

Electronic Supplementary Information (ESI)

Sustainable production of rare oleanane-type ginsenoside Ro with an
artificial glycosylation pathway in *Saccharomyces cerevisiae*

Shichao Ren¹, Qiuyan Sun¹, Liang Zhang¹, Wentao Sun¹, Yongxing Li¹, Xudong
Feng^{1*}, and Chun Li^{1,2*}

¹Key Laboratory of Medical Molecule Science and Pharmaceutical Engineering,
Ministry of Industry and Information Technology, Institute of Biochemical Engineering,
School of Chemistry and Chemical Engineering, Beijing Institute of Technology,
Beijing, 100081, China

²Key Lab for Industrial Biocatalysis, Ministry of Education, Department of Chemical
Engineering, Tsinghua University, Beijing 100084, China

*Corresponding author: Xudong Feng (xd.feng@bit.edu.cn)

Chun Li (lichun@mail.tsinghua.edu.cn)

Table S1 Strains and plasmids used in this study

Strains and [Plasmids]	Description	Sources
<i>S. cerevisiae</i> SynV	Derived from BY4741 with artificial chromosome V	(Xie et al., 2017)
OA01	<i>S. cerevisiae</i> SynV::HO::FBA1p-bAS-PGII1t-TDH2p-MtCPR-ADH1t-TEF1p-CYP716A12-CYC1t- HIS	This study
OA02	OA01::NTS::TPI1p-tHMG1-TPI1t- PYK1p-ERG1-PYK1t-LEU	This study
OA03	OA01::NTS::TPI1p-tHMG1-TPI1t-PGII1p-ERG9-PGII1t-PYK1p-ERG1-PYK1t-LEU	This study
CE01	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-OAGT-CYC1t	This study
CE02	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-OAGT1-CYC1t	This study
CE03	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-GuCsl-CYC1t	This study
CE04	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-GmCsl1-CYC1t	This study
CE05	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-GmCsl3-CYC1t	This study
CE06	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-Pn007110-CYC1t	This study
CE07	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-Pn007111-CYC1t	This study
CE08	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-Pn016322-CYC1t	This study
CE09	OA03::Gal80:: kanMX-ADH1t-HsUGDH-HXT7p-TEF1p-Pn022859-CYC1t	This study
CE10	OA03::Gal80:: kanMX-ADH1t-AtUGDH-HXT7p-TEF1p-GmCsl3-CYC1t	This study
CE14	OA03::Gal80:: kanMX-ADH1t-AtUGDH-HXT7p-TEF1p-Pn022859-CYC1t	This study
[pET-28a]	Kana resistance, used in prokaryote	Stored in lab
[pET-28a-11]	[pET-28a] with T7-UGT73C11-T7t cassette	Stored in lab
[pET-28a-33]	[pET-28a] with T7-UGT73C33-T7t cassette	Stored in lab
[pET-28a-24]	[pET-28a] with T7-UGT73F24-T7t cassette	Stored in lab
<i>E. coli</i> BL21(DE3)	<i>F-ompT hsdS(rB-mB-) gal dcm</i> (DE3)	Purchased from Biomed
DE3-11	<i>E. coli</i> BL21(DE3) with [pET-28a-11]	Stored in lab
DE3-33	<i>E. coli</i> BL21(DE3) with [pET-28a-33]	Stored in lab
DE3-24	<i>E. coli</i> BL21(DE3) with [pET-28a-24]	Stored in lab
[pET-28a-Gm2]	[pET-28a] with T7-GmsGT2-T7t cassette	This study

[pET-28a-F3]	[pET-28a] with <i>T7-UGT73F3-T7t</i> cassette	This study
[pET-28a-40]	[pET-28a] with <i>T7-UGT73P40-T7t</i> cassette	This study
[pET-28a-171]	[pET-28a] with <i>T7-17171-T7t</i> cassette	This study
[pET-28a-177]	[pET-28a] with <i>T7-17177-T7t</i> cassette	This study
[pET-28a-225]	[pET-28a] with <i>T7-023225-T7t</i> cassette	This study
[pET-28a-560]	[pET-28a] with <i>T7-027352560-T7t</i> cassette	This study
[pET-28a-971]	[pET-28a] with <i>T7-020234971-T7t</i> cassette	This study
[pET-28a-023]	[pET-28a] with <i>T7-15023-T7t</i> cassette	This study
[pET-28a-440]	[pET-28a] with <i>T7-003623440-T7t</i> cassette	This study
[pET-28a-389]	[pET-28a] with <i>T7-98389-T7t</i> cassette	This study
DE3-Gm2	<i>E. coli</i> BL21(DE3) with [pET-28a-Gm2]	This study
DE3-F3	<i>E. coli</i> BL21(DE3) with [pET-28a-F3]	This study
DE3-40	<i>E. coli</i> BL21(DE3) with [pET-28a-40]	This study
DE3-171	<i>E. coli</i> BL21(DE3) with [pET-28a-171]	This study
DE3-177	<i>E. coli</i> BL21(DE3) with [pET-28a-177]	This study
DE3-225	<i>E. coli</i> BL21(DE3) with [pET-28a-225]	This study
DE3-560	<i>E. coli</i> BL21(DE3) with [pET-28a-560]	This study
DE3-971	<i>E. coli</i> BL21(DE3) with [pET-28a-971]	This study
DE3-023	<i>E. coli</i> BL21(DE3) with [pET-28a-023]	This study
DE3-440	<i>E. coli</i> BL21(DE3) with [pET-28a-440]	This study
DE3-389	<i>E. coli</i> BL21(DE3) with [pET-28a-389]	This study
R101	CE14:: <i>YPRCτ3</i> :: <i>TEF1p-UGT73P40-CYC1t-URA3</i>	This study
IVA01	CE14:: <i>YPRCτ3</i> :: <i>HXT7p-UGT73F3-ADH1t-URA3</i>	This study
Ro01	CE14:: <i>YPRCτ3</i> :: <i>CYC1t-UGT73P40-Gal1/10-UGT73F3-ADH1t-URA3</i>	This study
Ro02	CE14:: <i>YPRCτ3</i> :: <i>CYC1t-PjmUGT1-Gal1/10-PjmUGT2-ADH1t-URA3</i>	This study
Ro03	CE14:: <i>YPRCτ3</i> :: <i>HXT7p-UGT73P40-Linker-UGT73F3 -ADH1t-URA3</i>	This study
Ro04	CE14:: <i>YPRCτ3</i> :: <i>TEF1p-UGT73F3-Linker-UGT73P40 -ADH1t-URA3</i>	This study
Ro05	Ro01:: <i>YORWΔ22</i> :: <i>TEF1p-BleoR-CYC1t-Gal1/10p-Pn022859-PGII t</i>	This study
Ro06	Ro01:: <i>YORWΔ22</i> :: <i>TEF1p-BleoR-CYC1t-ADH1t-AtUGDH-Gal1/10p-Pn022859-PGII t</i>	This study
Ro07	Ro01:: <i>YORWΔ22</i> :: <i>TEF1p-BleoR-CYC1t-ADH1t-Pn022859-Gal1/10p-UGT73F3-PGII t</i>	This study
Ro08	Ro01:: <i>YORWΔ22</i> :: <i>TEF1p-BleoR-CYC1t-ADH1t-AtUGDH-Gal1/10p</i>	This study
Ro09	Ro01:: <i>YORWΔ22</i> :: <i>TEF1p-BleoR-CYC1t-ADH1t-UGT73F3-Gal1/10p-AtUGDH-PGII t</i>	This study
Ro10	Ro05:: <i>YHRCΔ14</i> :: <i>MET17p-MET17-MET17t-Gal2p-Pn022859-PGII t</i>	This study

Table S2 Primers used in this study

Names	From 5'-3'	Usage
P1A-F	TGAATTCGAGAATTATCCTGGGCACGAGTGAAAC	GgbAS cassette
P1A-R	CAGTATTGTTATACGCCATTTTAAGTCCAAAGGCA CAAT	GgbAS cassette
P1B-F	TAAAATGGCGTATAACAATACTGACAGTACTAAAT AATTGCCTACTTGG	GgbAS cassette
P1B-Rx	GAGCGATTTGTTTAAGTCAAACAAACTGGAGTAG ATGGCAATG	GgbAS cassette
P1D-F	CACTAAACTACATTGGTTTTTCCAGTGAATGATTA TTTGTCG	GgbAS cassette
P1D-R	TGCATGCCTGCAGGCAGGTCTGCGATCACTGG	GgbAS cassette
P1PGI1t-F	TTTGACTTAAACAAATCGCTCTTAAATATATACCTA AAGAACATTAAAGCT	GgbAS cassette
P1PGI1t-R	GGAAAAACCAATGTAGTTTAGTGTTTTTCTTCCA GTGCGAG	GgbAS cassette
VP1-NF	TATCCTGGGCACGAGTGAAACAAAG	GgbAS cassette
VP1-R	CAGGTCTGCGATCACTGGGTAC	GgbAS cassette
P2A-F	GTGAATTCGAGGTTAATTCAAATTAATTGATATAG TTTTTTAATGAGTATTGAATCTGT	MtCPR cassette
P2A-R	GAAAAACCAATGCTATCAAAAACGATAGATCGAT TAGGATGACT	MtCPR cassette
VP2-F	CATTGGTTTTTCCAGTGAATGATTATTTGTCGTTAC C	MtCPR cassette
P2B-F	TTTTTGATAGCATTGGTTTTTCCAGTGAATGATTAT TTGTCGT	MtCPR cassette
P2B-R	AGAAGATGTCATTTTGTTTTGTTTGTGTGTGAT GAATTTAATTTGAACTAAG	MtCPR cassette
P2C-F	AACAAAACAAAATGACATCTTCTAATTCAGACTT GGTTAGAACA	MtCPR cassette
P2C-R	TCGCGGTCGAGCGACCTCATGCTATACCT	MtCPR cassette
P2D-F	GTCGCTCGACCGCGAATCCTTACATCAC	MtCPR cassette
VP2-R	TTTGTAAATAAAACCTTAGATTAGATTGCTATGCTTT CTTTCTAATGAGC	MtCPR cassette
P2D-R	TGCCTGCAGTTTGTAATTA AAACTTAGATTAGATT GCTATGCTTTCTTTCTAATG	MtCPR cassette
P3A-F	TGAATTCGAGGCTCTTAATTAACAATTCTTCGCCA GAGG	CYPs cassette
VP3-F	CTTCGCCAGAGGTTTGGTCAAGTC	CYPs cassette
P3A-R	TCGCGGTCGAGCGACCTCATGCTATACCT	CYPs cassette
P3B-F	GTCGCTCGACCGCGAATCCTTACATCAC	CYPs cassette
CYP-F	TAGCATGAGGTCGCTCCACACACCATAGCTTC	CYPs cassette
P3B-R	ATTTGGCTCCATTTTGTAAATTA AAACTTAGATTAG ATTGCTATGCTTTCTTT	CYPs cassette

P3C-F	TTAATTACAAAATGGAGCCAAATTTCTACTTGTCT TTATTGT	CYPs cassette
CYP-R	TATAGAGTGTACTAGCTTCGAGCGTCCCAAACC TTCTCAAG	CYPs cassette
P3C-R	GTGTACTAGCTTCGAGCGTCC	CYPs cassette
P3D-F	GCTCGAAGCTAGTACACTCTATATTTTTTTATGCCT CGGTAATG	CYPs cassette
VP3-R	GACGTTAATCACTTGCGATTGTGTGG	CYPs cassette
P3D-R	GCCTGCAGCTTGCCAGAGCATGTATCA	CYPs cassette
P4B-R	TCCCCAGCTGATGCGGATTTTCTCCTTACGC	HIS marker
P4A-F	GAATTCGAGGCTTGGTACCGCGGCTAGCT	HIS marker
VP4-F	GAGTTAGACAACCTGAAGTCTAGGTCCC	HIS marker
P4A-R	GTGTACTAGCTTCGAGCGTCC	HIS marker
P4B-F	GCTCGAAGCTAGTACACTCTATATTTTTTTATGCCT CGGTAATG	HIS marker
P4C-F	GCATCAGCTGGGGGAACAACCTTCACAGAATGT	HIS marker
VP4-R	CTTGCAATTTATACAGTGATGACCGCTG	HIS marker
VP4-NR	CTGTAAGATTCCGCCACATTTTATACACTCT	HIS marker
P4C-R	GCCTGCAGACTGTAAGATTCGCCACATTTTATAC AC	HIS marker
P5A-F	GTGAATTCGAGGATAGTTTAAACGGAAACGCAGGT GATATG	tHMG1 cassette
VP5-F	CAGACATGTTTCAGTAGGTGGGAGTG	tHMG1 cassette
P5A-R	TTCTAGATATAGAAGTACCTCCCAACTACTTTTC C	tHMG1 cassette
P5B-F	GGAGGTACTTCTATATCTAGGAACCCATCAG	tHMG1 cassette
P5B-R	CAATTGGTCCATTTTTAGTTTATGTATGTGTTTTTT GTAGT	tHMG1 cassette
P5C-F	CATAAACTAAAAATGGACCAATTGGTGAAAACCTG AAG	tHMG1 cassette
P5C-R	AATTATATTAATCTTAGGATTTAATGCAGGTGACG GACC	tHMG1 cassette
P5D-F	ATTAAATCCTAAGATTAATATAATTATATAAAAATAT TATCTTCTTTTCTTTATATCTAGTG	tHMG1 cassette
VP5-R	CATCTTCTCAACGCGAAAATGACGC	tHMG1 cassette
P5D-R	ATGCCTGCAGTATATAACAGTTGAAATTTGGATAA GAACATCTTCTCAACGC	tHMG1 cassette
P6A-F	AGTGAATTCGAGGATTAATATAATTATATAAAAATA TTATCTTCTTTTCTTTATATCTAGTG	ERG9 cassette
VP6-F	CGTGAATGTTCTTGTAAGGGACGG	ERG9 cassette
P6A-R	TGATTTTTGTACTATATAACAGTTGAAATTTGGAT AAGAACATCTTCTCAACGC	ERG9 cassette
P6B-F	AACTGTTATATAGTAACAAAATCACGATCTGGGT GGG	ERG9 cassette

P6-NV-R	GCTTTCCCATTTTTAGGCTGGTATCTTGATTCTAAA TCG	ERG9 cassette
P6-9-F	CCAGCCTAAAAATGGGAAAGCTATTACAATTGGC	ERG9 cassette
P6-9-R	CGATTTGTTACGCTCTGTGTAAAGTGATATATA ATAAAACCCAAG	ERG9 cassette
P6-NV-F	CACAGAGCGTGAACAAATCGCTCTTAAATATATAC CTAAAGAAC	ERG9 cassette
VP6-R	GAGTTATGTCCTTCGCGCACTGATTC	ERG9 cassette
P6D-R	TGCCTGCAGGTAGTTTAGTGTTTTTCTTCCAGTGC GAG	ERG9 cassette
P7A-F	GAATTCGAGACAAATCGCTCTTAAATATATACCTA AAGAACATTAAAGC	ERG1 cassette
VP7-F	GAACATTAAAGCTATATTATAAGCAAAGATACGTA AATTTTGC	ERG1 cassette
P7A-R	CGACGATCCAGTTTAGTGTTTTTCTTCCAGTGCG	ERG1 cassette
P7B-F	CACTAAACTGGATCGTGGTGTGATACGG	ERG1 cassette
P7B-Rx	CGTTAACAGCAGACATTGTGATGATGTTTTATTTG TTTTGATTGGTGTCTTG	ERG1 cassette
P7C-F	CATCATACAATGTCTGCTGTTAACGTTGCACCTG	ERG1 cassette
P7C-Rx	CATGATTCTTTTTTAAACCAATCAACTCACCAAAC AAAAATGG	ERG1 cassette
P7D-Fx	GTTGATTGGTTAAAAAAGAATCATGATTGAATG AAGATATTATTTTTTGAATTATATTTT	ERG1 cassette
VP7-R	CATTTATGTACCCATGTATAACCTTCCAAGTTAGTT C	ERG1 cassette
P7D-R	TGCCTGCAGCATTTATGTACCCATGTATAACCTTC CAAG	ERG1 cassette
V8-R	TGATTCTTTTTCTCGAATTCACTGGCCGTCG	LEU marker
P8A-F	GTGAATTCGAGAAAAAGAATCATGATTGAATGAA GATATTATTTTTTGAATTATATTTT	LEU marker
VP8-F	CTCCCATCGCAAGATAATCTTCCCTC	LEU marker
P8A-R	TTCCACAGTTCATTTATGTACCCATGTATAACCTT CCAAGTTAGT	LEU marker
P8B-F	TACATAAATGAACTGTGGGAATACTCAGGTATCGT	LEU marker
P8C-F	CGCGCGACGTTGCAAAGATGGGTTGAAAGAG	LEU marker
VP8-R	CTGTAGTTAAGCTGGTAAGAGCCTGAC	LEU marker
P8C-R	GCCTGCAGTTGCGGCCATATCTACCAGAAAGC	LEU marker
K-Gal80-V-R	CAGGTATCCGATACCGAGCTCGAATTCACTG	Csls and UGDH cassette
Gal80L-F	GAGCTCGGTATCGGATACCTGCACAAGC	Csls and UGDH cassette
K-Gal80-S-F	GGATACCTGCACAAGCAATTTGGCACCTG	Csls and UGDH cassette

G418-F	CGGCGCACTCTCGATGGTAGACAACC	Csls and UGDH cassette
G4-UD-GT-R	CACTAGTATCGATGATCTGAGCGAATTTCTTATGATTTATG	Csls and UGDH cassette
UDGH-R	CATAAGAAATTCGCTCAGATCATCGATACTAGTGCGGC	Csls and UGDH cassette
UDGH-F	TTTAATCAAAAAATGTTTCGAGATCAAGAAGATTTGTTGCATTGG	Csls and UGDH cassette
G4-UD-GT-F	CAACAAATCTTCTTGATCTCGAACATTTTTTTGATTAAAATTAATAAAAAAC	Csls and UGDH cassette
GTs-R	GCTGCACTCTTCGAGCGTCCC	Csls and UGDH cassette
K-Gal80-S-R	CACTTCATGAACCTGTTGGGAAGCCTTGG	Csls and UGDH cassette
K-Gal80R-R	CAAGCTTGCATCACTTCATGAACCTGTTGGGAAG	Csls and UGDH cassette
K-Gal80-V-F	GTTTCATGAAGTGATGCAAGCTTGGCGTAATCATG	Csls and UGDH cassette
G4-AtUD-GT-R	CCTGCTGTTGCTTAAGCGAATTTCTTATGATTTATG	Csls and UGDH cassette
AtUGDH-R	CATAAGAAATTCGCTTAAGCAACAGCAGGCATATC	Csls and UGDH cassette
G4-AtUD-GT-F	CCAATACAACAAATTTAACCATGCCGGTTTTTTGATTAAAATTAATAAAAAAC	Csls and UGDH cassette
AtUGDH-F	CAAAAACCGGCATGGTTAAAATTTGTTGTATTGG	Csls and UGDH cassette
YPRC3L-F	GCTCGGTAAAAGGAGGTGCACGCATTATGG	UGT73F3 cassette
YPL-HXT-R	CTCATTTCCACGGTTCCAAGGAGGTGAAGAAC	UGT73F3 cassette
YPL-HXTF3-F	CTCCTTGGAACCGTGGAAATGAGGGGTATG	UGT73F3 cassette
YP-F3-URA-R	GATGAATTGAAGAGCGACCTCATGCTATACCTGAGAAAGC	UGT73F3 cassette
YP-F3-URA-F	CTCAGGTATAGCATGAGGTCGCTCTTCAATTCATCATTTTTTTTTTATTC	UGT73F3 cassette
YPRC3R-24-URA-R	GCTGACGTCCCATCCGCGAATTTTAACAAAATATTAAACGTTTAC	UGT73P40 cassette
URA-YPRC3R-F	GTAAAATTCGCGGATGGGACGTCAGCACTG	UGT73P40 cassette
YPRC3R-R	CAAGCTTGCATCGGTATTACTCGAGCCCCTAATAC	UGT73P40 cassette
YPRC3L-TEF-R	GAAGCTATGGTGTGTGTTCCAAGGAGGTGAAGAAC	UGT73P40 cassette

YPRC3L-24-F	CCTCCTTGGAACACACACCATAGCTTCAAAATGT TTCTACTC	UGT73P40 cassette
YPRC3-24-R	GATGAATTGAACTTCGAGCGTCCCAAACCTTCT C	UGT73P40 cassette
YPRC3-24- UTA-F	GGTTTTGGGACGCTCGAAGTTCAATTCATCATTTT TTTTTTATTCTTTTTTTTIG	UGT73P40 cassette
F3Gal24-F	CCTCCTTGGAACTTCGAGCGTCCCAAACC	UGT73F3 and UGT73P40 cassette
YPL-HXT-R	CTCATTTCCACGGTTCCAAGGAGGTGAAGAAC	UGT73F3 and UGT73P40 cassette
YPL-HXTF3-F	CTCCTTGGAACCGTGGAAATGAGGGGTATG	UGT73F3 and UGT73P40 cassette
F3-L-67-R	ACCAGAACCACCAGAACCATTAACCTTAAACATCC AAAGCTTTAGACTTC	Fusion protein cassette
24-L-F3-F	GGTTCTGGTGGTTCTGGTATGGAAGGGGTGAAG TTGAAC	Fusion protein cassette
F3-L-67-F	GGATGATGGTTCTGGTGGTTCTGGTATGATGGAAT CTCAACAATTGC	Fusion protein cassette
L-his-67-R	GAACCATGATGATGATGATGATGGTCGACGGCGC TATTATTAACCTTAAACATCCAAAGC	Fusion protein cassette
24-L-F3-R	ACCAGAACCACCAGAACCATCATCCAATTTGAGG TCTCTCAATC	Fusion protein cassette
lac-YOL-F	GAGCTCGGTAGGTGTGGTTTTCGGCATTATC	Pn022859 overexpression
TE-YOL-R	GATCTGATATGATTACCAGCTTGATTCGCGGGCTG	Pn022859 overexpression
YOL-TE-F	CTGGTAATCATATCAGATCGATCCCCACACACCA TAG	Pn022859 overexpression
G1-CY-R	AACCTAGTGGCAAATTAAGCCTTCGAGCGTC	Pn022859 overexpression
CY-G1-F	GCTTTAATTTGCCACTAGGTTTTTCTCCTTGACG TTAAAGTATAG	Pn022859 overexpression
859-G1-R	GGATCCTTTCAAAAATTCTTACTTTTTTTTTGGAT GGACGCAAAG	Pn022859 overexpression
G1-859-F	GTAAGAATTTTTGAAAGGATCCGGATCCATGGCTT CTACTACATC	Pn022859 overexpression
PG-859-R	GAGCGATTTGTCTTAATTACCTTTTCTAGTCAACA TACCACCAATAATTGG	Pn022859 overexpression
859-PG-F	GACTAGAAAAGGTAATTAAGACAAATCGCTCTTA AATATATACCTAAAGAACATTAAGC	Pn022859 overexpression

YOR-PG-R	GATGATAGTTGGTCCGTAGTTT CAGTGC	Pn022859 overexpression
PG-YoR-F	GGAAGAAAAACACTAACTACGGACCACTATC ATCCGCTAATTAC	Pn022859 overexpression
YOR-R	CAAGCTTGCATGAAGTAGCGATCAGACTACG	AtUGDH overexpression
AD-CY-R	GTCGCTCCTAGTGGCAAATTAAGCCTTCGAGC	AtUGDH overexpression
CY-AD-F	CTTTAATTTGCCACTAGGAGCGACCTCATGCTATA CC	AtUGDH overexpression
G1-AtU-R	GGAGAAAAAACTCTGTGATGGTTAAAATTTGTTG TATTGGTGCTGG	AtUGDH overexpression
AtU-G1-F	CCATCACAGAGTTTTTCTCCTTGACGTTAAAGTA TAGAGG	AtUGDH overexpression
YoR-Gal-R	GCGGATGATAGTTGGTCCTTTCAAAAATTCTTACT TTTTTTTTGGATGG	AtUGDH overexpression
Gal-YoR-F	GTAAGAATTTTTGAAAGGACCAACTATCATCCGC TAATTACTGAC	AtUGDH overexpression
Gal-F3-F	GTAAGAATTTTTGAAAGGATCCGGATCCATGGAA GGGGTTG	UGT73F3 overexpression
PG-F3-R	GCGATTTGTCTTAATCATCCAATTTGAGGTCTCTC AATCTC	UGT73F3 overexpression
F3-PG-F	CCTCAAATTGGATGATTAAGACAAATCGCTCTTAA ATATATACCTAAAGAACATTAAGC	UGT73F3 overexpression
Gal-F3-R	CAAGGAGAAAAAACTCTGTGATGGAAGGGGTTG AAGTTG	UGT73F3 overexpression
YHL-F	GAATTCGAGCTCTGCAACTTGTCAGCCTTTG	Pn022859 overexpression
M17p-YHL-R	CCTCAAGAGAAAAAGCAAAAATAATCACTGAG CTTCTGTTGCTG	Pn022859 overexpression
YHL-M17p-F	GCTCAGTGATTATTTTTGCTTTTTCTCTTGAGGTC ACATGATCG	Pn022859 overexpression
YHR-M17t-R	CCACTGTATGTATGTGCCCTCGCTTTTATATTTAAT AACTAACATTTTTATCAACG	Pn022859 overexpression
M17t-YHR-F	GTTATTAATATAAAAAGCGAGGGCACATACATACA GTGGGACATTATGAG	Pn022859 overexpression
YHR-R	GCTTGCATTCAGCAGAGGCGTAAAAGTGC	Pn022859 overexpression
G2-M17t-R	CCTTGGATTAGTACAGCTTTTATATTTAATAACTAA CATTTTTATCAACGATGGAATTCC	Pn022859 overexpression
M17t-G2-F	GTTAGTTATTAATATAAAAAGCTGTACTAATCCAA GGAGGTTTACGGACCAGG	Pn022859 overexpression
859-G2-R	GGATCCGGATCCTATGAAAGAATTATTTTTTTTATT ATGTTAATCTTGTGTTTACTTAAC	Pn022859 overexpression

G2-859-F	CTTTCATAGGATCCGGATCCATGGCTTCTACTACA TCTTTG	Pn022859 overexpression
YHR-PGI1t-R	CTCGCGTAGTTTAGTGTTTTTCTTCCAGTGCGAG	Pn022859 overexpression
PGI1t-YHR-F	GAAAAACACTAAACTACGCGAGGGCACATACATA CAG	Pn022859 overexpression

Table S3 Constituents of trace metal solution

Composition	Concentration (g/L)
EDTA	15
ZnSO ₄ • 7H ₂ O	6.25
MnCl ₂ • 4H ₂ O	0.32
CuSO ₄ • 5H ₂ O	0.8
CoCl ₂ • 6H ₂ O	0.5
Na ₂ MnO ₄ • 2H ₂ O	0.5
CaCl ₂ • 2H ₂ O	2.9
FeSO ₄ • 7H ₂ O	2.8

Table S4 Constituents of vitamin solution

Constituents	Concentration (g/L)
Biotin	0.05
Calcium pantothenatel	1
Nicotinic acid	1
Myo-inositol	25
Thinmine-HCl	1
Pyridoxal-HCl	1
<i>p</i> -Aminobenzoic acid	0.2

Table S5 Enzymes used for phylogenetic tree analysis and sequence alignment

No.	Enzymes	Species	Accession number
1	GuCSyGT (GuCsl)	<i>Glycyrrhiza uralensis</i>	LC500232
2	GmCSyGT1 (GmCsl1)	<i>Glycine max</i>	LC500227
3	GmCSyGT2	<i>Glycine max</i>	LC500228
4	GmCSyGT3 (GmCsl3)	<i>Glycine max</i>	LC500229
5	LjCSyGT	<i>Lotus japonicus</i>	LC500233
6	GmCSIM1	<i>Glycine max</i>	LC500230
7	GmCSIM2	<i>Glycine max</i>	LC500231
8	PttCesA8	<i>Populus tremula x Populus tremuloides</i>	6WLB_A
9	AAT09895	<i>Populus tremula x Populus tremuloides</i>	AAT09895
10	WP_012408188	<i>Nostoc punctiforme</i>	WP_012408188
11	KAF3445492	<i>Rhamnella rubrinervis</i>	KAF3445492
12	At1G55850	<i>Arabidopsis thaliana</i>	NP_175981
13	At4G23990	<i>Arabidopsis thaliana</i>	NP_194130
14	At4G24010	<i>Arabidopsis thaliana</i>	NP_194132
15	At4G24000	<i>Arabidopsis thaliana</i>	NP_567692
16	At1G32180	<i>Arabidopsis thaliana</i>	NP_001321681
17	At1G02730	<i>Arabidopsis thaliana</i>	NP_171773
18	At3G03050	<i>Arabidopsis thaliana</i>	NP_186955
19	At2G33100	<i>Arabidopsis thaliana</i>	NP_001324935
20	At4G38190	<i>Arabidopsis thaliana</i>	NP_195532

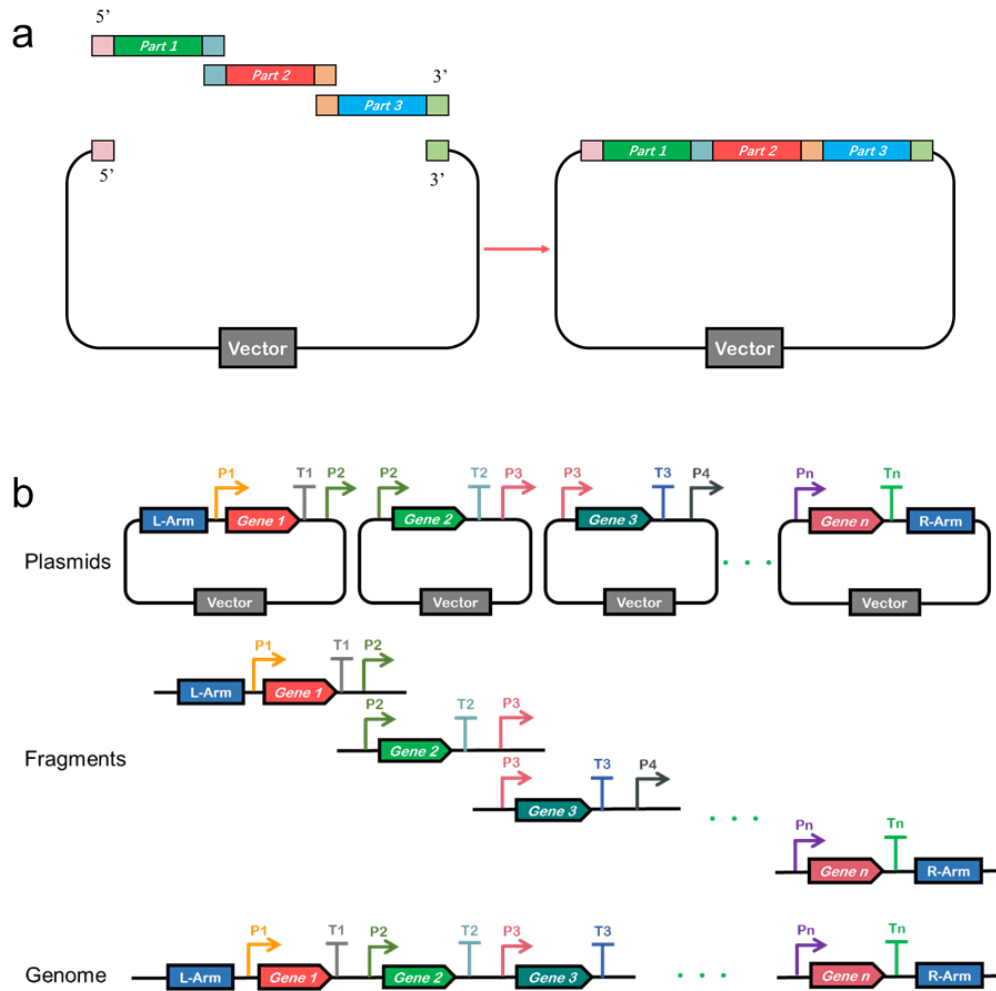


Fig. S1 Diagram of the construction strategy of engineered strains. **(a)** Plasmids construction. Gene elements were amplified with primers having 15-25 bp overlap by PCR and assembled into a plasmid by Gibson assembly. **(b)** Genome integration. Gene cassettes with homologous region were amplified and integrated into the genome through homologous recombination.

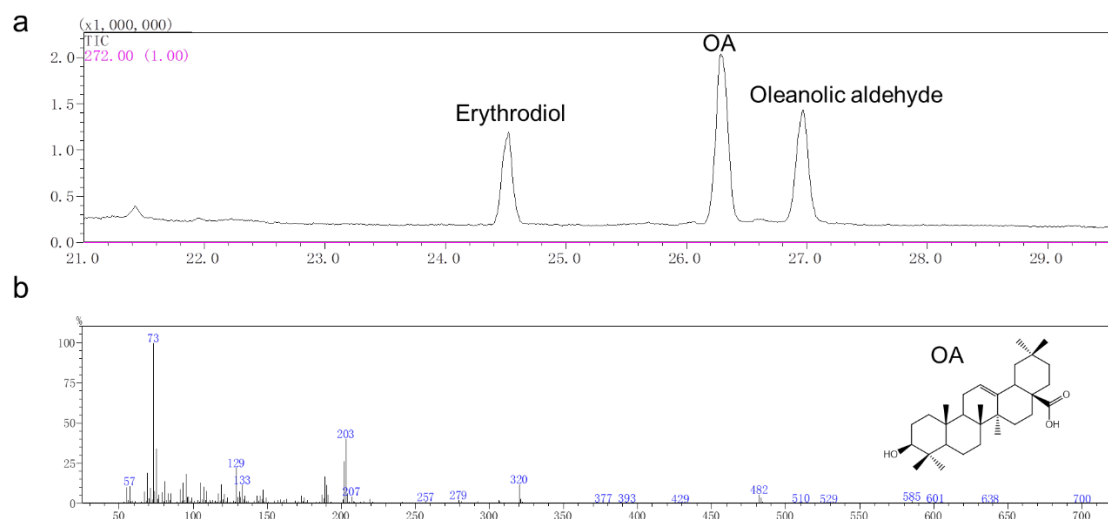


Fig. S2 GC-MS analysis of the fermentation products of strain OA01. (a) Chromatogram result of fermentation products of strain OA01. (b) Mass spectrum of OA. The GC-MS analysis was performed through a GCMS-QP2010 Ultra (Shimadzu Scientific Instruments) equipped with a SH-Rxi-5Sil MS (30 m × 0.25 mm × 0.25 μm, SHIMADZU, Japan) GC column. The flow rate of carrier gas helium was 1.2 mL/min and the split ratio was 5:1. The compound separation procedure was as follows: injection volume of 1 μL, injector temperature of 250 °C. Column temperature started with 80 °C and was kept for 1 min, then raised to 290 °C with increase of 20 °C/min and kept at 290 °C for 20 min. Ion source temperature was set 200 °C and interface temperature was set 290 °C. Scan mode was used and the m/z ranging from 50 to 700 was monitored.

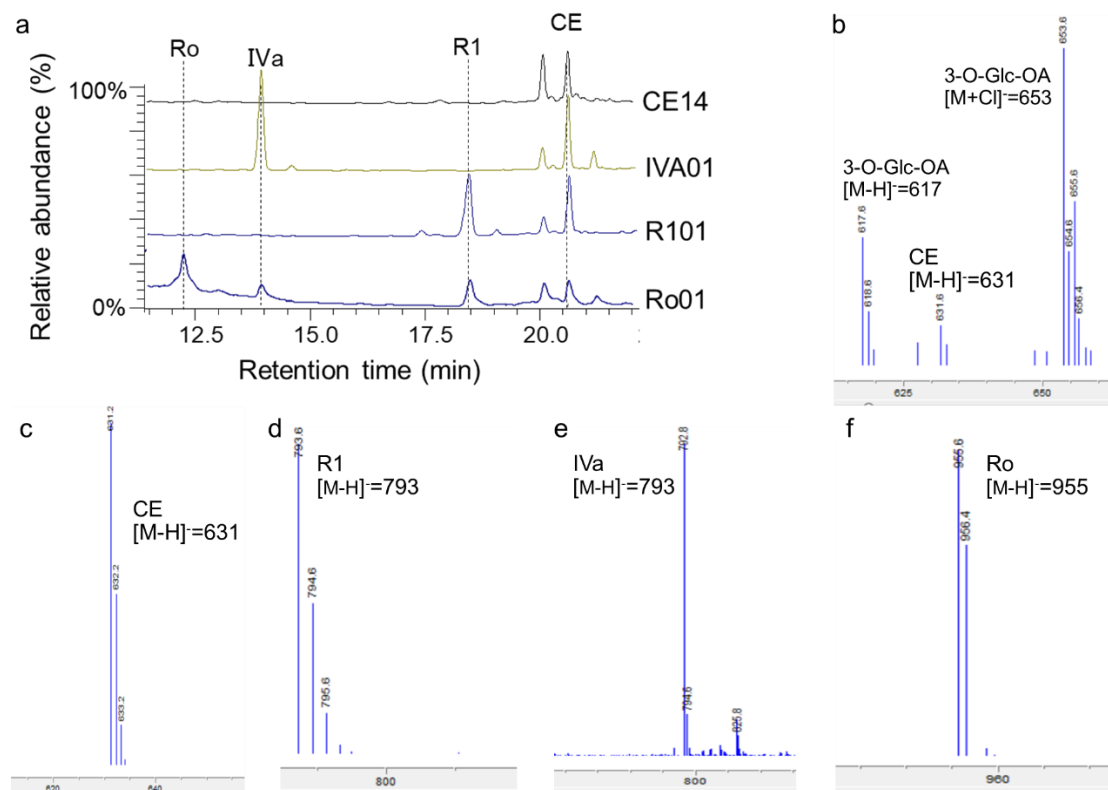


Fig. S3 UPLC/LC-MS results of the fermentation products of engineered strains. **(a)** UPLC results of the fermentation products of strains CE14, IVA01, R101 and Ro01. **(b)** LC-MS result of fermentation products of strain CE01. **(c)** LC-MS result of fermentation products of strain CE09. **(d)** LC-MS result of fermentation products of strain R101. **(e)** LC-MS result of fermentation products of strain IVA01. **(f)** LC-MS result of fermentation products of strain Ro01.

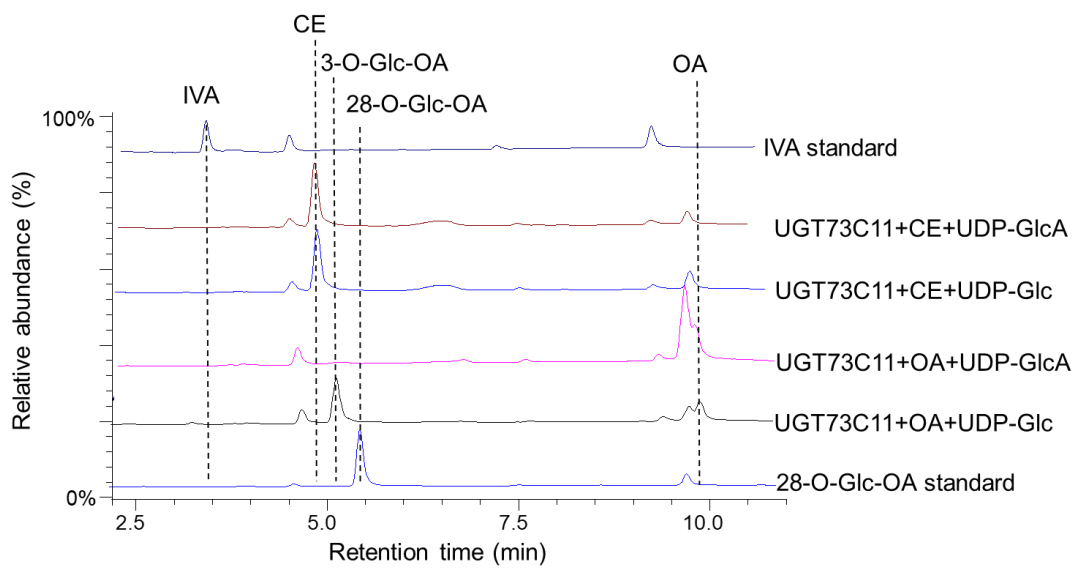


Fig. S4 UPLC results of the function of UGT73C11 toward OA and CE.

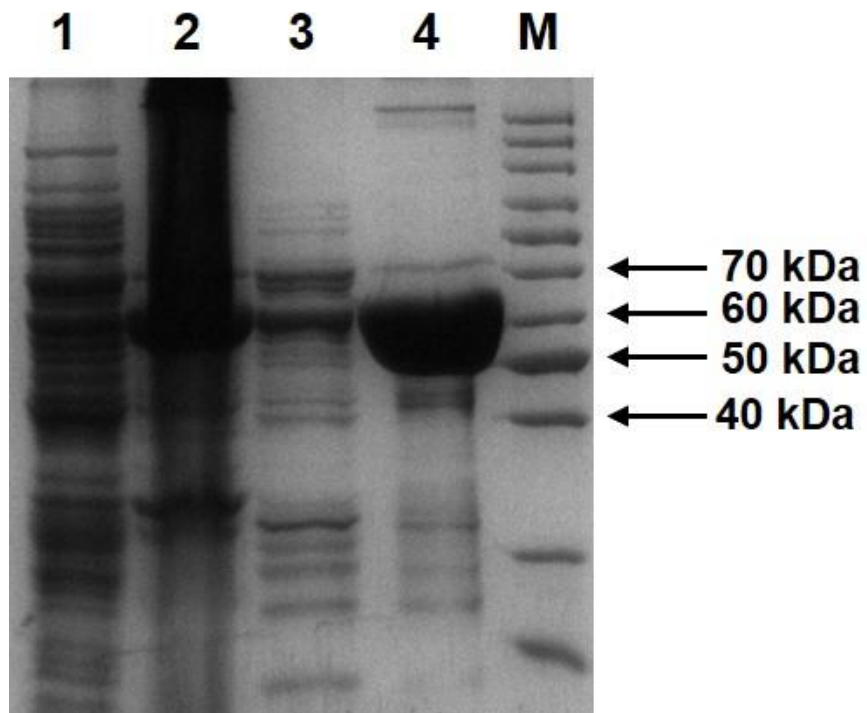


Fig. S5 SDS-PAGE analysis of UGT73F3 (1: supernatant, 2: sediment, 3: 5% imidazole elution, 4: 20% imidazole elution, M: marker).

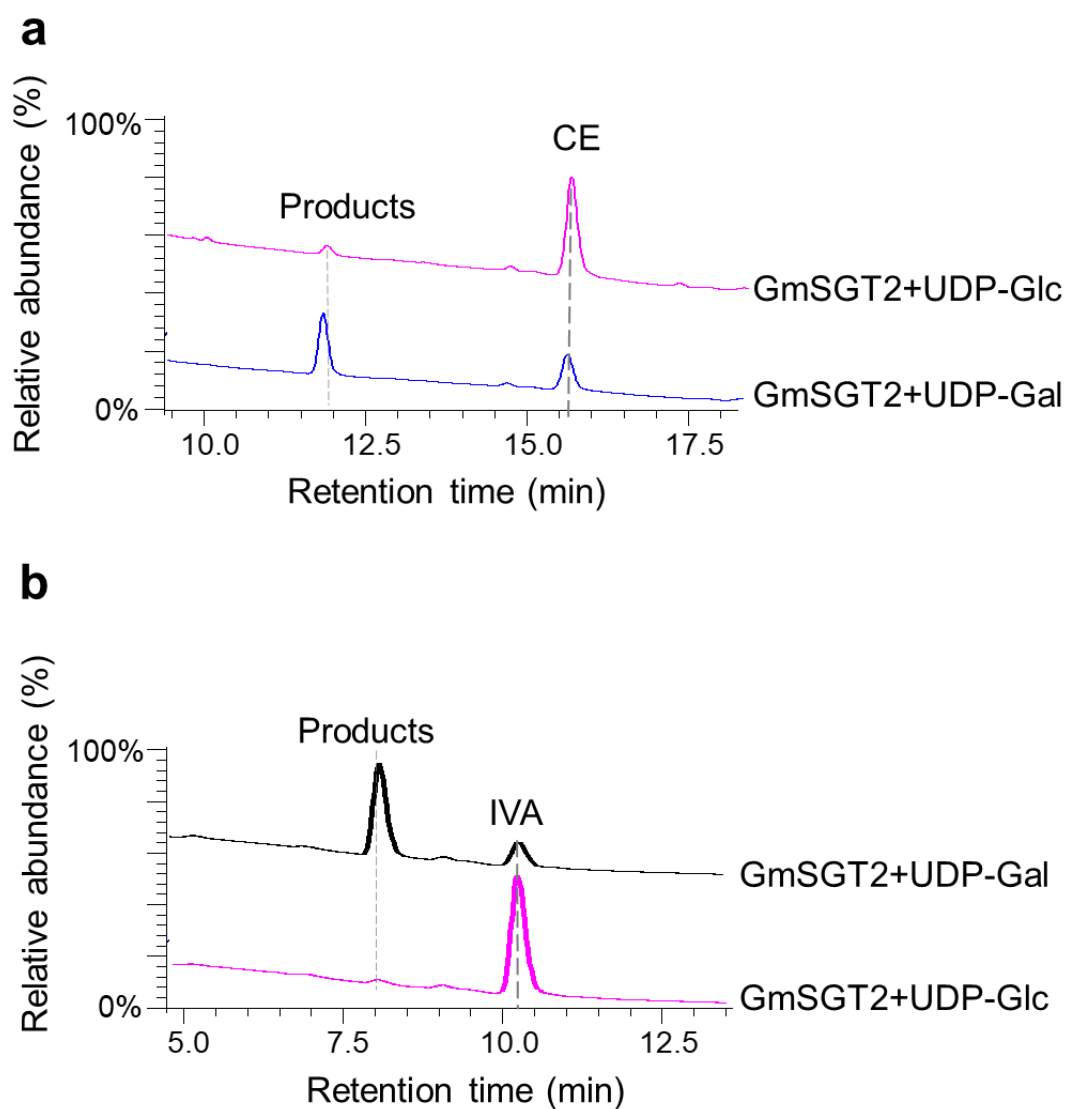


Fig. S6 UPLC results of GmSGT2 enzymatic assays. **(a)** UPLC results of the reaction of CE with UDP-Glc/UDP-Gal catalyzed by GmSGT2. **(b)** UPLC results of IVA with UDP-Glc/UDP-Gal catalyzed by GmSGT2.

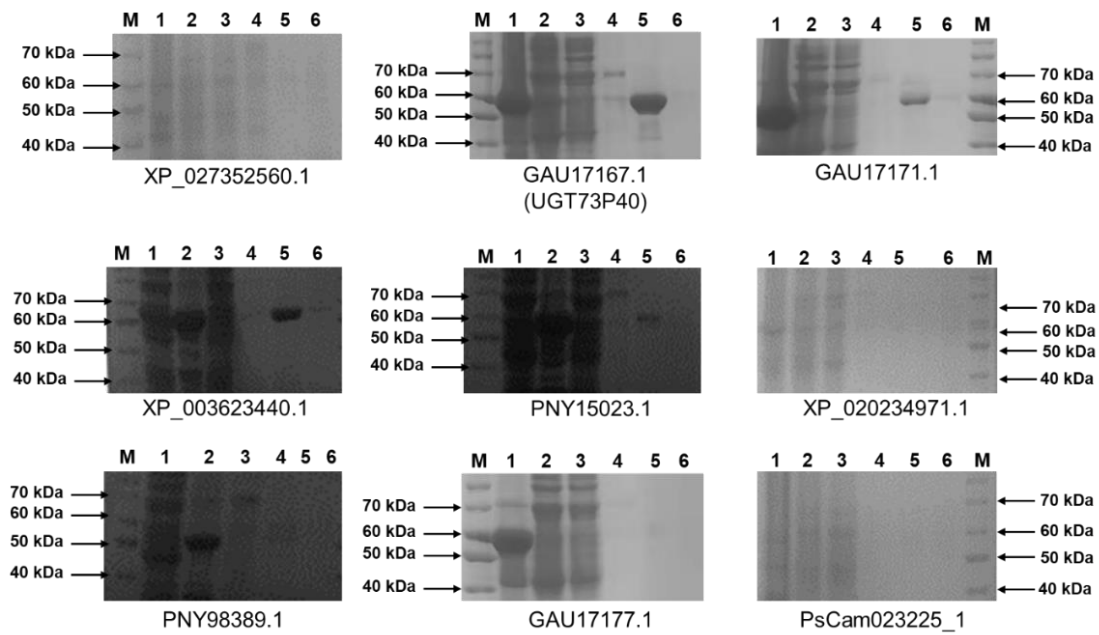


Fig. S7 SDS-PAGE analysis of nine UGTs candidates (M: marker, 1: sediment, 2: supernatant, 3: flow through, 4: 5% imidazole elution, 5: 20% imidazole elution, 6: 100% imidazole elution).

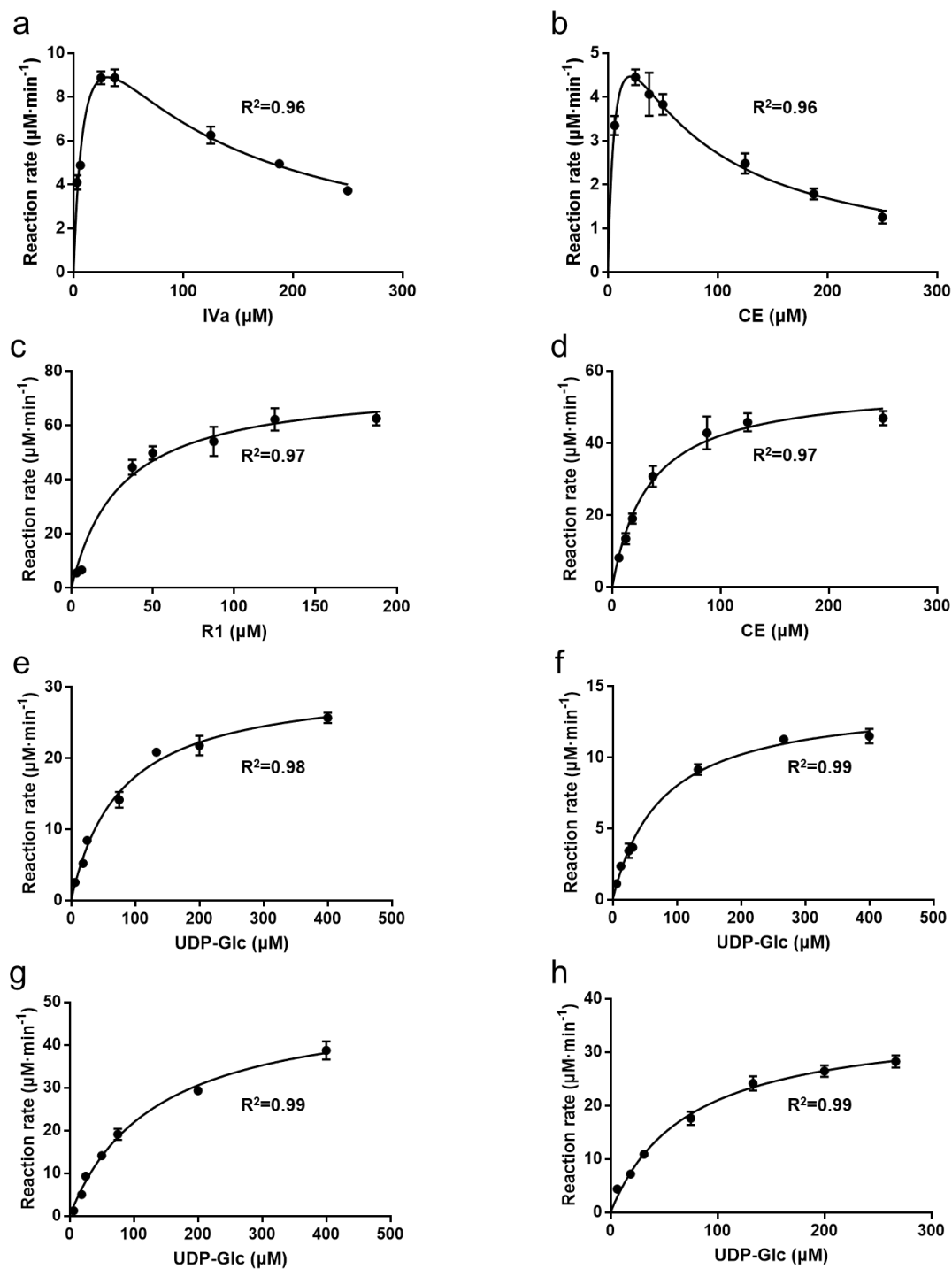


Fig. S8 Michaelis-Menten kinetics characterization of UGT73P40 and UGT73F3. **(a)** UGT73P40 towards IVa. **(b)** UGT73P40 towards CE. **(c)** UGT73F3 towards R1. **(d)** UGT73F3 towards CE. **(e)** UGT73P40 towards UDP-Glc with IVa as the acceptor. **(f)** UGT73P40 towards UDP-Glc with CE as the acceptor. **(g)** UGT73F3 towards UDP-Glc with R1 as the acceptor. **(h)** UGT73F3 towards UDP-Glc with CE as the acceptor. All data are presented as means \pm SD of three independent repeats.

Gene sequences obtained in this study:

Pn007110

ATGGCAACCCTTTCTCTTCATACATGCACTGTTTCAGCAACCTCGTGCAATAATCAACAG
AATTCATATATTTTTCCACTTCATAGCCATAATTGGCCTAGTCTATTACAGAATCTCTAAC
ATCTTTTATGGTGATGGAAATGTACCAGTTTTAGCATGGGGCCTTATGACCACGGCTGA
GCTTATATTTACCTTCCTTTGGTTTTCTCACGCAAGCATTTCGCTGGAGGCCAGTGGCTCG
GACAATTTATCCAGAAAATCTTCCGGGGGATGAGGAGTTACCAGGGGTTGATGTGTTC
ATATGCACAGCTGATCCCAATAAAGAGCCAACCTATAGAGGTGATGAACACGGTGTGTC
AGCTATGGCACTTGACTATCCACCGGATAAGCTCACCGTGTATCTGTCTGATGATGGTG
GTGCAGCCTCGACACTGGATGCGATAACAGAAGCATGTTCAATTTGCAAAGTCTTGGCT
ACTTTTTTGTAGAAAATATGGGATCAAGACCAGATGTCCTGAAGCATATTTTTCTAAATT
TGGCGATGATGAACGTCTTCTGAGGAGCGATGAATTC AAGGCAGAGGAAGAAAAGAT
TAAGTCGAAATATGATCTGTTCAAGATAAATGTTGAAAAAGCTGCCATGGAAGATGCCG
TTGTTTCATGATCGACCCCTCACATCAAGGTAATACATGATAATAGCAAGGATGTAAAC
AATGATGACGTCCAAAATAAGATGCCTCTTCTTGTATGTATCCCGTGAGAAAAGACC
ATCTCGTCTCATCGTTTCAAGGCTGGCGCCCTCAATGCGCTTCTTCGAGTTTCTGGCA
TAATGAGCAATGCCCCCTACATAATAGTGTTAGATTGTGATATGTA CTGCAATGACCCTG
CCTCAGCCC GACAAGCAATGTGCTTTCATCTGGACCCTGAGATGTCTCGAACTCTCTCT
TATGTACAATACCCTCAAATTTTTCTATAATGTTAGTGAAAATGATATATACGACGGCCAA
GCAAGATCAGCATAACAAGACAAGGTATCAAGGTATGGATGGACAACGAGGTACAGTTT
GTGCAGGA ACTGGTTATTACTTGAACAAAAAGGCATTATATAGTAGTCCCAATCAAGAG
GTGAAA ACTCGTCAATGTTTATATAGCTTTCCTTTTCCGTTGATTTACAGACAAGTTT
CTGCTCGAAC CAGAAAAGAATTTTGGTATATCTAGCGTGTTTCATTGCCTCACTAAAAGG
TAGTGACAAAGAAGATACATCTGGAGTGAGAGATTTATCAGAAGCAATTATAGAAGAG
GCCAAAATTCTGGCCTCTTGTGACTTTGAGAAAAACACACAATGGGGTAAAGAGATTG
GCTACTCCTATGATTGTTTGTGGAGAGCACATTCACGGGGTATCTTTTGCACAACAAA
GGATGGCGTTCTGTCTATCTTTATCCAAAAAGGCCATGTTTCTTGGGTTGCACCACCATT
GACATGAAGGATGCTTTGGTCCA ACTCATGAAATGGTCTCTGGATTGCTTCAAGTTGG
TCTCTCTAAGTTTAGCCCTGCCATATATGGTATGTCAAGGATGACCATTCTCCAAACTAT
GTGCTACGGGTATTTACGTTTTTACCCCTTTTCTCAGTTGCTTGCCTATTATATGGAATT
ATTCCTCAGCTATGTTTCTTGAGTGGCACTCCCTTGTACCCAAAGGTCTCAGACCCATG
GTTTTTGGCATTGCAACAATTTATACATCATCCCTTTGTCAGCATTGTACGAAGTCCT
CTCTAGCGGCGGCACAATTAGGACATGGTGGAAATGAACAAAGAATTTTTATGATAAAGT
CAGTTACAGCTTTGCTGTTTGGATGCCTAGATGTTCTAATGAAGTGGGTGGGCATAACA
AAAGCGAGTTTCAGGCTGACAAATAAAGCCGTCGACCAAGAAAAGCTCGAAAAATAT
ATGAAAGGCAAATTTGACTTCAGTGGTGCAAAGATGTTTCATGATTCCACTGACATTATT
AGTCATACTGAACGTGCTATGCTTTATAGGTGGAGTGAAAAGGGTAATAGGCGAGAAA
AACTTTGAAGAGATGTTTGTCTCAGGCTTTTCTTTCTTCTTCTTCAATTCTAGTTCTCAGTTTT
CCTATACTCAAGGGGCTCATACCAAAGAGAAACAAGTGA

Pn007111

ATGAGCAAGGCAATGGAAAATATTCTTCTCCTCCTCCATTCATGCCAAGTTCATAAGCTTCC
TCTAATTATCAACAGATTCATTCCCTTATTCCATTCATAGCGTTATCATTCCTAGTCTATT
ATAGAGTTTCCTTTATCTCCCAAACCCTAGGAACAAAAACACACCAATATTATCCCAT
CTTCTTGTTTTTGCTTCTGAGATCATACTCTCCTTTATTTGGCTCCTCAACCAAGCTTAC
CGATGGCGTCCCGTCTCCCGGAAGGTCTCCCGGATAGATTACCGGAAGATAAAGAAC
TTCCGGCCATCGACGTGTTTATATGTACTGCTGATCCCAATAAGGAGCCAACCTTTGGGA
GTGATGAATACTCTTATACTGCCATGGCGTTGGATTATCCAGCTGATAAACTCAATGTT
TATCTTTCCGATGATGGGGGTTCTCCGGTGACTTTGAAAGCCATGAAGGAGGCTTGGAA
AATTTGCCAAGTGGTGGCTGCCCTTCTGTAGGAGATATGACATTAAGACTAGCTGCCCT
GAGGCTTATTTTTCGGCTTTGGAGGAGCATGATAATTATGGGGATTTTCATGGCAGTGA
ATTTGTAGCTGACAGAAAAATTATCAAGGAAAAATATGAAGCGTTCACGGAGAGTGTG
ATGAGAATTAGAGACAGTGAAATCACCATACTTAATCGCGATCATTCTTCTGTTGTTGA
GGTGATACATGATAAAGCCGCCTTAGACCCAAGTGGGATGCCTCTTCTTATTTATGTATC
GCGTGAAAAAGACCAGTTCATCATCATCATTTTAAAGCTGGAGCTCTGAATGTTCTTC
TAAGAGTCTCGAATATGATAAGCAATTCGCCATACATTCTAGGCCTGGACTGTGACATTA
AGACGGACATATACGATAGTCAGTTGAGAACCTATTTTGAGGTACTATGGCTAGGTATG
GATGGACTAAAAGGACCTCCCATATCTGGTACATGCTATTATATAAAGAGGGAGGTATTC
TATGGAAATCAATGCAAGATGATTTTGACGTCAAGGAACTTAAACATATTTTTGGTAC
ATCAAATGACTTCATCAAATCCCTTACTGGGAACAATAAAGCTAGTACCATTGATGATC
AGAACTTAAACCAGTGCATTGCTAGAAGAGAGCAAACCTTGTGGCCTCTTGTGCTTACGA
GAATGATACAAATTGGGGCAAAGAGGCAAGTTTCCGGTATTTTTCCGTGGTAGAAGATT
ACTTTACAGGATTCAACTTGCATTGCCAAGGTTGGAGATCTGTGTATTTGATCCACCA
AGGCCAGCATTTTTGGGTGCCGGCACCAAAATCTTAGTGATGTATTGGTTCAAAGCAC
GCGATGGAGTGCCGGGTTGGTGGAAGTTGCCATTTCAAAGTTCCTCCCTCTTGTTTATG
GTTCACTACGAATGCCTATTCTTGAGGCCATGTGTTATGCAGAGCTTGCCTCCTTCCCT
TCTATTTCTTCCCCCTATGGTGTCTAGCTACCATCCCTCCATTATGTCTACTTAATGACAT
TCCTTTGTTCCCAAAGTTTCAAGTTCATTTTTCGTCCTATTCTCATTAAATTTGTCTCT
TCGCAATTGAAACACATAAAGAGATTTTCTCTAATGGAGGGTCATTCCAACTTGGAG
GAATGAACAGAGAATGTGGATGATTAAGTCCGTCACATGTCACCTTTATGGAAGCTTGA
ATGCAATTATGGAGAAAATGGGTATGAGAGAAGCTAGCTTCTTACCAACAAACAAAGT
CGTTGATGACGAACAACCTTAAACGATACCAAATGGGCATATTTGATTTCCAAGCATCGA
CTATGTTGCTTGCTCCGCTTTGTACCTTAAATCGCCTTAAATTTAGCTGCTGTAGTTGGGG
GAGTTGCAAAGCAGTTCTTGCAAGGAGCTGCAATGAGATGTTTATTCAAATTTTCCCT
TCATTTTTCGTGGTGACTTTGAATTATCCGATCATTGAAGGTATGATACTGAGGAAGGAT
AAAGGTCGCATTCACCTCCGTTACTCAATTATCTATCGCTATTACCATAATTATCTTGA
CATGTGGAAAGTTTATTCTCAAGTGA

Pn016322

ATGAATGAGGCAATGGGAATGGATAAACTCTTCTCCTCCTCCATTCATGCCAAGTGCAGAA
GCTTTCTCTAATCATCAACAGGCTCCATACTTTCTCCATTTTGTAGCTCTATTATTTTAT
TTCATTACAGAGTTTCTTTCATCTCCCAAACCCTAAGAACAAAAACACACGCCAATCTTA
CCCATGTTCTTGTGTTGCTTCTGAGCTCATACTCTTTTTATTTGGCTCCTCAACCAA

GCTTACCGATGGCGCCCCGTCTCTCGAAAGGTTTTCCCCGAGCGATTGCCGGAAGATA
AACAACTTCCGGCCATCGACGTGTTTATATGTACGGCTGATCCCAATAAGGAGCCA
TTGGGAGTGATGAATACTCTTATATCTGCCATGGCGTTGGATTATCCGGCTGATAAGCTG
AATGTTTATCTTTCGGATGATGGGGTTCTAAGGTGACTTTGAAAGCCATGAAGGAGGC
TTGGAGATTGCCAGGTGGTGGCTGCCCTTCTGTAAGAGATTGGTGTAAAGACTAGGT
GCCCTGAGGCTTATTTTTCTGCAACAGAGGAGCAAGATGATGGTAATTATGGGGATTTT
CTTACTAGCGAGTTTGTAGCCGAGAAAAATATTGTTTCAGGAAAAATATGAGGAGTTCAC
AGGGAGTGTGATTAGAATTAGAGAAAAATGAAATCCCCATACTTAATATAGATCATCCTTC
GGTTGTTGAGGTGATCCATGACAAATCCTTGGACCAAATTGTGATGCCTCTCCTTGTTT
ATGTTTCGCGTGAGAAAAGACCAGCTTATCCTCACCATTTCAAAGCTGGAGCACTGAA
CGTCTTCTGAGAGTCTCGAGCATGATAACAATTCAGCATACTACTAGGGCTGGATT
GCGACATGTATTGCAATGATCCAATTCAGCACGCCAAGCAATGTGTTTCCTTCTTGAC
CCCAAGATATCTCCCTCATTATCTCATGTCCAATTCCTCAAAAATTCCACAATATTAGC
GAGACCGACATCTACGATAGTCAACTGAGAGACGTTTTCAAGGTACGATGGCCGGGTA
TGATGGACTCAAGGGACCAACCATAAGTGGTACATGCTATTATATAAAGAGAGAGAC
GTTCTATAAAAGATCTATGCAAGATGATATAAATGATATCGATCTCAAGGAACTCAAACA
TATTTTTGGTACATCAAATGAGTTCATCAAATCCCTTCACCGGAAATACAACCCTAGTAT
CACTGATGATGGAAAATTCACCAACGTATTGCTAGAAGAGAGCAAGCTTCTAGCCTCTT
GTGCCTACGAAACCAATACCAAATGGGGAAAAGAGGCAAGTTTTTCGATATTTTTCCGT
GATAGAAGATTATTTTACAGGATTCAATTTGCACTGCAAGGGTTGGAGATCTGTGTATCT
TGATCCACCAAGACCAGCATTTTTGGGTGCTGGCACCACAAACCTTGACGATGTCTTG
GTTCAAGGCACGCGATGGACCGCTGGGCTGGTGGAAAGTTGCCATTTCAAAGTTCAACC
CTCTTATTTATGGTTCACTCAGAATGTCTATTCTTGAGGCCATGTGTTATGCAGAGCTTG
CTTTATACCCCTTCTATTTTTTCTCACTATGGTGTCTAGCTACCATTCCCTCAGCTCTGTCT
ACTCAATGACATCTCTCTGTACCCAGAGGTTTCAAGTTCATTTTTTCGTTCGTATTCTCATT
AATATTCGTCTCATCCCAATTTAAACACATACAAGAAATTTCTTTACCGGAGGGTCGTT
CCAAACTTGGAGGAATGAACAGAGGATGTGGATGATAAAGTCCATTACATGTTATTTTT
ATGGAAGCTTGAATGCGATTATGGAGAAAATTGGTATGAGAGAAGCCAGCTTTTTACCA
ACCAACAAGGTTGTGGATGATGAACAACCTTAAACGATACCAAATAGGTAAGTTTGATTT
TCAAGCATCGACTATGTTACTTGCTCCACTATGTACCTTAGTCACCCTAAATTTGGCTTC
CTTAGTCGGGGGAGTTGTGAAAGCGATTCTTACAAAAGTTGCAATGAGATGTTTATTC
AAATTTTCTCTCATTTTTTTGTGGCAACTCTGAGCTATCCAATCATTGAAGGTATGATATT
GAGAAAAGATATGGGTCGCATTCCACCTTCTGTTACTCGATTATGTATCGTTTTAACTAT
ACTTTTGTGACTTCTGGAAAATTTATTGTCAATTGA

Pn022859

ATGGCTTCCACAACCTCTCTCCTCCACACAGCCACCGTACATAAACTTCATGCCACCAT
GAACAGAACTCACATCCTCTTCCACTTCTTCTCATCCTTTCCGTTTTTTACTATAGACT
ATTCTCTCTCCTCAACTCCACGCCGGACCACAACCTTCCGGCGACGTCATGGATTTTCA
TGATTATCGCCGAACCTACTAACCTTTGCATGGTCTTTCGGTCAATCTTTCGGTGGA
GGCCGGTGTCCCGGTCCGTGCACCCGGAAGACTGCCCGGAGACAAATCGCTGCCCG
GGGTGGATGTTTTTATCTGTACGGCTGATCCGAAGAAGGAGCCAACCCTAGAGGTGAT
GAATACGGTGATATCGGCTATGTGCTTGGACTACCCGCCGGAGAAGCTTGCCGTCTATC

TTTCCGATGACGGTGGCGCTGCGGTGACCTTGACTGGAATGAGGGAAGCTGCTGTGTT
TGCTAGGTGGTGGATTCCGTTTTGCCGGAAGTACAATATCAAGACAAGGTGCCCCGGG
GCTTATTTTTTCAGCTTTAACCGACGGCAATGAACAAGATATTCGTGGTGGTGAATTTCA
AGTTGAACAAGACAAAATCAAGTTAAAGTATGAACATTTCAAGGAAGTTGTGGAGAG
TGGAGGAGATGATGCTCAAGCTGATGACATGAACAACATTACCGGTGATGGCCGTGCA
CCTCTAGTTCAGGTGATTCATGATAGTAGTAGAGGCAATGAAGATGATGAGAATGATGA
GAATGATGAGCAGATGCCACTCCTTGTATATGTTTCTCGTGAAAGAAGACCATCTTACC
CTCATCGTTTTCAAGGCCGGAGCTCTCAATGCCCTTCTTCGAGTGTCTGGAATAATTAGT
AATTTCCCATACGTAATGGTATTAGATTGCGACATGTATTGCAATGACCCAACCTCGGCT
AAGCAAGCAATGTGCTTCCACCTTGATCCAAACATGTCTTCCGAATTAGCTTTCGTTCA
ATACCCTCAAGTCTTCTACAATGTCACCAGCAATGACATCTATGATGGCCAAGCGAGAT
CTGCCTTCAAGACCAAGTGGCAAGGTATGGATGGACTAAGAGGACCTGTTTTGACTGG
CACTGGCTACTACCTAAAGAGAAAGGCATTATTCGGCAGCCCTAATGATCAAGAAGTTA
CTTATCTCCATCAACCTGCATTAAGTTTTGGTAATTCGGACATGTTCAATTGCGTTCGCTAC
ATGGTAACAACCAGAAACATATTGTCAATAACGGAGAAACAACGGAGGAGCAGCAGC
TAAAGGAAGCACATGTGTTGGCAAGTTGCAGCTTTGAGAAAGGCACAAAATGGGGCT
CCGAGATAGGTTACACGTACGAGTGCCTGTTGGAGAGCACTTTCACCGGCTATTATCTG
CAATGCAAGGGATGGAATCAGTGTACTGCTACCCAAAGAAACCAGCATTCTTGGGTT
GTGCCACAATTGATATGAAGGATGCATTGGTTCAACTAAGGAAATGGAATGGAAATCTC
CTAGAACTTGGATTCTCAAATTCAGTCCACTGACCTATGCAATTTCAAGAATGTCATT
TCTTCATAGCATGTGTTACGCATACTTTGCCTTCCAGCCCTTGACTTCCCTTTCACTCCT
CATATATGGCATTGTTCCCTCAGTTTTGCTTCTTCAAAGGCATCGCTCTTTTCCCAAAGGT
AACAGATCCATGGATTGCGGCGTTTTGCATTTGTATGGGTATCTTCACTTTTGCAGCACTT
GTATGAAGTTTTGTCGAGTGGAGGCTCATTACGGACATGGTGAACGAACAAAGAATT
TGGGTGATCAAAGTTGTGTCCGGAAGCTTATTCGGATGCATGGATTGTGTTAGGAAGA
AGATTGGCATTTCAAAACACTGTGTTAGGTTGACAAATAAAGCTATTGAAAAAGAGAA
ACTCGAGAATTATGAGAAGGGAAAATTTGATTTCCAAGGGGCACAGATCTTCATGGTT
CCATTGGTTGGACTAGTCATTTGGAACATGGCTTGTTTGATAGGGGGGATATGGAGAAT
GTTTGTGGGGGAAATAGTAGTAAATGTTTGGACAACTTTTCCTCTCTTCCTTCCTTT
CACTCGTTAGTTATCCCATCATAGGAGGGATGCTGACCAGAAAAGGAAATTA

PsCam023225_1 (UGT73P48, *E. coli* codon-optimized)

ATGGATAGTCAGCAGCTGCATGTGGTGTGTTGTGCCGGCCTGACCCCGGGCCATATGAT
TCCGATGATTGATAACCGCGCGCCTGTTTGCGAAACATGGCGTGAACGTGACCATTATTA
ACACCCATGCGAACGCGAGCACCTTTCAGAAAACCATTGATAGCGATTTTAAACAGCGG
CTATAGCATTAAAACGCAGCTGATTCATTTCCGAGCGCGCAAGTGGGTTTACCGGATG
GCATTGAAAACCTTAAAGATGCGACGAGCCTGGAAATTTTTGGCAAATGATGCAAGG
CAACCTGATGTTACAAGATCAGATTGAAATCTGCTGCAAGATCTGCAGCCGGATTGCA
TTATTAGCGATATGACCTATCCGTGGACCGTGGAAAGCGCGGCGAAACTGAACATTCCG
CGCATTTATTTTTATAGCAGTAGCCTGTTTAGCAACTGCGCGAGCTATTTTCTGCGCAA
TATCGCCCGCATGATGATCTGGCGAGCGATAACGAGAAATTTACCATTCCGTGCCTGCC
GCATACCATTTGAAATTACGAGCCTGCAGCTGGCGCGCAGCACCCGCGTGAGCAATAGC
ACCACGAGCGCGTTTTGGCGAAATGTTTGAAGCGAAAAACGCAGCTATGGCACCTG

TATAACAGCTTTCATGAACTGGAAAGCGATTATGAAAAACTGAGTCAGACCACCATTG
GCATTAAGCTGGAGCGTGGGCCCGGTGAGCACCTGGATTAACAAAGATGATGAAA
ACAAAGCGAACCCGCCCATATTGAAAAAGAAGGCGAACTGCTGAACTGGCTGAACA
GCAAAGCAACGAAAGCGTGCTGTATGTGAGCTTTGGCAGCCTGACCCGTGTGAGTC
ATGAGCAGCTGGTGGAAATTGCGTATGGCCTGGAAAACAGCAGCCATAACTTTATTTG
GGTGGTGC GCGAAAAAGATAAAGATGAAGATGGCGAAGGCTTTCTGCAAGATTTTGA
AGAACGCATGAAAGAAAGCAAAAAAGGCTATATTATTTGAACTGGGCGCCGAGCTG
CTGATTCTGGATCATCCGGCGACCGGCGGCATTGTGACCCATTGCGGCTGGAACAGCA
TTCTGGAAAGCCTGAGCGTGGGCCTGCCGATGATCACCTGGCCGATGTTTGC GGAACA
GTTTTATAACGAAAAACTGCTGGTGGATGTGCTGAAAATTGGCGTGAGCGTGGGCAGC
AAAGTGAACAAATTTTGGCTGAGCATTGGCGAAGATCATGTGGTTGTGCGCCGCGAAG
ATATTGCGAAAGCGGTGGAAATTCTGATGGGCAGCAGCCAAGAAAGCAAAGAAATGC
GCATTGCGCGGAAAAACTGGGCGATGCGGCGAAACGCACCATTGAAGAAGGCGGGCG
ATAGCTATTATAACCTGATTGAGCTGATTGATGAACTGAAAAGCCTGAAAAAAGCAA
AGCGCTGGGCGAAAAAGCGGAT

Reference

Xie, Z. X., Li, B. Z., Mitchell, L. A., Wu, Y., Qi, X., Jin, Z., Jia, B., Wang, X., Zeng, B. X., Liu, H. M., Wu, X. L., Feng, Q., Zhang, W. Z., Liu, W., Ding, M. Z., Li, X., Zhao, G. R., Qiao, J. J., Cheng, J. S., Zhao, M., Kuang, Z., Wang, X., Martin, J. A., Stracquadiano, G., Yang, K., Bai, X., Zhao, J., Hu, M. L., Lin, Q. H., Zhang, W. Q., Shen, M. H., Chen, S., Su, W., Wang, E. X., Guo, R., Zhai, F., Guo, X. J., Du, H. X., Zhu, J. Q., Song, T. Q., Dai, J. J., Li, F. F., Jiang, G. Z., Han, S. L., Liu, S. Y., Yu, Z. C., Yang, X. N., Chen, K., Hu, C., Li, D. S., Jia, N., Liu, Y., Wang, L. T., Wang, S., Wei, X. T., Fu, M. Q., Qu, L. M., Xin, S. Y., Liu, T., Tian, K. R., Li, X. N., Zhang, J. H., Song, L. X., Liu, J. G., Lv, J. F., Xu, H., Tao, R., Wang, Y., Zhang, T. T., Deng, Y. X., Wang, Y. R., Li, T., Ye, G. X., Xu, X. R., Xia, Z. B., Zhang, W., Yang, S. L., Liu, Y. L., Ding, W. Q., Liu, Z. N., Zhu, J. Q., Liu, N. Z., Walker, R., Luo, Y., Wang, Y., Shen, Y., Yang, H., Cai, Y., Ma, P. S., Zhang, C. T., Bader, J. S., Boeke, J. D., Yuan, Y. J., 2017. "Perfect" designer chromosome V and behavior of a ring derivative. *Science*. 355, 4704-4710.