Journal of Bioscience and Bioengineering
VOL. 135 No. 4, 259–265, 2023

www.elsevier.com/locate/jbiosc

REVIEW

Current advancements in systems and synthetic biology studies of *Saccharomyces cerevisiae*

Tiew-Yik Ting, YaDong Li, Hamidun Bunawan, Ahmad Bazli Ramzi, and Hoe-Han Goh*

Institute of Systems Biology, University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

Received 27 November 2022; accepted 26 January 2023
Available online 16 February 2023

***Saccharomyces cerevisiae* has a long-standing history of biotechnological applications even before the dawn of modern biotechnology. The field is undergoing accelerated advancement with the recent systems and synthetic biology approaches. In this review, we highlight the recent findings in the field with a focus on omics studies of *S. cerevisiae* to investigate its stress tolerance in different industries. The latest advancements in *S. cerevisiae* systems and synthetic biology approaches for the development of genome-scale metabolic models (GEMs) and molecular tools such as multiplex Cas9, Cas12a, Cpf1, and Csy4 genome editing tools, modular expression cassette with optimal transcription factors, promoters, and terminator libraries as well as metabolic engineering. Omics data analysis is key to the identification of exploitable native genes/proteins/pathways in *S. cerevisiae* with the optimization of heterologous pathway implementation and fermentation conditions. Through systems and synthetic biology, various heterologous compound productions that require non-native biosynthetic pathways in a cell factory have been established via different strategies of metabolic engineering integrated with machine learning.**

[**Key words:** Cell factory; Genome-scale metabolic model; Metabolic engineering; Omics; Synthetic biology]

The complete genome sequences of humans and other species have ushered the biological field into the post-genomic era with systems biology, contrasting with reductionist approaches. Systems biology is an integrated and holistic study of biological components, such as genomics, transcriptomics, proteomics, and metabolomics using computational and mathematical methods. Baker's yeast *Saccharomyces cerevisiae* is one of the earliest model eukaryotes to have its genome completely sequenced and published in 1996. Even before the dawn of modern biotechnology, *S. cerevisiae* has been extensively utilized in winemaking, baking, and other fermentation industries. This review summarizes the recent advancements in systems biology studies of *S. cerevisiae* with relevance to its industrial applications.

OMICS APPROACHES TO STUDY *S. CEREVISIAE* STRESS TOLERANCE

The history of winemaking can be dated over 7000 years ago, although it was only established in the late 1850s that *S. cerevisiae* is responsible for the fermentation of sugar into ethanol (1). *S. cerevisiae* carries out a fermentative pathway at a high sugar level to grow faster even in the presence of oxygen. This converts sugar into pyruvate via glycolysis rather than producing biomass via the tricarboxylic acid (TCA) cycle. The pyruvate is then decarboxylated into acetaldehyde and eventually reduced into ethanol. Furthermore, *S. cerevisiae* is also important as a leavening agent in the bakery industry. *S. cerevisiae* improves the bread structure and texture via increased dough volume by gas generation during sugar conversion into ethanol and carbon dioxide. However, fermentation processes for winemaking and baking are stressful for *S. cerevisiae* as they introduce high osmotic pressure, low pH, low O_2, high sugar, ethanol toxicity, and oxidative stress. These stresses can hinder the yeast fermentation ability and damage the yeast cells, which lowers industrial productivity. Some of the findings from recent omics studies to identify the targets and mechanisms of various stress tolerance in *S. cerevisiae* (2) are depicted in Fig. 1.

S. cerevisiae is used in lignocellulosic biomass fermentation for biofuel production. One major challenge is the acetic acid product from the pre-treatment of lignocellulosic biomass, which is an inhibitor of fermentation. Therefore, acetic acid tolerance is desired for efficient biofuel production. Genome analysis has identified genes responsible for acetic acid tolerance, namely *ASG1*, *ADH3*, *SKS1*, and *GIS4* (3). Adding a second copy of *HAA1* or the overexpression of $HAA1^{S135F}$ improved the acetic acid tolerance in CEN.PK113-7D strain (4). Furthermore, the overexpression of *PRS3* and *HAA1* related to cell wall integrity in xylose-consuming chassis enhanced acetic acid tolerance, indicating the importance of a robust cell wall (5). Another study identified a new strain of *S. cerevisiae* mutant (AFb.01) with enhanced growth and fermentation ability under acetic acid and furfural stresses due to the upregulation of thioredoxin (TRX1), stress response proteins (Hsp26p, Fmp16p), and increased metabolite levels of trehalose, fatty acids, gamma-aminobutyric acid (GABA), and putrescine (6). A comparative proteomics study of two natural *S. cerevisiae* isolates (J11 and YI38) with intermediate and extreme tolerance towards lignocellulosic inhibitors found differentially accumulated proteins

* Corresponding author.
E-mail address: gohhh@ukm.edu.my (H.-H. Goh).

1389-1723/$ – see front matter © 2023, The Society for Biotechnology, Japan. All rights reserved.
https://doi.org/10.1016/j.jbiosc.2023.01.010

involved in detoxification and redox homeostasis to be a shared proteome response upon the exposure to lignocellulosic inhibitors (7). Proteins related to energy management were enriched in YI83 strain with extreme tolerance towards lignocellulosic inhibitor, suggesting a detoxification mechanism with enhanced co-factor supply and energy management of cellular processes. These studies provide fundamental knowledge to improve strain for lignocellulose biofuel production.

Ethanol toxicity is another stress common in the bioethanol-producing strain. Multi-omics studies have unveiled the differentially expressed proteins and shared responses between ethanol-adapted and their ancestral strains (8). A gas chromatography-mass spectrometry (GC–MS) analysis of metabolic response against ethanol stress suggests that growth inhibition is linked to the TCA cycle (9). Another study revealed that the production of fatty alcohol induces cell wall stress (10). Other stresses such as low pH, hypoxia, and thermal stress have also been studied via multi-omics approaches (11–13) as summarized in Fig. 1.

Transcriptomics studies are important to identify the genes and pathways responsible for stress response and tolerance that can be targeted for the optimization of industrial process conditions and the engineering of stress tolerance strains suitable for different applications. Transcriptomics analysis shows that the upregulation of genes related TCA cycle and oxidative phosphorylation help to improve the specific growth rate with glycerol while genes related to the pentose phosphate pathway were downregulated (14). This knowledge is used to create an *S. cerevisiae* strain with improved glycerol assimilation rate by overexpressing *HAP4* for upregulating TCA cycle genes, *STL1* encoding glycerol/H^+ symporter for improved growth on glycerol, and disrupting *RIM15*, a kinase for transcription factor Gis1p upstream of three genes related to the pentose phosphate pathway (14). Recently, a transcriptomics study on *S. cerevisiae* strain YPH499 resulted from random mutagenesis using clustered regularly interspaced short palindromic repeats/CRISPR-associated enzyme (CRISPR/Cas) system contributes towards 2,3-butanediol resistance (15). Likewise, genome-wide association studies (GWAS) and protein–protein interaction (PPI) networks have been applied to identify novel gene targets for stress resistance engineering in the industrial strains of *S. cerevisiae* via multi-omics approaches (13). Proteome responses of *S. cerevisiae* towards different carbon sources such as glucose, galactose, maltose, and trehalose have been investigated (16). It is found that the 1887 interactions of the 1850 transcription factors and 37 chaperones in the central carbon metabolism (CCM) are the major changes in the proteomes of *S. cerevisiae*. The authors proposed that this knowledge can contribute to genome-scale metabolic (GEMs) modelling with a proteomics constraint.

SYSTEMS BIOLOGY IN *S. CEREVISIAE* STRAIN ENGINEERING

An overview of the applications of systems and synthetic biology in the bioengineering of baker's yeast for optimized industrial productions is shown in Fig. 2. Systems biology captures and provides comprehensive descriptions of cellular processes, such as the rate of transcription, translation, and protein synthesis in *S. cerevisiae* through quantitative omics analysis. In other words, systems biology can address the tedious design-build-test-learn (DBTL) cycles by providing an understanding of the genotype–phenotype relationship (17). A kinetic model which uses reaction kinetics to model time-dependent dynamics is a mathematical approach to study metabolism and rational design of *S. cerevisiae* as a cell factory. A recent review on the development for kinetics models of CCM in *S. cerevisiae* suggests that the metabolomics, fluxomics, and proteomics datasets can be used for re-fitting of parameter values or validation of the kinetic model simulations (18). For instance, optimization and risk analysis of complex living entities (ORCALE) framework that utilized metabolomics and fluxomics with metabolic control analysis (MCA) can generate kinetic models with improved predictive capabilities (19). Besides, flux balance analysis (FBA) is another popular mathematical method used to simulate the

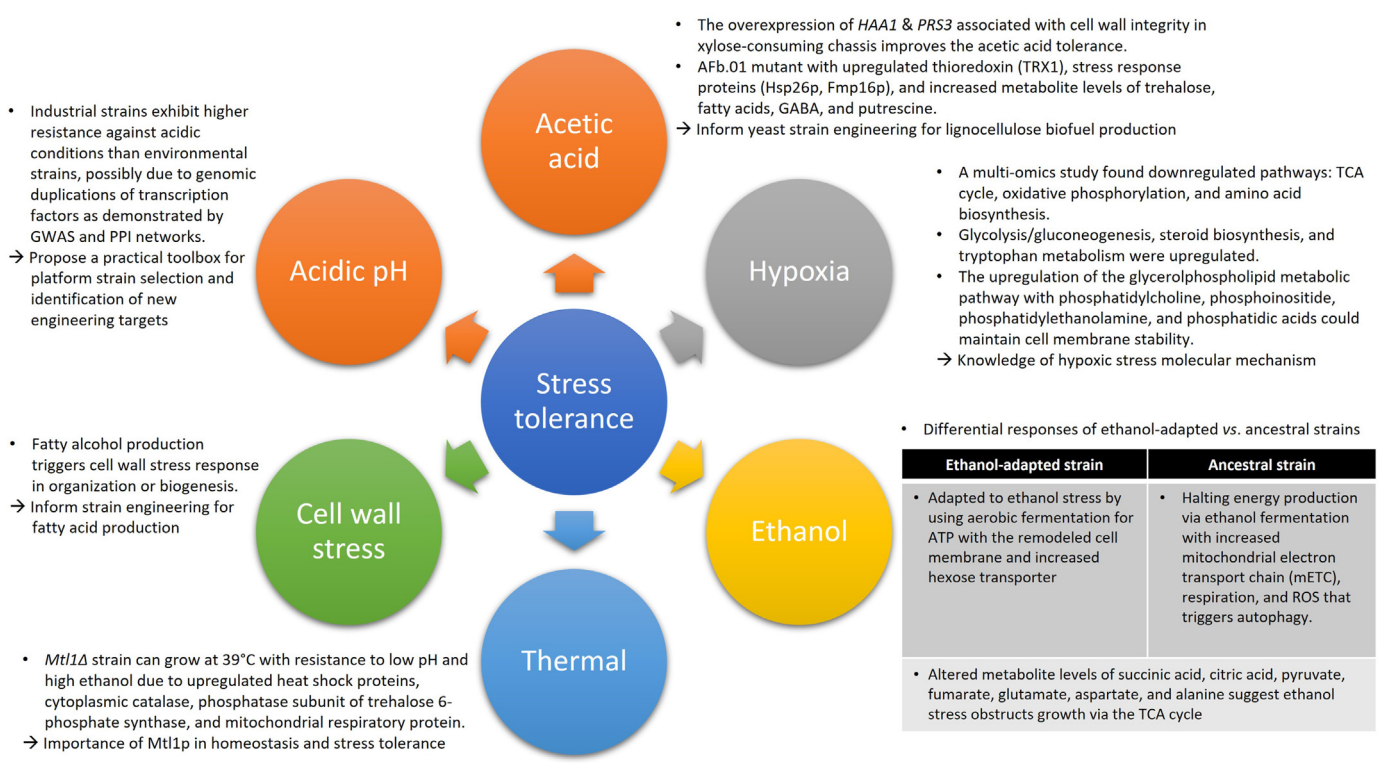

FIG. 1. Findings from selected omics studies of stress tolerance in *S. cerevisiae*.

flow of metabolites in metabolic network based on stoichiometry constraints where the sum of fluxes into a metabolite pool equals that of leaving the pool. The objective function, which is the proportion of metabolite of interest, is defined in FBA to allow the prediction of biological processes, such as the rate of substrate consumption, biotechnologically important metabolite synthesis, or the growth of *S. cerevisiae*. For instance, FBA simulation guides the systematic metabolic engineering of *S. cerevisiae* WRY2 strain by introducing ATP citrate lyase (*ACL*) gene from *Yarrowia lipolytica* and knocking out glycerol-3-phosphate (*GPD1*) gene for a higher flux of the acetyl-CoA that eventually improve the fatty acids production by ~70% (20).

The integration of the yeast genome sequence with FBA allows the reconstruction of GEMs. The first GEM of yeast was constructed in 2003 and several updates have been completed over the years (21). The latest GEMs are Yeast 8 and its iterations ecYeast8 with integrated enzyme constraints and proYeast8^{DB} with integrated protein 3D structures. Furthermore, panYeast8 (with pangenome annotation integration) and strain-specific GEMs (ssGEMs) based on the gene presence matrix of 1011 strains lead to the creation of coreYeast8 based on shared reactions, metabolites, and genes of different strains. Like the ecYeast8 with an enzyme-constraint component, GEM with Enzymatic Constraints using Kinetic and Omics data (GECKO) or ecYeast7, whole-cell model (WM_S288C), and Expression and Thermodynamics Flux for *S. cerevisiae* (yETFL) have been developed over the years. The latest models are proteome-constrained (pcYeast), proteome-constrained genome-scale protein secretory model (pcSecYeast), ecYeastGEM of Yeast8 with the integration of mitochondrial component (ecMitoYeast), and Yeast8 with the integration of enzyme cofactors (Cofactor-Yeast). Details can refer to a recent review of *S. cerevisiae* GEMs and references therein (21).

GEMs serve as a knowledge basis for all known biochemical conversion potentials of *S. cerevisiae.* It allows quantitative predictions of metabolism and *in silico* simulation of phenotype-genotype translation of *S. cerevisiae*. GEMs with FBA can be useful for the design of yeast strains through rational engineering, mutagenesis, directed evolution, and strain screening. A recent study demonstrated the usefulness of the GEMs to enhance heme production 70-fold by using Yeast 8 and ecYeast8 for predicting gene targets for CRISPR/Cas9 genome editing in *S. cerevisiae* (22). GEMs and FBA were also used to enhance the yield of L-phenylacetylcarbinol (L-PAC) production in *S. cerevisiae* by predicting novel genes deletion targets (*rpe1*, *pda1*, *adh1*, *zwf1,* and *pdc1*) for increased L-PAC production. A mutant strain with *zwf1* knockout yielded 2.48 g/L L-PAC from 2 g/L benzaldehyde, which is close to 88% of the theoretical yield (23).

MOLECULAR APPROACHES FOR THE MANIPULATION OF *S. CEREVISIAE*

Emergent of the revolutionary molecular tool, CRISPR system has quickly become the gold standard of genetic manipulation in *S. cerevisiae.* The first study that adopted CRISPR for genome editing in *S. cerevisiae* can be traced back to the year 2013, which showed that co-transformation of CRISPR/Cas9 and guide RNA (gRNA) plasmids can deliver almost 100% donor DNA recombination frequency at targeted *CAN1* locus, which is five times higher than the *delitto perfetto* method (24). More recently, CRISPR/Cas9 is used for the genetic engineering of a yeast strain to reduce urea production in commercial wine strains by the knockout of the *CAN1* gene that encodes for arginine permease (25). Furthermore, the genome shuffling method using CRISPR/Cas has managed to produce a thermotolerant strain that grows well under 39 °C (26). CRISPR/Cas9 system is also used in (CRISPR)-PCR-mediated chromosomal deletion (PCD) (CRISPR-PCD) and PCR-mediated chromosomal replacement (CRISPR-PCRep) technologies to delete two

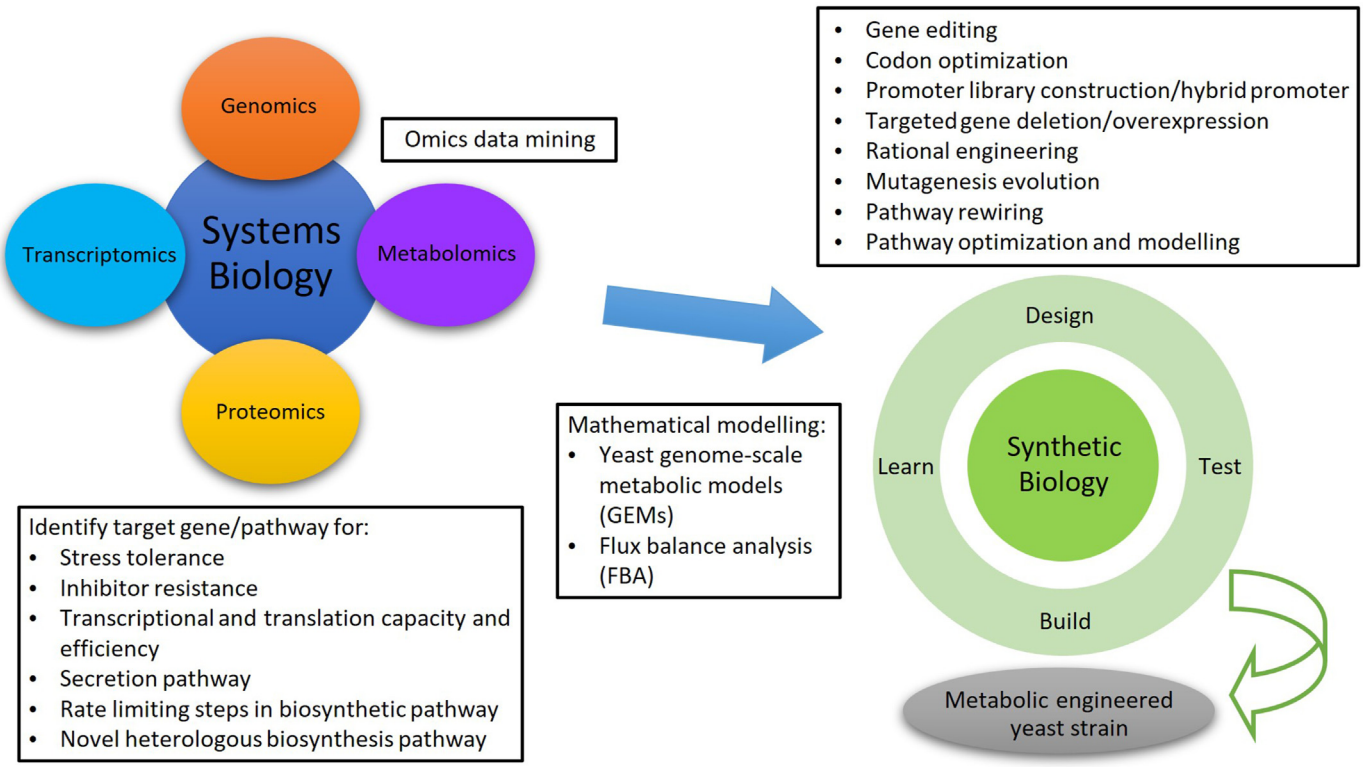

FIG. 2. An overview of systems and synthetic biology applications in the metabolic engineering of yeast strains for industrial productions.

chromosomal regions and replace three chromosomal regions in a single transformation, with efficiency from 83% to 100% (27).

Multiplex genome editing using CRISPR/Cas9 system accelerates strain engineering of *S. cerevisiae*. Highly efficient (80–100% efficiency) one-step multiplex CRISPR/Cas9 knock out of three genes (*ADH2*, *GPD1*, and *ALD4*) enhanced ethanol production ability by 1.41-fold (28). Furthermore, a reprogramming endonuclease-deficient Cas9 (dCas9) protein has been fused with the transcriptional activation domain of VP64, p65, and Rta (dCas9-VPR) for more efficient multiplex gene regulation and configurable up- or downregulation by changing the target position of the dCas9-activator at the gene promoter or open reading frame (ORF) (29). This study also demonstrated the use of a single dCas9-VPR with Pol II promoter ribozyme-sgRNA-ribozyme expression cassette for the expression of multiple sgRNAs in activating or repressing multiple genes (29). Apart from CRISPR/Cas random mutagenesis for 2,3-butanediol resistance (15) as mentioned above, Morita et al. (30) used dCas9 system to improve the production of 2,3-butanediol by suppressing *PDC1* promoter in ethanol synthetic pathway. Another study demonstrated that bacterial endoribonuclease Csy4 can produce multiple gRNAs from a single transcript by cleaving mRNA with gRNA sequences flanked with Csy4 recognition motifs for multiplex genome editing or CRISPR interference (CRISPRi) in *S. cerevisiae* (31).

Cas12a, also known as Cpf1 is a new class II/type V CRISPR-Cas system used for genome editing in *S. cerevisiae*. One major advantage of CRISPR/Cas12a is its endoribonuclease domain with pre-crRNA processing activity that auto-process own CRISPR RNA (crRNA), unlike Cas9 which requires a trans-activating CRISPR RNA (tracrRNA). This allows the expression of multiple gRNAs from one simple crRNA array for CRISPR/Cas12a multiplex genome editing, in contrast to CRISPR/Cas9 which requires complex or multiple gRNA expression constructs. A comparison of Cpf1 orthologs *As*Cpf1, *Lb*Cpf1, and *Fn*Cpf1 shows that *Lb*Cfp1 and *Fn*Cpf1 have comparable efficiencies as CRISPR/Cas9 (32). Multiplex genome editing using CRISPR/*Fn*Cpf1 (*Francisella novicida*) on four loci of the *S. cerevisiae* genome was accomplished with 100% efficiency (33). Later, Li et al. (34) developed self-cloning CRISPR/Cas12a (scCRISPR/Cpf1) for singleplex (80% efficiency) and tripleplex (32% efficiency) that enable genomic integration of β-carotene pathway and strain engineering of *S. cerevisiae* for patchoulol production without cloning steps.

One of the challenges in CRISPR-based genome editing of *S. cerevisiae* is due to the lack of specific protospacer adjacent motif (PAM) in the target sequence: *Sa*Cas9 (PAM = NNGRRT), *Sp*Cas9 (PAM = NGG), and *As*Cas12a (PAM = TTTV). Therefore, the joint use of these CRISPR/Cas systems will increase the availability of PAM in the target sequence and improve the range of editable targets in the *S. cerevisiae* genome. Recently, Okada et al. (35) have designed a series of CRISPR-*Sp*Cas9/*Sa*Cas9/*As*Cas12a backbone vectors with an URA3 marker and GAL1 promoter that is compatible with Golden Gate Assembly. These backbones which are available at NBRP Yeast Resource Center can shorten the genome-editing process of *S. cerevisiae* to 2 weeks.

The promoter is one of the most essential regulatory components in gene expression. Native promoters that exist in the *S. cerevisiae* genome have been identified and characterized for their promoter strengths. This library of the promoter is useful for optimizing the expression of the desired gene for industrial purposes (36). The effect of promoter on xylose-utilizing strain was investigated by expressing the lacZ reporter gene with 29 different promoters in different culture conditions (aerobic/microaerobic in xylose or glucose-containing media) (37). A more recent study has selected 30 out of 66 native *S. cerevisiae* promoters for (*2S*)-naringenin biosynthesis pathway optimization and managed to improve the titer to 1.21 g/L, the highest record to date (38). Promoter engineering of the *RIM15* gene (encodes for protein kinase RIM15p that phosphorylates GIS1p transcription factor to activate genes in glycerol and acetate metabolism, and genes for the survival during quiescent stage) by inserting promoter of gluconeogenic gene (*PCK1*) to 5′-untranslated regions of *RIM15* successful repressed *RIM15* gene in the early glucose-rich stage and induce *RIM15* gene in the late stage of alcohol fermentation and outperformed *RIM15*-deleted strain under severe alcohol fermentation process (39). Another important regulatory element for gene expression is the terminators. In 2017, Wei et al. (40) identified and characterized the strength of 100 native *S. cerevisiae* terminators (45 strong, 31 moderate, and 24 weak) and showed that strong terminators increase mRNA level and protein expression, although the effect is reduced if paired with strong promoters.

Furthermore, the development of the toolkits for gene expression in *S. cerevisiae* also contributes to advancement in the metabolic engineering of *S. cerevisiae*. For instance, the modular gene expression toolkit for *S. cerevisiae* with a set of activator/repressor transcription factors (TFs) with TF-dependent and native promoters can control heterologous expression and allow switching between two alternative metabolic branches (41). Moreover, a CRISPR/Cas9-based toolkit for *S. cerevisiae* gene expression which includes sgRNA plasmids, promoters with different strength and protein tags, together with a Cas-Designer web tool for primer design successfully improved the expression of taxadiene synthase by 25-fold (42). Besides, many computational methods and tools have been published recently to help the engineering of *S. cerevisiae* strains (Table 1) (43–49).

PATHWAY ENGINEERING STRATEGIES

S. cerevisiae has great potential for heterologous productions that require non-native biosynthesis pathways, such as plant phenylpropanoids (50,51) and terpenoids (52). Multi-omics studies identified novel genes and biosynthetic pathways for metabolic engineering or synthetic biology modifications of *S. cerevisiae*. DBTL cycles are applied to redesign and optimize the heterologous biosynthesis pathway in *S. cerevisiae* by introducing efficient biosynthetic enzymes (52,53). This section will discuss recent strategies used in metabolic pathway engineering of *S. cerevisiae*, including pathway refactoring, orthogonal pathway, combinatorial pathway, cofactor and metabolic flux redirection, and subcellular engineering.

Advances in synthetic biology allow the implementation of different strategies for pathway engineering in *S. cerevisiae*. With the help of DNA assembly technologies such as Golden Gate Assembly and gene synthesis services, multigene construction for pathway refactoring is simplified. Nowadays, synthetic genes can be cloned into plasmid constructs with multiple promoters and terminators to create a series of expression cassettes. The constructs can later be assembled into a fully refactored pathway using hierarchical Golden Gate Assembly. This strategy is not only highly effective in refactoring the zeaxanthin biosynthetic pathway in *S. cerevisiae* but also has high modularity that allows gene deletion or exchange using spacer plasmids if desired (54).

The orthogonal pathway is a recent strategy used to improve limonene synthesis in *S. cerevisiae* (55). In the study, a more efficient orthogonal limonene biosynthetic pathway with neryl diphosphate synthase 1 (SINDPS1) and plant limonene synthase improves the limonene yield by 6-folds when compared to conventional limonene biosynthetic pathway. The bottleneck present in the conventional limonene biosynthetic pathway is due to heterologous limonene synthase that produces limonene using endogenous substrate geranyl diphosphate (GPP) that is also involved in the yeast sterol pathway. In orthogonal limonene biosynthetic pathway,

TABLE 1. List of example computational methods and tools for *S. cerevisiae* systems and synthetic biology.

Tool	Type	Description	Reference
Experiment data depot (EDD)	Online tool	Repository of experiment data and metadata, visualization of data, and standardization of export data. Can be used for the storage of characterized synthetic biology parts, leveraging proteomics data to improve biofuel yield, and using extracellular metabolite concentrations to predict intracellular metabolic fluxes.	43
Network-based gene enrichment analysis (NETGE-PLUS)	Webserver	Standard functional interpretation for gene sets of model organisms including *S. cerevisiae*. Functional enrichment of genes and exploration of functional relationships in KEGG pathway via KEGG-NET resource.	44
Network propagation for functions prediction (NPF)	Protein annotation method	Function prediction using network propagation complement with multi-omics data such as domain annotation and protein complex information. The comprehensive evaluation shows NPF delivers better performance than the competing method.	45
YEASTRACT	Database and search tool	YEAst Search for Transcriptional Regulators and Consensus Tracking (YEASTRACT) can be used for the analysis and prediction of transcription factors and the target genes in *S. cerevisiae*.	46
Yeast-GEM	GEM database	Database of up-to-date consensus GEMs including Yeast 8 (https://github.com/SysBioChalmers/yeast-GEM)	47
GECKO 2.0	GEM Toolbox	Updated GECKO toolbox for construction of enzyme-constraint GEMs model (ecModels) for *S. cerevisiae* and other organisms. (https://github.com/SysBioChalmers/GECKO/releases/tag/v2.0.2)	48
Cas-Designer	CRISPR/Cas9 design tool	Design possible guide RNA sequences based on DNA sequence input and their potential off-target sites in the target genome and aid in choosing suitable gene knockout sites.	49

the SINDPS1 produces neryl pyrophosphate (NPP) (*cis*-GPP) instead of GPP from isopentenyl pyrophosphate and dimethylallyl pyrophosphate (DMAPP) and the plant limonene synthase converts NPP to limonene, which in turn bypassing the bottleneck present in conventional limonene biosynthetic pathway (55).

Pathway engineering using the combinatorial design is another strategy to improve the production of *S. cerevisiae*. Combinatorial approach for pathway engineering in *S. cerevisiae* can be achieved through a combination of enzyme homologs with various characteristics or a combination of pathways with similar functions (56). The latter approach is used to improve the production of geranylgeraniol (GGOH) to nearly 9-fold, in which two different branches of GGOH biosynthetic pathways were combined to overproduce geranylgeranyl pyrophosphate (GGPP) substrate for yeast endogenous phosphatase to produce GGOH (56).

Altering cofactor availability and redirecting metabolic flux can also be used in the metabolic engineering of *S. cerevisiae*. For instance, the production of protopanaxadiol (PPD) in engineered strain is significantly improved (>11-fold) by redirecting metabolic flux to increase NADPH availability, coupled with optimized expression balance of genes in the PPD pathway with different combinations of promoters (57). Combinational use of chromosomal integration of endogenous genes (*tHMG1*, *ERG1*, *ERG20*, *ERG9*, *POS5* or *UPC2.1*), knockout of genes affecting mevalonate (MVA) pathway using CRISPR/Cas9 (*BTS1*, *ROX1*, *YPL064w*, and *YJL062w*), the introduction of high efficiency heterologous *Tw*OSC^{T502E} friedelin synthase enzyme produces GQ1 strain, which hold the highest friedelin titer record of 2.45 mg/L, 65-fold higher than the wild-type strain BY4741 under optimized culture media (58).

Another example of metabolic engineering is the study by Yamada et al. (59) that optimized the expression of 12 genes involved in glycolysis and heterologous *d-LDH* gene in *S. cerevisiae* to produce a stable high D-lactic acid production strain with high glucose metabolism and low ethanol production. Lycopene production in *S. cerevisiae* was improved to 115.64 mg/mL, 2689-folds higher than the starting strain via pathway and chassis metabolism optimization. This is achieved by increasing the pool of rate-limiting step metabolite GGPP and expanding MVA pathway under the control of a constitutive promoter and CEN.PK2−1C as a chassis with enhanced terpenoid precursor synthesis (60). Furthermore, an increased copy number of genes in (*2S*)-naringenin metabolism pathway (4CL and *CHS*) eliminated the feedback inhibition of tyrosine and downregulated competitive pathways in a high titer *p*-coumaric acid strain and yielded 149.8 mg/L of (*2S*)-naringenin (61).

Apart from pathway assembly in the cytosol, subcellular engineering is getting popular for metabolic engineering in recent years (62). Subcellular engineering is engineering of pathways in the organelles such as the mitochondria, peroxisomes, endoplasmic reticulum (ER), Golgi apparatus, vacuoles, and cell wall. For recombinant protein expression in *S. cerevisiae,* one obstacle is low yield due to low protein secretion and intracellular retention of the protein. Hence, the ER and Golgi apparatus pathway of *S. cerevisiae* is engineered to improve protein production. Deleting *VPS5* and *VPS17* that participate in endosome-to-Golgi trafficking significantly improves heterologous protein secretion and reduces intracellular retention of the protein. This modification improved protein secretion from milligrams per liter scale to 2.5 g/L of α-amylase (63).

ROLE OF MACHINE LEARNING IN *S. CEREVISIAE* METABOLIC ENGINEERING

One of the challenges in systems biology is to transform a huge amount of data obtained from omics studies into knowledge. For example, it is difficult to understand the genotype−phenotype relationship due to a poor correlation of metabolome to gene expression. Machine learning (ML) can analyze and learn the patterns from huge experiment data obtained from systems biology studies and use the pattern to guide the metabolic engineering of *S. cerevisiae*. A recent ML study accurately predicted *S. cerevisiae* metabolome phenotype and essential genes in metabolic regulation based on enzyme abundance by mapping regulatory enzyme expression patterns of 97 *S. cerevisiae* strains with kinase knockout (64). Another study extensively evaluated 27 ML techniques to combine metabolic flux data captured from GEMs created by genome-wide expression profiles and transcriptomes of 1143 *S. cerevisiae* mutant strains to generate a multimodal artificial neural network (MMANN) that can predict and characterize cell growth of another 86 strains (65). Another neural network algorithm predicted ethanol yields from *S. cerevisiae* fermentation within 30- and 60-min using yeast morphological data with a coefficient of determination >0.9 (66).

ML has been used in the metabolic engineering of *S. cerevisiae* strain for industrial applications. GEMs were used for identifying gene targets (*PCK1*, *TKL1*, *TAL1*, *CDC19*, and *PFK1*) that perturb tryptophan synthesis and search transcriptomes to find 30 promoters for a total of 7776 combinational designs (67). Two ML models adopting the automated recommendation tool (ART) or Bayesian optimization inspired EVOLVE algorithm from TeselaGen

were trained using datasets from high-throughput analysis of tryptophan accumulation using engineered tryptophan biosensor. EVOLVE allows predictive engineering of *S. cerevisiae* by predicting the outcome of metabolic engineering based on the strain design as input. Based on the ART and EVOLVE prediction, *CDC19* and *PFK1* genes were knock-downed to produce SP606 strain with improved tryptophan titer and productivity up to 74% and 43%, respectively (67). ML workflow in conjunction with the YeastFab Assembly strategy (MiYA) is another example of ML application in the DBTL cycle of yeast strain engineering. MiYA can aid in optimizing combinatorial metabolic pathways by capturing rate-limiting steps and decreasing the accumulation of undesired byproducts or intermediates (68). MiYA uses an artificial neural network (ANN) ensemble trained with a small dataset from the compilation of genotypic and phenotypic information from representative *S. cerevisiae* strains. This ANN ensemble can predict the best genotypic design for experimental validation. For instances, MiYA is used by the same study to predict a strain with empirically verified 2.42-fold higher violacein production amongst 3125 designs fed to MiYA. A recent study used a model trained using sequences of fungal and bacterial sugar transporters together with a comparative genomics approach of 182 yeast phylogenomic analysis successfully identified three new yeast xylose transporters (Spx, SpH, and SpG) (69). Wet lab analysis showed that the hexose transporter deficient EBY_Xyl1 strain with a xylose consumption pathway harboring one of these transporters outperformed the control strain in cell growth as well as the uptake and consumption of xylose (69).

Other than strain engineering, the ML approach is also adapted for predicting yield and optimization of the fermentation process. For instance, ML using an optimized decision-making system algorithm is used to improve glucose consumption and ethanol production by optimizing fermentation parameters such as shaking speed, temperature, time, as well as pH and biomass (cellulose, hemicellulose, and lignin) content and can predict yield based on parameters used with 95% accuracy (70). These studies demonstrate how ML can be employed in an industrial setting to scale up bioproduction using *S. cerevisiae*.

CONCLUSION

This review presents the systems and synthetic biology studies of *S. cerevisiae* in the context of its industrial applications. In short, omics studies on the molecular mechanisms of *S. cerevisiae* provide fundamental resources for applied metabolic engineering and translational research. The rapid and continuous development of GEMs with molecular and computational tools facilitate systematic and efficient metabolic engineering of optimal *S. cerevisiae* strains for different industrial applications. Lastly, increasing studies of ML in recent years show that ML could be the next breakthrough for accelerated metabolic engineering of *S. cerevisiae*.

ACKNOWLEDGMENTS

We acknowledge the funding provided by Universiti Kebangsaan Malaysia (DIP-2020-005) for supporting the research group.

References

1. **García-Ríos, E. and Guillamón, J. M.:** Mechanisms of yeast adaptation to wine fermentations, Prog. Mol. Subcell. Biol., 37–59 (2019).
2. **Taymaz-Nikerel, H., Cankorur-Cetinkaya, A., and Kirdar, B.:** Genome-wide transcriptional response of *Saccharomyces cerevisiae* to stress-induced perturbations, Front. Bioeng. Biotechnol., **4**, 17 (2016).
3. **González-Ramos, D., De Vries, Grijseels, S. S., Van Berkum, Swinnen, S., Van Den Broek, M., Nevoigt, E., Daran, J.-M. G., Pronk, J. T., and Van Maris, A. J.:** A new laboratory evolution approach to select for constitutive acetic acid tolerance in *Saccharomyces cerevisiae* and identification of causal mutations, Biotechnol. Biofuels, **9**, 173 (2016).
4. **Swinnen, S., Henriques, S. F., Shrestha, R., Ho, P. W., Sá-Correia, I., and Nevoigt, E.:** Improvement of yeast tolerance to acetic acid through Haa1 transcription factor engineering: towards the underlying mechanisms, Microb. Cell Fact., **16**, 7 (2017).
5. **Cunha, J. T., Costa, C. E., Ferraz, L., Romaní, A., Johansson, B., Sá-Correia, I., and Domingues, L.:** *HAA1* and *PRS3* overexpression boosts yeast tolerance towards acetic acid improving xylose or glucose consumption: unravelling the underlying mechanisms, Appl. Microbiol. Biotechnol., **102**, 4589–4600 (2018).
6. **Unrean, P., Gätgens, J., Klein, B., Noack, S., and Champreda, V.:** Elucidating cellular mechanisms of *Saccharomyces cerevisiae* tolerant to combined lignocellulosic-derived inhibitors using high-throughput phenotyping and multiomics analyses, FEMS Yeast Res., **18**, foy106 (2018).
7. **de Witt, R. N., Kroukamp, H., and Volschenk, H.:** Proteome response of two natural strains of *Saccharomyces cerevisiae* with divergent lignocellulosic inhibitor stress tolerance, FEMS Yeast Res., **19**, foy116 (2019).
8. **Šoštarić, N., Arslan, A., Carvalho, B., Plech, M., Voordeckers, K., Verstrepen, K. J., and van Noort, V.:** Integrated multi-omics analysis of mechanisms underlying yeast ethanol tolerance, J. Proteome Res., **20**, 3840–3852 (2021).
9. **Ming, M., Wang, X., Lian, L., Zhang, H., Gao, W., Zhu, B., and Lou, D.:** Metabolic responses of *Saccharomyces cerevisiae* to ethanol stress using gas chromatography-mass spectrometry, Mol. Omics, **15**, 216–221 (2019).
10. **Dahlin, J., Holkenbrink, C., Marella, E. R., Wang, G., Liebal, U., Lieven, C., Weber, D., McCloskey, D., Ebert, B. E., Herrgård, M. J., Blank, L. M., Borodina, I., and Wang, H. L.:** Multi-omics analysis of fatty alcohol production in engineered yeasts *Saccharomyces cerevisiae* and *Yarrowia lipolytica*, Front. Genet., **10**, 747 (2019).
11. **Xia, Z., Zhou, X., Li, J., Li, L., Ma, Y., Wu, Y., Huang, Z., Li, X., Xu, P., and Xue, M.:** Multiple-omics techniques reveal the role of glycerophospholipid metabolic pathway in the response of *Saccharomyces cerevisiae* against hypoxic stress, Front. Microbiol., **10**, 1398 (2019).
12. **Martínez-Matías, N., Chorna, N., González-Crespo, S., Villanueva, L., Montes-Rodríguez, I., Melendez-Aponte, L. M., Roche-Lima, A., Carrasquillo-Carrión, K., Santiago-Cartagena, E., Rymond, B. C., Babu, M., Stagljar, I., and Rodríguez-Medina, J. R.:** Toward the discovery of biological functions associated with the mechanosensor Mtl1p of *Saccharomyces cerevisiae* via integrative multi-OMICs analysis, Sci. Rep., **11**, 7411 (2021).
13. **Kang, K., Bergdahl, B., Machado, D., Dato, L., Han, T.-L., Li, J., Villas-Boas, S., Herrgård, M. J., Förster, J., and Panagiotou, G.:** Linking genetic, metabolic, and phenotypic diversity among *Saccharomyces cerevisiae* strains using multi-omics associations, GigaScience, **8**, giz015 (2019).
14. **Kawai, K., Kanesaki, Y., Yoshikawa, H., and Hirasawa, T.:** Identification of metabolic engineering targets for improving glycerol assimilation ability of *Saccharomyces cerevisiae* based on adaptive laboratory evolution and transcriptome analysis, J. Biosci. Bioeng., **128**, 162–169 (2019).
15. **Mizobata, A., Mitsui, R., Yamada, R., Matsumoto, T., Yoshihara, S., Tokumoto, H., and Ogino, H.:** Improvement of 2,3-butanediol tolerance in *Saccharomyces cerevisiae* by using a novel mutagenesis strategy, J. Biosci. Bioeng., **131**, 283–289 (2021).
16. **Garcia-Albornoz, M., Holman, S. W., Antonisse, T., Daran-Lapujade, P., Teusink, B., Beynon, R. J., and Hubbard, S. J.:** A proteome-integrated, carbon source dependent genetic regulatory network in *Saccharomyces cerevisiae*, Mol. Omics, **16**, 59–72 (2020).
17. **Ramzi, A. B., Baharum, S. N., Bunawan, H., and Scrutton, N. S.:** Streamlining natural products biomanufacturing with omics and machine learning driven microbial engineering, Front. Bioeng. Biotechnol., **8**, 608918 (2020).
18. **Lao-Martil, D., Verhagen, K. J. A., Schmitz, J. P. J., Teusink, B., Wahl, S. A., and van Riel, N. A. W.:** Kinetic modeling of *Saccharomyces cerevisiae* central carbon metabolism: achievements, limitations, and opportunities, Metabolites, **12**, 74 (2022).
19. **Miskovic, L., Alff-Tuomala, S., Soh, K. C., Barth, D., Salusjärvi, L., Pitkänen, J.-P., Ruohonen, L., Penttilä, M., and Hatzimanikatis, V.:** A design–build–test cycle using modeling and experiments reveals interdependencies between upper glycolysis and xylose uptake in recombinant *S. cerevisiae* and improves predictive capabilities of large-scale kinetic models, Biotechnol. Biofuels, **10**, 166 (2017).
20. **Ghosh, A., Ando, D., Gin, J., Runguphan, W., Denby, C., Wang, G., Baidoo, E. E., Shymansky, C., Keasling, J. D., and García Martín, H.:** ^{13}C metabolic flux analysis for systematic metabolic engineering of *S. cerevisiae* for overproduction of fatty acids, Front. Bioeng. Biotechnol., **4**, 76 (2016).
21. **Chen, Y., Li, F., and Nielsen, J.:** Genome-scale modeling of yeast metabolism: retrospectives and perspectives, FEMS Yeast Res., **22**, foac003 (2022).
22. **Ishchuk, O. P., Domenzain, I., Sánchez, B. J., Muñiz-Paredes, F., Martínez, J. L., Nielsen, J., and Petranovic, D.:** Genome-scale modeling drives 70-fold improvement of intracellular heme production in *Saccharomyces cerevisiae*, Proc. Natl. Acad. Sci. USA, **119**, e2108245119 (2022).
23. **Iranmanesh, E., Asadollahi, M. A., and Biria, D.:** Improving L-phenylacetylcarbinol production in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering, J. Biotechnol., **308**, 27–34 (2020).

24. **DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M.:** Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems, Nucleic Acids Res., **41**, 4336–4343 (2013).
25. **Vigentini, I., Gebbia, M., Belotti, A., Foschino, R., and Roth, F. P.:** CRISPR/Cas9 system as a valuable genome editing tool for wine yeasts with application to decrease urea production, Front. Microbiol., **8**, https://doi.org/10.3389/fmicb.2017.02194 (2017).
26. **Mitsui, R., Yamada, R., and Ogino, H.:** Improved stress tolerance of *Saccharomyces cerevisiae* by CRISPR-cas-mediated genome evolution, Appl. Biochemi. Biotechnol., **189**, 810–821 (2019).
27. **Easmin, F., Sasano, Y., Kimura, S., Hassan, N., Ekino, K., Taguchi, H., and Harashima, S.:** CRISPR-PCD and CRISPR-PCRep: two novel technologies for simultaneous multiple segmental chromosomal deletion/replacement in *Saccharomyces cerevisiae*, J. Biosci. Bioeng., **129**, 129–139 (2020).
28. **Liu, K., Yuan, X., Liang, L., Fang, J., Chen, Y., He, W., and Xue, T.:** Using CRISPR/Cas9 for multiplex genome engineering to optimize the ethanol metabolic pathway in *Saccharomyces cerevisiae*, Biochem. Eng. J., **145**, 120–126 (2019).
29. **Deaner, M., Mejia, J., and Alper, H. S.:** Enabling graded and large-scale multiplex of desired genes using a dual-mode dCas9 activator in *Saccharomyces cerevisiae*, ACS Synth. Biol., **6**, 1931–1943 (2017).
30. **Morita, K., Seike, T., Ishii, J., Matsuda, F., and Shimizu, H.:** Improvement of 2,3-butanediol production by dCas9 gene expression system in *Saccharomyces cerevisiae*, J. Biosci. Bioeng., **133**, 208–212 (2022).
31. **Ferreira, R., Skrekas, C., Nielsen, J., and David, F.:** Multiplexed CRISPR/Cas9 genome editing and gene regulation using Csy4 in *Saccharomyces cerevisiae*, ACS Synth. Biol., **7**, 10–15 (2018).
32. **Verwaal, R., Buiting-Wiessenhaan, N., Dalhuijsen, S., and Roubos, J. A.:** CRISPR/Cpf1 enables fast and simple genome editing of *Saccharomyces cerevisiae*, Yeast, **35**, 201–211 (2018).
33. **Swiat, M. A., Dashko, S., den Ridder, M., Wijsman, M., van der Oost, J., Daran, J. M., and Daran-Lapujade, P.:** FnCpf1: a novel and efficient genome editing tool for *Saccharomyces cerevisiae*, Nucleic Acids Res., **45**, 12585–12598 (2017).
34. **Li, Z.-H., Wang, F.-Q., and Wei, D.-Z.:** Self-cloning CRISPR/Cpf1 facilitated genome editing in *Saccharomyces cerevisiae*, Bioresour. Bioprocess., **5**, 36 (2018).
35. **Okada, S., Doi, G., Nakagawa, S., Kusumoto, E., and Ito, T.:** Simple-to-use CRISPR-SpCas9/SaCas9/AsCas12a vector series for genome editing in *Saccharomyces cerevisiae*, G3 Genes Genomes Genet., **11**, jkab304 (2021).
36. **Yuan, T., Guo, Y., Dong, J., Li, T., Zhou, T., Sun, K., Zhang, M., Wu, Q., Xie, Z., Cai, Y., Cao, L., and Dai, J.:** Construction, characterization and application of a genome-wide promoter library in *Saccharomyces cerevisiae*, Front. Chem. Sci. Eng., **11**, 107–116 (2017).
37. **Nambu-Nishida, Y., Sakihama, Y., Ishii, J., Hasunuma, T., and Kondo, A.:** Selection of yeast *Saccharomyces cerevisiae* promoters available for xylose cultivation and fermentation, J. Biosci. Bioeng., **125**, 76–86 (2018).
38. **Gao, S., Zhou, H., Zhou, J., and Chen, J.:** Promoter-library-based pathway optimization for efficient (2*S*)-naringenin production from *p*-coumaric acid in *Saccharomyces cerevisiae*, J. Agric. Food Chem., **68**, 6884–6891 (2020).
39. **Watanabe, D., Kaneko, A., Sugimoto, Y., Ohnuki, S., Takagi, H., and Ohya, Y.:** Promoter engineering of the *Saccharomyces cerevisiae RIM15* gene for improvement of alcoholic fermentation rates under stress conditions, J. Biosci. Bioeng., **123**, 183–189 (2017).
40. **Wei, L., Wang, Z., Zhang, G., and Ye, B.:** Characterization of terminators in *Saccharomyces cerevisiae* and an exploration of factors affecting their strength, Chembiochem, **18**, 2422–2427 (2017).
41. **Rantasalo, A., Kuivanen, J., Penttilä, M., Jäntti, J., and Mojzita, D.:** Synthetic toolkit for complex genetic circuit engineering in *Saccharomyces cerevisiae*, ACS Synth. Biol., **7**, 1573–1587 (2018).
42. **Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N. J., Keasling, J. D., and Mukhopadhyay, A.:** A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*, Nucleic Acids Res., **45**, 496–508 (2017).
43. **Morrell, W. C., Birkel, G. W., Forrer, M., Lopez, T., Backman, T. W., Dussault, M., Petzold, C. J., Baidoo, E. E., Costello, Z., and D. Ando:** The experiment data depot: a web-based software tool for biological experimental data storage, sharing, and visualization, ACS Synth. Biol., **6**, 2248–2259 (2017).
44. **Bovo, S., Martelli, P. L., Di Lena, P., and Casadio, R.:** NETGE-PLUS: standard and network-based gene enrichment analysis in human and model organisms, J. Proteome Res., **19**, 2873–2878 (2020).
45. **Zhao, B., Zhang, Z., Jiang, M., Hu, S., Luo, Y., and Wang, L.:** NPF: network propagation for protein function prediction, BMC Bioinformatics, **21**, 355 (2020).
46. **Teixeira, M. C., Monteiro, P. T., Palma, M., Costa, C., Godinho, C. P., Pais, P., Cavalheiro, M., Antunes, M., Lemos, A., Pedreira, T., and Sá-Correia, I.:** YEASTRACT: an upgraded database for the analysis of transcription regulatory networks in *Saccharomyces cerevisiae*, Nucleic Acids Res., **46**, D348–D353 (2018).
47. **Lu, H., Li, F., Sánchez, B. J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P. M., Lappa, D., and Lieven, C.:** A consensus *S. cerevisiae* metabolic model Yeast 8 and its ecosystem for comprehensively probing cellular metabolism, Nat. Commun., **10**, 3586 (2019).
48. **Domenzain, I., Sánchez, B., Anton, M., Kerkhoven, E. J., Millán-Oropeza, A., Henry, C., Siewers, V., Morrissey, J. P., Sonnenschein, N., and Nielsen, J.:** Reconstruction of a catalogue of genome-scale metabolic models with enzymatic constraints using GECKO 2.0, Nat. Commun., **13**, 3766 (2022).
49. **Park, J., Bae, S., and Kim, J.-S.:** Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites, Bioinformatics, **31**, 4014–4016 (2015).
50. **Bahaudin, K. N. A. K., Sabri, S., Ramzi, A. B., Chor, A. L. T., Tencomnao, T., and Baharum, S. N.:** Current progress in production of flavonoids using systems and synthetic biology platforms, Sains Malays., **47**, 3077–3084 (2018).
51. **Shah, F. L. A., Ramzi, A. B., Baharum, S. N., Noor, N. M., Goh, H.-H., Leow, T. C., Oslan, S. N., and Sabri, S.:** Recent advancement of engineering microbial hosts for the biotechnological production of flavonoids, Mol. Biol. Rep., **46**, 6647–6659 (2019).
52. **Belcher, M. S., Mahinthakumar, J., and Keasling, J. D.:** New frontiers: harnessing pivotal advances in microbial engineering for the biosynthesis of plant-derived terpenoids, Curr. Opin. Biotechnol., **65**, 88–93 (2020).
53. **Ramzi, A. B.:** Metabolic engineering and synthetic biology, in: Aizat, W., Goh, H. H. and Baharum, S. (Eds.), Omics applications for systems biology. Advances in experimental medicine and biology, vol. 1102. Springer, Cham (2018).
54. **Ren, H., Hu, P., and Zhao, H.:** A plug-and-play pathway refactoring workflow for natural product research in *Escherichia coli* and *Saccharomyces cerevisiae*, Biotechnol. Bioeng., **114**, 1847–1854 (2017).
55. **Cheng, S., Liu, X., Jiang, G., Wu, J., Zhang, J.-l., Lei, D., Yuan, Y.-J., Qiao, J., and Zhao, G.-R.:** Orthogonal engineering of biosynthetic pathway for efficient production of limonene in *Saccharomyces cerevisiae*, ACS Synth. Biol., **8**, 968–975 (2019).
56. **Song, T.-Q., Ding, M.-Z., Zhai, F., Liu, D., Liu, H., Xiao, W.-H., and Yuan, Y.-J.:** Engineering *Saccharomyces cerevisiae* for geranylgeraniol overproduction by combinatorial design, Sci. Rep., **7**, 14991 (2017).
57. **Kim, J.-E., Jang, I.-S., Sung, B. H., Kim, S. C., and Lee, J. Y.:** Rerouting of NADPH synthetic pathways for increased protopanaxadiol production in *Saccharomyces cerevisiae*, Sci. Rep., **8**, 15820 (2018).
58. **Gao, H.-Y., Zhao, H., Hu, T.-Y., Jiang, Z.-Q., Xia, M., Zhang, Y.-F., Lu, Y., Liu, Y., Yin, Y., Chen, X.-C., Luo, Y.-F., Zhou, J.-W., Wang, J.-D., Gao, J., Gao, W., and Huang, L.-Q.:** Metabolic engineering of *Saccharomyces cerevisiae* for high-level friedelin via genetic manipulation, Front. Bioeng. Biotechnol., **10**, https://doi.org/10.3389/fbioe.2022.805429 (2022).
59. **Yamada, R., Wakita, K., Mitsui, R., and Ogino, H.:** Enhanced D-lactic acid production by recombinant *Saccharomyces cerevisiae* following optimization of the global metabolic pathway, Biotechnol. Bioeng., **114**, 2075–2084 (2017).
60. **Li, X., Wang, Z., Zhang, G., and Yi, L.:** Improving lycopene production in *Saccharomyces cerevisiae* through optimizing pathway and chassis metabolism, Chem. Eng. Sci., **193**, 364–369 (2019).
61. **Li, H., Gao, S., Zhang, S., Zeng, W., and Zhou, J.:** Effects of metabolic pathway gene copy numbers on the biosynthesis of (2*S*)-naringenin in *Saccharomyces cerevisiae*, J. Biotechnol., **325**, 119–127 (2021).
62. **Hammer, S. K. and Avalos, J. L.:** Harnessing yeast organelles for metabolic engineering, Nat. Chem. Biol., **13**, 823–832 (2017).
63. **Huang, M., Wang, G., Qin, J., Petranovic, D., and Nielsen, J.:** Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables improved protein production, Proc. Natl. Acad. Sci. USA, **115**, E11025–E11032 (2018).
64. **Zelezniak, A., Vowinckel, J., Capuano, F., Messner, C. B., Demichev, V., Polowsky, N., Mülleder, M., Kamrad, S., Klaus, B., and Keller, M. A.:** Machine learning predicts the yeast metabolome from the quantitative proteome of kinase knockouts, Cell Syst., **7**, 269–283.e266 (2018).
65. **Culley, C., Vijayakumar, S., Zampieri, G., and Angione, C.:** A mechanism-aware and multiomic machine-learning pipeline characterizes yeast cell growth, Proc. Natl. Acad. Sci. USA, **117**, 18869–18879 (2020).
66. **Itto-Nakama, K., Watanabe, S., Kondo, N., Ohnuki, S., Kikuchi, R., Nakamura, T., Ogasawara, W., Kasahara, K., and Ohya, Y.:** AI-based forecasting of ethanol fermentation using yeast morphological data, Biosci. Biotechnol. Biochem., **86**, 125–134 (2021).
67. **Zhang, J., Petersen, S. D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A., Abeliuk, E., Sánchez, B. J., Costello, Z., Chen, Y., Fero, M. J., Martin, H. G., Nielsen, J., Keasling, J. D., and Jensen, M. K.:** Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism, Nat. Commun., **11**, 4880 (2020).
68. **Zhou, Y., Li, G., Dong, J., Xing, X. H., Dai, J., and Zhang, C.:** MiYA, an efficient machine-learning workflow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*, Metab. Eng., **47**, 294–302 (2018).
69. **Fiamenghi, M. B., Bueno, J. G. R., Camargo, A. P., Borelli, G., Carazzolle, M. F., Pereira, G. A. G., dos Santos, L. V., and José, J.:** Machine learning and comparative genomics approaches for the discovery of xylose transporters in yeast, Biotechnol. Biofuels Bioprod., **15**, 57 (2022).
70. **Vinitha, N., Vasudevan, J., and Gopinath, K. P.:** Bioethanol production optimization through machine learning algorithm approach: biomass characteristics, saccharification, and fermentation conditions for enzymatic hydrolysis, Biomass Convers. Biorefin., https://doi.org/10.1007/s13399-022-03163-z (2022).