Supplementary materials for

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

Yang Gu[a,b], Jingbo Ma[a], Yonglian Zhu[b], Peng Xu[a,*]

[a] Department of Chemical, Biochemical and Environmental Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250, USA

[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)

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

| Names | Characteristics | Reference |
|---|---|---|
| **Strains** | | |
| po1g | Wild-type strain W29 (ATCC20460) derivate, W29 *ΔmatA*, *Δxpr*2-332, *Δaxp*-2, *Δleu*2-270, pBR platform | [1] |
| po1f | po1g derivate, po1g *Δura*3 | [1] |
| **po1fk** | po1f derivate, po1f *Δku70::loxP* | This work |
| **po1g pYXLP'** | po1g derivate, po1g with the empty plasmid pYXLP' | This work |
| **po1gP1** | po1g derivate, overexpression of genes *yliARO8* by plasmid pYLXP'-*ylARO8* | This work |
| **po1gP2** | po1g derivate, overexpression of genes *yliARO9* by plasmid pYLXP'-*ylARO9* | This work |
| **po1gP3** | po1g derivate, overexpression of genes *yliARO10* by plasmid pYLXP'-*ylARO10* | This work |
| **po1gP4** | po1g derivate, overexpression of genes *yliARO8* and *ylARO10* by plasmid pYLXP'-*ylARO10-ylARO8* | This work |
| **po1gP5** | po1g derivate, overexpression of genes *yliARO9* and *ylARO10* by plasmid pYLXP'-*ylARO10-ylARO9* | This work |
| **po1gP6** | po1g derivate, overexpression of genes *C18645* by plasmid pYLXP'-*ylARO10-C18645* | This work |
| **P4-1** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *BAP* by plasmid pYLXP'-*ylARO10-ylARO8-BAP* | This work |
| **P4-2** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *phep* by plasmid pYLXP'-*ylARO10-ylARO8-phep* | This work |
| **P4-3** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *Gap1* by plasmid pYLXP'-*ylARO10-ylARO8-Gap1* | This work |
| **P4-4** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *GapY1* by plasmid pYLXP'-*ylARO10-ylARO8-GapY1* | This work |
| **P4-5** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *GapY2* by plasmid pYLXP'-*ylARO10-ylARO8-GapY2* | This work |
| **P4-6** | po1g derivate, overexpression of genes *yliARO8, ylARO10* and *GapY3* by | This work |

| | | |
|---|---|---|
| | 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 pYLXP'-*ADH8* | This work |
| **ADH9** | po1g derivate, overexpression of genes *SFA1* by plasmid pYLXP'-*ADH9* | This work |
| **po1gP7-1** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and ADH2* by plasmid pYLXP'-*ARO10-ARO8-GapY3-ADH2* | This work |
| **po1gP7-2** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and ADH3* by plasmid pYLXP'-*ARO10-ARO8-GapY3-ADH3* | This work |
| **po1gP7-3** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR1* by plasmid pYLXP'-*ARO10-ARO8-GapY3-PAR1* | This work |
| **po1gP7-4** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR2* by plasmid pYLXP'-*ARO10-ARO8-GapY3-PAR2* | This work |
| **po1gP7-5** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR3* by plasmid pYLXP'-*ARO10-ARO8-GapY3-PAR3* | This work |
| **po1gP7-6** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR4* by plasmid pYLXP'-*ARO10-ARO8-GapY3-PAR4* | This work |
| **po1gP8-1** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and ADH2* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-ADH2* at pBR platform | This work |
| **po1gP8-2** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and ADH3* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-ADH3* at pBR platform | This work |
| **po1gP8-3** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR1* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-PAR1* at pBR platform | This work |
| **po1gP8-4** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR2* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-PAR2* at pBR platform | This work |
| **po1gP8-5** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR3* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-PAR3* at pBR platform | This work |
| **po1gP8-6** | po1g derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR4* by integrating linearized pYLXP'-*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 pYLXP'-*ARO10-ARO8-GapY3-PAR4*at pBR platform | This work |

| | | |
|---|---|---|
| **po1fk2** | po1fk1 derivate, po1fk1 *ΔPHA2::loxP* | This work |
| **po1fk2P** | po1fk2 derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR4* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-PAR4*at pBR platform | This work |
| **po1fk3** | po1fk3 derivate, po1fk3 *ΔYALI0B21846g::loxP* | This work |
| **po1fk3P** | po1fk4 derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR4* by integrating linearized pYLXP'-*ARO10-ARO8-GapY3-PAR4*at pBR platform | This work |
| **po1fk4** | po1fk3 derivate, overexpression of genes *ARO10, ARO8, GapY3, and PAR4* by integrating genes *ARO10, ARO8, GapY3, PAR4* at 26s sDNA sites | This work |
| **po1fk4P1** | po1fk4 derivate, overexpression of genes *citB and scIDP2* by plasmid pYLXP'-*citB-scIDP2* | This work |
| **po1fk4P2** | po1fk4 derivate, overexpression of genes *citB and ylIDP2* by plasmid pYLXP'-*citB-ylIDP2* | This work |
| **po1fk4P3** | po1fk4 derivate, overexpression of genes *ancA and scIDP2* by plasmid pYLXP'-*ancA-*sc*IDP2* | This work |
| **po1fk4P4** | po1fk4 derivate, overexpression of genes *ancA and ylIDP2* by plasmid pYLXP'-*ancA-ylIDP2* | This work |
| **po1fk4P5** | po1fk4 derivate, overexpression of genes *ancA, ylIDP2 and ylODC* by plasmid pYLXP'-*ancA-ylIDP2-ylODC* | This work |
| **po1fk5** | po1fk4 derivate, po1fk4 *ΔDGA1::loxP* | This work |
| **po1fk5P** | po1fk5 derivate, overexpression of genes *ancA and ylIDP2* by plasmid pYLXP'-*ancA-ylIDP2* | This work |
| **po1fk6** | po1fk4 derivate, po1fk4 *ΔDGA2::loxP* | This work |
| **po1fk6P** | po1fk6 derivate, overexpression of genes *ancA and ylIDP2* by plasmid pYLXP'-*ancA-ylIDP2* | This work |
| **po1fk7** | po1fk5 derivate, po1fk5 *ΔDGA2::loxP* | This work |
| **po1fk7P** | po1fk7 derivate, overexpression of genes *ancA and ylIDP2* by plasmid pYLXP'-*ancA-ylIDP2* | This work |

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Plasmids**

| | | |
|---|---|---|
| pYLXP' | YaliBrick plasmid | |
| pYLXP'-*loxP-ura* | pYLXP' containing the *loxP-URA-loxP* cassette | [2] |
| pYLXP'-*Cre* | pYLXP' containing gene *Cre* | [2] |
| pYLXP'-*loxP-ura-Δku70* | pYLXP'-loxP-ura containing the deletion cassette of gene *ku70* | This work |
| pYLXP'-*loxP-ura-ΔALD2* | pYLXP'-loxP-ura containing the deletion cassette of gene *ALD2* | This work |
| pYLXP'-*loxP-ura-ΔALD3* | pYLXP'-loxP-ura containing the deletion cassette of gene *ALD3* | This work |
| pYLXP'-*loxP-ura-ΔPHA2* | pYLXP'-loxP-ura containing the deletion cassette of gene *PHA2* | This work |
| pYLXP'-*loxP-ura-ΔYALI0B21846g* | pYLXP'-loxP-ura containing the deletion cassette of gene *YALI0B21846g* | This work |
| pYLXP'-*loxP-ura-ΔDGA1* | pYLXP'-loxP-ura containing the deletion cassette of gene *DGA1* | This work |
| pYLXP'-*loxP-ura-ΔDGA2* | pYLXP'-loxP-ura containing the deletion cassette of gene *DGA2* | This work |
| pYLXP'-*ylARO8* | pYLXP' containing gene *ylARO8* | This work |
| pYLXP'-*ylARO9* | pYLXP' containing gene *ylARO9* | This work |
| pYLXP'-*ylARO10* | pYLXP' containing gene *ylARO10* | This work |
| pYLXP'-*ylARO10-ylARO8* | pYLXP' containing genes *ylARO10 and ylARO8* | This work |
| pYLXP'-*ylARO10-ylARO9* | pYLXP' containing genes *ylARO10 and ylARO9* | This work |
| pYLXP'-*ylARO10-C18645* | pYLXP' containing genes *ylARO10 and C18645* | This work |
| pYLXP'-*ylARO10-ylARO8-BAP* | pYLXP' containing genes *ylARO10, ylARO8 and BAP* | This work |
| pYLXP'-*ylARO10-ylAR* | pYLXP' containing genes *ylARO10, ylARO8 and phep* | This work |

*O8-phep*

| | | |
|---|---|---|
| pYLXP'-*ylARO10-ylARO8-Gap1* | pYLXP' containing genes *ylARO10, ylARO8 and Gap1* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY1* | pYLXP' containing genes *ylARO10, ylARO8 and GapY1* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY2* | pYLXP' containing genes *ylARO10, ylARO8 and GapY2* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3* | pYLXP' containing genes *ylARO10, ylARO8 and GapY3* | This work |
| pYLXP'-*ylARO10-ylARO9-BAP* | pYLXP' containing genes *ylARO10, ylARO9 and BAP* | This work |
| pYLXP'-*ylARO10-ylARO9-phep* | pYLXP' containing genes *ylARO10, ylARO0 and phep* | This work |
| pYLXP'-*ylARO10-ylARO9-Gap1* | pYLXP' containing genes *ylARO10, ylARO9 and Gap1* | This work |
| pYLXP'-*ylARO10-ylARO9-GapY1* | pYLXP' containing genes *ylARO10, ylARO9 and GapY1* | This work |
| pYLXP'-*ylARO10-ylARO9-GapY2* | pYLXP' containing genes *ylARO10, ylARO9 and GapY2* | This work |
| pYLXP'-*ylARO10-ylARO9-GapY3* | 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-ylARO8-GapY3-ADH2* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *ADH2* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3-ADH3* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *ADH3* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3-PAR1* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *PAR1* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3-PAR2* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *PAR2* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3-PAR3* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *PAR3* | This work |
| pYLXP'-*ylARO10-ylARO8-GapY3-PAR4* | pYLXP' containing genes *ylARO10, ylARO8, GapY3,* and *PAR4* | This work |
| pYLXP'-*loxP-ura-26s DNA-ylARO10-ylARO8-GapY3-PAR4* | pYLXP' containing the 26s rDNA genomic integration cassette of genes *ylARO10, ylARO8, GapY3,* and *PAR4* | This work |
| pYLXP'-*citB-scIDP2* | pYLXP' containing genes *citB* and *scIDP2* | This work |
| pYLXP'-*citB-ylIDP2* | pYLXP' containing genes *citB* and *ylIDP2* | This work |
| pYLXP'-*ancA-scIDP2* | pYLXP' containing genes *ancA* and *scIDP2* | This work |
| pYLXP'-*ancA-ylIDP2* | pYLXP' containing genes *ancA* and *ylIDP2* | This work |

pYLXP'-*ancA-ylIDP2-y*    pYLXP' containing genes *ancA*, *ylIDP2* and *ylODC*    This work

*lODC*

**Table 2. Primers used in this study**

| Primers | Sequence |
| --- | --- |
| Ku70_DwF | gtccggagcggccgcGCATGCaagtcgacaCTAGGGAGGCACATCTAAACGAATAACG |
| Ku70_DwR | gttacatcctttatcagacatacctaggAGTGAACGACCAAGACTAAAGGGTG |
| ku70_UpF | ccctaaatttgatgaaagcctaggCGACTTGATGTTTAGAGTGTCCAGATCC |
| ku70_UpR | taatgtatgctatacgaagttatTTTCAAAAAGCGGCGGTTCGTG |
| D07062-F | cgaccagcactttttgcagtactaaccgcagTCTACTACACCCACCACTCTCATCACT |
| D07062-R | gacaggccatggaactagtcggtaccTTAGTTCTGGAACGACAGCCATTGG |
| F24937-F | cgaccagcactttttgcagtactaaccgcagCCTACCACCCTTGTTACTGGAGCTAC |
| F24937-R | cgtggggacaggccatggaactagtcggtaccCTACTGAGACTTAATCCACTGCTCAGC |
| F09097-F | ccgaccagcactttttgcagtactaaccgcagTCCATTCTTGTCACCGGCGC |
| F09097-R | ggacaggccatggaactagtcggtaccCTACTCCTCCTTCTTAACTTCCAACCAC |
| D08844-F | cgaccagcactttttgcagtactaaccgcagCTCGTTCTTCTCACCGGCG |
| D08844-R | gggacaggccatggaactagtcggtaccTTACTCGAGAGTATTGAGCTGCTTGGC |
| D12386-F | cgaccagcactttttgcagtactaaccgcagTCTACAGTTCTCGTCACTGGAGCTAC |
| D12386-R | gacaggccatggaactagtcggtaccCTAAATATTCCACCAATGCTTCTTCTTCTTCTC |
| C20251-F | gaccagcactttttgcagtactaaccgcagCCCACTACACTCGTCACTGGAG |
| C20251-R | gggacaggccatggaactagtcggtaccCTACTTATGCTGCTCGTACCACTGG |
| D11616-F | cgaccagcactttttgcagtactaaccgcagCACCAACTTCTCTCGTCACCG |
| D11616-R | ggacaggccatggaactagtcggtaccCTAGTCGTTCCGCTCGTAAGAAAG |
| D08778-F | gaccagcactttttgcagtactaaccgcagCCCACTACTCTTGTTACCGGAGC |
| D08778-R | ggacaggccatggaactagtcggtaccTCACTGCTTGATCTCGAGCCATTG |
| ALD3_Dw-F | gctagcgagacaataacggaggaAGGCCGTCCACATTAACCTGG |
| ALD3_Dw-R | gttacatcctttatcagacataCTGCTGCAACCAGCCCTACAAA |
| ALD3_Up-F | gcatccctaaatttgatgaaagCAAGAAGGGATAAAAATGGAAACTCGGTCT |
| ALD3_Up-R | taatgtatgctatacgaagttatACTTGCACTAGGTTAGCAGCGAC |
| ALD2_Dw-F | gctagcgagacaataacggaggaACGATGAGCGAACGAATCGTCT |
| ALD2_Dw-R | gttacatcctttatcagacataTCTGTTGGATTCTAGGGAACTGTTTCTG |
| ALD2_Up-F | ggcatccctaaatttgatgaaagCCGCTCTCAAGTGTCTGAAAGTTGAAT |

| | |
|---|---|
| ALD2_Up-R | taatgtatgctatacgaagttatATATTTAGAGTTCGGGATAAAGTTCAATGT |
| ALD2_Cas-F | CCGCTCTCAAGTGTCTGAAAGTTGAAT |
| ALD2_Cas-R | TCTGTTGGATTCTAGGGAACTGTTTCTG |
| ALD3_Cas-F | CAAGAAGGGATAAAAATGGAAACTCGGTCT |
| ALD3_Cas-R | CTGCTGCAACCAGCCCTACAAA |
| ALD2_DwChk-R | ACTCCTCTCTAGACTCCTCCTGTTC |
| ALD2_UpChk-F | GTGTCTCCATCACATGACCACAATC |
| ALD3_DwChk-R | TAGCCTCGTTAATGCACCGAGT |
| ALD3_UpChk-F | GGGAATGCTCCATTGAGATGATGGA |
| ADH1-F | accagcactttttgcagtactaaccgcagACCACCATCCCCAAGACCCA |
| ADH1-R | aggccatggaactagtcggtaccTTACTTGTAAGTGTCCAGAACGTATCTGCC |
| ADH2-F | ccagcactttttgcagtactaaccgcagTCTGCTCCCGTCATCCCCA |
| ADH2-R | gggacaggccatggaactagtcggtaccTTACTTGGAGGTGTCCAGAACGTATCG |
| ADH3-F | cagcactttttgcagtactaaccgcagACCACCATCCCCAAGACCC |
| ADH3-R | ggacaggccatggaactagtcggtaccTTACTTGGAGCAGTCCAGAACGTATC |
| ADH4-F | ccagcactttttgcagtactaaccgcagACAATCCCCAAGACCCAGAAAGC |
| ADH4-R | ggacaggccatggaactagtcggtaccTTACTTGCTGGTATCGACAACATACCGA |
| ADH5-F | ccagcactttttgcagtactaaccgcagAGCGACGTTCCCAAGACAC |
| ADH5-R | ggacaggccatggaactagtcggtaccCTACTTGCTGTTATCAACAACGACACGA |
| ADH6-F | accagcactttttgcagtactaaccgcagGTCAAAACGTCGACGAACCTTGC |
| ADH6-R | gacaggccatggaactagtcggtaccTTAGTACTCCCAAGCGTCCTCAAGAAT |
| ADH7-F | ccagcactttttgcagtactaaccgcagAGAGCCGCTTTCACCACC |
| ADH7-R | ggccatggaactagtcggtaccTTACTGCTTGAACTCCTTATCAAGAGTATCATCC |
| Gap1Y3-F | cactttttgcagtactaaccgcagAACGACATTGAAAAAAACCAAGACGC |
| Gap1Y3-R | acaggccatggaactagtcggtaccCTAACACCAGGCCTTGTAGACTCG |
| Gap1Y2-F | cactttttgcagtactaaccgcagGAAAAAAACGAAGTGGTATCAACCAACTCT |
| Gap1Y2-R | caggccatggaactagtcggtaccTTAGCACCAGAACTTGTAAACTCGGTAGT |
| Gap1Y1-F | cagcactttttgcagtactaaccgcagCAAATGCTCAACCGAACAAGGAAGG |
| Gap1Y1-R | caggccatggaactagtcggtaccCTAACACCAGAAATTGTATGTTCGATAAAGGAAG |

| Gap1p-F | cactttttgcagtactaaccgcagAGTAATACTTCTTCGTACGAGAAGAATAATCCAGA |
| Gap1p-R | caggccatggaactagtcggtaccTTAACACCAGAAATTCCAGATTCTATACCATCTT |
| pheP-F | gcactttttgcagtactaaccgcagaaaaacgcgtcaaccgtatcgg |
| pheP_R | gacaggccatggaactagtcggtaccttatttccgacgcagcgttttaaatgc |
| PHA2_DwChkR | AGAATACCTTGATTCTGGCCACCG |
| PHA2_UpChkF | TCTGGCCGAGTTCAAGCTCCA |
| PHA2_CasF | TACCATCACCGACCCAGAGACCA |
| PHA2_CasR | TGCCGCATGCCATTCAGCTA |
| PHA2_DwF | gctagcgagacaataacggaggaCTTAGACGGTTCAGCGTTTCTGT |
| PHA2_DwR | tacatccttttatcagacataTGCCGCATGCCATTCAGCTA |
| PHA2_UpF | atccctaaatttgatgaaagTACCATCACCGACCCAGAGACCA |
| PHA2_UpR | atgtatgctatacgaagttatGATGTGTAATGTGTGTGATCAAGTGTGC |
| ScBAP2_F | cagcactttttgcagtactaaccgcagCTATCTTCAGAAGATTTTGGATCTTCTGGGA |
| ScBAP2_R | caggccatggaactagtcggtaccTTAACACCAGAAATGATAAGCTTTTCTCATCAAAG |
| YliB21846_CheckF | GCTACTTGTTGTAGACGTAGACATACACT |
| YliB21846_CheckR | TATTCGCACGACACTGACATTTAAGGC |
| YliB21846_CasF | CCGTACGTAAGAGACTGCCATAAGT |
| YliB21846_CasR | CTTCCCCTCCTGCTCACCCC |
| YliB21846_Dw_F | cgagacaataacggaggaTCATCTCTAGAGACGAGGCGTGC |
| YliB21846_Dw_R | agcttgcctatgttacatccttttatcagacataCTTCCCCTCCTGCTCACCCC |
| YliB21846_Up_F | ccctaaatttgatgaaagCCGTACGTAAGAGACTGCCATAAGT |
| YliB21846_Up_R | atgctatacgaagttatGTTGTTGGTGGTTATTTGTTTGTGTGTCA |
| yliODC_F | actttttgcagtactaaccgcagTCTGACCAGAAGCCCCTCCCCTTCATCTACCAATTC |
| yliODC_R | caggccatggaactagtcggtaccTTATTTCTTACCGTCGTGGATTCCTCG |
| yliGDH_F | cactttttgcagtactaaccgcaggacgccattgacgtcaagggcttc |
| yliGDH_R | aggccatggaactagtcggtaccctattgcttcttagtcttctcagccatgt |
| YliDGA1a_ChkF | GTTTATGCATTCTGTTGGACCTTAGTCTG |
| YliDGA1a_ChkR | GTTATCTACCACGATTTTTTGGTTTCTGAGGC |
| YliDGA1b_ChkF | CTATCGCCCCAAAGTGTTTCTAGCA |

| | |
|---|---|
| YliDGA1b_ChkR | GAGATGGCATGCCAACGTTGAC |
| YliDGA1b_DwF | gctagcgagacaataacggaggaCATAACACTCATCAGTAGCCTTTACAGTGAT |
| YliDGA1b_DwR | ccttttatcagacatagcggccgcTTGCTCTTGTAATTCCATAGATAATATATACGAAA |
| YliDGA1b_UpF | aaatttgatgaaaggcggccgcCTTGGGAGTGTATTTGGAAAATGACTTGG |
| YliDGA1b_UpR | tgtatgctatacgaagttatTTTGCGGGCGGTACGGGTACA |
| YliDGA1a_DwF | gctagcgagacaataacggaggaGGAAAACTGCCTGGGTTAGGCAAAT |
| YliDGA1a_DwR | atcagacatagcggccgcTCTCTGATGGCCTGGAGCGAG |
| YliDGA1a_UpF | aatttgatgaaaggcggccgcATGCTGCGGGCGGATCCTGG |
| YliDGA1a_UpR | tgtatgctatacgaagttatAGCTTTTGTTTTGTGTGACTTGTCTGT |
| BsuCtiB_F | agcacttttttgcagtactaaccgcaggcaaacgagcaaaaaactgcagca |
| BsuCtiB_R | gggacaggccatggaactagtcggtacctcaggactgcttcattttttcacga |
| EcoAncA_F | ccagcacttttttgcagtactaaccgcagtcgtcaaccctacgagaagcca |
| EcoAncA_R | ccatggaactagtcggtaccttacttcaacatattacgaatgacataatgcaaaatgcc |
| ylIDP2-F | ccagcacttttttgcagtactaaccgcagTCCACCACCGCTACTCGAGG |
| ylIDP2_R | gggacaggccatggaactagtcggtaccCTAAGCCAGGTCCTTCTTCAGTCTG |
| ScIDP2-F | agcacttttttgcagtactaaccgcagACAAAGATTAAGGTAGCTAACCCCATTGT |
| ScIDP2_R | ggacaggccatggaactagtcggtaccTTACAATGCAGCTGCCTCGAACT |
| ARO9_F | ccagcacttttttgcagtactaaccgcagtccttccacgacaagttcatttctgac |
| ARO9_R | gggacaggccatggaactagtcggtaccctacttctgcttaatctcaccctcctc |
| ARO8_F | cagcacttttttgcagtactaaccgcagacctccgacaacaagcctct |
| ARO8_R | cgtggggacaggccatggaactagtcggtaccttaaagctggaactctcgcttgacga |
| C18645_F | cgaccagcacttttttgcagtactaaccgcagttcaactcagtcaacaaggaggaggacca |
| C18645_R | gggacaggccatggaactagtcggtacctcaactaccgttcttcttctgcatcttc |

*Table 3. Screening of high-performance strains*

| Strains | Clone numbers (2PE mg/L) | | | | | |
|---|---|---|---|---|---|---|
| po1gP1 | 549.12 | 487.23 | 654.21 | - | - | - |
| po1gP2 | 512.21 | 612.12 | 559.21 | - | - | - |
| po1gP3 | 498.21 | 719.23 | 871.21 | - | - | - |
| po1gP4 | 132.68 | 42.92 | 543.55 | 432.64 | 371.12 | 538.25 |
| po1gP5 | 491.18 | 626.17 | 217.01 | 575.09 | 199.42 | 106.75 |
| po1gP6 | 58.43 | 323.79 | 305.29 | 388.76 | 399.89 | 315.13 |
| P4-2 | 698.12 | 623.41 | 773.97 | 842.26 | 770.52 | 625.62 |
| P4-1 | 923.19 | 906.22 | 921.28 | 996.46 | 962.23 | 992.63 |
| P4-3 | 771.19 | 724.11 | 890.59 | 956.47 | 901.14 | 917.82 |
| P4-4 | 985.15 | 737.92 | 733.79 | 663.11 | 664.26 | 890.78 |
| P4-5 | 730.15 | 946.59 | 1012.96 | 816.94 | 815.21 | 862.20 |
| P4-6 | 678.46 | 840.43 | 616.79 | 902.48 | 790.57 | 725.55 |
| P5-2 | 735.52 | 781.55 | 732.93 | 843.31 | 750.48 | 830.17 |
| P5-1 | 992.15 | 900.95 | 963.66 | 992.76 | 998.86 | 907.37 |
| P5-3 | 745.97 | 702.19 | 716.91 | 692.29 | 721.81 | 701.65 |
| P5-4 | 921.85 | 943.05 | 724.78 | 842.06 | 717.30 | 677.69 |
| P5-5 | 908.62 | 668.39 | 911.21 | 629.55 | 900.56 | 726.22 |
| P5-6 | 949.47 | 1196.80 | 1179.34 | 711.45 | 638.18 | 681.53 |
| PAR1 | 523.77 | 499.89 | 397.57 | 449.35 | - | - |
| PAR2 | 570.86 | 429.21 | 677.02 | 225.71 | - | - |
| PAR3 | 526.74 | 586.87 | 424.32 | 568.46 | - | - |
| PAR4 | 780.40 | 154.27 | 830.46 | 631.56 | - | - |
| PAR5 | 426.24 | 399.20 | 376.76 | 217.27 | - | - |
| PAR6 | 365.15 | 417.90 | 518.30 | 475.34 | - | - |
| PAR7 | 371.96 | 338.97 | 644.89 | 398.05 | - | - |
| PAR8 | 477.16 | 427.39 | 567.60 | 494.52 | - | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| PARL | 411.57 | 281.05 | 353.74 | 213.82 | - | - |
| ADH1 | 208.36 | 604.61 | 203.37 | 238.95 | 668.20 | 576.52 |
| ADH2 | 635.49 | 413.87 | 369.66 | 415.69 | 251.99 | 420.20 |
| ADH3 | 459.32 | 564.77 | 745.12 | 668.20 | 452.32 | 325.12 |
| ADH4 | 277.98 | 184.96 | 580.73 | 732.83 | 123.00 | 439.38 |
| ADH5 | 476.59 | 452.32 | 439.47 | 181.79 | 412.81 | 283.06 |
| ADH6 | 40.43 | 179.59 | 69.68 | 254.20 | 213.82 | 234.25 |
| ADH7 | 263.88 | 307.32 | 144.20 | 81.86 | 140.36 | 152.25 |
| ADH8 | 277.21 | 306.56 | 281.05 | 345.78 | 612.00 | 228.30 |
| ADH9 | 145.06 | 154.46 | 200.59 | 108.72 | 236.65 | 183.52 |
| po1gP7-3 | 1586.92 | 1430.22 | 816.84 | 647.00 | 633.58 | 994.06 |
| | 366.88 | 669.92 | 708.28 | 1019.48 | 658.70 | 759.97 |
| po1gP7-5 | 848.30 | 750.96 | 886.37 | 800.83 | 531.06 | 679.51 |
| | 826.43 | 1058.32 | 1028.88 | 865.65 | 1082.87 | - |
| po1gP7-6 | 776.28 | 1052.28 | 1062.15 | 877.26 | 833.91 | 1326.26 |
| | 1554.41 | 525.78 | 1014.01 | 499.32 | 1030.70 | 878.70 |
| po1gP7-1 | 919.07 | 1026.67 | 955.32 | 917.25 | 1018.23 | 1107.13 |
| | 1088.62 | 918.40 | 924.15 | 917.06 | 940.36 | 941.32 |
| po1gP7-2 | 927.51 | 951.96 | 1070.40 | 925.11 | 974.31 | 986.30 |
| | 995.89 | 930.48 | 917.34 | 883.40 | 901.81 | 985.72 |
| po1gP7-4 | 722.19 | 713.08 | 1062.34 | 746.93 | 778.96 | 710.49 |
| | 1222.31 | 872.85 | 912.26 | 903.05 | 810.99 | 909.67 |
| po1gP8-1 | 812.14 | 954.94 | 758.05 | 874.28 | 497.21 | 739.55 |
| | 783.85 | 1067.14 | 744.92 | 877.83 | 736.86 | 727.85 |
| | 788.93 | 837.84 | 767.74 | 925.11 | 897.01 | - |
| po1gP8-2 | 920.99 | 622.93 | 1027.44 | 976.03 | 1061.96 | 858.08 |
| | 987.54 | 851.08 | 630.89 | 944.96 | 783.08 | 965.97 |
| | 773.88 | 736.76 | 747.22 | 919.65 | 729.38 | 1007.87 |
| po1gP8-3 | 855.78 | 756.71 | 973.16 | 933.36 | 955.42 | 816.36 |

|  | | | | | |
|---|---|---|---|---|---|
| | 713.56 | 786.25 | 995.79 | 938.92 | 845.99 | 644.51 |
| | 711.93 | 803.70 | 1110.10 | 587.54 | 811.85 | 461.82 |
| po1gP8-4 | 660.52 | 907.18 | 385.96 | 908.71 | 775.32 | 880.33 |
| | 864.31 | 606.44 | 829.50 | 705.40 | 676.35 | 617.85 |
| | 603.37 | 635.11 | 758.34 | 836.69 | 520.03 | 881.38 |
| po1gP8-5 | 1067.33 | 691.31 | 474.38 | 736.96 | 952.35 | 804.95 |
| | 820.68 | 768.70 | 1159.11 | 745.59 | 684.02 | 920.70 |
| | 269.25 | 433.05 | 854.63 | 749.14 | 661.96 | 890.97 |
| po1gP8-6 | 725.74 | 864.98 | 801.69 | 969.32 | 89.06 | 1137.48 |
| | 340.31 | 735.13 | 800.92 | 935.37 | 887.81 | - |
| | 891.07 | 845.51 | 822.50 | 723.53 | 895.96 | 772.73 |
| po1fk1P | 621.40 | 267.91 | 560.02 | 340.60 | 280.66 | 720.08 |
| | 770.71 | 565.68 | 796.61 | 335.14 | 547.07 | 647.67 |
| po1fk2P | 45.80 | 509.67 | 368.99 | 924.73 | 600.30 | 532.21 |
| | 288.62 | 201.55 | 734.17 | 679.99 | 408.02 | 775.32 |
| po1fk3P | 149.38 | 579.30 | 63.35 | 598.00 | 184.96 | 211.42 |
| | 229.84 | 479.27 | 606.92 | 249.98 | 659.85 | 367.45 |
| po1fk4P1 | 1023.41 | 1022.26 | 1034.82 | 1132.45 | 1391.67 | 1056.40 |
| po1fk4P2 | 1008.45 | 1062.82 | 1619.81 | 1284.74 | 1047.67 | 1071.17 |
| po1fk4P3 | 981.41 | 1104.16 | 993.39 | 1030.22 | 1098.98 | 934.41 |
| po1fk4P4 | 1092.36 | 1123.82 | 980.93 | 1110.68 | 982.84 | - |
| po1fk4P5 | 568.75 | 603.08 | 1447.58 | 1042.40 | 985.15 | - |
| po1fk5P | 1595.07 | 1103.68 | 1301.33 | 1230.75 | 1359.54 | - |
| po1fk6P | 1081.05 | 1108.47 | 1593.25 | 1175.41 | - | - |
| po1fk7P | 797.95 | 1802.69 | 968.94 | 1011.42 | - | - |

**Supplementary Note**

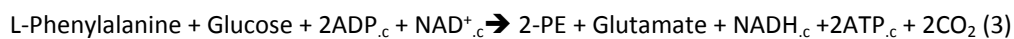**Stoichiometric model to assess Ehrlich pathway efficiency**

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

$$\text{L-Phenylalanine} + \text{2-Oxoglutarate} + \text{NADH}_c \rightarrow \text{2-PE} + \text{Glutamate} + \text{NAD}^+_c + CO_2 \text{ (1)}$$
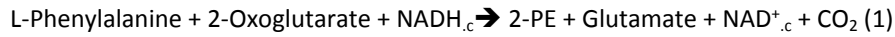
Here, 2-oxoglutarate is synthesized by the central carbon metabolism with the glucose as substrate (Fig 1b):

$$\text{Glucose} + \text{2ADP}_c + \text{2NAD}^+_c \rightarrow \text{2-Oxoglutarate} + \text{2NADH}_c + \text{+2ATP}_c + CO_2 \text{ (2)}$$
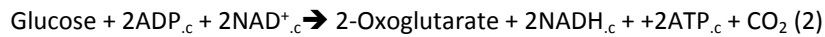
Thus, the overall stoichiometry of 2-PE biosynthesis is:

$$\text{L-Phenylalanine} + \text{Glucose} + \text{2ADP}_c + \text{NAD}^+_c \rightarrow \text{2-PE} + \text{Glutamate} + \text{NADH}_c + \text{2ATP}_c + 2CO_2 \text{ (3)}$$

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 $CO_2$. 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 $(RC_2H_4O_2N)_n$, 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

$$\text{x L-Phe} (C_9H_{11}NO_2) + \text{y Glucose} (C_6H_{12}O_6) + (8.5x+6y-10)\ O_2 \rightarrow \text{2-PE} (C_8H_{10}O) + \text{x Proteins} (C_2H_5O_2N)_n + (7x+6y-8)\ CO_2 +$$

$$(3x+6y-5)\ H_2O \text{ (4)}$$
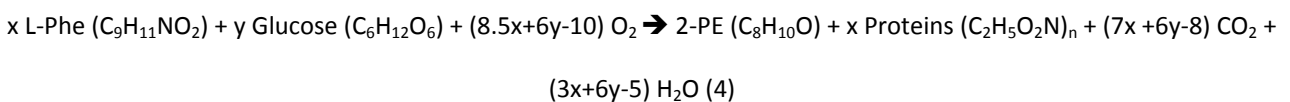
The yield of 2-PE (g/g L-phenylalanine) is

$$Y_{2\text{-PE}} = \frac{122}{165}x \text{ (5)}$$

Furthermore, introduction of the respiratory quotient (RQ):

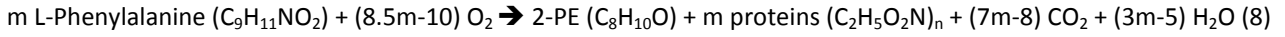$$RQ = \frac{7x + 6y - 8}{8.5x + 6y - 10}\ \text{(6)}$$

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.

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

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
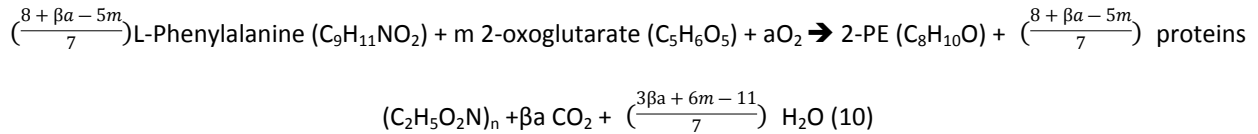
2-oxoglutarate supplementation.

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 L-Phenylalanine $(C_9H_{11}NO_2)$ + (8.5m-10) $O_2$ ➜ 2-PE $(C_8H_{10}O)$ + m proteins $(C_2H_5O_2N)_n$ + (7m-8) $CO_2$ + (3m-5) $H_2O$ (8)

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

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

As Shown in the metabolic model (Fig. S1), $Y_{2-PE}$ 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 $Y_{2-PE}$ 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

$(\frac{8+\beta a-5m}{7})$L-Phenylalanine $(C_9H_{11}NO_2)$ + m 2-oxoglutarate $(C_5H_6O_5)$ + a$O_2$ ➜ 2-PE $(C_8H_{10}O)$ + $(\frac{8+\beta a-5m}{7})$ proteins

$(C_2H_5O_2N)_n$ +βa $CO_2$ + $(\frac{3\beta a+6m-11}{7})$ $H_2O$ (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 $_{L-phenylalanine}$) is

$$Y_{2-PE}=\frac{122}{165*\frac{8-5m+\beta*\frac{29m+4}{17\beta-7}}{7}}\ (11)$$

If β=1, $Y_{2-PE}=\frac{122}{165*(1.2-0.3m)}$ (12);

If β=2, $Y_{2-PE}=\frac{122}{165*\frac{224+53m}{189}}$ (13);

If β=10, $Y_{2-PE}=\frac{122}{165*\frac{1344-525m}{1141}}$ (14);

Analysis of the above equation (Fig. S1) indicates that increasing aKG supplementation will significantly improve 2PE yield $Y_{2-PE}$ under all different β 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.
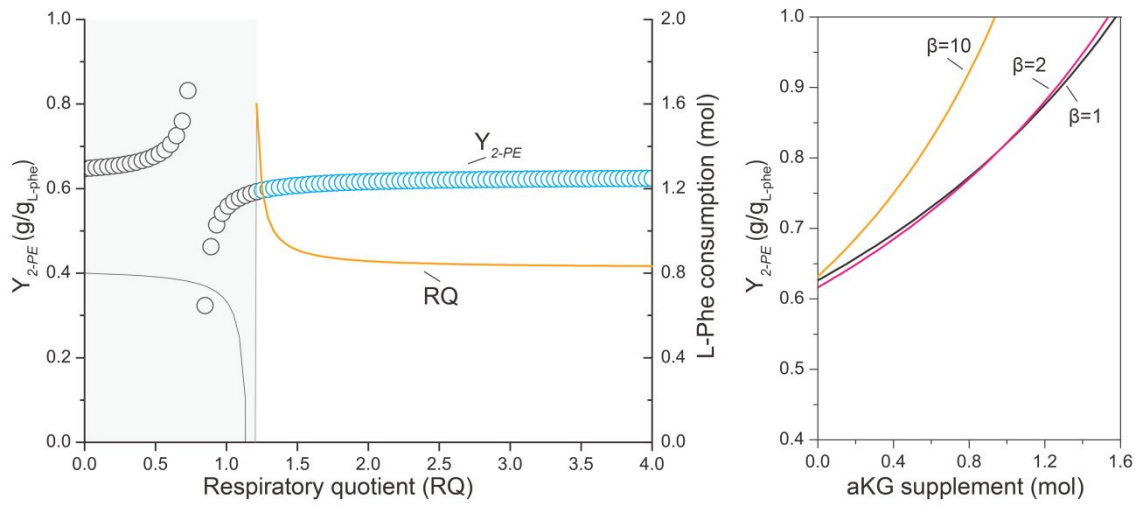
Fig. S1. Stoichiometric models reveal that 2-PE yield is driven by selectivity of Ehrlich pathway and supply of 2-oxoglutarate (aKG).
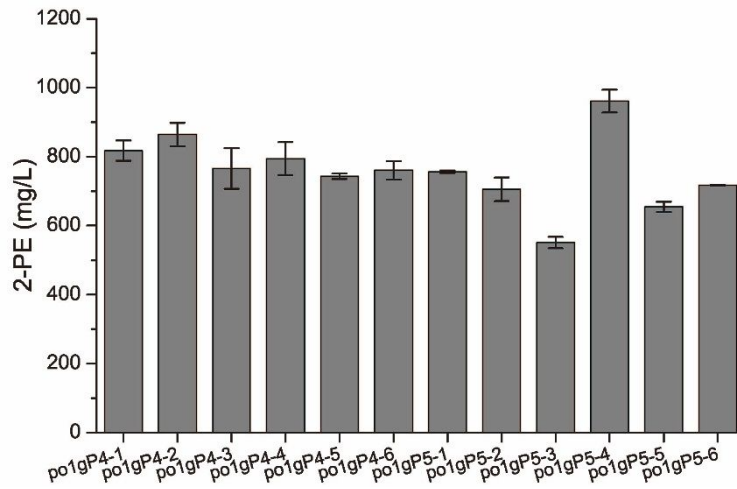
**Supplementary Figure**



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.
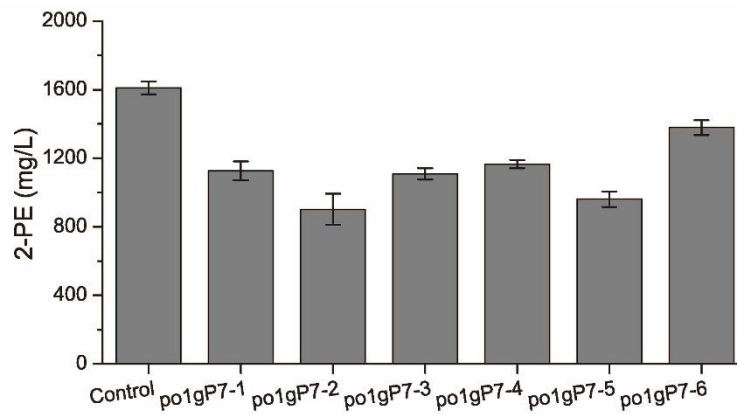


Fig. S3. 2-PE titers of strain strains **po1gP7-1, po1gP7-2**, **po1gP7-3, po1gP7-4, po1gP7-5,** and **po1gP7-6** in shake cultivations.

**References**

(1) Madzak, C., Treton, B., and Blanchin-Roland, S. (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica. J Mol Microbiol Biotechnol 2*, 207-216.

(2) Lv, Y., Edwards, H., Zhou, J., and Xu, P. (2019) Combining 26s rDNA and the Cre-loxP system for iterative gene integration and efficient marker curation in *Yarrowia lipolytica*. *ACS Synth. Biol. 8* (3), 568-576.