December, 2019

Beneficial mutations for carotenoid production identified from laboratory-evolved Saccharomyces cerevisiae

Avinash Godara, *Texas A&M University*
Maria Alejandra Gomez Rodriguez, *Texas A&M University*
Joshua D. Weatherston, *Texas A&M University*
George L. Peabody, *Oak Ridge National Laboratory*
Hung-Jen Wu, *Texas A&M University*, et al.

Beneficial mutations for carotenoids production identified from laboratory-evolved *Saccharomyces cerevisiae*

Avinash Godara[1], Maria Alejandra Gomez Rodriguez[1], Joshua D. Weatherston[1], George L. Peabody V[2], Hung-Jen Wu[1], Katy C. Kao[1,3,#]

[1]Department of Chemical Engineering, Texas A&M University, College Station, Texas, United States

[2]Currently at Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States

[3]Department of Chemical and Materials Engineering, San Jose State University, San Jose, CA, United States

[#]Corresponding author. Email: kao.katy@gmail.com. One Washington Square, San Jose, CA 95129, United States.

Abstract

Adaptive laboratory evolution (ALE) is a powerful tool used to increase strain fitness in the presence of environmental stressors. If production and strain fitness can be coupled, ALE can be used to increase product formation. In earlier work, carotenoids hyperproducing mutants were obtained using an ALE strategy. Here, *de novo* mutations were identified in hyperproducers, and reconstructed mutants were explored to determine the exact impact of each mutation on production and tolerance. A single mutation in *YMRCTy1-3* conferred increased carotenoid production, and when combined with other beneficial mutations led to further increased β-carotene production. Findings also suggest that the ALE strategy selected for mutations that confer increased carotenoids production as primary phenotype. Raman spectroscopy analysis and total lipid quantification revealed positive correlation between increased lipid content and increased β-carotene production. Finally, we demonstrated that the best combinations of mutations identified for β-carotene production were also beneficial for production of lycopene.

Keywords

Adaptive laboratory evolution; Raman spectroscopy; Carotenoids; Yeast; Oxidative stress;

1. Introduction

Carotenoids belong to an important class of isoprenoids naturally synthesized by plants and some microbes [10,14,27,46,39,2,43]. Due to their pigmented and antioxidant properties, carotenoids have applications in food and nutraceutical industries. All isoprenoids are secondary metabolites that are derived from C_5 precursor isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP) [4,6]. Due to the structural complexity of these compounds, there has been long-term interest in using microbial fermentation as an alternative to chemical synthesis or plant extraction.

In addition to native microbial producers such as *Rhodotrula glutinis* [11,36,31], non-native producers such as *Escherichia coli* and *Saccharomyces cerevisiae* have been engineered to produce these carotenoids via metabolic engineering [15,60,37,40,56,59]. Prior studies have shown that genes outside biosynthetic pathway can influence product formation [1,17]. However, due to the intricate nature of biological systems, genetic determinants for product formation outside of biosynthetic pathways are generally not known. Thus, complementary genome-scale techniques, such as screening deletion or overexpression libraries [45,5,54] and adaptive laboratory evolution (ALE)-based strategies [44,48] can help to identify additional gene and/or pathway targets related with increasing product formation. In our prior work, we leveraged the antioxidant properties of carotenoids to develop an ALE-based strategy to evolve strains for increased productivity. Starting with a parental *S. cerevisiae* strain engineered to produce β-carotene (and lacking the cytosolic catalase gene), ALE was conducted using hydrogen peroxide as a selective pressure. Several evolved mutants with enhanced β-carotene production were isolated after several rounds of ALE selection [48].

In this work, we sequenced the genomes of hyperproducing mutants from one evolved population and identified mutations responsible for enhanced β-carotene yield. Results showed that mutations selected for in the adaptive laboratory evolution experiment mainly contributed to increased oxidative stress tolerance that is dependent on increased production of β-carotene. Phenotypic characterization using Raman spectroscopy revealed that the increase in β-carotene production level in mutations is related to increase in total lipid content of the cell, which was further validated using total lipid quantification. Finally, we demonstrated that mutations associated with increased production of β-carotene also confer increased production of another carotenoid lycopene.

2. Materials and methods

2.1 Strains, plasmids and growth conditions

All yeast strains used in this work are derivatives of S288c and listed in **Supplementary Table S1**. *S. cerevisiae* strain YLH2 [48] and FY2 (MATα, ura3-52, isogenic to S288C) [23] are used as the base strains in this study. Unless otherwise specified, yeast strains are cultured in Yeast Peptone Dextrose (YPD) media at 30 ºC and *E. coli* strains used for subcloning were cultured at 37 ºC in Luria Broth (LB) supplemented with appropriate antibiotics. Cytosolic catalase T (*CTT1*) gene was deleted in FY2 strain as described in our earlier work [48]. For carotenoid quantification, yeast strains were cultured in YPD at 30 °C and 200 rpm for 72 h. For Raman Spectroscopy the strains were grown for 36 hours in YPD at 30 °C and 200 rpm.

2.2 Whole genome resequencing

Genomic DNA of isolated mutants and parental strain YLH2 were extracted using YeaStar DNA extraction kit (Zymo Research). Library preparations and NGS sequencing were performed by the Texas A&M Genomics Center for sequencing on the Illumina HiSeq 2500 platform using 100-bp single-end reads. An average coverage of >75-fold was obtained for each clone. The raw sequencing data were deposited in SRA database (https://www.ncbi.nlm.nih.gov/sra) with accession number SRP147744. The sequencing data was aligned to *S. cerevisiae* S288c reference genome with breseq v0.29 [18]. *De novo* mutations in isolated mutants were identified by comparing against the YLH2 parental strain, and verified via Sanger sequencing (mutations listed in **Table 1**).

2.3 Reconstructing mutations into YLH2 and FY2 strain

CRISPR-Cas9 using the one plasmid-system with pCRCT developed by [8] was used for site-directed mutagenesis to reconstruct identified mutations into YLH2. Briefly, 120 bp sequence including donor sequence and guide sequence was chemically synthesized by Integrated DNA Technologies, USA (**Supplementary Table S2**), then introduced into the pCRCT plasmid using Golden Gate reaction using BsaI restriction sites [20], and transformed into *E. coli* cells, and plated on LB+Xgal for blue/white screening and incubated overnight at 37 °C. The bacterial colony with correct plasmid construction was verified by restriction digestion. The constructed CRISPR-Cas9 plasmid was then transformed into FY2 and YLH2 strains and selected on SC -uracil plates. Colonies were picked and target mutations were first verified using PCR amplification refractory mutation system (ARMS) **(**see **Supplementary Table S3** for primer sequences) as described by [33] followed by confirmation via Sanger sequencing. Cas9

plasmid was cured by serially passaging the strain in YPD three times and verified by PCR amplification using primers curing_f and curing_r (**Table S3**).

2.4 Quantification of carotenoid production

Quantification of β-carotene production was performed as described by [48]. Detailed method and validation with HPLC are provided in Supplementary Information. All samples were blank corrected using a dodecane blank. The same procedure was used to quantify lycopene using OD_{476}. β-carotene and lycopene used for standard curves were purchased from Sigma. At least three biological replicates per strain were used for analysis.

2.5 Growth kinetics measurements

Microplate reader (TECAN Infinite ® M Nano) was used to measure growth curve for reconstructed strains. Cells were grown for 24 h in YPD, then normalized to OD_{600} ~0.05 in 200 µL final volume in YPD in 96-well plates. Cells were cultured in the microplate reader for 72 h with orbital shaking at constant intervals. Shaking was performed using 3mm amplitude using kinetic cycles of 2 min incubation time, then orbital shaking for 3 min at 198 rpm followed by OD measurements. To obtain specific growth rate, the OD_{600} observed was plotted on natural logarithmic scale versus time (h). The specific growth rate µ (slope of the linear region for each sample), duration of growth lag (length of time before the start of the linear region on the semi-log plot), and the maximum OD_{600} achieved were calculated for each sample. Two runs with three biological replicates each for a total of 6 biological replicates per strain were used for analysis.

123 2.6 Oxidative stress tolerance using hydrogen peroxide

124 Reconstructed mutants were treated with 1 M hydrogen peroxide for 1 h in a shaking
125 incubator at 30ºC. After the shock treatment, 10-fold serial dilutions were performed in
126 phosphate-buffered saline (PBS) and spotted on YPD plates. Three biological replicates per
127 strain were used for analysis. Detailed method is provided in Supplementary Information.

128

129 2.7 Raman spectroscopy analysis

130 To prepare samples for Raman spectroscopy, strains were first grown in YPD from single
131 colonies for 36 h at 30 ºC and the yields of β-carotene or lycopene were calculated as described
132 above. Three biological replicates per strain were used for analysis. One mL of the grown
133 culture was then centrifuged and the supernatant was vacuum aspirated. The cell pellet was
134 washed three times with 1 mL of PBS. Two µL of each sample was pipetted and gently spread
135 onto an aluminum-covered microscope slide, yielding a thin film of intact cells. The thin film
136 was allowed to dry fully before Raman spectral acquisition. 16 spectra were collected for each
137 biological replicate (48 spectra per strain) using a DXR Raman microscope (Thermo Scientific)
138 equipped with a high-brightness 780 nm diode laser, a 50x objective lens (N.A. 0.5, spot size 1.6
139 µm), a 25 µm confocal slit, a diffraction grating (400 line/mm, spectral resolution ~2.4-4.4 cm^{-1}),
140 and a CCD detector. Laser power was set at 20 mW and spectra were collected over the spectral
141 range 150 – 3200 cm^{-1}, with 10 accumulations of 12 s each (120 s total laser exposure per
142 spectrum). Each spectrum was processed post-collection for artifact removal, baseline correction
143 and noise reduction. After processing, the spectral intensity was tabulated (**Table S4**) for several
144 characteristic peaks corresponding to the carotenoids, lipids, and proteins within the cells.
145 Absolute peak intensities varied widely within samples due to inhomogeneity of the cell films.

Therefore, ratios of the characteristic peaks were used to compare the relative abundance of chemical compounds across strains and mutations. The ratio of total fatty acid peak and carotenoid peak was multiplied by the total carotenoid content, which is obtained via quantification, to get a relative lipid indicator.

$$Lipid\ indicator = \frac{Total\ fatty\ acid}{Carotenoid} * absolute\ carotenoid\ content(\frac{mg}{gdcw})$$

2.8 Quantification of lipid using classical Bligh and Dyer method

Yeast cells were streaked from frozen stock and an individual colony was grown in 5 ml YPD overnight at 30 °C and 250 rpm. The resulting culture was normalized to initial OD_{600} ~0.05 in 100 ml YPD for further growth for an additional 36 h. Cells were harvested via centrifugation and the cell pellet was dried for 48 h at room temperature. One gram of cell pellet was mixed with 1:2:0.8 chloroform:methanol:water (v/v/v) and homogenized for 10 min using Branson Digital Sonifier® 250 cell disruptor with 40% amplitude and pulse on and off every five seconds. Water containing 0.8% KCl and chloroform was added to bring the solvent ratio to 2:2:1.8 with total volume added of 10ml:10ml:9ml chloroform:methanol:water (v/v/v) and homogenized for another 10 min. The two liquid phases were separated for two hours at room temperature, and the total volume of organic layer (bottom phase) was measured. One ml of organic layer was spread on pre-weighed aluminum pans, and the sample was left to evaporate in chemical hood overnight. Weight was recorded after overnight evaporation using three technical replicates for each of three biological replicates per strain [12,13].

2.9 Bioreactor study

168 The seed culture for fermentation was started from a single colony inoculated in 5 mL of
169 YPD cultured overnight at 30ºC, then subcultured into 50 mL of YPD media. Seed culture was
170 grown for 24 h at 30 ºC and used to inoculate 1 L of fresh YPD media in a 3 L bioreactor
171 (Applikon®). The bioreactor was maintained at 30 ºC with continuous air feed at 2.6 L/min and
172 800 rpm. Online dissolved oxygen (DO) sensor was used to estimate DO levels. The pH of the
173 bioreactor was maintained at pH ~4.0 using 2 M HCl and 2 M NaOH. Batch bioreactor
174 experiments were run for ~64 hours.
175
176
177
178 **3. Results and Discussion**
179
180 3.1 Genome resequencing revealed mutations in genes not known to be related with isoprenoids
181 biosynthesis in hyperproducers
182 Previously, we evolved an engineered *S. cerevisiae* strain (YLH2) for improved β-
183 carotene production using hydrogen peroxide as the selective pressure [48]. Adaptive mutants
184 with enhanced carotenoids yields were isolated from two parallel populations at various time
185 points throughout the evolution [48]. To determine the underlying adaptive mutations
186 responsible for improved production, isolates SM12, SM13, and SM14 from one population and
187 the ancestral strain YLH2 were subjected to whole genome resequencing in this work. Four to
188 six *de novo* mutations were identified from each isolated mutant (see **Table 1**). Several identical
189 mutations (*e.g. HIS7 389*, *SRO9/GFD2* int) were found in more than one isolated mutant, which
190 was not surprising since these mutants originate from the same evolving population.

Among the *de novo* mutations that are identified in more than one isolate, a 24 bp mitochondrial deletion was found 152bp downstream of the mitochondrial proline tRNA gene *tP(UGG)Q* in mutants SM12 and SM13. Since this is a mutation in the mitochondria and difficult to analyze, this mutation was omitted from further characterization in this study. A missense mutation in *HIS7 389* was identified in all sequenced mutants, indicating that this mutation arose early on during the evolution experiment. Using kinetic quantitative-PCR (see Supplementary Info), the relative frequency of this mutation was estimated in population samples (revived from frozen stocks) from the evolution experiment. The results showed that the *HIS7 389* mutation was present in ~99% of the population very early on during the evolution, after the first cycle of H_2O_2 challenge [48], strongly suggesting this to be a jackpot mutation [35], which is a mutation that arose very early on in the experiment, likely in the initial inoculum used to seed the ALE experiment. *HIS7* encodes the imidazoleglycerol phosphate synthase gene, which is involved in histidine biosynthesis, and does not have an obvious role in β-carotene production. Indeed, subsequent verification experiments revealed that the mutation in *HIS7 389* confers no benefit to general fitness nor oxidative stress tolerance (see below). A common intergenic mutation between *SRO9/GFD2* was identified in 2/3 sequenced mutants. Sro9 is a La-motif-containing protein and is known to be involved in translation [52], and Gfd2 is currently a protein of unknown function. Mutations in the intergenic region of these two genes may potentially impact the expression of one or both genes; and in the case of Sro9, potentially lead to global changes in gene expression if impacted.

A total of 8 unique mutations were identified among the sequenced strains: 5 non-conservative missense mutations (*TYE7 86* in SM12, *FLO1 925* in SM13, and *EPL1 1754*, *YMRCTy1-3 1078*, and *ALG6 1411* in SM14), 1 intergenic mutation (between *DAK2/AQY3* in

214 SM13), 1 frame shift mutation (in *MDS3* in SM14) and 1 synonymous point mutation (in *SCY1*
215 in SM13). The intergenic mutation in *DAK2/AQY3* may play a role in stress adaptation as Dak2
216 is involved in stress adaptation and is required for detoxification of dihydroxyacetone [38], and
217 Aqy3 is reported to play a role in passive glycerol diffusion [41]. The mutation in *TYE7* can
218 potentially lead to changes in gene expression as it is a regulator and acts as a transcriptional
219 activator in Ty1-mediated gene expression [34]. Epl1 is involved in epigenetic regulation, and is
220 known to control the expression of genes involved in DNA repair, ubiquitylation and other stress
221 responses [19,24,53], thus a mutation in *EPL1* can potentially impact general adaptation. Flo1
222 promotes adhesion of yeast cells and formation of cell clumps, which can protect cells from
223 environmental stressors [51]; however, based on microscopic inspection of SM13 (which
224 contains the *FLO1* mutation), no obvious clumping of the cells was observed (data not shown).
225 The mutation in *SCY1* is a synonymous point mutation, while we do not expect the function of
226 the gene to be altered, synonymous mutations have been shown to impact gene expression
227 [7,32]. Scy1 plays a role in sterol transport [55], thus a synonymous mutation in this gene can
228 potentially alter lipid content in the cell. Alg6 encodes an endoplasmic glucosyltransferase [47]
229 and is involved in cell wall assembly [3]. The mutation in *MDS3* identified in SM14 is a
230 frameshift mutation near the N-terminus of the protein, strongly suggesting this to be an
231 inactivating mutation. Mds3 regulates sporulation in *S. cerevisiae [9]* and deletion in *MDS3* has
232 been shown to negatively impact resistance to oxidants [26]; thus the reason this mutation was
233 selected for during oxidative stress challenge needs further investigation. The *YMRCTy1-3*
234 encodes components of the Ty1 retrotransposon; Ty1 insertion mutations are known to impact
235 fitness and are observed in laboratory evolution experiments [58,25]; thus, mutation in
236 *YMRCTy1-3* can potentially impact fitness.

237

238 3.2 Assessing the impact of identified mutations on general fitness and β-carotene yield in
239 reconstructed mutants in the carotenogenic ancestral strain YLH2

240 Mutations that may be selected for in our prior ALE experiment include those that lead to
241 increased H_2O_2 resistance (*e.g.* via increased production of the antioxidant β-carotene) or confer
242 benefits on general fitness. To assess impacts of each mutation, all mutations (except for the
243 mitochondrial mutation) identified in SM12, SM13 and SM14 were introduced into YLH2 strain
244 individually to generate strains YAG01 to YAG10 (**Table S1**). The growth kinetics results
245 showed no significant differences in the specific growth rates (μ) of most of the reconstructed
246 mutants compared with the parental strain YLH2 (See **Table 2**). However, all of the mutations
247 led to significant reductions in the duration of the growth lag from ~4 h in parental strain YLH2
248 to between ~0.4 h – 2 h in the reconstructed mutants, which likely contributed to their selection
249 in the ALE experiment.

250 The impacts of each mutation on carotenoids production was evaluated. Overall, the
251 results showed that the majority of mutations alone conferred a positive impact on carotenoids
252 production (see **Figure 1A**). The intergenic mutation between *SRO9/GFD2* (YAG02) led to a
253 significant increase in β-carotene content, with yield of 17.9±0.1 mg/g DCW in YAG02
254 compared with 9.8± 0.3 mg/g DCW in YLH2 and both evolved mutants that contain this
255 mutation (SM12 [15.1± 0.2 mg/g DCW] SM13 [14.2± 0.1 mg/g DCW]). In SM14, the single
256 mutation in *YMRCTy1-3* alone resulted in β-carotene yield of 20.5±0.2 mg/g DCW (YAG10),
257 slightly exceeding the 19±0.1 mg/g DCW observed in strain SM14 (p-value ~ 0.003). These
258 results suggest that one or more of the other mutations present in SM12, SM13 and SM14 were
259 detrimental to product yield. Indeed, the jackpot mutation in *HIS7 389* that is present in all three

sequenced mutants caused a significant decrease in β-carotene yield when introduced into YLH2 (3.7±0.6 mg/g DCW).

3.3 Assessing the effects of combinations of mutations on β-carotene production.

All mutations (except for the mitochondrial mutation) identified in each sequenced evolved mutant were reconstructed in the YLH2 parental background to determine whether the β-carotene yield in the isolated hyperproducers can be fully recapitulated. The results for the resulting strains YAG11 (SM12), YAG12 (SM13), and YAG13 (SM14) are shown in **Figure 1B**. Surprisingly, combining the selected mutations led to significant decrease in product yield (~8 mg/g DCW in all three reconstructed mutants) compared with SM12, SM13, and SM14. The results suggest that the mitochondrial mutation in SM12 and SM13 likely plays a role on β-carotene yield and/or that there are additional undetected beneficial mutations in the hyperproducers.

Since all three hyperproducers contain the detrimental mutation in *HIS7*, we reconstructed SM12, SM13, and SM14 strains without the *HIS7* mutation (resulting in strains YAG14, 15, and 16, respectively) to assess the extent of the negative impacts of this mutation in the multiple-mutations backgrounds. Results are shown in **Figure 1B**. YAG14-16 contain the same mutation combinations as that of YAG11-13, respectively, with the exception of the *HIS7 389* mutation. Interestingly, strains YAG14 and YAG15 produced more β-carotene than SM12 and SM13 (p-value < 0.01), respectively, and YAG16 produced nearly as much as SM14. It is likely that due to the asexual nature of the ALE experiment, this detrimental *HIS7* mutation for β-carotene production was not able to be removed from the hyperproducers, leading to a reduction in the overall β-carotene yield in these strains.

Since multiple mutated genes were identified to have beneficial effects on β-carotene yield, we generated mutants containing combinations of mutations identified in different evolved strains (resulting strains YAG17 to YAG30 in **Table S1**) to determine whether superior combinations can be generated from these single mutations. All the superior recombinants identified (YAG22, YAG23, YAG26-YAG29) contain the *AQY3/DAK2* intergenic mutation, with the exception of YAG30 which contains the *SRO9*/*GFD2* intergenic mutation and the *YMRCTy1-3* mutation. The best combination comprises of the *YMRCTy1-3 1078*, *ALG6 1411*, and the *AQY3/DAK2 int* mutations (YAG28), which exhibited a β-carotene yield of 24.9 ± 0.5 mg/g DCW (**Figure 1C**). This superior strain exceeded the yield of the best evolved mutant SM14 by ~30%. The performance of YAG28 was further quantified using a benchtop bioreactor, and a β-carotene yield of 37.8±1.9 mg/g DCW, which is ~82% higher compared with SM14 (20.7±3.1 mg/g DCW), was observed. YAG28 also showed higher growth in the fermenter, reaching an OD_{600} around 20 as compared to SM14 which only reaches OD_{600} around 17 after 64 hours (p-value = 0.015 using two-tailed student t-test).

3.4 Decoupling the impacts of each mutation on hydrogen peroxide stress tolerance and β-carotene production

We had previously hypothesized that oxidative stress tolerance can be used as a selectable phenotype for increased β-carotene production in our prior ALE experiment, which led to the use of oxidative stress as the selective pressure [48]. It is possible that the mutations identified in the isolated hyperproducers conferred oxidative stress tolerance independent of β-carotene production, and that the increased productivity was a secondary effect. To determine whether this is the case, all single mutations (except for the mitochondrial mutation) were

306 reconstructed in a wild-type S288c strain (FY2) and subjected to short-term H_2O_2 challenge.
307 Since β-carotene itself is an antioxidant and protects against oxidative stress damage, we
308 reconstructed these mutations in a strain that does not contain the carotenogenic cassette to
309 decouple oxidative stress tolerance conferred by the mutation alone from that conferred from
310 production of β-carotene.

311 As shown in **Table 3 and Supplementary Figures S1 and S2**, mutations in *MDS3*
312 (YAG39), *EPL1* (YAG37), *FLO1* (YAG34), and *ALG6* (YAG38) conferred increased tolerance
313 to H_2O_2 challenge in the non-carotenogenic background. On the other hand, mutations in *SCY1*
314 (YAG36), *YMRCTy1-3* (YAG40), and the intergenic mutations in *DAK2/AQY3* (YAG35) and
315 *SRO9/GFD2* (YAG32) resulted in decreased tolerance to hydrogen peroxide stress. This result
316 runs counter to prior study showing that null mutation in *MDS3* led to decreased tolerance to
317 H_2O_2 [26]. One possible explanation for this discrepancy is the level of hydrogen peroxide used.
318 In Higgins *et al*, null mutation in *MDS3* was shown to exhibit reduced growth in the presence of
319 4 mM of H_2O_2 in the culture media [26] whereas in our study, the *MDS3 ins* mutation was found
320 to exhibit increased viability after 1 h exposure to 1 M H_2O_2. Prior study found *MDS3* null
321 mutant to exhibit reduced cell death in the presence of antifungal compounds (tunicamycin and
322 dithiothreitol) that disrupt proteins/enzymes in the endoplasmic reticulum (ER) [29], suggesting
323 that Mds3 is a death-promoting factor induced by ER stress. It may be possible that Mds3 has a
324 death-promoting effect in the presence of high concentration of H_2O_2 similar to exposure to
325 tunicamycin or dithiothreitol, and the inactivating *MDS3 ins* mutation led to increased viability.
326 Interestingly, *ALG6* null mutant exhibits increased resistance to fluconazole, which disrupts a
327 protein localized in the ER [28]. It is possible that the *ALG6 1411* mutation led to a similar

328 tolerance mechanism as *MDS3 ins*; however, it is not known if the missense mutation is an
329 inactivating mutation.

330 The intergenic mutation between *SRO9/GFD2* and the missense mutation in *YMRCTy1-3*
331 that led to the highest increase in β-carotene yields in YLH2 background (YAG02 and YAG10,
332 respectively), both caused significant reductions in hydrogen peroxide tolerance in the non-
333 carotenogenic background (strains YAG32 and YAG40, respectively); suggesting that the two
334 most beneficial mutations for β-carotene production were not selected due to beneficial effects in
335 H_2O_2 tolerance that are independent of carotenoids production.

336 To test the validity of the original hypothesis that increased β-carotene production
337 enhances H_2O_2 tolerance, the β-carotene biosynthesis cassette was introduced into each of the
338 reconstructed mutants in FY2 to generate carotenogenic versions of each mutant. The resulting
339 strains are YAG46-YAG55 (**Table S1**) with wild-type control YAG45. The β-carotene yields
340 from the FY2 reconstructed mutants (**Table 3**) were, in general, lower compared with those in
341 the YLH2 background, but the trend as to which mutations conferred the highest increase in
342 production was consistent. The two mutations (*SRO9/GFD2 int* and *YMRCTy1-3 1078*) that
343 caused significant decreases in H_2O_2 tolerance in the non-carotenogenic background (YAG32
344 and YAG40) showed the highest level of β-carotene yield in their carotenogenic counterparts
345 (YAG47 and YAG55). YAG47 and YAG55 also exhibited significant increases in H_2O_2
346 tolerance, demonstrating that increased production of β-carotene is the cause of increased
347 oxidative stress tolerance in these strains. Two of the mutations (*SCY1 1836* and the
348 *DAK2/AQY3 int*) that exhibited decreased H_2O_2 tolerance compared with the wild-type control in
349 the non-carotenogenic background both exhibited no difference in H_2O_2 tolerance relative to
350 wild-type control in the carotenogenic versions along with modest increases in β-carotene yield;

suggesting that increased β-carotene production helped to overcome the negative impacts of *SCY1 1836* and *DAK2/AQY3 int* mutations on oxidative stress tolerance. For the mutations that showed enhanced H_2O_2 tolerance in the non-carotenogenic background, the level of H_2O_2 tolerance appears to be independent of β-carotene production.

We further generated a Δ*ctt1* strain to determine whether the presence of the cytosolic catalase gene significantly influences the relative H_2O_2 tolerance and β-carotene production in these mutants. Results showed very similar trends in both relative oxidative stress tolerance and β-carotene yields in the wild-type versus Δ*ctt1* strains (see **Table 3 and Figure S3**), suggesting that the conclusions drawn regarding the impacts of each mutation on oxidative stress tolerance and β-carotene production are valid in both the *CTT1* wild-type and Δ*ctt1* backgrounds.

Overall, the data strongly suggest that the ALE strategy selected for few mutations that primarily increased H_2O_2 tolerance independent of β-carotene production (*e.g. MDS3 ins*), several mutations that conferred increased H_2O_2 tolerance both independent and dependent of increased β-carotene production (*EPL1 1754*, *FLO1 925*, and *ALG6 1411*), and mutations that conferred increased oxidative stress tolerance solely due to increased β-carotene production (*SCY1 1836*, *DAK2/AQY3 int*, *TYE7 86*, *SRO9/GFD2 int*, and *YMRCTy1-3 1078*). The mutations that conferred the highest increases in β-carotene yield were not selected for as a result of impacts other than increased product formation, suggesting that the oxidative stress challenge strategy used in the ALE experiments primarily selected for mutants with higher product yield as intended.

3.5 Total lipid content in reconstructed strains

373 Previous global gene expression profiling of a subset of the isolated adaptive mutants
374 revealed upregulation of key genes involved in lipid biosynthesis [48], suggesting lipid
375 metabolism may be perturbed in the carotenoids hyperproducers. Indeed, when cultured under
376 low nitrogen conditions, which has been shown to increase lipid content of oleaginous yeasts
377 [50,21,30] and algae [49,22], strain YLH2 showed increased β-carotene yield (by ~298%) and
378 total fatty acid content [42]. To determine whether any of the mutations analyzed here impact
379 lipid content, the relative lipid contents of these mutants were analyzed using Raman
380 spectroscopy. Based on prior reports [16,57] and comparison between the spectra of
381 carotenogenic strain YLH2 and a non-carotenoid producing strain FY2, Raman peaks were
382 identified for total fatty acid at 1449 cm^{-1}, unsaturated fatty acid at 1663 cm^{-1}, and carotenoid at
383 1156 cm^{-1}. Raman analysis was conducted for YLH2, SM14, all reconstructed single mutants
384 (YAG01 to YAG10), and the best producers generated YAG22 and YAG28.
385 Results showed no significant differences in unsaturated fatty acid content in any of the
386 reconstructed mutants compared with parental strain YLH2 (**Figure 2**). Thus, to compare
387 between various mutants, the total fatty acid signal of each strain was normalized by its carotene
388 signal and weighted by its carotene yield (equation shown in Materials and Methods). This yields
389 a lipid indicator that is proportional to the quantity of total fatty acid per cell dry weight for each
390 strain. All Raman data was further normalized against the carotenogenic parental strain, YLH2,
391 in **Figure 2**. While the majority of strains tested showed relative lipid indicator values > 1,
392 YAG10 and YAG28 exhibited the highest lipid content with relative lipid indicators ~4.5-fold
393 and ~3.5-fold greater than YLH2, respectively. Both YAG10 and YAG28 showed higher lipid
394 indicators compared with SM14 (~2). These data suggest that increased β-carotene yield

observed in strains YAG10 and YAG28 were likely partially due to increased lipid content of the strains.

To validate the Raman spectral results, the total lipid content of strains YLH2, SM14, YAG10 and YAG28 was measured. Results are shown in **Figure 3**. The lowest amount of total liquid was observed in YLH2 with 48.7 ± 1.8 mg/g DCW. Results for SM14 showed 88.4 ± 2.8 mg/g DCW total lipid. Strains YAG10 and YAG28, which showed the highest lipid indicators, contained 120.5 ± 4.6 mg/g DCW and 103.0 ± 2.7 mg/g DCW of total lipid, respectively. The results from the gravimetric method and the lipid indicator values from Raman spectroscopy analysis showed linear correlation (R^2 value =0.9153) (**Figure 3A**). The level of increase in total lipid also positively correlates with increase in amount of β-carotene produced (**Figure 3B**). Overall, the increased lipid content in strains with higher β-carotene yields based on lipid indicators from Raman spectroscopy analysis is supported by total lipid quantification.

3.6 Mutations that confer increased β-carotene yield also increase lycopene production

To determine if the identified beneficial mutations for β-carotene production can also benefit the production of other carotenoids, we constructed the three best combinations of mutations ([*ALG6 1411* and *AQY3/DAK2 int*], [*ALG6 1411, EPL1 1754* and *AQY3/DAK2 int*] and [*ALG6*, *AQY3/DAK2 int* and *YMRCTy1-3 1078*]) in non-carotenogenic strain FY2, to generate strains YAG42, YAG43 and YAG44, respectively. The heterologous genes *crtE*, *crtB* and *crtI* for lycopene biosynthesis were integrated into these strains to generate lycopene-producing versions of these mutants. A control strain (YAG41) with just the lycopene biosynthetic genes integrated into FY2 was also constructed. As shown in **Figure 4**, the control strain produced 30.2 ± 0.3 mg/g DCW of lycopene. In strain YAG42, a 26% increase in

lycopene production was observed, with total lycopene yield of 38.2 ± 0.3 mg/g DCW. The best combination of mutations for β-carotene production (*ALG6 1411*, *AQY3/DAK2 int* and *YMRCTy1-3 1078*) was also the best combination observed for lycopene production, with strain YAG44 producing 42.5 ± 0.3mg/g DCW. Based on knowledge gained from ALE experiments, we successfully generated a strain that achieved ~40% increase in lycopene yield compared with the wild-type YAG41.

These engineered lycopene hyperproducers also exhibited increased lipid content compared to the control strain according to Raman spectroscopy analysis (Supplementary **Figure S4**).

Conclusions

In our prior work, we hypothesized that increased productivity of β-carotene, a compound with antioxidant properties, is correlated with tolerance to oxidative stress. Using hydrogen peroxide as the selective pressure in adaptive laboratory evolution experiments, mutants with enhanced production of β-carotene were successfully isolated. In ALE experiments, the phenotypes of the mutants selected are dependent on the selection strategy. While we successfully isolated hyperproducers, it was not clear if β-carotene production was the primary phenotype being selected for, or if it was a secondary consequence of increased oxidative stress tolerance during the ALE experiment. In this work, we sequenced the β-carotene hyperproducers from the ALE experiments, identified beneficial mutations for β-carotene production, and demonstrated that the most impactful beneficial mutations for β-carotene production did not confer enhanced tolerance to hydrogen peroxide in the absence of carotenoids production. The results strongly suggest that the increased production of the desired

441 antioxidant compound was indeed the phenotype selected for using this ALE strategy. By
442 generating combinations of mutations identified from different β-carotene hyperproducers, we
443 identified a superior combination of mutations that led to ~30% increase in β-carotene yield in
444 test tubes and ~82% using a benchtop bioreactor compared with the best evolved hyperproducer
445 SM14. Using Raman spectroscopy, total lipid quantification, and qRT-PCR, we found that the
446 best β-carotene producers exhibit increased lipid biosynthesis. Finally, to demonstrate the broad
447 utility of the findings from this work, we rationally constructed an improved lycopene producer
448 using the best combinations of mutations identified for β-carotene, resulting in a strain that
449 produces ~40% more lycopene (42.5±0.3 mg/g DCW) compared with the baseline strain. In
450 conclusion, we have successfully identified additional genes important for increasing carotenoids
451 yields in yeast, which cannot currently be determined using rational approaches, and we have
452 demonstrated that these mutations confer benefits to the production of carotenoids other than β-
453 carotene.
454
455 **Acknowledgements**
456 This work was partially financially supported by NSF CBET-1605347. We thank Lilliana
457 Andrea Morales Laverde, Katherine Miller and Ibrahim Hilaly for assistance with experiments.
458

References

1. Alper H, Miyaoku K, Stephanopoulos G (2005) Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. Nat Biotechnol 23:612-616. doi:10.1038/nbt1083
2. An GH, Jang BG, Cho MH (2001) Cultivation of the carotenoid-hyperproducing mutant 2A2N of the red yeast *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) with molasses. J Biosci Bioeng 92:121-125. doi:10.1016/S1389-1723(01)80211-5
3. Arias P, Diez-Muniz S, Garcia R, Nombela C, Rodriguez-Pena JM, Arroyo J (2011) Genome-wide survey of yeast mutations leading to activation of the yeast cell integrity MAPK pathway: novel insights into diverse MAPK outcomes. BMC Genomics 12:390. doi:10.1186/1471-2164-12-390
4. Arigoni D, Sagner S, Latzel C, Eisenreich W, Bacher A, Zenk MH (1997) Terpenoid biosynthesis from 1-deoxy-D-xylulose in higher plants by intramolecular skeletal rearrangement. Proc Natl Acad Sci U S A 94:10600-10605. doi:10.1073/pnas.94.20.10600
5. Arlt H, Perz A, Ungermann C (2011) An Overexpression Screen in *Saccharomyces cerevisiae* Identifies Novel Genes that Affect Endocytic Protein Trafficking. Traffic 12:1592-1603. doi:10.1111/j.1600-0854.2011.01252.x
6. Asadollahi MA, Maury J, Moller K, Nielsen KF, Schalk M, Clark A, Nielsen J (2008) Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. Biotechnol Bioeng 99:666-677. doi:10.1002/bit.21581
7. Bailey SF, Hinz A, Kassen R (2014) Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. Nature Communications 5:4076. doi:10.1038/ncomms5076
8. Bao Z, Xiao H, Liang J, Zhang L, Xiong X, Sun N, Si T, Zhao H (2015) Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. ACS synthetic biology 4:585-594. doi:10.1021/sb500255k
9. Benni ML, Neigeborn L (1997) Identification of a new class of negative regulators affecting sporulation-specific gene expression in yeast. Genetics 147:1351-1366
10. Bhosale P, Bernstein PS (2004) Beta-carotene production by *Flavobacterium multivorum* in the presence of inorganic salts and urea. J Ind Microbiol Biotechnol 31:565-571. doi:10.1007/s10295-004-0187-9
11. Bhosale P, Gadre RV (2001) Production of beta-carotene by a *Rhodotorula glutinis* mutant in sea water medium. Bioresour Technol 76:53-55. doi:10.1016/S0960-8524(00)00075-4
12. Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Canadian journal of biochemistry and physiology 37:911-917. doi:10.1139/o59-099
13. Breil C, Abert Vian M, Zemb T, Kunz W, Chemat F (2017) "Bligh and Dyer" and Folch Methods for Solid-Liquid-Liquid Extraction of Lipids from Microorganisms. Comprehension of Solvatation Mechanisms and towards Substitution with Alternative Solvents. International journal of molecular sciences 18. doi:10.3390/ijms18040708
14. Chen D, Han Y, Gu Z (2006) Application of statistical methodology to the optimization of fermentative medium for carotenoids production by *Rhodobacter sphaeroides*. Process Biochem 41:1773-1778. doi:10.1016/j.procbio.2006.03.023
15. Chen Y, Xiao W, Wang Y, Liu H, Li X, Yuan Y (2016) Lycopene overproduction in *Saccharomyces cerevisiae* through combining pathway engineering with host engineering. Microbial cell factories 15:113. doi:10.1186/s12934-016-0509-4
16. Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M (2015) Raman spectroscopy of lipids: a review. Journal of Raman Spectroscopy 46:4-20. doi:10.1002/jrs.4607

505 17. de Ruijter JC, Jurgens G, Frey AD (2017) Screening for novel genes of *Saccharomyces cerevisiae*
506 involved in recombinant antibody production. FEMS Yeast Res 17. doi:10.1093/femsyr/fow104
507 18. Deatherage DE, Barrick JE (2014) Identification of mutations in laboratory-evolved microbes from
508 next-generation sequencing data using breseq. Methods in molecular biology (Clifton, NJ)
509 1151:165-188. doi:10.1007/978-1-4939-0554-6_12
510 19. Dohmen RJ, Willers I, Marques AJ (2007) Biting the hand that feeds: Rpn4-dependent feedback
511 regulation of proteasome function. Biochim Biophys Acta 1773:1599-1604.
512 doi:10.1016/j.bbamcr.2007.05.015
513 20. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high
514 throughput capability. PloS one 3:e3647. doi:10.1371/journal.pone.0003647
515 21. Evans CT, Ratledge C (1984) Effect of Nitrogen Source on Lipid Accumulation in Oleaginous Yeasts.
516 Microbiology 130:1693-1704. doi:10.1099/00221287-130-7-1693
517 22. Fakhry EM, El Maghraby DM (2015) Lipid accumulation in response to nitrogen limitation and
518 variation of temperature in *Nannochloropsis salina*. Botanical studies 56:6-6.
519 doi:10.1186/s40529-015-0085-7
520 23. Fred W, Catherine D, L. R-HS (1995) Construction of a set of convenient *Saccharomyces cerevisiae*
521 strains that are isogenic to S288C. Yeast 11:53-55. doi:10.1002/yea.320110107
522 24. Goossens KVY, Ielasi FS, Nookaew I, Stals I, Alonso-Sarduy L, Daenen L, Van Mulders SE, Stassen C,
523 van Eijsden RGE, Siewers V, Delvaux FR, Kasas S, Nielsen J, Devreese B, Willaert RG (2015)
524 Molecular Mechanism of Flocculation Self-Recognition in Yeast and Its Role in Mating and
525 Survival. mBio 6:e00427-00415. doi:10.1128/mBio.00427-15
526 25. Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ
527 (2008) The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited
528 Environments in Yeast. PLoS Genet 4:e1000303. doi:10.1371/journal.pgen.1000303
529 26. Higgins VJ, Alic N, Thorpe GW, Breitenbach M, Larsson V, Dawes IW (2002) Phenotypic analysis of
530 gene deletant strains for sensitivity to oxidative stress. Yeast 19:203-214. doi:10.1002/yea.811
531 27. Jeon Y-C, Cho C-W, Yun Y-S (2006) Combined effects of light intensity and acetate concentration on
532 the growth of unicellular microalga *Haematococcus pluvialis*. Enzyme Microb Technol 39:490-
533 495. doi:10.1016/j.enzmictec.2005.12.021
534 28. Kapitzky L, Beltrao P, Berens TJ, Gassner N, Zhou C, Wuster A, Wu J, Babu MM, Elledge SJ, Toczyski D,
535 Lokey RS, Krogan NJ (2010) Cross-species chemogenomic profiling reveals evolutionarily
536 conserved drug mode of action. Mol Syst Biol 6:451. doi:10.1038/msb.2010.107
537 29. Kim H, Kim A, Cunningham KW (2012) Vacuolar H+-ATPase (V-ATPase) promotes vacuolar membrane
538 permeabilization and nonapoptotic death in stressed yeast. J Biol Chem 287:19029-19039.
539 doi:10.1074/jbc.M112.363390
540 30. Kolouchova I, Matatkova O, Sigler K, Masak J, Rezanka T (2016) Lipid accumulation by oleaginous and
541 non-oleaginous yeast strains in nitrogen and phosphate limitation. Folia Microbiol (Praha)
542 61:431-438. doi:10.1007/s12223-016-0454-y
543 31. Kot AM, Błażejak S, Kurcz A, Gientka I, Kieliszek M (2016) *Rhodotorula glutinis*-potential source of
544 lipids, carotenoids, and enzymes for use in industries. Appl Microbiol Biotechnol 100:6103-6117.
545 doi:10.1007/s00253-016-7611-8
546 32. Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene
547 expression in *Escherichia coli*. Science (New York, NY) 324:255-258.
548 doi:10.1126/science.1170160
549 33. Little S (2001) Amplification-refractory mutation system (ARMS) analysis of point mutations. Current
550 protocols in human genetics Chapter 9:Unit 9.8. doi:10.1002/0471142905.hg0908s07

34. Lohning C, Ciriacy M (1994) The *TYE7* gene of *Saccharomyces cerevisiae* encodes a putative bHLH-LZ transcription factor required for Ty1-mediated gene expression. Yeast 10:1329-1339. doi:10.1002/yea.320101010
35. Luria SE, Delbruck M (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics 28:491-511
36. Malisorn C, Suntornsuk W (2008) Optimization of β-carotene production by *Rhodotorula glutinis* DM28 in fermented radish brine. Bioresour Technol 99:2281-2287. doi:10.1016/j.biortech.2007.05.019
37. Mata-Gómez LC, Montañez JC, Méndez-Zavala A, Aguilar CN (2014) Biotechnological production of carotenoids by yeasts: an overview. Microbial cell factories 13:12-12. doi:10.1186/1475-2859-13-12
38. Molin M, Norbeck J, Blomberg A (2003) Dihydroxyacetone kinases in *Saccharomyces cerevisiae* are involved in detoxification of dihydroxyacetone. J Biol Chem 278:1415-1423. doi:10.1074/jbc.M203030200
39. Nanou K, Roukas T, Kotzekidou P (2007) Role of hydrolytic enzymes and oxidative stress in autolysis and morphology of *Blakeslea trispora* during beta-carotene production in submerged fermentation. Appl Microbiol Biotechnol 74:447-453. doi:10.1007/s00253-006-0666-1
40. Nishizaki T, Tsuge K, Itaya M, Doi N, Yanagawa H (2007) Metabolic Engineering of Carotenoid Biosynthesis in *Escherichia coli* by Ordered Gene Assembly in *Bacillus subtilis*. Appl Environ Microbiol 73:1355. doi:10.1128/AEM.02268-06
41. Oliveira R, Lages F, Silva-Graca M, Lucas C (2003) Fps1p channel is the mediator of the major part of glycerol passive diffusion in *Saccharomyces cerevisiae*: artefacts and re-definitions. Biochim Biophys Acta 1613:57-71. doi:10.1016/S0005-2736(03)00138-X
42. Olson ML, Johnson J, Carswell WF, Reyes LH, Senger RS, Kao KC (2016) Characterization of an evolved carotenoids hyper-producer of *Saccharomyces cerevisiae* through bioreactor parameter optimization and Raman spectroscopy. J Ind Microbiol Biotechnol 43:1355-1363. doi:10.1007/s10295-016-1808-9
43. P. B, R.V. G (2001) Optimization of carotenoid production from hyper-producing *Rhodotorula glutinis* mutant 32 by a factorial approach. Lett Appl Microbiol 33:12-16. doi:doi:10.1046/j.1472-765X.2001.00940.x
44. Portnoy VA, Bezdan D, Zengler K (2011) Adaptive laboratory evolution--harnessing the power of biology for metabolic engineering. Curr Opin Biotechnol 22:590-594. doi:10.1016/j.copbio.2011.03.007
45. Prelich G (2012) Gene overexpression: uses, mechanisms, and interpretation. Genetics 190:841-854. doi:10.1534/genetics.111.136911
46. Raja R, Hemaiswarya S, Rengasamy R (2007) Exploitation of *Dunaliella* for beta-carotene production. Appl Microbiol Biotechnol 74:517-523. doi:10.1007/s00253-006-0777-8
47. Reiss G, te Heesen S, Zimmerman J, Robbins PW, Aebi M (1996) Isolation of the *ALG6* locus of *Saccharomyces cerevisiae* required for glucosylation in the N-linked glycosylation pathway. Glycobiology 6:493-498
48. Reyes LH, Gomez JM, Kao KC (2014) Improving carotenoids production in yeast via adaptive laboratory evolution. Metab Eng 21:26-33. doi:10.1016/j.ymben.2013.11.002
49. Richardson B, Orcutt DM, Schwertner HA, Martinez CL, Wickline HE (1969) Effects of nitrogen limitation on the growth and composition of unicellular algae in continuous culture. Applied microbiology 18:245-250
50. Sitepu IR, Sestric R, Ignatia L, Levin D, German JB, Gillies LA, Almada LA, Boundy-Mills KL (2013) Manipulation of culture conditions alters lipid content and fatty acid profiles of a wide variety of

known and new oleaginous yeast species. Bioresour Technol 144:360-369. doi:10.1016/j.biortech.2013.06.047
51. Smukalla S, Caldara M, Pochet N, Beauvais A, Guadagnini S, Yan C, Vinces MD, Jansen A, Prevost MC, Latge JP, Fink GR, Foster KR, Verstrepen KJ (2008) *FLO1* is a variable green beard gene that drives biofilm-like cooperation in budding yeast. Cell 135:726-737. doi:10.1016/j.cell.2008.09.037
52. Sobel SG, Wolin SL (1999) Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. Molecular biology of the cell 10:3849-3862. doi:10.1091/mbc.10.11.3849
53. Stankunas K, Berger J, Ruse C, Sinclair DA, Randazzo F, Brock HW (1998) The enhancer of polycomb gene of Drosophila encodes a chromatin protein conserved in yeast and mammals. Development 125:4055-4066
54. Stevenson LF, Kennedy BK, Harlow E (2001) A large-scale overexpression screen in *Saccharomyces cerevisiae* identifies previously uncharacterized cell cycle genes. Proceedings of the National Academy of Sciences 98:3946-3951. doi:10.1073/pnas.051013498
55. Sullivan DP, Georgiev A, Menon AK (2009) Tritium suicide selection identifies proteins involved in the uptake and intracellular transport of sterols in *Saccharomyces cerevisiae*. Eukaryot Cell 8:161-169. doi:10.1128/ec.00135-08
56. Verwaal R, Wang J, Meijnen JP, Visser H, Sandmann G, van den Berg JA, van Ooyen AJ (2007) High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. Appl Environ Microbiol 73:4342-4350. doi:10.1128/aem.02759-06
57. Weatherston JD, Seguban RKO, Hunt D, Wu H-J (2018) Low-Cost and Simple Fabrication of Nanoplasmonic Paper for Coupled Chromatography Separation and Surface Enhanced Raman Detection. ACS Sensors 3:852-857. doi:10.1021/acssensors.8b00098
58. Wilke CM, Adams J (1992) Fitness effects of Ty transposition in *Saccharomyces cerevisiae*. Genetics 131:31-42
59. Wu T, Ye L, Zhao D, Li S, Li Q, Zhang B, Bi C, Zhang X (2017) Membrane engineering - A novel strategy to enhance the production and accumulation of beta-carotene in *Escherichia coli*. Metab Eng 43:85-91. doi:10.1016/j.ymben.2017.07.001
60. Yang J, Guo L (2014) Biosynthesis of beta-carotene in engineered *E. coli* using the MEP and MVA pathways. Microbial cell factories 13:160. doi:10.1186/s12934-014-0160-x

Table and Figure captions:

Table 1. List of identified mutations.

Table 2. Growth kinetics of reconstructed mutants. Two-tailed student t-test was used to calculate p-values comparing with YLH2. Bold: p-value < 0.001 compared with YLH2 using two-tailed Student t-test.

Table 3. Impacts of mutation on relative H2O2 tolerance in non-carotenogenic, carotenogenic, and carotenogenic strains lacking *CTT1*.

Figure 1. β-carotene production of evolved and reconstructed mutants. A) Single mutations; asterisks: p-value < 0.05 using 2 tailed Student t-test against YLH2. B) Reconstructed evolved mutants. C) Multiple mutations; asterisks: higher production compared with SM14 with p-value < 0.05 using 2 tailed Student t-test. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078*.

Figure 2. Relative Raman peak intensity ratios compared with YLH2. FA: fatty acid. Error bars represent a confidence interval about the sample mean, with p = 0.90. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078*.

Figure 3: Lipid analysis using Raman and total lipid quantification. A) Correlation between lipid indicator obtained from Raman spectroscopy analysis and total lipid quantification results using the Bligh and Dyer method. Error bars are standard errors based on three biological replicates and three technical replicates each for lipid quantification results and three biological replicates and sixteen technical replicates for Raman spectroscopy analysis. YAG10 contains the *YMRCTy1-3 1078* mutation (j) and YAG28 contains *ALG6 1411* (h), *DAK2/AQY3 int* (e), and *YMRCTy1-3 1078* (j) mutations. B) Positive correlation between levels of increase in total lipid and β-carotene production of hyperproducers relative to YLH2.

Figure 4: Effect of beneficial mutations on lycopene production. Asterisk: p-value < 0.05 using two-tailed student t-test compared with YAG41. Mutations: (e) *DAK2/AQY3 int*, (g) *EPL1 1754*, (h) *ALG6 1411*, (j) *YMRCTy1-3 1078*.

664 Supplementary tables and figures:

665 Table S1. List of strains used.

666 Table S2. Guide RNA (gRNA) sequences for site-directed mutagenesis using CRISPR/Cas9.

667 Table S3. Primer sequences.

668 Table S4. Characteristic wavelengths for Raman spectroscopy and their corresponding species.

669 Table S5. qRT-PCR results for genes flanking up and downstream of identified intergenic or
670 Ty1 mutations. Relative to YLH2.

671 Table S6. qRT-PCR results for relative expression levels of select genes involved in lipid
672 biosynthesis, sterol biosynthetic pathway and oxidative stress response. Bold: p-value < 0.05
673 using a student t-test compared with YLH2.

674 Table S7. Effect of promoters on lycopene yield. Three biological replicates were used for
675 analysis.

676 Table S8. Primers used for small RNA detection

677 Figure S1: Oxidative stress tolerance of reconstructed single mutants in FY2 subjected to 1 h
678 exposure in 1 M H_2O_2. Strains are listed in increasing order of beta-carotene observed in their
679 YLH2 counterparts.

680 Figure S2: Oxidative stress tolerance of reconstructed single mutants in carotenogenic strain of
681 FY2 subjected to 1 h exposure in 1 M H_2O_2. Strains are listed in increasing order of β-carotene
682 production.

Figure S3: Oxidative stress tolerance of reconstructed single mutants in carotenogenic FY2 Δ*ctt1* strain subjected to 30 min exposure in 100 mM H_2O_2. Strains are listed in increasing order of beta-carotene production.

Figure S4. Raman results for lycopene producers. Relative peak intensity ratios compared with YAG41. FA: fatty acid.

Figure S5: Correlation between β-carotene quantification using HPLC and using absorbance at OD_{454}. A) and C) HPLC calibration curve using pure compounds for β-carotene and lycopene, respectively; peak area under the curve is shown. B) Standard curve for β-carotene using absorbance measurements at OD_{454}. D) Amount of lycopene and β-carotene in strain YLH2, SM14 and YAG28 based on HPLC quantification. E) Correlation between HPLC and absorbance-based quantification of β-carotene for strains tested. Error bars are standard deviations based on three biological replicates per strain.

Figure S6. Epistatic interactions between different mutations on β-carotene production. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078*.

699 Table 1. List of identified mutations.

Strain	Chr	Location	Type	Gene/Mutation	Amino acid change	Gene function
SM12	2	716,077	G->A	HIS7 389 (a)	S130L	Histidine biosynthesis
	3	58,825	T->A	SRO9/GFD2 int (b)	intergenic	
	15	977,984	T->G	TYE7 86 (c)	D29A	Putative regulator of Ty1
	mito	954	-24Δbp	tP(UGG)Q/ST15Sr RNA	deletion in noncoding sequence	
SM13	1	204,327	G->A	FLO1 925 (d)	V309I	Protein involved in flocculation
	2	716,077	G->A	HIS7 389 (a)	S130L	Histidine biosynthesis
	3	58,825	T->A	SRO9/GFD2 int (b)	intergenic	
	6	22,874	A->G	DAK2/AQY3 int (e)	intergenic	
	7	354,893	C->T	SCY1 1836 (f)	I612I	Putative kinase
	mito	954	-24Δbp	tP(UGG)Q/ST15Sr RNA	deletion in noncoding sequence	
SM14	6	88592	T>G	EPL1 1754 (g)	L585W	Subunit of histone acetytransferase
	7	124954^ 124955	insA	MDS3 insA (i)	nonsense mutation	Putative regulator
	13	361550	G>A	YMRCTy1-3 1078 (j)	H360Y	Ty1 element
	15	330827	A>C	ALG6 1411 (h)	S471R	Glucosyltransferase
	2	716077	G>A	HIS7 389 (a)	S130L	Histidine biosynthesis

700

701

Table 2. Growth kinetics of reconstructed mutants. Two-tailed student t-test was used to calculate p-values comparing with YLH2. Bold: p-value < 0.001 compared with YLH2 using two-tailed Student t-test.

Strain	Mutation	Specific growth rate (per hour)	Lag phase (hours)	OD after 24hr (Avg±Stdev)
YLH2		0.28±0.02	4.29±0.23	0.98±0.05
YAG01	HIS7 389 (a)	**0.23±0.01**	**0.76 ± 0.17**	0.82±0.06
YAG02	SRO9/GFD2 int (b)	0.28±0.00	**0.96 ± 0.12**	0.94±0.02
YAG03	TYE7 86 (c)	0.26±0.02	**0.66 ± 0.23**	0.91±0.03
YAG04	FLO1 925 (d)	**0.22±0.01**	**0.45 ± 0.17**	0.90±0.05
YAG05	DAK2/AQY3 int (e)	**0.23±0.01**	**0.91 ± 0.19**	1.01±0.09
YAG06	SCY1 1836 (f)	0.27±0.01	**0.45 ± 0.17**	0.97±0.03
YAG07	EPL1 1754 (g)	**0.20±0.00**	**1.87 ± 0.12**	0.94±0.04
YAG08	ALG6 1411 (h)	0.24±0.02	**1.16 ± 0.23**	1.00±0.05
YAG09	MDS3 ins (i)	0.27±0.01	**0.35 ± 0.12**	0.96±0.05
YAG10	YMRCTy1-3 1078 (j)	0.27±0.01	**0.40 ± 0.16**	1.11±0.12

Table 3. Impacts of mutation on relative H_2O_2 tolerance in non-carotenogenic, carotenogenic, and carotenogenic strains lacking *CTT1*.

	FY2		FY2+carotenogenic			FY2+Δctt1+carotenogenic		
Mutation	Strain	relative H_2O_2 tolerance	Strain	relative H_2O_2 tolerance	β-carotene (mg/g DCW)	Strain	relative H_2O_2 tolerance	β-carotene (mg/g DCW)
wild-type	FY2		YAG45		7.2 ± 0.1	YAG56		6.8 ± 0.2
HIS7 389 (a)	YAG31	ND	YAG46	ND	**2.5 ± 0**	YAG57	ND	**2.4 ± 0.3**
SCY1 1836 (f)	YAG36	-	YAG51	ND	**9.4 ± 0.9**	YAG62	ND	**9.3 ± 0.1**
MDS3 ins (i)	YAG39	+	YAG54	+	**9.7 ± 0.3**	YAG65	+	**9.3 ± 0.1**
EPL1 1754 (g)	YAG37	+	YAG52	+	**10.5 ± 0.1**	YAG63	+	**10.3 ± 0.2**
DAK2/AQY3 int (e)	YAG35	-	YAG50	ND	**10.5 ± 0**	YAG61	ND	**10.3 ± 0.2**
FLO1 925 (d)	YAG34	+	YAG49	+	**11.2 ± 0.2**	YAG60	+	**10.8 ± 0.5**
ALG6 1411 (h)	YAG38	+	YAG53	ND	**11.4 ± 0.1**	YAG64	ND	**11.3 ± 0.1**
TYE7 86 (c)	YAG33	ND	YAG48	ND	**11.2 ± 0.1**	YAG59	ND	**11.2 ± 0.2**
SRO9/GFD2 int (b)	YAG32	-	YAG47	+	**15.5 ± 0**	YAG58	+	**15.3 ± 0.2**
YMRCTy1-3 1078 (j)	YAG40	-	YAG55	+	**16.4 ± 0.1**	YAG66	+	**15.7 ± 0.3**

ND: not different from wild-type.
+: ~10X higher survival compared to wild-type.
-: ~10X lower survival compared to wild-type.
Samples were treated with 1 M H_2O_2 for one hour. The Δ*ctt1* samples were treated with 100 mM H_2O_2 for 30 minutes.
Bold: p-value < 0.005 using two-tailed Student t-test compared with wild-type.

Figure 1. β-carotene production of evolved and reconstructed mutants. A) Single mutations; asterisks: p-value < 0.05 using two-tailed Student t-test against YLH2. B) Reconstructed evolved mutants. C) Multiple mutations; asterisks: higher production compared with SM14 with p-value < 0.05 using 2 tailed Student t-test. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078*.

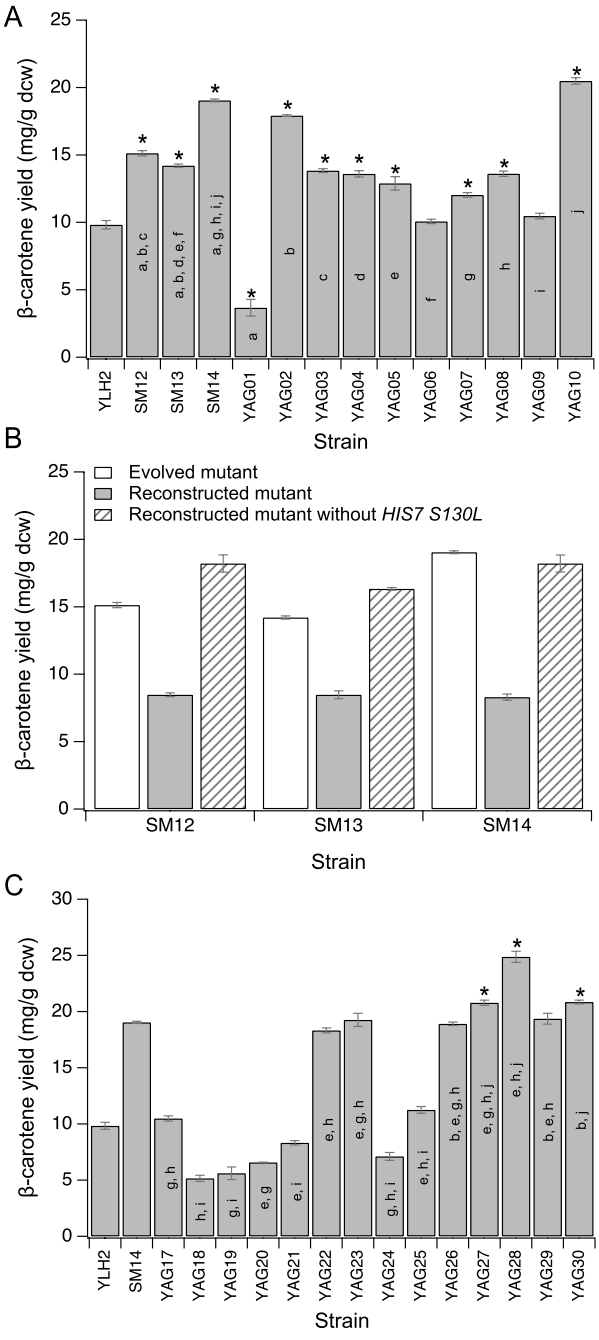


Figure 2. Relative Raman peak intensity ratios compared with YLH2. FA: fatty acid. Error bars represent a confidence interval about the sample mean, with p = 0.90. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078*.

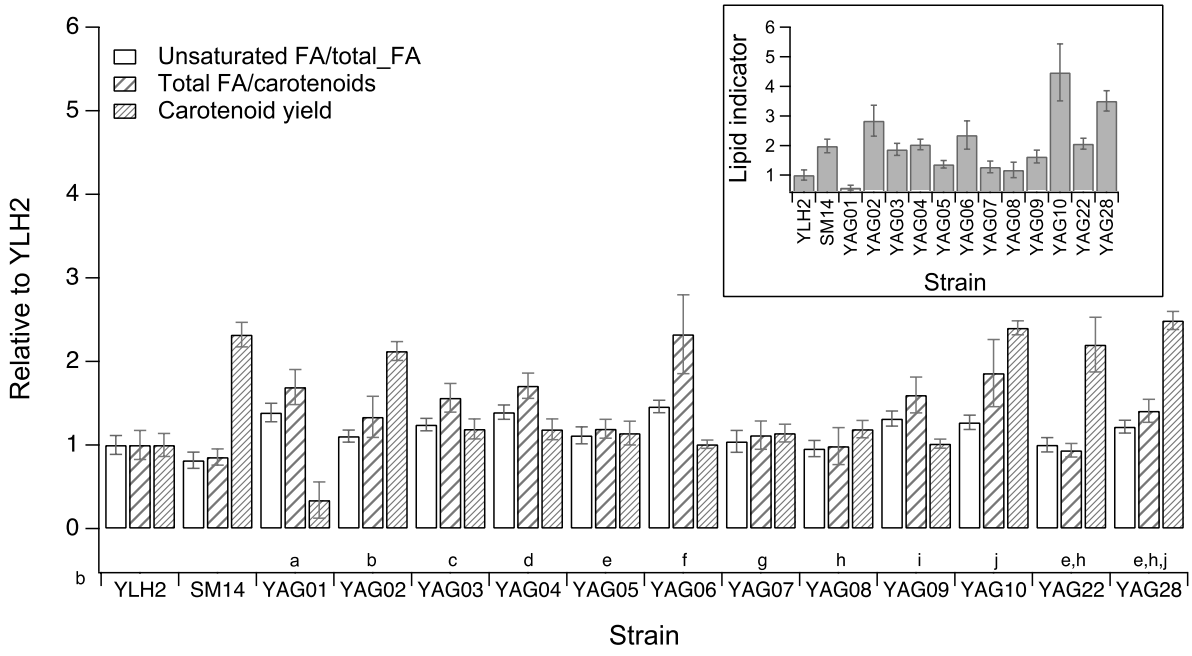
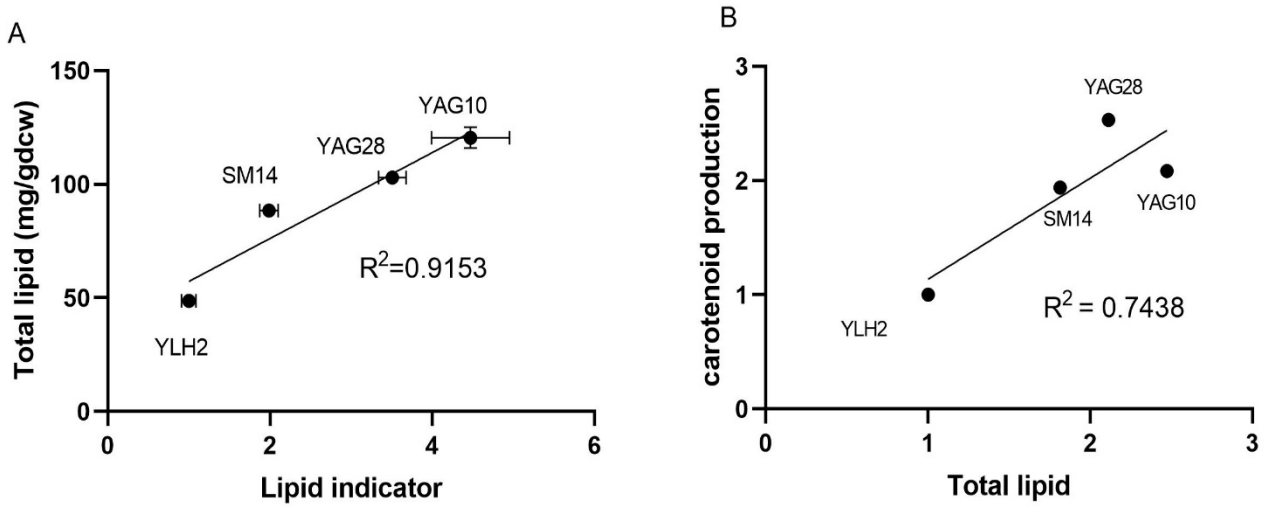
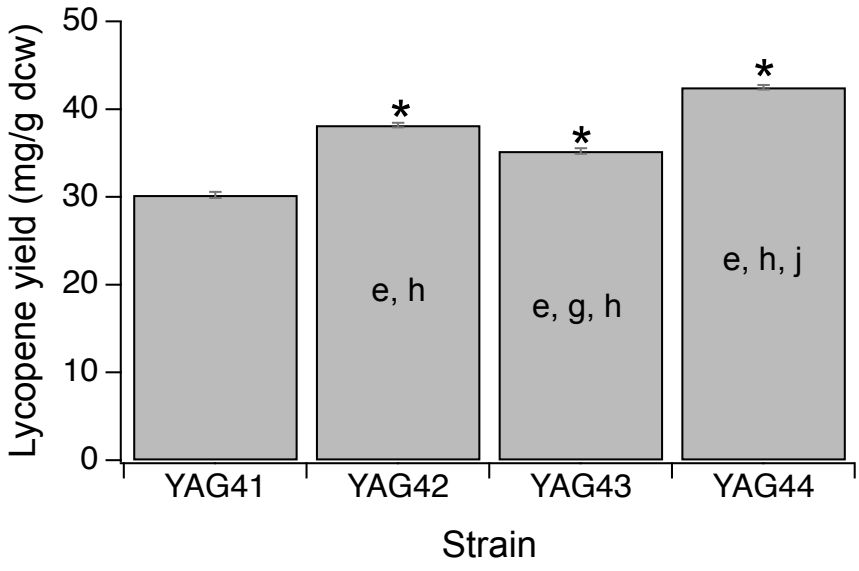


Figure 3: Lipid analysis using Raman and total lipid quantification. A) Correlation between lipid indicator obtained from Raman spectroscopy analysis and total lipid quantification results using the Bligh and Dyer method. Error bars are standard errors based on three biological replicates and three technical replicates each for lipid quantification results and three biological replicates and sixteen technical replicates for Raman spectroscopy analysis. YAG10 contains the *YMRCTy1-3 1078* mutation (j) and YAG28 contains *ALG6 1411* (h), *DAK2/AQY3 int* (e), and *YMRCTy1-3 1078* (j) mutations. B) Positive correlation between levels of increase in total lipid and β-carotene production of hyperproducers relative to YLH2.



751 Figure 4: Effect of beneficial mutations on lycopene production. Asterisk: p-value < 0.05 using
752 2-tailed student t-test compared with YAG41. Mutations: (e) *DAK2/AQY3 int*, (g) *EPL1 1754*,
753 (h) *ALG6 1411*, (j) *YMRCTy1-3 1078*.

754



755