CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Online Research Database In Technology

Technical University of Denmark

DTU

Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in Saccharomyces cerevisiae

Strucko, Tomas; Zirngibl, Katharina; Pereira, Filipa; Kafkia, Eleni; Tharwat Tolba Mohamed, Elsayed; Rettel, Mandy; Stein, Frank; Feist, Adam; Jouhten, Paula; Patil, Kiran Raosaheb; Förster, Jochen

Published in:
Metabolic Engineering

Link to article, DOI:
10.1016/j.ymben.2018.03.006

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Strucko, T., Zirngibl, K., Pereira, F., Kafkia, E., Mohamed, E. T., Rettel, M., ... Forster, J. (2018). Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in Saccharomyces cerevisiae. Metabolic Engineering. DOI: 10.1016/j.ymben.2018.03.006

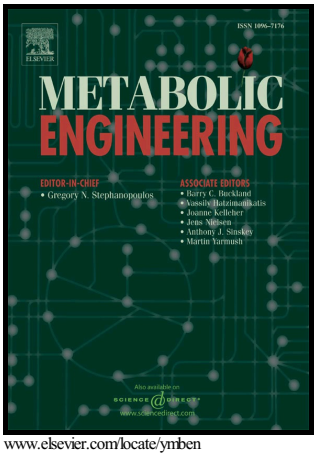
DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Author's Accepted Manuscript



www.elsevier.com/locate/ymben

Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in *Saccharomyces cerevisiae*

Tomas Strucko, Katharina Zirngibl, Filipa Pereira, Eleni Kafkia, Elsayed T. Mohamed, Mandy Rettel, Frank Stein, Adam M. Feist, Paula Jouhten, Kiran Raosaheb Patil, Jochen Forster

PII: S1096-7176(17)30403-2
DOI: https://doi.org/10.1016/j.ymben.2018.03.006
Reference: YMBEN1360

To appear in: *Metabolic Engineering*

Received date: 1 November 2017
Revised date: 6 March 2018
Accepted date: 6 March 2018

Cite this article as: Tomas Strucko, Katharina Zirngibl, Filipa Pereira, Eleni Kafkia, Elsayed T. Mohamed, Mandy Rettel, Frank Stein, Adam M. Feist, Paula Jouhten, Kiran Raosaheb Patil and Jochen Forster, Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in *Saccharomyces cerevisiae*, *Metabolic Engineering,* https://doi.org/10.1016/j.ymben.2018.03.006

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in *Saccharomyces cerevisiae*

Tomas Strucko[1,2], Katharina Zirngibl[2], Filipa Pereira[2], Eleni Kafkia[2], Elsayed T. Mohamed[1], Mandy Rettel[2], Frank Stein[2], Adam M. Feist[1,3], Paula Jouhten[2], Kiran Raosaheb Patil[2,*], Jochen Forster[1,1]

[1]Technical University of Denmark, Novo Nordisk Foundation Center for Biosustainability, Hørsholm, Denmark

[2]European Molecular Biology Laboratory, Structural and Computation Biology Unit, Heidelberg, Germany

[3]Department of Bioengineering, University of California, 9500 Gilman Drive La Jolla, San Diego, CA 92093, USA.

[*]Corresponding author. patil@embl.de

Accepted manuscript

Abstract

Most microbial species, including model eukaryote *Saccharomyces cerevisiae*, possess genetic capability to utilize many alternative nutrient sources. Yet, it remains an open question whether these manifest into assimilatory phenotypes. Despite possessing all necessary pathways, *S. cerevisiae* grows poorly or not at all when glycerol is the sole carbon source. Here we discover, through multiple evolved lineages, genetic determinants underlying glycerol catabolism and the associated fitness trade-offs. Most evolved lineages adapted through mutations in the HOG pathway, but showed hampered osmotolerance. In the other lineages, we find that only three mutations cause the improved phenotype. One of these contributes counter-intuitively by decoupling the TCA cycle from oxidative phosphorylation, and thereby hampers ethanol utilization. Transcriptomics, proteomics and metabolomics analysis of the re-engineered strains affirmed the causality of the three mutations at molecular level. Introduction of these mutations resulted in

[1] Current address: Carlsberg A/S, Carlsberg Research Laboratory, Copenhagen, Denmark

ACCEPTED MANUSCRIPT

improved glycerol utilization also in industrial strains. Our findings not only have a direct relevance for improving glycerol-based bioprocesses, but also illustrate how a metabolic pathway can remain unexploited due to fitness trade-offs in other, ecologically important, traits.

Keywords

Adaptive laboratory evolution; causal mutations; multi-omics analysis

Introduction

Glycerol is a major by-product of *Saccharomyces cerevisiae* metabolism and can also serve as a carbon source. In complex growth media that include supplements such as amino acids, yeast extract or peptone, *S. cerevisiae* can assimilate glycerol (Merico et al., 2011; Swinnen et al., 2013) using dedicated transporters and catabolic pathways (**Figure 1a**). However, many *S. cerevisiae* isolates, including popular laboratory strains, show no or very poor growth in minimal glycerol (MG) medium (Swinnen et al., 2013). This is in contrast with the capabilities of yeast as reflected in its metabolic network. Several efforts have been made towards realizing this capability, mainly due to interest in glycerol as an alternative carbon source for biotechnological applications (da Silva et al., 2009; Ho et al., 2017; Klein et al., 2016a, 2016b; Merico et al., 2011; Ochoa-Estopier et al., 2011; Swinnen et al., 2016, 2013). Adaptive laboratory evolution (ALE) experiments have been particularly successful in showing that yeast metabolism can evolve to grow efficiently on glycerol (Ho et al., 2017; Merico et al., 2011; Ochoa-Estopier et al., 2011). The underlying mechanisms, however, remain unclear. Furthermore, this evolvability, in light of the poor performance of the wild-type *S. cerevisiae* in MG medium, raises the fundamental question of whether improved growth on glycerol compromises other ecologically important traits. To address this, we combined ALE with genetic crossing and omics analyses to reveal the causal mutations for glycerol metabolism and the associated trade-offs.

Accepted manuscript

ACCEPTED MANUSCRIPT

Results and Discussion

We started with ALE experiments and generated a collection of yeast strains that could efficiently grow in MG medium containing 1 % glycerol (**Figure 1b**). Two strain backgrounds were used for evolution, a wild-type (WT) strain (CEN.PK113-7D) and a NOX knock-in mutant (CEN.PK113-1A based) expressing a *nox* gene encoding for cytosolic water forming NADH oxidase (**Methods**). The NOX mutant was chosen as it was previously hypothesized that unbalanced levels of intracellular NADH hamper glycerol catabolism in MG media (Merico et al., 2011). While we did not observe improvements in the final growth performance in the evolved NOX lineages over those started with WT, the adaptation rates were faster (**Supplementary Fig 1**). We performed ALE experiments in two different modes: mode-I, short-term (up to 80 generations, manual transfers) and, mode-II, long-term (>300 generations, automated transfers). In mode-I we used two NOX based replicates, whereas in mode-II five parallel lineages of each WT and NOX were evolved. As it is generally observed that adaptation rates are highest in the early stages (Barrick and Lenski, 2013; Wiser et al., 2013), we also characterized the intermediate lineages from mode-I. Both ALE modes resulted in lineages with similar high efficiency growth on MG medium (mean $\mu_{max} \approx 0.22$); the best performing isolate being ALE7 with $\mu_{max} = 0.229 \pm 0.002$ (**Figure 1c, d, e,** and **Supplementary Fig 2**). Notably, the evolved lineages show a four-fold faster growth rate in MG medium compared to the parental strains in a rich medium containing amino acids.

Next, we sequenced intermediate and endpoint lineages of the mode-I, and the all endpoint lineages from the mode-II ALE experiments. As the mutation landscape is dynamic in evolving cultures and usually results in heterogeneous populations (Kao and Sherlock, 2008; Lang et al., 2013), whole population sequencing (as opposed to clonal) was chosen so as to capture the spectrum of mutations that occurred during the ALE process. Overall, 93 unique mutations were found across the 17 sequenced lineages; the most abundant were single nucleotide variations (SNV) and one nucleotide insertions or deletions (**Figure 1f** and **Supplementary Table 1**). Among these, we identified 32 gene-coding regions that contained non-synonymous mutations resulting in amino acid substitutions, translation frameshifts and/or premature stop codons (i.e., structural

Accepted manuscript

ACCEPTED MANUSCRIPT

mutations). These genes code for proteins involved in a range of functions, from metabolism and regulation to signaling. Interestingly, three genes were found to be frequently mutated in simultaneously evolved lineages (**Figure 1f**). Consistent with its known function in glycerol utilization, mutations in the glycerol kinase encoding gene, *GUT1,* were observed in most of the cases (**Supplementary Table 1**). Similarly, glycerol kinase gene was found to be a common target for mutations in *E. coli* (Herring et al., 2006) and *S. cerevisiae* (Ho et al., 2017) during adaptation to glycerol medium. The other two genes (*PBS*2 and *HOG1*) were found to be frequently mutated across different lineages. The *PBS*2 and *HOG1* are coding for the key proteins of the high-osmolarity glycerol (HOG) pathway. In agreement to our observations, (Kvitek and Sherlock, 2013) have reported that the major signaling pathways, including HOG, of *S. cerevisiae* were affected earliest during laboratory evolution. The latter is also true in the mode-I ALE experiment where the *PBS2* gene mutated as early as after five passages of the adaptation process (**Supplementary Table 1**). To confirm the role of the HOG pathway in glycerol utilization, we reintroduced the corresponding single and double mutations in the wild-type strains (**Figure 2a**, **Supplementary Fig 3** and **Supplementary Fig 4**). Although we could experimentally demonstrate the positive effect of these mutations, the full phenotype of the endpoint mode-I lineages could not be entirely recapitulated in any of the two background strains (WT or NOX, **Supplementary Fig 5**). Nevertheless, a clear positive interaction between *GUT1* and *PBS2* or *HOG1* was observed (**Figure 2b** and **Supplementary Fig 4**). Moreover, exclusively all HOG pathway related mutations caused loss-of-function phenotypes as these lineages displayed elevated sensitivity to osmotic stress (**Supplementary Fig 6**). To prove that the latter was attributed to single SNVs in the ALE lineages, we individually assessed the osmosensitivity of several variants of *HOG1* and *PBS2* that were re-engineered into the wild-type CEN.PK. In all tested cases, a single nucleotide substitution in the genes led to the osmosensitive phenotype (**Figure 2c**). As sensitivity to osmotic stresses is not a desired feature in biotechnology settings, we did not pursue in-depth analysis of any lineage (mode-I and mode-II) that harbored mutations affecting the genes of the HOG pathway.

To further dissect and identify causal mutations underlying the evolved glycerol phenotype, we focused on the endpoint ALE2 lineage from mode-II, as it displayed high growth on glycerol (μ_{max} =

Accepted manuscript

ACCEPTED MANUSCRIPT

0.220 ± 0.004) and did not exhibit osmosensitive phenotype (**Supplementary Fig 6**). Notably, this lineage had three-fold more mutations (12 ORFs affected) as compared to the other evolved lineages (~4 ORFs affected per lineage). To narrow down on the causative mutations among these, we backcrossed ALE2 with the parental strain (**Figure 2a** and **Methods**). The hybrid strains retaining the ALE2 phenotype were further crossed with the parental strain and this procedure was repeated three times. The mutation and glycerol phenotype segregation patterns were followed by sequencing the genomes of the three tetrads of the first generation (**Supplementary Table 1**). These 12 strains showed phenotypic heterogeneity with growth characteristics segregating in a non-Mendelian manner (not a 2:2 segregation). Moreover, growth traits continued to re-appear in the second and third generation crosses, supporting the complex nature of the glycerol utilization phenotype (**Figure 2d** and **Supplementary Fig 7**). We did not detect other than SNV types of mutations (i.e. gene copy variations or structural genome rearrangements) in ALE2 lineage based on DNA sequencing data (**Methods**) as well as on 2:2 segregation of mutations in subsequent tetrad analysis. One of the third-generation hybrid strains retained the ALE2 phenotype and harbored only four mutations. Two of the affected genes code for metabolic proteins; Gut1 – glycerol kinase and Kgd1 – alpha-ketoglutarate dehydrogenase, and the other two for globally acting signaling/regulatory proteins; Ubc13 – E2 ubiquitin-conjugating enzyme and Ino80 – nucleosome spacing factor (**Figure 2e** and **Supplementary Table 1**). To identify the causal mutations, we re-engineered the single, double and triple combinations of the mutated genes into the parental strain. None of the single reconstructed mutants grew in the MG medium after more than 80 hours of cultivation (**Supplementary Fig 3**). Although the gene pairs (except $UBC13^{ALE2}$-$INO80^{ALE2}$ and $GUT1^{ALE2}$-$INO80^{ALE2}$) showed synergistic growth improvement, none recapitulated the efficiency of the ALE2 lineage (**Supplementary Fig 4**). In contrast, the triple mutants not only recapitulated the phenotypic diversity of the backcrossed isolates, but also fully restored the ALE2 growth phenotype in the case of a mutant harboring $GUT1^{ALE2}$, $KGD1^{ALE2}$ and $UBC13^{ALE2}$ variants (hereafter referred to as the R-GKU strain) (**Figure 2d-f** and **Supplementary Fig 4**). At least three interacting mutations are thus necessary for efficient glycerol uptake in MG medium.

Accepted manuscript

ACCEPTED MANUSCRIPT

To investigate the effect of causal mutations at molecular level, we analyzed transcriptomics and proteomics profiles of selected re-engineered strains grown in well-controlled reactors (see **Methods**). While 271 genes showed differential expression when comparing the R-GKU strain with the ALE2 endpoint lineage, abundance of only three proteins had changed significantly (**Figure 2h**, **Supplementary Table 2** and **Supplementary Table 3**). The majority of the transcriptomics changes could be attributed (GO terms) to the opposite mating types of the R-GKU and ALE2 strains (**Supplementary Table 4**). Together with the growth physiology, the limited changes in gene/protein expression confirmed that the re-engineered R-GKU strain captures the major genetic contributors of efficient glycerol utilization.

As Gut1 catalyzes the first essential step in glycerol metabolism (Ho et al., 2017; Swinnen et al., 2013), *GUT1* is an expected target for adaptive mutations (**Figure 1a**). The *GUT1*^{ALE2} variant found in the ALE2 lineage has a single amino acid residue change (E572Q), and showed a substantial positive effect on glycerol utilization when combined with the other mutations (*KGD1*^{ALE2} and/or *UBC13*^{ALE2}) (**Supplementary Fig 4**). In contrast to *GUT1*, the roles of *UBC13* and *KGD1* were less intuitive and their association with glycerol metabolism has not been previously reported. Ubc13 is a E2 ubiquitin-conjugating enzyme involved in the error-free DNA post-replication repair pathway (Hofmann and Pickart, 1999).The *UBC13*^{ALE2} variant likely results in loss-of-function as the 476ΔG mutation caused a translation frame shift with a premature stop codon after the 71st amino acid. The variant thus lacks its well-described functional regions (VanDemark et al., 2001) (**Supplementary Fig 8**). Surprisingly, transcriptomics and proteomics analysis of R-GK (*GUT1*^{ALE2}-*KGD1*^{ALE2}), compared to R-GKU, showed no broad effect of the *UBC13*^{ALE2} mutation except for minor gene expression differences (**Figure 2j** and **Supplementary Table 2**). Furthermore, Ubc13 protein levels were noticeably lower in the R-GKU as compared to R-GK. This could also explain the considerably higher number of mutations that accumulated in the ALE2 lineage as it is known that the *Δubc13* mutant has an increased mutation rate (Brusky et al., 2000). In recent studies it was demonstrated that a truncated version of the Ubr2 (ubiquitin-protein ligase (E3)) protein had a positive effect in glycerol utilization (Ho et al., 2017; Swinnen et al., 2016). These studies, together with our findings, suggest a possibility that the decreased rates of protein ubiquitylation might be involved in the

Accepted manuscript

ACCEPTED MANUSCRIPT

glycerol utilization. However, the mode of action of Ubc13 and Ubr2 are thought to be different (Finley et al., 2012), and, thus, different mechanisms might be at play in these two cases. Nevertheless, the beneficial role of the *UBC13*^{ALE2} allele in the glycerol phenotype points to a potential role of Ubc13 in glycerol catabolism.

The third causal mutation in R-GKU resides in one of the genes (*KGD1)* coding for the α-ketoglutarate dehydrogenase complex, which carries out a key step in the tricarboxylic acid (TCA) cycle. The mutation, C2969A, caused an A990D substitution that resulted in a loss of enzymatic activity as the strains with the *KGD1*^{ALE2} variant resembled the *Δkgd1* phenotype with impaired diauxic shift (no growth on ethanol (Repetto and Tzagoloff, 1989)). We have also confirmed that solely *KGD1*^{ALE2} variant was responsible for this deficiency (**Figure 2g**). In further support, the *Δkgd1* phenotype was segregated in a 2:2 manner during our back-crossing experiments (**Supplementary Fig 9**). At the level of protein expression, the *KGD1*^{ALE2} allele did not impact cellular processes other than those closely linked to TCA cycle (**Figure 2i** and **Supplementary Table 2**). To understand the role of the loss-of-function of Kgd1 in improving glycerol metabolism, we compared, using a genome-scale metabolic model of *S. cerevisiae* (Förster et al., 2003)**,** flux distributions (including flux variability) between growth on glucose and glycerol (**Methods**). Interestingly, optimal growth on glycerol (with minimum changes to flux distribution corresponding to optimal glucose utilization) required down-regulation of the Kgd1 flux concomitant with an up-regulation of oxidative phosphorylation. This implies decoupling of the relative activities of TCA cycle and oxidative phosphorylation (**Figure 3a**). Lower relative TCA cycle activity is required for efficient glycerol utilization since glycerol carbons are more reduced than glucose (or other common carbon sources like ethanol) carbons and yield more NADH per C-mole when metabolized. Oxidative phosphorylation is yet required to regenerate NAD^+. The ability of the *KGD1*^{ALE2} allele mutants to metabolize glycerol, which requires oxidative phosphorylation, combined with impaired ethanol utilization shows that this decoupling is indeed the case. Supporting this, the transcriptomics and proteomics data (R-GKU vs R-GU) showed decreased expression of *CIT3* (encoding citrate synthase, an initial TCA cycle enzyme) and Dld3 (2-hydroxyglutarate transhydrogenase, on both gene and protein level). The latter enzyme is also acting on the Kgd1 substrate α-ketoglutarate and

Accepted manuscript

ACCEPTED MANUSCRIPT

supports the modelling results predicting an increased flux through the GABA shunt. When the GABA shunt is active (*i.e.* used for L-glutamate degradation) a pathway intermediate has been suggested to trigger the repression of *DLD3* (Liu and Butow, 1999). This GABA shunt activation as a TCA cycle bypass could enhance glycerol utilization by generating NADPH (in succinate semialdehyde dehydrogenase reaction) instead of NADH (in α-ketoglutarate dehydrogenase reaction). Lastly, the modelled metabolic flux alterations were supported by changes observed in the TCA cycle metabolites (**Figure 3a** and **Supplementary Table 5**). The identified Kgd1 loss-of-function mutation thus reveals a novel and non-intuitive link between the TCA cycle operation and glycerol metabolism.

We next asked whether the identified mutations would be causative beyond the laboratory CEN.PK strain. To test this, we re-engineered the three causal mutations of ALE2 in two industrial strains originating from different geographic locations, namely wine yeast L.1528 from Chile and beer yeast CLIB382 from Ireland. Both wild-type strains could sustain some growth in MG medium, but only after an extensive 100+ hours of lag phase. Introduction of the three mutations noticeably improved the growth of both strains (**Figure 3b**), showing the importance of the *GUT1*, *KGD1* and *UBC13* modulations for glycerol utilization also in industrial yeasts.

Conclusions

Taken together, the results of this ALE-driven study show that efficient growth on glycerol as a sole carbon source is a complex trait requiring synergistic interactions between, in this case, at least three genes, including those in metabolic pathways and regulatory processes. From a biotechnological perspective, the identified causal mutations and the corresponding mechanisms could be readily applied for improving glycerol-based bioprocesses. Interestingly, glycerol utilization efficiency appears to be associated with fitness trade-offs with regards to osmosensitivity or ethanol utilization. These trade-offs offer an explanation to why yeast isolates show sub-optimal or no capacity for glycerol metabolism. From an evolutionary standpoint, these would heavily compromise *S. cerevisiae* survival strategies, as growth in high glucose concentration environments as well as the subsequent utilization of ethanol are essential in its natural habitats (Jouhten et al., 2016; Zhou

Accepted manuscript

ACCEPTED MANUSCRIPT

et al., 2017). In a broader perspective, these results show how metabolic capability of a species can remain latent, and how it can be uncovered through laboratory evolution.

Materials and Methods

Molecular cloning

All DNA fragments that constituted genetic elements (*i.e.*, promoters, open reading frames (ORFs) and terminators) or plasmids backbone sequences were amplified by PCR using PfuX7 (Nørholm, 2010) polymerase with specific primers denoted in the **Supplementary Table 6**. The *TEF1* promoter sequence was amplified from the pSP-GM1 (Partow et al., 2010) plasmid, and the *nox* gene from gDNA of *Streptococcus pneumoniae* SV1.

Integrative plasmid pTS1 targeting the specific chromosomal site X-3 (described earlier (Mikkelsen et al., 2012)) was assembled by the uracil-specific excision reagent (USER™) following well established protocols (Geu-Flores et al., 2007). Specifically, the *TEF1* promoter and the *nox* was USER-cloned into *AsiSI/Nb.BsmI* linearized pCfB2223 (Stovicek et al., 2015a) plasmid backbone resulting in pTS1.

The construction of plasmids harboring single guide RNAs (gRNAs) expression cassette targeting specific genetic loci were constructed as follows. Unique linear fragments were obtained by PCR amplifying the pCfB2311 (Stovicek et al., 2015b) plasmid backbone with the generic 5`-phosphorylated primer (TS109) in combination with the specific primer for each genetic target (see **Supplementary Table 6**). Thereafter, each fragment was independently circularized via blunt end ligation by T4 ligase (Thermo Fisher Scientific) according to the manufacturer's recommendations.

The construction of pTS83 vector encoding three gRNAs cassettes was assembled by USER-cloning as described previously (Jakočiūnas et al., 2015). In short, three gRNA cassettes targeting *GUT1, UBC13* and *KGD1* loci were independently amplified from plasmids pTS53, pTS56 and pTS70, respectively. Subsequently, the resulting PCR fragments were USER cloned into *AsiSI/Nb.BsmI* linearized pTAJAK71 (Jakočiūnas et al., 2015) vector.

Accepted manuscript

ACCEPTED MANUSCRIPT

All genetic constructs were validated using Mix2Seq sequencing (Eurofins Genomics). For the full list of all plasmids used and constructed in this work see **Supplementary Table 7**.

Strains and cultivation media

Escherichia coli DH5α strain used for maintenance and amplification of cloned plasmids, and was propagated in 2xYT medium (Sigma) supplemented with 100 mg/L of Ampicillin (Sigma). *Saccharomyces cerevisiae* strains used in this study were prototrophic laboratory haploid strains CEN.PK113-1A and CEN.PK113-7D, and industrially relevant diploid strains L.1528 and CLIB382, see **Supplementary Table 8**.

For maintenance and genetic transformation of yeast strains a yeast extract peptone dextrose (YPD) medium containing 10 g/L of yeast extract, 20 g/L of peptone and 20 g/L of glucose was used. Solid YPD medium was prepared by addition of 20 g/L of agar prior autoclavation. For selection of yeast strains with dominant markers NatMX, KanMX or HphMX, YPD medium was supplemented (after autoclavation) by 100 mg/L of nourseothricin (ClonNat, Werner BioAgents), 200 mg/L of G418 disulfate salt (Sigma) or 200 mg/L hygromycin B (Sigma), respectively. For high osmotic stress sensitivity assays the YPD medium was supplemented with of potassium chloride to a final concentration of 0.5 moles/L. Plates with sporulation (SPO) medium were prepared as described elsewhere (Sherman, 2002).

Adaptive laboratory evolution and strain characterization was done in a well-defined mineral (M) media described by (Verduyn et al., 1992) containing 5 g/L $(NH_4)_2SO_4$, 3 g/L KH_2PO_4, 0.75 g/L Mg_2SO_4, 1.5 mL/L trace metal solution and 1.5 mL/L vitamins solution. The composition of the trace metal solution is 3 g/L $FeSO_4 \cdot 7H_2O$, 4.5 g/L $ZnSO_4 \cdot 7H_2O$, 4.5 g/L $CaCl_2 \cdot 6H_2O$, 0.84 g/L $MnCl_2 \cdot 2H2O$, 0.3 g/L $CoCl_2 \cdot 6H_2O$, 0.3 g/L $CuSO_4 \cdot 5H_2O$, 0.4 g/L $NaMoO_4 \cdot 2H_2O$, 1 g/L H_3BO_3, 0.1 g/L KI and 15 g/L $Na_2EDTA \cdot 2H_2O$. The vitamin solution includes 50 mg/L d-biotin, 200 mg/L para-amino benzoic acid, 1.0 g/L nicotinic acid, 1.0 g/L Ca-pantothenate, 1.0 g/L pyridoxine-HCl, 1.0 g/L thiamine-HCl and 25 mg/L minositol. The carbon source in the M medium was either 10 mL/L of glycerol or 30 g/L of glucose resulting in MG or MD media, respectively. The pH was adjusted with KOH/H_2SO_4 to 4.2 for the MG and to 6.5 for the MD medium. For the initial stage of the adaptive

Accepted manuscript

ACCEPTED MANUSCRIPT

laboratory evolution MG medium was additionally supplemented with 1.92 g/L of Y1501 amino acid mix (Sigma) and was denoted as MG+ medium. For small scale cultivations, the media was filter-sterilized by a bottle-top (0.45 µm pore size) filter (VWR). For the 1L batch fermentation experiments medium was heat-sterilized, sterile vitamin solution and glycerol were added after medium cooled down to 30 °C.

Yeast genetic transformation

All genetic modifications of *S. cerevisiae* laboratory and industrial strains were done using well described (Lithium acetate, PEG and ssDNA) transformation protocol (Gietz and Schiestl, 2007). Routinely, 200-500 ng of plasmid DNA and 0.5-1 µg of linear DNA was used per transformation. For CRISPR/Cas9 genome editing purpose, 2.5 mM of 90-bp long dsOligo was used as a repair template. For industrial cerevisiae strains the amount of 90-bp dsOligo was doubled. All transformants were selected on YPD medium plates supplemented accordingly with appropriate selection drug/s. Finally, single colonies were streak-purified on a selection medium prior further analyses and subsequent genetic manipulations.

Adaptive laboratory evolution

Two starting *S. cerevisiae* strains, the wild-type CEN.PK113-7D (WT) and the TS29 (NOX) strain expressing water forming NADH oxidase from *S. pneumoniae* were used in adaptive laboratory evolution (ALE) experiments. Moreover, ALE was done in two modes – I) manual ALE and II) using automated robotic set-up.

The mode-I ALE was performed in 500 mL shake flasks with 50 mL of medium using only the *NOX* expressing strain. Prior to ALE, two cultures of NOX strain were preconditioned by 72 hours cultivation in MG+ medium at 30 °C and constant agitation at 250 rpm. Subsequently, cultures were re-inoculated into the 50:50 of MG+ and MG media mix and cultivated until the stationary growth phase was reached. Thereafter, ALE was done exclusively in the MG medium by serial transfer of yeast cultures into fresh medium at late-exponential/stationary growth phase. The fresh cultures were inoculated to a starting OD_{600} 0.1 to 0.3. The ALE experiment lasted for up to 80 cumulative

Accepted manuscript

ACCEPTED MANUSCRIPT

generations in MG media. Cryostocks of the intermediate cultures were prepared at regular intervals.

For the for the mode-II ALE experiment, two strains WT and NOX were pre-cultured in two separate 500 mL shake flasks with 50 mL of MG+ medium. Five replicates with 15 mL of MG+ medium were inoculated to a starting OD_{600} of 0.3 per each strain. The tubes were cultured at 30 °C with constant agitation at 1000 rpm. A total of 900 µL of culture was serially passaged to fresh medium during early-exponential phase. The OD_{600} was automatically measured at regular intervals to assess cultures growth state. The growth medium composition was gradually changed from the MG+ to the MG medium (mixing from 100:0 to 0:100) during the ALE experiment. The last 300+ generations were evolved purely in the MG medium. Cryostocks of the intermediate cultures were prepared at regular intervals, however, only the final ALE lineages were used in further analyses.

Classical genetics techniques

Classical genetic techniques were done according to standard protocols (Sherman, 2002) with slight modifications. In brief, diploids were generated by combining two medium sized single colonies of the haploid strains with opposite mating type in 200 µL of sterile water in a 1.5 mL Eppendorf tube and vortexing vigorously. Thereafter, 10 µL of the suspension was plated onto YPD plate and incubated for 4~6 hours at 30 °C. The cell mix was then scraped out from the YPD plate and resuspended in 400 µL of sterile water. 10 µL of the suspension is plated on YPD plate and formed zygotes were isolated using spore dissection microscope. The ploidy of and mating type was confirmed by multiplex colony PCR on *MAT* locus (Illuxley et al., 1990) using the primers MAT_R, MATα_F and MATa_F.

The sporulation was induced by plating diploid cells on SPO plates and incubating at 30 °C for up to 6 days depending on sporulation efficiency. After confirming the presence of tetrads on SPO plates a small portion of biomass with spores was resuspended in 50 µL of (2.5 mg/mL) sterile Glucanex® (Thermo Fisher Scientific) solution and digested for up to 15 min at approx. 25 °C. The reaction was stopped by adding 450 µL of sterile MilliQ water and up to 5 µL of the resulting suspension was carefully transferred to a dissection plate. Tetrads were dissected using Axio Scope.A1(Carl Zeiss)

Accepted manuscript

ACCEPTED MANUSCRIPT

microscope equipped with dissection platform. Plates with dissected spores were incubated at 30 °C for 2 days at 30 °C.

Re-engineering mutations found in evolved strains

In order to re-engineer the mutations of interest in the wild type strains the CRISPR-Cas9 techniques (Jakočiūnas et al., 2015; Stovicek et al., 2015b) optimized for the *S. cerevisiae* were used. For each modification, a specific gRNA sequence (**Supplementary Table 6**) targeting the Cas9 nuclease to the appropriate genetic locus was designed. The quality and specificity of gRNA was assessed using a *CRISPRdirect* online tool developed by Naito et al. (Naito et al., 2015). In order to repair a DNA double strand break introduced by Cas9 repair templates (90-bp dsOligo flanking 45 bp up- and 45 bp downstream of the specific cut site) were designed for each locus of interest. Each repair template contained a specific mutation and a silent mutation that would disturb (NGG) PAM site motive (**Supplementary Table 6**).

First, the two laboratory strains CEN.PK113-7D and CEN.PK113-1A were transformed (as described above) with the Cas9 expressing plasmid pCfB2312 resulting in strains 7D_Cas9 and 1A_Cas9, respectively. Yeast transformants were selected on YPD+G418 plates. Subsequently, the Cas9 expressing strains were individually transformed with different single gRNA expressing plasmids (see **Supplementary Table 7**) and resulting cells were selected on YPD+G418+CloNat. To confirm that the gene editing was successful routinely 5 colonies per edit were tested. A 500 bp long DNA fragment flanking the locus of an edit was amplified by colony-PCR using OneTaq® 2X Master Mix (New England Biolabs) with specific primers (**Supplementary Table 6**), column purified using NucleoSpin® kit (MACHEREY-NAGEL) and sent for sequencing (Eurofins Genomics). Each engineered strain harboring a correct genetic edit was streaked on YPD+G418 plates and incubated for 2-3 days at 30 °C. Subsequently, yeast strains were replica-plated on YPD+G418+CloNat and YPD+G418 media in order to select for the mutants that have lost gRNA expressing plasmid. Next, yeast cells with a single genetic edit (without the corresponding gRNA plasmid) were transformed with a new gRNA expressing plasmid targeting a different locus. Subsequently, gRNA plasmids were “kicked out” from the correct strains harboring a double genetic edit. The transformation cycle

Accepted manuscript

ACCEPTED MANUSCRIPT

was repeated for generation of the strains containing triple gene mutations (re-engineered strains are listed in the **Supplementary Table 8**).

Characterization of growth in microtiter scale

All evolved and re-engineered yeast strains were characterized in microtiter scale setting using the Growth Profiler 1152 systems (Enzyscreen). An overnight pre-culture was prepared by inoculating each strain into a well of 24-deep well plate (Porvair Sciences) filled with 3 mL of YPD medium and incubating at 30 °C with 300 rpm shaking. Next day, the plate with the pre-culture was spun-down at 2200 g for 5 min and resuspended in 3 mL of MG medium. 200 µL aliquots of resuspended pre-culture was transferred to an according volume of fresh MG medium to result in a pre-inoculum suspension of OD_{600} 4.5. Finally, 50 µL of each pre-inoculum was inoculated into a separate well of Krystal 24-well clear bottom white microplate (Porvair Sciences) prefilled with 700 µL per well of MG media and incubated for minimum of 80 hours at 30 °C with 225 rpm shaking. Cell growth (green color intensity, G-value) was monitored by scanning the bottom of the plates in 30 min intervals. G-value was converted to an OD_{600} equivalent by using a spline fitted calibration curve using the data from (Bergdahl, **Unpublished data**). For growth characterization in MD medium the procedure was done exactly as described above. Growth rates were estimated by calculating the maximum slope values of best linear fit on log-transformed OD values (minimum of 10 data points within values 1 < OD < 9 were used).

Controlled batch fermentation

The evolved lineage ALE2 and re-engineered strains TS154(R-GU), TS170(R-GK) and TS177(R-GKU) from YPD plate were inoculated to 0.5 L shake flasks with 100 mL of MG (pH 4.2). Pre-cultures were incubated in an orbital shaker set to 200 rpm at 30 °C until late-exponential phase (OD_{600} 5 ~ 7). Cell suspension was up-concentrated by centrifugation and resuspension in fresh MG medium and used for inoculation. Batch cultivations were performed under aerobic conditions in one liter Sartorius fermenters equipped with continuous data acquisition (Braun Biotech International). Each fermenter was inoculated to an initial OD_{600} of 0.2. Cell culture aeration was ensured by constant airflow of 1.5 v.v.m. (80 L/h) and stirring speed of 1000 rpm. The temperature was

Accepted manuscript

ACCEPTED MANUSCRIPT

maintained at 30 °C during the fermentation and pH (4.2) level was controlled by automatic addition of 2M NaOH solution. The exhaust gas composition was constantly monitored by off gas analyzer 1311 Fast response triple gas (Innova) combined with Mass Spectrometer Prima Pro Process MS (ThermoFisher Scientific).

The batch cultures were sampled in regular intervals for estimation of OD_{600}, cell dry weight (CDW) and extracellular metabolites. Samples for transcriptomic, proteomics and intracellular metabolomics were taken at the early-exponential growth state (OD_{600} 2~5). All experiments were done in triplicates except for the strain TS154 (R-GU), which was done in duplicate.

Cell dry weight sampling

The biomass concentration was determined by measuring CDW as previously described (Olsson and Nielsen, 1997) using polyethersulfone filters with a pore size of 0.45 µm Montamil® (Membrane Solutions, LLC). The filters were pre-dried in a microwave oven at 150 W for 20 min and weighted on analytical scales. 5 mL of cultivation broth was filtered and then washed with three volumes of distilled water. Thereafter, the filters with biomass were dried in the microwave oven at 150 W for 20 min and cooled down in a desiccator for a minimum of 2 hours. The filters with dried biomass were weighed in order to determine the CDW.

Extracellular metabolite analysis

Samples for quantification of extracellular metabolites were prepared by filtering (0.20 µm Phenex-RC, Phenomenex) 1.5 mL of fermentation broth into glass vials and stored at -20 °C until further analyses. Glycerol and several metabolites of the central carbon metabolism were analyzed by high performance liquid chromatography (HPLC system, Waters) equipped with Rezex ROA-Organic Acid column (Phenomenex). The column temperature was set to 65 °C and elution was done by sulfuric acid 0.5 mM with constant flow-rate 0.5 mL/min. Metabolites were detected by RI differential refractometer (Waters) and PDA detector (Waters) at the 210 nm wavelength.

Accepted manuscript

ACCEPTED MANUSCRIPT

Genomic DNA sample preparation and analysis of genomic variants

Genomic DNA of *S. cerevisiae* strains was isolated using a ZR Fungal/Bacterial DNA MiniPrep™ kit (Zymoresearch). DNA was extracted following the manufacturer's recommendations, except that yeast cells were disrupted by five cycles of 1 min vortexing and 1 min on ice. The quality and the concentration of extracted DNA was assessed with the spectrophotometer NanoPhotometer® P-Class (IMPLEN). 150 bp pair-end DNA libraries were prepared using TruSeq Nano DNA HT Library Prep Kit and sequenced using Miseq™ platform (Illumina). The CLC Genomics and an in house developed pipeline was used to map sequencing reads and identify mutations relative to the *S. cerevisiae* CEN.PK113-7D genome (Nijkamp et al., 2012). In short all sequencing reads were passed through quality control with FastQC (Andrews, 2010) (version 0.11.3), followed by adapter trimming using cutadapt (Martin, 2011) (version 1.9.1) with default options. Subsequent quality trimming and filtering was performed with FaQCs (Lo and Chain, 2014) (version 1.34) using default parameters. Picard Tools (http://broadinstitute.github.io/picard/) (version 1.129) were used for sequencing file formatting and the removal of read duplicates. The genome-wide detection of single-nucleotide variants (SNVs) and small structural variants (SVs) was performed with the GATK HaplotypeCaller (McKenna et al., 2010) (version 3.3.0) using default settings. Post-processing and manual filtering of the raw vcf files was conducted according to the GATK Best Practices (DePristo et al., 2011) recommendations for germline variant detection. In addition, while analyzing the tetrad genomes, only the variants with a 2:2 segregation pattern were kept. Structural variants (SV) were investigated with delly2 (Rausch et al., 2012) (version 0.7.2). All samples had an average mapped coverage of at least 40 reads. Mutations of interest were confirmed by Sanger sequencing 500 bp long fragments of the loci of interest obtained by PCR with specific primers.

RNA-seq sample preparation and differential expression analysis

All RNA samples were prepared as follows, 10 mL of fermentation broth was sprayed into 50 mL Falcon® tube filled with ice and immediately centrifuged at 10000 x *g* for 5 min at 4 °C. After centrifugation supernatant was discarded and cell pellet was frozen by placing the tube into dry-ice

Accepted manuscript

ACCEPTED MANUSCRIPT

bath. Tubes with frozen biomass were kept at -80 °C until extraction. Total RNA of each sample was isolated using RNAeasy kit (Qiagen) by following manufacturer's recommendations. Briefly, 594 µL of RLT buffer plus 6 µL of β-mercaptoethanol were added to the Falcon® tube containing the frozen cell pellet and let it unfreeze on ice. Cell suspension was transferred to an ice-cold FastPrep Cap tube containing 600 µL of glass beads (400 nm acid washed, Sigma). Cells were disrupted using FastPrep (2 cycles with the following conditions: 10 seconds at speed 6, 15 seconds on ice). Cell lysate was transferred to a new tube and centrifuged 2 min at full speed in microcentrifuge (Eppendorf). Supernatant was careful mixed with 1 volume of 70 % HPLC-grade ethanol. Sample was transferred to an RNAeasy column and washed according to the manufacturer's instructions. RNA was then eluted with 60 µL of RNase-free water. Eluted sample was digested with Turbo DNAse (Invitrogen Ambion) accordingly to manufacture instructions followed by RNA clean-up (RNAeasy kit, Qiagen).

RNA library was prepared using the Illumina TruSeq Stranded mRNA LT sample prep kit starting with 500 ng of total RNA, following manufactures instructions using Beckman Biomek FX Laboratory automation station. Samples were sequences using Hiseq2000 instruments in the 50 bp single read mode and loaded 8 pM onto the flowcell at the Genomics Core Facility of EMBL (Heidelberg, Germany).

The quality of the raw RNA sequencing reads was assessed using FastQC (Andrews, 2010) (version 0.11.3). Prior to the alignment, adapter trimming was performed using cutadapt (Martin, 2011) (version 1.9.1) with default options providing the standard lllumina TrueSeq Index adapters. Subsequent quality trimming and filtering was performed with FaQCs (Lo and Chain, 2014) (version 1.34) using the following parameters: `-q 20 -min_L 25 -n 5 -discard 1`. The total reads per sample after trimming and filtering ranged from 17.5 to 27 million. The sequencing reads were aligned to the reference genome of *S. cerevisiae* CEN.PK113-7D (Nijkamp et al., 2012). using tophat2 (D. Kim et al., 2013) (version 2.0.10) with the following parameter: `-G -T -x 20 -M --microexon-search --no-coverage-search --no-novel-juncs -a 6`. Only reads with unique mappings were considered for differential expression analysis. Gene level count tables were

Accepted manuscript

obtained using the count script of the HTSeq (Anders et al., 2015) python library (version 0.6.1p1.) with default options. All reads mapped in total to about 5400 genes. The following statistical analysis was performed using the Bioconductor package DESeq2 (Love et al., 2014) (version 1.12.4). Size-factor based normalization to control for batch effects and inter-sample variability and dispersion estimation were conducted using the package defaults. The differential expression analysis was again performed with the package defaults, which include multiple testing correction, independent filtering and cooks cutoff (Anders and Huber, 2010) for outlier. Raw P-values as returned by DESeq2 were used as input to fdrtool (Strimmer, 2008) (version 1.2.15) in order to compute q-values (false discovery rates (FDRs)). Genes with a FDR < 0.1 were considered as significantly differentially expressed. Biostatistical analyses were conducted using R V.3.3.1 (R Development Core Team).

Proteomics sample preparation and data analysis

For proteomics analysis 10 mL of fermentation broth was transferred into ice-cold 15 mL Falcon® and immediately centrifuged at 10000 x *g* for 2 min at 4 °C. After centrifugation supernatant was discarded and cell pellet was washed once with PBS buffer. Pellet was frozen by placing the tube into dry-ice bath. Frozen samples were kept at -80 °C until extraction. Cell pellets were lysed using 0.1 % RapiGest in 100 mM ammonium bicarbonate. Three cycles of sonication (Cell disruptor, Sonifier, Branson) were applied to the lysate (1 cycle: 15 seconds sonication, 15 seconds on ice), followed by 15 min bead beating using Precellys Lysing Kit (KT0361-1-004.2). Cell lysate was transferred into a new tube after centrifugation (5 min, 5000 x *g*) and incubated at 80 °C for 15 min. Benzonase (25U, Merck) was added to the lysate for 30 min at 37 °C. Cysteines were reduced using dithiothreitol (56 °C, 30 min, 10 mM). The sample was cooled to 24 °C and alkylated with iodacetamide (room temperature, in the dark, 30 min, 10 mM). Proteins were TCA precipitated, TCA pellet was washed by acetone and dried. The proteins were digested in 50 mM HEPES (pH 8.5) using LysC (Wako) with an enzyme to protein ration 1:50 at 37 °C for 4 hours, followed by trypsin (Promega) with an enzyme to protein ratio 1:50 at 37 °C overnight. TMT10plex™ Isobaric Label

Accepted manuscript

ACCEPTED MANUSCRIPT

Reagent (ThermoFisher) was added to the samples according the manufacturer's instructions. Labeled peptides were cleaned up using OASIS® HLB µElution Plate (Waters).

Offline high pH reverse phase fractionation was performed using an Agilent 1200 Infinity high-performance liquid chromatography (HPLC) system, equipped with a Gemini C18 column (3 µm, 110 Å, 100 x 1.0 mm, Phenomenex). The solvent system consisted of 20 mM ammonium formate (pH 10.0) as mobile phase (A) and 100 % acetonitrile as mobile phase (B).

Peptide analysis by LC-MS/MS

Peptides were separated using the UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5 µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (Acclaim PepMap 100 75 µm x 50 cm C_{18}, 3 µm, 100 Å). The outlet of the analytical column was coupled directly to a QExactive plus (Thermo) using the proxeon nanoflow source in positive ion mode. Solvent A was water, 0.1 % formic acid and solvent B was acetonitrile, 0.1 % formic acid. Trapping time was 6 min with a constant flow of solvent A at 30 µL/min onto the trapping column. Peptides were eluted via the analytical column a constant flow of 0.3 µL/min. During the elution step, the percentage of solvent B increased in a linear fashion from 2 % to 4 % B in 4 min, from 4 % to 8 % in 2 min, then 8 % to 28 % for a further 96 min, and finally from 28 % to 40 % in another 10 min. Column cleaning at 80 % B followed, lasting 3 min, before returning to initial conditions for the re-equilibration, lasting 10 min.

The peptides were introduced into the mass spectrometer (QExactive plus, ThermoFisher) via a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (New Objective) and a spray voltage of 2.3 kV was applied. The capillary temperature was set at 320 °C. Full scan MS spectra with mass range 350-1400 *m/z* were acquired in profile mode in the FT with resolution of 70,000. The filling time was set at maximum of 100 ms with a limitation of $3x10^6$ ions. DDA was performed with the resolution of the Orbitrap set to 35000, with a fill time of 120 ms and a limitation of $2x10^5$ ions. Normalized collision energy of 32 was used. A loop count of 10 with count 1 was used and a minimum AGC trigger of $2e^2$ was set. Dynamic exclusion time of 30 seconds was applied. The peptide match algorithm was set to 'preferred' and charge exclusion 'unassigned', charge states 1, 5 - 8 were

Accepted manuscript

ACCEPTED MANUSCRIPT

excluded. Isolation window was set to 1.0 *m/z* and 100 *m/z* set as the fixed first mass. MS/MS data was acquired in profile mode.

Intracellular metabolome analysis

For intracellular metabolomics analysis, cells were harvested using a fast filtration protocol properly adapted from (S. Kim et al., 2013). Briefly, 5 mL of culture were sampled at mid-exponential growth phase and were vacuum-filtered through nylon membrane filters (0.45 µm, Whatman^{TM}), followed by three rapid washing steps with 5 mL of PBS to ensure no contamination from extracellular metabolites. The polar metabolites were extracted by adding the cell-containing filter in 5 mL of cold (-20 ^{o}C) HPLC-grade methanol (Biosolve Chimie, France)/MilliQ water (1:1, v/v) and incubating for 1 h at -20 ^{o}C. The mixture of metabolites and cell debris was centrifuged at 10000 rpm and 0 ^{o}C for 10 min, and the supernatants were collected and dried with speed-vac. The dried metabolites were derivatized to their (MeOx)TMS-derivatives through reaction with 100 µL of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, UK) solution in pyridine (Sigma-Aldrich) for 90 min at 40 ^{o}C, followed by reaction with 200 µL N-methyl-trimethylsilyl-trifluoroacetamide (MSTFA) (Alfa Aesar, UK) for 10 hours at room temperature, as justified in (Kanani and Klapa, 2007). The metabolic profile of each sample was measured thrice using a Shimadzu TQ8040 GC-(triple quadrupole) MS system (Shimadzu Corp.). The gas chromatograph was equipped with a 30 m x 0.25 mm x 0.25 µm DB-50MS capillary column (Phenomenex, USA). The detector was operated both in scan mode recording in the range of 50-600 *m/z*, as well as in MRM mode for the mentioned metabolites. The metabolite quantification was carried out by calculating the peak areas of the identified marker ions of each metabolite (**Supplementary Table 9**). For glucose, the smaller of the two derivative peaks was used for quantification.

Genome scale metabolic modeling

Metabolic reactions that likely become re-regulated during adaptive evolution of *S. cerevisiae* for glycerol utilization were identified using model simulations. Specifically, we used a mixed-integer linear programming routine to identify a minimum number of reactions, in the genome-scale metabolic model of *S. cerevisiae* (iFF708 (Förster et al., 2003)), that have to be re-regulated to

Accepted manuscript

ACCEPTED MANUSCRIPT

achieve optimal glycerol utilization. As a reference metabolic state for re-regulation we used the distribution of fluxes during respiratory growth on glucose. A set of minimum number of reactions were then identified whose (absolute) flux have to change for optimal glycerol utilization by >25 % beyond the (absolute) flux range extremes (Mahadevan and Schilling, 2003) (± 0.001 mmol/(g CDW h) when 6 C-moles of carbon source were converted to biomass) under the reference metabolic state (**Supplementary Table 10**). Equal C-molar conversion of carbon source to biomass was considered in the reference metabolic state and glycerol utilization. All the simulations were performed with Matlab R2015a v. 8.5.0 using IBM ILOG CPLEX v. 12.6.1 functions `'cplexlp'` and `'cplexmilp'`.

Spot assays

To assess osmosensitivity, yeast cells were inoculated in 3 mL of YPD medium and grown overnight at 30 °C with constant shaking 250 rpm. Next day, 100 µL aliquots of each culture were resuspended in sterile MiliQ water to OD_{600} = 2, and 3 µL of 10-fold dilution series were plated on YPD plates with and without KCl. Plates were incubated at 30 °C for 3 days and cell growth was monitored once a day.

Accepted manuscript

Acknowledgements

Authors would like to acknowledge Uffe H. Mortensen and Tina Johansen for facilitating the 1 L fermentation experiments, Anna Koza for technical assistance with MiSeq™ re-sequencing, EMBL Core Facilities (EMBL, Heidelberg) for RNA sequencing and protein mass-spectrometry. Authors also are thankful to Sheila I. Jensen for providing the genomic DNA of *Streptococcus pneumoniae* SV1. This work was sponsored by the Novo Nordisk Foundation, the German Ministry of Education and Research (BMBF, Grant number 031A343A), and The Danish Agency for Science, Technology and Innovation (DASTI) as part of the ERA-IB project "*DeYeastLibrary*". Funding for A.M.F. and E.T.M. was provided by The Novo Nordisk Foundation Grant Numbers NNF14OC0011269 and NNF10CC1016517.

ACCEPTED MANUSCRIPT

Conflict of interest

T.S, J.F, K.R.P and K.Z are co-inventors of a pending patent application.

References

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106. https://doi.org/10.1186/gb-2010-11-10-r106

Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq – a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed 1.29.18).

Barrick, J.E., Lenski, R.E., 2013. Genome dynamics during experimental evolution. Nat. Rev. Genet. 14, 827–839. https://doi.org/10.1038/nrg3564

Brusky, J., Zhu, Y., Xiao, W., 2000. *UBC13*, a DNA-damage-inducible gene, is a member of the error-free postreplication repair pathway in *Saccharomyces cerevisiae*. Curr. Genet. 37, 168–174. https://doi.org/10.1007/s002940050515

da Silva, G.P., Mack, M., Contiero, J., 2009. Glycerol: A promising and abundant carbon source for industrial microbiology. Biotechnol. Adv. 27, 30–39. https://doi.org/10.1016/j.biotechadv.2008.07.006

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V, Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498. https://doi.org/10.1038/ng.806

Finley, D., Ulrich, H.D., Sommer, T., Kaiser, P., 2012. The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. Genetics 192, 319–360. https://doi.org/10.1534/genetics.112.140467

Förster, J., Famili, I., Fu, P., Palsson, B.Ø., Nielsen, J., 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res. 13, 244–253. https://doi.org/10.1101/gr.234503

Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T., Halkier, B.A., 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 35, e55. https://doi.org/10.1093/nar/gkm106

Gietz, R.D., Schiestl, R.H., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 31–34. https://doi.org/10.1038/nprot.2007.13

Herring, C.D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M.K., Joyce, A.R., Albert, T.J., Blattner,

Accepted manuscript

ACCEPTED MANUSCRIPT

F.R., van den Boom, D., Cantor, C.R., Palsson, B.Ø., 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. 38, 1406–1412. https://doi.org/10.1038/ng1906

Ho, P.-W., Swinnen, S., Duitama, J., Nevoigt, E., 2017. The sole introduction of two single-point mutations establishes glycerol utilization in *Saccharomyces cerevisiae* CEN.PK derivatives. Biotechnol. Biofuels 10, 10–24. https://doi.org/10.1186/s13068-016-0696-6

Hofmann, R.M., Pickart, C.M., 1999. Noncanonical *MMS2*-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. Cell 96, 645–653. https://doi.org/10.1016/S0092-8674(00)80575-9

Illuxley, C., Green, E.D., Dunbam, I., 1990. Rapid assessment of *S. cerevisiae* mating type by PCR. Trends Genet. 6, 236–236. https://doi.org/10.1016/0168-9525(90)90190-H

Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., Keasling, J.D., 2015. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. Metab. Eng. 28, 213–222. https://doi.org/10.1016/j.ymben.2015.01.008

Jouhten, P., Ponomarova, O., Gonzalez, R., Patil, K.R., 2016. *Saccharomyces cerevisiae* metabolism in ecological context. FEMS Yeast Res. 16, fow080. https://doi.org/10.1093/femsyr/fow080

Kanani, H., Klapa, M., 2007. Data correction strategy for metabolomics analysis using gas chromatography–mass spectrometry. Metab. Eng. 9, 39–51. https://doi.org/10.1016/j.ymben.2006.08.001

Kao, K.C., Sherlock, G., 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. Nat. Genet. 40, 1499–1504. https://doi.org/10.1038/ng.280

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36. https://doi.org/10.1186/gb-2013-14-4-r36

Kim, S., Lee, D.Y., Wohlgemuth, G., Park, H.S., Fiehn, O., Kim, K.H., 2013. Evaluation and optimization of metabolome sample preparation methods for *Saccharomyces cerevisiae*. Anal. Chem. 85, 2169–2176. https://doi.org/10.1021/ac302881e

Klein, M., Carrillo, M., Xiberras, J., Islam, Z., Swinnen, S., Nevoigt, E., 2016a. Towards the exploitation of glycerol's high reducing power in *Saccharomyces cerevisiae*-based bioprocesses. Metab. Eng. 38, 464–472. https://doi.org/10.1016/j.ymben.2016.10.008

Klein, M., Islam, Z., Knudsen, P.B., Carrillo, M., Swinnen, S., Workman, M., Nevoigt, E., 2016b. The expression of glycerol facilitators from various yeast species improves growth on glycerol of *Saccharomyces cerevisiae*. Metab. Eng. Commun. 3, 252–257. https://doi.org/10.1016/j.meteno.2016.09.001

Kvitek, D.J., Sherlock, G., 2013. Whole genome, whole population sequencing reveals that loss of signaling

Accepted manuscript

ACCEPTED MANUSCRIPT

networks is the major adaptive strategy in a constant environment. PLoS Genet. 9, e1003972. https://doi.org/10.1371/journal.pgen.1003972

Lang, G.I., Rice, D.P., Hickman, M.J., Sodergren, E., Weinstock, G.M., Botstein, D., Desai, M.M., 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500, 571–574. https://doi.org/10.1038/nature12344

Liu, Z., Butow, R.A., 1999. A transcriptional switch in the expression of yeast tricarboxylic acid cycle genes in response to a reduction or loss of respiratory function. Mol. Cell. Biol. 19, 6720–6728. https://doi.org/10.1128/MCB.19.10.6720

Lo, C.-C., Chain, P.S.G., 2014. Rapid evaluation and quality control of next generation sequencing data with FaQCs. BMC Bioinformatics 15, 366–373. https://doi.org/10.1186/s12859-014-0366-2

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550–570 . https://doi.org/10.1186/s13059-014-0550-8

Mahadevan, R., Schilling, C.H., 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. 5, 264–276. https://doi.org/10.1016/j.ymben.2003.09.002

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12. https://doi.org/10.14806/ej.17.1.200

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

Merico, A., Ragni, E., Galafassi, S., Popolo, L., Compagno, C., 2011. Generation of an evolved *Saccharomyces cerevisiae* strain with a high freeze tolerance and an improved ability to grow on glycerol. J. Ind. Microbiol. Biotechnol. 38, 1037–1044. https://doi.org/10.1007/s10295-010-0878-3

Mikkelsen, M.D., Buron, L.D., Salomonsen, B., Olsen, C.E., Hansen, B.G., Mortensen, U.H., Halkier, B.A., 2012. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab. Eng. 14, 104–111. https://doi.org/10.1016/j.ymben.2012.01.006

Naito, Y., Hino, K., Bono, H., Ui-Tei, K., 2015. CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics 31, 1120–1123. https://doi.org/10.1093/bioinformatics/btu743

Nijkamp, J.F., van den Broek, M., Datema, E., de Kok, S., Bosman, L., Luttik, M.A., Daran-Lapujade, P., Vongsangnak, W., Nielsen, J., Heijne, W.H.M., Klaassen, P., Paddon, C.J., Platt, D., Kötter, P., van Ham, R.C., Reinders, M.J.T., Pronk, J.T., de Ridder, D., Daran, J.-M., 2012. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. Microb. Cell Fact. 11, 36–52. https://doi.org/10.1186/1475-2859-11-36

Accepted manuscript

ACCEPTED MANUSCRIPT

Nørholm, M.H.H., 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol. 10, 21–27. https://doi.org/10.1186/1472-6750-10-21

Ochoa-Estopier, A., Lesage, J., Gorret, N., Guillouet, S.E., 2011. Kinetic analysis of a *Saccharomyces cerevisiae* strain adapted for improved growth on glycerol: Implications for the development of yeast bioprocesses on glycerol. Bioresour. Technol. 102, 1521–1527. https://doi.org/10.1016/j.biortech.2010.08.003

Olsson, L., Nielsen, J., 1997. On-line and in situ monitoring of biomass in submerged cultivations. Trends Biotechnol. 15, 517–522. https://doi.org/10.1016/S0167-7799(97)01136-0

Partow, S., Siewers, V., Bjørn, S., Nielsen, J., Maury, J., 2010. Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. Yeast 27, 955–964. https://doi.org/10.1002/yea.1806

Rausch, T., Zichner, T., Schlattl, A., Stutz, A.M., Benes, V., Korbel, J.O., 2012. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 28, i333–i339. https://doi.org/10.1093/bioinformatics/bts378

Repetto, B., Tzagoloff, A., 1989. Structure and regulation of *KGD1*, the structural gene for yeast alpha-ketoglutarate dehydrogenase. Mol. Cell. Biol. 9, 2695–2705.

Sherman, F., 2002. Getting started with yeast. Methods Enzymol. 350, 3–41. https://doi.org/10.1016/S0076-6879(02)50954-X

Stovicek, V., Borja, G.M., Forster, J., Borodina, I., 2015a. EasyClone 2.0: expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. J. Ind. Microbiol. Biotechnol. 42, 1519–1531. https://doi.org/10.1007/s10295-015-1684-8

Stovicek, V., Borodina, I., Forster, J., 2015b. CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. Metabolic Engineering Communications 2, 13–22 . https://doi.org/10.1016/j.meteno.2015.03.001

Strimmer, K., 2008. A unified approach to false discovery rate estimation. BMC Bioinformatics 9, 303–316. https://doi.org/10.1186/1471-2105-9-303

Swinnen, S., Ho, P.-W., Klein, M., Nevoigt, E., 2016. Genetic determinants for enhanced glycerol growth of *Saccharomyces cerevisiae*. Metab. Eng. 36, 68–79. https://doi.org/10.1016/j.ymben.2016.03.003

Swinnen, S., Klein, M., Carrillo, M., McInnes, J., Nguyen, H.T.T., Nevoigt, E., 2013. Re-evaluation of glycerol utilization in *Saccharomyces cerevisiae*: characterization of an isolate that grows on glycerol without supporting supplements. Biotechnol. Biofuels 6, 157–168. https://doi.org/10.1186/1754-6834-6-157

VanDemark, A.P., Hofmann, R.M., Tsui, C., Pickart, C.M., Wolberger, C., 2001. Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer. Cell 105, 711–720. https://doi.org/10.1016/S0092-8674(01)00387-7

Verduyn, C., Postma, E., Scheffers, W.A., Van Dijken, J.P., 1992. Effect of benzoic acid on metabolic fluxes in

Accepted manuscript

ACCEPTED MANUSCRIPT

yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast 8, 501–517. https://doi.org/10.1002/yea.320080703

Wiser, M.J., Ribeck, N., Lenski, R.E., 2013. Long-term dynamics of adaptation in asexual populations. Science 342, 1364–1367. https://doi.org/10.1126/science.1243357

Zhou, N., Swamy, K.B.S., Leu, J.-Y., McDonald, M.J., Galafassi, S., Compagno, C., Piškur, J., 2017. Coevolution with bacteria drives the evolution of aerobic fermentation in *Lachancea kluyveri*. PLoS One 12, e0173318. https://doi.org/10.1371/journal.pone.0173318

Figure 1. Improving glycerol catabolism in *S. cerevisiae*. (a) Major pathway for glycerol catabolism: Stl1 – glycerol/H^+ symporter, Fps1 – aquaglyceroporin, Gut1 – glycerol kinase, Gut2 – FAD^+ – dependent glycerol 3-phosphate (G3P) dehydrogenase localized in mitochondrial membrane, DHAP – dihydroxyacetone phosphate. (b) Overview of the adaptive laboratory evolution (ALE) experiment and resulting analyses, see Methods for details. (c) Growth curves of NOX lineage 1 from mode-I ALE. (d) Growth curves of the mode-II lineages. Note, that only three out five wild-type lineages evolved to grow on MG medium. (e) Boxplot maximum growth rates of WT and NOX lineages of mode-II experiment. (f) Summary of the genome sequencing results of 17 lineages. *HOG1* – (High Osmolarity Glycerol response) gene coding for mitogen-activated protein kinase, *PBS2* – gene coding for MAP kinase kinase of the HOG signaling pathway. Each growth curve represents two biological replicates plotted using R language '*loess'* function.

Figure 2. Investigating causal mutations and their functional signatures. (a) Workflow for genomic strain analysis. (b) Re-engineering the glycerol phenotype achieved in the mode-I evolution experiment. (c) Spot assay experiment confirming the trade-off phenotypes conferring a sensitivity to osmotic stress. (d) Different growth phenotypes observed in tetrad analysis. Strain names correspond to the names in the panel e. (e) Mutation mapping during backcross experiment. (f) Growth phenotypes of re-engineered triple mutants. R-GUI (*GUT1*^{ALE2}, *UBC13*^{ALE2} and *INO80*^{ALE2}), R-GKI (*GUT1*^{ALE2}, *KGD1*^{ALE2} and *INO80*^{ALE2}) and R-GKU (*GUT1*^{ALE2}, *KGD1*^{ALE2} and *UBC13*^{ALE2}). (g) Confirming the trade-off phenotype of the *KGD1*^{ALE2} allele conferring the absence of respiratory metabolism when cultivated on glucose minimal medium. (h, i, j) Log2 fold changes (log2FC) of expression and translation of genes from the union of transcriptomic and proteomic data; (h) of R-GKU vs. ALE2 (i) of R-GU vs. R-GKU and (j) of R-GK vs. R-GKU. Blue dots depict only significantly changed transcripts ($p_{adj} < 0.1$), green – only significantly changed levels of proteins ($p_{adj} < 0.1$), and red – significantly altered levels in both protein and transcript ($p_{adj} < 0.1$).

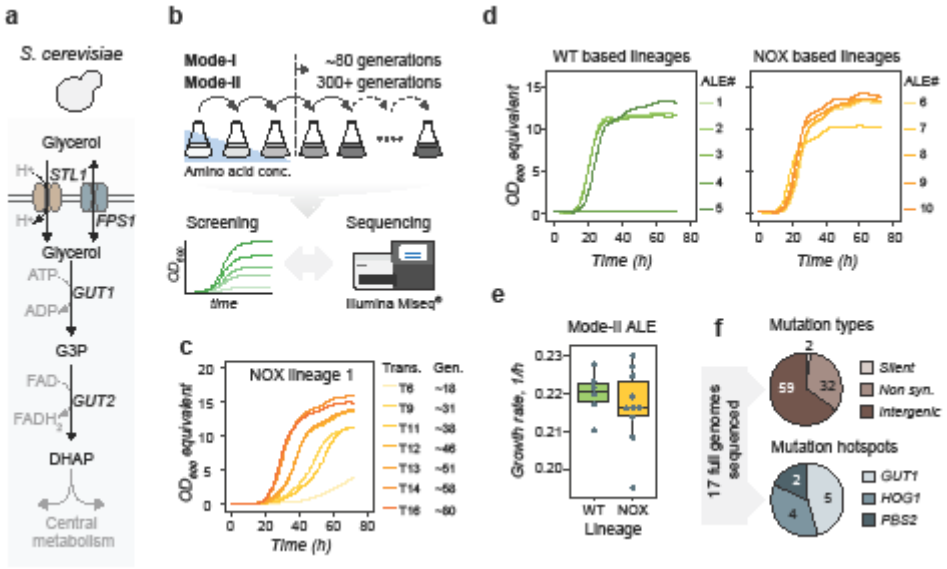
Figure 3. Overview of the causal mutations and their role in improving glycerol catabolism. (a) Metabolic map with the model-predicted and experimentally estimated changes. Colored dashed arrows represent model-predicted flux changes when carbon source is shifted from glucose to glycerol. Highlighted arrows depict estimated flux changes based on ratios of altered metabolite levels in the *KGD1*^{ALE2} (R-GKU) vs. *KGD1*^{WT} (R-GU) mutants grown in MG medium. GABA – γ-aminobutyric acid, SSA – succinate semialdehyde. (b) Growth phenotypes of two

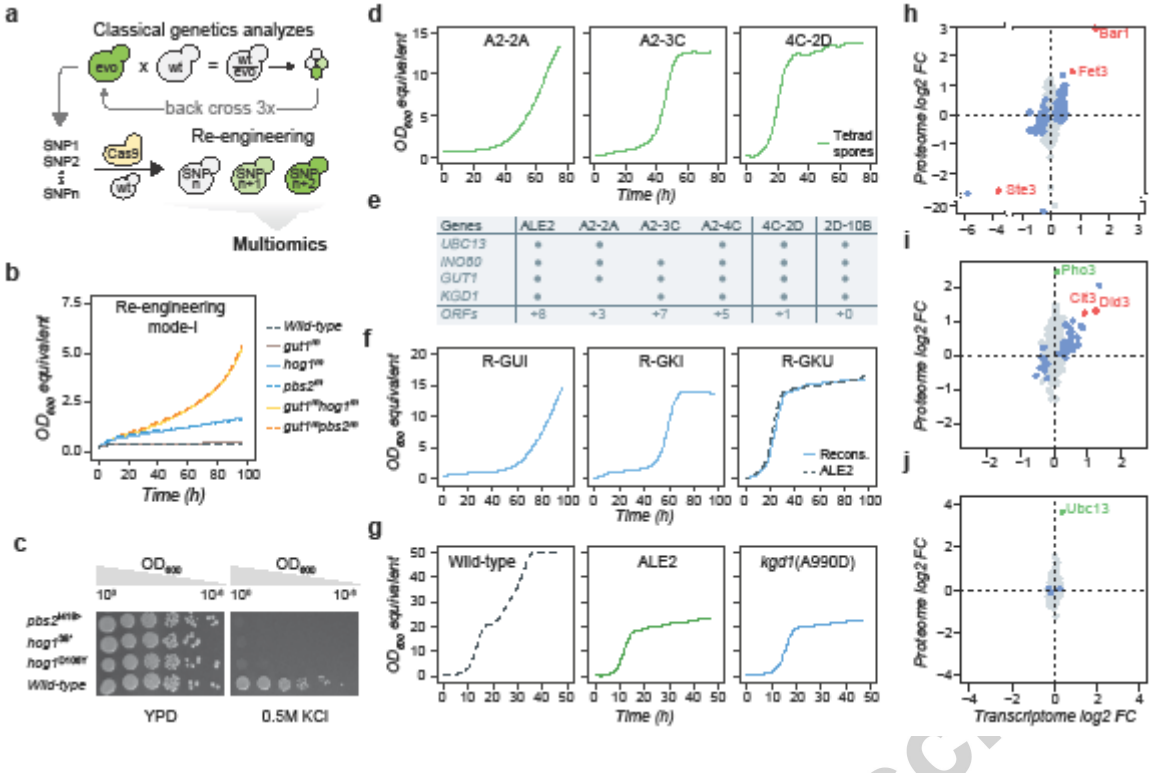
Accepted manuscript

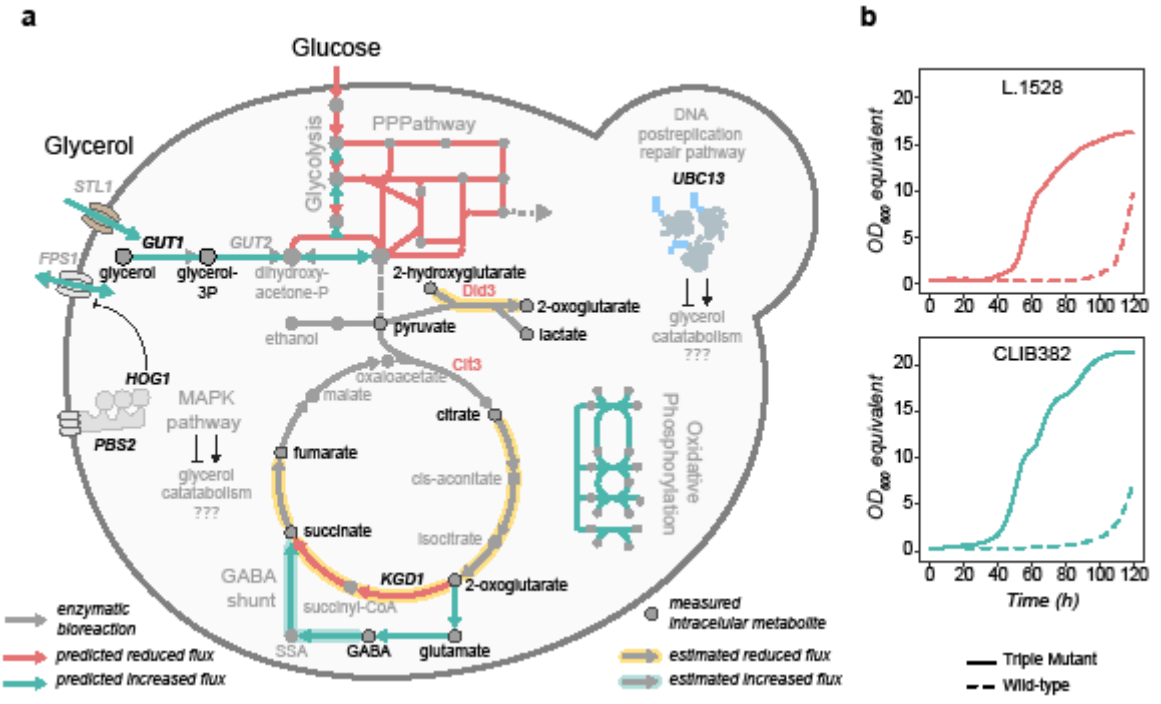
industrially relevant *S. cerevisiae* isolates with and without re-engineered triple mutations (*GUT1*^{ALE2}, *KGD1*^{ALE2} and *UBC13*^{ALE2}). Growth curves represent two biological replicates plotted using R language '*loess'* function.

Highlights

- Mutations underlying efficient glycerol utilization by yeast identified
- Only three single nucleotide variants dramatically improve glycerol utilization
- Multi-omics analysis and metabolic modeling affirms causality of mutations
- Industrial yeasts with improved glycerol uptake constructed
- Operational decoupling of TCA cycle from oxidative phosphorylation observed




a
Classical genetics analyzes
evo
x
wt
=
back cross 3x
SNP1
SNP2
SNPn
Cas9
Re-engineering
Multiomics
b
Re-engineering mode-I
OD600 equivalent
Time (h)
Wild-type
gut1
hog1
pbs2
gut1hog1
gut1pbs2
c
OD600
YPD
0.5M KCl
Wild-type
d
A2-2A
A2-3C
4C-2D
Tetrad spores
Time (h)
e
Genes
ALE2
A2-2A
A2-3C
A2-4C
4C-2D
2D-10B
UBC13
INO80
GUT1
KGD1
ORFs
+8
+3
+7
+5
+1
+0
f
R-GUI
R-GKI
R-GKU
Recons.
ALE2
g
Wild-type
ALE2
kgd1(A990D)
h
Bar1
Fet3
Ste3
Proteome log2 FC
i
Pho3
Clt3
Dld3
j
Ubc1
Transcriptome log2 FC


a
Glucose
Glycerol
STL1
FPS1
GUT1
GUT2
glycerol
glycerol-3P
dihydroxy-acetone-P
Glycolysis
PPPathway
2-hydroxyglutarate
Dld3
2-oxoglutarate
pyruvate
lactate
ethanol
HOG1
PBS2
MAPK pathway
glycerol catabolism ???
Clt3
oxaloacetate
malate
citrate
fumarate
cis-aconitate
succinate
isocitrate
KGD1
GABA shunt
succinyl-CoA
SSA
GABA
glutamate
DNA postreplication repair pathway
UBC13
glycerol catabolism ???
Oxidative Phosphorylation
enzymatic bioreaction
predicted reduced flux
predicted increased flux
measured intracellular metabolite
estimated reduced flux
estimated increased flux
b
L.1528
CLIB382
OD600 equivalent
Time (h)
Triple Mutant
Wild-type