This work is on a Creative Commons Attribution 4.0 International (CC BY 4.0) license, https://creativecommons.org/licenses/by/4.0/. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it’s important to you. Thank you.
Supplementary materials for

Refactoring of Ehrlich pathway for high-yield 2-phenylethanol production in *Yarrowia lipolytica*

Yang Gu\textsuperscript{a,b}, Jingbo Ma\textsuperscript{a}, Yonglian Zhu\textsuperscript{b}, Peng Xu\textsuperscript{a,*}

\textsuperscript{a} Department of Chemical, Biochemical and Environmental Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250, USA

\textsuperscript{b} Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China

*Corresponding author.

E-mail addresses: pengxu@umbc.edu (P. Xu)
## Supplementary Table

### Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Names</th>
<th>Characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>po1g</td>
<td>Wild-type strain W29 (ATCC20460) derivate, W29 ΔmatA, Δxpr2-332, Δaxp-2, Δleu2-270, pBR platform</td>
<td>¹</td>
</tr>
<tr>
<td>po1f</td>
<td>po1g derivate, po1g Δura3</td>
<td>¹</td>
</tr>
<tr>
<td>po1fk</td>
<td>po1f derivate, po1f Δku70::loxP</td>
<td>This work</td>
</tr>
<tr>
<td>po1g pYXLP'</td>
<td>po1g derivate, po1g with the empty plasmid pYLXP'</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP1</td>
<td>po1g derivate, overexpression of genes yliARO8 by plasmid pYLXP'-ylARO8</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP2</td>
<td>po1g derivate, overexpression of genes yliARO9 by plasmid pYLXP'-ylARO9</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP3</td>
<td>po1g derivate, overexpression of genes yliARO10 by plasmid pYLXP'-ylARO10</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP4</td>
<td>po1g derivate, overexpression of genes yliARO8 and ylARO10 by plasmid pYLXP'-ylARO10-ylARO8</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP5</td>
<td>po1g derivate, overexpression of genes yliARO9 and ylARO10 by plasmid pYLXP'-ylARO10-ylARO9</td>
<td>This work</td>
</tr>
<tr>
<td>po1gP6</td>
<td>po1g derivate, overexpression of genes C18645 by plasmid pYLXP'-ylARO10-C18645</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-1</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and BAP by plasmid pYLXP'-ylARO10-ylARO8-BAP</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-2</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and phep by plasmid pYLXP'-ylARO10-ylARO8-phep</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-3</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and Gap1 by plasmid pYLXP'-ylARO10-ylARO8-Gap1</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-4</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and GapY1 by plasmid pYLXP'-ylARO10-ylARO8-GapY1</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-5</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and GapY2 by plasmid pYLXP'-ylARO10-ylARO8-GapY2</td>
<td>This work</td>
</tr>
<tr>
<td><strong>P4-6</strong></td>
<td>po1g derivate, overexpression of genes yliARO8, ylARO10 and GapY3 by plasmid pYLXP'-ylARO10-ylARO8-GapY3</td>
<td>This work</td>
</tr>
</tbody>
</table>
plasmid pYLXP'-ylARO10-ylARO8-GapY3

P5-1 po1g derivate, overexpression of genes yliARO9, ylARO10 and BAP by plasmid pYLXP'-ylARO10-ylARO9-BAP

This work

P5-2 po1g derivate, overexpression of genes yliARO9, ylARO10 and phep by plasmid pYLXP'-ylARO10-ylARO9-phep

This work

P5-3 po1g derivate, overexpression of genes yliARO9, ylARO10 and Gap1 by plasmid pYLXP'-ylARO10-ylARO9-Gap1

This work

P5-4 po1g derivate, overexpression of genes yliARO9, ylARO10 and GapY1 by plasmid pYLXP'-ylARO10-ylARO9-GapY1

This work

P5-5 po1g derivate, overexpression of genes yliARO9, ylARO10 and GapY2 by plasmid pYLXP'-ylARO10-ylARO9-GapY2

This work

P5-6 po1g derivate, overexpression of genes yliARO9, ylARO10 and GapY3 by plasmid pYLXP'-ylARO10-ylARO9-GapY3

This work

PAR1 po1g derivate, overexpression of genes PAR1 by plasmid pYLXP'-PAR1

This work

PAR2 po1g derivate, overexpression of genes PAR2 by plasmid pYLXP'-PAR2

This work

PAR3 po1g derivate, overexpression of genes PAR3 by plasmid pYLXP'-PAR3

This work

PAR4 po1g derivate, overexpression of genes PAR4 by plasmid pYLXP'-PAR4

This work

PAR5 po1g derivate, overexpression of genes PAR5 by plasmid pYLXP'-PAR5

This work

PAR6 po1g derivate, overexpression of genes PAR6 by plasmid pYLXP'-PAR6

This work

PAR7 po1g derivate, overexpression of genes PAR7 by plasmid pYLXP'-PAR7

This work

PAR8 po1g derivate, overexpression of genes PAR8 by plasmid pYLXP'-PAR8

This work

PARL po1g derivate, overexpression of genes PARL by plasmid pYLXP'-PARL

This work

ADH1 po1g derivate, overexpression of genes ADH1 by plasmid pYLXP'-ADH1

This work

ADH2 po1g derivate, overexpression of genes ADH2 by plasmid pYLXP'-ADH2

This work

ADH3 po1g derivate, overexpression of genes ADH3 by plasmid pYLXP'-ADH3

This work

ADH4 po1g derivate, overexpression of genes ADH4 by plasmid pYLXP'-ADH4

This work

ADH5 po1g derivate, overexpression of genes ADH5 by plasmid pYLXP'-ADH5

This work

ADH6 po1g derivate, overexpression of genes ADH6 by plasmid pYLXP'-ADH6

This work

ADH7 po1g derivate, overexpression of genes ADH7 by plasmid pYLXP'-ADH7

This work
ADH8 po1g derivate, overexpression of genes ADH8 by plasmid pYLX'p-ADH8 This work

ADH9 po1g derivate, overexpression of genes SFA1 by plasmid pYLX'p-ADH9 This work

po1gP7-1 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and ADH2 by plasmid pYLX'p-ARO10-ARO8-GapY3-ADH2 This work

po1gP7-2 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and ADH3 by plasmid pYLX'p-ARO10-ARO8-GapY3-ADH3 This work

po1gP7-3 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR1 by plasmid pYLX'p-ARO10-ARO8-GapY3-PAR1 This work

po1gP7-4 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR2 by plasmid pYLX'p-ARO10-ARO8-GapY3-PAR2 This work

po1gP7-5 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR3 by plasmid pYLX'p-ARO10-ARO8-GapY3-PAR3 This work

po1gP7-6 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by plasmid pYLX'p-ARO10-ARO8-GapY3-PAR4 This work

po1gP8-1 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and ADH2 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-ADH2 at pBR platform This work

po1gP8-2 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and ADH3 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-ADH3 at pBR platform This work

po1gP8-3 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR1 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-PAR1 at pBR platform This work

po1gP8-4 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR2 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-PAR2 at pBR platform This work

po1gP8-5 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR3 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-PAR3 at pBR platform This work

po1gP8-6 po1g derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-PAR4 at pBR platform This work

po1fk1 po1fk derivate, po1fk ΔALD2 ΔALD3::loxP This work

po1fk1P po1fk1 derivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by integrating linearized pYLX'p-ARO10-ARO8-GapY3-PAR4 at pBR platform This work
This work

po1fk2 deivate, po1fk1 ΔPHA2::loxP

po1fk2P deivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by integrating linearized pYLXP'-ARO10-ARO8-GapY3-PAR4 at pBR platform

po1fk3 deivate, po1fk3 ΔYAL10821846::loxP

po1fk3P deivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by integrating linearized pYLXP'-ARO10-ARO8-GapY3-PAR4 at pBR platform

po1fk4 deivate, overexpression of genes ARO10, ARO8, GapY3, and PAR4 by integrating genes ARO10, ARO8, GapY3, PAR4 at 26s sDNA sites

po1fk4P1 deivate, overexpression of genes citB and scIDP2 by plasmid pYLXP'-citB-scIDP2

po1fk4P2 deivate, overexpression of genes citB and ylIDP2 by plasmid pYLXP'-citB-ylIDP2

po1fk4P3 deivate, overexpression of genes ancA and scIDP2 by plasmid pYLXP'-ancA-scIDP2

po1fk4P4 deivate, overexpression of genes ancA and ylIDP2 by plasmid pYLXP'-ancA-ylIDP2

po1fk4P5 deivate, overexpression of genes ancA, ylIDP2 and ylODC by plasmid pYLXP'-ancA-ylIDP2-ylODC

po1fk5 deivate, po1fk4 ΔDGA1::loxP

po1fk5P deivate, overexpression of genes ancA and ylIDP2 by plasmid pYLXP'-ancA-ylIDP2

po1fk6 deivate, po1fk4 ΔDGA2::loxP

po1fk6P deivate, overexpression of genes ancA and ylIDP2 by plasmid pYLXP'-ancA-ylIDP2

po1fk7 deivate, po1fk5 ΔDGA2::loxP

po1fk7P deivate, overexpression of genes ancA and ylIDP2 by plasmid pYLXP'-ancA-ylIDP2

*******************************************************************************************

Plasmids
pYLP'  YaliBrick plasmid
pYLP'-loxP-ura  pYLP' containing the loxP-URA-loxP cassette
pYLP'-Cre  pYLP' containing gene Cre
pYLP'-loxP-ura-Δku  pYLP'-loxP-ura containing the deletion cassette of gene ku70
pYLP'-loxP-ura-ΔALD2  This work
pYLP'-loxP-ura-ΔALD3  This work
pYLP'-loxP-ura-ΔPH2A  This work
pYLP'-loxP-ura-ΔYA  This work
pYLP'-loxP-ura-ΔDG1A1  This work
pYLP'-loxP-ura-ΔDG2A2  This work
pYLP'-ylARO8  pYLP' containing gene ylARO8
pYLP'-ylARO9  pYLP' containing gene ylARO9
pYLP'-ylARO10  pYLP' containing gene ylARO10
pYLP'-ylARO10-ylAR  pYLP' containing genes ylARO10 and ylARO8
pYLP'-ylARO10-ylAR08  This work
pYLP'-ylARO10-ylAR09  This work
pYLP'-ylARO10-ylAR08-BAP  This work
pYLP'-ylARO10-ylAR08-BAP  This work
pYLP'-ylARO10-ylAR08-BAP  This work
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8 and Gap1

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8 and GapY1

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8 and GapY2

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8 and GapY3

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and BAP

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and phen

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and Gap1

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and GapY1

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and GapY2

This work

pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO9 and GapY3

This work

pYLXP'-PAR1 pYLXP' containing gene PAR1

This work

pYLXP'-PAR2 pYLXP' containing gene PAR2

This work

pYLXP'-PAR3 pYLXP' containing gene PAR3

This work

pYLXP'-PAR4 pYLXP' containing gene PAR4

This work

pYLXP'-PAR5 pYLXP' containing gene PAR5

This work

pYLXP'-PAR6 pYLXP' containing gene PAR6

This work

pYLXP'-PAR7 pYLXP' containing gene PAR7

This work

pYLXP'-PAR8 pYLXP' containing gene PAR8

This work
pYLXP'-PARL pYLXP' containing gene PARL This work
pYLXP'-ADH1 pYLXP' containing gene ADH1 This work
pYLXP'-ADH2 pYLXP' containing gene ADH2 This work
pYLXP'-ADH3 pYLXP' containing gene ADH3 This work
pYLXP'-ADH4 pYLXP' containing gene ADH4 This work
pYLXP'-ADH5 pYLXP' containing gene ADH5 This work
pYLXP'-ADH6 pYLXP' containing gene ADH6 This work
pYLXP'-ADH7 pYLXP' containing gene ADH7 This work
pYLXP'-ADH8 pYLXP' containing gene ADH8 This work
pYLXP'-ADH9 pYLXP' containing gene SFA1 This work
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and ADH2 This work
08-GapY3-ADH2
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and ADH3 This work
08-GapY3-ADH3
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and PAR1 This work
08-GapY3-PAR1
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and PAR2 This work
08-GapY3-PAR2
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and PAR3 This work
08-GapY3-PAR3
pYLXP'-ylARO10-ylAR pYLXP' containing genes ylARO10, ylARO8, GapY3, and PAR4 This work
08-GapY3-PAR4
pYLXP'-loxP-ura-26s pYLXP' containing the 26s rDNA genomic integration cassette of genes This work
DNA-ylARO10-ylAR ylARO10, ylARO8, GapY3, and PAR4
8-GapY3-PAR4
pYLXP'-citB-sclDP2 pYLXP' containing genes citB and sclDP2 This work
pYLXP'-citB-ylIDP2 pYLXP' containing genes citB and ylIDP2 This work
pYLXP'-ancA-sclDP2 pYLXP' containing genes ancA and sclDP2 This work
pYLXP'-ancA-ylIDP2 pYLXP' containing genes ancA and ylIDP2 This work
This work

pYLXP\(^\prime\)-ancA-ylIDP2-y  pYLXP\(^\prime\) containing genes ancA, ylIDP2 and ylODC

IODC
Table 2. Primers used in this study

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ku70_DwF</td>
<td>gtccggagcgccgcGCATGCaagtcgacaCTAGGGAGGCACATCTAAACGAATAACG</td>
</tr>
<tr>
<td>Ku70_DwR</td>
<td>gttacatctttttcatagacacatctcggAGTGACAGGGAAGAAGGAGG</td>
</tr>
<tr>
<td>ku70_UpF</td>
<td>cccataatgtggaagacgtagCGACTTGATTTAGAGTGTCAGCAGTCC</td>
</tr>
<tr>
<td>ku70_UpR</td>
<td>tatgtagctatatcgaagattattTTAAAAAGCGGCAGGTCGTG</td>
</tr>
<tr>
<td>D07062-F</td>
<td>cgaccagcactttttgcaagacacatcaccGACTCATACACACCACCACGTCATACACT</td>
</tr>
<tr>
<td>D07062-R</td>
<td>gacaggccatggaactagtgcgtaccTCTATCTGGAACGACAGCCATTGG</td>
</tr>
<tr>
<td>F24937-F</td>
<td>cgaccagcactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>F24937-R</td>
<td>cggcagcactttttgccagcaactaacccgactCTAGTTCTGGAACGACAGCCATTGG</td>
</tr>
<tr>
<td>F09097-F</td>
<td>ccgaccagcactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>F09097-R</td>
<td>ggacaggccatggaactagtgcgtaccCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>D08844-F</td>
<td>cgaccagcactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>D08844-R</td>
<td>gggcagcactttttgccagcaactaacccgactCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>D12386-F</td>
<td>cgaccagcactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>D12386-R</td>
<td>gacaggccatggaactagtgcgtaccCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>C02051-F</td>
<td>gcagccgactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>C02051-R</td>
<td>gggcagcactttttgccagcaactaacccgactCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>D11616-F</td>
<td>cgaccagcactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>D11616-R</td>
<td>ggacaggccatggaactagtgcgtaccCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>D08778-F</td>
<td>gcagccgactttttgccagcaactaacccgactCCTGCCAGCTTCTCAGTCC</td>
</tr>
<tr>
<td>D08778-R</td>
<td>ggacaggccatggaactagtgcgtaccCTACTCCTCCTCTTTACCTTCCCAACAC</td>
</tr>
<tr>
<td>ALD3_Dw-F</td>
<td>gctagcgagcaataacggaggaAGGCCGTCCACATTAACCTG</td>
</tr>
<tr>
<td>ALD3_Dw-R</td>
<td>gttacatctttttcatagacataCTGCGCAACCACCCTTACCAAA</td>
</tr>
<tr>
<td>ALD3_Up-F</td>
<td>gctagcgagcaataacggaggaAGGCCGTCCACATTAACCTG</td>
</tr>
<tr>
<td>ALD3_Up-R</td>
<td>taatgtagctatacagagttatACTTGCACTAGGTTAGCAGCGAC</td>
</tr>
<tr>
<td>ALD2_Dw-F</td>
<td>gctagcgagcaataacggaggaAGGCCGTCCACATTAACCTG</td>
</tr>
<tr>
<td>ALD2_Dw-R</td>
<td>gttacatctttttcatagacataCTGCGCAACCACCCTTACCAAA</td>
</tr>
<tr>
<td>ALD2_Up-F</td>
<td>gctagcgagcaataacggaggaAGGCCGTCCACATTAACCTG</td>
</tr>
</tbody>
</table>
ALD2_Up-R  taatgtatctatacgaagttatatATATTTAGAGGTTGGAATCTGTTATATGTAAT
ALD2_Cas-F  CCGCTCTCTCAAGTGTCTGAAAGTTGAAT
ALD2_Cas-R  TCTGTTGATTTCTAGGGAACGTGTCT
ALD3_Cas-F  CAAGAAGGATAAAAAATGGAAACTCGGTCT
ALD3_Cas-R  CTGCTGCAACCAGCCTACAAAA
ALD2_DwChk-R  ACTCCTCTCTAGACTCCTCCTGTCT
ALD2_UpChk-F  GTGTCCATCATCACATGACCACAATC
ALD3_DwChk-R  TAGCCTCGTTAATGCACCAGAGTTGAAT
ALD3_UpChk-F  GGGAAATGCTCCATCATGAGTATGGA
ADH1-F  accagcactttttgcagtaaccgcagACACCATCCCCAAAGACACCA
ADH1-R  aggccatgaactagctgtaaccTTACTTGTAAAGTGTCAGAAGCTATCTGCC
ADH2-F  cccagcactttttgcagtaaccgcagTCTGCTCCGGTCAACCTACCA
ADH2-R  ggacagccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
ADH3-F  cagccacattttttgcagtaaccgcagACCACCATCCCCACAGACCC
ADH3-R  ggacagccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
ADH4-F  cccagcactttttttgcagtaaccgcagCACAATCCCCACAGGACCAAGC
ADH4-R  ggacagccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
ADH5-F  cccagcactttttttgcagtaaccgcagAGCGACGTCCCACAGACCA
ADH5-R  ggacagccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
ADH6-F  accagcactttttttgcagtaaccgcagGTCAAAACGTCGACGAACCCTG
ADH6-R  gacagccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
ADH7-F  cccagcactttttttgcagtaaccgcagAGAGCCGCTTTCACACC
ADH7-R  ggccatgaactagctgtaaccTTACTTGGAGGTCGAGGAAACGTATCG
Gap1Y3-F  acttttttcgagctataacccgagAAGCAATGGAAAAACGAACGAC
Gap1Y3-R  acaggctgagctataacccgagATCAACACCATTTGGTACTCG
Gap1Y2-F  cccggttccgagctataacccgagAAAAACGAGTGTATCAACCAACTCT
Gap1Y2-R  cagccagccatgaactagctgtaaccTTAGCAGGAAACTTGGTACTCG
Gap1Y1-F  cagccacattttttttgcagtaaccgcagCAATGGTCTACACCCAGAAGGAGG
Gap1Y1-R  cagccagccatgaactagctgtaaccCTAACCAGAATTTGTATGTTCCATTAAAGGAAG
YliDGA1b_ChkR  GAGATGGGCATGCCAACGTGAC
YliDGA1b_DwF   gctagcgagacaataacggaggaCATAACACTCATCAGTAGCCTTTACAGTGAT
YliDGA1b_DwR   ccttttatcagatacagcggccgcTTGCTCTTTGTAATTCCATAGATAATATAAGGAA
YliDGA1b_UpF   aaatttgatgaaaggccgacgcCTTGGGAGTGTATTTGGAAAATGACTTGG
YliDGA1b_UpR   tgtatgctatacgaagttaATTTGCCGGCGTGACGGTTAC
YliDGA1a_DwF   gctagcgagacaataacggaggaGGAAAACTGCCTGGGTTAGGCAAAT
YliDGA1a_DwR   atcagacatagcggccgcTCTCTGATGGCCTGGAGCGAG
YliDGA1a_UpF   aatttgatgaaaggccgacgcATGCTGCGGGCGGATCGCT
YliDGA1a_UpR   tgtatgctatacgaagttaATTTGCCGGCGTGACGGTTAC
BsuCtiB_F      agcactttttgcagtactaaccgcagtcgtcaaccctacgagaagcca
BsuCtiB_R      gggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
EcoAncA_F      cccagcactttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG
EcoAncA_R      ccatggaactagtcggtacttttcaacatacttagcataagcataatgcaaaaaagcc
ylIDP2-F       cccagcactttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG
ylIDP2_R       gggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
ScIDP2-F       agcactttttgcagtactaaccgcagACAAAGATTAAGGTAGCTAACCCCATTGT
ScIDP2_R       gggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
ARO9_F         cccagcactttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG
ARO9_R         gggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
ARO8_F         cccagcactttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG
ARO8_R         cgtggggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
C18645_F       cgcacagcactttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG
C18645_R       gggacagcagccatgtaacgctgcgtacctcagactgtctttttcaacga
<table>
<thead>
<tr>
<th>Strains</th>
<th>Clone numbers (2PE mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>po1gP1</td>
<td>549.12 487.23 654.21</td>
</tr>
<tr>
<td>po1gP2</td>
<td>512.21 612.12 559.21</td>
</tr>
<tr>
<td>po1gP3</td>
<td>498.21 719.23 871.21</td>
</tr>
<tr>
<td>po1gP4</td>
<td>132.68 42.92 543.55 432.64 371.12 538.25</td>
</tr>
<tr>
<td>po1gP5</td>
<td>491.18 626.17 217.01 575.09 199.42 106.75</td>
</tr>
<tr>
<td>po1gP6</td>
<td>58.43 323.79 305.29 388.76 399.89 315.13</td>
</tr>
<tr>
<td>P4-2</td>
<td>698.12 623.41 773.97 842.26 770.52 625.62</td>
</tr>
<tr>
<td>P4-1</td>
<td>923.19 906.22 921.28 996.46 962.23 992.63</td>
</tr>
<tr>
<td>P4-3</td>
<td>771.19 724.11 890.59 956.47 901.14 917.82</td>
</tr>
<tr>
<td>P4-4</td>
<td>985.15 737.92 733.79 663.11 664.26 890.78</td>
</tr>
<tr>
<td>P4-5</td>
<td>730.15 946.59 1012.96 816.94 815.21 862.20</td>
</tr>
<tr>
<td>P4-6</td>
<td>678.46 840.43 616.79 902.48 790.57 725.55</td>
</tr>
<tr>
<td>P5-2</td>
<td>735.52 781.55 732.93 843.31 750.48 830.17</td>
</tr>
<tr>
<td>P5-1</td>
<td>992.15 900.95 963.66 992.76 998.86 907.37</td>
</tr>
<tr>
<td>P5-3</td>
<td>745.97 702.19 716.91 692.29 721.81 701.65</td>
</tr>
<tr>
<td>P5-4</td>
<td>921.85 943.05 724.78 842.06 717.30 677.69</td>
</tr>
<tr>
<td>P5-5</td>
<td>908.62 668.39 911.21 629.55 900.56 726.22</td>
</tr>
<tr>
<td>P5-6</td>
<td>949.47 1196.80 1179.34 711.45 638.18 681.53</td>
</tr>
<tr>
<td>PAR1</td>
<td>523.77 499.89 397.57 449.35</td>
</tr>
<tr>
<td>PAR2</td>
<td>570.86 429.21 677.02 225.71</td>
</tr>
<tr>
<td>PAR3</td>
<td>526.74 586.87 424.32 568.46</td>
</tr>
<tr>
<td>PAR4</td>
<td>780.40 154.27 830.46 631.56</td>
</tr>
<tr>
<td>PAR5</td>
<td>426.24 399.20 376.76 217.27</td>
</tr>
<tr>
<td>PAR6</td>
<td>365.15 417.90 518.30 475.34</td>
</tr>
<tr>
<td>PAR7</td>
<td>371.96 338.97 644.89 398.05</td>
</tr>
<tr>
<td>PAR8</td>
<td>477.16 427.39 567.60 494.52</td>
</tr>
<tr>
<td></td>
<td>PARL</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>411.57</td>
</tr>
<tr>
<td></td>
<td>281.05</td>
</tr>
<tr>
<td></td>
<td>353.74</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>po1gP7-3</td>
<td>366.88</td>
</tr>
<tr>
<td>po1gP7-5</td>
<td>366.88</td>
</tr>
<tr>
<td>po1gP7-6</td>
<td>366.88</td>
</tr>
<tr>
<td>po1gP7-1</td>
<td>919.07</td>
</tr>
<tr>
<td>po1gP7-2</td>
<td>919.07</td>
</tr>
<tr>
<td>po1gP7-4</td>
<td>919.07</td>
</tr>
<tr>
<td>po1gP7-4</td>
<td>919.07</td>
</tr>
<tr>
<td>po1gP8-1</td>
<td>812.14</td>
</tr>
<tr>
<td>po1gP8-2</td>
<td>812.14</td>
</tr>
<tr>
<td>po1gP8-3</td>
<td>812.14</td>
</tr>
<tr>
<td></td>
<td>713.56</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>711.93</td>
</tr>
<tr>
<td>po1gP8-4</td>
<td>660.52</td>
</tr>
<tr>
<td></td>
<td>864.31</td>
</tr>
<tr>
<td></td>
<td>603.37</td>
</tr>
<tr>
<td>po1gP8-5</td>
<td>1067.33</td>
</tr>
<tr>
<td></td>
<td>820.68</td>
</tr>
<tr>
<td></td>
<td>269.25</td>
</tr>
<tr>
<td>po1gP8-6</td>
<td>725.74</td>
</tr>
<tr>
<td></td>
<td>340.31</td>
</tr>
<tr>
<td></td>
<td>891.07</td>
</tr>
<tr>
<td>po1fk1P</td>
<td>621.40</td>
</tr>
<tr>
<td></td>
<td>770.71</td>
</tr>
<tr>
<td>po1fk2P</td>
<td>45.80</td>
</tr>
<tr>
<td></td>
<td>288.62</td>
</tr>
<tr>
<td>po1fk3P</td>
<td>149.38</td>
</tr>
<tr>
<td></td>
<td>229.84</td>
</tr>
<tr>
<td>po1fk4P1</td>
<td>1023.41</td>
</tr>
<tr>
<td>po1fk4P2</td>
<td>1008.45</td>
</tr>
<tr>
<td>po1fk4P3</td>
<td>981.41</td>
</tr>
<tr>
<td>po1fk4P4</td>
<td>1092.36</td>
</tr>
<tr>
<td>po1fk4P5</td>
<td>568.75</td>
</tr>
<tr>
<td>po1fk5P</td>
<td>1595.07</td>
</tr>
<tr>
<td>po1fk6P</td>
<td>1081.05</td>
</tr>
<tr>
<td>po1fk7P</td>
<td>797.95</td>
</tr>
</tbody>
</table>
Supplementary Note

Stoichiometric model to assess Ehrlich pathway efficiency

To systematically assess the Ehrlich pathway, we analyzed the stoichiometrics of 2-PE biosynthesis using L-phenylalanine as substrate, listed below:

\[
\text{L-Phenylalanine + 2-Oxoglutarate + NADH} \rightarrow \text{2-PE + Glutamate + NAD}^+ + \text{CO}_2 \tag{1}
\]

Here, 2-oxoglutarate is synthesized by the central carbon metabolism with the glucose as substrate (Fig 1b):

\[
\text{Glucose + 2ADP} + 2NAD^+ \rightarrow 2-\text{Oxoglutarate + 2NADH} + 2\text{ATP} + \text{CO}_2 \tag{2}
\]

Thus, the overall stoichiometry of 2-PE biosynthesis is:

\[
\text{L-Phenylalanine + Glucose + 2ADP} + \text{NAD}^+ \rightarrow \text{2-PE + Glutamate + NADH} + 2\text{ATP} + 2\text{CO}_2 \tag{3}
\]

On the basis of the stoichiometry, producing 1 mol 2-PE will consume 1 mol of L-phenylalanine and glucose, with the formation of 1 mol cytosolic NADH, 1 mol of glutamate, 2 mol of cytosolic ATP and 2 mol of CO\textsubscript{2}. This overall stoichiometrics suggests that NADH is not the limiting factor of 2-PE production.

Next, to rationally predict the engineering targets, a global stoichiometric model was established. Herein, for production of 1 mol of 2-PE, x mol of L-phe and y mol of glucose will be consumed. The final outcome of nitrogen metabolism in microbes should be some proteins or metabolisms with N elements. Herein, we assumed that the final outcome of nitrogen metabolism is proteins and the formula is (RC\textsubscript{2}H\textsubscript{4}O\textsubscript{2}N)\textsubscript{n}, the general linear formula of proteins (R represents the side chain). For the convenience of the following deduction, we further assumed that the R group is H.

Thus, the overall stoichiometrics will be

\[
x \text{ L-Phe (C}_9\text{H}_{11}\text{NO}_2) + y \text{ Glucose (C}_6\text{H}_{12}\text{O}_6) + (8.5x+6y-10) \text{ O}_2 \rightarrow \text{2-PE (C}_8\text{H}_{10}\text{O}) + x \text{ Proteins (C}_2\text{H}_5\text{O}_2\text{N})_n + (7x + 6y - 8) \text{ CO}_2 + (3x + 6y - 5) \text{ H}_2\text{O} \tag{4}
\]

The yield of 2-PE (g/g L-phenylalanine) is

\[
Y_{2-\text{PE}} = \frac{122}{165} \tag{5}
\]

Furthermore, introduction of the respiratory quotient (RQ):

\[
\text{RQ} = \frac{7x + 6y - 8}{8.5x + 6y - 10} \tag{6}
\]

Solve the above equation (6) and substitute x into equation (5), we will get a general yield which depends on both RQ and the amount of glucose (y) consumed.

\[
Y_{2-\text{PE}} = \frac{122}{165} * \frac{8.5RQ - 7}{(6 - 6RQ)y + 10RQ - 8} \tag{7}
\]

However, the above mathematical model with three unknown variables is obscure and esoteric. Thus, we decided to reduce the degree of freedom of the above model and divided the yield into two parts, namely L-phe consumption and
2-oxoglutarate supplementation.

To deduce the boundary of above model (Eqn.7), we assumed that production of 1 mol of 2-PE will only require m mol of L-phe without glucose (y=0), which means L-phe is used as both N and C sources for cell maintenance and product production. Thus, the metabolic model was

\[ m \text{ L-Phenylalanine (C}_{9}\text{H}_{11}\text{NO}_{2}) + (8.5m-10) \text{O}_{2} \rightarrow 2\text{-PE (C}_{8}\text{H}_{10}\text{O}) + m \text{ proteins (C}_{2}\text{H}_{5}\text{O}_{2}\text{N})_n + (7m-8) \text{ CO}_{2} + (3m-5) \text{ H}_{2}\text{O} \] (8)

As a result, the yield of 2-PE (g/g L-phenylalanine) and the respiratory quotient (RQ) were \( Y_{2\text{-PE}} = \frac{122}{165m} \) and \( RQ = \frac{7m-8}{8.5m-10} \) respectively. And, the yield of 2-PE could be solved as

\[ Y_{2\text{-PE}} = \frac{122}{165} * \frac{8.5RQ - 7}{10RQ - 8} \] (9)

As Shown in the metabolic model (Fig. S1), \( Y_{2\text{-PE}} \) increases as RQ increases, whereas RQ increases as L-phe consumption (m) decreases. High RQ value represents less L-phe is oxidized to maintain cell metabolism, and more L-phe is used to synthesize 2-PE through the Ehrlich pathway. In addition, the stoichiometrics (Eqn. 9) suggest that the theoretically maximum \( Y_{2\text{-PE}} \) is 0.739 g/g L-phenylalanine. On the other hand, to account for the cofactor aKG, we assumed that RQ value is a certain value (β) and production of 1 mol of 2-PE will require m mol of aKG. Thus, the metabolic model was

\[ (\frac{8 + \beta a - 5m}{7}) \text{L-Phenylalanine (C}_{9}\text{H}_{11}\text{NO}_{2}) + m \text{ 2-oxoglutarate (C}_{4}\text{H}_{6}\text{O}_{5}) + a\text{O}_{2} \rightarrow 2\text{-PE (C}_{8}\text{H}_{10}\text{O}) + (\frac{8 + \beta a - 5m}{7}) \text{ proteins} \]

\[ \text{(C}_{2}\text{H}_{5}\text{O}_{2}\text{N})_n + \beta a \text{ CO}_{2} + (\frac{3\beta a + 6m - 11}{7}) \text{ H}_{2}\text{O} \] (10)

According to mass balance of element O, ‘a’ could be further simplified, which is \( \frac{29m + 4}{17\beta - 7} \). As a result, the yield of 2-PE \( (g/g \text{ L-phenylalanine}) \) is

\[ Y_{2\text{-PE}} = \frac{122}{165} \frac{8.5RQ - 7}{10RQ - 8} \frac{29m + 4}{17\beta - 7} \] (11)

If \( \beta = 1 \), \( Y_{2\text{-PE}} = \frac{122}{165} \frac{8.5RQ - 7}{10RQ - 8} \frac{30}{17} \) (12);

If \( \beta = 2 \), \( Y_{2\text{-PE}} = \frac{122}{165} \frac{8.5RQ - 7}{10RQ - 8} \frac{62}{17} \) (13);

If \( \beta = 10 \), \( Y_{2\text{-PE}} = \frac{122}{165} \frac{8.5RQ - 7}{10RQ - 8} \frac{1344}{1141} \) (14);

Analysis of the above equation (Fig. S1) indicates that increasing aKG supplementation will significantly improve 2PE yield \( Y_{2\text{-PE}} \) under all different \( \beta \) values. Thus, as suggested by the mathematical model, 2-PE yield is more sensitive to the supply of aKG. In the following section, we will focus our work to improve the pathway selectivity toward L-phe and enhance the availability of aKG.
Fig. S1. Stoichiometric models reveal that 2-PE yield is driven by selectivity of Ehrlich pathway and supply of 2-oxoglutarate (aKG).
Supplementary Figure

Fig. S2. 2-PE titers of screening L-phenylalanine specific permeases with the addition of final concentration of 4g/L L-phenylalanine into CSM fermentation medium (48 h) in shake cultivations.

Fig. S3. 2-PE titers of strains po1gP7-1, po1gP7-2, po1gP7-3, po1gP7-4, po1gP7-5, and po1gP7-6 in shake cultivations.
References
