

# Screening non-conventional yeasts for acid tolerance and engineering *Pichia occidentalis* for production of muconic acid

Received: 24 April 2023

Accepted: 22 August 2023

Published online: 31 August 2023

 Check for updates

Michael E. Pyne<sup>1,2,5,10</sup>, James A. Bagley<sup>1,6,10</sup>, Lauren Narcross<sup>1,2,7</sup>, Kaspar Kevvai<sup>1,2,8</sup>, Kealan Exley<sup>1,2,9</sup>, Meghan Davies<sup>1,2,3</sup>, Qingzhao Wang<sup>4</sup>, Malcolm Whiteway<sup>1,2</sup> & Vincent J. J. Martin<sup>1,2</sup> ✉

*Saccharomyces cerevisiae* is a workhorse of industrial biotechnology owing to the organism's prominence in alcohol fermentation and the suite of sophisticated genetic tools available to manipulate its metabolism. However, *S. cerevisiae* is not suited to overproduce many bulk bioproducts, as toxicity constrains production at high titers. Here, we employ a high-throughput assay to screen 108 publicly accessible yeast strains for tolerance to 20 g L<sup>-1</sup> adipic acid (AA), a nylon precursor. We identify 15 tolerant yeasts and select *Pichia occidentalis* for production of *cis,cis*-muconic acid (CCM), the precursor to AA. By developing a genome editing toolkit for *P. occidentalis*, we demonstrate fed-batch production of CCM with a maximum titer (38.8 g L<sup>-1</sup>), yield (0.134 g g<sup>-1</sup> glucose) and productivity (0.511 g L<sup>-1</sup> h<sup>-1</sup>) that surpasses all metrics achieved using *S. cerevisiae*. This work brings us closer to the industrial bio-production of AA and underscores the importance of host selection in bioprocessing.

Historically, brewer's yeast (*Saccharomyces cerevisiae*) has been the preferred microbial host for the overproduction of fuels, chemicals, and pharmaceuticals. However, owing to the low value of many bulk chemicals and biofuels, concentrations required for commercial scale production are often toxic to *S. cerevisiae* and other model organisms. For instance, *S. cerevisiae* is unable to sustain growth in low to moderate concentrations of butanol (2%)<sup>1</sup>, vanillin (0.05%)<sup>2</sup>, benzaldehyde (0.1–0.2%)<sup>3</sup>, and many organic acids (0.25–2.5%)<sup>4–6</sup>. Organic acids pose a distinct bioprocessing challenge, as an ideal organic acid bioprocess would be performed at acidic pH, which simplifies downstream recovery of the undissociated form and circumvents the requirement of maintaining pH during fermentation. However, acid toxicity is inversely proportional to pH, as undissociated acids (pH < pK<sub>a</sub>) freely diffuse into cells and dissociate in the cytosol (pH 5–7)<sup>7</sup>. Accordingly, to circumvent product toxicity, cultivation of most organic-acid-

producing strains of *S. cerevisiae* and other microbes is performed at pH > pK<sub>a</sub><sup>8–10</sup>.

Adipic acid is a C6 dicarboxylic acid used in the production of nylon 6,6. With an annual production of three million tonnes and a global market of six billion USD<sup>11</sup>, adipic acid has been classified as a 'super-commodity' by the U.S. Department of Energy<sup>12</sup>. Global demand for adipic acid is presently met via a chemical process involving cyclohexane and nitric acid, but this method utilizes petrochemical resources and liberates nitrous oxide, a potent greenhouse gas, prompting the investigation of more sustainable routes to its production. There are no natural routes for adipic acid biosynthesis but at least eight engineered biochemical pathways have been proposed for its bioproduction<sup>11</sup>. Of these pathways, major progress has been attained using the *cis,cis*-muconic acid (CCM) route. This strategy employs three heterologous enzymes, 3-dehydroshikimate dehydratase, protocatechuic acid

<sup>1</sup>Department of Biology, Concordia University, Montréal, QC H4B 1R6, Canada. <sup>2</sup>Centre for Applied Synthetic Biology, Concordia University, Montréal, QC H4B 1R6, Canada. <sup>3</sup>BenchSci, Toronto, ON, Canada. <sup>4</sup>bp Biosciences Center, San Diego, California, USA. <sup>5</sup>Present address: Department of Biology, University of Western Ontario, Ontario, Canada. <sup>6</sup>Present address: Centre for Applied Synthetic Biology, Concordia University, Montréal, QC H4B 1R6, Canada. <sup>7</sup>Present address: Amyris, Inc., Emeryville, CA, USA. <sup>8</sup>Present address: Pivot Bio, Berkeley, CA, USA. <sup>9</sup>Present address: Novo Nordisk Foundation Center for Biosustainability, Lyngby, Denmark. <sup>10</sup>These authors contributed equally: Michael E. Pyne, James A. Bagley. ✉e-mail: [vincent.martin@concordia.ca](mailto:vincent.martin@concordia.ca)

decarboxylase, and catechol dioxygenase, for synthesis of CCM from 3-dehydroshikimate, an intermediate of the shikimate pathway. CCM is converted to adipic acid via hydrogenation by enoate reductase enzymes, yet bacterial variants characterized to date function poorly in yeast hosts<sup>13</sup>, achieving <0.9% molar conversion of CCM to adipic acid<sup>14</sup>. In the absence of a suitable yeast-active enoate reductase, *S. cerevisiae* has been engineered to synthesize CCM at a titer, yield, and productivity of up to 22.5 g L<sup>-1</sup>, 0.1 g g<sup>-1</sup> glucose, and 0.21 g L<sup>-1</sup> h<sup>-1</sup>, respectively<sup>9,10</sup>. Despite these advancements, yeast CCM processes are performed at pH 5–6, which is well above the pK<sub>a</sub> of CCM (3.87) and maintains the acid in its dissociated form to limit product toxicity.

In response to the poor tolerance of *S. cerevisiae* to many bulk bioproducts, focus has shifted to screening non-conventional yeast species for phenotypes better suited for industrial applications<sup>15</sup>. Strain selection is a critical and often overlooked aspect of bioprocess development, as selecting a microbial host that is tailored to the target end product is likely to reduce strain engineering interventions, simplify downstream separation, limit product toxicity, and improve overall process metrics<sup>16</sup>. For instance, *Yarrowia lipolytica* is an oleaginous and osmotolerant host that is well-suited for the production of fatty acids and alkanes<sup>17</sup>, *Kluyveromyces marxianus* is a rapid-growing thermotolerant yeast<sup>18</sup>, and *Pichia kudriavzevii* is a robust acid tolerant host utilized for the production of diverse organic acids<sup>19</sup>. Approximately 1500 species of yeast have been classified across more than 100 genera and public culture repositories contain thousands of distinct yeast strains. The majority of these yeast species can be cultivated in the laboratory with standard media<sup>20</sup> and major advancements in genome sequencing, systems biology, and strain engineering have dramatically reduced the cost and labor involved in the genetic domestication of non-conventional hosts. Sophisticated genetic toolkits have been developed for diverse yeast species, including *Y. lipolytica*<sup>21</sup>, *K. marxianus*<sup>18</sup>, and *P. pastoris*<sup>22</sup>, providing new opportunities for host selection beyond *Saccharomyces*.

In this study, we screen 108 distinct strains of yeast from public culture collections for natural tolerance to high concentrations of adipic acid and other dicarboxylic acids. Our screen identifies several species of *Pichia* as prospective industrial hosts to produce diverse organic acids. Based on antibiotic susceptibility, plasmid transformation, and CRISPR-Cas9 editing assays, we select *P. occidentalis* for high level production of CCM, the direct precursor to adipic acid. Under fed-batch conditions, engineered *P. occidentalis* synthesizes 38.8 g L<sup>-1</sup> CCM with a maximum yield and productivity that surpasses all present metrics achieved using *S. cerevisiae*.

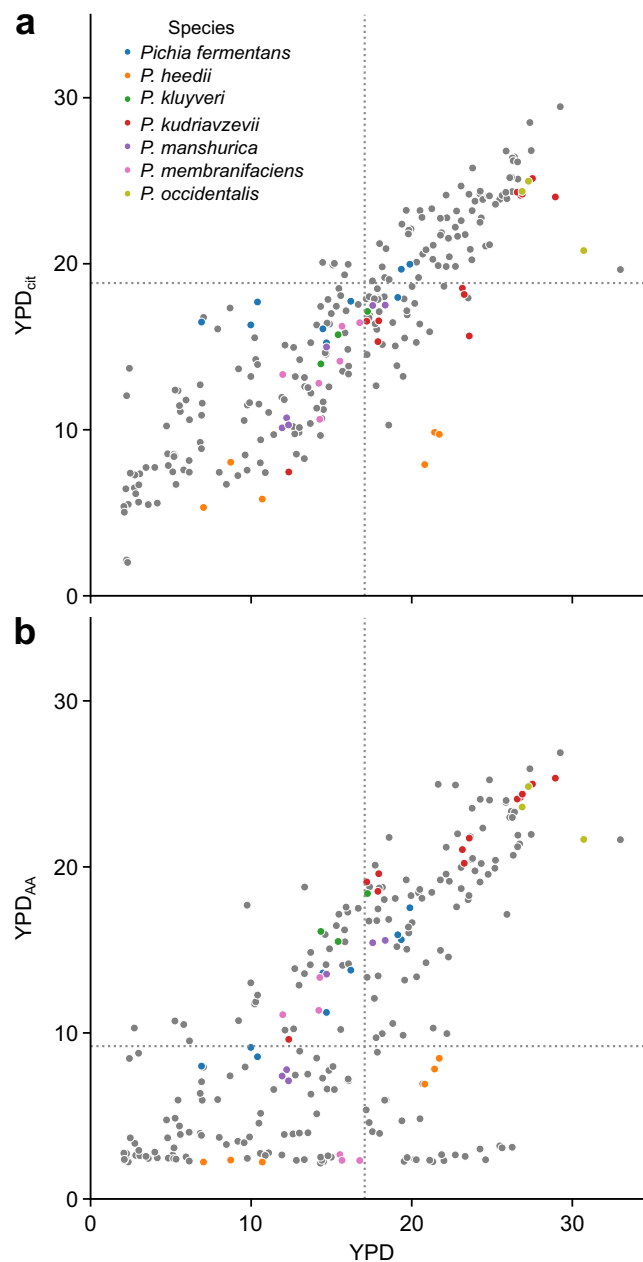
## Results

### Screening non-conventional yeast strains for organic acid tolerance

To build a library of non-conventional yeasts for screening prospective acid tolerant hosts, we surveyed public culture repositories for non-*Saccharomyces* strains, selecting strains with reported acid tolerance where available. In total, 153 strains from 83 distinct species were obtained, of which 124 were successfully cultured in YM medium at 30 °C and 108 were successfully screened in our assay, with the remainder being unable to grow reliably in YPD. The final 108 strains spanned 66 species from 28 genera (Supplementary Data 1).

Subsequent screening involved cultivating strains in YPD containing 20 g L<sup>-1</sup> citric acid (0.1 M; YPD<sub>cit</sub>; Fig. 1a) or YPD containing 20 g L<sup>-1</sup> adipic acid (0.137 M; YPD<sub>AA</sub>; Fig. 1b). Addition of adipic acid dropped the pH of YPD to -3.7 (pK<sub>a1</sub> = 4.41, pK<sub>a2</sub> = 5.41) and addition of citric acid dropped the pH of YPD to -3.0 (pK<sub>a1</sub> = 3.13, pK<sub>a2</sub> = 4.76, pK<sub>a3</sub> = 6.39). YPD<sub>AA</sub> medium yielded a wide diversity of growth phenotypes. Despite a higher pH, the correlation between growth in YPD and YPD<sub>AA</sub> ( $R^2 = 0.685$ ) was lower than that of growth in YPD and YPD<sub>cit</sub> ( $R^2 = 0.905$ ), indicating that adipic acid impacted the growth of more strains.

We employed two criteria for selecting strains for subsequent rounds of screening. First, we defined adipic acid tolerance as a growth rate in YPD<sub>AA</sub> (estimated as area under the curve at OD<sub>600</sub>) that is no more than 20% slower compared to growth in YPD. Second, to ensure robust growth of strains under standard laboratory conditions, we selected strains with growth rates in YPD (measured as area under strain growth curves at OD<sub>600</sub>) that were no more than 20% slower compared to the control strain (*S. cerevisiae* CEN.PK 113-7D). Despite the use of rich YPD medium, some strains were unable to grow to an OD<sub>600</sub> >1.0 in the allotted time in the absence of acid inhibitors.



**Fig. 1 | Growth of 83 candidate yeast strains in YPD supplemented with citric and adipic acids. a** Growth of yeast strains in YPD compared to YPD supplemented with 0.1 M citric acid (YPD<sub>cit</sub>). **b** Growth of yeast strains in YPD compared to YPD supplemented with 0.15 M adipic acid (YPD<sub>AA</sub>). Growth was assessed by measuring area under the curve (AUC) of growth curves. Each data point represents a single replicate of  $n = 3$  independent biological replicates for each yeast strain. Dashed lines show average AUCs for *S. cerevisiae* CEN.PK113-7D in each growth condition. Source data are provided as a Source Data file.

Applying our screening criteria resulted in the identification of 15 fast growing and adipic acid tolerant strains from the original set of 108 candidates (Table 1). Notably, strains from the genus *Pichia* were enriched among the adipic acid tolerant candidates; of the fifteen strains, seven were *Pichia* (47%), while *Pichia* represented only 10% of the original strain collection.

**Table 1 | Summary of 15 adipic acid tolerant yeast identified from a preliminary screen**

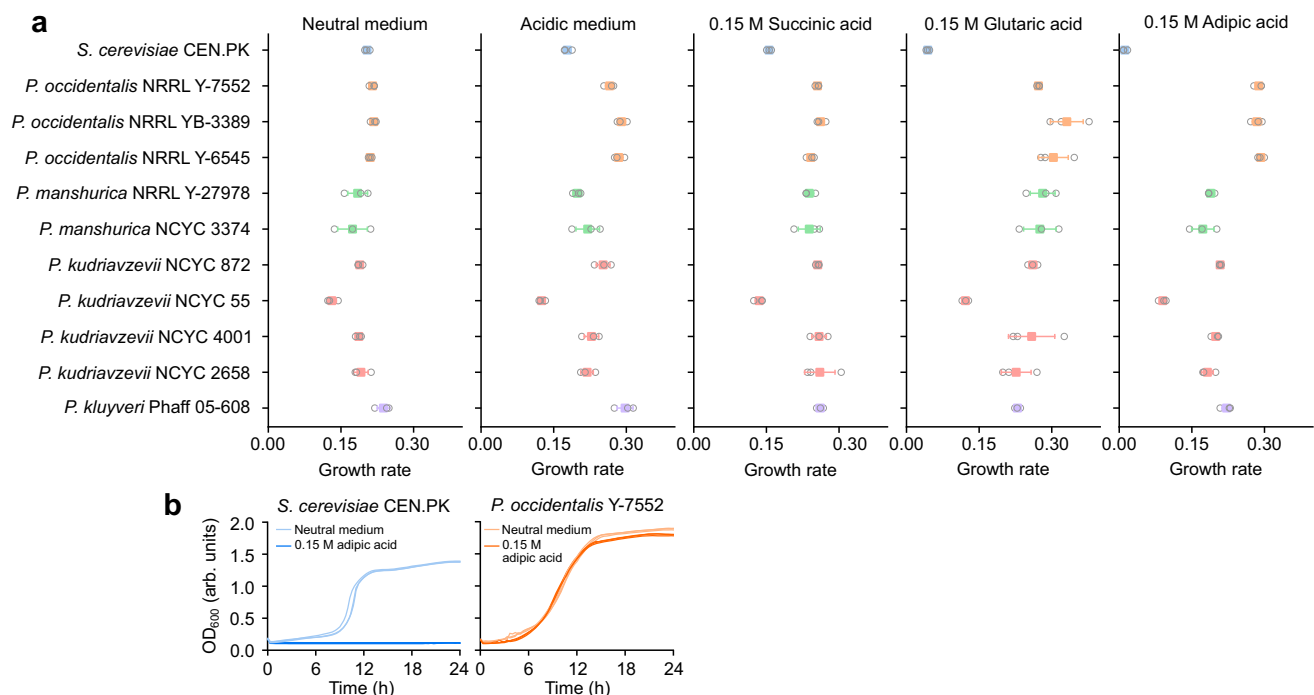
Strain ID	Depositor ID	Species	YPD AUC <sup>a</sup>	YPD <sub>AA</sub> AUC <sup>a</sup>
O0E	3092 CBS	<i>Candida boidinii</i>	21.6	19.7
O19	Y-7921 NRRL	<i>Candida sorbophila</i>	28.5	24.7
O1L	8880 CBS	<i>Cyberlindnera saturnus</i>	18.7	18.7
O1R	2530 NCYC	<i>Debaryomyces hansenii</i>	19.9	20.9
O1X	1478 NCYC	<i>Kazachstania exigua</i>	25.6	21.4
O2O	971 NCYC	<i>K. unispora</i>	17.8	17.4
O2G	562 NCYC	<i>Pichia fermentans</i>	20.4	16.6
O1W	4001 NCYC	<i>P. kudriavzevii</i>	30.5	26.5
O2M	2658 NCYC	<i>P. kudriavzevii</i>	26.6	22.2
O1V	55 NCYC	<i>P. kudriavzevii</i>	19.5	20.5
O2N	872 NCYC	<i>P. kudriavzevii</i>	18.2	19.9
O2Q	Y-27978 NRRL	<i>P. manshurica</i>	17.3	14.8
O2W	Y-7552	<i>P. occidentalis</i>	29.7	25.0
O3P	Y-27308 NRRL	<i>Tetrapisispora arboricola</i>	22.6	19.5
O3V	432 NCYC	<i>Wickerhamomyces anomalus</i>	27.1	26.1

<sup>a</sup>Relative growth rates were compared by measuring area under the curve (AUC).

The over-representation of *Pichia* strains prompted further probing of the *Pichia* clade for acid tolerant hosts. An additional eight *Pichia* strains were added, three of which were from species not represented in the initial screen (*Candida sorboxylosa*, *Pichia norvegensis* and *Pichia exigua*). Despite its name, *C. sorboxylosa* has been placed in the *Pichia* genus by whole genome phylogenetic analysis<sup>23</sup>. We also added two new strains of *P. occidentalis* (Y-6545 and YB-3389) based on the adipic acid tolerance of *P. occidentalis* Y-7552 in our initial screen.

Subsequent acid tolerance screening was performed in defined Yeast Nitrogen Base (YNB) medium to identify strains with robust growth in both defined and rich media. Intriguingly, certain *P. occidentalis* strains displayed a growth defect in YNB. Specifically, *P. occidentalis* Y-6545 and YB-3389 were unable to reach an OD<sub>600</sub> of 1.0 in the allotted time (24 h) in 1× YNB compared to an OD<sub>600</sub> of ~1.75 in YPD (Supplementary Fig. 1). An increase in the concentration of YNB from 1× (1.7 g L<sup>-1</sup>) to 3× (5.1 g L<sup>-1</sup>) alleviated this limitation. Since YNB is composed of a mixture of 20 vitamins and salts, the nature of this apparent nutrient limitation is not known. To ensure robust growth of all yeast strains, subsequent screening experiments were performed in 3× YNB.

To further characterize acid tolerant phenotypes, we performed a subsequent round of screening in defined medium to evaluate the ability of 10 candidate *Pichia* strains to grow on three different dicarboxylic acids of varying chain lengths, namely succinic (C4), glutaric (C5), and adipic acid (C6) (Fig. 2a). Acids were added to each defined medium at a concentration of 0.15 M and a pH of 2.8. For the dicarboxylic acids tested, the control *S. cerevisiae* CEN.PK 113-7D strain exhibited a steady decrease in growth rate with increasing acid chain length assuming a comparable level of dissociation between the dicarboxylic acids. Adipic acid, the longest dicarboxylic acid screened,



**Fig. 2 | Growth of 10 candidate *Pichia* strains in various dicarboxylic acids.** **a** Maximum growth rates of candidate strains in 3× YNB supplemented with succinic, glutaric, or adipic acids (0.15 M each). All dicarboxylic-acid-containing media were pH adjusted to 2.8 after the addition of 0.15 M of the respective dicarboxylic acid. Acidic medium is 3× YNB medium with the pH adjusted to 2.8 and neutral medium is 3× YNB medium without pH adjustment. Error bars represent the mean ±

s.d. of  $n = 3$  independent biological samples. **b** Representative growth curves of *S. cerevisiae* CEN.PK and *P. occidentalis* strain Y-7552 in 3× YNB supplemented with 0.15 M (22 g L<sup>-1</sup>) adipic acid. Growth was assessed by measuring OD<sub>600</sub> in arbitrary units (arb. units). Three independent biological replicates are overlaid for each growth condition. Source data are provided as a Source Data file.

completely inhibited growth of *S. cerevisiae* at a concentration of 0.15 M (21.9 g L<sup>-1</sup>) (Fig. 2b). This general trend was also observed with most of the *Pichia* strains, except for the three strains of *P. occidentalis*, where the growth rate generally increased with chain length ( $P < 0.001$ ) (Fig. 2a). Overall, one or multiple strains of *P. kluyveri*, *P. kudriavzevii*, *P. manshurica*, and *P. occidentalis* exhibited notable tolerance to multiple dicarboxylic acids.

We also screened 93 yeast strains in YPD medium and YPD containing 0.1 M citrate buffer, pH 3.0 (YPD<sub>cit</sub>). However, this condition was not sufficiently inhibitory, as most strains grew to comparable levels in YPD and YPD<sub>cit</sub> (Fig. 1a). It has been reported that citric acid is more toxic to yeast at pH 4.5 than pH 3.0 owing to the propensity of the dissociated form to bind divalent cations, providing a possible explanation for our observations.

In addition to dicarboxylic acids, we screened 26 *Pichia* strains on several monocarboxylic acids using the same conditions (defined 3× YNB medium containing 0.15 M of acid, pH 2.8). Only one strain grew in acetic acid (*P. manshurica* Y-27978) and none of the tested strains grew in propionic, butyric or valeric acids. Assuming comparable levels of dissociation between monocarboxylic acids, these findings indicate a higher degree of toxicity caused by propionic, butyric and valeric acids at equimolar concentrations.

### Additional traits used for selection

Based on our screen for tolerance to adipic acid, we selected eight candidate strains of *Pichia* as potential CCM production hosts. We first assessed carbon source utilization of the eight selected strains by testing the capacity to utilize xylose, an abundant C5 sugar in plant lignocellulose. Screening xylose utilization by patching colonies onto agar plates containing xylose as sole carbon source revealed that *P. fermentans* 562 NCYC was the only adipic acid tolerant *Pichia* strain capable of xylose utilization (Fig. 3a), which is consistent with yeast phenotypic characterization data<sup>20</sup> (Supplementary Table 1). In addition to carbon source utilization, strain suitability for genetic and metabolic engineering requires the ability to use common selectable markers and introduce foreign genetic material. Accordingly, the eight selected *Pichia* strains were screened for susceptibility to high concentrations of G418 (1000 μg mL<sup>-1</sup>), hygromycin (600 μg mL<sup>-1</sup>), nourseothricin (100 μg mL<sup>-1</sup>), and zeocin (300 μg mL<sup>-1</sup>) (Fig. 3a). Except for *P. membranifaciens* NCYC 55 and *P. manshurica* Y-27978, all strains were resistant to 300 μg mL<sup>-1</sup> zeocin. In contrast, nearly all strains were susceptible to nourseothricin and hygromycin at the concentrations tested. G418 was effective against roughly half of the strains assayed. Of the three strains of *P. occidentalis*, only strain Y-7552 was susceptible to G418, while nourseothricin and hygromycin were effective against all three strains.

Having identified effective concentrations of antibiotics, we screened the eight candidate *Pichia* strains for transformation using vectors expressing a hygromycin-resistance marker from the *Ashbya gossypii* *TEF1* promoter (*P*<sub>TEF1</sub>). We confirmed successful transformation of six of the eight strains of *Pichia* (Fig. 3a). The highest transformation efficiency was attained using *P. fermentans*, the least adipic acid tolerant of the eight candidate *Pichia* strains. Transformants were obtained for all three strains of *P. occidentalis*, the most adipic acid tolerant strains identified in this study. Transformation efficiency of *P. occidentalis* Y-7552 (10<sup>2</sup> CFU μg<sup>-1</sup> DNA) was higher than strains Y-6545 and YB-3389 (10<sup>1</sup> CFU μg<sup>-1</sup> DNA).

Based on antibiotic susceptibility and plasmid transformation assays, we selected *P. kluyveri* and all three strains of *P. occidentalis* for genome editing assays, which was aided by recently reported draft genome sequences for these strains<sup>24</sup>. We selected the *ADE2* gene for deletion since its inactivation leads to accumulation of a colored intermediate for visual identification of deletion mutants (Fig. 3b). We anticipated a low level of homologous recombination in *Pichia* strains<sup>25,26</sup>, prompting us to pursue a CRISPR system in which

homology-directed repair templates were precloned into pCas delivery vectors (Fig. 3c), thus circumventing the need to assemble multiple overlapping DNA fragments in *Pichia*.

We first assessed the capacity of an in-house CRISPR-Cas9 vector (pCas-Hyg-CEN6ARS4) possessing the *S. cerevisiae* CEN6/ARS4 origin, as well as gRNA and *cas9* expression cassettes (Fig. 3c), to function in *Pichia*. We designed *ade2Δ* homology repair templates for all four candidate *Pichia* strains by fusing -800 bp homology arms flanking the *ADE2* coding sequence and small portions (-100–150 bp) of the promoter and terminator elements (Fig. 3b). Using this design, homology-directed repair of an *ADE2* double-strand break results in a large deletion (-2 kb) of the entire *ADE2* coding sequence. Following assembly of *ade2Δ* pCas vectors in *S. cerevisiae*, vectors were linearized with BglII and used to transform the relevant *Pichia* strain along with an overlapping *ADE2* gRNA for assembly via gap repair in *Pichia* (Fig. 3d). Using this approach, we obtained Hyg<sup>R</sup> colonies for all four strains of *Pichia* (Fig. 3e), signifying successful vector assembly via gap repair. Following screening of the *ADE2* locus, we identified *ade2Δ* colonies in all three strains of *P. occidentalis*, but we were unable to identify *ade2Δ* mutants from Hyg<sup>R</sup> colonies of *P. kluyveri* (Fig. 3f). Editing efficiencies varied from 50% to 69% based on screening of 10–13 Hyg<sup>R</sup> transformants selected at random.

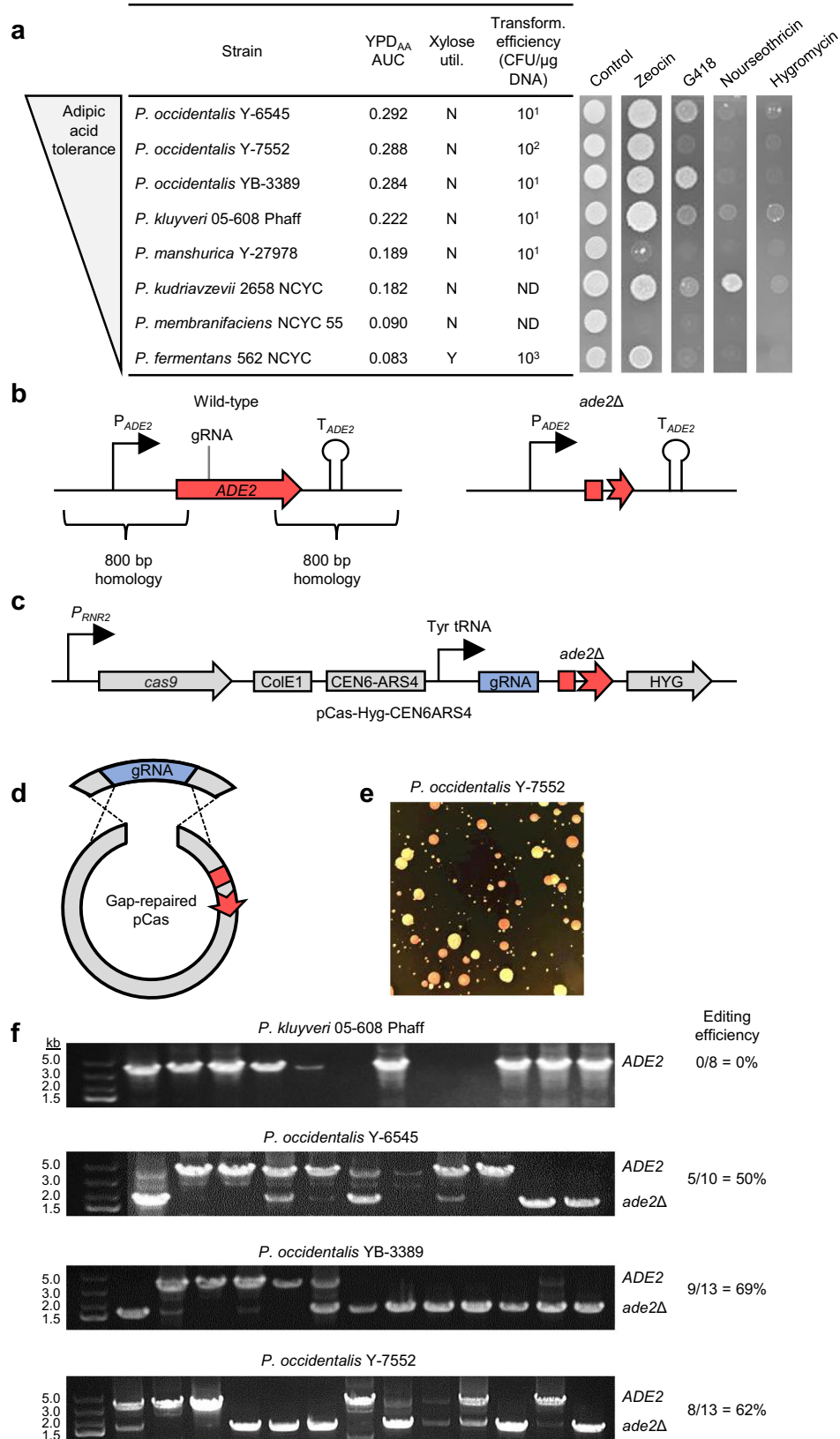
Taken together, *P. occidentalis* Y-7552 was selected as the final host for production of CCM. Strain Y-7552 is sensitive to at least three antibiotics (G418, hygromycin, and nourseothricin), yields the highest transformation efficiency of the three *P. occidentalis* strains (Fig. 3a), and does not exhibit the growth defect in 1× YNB observed for strains Y-6545 and YB-3389 (Supplementary Fig. 1). Strain Y-7552 is diploid and the type strain of *P. occidentalis*<sup>27</sup>. The strain is able to utilize glucose, glycerol, and ethanol as sole carbon sources, yet does not grow on sucrose, galactose, lactose, xylose, starch, or cellobiose<sup>20</sup> (Supplementary Table 1). In addition to *P. kudriavzevii*, *P. occidentalis* grows in medium devoid of common yeast vitamins, adding to the strain's potential for industrial production of high value organic acids.

### CRISPR-Cas9 editing and antibiotic marker recycling in *P. occidentalis* Y-7552

Having successfully deleted the *ADE2* gene in *P. occidentalis* strain Y-7552 using CRISPR-Cas9, we wished to cure cells of the pCas-Hyg-ADE2 gRNA vector, which is a requisite step for performing iterative rounds of CRISPR-Cas9 genome editing. Because the CEN6/ARS4 replication module is derived from *Saccharomyces cerevisiae*, we anticipated chromosomal integration of the pCas-Hyg-CEN6ARS4 vector in *Pichia*. Accordingly, the vector could not be cured by subculturing up to six times in non-selective YPD medium, and other methods of plasmid curing were unsuccessful, including culturing at an elevated temperature (37 °C), heat-shocking cells at 42 °C in a mock lithium-acetate-PEG transformation, or attempting to displace the original Hyg<sup>R</sup>-CEN6-ARS4 vector by introducing a G418<sup>R</sup>-CEN6-ARS4 or Nat<sup>R</sup>-CEN6-ARS4 vector to the *ade2Δ* host (Supplementary Fig. 2). We also constructed pCas-G418-ADE2 vectors containing the *S. cerevisiae* 2μ origin (pCas-G418-2μ-ADE2) and the broad host panARS origin (pCas-G418-panARS-ADE2), and while both vectors facilitated deletion of *ADE2* based on color phenotype and colony PCR, all transformants exhibited stable integration of the selectable markers (Supplementary Fig. 3).

To confirm vector integration, we constructed a pCas derivative lacking a functional plasmid origin (pCas-G418-NoOri-ADE2). Introduction of pCas-G418-NoOri-ADE2 to *P. occidentalis* generated G418-resistant transformants, including red *ade2Δ* colonies (Supplementary Fig. 4), indicating chromosomal integration as the vector lacks a functional replication origin. Collectively, these data are consistent with chromosomal integration of pCas vectors in *P. occidentalis*.

Chromosomal integration of vectors poses a challenge for strain engineering, as chromosomal markers are difficult to excise from cells,



thus hindering subsequent rounds of genome editing. To enable the recycling of antibiotic-resistance markers in *P. occidentalis*, we explored CRISPR-Cas9 as a means of inactivating or swapping chromosomal antibiotic resistance markers. We first deleted the *ADE2* gene using pCas-G418-CEN6ARS4-ADE2, yielding an *ade2Δ* G418<sup>R</sup> mutant. In a subsequent editing event, we introduced to the *ade2Δ* G418<sup>R</sup> strain a

pCas-Hyg-CEN6ARS4 vector harboring a gRNA targeting the G418-resistance marker. Targeting the chromosomal G418<sup>R</sup> marker yielded a substantial increase in the number of Hyg<sup>R</sup> transformants compared to a control transformation in which the G418<sup>R</sup> gRNA was replaced with a non-targeting gRNA (Supplementary Fig. 5). While we did not include an exogenous donor for repair of double-stranded breaks (DSBs)

**Fig. 3 | Xylose utilization, antibiotic susceptibility, plasmid transformation, and CRISPR-Cas9 editing of adipic acid tolerant *Pichia*.** **a** Xylose utilization, antibiotic susceptibility and transformation efficiency of adipic acid tolerant *Pichia* spp. Strains are sorted by adipic acid tolerance (AUC in YPD containing 20 g L<sup>-1</sup> adipic acid) and xylose utilization and susceptibility to four common yeast antibiotics is shown. Transformation efficiency using a plasmid conferring Hyg<sup>R</sup> is provided rounded to the nearest order of magnitude. Transformants of *P. kudriavzevii* 2658 NCYC and *P. membranifaciens* NCYC55 were not detected (ND). Repeating plasmid transformation and antibiotic susceptibility assays routinely yielded similar results. **b** *ADE2* deletion strategy utilized in adipic acid tolerant *Pichia*. Large homology regions (~800 bp) were designed flanking the *ADE2* coding sequence. **c** Structure of *ADE2* pCas CRISPR-Cas9 vectors containing precloned

*ade2Δ* donor cassettes. Donor cassettes were precloned into the BglII site of pCas vectors. **d** Gap repair strategy of pCas vector assembly in adipic acid tolerant *Pichia*. pCas vectors harboring precloned *ade2Δ* donor cassettes were linearized by digestion with BsaI or NotI and used to transform *Pichia* species along with an overlapping *ADE2* gRNA PCR product. **e** Deletion of *ADE2* yields pink colonies in adipic acid tolerant *Pichia*. Transformants were plated onto YPD agar plates containing hygromycin without adenine supplementation. A representative YPD agar plate containing *P. occidentalis* Y-7552 is shown. **f** Screening *ADE2* deletion in adipic acid tolerant *Pichia*. A pCas-Hyg-CEN6ARS4 vector possessing an *ADE2* gRNA and repair donor was transferred to *P. kluyveri* and three strains of *P. occidentalis*. The *ADE2* locus was screened for deletion in randomly selected Hyg<sup>R</sup> transformants prior to color development. Source data are provided as a Source Data file.

delivered to the chromosomal G418 marker, we suspected that the newly introduced pCas-Hyg-CEN6ARS4 vector provided an efficient DSB repair donor. Specifically, pCas-Hyg-CEN6ARS4 harbors a vector-encoded P<sub>TEFI</sub>-Hyg-T<sub>TEFI</sub> donor for repair of the DSB delivered to the chromosomal P<sub>TEFI</sub>-G418-T<sub>TEFI</sub> cassette within the *ade2Δ* G418<sup>R</sup> strain. Loss of G418 resistance was confirmed by absence of growth on G418-containing YPD agar plates and by PCR using primers specific to the Hyg<sup>R</sup> marker (Supplementary Fig. 5).

Having established a general strategy for recycling antibiotic markers, we tested this approach for performing iterative rounds of CRISPR editing in *P. occidentalis* by using our *ade2Δ* G418<sup>R</sup> mutant to delete a new target gene and concurrently inactivate the pre-existing G418<sup>R</sup> marker. For gene deletion, we selected *FCY1*, which encodes cytosine deaminase, as its inactivation confers resistance to 5-fluorocytosine, a toxic analog of cytosine (Supplementary Fig. 6). Deleting *FCY1* and eliminating G418<sup>R</sup> requires the introduction of two gRNAs from a single pCas-Hyg-CEN6ARS4 vector. To test the delivery and efficiency of two gRNA species, we transformed the *ade2Δ* G418<sup>R</sup> mutant with linearized pCas-Hyg-CEN6ARS4 vector along with *FCY1* and G418 gRNAs, each possessing overlap to pCas-Hyg-CEN6ARS4. Because only one gRNA species is required for gap repair, Hyg<sup>R</sup> colonies possess three possible genotypes: (1) *FCY1* G418<sup>S</sup> single mutants; (2) *fcy1Δ* G418<sup>R</sup> single mutants; or (3) *fcy1Δ* G418<sup>S</sup> double mutants. To assay the proportion of these genotypes, we introduced equal amounts of *FCY1* and G418 gRNAs and screened resulting Hyg<sup>R</sup> colonies for *FCY1* deletion and G418 susceptibility. Screening 18 Hyg<sup>R</sup> colonies in this manner yielded 11 G418<sup>S</sup> colonies, but no *fcy1Δ* mutants, indicating that gap repair occurred exclusively using the G418 gRNA despite the presence of equal amounts of G418 and *FCY1* gRNAs (Supplementary Fig. 7).

Because targeting a chromosomal G418<sup>R</sup> marker yields a major increase in the number of Hyg<sup>R</sup> transformants (Supplementary Fig. 5), we surmised that the transformation advantage afforded by the G418 gRNA obscured *fcy1Δ* mutant colonies, yielding exclusively *FCY1* G418<sup>S</sup> single mutants. To overcome this imbalance, we titrated the relative amount of G418 and *FCY1* gRNA species. We transformed the *ade2Δ* G418<sup>R</sup> mutant with a 20-, 10-, 5-, and 3-fold excess of *FCY1* gRNA relative to G418 gRNA to offset the advantage in transformation afforded by the G418 gRNA. Screening resultant Hyg<sup>R</sup> colonies for *FCY1* deletion and G418 susceptibility yielded several colonies possessing both *fcy1Δ* G418<sup>S</sup> editing events (Supplementary Fig. 8). The greatest proportion of *fcy1Δ* G418<sup>S</sup> colonies (37.5%) was attained by introducing a 10-fold excess of *FCY1* gRNA, while a five-fold and three-fold excess each yielded 27% and 25% double-edited colonies, respectively. A 20-fold excess of *FCY1* gRNA favored *FCY1* deletion, as only 12.5% of colonies possessed both *fcy1Δ* G418<sup>S</sup> editing events.

### Characterization of *P. occidentalis* promoters and terminators for strain engineering

To develop further tools for engineering *P. occidentalis* for production of heterologous metabolites, we characterized a panel of native promoter and terminator elements. We selected a suitable fluorescent

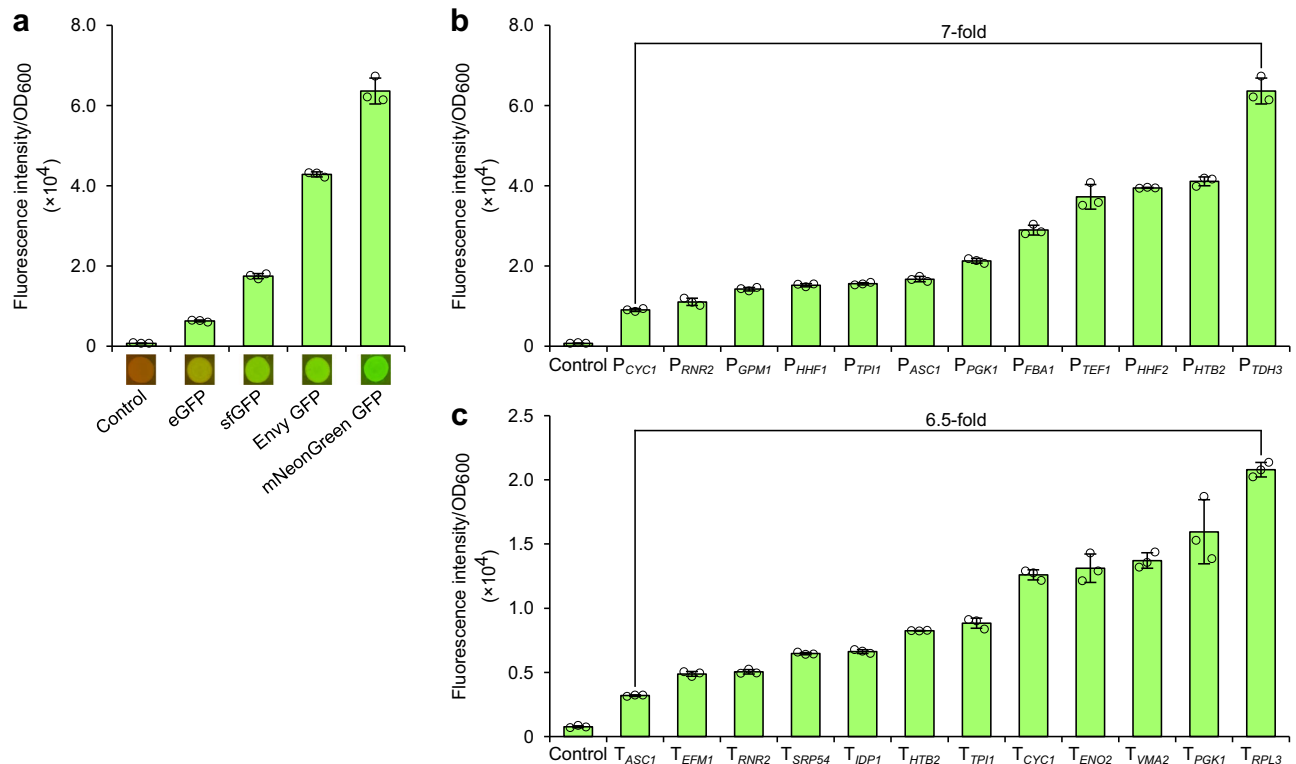
reporter by chromosomally integrating a set of four green fluorescent protein (GFP) gene variants using the *P. occidentalis* *TDH3* promoter (P<sub>TDH3</sub>), the ortholog of the strongest native promoter in *S. cerevisiae*<sup>28,29</sup>. Enhanced GFP (eGFP), superfolder GFP (sfGFP), Envy GFP, and mNeonGreen GFP were integrated into the *FCY1* locus of *P. occidentalis* using pCas-Hyg-CENARS4-*FCY1*. Cultures harboring mNeonGreen generated the greatest fluorescence followed by Envy, sfGFP, and eGFP (Fig. 4a).

We then selected a panel of 12 *P. occidentalis* promoters and terminators based on those characterized in *S. cerevisiae*<sup>28–30</sup> (Supplementary Data 2). Promoter constructs were cloned upstream of mNeonGreen with a fixed *RPL3* terminator, while terminators were cloned downstream of mNeonGreen with a fixed *GPMT1* promoter. Full mNeonGreen expression cassettes were cloned into vector pCas-Hyg-CEN6ARS4 flanked by homology to the *P. occidentalis* *FCY1* gene and integrated into the *FCY1* locus. The fluorescence output from mNeonGreen promoter and terminator integrants was quantified from 12-h exponential-phase cultures. Fluorescence output varied by a factor of 7 and 6.5 across the 12 selected promoters and terminators, respectively (Fig. 4b, c), and by a factor of nearly 20 across all 24 constructs, demonstrating effective modulation of gene expression in *P. occidentalis*. In line with *S. cerevisiae* and other organisms, expression of mNeonGreen from P<sub>TDH3</sub> yielded the highest fluorescence. T<sub>RPL3</sub> and T<sub>VMA2</sub> facilitated high-level expression in *P. occidentalis* and were among the top-performing elements of all 5302 terminators assayed in *S. cerevisiae*<sup>30</sup>.

### Engineering *P. occidentalis* for production of *cis,cis*-muconic acid

The techniques for CRISPR-Cas9 genome editing, antibiotic marker recycling, and heterologous gene expression in *P. occidentalis*, allowed us to engineer the host for heterologous production of toxic acid products. Although *P. occidentalis* was selected for its tolerance to adipic acid, yeast adipic acid synthesis is currently poor (2.59 mg L<sup>-1</sup>)<sup>14</sup> owing to an inability to functionally express bacterial enoate reductase enzymes in eukaryotic hosts<sup>13,14</sup>. In the absence of such a yeast-active enoate reductase enzyme, we opted to engineer *P. occidentalis* for synthesis of CCM, the direct precursor to adipic acid.

We first engineered *P. occidentalis* Y-7552 to produce protocatechuic acid (PCA), the committed heterologous metabolite in the proposed CCM synthesis pathway (Fig. 5a). We integrated *AROZ* encoding DHS dehydratase from *Podospira anserina* (*PaAROZ*) into the *ARO4* locus of *P. occidentalis*. The resulting strain (LP620) produced 1.0 g L<sup>-1</sup> PCA in batch microtiter plate cultivations from 20 g L<sup>-1</sup> glucose (Fig. 5b). We next integrated a feedback-resistant mutant of *P. occidentalis* 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase (*ARO4*<sup>K225L</sup>)<sup>31</sup> into LP620 (yielding LP622) which increased the PCA titer to 1.9 g L<sup>-1</sup>. Subsequent integration of *aroB* and *aroD* from *E. coli*, encoding 3-dehydroquinate synthase and 3-dehydroquinate dehydratase, respectively, into the *P. occidentalis* *ARO3* locus of strain LP622 (yielding LP630) facilitated a further increase in PCA titer to 2.0 g L<sup>-1</sup> in batch microtiter plate cultivations (Fig. 5b).



**Fig. 4 | Characterization of *P. occidentalis* promoters and terminators.** **a** OD-normalized fluorescence intensity of *P. occidentalis* Y-7552 strains expressing common GFP gene variants. GFP-encoding variants were expressed using P<sub>TDH3</sub> and T<sub>RPL3</sub> from *P. occidentalis* Y-7552. GFP-expressing colonies are shown with UV exposure. **b** OD-normalized fluorescence intensity of mNeonGreen expressed from

12 *P. occidentalis* gene promoters with the RPL3 terminator. **c** OD-normalized fluorescence intensity of mNeonGreen expressed from the GPM1 promoter and 12 *P. occidentalis* gene terminators. Error bars represent the mean  $\pm$  s.d. of  $n = 3$  independent biological samples. All GFP constructs were integrated into the FCY1 locus of *P. occidentalis*. Source data are provided as a Source Data file.

In the heterologous pathway, PCA decarboxylase (AroY) converts PCA to catechol, which is subsequently converted to CCM by 1,2-catechol dioxygenase (HQD2). We screened two bacterial AroY variants (AroY from *Klebsiella pneumoniae* and EcdC from *Enterobacter cloacae*) for conversion of PCA to catechol in strain LP630. PCA decarboxylases require a prenylated FMN cofactor that is synthesized by a flavin prenyltransferase in *S. cerevisiae* (ScPad1)<sup>32</sup>. We were unable to identify a candidate PADI-encoding gene in the genome of *P. occidentalis*, prompting us to investigate pairing of *aroY* and *ecdC* with ScPADI. Analysis of PCA decarboxylase strains revealed production of 1.7 g L<sup>-1</sup> (*aroY* + ScPADI) and 1.3 g L<sup>-1</sup> (*ecdC* + ScPADI) catechol and near-complete consumption of PCA (Fig. 5b; Supplementary Fig. 9). By contrast, omission of ScPADI failed to alter PCA levels relative to the control strain (Supplementary Fig. 9), and pairing *aroY* or *ecdC* with HQD2 from *Candida albicans* (CaHQD2) in the absence of ScPADI did not yield catechol or CCM.

To complete the heterologous CCM production pathway, we introduced CaHQD2 to our catechol production strain expressing *aroY* + ScPADI (LP632). Because *S. cerevisiae* CCM production strains accumulate PCA under fed-batch conditions<sup>9,10</sup>, we included an additional copy of *aroY* with CaHQD2, yielding strain LP635. The introduction of *aroY* + CaHQD2 resulted in the production of 2.5 g L<sup>-1</sup> CCM with a minor quantity of PCA (0.06 g L<sup>-1</sup>) and no detectable catechol (Fig. 5b).

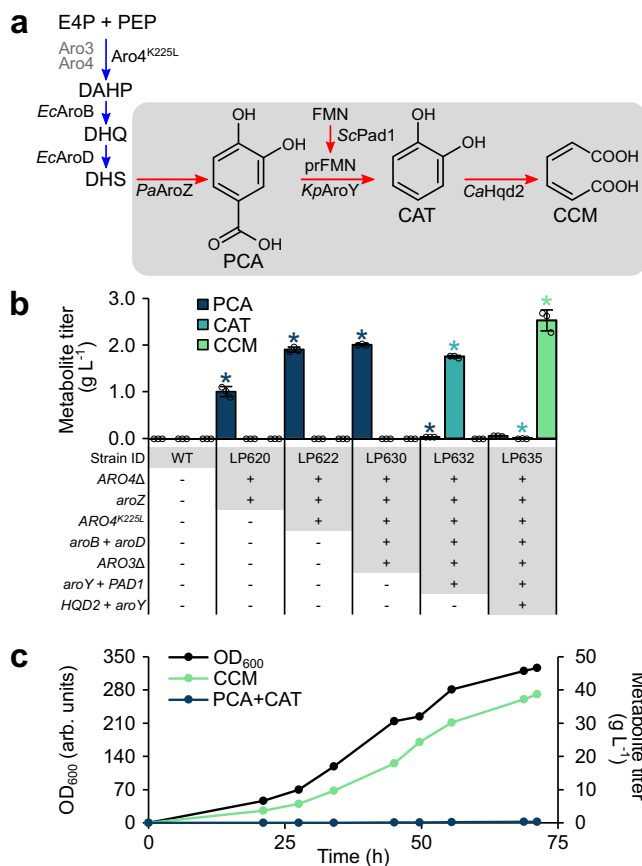
#### Fed-batch production of *cis,cis*-muconic acid

After establishing CCM synthesis in 96-well plate format, we characterized our *P. occidentalis* CCM production strain (LP635) in fed-batch fermentation. We first assessed CCM production under cultivation at pH 6.0 to benchmark our *P. occidentalis* host against prior studies involving analogous engineered strains of *S. cerevisiae*<sup>9,10</sup>.

Under these conditions, strain LP635 synthesized 38.8 g L<sup>-1</sup> CCM with a yield and productivity of 0.134 g g<sup>-1</sup> glucose and 0.511 g L<sup>-1</sup> h<sup>-1</sup>, respectively (Fig. 5c). A small amount of PCA and catechol was detected (<0.4 g L<sup>-1</sup> combined concentration).

Having established high-level production of CCM at intermediate pH, we next characterized our CCM production strain without pH control. Under these conditions, the pH rapidly dropped to 2.0 during the batch phase and growth stalled, as evidenced by a decline in CO<sub>2</sub> production (Supplementary Fig. 10). Growth could be transiently restored in a pH-dependent manner through stepwise increase in pH from 2.0 to 4.5 in 0.5-unit increments. For instance, increasing culture pH from 3.0 to 3.5 restored growth, as evidenced by a near instantaneous restoration of CO<sub>2</sub> production. However, growth was transient, as CO<sub>2</sub> concentration began to decline approximately 6 h following the increase in pH. Growth could again be restored by increasing pH from 3.5 to 4.0. Under these conditions, strain LP635 synthesized 7.2 g L<sup>-1</sup> CCM. Glycerol accumulated throughout the cultivation, reaching a concentration of 12.4 g L<sup>-1</sup> in late-stage samples.

Although *P. occidentalis* was selected for its robust tolerance to adipic acid at low pH, we surmised that poor growth of our CCM production strain at low pH was a consequence of CCM toxicity. To test this hypothesis, we first compared growth curves of wild-type *P. occidentalis* and strains engineered for PCA and CCM production in 3 $\times$  YNB (Supplementary Fig. 11). Our engineered PCA-producing strain exhibited a decline in maximum growth rate compared to wild-type *P. occidentalis* ( $\mu_{\max}$  0.23 h<sup>-1</sup> and 0.32 h<sup>-1</sup>, respectively), yet final biomass concentration was largely unaltered between the two strains (OD<sub>max</sub> 1.8 and 1.9, respectively). The same engineered strain also exhibited a prolonged lag phase relative to the wild-type strain. In contrast, engineering *P. occidentalis* for CCM production caused a 62% and 32% reduction in maximal growth rate and biomass concentration,



**Fig. 5 | Engineering *P. occidentalis* Y-7552 for efficient synthesis of CCM.**

**a** Heterologous pathway for synthesis of CCM from endogenous 3-dehydroshikimate (DHS). The native yeast shikimate pathway is shown in blue, while the heterologous CCM synthesis pathway is shaded and shown in red. **b** Titers of CCM pathway intermediates in culture supernatants of successive engineered *P. occidentalis* production strains. Metabolite titers are depicted in  $\text{g L}^{-1}$  for PCA, CAT, and CCM. Asterisks (\*) denote a significant increase or decrease in metabolite titer relative to the respective parent strain ( $P < 0.05$ ). Statistical differences between parent and derivative strains were tested using two-tailed Student's *t*-test. Error bars represent the mean  $\pm$  s.d. of  $n = 3$  independent biological samples. **c** Cultivation of a CCM-producing *P. occidentalis* strain (LP635) in a fed-batch fermentor using a mineral medium at pH 6.0. Growth of biomass ( $\text{OD}_{600}$ ) and accumulation of CCM pathway metabolites in the culture medium were measured during cultivation.  $\text{OD}_{600}$  is reported in arbitrary units (arb. units). CAT catechol, CCM *cis,cis*-muconic acid, DAHP 3-deoxy-D-arabinoheptulosonate 7-phosphate, DHQ 3-dehydroquininate, DHS 3-dehydroshikimate, E4P erythrose 4-phosphate, FMN flavin mononucleotide, PCA protocatechuic acid, PEP phosphoenolpyruvate, prFMN prenylated flavin mononucleotide. Source data are provided as a Source Data file.

respectively ( $\mu_{\text{max}}$  0.12  $\text{h}^{-1}$  and  $\text{OD}_{\text{max}}$  1.3) relative to wild-type *P. occidentalis*. While these findings are indicative of CCM toxicity, our CCM production strain is more extensively genetically modified than our PCA-producing strain, raising the possibility that growth inhibition could result from effects of genome editing, such as vector integration or metabolic burden from increased expression of *cas9*.

To better establish the effect of CCM on *P. occidentalis* fitness, we compared growth curves of wild-type *P. occidentalis* in 3 $\times$  YNB saturated with CCM ( $< 2 \text{ g L}^{-1}$ ) or adipic acid ( $22 \text{ g L}^{-1}$ ). In both instances the pH of the medium was  $< 3.5$ . While growth of *P. occidentalis* was largely unaffected in the presence of  $22 \text{ g L}^{-1}$  adipic acid ( $\mu_{\text{max}}$  0.28  $\text{h}^{-1}$  and  $\text{OD}_{\text{max}}$  1.8) relative to media without acid supplementation ( $\mu_{\text{max}}$  0.32  $\text{h}^{-1}$  and  $\text{OD}_{\text{max}}$  1.9), supplementation with  $< 2 \text{ g L}^{-1}$  CCM caused a 38% and 42% reduction in maximal growth rate and  $\text{OD}_{600}$  ( $\mu_{\text{max}}$

0.20  $\text{h}^{-1}$  and  $\text{OD}_{\text{max}}$  1.1) (Supplementary Fig. 11), demonstrating significant toxicity of CCM relative to adipic acid. CCM toxicity was not specific to *P. occidentalis* Y-7552, as all adipic acid tolerant *Pichia* species and strains assayed exhibited marked toxicity to  $< 2 \text{ g L}^{-1}$  CCM (Supplementary Fig. 12) despite notable tolerance to adipic acid (Fig. 3a, Table 1). Growth of *P. fermentans* 562 NCYC and the *S. cerevisiae* control was completely inhibited in the presence of  $< 2 \text{ g L}^{-1}$  CCM.

## Discussion

In this study we screened 108 yeast strains across 28 genera and identified *P. occidentalis* as a non-conventional host with exceptional tolerance to several dicarboxylic acids, notably adipic acid. We developed a comprehensive strain engineering toolkit for domestication of *P. occidentalis*, including techniques for CRISPR-Cas9 genome editing, antibiotic marker recycling and high-level expression of heterologous genes. We leveraged our toolkit to engineer the host for high-level production of CCM, the direct precursor to adipic acid, using a simple mineral medium in a controlled fed-batch cultivation. Under these conditions, production of CCM by our strain ( $38.8 \text{ g L}^{-1}$  at  $0.511 \text{ g L}^{-1} \text{ h}^{-1}$ ) outperformed all prior *S. cerevisiae* CCM strains spanning more than 10 studies and achieving up to  $22.5 \text{ g L}^{-1}$  CCM at  $0.21 \text{ g L}^{-1} \text{ h}^{-1}$  (ref. 10). Our strain also outcompeted a recent evolved and engineered *Pseudomonas putida* host with a CCM output of  $33.7 \text{ g L}^{-1}$  CCM at  $0.18 \text{ g L}^{-1} \text{ h}^{-1}$  (ref. 8). Commercial-scale production of biobased acids typically demands titers of at least  $50\text{--}100 \text{ g L}^{-1}$  at a rate of more than  $1 \text{ g L}^{-1} \text{ h}^{-1}$  (refs. 33,34) and, accordingly, our efforts provide a promising foundation for industrial production of CCM and adipic acid using non-conventional yeasts, such as *P. occidentalis*.

CCM is converted to adipic acid via hydrogenation by bacterial enoate reductase enzymes<sup>35</sup>, which presently function poorly in yeast<sup>13</sup>. Accordingly, *P. occidentalis* was selected for its robust tolerance to adipic acid, yet we engineered the strain to produce CCM due to the lack of a suitable enoate reductase. Unexpectedly, CCM was substantially more toxic to *P. occidentalis* than adipic acid, even when adipic acid was supplied at a 10-fold higher concentration ( $22 \text{ g L}^{-1}$  compared to  $< 2 \text{ g L}^{-1}$ ). CCM toxicity was pH-dependent, as growth of strain LP635 could not be sustained between pH 2.0 and 3.5, well below the  $\text{pK}_a$  of CCM (3.87), yet was largely unaffected at pH 6.0 ( $\text{OD}_{600}$  of 327 in fed-batch fermentation), enabling highly productive synthesis of CCM ( $38.8 \text{ g L}^{-1}$ ). A previous fed-batch study employing engineered *S. cerevisiae* also reported growth arrest and incomplete sugar utilization at CCM concentrations above  $2.5 \text{ g L}^{-1}$  (ref. 6). In light of these findings, we screened our most adipic acid tolerant yeasts for CCM tolerance, which revealed that all three *P. occidentalis* isolates were the most tolerant strains to both acids, suggesting similar mechanisms of toxicity between CCM and adipic acid. Owing to the unexpected toxicity of CCM against all yeasts assayed, extending the heterologous pathway to adipic acid is likely to improve strain fitness and metabolite production at low pH. Therefore, future efforts should focus on the identification or engineering of alternative enoate reductase variants with improved functionality in yeast hosts. Meanwhile, in situ extraction is a promising avenue to mitigate CCM toxicity, as a wide array of biocompatible extractants have been evaluated for efficient removal of CCM and related metabolites<sup>6,36</sup>.

Other groups have shown that certain strains of *Pichia* are tolerant to high concentrations of organic acids<sup>37,38</sup>, yet the numbers of strains assayed were comparatively small and comprised largely of environmental isolates rather than common type strains. Aside from one study investigating lactic acid production using a novel isolate<sup>37</sup>, *P. occidentalis* has not been widely utilized as a metabolic engineering host and investigations into its acid tolerance have focussed largely on issues with food contamination<sup>39,40</sup>. It is thus unlikely that *P. occidentalis* would have been selected for screening had we reduced the breadth of our assay or selected strains based on a simple literature search. Moreover, our screen identified three strains of *P. occidentalis* as the most tolerant



to adipic acid of all 108 strains screened. Using *P. occidentalis* Y-7552, we developed methods for transformation and CRISPR-mediated genome editing. Owing to the low transformation efficiency and homologous recombination of *Pichia* spp., we designed a CRISPR-Cas9 delivery system in which homology-directed repair donors are pre-cloned into pCas vectors. Relatedly, the efficiency of CRISPR-mediated genome editing in *Pichia pastoris* was improved up to 25-fold by including an ARS within DNA repair donors<sup>41</sup>. Using this system, repair donors are stably inherited by daughter cells and are resistant to DNA exonucleases, thus increasing the likelihood of targeted editing events<sup>42</sup>. While we succeeded in inactivating the *ADE2* and *FCY1* loci, yielding mutants with screenable phenotypes (red pigmentation and 5-FC resistance, respectively), we were unable to identify a replicative plasmid origin for use in *P. occidentalis* despite the screening of several plasmids, including the broad *panARS* origin<sup>43</sup>. To address this challenge, we developed a CRISPR-mediated marker swapping strategy, which we leveraged to perform five successive rounds of genome editing in the construction of our final CCM production strain. To add to our genetic toolkit for *P. occidentalis*, we characterized a panel of 12 promoters and 12 terminators, which enabled us to modulate enzyme activity by a factor of 20. A genome-wide assay of all 5,302 *S. cerevisiae* terminators revealed an 81-fold difference in terminator activity<sup>30</sup>, while the relative strength of *P. occidentalis* promoters ( $P_{TPH} < P_{PGKI} < P_{TEFI} < P_{TDH3}$ ) followed a similar overall pattern as those from *P. pastoris* ( $P_{PGKI} < P_{TPH} < P_{TEFI} < P_{TDH3}$ )<sup>44</sup>. Taken together, our CRISPR-Cas9 editing technique and gene expression toolkit enables efficient strain engineering of *P. occidentalis* and contributes to the growing collection of non-conventional yeasts, such as *Yarrowia lipolytica*, *Kluyveromyces marxianus*, and *Komagataella phaffi* (*P. pastoris*)<sup>45</sup>, that have been domesticated for industrial applications.

In summary, screening culture collections for prospective adipic acid tolerant yeasts identified *P. occidentalis* as a non-conventional and untapped host for industrial production of CCM and adipic acid. The strain engineering tools and techniques developed herein provide a solid foundation for engineering acid tolerant *Pichia* and will enable the construction of highly robust organic acid production strains.

## Methods

### Plasmids, strains, and growth media

All 153 non-*Saccharomyces* yeast strains ordered in this study are listed in Supplementary Data 1. *Saccharomyces cerevisiae* CEN.PK2-7D (*MATa*) was utilized for assembling plasmids in vivo via gap repair. Unless stated otherwise all yeast cultures were grown in YPD (10 g L<sup>-1</sup> Bacto Yeast Extract, 20 g L<sup>-1</sup> Bacto peptone, 20 g L<sup>-1</sup> glucose). Recombinant *S. cerevisiae* strains were selected on YPD agar containing 400 µg mL<sup>-1</sup> G418 or 200 µg mL<sup>-1</sup> hygromycin B. Recombinant *Pichia* strains were selected on YPD agar containing 1200 µg mL<sup>-1</sup> G418, 600 µg mL<sup>-1</sup> hygromycin B, or 100 µg mL<sup>-1</sup> nourseothricin. Plasmids and engineered strains employed in this work are provided in Tables 2 and 3, respectively. Oligonucleotides utilized in this work are provided in Supplementary Data 2.

### Microtiter plate assay for tolerance to various acids

Strains were received from repositories and used to inoculate liquid yeast malt (YM) medium, containing 5 g L<sup>-1</sup> of peptone, 3 g L<sup>-1</sup> of yeast extract, 3 g L<sup>-1</sup> of malt extract and 10 g L<sup>-1</sup> of dextrose. Strains were grown at 30 °C until fully saturated (up to seven days) and subsequently stored in 15% glycerol at -80 °C.

For high-throughput acid tolerance assays, frozen strains were revived in batches of 30 in YPD medium at 30 °C. Once all cultures were visibly turbid, strains were back-diluted (1:20) into 180 µL of YPD in a shallow 96-well plate and left to grow for 8 h or until cultures reached an OD<sub>600</sub> of 1.8. Following the second incubation, OD<sub>600</sub> was measured on a plate reader and a custom script was used to create a CSV file with the absolute volumes of diluent and cell culture necessary to dilute each culture to an OD<sub>600</sub> of 1.8. Dilutions were then carried out using the Span-8 of a Biomek FXP liquid handling system to create an intermediate plate with 30 strains in triplicate, including *S. cerevisiae* CEN.PK 113-7D as a control. From the intermediate plates, a 10 µL inoculum was drawn and dispensed into an experimental plate containing 170 µL of the relevant growth medium. The final volume of each well was 180 µL and the initial OD<sub>600</sub> of all strains was 0.1. Initial tests for organic acid tolerance were conducted in YPD medium, YPD<sub>cit</sub>

**Table 2 | Plasmids utilized in this study**

Plasmid	Description	Source or reference
pCAS-G418-2µ	<i>P<sub>RNR2</sub>-Cas9<sub>NLS</sub>-TCYCI</i> , pUC, 2µ, <i>P<sub>rRNA_Tyr</sub></i> -3HDV-gRNA-Scaffold- <i>T<sub>SNR52</sub></i> , <i>P<sub>TEFI</sub>-kanMX-T<sub>TEFI</sub></i>	46
pCAS-Hyg-2µ	<i>P<sub>RNR2</sub>-Cas9<sub>NLS</sub>-TCYCI</i> , pUC, 2µ, <i>P<sub>rRNA_Tyr</sub></i> -3HDV-gRNA-Scaffold- <i>T<sub>SNR52</sub></i> , <i>P<sub>TEFI</sub>-HphNTI-T<sub>TEFI</sub></i>	49
pCas-G418-CEN6ARS4	CEN6ARS4 derivative of pCAS-G418-2µ	This study
pCas-Hyg-CEN6ARS4	CEN6ARS4 derivative of pCAS-Hyg-2µ	This study
pCas-G418-panARS	panARS derivative of pCAS-G418-2µ	This study
pCas-G418-CEN6ARS4-PoADE2	<i>P. occidentalis ade2Δ</i> donor cloned into pCas-G418-CEN6ARS4	This study
pCas-Hyg-CEN6ARS4-PoADE2	<i>P. occidentalis ade2Δ</i> donor cloned into pCas-Hyg-CEN6ARS4	This study
pCas-G418-2µ-PoADE2	<i>P. occidentalis ade2Δ</i> donor cloned into pCAS-G418-2µ	This study
pCas-Hyg-2µ-PoADE2	<i>P. occidentalis ade2Δ</i> donor cloned into pCas-Hyg-2µ	This study
pCas-G418-panARS-PoADE2	<i>P. occidentalis ade2Δ</i> donor cloned into pCas-G418-panARS	This study
pCas-G418-NoOri-PoADE2	Deletion of the 2µ origin from pCAS-G418-2µ-PoADE2	This study
pCas-Hyg-CEN6ARS4-PKIAD2	<i>P. kluyveri ade2Δ</i> donor cloned into pCas-Hyg-CEN6ARS4	This study

**Table 3 | Engineered *Pichia occidentalis* strains utilized in this study**

ID	Description	Parent	Manipulation	Reference
LP618	Deletion of <i>ADE2</i>	WT <i>P. occidentalis</i>	<i>ade2Δ</i>	This study
LP620	Integration of <i>PaAROZ</i>	WT <i>P. occidentalis</i>	<i>ARO4::PaAROZ</i>	This study
LP622	Integration of <i>PoARO4</i> <sup>K225L</sup>	LP620	01b- <i>PoARO4</i> <sup>K225L</sup> -01b	This study
LP630	Integration of <i>Ec.aroB</i> and <i>Ec.aroD</i>	LP622	<i>ARO3::Ec.aroB-Ec.aroD</i>	This study
LP632	Integration of <i>Kp.aroY</i> and <i>ScPAD1</i>	LP630	03b- <i>Kp.aroY-ScPAD1</i> -03b	This study
LP635	Integration of <i>CaHQD2</i> and <i>Kp.aroY</i>	LP632	03a- <i>CaHQD2-Kp.aroY</i> -03a	This study

(YPD containing 0.1 M citric acid added to decrease the pH to 3.0), or YPD<sub>AA</sub> (YPD containing 20 g L<sup>-1</sup> of adipic acid, resulting in a pH of 3.7). Organic acid tests conducted in minimal media used a common base of 20 g L<sup>-1</sup> of glucose, 5.1 g L<sup>-1</sup> of YNB without amino acids and without ammonium sulfate, 5.0 g L<sup>-1</sup> of ammonium sulfate, and 0.15 M of relevant organic acid. HCl was added dropwise to a final pH of 2.8. After inoculation, plates were sealed with parafilm and incubated at 30 °C in Tecan Sunrise plate readers for 48 h. Plates were orbitally shaken for 15 min, allowed to rest for 5 min prior to OD<sub>600</sub> readings, and shaking was restarted.

For CCM tolerance assays, growth curves were generated in triplicate in 96-well microtiter plates containing 180 µl of medium. Strains engineered for PCA or CCM production were assayed in 3× YNB medium containing 2% glucose. For tolerance to exogenous CCM, CCM was added to 3× YNB at 2 g L<sup>-1</sup>, heated gently at 50 °C for 4–6 hours and the resulting partially dissolved slurry was filtered. For tolerance to exogenous adipic acid, adipic acid was dissolved at 22 g L<sup>-1</sup> (0.15 M) with gentle heating. Cultures were inoculated using saturated overnight cultures to an initial OD<sub>600</sub> of roughly 0.1 and wrapped in Parafilm to minimize evaporation. Absorbance was measured at 600 nm every 20 min with a Sunrise absorbance microplate reader (Tecan) over the course of 30 h. Maximum specific growth rates ( $\mu_{\max}$  h<sup>-1</sup>) were determined from triplicate cultures and were based on OD<sub>600</sub> readings.

### Yeast strain construction

Yeast genetic modifications were made via CRISPR-Cas9-mediated genomic integration<sup>46,47</sup> and in vivo DNA assembly<sup>48</sup>. Sequences of integration loci, gRNAs, *ADE2*, promoters, and terminators from adipic acid tolerant *Pichia* spp. were obtained from reported draft genome sequences<sup>24</sup>. Cas9 and gRNAs were delivered to *Pichia* strains using pCas-G418-CEN6ARS4 (ref. 46) or pCas-Hyg-CEN6ARS4 (ref. 49). For chromosomal integration in *Pichia*, homology repair donors were first cloned into pCas vectors via gap repair in *S. cerevisiae*. Homology repair donors were designed to possess roughly 800 bp of flanking homology to the *Pichia* genome, in addition to 40–50 bp of internal overlap and 40–50 bp of external overlap to BglII-digested pCas-G418-CEN6ARS4 for assembly in *S. cerevisiae*. pCas vectors harboring repair donor templates were recovered from *S. cerevisiae*, linearized with BsaI or NotI and 200 ng was used to transform *Pichia* strains with a five- to tenfold excess of a gRNA species targeting the *Pichia* genomic integration locus (600 ng) relative to a gRNA targeting chromosomal G418- or Hyg-resistance markers (120 ng) in a standard 50 µl lithium acetate transformation. gRNAs were retargeted using a two-step SOE PCR. Cells were heat-shocked at 42 °C for 40 min, recovered for 16 h without shaking, and plated onto YPD agar plates containing appropriate antibiotics. Selection of gRNAs targeting *Pichia* loci was performed using CCTop<sup>50</sup>. gRNA target sites utilized in this study are provided in Table 4. Large intergenic regions (>1.5 kb) were selected as loci for DNA integration in *P. occidentalis*. *P. occidentalis* promoters (0.6–1.5 kb) were identified by selecting the full sequence flanked by the target and upstream coding sequences. *P. occidentalis* terminators (0.3–1.2 kb) were identified by selecting the full sequence flanked by the target and downstream coding sequences. *Pichia* expression cassettes are provided in Table 5. Sequences of *P. occidentalis* promoters and terminators, and integration loci are provided in Supplementary Data 3 and 4, respectively. Genes and synthetic DNAs are provided in Supplementary Data 5.

### Microtiter plate assay of GFP-producing strains

*Pichia* colonies expressing GFP variants were picked in triplicate into standard V-bottom 96-well microtiter plates containing 0.15 mL YNB medium and 20 g L<sup>-1</sup> glucose. Cultures were grown overnight and diluted into 0.8 mL fresh YNB medium in deep 96-well plates to an initial OD<sub>600</sub> of approximately 0.1. Cultures were

incubated with shaking for 12 h prior to OD<sub>600</sub> and fluorescence measurements. Fluorescence was measured using the M200 plate reader (Tecan) using an excitation wavelength of 485 nm and an emission wavelength of 525 nm. Gain was adjusted manually for

**Table 4 | *Pichia* target sites utilized in this study**

Organism	Target site ID	Target site sequence <sup>a</sup>	Reference
<i>P. kluyveri</i>	ADE2	TAGAGAAATGTCTGATTAATGG	This study
<i>P. occidentalis</i>	ADE2	TTAATGAACCCGGTTGTGATGG	This study
	FCY1	TAAGTTTCTGTTCGCCCTGGG	This study
	ARO3	ATCAAGGGCAAACCTCTGTGG	This study
	ARO4	TGACATCTTGTGAATCTGTGTGG	This study
	O1b	TTTTGGTAGCGTGCCTGTACGG	This study
	O3a	TGACTATACACTCTGTCTACAGG	This study
	O3b	TTTGTGTAAACTATTAAGCGGG	This study
Other	G418 <sup>R</sup>	TCCGACTCCTGATGATGCATGG	This study
	Hyg <sup>R</sup>	TCTCGATGAGCTGATGCTTTGGG	This study

<sup>a</sup>PAMs are underlined.

**Table 5 | Expression cassettes utilized in this study**

Description	Cassette <sup>a</sup>	Locus
eGFP	P <sub>TDH3</sub> -eGFP-T <sub>RPL3</sub>	FCY1
SuperfolderGFP (sfGFP)	P <sub>TDH3</sub> -sfGFP-T <sub>RPL3</sub>	FCY1
Envy GFP	P <sub>TDH3</sub> -Envy-T <sub>RPL3</sub>	FCY1
mNeonGreen GFP	P <sub>TDH3</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>ASC1</sub> -mNeonGreen	P <sub>ASC1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>CYC1</sub> -mNeonGreen	P <sub>CYC1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>FBA1</sub> -mNeonGreen	P <sub>FBA1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>GPM1</sub> -mNeonGreen	P <sub>GPM1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>HHF1</sub> -mNeonGreen	P <sub>HHF1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>HHF2</sub> -mNeonGreen	P <sub>HHF2</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>HTB2</sub> -mNeonGreen	P <sub>HTB2</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>PGK1</sub> -mNeonGreen	P <sub>PGK1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>RNR2</sub> -mNeonGreen	P <sub>RNR2</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>TEF1</sub> -mNeonGreen	P <sub>TEF1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
P <sub>TPH1</sub> -mNeonGreen	P <sub>TPH1</sub> -mNeonGreen-T <sub>RPL3</sub>	FCY1
mNeonGreen-T <sub>ASC1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>ASC1</sub>	FCY1
mNeonGreen-T <sub>CYC1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>CYC1</sub>	FCY1
mNeonGreen-T <sub>EFM1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>EFM1</sub>	FCY1
mNeonGreen-T <sub>ENO2</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>ENO2</sub>	FCY1
mNeonGreen-T <sub>HTB2</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>HTB2</sub>	FCY1
mNeonGreen-T <sub>IDP1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>IDP1</sub>	FCY1
mNeonGreen-T <sub>PGK1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>PGK1</sub>	FCY1
mNeonGreen-T <sub>RNR2</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>RNR2</sub>	FCY1
mNeonGreen-T <sub>SRP54</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>SRP54</sub>	FCY1
mNeonGreen-T <sub>TPH1</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>TPH1</sub>	FCY1
mNeonGreen-T <sub>VMA2</sub>	P <sub>GPM1</sub> -mNeonGreen-T <sub>VMA2</sub>	FCY1
PaAROZ	P <sub>TDH3</sub> -PaAROZ-T <sub>RPL3</sub>	ARO4
PoARO4 <sup>K225L</sup>	P <sub>TDH3</sub> -PoARO4 <sup>K225L</sup> -T <sub>EFM1</sub>	O1b
<i>Ec.aroB</i> - <i>Ec.aroD</i>	P <sub>TDH3</sub> - <i>Ec.aroB</i> -T <sub>IDP1</sub> -LTP1-P <sub>GPM1</sub> - <i>Ec.aroD</i> -T <sub>VMA2</sub>	ARO3
<i>Kp.aroY</i> - <i>ScPAD1</i>	P <sub>TDH3</sub> - <i>Kp.aroY</i> -T <sub>SRP54</sub> -LTP1-P <sub>TEF1</sub> - <i>ScPAD1</i> -T <sub>RPL15A</sub>	O3b
<i>CaHQD2</i> - <i>Kp.aroY</i>	P <sub>TDH3</sub> - <i>CaHQD2</i> -T <sub>SRP54</sub> -LTP1-P <sub>TEF1</sub> - <i>Kp.aroY</i> -T <sub>RPL15A</sub>	O3a

<sup>a</sup>LTP1 is a synthetic 40 bp linker (5'-CCGGTCTTAGAAAACGCATAAACATACAAGTGGACAGATG-3') used to join expression cassettes.

each sample. A strain lacking GFP was included as a control for autofluorescence.

### Microtiter plate assay

*Pichia* colonies synthesizing heterologous metabolites were picked in triplicate into deep 96-well plates containing 0.5 mL YPD medium. Cultures were grown overnight, diluted 50× into fresh YPD medium, and incubated for 48–72 h. OD<sub>600</sub> of cultures was measured using the M200 plate reader (Tecan) and samples were frozen at −20 °C prior to HPLC or LC-MS analysis.

### Fed-batch cultivation of *P. occidentalis* strains

Controlled fed-batch fermentations were carried out in 3 L BioBundle fermentors (Applikon). Cultivation temperature was maintained at 30 °C. Fermentations were performed at pH 6.0 by automated titration with 12.5% NH<sub>3</sub> or without pH control and stepwise adjustment of pH during cultivation with 4 N NaOH to rescue growth of CCM-producing strains at low pH (pH 2.0–3.5). Dissolved oxygen was maintained at 30% of air saturation (aeration rate 1.5 L min<sup>−1</sup>). Off-gas composition (concentration of O<sub>2</sub> and CO<sub>2</sub>) was analyzed using a Tandem Multiplex gas analyzer (Magellan BioTech). Bioreactor inoculum was generated in a 500 mL shake flask containing 50 mL of batch medium (40 g glucose, 2.5 g KH<sub>2</sub>PO<sub>4</sub>, 6.0 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.0 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 mL vitamin stock, and 5 mL trace element stock per liter) and grown for 36–48 h at 30 °C. Composition of the vitamin and trace element stock solutions are described in Pyne et al.<sup>51</sup> Cells were washed and suspended in 0.9% NaCl and used to inoculate (OD<sub>600</sub> = −0.3) 1 L of batch medium. The culture was operated in batch mode until glucose was exhausted, followed by fed-batch phase with automated addition of feed. Feeding medium for cultivation at pH 6.0 contained 600 g glucose, 20.8 g KH<sub>2</sub>PO<sub>4</sub>, 5 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 8.3 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 g K<sub>2</sub>SO<sub>4</sub>, 20 mL vitamin stock, and 20 mL trace element stock per liter. Feeding medium for cultivation without pH control contained 360 g glucose, 7.5 g KH<sub>2</sub>PO<sub>4</sub>, 30 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 3 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 7.5 mL vitamin stock, and 7.5 mL trace element stock per liter. Samples were collected every 12 h for a total of three days. Cell dry weight (g L<sup>−1</sup>) was calculated using a conversion factor of 0.36 g L<sup>−1</sup> per OD<sub>600</sub> (determined gravimetrically). Bioreactor data were collected and analyzed using BioXpert 1.3 (Applikon).

### LC-MS and HPLC-UV analysis

LC-MS was utilized to quantify PCA and CCM from cultures grown in deep 96-well plates. To extract metabolites, 7 µL cell broth was combined with 27 µL 100% acetonitrile and samples were shaken vigorously for 5 min, followed by addition of 146 µL of 0.123% formic acid (0.1% final concentration). Extracts were centrifuged at 4000 RCF for 10 min and supernatants were further diluted 10-fold in 15% acetonitrile containing 0.1% formic acid prior to LC-MS analysis. Ten µL of diluted extract was separated on a 1290 Infinity II LC system (Agilent Technologies) with a Zorbax Rapid Resolution HT C18 column (50 × 2.1 mm, 1.8 µm; Agilent Technologies). Metabolites were separated using the following gradient: 2% B to 10% B from 0 to 4 min (0.3 mL min<sup>−1</sup>), 10% B to 85% B from 4 to 6 min (0.3 mL min<sup>−1</sup>), held at 85% B from 6 to 7 min (0.3 mL min<sup>−1</sup>), 85% B to 2% B from 7 to 7.1 min (0.3 mL min<sup>−1</sup>), and held at 2% B from 7.1 to 9 min (0.45 mL min<sup>−1</sup>). Solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in 100% ACN. Following separation, eluent was injected into an Agilent 6545 quadrupole time-of-flight MS (QTOF-MS; Agilent Technologies) in negative mode. The sample tray and column compartment were set to 4 °C and 30 °C, respectively. The sheath gas flow rate and temperature were adjusted to 10 L min<sup>−1</sup> and 350 °C, respectively, while drying and nebulizing gases were set to 12 L min<sup>−1</sup> and 55 psig, respectively. The drying gas temperature was set to 325 °C. QTOF data was processed and manipulated using MassHunter Workstation software version B.06.00 (Agilent Technologies).

Glucose, glycerol, PCA, catechol, and CCM from fed-batch fermentation broth samples were analyzed and quantified using HPLC-UV and RID. For analysis of CCM, culture broth was combined with an equal volume of 10 N NaOH and incubated for 10 min with shaking. Cell debris was centrifuged at maximum speed for 10 min and supernatant was diluted in 40 mM PBS, pH 7.4. Alternatively, CCM was analyzed by diluting culture supernatant 50-fold directly in water or 40 mM PBS, pH 7.4. For analysis of glucose, glycerol, PCA, and catechol, culture broth was diluted in 40 mM PBS, pH 7.4 and centrifuged at maximum speed for 10 minutes. Five µL of extracted broth was separated on an Agilent 1200 HPLC system equipped with an Aminex Fast Acid Analysis column. Metabolites were separated isocratically using 10 mM H<sub>2</sub>SO<sub>4</sub> at a flow rate of 0.8 mL min<sup>−1</sup>. PCA and catechol were detected at 210 nm. CCM was detected at 260 nm. Traces were collected and analyzed using ChemStation version B.02.00 (Agilent Technologies).

### Statistical analyses

All numerical values are depicted as means ± s.d. Statistical differences between control and engineered strains were assessed via two-tailed Student's *t*-tests assuming equal variances. In all cases, *P*-values < 0.05 were considered significant. Statistical significance of dicarboxylic acid chain length was assessed via OLS linear regression with hourly growth rates as the response variable. Statistical analyses were performed using Excel 2306 Build 16.0.16529.20164 (Microsoft) or Prism version 9.5.1 (GraphPad).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Draft genome sequences of relevant acid tolerant *Pichia* strains were accessed under the BioProject accession [PRJNA870168](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA870168). Source data are provided with this paper.

### References

- Knoshaug, E. P. & Zhang, M. Butanol tolerance in a selection of microorganisms. *Appl. Biochem. Biotechnol.* **153**, 13–20 (2009).
- Hansen, E. H. et al. De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Appl. Environ. Microbiol.* **75**, 2765–2774 (2009).
- Long, A. & Ward, O. Biotransformation of benzaldehyde by *Saccharomyces cerevisiae*: characterization of the fermentation and toxicity effects of substrates and products. *Biotechnol. Bioeng.* **34**, 933–941 (1989).
- Fletcher, E., Mercurio, K., Walden, E. A. & Baetz, K. A yeast chemogenomic screen identifies pathways that modulate adipic acid toxicity. *IScience* **24**, 102327 (2021).
- Narendranath, N., Thomas, K. & Ingledew, W. Effects of acetic acid and lactic acid on the growth of *Saccharomyces cerevisiae* in a minimal medium. *J. Ind. Microbiol. Biotechnol.* **26**, 171–177 (2001).
- Nicolaï, T., Deparis, Q., Foulquié-Moreno, M. R. & Thevelein, J. M. In-situ muconic acid extraction reveals sugar consumption bottleneck in a xylose-utilizing *Saccharomyces cerevisiae* strain. *Microb. Cell Factories* **20**, 1–19 (2021).
- Jarboe, L. R., Royce, L. A. & Liu, P. Understanding biocatalyst inhibition by carboxylic acids. *Front. Microbiol.* **4**, 272 (2013).
- Ling, C. et al. Muconic acid production from glucose and xylose in *Pseudomonas putida* via evolution and metabolic engineering. *Nat. Commun.* **13**, 1–14 (2022).
- Wang, G. et al. Improvement of *cis*, *cis*-muconic acid production in *Saccharomyces cerevisiae* through biosensor-aided genome engineering. *ACS Synth. Biol.* **9**, 634–646 (2020).
- Wang, G. et al. An integrated yeast-based process for *cis*, *cis*-muconic acid production. *Biotechnol. Bioeng.* **119**, 376–387 (2022).

11. Skoog, E., Shin, J. H., Saez-Jimenez, V., Mapelli, V. & Olsson, L. Biobased adipic acid—The challenge of developing the production host. *Biotechnol. Adv.* **36**, 2248–2263 (2018).
12. Werpy, T. & Petersen, G. Top value added chemicals from biomass: volume I—results of screening for potential candidates from sugars and synthesis gas. (National Renewable Energy Lab., Golden, CO, 2004).
13. Biz, A. & Mahadevan, R. Overcoming challenges in expressing iron–sulfur enzymes in yeast. *Trends Biotechnol.* **39**, 665–677 (2021).
14. Raj, K. et al. Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* **6**, 28–32 (2018).
15. Wagner, J. M. & Alper, H. S. Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances. *Fungal Genet. Biol.* **89**, 126–136 (2016).
16. Thorwall, S., Schwartz, C., Chartron, J. W. & Wheeldon, I. Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. *Nat. Chem. Biol.* **16**, 113–121 (2020).
17. Bankar, A. V., Kumar, A. R. & Zinjarde, S. S. Environmental and industrial applications of *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol.* **84**, 847–865 (2009).
18. Cernak, P. et al. Engineering *Kluyveromyces marxianus* as a robust synthetic biology platform host. *MBio* **9**, e01410–e01418 (2018).
19. Sun, W. et al. Metabolic engineering of an acid-tolerant yeast strain *Pichia kudriavzevii* for itaconic acid production. *Metab. Eng. Commun.* **10**, e00124 (2020).
20. Kurtzman, C., Fell, J. W. & Boekhout, T. *The Yeasts: A Taxonomic Study* (Elsevier, 2011).
21. Schwartz, C. et al. Validating genome-wide CRISPR-Cas9 function improves screening in the oleaginous yeast *Yarrowia lipolytica*. *Metab. Eng.* **55**, 102–110 (2019).
22. Gao, J. et al. Synthetic biology toolkit for marker-less integration of multigene pathways into *Pichia pastoris* via CRISPR/Cas9. *ACS Synth. Biol.* **11**, 623–633 (2022).
23. Shen, X.-X. et al. Tempo and mode of genome evolution in the budding yeast subphylum. *Cell* **175**, 1533–1545.e1520 (2018).
24. Bagley, J. et al. Genome sequencing of 15 acid tolerant yeasts. *Microbiol. Resour. Announc.* (In press).
25. Cai, P. et al. Recombination machinery engineering facilitates metabolic engineering of the industrial yeast *Pichia pastoris*. *Nucleic Acids Res.* **49**, 7791–7805 (2021).
26. Gao, J. et al. Enhancing homologous recombination efficiency in *Pichia pastoris* for multiplex genome integration using short homology arms. *ACS Synth. Biol.* **11**, 547–553 (2022).
27. Kurtzman, C. & Smiley, M. Heterothallism in *Pichia kudriavzevii* and *Pichia terricola*. *Antonie Van. Leeuwenhoek* **42**, 355–363 (1976).
28. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
29. Reider Apel, A. et al. A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **45**, 496–508 (2016).
30. Yamanishi, M. et al. A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a “terminatome” toolbox. *ACS Synth. Biol.* **2**, 337–347 (2013).
31. Luttik, M. et al. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. *Metab. Eng.* **10**, 141–153 (2008).
32. Weber, H. E. et al. Requirement of a functional flavin mononucleotide prenyltransferase for the activity of a bacterial decarboxylase in a heterologous muconic acid pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **83**, e03472–03416 (2017).
33. Wang, J., Lin, M., Xu, M. & Yang, S.-T. Anaerobic fermentation for production of carboxylic acids as bulk chemicals from renewable biomass. *Anaerobes Biotechnol.* **156**, 323–361 (2016).
34. Warnecke, T. & Gill, R. T. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. *Microb. Cell Factories* **4**, 1–8 (2005).
35. Joo, J. C. et al. Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chem. Sci.* **8**, 1406–1413 (2017).
36. De Brabander, P. et al. In situ product recovery of bio-based industrial platform chemicals: a guideline to solvent selection. *Ferment* **7**, 26 (2021).
37. Park, H. J. et al. Low-pH production of d-lactic acid using newly isolated acid tolerant yeast *Pichia kudriavzevii* NG7. *Biotechnol. Bioeng.* **115**, 2232–2242 (2018).
38. Xiao, H., Shao, Z., Jiang, Y., Dole, S. & Zhao, H. Exploiting *Issatchenkia orientalis* SD108 for succinic acid production. *Micro Cell Factories* **13**, 1–11 (2014).
39. Arroyo Lopez, F., Duran Quintana, M. & Fernández, A. G. Microbial evolution during storage of seasoned olives prepared with organic acids with potassium sorbate, sodium benzoate, and ozone used as preservatives. *J. Food Prot.* **69**, 1354–1364 (2006).
40. López, F. A., Quintana, M. D. & Fernández, A. G. Modelling of the growth–no growth interface of *Issatchenkia occidentalis*, an olive spoiling yeast, as a function of the culture media, NaCl, citric and sorbic acid concentrations: Study of its inactivation in the no growth region. *Int. J. Food Microbiol.* **117**, 150–159 (2007).
41. Weninger, A. et al. Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. *J. Cell. Biochem.* **119**, 3183–3198 (2018).
42. Carroll, D. & Beumer, K. J. Genome engineering with TALENs and ZFNs: repair pathways and donor design. *Methods* **69**, 137–141 (2014).
43. Liachko, I. & Dunham, M. J. An autonomously replicating sequence for use in a wide range of budding yeasts. *FEMS Yeast Res.* **14**, 364–367 (2014).
44. Obst, U., Lu, T. K. & Sieber, V. A modular toolkit for generating *Pichia pastoris* secretion libraries. *ACS Synth. Biol.* **6**, 1016–1025 (2017).
45. Malci, K. et al. Standardization of synthetic biology tools and assembly methods for *Saccharomyces cerevisiae* and emerging yeast species. *ACS Synth. Biol.* **11**, 2527–2547 (2022).
46. Ryan, O. W. et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. *Elife* **3**, e03703 (2014).
47. Horwitz, A. A. et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Syst.* **1**, 88–96 (2015).
48. Shao, Z., Zhao, H. & Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **37**, e16–e16 (2009).
49. Pyne, M. E. et al. A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nat. Commun.* **11**, 1–10 (2020).
50. Labuhn, M. et al. Refined sgRNA efficacy prediction improves large- and small-scale CRISPR–Cas9 applications. *Nucleic Acids Res.* **46**, 1375–1385 (2017).
51. Pyne, M. E. et al. An engineered Aro1 protein degradation approach for increased *cis*, *cis*-muconic acid biosynthesis in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **84**, e01095–01018 (2018).

## Acknowledgements

We thank Nicholas Gold and the Concordia Genome Foundry for assistance with high-throughput screening. We also thank Marcos DiFalco for assistance with LC-MS analyses. This study was financially supported by an NSERC-Industrial Biocatalysis Network (IBN) grant, an NSERC Discovery grant, and by bp Biosciences Center. M.E.P. was supported by an NSERC Postdoctoral Fellowship, K.K. was supported by a Concordia University Horizon Postdoctoral Fellowship, and L.N. was supported by a FQRNT DE Doctoral Research Scholarship for Foreign Students. V.J.J.M. is supported by a Concordia University Research Chair. M.W. is supported by a Canada Research Chair.

## Author contributions

M.E.P., J.A.B., L.N., K.K., Q.W., and V.J.J.M. designed the research. M.E.P., J.A.B., and L.N. performed the experiments. K.E. and M.D. assisted in preliminary studies and fed-batch cultivation. V.J.J.M. and Q.W. supervised the research. M.E.P., J.A.B., M.W. and V.J.J.M. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-41064-5>.

**Correspondence** and requests for materials should be addressed to Vincent J. J. Martin.

**Peer review information** *Nature Communications* thanks Nuno Mira, Tsutomu Tanaka and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023