CHALMERS
UNIVERSITY OF TECHNOLOGY

Improved production of Taxol® precursors in S. cerevisiae using combinatorial in silico design and metabolic engineering

Microbial Cell Factories

RESEARCH

Open Access

Improved production of Taxol® precursors in *S. cerevisiae* using combinatorial in silico design and metabolic engineering

Koray Malcı[1,2,9*], Rodrigo Santibáñez[3], Nestor Jonguitud-Borrego[1,2], Jorge H. Santoyo-Garcia[1,2], Eduard J. Kerkhoven[4,5,6] and Leonardo Rios-Solis[1,2,7,8*]

Abstract

Background Integrated metabolic engineering approaches that combine system and synthetic biology tools enable the efficient design of microbial cell factories for synthesizing high-value products. In this study, we utilized in silico design algorithms on the yeast genome-scale model to predict genomic modifications that could enhance the production of early-step Taxol® in engineered *Saccharomyces cerevisiae* cells.

Results Using constraint-based reconstruction and analysis (COBRA) methods, we narrowed down the solution set of genomic modification candidates. We screened 17 genomic modifications, including nine gene deletions and eight gene overexpressions, through wet-lab studies to determine their impact on taxadiene production, the first metabolite in the Taxol® biosynthetic pathway. Under different cultivation conditions, most single genomic modifications resulted in increased taxadiene production. The strain named KM32, which contained four overexpressed genes (*ILV2*, *TRR1*, *ADE13*, and *ECM31*) involved in branched-chain amino acid biosynthesis, the thioredoxin system, de novo purine synthesis, and the pantothenate pathway, respectively, exhibited the best performance. KM32 achieved a 50% increase in taxadiene production, reaching 215 mg/L. Furthermore, KM32 produced the highest reported yields of taxa-4(20),11-dien-5α-ol (T5α-ol) at 43.65 mg/L and taxa-4(20),11-dien-5-α-yl acetate (T5αAc) at 26.2 mg/L among early-step Taxol® metabolites in *S. cerevisiae*.

Conclusions This study highlights the effectiveness of computational and integrated approaches in identifying promising genomic modifications that can enhance the performance of yeast cell factories. By employing in silico design algorithms and wet-lab screening, we successfully improved taxadiene production in engineered *S. cerevisiae* strains. The best-performing strain, KM32, achieved substantial increases in taxadiene as well as production of T5α-ol and T5αAc. These findings emphasize the importance of using systematic and integrated strategies to develop efficient yeast cell factories, providing potential implications for the industrial production of high-value isoprenoids like Taxol®.

Keywords Computational metabolic engineering, Genome-scale modelling, in silico design, Synthetic biology, Systems biology, Mevalonate pathway, *Saccharomyces cerevisiae*, Taxadiene, Taxol

*Correspondence:
Koray Malcı
k.malci@imperial.ac.uk
Leonardo Rios-Solis
leo.rios@ucl.ac.uk
Full list of author information is available at the end of the article

Introduction

Microbial chassis have been extensively investigated for the sustainable and economically viable production of industrially important compounds, including pharmaceuticals [1], biofuels [2], enzymes [3] and polymers [4] in sustainable and economically attractive ways. Among these microbial hosts, *Saccharomyces cerevisiae,* commonly known as baker's yeast, has emerged as the most widely studied eukaryotic synthetic biology chassis [5]. *S. cerevisiae* offers advantages such as high biomass production in cost-effective media [6] and efficient control and scalability of fermentation processes [7]. Furthermore, the development of numerous synthetic biology tools and well-characterized genetic parts have accelerated the design of engineered yeast strains with improved performance and reliability [8, 9].

S. cerevisiae has been harnessed as a platform to produce high-value biopharmaceuticals, ranging from recombinant therapeutic proteins to plant-derived natural products [10–14]. By integrating heterologous plant-derived genes, early-step precursors of Taxol® (paclitaxel), a leading anticancer drug with a market size over billions of USD [15], have been produced by yeast cell factories [16–20]. The biosynthesis of Taxol® initiates with the conversion of geranylgeranyl diphosphate (GGPP), a product of the yeast mevalonate pathway, into taxadiene (taxa-4(5),11(12)-diene) catalyzed by taxadiene synthase [21]. Subsequently, taxadiene undergoes hydroxylation by a class II cytochrome P450 hydroxylase, taxadiene-5α-hydroxylase (T5αOH) with the assistance of cytochrome P450 reductase (CPR), and an acylation catalyzed by taxadiene-5α-ol-O-acetyltransferase (TAT) [15]. Although these enzymes have been expressed, and their products have been successfully produced in yeast [22, 23], there is still room to improve their titers for economically feasible production of the subsequent Taxol® precursors. Figure 1 illustrates the biochemical reactions in the mevalonate and Taxol® pathways.

As omics technologies are evolving and sequencing platforms are becoming more accessible, the reconstruction of genome-scale metabolic models (GEMs) for various organisms has gained momentum, enabling the conversion of integrated omics data into valuable system-level information [24]. Since the first yeast GEM was reconstructed in 2003, several iterations and enhancements of *S. cerevisiae* GEMs have been developed by adding and connecting more genes, reactions, and compartments [25–27]. The latest consensus yeast GEM, yeast 8 [27], serves as the basis for constructing more refined and up-to-date versions that better mimic yeast metabolism. Among the yeast 8 derivatives, yeast 8.5.0 comprises 4055 reactions, 2742 metabolites, and 1151 genes in 14 cellular compartments [28]. Leveraging yeast GEMs as a bottom-up systems biology tool, they have been employed in diverse applications, from designing yeast cell factories to optimizing culture conditions [29]. In recent years, the coupling of GEMs and constraint-based reconstruction and analysis (COBRA) methods [30] has found widespread usage in biological applications. COBRA is an integrative analysis framework that can be applied to biochemical systems to demonstrate and predict the connections from phenotype to genotype by imposing constraints considering factors such as physicochemical laws, environmental conditions, and genetic information [31, 32].

Several strain optimization algorithms and programs have been developed utilising COBRA methods and mathematical modelling [33]. For example, OptKnock [34] employs bilevel mixed-integer optimisation to predict gene deletions and elimination of reactions from the host's metabolism, thereby enhancing the production of target compounds. Another tool, OptGene [35], utilises genetic algorithms [36] to identify gene knockouts with the aim of improving bio-production yields. Leveraging bilevel programming, Ranganathan et al. (2010) introduced the OptForce framework to identify a set of reactions where fluxes need to be modified to achieve overproduction of the target compound [37]. These modifications can involve zeroing (deletion), decreasing (downregulation), or increasing (upregulation) of the fluxes associated with the target reactions, thus exploring all possible modifications [37].

Previously, we reported the production of early-step Taxol® precursors through the improved mevalonate pathway using our engineered yeast strains [22, 23]. In the present study, we used a combinatorial in silico system biology approach to further improve the flux towards GGPP in the mevalonate pathway to enhance the titers of the taxadiene and the next precursors in the Taxol® pathway. We utilised three strain design frameworks,

(See figure on next page.)

Fig. 1 **A** The native biochemical reactions in the mevalonate pathway leading to the production of GGPP in yeast**. B** The Taxol® biosynthesis from GGPP in yew trees. HMG-CoA: hydroxymethylglutaryl-CoA, mevalonate-P: R-5-hosphomevalonate, mevalonate-PP: R-5 diphosphomevalonate, IPP: isopentenyl diphosphate, DMAPP: dimethylallyl diphosphate, GPP: geranyl diphosphate, FPP: farnesyl diphosphate, GGPP: geranylgeranyl diphosphate, T5α-ol: taxa-4(20),11-dien-5α-ol, T5αAc: taxa-4(20),11-dien-5-α-yl acetate. The dashed arrows represent naturally occurring multiple steps

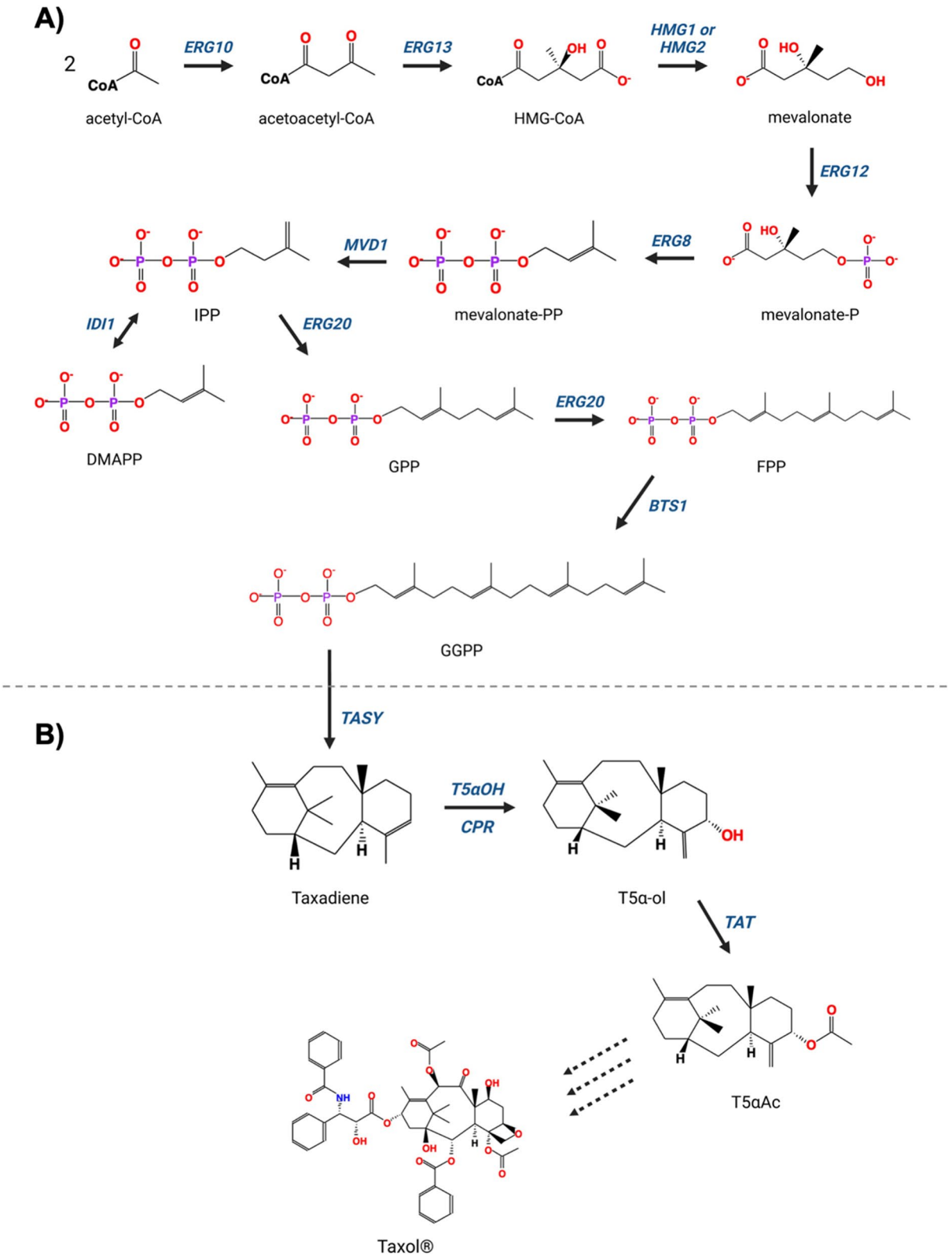

Fig. 1 (See legend on previous page.)

namely OptKnock, OptGene and OptForce in conjunction with yeast GEM, to predict gene and reaction candidates for knockout or overexpression. The genomic modifications determined by in silico analyses were implemented using the CRISPR-based ACtivE toolkit [38] to design yeast strains. Overall, we tested the knockout of nine genes and the overexpression of eight genes, both individually and in various combinations. Following the screening of the strains, the best-performing one was also evaluated in 250 mL in a mini-scale bench-top bioreactor. We achieved a 50% increase in taxadiene yield compared to the parent strain when four reactions were upregulated by integrating an extra copy of *ILV2, TRR1, ADE13* and *ECM31* genes driven by a galactose-inducible promoter. Using a micro-scale high-throughput bioreactor platform, i.e. BioLector, we achieved the detection of 215 mg/L of taxadiene, 35.4 mg/L of T5α-ol and 26.2 mg/L of T5αAc in complete synthetic media (CSM). These are the highest titers reported until now in *S. cerevisiae*. Our findings highlight the applicability of the genomic modifications employed in this study to enhance the metabolic flux towards the yeast mevalonate pathway for the production of high-value terpenoids.

Materials & methods

in silico design and analyses

In three design algorithms, OptKnock, OptGene and OptForce, cytosolic acetyl-CoA and geranylgeranyl diphosphate (GGPP) were separately selected as target metabolites to increase production in glucose or galactose-containing media. During simulations, uptake routes for key molecules such as inorganic phosphate, sulphate, ammonia, and oxygen were unconstrained, while secretion routes for acetate, carbon dioxide, ethanol, glycolaldehyde, diphosphate, water, and glycerol and acetaldehyde were enabled. When galactose exchange was enabled, glucose exchange was constrained or vice versa. While *S. cerevisiae* cultivations (see below) used complete synthetic medium (CSM, Additional file 1: Table S6 and Table S7), it was decided not to allow for in silico amino acid uptake, due to the difficulties to provide realistic constraints for these nutrients.

The design algorithms underwent up to five iterative runs to generate a set of predictions, each containing a maximum of three candidates. Subsequently, we pooled all recommended genomic modification candidates and through various approaches we prioritized high-potential modifications from this collective pool, as described in the Additional file 1.

The COBRA Toolbox v3.0 [32, 39] was used in MATLAB R2019a for in silico studies, with Gurobi Optimizer (9.5.0) as the solver. Yeast GEM 8.5.0 [28] was used to represent yeast metabolism. The ECDF Linux Compute Cluster (Eddie), the University of Edinburgh's research computing cluster, was employed to find to run OptForce. Escher [40] was used on the iMM904 yeast model to visualize the metabolic maps since Yeast 8 does not have a compatible Escher map.

Oligonucleotides and Reagents

All primers used in the study are listed in Additional file 1: Table S1 and Table S3. The primers were ordered from Integrated DNA Technologies (IDT) as standard DNA oligos. Synthetic gRNA cassettes (Additional file 1: Table S2) were ordered from Twist Bioscience. Phusion Flash High-Fidelity PCR Master Mix (Thermo Scientific™) and PrimeSTAR® GXL DNA Polymerase (TaKaRa) were used for PCR reactions to produce the DNA parts for genome modifications. GeneJET PCR Purification Kit (Thermo Scientific™) was used for PCR clean-up. DreamTaq Green PCR Master Mix (Thermo Scientific™) was used for colony PCR.

Strains and Media

The original *S. cerevisiae* strain used in this study is a CEN.PK2-1C-originated yeast strain, *LRS6* (*MATa, leu2-3, 112::HIS3MX6-GAL1p-ERG19/GAL10p-ERG8; ura3-52::URA3-GAL1p-MvaS^{A110G}/GAL10p-MvaE*; *his3Δ1::hphMX4-GAL1p-ERG12/GAL10p-IDI1; trp1-289::TRP1_GAL1p-CrtE (X. dendrorhous)/GAL10p ERG20; YPRCdelta15::NatMX-GAL1p-CrtE/GAL10p-CrtE*; *ARS1014::GAL1p-TASY-GFP*; *ARS1622b::GAL1p-MBP-TASY-ERG20*; *ARS1114a::TDH3p-MBP-TASY-ERG20; ARS511b::GAL1p-T5αOH/GAL3-CPR; RKC3::GAL1-TAT*) [23]. Briefly, the original strain contains additional copies of native and heterologous genes in the mevalonate pathway to improve GGPP production, along with codon-optimized heterologous genes encoding the enzymes responsible for early steps in Taxol® biosynthesis (Figs. 1, 3A). An additional *GAL1* promoter-driven *TAT* gene was integrated to the *RKC4* location into the *LRS6* genome, which was used as the parent strain in the subsequent studies. The other strains derived from the parent strain are shown in Additional file 1: Table S4.

Complete synthetic medium (CSM) consisting of 0.67% (w/v) yeast nitrogen base without amino acids (Alfa Aesar™), 0.08% (w/v) complete supplement mixture (MP Biomedicals™) and 2% (w/v) glucose (Alfa Aesar™) or 2% (w/v) galactose (ACROS Organics™) was used for the cultivations. The compositions of the yeast nitrogen base and the complete supplement mixture are given in Additional file 1: Table S6 and Table S7, respectively. Pre-cultures were grown in standard rich medium, YPD medium, 1% (w/v) yeast extract (Fisher BioReagents™), 2% (w/v) peptone (Merck, Millipore®), 2% (w/v) glucose.

To select the positive transformants, selective media, CSM-URA, 0.67% (w/v) yeast nitrogen base without amino acids, 0.077% (w/v) complete supplement mixture without uracil, 2% glucose, 2% agar (ACROS Organics™) was used.

Yeast transformation and strain construction

The chemicals and reagents were obtained from Sigma-Aldrich unless otherwise stated. Transformations were carried out according to LiAc/PEG heat-shock method [41]. Briefly, fresh cultures were prepared to obtain the cells in the exponential phase following an overnight culture. The cells were then washed using sterile water and were pelleted by centrifugation. The transformation mix, 240 µL PEG (50%(w/v)), 36 µL 1.0 M lithium acetate (LiAc) and 50 µL single-stranded carrier DNA (2.0 mg/mL, herring sperm DNA, Promega), was added onto the cell pellet. Following this, DNA fragments and water were added until the volume was made up to 360 µL. After homogenous transformation mixes were obtained, the cells were incubated for 45 min at 42 °C. Finally, the transformation mix was removed, the cells were plated onto the selective media, and the plates were incubated for 2–3 days at 30 °C.

Genomic modifications, gene integrations or deletions, were carried out using the modular ACtivE toolkit and method that was recently developed by our group [38]. Additional file 1: Figure S1, S2 outline the overall working principle of the ACtivE method. Briefly, 50 fmol equivalent molarity of each plasmid module was used to assemble a single all-in-one CRISPR plasmid. Four plasmid modules, Cas9 cassette, selection marker (*URA3*), storage part (bacteria ORI and Amp^R) and yeast origin of replication (2 µ) were combined with the corresponding gRNA cassettes according to the ACtivE method [38]. ARS209, ARS306, ARS727, ARS1531 and ARS1603 were used as integration sites for the genomic integrations [38]. 500 ng–1000 ng from each donor DNA part containing overlapping fragments with their neighbor parts was used. The upstream homology arm (UHA), promoter, coding sequence (with terminator) and downstream homology arm (DHA) were designed for integrations, whereas only UHA and DHA were used for deletions. The potential crRNA sequences on each region were scored using CRISPOR [42] (http://crispor.tefor.net), an online gRNA selection tool giving sequence-based scores using sequence prediction algorithms.

The colonies were first screened using colony PCR to detect the genomic modifications. Genomic DNAs of the positive transformants were extracted using Pierce Yeast DNA Extraction Reagent Kit (Thermo Scientific™). The target regions were confirmed by Sanger sequencing performed at GENEWIZ, Inc (Leipzig, Germany).

Before the cultivations, *URA3* containing CRISPR plasmids were removed using the 5-Fluoroorotic Acid (5-FOA) (Thermo Fisher Scientific) counter-selection method with a synthetic defined medium supplemented with 0.1% (w/v) 5-FOA.

High-throughput strain screening

To determine the best producers of the Taxol® precursors, the strains containing single genomic modifications (integration or deletion) were screened using V-shaped, 24-deep well-plates (Axygen™). A 20% dodecane (ACROS Organics™) overlay was added to set a working volume of 2 ml in each well, and CSM with glucose or galactose was used for the cultivations. The initial OD_{600} was set to 1.0 for each well by diluting pre-cultures. The plates were incubated with shaking at 350 rpm at 30 °C for 72 h on thermomixers. Gas permeable adhesive plate seals (Thermo Scientific™) ensured oxygen transfer. The total biomass in each well was measured at 600 nm wavelength using the Nanodrop™ 2000c spectrophotometer (Thermo Scientific™).

In the second-level screening, the best performing strains with single genomic modifications and the strains containing multiple modifications were screened employing a BioLector Pro (mp2-labs) microbioreactor-screening platform. A flower-shaped, transparent bottom 48 well-plate (mp2-labs) containing pH and dissolved oxygen (DO) optodotes was used for online monitoring of biomass, pH and DO. Similar to the first screening, glucose or galactose-containing CSM was used as the medium, the initial OD_{600} was set to 1.0, and a 20% dodecane overlay was used in 1 mL working volume. The plate was covered using a gas-permeable sealing foil with an evaporation reduction layer (mp2-labs), and the temperature was maintained at 30 °C under the agitation of 1000 rpm. Biomass absorbance units were measured with the gain = 6.

After cultivations in the plates, dodecane layers were collected following a centrifuge step, and the taxane (taxadiene and derivatives) production was analysed via gas chromatography-mass spectrometry (GC–MS).

Bioreactor cultivation

MiniBio500 bioreactor (Applikon Biotechnology) was used for larger scale cultivation for the best producer strain. 250 mL total reaction volume containing a 20% dodecane layer was used for the cultivations. Similar to the other cultivations, the initial OD_{600} was adjusted to 1.0 by diluting overnight cultures. The same medium compositions as the previous experiments, galactose containing CSM, were used. To mitigate the foam formation, polypropylene glycol P2000 0.01% (v/v) (Alfa Aesar), was used in the medium as anti-foam. pH, DO, and

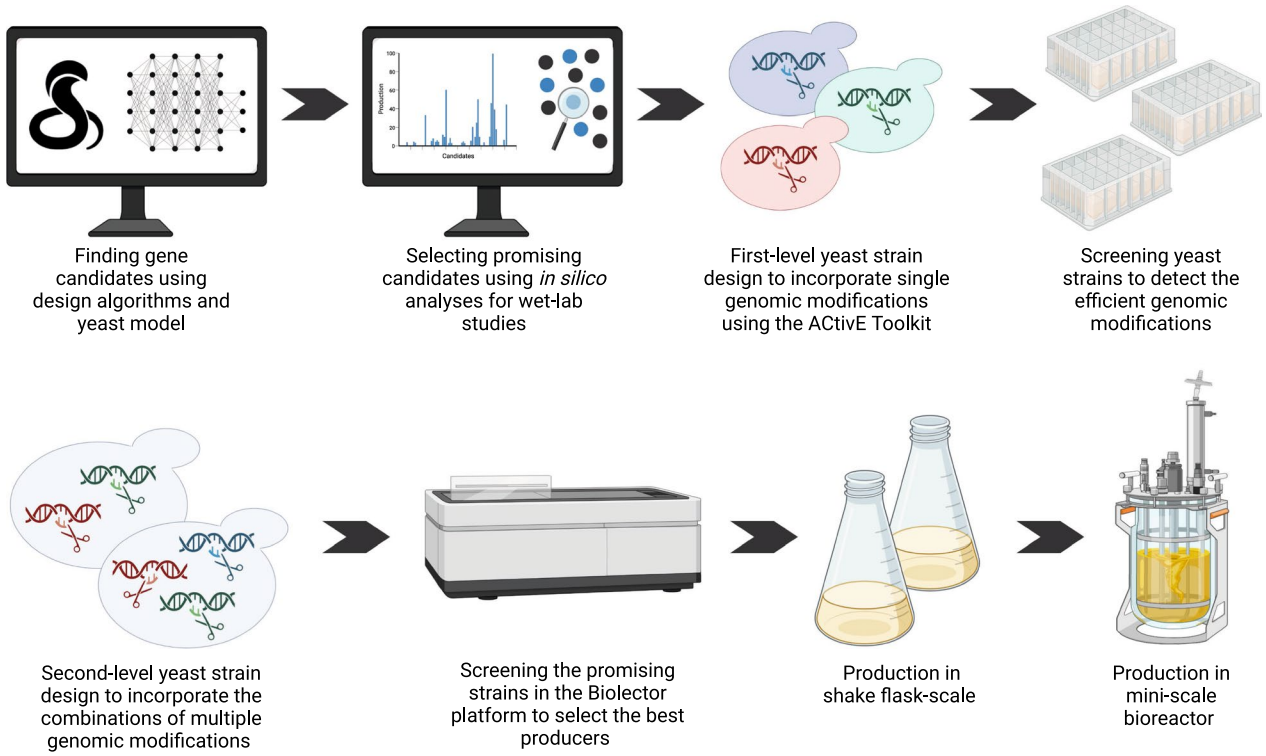

Fig. 2 The workflow followed in this study from in silico design to lab bench scale production. Three design algorithms, OptKnock, OptGene and OptForce, were used on Yeast 8.5.0 to find gene candidates (Additional file 1: Table S5). 17 of these gene candidates were prioritised through further in silico predictions (Figs. 4, 5). After selecting the gene candidates, the yeast strains containing single genomic modifications were designed (Table 1). To engineer the yeast genome, the ACtivE Toolkit was used. The strains were then screened in deep-well plates to detect the good performance genomic modifications (Fig. 6). These modifications were then combined to design second-level yeast strains containing multiple genomic modifications (Table 2). The best-performing strains were then screened in the Biolector platform with real-time pH, DO and biomass monitoring (Fig. 7 and Fig. 8). Finally, the best-producer was used for the productions a lab bench scale bioreactor (Fig. 9)

temperature were measured online using the my-control system (Applikon Biotechnology). Temperature was set to 30 °C. A setpoint 70% saturation of dissolved oxygen (DO) was applied, an air sparger was used automatically to maintain O_2 level. pH was maintained in a particular range ($5.0 < pH < 6.5$) and 1 M of NaOH was added when the pH was below the threshold. Biomass was measured online using the Optura system with a BE2100 OD sensor (BugLab). Samples were taken daily for offline biomass measurement and quantification of the metabolite concentrations.

An inverted microscope (Leica Microsystems) was used following all cultures in this study to detect if there was microbial contamination. The cells were monitored under 100X lens with immersion oil (Nikon).

Metabolite identification and quantification

The dodecane layer collected at the end of the cultivations was analyzed by GC–MS as described previously [22]. A GC system, Trace™ 1300 Series (Thermo Scientific™), equipped with TraceGOLD™ TG-SQC GC column, was used. The mass spectra of 50–650 m/z were recorded on an ISQ™ Series Single Quadrupole MS (Thermo Scientific™) using EI ionization mode and a scan time of 0.204 s. The GC–MS data were processed using the Xcalibur™ software. Pure taxadiene provided by the Baran Lab (The Scripps Research Institute) and GGOH (Sigma Aldrich) were used as standards. The concentrations of the additional compounds were calculated relative to the taxadiene standard.

Statistical analysis

Strain screening experiments were conducted in at least three replicates. The well layouts in both 24-well plates and 48 well-plate were randomized to mitigate the plate effects and possible errors. The error bars represent the standard deviations of different experiments. The one-way analysis of variance (ANOVA) was used to determine if there was a statistically significant difference between the samples or experiments. The null hypothesis considered no significant difference between the samples/runs; the null hypothesis was rejected if the p-value ≤ 0.05.

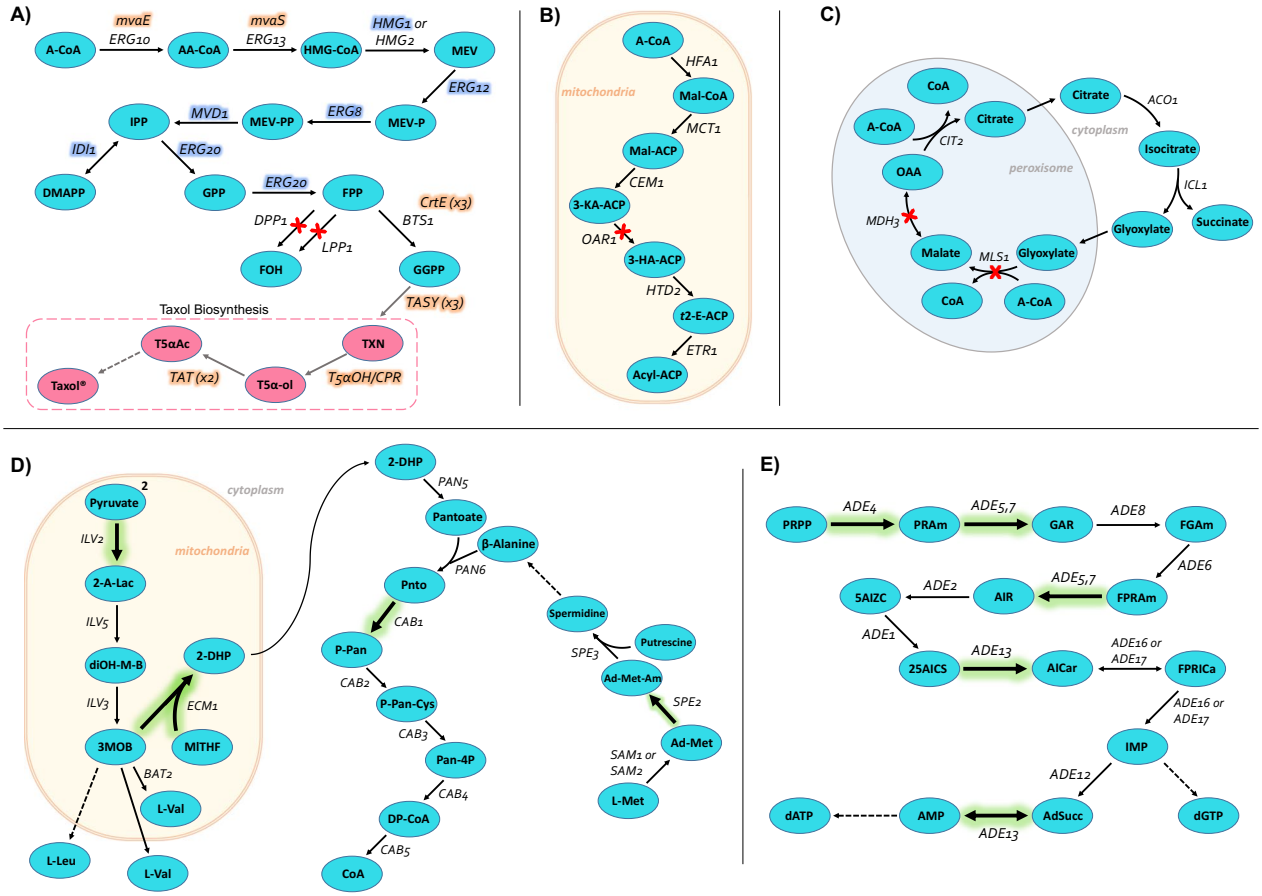

Fig. 3 The pathways, genes and reactions modified in this study. **A)** The mevalonate pathway. The additional copies of native genes (blue shaded) and the heterologous genes (orange shaded) integrated in a previous study to construct the parent strain LRS6 [23] are highlighted. **B** Fatty acid biosynthesis in mitochondria **C** Glyoxylate cycle **D** Connected pathways; branched-chain amino acid biosynthesis (from two pyruvate molecules to L-valine or L-leucine), pantothenate pathway (from 3-methyl-2-oxobutanoate a 5,10-methylenetetrahydrofolate to pantothenate) and the spermidine biosynthesis (from *S*-adenosyl-L-methionine to spermidine) **E** de novo purine biosynthesis. The red crosses indicate corresponding gene deletions suggested by the design algorithms; thick arrows highlighted with green colour indicate the corresponding overexpressed genes suggested by the design algorithms. The arrows show the direction of the fluxes. Dashed arrows indicate multiple reactions. A-CoA: acetyl-CoA, AA-CoA: acetoacetyl-CoA, HMG-CoA: hydroxymethylglutaryl-CoA, MEV: mevalonate, MEV-P: R-5 phosphomevalonate, MEV-PP: R-5-diphosphomevalonate, IPP: isopentenyl diphosphate, DMAPP: dimethylallyl diphosphate, GPP: geranyl diphosphate, FPP: farnesyl diphosphate, FOH: farnesol, GGPP: geranylgeranyl diphosphate, TXN: taxadiene, T5α-ol: taxa-4(20),11-dien-5α-ol, T5αAc: taxa-4(20),11-dien-5-α-yl acetate, Mal-CoA: malonyl-CoA, Mal-ACP: malonyl-ACP, 3-KA-ACP: 3-ketoacyl-ACP, 3-hydroxyacyl-ACP, *t*2-E-ACP: *trans*-2-enoyl-ACP, OAA: oxaloacetic acid, 2-A-Lac: 2-acetolactate, diOH-M-B: 2,3-dihydroxy-3 methylbutanoate, 3MOB: 3-methyl-2-oxobutanoate, 2-DHP: 2-dehydropantoate, Pnto: R-pantothenate, P-Pan: R- 4′ phosphopantothenate, P-Pan-Cys: R- 4′ phosphopantothenoyl-L-cysteine, Pan-4P: 4′-phosphopantetheine, DP-CoA: 3′-dephospho-CoA, PRPP: 5-Phospho-alpha-D-ribose 1-diphosphate, PRAm: 5-Phospho-beta-D-ribosylamine, GAR: N1-(5-Phospho-D-ribosyl) glycinamide, FGAm: N2-Formyl-N1-(5-phospho-D-ribosyl) glycinamide, FPRAm: 2-(Formamido)-N1-(5-phospho-D-ribosyl) acetamidine, AIR: 5-amino-1-(5-phospho-D-ribosyl)imidazole, 5AIZC: 5-amino-1-(5-phospho-D-ribosyl) imidazole-4-carboxylate, 25AICS: (S)-2-[5-Amin o-1-(5-phospho-D-ribosyl) imidazole-4-carboxamido]succinate, AICar: 5-Amino-1-(5-Phospho-D-ribosyl)imidazole-4-carboxamide, FPRICa: 5-Formamido-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide

Results & discussion

The main objective of the study was to improve metabolic fluxes towards the mevalonate pathway, thus, towards GGPP which is a common precursor of diterpenes [43], to enhance the production of the early step Taxol® precursors synthesized by our engineered yeast strains. To this end, computer-aided design was used to predict/identify potential genomic modifications that would be difficult to determine intuitively. Strain design algorithms were used on yeast GEM 8.5.0. The gene deletion or integration candidates were simulated using COBRA tools. The strains containing single genomic modifications were then designed using the ACtivE toolkit and were screened in V-shaped deep-well plates. The promising

Table 1 Selected gene candidates predicted by different design algorithms in various media conditions

Target Gene	Design algorithm	Target compound	Carbon source	Intervention
DPP1 (YDR284C)	OptGene	Acetyl-CoA	Glucose	Knock-out
OAR1 (YKL055C)	OptGene	GGPP	Galactose	Knock-out
MDH3 (YDL078C)	OptKnock & OptForce	Acetyl-CoA	Glucose	Knock-out
ACH1 (YBL015W)	OptKnock & OptGene	Acetyl-CoA	Glucose	Knock-out
MLS1 (YNL117W)	OptKnock	Acetyl-CoA	Galactose	Knock-out
DIT1 (YDR403W)	OptGene	Acetyl-CoA	Glucose	Knock-out
LPP1 (YDR503C)	OptGene	GGPP	Glucose & galactose	Knock-out
ISC1 (YER019W)	OptGene	Acetyl-CoA	Galactose	Knock-out
DTR1 (YBR180W)	OptGene	GGPP	Glucose	Knock-out
ILV2 (YMR108W)	OptForce	Acetyl-CoA	Galactose	Overexpression
TRR1 (YDR353W)	OptForce	Acetyl-CoA	Galactose	Overexpression
ADE4 (YMR300C)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression
ADE5,7 (YGL234W)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression
ADE13 (YLR359W)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression
ECM31 (YBR176W)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression
CAB1 (YDR531W)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression
SPE2 (YOL052C)	OptForce	Acetyl-CoA	Glucose & galactose	Overexpression

single genomic modifications and their combinations were tested in a BioLector microbioreactor system. Following this, the best producer was also tested in a lab bench bioreactor. Figure 2 demonstrates the whole process explained here.

in silico design

Our combinatorial approach benefited from three design algorithms, OptKnock, OptGene and OptForce on yeast GEM 8.5.0. Due to the incomplete knowledge on the Taxol biosynthetic pathway, including the exact reactions that are catalysed by its enzymes, we focused on acetyl-CoA and GGPP as production targets (after addition of their respective exchange reactions), which are the first and last metabolites in the mevalonate pathway, respectively (Fig. 3A).Since the *S. cerevisiae* strain used for genomic modifications employed inducible galactose promoters, the simulations were carried out with galactose or glucose as carbon source, separately. Nonessential reactions and genes were first determined for OptKnock and OptGene using flux balance analysis (FBA) [44]. Minimum growth rate was set to 50% of wild-type growth when targeting acetyl-CoA for OptKnock deletions, however, it was set to 20% for GGPP as OptKnock could not predict potential gene deletions in favour of GGPP maintaining higher growth rate. OptGene was run over 500 and 1000 generations, and all predictions were considered. The genomic modifications suggested by both the first "MUST set" and the second "MUST set" of OptForce runs [37] were considered. Since the design algorithms, OptKnock, OptGene and OptForce, used in this study prioritise growth-coupled solutions, all solutions taken into account were growth-coupled. Table 1 compiles the selected target genes and corresponding design algorithms.

Using different conditions and design algorithms as highlighted above, various gene and reaction candidates were found for either acetyl-CoA or GGPP overproduction. We intended to use these in silico predicted targets to implement in vivo by constructing relatively simple single-mutation strains. To therefore prioritize the most promising ones among the predicted genomic modifications (Additional file 1: Table S5), we employed in silico simulations to validate their effect. In these simulations, the lower bound of the biomass function was set to a value of 0.1, to ensure that the optimised model would still have active fluxes toward the biomass reaction. Then, by going through one candidate reaction at a time, initial flux of these reactions were first fixed to an arbitrary value of 0.01 in an initial simulation, while in a subsequent simulation of a mutant strain the fluxes were fixed to either 0 or 0.02 to mimic deletion (reaction knockout) and overexpression (gene integration), respectively. Then, in each of the model variants, the maximum production fluxes of acetyl-CoA, GGPP and biomass were determined, and the effect of the genomic modification relative to the initial simulation was plotted (Fig. 4). The genes that showed the largest increase in maximum flux values were prioritized for further analysis and wet-lab experiments.

Although clear increases were found with approximately two-fold overproduction in acetyl-CoA when

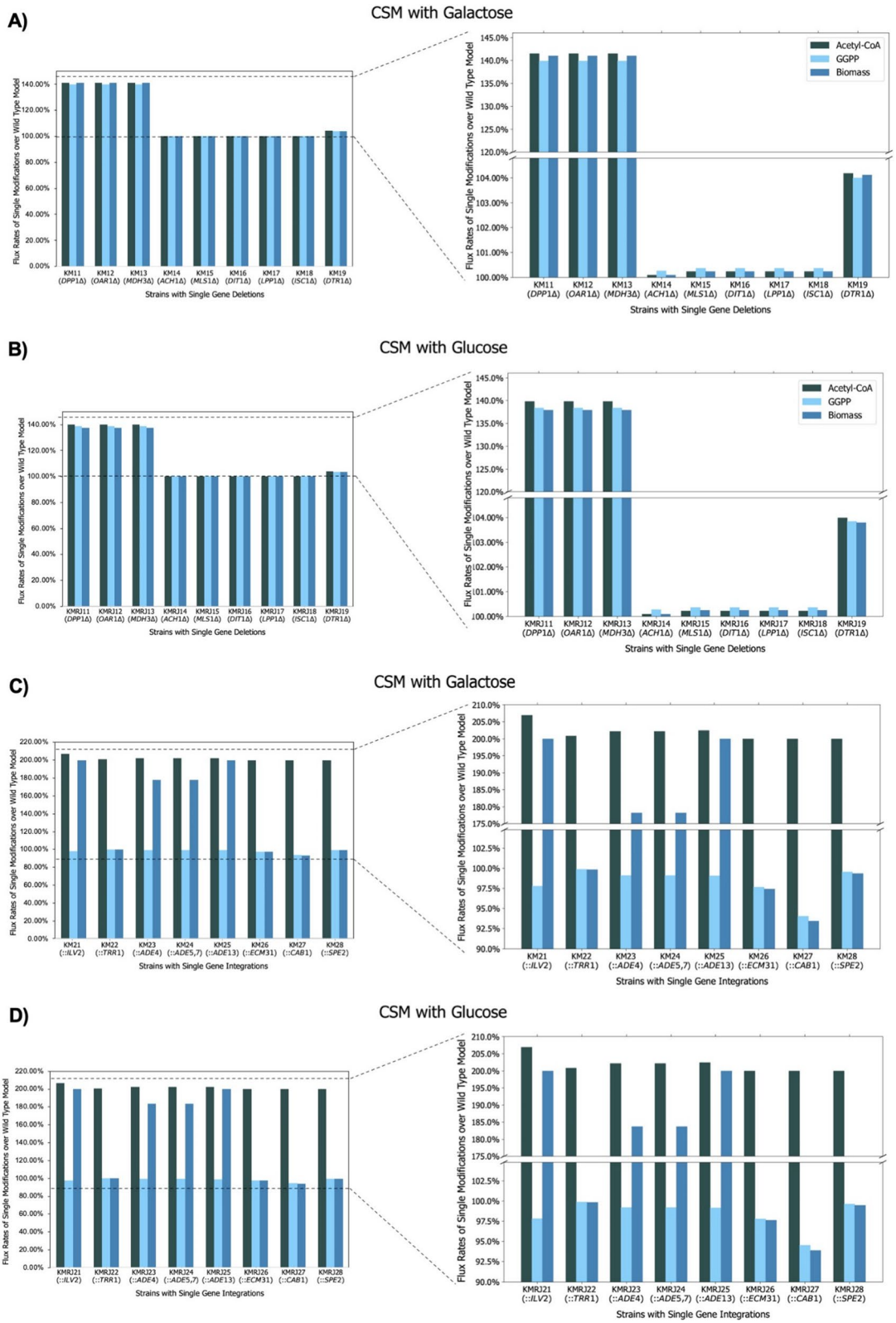

Fig. 4 Comparison of maximum in silico fluxes for corresponding objectives, acetyl-CoA, GGPP and biomass reactions. For simplicity, the initial state (flux value was 0.01) was referred to as wild-type, while the final state (flux value is 0.00 for deletion and 0.02 for integration) was referred to as modification. **A**, **C** were simulated in galactose containing CSM, **B**, **D** were simulated in glucose-containing CSM. Δ; gene deletion,:; gene integration

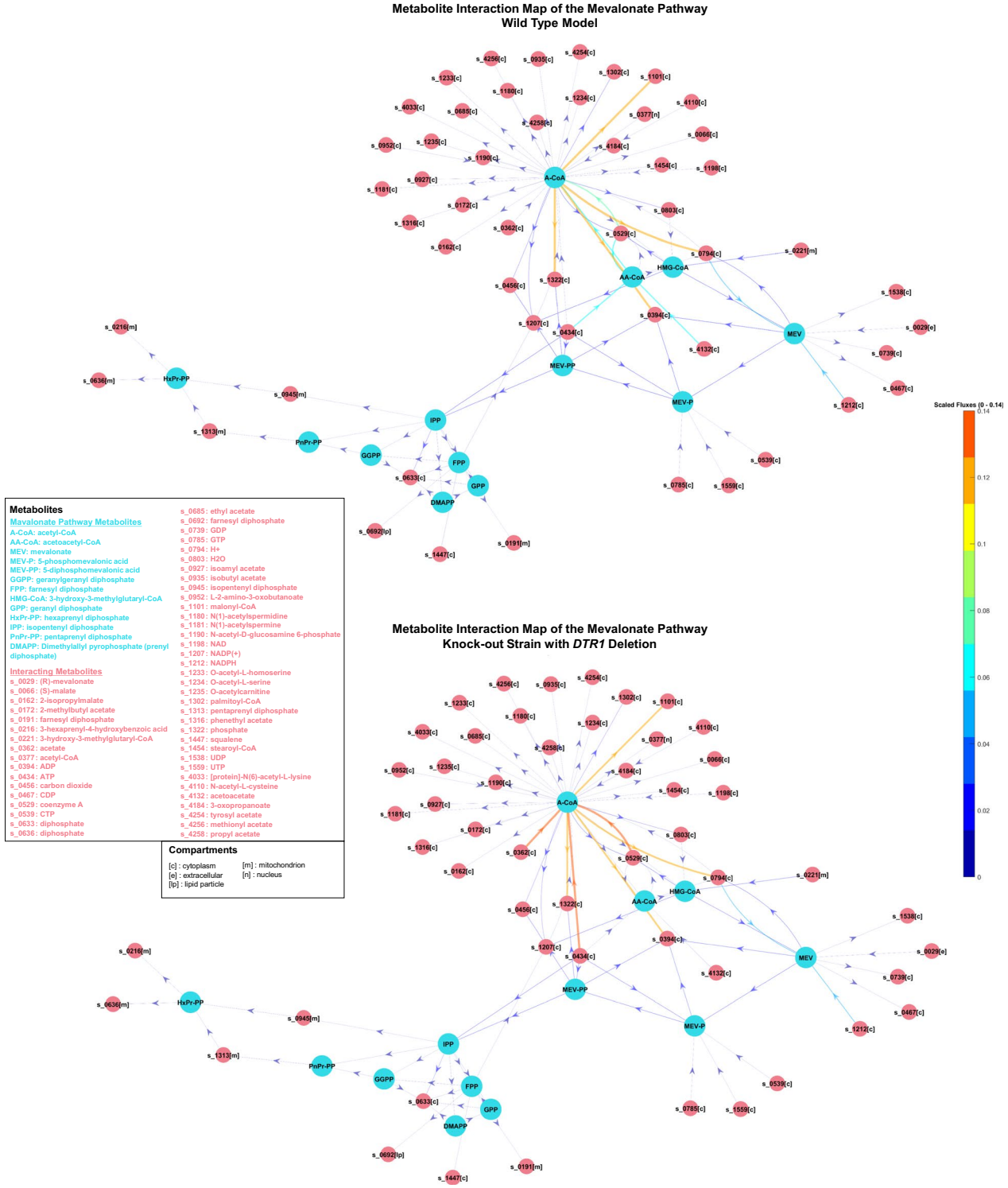

Fig. 5 Metabolite interactions with corresponding fluxes in the mevalonate pathway in the wild-type model and the *DTR1* deleted knock-out model. Metabolite interactions were constructed using metabolite-metabolite interaction network in COBRA Toolbox [45, 46]

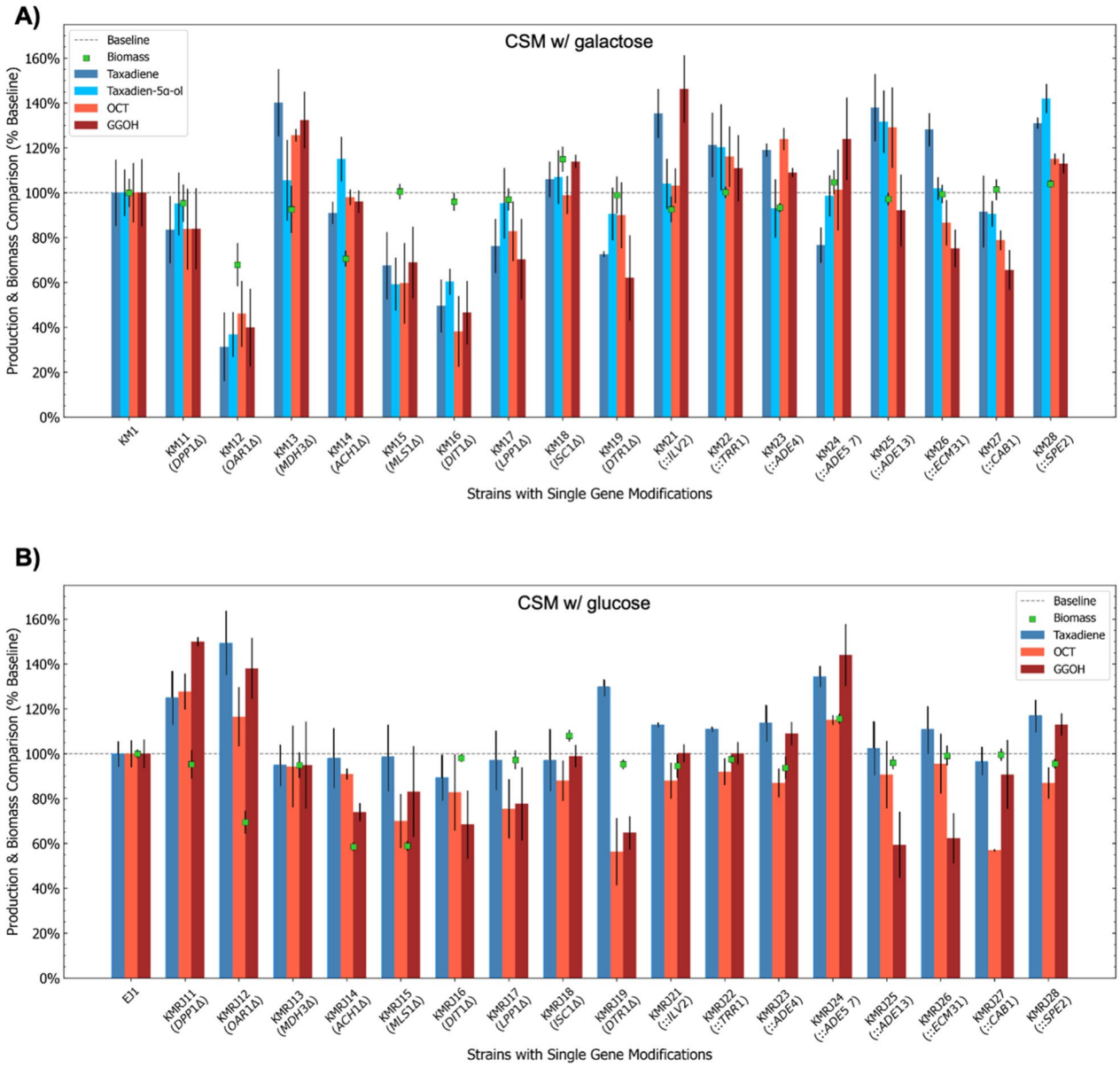

Fig. 6 Comparison of biomass and production of early step Taxol® precursors and side products. **A** KM1-derived strains in the galactose-containing CSM. **B** EJ1-derived strains in the glucose-containing CSM. Relative values are shown considering 100% baseline (dashed line) of the parent strains, KM1 or EJ1. Error bars indicate standard deviations between three replicates. Δ; gene deletion, gene integration

particular genes were overexpressed in silico, the overproduction rates were relatively low in gene deletions except for *DPP1, OAR1* and *MDH3* that showed ~ 40% increase (Fig. 4). To understand the lower increases better, metabolite interaction maps were constructed for the gene deletions including *ACH1, MLS1, DIT1, LPP1, ISC1* and *DTR1* to investigate the effects of these gene knockouts. First, wild-type and knock-out models were optimised by FBA in glucose or galactose containing media, then the metabolite interactions in the mevalonate pathway were determined with their fluxes, as shown in Fig. 5. The acetyl-CoA-centered and GGPP-centered interactions were also investigated from this map (Additional file 1: Figure S3–S9), and flux differences between the wild-type and knock-out models were compared (Fig. 5). We could not detect flux increase towards GGPP in any of the knock-out models even though the objective functions were removed, as explained in the methodology section. Therefore, only acetyl-CoA increases were considered. The gene deletions that were predicted to

increase the possible maximum fluxes (Fig. 4A & B) and in metabolite interaction maps (Fig. 5, Additional file 1: Figure S3–S9) were selected for the strain construction.

For overexpressions, only the reactions that increased the fluxes towards acetyl-CoA production were considered (Table 1). This was because OptForce suggested to upregulate particular reactions in the mevalonate pathway (Additional file 1: Table S5) for GGPP overproduction, and these were already improved in our engineered strain, LRS6. Therefore, the initial metabolic state in the parent strain was as if OptForce's predictions were implemented in the mevalonate pathway. OptForce did not find upregulation of early step reactions (the reactions of *ERG10* and *ERG13* genes) useful in the mevalonate pathway; the possible maximum flux of the GGPP production did not change compared to the initial state. It was a consistent result found by in silico simulations as a previous experimental study reported that early step mevalonate pathway genes of *S. cerevisiae* did not show efficient production of mevalonate compared to bacterial-sourced equivalents [47]. OptForce did not suggest any reaction in other pathways to be upregulated to enhance GGPP production.

Strain construction and screening of the single genomic modifications

As our engineered yeast strains contained additional genes (Fig. 3) driven by galactose-inducible promoters, the parent strain KM1 (Additional file 1: Table S4), was not able to induce the integrated mevalonate pathway genes using glucose as sole carbon source. Therefore, we first deleted the *GAL80* gene yielding with EJ1 strain (Additional file 1: Table S4) to allow glucose utilization as Gal80 protein inhibits the transcription of the galactose-inducible genes in the absence of galactose [48]. These two strains were then designed to incorporate the genomic modifications, single-gene deletion or integration (overexpression), predicted by the design algorithms. When the additional copies of native yeast genes were integrated for overexpression, native promoters were simply changed with *GAL1* promoter using the ACtivE method (Additional file 1: Figure S2) [38] to enhance their expression rates [22, 49]. Additional copies were integrated into an intergenic region close to the autonomously replication sequence (ARS) 1603 [38] in each strain to screen the effect of single gene overexpression.

In this study, nine gene knock-outs and eight gene overexpressions were tested to determine whether these modifications affect biomass and the production of early steps Taxol® precursors, taxadiene, taxa-4(20),11-dien-5α-ol (T5α-ol), and taxa-4(20),11-dien-5-α-yl acetate (T5αAc), via the mevalonate pathway. To screen single genomic modifications, the strains containing single gene deletions or integrations were cultured for three days in galactose or glucose-containing CSM media as simulated previously. In addition to the Taxol® precursors, some side compounds such as verticillene, iso-taxadiene, geranylgeraniol (GGOH), 5(12)-oxa-3(11)-cyclotaxane (OCT), iso-OCT, and additional diterpenes were also detected by GC/MS analyses as previously reported in more detail [23]. Among them, GGOH and OCT were also quantified to compare the production of these side products. Two Taxol® precursors (taxadiene and T5α-ol), two side-products (GGOH and OCT) and biomass were considered in a micro-scale screening in 2 mL cultures to characterize the modified strains. Figure 6 shows the comparison of the production of these metabolites and biomass between the parent strains and designed strains.

Differences in production between galactose-containing (CSM w/ galactose) and glucose-containing (CSM w/ glucose) media were detected. Taxadien-5α-ol productions varying from ~ 4 mg/L to ~ 8 mg/L were observed in galactose-containing media with KM1-derived strains; however, only weak peaks were seen in the gas chromatogram for T5α-ol when the glucose-containing media was used with EJ1-derived strains that did not contain the *GAL80* gene. This was most probably because the expression rates of the galactose-inducible promoters were still higher in galactose-containing media, even though EJ1-derived strains could produce taxadiene under glucose only conditions (Fig. 6). Additional regulatory elements might have had an impact on the activation of galactose-inducible promoters in the presence of glucose. In the galactose-containing CSM media, mainly integrations increased the downstream production from GGPP, whereas most of the deletions and integrations showed positive impacts on the production of the precursors in glucose-containing media.

DPP1 and *LPP1* genes, encode two Mg^{+2}-independent phosphatidate phosphatases found in yeast [50]. Deletion of these genes might result in higher farnesyl diphosphate (FPP) and GGPP titers as these enzymes are also responsible for converting FPP to farnesol (Fig. 3A). However, in the galactose-containing medium, both single deletions showed a similar pattern with lower average ($p > 0.05$) taxadiene production (Fig. 6A). On the other hand, a ~ 20% increase ($p < 0.05$) in taxadiene production with *DPP1* deletion was observed in the glucose-containing medium (Fig. 6B), while *LPP1* deletion resulted in almost the same taxadiene production ($p > 0.05$). No significant growth phenotype was observed in any condition (Fig. 6). in silico simulations suggested that *DPP1* deletion could lead to acetyl-CoA overproduction in a glucose-containing medium (Table 1 and Fig. 4B), and experimental results (Fig. 6B) supported this.

The protein encoded by *OAR1* is an NADPH-dependent 3-oxoacyl-(acyl-carrier-protein) reductase, which is involved in fatty acid biosynthesis in yeast [51, 52] (as illustrated in Fig. 3B). *OAR1* deficient yeast strains could survive on fermentable carbon sources like glucose but not on non-fermentable carbon sources such as glycerol [51, 52]. Approximately 30% reduction in biomass was observed in the glucose-containing and the galactose-containing CSMs (Fig. 6). Also, the production rate of the Taxol® precursors obtained from KM12 was less than 40% of the parent strain. In contrast, more than 40% increase was noted in glucose-containing CSM using KMJR12 (Fig. 6B).

Both OptKnock and OptForce found *MDH3* as a good target to increase acetyl-CoA (Table 1). *MDH3* encodes a peroxisomal malate dehydrogenase [53] responsible for the interconversion of malate and oxaloacetate using NAD^+ or NADH depending on the reaction in the glyoxylate cycle (Fig. 3C) [54, 55]. OptKnock also suggested deleting the *MLS1* gene that encodes another enzyme, malate synthase, involved in the glyoxylate cycle [54]. Malate synthase catalyses the formation of malate from glyoxylate and acetyl-CoA in either cytosol or peroxisome, depending on the carbon source (Fig. 3C) [56]. Cytosolic and peroxisomal acetyl-CoA is an essential compound in the glyoxylate cycle. However, the glyoxylate cycle is activated when two-carbon compounds such as ethanol and acetate are used as the sole carbon sources or β-oxidation of fatty acids is enabled [56]. Although we used only six-carbon sources, galactose or glucose, we detected significant effects of *MDH3* or *MLS1* deletions in our study, as shown in Fig. 6. *MLS1* deleted strain KMRJ15 produced less biomass, ~ 60% of the reference strain, in glucose-containing CSM. A similar growth trend was previously reported [57]. On the other hand, KM13 (*MDH3*Δ) showed a significant increase with ~ 40% in taxadiene production when galactose was used as carbon source even though it was unexpected as *MDH3* might not be functional in the galactose-containing medium. In contrast, *MLS1* deletion resulted in ~ 30% lower taxadiene in KM15. These results might indicate that the glyoxylate cycle genes might still be functional under different carbon sources, and that blocking particular reactions in this cycle can affect cytosolic acetyl-CoA synthesis even if the carbon source is not a two-carbon compound or fatty acid.

In mitochondria, acetyl-CoA hydrolase (Ach1) encoded by *ACH1* can catalyse the conversion of acetyl-CoA into acetate and CoA. Ach1 can also reversibly transfer CoA from succinyl-CoA to acetate [58]. OptKnock and OptGene proposed *ACH1* deletion to increase acetyl-CoA concentration in glucose-containing media. Although we detected significantly lower biomass ($p < 0.05$) in both KM14 and KMRJ14 strains (Fig. 6), taxadiene concentrations were comparable with the reference strains, indicating that production of Taxol® precursors was improved per OD with *ACH1* deletion even if the total production could not be enhanced.

Perhaps, the most surprising findings of OptGene were the deletions of *DIT1* and *DTR1* genes that take part in the sporulation process [59, 60]. Dit1 plays a role in forming the dityrosine layer of the spore wall, and its expression rate increases during spore wall maturation [61]. Dit1 expression might be active in vegetative cells; yet, its cellular activities of are not fully elucidated [62]. *DTR1* is a member of the gene network controlling the assembly of the spore wall [63]. Dtr1 protein can also increase the resistance against some growth inhibitors in vegetative yeast cells [60]. The deletions of these genes did not show a significant decrease in biomass in either glucose-containing or galactose-containing CSM (Fig. 6). However, *DIT1* deletion in KM16 caused a dramatically lower production of the Taxol precursors in the galactose-containing medium. In contrast, *DTR1* deletion increased taxadiene concentration in the glucose-containing medium. It is very likely that deletion of these sporulation-specific genes directly makes an impact on GGPP concentration, as it was reported that geranylgeranyl diphosphate synthase has a critical role in the sporulation process in fission yeast, and related genes have functional similarity with *S. cerevisiae* [64].

ISC1 encodes inositol phosphosphingolipase C that hydrolyses sphingolipids to produce ceramide, a compound involved in regulations of cell growth, death and stress response [65]. There is no direct relation between *ISC1* and acetyl-CoA or GGPP, and it was shown that *ISC1* deletion might lead to a higher budding pattern than the wild-type strains [66]. Consistently, *ISC1* deletions resulted in higher average biomass in galactose or glucose-containing media (Fig. 6). OptGene suggested *ISC1* deletion to increase acetyl-CoA concentration in a galactose-containing CSM. Parallel to this, we obtained a higher average production for the target metabolites with a ~ 6% increase in average in taxadiene and T5α-ol in the galactose-containing medium, although it was not statistically significant ($p > 0.05$).

Majority of the OptForce predictions were involved in the biosynthesis of metabolic precursors in yeast. *ILV2* encodes acetolactate synthase that catalyses the conversion of two pyruvate molecules to 2-acetolactate in isoleucine and valine biosynthesis pathways [67]. From 3-methyl-2-oxobutanoate produced in valine biosynthesis, coenzyme-A precursors are produced in the phosphopantothenate biosynthetic pathway, where *ECM31* and *CAB1* catalyse the first and fourth steps, respectively [68]. The product of *SPE2* plays a critical role in

spermidine biosynthesis [69]. Although, these genes and pathways might look irrelevant and independent from each other, connecting the intermediate or final products in these pathways (as shown in Fig. 3D) reveals their direct impacts on CoA production and possibly enhanced acetyl-CoA production. *ADE4, ADE5,7* and *ADE13* are responsible for five reactions in de novo purine synthesis as shown in Fig. 3E [70]. while TRR1 regulates the thioredoxin system, protecting yeast against oxidative stress [71].

Integration of *ADE5,7* encoding a bifunctional protein responsible for the second and fifth reaction in de novo purine synthesis pathway resulted in lower taxediene production (~23% decrease, $p < 0.05$) in galactose-containing CSM with strain KM24. In contrast, taxediene production was significantly higher in the glucose-containing medium. Apart from this, overexpression of the *CAB1* gene could not increase the production even though an increased flux in the phosphopantothenate pathway could potentially increase acetyl-CoA concentration. Olzhausen et al. (2021) reported that the native *CAB1* is relatively inefficient for the production of coenzyme A compared to other phosphopantothenate pathway genes, and the researchers dramatically increased coenzyme A titer using mutant *CAB1* W331R [72]. This might be the reason behind non-effective taxane production when *CAB1* was overexpressed. OptForce suggested upregulation of early-stage and mid-stage reactions in de novo purine synthesis that have relatively lower fluxes than the late-stage reactions, as shown in Additional file 1: Figure S11 that seems like an effective strategy to enhance the flux towards the downstream part of the pathway. Still, the relation of de novo purine synthesis pathway and thioredoxin system with acetyl-CoA should be further investigated as increased fluxes in these systems can potentially increase acetyl-CoA concentration or the fluxes in the mevalonate pathway towards GGPP.

In Fig. 3, we have highlighted genomic modifications spanning various metabolic pathways. While certain modifications, such as the deletions of the *LPP1* and *DPP1* genes, may appear intuitively predictable, it is important to note that a significant portion of these genomic alterations, including those not depicted in Fig. 3, such as *DTR1* and *DIT1* deletions, or *TRR1* overexpression, are considerably challenging to anticipate. These findings show that mathematical modelling and optimization of the microbial systems can be beneficial in finding useful genomic modifications that are very difficult to be intuitively predicted. Indeed, detecting small changes using a relatively low number of replicates is quite challenging as the noise deviations can make the screening and evaluations harder. Nevertheless, statistically significant impacts of the genomic alterations were detected when single modifications were tested. Here, we coupled the native genes with strong galactose-inducible promoters. Therefore, alternative approaches such as protein engineering or integrating heterologous genes encoding higher performance enzymes for the target reaction could also be used to enhance the fluxes in the pathways of interest.

Next level screening using high-throughput microscale bioreactor system

To investigate the impact of multiple modifications on taxane production, we incorporated the empirical findings from the first screening experiments and strategically combined the most promising modifications to engineer *next-level* strains. For galactose-containing CSM, five genomic integrations and one gene deletion were combined to construct different combinations of KM1-derived strains, while three gene deletions and two genomic integrations were used to produce EJ1-derived strains to be used in glucose-containing CSM, as shown in Table 2.

The *first-level* strain screening revealed promising genomic modifications enhancing GGPP and taxadiene productions, as discussed above. However, the next compound, T5α-ol, in the Taxol® pathway could

Table 2 The strains containing multiple modifications

Strain	Genotype
KM31	KM1, *ARS1603::ILV2, ARS209::TRR1*
KM32	KM1, *ARS1603::ILV2, ARS209::TRR1, ARS306::ADE13, ARS727::ECM31*
KM33	KM1, *ARS1603::ILV2, ARS209::TRR1, ARS306::ADE13, ARS727::ECM31, MDH3Δ*
KM34	KM1, *ARS1603::ILV2, ARS209::TRR1, ARS306::ADE13, ARS727::ECM31, ARS1531::SPE2, MDH3Δ*
KMRJ31	EJ1, *OAR1Δ, DTR1Δ*
KMRJ32	EJ1, *OAR1Δ, DTR1Δ, ARS1603::ADE5,7*
KMRJ33	EJ1, *OAR1Δ, DTR1Δ, ARS1603::ADE5,7, DPP1Δ*
KMRJ34	EJ1, *OAR1Δ, DTR1Δ, ARS1603::ADE5,7, DPP1Δ, ARS1531::SPE2*

not be effectively detected, and only KM25 and KM28 showed a significant increase (p-value < 0.05) in T5α-ol titer compared to the parent strain (Fig. 6). In addition, KMRJ strains could not form a clear T5α-ol peak in the GC chromatogram when they were used in glucose-containing media. The reason behind it was probably the limited agitation and air transfer in deep-well plates because of the shape of the wells and the shaking frequency (350 rpm) as it was proven that oxygen supply is critical for oxygenation reactions for T5α-ol synthesis [23]. Therefore, in the second-level screening, the BioLector Pro microbioreactor-screening platform was utilized with a flower-shaped plate and a higher shaking frequency of 1000 rpm to ensure sufficient oxygen transfer since the flower shape geometry can provide a better oxygen supply [73].

To this end, we primarily considered taxadiene yields, as it is the direct product derived from GGPP. In contrast, T5α-ol titers were significantly lower, reaching a maximum of only 8 mg/L compared to taxadiene, which reached a maximum of 95 mg/L. For the galactose-containing medium, the three best taxadiene producer strains containing single modifications (KM21, KM25 and KM13) and four strains containing multiple modifications (KM31, KM32, KM33 and KM34) were screened while biomass, pH and DO were monitored in real-time using the BioLector. Interestingly, all multiple modification strains resulted in higher biomass compared to the parent strain, KM1, after three days. It seems increasing fluxes in particular reactions and pathways ended in favor of biomass rather than a metabolic burden on the cell although overexpression of *ILV2* (KM21) or deletion of *MDH3* (KM13) led to lower biomass compared to the parent strain KM1. Most of the strains showed a similar pattern for pH change during three days of cultivation. The initial pH between 5.50 and 5.75 dropped to ~ 5.0 and slightly increased to around 5.25, as shown in Fig. 7B. *S. cerevisiae* tends to acidify the culture pH [74], therefore

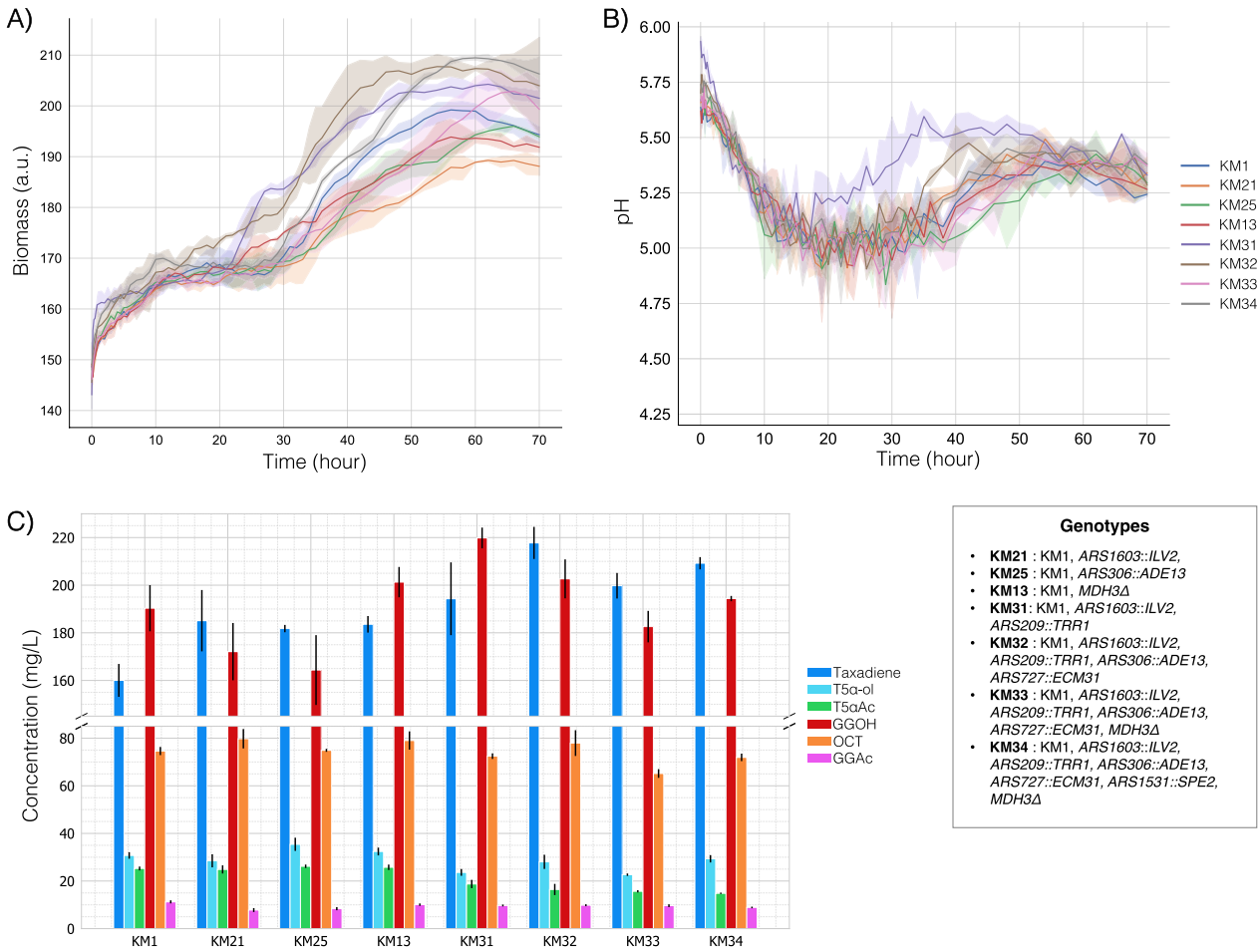

Fig. 7 Performances of the selected KM1-derived strains in galactose-containing CSM in the high-throughput microbioreactor system. **A** Biomass trend **B** Change in culture pH **C** Production of the target molecules and side-products Error bars represent the standard deviation of three replicates

it was an expected behavior. However, KM31 containing overexpressed *ILV2* and *TRR1* genes showed a higher pH between 20 and 50 h of culture. To further elucidate the reason behind this, the pH pattern of KM22 cultures should also be tested.

Although KM21, KM25 and KM13 strains increased taxadiene production by 35%–40% in the deep well-plates compared to KM1 (Fig. 6A), taxadiene concentration was only enhanced by ~ 15% (p-value < 0.05) by these strains in the BioLector (Fig. 7C). However, the highest taxadiene concentration was ~ 110 mg/L in deep-well plates and it was doubled in the BioLector system as ~ 215 mg/L production was recorded with KM32. This is a remarkable improvement as approximately 1.6-fold higher production was achieved than the previously reported maximum, which was 137 mg/L in *S. cerevisiae* cell factories [23]. It is worth noting that previous studies on taxadiene synthesis used richer media, which likely contributed to higher production yields. In this study, complete synthetic defined media were used, as all in silico illustrations were conducted using defined media. While *E. coli* achieved higher taxadiene production at 1 g/L in a previous attempt [75], the expression of the next step enzymes, T5αOH and CPR, completely abolished taxadiene synthesis in *E. coli* [75]. In contrast, KM32 increased taxadiene production while also expressing the downstream genes responsible for the second and third steps, namely T5αOH, CPR, and TAT.

Dissolved oxygen (DO) was maintained between 75%–100% for all the strains used in the BioLector system (Additional file 1: Figure S12 and S13). Accordingly, greater T5α-ol was synthesized by all KM1-derived strains with at least 23 mg/L concentration in KM33 (Fig. 7C). This production was almost three-fold higher than the maximum T5α-ol production observed in deep-well plates, proving that oxygen supply was the critical factor for T5α-ol production. Also, the acetylated precursor, T5αAc, was detected in all KM1-derived strains with concentrations ranging from 15 to 26 mg/L. Nevertheless, all multi-modification strains KM31, KM32, KM33 and KM34, produced less T5αAc (p-value < 0.05) than the parent strain as shown in Fig. 7C even though a significant increase (p-value < 0.05) was observed in taxadiene production in KM32, KM33 and KM34 and single-modification strains (KM21, KM25, KM13). This indicates that to enhance the yield of T5α-ol and the following compound, T5αAc, further improvement in the second reaction catalyzed by T5αOH and CPR is a necessity. KM25 was the best producer of T5α-ol and T5αAc with 35.4 mg/L and 26.2 mg/L, respectively. Yet, no statistically significant difference (p-value > 0.05) was observed for T5αAc titer between KM1, KM21, KM25 and KM13. In addition to taxadiene, these are the highest yields for T5α-ol and T5αAc production reported so far using *S. cerevisiae* as a cell factory. It is noteworthy that the previous study conducted in *E. coli* achieved a T5α-ol production of 23 mg/L without the expression of the subsequent gene, TAT, in the pathway [75]. In addition, this *E. coli* strain was unable to accumulate taxadiene, as mentioned earlier. Taking these findings into account, it is anticipated that further enhancements will lead to an increased metabolic flux towards T5αAc in our *S. cerevisiae* strains.

In the glucose-containing CSM, KMRJ12, KMRJ19 and KMRJ24 strains were used with multi-modification strains KMRJ31, KMRJ32, KMRJ33, and KMRJ34 (Table 2) in the BioLector platform. In parallel to the deep-well plates, all strains with the *OAR1Δ* genotype produced less biomass in the BioLector system as shown in Fig. 8A. A correlation was also observed between pH change and biomass and/or *OAR1Δ* (Fig. 8B). Like in the galactose-containing medium, after a sharp drop, the pH reached around 5.25 for the strains expressing the *OAR1* gene (EJ1, KMRJ19, KMRJ24). The pH stabilised at around 5.0 for the other strains, this might be because of lower biomass or the *OAR1* deletion.

Another interesting finding with *OAR1* deletion was the dramatic reductions in the titers of the oxygenated product, T5α-ol, and naturally, the next product T5αAc as seen in Fig. 8C. Also, GGOH production dropped at least twofold in all *OAR1*-deleted strains. GGOH is indeed a valuable compound that can be used in the pharmaceutical and cosmetic industries. Metabolic engineering studies have been carried out to improve GGOH production in *S. cerevisiae* [76, 77]. Although the GGOH synthesis is not fully identified in yeast, native phosphatases are thought to be involved in GGOH synthesis from GGPP. Considering its chemical structure, oxygenation reactions should also take a part in GGOH synthesis [78]. Taking the reductions in T5α-ol and GGOH into consideration, it is likely that *OAR1* deletion affected the oxygenation capability of *S. cerevisiae,* and it might also be one of the reasons behind the decrease in cell fitness. On the other hand, still, the strains with the *OAR1Δ* genotype showed greater taxadiene production than the reference strain (Fig. 8C), meaning that *OAR1* deletion led to the highest taxadiene production per OD. However, it should be also noted that GGOH reduction might have improved the taxadiene concentration.

The combination of *OAR1Δ* and *DTR1Δ* in KMRJ31 resulted in the maximum taxadiene production of 128 mg/L, which is still a remarkable titer for taxadiene considering the similar studies reported previously [22, 79]. On the other hand, even if *OAR1Δ* genotypes were excluded, KMRJ strains produced ~ 7 mg/L of T5α-ol from each 100 mg/L taxadiene on average (7% yield), whereas this amount was ~ 18 mg/L for KM strains (18%

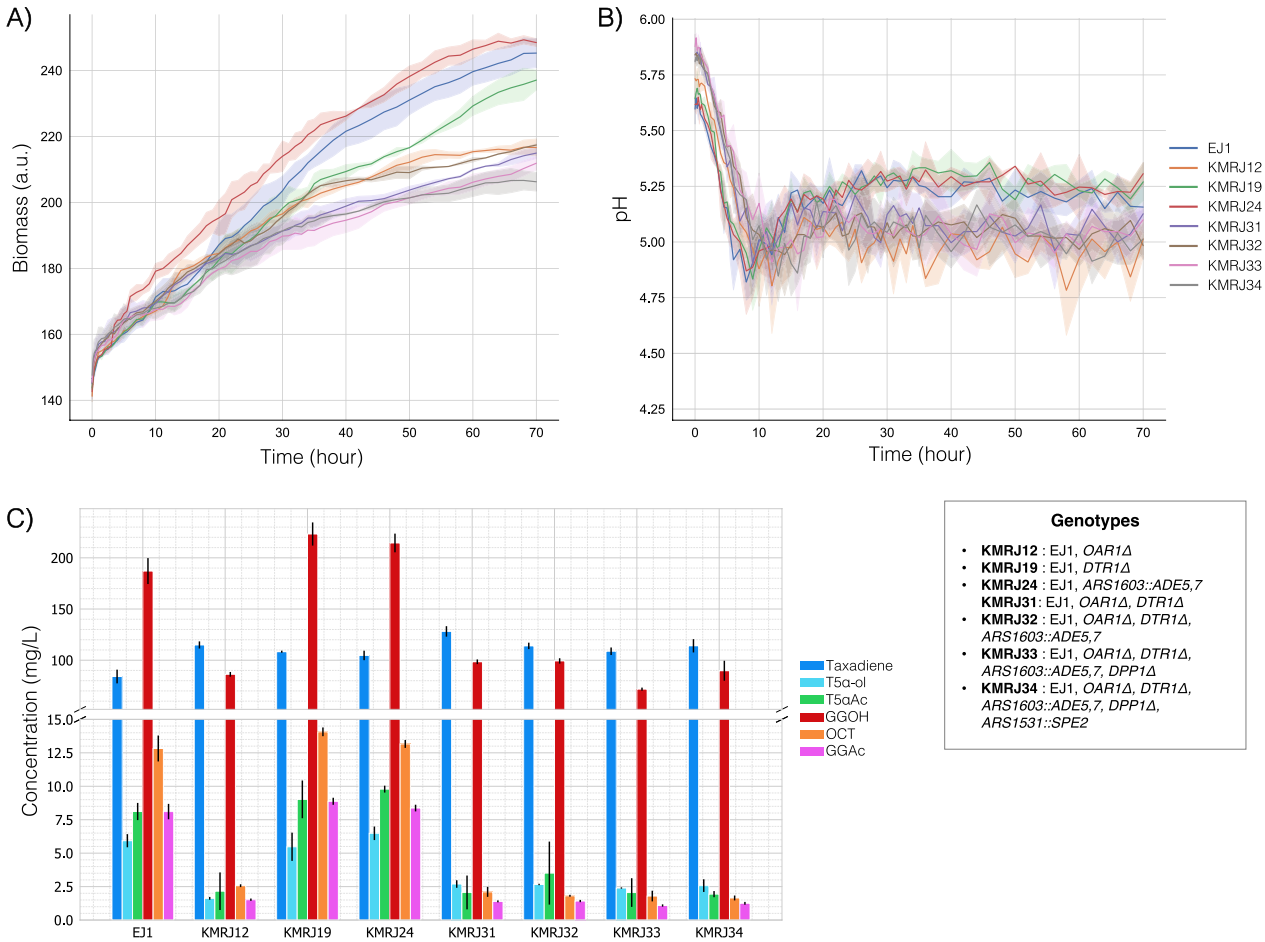

Fig. 8 Performances of the selected EJ1-derived strains in glucose-containing CSM in the high-throughput microbioreactor system. **A** Biomass trend **B** Change in culture pH **C** Production of the target molecules and side-products Error bars represent the standard deviation of three replicates

yield). This is probably because the galactose-inducible promoters could not express T5αOH in glucose-containing CSM as efficiently as in galactose-containing CSM, even though the *GAL80* was deleted.

In addition to Taxol® precursors, KMRJ19 with *DTR1* deletion produced the highest GGOH with 221 mg/L among all the strains used in this study, it was followed by KMRJ24 and KM31in the galactose-containing medium with the 219 mg/L (p-value > 0.05). These production yields of GGOH are comparable with those in similar studies [76, 77]. For this reason, the related modifications, *DTR1* deletion, *ADE 5,7* overexpression or simultaneous overexpression of *ILV2* and *TRR1*, are promising genomic modifications to increase GGOH production.

While we have tested three single modifications in the Biolector platform, there are several other promising candidates, such as KM28, KM26, and KM23, which warrant further investigation in well-controlled conditions. Specifically, understanding the impact of overexpressing the *SPE2, ECM31,* and *ADE4* genes is crucial to gain deeper insights into their potential contributions to the overall process. Therefore, future experiments focusing on these candidates could provide valuable data to complement and enrich our current findings.

Scale-up of the production of the Taxol® precursors using the best strain

Finally, the production was scaled-up using the best taxadiene producer, KM32, in a 2% (w/v) galactose-containing CSM medium in 250 mL reaction volume in a mini-scale bioreactor. To maintain similar conditions as in the BioLector system and to mimic the in silico simulations where the oxygen was not a limiting factor, the minimum %DO was set to 75% and O_2 was supplied with air when it was below the threshold. Yet, it is important to note that in an industrial-scale setting, the availability of oxygen can potentially act as a limiting factor that may impact production yields. Likewise, the pH was set to 6.0 as it is a suitable pH considering both yeast growth and the enzymatic activities in the Taxol® pathway [23].

A 1 M NaOH was added as needed when the pH was less than 6.0. Figure 9 shows the reactor parameters and the production of Taxol® precursors during the five days of cultivation.

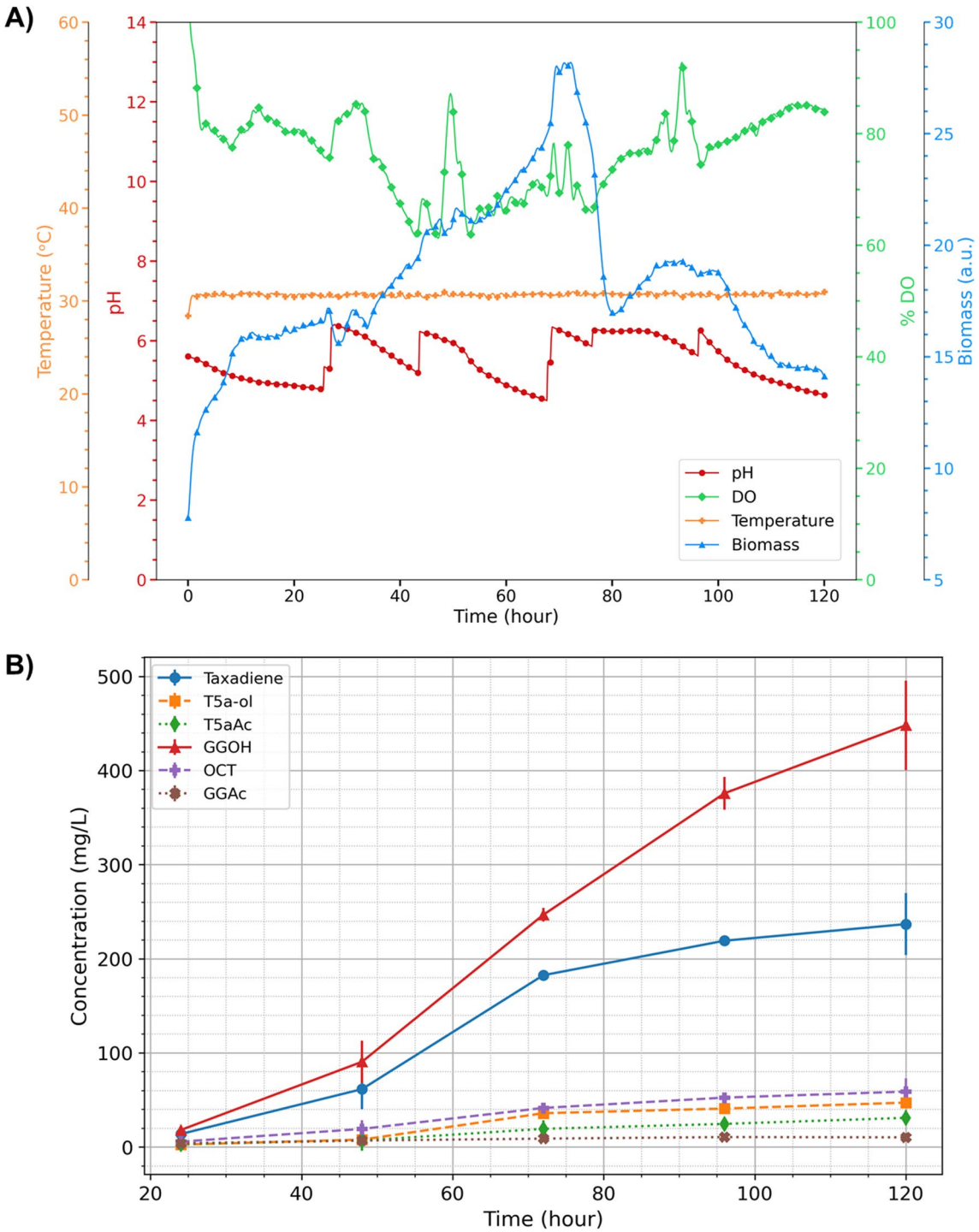

Fig. 9 **A** The trends of the bioreactor parameters in KM32 culture **B** Metabolite concentrations during five days of cultivation. Error bars represent the standard deviation of two measurements taken each time. Total working volume was set to 250 mL: 200 mL of the culture and 50 mL of the dodecane layer

Considering the first three days, the growth showed a similar pattern as a decrease was observed towards the third day and then dramatically dropped down (Fig. 9A). Although this decrease was expected, the sharp drop between 72nd and 80th hours was probably a temporary a sensor problem. On the other hand, the production yield was relatively low for the Taxol® precursors in the first three days. In 72 h, 182.5 mg/L taxadiene, 35.8 mg/L of T5α-ol and 19.2 mg/L of T5αAc were produced. These reached to 236.7, 47.1 and 31.2 mg/L, respectively on the fifth day (Fig. 9B). Although the total titers of the target molecules were higher at the end of cultivation, on the fifth day, the yield on the third day was lower compared to the BioLector system.

In contrast, the concentration of GGOH was significantly higher than that of taxadiene (Fig. 9B). The accumulation of GGOH is primarily attributed to the promiscuous activity of TASY enzymes, as previously reported [19]. While adjusting the media composition and conditions can reduce the synthesis of this undesired side product [19], it appears necessary to engineer the TASY enzymes to enhance their specificity in order to mitigate GGOH accumulation and improve taxadiene titer.

pH and DO showed more fluctuations in the bioreactor-scale. This might be the reason behind the decrease in production yield. Consequently, optimising the bioprocess conditions and implementing better pH and DO control measures may further enhance the titers of Taxol® precursors on larger scales.

Comparison of the in silico predictions and wet-lab results

The genome-scale model utilised in this study, yeast 8.5.0, represents a significant advancement in *S. cerevisiae* S288C's metabolic understanding [27]. While it is important to acknowledge that certain discrepancies may arise due to the specific genotype of the parent strain employed in our research, we observed enhancements in taxadiene production, and by extension, the likely increase in GGPP production, meaning that our wet-lab validations clearly demonstrated the potency of mathematical models and computational frameworks in identifying valuable targets that may elude intuitive prediction. Considering Table 1, Fig. 4, and the experimental results, however, there are also exceptions. *OAR1* deletion was predicted to increase GGPP concentration in the galactose-containing CSM, and *MDH3* deletion was suggested to increase acetyl-CoA production in the cytosol in the glucose-containing medium. Nevertheless, these deletions showed higher productions in the opposite carbon source (Fig. 6). According to the in silico predictions, *MLS1* deletion should have increased the acetyl-CoA production in the galactose-containing medium, but in contrast, this deletion decreased GGPP concentration and probably acetyl-CoA concentration in the galactose-containing medium and could not make any difference in glucose-containing CSM (Fig. 6). The other deletions showed either predicted results with overproductions or similar production yields to the reference strains (p-value > 0.05). It is also possible that more replicates could have given more precise results and could have mitigated the analytic measurement deviations.

Based on these findings, the in silico predictions provided a refined set of solutions, taking into account the comprehensive genomic context of the host organism. We further refined our selection process by detecting the possible maximum fluxes in conjunction with metabolite interaction maps, enabling us to prioritize single genomic modifications from an initial pool of over 60 predictions (Additional file 1: Table S5). It is worth to notion that the three algorithms are able to predict *sets* of genomic modifications that together have a beneficial effect. Both in silico and experimental validation of *single* genomic modifications from these sets can therefore lead to somewhat false negatives, as some of these modifications are only beneficial in combination with others. Regardless, this strategy was designed to match our desire to experimentally validate single genomic modifications, and this prioritization strategy indeed resulted in a strain with improved taxadiene productions.

Conclusion

Computational tools, in silico design frameworks and genome-scale models were used to rationally design yeast cell factories in a way that intuitive estimations are very difficult. In specific, in this study, a combinatorial design approach was used with three in silico algorithms, OptKnock, OptGene and OptForce, on the yeast-GEM 8.5.0. A set of 17 genomic modifications were predicted by the simulations (nine gene knockouts, eight gene overexpressions and their combinations via extra copies and strong promoters) and tested for the increased production of Taxol® precursors, through GGPP overproduction. Therefore, this study is one of the most comprehensive studies reported until now in terms of the amount and the variability of the genomic manipulations. Without a doubt, the ACtivE toolkit and method facilitated this process with an accelerated genome editing process. The findings of this work showed that simultaneous overexpression of four genes, *ILV2, TRR1, ADE13* and *ECM31,* related to upregulations in the pantothenate pathway, branched-amino acid biosynthesis pathways and de novo purine synthesis pathway, could enhance taxadiene production by ~ 50% through GGPP overproduction. Also, single

modifications could increase taxadiene yield from 15 to 40% depending on the cultivation conditions and carbon source. As the oxygen supply is crucial for the second step in the Taxol® biosynthesis pathway, higher yields in T5α-ol and T5αAc productions were reported with the conditions allowing better oxygen access. Using the best-performing strain KM32 containing additional copies of the above-mentioned four genes, we achieved 215 mg/L of taxadiene, 43.65 mg/L of T5α-ol and 26.2 mg/L of T5αAc titers that are the highest production yields reported until now in *S. cerevisiae.* The genomic modifications reporting higher GGPP synthesis can be also used to increase the production of other high-value isoprenoids through the same precursor. In addition to the mevalonate pathway, similar integrated approaches combining different design algorithms, genome engineering and bioprocessing studies can be used for enhancing metabolic fluxes towards the target native or heterologous pathways in *S. cerevisiae.* Developing alternative in silico prediction and design tools, constructing more accurate genome-scale models would improve the efficiency of this process.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12934-023-02251-7.

Additional file 1: Table S1. The primer list used to produce the donor DNA parts for the gene deletions. **Table S2.** crRNA sequences used to target the corresponding genes for deletions. **Table S3.** The primer list used to produce the donor DNA parts for the genomic integrations. **Table S4.** The yeast strains used in the study. **Table S5.** The gene candidates predicted by the design algorithms. **Table S6.** The components of the yeast nitrogen base used in the study. **Table S7.** The components of the complete supplement mixture used in the study. **Figure S1.** The ACtivE method for gene deletion. **Figure S2.** The ACtivE method for genomic integration. **Figure S3.** Acetyl-CoA-centred sub-metabolite interaction map of the wild-type model. **Figure S4.** Acetyl-CoA-centred sub-metabolite interaction map of *DTR1* deleted model. **Figure S5.** Acetyl-CoA-centred sub-metabolite interaction map of *ACH1* deleted model. **Figure S6.** Acetyl-CoA-centred sub-metabolite interaction map of *MLS1* deleted model. **Figure S7.** Acetyl-CoA-centred sub-metabolite interaction map of *DIT1* deleted model. **Figure S8.** Acetyl-CoA-centred sub-metabolite interaction map of *LPP1* deleted model. **Figure S9.** Acetyl-CoA-centred sub-metabolite interaction map of *ISC1* deleted model. **Figure S10.** de novo purine nucleotide biosynthesis pathway. **Figure S11.** Phosphopantothenate biosynthesis pathway. **Figure S12.** Dissolved oxygen (DO) concentrations of three-day cultures of KM1-derived strains measured by the BioLector microbioreactor system in the galactose-containing CSM. **Figure S13.** Dissolved oxygen (DO) concentrations of three-day cultures of EJ1-derived strains measured by the BioLector microbioreactor system in the glucose-containing CSM. **Figure S14.** Gas chromatography shows the compounds' peaks produced by KM32 and the mass spectrum of taxadiene. **Figure S15.** Gas chromatography shows the compounds' peaks produced by KM32 and the mass spectrum of T5α-ol. The retention time of the T5α-ol peak was at 8.20th minutes. **Figure S16.** Gas chromatography shows the compounds' peaks produced by KM32 and the mass spectrum of T5αAc. The retention time of the T5αAc peak was at 8.61st minutes.

Acknowledgements
We would like to thank Mr Stuart Martin, Mr Mark Lauchlan and Miss Katalin Kis at the School of Engineering, University of Edinburgh, UK for their technical support with GC-MS analysis and laboratory operations. We would like to thank Dr David Gomez-Cabeza at the School of Engineering, University of Edinburgh, UK for their kind assistance with the inverted microscope.
This work was supported by the YLSY program of the Ministry of National Education of Turkey, the Mexican government's dependence CONACyT (Mexican National Council for Science and Technology) CVU:675492 and CVU: 537962, Novo Nordisk Foundation (Grant Number NNF20CC0035580), the Engineering and Physical Sciences Research Council (Grant number EP/R513209/1), the Royal Society (Grant Number RSG\R1\180345) and the British Council (Grant Number: 527429894).

Author contributions
KM: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Writing—Original Draft. RS: Methodology, Software, Writing—Reviewing and Editing. NJ-B: Investigation, Resources. JS-G: Investigation. EJK: Supervision, Writing—Reviewing and Editing. LR-S: Conceptualization, Supervision, Writing—Reviewing and Editing, Project administration.

Availability of data and materials
MATLAB scripts are available on https://github.com/kmalci/yeast_modelling. Plasmids and other data will be made available on request.

Declarations

Competing interests
Authors declare that they have no competing interests.

Author details
[1]Institute for Bioengineering, School of Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3BF, UK. [2]Centre for Engineering Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3BF, UK. [3]Department of Pediatrics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0760, USA. [4]Department of Life Sciences, Chalmers University of Technology, Kemivägen 10, SE-412 96, Gothenburg, Sweden. [5]SciLifeLab, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden. [6]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs, Lyngby, Denmark. [7]School of Natural and Environmental Sciences, Molecular Biology and Biotechnology Division, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK. [8]Department of Biochemical Engineering, The Advanced Centre for Biochemical Engineering, University College London, Gower Street, London WC1E 6BT, UK. [9]Present Address: Department of Bioengineering, Imperial College London, London SW7 2AZ, UK.

Received: 15 June 2023 Accepted: 14 November 2023
Published online: 29 November 2023

References

1. Wong J, Rios-Solis L, Keasling JD. Microbial production of isoprenoids. consequences microb interact with hydrocarb oils. Lipids Prod Fuels Chem. 2017. https://doi.org/10.1007/978-3-319-50436-0_219.
2. Walls LE, Otoupal P, Ledesma-Amaro R, et al. Bioconversion of cellulose into bisabolene using *Ruminococcus flavefaciens* and *Rhodosporidium toruloides*. Bioresour Technol. 2023;368:128216. https://doi.org/10.1016/J.BIORTECH.2022.128216.
3. Malcı K, Kurt-Gür G, Tamerler C, Yazgan-Karatas A. Combinatorial decolorization performance of Pycnoporus sanguineus MUCL 38531 sourced recombinant laccase/mediator systems on toxic textile dyes. Int J Environ Sci Technol. 2022. https://doi.org/10.1007/S13762-022-04080-4.
4. Jung YK, Kim TY, Park SJ, Lee SY. Metabolic engineering of *Escherichia coli* for the production of polylactic acid and its copolymers. Biotechnol Bioeng. 2010;105:161–71. https://doi.org/10.1002/BIT.22548.

5. Malcı K, Watts E, Roberts TM, et al. Standardization of synthetic biology tools and assembly methods for *Saccharomyces cerevisiae* and emerging yeast species. ACS Synth Biol. 2022;11:2527–47. https://doi.org/10.1021/ACSSYNBIO.
6. Gómez-Pastor R, Pérez-Torrado R, Garre E, Matallana E. Recent advances in yeast biomass production. Biomass Detect Prod Usage. 2011. https://doi.org/10.5772/19458.
7. Teworte S, Malcı K, Walls LE, et al. Recent advances in fed-batch microscale bioreactor design. Biotechnol Adv. 2022;55:107888. https://doi.org/10.1016/J.BIOTECHADV.2021.107888.
8. Jensen MK, Keasling JD. Recent applications of synthetic biology tools for yeast metabolic engineering. FEMS Yeast Res. 2014. https://doi.org/10.1111/1567-1364.12185.
9. Malcı K, Walls LE, Rios-Solis L. Multiplex genome engineering methods for yeast cell factory development. Front Bioeng Biotechnol. 2020;8:1264. https://doi.org/10.3389/FBIOE.2020.589468.
10. Wong J, de Rond T, d'Espaux L, et al. High-titer production of lathyrane diterpenoids from sugar by engineered Saccharomyces cerevisiae. Metab Eng. 2018;45:142–8. https://doi.org/10.1016/J.YMBEN.2017.12.007.
11. Kulagina N, Besseau S, Godon C, et al. Yeasts as biopharmaceutical production platforms. Front Fungal Biol. 2021. https://doi.org/10.3389/FFUNB.2021.733492.
12. Cao X, Yu W, Chen Y, et al. Engineering yeast for high-level production of diterpenoid sclareol. Metab Eng. 2023;75:19–28. https://doi.org/10.1016/J.YMBEN.2022.11.002.
13. Zhou YJ, Gao W, Rong Q, et al. Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. J Am Chem Soc. 2012;134:3234–41. https://doi.org/10.1021/JA2114486
14. Zhang C, Ju H, Lu CZ, et al. High-titer production of 13R-manoyl oxide in metabolically engineered Saccharomyces cerevisiae. Microb Cell Fact. 2019;18:1–9. https://doi.org/10.1186/S12934-019-1123-Z
15. McElroy C, Jennewein S. Taxol® biosynthesis and production: from forests to fermenters. Biotechnol Nat Prod. 2017. https://doi.org/10.1007/978-3-319-67903-7_7
16. Nowrouzi B, Lungang L, Rios-Solis L. Exploring optimal taxol® CYP725A4 activity in *Saccharomyces cerevisiae*. Microb Cell Fact. 2022;21:1–24. https://doi.org/10.1186/S12934-022-01922-1
17. Walls LE, Martinez JL, del Rio Chanona EA, Rios-Solis L. Definitive screening accelerates taxol biosynthetic pathway optimization and scale up in Saccharomyces cerevisiae cell factories. Biotechnol J. 2022. https://doi.org/10.1002/BIOT.202100414.
18. Walls LE, Martinez JL, Rios-Solis L. Enhancing *Saccharomyces cerevisiae* taxane biosynthesis and overcoming nutritional stress-induced pseudohyphal Growth. Microorganisms. 2022. https://doi.org/10.3390/MICROORGANISMS10010163/S1.
19. Santoyo-Garcia JH, Walls LE, Valdivia-Cabrera M, et al. The synergetic effect from the combination of different adsorption resins in batch and semi-continuous cultivations of *S. cerevisiae* cell factories to produce acetylated Taxanes precursors of the anticancer drug Taxol. Biotechnol Bioeng. 2023. https://doi.org/10.1002/bit.28487
20. Nowrouzi B, Torres-Montero P, Kerkhoven EJ, et al. Rewiring Saccharomyces cerevisiae metabolism for optimised Taxol® precursors production. Biorxiv. 2023. https://doi.org/10.1101/2023.06.03.543533.
21. Köksal M, Jin Y, Coates RM, et al. (2010) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. Nat. 2010;4697328(469):116–20. https://doi.org/10.1038/nature09628.
22. Nowrouzi B, Li RA, Walls LE, et al. (2020) Enhanced production of taxadiene in *Saccharomyces cerevisiae*. Microb Cell Factories. 2020;191(19):1–12. https://doi.org/10.1186/S12934-020-01458-2.
23. Walls LE, Malcı K, Nowrouzi B, et al. Optimizing the biosynthesis of oxygenated and acetylated Taxol precursors in Saccharomyces cerevisiae using advanced bioprocessing strategies. Biotechnol Bioeng. 2020;118:279–93. https://doi.org/10.1002/BIT.27569.
24. Gu C, Kim GB, Kim WJ, et al. Current status and applications of genome-scale metabolic models. Genome Biol. 2019. https://doi.org/10.1186/S13059-019-1730-3.
25. Lopes H, Rocha I. Genome-scale modeling of yeast: chronology, applications and critical perspectives. FEMS Yeast Res. 2017;17:50. https://doi.org/10.1093/FEMSYR/FOX050.
26. Oftadeh O, Salvy P, Masid M, et al. (2021) A genome-scale metabolic model of *Saccharomyces cerevisiae* that integrates expression constraints and reaction thermodynamics. Nat Commun. 2021;121(12):1–10. https://doi.org/10.1038/s41467-021-25158-6.
27. Lu H, Li F, Sánchez BJ, et al. A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. Nat Commun. 2019. https://doi.org/10.1038/s41467-019-11581-3.
28. Sánchez B, Kerkhoven E, et al. SysBioChalmers/yeast-GEM: yeast. 2021. https://doi.org/10.5281/ZENODO.5062615.
29. Chen Y, Li F, Nielsen J. Genome-scale modeling of yeast metabolism: retrospectives and perspectives. FEMS Yeast Res. 2022;22:1–9. https://doi.org/10.1093/FEMSYR/FOAC003.
30. Palsson BØ. Systems biology constraint-based reconstruction and analysis. Cambridge University Press; 2015.
31. O'Brien EJ, Monk JM, Palsson BO. Using genome-scale models to predict biological capabilities. Cell. 2015;161:971. https://doi.org/10.1016/J.CELL.2015.05.019.
32. Heirendt L, Arreckx S, Pfau T, et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat Protoc. 2019;14:639–702. https://doi.org/10.1038/s41596-018-0098-2.
33. Fernández-Castané A, Fehér T, Carbonell P, et al. Computer-aided design for metabolic engineering. J Biotechnol. 2014. https://doi.org/10.1016/J.JBIOTEC.2014.03.029.
34. Burgard AP, Pharkya P, Maranas CD. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioeng. 2003;84:647–57. https://doi.org/10.1002/BIT.10803.
35. Patil KR, Rocha I, Förster J, Nielsen J. Evolutionary programming as a platform for in silico metabolic engineering. BMC Bioinformatics. 2005;6:1–12. https://doi.org/10.1186/1471-2105-6-308
36. Yang X-S. Genetic algorithms. Nature-Inspired Optim Algorithms. 2021. https://doi.org/10.1016/B978-0-12-821986-7.00013-5.
37. Ranganathan S, Suthers PF, Maranas CD. OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLOS Comput Biol. 2010;6: e1000744. https://doi.org/10.1371/JOURNAL.PCBI.1000744.
38. Malcı K, Jonguitud-Borrego N, Van Der Straten WH, et al. ACtivE: assembly and crispr-targeted in vivo editing for yeast genome engineering using minimum reagents and time. ACS Synth Biol. 2022;2022:3629–43. https://doi.org/10.1021/acssynbio.2c00175.
39. openCOBRA Home Page—The COBRA Toolbox. https://opencobra.github.io/cobratoolbox/stable/index.html. Accessed 26 Feb 2023
40. King ZA, Dräger A, Ebrahim A, et al. Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways. PLOS Comput Biol. 2015;11:e1004321. https://doi.org/10.1371/JOURNAL.PCBI.1004321.
41. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007;2:31–4. https://doi.org/10.1038/nprot.2007.13.
42. Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 2018;46:W242–5. https://doi.org/10.1093/NAR/GKY354.
43. Toyomasu T, Sassa T. Diterpenes. Compr Nat Prod II Chem Biol. 2010;1:643–72. https://doi.org/10.1016/B978-008045382-8.00006-X.
44. Orth JD, Thiele I. Palsson BO (2010) What is flux balance analysis? Nat Biotechnol. 2010;283(28):245–8. https://doi.org/10.1038/nbt.1614.
45. Kocabaş K, Arif A, Uddin R, Çakır T. Dual transcriptome based reconstruction of Salmonella-human integrated metabolic network to screen potential drug targets. PLoS ONE. 2022. https://doi.org/10.1371/JOURNAL.PONE.0268889.
46. openCOBRA - createMetIntrcNetwork COBRA.tutorials/visualization/createMetIntrcNetworkTutorial at master · opencobra/COBRA.tutorials · GitHub. https://github.com/opencobra/COBRA.tutorials/tree/master/visualization/createMetIntrcNetworkTutorial. Accessed 26 Feb 2023
47. Yoon SH, Lee SH, Das A, et al. Combinatorial expression of bacterial whole mevalonate pathway for the production of beta-carotene in *E. coli*. J Biotechnol. 2009;140:218–26. https://doi.org/10.1016/J.JBIOTEC.2009.01.008.

48. Pilauri V, Bewley M, Diep C, Hopper J. Gal80 dimerization and the yeast GAL gene switch. Genetics. 2005;169:1903. https://doi.org/10.1534/GENETICS.104.036723.
49. Lee ME, DeLoache WC, Cervantes B, Dueber JE. A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth Biol. 2015;4:975–86. https://doi.org/10.1021/sb500366v.
50. George CM. Discoveries of the phosphatidate phosphatase genes in yeast published in the journal of biological chemistry. J Biol Chem. 2019. https://doi.org/10.1074/jbc.TM118.004159.
51. Schneider R, Brors B, Bürger F, et al. (1997) Two genes of the putative mitochondrial fatty acid synthase in the genome of *Saccharomyces cerevisiae*. Curr Genet. 1997;326(32):384–8. https://doi.org/10.1007/S002940050292.
52. Venkatesan R, Sah-Teli SK, Awoniyi LO, et al. (2014) Insights into mitochondrial fatty acid synthesis from the structure of heterotetrameric 3-ketoacyl-ACP reductase/3R-hydroxyacyl-CoA dehydrogenase. Nat Commun. 2014;51(5):1–12. https://doi.org/10.1038/ncomms5805.
53. Steffan JS, McAlister-Henn L. Isolation and characterization of the yeast gene encoding the MDH3 isozyme of malate dehydrogenase. J Biol Chem. 1992;267:24708–15. https://doi.org/10.1016/S0021-9258(18)35822-8.
54. Kunze M, Pracharoenwattana I, Smith SM, Hartig A. A central role for the peroxisomal membrane in glyoxylate cycle function. Biochim Biophys Acta - Mol Cell Res. 2006;1763:1441–52. https://doi.org/10.1016/J.BBAMCR.2006.09.009.
55. Rottensteiner H, Theodoulou FL. The ins and outs of peroxisomes: co-ordination of membrane transport and peroxisomal metabolism. Biochim Biophys Acta - Mol Cell Res. 2006;1763:1527–40. https://doi.org/10.1016/J.BBAMCR.2006.08.012.
56. Kunze M, Hartig A. Permeability of the peroxisomal membrane: lessons from the glyoxylate cycle. Front Physiol. 2013. https://doi.org/10.3389/FPHYS.2013.00204.
57. Chen Y, Siewers V, Nielsen J. Profiling of cytosolic and peroxisomal Acetyl-CoA metabolism in *Saccharomyces cerevisiae*. PLoS ONE. 2012;7: e42475. https://doi.org/10.1371/JOURNAL.PONE.0042475.
58. Chen Y, Zhang Y, Siewers V, Nielsen J. Ach1 is involved in shuttling mitochondrial acetyl units for cytosolic C2 provision in Saccharomyces cerevisiae lacking pyruvate decarboxylase. FEMS Yeast Res. 2015;15:1–8. https://doi.org/10.1093/FEMSYR/FOV015.
59. Friesen H, Hepworth SR, Segall J. An Ssn6-Tup1-dependent negative regulatory element controls sporulation-specific expression of DIT1 and DIT2 in Saccharomyces cerevisiae. Mol Cell Biol. 1997;17:123–34. https://doi.org/10.1128/MCB.17.1.123.
60. Felder T, Bogengruber E, Tenreiro S, et al. Dtr1p, a multidrug resistance transporter of the major facilitator superfamily, plays an essential role in spore wall maturation in Saccharomyces cerevisiae. Eukaryot Cell. 2002;1:799–810. https://doi.org/10.1128/EC.1.5.799-810.
61. Briza P, Eckerstorfert M, Breitenbach M. The sporulation-specific enzymes encoded by the DITl and DIT2 genes catalyze a two-step reaction leading to a soluble LL-dityrosine-containing precursor of the yeast spore wall (*Saccharomyces cereviuiae*/epimerlzation). Genetics. 1994;91:4524–8.
62. Basiony M, Yang Y, Liu G, et al. Studies on the properties of the sporulation specific protein dit1 and its product formyl tyrosine. J Fungi. 2020;6:1–15. https://doi.org/10.3390/JOF6020077.
63. Lin CPC, Kim C, Smith SO, Neiman AM. A highly redundant gene network controls assembly of the outer spore wall in *S cerevisiae*. PLoS Genet. 2013;9:1003700. https://doi.org/10.1371/JOURNAL.PGEN.1003700.
64. Ye Y, Fujii M, Hirata A, et al. Geranylgeranyl diphosphate synthase in fission yeast is a heteromer of farnesyl diphosphate synthase (FPS), Fps1, and an FPS-like protein, Spo9, essential for sporulation. Mol Biol Cell. 2007;18:3568–81. https://doi.org/10.1091/MBC.E07-02-0112
65. Matmati N, Hannun YA. Thematic review series: sphingolipids ISC1 (inositol phosphosphingolipid-phospholipase C), the yeast homologue of neutral sphingomyelinases. J Lipid Res. 2008;49:922. https://doi.org/10.1194/JLR.R800004-JLR200.
66. Matmati N, Hassan BH, Ren J, et al. Yeast sphingolipid phospholipase gene ISC1 regulates the spindle checkpoint by a CDC55-dependent mechanism. Mol Cell Biol. 2020. https://doi.org/10.1128/MCB.00340-19.
67. Brat D, Weber C, Lorenzen W, et al. Cytosolic re-localization and optimization of valine synthesis and catabolism enables inseased isobutanol production with the yeast *Saccharomyces cerevisiae*. Biotechnol Biofuels. 2012;5:1–16. https://doi.org/10.1186/1754-6834-5-65
68. Olzhausen J, Schübbe S, Schüller HJ. Genetic analysis of coenzyme A biosynthesis in the yeast *Saccharomyces cerevisiae*: Identification of a conditional mutation in the pantothenate kinase gene CAB1. Curr Genet. 2009;55:163–73. https://doi.org/10.1007/S00294-009-0234-1
69. Balasundaram D, Dinman JD, Tabor CW, Tabor H. SPE1 and SPE2: two essential genes in the biosynthesis of polyamines that modulate +1 ribosomal frameshifting in *Saccharomyces cerevisiae*. J Bacteriol. 1994;176:7126. https://doi.org/10.1128/JB.176.22.7126-7128.1994.
70. Rébora K, Laloo B, Daignan-Fornier B. Revisiting purine-histidine cross-pathway regulation in *Saccharomyces cerevisiae*: a central role for a small molecule. Genetics. 2005;170:61. https://doi.org/10.1534/GENETICS.104.039396.
71. Trotter EW, Grant CM. Thioredoxins are required for protection against a reductive stress in the yeast *Saccharomyces cerevisiae*. Mol Microbiol. 2002;46:869–78. https://doi.org/10.1046/J.1365-2958.2002.03216.X.
72. Olzhausen J, Grigat M, Seifert L, et al. Increased biosynthesis of acetyl-CoA in the yeast *Saccharomyces cerevisiae* by overexpression of a deregulated pantothenate kinase gene and engineering of the coenzyme a biosynthetic pathway. Appl Microbiol Biotechnol. 2021;105:7321. https://doi.org/10.1007/S00253-021-11523-4.
73. Funke M, Diederichs S, Kensy F, et al. The baffled microtiter plate: Increased oxygen transfer and improved online monitoring in small scale fermentations. Biotechnol Bioeng. 2009;103:1118–28. https://doi.org/10.1002/BIT.22341.
74. Walker GM, Stewart GG. (2016) Saccharomyces cerevisiae in the Production of Fermented Beverages. Beverages 2016, Vol 2, Page 30 2:30. https://doi.org/10.3390/BEVERAGES2040030
75. Ajikumar PK, Xiao W-H, Tyo KEJ, et al. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. Science. 2010;330:70–4. https://doi.org/10.1126/science.1191652.
76. Song TQ, Ding MZ, Zhai F, et al. (2017) Engineering *Saccharomyces cerevisiae* for geranylgeraniol overproduction by combinatorial design. Sci Reports. 2017;71(7):1–11. https://doi.org/10.1038/s41598-017-15005-4.
77. Tokuhiro K, Muramatsu M, Ohto C, et al. Overproduction of geranylgeraniol by metabolically engineered *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2009;75:5536. https://doi.org/10.1128/AEM.00277-09.
78. Paz JL, Rodrigues JAR. Preparation of aromatic geraniol analogues via a Cu(I)-mediated grignard coupling. J Braz Chem Soc. 2003;14:975–81. https://doi.org/10.1590/S0103-50532003000600014.
79. Santoyo-Garcia JH, Walls LE, Nowrouzi B, et al. In situ solid-liquid extraction enhances recovery of taxadiene from engineered *Saccharomyces cerevisiae* cell factories. Sep Purif Technol. 2022;290:120880. https://doi.org/10.1016/J.SEPPUR.2022.120880.

Publisher's Note