Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation

Raphael Ferreira, Christos Skrekas, Alex Hedin, Benjamín Sánchez, Verena Siewers, Jens Nielsen, and Florian David

Supplementary File

Materials and Methods

- Plasmid constructions
- Strain constructions
- Flux balance analysis
- Next-generation sequencing analysis
- Library design
- 3-Hydroxypropionic acid measurements
- R Shiny App

Supplementary Figures

Material and Methods

Plasmid constructions

pFDA09 (TEF1p-BS123-GFP w/ NLS-fapR) [1] constructed by David et al. was used as the malonyl-CoA biosensor plasmid and the positive control of the malonyl-CoA biosensor respectively (Table S1). All the plasmids in the present study were constructed using Gibson Assembly Master Mix (New England Biolabs), consisting of T5 exonuclease, DNA polymerase, and DNA ligase. All the fragments were purified, quantified with a Nanodrop and were incubated at 50°C for one hour with a molecular concentration ratio of insert:backbone 3:1. The reaction was used to transform competent *E. coli* DH5α cells and plasmid extraction and verification of the construct by digestion and sequencing followed.

The plasmid pDTU-113 was used for all the dCas9-based experiments [1]. pDTU-113 contains a gene encoding dCas9 fused with VPR activator, a TetR repressor gene and a gRNA cloning site under the control of the RNA pol III promoter pRPR1 with a *tetO* operator. The TetR-*tetO* system enables inducible gRNA expression and gRNAs transcription was induced by the addition of anhydrotetracycline (aTc) to the growth medium. The gRNA library oligos were ordered from Twist Bioscience (San Francisco, CA, USA) and were amplified with the primers gRNA-F and gRNA-R (Table S1). These primers contain a 100bp tail to increase recombination efficiency. The PCR amplified library was cloned directly in yeast cells following the high-efficiency transformation protocol developed by Benatuil and colleagues (Fig. S1a) [2]. Approximately 4 µg of digested backbone and 12 µg of PCR amplified gRNA library were used for transformation of a 100 mL fresh yeast culture of OD_{600} ~ 1.6. The library was grown on appropriate selective plates and the aim was to take a number of colonies at least three times larger than the library size, to ensure full coverage.

Strain constructions

CEN.PK113-11C (MATa *SUC2 MAL2-8c his3Δ1 ura3-52*; P. Kötter, University of Frankfurt Germany) [3] was utilized as the background strain for all the experiments (Table S1). HXT7p-ACC1** strain was obtained from Gossing and colleagues [4]. For constructing MCR01 the malonyl-CoA reductase, *mcr* gene, was received by Li et al. (Unpublished data). The expression of the bifunctional MCR has been separated into two subdomains. The C-terminal MCR was expressed downstream of a *HXT7* promoter and N-terminal MCR under a *TDH3* promoter with an UAS upstream of the promoter. The two subdomains were amplified with the primers

MCR_part1 (fw/rv) and MCR_part2 (fw/rv) (Table S1). The amplified MCR was genetic integrated into XI-3 in the *S. cerevisiae* genome, with CRISPR/Cas9 technique.

E. coli strain DH5α (*fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) [5] was used as a host for the construction of plasmids. *E. coli* strains were grown in lysogeny broth (LB) (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl and pH was set to 7.0 with NaOH). Selective LB Amp medium additionally contained 100 mg/L ampicillin. For LB and LB Amp plates, 16 g/L agar was added. Yeast extract peptone dextrose (YPD) was used as rich *S. cerevisiae* medium (20 g/L glucose, 10 g/L yeast extract, and 20 g/L peptone from meat). For the selection of kanMX marker containing cells, 200 mg/L of G418 was added. For YPD and YPD+G418 plates, 20 g/L agar was added. For selection for auxotrophy markers in yeast (*URA3*, *HIS3* or both), synthetic dextrose (SD) dropout medium was used (6.7 g/L yeast nitrogen base without amino acids, 0.77 g/L of Complete Supplement Mixture (CSM) (Sigma-Aldrich, St. Louis, MO, USA) without uracil or histidine or 0.75 g/L of CSM without both uracil and histidine and 20 g/L glucose). For SD dropout plates, 20 g/L agar were added. DELFT medium was used as a minimal synthetic medium for yeast growth. DELFT composition was 7.5 g/L $(NH_4)_2SO_4$, 14.4 g/L KH_2PO_4, 0.5 g/L $MgSO_4$-$7H_20$, 22 g/L dextrose, 2 mL/L trace metals solution, and 1 mL/L vitamins [6]. The pH was adjusted to 6. Sterilization of the media and all the other materials was done by autoclaving them at 121°C for 15 minutes. For the glucose-containing media, glucose was autoclaved separately. Standard growth temperatures for liquid cultures and plates were 30°C for *S. cerevisiae* and 37°C for *E. coli.* Liquid cultures of both microorganisms were grown aerobically at 30°C in shake flasks with a shaking speed of 200 rpm. In order to achieve aerobic conditions, the total volume of the culture was not exceeding 1/5 of the total volume of the shake flask. For small cultures of 5-10 mL 50 mL falcon tubes were used and for 1-2 mL cultures 15 mL culture tubes were used. To monitor the growth of *S. cerevisiae* cultures, the OD_{600} of the culture was measured. Plasmids from *E. coli* DH5α strains grown overnight in selective conditions were purified by using the GeneJET Plasmid Miniprep Kit (Thermo Scientific, Waltham, MA, USA). For PCR reactions that required high fidelity (e.g. promoters, ORFs, etc.) Phusion high-fidelity DNA polymerase (Thermo Scientific) or PrimeSTAR HS DNA polymerase (Takara, Clontech, Kusatsu, Japan) were used. For other PCR reactions, DreamTaq DNA Polymerase (Thermo Scientific) was used. For each PCR reaction the protocol of the enzyme that was used was followed and the melting temperature (T_m) of each primer pair was determined by using the NEB T_m calculator tool (New England Biolabs, Ipswich, MA, USA). The resulting PCR products were purified

with the GeneJET PCR Purification Kit (Thermo Scientific). Plasmids and/or PCR products were digested by using FastDigest restriction enzymes (Thermo Scientific). Visualization of the DNA was done by gel electrophoresis in 1% agarose, 1X TAE (Tris base, acetic acid, and EDTA) buffer, and 0.05 μl/mL GelRed (Thermo Scientific). Gel extraction of DNA bands was done by electrophoresis in 1% agarose, 1X TAE buffer, and 0.1 μl/mL GelGreen. The gel was displayed in Visi-Blue transilluminator (UVP, Upland, CA, USA), the desired band was excised and purified with the GeneJET PCR Purification Kit. All DNA concentrations were determined by using Nanodrop (Thermo Scientific). Sequencing of plasmids or PCR products was done by Eurofins Genomics (Ebersberg, Germany). Colony PCR for *S. cerevisiae* was performed by suspending a bit of biomass in 20-25 μL of 20 mM NaOH, incubating at 95^0C for 40-45 minutes and centrifuging at 13.000 rpm for 10 minutes. Then, 1 μL of the supernatant was used as a template for the PCR reaction. Maintenance of DNA and microbial strains. Isolated DNA (plasmids circular or linearized, PCR products etc.) was kept in -20°C diluted in TE buffer or sterilized dH_2O. For short term storage, transformation plates or plates containing streaks of the desired strain could be kept at 4°C for 2 months maximum. For long term storage, an overnight grown culture of the desired *E. coli* or *S. cerevisiae* strain with the selective pressure that was needed in each case was used for preparation of glycerol stock of final total volume 1000 mL and final concentration of glycerol of 15% v/v.

Flux balance analysis

Flux Balance Analysis (FBA) was performed using the consensus genome-scale metabolic model of yeast as it was revised by Aung et al. 2013 (Yeast v7.6) [7,8]. The analysis was performed for two carbon sources: glucose and ethanol. Once the maximum specific growth rate (μ_{max}) for each carbon source was identified, 11 suboptimal growth rates were applied to the model, in the interval $0.3*\mu_{max}$ to $0.8*\mu_{max}$ in increments of $0.05*\mu_{max}$. For each suboptimal growth rate, two analyses were performed, one for maximum cytosolic acetyl production and one for maximum cytosolic malonyl production. In all simulations parsimonious FBA (pFBA) was performed, meaning that the solution for each analysis was the one with the minimal sum of fluxes [9].

For each reaction and condition, a *k*-score was calculated as $k = v/v_{\mu max}$, in a fashion similar to previous studies [10]. This score compares the flux of each reaction in each simulation with the flux of the same reaction under maximum growth rate conditions. Therefore, reactions that show *k*-scores > 1 have an increase in the corresponding flux compared to the μ_{max} condition

and *k*-scores < 1 have a decrease in the corresponding flux. The maximum *k*-score was arbitrarily set to 100 for *k*-scores that are infinite (i.e. when $v_{\mu max} = 0$). Reactions that showed inconsistencies between the 11 different suboptimal growth rates (i.e. some k < 1 and some k > 1) were filtered out. Afterwards, reactions were associated with the genes that encode the enzyme(s) that catalyze each reaction. In terms of gene expression, it was assumed that a *k*-score > 1 represented upregulation and a *k*-score < 1 represented downregulation of the gene expression levels. For genes that take part in more than one reaction, average k scores were calculated. Genes that showed average *k*-scores between 0.549 and 1.001 were ruled out as non-significantly changed. The value of 0.549 was chosen because it corresponds to the value of a flux that decreases proportionally to the decrease in biomass production ((0.3+0.35+…+0.8)/11) and not to any additional increase towards acetyl/malonyl. Finally, FVA was computed for all cases to show if any genes were associated to reactions that exhibited flux variability, which leads to *k*-scores that vary depending on the flux variability and could correspond to false positives [11]. Said genes were marked as such in Table S2.

Library design

For each gene retrieved from the FBA, 15-21 gRNAs were designed using CRISPR-ERA and Yeast CRISPRi [12,13]. The binding site of each gRNA was identified by the distance from the transcription start site (TSS) [13]. For each gene, 2-5 gRNAs were designed to target a window from -50 to +50 from the TSS. The total library contained 3194 gRNAs in total.

Fluorescence measurements and sorting

For real-time monitoring of GFP expression, a BioLector ® was used (m2p-labs GmbH, Baesweiler, Germany). The cultures were grown in using FlowerPlates ®, the total volume of each culture was 1 mL and initial OD_{600} was 0.05-0.1. GFP expression levels were measured by the ratio of green fluorescence/biomass. For the FACS assays, the FACS Aria® cell sorter was used (BD Biosciences, San Jose, CA, USA). Cryostocks of the yeast library and control (no library) were grown in synthetic minimal media at an initial OD600 = 0.1. After 3 hours of culture, aTc (1000x) was added to induce gRNA expression. At t7h, sample from the culture was taken for OD_{600} measurement and FACS. An initial round of 20.000 events was performed to evaluate the overall fluorescence difference between the library and the control (Fig. S1b). Two gates were subsequently defined for sorting, namely P2 that contained cells that showed medium to high fluorescence (approximately top 2% of the population) and P3 that contained cells with extremely high fluorescence (approximately top 3% of the population) (Fig. S1b).

Next-generation sequencing analysis

After each round of sorting, plasmids of the sorted yeast libraries were extracted using Zymoprep Yeast Plasmid Miniprep (Zymo Research, CA, USA). Libraries were prepared based on Illumina DNA Nextera Sequencing and next-generation sequencing was performed on a MiSeq Benchtop Sequencer (Illumina, San Diego, CA) as described in Lee et al. [14]. Fastq sequences were analyzed using a custom R script. The sequences between the end of RPR1p (5'-CGATTGGCAG) and beginning of RPR1t (5'-GTTTTAGAGC) were probed and matched to the gRNA library. The number of hits, i.e. the number of times a gRNA is matched to the gRNA library, was quantified and normalized by the total number of NGS read counts for each run. The most enriched gRNAs were retrieved by taking the base 2 log of the ratio of the normalized gRNA hits of each sorting over the NGS from the preculture [15], and plotted using the R library ggplot2 3.3.0.

3-Hydroxypropionic acid measurements

All culture samples were centrifuged at 12,000g, and the supernatants were diluted 1:5 with 0.5 mM H_2SO_4. The concentration of 3-HP was conducted with high-performance liquid chromatography (HPLC, Dionex UltiMate 3000; Thermo Fisher Scientific, Waltham, MA, USA), equipped with an Aminex HPX-87H (Bio-Rad, Hercules, USA) column at 65°C. Each sample was analyzed with a mobile phase of 0.5 mM H_2SO_4 at a flow rate of 0.5 mL/ min for 35 min.

R Shiny App

An app was built using the R programming language and the R Shiny application framework (http://CRAN.R-project.org/package=shiny). The complete list of gRNAs was previously obtained from Yeast CRISPRi and integrated into the app [12]. A FBA was previously performed to optimize each metabolite present in Yeast 7.6, and the list k scores for each gene was integrated into the app.

Supplementary figures

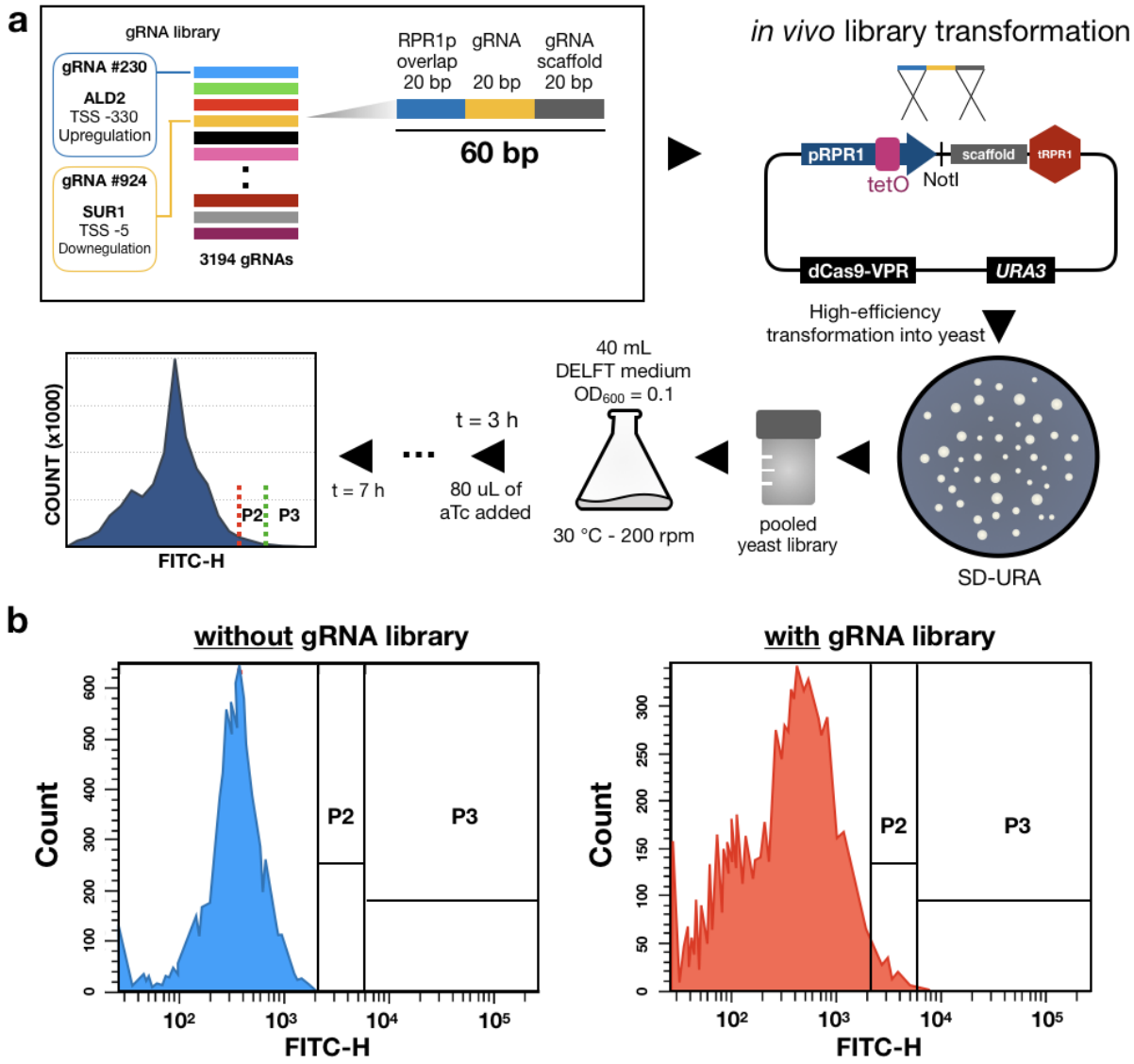

Figure S1. Experimental design of library generation and sorting. a. A 60 bp oligo library totaling 3194 gRNAs was cloned into a centromeric plasmid under control of a aTc RNA pol III inducible promoter *in vivo* (CEN.PK-11C + HXT7p-ACC1** + FapR-based malonyl-CoA biosensor) and the resulting transformants grown on appropriate selective plates. Around 10,000 grown colonies were pooled to ensure approximate three-time full coverage of the initial library. Aliquots of the pooled library were grown in DELFT medium and aTc was added after 3 h of growth to induce gRNA expression. After 7 h and 10 h of growth, the OD_{600} was measured and 200,000 cells were FACS-sorted based on two fluorescent gate values. **b.** Fluorescent gating and sorting strategy. Two gates were selected for sorting, one that contained cells that showed medium to high fluorescence (P2) and one that contained cells with extremely high fluorescence (P3). The gating was based on the Acc1** background strain with the library.

Here, the HXT7p-ACC1** strain without the library (left) shows few very few cells on P2 and P3.

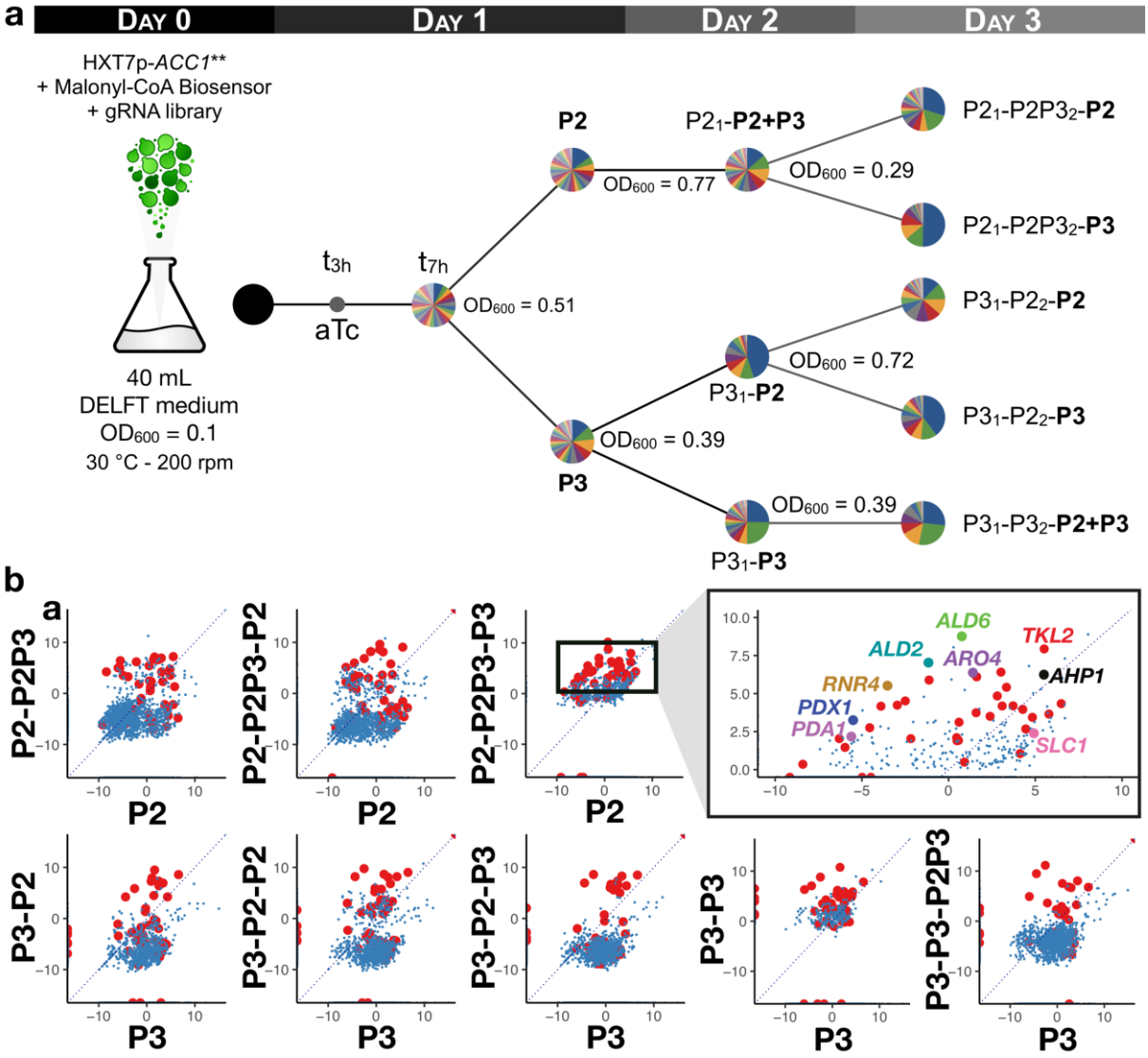

Figure S2. Three-day sorting strategy. a. Sorting at t_{7h}. The yeast library was subsequently sorted using two different gates, one for medium fluorescence, P2, and one for high fluorescence, P3, except for $P2_1$-P2+P3 where $P2_1$-P2 and $P2_1$-P3 were accidentally mixed together. Each pie chart represents the gRNA distribution of their respective sorting. **b.** Log2 fold-change gRNAs over the initial library (pre-cultured). Shown in red the most enriched gRNAs that were tested for 3-HP production.

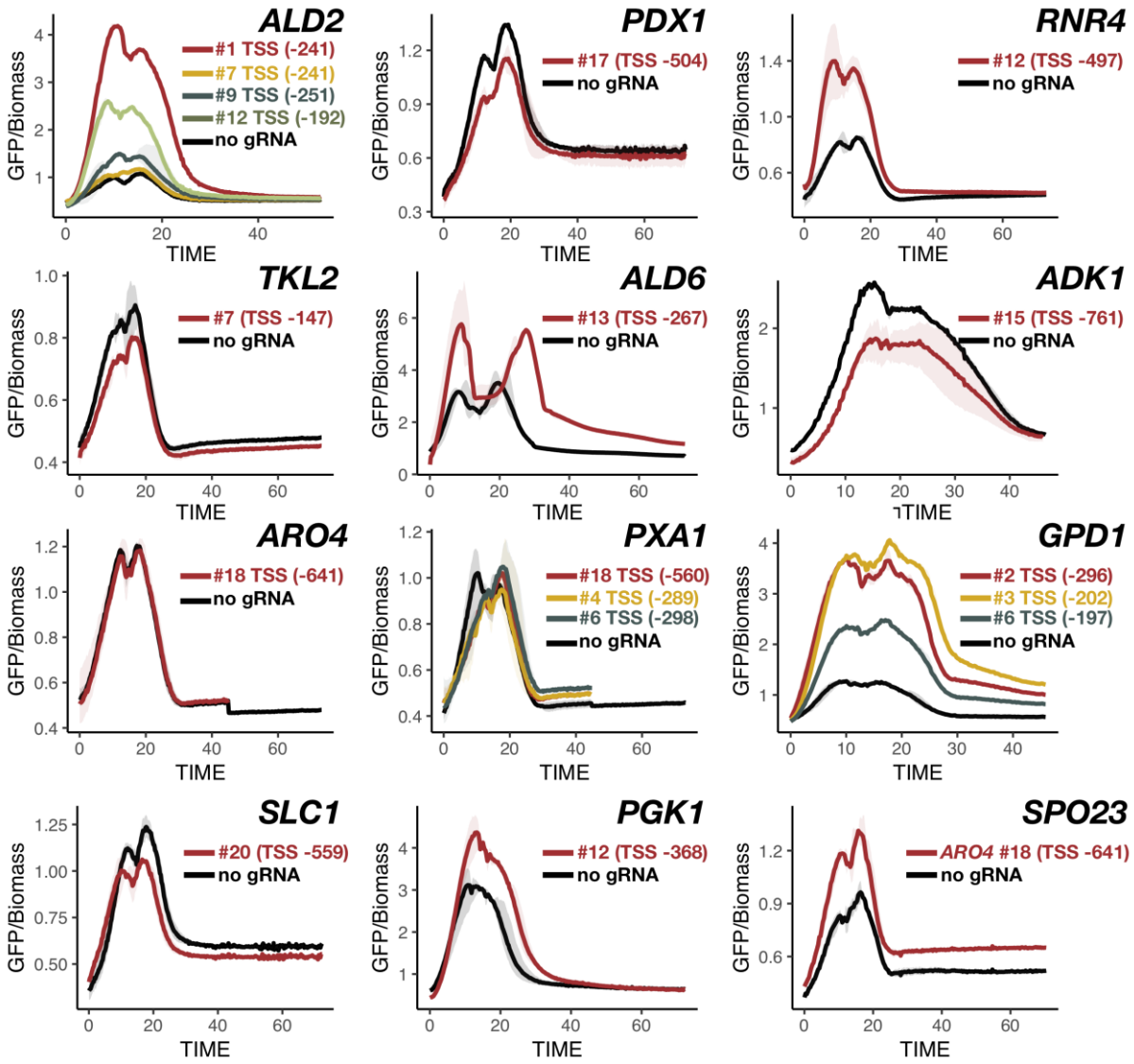

Figure S3. Transcriptional regulation. Mean fluorescence intensities (GFP/OD) were obtained from three biological replicates ± S.D. monitored with a BioLector (m2p-labs).

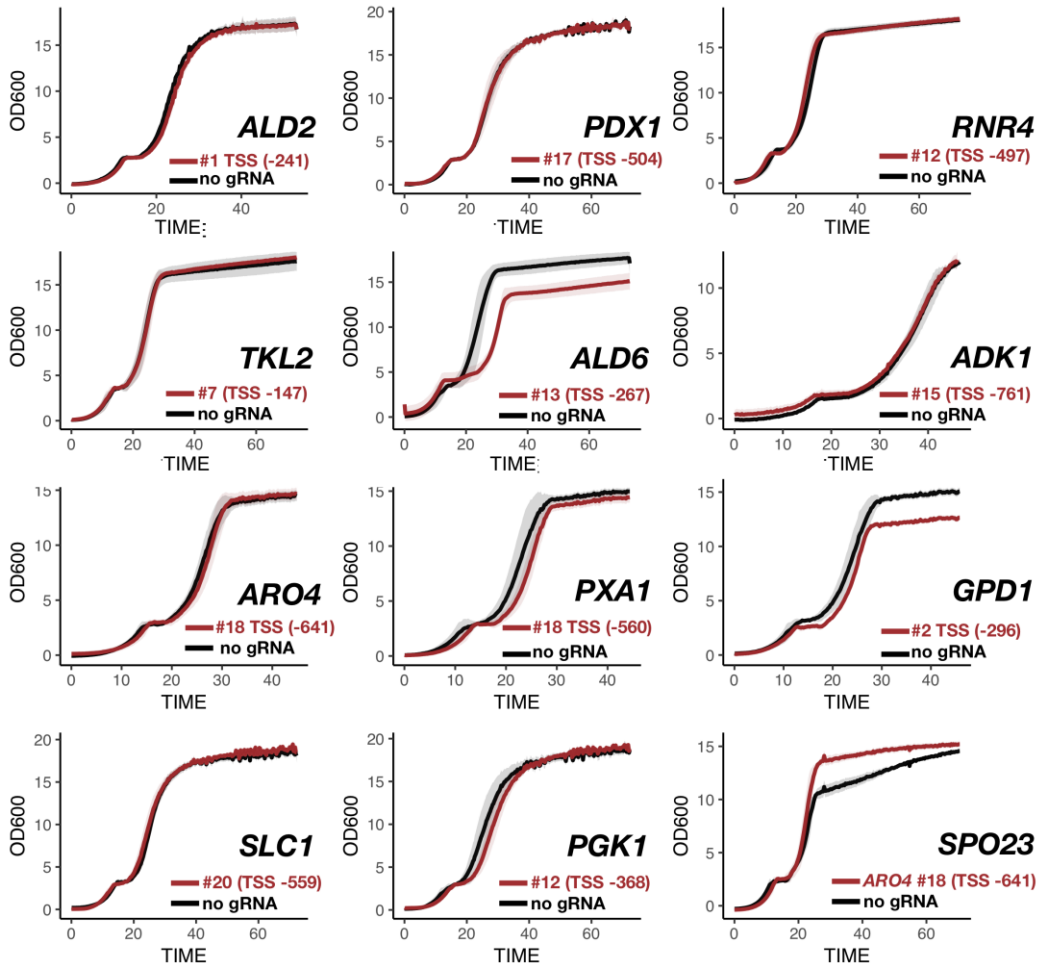

Figure S4. Effect of the gRNAs on growth. Optical densities (OD_{600}) were obtained from three biological replicates ± S.D. monitored with a BioLector.

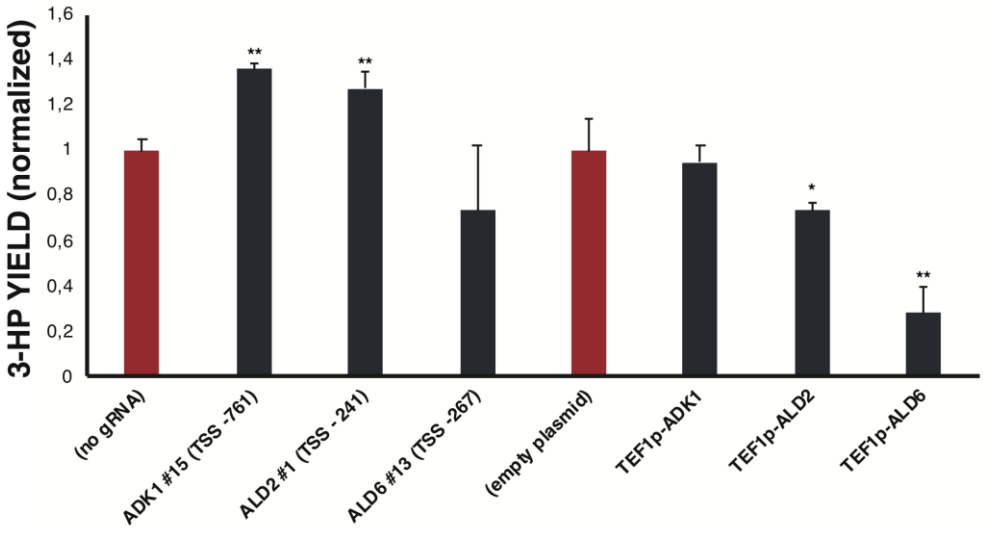

Figure S5. Fine-tuning comparison of three upregulating gRNAs against expressing the targeted genes under the strong constitutive TEF1p. Displayed are averages ± SD of triplicates grown in defined minimal medium with 20 g/L glucose and cultures were sampled after 72 h. Displayed are averages ± SD. ** *p* value < 0.01; * *p* value < 0.05; (Student's t-test: one-tailed, two-sample equal variance).

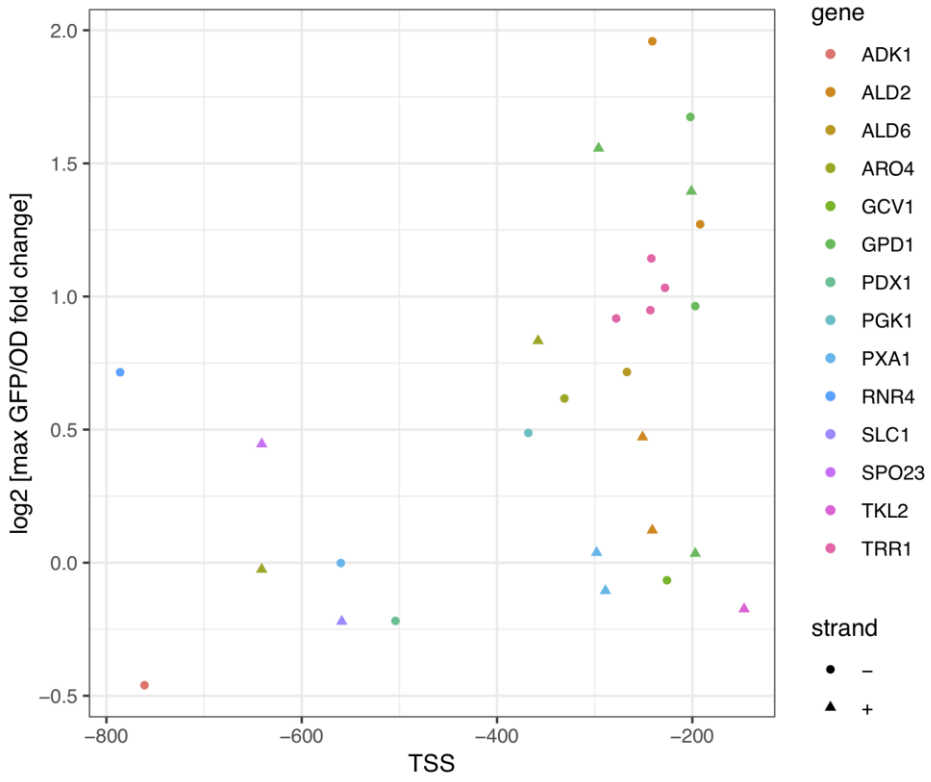

Figure S6. Effect of target location and strand on transcriptional regulation. Displayed in y axis, gRNA efficacy measured log2 of the average maximum value of GFP divided by OD_{600} normalized by control (no gRNA). In the x axis, the distance to TSS. The shape and the color of each points represent the strand and the gene targeted by the gRNA, respectively.

References

(1) Jensen, E. D.; Ferreira, R.; Jakočiūnas, T.; Arsovska, D.; Zhang, J.; Ding, L.; Smith, J. D.; David, F.; Nielsen, J.; Jensen, M. K.; et al. Transcriptional Reprogramming in Yeast Using DCas9 and Combinatorial GRNA Strategies. *Microb. Cell Factories* **2017**, *16* (1), 46.

(2) Benatuil, L.; Perez, J. M.; Belk, J.; Hsieh, C.-M. An Improved Yeast Transformation Method for the Generation of Very Large Human Antibody Libraries. *Protein Eng. Des. Sel.* **2010**, *23* (4), 155–159.

(3) van Dijken, J. P.; Bauer, J.; Brambilla, L.; Duboc, P.; Francois, J. M.; Gancedo, C.; Giuseppin, M. L. F.; Heijnen, J. J.; Hoare, M.; Lange, H. C.; et al. An Interlaboratory Comparison of Physiological and Genetic Properties of Four Saccharomyces Cerevisiae Strains. *Enzyme Microb. Technol.* **2000**, *26* (9), 706–714.

(4) Gossing, M.; Smialowska, A.; Nielsen, J. Impact of Forced Fatty Acid Synthesis on Metabolism and Physiology of Saccharomyces Cerevisiae. *FEMS Yeast Res.* **2018**, *18* (8).

(5) Taylor, R. G.; Walker, D. C.; McInnes, R. R. E. Coli Host Strains Significantly Affect the Quality of Small Scale Plasmid DNA Preparations Used for Sequencing. *Nucleic Acids Res.* **1993**, *21* (7), 1677–1678.

(6) Verduyn, C.; Postma, E.; Scheffers, W. A.; Dijken, J. P. V. Effect of Benzoic Acid on Metabolic Fluxes in Yeasts: A Continuous-Culture Study on the Regulation of Respiration and Alcoholic Fermentation. *Yeast* **1992**, *8* (7), 501–517.

(7) Aung, H. W.; Henry, S. A.; Walker, L. P. Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Ind. Biotechnol.* **2013**, *9* (4), 215–228.

(8) Orth, J. D.; Thiele, I.; Palsson, B. Ø. What Is Flux Balance Analysis? *Nat. Biotechnol.* **2010**, *28* (3), 245–248.

(9) Lewis, N. E.; Hixson, K. K.; Conrad, T. M.; Lerman, J. A.; Charusanti, P.; Polpitiya, A. D.; Adkins, J. N.; Schramm, G.; Purvine, S. O.; Lopez-Ferrer, D.; et al. Omic Data from Evolved E. Coli Are Consistent with Computed Optimal Growth from Genome-Scale Models. *Mol. Syst. Biol.* **2010**, *6* (1), 390.

(10) Choi, H. S.; Lee, S. Y.; Kim, T. Y.; Woo, H. M. In Silico Identification of Gene Amplification Targets for Improvement of Lycopene Production. *Appl. Environ. Microbiol.* **2010**, *76* (10), 3097–3105.

(11) Mahadevan, R., Schilling, C. H. (2003) The Effects of Alternate Optimal Solutions in Constraint-Based Genome-Scale Metabolic Models. *Metab. Eng.* 5 (4), 264 – 276

(12) Smith, J. D.; Suresh, S.; Schlecht, U.; Wu, M.; Wagih, O.; Peltz, G.; Davis, R. W.; Steinmetz, L. M.; Parts, L.; St.Onge, R. P. Quantitative CRISPR Interference Screens in Yeast Identify Chemical-Genetic Interactions and New Rules for Guide RNA Design. *Genome Biol.* **2016**, *17* (1), 45.

(13) Liu, H.; Wei, Z.; Dominguez, A.; Li, Y.; Wang, X.; Qi, L. S. CRISPR-ERA: A Comprehensive Design Tool for CRISPR-Mediated Gene Editing, Repression and Activation. *Bioinformatics* **2015**, *31* (22), 3676–3678.

(14) Lee, J. S.; Kallehauge, T. B.; Pedersen, L. E.; Kildegaard, H. F. Site-Specific Integration in CHO Cells Mediated by CRISPR/Cas9 and Homology-Directed DNA Repair Pathway. *Sci. Rep.* **2015**, *5*, 8572.

(15) Joung, J.; Konermann, S.; Gootenberg, J. S.; Abudayyeh, O. O.; Platt, R. J.; Brigham, M. D.; Sanjana, N. E.; Zhang, F. Genome-Scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening. *Nat. Protoc.* **2017**, *12* (4), 828–863.