

Supplementary materials for

**Sustainable supply of wilforic acid C via decompartmentalization of
peroxisomal acetyl-CoA and systematic engineering in yeast**

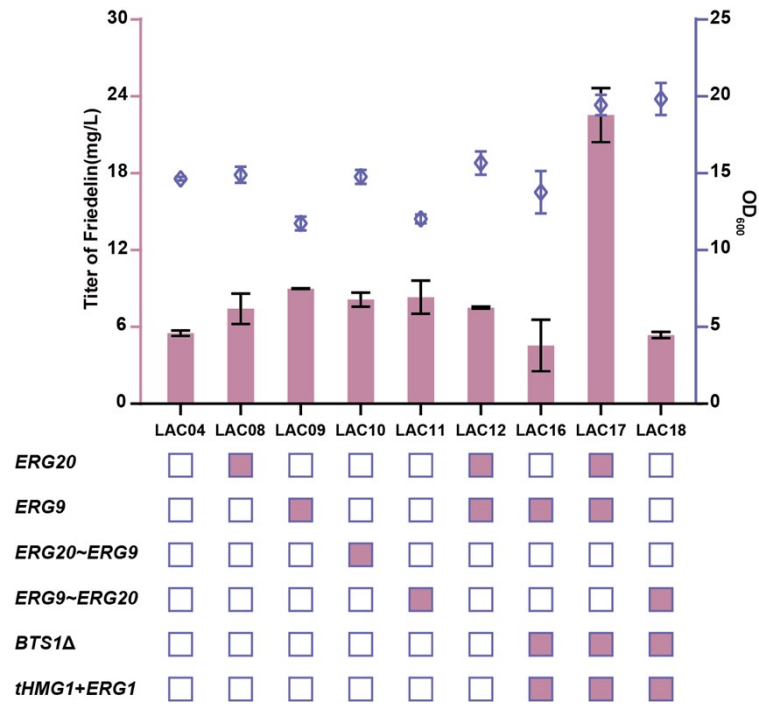


Fig. S1† The production of friedelin in engineered *S. cerevisiae* strains harboring MVA pathway modifications. OD₆₀₀ indicates the final yeast biomass after fermentation. “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. All data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

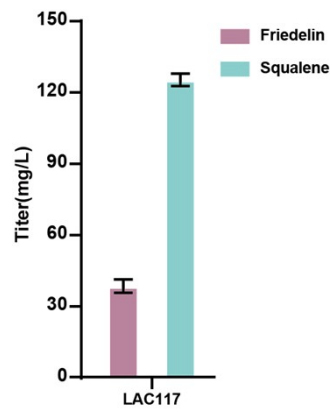


Fig. S2† The production of friedelin and squalene in strain LAC117. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

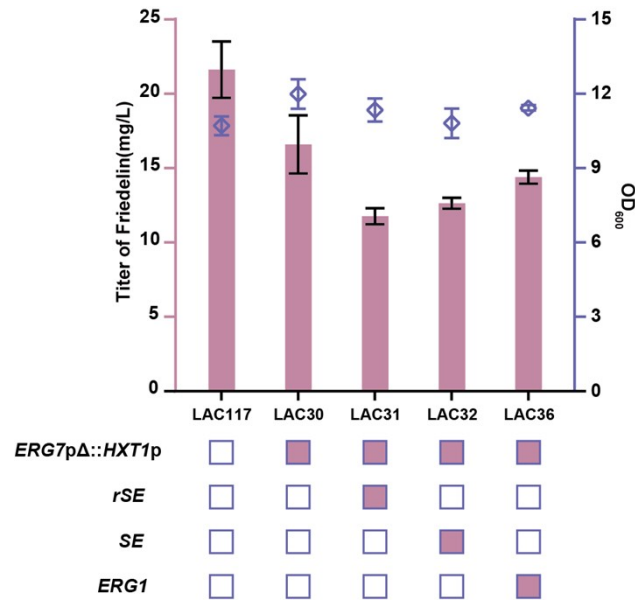


Fig. S3† The production of friedelin in strains with ERG7 promoter replacement and squalene epoxidase expression, including native ERG1 and squalene epoxidase in *Rattus norvegicus* (SE) and truncated SE(rSE). “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

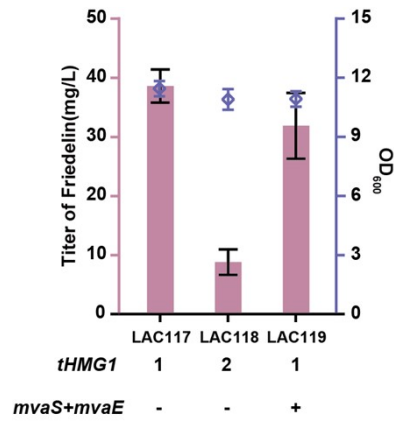


Fig. S4† Effect of increasing *tHMG1* copy number or introducing heterologous *mvaS* and *mvaE* in LAC117 on friedelin production. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

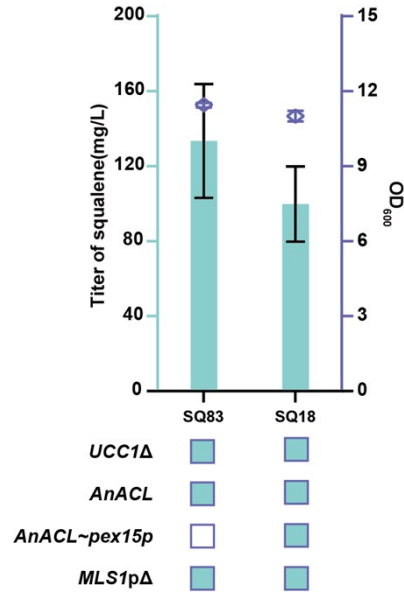


Fig. S5† The production of squalene in strains SQ18, with coexpression of *AnACL* and *AnACL~pex15p*, and deletion of *MLS1p*. “●” indicates that the specified genetic modification was applied, and “○” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

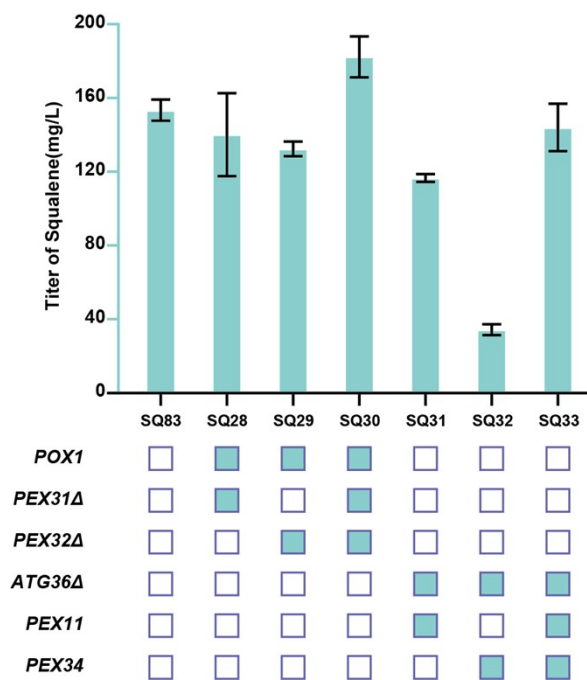


Fig. S6† Effects of peroxisomal metabolic and morphological engineering on squalene production. “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

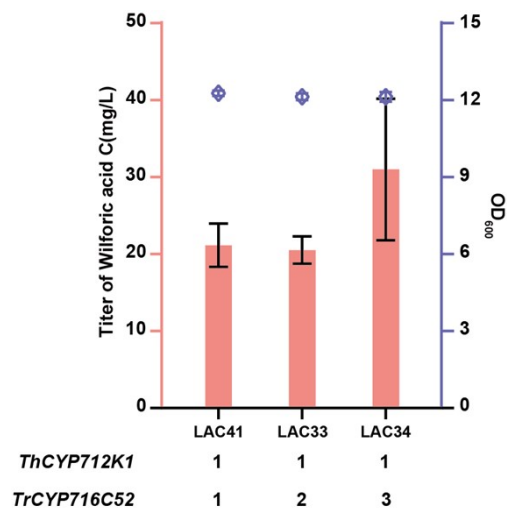


Fig. S7† Effect of increasing copy number of TrCYP716C52 in strain LAC41. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

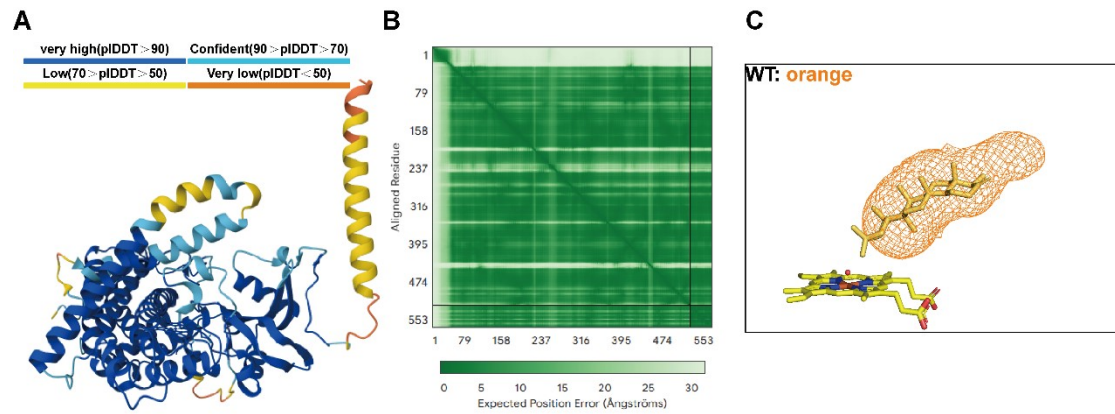


Fig. S8† AlphaFold3-predicted structures of wild-type ThCYP712K1 (A), expected position error (B), and substrate-channel analysis performed using CAVER 3.0 (C).

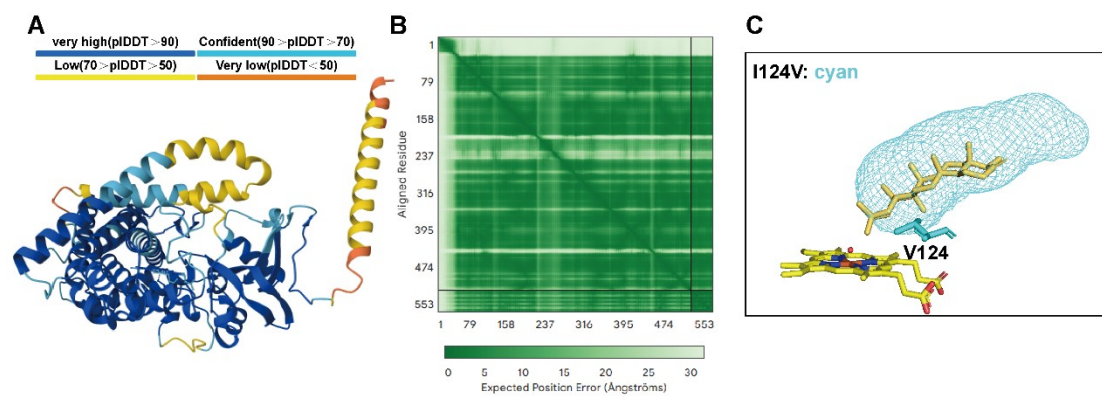


Fig. S9† AlphaFold3-predicted structures of ThCYP712K1^{I124V} (A), expected position error (B), and substrate-channel analysis performed using CAVER 3.0 (C).

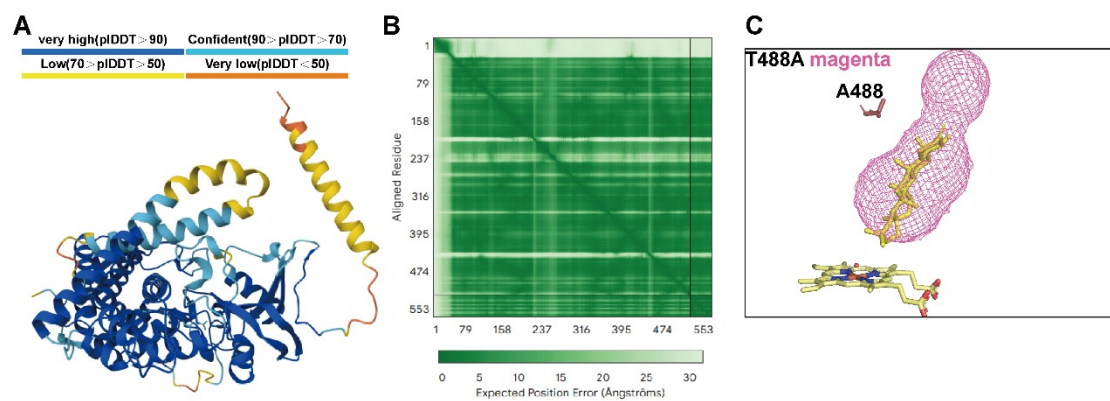


Fig. S10† AlphaFold3-predicted structures of ThCYP712K1^{T488A} (A), expected position error (B), and substrate-channel analysis performed using CAVER 3.0 (C).

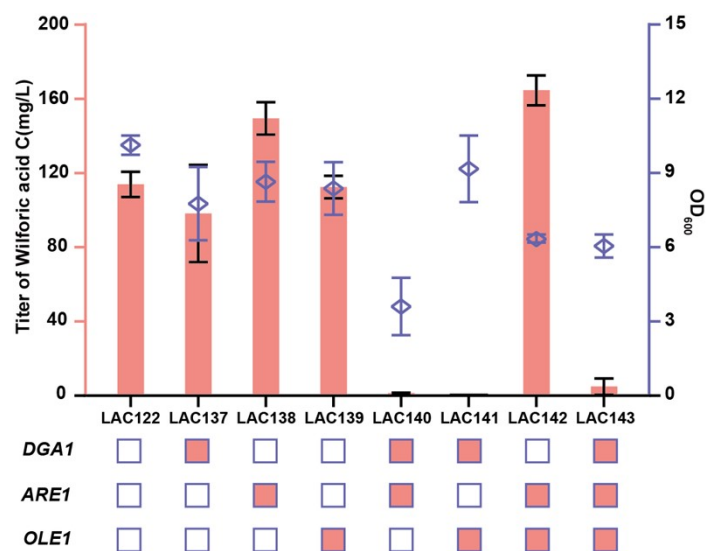


Fig. S11† Effect of LD formation on wilforic acid C production in LAC122. *DGA1*, *ARE1* and *OLE1* were overexpressed in strain LAC122 individually or in combination. “●” indicates that the specified genetic modification was applied, and “○” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

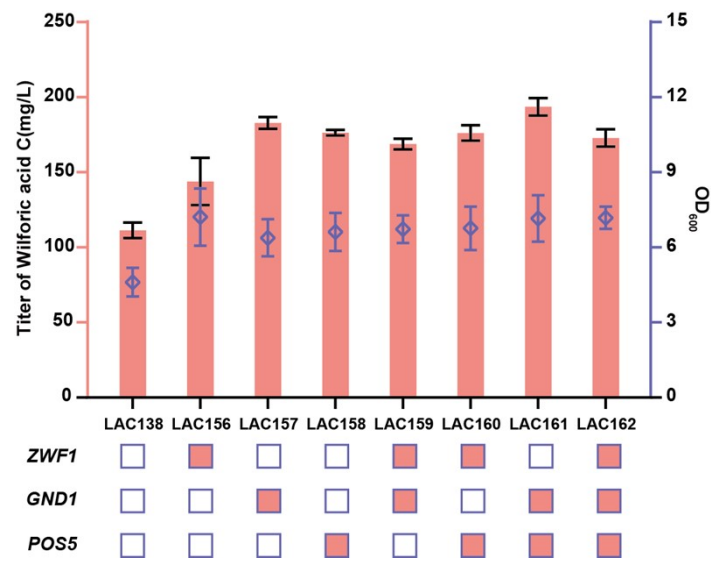


Fig. S12† Effect of NADPH supply on wilforic acid C production in LAC138. *ZWF1*, *GND1* and *POS5* were overexpressed in strain LAC138 individually or in combination. “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

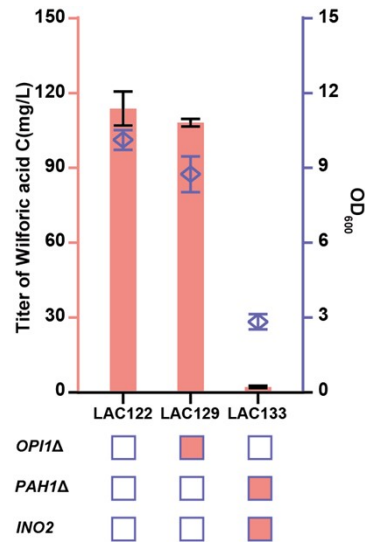


Fig. S13† Effect of ER morphology engineering on wilforic acid C production. The genes *OPI1*, encoding a transcriptional regulator which negatively regulates phospholipid biosynthesis, and *PAH1*, encoding Mg^{2+} dependent phosphatidate (PA) phosphatase were deleted. *INO2* encoding a transcription activator for phospholipid biosynthesis was overexpressed in strain LAC122. “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation.

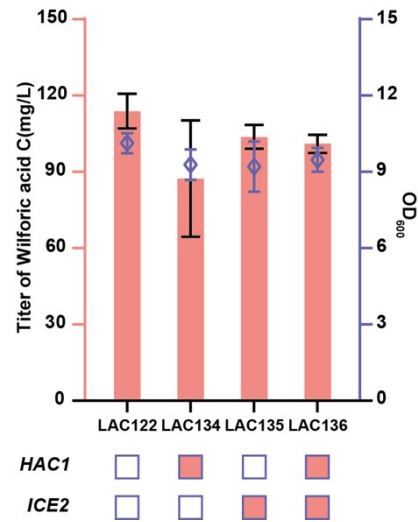


Fig. S14† Effect of maintaining ER stability on wilforic acid C production. *HAC1* and *ICE2* were overexpressed in strain LAC122 individually or in combination. “○” indicates that the specified genetic modification was applied, and “●” indicates no modification. OD₆₀₀ indicates the final yeast biomass after fermentation. Cells were grown in defined minimal medium, and cultures were sampled after 72 h of growth for detection. Data represent the mean of n=3 biologically independent samples, and error bars show standard deviation

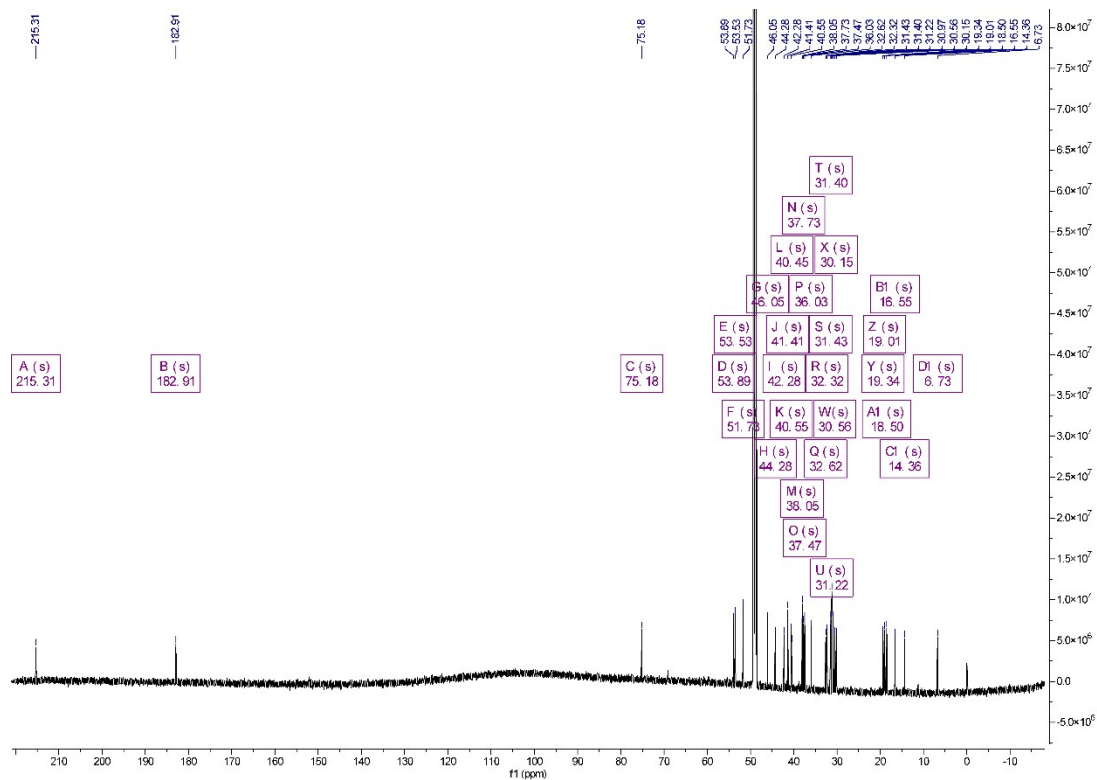


Fig. S16† The ^{13}C NMR spectrum of wilroric acid C (CD_3OD , 201 MHz). δ 215.3, 182.9, 75.2, 53.9, 53.5, 51.7, 46.1, 44.3, 42.3, 41.4, 40.6, 40.5, 38.1, 37.7, 37.5, 36.0, 32.6, 32.3, 31.4, 31.4, 31.2, 31.0, 30.6, 30.2, 19.3, 19.0, 18.5, 16.6, 14.4, 6.7.

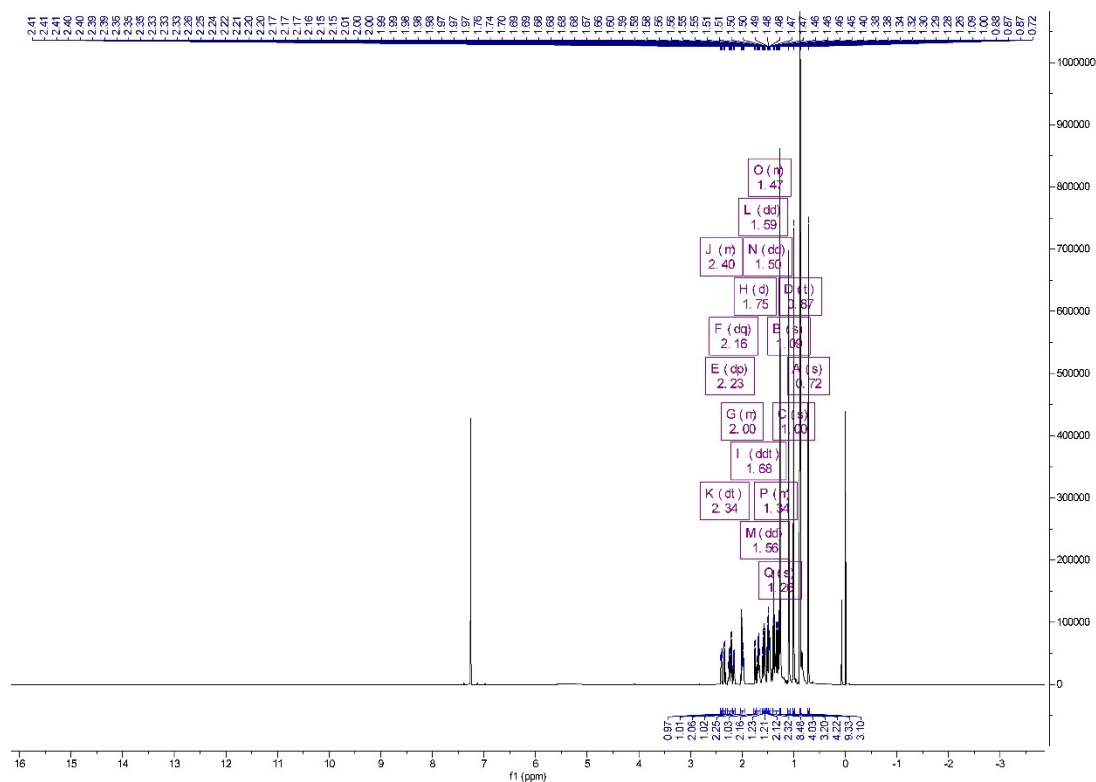


Fig. S17† The ^1H NMR spectrum of polypunonic acid (CDCl_3 , 800 MHz). δ 2.43 – 2.38 (m, 1H), 2.34 (dt, $J = 15.1, 2.2$ Hz, 1H), 2.23 (dp, $J = 27.9, 7.0$ Hz, 2H), 2.16 (dq, $J = 14.3, 2.3$ Hz, 1H), 2.03 – 1.96 (m, 2H), 1.75 (d, $J = 12.7$ Hz, 1H), 1.68 (ddt, $J = 15.8, 12.8, 6.3$ Hz, 2H), 1.59 (dd, $J = 15.1, 7.8$ Hz, 1H), 1.56 (dd, $J = 7.9, 2.6$ Hz, 1H), 1.50 (dd, $J = 12.8, 3.3$ Hz, 2H), 1.49 – 1.45 (m, 2H), 1.40 – 1.28 (m, 8H), 1.26 (s, 4H), 1.09 (s, 3H), 1.00 (s, 4H), 0.87 (t, $J = 5.5$ Hz, 9H), 0.72 (s, 3H).

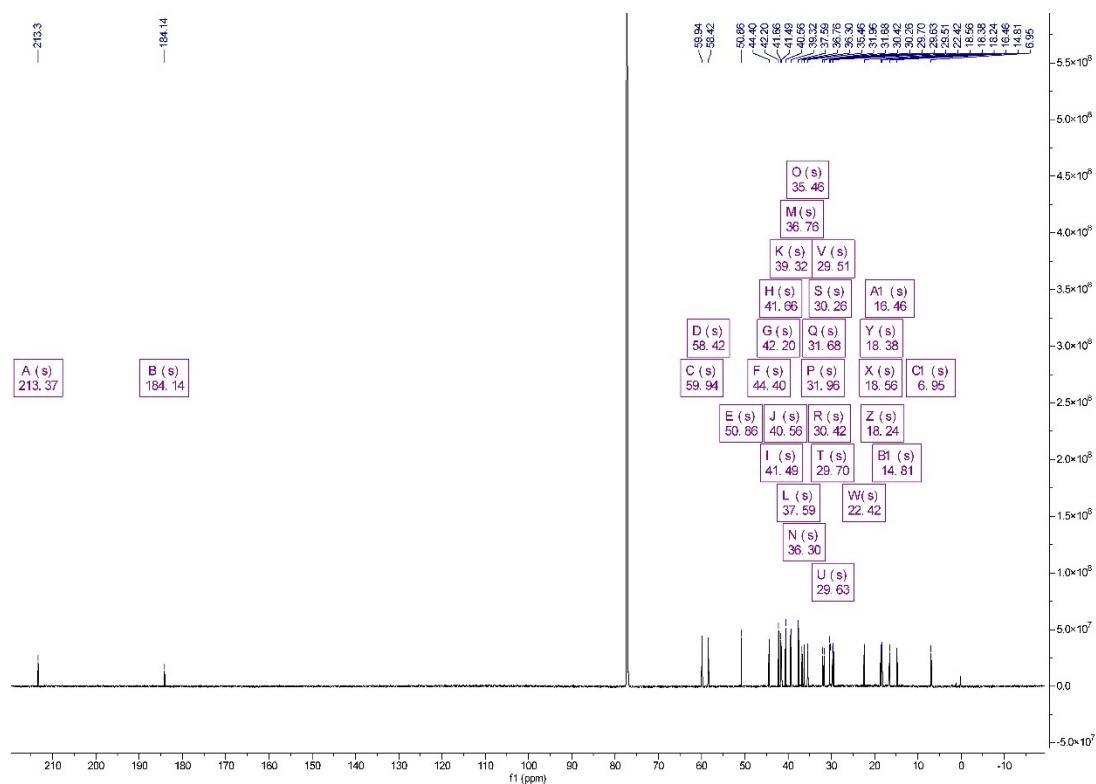


Fig. S18† The ^{13}C NMR spectrum of polpunonic acid (CDCl_3 , 201 MHz). δ 213.4, 184.1, 59.9, 58.4, 50.9, 44.4, 42.2, 41.7, 41.5, 40.6, 39.3, 39.3, 37.6, 36.8, 36.3, 35.5, 32.0, 31.7, 30.4, 30.3, 29.7, 29.6, 29.5, 22.4, 18.6, 18.4, 18.2, 16.5, 14.8, 7.0.

Table S1† List of yeast strains in this study

Strain	Parent strain	Genotype	Reference
CEN.PK113-5D		MATa ura3-52 trp1-289 leu2-3, 112 his3Δ1 MAL2-8c SUC2	Retained in the laboratory
THY02	CEN.PK113-5D	CAN1Δ:: pTEF1-CAS9-tCYC1, Gal80Δ	Retained in the laboratory
LAC01	THY02	X-2::IDP1t-MiFRS4-gal10,gall-TwCPR3-FBA1t	This study
LAC02	THY02	X-2::IDP1t-KdFRS-gal10,gall-TwCPR3-FBA1t	This study
LAC03	THY02	X-2::IDP1t-KdGLS-gal10,gall-TwCPR3-FBA1t	This study
LAC04	THY02	X-2::IDP1t-TwOSC1 ^{T502E} -gal10,gall-TwCPR3-FBA1t	This study
LAC05	THY02	X-2::TEF1p-IcOSC3-ADH1t-TDH3p-TwCPR3-IDP1t	This study
LAC08	LAC04	X-3:: TDH3p-ERG20-FBA1t	This study
LAC09	LAC04	X-3:: TDH3p-ERG9-FBA1t	This study
LAC10	LAC04	X-3:: TDH3p-ERG20~ERG9-FBA1t	This study
LAC11	LAC04	X-3:: TDH3p-ERG9~ERG20-FBA1t	This study
LAC12	LAC04	X-3:: TDH3p-ERG20-CYC1t-PGK1p-ERG9-FBA1t	This study
LAC14	LAC12	BTS1Δ	This study
LAC16	LAC09	BTS1Δ::THD3p-tHMG1-FBA1t-TEF1p-ERG1-IDP1t	This study
LAC17	LAC12	BTS1Δ::THD3p-tHMG1-FBA1t-TEF1p-ERG1-IDP1t	This study
LAC18	LAC11	BTS1Δ::THD3p-tHMG1-FBA1t-TEF1p-ERG1-IDP1t	This study
LAC116	LAC17	XII-1::TDH3p-ERG12-IDP1t-TEF1p-ERG8-FBA1t	This study
LAC117	LAC116	XII-5::TDH3p-ERG10-IDP1t-TEF1p-ERG13-FBA1t-tHXT7p	This study
LAC118	LAC116	XII-5::TDH3p-ERG10-IDP1t-TEF1p-ERG13-FBA1t-tHXT7p-tHMG1-ADH1t	This study
LAC119	LAC116	XII-5::TDH3p-mvaE-IDP1t-TEF1p-mvaS-FBA1t-tHXT7p	This study

LAC92	LAC117	UCC1Δ:: IDP1t-AnACLa-gal1,10-AnACLB-FBA1t; MLS1pΔ	This study
LAC93	LAC117	pex31Δ	This study
LAC114	LAC93	XII-2::gal1,10-TwOSC1 ^{T502E} -IDP1t	This study
LAC120	LAC93	XII-4::FBA1t-TwOSC1 ^{T502E} -gal1,10-TwOSC1T502E-IDP1t	This study
LAC121	LAC114	XII-4::FBA1t-TwOSC1 ^{T502E} -gal1,10-TwOSC1T502E-IDP1t	This study
LAC122	LAC121	XI-3::IDP1t-ThCYP712K-SptGal2-SkGal2-TrCYP716C52- FBA1t-gal7-TwCPR3-TDH2t	This study
LAC123	LAC122	int14::IDP1t-ThCYP712K1 ^{I124V} -SptGAL2-SkGAL2- TrCYP716C52-FBA1t	This study
LAC124	LAC123	YPRC delta15::IDP1t-ThCYP712K1 ^{I124V} -SptGAL2- SkGAL2-TrCYP716C52-FBA1t	This study
LAC125	LAC124	int17::IDP1t-ThCYP712K1 ^{I124V} -SptGAL2	This study
LAC126	LAC124	int17::IDP1t-ThCYP712K1 ^{I124V} -SptGAL2-SkGAL2- ThCYP712K1 ^{I124V} -FBA1t	This study
LAC127	LAC121	XI-3::IDP1t-ThCYP712K ^{I124V} -SptGal2-SkGal2- TrCYP716C52-FBA1t-gal7-TwCPR3-TDH2t	This study
LAC129	LAC122	OPIΔ	This study
LAC133	LAC122	PAH1Δ::TDH3p-INO2-IDP1t	This study
LAC134	LAC122	XI-2::TDH3p-HAC1-IDP1t	This study
LAC135	LAC122	XI-2::TEF1p-ICE2-FBA1t	This study
LAC136	LAC122	XI-2::TDH3p-HAC1-IDP1t-TEF1p-ICE2-FBA1t	This study
LAC137	LAC122	V-1::TDH3p-DGA1-FBA1t	This study
LAC138	LAC122	V-1::TEF1p-ARE1-FBA1t	This study
LAC139	LAC122	V-1::tHXT7p-OLE1-ADH1t	This study
LAC140	LAC122	V-1::TDH3p-DGA1-IDP1t-TEF1p-ARE1-FBA1t	This study
LAC141	LAC122	V-1::TDH3p-DGA1-IDP1t-TEF1p-OLE1-FBA1t	This study
LAC142	LAC122	V-1::TEF1p-ARE1-FBA1t-tHXT7p-OLE1-ADH1t	This study
LAC150	LAC127	V-1::TEF1p-ARE1-FBA1t	This study

LAC156	LAC138	III-2::TDH3p-ZWF1-IDP1t	This study
LAC157	LAC138	III-2::TEF1p-GND1-FBA1t	This study
LAC158	LAC138	III-2::tHXT7p-POS5-ADH1t	This study
LAC159	LAC138	III-2::TDH3p-ZWF1-IDP1t-TEF1p-GND1-FBA1t	This study
LAC160	LAC138	III-2::TDH3p-ZWF1-IDP1t-tHXT7p-POS5-ADH1t	This study
LAC161	LAC138	III-2::TEF1p-GND1-FBA1t-tHXT7p-POS5-ADH1t	This study
LAC162	LAC138	III-2::TDH3p-ZWF1-IDP1t-TEF1p-GND1-FBA1t-tHXT7p-POS5-ADH1t	This study
LAC168	LAC150	III-2::TEF1p-GND1-FBA1t-tHXT7p-POS5-ADH1t	This study
LAC24	LAC117	int14::gal1,10-TrCYP716C52-FBA1t	This study
LAC33	LAC24	XI-3::IDP1t-ThCYP712K1-gal1,10-TrCYP716C52-FBA1t	This study
LAC107	LAC33	int17::gal7-ATR1-TDH2t	This study
LAC108	LAC33	int17::gal7-MTR1-TDH2t	This study
LAC109	LAC33	int17::gal7-SmCPR1-TDH2t	This study
LAC110	LAC33	int17::gal7-TwCPR3-TDH2t	This study
LAC42	LAC33	int17::gal7-AtCYB5-TDH2t	This study
LAC30	LAC117	PERG7Δ::PHXT1	This study
LAC31	LAC117	ERG7pΔ::tHXT7p-rSE-CYC1t-HXT1p	This study
LAC32	LAC117	ERG7pΔ::tHXT7p-SE-CYC1t-HXT1p	This study
LAC36	LAC117	ERG7pΔ::tHXT7p-ERG1-CYC1t-HXT1p	This study
LAC41	LAC117	XI-3::IDP1t-ThCYP712K1-gal1,10-TrCYP716C52-FBA1t	This study
LAC25	LAC24	17 site::gal1,10-TrCYP716C52-FBA1t	This study
LAC34	LAC25	XI-3::IDP1t-ThCYP712K1-gal1,10-TrCYP716C52-FBA1t	This study
SQ17	LAC17	TwOSC1 ^{T502E} Δ	This study
SQ37	SQ17	UCC1Δ	This study
SQ76	SQ17	UCC1Δ:: IDP1t-AnACLa-gal1,10-AnACLB-FBA1t	This study
SQ77	SQ17	UCC1Δ:: IDP1t-AnACLa~pex15p-gal1,10-AnACLB~pex15p-FBA1t	This study
SQ78	SQ77	XI-5:: IDP1t-AnACLa-gal1,10-AnACLB-FBA1t	This study

SQ79	SQ76	MLS1pΔ::PCK1p	This study
SQ80	SQ76	MLS1pΔ::FBP1p	This study
SQ81	SQ76	MLS1pΔ::ICL2p	This study
SQ82	SQ76	MLS1pΔ::HXT13p	This study
SQ83	SQ76	MLS1pΔ	This study
SQ18	SQ78	MLS1pΔ	This study
SQ6	SQ83	CAT2Δ	This study
SQ14	SQ83	III-2::TDH3p-POX1-FBA1t	This study
SQ15	SQ83	III-2::TDH3p-POT1-FBA1t	This study
SQ16	SQ83	III-2::TDH3p-POT1-IDP1t-TEF1p-POX1-FBA1t	This study
SQ19	SQ83	Pex30Δ	This study
SQ20	SQ83	Pex31Δ	This study
SQ21	SQ83	Pex32Δ	This study
SQ22	SQ83	ATG36Δ	This study
SQ23	SQ83	int19::TDH3p-pex11-FBA1t	This study
SQ24	SQ83	int19::TDH3p-pex3-FBA1t	This study
SQ25	SQ83	int19::TDH3p-pex34-FBA1t	This study

Table S2† The nucleotide sequences of the genes identified in this study

Gene	Sequence
MiFRS4	ATGTGGAAAAATCAAAATCGCCGATAGAGGTAAGTCCCATACAACGAATACTTATACACT ACCAACGATTTCTGTCGGTAGACAAATCTGGGAATTCGATCCAACTCTGGTACATCTGAA GAATTGGCTGAAATTGAAGAGGTGAGAAGAAAGTTCACCGAAAACAGATACGAAGTTAA GCCAGCTTCTGATTTATTGTGGATGATGCAATTCCTAAGAAAGAACAACCTTCAAGCAAAC CATCCCACCATTAAGGATCGGTGAAAAGGAACAAGTCACCTACGAAGATGTCACTACTGC CTTGAGGAGAGCTTCTTTCTTTCAGTGCTTTGCAAGCCTCCGATGGTCACTGGCCAGCT GAAAATGCTGGTGTCTCTTTCTTCTTGCCACCATTCATCTTCTGTTGTACATTACCGGTCA TTTGAAGTCTATTATCACTCCAGAACATAGAAAGGAAATTTGAGATTCATTACAACCAC CAAAATGAAGACGGTGGTTGGGGTATCCACATCGAAGGTCACTCCACTGTTTTCGCTACT GCCTTCACTTACGTTTGTATGAGAATCTTAGGTGTTGGTCCGGATGAAGATGCTTGTGCTA

GAGCCAGAAAATGGATCTTGGACAGAGGTGGTATTACCTACATGGCTTCTCTGGGGTAAG
ACCTGGTTCTCTGTCTTGGGCATCTTTGACTGGTACGGTTGTAACCCAATGCCACCAGAAT
TCTGGATTTTGCCATCATATCTACCTATCCACCCAGCTAAAATGTGGTGTTACTGTAGAAT
GGTTTATATGCCAATGTCTTACTTATACGGTAAGAGATTCGTTGCCCAATTACCCCTTTG
ATCTTGCAACTTCGTGAAGAATTGCACACCCAACCATAACCATGAAATCGAATGGAGAAAAG
ATGAGACACAGATGTGCCAAGGAAGACTTGTACTTCCCACACTCTTTGATTCAAACTTCT
TGTGGGACTCCTTGTACGTGCATCCGAACCTTTGTTAACCAGATGGCCTTTCTCTAAGAT
CAGAGAACGTGCCTTGAAAAAGGCTATGGAACACATTCATTACGAAGATGAAAACTCTA
GATATATTACCATGGTTGTGTTGAAAAGGCCTTATGTATGTTGTGTTGGGTGGAAG
ACCCAAACGGTGAATATTTTAAGAAACACTTGGCCAGAATCCAGACTACCTCTGGGTTG
CTGAGGATGGTATGAAAGTTGTTTCCTTCGGTTCTCAATTGTGGGATGCCACCTTTGGTTT
CCAAGCTTTAGTCGCTTCCAACCTGACTGATGAAGTTGCTCCAACCTGGTTAAGGCTTAC
GATTTTCATCAAGAAGTGCCAGGTTAGAGATAATCCATCTGGTAACTTCGAAAAGATGTTT
CGGCATATCTCTAAGGGTTCCTGGACATTCTCCGACCAAGACCACGGTTGGCAATTGTCT
GACTGTACTGCTGAAGCTTTGAAGTGTTGTTGCTAGCTGCTACTATGCCAGAAGAATTG
GTCGGTGAAAAGTTGGATCCACAATGGATCTTTGAATCCGTCAACATCATCTTGTCTTGC
AAGAACCAAAGACTGGTGGTTTGGCTGGTTGGGAACCAGTTCGTGCTGGTCAATGGATG
GAAATGTTGAACCAATGGAATTTTGGAAAACATTGTGATTGAACACACTTACATTGAAT
GTACTGGTTCCTCTATTATCGCTTTCATTACTTTGAAGAAGTTGTTCCAGGTCACAGAACC
AAGGACATTGACAACCTTTATTGTTAACGCTATCAGATACTTGAAGACGAACAATACCTT
GACGGTTCCTTGGTACGGTAACTGGGGTATCTGTTTCATCTACTCTACTATGTTTGCTTTGG
GTGGGCTGGCAGCTGCTGGTAGAACTTACAAGAACTGCCAAGCTGTCAGAAGAGGTGTT
GACTTCTTGTTAATTAACCAATCTGACGACGGAGGTTGGGGTGAATCTTACATTTCTTGTC
CAAGAAAGAAGTACACCCCTCTAGAAGGTCGTAGATCCAATGTAGTCCAAACTGCATGG
GCTATGTTGGGTTTGTGTACGCTGGTCAAGCTGAAAGAGACCCAACCTCATTACACCGT
GGTGCTAAGTTGTTAATTAACCAATGGAAGAAGGTGGTTACCCACAACAAGAAATA
ACCGGTGTCTTCAAGATGAACTGTATGTTGCACTACCCAATCTACAGAAACGCTTTCCCAA
TCTGGGCTTTGGGTGAATACCGTAAGAGAGTTCCACTACCATCCAAGGGTAACTCCATGG
CCATGAAAATCAACTCCGCTTGA

KdFRS ATGTGGAAACTAAAGATTGCTGAAGGTGGTTCAGATCCATACATCTACACCACCAACAAC
TTTGTGGTAGACAAATCTGGGAATTTGATCCTCAAGCAACTGACCCACAACAATTGGCT
AAGGTTGAAGCTGCCAGATTGAACTTTTACAACCACAGACACAAGATTAAGCCATCCTCT
GACTTGTTGTGGAGATTACAATTTTTGGAAGAAAAGGATTCAGACAAAACATTGCTCAA
GTTAAGGTCGAAGACGGCGAAGAAGTCTCTTACGAAGCTGCTACTGCTGCTTTGAAGCG
GGGTGTTCACTTCTACTCTGCTTTGCAAGCCTCTGACGGTCACTGGCCAGCTGAAAATGCC
GGTCCAATGTTCTTCATGTCTCCACTCGTTATGTGTTTGTACATCACTGGTCACTTAAACAC
CATTTTCACTGAAGAACATAGAAGAGAAACATTGCGTTACATTTACTACCACCAAAACGA
AGATGGTGGTTGGGGTTTCCACATTGAAGGTCAATCCACCATGTTGGAAGTGTGTTTGA
CTACATCTGTATGAGATTGTTGGGTGAAGGTCCAGAAGGTGGTCAAGACAATGCTGTTTC
TAGAGGTAGAAAGTGGATTCTAGACCACGGTGGTGCTACCGCCATTCCATCTTGGGGTAA
GACCTGGTTGTCCATCATGGGTTTGTGTGACTGGTCTGGTTGTAATCCAATGCCACCAGA
ATTCTGGTTATTGCCATCCTATTTGCCAATGCACCCAGCTAAGATGTGGTGTACTGTAGA
ATGGTTTACATGCCCATGTCCTACTTGTACGGTAAAAGATTCACCACTCACATAACCCCTCT
AATCTTGCAATTGAGAGAGGAATTGCACACTCAACCATACGATCAAATTAAGTGAAGAA
GGTCCGTCACGTCTGTTGTAAGGAAGACACTTACTACCCACATCCAATCTTGCAAGATTG
ATCTGGGATACTCTCTACTTGACTACTGAACCATTGTTGACCAGATGGCCATTGAACAAGT
TGATCAGAGAAAAGAGCTTTGAAGAAGACTATGAAACACATTCACTACGAAGATGAAAAC
TCCAGATATATTGTTATCGGTGCCGTCGAAAAGGTCTTGTGTATGTTGGCCTGTTGGGTTG
AAGATCCAAACGGTGATTACTTCAAAAAGCACTTGGCAAGAGTACCAGATTACTTCTGGG
TTGCTGAAGATGGTATGAAGATCCAAAGTTTCGGTTCAGCACTGGGATACCGCTTTCTT
TGTTCAAGCCCTCTTGGCTTCTGACATGACTGACGAAATTAGAACTACCCTAGCAAAGGCT
CACGACTGTATCAAGAAAAGCCAAGTCAAGGACAACCCCTCCGGTGATTTCAGATCCATG
TACAGACATATCTCAAAGGTGCTTGGACTTCTCTGACCAAGATCACGGTTGGCAATTGT
CTGATTGTACTGCTGAAGGTTTGAAGTGCTGCTTGTGTTTCTTTGATGCAACCAGAAGT
TGTTGGTGAAGCTATGCCACCAGAACGTTTGTACGACTCGGTTAACGTTTTATTATATTTG
CAATCCAAGAACGGTGGTATGCCAGGTTGGGAACCAGCCGGTGAATCTGAATGGTTGGA
ATTGCTCAACCCAACCGAATTCTTCGAAAACATTGTCATCGAACACGAATACGTCGAATGT
ACCTCTTCTGCCGTCCAAGCTTTGGTCTTGTTCAAGAAATTGTACCCTTTACATAGACGAA

AGGAAGTTGAAAGATTCATTACCAACGGTGCTAAGTATTTGGAGGACATTCAAATGCCAG
ATGGCTCTTGGTACGGTAACTGGGGTGTTTGTTCACCTACGGTGCTTGGTTCGCCTTGGA
AGGTTTATCCGCCGCTGGTAAGACGTACAACAACGTGCTGCTGTTAGAAAGGGTGTTGA
TTTTTTGTTGAACATCCAATTAGAAGACGGTGTTGGGGTGAATCTTACCAATCTTGTTCCA
GACAAGAAGTACGTTCCATTGGAAGACAACAGATCTAACTTGGTCCAACTTCCTGGGCT
TTGATGGGCTTAATTTACGCTGGCCAAGCTGACAGAGACCCAACCTCCACTGCATAGAGCT
GCCAAGTTGTTAATCAACTCTCAATTGGAAGACGGTGACTTCCCACAACAAGAAATTGCT
GGTGTCTTCAAGATGAACTGCACCTTGCAATTCGCTGCTTACAGAAACATCTTCCCAATCT
GGGCTTTGGCTGTCTATCGTCGTTTCTGTAACCCAAATAGTGAAGCCATCTCGAAGCCATC
TAAGTGA

KdGLS ATGTGGAAGTTGAAGATTGCTGATGGTGGCTCCAACCCATACATTTTTACCACTAACAAC
TCGTCGGTAGACAAATCTGGGAATTCGATCCACAAGCCACCGATCCTCAACAATTGGCTA
AGGTTGAAGCTGCTCGTCTTGATTTCTACCACAACAGATACAAATTGAAGCCAAATTCTGA
TTTGTGTGGAGAATGCAATTCTTGAAGAAAAGGATTTAGACAAAACATCCCACAAGT
TAAGGTCGAGGATGGTGAAGAAGTTTCTTACGAAGCTGTCACCGCTGCTTTGCGTAGAG
GTGTTCACTTGACTCTGCTCTACAAGCTTCTGACGGACATTGGCCAGCTGAAAACGCTG
GTCCAATGTTCTTCATGCCACCAATGGTTATGTGTTTGTACATCACTGGCCATTTGAACGC
TATTTTCACTGAAGAACATAGATCTGAAACTCTAAGATACATCTACTACCACCAGAACGAA
GATGGTGGTTGGGGTTTCCACATTGAAGGTCACTCTACCATGTTCCGTACCGTCTTGAAC
ATATCTGTATGAGATTGTTGGGTGAAGGTCCAGAAGGTGGTCAAGACAATGCTGTCTCCA
GAGGTAGAAAATGGATTTTGGACCACGGTGGTGCTACTTCCATTCCATCTTGGGGTAAAA
CCTGGTTATCTATCATGGGTTTGTGCGATTGGTCTGGTTGTAACCCTATGCCACCTGAATT
CTGGTTGTTGCCATCCTACTTGCCAATGCACCCAGGTAAGATGTGGTGTACTGTCGTATG
GTCTACATGCCAATGTCTTACTTGACGGTAAGCGTTTCACCGCTAGAATTACTCCACTCA
TTTTGCAATTGAGAGAAGAAATTCACATTCAACCATACGATCAAATTGACTGGAAGAAGG
TCAGACACGTCTGTTGTAAGGAAGACATGTATTACCCACATCCTCTCTTGCAAGACTTGTT
GTGGGACACTTTATACTTGACCACCGAACCATTGTTGACTAGATGGCCATTAAACAAGTT
GATCAGACAAAGAGCTTTGCAAAAGACAATGAAGCACATTCATTACGAAGATGAAAATTC
GCGTTACATTACCATTGGTACTGTTGAAAAGGTTTTGTGTATGTTGGCTTGTGGGTTGAA

GACCCAAACGGTGACTACTTTAAGAAGCACCTGGCCAGAGTTCAGATTATTTCTGGGTT
GCTGAAGACGGCATGAAGATCCAATCTTTCGGTTCTCAACACTGGGATACGGTTTTCTCT
GCTCAAGCTTTGTTGGCTTCTGACATGGCTGACGAAATCGGTACTACTTTGGCCAAGGCT
CACTACTGTATCAAGGAATCTCAAGTCAAGGACAACCCATCTGGTGACTTCAGATCCATGT
ACAGACACATCTCCAAGGGTTCCTGGACTTTCTCAGACCAAGATCATGGTTGGCAATTGTC
CGACTGTACTGCTGAAGGTTTGAAATGCTGCTTATTGTTTTCTTGATGCAACCAGAAGTC
GTTGGTGAAGCTATGCCACCAGAAAGATTATTTGACTCCGTAAACATCTTGTTATATTAC
AATCCAAGAACGGTGGTATGCCTGGTTGGGAACCAGCCGGTGCCAGTGAATGGTTAGAA
TTGTTGAACCCAACCGAATTCTTCGAAAATATTGTCATTGAACACGAATACGTTGAATGTA
CCTCATCTGCGGTTCAAGCCTTGGTTTTGTTAAAAAGTTGCACCCAGGTCACAGAAGAAA
GGAAGTCGAAAGATTCATCACCAACGGTGCTAAGTACATCGAAGACATCCAAATGCCAG
ACGGTGCCTGGTACGGTAACTGGGGTGTGTTGTTTCACTTACGGTGCTTGGTTCGCTCTAG
GTGGTTTAGCCGCCGCTGGTAAGACCTACAACAACGTGCTGCCGTTAGAAAGGGTGTA
GATTTTTTGTACGGATTCAACTAGAAGATGGTGGTTGGGGTGAATCCTACCAATCTTGTC
CAGACAAGAAATACGTCCCATTGGAAGACAACAGATCTAACTTAGTCCACACTTCTTGGG
CTTTGATGGGTTTGTGTTCTGGTCAAGCCGACAGAGACCCGAACCCACTTCATCGTG
TGCAAAATTATTGATCAACTCCCAACTAGAAGATGGTGACTTCCCACAACAAGAAATCAC
CGGTGTGTTCAAGATGAACTGTATGTTGCACTTCGCCGTTACAGATCCATCTTCCAGTC
TGGGCTCTAGCAGAATACAAGAGATTCTGTAATCTGTCCTCTGAAGCTATCTCTAAGCCAT
CCAAGTGA

TwOSC1^T ATGTGGAAGTTGAAAGTTGCCGAAAGAGGTAACGCTCCATATTCTGAATACTTGTACACT
502E ACCAACGACTTCTCTGGTAGACAACTTGGGAATTTGATCCAAATGCTGGTACTCCACAA
GAATTGGCTAAAGTTGAAGAAGCCAGAAGAAAGTTCACCGAAGATAGACATACTGTAA
GCCAGCTTCAGATTTGTTGTGGATGATGCAATTCATGCGTGAGAAGAACTCAAGCAAAC
TATTCCACCAGTTAGGTTGGGTGAAGAGGAACAAGTTACTTACGAAGATTTGACTACTGC
TTTACTAGGACTACCAATTTCTTTACCGCTTTACAAGCTTCTGATGGTCATTGGCCAGCT
GAAAATGGTGGTGTTCCTTTTTCTTGCCACCATTCTTCAGCTTGACATTACTGGTCA
CCTGAACTCTATTATTACCCAGAGTACAGAAAAGAGATCCTGAGATTCATCTACAACCAC
CAAAATGAAGATGGTGGTTGGGGTATTCATATCGAAGGTCATTCTACTATGTTCCGGTACT

GCTTCTCTTACGTCTGCTTGAGAATTTTGGGTATCGAAGTTGATGGTGGTAAGGATAAT
GCTTGTGCTAGAGCTAGAAAGTGGATTTTGGATCATGGTGGTATTACCTATATGCCATCTT
GGGGTAAAACCTGGTTGTCTATTTTGGGTGTTTATGATTGGTACGGCTGTAATCCAATGC
CACCTGAATTTTGGTTGTTGCCATCTTACTTGCCAATTCATCCAGCTAAGATTTGGTGTTAC
TGCAGAATGGTTTACATGCCCATGTCTTACTTGTACGGTAAAAGATTCGTTGGTCCAATCA
CTCCATTGATCTTGCAATTGAGAGAAGAATTGCATACCCAACCATTCCACGAAATTCAATG
GCGTCAAACTAGACATAGATGTGCTAAAGAGGACTTGTACTACCCACACTCTTTGATTCAA
GATTTTCATCTGGGATTCTTGTACGTTGCTTCTGAACCATTATTGACTAGATGGCCATTGA
ACAAGATTAGAGAAAAGGCTTTGGCTAAGGCCATGGAACATATTCATTATGAGGACGAA
AACTCCAGGTACATTACCATTGGTTGTGTTGAAAAAGCCTTGTGTATGTTGTGTTGCTGGG
TTGAAGATCCAACTCTGACTACTTCAAGAAACACTTGGCTAGAATTCCAGATTACTTGTG
GGTTGCCGAAGATGGTATGAAGGTTCAATCTTTTGGTTCCCAATTGTGGGATGCTACTTTT
GGTTTTCAAGCTTTGGTTGCTTCTAACTTGACCGAAGATGAAGTTGGTCCAGCTTTGGCAA
AAGCTTACGATTTCATTAAGAAGTCCCAAGTCAAGGATAACCCATCTGGTGATTTTGAATC
CATGCATAGGCATATCTCTAAAGGTTCTTGACCTTCTCTGATCAAGATCATGGTTGGCAA
TTGTCTGATTGCACTGCTGAAGCTTTGAAGTGTTGTTTGGTGGCTGCTGAAATGCCACAAG
AAGTTGTTGGAGAAAAAATGAAGCCTGAATGGGTTTACGAAGCCATCAACATTATCTTGT
CCTTGCAATCTAAATCCGGTGGTTTGGCTGGTTGGGAACCTGTTAGAGCTGGTGAATGGA
TGGAATTTTGAACCAATGGAATTCTTGAGAACATCGTTATCGAACACACCTACGTTG
AATGTACCGGTTCTTCTATTATTGCCTTCGTGTCTCTGAAAAAGTTGTACCCAGGTCATAG
AACTAAGGACATCGATAACTTCATTAGAAACGCCATCAGGTACTTGGAAGATGTTCAATA
TCCAGATGGTTCTTGGTATGGTAACTGGGGTATTTGTTTCATCTACTCTACCATGTTTGCCT
TAGGTGGTTTAGCTGCTACTGGTAGAACTTACGATAATTGCCAAGCTGTTAGAAGAGGTG
TCGACTTCATATTGAAGAACCAGTCTGATGACGGTGGATGGGGTGAATCTTATTTGTCTT
GTCCAAGAAAGGTTTACACCCCATTGGATGGTAGAAGATCTAACGTTGTTCAAAGTCTT
GGGCTATGTTGGGTTTGTGTATGCTGGTCAAGCTGAAAGGGATCCAAGTCCATTGCATA
GAGGTGCAAAAGTTTTGATCAACTACCAGATGGAAGATGGCGGTTACCCACAACAAGAA
ATTACTGGTGTTTTCAAGATGAACTGCATGTTGCATTACCCATCTACAGAAATGCTTTTCC
AATTTGGGCTTTGGGTGAATACCGTAAAAGAGTTCCATTGCCATCTAAGGGTTACTAA

IcOSC3 ATGTGGAGATTGAAGGTCGGTGAAAGACCATCCAACCCACTGTTAAGATCTCCAAACGGT
TTCTTGGGTAGAGAAGTCTGGGAATTCGACCCAGCTGCCGGTACTCCAGAAGAAAGAGC
TGAGGTCGAAAGATTGAGACAAGAATACACCAGAAACAGATTCACTCAAAGAGAATGTG
GTGATTTATTGATGAGAATGCAATACACCAAGCAAACTTGACAAACACCAACTTGCAAG
CTATCAAGATCGATAAGGACTCCGATGTTACTGGGGAACTATCTTGATCGTTTTGAGAA
GAGCTCTTTCTCAATATTCTTCCCTACAAGGTCCAGACGGTCACTGGCCAGGTGACTACTC
CGGTATCTCTTTCATTTTGCCATTGATGATTTTCGCTCTACATGTCACCAGATCTTTGAACG
AAGTATTGTCTATTGAACACATCAGAGAAATTTGTAGATACATCTACAACATCCAAAATGA
AGACGGTGGTTGGGGTACCCACACTTTGGGCCCATCATCCATGTTCCGTACTTGTGTCAA
CTACGCCACCTTGAGAATTTGGGAGAAGTTTGGATGGTGAAAACGACGCTTTGTCCAA
GGGTAGAGCTTGGATCCTCTCTCACGGTTCTGCTACCGCCGCTCCACAATGGGCTAAAAT
CTACTTGTCTGTTATTGGTGCTTACAACCTGGGCCGGTAACAACGCTATCATCCCCGAATTG
TGGATGTTGCCACACTTCTTACCAATTCACCCAGGTAGATTCTGGTGTTTCTGTCGTATGG
TTTACATGCCAATGGCTTACATTTACGGAAAGAGATTTGTTGGTCCTATCACACCAACCAT
TTTGGCTATTAGAAACGAAATTTACAACATTCCGTACAATGAGATTAACCTGGGACAAAGC
TCGTAACCTCTTGCTAAGGAAGACTTGATCTACCGTCCATCTTGGTTGCAATCCATTGCT
ATTGCTTACTTGAACAAGTTCATCGAACCATTGTCTAATTTATGGCCGATGAACAAATTGA
GAGAACGTGCCATGTCTAACTTGATGGAACACATTCACTACGAAGATGAAACCTCAAACCT
ATGTTGGTCTTTGCCAATCAATAAGGCATTGAACATGATCTGTTGTTGGATTGAAAAGCC
AAACTCTAACGAATTCAGACAACATTTGCCAAGAATCCATGACTTCCTCTGGTTGGCTGAA
GATGGTATGAAATCTAAGGTCTACGTTGGTTGTCACTCTTGGGAACTGCTTTGATTGTTT
AAGCTTTCTGTTCTACTGGTTTGACTGAAGAATTTTGCTCTACCTTGCGAAAGGCTCACCA
ATTCTTGAAGAACGCTCAAGTCATCCAAAACATTCCAAACTACAAGTCTTACTACAGAGAA
CGTACCAAGGGTTCTTGGACTTTGTCCAACGGTGAAAACCTCTGGCCAATCGCTGACACT
AACGCTGAAGCTTTGAAAGCCATCTTGTTGTTGTCTAACATCCCATCCGAATTGGTCGGTA
ATCCAATTCGTCAAGAACGCCTATACGATGCCGTTGACTGTTTGTTATCATTCGTTAAACA
GGATGGTACTCTGTCTTCGCCGAATGTAAGAGAACCACCCCATGGGTTGAATTTATCAA
CCCTTCGAATCCTTCAGAAACATCATTGTCGACTACCCATACCCTGAATGTACTAGCTCCT
TGATTCAAGCCTTAATCTTATTACAGAGACAAGTGTCCAATGTACAGACACGAAGAAATTG

ACAAGATTATCAAGTCCGGTGCCTTATTTATTGAAAAGGTTCAAAGAAAGGACGGTTCCT
GGTACGGCTCTTGGGCTGTCTGTTTCACCTACGCTACTTTCTTGCTATTAAGGGTTTGGCT
GCCGCCGGTAGAACTTACCAAAACAGCCTGGCTATTCGTAAGGCTTGTGATTTCTTGCTGT
CTAAGCAATTGGGTACCGGTGGTTGGGGTGAAAATTATTTGGGTTGCCAAATCGAAGAA
TACGTCGATTCTGGTACCCACATGCTGTTAACACTGCTCAAGCTATGTTGGCTTTGCTTTA
TGCTGGTCAAGTTGGTCAAGATCCAGCTCCATTATACCGGGCAGCTAAGGAATTGATCAA
CATGCAAATGGAAAACGGTGAATTCACACAACAAGAAATCGTCGGTAACTTTAACTCTTC
TTTGTTCTTCAACTACCCAACTATAGAAATTTATCCCAATCTGGGCTTTGGGTGAATTCA
GAAGAAGATTGTTGGCGAAGAAGGGTTGA

mvaE ATGAAAACAGTTGTAATAATTGATGCTAAGAACTCCAATCGGCAAATATAAAGGAAGT
TTGTCCAAGTATCAGCTGTGGATTGGGCACCCATGTCACTACTCAGCTTCTAAAAAGAC
ATTCAACTATCTCAGAAGAAATCGATCAGGTTATTTTCGGTAATGTCCTTCAGGCCGGTAA
TGGACAAAACCCAGCTAGACAAATAGCTATTAATTCCGGCTTATCTCATGAAATACCAGC
AATGACTGTTAATGAAGTGTGTGGTTCAGGTATGAAAGCTGTTATACTGGCCAAGCAATT
GATTCAGCTTGGGGAAGCGGAAGTTTTGATAGCAGGTGGCATTGAAAACATGTCACAAG
CCCCCAAATTACAAAGGTTTAATTATGAAACTGAGTCATACGATGCACCTTTTTCTTCTATG
ATGTATGACGGTTTAACTGACGCATTTTCTGGACAAGCTATGGGTTTAACTGCTGAAAAC
GTAGCCGAAAAGTACCACGTGACAAGAGAAGAACAAGATCAATTTTCAGTGCATAGCCA
GTAAAAGCAGCTCAAGCACAAGCCGAAGGTATTTTGGCGATGAAATAGCCCCATTAGA
AGTATCTGGAACCTTAGTTGAGAAAGATGAGGGAATTAGACCAAATTCAGTGTTGAGA
AATTGGGGACGTAAAAACAGTTTTCAAAGAAGATGGTACTGTGACGGCTGGTAACGCTT
CAACTATTAATGATGGGGCTAGTGCTTTAATCATTGCCAGTCAAGAATATGCTGAAGCGC
ACGGCTTACCATATTTAGCGATAATACGTGACAGCGTCGAGGTTGGTATTGACCCTGCCT
ACATGGGCATAAGCCCGATCAAAGCGATCCAAAAATTATTGGCACGTAATCAACTAACTA
CAGAAGAAATAGACCTTTATGAGATCAACGAAGCTTTCGCAGCAACATCCATAGTTGTCC
AGAGAGAGTTAGCACTACCAGAAGAGAAAGTAAATATTTATGGTGGTGGTATAAGTCTT
GGACACGCAATAGGAGCTACTGGCGCAAGACTATTGACCTCATTATCTTACCAGTTGAAC
CAAAAGGAAAAAAAATACGGTGTTGCCTCACTATGCATAGGTGGTGGGTTAGGTTTAGC
CATGCTATTAGAAAGGCCTCAACAGAAAAAAAACCTCTAGGTTCTATCAAATGTCACCAGA

AGAGCGTCTTGCGAGCCTGTAAATGAAGGCCAAATTTAGCCGATACAAAGAAAGAATT
TGAAAATACCGCTCTAAGCAGTCAAATTGCGAACCACATGATTGAAAATCAAATTAGTGA
AACTGAAGTCCCTATGGGAGTCGGTTTGCACCTAACTGTGGATGAAACAGATTATCTAGT
TCCAATGGCTACTGAAGAACCAAGTGTTATTGCTGCATTGTCTAACGGTGCAAAAATAGC
CCAAGGTTTTAAGACTGTTAATCAACAAAGATTAATGAGAGGACAGATAGTCTTCTACGA
TGTGGCAGACGCTGAGTCCTTGATTGATGAGCTACAAGTTAGAGAGACTGAGATTTTTCA
GCAGGCTGAATTAAGCTATCCAAGTATTGTAAAGAGAGGGGGTGGCCTTAGAGACTTGC
AATACAGAGCTTTTGATGAAAGTTTCGTATCTGTTGATTTTTTGGTTGACGTCAAAGATGC
TATGGGCGCGAATATTGTTAATGCAATGTTAGAGGGCGTCGCAGAATTATTCGTGAATG
GTTCCGCCGAACAAAAGATCCTATTTAGTATATTGTCCAATATGCAACTGAATCAGTAGTA
ACAATGAAGACTGCAATCCCTGTGTCCCGTTTAAGTAAGGGTCTAATGGTAGAGAAATC
GCAGAGAAGATCGTTCCTTGCATCAAGATACGCTTCACTGGACCCGTATAGGGCAGTAACA
CATAACAAAGGTATTATGAATGGAATCGAAGCTGTTGTATTGGCTACAGGGAATGATACA
AGAGCTGTAAGCGCTTCCTGTCATGCATTTGCAGTCAAGGAGGGAAGGTATCAGGGGTT
AACCTCTTGGACTCTGGATGGTGAGCAATTAATTGGGGAAATTTCTGTACCTTTAGCTTTA
GCCACCGTTGGGGGGGCAACGAAAGTTTTGCCAAAGTCCCAGGCAGCAGCGGATCTATT
AGCAGTAACCGACGCCAAGGAATTGTCTAGAGTTGTTGCCGCAGTCGGTCTAGCACAAA
ATTTAGCTGCTCTGAGGGCACTAGTAAGTGAGGGTATTCAAAAAGGCCACATGGCCCTTC
AAGCTAGATCTTTAGCCATGACTGTAGGTGCTACTGGGAAGGAAGTCGAAGCTGTCGCA
CAACAGTTGAAAAGGCAAAAGACTATGAATCAAGATAGAGCGTTGGCAATCTTAAATGA
CTTGAGAAAGCAAtaa

mvaS

ATGACAATTGGCATCGATAAAATAAGTTTCTTTGTTCCACCATATTATATTGATATGACAG
CTTTAGCAGAAGCTCGTAACGTAGATCCAGGTAAGTTCCATATAGGGATAGGCCAAGACC
AAATGGCCGTAAATCCGATTTCCCAAGATATTGTTACATTCGCGGCAACGCGGCTGAAG
CCATCCTGACAAAGGAGGACAAAGAAGCCATTGATATGGTGATCGTCGGCACTGAATCA
AGTATTGATGAATCAAAAGCCGCCGAGTAGTCTTACACAGATTAATGGGAATACAACCT
TTCGCGAGATCATTCGAGATCAAGGAAGCATGTTATGGAGCTACCGCTGGTTTACAATTG
GCTAAGAATCACGTGGCATTGCACCCAGACAAAAAGTTTTGGTAGTCGCCGAGATATT
GCAAAAATATGGTTTGAATCCGGAGGAGAACCTACTCAAGGAGCTGGAGCTGTGGCTAT

GTTAGTCGCTTCTGAACCTAGGATACTTGCTTTAAAAGAGGACAATGTCATGTTGACTCAA
GACATATACGACTTTTGGAGACCAACCGGACATCCTTATCCTATGGTGGACGGTCCATTAT
CTAATGAAACGTACATTCAATCTTTGCACAAGTATGGGATGAACATAAGAAAAGAACTG
GTTTGGACTTCGCCGATTATGATGCATTGGCATTTACATACCATATACCAAGATGGGAAA
AAAAGCTCTGCTTGCCAAAATATCTGATCAAACAGAAGCAGAACAAGAAAGAATATTAGC
AAGATATGAAGAATCAATTATCTATTCAAGAAGGGTTGGCAATTTATACACCGGTTCACT
GTATTTAGGATTAATAAGCCTTCTTGAAAATGCAACTACACTAACTGCAGGTAATCAAATA
GGACTATTTTCATATGGATCTGGTGCTGTAGCAGAATTTTTCACCGGTGAATTAGTTGCCG
GTTATCAAAACCACTTACAAAAGGAGACTCATCTAGCTTTGTTGGATAACAGGACAGAAC
TAAGTATAGCTGAATACGAAGCTATGTTTGCAGAAACATTGGATACAGATATTGACCAGA
CTTTAGAGGATGAACTGAAATACAGCATTTCTGCTATCAACAATACTGTTTCGTAGCTACAG
GAACTaa

AnACLa ATGTCTGCCAAATCCATCTTCGAAGCCGATGGTAAGGCTATTTTGAATTACATTGACCA
GAGCCCCAGTCATTAAGCCTACTCCATTGCCTCCATCCAACACTCACAACCCACCTCCAAA
GCTGGCCTCTTTGACTTCCCAGACGATTTGTCCGTCAAGGACGTCTTGGATCAAGCTGAA
GTCATTACCCATGGTTGTTGACCCCAGGTTCTAAGTTTGTGCTAAGCCTGACCAATTGA
TAAAGAGAAGAGGTAAGTCTGGTCTATTAGCATTGAACAAGACCTGGGCAGAAGCTAGA
GAGTGGATTGAAGCTAGAGCTACCAAGGAACAACAAGTTGAAACCGTTGTGGGAGTTTT
GAGACACTTCTTGGTCGAACCATTGTTCCACACCCACAAGAACTGAATACTACATTAAT
ATCACTCTGTCAGAGAAGGTGACTGGATCTTATTTACCCACGAAGGTGGTGTGACGTT
GGTGATGTCGATGCCAAAGCTGAAAAGTTGTTGATTCCAGTTAACTGAAGAACTACCCA
TCTAACGAAGAAATTGCTTCTGCTTTATTGTCTAAAGTTCCAAAGGGTATCCATAACGTCT
TAGTTGATTTCTCTCTAGATTGTACGCTGTCTACGTTGACTGTCAATTCATTATTTGGAA
ATCAACCCATTGGTTGTCATCCCAAACGCTGATGCTACTTCTGCTGATGTTCACTTCTTGGA
CTTAGCTGCTAAGTTGGACCAAACCGCCGAATTCGAATGTGGTACCAAGTGGGCTGTTGC
TAGATCTCCAGCTAACTTGGGTTTAGCTGCTTTGCCAACTAGTGACAAGGTCAACATTGAC
GCCGGTCCACCAATGGAATTCCCAGCACCATTGGTAGAGAATTATCGAAGGAAGAAAA
GTTTCATCTCCGACATGGATGCTAAGACTGGTGCTAGCTTGAAATTGACTGTCTTAAACCCA
AACGGTCGTGTCTGGACTTTGGTGGCCGGTGGTGGTGCTTCCGTTGTTTACGCTGACGCT

ATCGCTTCTGCTGGTTTCGTTTCTGAATTGGCTAACTATGGTGAATACTCTGGTGCCCCAA
CCGAAACTCAAACATTCAACTACGCCAGAACCATTCTGGACTTGATGTTGCGTTCACCAAT
TCATCCAGACGGTAAAGTTTTGTTTATTGGTGGCGGTATCGCCAACTTCACCAACGTTGCC
TCCACCTTCAAGGGTGTTATCAGAGCTTTCGCTGAAGTTGCTCCGGTTCTTAATGAACACA
AGGTTCAAATCTGGGTTAGAAGAGCTGGTCCAAACTACCAAGAAGGTTTGAAGAACATC
AAGGCTGTTGGTGAAGAATTGGGTTTGAACATGCACGTTTACGGTCCAGAAATGCACGTC
TCCGGTATTGTCCATTGGCTTTGCAAGGTAAGCAAACCTGACATCAAGGAATTCGGTACT
GCTTGA

AnACLb ATGCCAGCCGCTCCATTGGTCAGCACGGCCAATGGTCCAAACGCTAACGACAACATCACA
AGATTTGAACCTCCATCTAGGGTCAGATCTCCATTTGCTGACGCTTGTTCACAACAAGA
CCAGATGTTTCGTCTACGGTATGCAACCAAGAGCCGTTCAAGGTATGTTGGATTTCGATTT
CATTTGTAAGAGATCCACTCCATCAGTTGCCGGTATTATTTACACCTTCGGTGGTCAATTT
GTCTCCAAGATGTACTGGGGTACCTCTGAACTTTACTACCTGTTTACCAAGATACAGCCA
AGGCTATGGCTAAGCACCCAGATGTCGACACTGTAGTTAACTTCGCTTCTTCTCGTTCGGT
TTACTCTTCTACCATGGAATTGATGCAATACCCACAAATCAAGTGCATTGCTATTATCGCT
GAAGGTGTCCCAGAGAGAAGAGCTAGAGAAATCTTGGTTACTGCTAAGGAAAAGGGTAT
TACCATTATCGGTCCAGCTACTGTTGGTGGCATCAAACCAGGTGCTTTCAAGATTGGTAAC
ACTGGTGGCATGATGGATAATATCGTTGCTTCCAAGTTGTACAGAAAGGGTTCAGTCGGT
TACGTTTCCAAGTCTGGTGGTATGTCTAACGAATTGAACAACATTATCTCCCAAACCTACTG
ATGGTGTTTATGAAGGTGTTGCTATCGGTGGTGACCGTTACCCAGGTACTACTTTTATTGA
CCACTTGTTGAGATACCAAGCTGAACCAGAATGTAAGATTTTGGTTTTGTTGGGTGAAGT
TGGTGGTGTGAAGAATACAGAGTCATTGAAGCCGTCAAGAACGGAGTTATCACCAAGC
CTATTGTCGCTTGGGCCATTGGTACCTGTGCTTCCATGTTCAAGACTGAAGTCCAATTCGG
TCACGCTGGTGCTTCTGCCAACTCTGACTTGGAACCGCAGTTGCCAAGAACAAGCTAT
GAGAGAAGCTGGTATCTACGTTCCAGACACTTTCGAAGACATGCCTGCTGTCTTGAAGAA
GGTTTACGAAGAACAAGTCCAAAACGGTGTTATCAAGCCACAACCAGAACCAGTCCCACC
AAAGATCCCAATTGATTATTCTTGGGCTCAAGAATTAGGTTTAATCAGAAAGCCAGCTGCT
TTCATCTCTACTATCTCTGACGACAGAGGTCAAGAATTGCTCTACGCTGGTATGCCAATCT
CTGATGTTTTCAAAGAAGACATCGGTATCGGTGGTGTATGTCCTTATTGTGGTTCAGAA

GGCGTTTGCCATCTTACGCTACCAAGTTCTTGAAATGGTCTTGATGTTGACCGCTGATCA
CGGTCCAGCTGTTTCTGGTGCTATGAACACTATCATCACCACCAGAGCTGGTAAGGATTT
GATTTCCGCTCTAGTGTGGGTTTGTGACTATTGGTTCCAGATTCGGTGGTGCCTTGGAT
GGTGCTGCTGAAGAATTCATAAAGCTTTCGACAAGGGTATGTCTCCAAGAGACTTCGTT
GACACCATGAGAAAGGAAAACAAGTTAATTCCAGGTATCGGGCACAGAATCAAATCCAG
AAACAACCCAGACTTGCGTGTTGAAGTAGTTAAGGAATATGTTAAGAAGCATTCCCAAG
TACCAAATTGCTGGACTACGCCATTGCTGTCGAAACCGTCACCACTTCTAAGAAGGACAA
CTTGATCTTGAACGTTGACGGTTGTATCGCTGTCTGTTTCGTTGATTTGATGAGAACTGT
GGTGCCTTCTCCGCTGAAGAATCCGAAGACTACATGAAGATGGGTGTGTTGAATGGTTTA
TTCGTATTGGGTAGATCCATAGGTTTGATTGCTCACTACTTAGACCAAAGCGTTTGAGAA
CCGTTTGTACAGACATCCATGGGATGACATTACTTACTTCTCCGCCTTGCAAAAGGG
TGGTCTGAAGGTAGAGTTGAAGTCAACGTCTGA

ThCYP71 ATGGCCACCATCACTGACATCCAATACTACGTCGTTTTGTTCTTCTTGTTGGTTGTTCTCTAC
2K1 TTTGTTGTTGCAATACGCTTTCAGAAGATCCACTAAGGGTGCTACCCACTTGAGATTGCCA
CCATCTCCACCATCTTGCCAGTCGTTGGTCACATCCACTTGTTGTCCAAGGACATCCACAT
CTGTTTCCAAACTTGGCCAGAAAGTACGGTCCATTGTTGTACTTGAGATTCGGTTCCTTC
AAGTGTTTGTTGATCTCCTCTGCCTCCGTCGCCACCGAAGTCTTCAAGTCCAATGACGTTG
CGTCTCTTCTAAGCCATTCTCTGTTTGGGTGATAGATTGATCTTCGGCAATACTGGCTTC
ATTATTTCTCAGTACGGTGATTACTGGAGATACATGAAGAAGTTGACTGTTACTGAATTGT
TGGGTGCTAGACAATTGGAAAGATCTAGAAAGGTTAGACAAGAAGAATTGCACAGATAC
TTGCAAAGAGTTTTCGAAAAGGCTGGTGCTGGTGAAGTTTTCGATGTTGGTTCTGAATTG
ATGAAGTTGACTAATAACACTATCTGCAGAATGGTTTTGTCTACCAGATGTTCTGAAGATG
ATAACGAAGCTGAAGAAGTTAAGAAGTTGGTTGAGGGCTCCTTCGATCTCGCTATGAAG
ATGACCATGGTTGCCATGTCTGGTCCATTGAAGAAGGTTGTTGGTAAGTACTACGAAAAG
GAAGATAAGGATATCAACAGAAGATGTGATGAATTGTTGGAAAGAATGTGGAAGGAAC
ACGAAGAAAGAGCTAAGAGAGAAGGTGTTGATAGAGAAGACAAGGATTTTCATGGATAT
TTTGTTGGAAGCTTACTACAACGACAAGGCTGAATTCAAGATTACCAGAAACCAAGTCAA
GGCCTTCATTTTGGACATTTTCATTGCTGGTACTTCTACTTCTGCTGATACTATGCAATGGG
TTATGGCTAACTTGATCAACCACCCAGATGTTTTGAAGAAGGTTAGACAAGAAATTGAAT

CTGTTGTTGGTAACTCTAAGAGAGTTGTTGAAGAATCTGATTGCCAACTTCCCATACTT
GCAAGCTGTTGTTAAGGAACTTTGAGATTGTACCCACCAGGTCCAGTTTTGCCAAGAAG
AACCAACGAAGACCGTAAGGTTTCTGGCTTCGATATCCCAAAGGATTGATTGTAGCGTT
CAACGTGTACGCCATCATGAGAGACCCAGAAGCTTGGGATAAGCCAAACGAATTCATCCC
AGAAAGATTCTTGTCTTCCACCAAGGAACAAAACACCCAATTGGTCAACTTCATGGCTTTC
GGTGCCGGTAGAAGATTGTGTCCAGGTTCTACTCTCGCTCTCACCTTGATGAATACTGCTA
TCGCTAACATGGTTCAATGTTTCGACTGGAAGGTCGGTATCGACGGTGATATCGTTAACA
TGAAGGCTGGTACTGGTTTCACCTTGGCTTTGGCCGAACCATTGATGTGTAGACCAGTCG
TTAGATTCAACCCATTCGCCGGTTAA

TrCYP71 ATGGCTTTCTACCAAACCTTGTCCGTTGATTCTATCTTGGAGCCAAAGGTGCTATTGTTGG
6C52 GTGCTATCTTCATCGTCTCCAACACTTCTTGTTCGAAGAGAATCTTCGGTACCAAGAAGTT
GCCACCAGGTTCTTTGGGTTGGCCAGTTGTTGGTGAAACCGTTCAATTCTTGCACAACCCA
GCTAAGTTCGTTGTTTCTAGAATGGGTAAGTACTCCAACGAAGTTTTCAAGACTAACATCT
TGGGTGAAAACACTGCTGTTTTGTGTGGTCCAGCTGGTAACAAGTTCATGTTCACTAACG
AACAAAAGTTGGTCAAGGTCTGGCACCCACACTCTACTCAAAGATTTTCTACTCCAACAA
GCCACAACAAGCTCCACCACCACCACCAACCCCAACCCCAAGAAGCTGGTAAGCAAGG
TGCTTTGAGATCTATGGGTTTCTTGAGACCAGAAGCTGTTATGAGATACTTGGACAAGAT
TGATTTGATCACCAAGCACCCTTGAACACTCACTGGGAAGGTAAGCAAGAAGTTAAGGT
TTACTCTTTGGTCAAGACCTACACTTTGGCCTTGGCTTGTCAATTGTTGTTGTTGTTAAC
GATCCAGACTTGGTTGAAAGATTGGTTGGTAAGTTCGATGAATTGTCTCACGGTATTAC
ACTATTGCTTGGAAACATTCCAGGTATGGCTTTCTACTCTGCTAAGAAGGCTGCCGTTGCTA
TTTCTGGTGAAATTCAACAAGTTATCAAGCAAAAGACTGATGCTATGGCTTCTGGTGTTCA
AATGCAAGACATCTTGTCTTACATGATCAAGGCTGGTCAAGAAGGTAAGTTCAAGGCCCC
AGCTGATATTGCTGAAGCTATCATGGGTTTGATCTCTGCTGGTTACGGTACTACTGCCATC
GCCTTGACCTTCGCCATGAAGTACATTGGTGAAAGACCAGATGTCTACCAAAAGATCTTG
GCTGAACAAATGGAAATCAAGTCCTCTGAAGGTTTCAACGGTGTTTTGGATTGGGATCAC
ATGCAAAAGATGAAGTACACTTGGAAACGTTGTTTCTGAAACCATGAGATTGAGAGCTCCA
TTGCAAGGTATGTTGAGAGAAGCCATGACTGACTTGACCTACGCTGGTTACACCATCCCA
AAGGGTTGGAAGTTGTACTGGACTGTTTCTTCTACTAACAAGAACCCAGACTACTTCCCAA

ACCCAGAAAAGTTCGATCCATCTAGATTCTGAAGATGGTTCTACTCCATACACTTTCAACCC
ATTCGGTGGTGGTCCAAGAATGTGTCCAGGTAAGGAATACGCTAGAGTTGTTGTTTTGAC
TTTCTTGACAACGTTGTTACTAAGTTCAAGTGGGATTTGTTGTGTCCAGATGAAAAGGTT
ATTGGTGACATGGCTCCAGCTCCAGAATTGGGTGTCCAATCAGATTGCACTCTGCCAACT
AA

ATR1 ATGACTTCTGCCTTGTATGCCTCTGATTTGTTCAAGCAATTGAAGTCCATTATGGGCACCG
ATTCTTTGTCTGATGATGTTGTTTTGGTTATCGCTACTACCTCTTTGGCTTTGGTTGCTGGT
TTTGTTGTTCTGTTGTGGAAAAAGACTACCGCTGATAGATCTGGTGAATTGAAACCATTG
ATGATCCCCAAATCTTTGATGGCCAAAGATGAAGATGATGACTTGGACTTAGGTTCTGGT
AAGACTAGAGTTTCATTTTCTCGGTACTCAAAGTGGTACTGCTGAAGGTTTTGCTAAAG
CTTTGTCCGAAGAAATCAAGGCCAGATACGAAAAAGCTGCCGTTAAGGTTATTGATTGG
ATGATTATGCTGCCGATGACGACCAATACGAAGAAAAGTTGAAGAAAGAAACCTTGCC
TTCTTCTGTGTTGCTACTTATGGTGATGGTGAACCTACTGATAATGCTGCTAGATTTTACA
AGTGGTTCACCGAAGAGAACGAAAGAGATATCAAGTTGCAACAATTGGCCTACGGTGTT
TTTGCTTTGGGTAATAGACAATACGAGCACTTCAACAAGATCGGTATCGTTTTGGATGAA
GAGTTGTGTAAAAAGGGTGCCAAGAGATTGATTGAAGTTGGTTTGGGTGATGATGACCA
GTCTATCGAAGATGATTTTAACGCCTGGAAAGAATCCTTGTTGGTCTGAATTGGATAAGTT
GTTGAAGGACGAAGATGACAAATCTGTTGCTACACCATACACTGCTGTTATTCCAGAGTA
TAGAGTTGTTACTCACGATCCAAGATTCACGACTCAAAAGTCTATGGAATCTAACGTTGCT
AACGGTAACACCACCATCGATATTCATCATCCATGTAGAGTTGATGTCGCCGTCCAAAAA
GAATTGCATACTCATGAATCCGACAGATCCTGCATTCATTTGGAATTCGATATTTCCAGAA
CCGGTATTACTTACGAAACCGGTGATCATGTTGGTGTTTACGCTGAAAATCACGTTGAAAT
CGTTGAAGAAGCCGGTAAGTTGTTAGGTCATTCATTGGATTGGTGTTCTCCATTCATGCC
GACAAAGAAGATGGTTCTCCTTTGGAATCTGCTGTTCCACCACCATTTCCAGGTCCATGTA
CTTTAGGTACTGGTTTGGCTAGATATGCTGACTTGTTGAATCCACCAAGAAAGTCTGCTTT
AGTTGCTTTGGCTGCTTATGCTACTGAACCATCTGAAGCCGAAAAATTGAAACATTTGACT
TCCCAGATGGTAAGGACGAATATTCTCAATGGATAGTTGCCTCTCAGAGGTCTTTGTTG
GAAGTTATGGCTGCTTTTCCATCTGCTAAACCACCATTTGGGTGTTTTTTTGTGCTGCTATTGC
TCCAAGATTGCAACCTAGGTATTACTCCATTTCTTCATCACCAGATTGGCCCCATCTAGA

GTTTCATGTTACATCTGCTTTGGTTTATGGTCCAACTCCAACTGGTAGAATTCATAAGGGTG
TTTGTTCTACCTGGATGAAGAACGCTGTTCCAGCTGAAAAATCTCATGAATGTTCTGGTGC
CCCAATTTTCATTAGAGCTTCTAATTTCAAGCTGCCAAGCAATCCATCTACTCCAATAGTTA
TGGTTGGTCCAGGTACAGGTTTAGCTCCTTTTAGAGGTTTCCTACAAGAAAGGATGGCCT
TGAAAGAGGATGGCGAAGAATTGGGTTCTCCTTGTTGTTTTTGGTTGCAGAAACAGAC
AGATGGATTTTCATCTATGAGGACGAGTTGAACAACCTCGTTGATCAAGGTGTTATCTCCG
AATTGATTATGGCCTTTTCTAGAGAAGGTGCCAGAAAGAATATGTCCAACATAAGATGA
TGGAAAAAGCCGCTCAAGTTTGGGACCTAATCAAAGAAGAAGGATACTTGTACGTTTGC
GGTGATGCTAAAGGTATGGCTAGAGATGTTTCATAGAACATTGCATACCATCGTCCAAGAA
CAAGAAGGTGTTTCATCTTCTGAAGCTGAAGCTATCGTTAAGAAGTTGCAAACTGAAGGT
AGATACTTGAGAGATGTCTGGTGA

MTR1 ATGACTTCTTCCAACTCTGACCTTGTTAGAACCATGAATCCGTTCTAGGTGTTAGTTTGG
GAGACTCCGTTTCCGACTCTGTTGTTTTGATCGTTACCACCTCTGCTGCTGTCATCATAGGT
TTGTTGGTCTTCTTGTTGAAGAAGTCTTCTGACAGAAGTAAGGAGCTAAAGCCAGTTATC
GTCCCAAAGTCCTTGGTCAAGGAAGAAGATGATGATGCTGACATCGCTGATGGTAAGAC
CAAGGTTACTGTCTTTTTCGGTACCCAAACCGGTACTGCCGAAGGTTTCGCGAAGGCCTT
GGCCGAAGAAATCAAGGCAAGATATGAAAAGGCTTTCGTCAAGGTTGTTGACATGGACG
ACTACGCCGCTGATGACGATCAATACGAGGAAAAATTAAAGAAGGAAACCTTAGCTTTTT
TCATGTTGGCCACTTACGGTGACGGTGAACCAACCGATAACGCTGCTAGATTCTACAAGT
GGTTCATGAAGGTAAGGACGAAAGAGGTACCTGGTTGCAACAATTGACTTACGGTGTTT
TCGGTTTAGGTAACAGACAATATGAACACTTCAACAAGATTGGTAAGGTCGTGGACGAC
GATTTGTCTGAACAAGGTGCCAAGAGATTAGTTCCATTGGGCATGGGTGATGACGACCA
ATCTATCGAAGACGACTTCAATGCTTGGAAGGAATCTTATGGCCAGAATTGGATCAATT
ATTGCGTGATGAAGATGATGTTAACTGTCTCTACTCCATACACTGCTGCCATCTCTGAA
TACAGAGTTGTTTTCCACGACCCAACGGTCACTCCATCTTACGAAAACCATTTCAATGCTG
CTAACGGTGGTGCTGTTTTTGATATTCACCACCCATGCAGAGCCAACGTTGCTGTAGAA
GAGAATTGCATAAGCCACAAAGCGACCGTTCCTGTATCCATTTGGAATTCGATGTCTCCG
GTACAGGTGTTACTTACGAAACCGGTGACCACGTCGGTGTCTACGCTGACAACTGTGATG
AAACTGTCAAAGAAGCTGGTAAGTTGTTGGGTCAAGATTTGGACTTGTTGTTCTCCTTGC

ACACCGACAACGAAGACGGTACTTCATTAGGTGGTTCTTTGTTGCCTCCATTCCCAGGTCC
ATGTACTGTCCGTACCGCTTTGGCTAGATACGCTGACTTGTTGAACCCACCTAGAAAGGCT
GCTTTGATTGCCCTTGCCGCTCACGCTTCCGAACCATCCGAAGCTGAAAGATTGAAATTCT
TGTCTCCCCACAAGGTAAAGACGAATATTCTAAATGGGTTGTCGGTTCACAGAACTTT
GCTAGAAGTTATGGCTGATTTCCCATCCGCTAAGCCACCATTGGGTGTCTTCTTTGCCGCA
ATTGCCCCAAGATTGCAACCAAGATACTACTCCATTTCTCTCTCCAAGATTGCTCCGCA
AAGAGTTCATGTCACCTGTGCTCTGGTCGAAGGTCCAACCCCAACTGGTCGTATTCACAA
GGGTGTTTGTTCACCTGGATGAAGAACGCTATCCCATCTGAAGAATCTAGAGACTGTTC
TTGGGCTCCAATTTTCATCAGACCTTCTAACTTCAAGTTGCCAGCTGATCCATCTATTCCAA
TTATCATGGTTGGTCCAGGTACTGGTTTGGCCCCCTTCAGAGGTTTCTTGCAAGAAAGATT
CGCTTTGAAGGAAGACGGGGTACAATTGGGTCCAGCTTTATTGTTCTTCGGTTGTAGAAA
CCGTCAAATGGATTTTATCTACGAAGAAGAATTGAACAACTTCGTTGAACAGGGTTCTCT
GAGTGAATTGATCGTTGCCTTTTCCAGAGAAGGTCCAGAAAAGGAATACGTTCAACACAA
GATGATGGACAAGGCTTCTTACTTCTGGTCTTTGATTTCTCAAGGTGGTACTTGTACGTC
TGTGGTGACGCTAAGGGCATGGCTAGAGATGTTACAGAACCTTGACACTATCGTTCAA
CAACAAGAAAATGCTGATTCCTCTAAGGCTGAAGCTACCGTTAAGAACTCCAAATGGAC
GGTAGATACTTGCGGGACGTTTGGTGA

SmCPR1 ATGGAACCATCTTCTAAGAAATTGTCTCCATTGGATTTCAATTACCGCCATTTTGAAGGGTG
ACATTGAAGGAGTCGCTCCAAGAGGTGTTGCTGCTATGTTGATGGAAAACAGAGATTTA
GCTATGGTTTTGACTACTTCCGTTGCCGTATTGATTGGTTGTGTCGTTGTCTTGGCTTGA
GAAGGACCGCTGGTTCGCTGGTAAGAAGCAACTGCAACCTCCGAAGTTAGTTGTTCCAA
AACCAGCTGCTGAGCCAGAAGAAGCTGAAGACGAAAAGACCAAGGTTTCTGTCTTCTCG
GTA CTCAAACCGGTACTGCTGAAGGTTTCGCTAAAGCCTTGCTGAAGAAGCTAAGGCTC
GTTACCCACAAGCTAAGTTCAAGGTCATTGACTTGGATGACTACGCCGCGGACGATGACG
AATATGAAGAAAAGCTAAAAAAGGAATCCTTGGCATTCTTCTTTTGGCTTCTACGGTGA
TGGTGAACCAACTGACAACGCTGCCAGATTCTACAAGTGGTTCACCGAAGGTAAGGACA
GAGAAGACTGGTTGAAGAACTGCAATACGGTGTTTTCGGTTTGGGTAACAGACAATATG
AACACTTCAACAAGATCGCCATCGTTGTTGACGATTTGATTACTGAACAAGGTGGTAAGA
AGTTAGTCCCAGTTGGTTTGGGTGATGATGACCAATGTATCGAAGATGACTTCTCTGCTT

GGAGAGAATTGGTCTGGCCAGAATTGGACAAGTTGTTGAGAAACGAAGACGACGCTACT
GTTGCTACCCCATACACCGCTGTTGTTCTACAATACAGAGTCGTCTTACACGATCAAAC TG
ATGGTTTGATCACTGAAAATGGTTCTCCAAACGGTCACGCTAACGGTAACACCATCTACG
ATGCCCAACATCCATGTCGTGCAAACGTCGCTGTCAGAAGAGAATTGCACACTCCAGCTT
CTGACAGATCTTGTACTCACCTGGAATTCGACACATCAGGGACTGGTTTAGTCTACGAAA
CTGGTGACCACGTTGGTGTTTACTGTGAAAACCTTTTGAAAAACGTTGAAGAAGCCGAAA
AATTGTTGAACTTGTCTCCTCAAACCTTACTTCTCCGTTCACTGATAACGAAGACGGTAC
CCCATTGTCCGGCTCTAGTTTACCACCACCATCCACCATGCACTTTGAGAACAGCGTTG
ACCAAGTACGCTGACTTGATTTCATGCCAAAGAAGTCCGTTTTAGTTGCTTAGCTGAAT
ACGCCTCTAACCAATCTGAAGCTGACCGTTTGAGATATTTGGCTTCCCAGATGGTAAGG
AAGAATACGCTCAATACATTGTTGCGTCTCAAAGAAGCTTATTGGAAGTTATGGCTGAAT
TTCCTTCTGCCAAGCCACCACTAGGTGTGTTTTTGGCGCTATTGCTCCACGTTTGCAACCA
AGATTCTACTCTATCTTCTTCTCCAAAGATCGCTCCAACCAGAGTTCATGTTACTTGTGC
TTTGGTCTACGACAAGACTCCTACCGGTAGAATCCACAAGGGTATCTGTTCTACCTGGATC
AAGAATGCTGTCCCATTGGAAGAGTCCTCCGACTGTTCTTGGGCTCCAATTTTCATTCGGA
ATTCCAACCTCAAGTTGCCAGCAGACCCAAAGGTTCCAATCATCATGGTCGGTCCAGGTA
CGGGTTTGGCCCCATTAGAGGTTTCTTGCAAGAAAGACTAGCGTTGAAGGAATCTGGG
GCAGAATTGGGTCCAGCTATCTTATTCTCGGTTGCAGAAACAGAAAGATGGATTTTCATCT
ACGAAGATGAATTAAACTCTTTCGTCAAGGTCGGTGCTATCTCTGAATTGATCGTCGCTTT
CTCCAGAGAAGGTCCAGCCAAGGAATACGTTCAACACAAAATGTCTCAGAGAGCTTCTGA
TATCTGGAAGATGATTTCTGACGGTGGTTACATGTACGTCTGTGGTGATGCTAAGGGTAT
GGCCAGAGACGTCCACAGAACCTTGCATACCATGCTCAAGAACAAGGTTCCCTCTCCTC
ATCCGAAGCTGAAGGTATGGTTAAGAACTTGCAAACCACTGGTAGATACTTGAGAGATGT
TTGGTGA

TwCPR3 ATGTCTAGCAGCTCTGACCTAGTTCAGTCCGTGGAATCTTTGCTAGGCATTAGTTTTGGTG
GGTCAGTGTCTGATTCAATGCTGTTAGTGGCCACCACGTCACTTGCTGTAATAGTCGGACT
TCTTGTGTTTTCTGGAAGAAGAGTTCTGACCGTGGCCGTGAAGTGAAACCCGTATTCTG
GCCAAGCCATTGTCTGTCAAGGATGATGAAGCCGAAGCCGAAGCTACAAGTGGAAGA
CGAGAATTACAATCTTCTCGGGACGCAGACGGGTACGGCCGAGGGTTTCGCGAAGGCC

TTAGCGGAAGAAATTAAGCCAGATATGAAAAAGCCGCTGTTAAGGTAGTTGACATGGA
TGATTATGCTGGCGACGACGACGAGTATGAAGAAAAGCTGAAGAAAGAAACCTTGGCTT
TCTTTATGGTCGCAACATATGGGGACGGCGAGCCGACTGACAACGCTGCGCGTTTCTACA
AATGGTTTACTGAAGGAGGCGAGAGAGGGAAGTGGTTGCAACAATTGACATACGGCGTC
TTTGGTCTTGGGAATAGACAGTATGAACACTTTAATAAGATAGCTAAAGTTGTGGATGAA
CAGCTAAGTGAGCAGGGTGCGAAGAGACTTATCCAGGTGGGCTTAGGGGACGATGACC
AGTGATCGAGGATGATTTAGCGCGTGGAGAGACAGTCTTTGGCCGAGTTAGACGCG
TTGCTATTTGACGAGGATGAGGCATCAGCGTCTACACCGTACACAGCGCGATCCCTGAG
TATAGAGTGGTCATCCACGATCCATCAGTGACCAGTTGTGGTGACAAAGATCTGAATATG
GTGAATGGCAACGCTTCTTTGACATACACCACCCGTGTAGAGTGAATGTAGCGGTGCAG
CGTGAGTTGCATAAACCCGTTAGTGATCGTAGTTGCATCCACCTTGAATTTGATATTCCGG
GGACAGGCATAACTTACGAAACCGGAGACCACGTTGGTGTCTTTGCAGATAATTGTGACG
AGACTGTAGAGGAAGCCGGGAGACTACTTGGCCAACCCCTGGACATGCTGTTCTCTGTGC
ATACAGATAATGAGGACGGGACCTCATTAGGGTCTAGTCTACCACCGCCCTTCCCAGGCC
CCTGTACTCTTAGAACCGCATTGGCACGTTATGCAGATCTTCTTAATGCGCCTAAGAAGGC
GGCACTGATTGCGTTAGCCGCTCACGCGACTGAGCCTAGTGAAGCGGAAAGGTTGAAAC
ATCTGTCATCCCCGCAAGGTAAAGAGGAGTATGCTCAATGGGTAGTCGGAAGCAAGCGT
TCTTTATTAGAGGTAATGCAGGAATCCCGAGCGCAAGGCCTCCTCTGGGTGTCTTTTTTG
CTGCAGTTGCTAGCCGTCTTCAGCCGAGGTACTATAGTATTTCAAGCTCCCCCGTTATGT
CCCGAACAGAGTGACGTAACCTGTGCCCTAGTGACGGACCCACCCCACTGGCCGTAT
CCATAGGGGTGTTTGCTCCACTTGGATGAAAAATGCAATCCCACTAGAGAAGTCACATGA
TTGTAGTTGGGCTCCCATCTTTATACGTCAGTCCAACCTTAAGTTACCTGCCGATCCATCAA
TACCTATAATTATGGTAGGGCCTGGTACCGGTCTTGCCCTTTAGGGGATTCTACAGG
AAAGGATGGCCCTGAAAGAAGATGGCGCTGAAGTGGGTCCGGCGCTATTGTTCTTTGGT
TGCAGATCCCGTCAGATGGATTTCAATTACGAAGACGAATTAATAACTTCGTGGAACAG
GGAGTGCTTAGCGAATTGGTAGTGGCCTTGTCAGGGAGGGACCAGAGAAAGTATATGT
CCAACACAAAATGATGAACAAAGCGGCCGAGATATGGTCCGTAATCTCTCAAGGCGGCT
ACTTTTATGTTTGCGGCGACGCGAAAGGTATGGCCAGGGACGTGCATCGTACGTTGCACA
CCATAGTACAAGAGCAGGGAAACATGGACTCTTCCAAGACGGAATCAATGGTCAAAAAA

CTACAGATGGAGGGAAGATATTTGAGGGATGTTTGtaa

AtCYB5 ATGGCTTCTGATCCAAAGAAGCTTCGTTTTGACGATATCTCTAAACACAACAAGACTAAGG
ATTGCTGGTTGATCATTGATGGTAAGGTTTACGATGTTACGCCCTTCATGGAAGATCATCC
AGGTGGTGATGAAGTTTTGTTGGCTGCTACTGGTAAAGATGCTACTGATGATTTTGAAGA
TGTTGGTCATTCCGATGATGCCAGAGAAATGATGCATAAGTACTACATTGGTGAAGTTGA
TAAGGCTACGGTTCCAAAAAAGAGCTTACGTTCCACCAGCTGATAGACATTATAATCC
AGATAAGACCCAAGACTTCATCGTCAAGATCCTGCAATTTTGGTTCCATTGGTCATTTTG
GGTTAGCTTTCGCTGTTAGGTCTTACACCAAAGAAAGATCTGCTtaa

SE ATGTGGACATTTTTAGGAATAGCAACATTTACATACTTTTACAAGAAATGTGGTGATGTAA
CATTAGCCAATAAGGAGCTTCTTTTGTGTGTGTTGGTTTTCTTGTCCTTGGCTTAGTGTTA
AGCTACAGATGTAGACATAGAAACGGTGGATTGCTTGGCAGGCATCAGAGTGGTAGCCA
GTTGCTGCATTCTCTGATATTTTATCCGCATTACCGCTTATTGGTTTCTTTGGGCCAAAT
CTCCTCCAGAAAGCGAAAAGAAAGAACTGGAGAGTAAAAGAAGACGTAAAGAAGT
TAACCTAAGCGAAACGACGCTTACAGGTGCTGCTACAAGTGTGAGCACATCCTCAGTTAC
CGATCCTGAAGTAATCATCATTGGTAGCGGGGTTCTTGGTTCAGCATTAGCCACTGTTTTG
TCAAGAGACGGGAGAACTGTTACAGTCATAGAACGTGACTTGAAAGAGCCGGATAGGAT
CTTAGGTGAATGTCTTCAACCAGGGGGATACAGAGTATTGAGGGAATTGGGTCTAGGCG
ATACTGTTGAATCTCTAAATGCTCATCATATTCATGGCTACGTCATACATGACTGCGAGTC
TAGGTCTGAGGTCCAAATACCATATCCAGTCTCAGAAAACAATCAAGTTCAATCCGGCGT
CGCATTTTCATCATGGTAAATTTATAATGAGTTTAAGGAAAGCAGCTATGGCTGAACCGAA
TGTTAAGTTCATCGAGGGTGTCGTGTTGAGATTACTAGAAGAAGATGACGCAGTGATAG
GTGTACAATACAAGGATAAAGAACTGGAGACACTAAAGAACTTCATGCTCCATTGACCG
TCGTTGCTGACGGTCTGTTCTCAAATTCAGGAAGAATTTGATTTCTAACAAGGTCTCAGT
CTCTTCTCATTTTGTGGGTTTATAATGAAGGATGCCCCACAATTTAAGGCGAATTTTGCT
GAGCTGGTATTAGTTGACCCATCTCCAGTACTGATTTACCAGATATCACCGTCCGAGACTA
GAGTACTTGTCGACATTAGAGGAGAGCTGCCAAGGAATTTGAGAGAATATATGACGGAA
CAGATATACCCGCAAATTCGGACCATCTTAAGGAATCATTTTGGGAAGCATGTCAAAC
GCTAGACTTAGGACAATGCCCGCTTCATTTCTACCGCCATCATCTGTTAATAAGAGAGGCG
TCTTGCTTTTGGGTGATGCCTATAATTTAAGACACCCCTAACAGGAGGTGGTATGACTGT

GGCCCTGAAAGATATCAAAATTTGGCGTCAGTTGTTGAAGGACATACCCGATCTATATGA
TGACGCGGCTATTTTTCAAGCTAAGAAATCTTTTTCTGGAGCAGAAAAAGAAGTCATAG
CTTTGTTGTGAACGTTTTAGCTCAGGCACTTTACGAACTGTTCTCTGCCACAGATGACTCA
TTAAGACAATTGCGTAAGGCTTGTTTTTTGTATTTCAAACCTGGTGGTGAATGTTTGACAG
GTCCGGTAGGATTACTAAGCATCTTAAGCCCTGACCCACTATTATTAATAAGGCATTTTTT
TAGTGTCGCAGTATATGCTACATATTTTTGCTTCAAGAGCGAACCTGGGGCTACTAAACCA
AGGGCTCTTTTCAGTTCAGGCGCCATCTTATATAAGGCTTGCTCTATTATATTCCCGTTGAT
TTACTCTGAAATGAAGTATCTGGTACATtaa

rSE

ATGAAAGAAGTTAACCTAAGCGAAACGACGCTTACAGGTGCTGCTACAAGTGTGAGCAC
ATCCTCAGTTACCGATCCTGAAGTAATCATCATTGGTAGCGGGTTCTTGTTTCAGCATT
GCCACTGTTTTGTCAAGAGACGGGAGAACTGTTACAGTCATAGAACGTGACTTGAAAGA
GCCGGATAGGATCTTAGGTGAATGTCTTCAACCAGGGGGATACAGAGTATTGAGGGAAT
TGGGTCTAGGCGATACTGTTGAATCTCTAAATGCTCATCATATTCATGGCTACGTCATACA
TGACTGCGAGTCTAGGTCTGAGGTCAAATACCATATCCAGTCTCAGAAAACAATCAAGT
TCAATCCGGCGTCGCATTTTCATCATGGTAAATTTATAATGAGTTTAAGGAAAGCAGCTATG
GCTGAACCGAATGTTAAGTTCATCGAGGGTGTCGTGTTGAGATTACTAGAAGAAGATGA
CGCAGTGATAGGTGTACAATACAAGGATAAAGAACTGGAGACACTAAAGAACTTCATG
CTCCATTGACCGTCGTTGCTGACGGTCTGTTCTCAAATTCAGGAAGAATTTGATTTCTAA
CAAGGTCTCAGTCTCTTCTCATTTTGTGTTGGGTTTATAATGAAGGATGCCCCACAATTTAAG
GCGAATTTTGCTGAGCTGGTATTAGTTGACCCATCTCCAGTACTGATTTACCAGATATCAC
CGTCCGAGACTAGAGTACTTGTCGACATTAGAGGAGAGCTGCCAAGGAATTTGAGAGAA
TATATGACGGAACAGATATACCCGCAAATTCGGGACCATCTTAAGGAATCATTTTTGGAA
GCATGTCAAAACGCTAGACTTAGGACAATGCCCGCTTCATTCTACCGCCATCATCTGTTA
ATAAGAGAGGCGTCTTGCTTTTGGGTGATGCCTATAATTTAAGACACCCCTAACAGGAG
GTGGTATGACTGTGGCCCTGAAAGATATCAAAATTTGGCGTCAGTTGTTGAAGGACATAC
CCGATCTATATGATGACGCGGCTATTTTTCAAGCTAAGAAATCTTTTTCTGGAGCAGAAA
AAGAAGTCATAGCTTTGTTGTGAACGTTTTAGCTCAGGCACTTTACGAACTGTTCTCTGCC
ACAGATGACTCATTAAGACAATTGCGTAAGGCTTGTTTTTTGTATTTCAAACCTGGTGGTG
AATGTTTGACAGGTCCGGTAGGATTACTAAGCATCTTAAGCCCTGACCCACTATTATTAAT

AAGGCATTTTTTTAGTGTGCGCAGTATATGCTACATATTTTTGCTTCAAGAGCGAACCTTGG
GCTACTAAACCAAGGGCTCTTTTCAGTTCAGGCGCCATCTTATATAAGGCTTGCTCTATTA
TATCCCCGTTGATTACTCTGAAATGAAGTATCTGGTACATtaa

Table S3† Tunnel statistics of ThCYP712K1 and its variant I124V by Caver Analyst 2.0 BETA.

ID	Avg_BR (Å)	Avg_L (Å)	Avg_C (Å)	Priority
ThCYP712K1	1.63	17.65	1.15	0.717
ThCYP712K1 ^{I124V}	1.95	16.59	1.16	0.808
ThCYP712K1 ^{T488A}	3.43	29.00	1.16	0.806

Abbreviations: Avg_BR: average bottleneck radius; Avg_L: average tunnel length; Avg_C: average tunnel curvature; Priority: tunnel priority calculated by averaging tunnel throughputs over all snapshots.