

Supplementary Information for

Elucidating aromatic acid tolerance at low pH in *Saccharomyces cerevisiae* using adaptive laboratory evolution

Rui Pereira, Elsayed T. Mohamed, Mohammad S. Radi, Markus J. Herrgård, Adam M. Feist, Jens Nielsen, Yun Chen

Corresponding author: Yun Chen

Email: yunc@chalmers.se

This PDF file includes:

Supplementary text Figure S1 Tables S1 to S8 SI References

Supplementary Information Text

Materials and Methods

Strains, media and DNA handling

The strains used and constructed in this work are listed in Table S6. The strain GL01 (1) was used for all the ALE experiments because the presence of the G418 resistance marker (KANMX) and a GFP gene (under the control of a *GAL1* promoter) allowed to check the evolving populations for possible contamination during the ALE. Reverse engineering of tolerance mutations was performed on the strain IMX585 (2), which has a copy of Cas9 integrated in a CEN.PK113-7D background. The coumaric acid production strains tested in this work were based on strains obtained from (3). *Escherichia coli* DH5α was used for plasmid isolation and maintenance using the competence and transformation procedures developed by (4) as described in (5).

The oligonucleotides used here are listed in Table S7 and were purchased from Eurofins Genomics. High fidelity PCR was performed with Phusion polymerase (ThermoFisher) and colony PCR was performed with DreamTaq DNA polymerase (ThermoFisher). The appropriate GeneJET kits from ThermoFisher were used for plasmid extraction and PCR product purification.

Plasmid selection and maintenance in *E. coli* was performed in LB medium containing 10 g/L of peptone, 10 g/L of NaCl, 5 g/L of yeast extract and supplemented with 100 mg/L of ampicillin sodium salt. Solid LB was prepared by adding 16 g/L of agar to the recipe above. The selection of *S. cerevisiae* strains expressing the KANMX marker was done in YPD medium containing 10 g/L yeast extract, 20 g/L of peptone, 20 g/L of glucose and supplemented with 200 mg/L G418 disulfate (Sigma-Aldrich). Selection of strains expressing the *amdSYM* marker was performed in medium with 0.6 g/L of acetamide as

the only nitrogen source as described in (6). URA⁺ strains were selected in Synthetic Dextrose (SD) medium containing 6.7 g/L of yeast nitrogen base without amino acids (ForMedium), 0.77 g/L of complete supplement mixture without uracil (ForMedium) and 20 g/L of glucose. Solid versions of each medium were prepared by adding 20 g/L of agar. ALE experiments and growth tests were performed in the minimal medium described by (7), which contained 20 g/L of glucose, 5 g/L of $(NH4)_2SO_4$, 3 g/L of KH_2PO_4 , 0.5 g/L of MgSO₄·7H₂O, 1 mL/L of vitamin solution and 1 mL/L of trace metal solution. The pH of the medium was adjusted to 3.5 with HCl and buffered at this pH by adding 140 mL/L of 0.5 M citrate solution and 60 mL/L of 1M $Na₂HPO₄$. For testing strains in alternative nitrogen sources, $(NH4)_2SO_4$ was replaced with 6.6 g/L of K_2SO_4 and the required amino acid was added in the appropriate concentration (5 g/L of L-phenylalanine, 5g/L of L-tryptophan or 0.5 g/L of L-tyrosine). Coumaric acid, ferulic acid, phenylacetic acid and indoleacetic acid were purchased from Sigma-Aldrich and added to the culture media in the appropriate concentration. The solid version of this medium was prepared by adding agarose to a final concentration of 10 g/L.

Strain construction

The list of plasmids used and constructed in this study is provided in Table S8. The reverse engineered strains containing deletions, point mutations and insertions were constructed using the single gRNA method described elsewhere (2). The primers used for the deletion of *ESBP6* and *PDR12* were designed using the Yeastriction tool (http://yeastriction.tnw.tudelft.nl). For each deletion a set of six oligonucleotides were ordered: two complementary 120 bp oligonucleotides with the desired 20 bp gRNA target sequence (without PAM site) flanked by two 50 bp sequences homologous to the gRNA site on plasmid pMEL13 (2); two complementary 120 bp oligonucleotides composed of 60 bp homologous to the region upstream of the coding sequence of the target gene and 60

bp homologous to the region downstream of the coding sequence of the target gene; and two colony PCR primers for verifying the desired region was removed for the genome. Each pair of complementary oligonucleotides were mixed in equimolar amounts and annealed by heating the mixture to 95 \degree C for five minutes and allowing it to cool down at room temperature. In order to generate plasmids expressing the desired gRNA, each double stranded gRNA insert was cloned into the plasmid pMEL13 (linearized with primers 1 and 2) using CPEC cloning (8), yielding plasmids pMEL13-ESBP6-KO and pMEL13- PDR12-KO. The sequence of each plasmid was confirmed by sequencing with primer 3. Single deletion strains for either gene were obtained by transforming 100 ng of the desired gRNA plasmid with 2 μg of double stranded repair fragment into the strain IMX585 using the high-efficiency protocol described by (9). The transformation mixture was spread on plates containing G418 and individual colonies tested for the deletion using colony PCR. The deletion of *ESBP6* was performed with the primers 4-7 and confirmed with the primers 8/9. *PDR12* was deleted with the primers 10-13 and tested using the flanking primers 14/15.

The introduction of single nucleotide substitutions into the strain IMX585 was done using the same protocol described for deletions but with the following modifications: the gRNA site was selected based on the proximity to the substitution site and the possibility of inactivating the PAM site using a synonymous substitution; and the two complementary oligonucleotides for the repair fragment were designed to introduce the desired mutation and inactivating the PAM site. The plasmids expressing the gRNAs for introducing each mutation were obtained using CPEC cloning (8) using linearized pMEL13 plasmid and the appropriate double stranded repair fragment, resulting in: pMEL13-ARO80-706, pMEL13- ARO80-910 and pMEL13-NRG1-226. The mutation *ARO80*S706T was introduced using primers 16-19, amplified using primers 20/21 and sequenced with primer 20. The *ARO80*E910D mutation was introduced using primers 22-25, amplified using primers 20/21

and sequenced with primer 21. The NRG1^{H226N} mutation was introduced using primers 26-29, amplified using primers 30/31 and sequenced with primer 30. The double mutants *ARO80*S706T+*ΔESBP6* and *ARO80*E910D+*ΔESBP6* were obtained by deleting the gene *ESBP6* on each of the *ARO80* point mutation strains using the protocol described above. For the integration of *ESBP6* under the control of a *TEF1* promoter on the site X-2 of chromosome X (10) the same protocol described for gene deletions was used but in this case using a gRNA plasmid designed to create a double strand break on the desired integration site (pQC059) (3). The integration cassette was obtained by amplifying the following DNA segments: the region downstream of the site X-2 was amplified from plasmid pX-2 (10) using primers 32/33, the *TEF1* promoter was amplified from plasmid pSPGM1 (11) using primers 34/35, the *ESBP6* coding sequence was amplified from the genomic DNA of CEN.PK113-7D using primers 36/37 and the *ADH1* terminator and upstream region of the site X-2 were amplified from plasmid pX-2 using primers 40/41. To integrate the *ESBP6* cassette into the strain IMX585, an equimolar amount of each DNA segment described above (total amount of 2 μg) and 200 ng of pQC059 were transformed into the target strain using the high-efficiency protocol described by (9). Isolated colonies from the transformation were tested by colony PCR using primers flanking the integration region (primers 40/41) and confirmed by sequencing.

The strains QL01_ESBP6 and QL38_ESBP6 were constructed by integrating *ESBP6* under the control of the *TEF1* promoter on site XI-5 of chromosome XI (10) in strains QL01 and QL38 (3) respectively. The integration cassette was obtained by amplifying the following DNA fragments: the region upstream of the site XI-5 together with the *ADH1* terminator were amplified from plasmid pIRP01 (12) with primers 42/43, the *TEF1* promotor together with the *ESBP6* coding sequence and *ADH1* terminator were amplified from the genomic DNA of strain ESBP6_OE with primers 44/45 and the *TEF1* promoter together with KlURA3 marker and the region downstream of the site XI-5 were amplified

from the plasmid pIRP01 (12) with the primers 46/47. Equimolar amounts of each purified fragment were mixed together for a total of 2 μg of DNA and transformed into QL01 or QL38 using the high-efficiency protocol described by (9). The transformation mix was plated on SD medium without uracil and a selection of clones was tested for the correct integration using primers 44/48. The integrated region was sequenced to confirm there were no errors. The strains QL01_URA and QL38_URA were constructed by integrating a cassette including only the KlURA3 marker on site XI-5 of chromosome XI in strains QL01 and QL38 (3) respectively. The integration cassette was obtained by amplifying a DNA fragment including the region upstream of the site XI-5 together with the KlURA3 marker and the region downstream of the site XI-5 with primers 42/46 from plasmid pXI-5 (10). 2 μg of the DNA cassette were transformed into QL01 or QL38 using the highefficiency protocol described by (9), the transformation mix was plated on SD medium without uracil and a selection of clones was tested for the correct integration using primers 44/48. The integrated region was sequenced to confirm there were no errors.

Adaptive laboratory evolution

Adaptive laboratory evolution experiments were conducted on an automated platform using a liquid handling robot as described previously (13, 14). Five independent experiments per aromatic acid were serially propagated (400 μL to 900 μL passage volume) in 17 mL (working volume) flasks of minimal medium supplemented with 10 g/L glucose, kept at 30°C through placement in a heat block and aerated by magnetic stirrers at 1200 rpm. Optical density (OD600nm) for each flask was monitored with a Tecan Sunrise plate reader as a proxy for cell density. Growth rates were calculated automatically using a best fit approximation for the growth using a three-phase model for lag, exponential growth, and stationary phase. Growth rates and final flask densities were recorded for each flask along with the quality of the fit for the growth model. The timing for increasing the concentration of each aromatic acid was decided according to the methodology described elsewhere (1).

Growth screening

The growth characterization of the post evolution mutants and reverse engineered strains was performed in a Growth profiler 960 (Enzyscreen) using 96-half-deep well microplates (with transparent bottom) with a total culture volume of 250 μl, agitation at 250 rpm, temperature controlled at 30 ºC and initial OD600 of 0.05. Three biological replicates (each with 2-3 technical replicates) for each strain were tested. Image scans were taken every 30 minutes and green pixels were converted to OD600 using a calibration curve created by filling ten wells with cell suspensions of known OD600 between zero and ten. The maximum growth rates for each strain were calculated by finding the highest slope in plots of the natural logarithm of the OD600 versus time.

gDNA extraction and sequencing

Chosen clones isolated from the evolved populations were cultivated overnight in 15 mL of YPD medium up to an OD600 close to four. The Blood & Cell Culture DNA Mini Kit (Qiagen) was used to extract genomic DNA from 3 mL of the overnight yeast culture (∼5 \times 10⁸ cells) using the protocol recommended by the manufacturer. Sequencing was performed with the NextSeq (High output kit) with 2x150 paired-end reads targeting a genome coverage of 100x per sample.

Sequencing data analysis

Mutations in evolved clones were identified using breseq 0.30.2 (15) with Bowtie 2.2.8 as aligner (16). Around nine million short reads for each strain (in fastq format) were aligned to the reference genome of CEN.PK113-7D (17) modified as follows: the GFP and KANMX

genes were added to chromosome XI, the annotations from the genome of *S. cerevisiae* S288c (R64-2-1) were transferred to the CEN.PK113-7D genome using RATT (18), the genome was polished by using breseq with the sequencing data for the strain GL01 and the identified mutations were patched to it. Each sample was analyzed by running breseq in consensus mode with the option "junction-alignment-pair-limit" set to 0 (no limit) to ensure all possible new junctions were evaluated. Reads from the reference strain (GL01) were also processed in the same conditions for the identification of false-positives. A custom-made script was used to find if the mutations detected in each mutant strain were false-positives by checking if there was any evidence for the same mutation occurring in the reference strain. To detect large-scale chromosome duplication/deletion, each chromosome coverage map generated by breseq was inspected visually.

Phenylacetic acid and phenylethanol quantification

Fusel acid and fusel alcohol production tests were performed in 100 mL shake-flasks containing 25 mL of the minimal medium described in the "Strains, media and DNA handling" section with 20g/L glucose or 10 g/L ethanol, the pH adjusted to 3.5 and 5 g/L of L-phenylalanine. Cultures were inoculated in triplicate with an initial OD600 of 0.05 and shaken at 200 rpm in an orbital incubator, at 30 °C for 96 hours. Extracellular concentration of phenylacetic acid and phenylethanol in cultures of a reference strain (IMX585) and a double deletion mutant (ΔPDR12 + ΔESBP6) by HPLC using a method previously described by Li et al. (3).

Coumaric acid production tests

Coumaric acid production tests were performed in 100 mL shake-flasks containing 20 mL of the minimal medium described in the "Strains, media and DNA handling" section without glucose and pH adjusted to 3.5. At the beginning of the culture six tablets of FeedBeads

(SMFB08001, Kuhner Shaker, Basel, Switzerland) were added and the culture inoculated with an initial OD600 of 0.05. Cultures were shaken at 200 rpm in an orbital incubator, at 30 ºC during 96 hours. To the medium used for strains QL38_URA and QL38_ESBP6 1% of galactose was also added to induce genes controlled by the GAL promoters. *p*-coumaric acid was quantified by mixing 1 mL of culture with 1 mL of absolute ethanol, thorough vortexing, centrifuging at 14000 x g and analyzing the supernatant using HPLC as described by Li et al. (3).

g/L of indoleacetic acid for a reference strain (IMX585) in comparison to a *PDR12* deletion mutant (ΔPDR12) and a double mutant lacking *PDR12* and displaying the overexpression of *ESBP6* (ΔPDR12 + ESBP6_OE). Cells were grown in 96-well plates with a volume of 250 μL in medium with 0.1 g/L of indoleacetic acid. $B - G$ rowth curves in medium with 0.5 g/L of phenylacetic acid and 0.5 g/L of L-tyrosine or 5 g/L of L-tryptophan as nitrogen source for a reference strain (IMX585) in comparison to a PDR12 deletion mutant (ΔPDR12). Cells were grown in 96-well plates with a volume of 250 μL in medium with 5 g/L of L-tryptophan or 0.5 g/L of L-tyrosine as nitrogen source and 0.5 g/L of phenylacetic acid. C – Growth curves in medium with 0.1 g/L of indoleacetic acid and 5 g/L of L-tryptophan as nitrogen source for a reference strain (IMX585) in comparison to a PDR12 deletion mutant (ΔPDR12) and a double deletion mutant lacking PDR12 and ESBP6 (ΔPDR12 + ΔESBP6). Cells were grown in 96-well plates with a volume of 250 μL in medium with 5 g/L of L-tryptophan as nitrogen source and 0.1 g/L of indoleacetic acid. All growth curves were performed for three biological replicates (each with two technical replicates) and the average OD600s are shown as solid lines with the interval encompassing the standard deviation represented as a shaded area.

Table S1. Growth-rates and maximum optical densities for a wild-type strain and individual mutants evolved in coumaric acid

Table S2. Growth-rates and maximum optical densities for a wild-type strain and individual mutants evolved in ferulic acid

Table S3. Mutations present in individual mutants evolved in coumaric acid

Table S5. Comparison between genes affected by increased read coverage in chromosome XIV and Aro80p targets

Table S7. List of primers used in this work

Table S8. List of plasmids used in this work

SI References

- 1. R. Pereira, *et al.*, Adaptive laboratory evolution of tolerance to dicarboxylic acids in Saccharomyces cerevisiae. *Metab. Eng.* 56, 130–141 (2019).
- 2. R. Mans, *et al.*, CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Res.* 15 (2015).
- 3. Q. Liu, *et al.*, Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. *Nat. Commun.* 10 (2019).
- 4. H. Inoue, H. Nojima, H. Okayama, High efficiency transformation of Escherichia coli with plasmids. *Gene* 96, 23–28 (1990).
- 5. J. Sambrook, *Molecular cloning : a laboratory manual* (Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001).
- 6. D. Solis-Escalante, *et al.*, *amdSYM* , a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 13, 126–139 (2013).
- 7. C. Verduyn, E. Postma, W. A. Scheffers, J. P. Van Dijken, Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8, 501–517 (1992).
- 8. J. Quan, J. Tian, Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLoS One* 4, e6441 (2009).
- 9. R. D. Gietz, R. A. Woods, "Yeast Transformation by the LiAc/SS Carrier DNA/PEG Method" in *Yeast Protocols*, (Humana Press, 2006), pp. 107–120.
- 10. M. D. Mikkelsen, *et al.*, Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* 14, 104– 111 (2012).
- 11. Y. Chen, S. Partow, G. Scalcinati, V. Siewers, J. Nielsen, Enhancing the copy number of episomal plasmids in Saccharomyces cerevisiae for improved protein production. *FEMS Yeast Res.* 12, 598–607 (2012).
- 12. J. López, *et al.*, Production of β-ionone by combined expression of carotenogenic and plant CCD1 genes in Saccharomyces cerevisiae. *Microb. Cell Fact.* 14, 84 (2015).
- 13. E. T. Mohamed, *et al.*, Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution. *Microb. Cell Fact.* 16, 204 (2017).
- 14. T. E. Sandberg, *et al.*, Evolution of Escherichia coli to 42 °C and Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations. *Mol. Biol. Evol.* 31, 2647–2662 (2014).
- 15. D. E. Deatherage, J. E. Barrick, "Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using breseq" in (Humana Press, New York, NY, 2014), pp. 165–188.
- 16. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359 (2012).
- 17. P. Jenjaroenpun, *et al.*, Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of Saccharomyces cerevisiae CEN.PK113- 7D. *Nucleic Acids Res.* (2018) https:/doi.org/10.1093/nar/gky014 (January 16, 2018).
- 18. T. D. Otto, G. P. Dillon, W. S. Degrave, M. Berriman, RATT: Rapid Annotation Transfer Tool. *Nucleic Acids Res.* 39, e57–e57 (2011).