Microbial Cell Factories

RESEARCH Open Access

Enabling malic acid production from corn-stover hydrolysate in *Lipomyces starkeyi* via metabolic engineering and bioprocess optimization

Jeffrey J. Czajka[1,2], Ziyu Dai[1,2], Tijana Radivojević[2,3,4], Joonhoon Kim[1,2,4], Shuang Deng[1,2], Teresa Lemmon[1], Marie Swita[1], Meagan C Burnet[2,5], Nathalie Munoz[2,5], Yuqian Gao[2,5], Young-Mo Kim[2,5], Beth Hofstad[1,2], Jon K. Magnuson[1,2,4], Hector Garcia Martin[2,3,4], Kristin E. Burnum-Johnson[2,5] and Kyle R. Pomraning[1,2*]

Abstract

Background *Lipomyces starkeyi* is an oleaginous yeast with a native metabolism well-suited for production of lipids and biofuels from complex lignocellulosic and waste feedstocks. Recent advances in genetic engineering tools have facilitated the development of *L. starkeyi* into a microbial chassis for biofuel and chemical production. However, the feasibility of redirecting *L. starkeyi* lipid flux away from lipids and towards other products remains relatively unexplored. Here, we engineer the native metabolism to produce malic acid by introducing the reductive TCA pathway and a C4-dicarboxylic acid transporter to the yeast.

Results Heterogeneous expression of two genes, the *Aspergillus oryzae* malate transporter and malate dehydrogenase, enabled *L. starkeyi* malic acid production. Overexpression of a third gene, the native pyruvate carboxylase, allowed titers to reach approximately 10 g/L during shaking flasks cultivations, with production of malic acid inhibited at pH values less than 4. Corn-stover hydrolysates were found to be well-tolerated, and controlled bioreactor fermentations on the real hydrolysate produced 26.5 g/L of malic acid. Proteomic, transcriptomic and metabolomic data from real and mock hydrolysate fermentations indicated increased levels of a *S. cerevisiae* hsp9/hsp12 homolog (proteinID: 101453), glutathione dependent formaldehyde dehydrogenases (proteinIDs: 2047, 278215), oxidoreductases, and expression of efflux pumps and permeases during growth on the real hydrolysate. Simultaneously, machine learning based medium optimization improved production dynamics by 18% on mock hydrolysate and revealed lower tolerance to boron (a trace element included in the standard cultivation medium) than other yeasts.

Conclusions Together, this work demonstrated the ability to produce organic acids in *L. starkeyi* with minimal byproducts. The fermentation characterization and omic analyses provide a rich dataset for understanding *L. starkeyi* physiology and metabolic response to growth in hydrolysates. Identified upregulated genes and proteins provide

*Correspondence:
Kyle R. Pomraning
kyle.pomraning@pnnl.gov

Full list of author information is available at the end of the article

potential targets for overexpression for improving growth and tolerance to concentrated hydrolysates, as well as valuable information for future *L. starkeyi* engineering work.

Keywords Oleaginous yeast, *Lipomyces starkeyi*, Malic acid production, Machine learning medium optimization

Introduction

Microbial production of commodity and specialty chemicals is increasingly important due to bioprocess potential to both reduce greenhouse gas emissions compared to traditional production means and provide more sustainable chemical production [1]. Oleaginous yeasts are particularly promising as microbial cell factories because they can naturally accumulate a large percentage (> 20%) of their biomass as lipids and are relatively easy to cultivate. Many oleaginous organisms have also been demonstrated to grow on a broad range of waste feedstocks, including hydrolysates [2]. *Lipomyces starkeyi* is a non-conventional oleaginous yeast with several additional native traits that make it an attractive biomanufacturing host. *L. starkeyi* naturally reaches lipid yields greater than 60% of its cell weight under certain conditions [3] and only displays a yeast morphology, reducing the risk of foaming and alleviating morphological complications during bioreactor fermentations [4]. While many *Lipomyces* species produce a polysaccharide slime layer, several "dry" strains, including *L. starkeyi* NRRL-11,558, have been identified [5]. These attributes have led to research exploring strategies for improving *L. starkeyi* lipid production on various waste feedstocks and hydrolysates, although previous engineering efforts have been limited due to a lack of genetic tools [6–10]. The relatively recent advances in *Lipomyces* genetic tools are beginning to enable the development of *L. starkeyi* into a microbial production chassis [11, 12].

Malic acid is a platform chemical used in a wide range of applications [13, 14], and several biochemical pathways have been identified and engineered to enable microbial production of the metabolite (previously reviewed and summarized [14, 15]),. The reductive TCA (rTCA) pathway is particularly attractive for malic acid synthesis due to its high theoretical yield and short pathway, which involves only two enzymatic steps: pyruvate carboxylase (PYC) and malic acid dehydrogenase (MDH, Fig. 1). Engineering the rTCA pathway along with a C4-dicarboxylic acid transporter into different microbes has been effective in achieving high malic acid titers [15–21]. While titers have reached over 200 g/L in some species, malic acid production is often accompanied by byproduct formation that complicates downstream separations (reviewed in [14]). For example, engineering the rTCA pathway (and deleting oxaloacetate acetyl hydrolase) in the filamentous fungus *Aspergillus niger* led to approximately 200 g/L of malic acid with the co-production of 28 g/L of citric acid in fed-batch fermentations [19]. Extensive genetic modifications (deleting two genes and overexpressing seven) were required to eliminate citric acid production while maintaining high titers [15, 22]. Thus, we were motivated to engineer *L. starkeyi* NRRL-11,558 to produce malic acid as a case study to investigate the feasibility of redirecting carbon flux away from lipid metabolism towards organic acid production, furthering the development of *L. starkeyi* into a microbial chassis for various biochemicals.

In this study, we engineered *L. starkeyi* NRRL-11,558 to produce malic acid and characterized production on both mock and corn-stover based hydrolysates. We collected transcriptomic, metabolomic, and proteomic data from the hydrolysate fermentations to gain insight into the engineered strain's metabolic responses and to provide a rich dataset for future engineering efforts. Finally, we utilized a recently developed medium optimization pipeline which leverages the Automated Recommendations Tool (ART) as a machine learning tool for medium optimization to improve production [23, 24].

Materials & methods

Chemicals, strains, and media

Chemicals used for routine culturing were purchased from Sigma-Aldrich (St. Louis, MO) unless noted otherwise. Hydrolysate was obtained from the National Renewable Energy Laboratory using a described deacetylation and disc refining enzymatic hydrolysis (DDR-EH, referred to as DDR throughout this manuscript) method on a corn-stover feedstock through a deacetylation and disc refining method followed by enzymatic hydrolysis, which was developed by the National Renewable Energy Laboratory [25–28]. Briefly, a dilute alkaline extraction is performed before acid pretreatment to remove acetate groups followed by enzymatic hydrolysis [27]. *L. starkeyi* strain NRRL Y-11,558 was obtained from American Type Tissue Culture (ATCC64135; Manassas, VA) and was used as the base (i.e., WT) strain for engineering. Cells were routinely grown on either YPD media (1% yeast extract, 2% peptone, 2% glucose) or a LPM salt media. The LPM salts formulation was as follows: per liter, 1.5 g KH_2PO_4, 1.43 g NH_4Cl, 0.5 g KCl, 0.5 g $MgSO_4$ * $7H_2O$, 1 mL of 1000 x trace element solution [2.25 g/L $ZnSO_4$, 11 g/L H_3BO_3, 5 g/L $MnCl_2$, 5 g/L $FeSO_4$, 1.7 g/L $CoCl_2$*$6H_2O$, 1.6 g/L $CuSO_4$*$5H_2O$, 0.085 g/L Na_2MoO_4, 5 g/L Na_4EDTA], and 1 mL 1000 x vitamin solution [1 g/L of the following components: biotin, pyridoxine, thiamine, riboflavin, para-aminobenzoic acid, nicotinic acid], pH adjusted with 5 M KOH. The mock hydrolysate

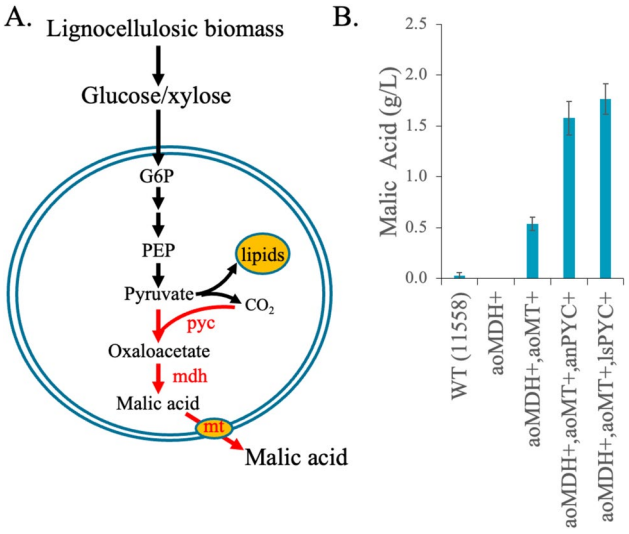

Fig. 1 Engineered malic acid biosynthetic pathway and production comparison among mutants. (**A**). The engineered rTCA pathway and the C4-dicarboxylic acid transporter in *L. starkeyi*. (**B**). Malic acid titers among different engineered *L. starkeyi* strains grown in glucose containing minimal media. Abbreviations: aoMDH, *A. oryzae* malate dehydrogenase; aoMT, *A. oryzae malate* transporter; anPYC, *A. niger* pyruvate carboxylase; lsPYC, *L. starkeyi* pyruvate carboxylase. Error bars represent standard deviations ($n=8$ for WT and $n=12$ for transgenic isolates constructs)

Table 1 Oligos used for vector construction

Oligo name	Oligo nucleotides
2288trpcF	gagaattaagggagtcacgaGGCTTAGTAGACGTGAGATTG
2289trpcR	aagtagtcatTGTATACTCGCTCTAAGGAATTTAAAG
2290nat1F	cgagtatacaATGACTACTTTGGATGACACTG
2291nat1R	ttagcgaaggGACCGAAGGCAAGTAGTGTAAG
2292tefF1	gccttcggtcCCTTCGCTAAGCCTCATCTATTAAC
2293tefR1	ctgccttgacTCCCTGTAGAAGTAAGCTGTTAGC
2294mdhF	tctacagggaGTCAAGGCAGCTGTGTTGGG
2295mdhR	ctgagcagggCTACTTAGGCGGAGGATTCTTAACG
2296tkuF	gcctaagtagCCCTGCTCAGCCACATAACG
2297tkuR	cgggcccatcgatgatcaggAGAGACGGTATTGCTGGTATTGG
2298mt1F	aattaaaccctcagcgagctCTAATCACTAACATCCTCATCCTTACC
2299mt1R	cacattcgcaATGCTGACACCTCCGAAGTTC
2300tdhF	gtgtcagcatTGCGAATGTGGATTAGAGTAAG
2301tdhR	acggccagtgaattcgagctGCAGATAGGTGAAAGGTCCAAATC
2302tefF2	tttcacctatctgcagctcgGCACAAGAATGCGGAGAAATC
2303tefR2	ggggagcagcTCCCTGTAGAAGTAAGCTGTTAGC
2304pycF	tctacagggaGCTGCTCCCCGCCAGCCCGA
2305pycR	gatgacaccaCTAGGCCTTGACGATCTTGCAGACAAGATCCTG
2306TtefF	caaggcctagTGGTGTCATCAAGTCCGTTATC
2307TtefR	ttgtaaaacgacggccagtgGAGGTCAATGAGGACGAAGAAG
2348tfR	agtcggtagaTCCCTGTAGAAGTAAGCTG
2349lsPCF	tctacagggaTCTACCGACTACTCGAATG
2350lsPCR	gatgacaccaCCCGGATCACTCATTACC
2351TtfF	gtgatccgggTGGTGTCATCAAGTCCGTTATC

formulation added glucose and xylose in an ~2:1 ratio, with the shaking flask medium screening using 26.5 g/L of glucose and 13.5 g/L of xylose. The hydrolysate fermentations utilized 52.7 g/L glucose and 27.3 g/L of xylose. DDR-hydrolysates added the appropriate amount of DDR to obtain a total sugar concentration similar to the mock hydrolysate (154 mL/L).

Genetic engineering

A previously described *Agrobacterium*-mediated transformation protocol for transforming *L. starkeyi* was used to generate strains in this study [11]. The transgene expression constructs were prepared by PCR and Gibson Assembly. The oligo pairs of 2288trpcF/2288trpcR and 2290nat1F/2291nat1R were used for isolation of *L. starkeyi* trpC promoter and the bacterial nourseothricin N-acetyl transferase (nat1) selection marker gene ([11], Table 1). The oligo pairs of 2292tefF/2293tefR, 2294mdhF/2295mdhR, and 2296tkuF/2297tkuR were applied to isolate the DNA fragments of *L. starkeyi tef1* promoter, *L. starkeyi mdh*, and the *L. starkeyi* transcriptional terminator of ku70, respectively. All five PCR DNA fragments were assembled into the pRF-Hu2 vector digested with *Hin*dIII/*Stu*I restriction enzymes to form pZD4078. The oligo pairs of 2298mt1F/2299mt1R and 2300tdhF/2301tdhR were used to isolate the DNA fragments of the *A. oryzae* malate transporter (*mt1*) with codon usage optimized for *L. starkeyi* and the *L. starkeyi tdh* promoter, which were inserted into pZD4078 linearized by the *Sac*I restriction enzyme to form pZD4079 via Gibson assembly.

The oligo pairs of 2302tefF2/2303tefR2, 2304pycF/2305pycR, and 2306TtefF/2307TtefR were used for the isolation of the *L. starkeyi tef1* promoter, the *A. niger* pyruvate carboxylase (*pyc*), and the *L. starkeyi tef1* transcriptional terminator. The DNA fragments of the *tef1* promoter, *pyc*, and the *tef1* transcriptional terminator were inserted into pZD4079 plasmid DNA linearized by the restriction enzyme *Eco*RI to form the plasmid DNA pZD4081. Similarly, the oligo pairs of 2302tefF2/2348tefR, 2349lspcF/2350lspcF, and 2351TtfF/2307TtefR were applied to isolate the *L. starkeyi* DNA fragments for *tef1* promoter, *pyc1* coding sequence, and *tef1* transcriptional terminator, which were cloned into pZD4079 at restriction enzyme *Eco*RI site to form pZD4082.

Growth conditions

Cells were maintained on YPD agar plates (1% yeast extract, 2% peptone, 2% glucose, 2% agar). Frozen cell stocks were maintained at -80 °C in 15% glycerol. Cultures were routinely grown in YPD media (same formulation as YPD agar, without the agar addition) or LPM salt media. For growth in shake flasks and bioreactor, cells

from the glycerol stocks were plated onto YPD plates and grown for 4–5 days at 28 °C until colonies were observed. For production assays, the following growth procedure was followed: 5–10 colonies were picked and cultured for 36 h in 2 mL YPD media, 1 mL of the YPD seed culture was transferred to 35 mL of the mock hydrolysate media with 40 g/L of total sugar. Initial screening assays occurred in glucose only cultures (40 g/L glucose). The seed culture was grown for two days and used to inoculate the specified medium condition to be tested at a starting OD_{600} of 0.05–0.10, with an exception of the initial OD_{600} of the lsPYC + and anPYC + strains reported in Fig. 2D which started at a value of 0.01.

Bioreactor cultivations

For bioreactor seed cultures, the same procedure was followed as the growth in shake flasks with the following modifications: 3 mL of YPD seed culture was grown for 4 days, and the mock hydrolysate seed was a 50 mL culture containing 80 g/L total sugar before being inoculated into the bioreactors at OD_{600} ~ 0.35. pH cultivation seed culture media contained 80 g/L of glucose as the sole carbon source. Bioreactor cultivations were performed in an Infors AG Sixfors Fermenter system with 675 mL vessels. The initial pH screening cultivations utilized 80 g/L of glucose and a working volume of 400 mL. The vessels were operated at 2.0 vvm, 400 rpm, and 30 °C. The hydrolysate fermentations were operated with 500 mL working volume, 2 vvm, 30 °C, and at a pH of 5.5. pH was controlled using 5 M KOH.

Omic sampling and analytical procedures

15 mL samples were collected from the bioreactors for omic analyses, with 1.5 mL of cell culture aliquoted into individual tubes for each individual omic analysis (proteomics, transcriptomics, metabolomics). Cells were

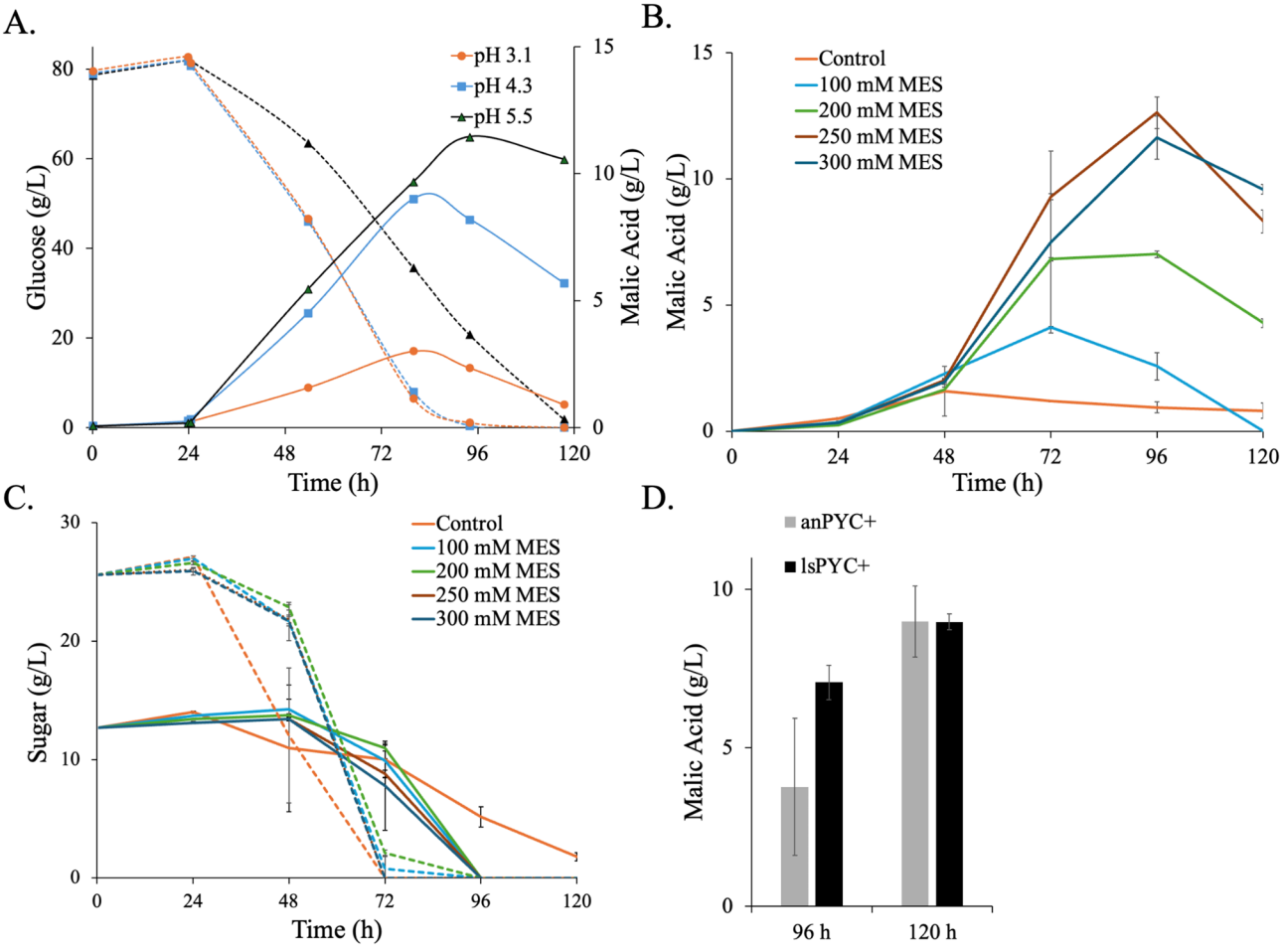

Fig. 2 Malic acid production dynamics in buffered mock hydrolysates. **(A).** 0.5 L bioreactor production at various pH setpoints. Dashed lines indicate glucose and solid lines indicate malic acid. **(B).** Malic acid time dynamics and **(C).** sugar consumption time dynamics in buffered shake flask conditions using mock hydrolysate. Dashed lines indicate glucose and solid lines indicate xylose. **(D).** Malic acid titers of the lsPYC + and anPYC strains in buffered shaking flask conditions. A long lag phase was observed in the growth due to a starting OD_{600} of 0.01. Error bars represent standard deviations ($n = 6$ for **B**, $n = 5$ for **C** up to 72 h, $n = 3$ after 72 h)

pelleted via centrifugation for 30 s at 13,000 x g, the supernatant was discarded, and the cells were washed once with PBS before being frozen in liquid nitrogen. For extracellular metabolite analysis, the supernatant/culture broth was filtered through 0.22 µm filters and frozen until analysis was conducted. For dry cell weight determination, 5–9 mL of culture was pelleted in a pre-weighed conical tube (5 min at 4,000 x g), washed once with water, and frozen at -20 °C. The frozen pellet was then lyophilized for 24–36 h.

Approximately 5 mg of lyophilized cell pellet was extracted using MPLEx to obtain the intracellular metabolites [29]. The dried extracts were derivatized and analyzed via GC-MS as previously described [30]. Proteomic data was processed, collected, and analyzed according to published methods described in our prior work [30]. Extracellular metabolites were quantified from filtered medium. These samples were analyzed using high performance liquid chromatography (HPLC) equipped with a Waters 2489 UV/Visible detector collecting signal at 210 nm. A Bio-Rad Aminex HPX-87 H ion exclusion column (300 mm × 7.8 mm) heated to 50 °C was used for analyte separation. Sulfuric acid (0.0045 M) was used as eluent at a flow rate of 0.55 mL/min. Calibration curves were built for each quantified metabolite using linear regression and used to determine the concentration in the analyzed samples.

RNA was extracted from samples using a Maxwell 16 LEV Plant RNA kit (Promega, Madison, WI) and sequenced on an Illumina platform. Sequences were mapped to the *L. starkeyi* NRRL-11,557 transcriptome [31] to quantify expression with featureCounts [32] and converted into reads per kilobase of transcript per million mapped reads (RPKM).

Omic data analysis

We utilized the RPKM values from transcriptomics to calculate Log_2FC differences and statistics between various conditions. The RPKM values were transformed by Log_2 for each condition. For comparing conditions, the Log_2(RPKM) values were subtracted from each other (giving Log_2(condition 1/condition 2). p-values were calculated using the t-test for two independent samples with identical variances for all the samples in the conditions. The Benjamin-Hochberg false discovery correction was applied to the list to obtain q-values. The same procedure was applied to the proteomics and metabolomics data using spectra counts.

For GSEA, the JGI annotated genome for NRRL-11,557 (https://mycocosm.jgi.doe.gov/pages/search-for-genes.jsf?organism=Lipst1_1) was used to obtain GO, KOG, KEGG, and IPR terms for each gene and protein. The python package GSEApy was utilized to perform GSEA [33].

Machine learning medium optimization

The ART algorithm and a previously described medium optimization pipeline were adopted for the workflow here [23, 24]. The specific code and input/output files utilized are stored at https://github.com/AgileBioFoundry/LstarkeyiMalicAcidProduction. Medium optimization was performed in the mock hydrolysate media containing 250 mM of 2-(N-morpholino) ethanesulfonic acid (MES) as a buffering agent (see Chemicals, strains and media). ART was allowed to recommend salt and micronutrient concentrations approaching toxicity levels reported for other yeasts [34–39]. Nitrogen and phosphorus media component concentrations were varied by approximately one order of magnitude to maintain C/N and C/P ratios around the Redfield Ratio ($C_{106}N_{16}P_1$). For each learning cycle, the inputs to ART were the individual media component concentrations with malic acid titers as the output to be learned. During initial design of an experiment, a weighting factor was applied to the ART sampling as detailed in the github repository, in the following notebook https://github.com/AgileBioFoundry/LstarkeyiMalicAcidProduction/blob/main/Lipomyces_Media_Opt_Cycle1.ipynb. A control condition was utilized in each DBTL cycle as previously described [24].

Results

rTCA cycle engineering enabled malic acid production

L. starkeyi NRRL-11,558 is a "dry" strain that does not secrete a polysaccharide slime layer and which was not found to produce significant amount of organic acid byproducts under our growth conditions (Figure S1). We engineered the rTCA cycle pathway, consisting of two genes, and a malic acid transporter into strain NRRL-11,558 (henceforth, referred to as the wildtype (WT)). Introducing the first rTCA enzymatic step, *Aspergillus oryzae* malate dehydrogenase (aoMDH), and integrating an *A. oryzae* malate transporter (aoMT) increased the limited malic acid production and secretion in a minimal medium (0.5 +/- 0.1 g/L). Overexpressing either the native PYC (lsPYC+) or a heterologous *A. niger* PYC (anPYC) into the engineered strain (aoMT+, aoMDH+) improved malic production over three-fold, with the highest titer of 1.8 +/- 0.2 g/L achieved in the lsPYC + strain (Fig. 1B).

Malic acid production has previously been shown to be optimal near neutral pH values in several microbes such as *Rhizopus oryzae* [40], *Myceliophthora thermophila* [16], and *Aspergillus oryzae* [17]. At acidic pH values, the permeability of undissociated acid through the membrane transporters significantly decreases, which can hinder production [41]. However, acidic production broths are essential for efficient and economic production and separation of organic acids at industrial scales [42]. Our initial cultivation medium had limited buffering

capabilities, and the pH value had dropped below 2.0 at the end of the shaking flask cultivations, potentially inhibiting production (Figure **S2**). Therefore, we evaluated production at three pH values in controlled fermenter vessels to determine if malic acid production occurred at acidic or slightly acidic pH values.

Buffered conditions improved production at mildly acidic pH values

The lsPYC + production strain was cultivated in 0.5 L bioreactors with precise pH measurement and control. The pH values of the reactors were set to 5.5, 4.3, and 3.1. Consistent malic acid production was observed at pH values of 5.5 and 4.3, with the highest malic acid titer (~ 12 g/L) obtained at a pH value of 5.5 (Fig. 2A). Although it was promising that production occurred at pH 4.3, we chose to use a pH of 5.5 in this work to investigate optimal production conditions in shake flasks and due to wider availability of buffering agents.

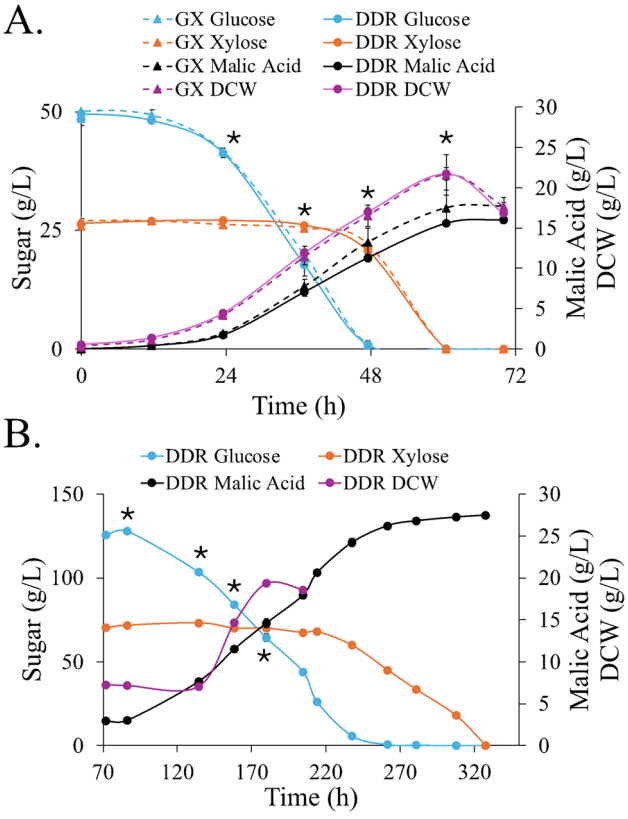

Fig. 3 Comparison of malic acid production and sugar consumption on mock and DDR hydrolysate in 0.5 L bioreactors. Malic acid titers, DCW, and glucose and xylose consumption in (**A**). Mock (GX– glucose & xylose) and low gravity DDR hydrolysates over 70 h and, (**B**). high gravity DDR hydrolysate from 70 h to over 320 h. The high gravity DDR had 3x the starting sugar concentration as the low gravity DDR. The * represents timepoints at which samples were collected for omic analyses. Error bars represent standard deviation ($n = 2$)

Next, we sought to replicate the obtained bioreactor titers in shake flask cultivations by supplementing the media with a buffering agent (Figure S2). As we were also interested in investigating production dynamics on hydrolysates, we switched the cultivation carbon source from glucose to a mock hydrolysate (glucose and xylose in a 2:1 ratio). We evaluated several potential buffers with the lsPYC + strain and identified 2-(N-morpholino) ethanesulfonic acid (MES) as a sufficient buffering agent that did not negatively affect biomass accumulation (Figs. 2B and S2). Malic acid production reached similar titers as in the bioreactors (12.6 +/- 0.6 g/L in the shake flasks), with secretion observed during both glucose and xylose consumption phases (Fig. 2B). *L. starkeyi* has been shown to be able to grow on malic acid [43], and its uptake was observed after exhaustion of the sugar substrates (Fig. 2, panels B and C). Therefore, we re-evaluated the lsPYC + and anPYC strains in the buffered media (using 250 mM MES) to ensure we had the optimal production strain. The lsPYC + strain demonstrated faster production kinetics and a small improvement in titer compared to the anPYC strain (Fig. 2D). From these results, the lsPYC + strain was selected for further work characterizing and improving malic acid production on hydrolysates.

Strain production on high gravity corn-stover hydrolysate improved titers

Mock hydrolysate media formulations can simulate strain growth on the major sugar components in real hydrolysates but often cannot fully capture the stresses associated with the real hydrolysates. Thus, a corn-stover based hydrolysate from high solids enzymatic hydrolysis (referred to as DDR hydrolysate) was used to test the growth, production, and sugar conversion on a real feedstock. The lsPYC + strain was grown on both the pure sugar mock and DDR hydrolysates in 0.5 L bioreactors (Fig. 3A), with initial cultivations occurring at relatively low carbon loadings (~ 75 g/L of total sugar). Production dynamics on the mock and DDR hydrolysates were similar, with 17.7 +/- 1.1 g/L and 16.0 +/- 0.4 g/L of malic acid achieved, respectively (p-values > 0.05). Biomass titers on both hydrolysates were also comparable (Fig. 3A). Similarly, the overall yield in the 0.5 L fermentations was not significantly different compared to the shaking flask conditions (0.297 +/- 0.025 vs. 0.295 +/- 0.014 Cmol malate/Cmol of sugar consumed), demonstrating limited inhibition effects from the low-density hydrolysate and as well as from a bioreactor environment.

Next, we evaluated production on more concentrated DDR. A single low-density DDR reactor was re-inoculated into a high gravity (less diluted, with a total sugar concentration of ~ 200 g/L) DDR hydrolysate. An initial delay of growth was observed in the concentrated conditions (Fig. 3B). However, after the lag phase, a final titer

of 26.5 g/L was achieved at sugar exhaustion. An examination of the product yields at the end of fermentation revealed a reduction of yield in the high gravity DDR compared to the low carbon loading DDR (0.121 Cmol malate/Cmol of sugar consumed vs. 0.190 +/- 0.005). The long lag phase and the reduced yield during fermentation likely indicate that the cells experienced a stressed condition which limited product yields. To further investigate the stress responses during the DDR and mock hydrolysate fermentations, omic samples were collected at timepoints indicated by asterisks (*) in Fig. 3. These omic measurements included transcriptomics, proteomics, and metabolomics and were previously used to construct a genome-scale model for the *Lipomyces* clade [43].

Omic analysis revealed cell stress responses and engineering targets

The molecular changes in the transcriptomic, proteomic, and metabolomic profiles can provide insight into cellular stress responses during hydrolysate growth. Principal Component Analysis (PCA) of the each omic data type indicated relatively similar profiles for the biological replicates (sampled from individual bioreactors) and highlighted a distinct metabolic and regulatory response to the high gravity DDR compared to the low gravity and mock hydrolysate samples (Fig. 4). The proteomic and transcriptomic PCA plots showed that samples collected from the same phase of fermentation (early, mid, or late stage) were closely grouped together irrespective of the cultivation media for mock and low gravity DDR hydrolysate (Fig. 4A and B). On the other hand, the metabolomic PCA plot indicated a closer grouping of the DDR hydrolysate samples (both low gravity and high gravity) than

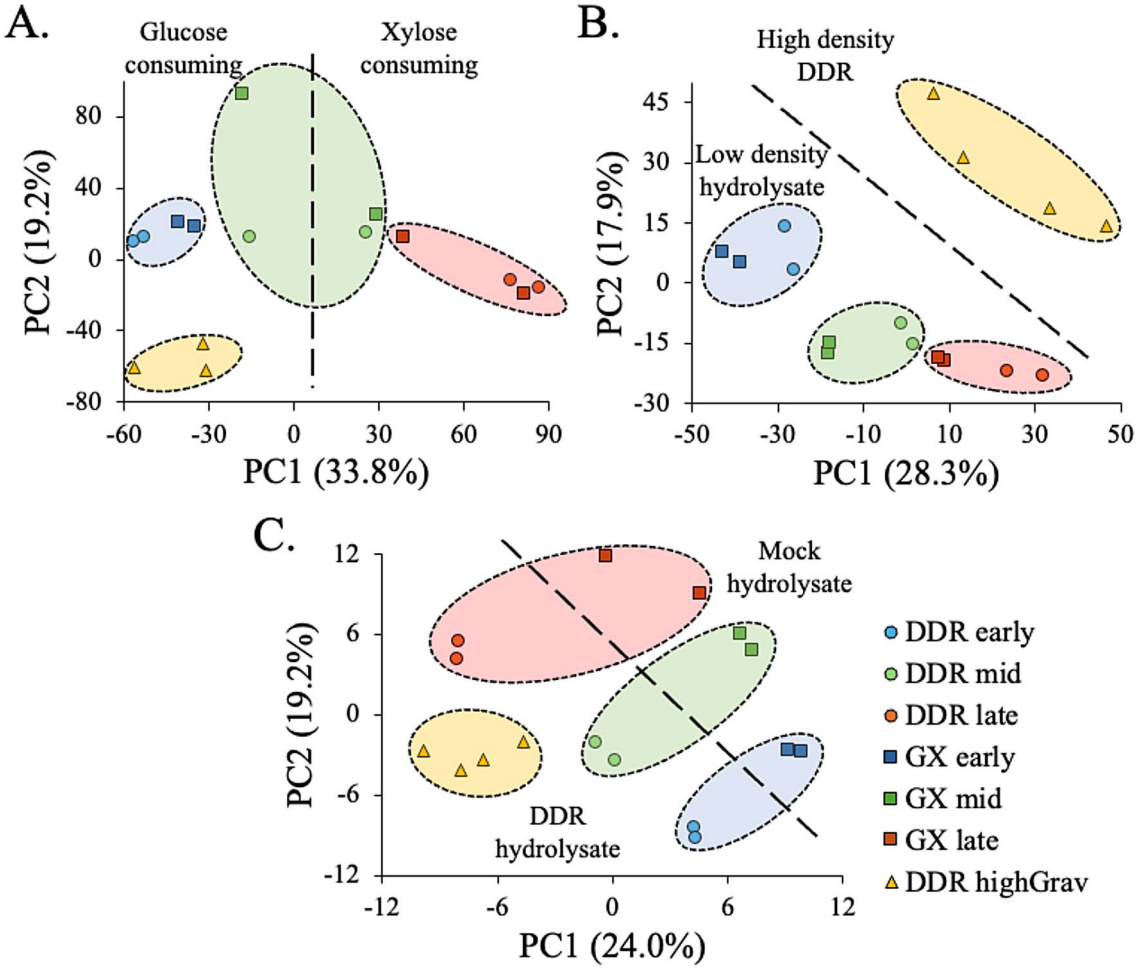

Fig. 4 PCA of omic data collected from the mock and DDR hydrolysate bioreactor cultivations. (**A**). Transcriptomic data with the outlier removed, (**B**). Proteomic data, and (**C**). Metabolomic data. The circles indicate samples collected from the same fermentation timepoints in both the mock and DDR hydrolysates. Blue - early phase, Green - mid phase, Red - late phase, Yellow - high gravity DDR fermentation. GX represented the glucose-xylose mock hydrolysate and DDR represented the real hydrolysate. Dashed lines indicate different fermentation regimes noted in the figure panels. Legend applies to all panels

the mock samples (Fig. 4C). However, one noted exception was the exclusion of the first transcriptomic sample from the high gravity DDR, which was collected after the transfer of the culture from the low-to-high gravity DDR (Figure S3). This sample was an outlier that may have reflected stress responses during re-inoculation handling and was thus excluded from downstream analysis.

Analysis of the individual omic data types revealed a significant number of upregulated and downregulated genes, proteins, and metabolites (p-values < 0.05 and an absolute $\log_2$ fold change > 2, Figures S4-S6, Additional Files 2 and 3). Amylases and catalases were highly upregulated on DDR compared to mock hydrolysates (each accounted for three of the top 20 most differentially expressed proteins) while maltose/isomaltose was accumulated intracellularly. It has been observed that the DDR process does not completely hydrolyze the glucan fraction to monomeric glucose and can leave up to ~15% as oligomers after hydrolysis [26–28]. Expression of amylases (glycosyl hydrolase family 13) that cleave α(1→4) glycosidic linkages in glycogen, starch, and related α-glucans, and intracellular accumulation of disaccharides suggests *L. starkeyi* is responding to and able to utilize the undigested oligomers. Five of the top 20 upregulated proteins where oxidoreductases, oxidases, and hydrolases associated with detoxification and antioxidant defenses, along with an unannotated protein (proteinID: 198908) with homology to polyprenyl-4-hydroxybenzoate decarboxylases in *Aspergillus* species. Four of the these stress response proteins were also found to be up-regulated on a corn-stover hydrolysate that underwent a different pre-treatment process (PCS [10]),, potentially indicating conserved stress response mechanisms, and potential targets for overexpression (proteinIDs: 5102, 6823, 69860, 198908). Many of the most highly upregulated genes and proteins beyond the top 20 were also annotated as oxidoreductases, catalases, and dehydrogenases (Additional File 2). This abundance of highly upregulated catalases suggests additional hydrogen peroxide stress during conversion of DDR. In yeast catalases are typically expressed in the peroxisome to mitigate H_2O_2 formation during oxidative processes [44]. A metabolomic comparison revealed the presence of the potentially toxic compounds 4-hydroxybenzoic acid and benzoic acid within the yeast during DDR growth and a large increase in several sugar alcohols including palatinitol, sorbitol, deoxyhexitol, threitol, and erythritol during growth (Additional File 3).

Gene Set Enrichment Analysis (GSEA, Figures S7-S9) using the transcriptomic data (based on GO terms) showed genes involved in membrane and oxidoreductase activity were enriched on DDR (fdr < 0.15), further corroborating potential stress responses. Additionally, there was a significant enrichment (fdr < 0.05) of processes associated with ATP-coupled transmembrane movement of substances and catalytic activity in the DDR media. On a more granular level, many upregulated genes in DDR were indicated to be membrane proteins and transporters, with several annotated as proteins in the major facilitator superfamily MFS-1 (Additional File 2). Several of the highly expressed membrane proteins had homology to transporters classified as efflux pumps in the Transporter Classification Database [45]. There was an also general decrease in most amino acids under DDR conditions (except for aspartate) and fatty acids in the DDR (linoleic acid, palmitoleic acid, oleic acid, etc.).

Increasing the DDR hydrolysate concentration (high gravity DDR) led to further changes in gene expression compared to growth in the low gravity DDR. The most highly upregulated protein on the high gravity DDR is a homolog of *S. cerevisiae* hsp9/hsp12 (*L. starkeyi* proteinID: 101453), a stress-inducible molecular chaperone that has been indicated to maintain membrane organization during stress response and to assist with protein folding [46, 47]. A chromatin assembly factor-I protein (proteinID: 328687) that helps protect DNA from damaging agents was also among the most highly expressed proteins [48]. The transcriptome data further indicated the upregulation of two glutathione dependent formaldehyde dehydrogenases (proteinIDs: 2047, 278215) which can act as reductases to catabolize toxic hydrolysate compounds [49]. GSEA analysis showed a trend of reductases and hydrolyses processes (GO-term based) being upregulated in higher DDR concentrations, and several glutathione-dependent formaldehyde deactivating transcripts highly expressed (Figures S7-S9, Additional File 2). Compared to the low-gravity DDR, expression of several zinc/iron permeases, ferric reductases, and proteins involved in oligopeptide transport were highly repressed (Additional File 2).

A comparison of the xylose-glucose consumption phases showed that enzymes and metabolites involved in xylose metabolism are specifically enriched. However, two putative xylose transporters, proteinIDs: 7313 and 74,605, were identified (highly upregulated and structurally similar to HxtB).

Machine learning medium optimization identified enhanced production medium

Oleaginous yeast metabolic changes are highly correlated with nutrient limitations [50–52]. In the model oleaginous yeast, *Y. lipolytica*, an acid producing phase has been induced by sulfur limitation [53] and iron concentrations [54]. These observations provided motivation for conducting malic acid medium optimization. Here, we utilized a medium optimization framework that leverages the automated recommendation tool (ART), a machine learning algorithm developed for biological applications.

The medium optimization pipeline was effective at identifying conditions that enhanced *Pseudomonas putida* flaviolin production [23, 24, 55].

We conducted three cycles of medium optimization to enhance malic acid production, targeting 18 media components across a wide concentrations range that approached toxicity for some trace elements [34–39]. We sampled conditions at 96 h, as this timepoint had the highest production in our control condition and was the point at which our carbon source was fully consumed. Initially, only 20% of conditions (10/50, excluding the control condition) displayed growth at 96 h (Fig. 5A). Over subsequent cycles, ART learned to avoid inhibitory concentrations, improving growth to 47% and 84% in cycles two and three, respectively (Fig. 5B). Malic acid production saw similar gains, increasing from 2% of conditions with improved titer in the first cycle, to 9% and 12% in the latter cycles.

A limitation of the large-scale experimental setup with single timepoint sampling approach was that the different medium conditions may have led to changes in sugar utilization rates. If the sugar was completely consumed before the sample, as was observed for several conditions, product consumption may have occurred (Fig. 2). Thus, we selected the 12 best performing conditions (based on malic acid titers) for a time-course experiment. Several conditions showed production dynamic shifts, with the best strain improving production rates by approximately one day compared to the control (10.4 +/- 0.1 g/L produced at 76 h versus 9.4 +/- 0.4 g/L in the control). While rates were improved, there was limited improvement of production titers in the best condition compared to the control (10.9 +/- 0.1 g/L and 10.7 +/-1 0.2 g/L, Fig. 5C). The improved rates corresponded with an increased sugar consumption data, with the improved rate condition having ~ 4.5 g/L more sugar consumed by the 76 h timepoint than the control (Figure S10). Sugar was completely consumed in this condition by the 85 h measurement. Exploring the wide range of concentrations revealed that *L. starkeyi* NRRL-11,558 has high

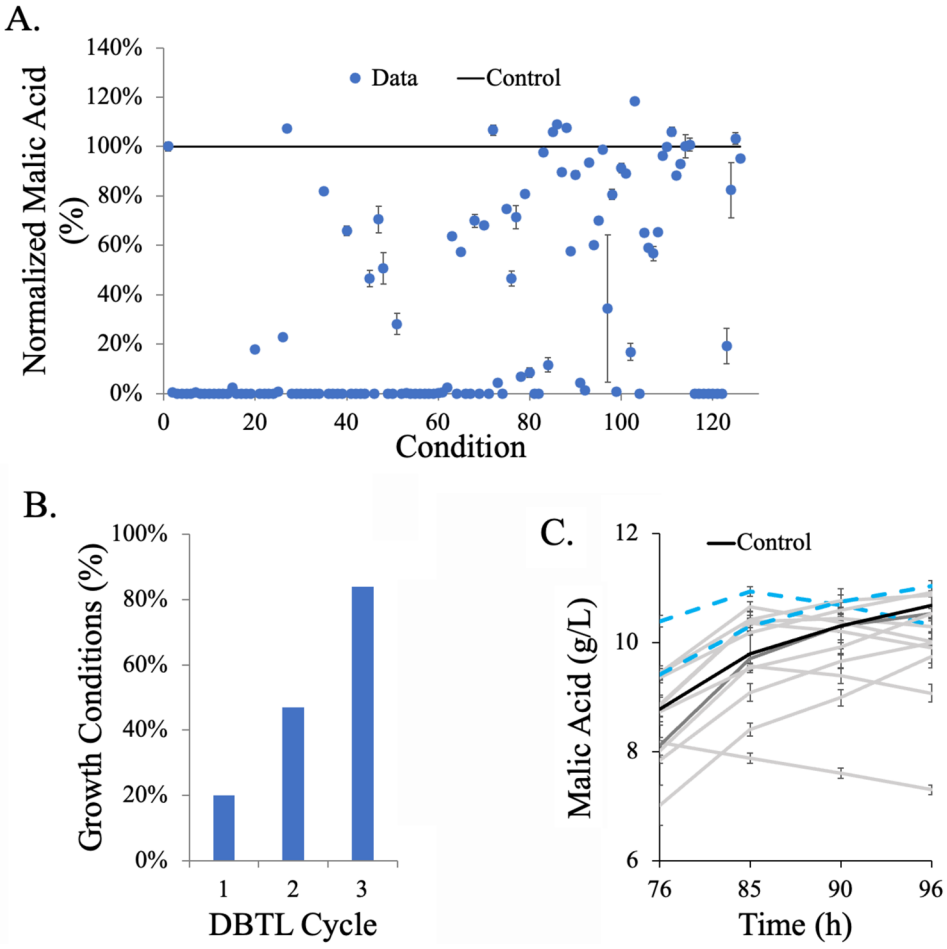

Fig. 5 ART medium optimization results. (**A**). Malic acid titers over the course of three DBTL cycles. (**B**). The percentage of media designs displaying growth across the cycles. (**C**). Dynamic sampling of malic acid production from the highest titer conditions identified in (**A**). Blue dotted lines represent the two highest titer conditions from the experiment. Error bars represent standard deviation ($n = 3$)

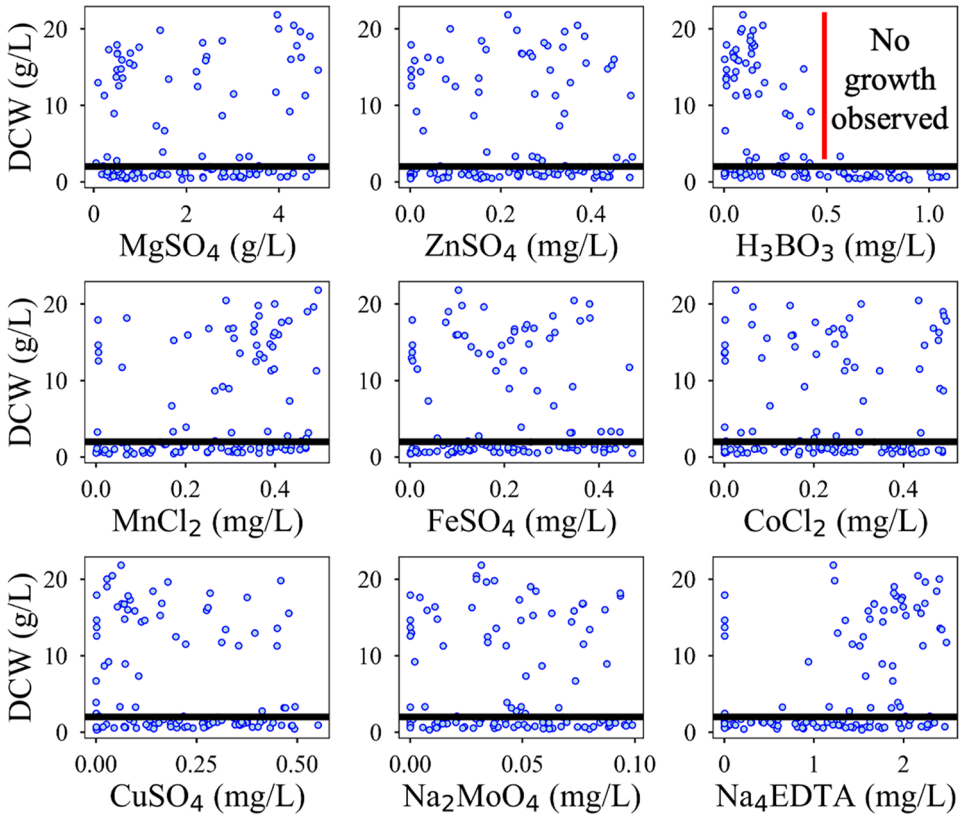

Fig. 6 Tolerances to trace elements and metal media components. Biomass DCW plotted versus the trace element concentration for each trace element varied in the media. Y-axes indicate the obtained DCW (g/L) and the X-axes represent the specified component concentrations. Black solid lines indicate the DCW threshold for growth/no growth in a condition. The majority of medium components had growth across their concentration range except for boric acid indicated with the red solid line delimitating point of inhibitory effect. Data points represent the average of triplicate flasks. Error bars have been omitted for clarity

tolerance to most trace elements, except for boron. Concentrations of 8–10 mM of boron inhibited growth, which is significantly lower than the ~80 mM tolerated by *Saccharomyces cerevisiae* ([56], Fig. 6).

Discussion

Metabolic engineering of the rTCA pathway significantly enhanced *L. starkeyi* malate production

The rTCA pathway and a C4-dicarboxylic acid transporter have been introduced into various microbial backgrounds and have led to significant malic acid production. In this study, engineering the rTCA cycle into *L. starkeyi* resulted in an accumulation of 26.5 g/L malic acid on a corn-stover hydrolysate. Although the obtained titer is significantly lower than the titers achieved in *Aspergillus* species (200 g/L of malic acid obtained in *A. niger* [15, 19] and >100 g/L in *A. oryzae* [57, 58]), it is a promising first step. Importantly, organic acid byproducts were not detected at significant levels in this strain (FigureS1) which can complicate downstream processing [15]. Furthermore, the malic acid titer in the study was in line with results obtained from organisms that were not initially prolific acid producers (unlike *Aspergillus* species). For example, ~40 g/L of malic acid was obtained in *Pichia pastoris* [21] and ~9 g/L in *Torulopsis glabrata* [18]. In previous work, the C4-dicarboxylic acid transporter has been demonstrated to be a crucial step for controlling malic secretion, which was further confirmed in our work [22]. Several prominent transporters have been identified as efficient transporters for enhancing malic acid production in addition to the *A. oryzae* transporter overexpressed in *L. starkeyi* here. Specifically, the *A. carbonarius c4t318* transporter and *Schizosaccharomyces pombe* have recently been used in several of the highest producing strains [15, 22, 57]. Utilizing one of these alternative transporters may lead to more efficient *L. starkeyi* malic transport and could improve overall production. These prior works have also demonstrated several engineering strategies, such as redox balancing and overexpression of glycolytic and TCA cycle enzymatic steps that could lead to further enhancement of malic acid titer in *L. starkeyi* (reviewed in [14]). Redox balancing may be important as additional analysis of the collected data from the medium optimization revealed a tradeoff between biomass titer and malate produced

(Figure S1), potentially hinting at the need for cofactor/redox balancing.

Omic insights revealed upregulated expression of transporters and oxidoreductases

Hydrolysates produced from waste lignocellulosic biomass typically contain compounds that pose challenges to cell growth. *L. starkeyi* has been demonstrated to grow robustly on a wide range of different hydrolysates and waste feedstocks [6–10], and our results demonstrated similar robust growth. The omics data did indicate that despite the robust growth, the yeast was under a stressed-oxidative state in the DDR hydrolysate with upregulation of oxidoreductases and a corresponding enrichment of intracellular sugar alcohols observed [59]. The identified a hsp9/hsp12 homolog (*L. starkeyi* proteinID: 101453) that was highly upregulated in high gravity DDR along with the glutathione-dependent formaldehyde-activating enzymes and upregulated oxidoreductases may be alleviating the stressed condition in *L. starkeyi*. Overexpression of these genes in *L. starkeyi* could improve production of malic acid on and tolerance to highly concentrated hydrolysates. Such strategies have been shown to increase tolerances to hydrolysates in both bacteria and yeast and have led to improved production [60–62]. Additionally, there were indicators of upregulation of many membrane transporter proteins in DDR (Additional File 2). The highly expressed transporters likely facilitate growth in hydrolysates by preventing the accumulation of intracellular toxic compounds and may also represent future targets for potential overexpression [63].

In addition to the stress response, we see an upregulation of carbohydrate, amino acid, and inorganic ion transporters on DDR, a corresponding downregulation of native amino acid biosynthesis routes, and an overall decrease of amino acid components. Being derived from organic sources, the hydrolysates may contain a broad range of oligo and monosaccharides present in trace amounts. The upregulation of many amylases and the accumulation intracellular disaccharides suggests the presence of undigested oligomer sugars, which has been previously reported in DDR hydrolysate [26, 27]. *L. starkeyi* may take advantage of the multiple substrates and simultaneous uptake many of them. Prior work had indicated that *L. starkeyi* may have less tightly regulated catabolic repression, as xylose uptake was observed before the complete exhaustion of glucose [10]. Our results supported this previous observed phenomenon, as our time-course groupings of the PCA plots (Fig. 3) revealed a transitory state with changes in gene expression, proteins present, and metabolomics before glucose exhaustion. As a microbe found in soils, it is likely *L. starkeyi* has evolved to have metabolic flexibility to take advantage of the presence of multiple substrates in the environment.

Takeaways from machine learning for biological media optimization

ART is a machine learning tool that has been shown to successfully guide the genetic engineering process by leveraging biological data [55]. It had been previously utilized to identify medium conditions that significantly enhance flaviolin production in *Pseudomonas putida* [24]. Employing ART and the medium optimization pipeline led to a condition with improvements in production dynamics and improved titers despite the enormous initial design space and a limited number of cycles. The rapid generation of media compositions that resulted in viable (non-zero) production and growth demonstrate the power of machine learning tools for biological medium optimization. This medium optimization strategy faced several challenges that could be overcome in future work. Specifically, initializing the model design space with tighter constraints can limit conditions exploring unfeasible (toxic) growth conditions. Robotic systems for liquid handling can increase the throughput of such experiments while simultaneously reducing the associated human error in medium preparation, which can improve repeatability and enhance the performance of the machine learning algorithm, as was demonstrated in the *P. putida* work [24]. Finally, the dynamic nature (secretion and consumption) of the malic acid production system led to issues with collecting optimal timepoints as production dynamics shifted over the medium conditions, with several conditions reaching maximum titers sooner than the control. Subsequent consumption of secreted malic acid after sugar exhaustion added difficulty in ensuring our data represented the max titers. Despite these challenges, ART was able to identify an enhanced condition. Products that are not re-consumed or which can be measured online during cultivation are likely to lead to better overall results and recommendations for optimal mediums.

Conclusion

In this study, we have successfully established malic acid production in the oleaginous yeast *L. starkeyi* with minimal formation of byproducts. The engineered malic production strain provides a platform for further research endeavors aimed at achieving high titers with limited loss of carbon to byproducts. Our comprehensive collection of omic datasets, encompassing both real and mock hydrolysate fermentations, provide a valuable and information-rich resource that can aid in the identification of future engineering targets and enhance our understanding *L. starkeyi* metabolism. Moreover, through our investigation, we have garnered practical insights into using machine learning-based medium optimization tools, which will prove instrumental in guiding future work. Overall, our work contributes significantly to the

advancement of *L. starkeyi* as a promising microbial production chassis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12934-025-02705-0.

Supplementary Material 1: Additional File 1. Full omic dataset and complete list of annotations.

Supplementary Material 2: Additional File 2. Log2FC results for transcriptomic and proteomic data.

Supplementary Material 3: Additional File 3. Log2FC results for metabolomic data.

Supplementary Material 4

Acknowledgements
This research was initially supported through the Fungal Biotechnology Core R&D awarded by DOE-EERE, Office of Bioenergy Technologies. Further research was a part of the Agile BioFoundry (https://agilebiofoundry.org) supported by the DOE-EERE through contract DE-AC02-05CH11231 (Lawrence Berkeley National Laboratory, LBNL) and DE-NL0030038 (Pacific Northwest National Laboratory, PNNL). A portion of this research was also performed on a project award (10.46936/reso.proj.2020.51637/60000235) from EMSL, a DOE SC-BER User Facility under Contract No. DE-AC05-76RL01830. In addition, a portion of the work was conducted in the DOE Joint BioEnergy Institute supported by the DOE SC-BER via contract DE-AC0205CH11231 to LBNL. JJC is grateful for support from a Linus Pauling Distinguished Fellowship by PNNL-Laboratory Directed Research and Development Program. PNNL is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830.

Author contributions
Project Conceptualization: JJM, KRP, ZD. Experimental Investigation: ZD, KRP, JJC. Analytical Data Collection: TL, MS, NMM, MCB, YG, YMK. Data Curation, Interpretation, Analysis: JJC, KRP, JKM. Machine Learning Software– Original Code: TR. Machine Learning Code Adaptions– JJC. Writing– Original Draft: JJC. Writing– Review & Editing: SD, KRP. Supervision: BH, YMK, HGM, KEBJ, KRP, JKM. All authors read and approved the manuscript.

Funding
This research was initially supported through the Fungal Biotechnology Core R&D awarded by DOE-EERE, Office of Bioenergy Technologies. Further research was a part of the Agile BioFoundry (https://agilebiofoundry.org) supported by the DOE-EERE through contract DE-AC02-05CH11231 (Lawrence Berkeley National Laboratory, LBNL) and DE-NL0030038 (Pacific Northwest National Laboratory, PNNL). A portion of this research was also performed on a project award (https://doi.org/10.46936/reso.proj.2020.51637/60000235) from EMSL, a DOE SC-BER User Facility under Contract No. DE-AC05-76RL01830. In addition, a portion of the work was conducted in the DOE Joint BioEnergy Institute supported by the DOE SC-BER via contract DE-AC0205CH11231 to LBNL. JJC is grateful for support from a Linus Pauling Distinguished Fellowship by PNNL-Laboratory Directed Research and Development Program. PNNL is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830.

Data availability
Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA
[2]DOE Agile BioFoundry, Emeryville, CA 94608, USA
[3]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]DOE Joint BioEnergy Institute, Emeryville, CA 94608, USA
[5]Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA

Received: 11 October 2024 / Accepted: 24 March 2025
Published online: 21 May 2025

References

1. Caporusso A, Capece A, De Bari I. Oleaginous yeasts as cell factories for the sustainable production of microbial lipids by the valorization of Agri-Food wastes. Ferment [Internet]. 2021; 7(2).
2. Abeln F, Chuck CJ. The history, state of the Art and future prospects for oleaginous yeast research. Microb Cell Fact. 2021;20(1):221.
3. Juanssilfero AB, Kahar P, Amza RL, Yopi, Sudesh K, Ogino C, et al. Lipid production by lipomyces starkeyi using Sap squeezed from felled old oil palm trunks. J Biosci Bioeng. 2019;127(6):726–31.
4. Smith MT. 37 - Lipomyces lodder & Kreger-van Rij. In: Kurtzman CP, Fell JW, editors. The yeasts (Fourth Edition). Amsterdam: Elsevier; 1998. pp. 248–53.
5. Dai Z, Pomraning KR, Deng S, Hofstad BA, Panisko EA, Rodriguez D, et al. Deletion of the KU70 homologue facilitates gene targeting in lipomyces starkeyi strain NRRL Y-11558. Curr Genet. 2019;65(1):269–82.
6. Zhang L, Lim EY, Loh K-C, Dai Y, Tong YW. Two-Stage fermentation of lipomyces starkeyi for production of microbial lipids and biodiesel. Microorganisms [Internet]. 2021; 9(8).
7. Di Fidio N, Dragoni F, Antonetti C, De Bari I, Raspolli Galletti AM, Ragaglini G. From paper mill waste to single cell oil: enzymatic hydrolysis to sugars and their fermentation into microbial oil by the yeast lipomyces starkeyi. Bioresour Technol. 2020;315:123790.
8. Islam MA, Yousuf A, Karim A, Pirozzi D, Khan MR, Wahid ZA. Bioremediation of palm oil mill effluent and lipid production by lipomyces starkeyi: A combined approach. J Clean Prod. 2018;172:1779–87.
9. Huang C, Chen X-F, Yang X-Y, Xiong L, Lin X-Q, Yang J, et al. Bioconversion of corncob acid hydrolysate into microbial oil by the oleaginous yeast lipomyces starkeyi. Appl Biochem Biotechnol. 2014;172(4):2197–204.
10. Pomraning KR, Collett JR, Kim J, Panisko EA, Culley DE, Dai Z, et al. Transcriptomic analysis of the oleaginous yeast lipomyces starkeyi during lipid accumulation on enzymatically treated corn Stover hydrolysate. Biotechnol Biofuels. 2019;12(1):162.
11. Dai Z, Deng S, Culley DE, Bruno KS, Magnuson JK. Agrobacterium tumefaciens-mediated transformation of oleaginous yeast lipomyces species. Appl Microbiol Biotechnol. 2017;101(15):6099–110.
12. Calvey CH, Willis LB, Jeffries TW. An optimized transformation protocol for lipomyces starkeyi. Curr Genet. 2014;60(3):223–30.
13. Vinoth Kumar R, Pakshirajan K, Pugazhenthi G. Chapter 9 - Malic and succinic acid: potential C4 platform chemicals for polymer and biodegradable plastic production. In: Kaur Brar S, Jyoti Sarma S, Pakshirajan K, editors. Platform chemical biorefinery. Amsterdam: Elsevier; 2016. pp. 159–79.
14. Wei Z, Xu Y, Xu Q, Cao W, Huang H, Liu H. Microbial biosynthesis of L-Malic acid and related metabolic engineering strategies: advances and prospects. Front Bioeng Biotechnol. 2021;9:765685.
15. Xu Y, Zhou Y, Cao W, Liu H. Improved production of malic acid in Aspergillus Niger by abolishing citric acid accumulation and enhancing glycolytic flux. ACS Synth Biol. 2020;9(6):1418–25.
16. Li J, Lin L, Sun T, Xu J, Ji J, Liu Q, et al. Direct production of commodity chemicals from lignocellulose using myceliophthora thermophila. Metab Eng. 2020;61:416–26.
17. Brown SH, Bashkirova L, Berka R, Chandler T, Doty T, McCall K, et al. Metabolic engineering of Aspergillus oryzae NRRL 3488 for increased production of l-malic acid. Appl Microbiol Biotechnol. 2013;97(20):8903–12.
18. Chen X, Xu G, Xu N, Zou W, Zhu P, Liu L, et al. Metabolic engineering of torulopsis glabrata for malate production. Metab Eng. 2013;19:10–6.

19. Xu Y, Shan L, Zhou Y, Xie Z, Ball AS, Cao W, et al. Development of a Cre-loxP-based genetic system in Aspergillus Niger ATCC1015 and its application to construction of efficient organic acid-producing cell factories. Appl Microbiol Biotechnol. 2019;103(19):8105–14.
20. Zelle Rintze M, de Hulster E, van Winden Wouter A, de Waard P, Dijkema C, Winkler Aaron A, et al. Malic acid production by Saccharomyces cerevisiae: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. Appl Environ Microbiol. 2008;74(9):2766–77.
21. Zhang T, Ge C, Deng L, Tan T, Wang F. C4-dicarboxylic acid production by overexpressing the reductive TCA pathway. FEMS Microbiol Lett. 2015;362(9):fnv052.
22. Yang L, Christakou E, Vang J, Lübeck M, Lübeck PS. Overexpression of a C4-dicarboxylate transporter is the key for rerouting citric acid to C4-dicarboxylic acid production in Aspergillus Carbonarius. Microb Cell Fact. 2017;16(1):43.
23. Radivojević T, Costello Z, Workman K, Garcia Martin H. A machine learning automated recommendation tool for synthetic biology. Nat Commun. 2020;11(1):4879.
24. Zournas A, Incha M, Radivojevic T, Blay V, Marti J, Costello Z et al. Machine learning-led semi-automated medium optimization reveals salt as key for Flaviolin production in Pseudomonas Putida. 2024.
25. Chen X, Jennings E, Shekiro J, Kuhn EM, O'Brien M, Wang W, et al. editors. Novel DDR Processing of Corn Stover Achieves High Monomeric Sugar Concentrations from Enzymatic Hydrolysis (230 g/L) and High Ethanol Concentration (10% v/v) During Fermentation2015 2015-04-03; United States. Research Org.: National renewable energy lab. (NREL), golden, CO (United States): Fairfax, VA: society for industrial microbiology and biotechnology; sponsor org.: USDOE Office of Energy Efficiency and Renewable Energy (EERE), Transportation Office. Bioenergy Technologies Office.
26. Chen X, Kuhn E, Jennings EW, Nelson R, Tao L, Zhang M, et al. DMR (deacetylation and mechanical refining) processing of corn Stover achieves high monomeric sugar concentrations (230 g L– 1) during enzymatic hydrolysis and high ethanol concentrations (> 10% v/v) during fermentation without hydrolysate purification or concentration. Energy Environ Sci. 2016;9(4):1237–45.
27. Chen X, Wang W, Ciesielski P, Trass O, Park S, Tao L, et al. Improving sugar yields and reducing enzyme loadings in the deacetylation and mechanical refining (DMR) process through multistage disk and Szego refining and corresponding Techno-Economic analysis. ACS Sustain Chem Eng. 2016;4(1):324–33.
28. Kuhn EM, Chen X, Tucker MP. Deacetylation and mechanical refining (DMR) and deacetylation and dilute acid (DDA) pretreatment of corn Stover, Switchgrass, and a 50:50 corn Stover/Switchgrass blend. ACS Sustain Chem Eng. 2020;8(17):6734–43.
29. Bilbao A, Munoz N, Kim J, Orton DJ, Gao Y, Poorey K, et al. PeakDecoder enables machine learning-based metabolite annotation and accurate profiling in multidimensional mass spectrometry measurements. Nat Commun. 2023;14(1):2461.
30. Pomraning KR, Dai Z, Munoz N, Kim YM, Gao Y, Deng S, et al. Integration of proteomics and metabolomics into the design, build, test, learn cycle to improve 3-Hydroxypropionic acid production in Aspergillus pseudoterreus. Front Bioeng Biotechnol. 2021;9:603832.
31. Riley R, Haridas S, Wolfe KH, Lopes MR, Hittinger CT, Göker M et al. Comp Genomics Biotechnologically Important Yeasts. 2016;113.
32. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
33. Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods. 2018;15(7):475–6.
34. Mowll JL, Gadd GM. Zinc uptake and toxicity in the yeasts sporobolomyces roseus and Saccharomyces cerevisiae. Microbiology. 1983;129(11):3421–5.
35. Takano J, Kobayashi M, Noda Y, Fujiwara T. Saccharomyces cerevisiae Bor1p is a Boron exporter and a key determinant of Boron tolerance. FEMS Microbiol Lett. 2007;267(2):230–5.
36. Yang M, Jensen LT, Gardner AJ, Culotta VC. Manganese toxicity and Saccharomyces cerevisiae Mam3p, a member of the ACDP (ancient conserved domain protein) family. Biochem J. 2005;386(Pt 3):479–87.
37. Chen OS, Hemenway S, Kaplan J. Genetic analysis of iron citrate toxicity in yeast: Implications for mammalian iron homeostasis. Proceedings of the National Academy of Sciences. 2002;99(26):16922-7.
38. Pimentel C, Caetano SM, Menezes R, Figueira I, Santos CN, Ferreira RB, et al. Yap1 mediates tolerance to Cobalt toxicity in the yeast Saccharomyces cerevisiae. Biochimica et biophysica acta (BBA) -. Gen Subj. 2014;1840(6):1977–86.
39. Adamo GM, Brocca S, Passolunghi S, Salvato B, Lotti M. Laboratory evolution of copper tolerant yeast strains. Microb Cell Fact. 2012;11(1):1.
40. Naude A, Nicol W. Malic acid production through the whole-cell hydration of fumaric acid with immobilised rhizopus oryzae. Biochem Eng J. 2018;137:152–61.
41. Côrte-Real M, Leão C. Transport of malic acid and other Dicarboxylic acids in the yeast Hansenula anomala. Appl Environ Microbiol. 1990;56(4):1109–13.
42. Bhagwat SS, Li Y, Cortés-Peña YR, Brace EC, Martin TA, Zhao H, et al. Sustainable production of acrylic acid via 3-Hydroxypropionic acid from lignocellulosic biomass. ACS Sustain Chem Eng. 2021;9(49):16659–69.
43. Czajka JJ, Han Y, Kim J, Mondo SJ, Hofstad BA, Robles A, et al. Genome-scale model development and genomic sequencing of the oleaginous clade lipomyces. Front Bioeng Biotechnol. 2024;12:1356551.
44. Sibirny AA. Yeast peroxisomes: structure, functions and biotechnological opportunities. FEMS Yeast Res. 2016;16(4):fow038.
45. Saier MH Jr., Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, et al. The transporter classification database (TCDB): 2021 update. Nucleic Acids Res. 2021;49(D1):D461–7.
46. Girstmair H, Tippel F, Lopez A, Tych K, Stein F, Haberkant P, et al. The Hsp90 isoforms from S. cerevisiae differ in structure, function and client range. Nat Commun. 2019;10(1):3626.
47. Welker S, Rudolph B, Frenzel E, Hagn F, Liebisch G, Schmitz G, et al. Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. Mol Cell. 2010;39(4):507–20.
48. Linger J, Tyler JK. The yeast histone chaperone chromatin assembly factor 1 protects against double-strand DNA-damaging agents. Genetics. 2005;171(4):1513–22.
49. Jayakody LN, Jin Y-S. In-depth Understanding of molecular mechanisms of aldehyde toxicity to engineer robust Saccharomyces cerevisiae. Appl Microbiol Biotechnol. 2021;105(7):2675–92.
50. Hassan M, Blanc PJ, Granger L-M, Pareilleux A, Goma G. Influence of nitrogen and iron limitations on lipid production by Cryptococcus curvatus grown in batch and fed-batch culture. Process Biochem. 1996;31(4):355–61.
51. Wu S, Hu C, Jin G, Zhao X, Zhao ZK. Phosphate-limitation mediated lipid production by Rhodosporidium toruloides. Bioresour Technol. 2010;101(15):6124–9.
52. Chopra J, Sen R. Process optimization involving critical evaluation of oxygen transfer, oxygen uptake and nitrogen limitation for enhanced biomass and lipid production by oleaginous yeast for biofuel application. Bioprocess Biosyst Eng. 2018;41(8):1103–13.
53. Morgunov IG, Kamzolova SV, Lunina JN. The citric acid production from Raw glycerol by Yarrowia lipolytica yeast and its regulation. Appl Microbiol Biotechnol. 2013;97(16):7387–97.
54. Kamzolova SV, Samoilenko VA, Lunina JN, Morgunov IG. Effects of medium components on isocitric acid production by Yarrowia lipolytica yeast. Ferment [Internet]. 2020; 6(4).
55. Zhang J, Petersen SD, Radivojevic T, Ramirez A, Pérez-Manríquez A, Abeliuk E, et al. Combining mechanistic and machine learning models for predictive engineering and optimization of Tryptophan metabolism. Nat Commun. 2020;11(1):4880.
56. Uluisik I, Kaya A, Unlu ES, Avsar K, Karakaya HC, Yalcin T, et al. Genome-wide identification of genes that play a role in Boron stress response in yeast. Genomics. 2011;97(2):106–11.
57. Liu J, Xie Z, Shin H-d, Li J, Du G, Chen J, et al. Rewiring the reductive Tricarboxylic acid pathway and L-malate transport pathway of Aspergillus oryzae for overproduction of L-malate. J Biotechnol. 2017;253:1–9.
58. Liu J, Li J, Liu Y, Shin H-d, Ledesma-Amaro R, Du G, et al. Synergistic rewiring of carbon metabolism and redox metabolism in cytoplasm and mitochondria of Aspergillus oryzae for increased l-Malate production. ACS Synth Biol. 2018;7(9):2139–47.
59. Singh M, Kumar J, Singh S, Singh VP, Prasad SM. Roles of osmoprotectants in improving salinity and drought tolerance in plants: a review. Reviews Environ Sci Bio/Technology. 2015;14(3):407–26.
60. Jilani SB, Prasad R, Yazdani SS. Overexpression of oxidoreductase YghA confers tolerance of furfural in ethanologenic Escherichia coli strain SSK42. Appl Environ Microbiol. 2021;87(23):e0185521.
61. Liu C-G, Cao L-Y, Wen Y, Li K, Mehmood MA, Zhao X-Q, et al. Intracellular redox manipulation of zymomonas mobilis for improving tolerance against lignocellulose hydrolysate-derived stress. Chem Eng Sci. 2020;227:115933.

62. Fu H, Hu J, Guo X, Feng J, Yang S-T, Wang J. Butanol production from Saccharina Japonica hydrolysate by engineered clostridium tyrobutyricum: the effects of pretreatment method and heat shock protein overexpression. Bioresour Technol. 2021;335:125290.
63. Alriksson B, Horváth IS, Jönsson LJ. Overexpression of Saccharomyces cerevisiae transcription factor and multidrug resistance genes conveys enhanced resistance to lignocellulose-derived fermentation inhibitors. Process Biochem. 2010;45(2):264–71.

Publisher's note