R., Ward, J. M., Keep, N. H. & Hailes, H. C. Pictet-Spenglerases in alkaloid biosynthesis: future applications in biocatalysis. *Curr. Opin. Chem. Biol.* **55**, 69–76 (2020).
17. Galanie, S. & Smolke, C. D. Optimization of yeast-based production of medicinal protoberberine alkaloids. *Micro. Cell Fact.* **14**, 144 (2015).
18. Yee, D. A. et al. Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metab. Eng.* **55**, 76–84 (2019).
19. Beaudoin, G. A. & Facchini, P. J. Benzylisoquinoline alkaloid biosynthesis in opium poppy. *Planta* **240**, 19–32 (2014).
20. Knoblauch, B. et al. An ER-peroxisome tether exerts peroxisome population control in yeast. *EMBO J.* **32**, 2439–2453 (2013).
21. Srinivasan, P. & Smolke, C. D. Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* **585**, 614–619 (2020).
22. Bernauer, L., Radkohl, A., Lehmayr, L. G. K. & Emmerstorfer-Augustin, A. *Komagataella phaffii* as emerging model organism in fundamental research. *Front. Microbiol.* **11**, 607028 (2020).
23. Gao, J. et al. Biosynthesis of catharanthine in engineered *Pichia pastoris*. *Nat. Synth.* **2**, 231–242 (2023). **Here, the authors engineered the methylotrophic yeast *Pichia pastoris* for complete biosynthesis of catharanthine through the selecting stable integration sites, engineering of enzymes with high activity and/or specificity, and increasing production of the key enzymes.**
24. Uruï, M. et al. Establishment of a co-culture system using *Escherichia coli* and *Pichia pastoris* (*Komagataella phaffii*) for valuable alkaloid production. *Micro. Cell Fact.* **20**, 200 (2021).
25. Yurimoto, H., Oku, M. & Sakai, Y. Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. *Int. J. Microbiol.* **2011**, 101298 (2011).
26. Winter, J. M. & Tang, Y. Synthetic biological approaches to natural product biosynthesis. *Curr. Opin. Biotechnol.* **23**, 736–743 (2012).
27. Naseri, G., Behrend, J., Rieper, L. & Mueller-Roeber, B. COMPASS for rapid combinatorial optimization of biochemical pathways based on artificial transcription factors. *Nat. Commun.* **10**, 2615 (2019).
28. Naseri, G., Prause, K., Hamdo, H. H. & Arenz, C. Artificial transcription factors for tuneable gene expression in *Pichia pastoris*. *Front. Bioeng. Biotechnol.* **9**, 676900 (2021).
29. Sanford, A., Kiriakov, S. & Khalil, A. S. A toolkit for precise, multi-gene control in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **11**, 3912–3920 (2022).
30. Ito, Y. et al. Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. *Nucleic Acids Res.* **48**, 13000–13012 (2020).
31. Gao, J., Jiang, L. & Lian, J. Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products. *Synth. Syst. Biotechnol.* **6**, 110–119 (2021).
32. Dusseaux, S., Wajn, W. T., Liu, Y., Ignea, C. & Kampranis, S. C. Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl Acad. Sci. USA* **117**, e31789–e31799 (2020).
33. Freestone, T. S. & Zhao, H. Combinatorial pathway engineering for optimized production of the anti-malarial FR900098. *Biotechnol. Bioeng.* **113**, 384–392 (2016). **The authors developed a combinatorial optimization method that allowed, by searching only 3% of the possible samples in the anti-malarial FR900098 pathway library, a top-producing strain was identified.**
34. Woodley, J. M. Towards the sustainable production of bulk-chemicals using biotechnology. *Nat. Biotechnol.* **59**, 59–64 (2020).
35. Fast, A. G. & Papoutsakis, E. T. Stoichiometric and energetic analyses of non-photosynthetic CO₂-fixation pathways to support synthetic biology strategies for production of fuels and chemicals. *Curr. Opin. Chem. Eng.* **1**, 380–395 (2012).
36. Bennett, R. K., Steinberg, L. M., Chen, W. & Papoutsakis, E. T. Engineering the bioconversion of methane and methanol to fuels and chemicals in native and synthetic methylotrophs. *Curr. Opin. Biotechnol.* **50**, 81–93 (2018).
37. Couderc, R. & Baratti, J. Immobilized yeast cells with methanol oxidase activity: preparation and enzymatic properties. *Biotechnol. Bioeng.* **22**, 1155–1173 (1980).
38. Gassler, T. et al. The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO₂. *Nat. Biotechnol.* **38**, 210–216 (2020). **The authors established autotrophic *Pichia pastoris* to utilize CO₂ as the only carbon source to grow.**
39. Baumschabl, M. et al. Conversion of CO₂ into organic acids by engineered autotrophic yeast. *Proc. Natl Acad. Sci. USA* **119**, e2211827119 (2022).
40. Papapetridis, I. et al. Optimizing anaerobic growth rate and fermentation kinetics in *Saccharomyces cerevisiae* strains expressing Calvin-cycle enzymes for improved ethanol yield. *Biotechnol. Biofuels* **11**, 17 (2018).
41. Gong, F., Cai, Z. & Li, Y. Synthetic biology for CO₂ fixation. *Sci. China Life Sci.* **59**, 1106–1114 (2016).
42. Li, Y. J. et al. Engineered yeast with a CO₂-fixation pathway to improve the bio-ethanol production from xylose-mixed sugars. *Sci. Rep.* **7**, 43875 (2017).
43. Yang, F., Zhang, J., Cai, Z., Zhou, J. & Li, Y. Exploring the oxygenase function of Form II RuBisCo for production of glycolate from CO₂. *AMB Express* **11**, 65 (2021).
44. Kildegaard, K. R. et al. Engineering and systems-level analysis of *Saccharomyces cerevisiae* for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway. *Micro. Cell Fact.* **15**, 53 (2016).
45. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S. & Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900–904 (2016).
46. Miller, T. E. et al. Light-powered CO₂ fixation in a chloroplast mimic with natural and synthetic parts. *Science* **368**, 649–654 (2020).
47. Sundaram, S. et al. A modular in vitro platform for the production of terpenes and polyketides from CO₂. *Angew. Chem. Int. Ed. Engl.* **60**, 16420–16425 (2021).
48. Cai, P. et al. Methanol biotransformation toward high-level production of fatty acid derivatives by engineering the industrial yeast *Pichia pastoris*. *Proc. Natl Acad. Sci. USA* **119**, e2201711119 (2022).
49. Gassler, T., Baumschabl, M., Sallaberger, J., Egermeier, M. & Mattanovich, D. Adaptive laboratory evolution and reverse

- engineering enhances autotrophic growth in *Pichia pastoris*. *Met. Eng.* **69**, 112–121 (2022). **The authors enhanced the growth of the autotrophic *Pichia pastoris* by reverse engineering and adaptive laboratory evolution, an essential step toward the efficient conversion of CO₂ to value-added products.**
50. Shlosberg, Y., Schuster, G. & Adir, N. Harnessing photosynthesis to produce electricity using cyanobacteria, green algae, seaweeds and plants. *Front. Plant Sci.* **13**, 955843 (2022).
51. Cheng, Y. C. et al. Conversion of *Escherichia coli* into mixotrophic CO₂ assimilation with malate and hydrogen based on recombinant expression of 2-oxoglutarate:ferredoxin oxidoreductase using adaptive laboratory evolution. *Microorganisms* **11**, 253 (2023).
52. Zhang, P. et al. Chem-bio interface design for rapid conversion of CO₂ to bioplastics in an integrated system. *Chem* **8**, 3363–3381 (2022).
53. Khakhar, A., Bolten, N. J., Nemhauser, J. & Klavins, E. Cell-cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors. *ACS Synth. Biol.* **5**, 279–286 (2016).
54. Charubin, K. & Papoutsakis, E. T. Direct cell-to-cell exchange of matter in a synthetic *Clostridium syntrophy* enables CO₂ fixation, superior metabolite yields, and an expanded metabolic space. *Metab. Eng.* **52**, 9–19 (2019).
55. Kang, C. W. et al. Circuit-guided population acclimation of a synthetic microbial consortium for improved biochemical production. *Nat. Commun.* **13**, 6506 (2022).
56. Sudheer, S., Bai, R. G., Usmani, Z. & Sharma, M. Insights on engineered microbes in sustainable agriculture: biotechnological developments and future prospects. *Curr. Genomics* **21**, 321–333 (2020).
57. Johns, N. I., Blazejewski, T., Gomes, A. L. & Wang, H. H. Principles for designing synthetic microbial communities. *Curr. Opin. Microbiol.* **31**, 146–153 (2016).
58. Molitor, B., Mishra, A. & Angenent, L. T. Power-to-protein: converting renewable electric power and carbon dioxide into single cell protein with a two-stage bioprocess. *Energy Environ. Sci.* **12**, 3515–3521 (2019).
59. Hu, P. et al. Integrated bioprocess for conversion of gaseous substrates to liquids. *Proc. Natl Acad. Sci. USA* **113**, e3773–e3778 (2016).
60. Nguyen, P. Q. et al. Wearable materials with embedded synthetic biology sensors for biomolecule detection. *Nat. Biotechnol.* **39**, 1366–1374 (2021).
61. Vaiana, C. A. et al. Characterizing chemical signaling between engineered “microbial sentinels” in porous microplates. *Mol. Syst. Biol.* **18**, e10785 (2022).
62. Wang, L. et al. Engineering consortia by polymeric microbial swarmbots. *Nat. Commun.* **13**, 3879 (2022).
63. Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. Design and analysis of synthetic carbon fixation pathways. *Proc. Natl Acad. Sci. USA* **107**, e8889–e8894 (2010).
64. Durall, C., Lindberg, P., Yu, J. & Lindblad, P. Increased ethylene production by overexpressing phosphoenolpyruvate carboxylase in the cyanobacterium *Synechocystis* PCC 6803. *Biotechnol. Biofuels* **13**, 16 (2020).
65. Cournoyer, J. E. et al. Engineering artificial photosynthetic life-forms through endosymbiosis. *Nat. Commun.* **13**, 2254 (2022).
66. Borodina, I. & Nielsen, J. Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. *Biotechnol. J.* **9**, 609–620 (2014).
67. Naseri, G. et al. Plant-derived transcription factors for orthologous regulation of gene expression in the yeast *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **6**, 1742–1756 (2017).
68. Dvorak, P. et al. Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Micro. Cell Fact.* **14**, 201 (2015).
69. Naseri, G. & Koffas, M. A. G. Application of combinatorial optimization strategies in synthetic biology. *Nat. Commun.* **11**, 2446 (2020).
70. Machens, F., Balazadeh, S., Mueller-Roeber, B. & Messerschmidt, K. Synthetic promoters and transcription factors for heterologous protein expression in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* **5**, 63 (2017).
71. Hochrein, L., Machens, F., Messerschmidt, K. & Mueller-Roeber, B. PhiReX: a programmable and red light-regulated protein expression switch for yeast. *Nucleic Acids Res.* **45**, 9193–9205 (2017).
72. Benzinger, D. & Khammash, M. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nat. Commun.* **9**, 3521 (2018). **In this article, the authors generated a collection of optogenetic promoters in *Saccharomyces cerevisiae* and demonstrated graded multi-gene regulation at fixed expression ratios using pulsatile signaling, regardless of the differences in promoter dose-response characteristics.**
73. Montano Lopez, J., Duran, L. & Avalos, J. L. Physiological limitations and opportunities in microbial metabolic engineering. *Nat. Rev. Microbiol.* **20**, 35–48 (2022).
74. Cao, C. et al. Global metabolic rewiring of yeast enables overproduction of sesquiterpene (+)-valencene. *J. Agric Food Chem.* **70**, 7180–7187 (2022).
75. Huang, M., Bao, J., Hallstrom, B. M., Petranovic, D. & Nielsen, J. Efficient protein production by yeast requires global tuning of metabolism. *Nat. Commun.* **8**, 1131 (2017).
76. Shi, Y. et al. Production of plant sesquiterpene lactone parthenolide in the yeast cell factory. *ACS Synth. Biol.* **11**, 2473–2483 (2022).
77. Choi, K. R., Ahn, Y.-J. & Lee, S. Y. Bacterial conversion of CO₂ to organic compounds. *J. CO₂ Util.* **58**, 101929 (2022).
78. Srinivasan, P. & Smolke, C. D. Engineering cellular metabolite transport for biosynthesis of computationally predicted tropane alkaloid derivatives in yeast. *Proc. Natl Acad. Sci. USA* **118**, e2104460118 (2021).
79. Liu, J., Zhang, W., Du, G., Chen, J. & Zhou, J. Overproduction of geraniol by enhanced precursor supply in *Saccharomyces cerevisiae*. *J. Biotechnol.* **168**, 446–451 (2013).
80. Scalcinati, G. et al. Combined metabolic engineering of precursor and co-factor supply to increase alpha-santalene production by *Saccharomyces cerevisiae*. *Micro. Cell Fact.* **11**, 117 (2012).
81. Campbell, R. & Robert, J. S. The structure of evolution by natural selection. *Biol. Philos.* **20**, 673–696 (2005).
82. Simon, A. J., d’Oelsnitz, S. & Ellington, A. D. Synthetic evolution. *Nat. Biotechnol.* **37**, 730–743 (2019).
83. Hérisson, J. et al. The automated Galaxy-SynBioCAD pipeline for synthetic biology design and engineering. *Nat. Commun.* **13**, 5082 (2022). **The authors introduced the Galaxy-SynBioCAD portal that has a set of synthetic biology tools for easily metabolic engineering designs.**
84. Zanger, U. M. & Schwab, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharm. Ther.* **138**, 103–141 (2013).
85. Forman, V., Bjerg-Jensen, N., Dyekjaer, J. D., Moller, B. L. & Pateraki, I. Engineering of CYP76AH15 can improve activity and specificity towards forskolin biosynthesis in yeast. *Micro. Cell Fact.* **17**, 181 (2018).
86. Davies, M. E., Tsyplenkov, D. & Martin, V. J. J. Engineering yeast for de-novo synthesis of the insect repellent - nepetalactone. *ACS Synth. Biol.* **10**, 2896–2903 (2021).

87. Ignea, C. et al. Expanding the terpene biosynthetic code with non-canonical 16 carbon atom building blocks. *Nat. Commun.* **13**, 5188 (2022).
88. Hossain, G. S., Saini, M., Miyake, R., Ling, H. & Chang, M. W. Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends Biotechnol.* **38**, 797–810 (2020).
89. d’Oelsnitz, S. et al. Using fungible biosensors to evolve improved alkaloid biosyntheses. *Nat. Chem. Biol.* **18**, 981–989 (2022).
90. d’Oelsnitz, S., Nguyen, V., Alper, H. S. & Ellington, A. D. Evolving a generalist biosensor for bicyclic monoterpenes. *ACS Synth. Biol.* **11**, 265–272 (2022).
91. Beltran, J. et al. Rapid biosensor development using plant hormone receptors as reprogrammable scaffolds. *Nat. Biotechnol.* **40**, 1855–1861 (2022).
92. Jang, S. et al. Development of artificial riboswitches for monitoring of naringenin in vivo. *ACS Synth. Biol.* **6**, 2077–2085 (2017).
93. Castillo-Hair, S. M. et al. FlowCal: a user-friendly, open source software tool for automatically converting flow cytometry data from arbitrary to calibrated units. *ACS Synth. Biol.* **5**, 774–780 (2016).
94. Steel, H., Habgood, R., Kelly, C. L. & Papachristodoulou, A. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLoS Biol.* **18**, e3000794 (2020).
95. Wong, B. G., Mancuso, C. P., Kiriakov, S., Bashor, C. J. & Khalil, A. S. Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nat. Biotechnol.* **36**, 614–623 (2018).
96. Bertaux, F. et al. Enhancing bioreactor arrays for automated measurements and reactive control with ReacSight. *Nat. Commun.* **13**, 3363 (2022). **The authors established the generic and flexible ReacSight strategy to enhance multi-bioreactor setups for automated measurements and reactive experiment control.**
97. Zhou, Q., Ning, S. & Luo, Y. Coordinated regulation for nature products discovery and overproduction in *Streptomyces*. *Synth. Syst. Biotechnol.* **5**, 49–58 (2020).
98. Bratovic, M. et al. Bridge helix arginines play a critical role in Cas9 sensitivity to mismatches. *Nat. Chem. Biol.* **16**, 587–595 (2020).
99. Jariani, A. et al. A new protocol for single-cell RNA-seq reveals stochastic gene expression during lag phase in budding yeast. *eLife* **9**, e55320 (2020).
100. Chen, G., Ning, B. & Shi, T. Single-cell RNA-seq technologies and related computational data analysis. *Front. Genet.* **10**, 317 (2019).
101. Amarasinghe, S. L. et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **21**, 30 (2020).
102. Kumar, S. S., Indiran, T., Itty, G. V., Shettigar, J. P. & Paul, T. V. Development of a nonlinear model predictive control-based nonlinear three-mode controller for a nonlinear system. *ACS Omega* **7**, 42418–42437 (2022).
103. O’Connor, O. M., Alnahhas, R. N., Lugagne, J. B. & Dunlop, M. J. DeLTA 2.0: a deep learning pipeline for quantifying single-cell spatial and temporal dynamics. *PLoS Comput. Biol.* **18**, e1009797 (2022).
104. Avsec, Z. et al. Effective gene expression prediction from sequence by integrating long-range interactions. *Nat. Methods* **18**, 1196–1203 (2021).
105. Xu, Y. et al. Deep dive into machine learning models for protein engineering. *J. Chem. Inf. Model* **60**, 2773–2790 (2020). **The authors evaluated the performance of prediction models built using an array of machine learning methods and protein descriptor types, including single amino acid descriptors and one structure-based three-dimensional descriptor, for protein redesign and engineering.**
106. Skinnider, M. A. et al. Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. *Nat. Commun.* **11**, 6058 (2020).
107. Sorokina, M. & Steinbeck, C. Review on natural products databases: where to find data in 2020. *J. Cheminform* **12**, 20 (2020).
108. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
109. Hillson, N. et al. Building a global alliance of biofoundries. *Nat. Commun.* **10**, 2040 (2019).
110. FAO | United Nations organization. World Summit on Food Security. *Wikipedia*, (2009).
111. Liew, F. E. et al. Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat. Biotechnol.* **40**, 335–344 (2022).
112. Lee, H., Bae, J., Jin, S., Kang, S. & Cho, B. K. Engineering acetogenic bacteria for efficient one-carbon utilization. *Front. Microbiol.* **13**, 865168 (2022).
113. Malci, K. et al. Standardization of Synthetic biology tools and assembly methods for *Saccharomyces cerevisiae* and emerging yeast species. *ACS Synth. Biol.* **11**, 2527–2547 (2022).
114. Dixon, T. A., Freemont, P. S., Johnson, R. A. & Pretorius, I. S. A global forum on synthetic biology: the need for international engagement. *Nat. Commun.* **13**, 3516 (2022).
115. UNLOCK. Wageningen University & Research. <https://m-unlock.nl/> (2022).
116. Teixeira, L. V., Moutinho, L. F. & Romão-Dumaresq, A. S. Gas fermentation of C₁ feedstocks: commercialization status and future prospects. *Biofuels Bioprod. Biorefin.* **12**, 1103–1117 (2018).
117. Kopke, M. & Simpson, S. D. Pollution to products: recycling of ‘above ground’ carbon by gas fermentation. *Curr. Opin. Biotechnol.* **65**, 180–189 (2020).
118. Takors, R. et al. Using gas mixtures of CO, CO₂ and H₂ as microbial substrates: the do’s and don’ts of successful technology transfer from laboratory to production scale. *Micro. Biotechnol.* **11**, 606–625 (2018).

Acknowledgements

The author thanks the Max Planck Society for its financial support (to Marc Erhardt, Max Planck Fellow, Max Planck Unit for the Science of Pathogens), and Emmanuelle Charpentier and Marc Erhardt for their insightful discussions.

Author contributions

G.N. conceived and wrote the article.

Competing interests

The author declares no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Gita Naseri.

Peer review information *Nature Communications* thanks Pablo Carbone, Oliver Kayser and Michael Köpke for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023