Supporting Information

Biosynthesis of monomers for plastics from renewable oils

Wenhua Lu, Jon E. Ness, Wenchun Xie, Xiaoyan Zhang, Jeremy Minshull and Richard

A. Gross

Materials and Methods

Deletion of genes from the *C. tropicalis* genome

DNA constructs for modification of the *C. tropicalis* genome were based on the SAT1-flipper¹ and had the general form shown in Supplementary Figure 1. The targeting sequences in the deletion constructs were regions ranging in size from 198 bp to 300 bp, identical to the genomic sequences flanking the DNA to be deleted from *C. tropicalis*. For some of the sequences these were based on sequences that had been reported previously:^{2,3} CYP52A17 (GenBank accession AY230504), CYP52A13 (AY230499), CYP52A18 (AY230505), CYP52A14 (AY230500), FAO1 (AY538780), FAO2A (AY538781), FAO2B (AY538782), and CYP52A12 (AY230498).

Each deletion construct was linearized with restriction enzymes such that the 5' and 3' gene targeting sequences were at each end of the linearized DNA. Desalted DNA (1--3- µg) was transformed into *C. tropicalis* prepared by the method of Reuß et al¹ by electroporation at 1.8 kV using a Bio-Rad *E. coli* Pulser. D-Sorbitol (1 mL of 1M solution) was added to the electoporated yeast, the cells were transferred to a 14 mL culture tube and 1 mL of 2 x YPD broth was added. Cells were then rolled on a Rollerdrum for 1 hour at 37°C before spreading 100 µL on 100 mm diameter plates containing YPD Agar + 200 µg/mL nourseothricin. Plates were incubated for 2-4 days at 30°C. Large colonies (8-16) were individually streaked onto a YPD Agar plate to purify. A single colony from each streak was patched to a YPD agar stock plate and incubated overnight at 30°C.

The chromosomal location of each integrant was tested by PCR amplification from genomic DNA. Each nourseothricin-resistant isolate was inoculated into 2 mL of YP

Broth and rolled overnight at 30°C on a Rollerdrum. Genomic DNA from a 0.5 mL sample of each culture was isolated using Zymo Research's YeaStar genomic DNA isolation kit according to the manufacturer's instructions, eluting the DNA in 120 µL of TE, pH 8.0. The resulting gDNA (2.5 µL) was used as template in a 50 µL PCR amplification reaction. As a control for each analysis, genomic DNA was prepared from the parental strain that was transformed with the targeting construct. Typically one oligonucleotide primer was complementary to sequences within the targeting construct (for example within the nourseothricin-resistance gene) and a second was complementary to the genomic sequence flanking the desired integration location. The size of amplicons was determined by running 5- 10 µL of the completed PCR reaction on a 1% Agarose-TBE gel, this was then used to determine which strain(s) possessed the desired genomic structure.

Strains carrying a genomic targeting construct to be excised were treated as described by according to the protocol described by Reuß et al.¹ Briefly, the strain was inoculated from a YPD agar stock plate into 2 mL YP (YPD without dextrose) broth + 2% maltose in a 14 mL culture tube. The culture tubes were rolled for ~48 hours at 30°C on a rollerdrum. Growth with maltose induced production of Flp recombinase in the host strain from the integrated targeting construct. The Flp recombinase then acted at Frt sites located near the ends of the targeting construct (between the targeting sequences) to excise the sequences between the Frt sites, including the genes encoding Flp recombinase and conferring nourseothricin resistance. The culture was then diluted in serial 10-fold dilutions from 10-fold to 10,000-fold. Aliquots (100 µL) of 100, 1,000 and 10,000-fold dilutions were spread onto YPD agar plates.

Putative excisants were identified by replica-plating colonies on the YPD agar plates from the dilution series (the most useful plates for this purpose were those with 50-500 colonies) to a YPD agar + 200 µg/mL nourseothricin plates and then to a YPD agar plate. Putative excisants were identified as colonies that grow on YPD agar, but not YPD agar + 200 µg/mL nourseothricin following overnight incubation at room temperature. Putative excisants were streaked for single colonies to a YPD agar plate and incubated overnight at 30°C. Putative excisants were tested for excision of the integrant by PCR amplification from genomic DNA, typically using a pair of primers complementary to sequences on either side of the region to be deleted.

Identification and deletion of *C. tropicalis* ADH genes

Two degenerate PCR primers based on the DNA sequences of the structural genes for ADH1A, ADH1B, ADH2A and ADH2B of *C. albicans* (from <http://www.candidagenome.org/>) were designed to PCR amplify a ~1000 bp fragment corresponding to homologous ADH genes from *C. tropicalis*. These were primers 5'ADH (ACTCAAAAAGCYGTTTCTTYGAWACCAAYGGTGG) and 3'ADH (GTCCAAKACRTATCTACYCAAGATTTTACCTTCTTC). The primers were used to amplify from genomic DNA isolated from strain H5343 (using the YeaStar™ Genomic DNA Kit from Zymo Research). The resulting amplicon was purified using the QIEXII Gel Purification Kit from Qiagen and cloned. Plasmids were prepared from 96 transformants using Macherey–Nagel's Nucleospin Robot 96 plasmid kit and sequenced on an ABI 3730 automated sequencer. For deletion of the first allele of each ADH gene, a targeting construct was created as described above (Supplementary Figure 1). For the second allele of each, the nested strategy (Supplementary Figure 2) was attempted.

To delete the second alleles of CYP52A12, ADH-A4 and ADH-B4, "nested" constructs were designed with targeting sequences identical to target gene sequences that would have been removed during deletion of the first allele (Supplementary Figure 2).

The nested strategy was unsuccessful in deleting the second alleles of FAO1, ADH-A10 and ADH-B11. For deletion of these genes, we identified the second allele by PCR amplification from a strain in which the known allele of the gene had been deleted. The resulting PCR product was sequenced. The sequences of the second alleles identified in this way were ~ 95% identical to the first allele sequenced. GU056288 is the second allele of the previously reported FAO1, we have designated it FAO1B. GU056284 is the second allele of ADH-A10, we have designated it ADH-A10B. We used the sequence of the 5' and 3' ends of the second alleles as targeting sequences to eliminate these genes from the *C. tropicalis* genome.

Expression of synthetic P450s from pXICL

Expression vector pXICL comprised the *C. tropicalis* ICL promoter,⁴ an NcoI site, a stuffer, a NotI site, the *C. tropicalis* ICL transcriptional terminator region, the TEF1 promoter from *S. cerevisiae*, the EM7 prokaryotic promoter, the *Sh ble* gene (which codes for ZeocinTM resistance) recoded for *C. tropicalis*, and the CYC1 transcriptional terminator from *S. cerevisiae*. The sequence of pXICL has been submitted to Genbank (GU046314). Vector pXICL can be propagated in *E. coli* in LB Lennox medium using ZeocinTM at 25 µg/mL. Genes for integration into the *C. tropicalis* genome under control of the ICL promoter were optimized for expression in *C. tropicalis* (which uses CUG to encode serine not leucine) and had internal NcoI, NotI and BsiWI sites removed in GeneDesigner.⁵ Prior to transformation into *C. tropicalis*, the vector or its derivatives

were linearized at a unique BsiWI site in the ICL promoter region. *C. tropicalis* transformants were selected on YPD agar with Zeocin™ at 200 µg/mL. Transformants were screened for integration at ICL by PCR amplification from genomic DNA. Positives were expressed by culturing as described below.

Reactions with purified alcohol dehydrogenase proteins

Activities of alcohol dehydrogenase from *Saccharomyces cerevisiae* (Cat #: A3263) (from Sigma [St. Louis, MO]) on 12-hydroxydodecanoic acid (Cat #: 198781) and 16-hydroxyhexadecanoic acid (Cat #: H2640) (from Sigma [St. Louis, MO]) were measured by a modification of a previously described procedure.⁶ To a total reaction volume of 6 mL 0.1 M Tris-HCl buffer (pH 9.0) were added 8 mg enzyme, 1 µmol of NAD and 500 µg 12-hydroxydodecanoic acid or 16-hydroxyhexadecanoic acid dissolved in 50 µL of acetone. The reaction was incubated for 1 h at 37 °C. Control reactions without addition of enzyme were performed in parallel. The reaction was terminated by addition of 5 mL 0.1 N HCl. Products were extracted into diethyl ether and formation of fatty aldehydes was determined by ¹H-NMR and liquid chromatography-mass spectrometry (LC-MS).

Shake-flask oxidations of methyl tetradecanoate and ω-hydroxyfatty acids by *C. tropicalis*

C. tropicalis strains were precultured in 30 mL YPD medium containing (per liter): yeast extract, 10 g; peptone, 20 g; glucose, 20 g and shaken at 250 rpm, 30 °C for 20 h in 500 mL flask. The cells were inoculated at 10 % (v/v) to 30 mL culture medium consisting of (per liter): peptone, 3 g; yeast extract, 6 g; yeast nitrogen base, 6.7 g; acetic acid, 3 g; K₂HPO₄, 7.2 g; KH₂PO₄ 9.3 g; glucose, 20 g in 500 mL flask. After 12 h of cultivation at 250 rpm and 30 °C, the biotransformation phase was begun by addition of

30 g/L methyl tetradecanoate. pH was adjusted to 7.5 by addition of 2 mol/L NaOH solution. During the biotransformation, glycerol (50%) was fed (2.5 % per day) as co-substrates and the pH was maintained at ~7.5 by addition of 2 mol/L NaOH solution. For oxidation of ω -hydroxyfatty acids, the preculture of selected *C. tropicalis* strains was diluted in fresh culture medium to $A_{600\text{nm}} = 1.0$, and then continued to cultivate until the $A_{600\text{nm}}$ reached to ~5.0. Biotransformation was initiated by adding 5 mL culture to a 125 mL flask together with 50 mg of ω -hydroxyfatty acid and the pH was adjusted to ~7.5 by addition of 2 mol/L NaOH solution. Samples were taken and acidified to pH ~ 1.0 by addition of 6 N HCl. Products were extracted into diethyl ether and the concentrations of products were measured by LC-MS.

Fermentation procedure

Selected *C. tropicalis* strains were precultured in 50 mL YPD medium and shaken at 250 rpm, 30 °C for 20 h in 500 mL flask. The cells were inoculated at 10 % (v/v) into 150 mL of conversion medium containing (per liter): glucose, 30 g; $(\text{NH}_4)_2\text{SO}_4$, 7 g; KH_2PO_4 , 5.1 g; MgSO_4 , 0.5 g; CaCl_2 , 0.1 g; citric acid, 0.06 g; biotin, 0.0002 g and 1 mL/L of a trace elements solution (H_3BO_3 , 0.9; CuSO_4 , 0.07; KI, 0.18 g; FeCl_3 , 0.36; MnSO_4 , 0.72; Na_2MoO_4 , 0.36 g; ZnSO_4 , 0.72 g) in 400 mL DASGIP parallel magnetic-stirrer fermentors (Fig 4B), or in 700 mL of conversion medium in 1.5 L DASGIP parallel overhead driver fermentors (Fig 4C). The cultures were grown at 30 °C at an aeration rate of 1.0 vvm up to 16 hrs as growth phase. The pH was maintained at 6.0 by automatic addition of 6 mol/L NaOH or 2 mol/L H_2SO_4 solutions. Dissolved oxygen was maintained at 70% saturation by agitation and O_2 -cascade control mode. After 6 h growth, ethanol was fed into the cell culture to 5 g/L. During the biotransformation

phase, 80% glycerol was fed as co-substrate by dissolved oxygen-stat control mode (the high limit of dissolved oxygen was 75% and low limit of dissolved oxygen was 70%). Saturated fatty acids (C14:0 to C18:0) are not suitable substrates for whole-cell biotransformation due to their high melting points (m.p. from 52 to 72 °C). Instead, their methyl esters (m.p. from 28 to 35 °C) were used as substrates during biotransformation. Temperature was maintained at 35°C for conversion of methyl octadecanoate (C18:0) and 30 °C for the rest of fatty acid substrates (methyl tetradecanoate, methyl hexadecanoate, *cis*-9-octadecenoic acid and *cis,cis*-9,12-octadecadienoic acid). Every 12 h, ethanol was added into cell culture to 2 g/L, and substrate was added to 20 g/L or 40 g/L until its total concentration added reached to the predetermined value. To measure the substrate and product concentrations, 2 mL samples were acidified to pH ~ 1.0 by addition of 6 N HCl, extracted into diethyl ether and analyzed by LC-MS. The product concentration (g/L), conversion efficiency (mol-% conversion of substrate to ω -hydroxyfatty acid or α,ω -dicarboxylic acid) and productivity (g/L/h) data were normalized to the initial culture volume by correcting volume change associated with addition of acid, base, glycerol and substrate to the fermentor, to allow an accurate comparison between fermentations.

Preparation of ω -hydroxyfatty acids and α,ω -dicarboxylic acids.

Commercially unavailable ω -hydroxyfatty acids and α,ω -dicarboxylic acids used in this study for over-oxidation testing as well as reference standards for LC-MS analysis were prepared from biotransformations of fatty acids by selected *C. tropicalis* strains. The crude product extracted from fermentation broth was separated by a Biotage SP1™ Flash Chromatography system (Biotage Inc., A Dynax Corp. Company, USA) using a Biotage SI 40+M column (150 mm × 40 mm ID; 100 g KP-Sil™ silica; 40–63 μ m

particle size; flow rate: 25-50 mL/min) and eluted with n-hexane and diethyl ether mixtures by running a gradient from 20% to 100% of diethyl ether in 10-15 column volumes (CV) at flow rate 30 mL/min. Fractions were monitored at 220 nm and 280 nm by UV detector. The structures of purified ω -hydroxyfatty acids and α,ω -dicarboxylic acids (>99% pure) were characterized by gas chromatography-mass spectrometry (GC-MS) and NMR.

Analytical methods:

Gas Chromatography-Mass Spectrometry: GC-MS analysis was performed at 70 eV using a ThermoFinnigan TraceGC Ultra gas chromatograph coupled with a Trace DSQ mass spectrometer. Purified products were esterified with BF_3 in methanol (10%, w/w) at 70 °C for 20 min. Silylation was performed with HMDS/TMCS/Pyridine at 70 °C for 10 min. GC-MS experiments were performed with injector, ion source and interface temperatures of 200, 250 and 280 °C, respectively. Samples in hexane (1 μL) were injected in PTV split mode and run on a capillary column (Varian CP8944 VF-5MS, 0.25 mm x 0.25 μm x 30 m). The oven temperature was programmed at 120 °C for 1 min increasing to 260 °C at 20 °C/min, and then to 280 °C at 4.0 °C/min.

NMR: Proton (^1H) and carbon (^{13}C) NMR spectra were recorded on a Bruker DPX300 NMR spectrometer. The chemical shifts (ppm) reported were referenced relative to internal tetramethylsilane (TMS, 0.00 ppm) or to the solvent resonance at the appropriate frequency.

Liquid Chromatography-Mass Spectrometry: The concentrations of fatty acids, ω -hydroxyfatty acids and α,ω -dicarboxylic acids were measured by LC-MS with authentic reference compounds (obtained from Sigma [St. Louis, MO]) or purified products as

standards (Purification procedure is available in Supporting Information). Suberic acid was used as internal standard. The solvent delivery system was a Waters Alliance 2795 Separation Module (Milford, MA, USA) coupled with a Waters 2996 photodiode array detector and Waters ZQ detector with an electron spray ionization mode. The separation was carried on a reversed-phase column with dimensions of 150 x 4.6 mm and particle size 5 μ m. The mobile phase used for separation contained 10% H₂O, 5% acetonitrile, 5% formic acid solution (1% in water) and 80% methanol.

Candida tropicalis strains used

Strain Name	Genotype
H5343 = DP1	ura3A/ura3B pox5A::ura3A/pox5B::ura3A pox4A::ura3A/pox4B::URA3A
DP174	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14
DP283 DP284	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b
DP390	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b Δadh-a4/Δadh-a4b
DP415	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b Δadh-a4/Δadh-a4b Δadh- b4/Δadh-b4b adh-a10::SAT1
DP416	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b Δadh-a4/Δadh-a4b Δadh- b4/Δadh-b4b Δadh-a10
DP417	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b Δadh-a4/Δadh-a4b Δadh- b4/Δadh-b4b Δadh-a10 adh-b11::SAT1
DP421	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b Δfao2a/Δfao2b Δcyp52a12/Δcyp52a12b Δadh-a4/Δadh-a4b Δadh- b4/Δadh-b4b Δadh-a10 Δadh-b11
DP423	DP1 Δcyp52a17/Δcyp52a18 Δcyp52a13/Δcyp52a14 Δfao1/Δfao1b

DP424	Δ fao2a/ Δ fao2b Δ cyp52a12/ Δ cyp52a12b Δ adh-a4/ Δ adh-a4b Δ adh-b4/ Δ adh-b4b Δ adh-a10/adh-a10b::SAT1 Δ adh-b11
DP427 DP428	DP1 Δ cyp52a17/ Δ cyp52a18 Δ cyp52a13/ Δ cyp52a14 Δ fao1/ Δ fao1b Δ fao2a/ Δ fao2b Δ cyp52a12/ Δ cyp52a12b Δ adh-a4/ Δ adh-a4b Δ adh-b4/ Δ adh-b4b Δ adh-a10 Δ adh-b11 pXICL::CYP52A17
DP431	DP1 Δ cyp52a17/ Δ cyp52a18 Δ cyp52a13/ Δ cyp52a14 Δ fao1/ Δ fao1b Δ fao2a/ Δ fao2b Δ cyp52a12/ Δ cyp52a12b Δ adh-a4/ Δ adh-a4b Δ adh-b4/ Δ adh-b4b Δ adh-a10/ Δ adh-a10b Δ adh-b11
DP522 DP523	DP1 Δ cyp52a17/ Δ cyp52a18 Δ cyp52a13/ Δ cyp52a14 Δ fao1/ Δ fao1b Δ fao2a/ Δ fao2b Δ cyp52a12/ Δ cyp52a12b Δ adh-a4/ Δ adh-a4b Δ adh-b4/ Δ adh-b4b Δ adh-a10 Δ adh-b11 pXICL::CYP52A13
DP526 DP527	DP1 Δ cyp52a17/ Δ cyp52a18 Δ cyp52a13/ Δ cyp52a14 Δ fao1/ Δ fao1b Δ fao2a/ Δ fao2b Δ cyp52a12/ Δ cyp52a12b Δ adh-a4/ Δ adh-a4b Δ adh-b4/ Δ adh-b4b Δ adh-a10 Δ adh-b11 pXICL::CYP52A12

***Candida tropicalis* FAO1B (GU056288)**

CTTATGTTATTATGTGACGGGATCATCCACGAAACCACCGTCGACCAAATCAAAGACGTTATTGCTCCTGA
CTTCCCTGCTGACAAGTACGAAGAGTACGTCAGGACATTACCAAACCCTCCGAAACCCAGGGTTCAGGG
AAACCGTCTACAACACAGTCAACGCAAACACCACGGACGCAATCCACCAGTTCATTATCTTGACCAATGTT
TTGGCATCCAGGGTCTTGGCTCCAGCTTTGACCAACTCGTTGACGCCTATCAAGGACATGAGCTTGGAAGA
CCGTGAAAAATTGTTGGCCTCGTGGCGCGACTCCCCAATCGCTGCCAAAAGGAAATTGTTTCAGGTTGGTTT
CCACGCTTACCTTGGTTACTTTACGAGATTGGCCAATGAGTTGCATTTGAAAGCCATTCACTATCCAGGA
AGAGAAGACCGTGAAAAGGCTTATGAAACCCAGGAGATTGACCCTTTCAAGTACCAGTTTATGGAAAAGCC
AAAGTTTGACGGCGCTGAGTTGTACTTGCCAGATATTGATGTTATCATTATTGGATCTGGTGCCGGTGCTG
GTGTTGTGGCCACACTTTGGCCAACGATGGCTTCAAGAGTTTGGTTTTGGAAAAGGGCAAATACTTTAGC
AACTCCGAGTTGAACTTTGATGACAAGGACGGCGTTCAAGAATTATACCAAAGTGGAGGTACTTTGACTAC
AGTCAACCAACAGTTGTTTGTCTTGCTGGTTCCACTTTTGGTGGCGGTACCACTGTCAATTGGTCAGCCT
GTCTTAAGACGCCATTCAAGGTGCGTAAGGAATGGTATGATGAGTTTGGTGTGACTTTGCTGCTGATGAA
GCATACGATAAAGCGCAGGATTATGTTTGGCAGCAAATGGGAGCTTCTACCGAAGGCATCACCCACTCTTT
GGCTAACGAGATTATTATTGAAGGTGGTAAGAAATTAGGTTACAAGGCCAAGGTATTAGACCAAAACAGCG
GTGGTCATCCTCAGCACAGATGCGGTTTCTGTTATTTGGGCTGTAAGCACGGTATCAAGCAGGGTTCTGTT
AATAACTGGTTTAGAGACGCAGCTGCCCACGGTTCCCAGTTCATGCAACAGGTTAGAGTTTTGCAAATACT
TAACAAGAAGGGGATCGCTTACGGTATCTTGTGTGAGGATGTTGTAACGGCGCCAAGTTCACCATTACTG
GCCCCAAAAGTTTGTTGTTGCTGCCGGTGCTTTGAACACTCCATCTGTGTTGGTCAACTCCGGCTTCAAG
AACAAGAACATCGGTAAGAACTTAACTTTGCACCCAGTTTCTGTGCTGTTTGGTGATTTTGGCAAAGACGT
TCAAGCAGACCACTTCCACAACCTCCATCATGACTGCCCTTTGTTTCAAGCCGCTGATTTAGACGGCAAGG
GCCATGGATGCAGAATTGAAACCATCTTGAACGCTCCATTATCCAGGCTTCATTCTTACCATGGAGAGGT
AGTAACGAGGCTAGACGAGACTTGTTGCGTTACAACAACATGGTGGCGATGTTGCTCCTTAGTCGTGACAC
CACCAGTGGTTCCGTTTCTGCTCATCCAACCAAACCTGAAGCTTTGGTTGTCGAGTACGACGTGAACAAGT
TTGACAGAACTCGATCTTGCAGGCATTGTTGGTCACTGCTGACTTGTTGTATATCCAAGGTGCCAAGAGA
ATCCTTAGTCCACAGGCATGGGTGCCAATTTTTGAATCCGACAAGCCAAAGGATAAGAGATCAATCAAGGA
CGAGGACTATGTCGAATGGAGAGCCAAGGTTGCCAAGATTCTTTTCGACACCTACGGCTCACCTTATGGTT
CGGCACATCAAATGTCTTCTTGCCGTATGTCAGGTAAGGGTCCTAAATACGGTGCTGTTGACACCGATGGT

AGATTGTTTGAATGTTTGAATGTTTATGTTGCCGATGCAAGTCTTTTGCCAACTGCAAGCGGTGCTAATCC
TATG

***Candida tropicalis* ADH-A4 (GU056282)**

GAATTAGAATACAAAGATATCCCAGTGCCAACCCCAAAGGCCAACGAATTGCTCATCAACGTCAAATACTC
GGGTGTCTGCCACACTGATTTGCACGCCTGGAAGGGTGACTGGCCATTGGCCACCAAGTTGCCATTGGTTG
GTGGTCACGAAGGTGCTGGTGTCGTTGTCGGCATGGGTGAAAACGTCAAGGGCTGGAAGATTGGTGACTTC
GCCGGTATCAAATGGTTGAACGGTTCCTGTATGTCCTGTGAGTTCTGTCAACAAGGTGCTGAACCAAACCTG
TGGTGAGGCCGACTTGTCTGGTTACACCCACGATGGTTCTTTTGAACAATACGCCACTGCTGATGCTGTTT
AAGCCGCCAGAATCCCAGCTGGTACTGATTTGGCCGAAGTTGCCCAATCTTGTGTGCGGGTGTACCGTTC
TACAAAGCCTTGAAGACTGCCGACTTGGCCGCTGGTCAATGGGTGCTATCTCCGGTGCTGGTGGTGGTTT
GGGTTCTTGGCTGTCCAATACGCCGTGCGCATGGGCTTGAGAGTCGTTGCCATTGACGGTGGTGACGAAA
AGGGTGCCCTTTGTCAAGTCCTTGGGTGCTGAAGCCTACATTGATTTCTCAAGGAAAAGGACATTGTCTCT
GCTGTCAAGAAGGCCACCGATGGAGGTCCACACGGTGCTATCAATGTTTCCGTTTCCGAAAAAGCCATTGA
CCAATCCGTCGAGTACGTTAGACCATTGGGTAAGGTTGTTTTGGTTGGTTTGCCAGCTGGCTCCAAGGTCA
CTGCTGGTGTTTTTGAAGCCGTTGTCAAGTCCATTGAAATCAAGGGTTCCTATGTGCGGTAACAGAAAGGAT
ACCGCCGAAGCCGTTGACTTTTTCTCCAGAGGCTTGATCAAGTGTCCAATCAAGATTGTTGGCTTGAGTGA
ATTGCCACAGGTCTTCAAGTTGATG

***Candida tropicalis* ADH-A10 (GU056283)**

AAGTTAGAATACAAAGACGTGCCGGTCCCTGTCCCTAAACCCAACGAATTGCTTGTCAACGTCAAGTACTC
GGGTGTGTGTCATTCTGACTTGTCATGTCTGGAAAGGCGACTGGCCCATTCCTGCCAAGTTGCCCTTGGTGG
GAGGTCACGAAGGTGCTGGTGTCGTTGTCGGCATGGGTGACAACGTCAAGGGCTGGAAGGTGGGGGACTTG
GCTGGTATCAAGTGGTTGAATGGTTCGTGTATGAACTGTGAGTTTTGCCAACAGGGCGCAGAACCTAACTG
TTCAAGAGCCGACATGTCTGGGTATACCCACGATGGAACTTTCCAACAATACGCCACTGCTGATGCTGTCC
AAGCTGCCAAGATCCCAGAAGGCGCCGACATGGCTAGTATCGCCCCGATCTTGTGCGCTGGTGTGACCGTG
TACAAGGCTTTGAAGAACGCCGACTTGTTGGCTGGCCAATGGGTGGCTATCTCTGGTGCTGGTGGTGGTTT
GGGCTCCTTGGGTGTGCAGTACGCTAAAGCCATGGGTTACAGAGTGTTGGCTATCGACGGTGGTGACGAGA

GAGGAGAGTTTGTCAAGTCCTTGGGCGCCGAAGTGACATTGACTTCCTTAAGGAACAGGACATCGTTAGT
GCTATCAGAAAGGCAACTGGTGGTGGTCCACACGGTGTTATTAACGTGTCAGTGTCCGAAAAGGCAATCAA
CCAGTCGGTGGAGTACGTGAGAACTTTGGGGAAAGTGGTTTTAGTTAGCTTGCCGGCAGGTGGTAAACTCA
CTGCTCCTCTTTTCGAGTCTGTTGCTAGATCAATCCAGATTAGAACTACGTGTGTTGGCAACAGAAAGGAT
ACTACTGAAGCTATTGATTTCTTTGTTAGAGGGTTGATCGATTGCCCAATTAAAGTCGCTGGTTTTAAGTGA
AGTGCCAGAGATTTTTGACTTGATG

***Candida tropicalis* ADH-B2 (GU056285)**

CCATTGCAATACATCGATATTCCAGTCCCAGTCCCTAAGCCAAACGATTTGCTCGTCAATGTCAAATACTC
CGGTCTTTGTCACTCAGATATACACCTCTGGAAGGGTGACTGATTCCCAGCATCAAAATTGCCAGTTGTTG
GTGGTCACGAAGGTGCCAGTGTTGTCGTTGCTATTGGTGAAAACGTCCAGGGCTGGAAAGTAGGTGCCTTG
GCGGGCATAAAGATGTTGAATGGTTCCTGTATGAACTGTGAATTCTGTCAACAAAGTGCTTAACCAAGCTG
TCCCCATGCTGATGTCTCGGGTTACTCCCACGACGGCACTTTCCAACAGTACGCTACCGCTGATGCTGCTC
AAGCTGCTAAATTCCCAGCTGGTTCTGATTTAGCTAGCATCGCACCTATATCCTGTGCCGGTGTTACTGTT
TACAAAGCATTGAAGACTGCTGGCTTGATCCGGGCCAATGGGTTGCCATCTCCGATGCTGGTGGTGGTTT
GGGTTCTTTGGCCGTGCAATACGCCAAGGCCATGGGCTACAGAGTGGTGGCCATTGACTGCGGCGGCGAAA
ATGGAGTGTTTGTGAGATCGTTGGGTACTGAAGCTTTGTTGATTCCACCAAGGAGGCCAATGTCTCTGAG
GCTATCATCAAGGCTACCGACGGTGGTGTCCATGGTGTGTCATCAACGTTTCCATTTCTGAAAAGCCATCAA
CCAGTCTGTTGAAAATGTCAGAACTTTGGGTACTGTTGTCTTGGTTGGTTTGCCAGCTGGTGCCAAGCTCG
AAGCACCTATCTTCAATGCCGTTGCCAAATCCATCTAAATCAAGGATTCTTACGTGGGTAACCGAAGAGAC
ACTGCTGAGGCTGTTGATTTCTTCGCGAAAGGTTTGGTCAAGTGTCCAATTAAGGTTGTTGAGTTGAGTGA
ATTGCCAGAGATTTTCAAATTGTTG

***Candida tropicalis* ADH-B4 (GU056286)**

AAATTAGAATACAAGGACATCCCAGTTCCAAAGCCAAAGCCAAACGAATTGCTCATCAACGTCAAGTACTC
CGGTGTCTGCCACACTGATTTACACGCCTGGAAGGGTGACTGGCCATTGGACACCAAGTTGCCATTGGTGG
GTGGTCACGAAGGTGCTGGTGTGTTGTTGGCATTGGTGACAATGTCAAGGGATGGAAGGTCGGTGATTTG
GCCGGTGTCAAGTGGTTGAACGGTTCCTGTATGAACTGTGAGTACTGTCAACAGGGTGCCGAACCAAAGT

TCCACAGGCTGACTTGTCTGGTTACACCCACGACGGTTCTTTCCAGCAATACGCCACTGCAGATGCCGTGC
AAGCCGCTAGAATTCCAGCTGGTACTGATTTAGCCAACGTTGCCCCATCTTGTGTGCTGGTGTCACTGTT
TACAAGGCCTTGAAGACCGCCGACTTGCAGCCAGGTCAATGGGTCGCCATTTCCGGTGCCGCTGGTGGTTT
GGGTTCTTTGGCCGTTCAATACGCCAAGGCCATGGGCTACAGAGTTGTCGCCATCGATGGTGGTGCCGACA
AGGGTGAGTTCGTCAAGTCTTTGGGCGCTGAGGTCTTTGTTGATTTCTCAAGGAAAAGGACATTGTTGGT
GCTGTCAAGAAGGCAACCGATGGTGGCCACACGGTGCCGTTAACGTTTCCATCTCCGAAAAGGCCATCAA
CCAATCTGTGCGACTACGTTAGAACCTTGGGTAAGGTTGTCTTGGTCGGTTTGCCAGCTGGCTCCAAGGTTT
CTGCTCCAGTCTTTGACTCCGTCGTCAAGTCCATCCAAATCAAGGGTTCCTATGTCGGTAACAGAAAGGAC
ACTGCCGAAGCTGTTGACTTTTTCTCCAGAGGCTTGATCAAGTGTCCAATCAAGGTTGTCGGTTTGAGTGA
ATTGCCAGAAGTCTACAAGTTGATG

***Candida tropicalis* ADH-B11 (GU056287)**

CCATTGCAATACACCGATATCCCAGTTCAGTCCCTAAGCCAAACGAATTGCTCGTCCACGTCAAATACTC
CGGTGTTTGTCACTCAGATATACACGTCTGGAAGGGTGACTGGTTCCCAGCATCGAAATTGCCCGTTGTTG
GTGGTCACGAAGGTGCCGGTGTGTCGTTGCCATTGGTGAAAACGTCCAAGGCTGGAAAGTAGGTGACTTG
GCAGGTATAAAGATGTTGAATGGTTCCTGTATGAACTGTGAATACTGTCAACAAGGTGCTGAACCAAACCTG
TCCCCACGCTGATGTCTCGGGTTACTCCCACGACGGTACTTTCCAACAGTACGCTACCGCCGATGCTGTTT
AAGCTGCTAAATTCCCAGCTGGTTCTGATTTAGCTAGCATCGCACCTATATCCTGCGCCGGTGTTACTGTT
TACAAAGCATTGAAAACGTCAGGCTTGCAGCCAGGTCAATGGGTTGCCATCTCTGGTGCAGCTGGTGGTTT
GGGTTCTTTGGCTGTGCAATACGCCAAGGCCATGGGTTTGAGAGTCGTGGCCATTGACGGTGGTGACGAAA
GAGGAGTGTTTGTCAAATCGTTGGGTGCTGAAGTTTTCGTTGATTTACCAAAGAGGCCAATGTCTCTGAG
GCTATCATCAAGGCTACCGACGGTGGTGCCCATGGCGTCATCAACGTTTCCATTTCTGAAAAAGCCATCAA
CCAGTCTGTTGAATATGTTAGAACTTTGGGAACGTTGTCTTGGTTGGTTTGCCAGCTGGTGCAAAGCTCG
AAGCTCCTATCTTCAATGCCGTTGCCAAATCCATCCAAATCAAAGGTTCTTACGTGGGAAACAGAAGAGAC
ACTGCTGAGGCTGTTGATTTCTTCGCTAGAGGTTTGGTCAAATGTCCAATTAAGGTTGTTGGGTTGAGTGA
ATTGCCAGAGATTTTCAAATTGTTG

References

- (1) Reuß, O.; Vik, R.; Kolter, R.; Morschhäuser, J. *Gene* **2004**, *341*, 119-127.
- (2) Eirich, L. D.; Craft, D. L.; Steinberg, L.; Asif, A.; Eschenfeldt, W. H.; Stols, L.; Donnelly, M. I.; Wilson, C. R. *Appl. Environ. Microb.* **2004**, *70*, 4872-4879.
- (3) Craft, D. L.; Madduri, K. M.; Eshoo, M.; Wilson, C. R. *Appl. Environ. Microb.* **2003**, *69*, 5983-5991.
- (4) Umemura, K.; Atomi, H.; Kanai, T.; Teranishi, Y.; Ueda, M.; Tanaka, A. *Appl. Microbiol. Biot.* **1995**, *43*, 489-492.
- (5) Villalobos, A.; Ness, J. E.; Gustafsson, C.; Minshull, J.; Govindarajan, S. *BMC bioinformatics* **2006**, *7*, 285.
- (6) Bjorkhem, I. *Eur. J. Biochem.* **1972**, *30*, 441-451.

Supplementary Figures

Supplementary Figure 1 shows a schematic representation of a DNA "genomic targeting" construct for modifying the genome of yeasts¹⁷. The general structure is that the construct has two targeting sequences that are homologous to the sequences of two regions of the target yeast chromosome. Between these targeting sequences are two sites recognized by a site-specific recombinase (in this case *frt* sites recognized by the *flp* recombinase). Between the two site specific recombinase sites are sequence elements, one of which encodes a selective marker (in this case the gene encoding resistance to the Nourseothricin resistance marker from transposon Tn1825 and the other of which encodes the site-specific recombinase (in this case *flp*).

Supplementary Figure 2 shows the nested deletion approach. The first targeting construct eliminates all DNA between the targeting sequences (deleted DNA is shown with vertical hash marks). A second targeting construct uses targeting sequences (checkerboard fill) that are "nested" within the deleted region. This will remove a smaller region from within the second allele of a gene, since these regions have already been eliminated from the first allele.

Supplementary Figure 1

Supplementary Figure 2