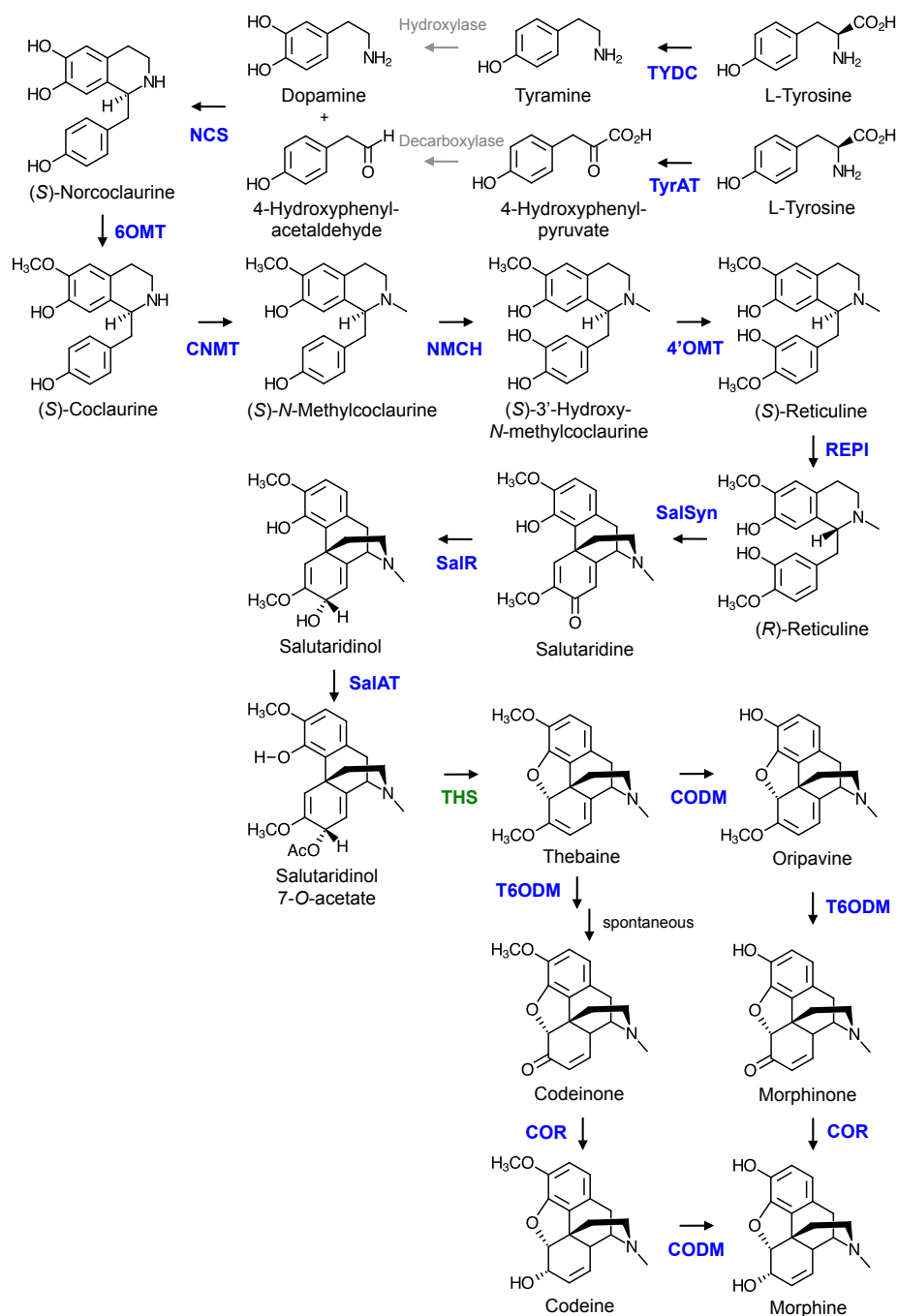


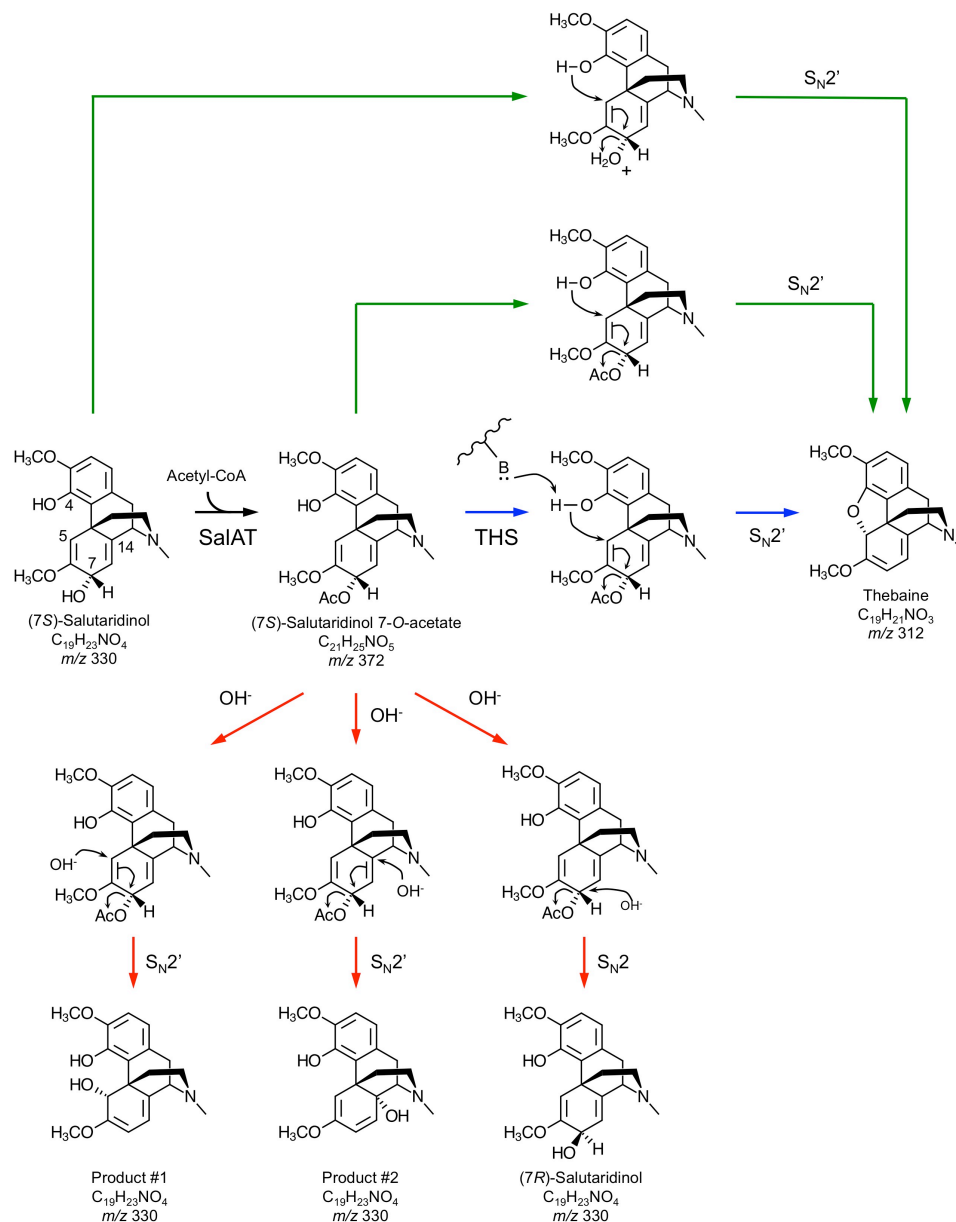
Supplementary Information

A pathogenesis-related 10 protein catalyzes the final step in thebaine biosynthesis

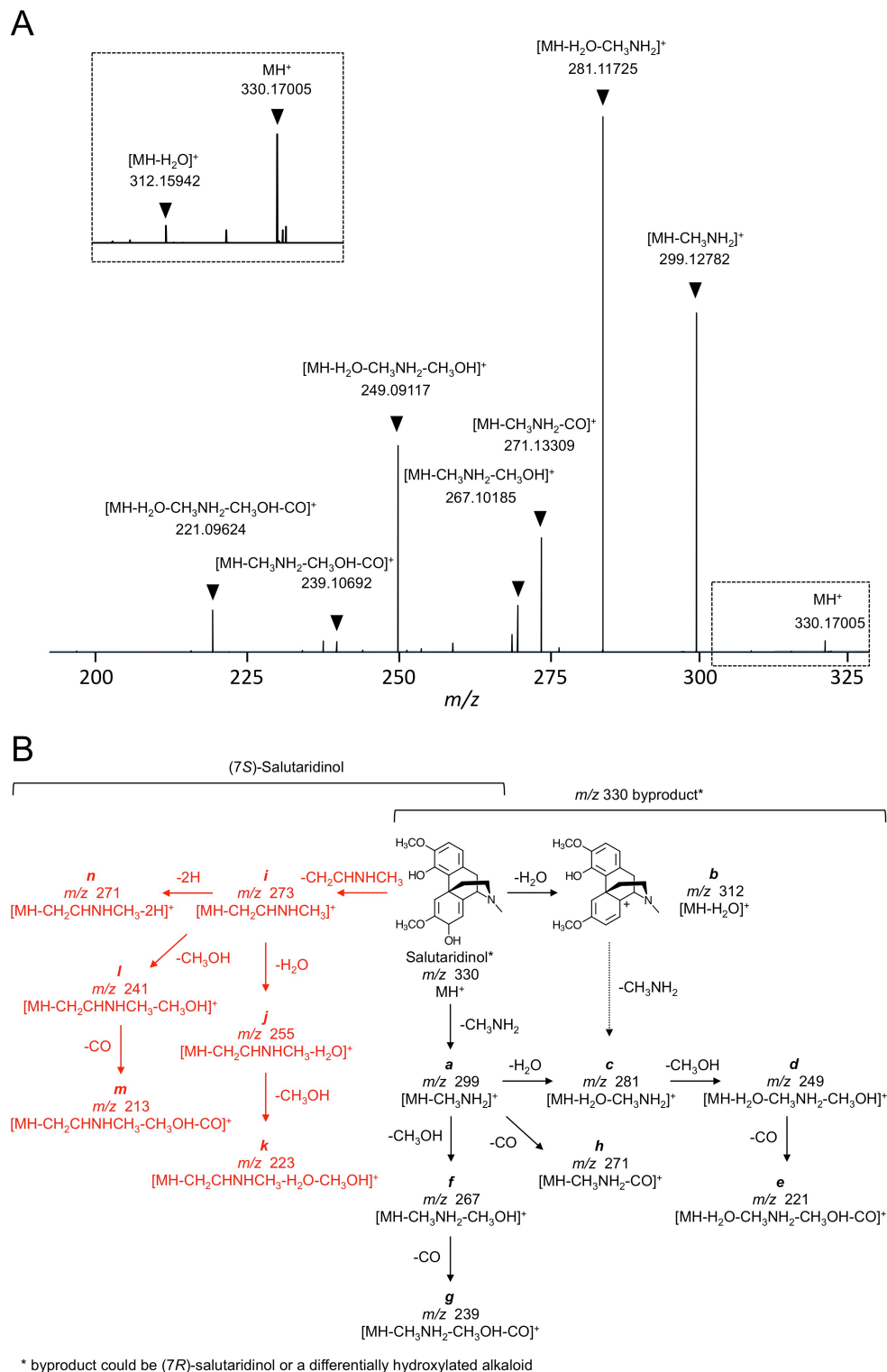
Xue Chen^{1,2}, Jillian M. Hagel^{1,2}, Limei Chang^{1,2}, Joseph E. Tucker^{2,3}, Stacey A. Shiigi⁴, Yuora Yelapaala⁴, Hsiang-Yun Chen⁴, Rodrigo Estrada⁴, Jeffrey Colbeck⁴, Maria Enquist-Newman⁴, Ana B. Ibáñez⁴, Guillaume Cottarel⁴, Genevieve M. Vidanes⁴ and Peter J. Facchini^{1,2,*}



Supplementary Figure 1. Schematic representation of benzylisoquinoline alkaloid (BIA) metabolism leading to opiate production in opium poppy. Enzyme names are shown in blue except for thebaine synthase (THS), which is shown in green. Other abbreviations: 4'OMT, 3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase; 6OMT, norcoclaurine 6-O-methyltransferase; CNMT, coclaurine N-methyltransferase; CODM, codeine O-demethylase; COR, codeinone reductase; NCS, norcoclaurine synthase; NMCH, N-methylcoclaurine 3'-hydroxylase; REPI, reticuline epimerase; SalAT, salutaridinol 7-O-acetyltransferase; SalR, salutaridine reductase; SalSyn, salutaridine synthase; T6ODM, thebaine 6-O-demethylase; TYDC, tyrosine decarboxylase; TyrAT, tyrosine aminotransferase.



Supplementary Figure 2. Proposed mechanisms underlying THS activity and the spontaneous formation of hydroxylated byproduct. At pH <5 *in vitro*, (7S)-salutaridinol undergoes spontaneous, intramolecular allylic rearrangement and dehydration yielding thebaine. However, this process is not spontaneous *in vivo*. First, SalAT must O-acetylate (7S)-salutaridinol to create a better leaving group. Second, THS is required to (i) promote thebaine formation beyond the small amount that forms spontaneously (e.g. <10% at pH 7.0) and (ii) outcompete, or shield against, spontaneous degradation of (7S)-salutaridinol 7-O-acetate yielding the m/z 330-byproduct. One possible mechanism for THS activity involves deprotonation of the C4 hydroxyl group of salutaridinol 7-O-acetate by a catalytic residue acting as a general base, yielding a phenolate anion nucleophile. This newly formed nucleophile would then attack at C5 to facilitate S_N2' *syn* displacement of the O-acetyl leaving group. Alternatively, THS may not require a catalytic base, and could simply ‘shield’ the labile acetate from intermolecular nucleophilic attack under aqueous conditions. High-resolution MSⁿ analysis supported the formation of a hydroxylated alkaloid with m/z 330 and structural similarity to (7S)-salutaridinol (Supplementary Table 1; Supplementary Figure 3). Three possible structures are proposed resulting from either S_N2' or S_N2 elimination of the acetate.

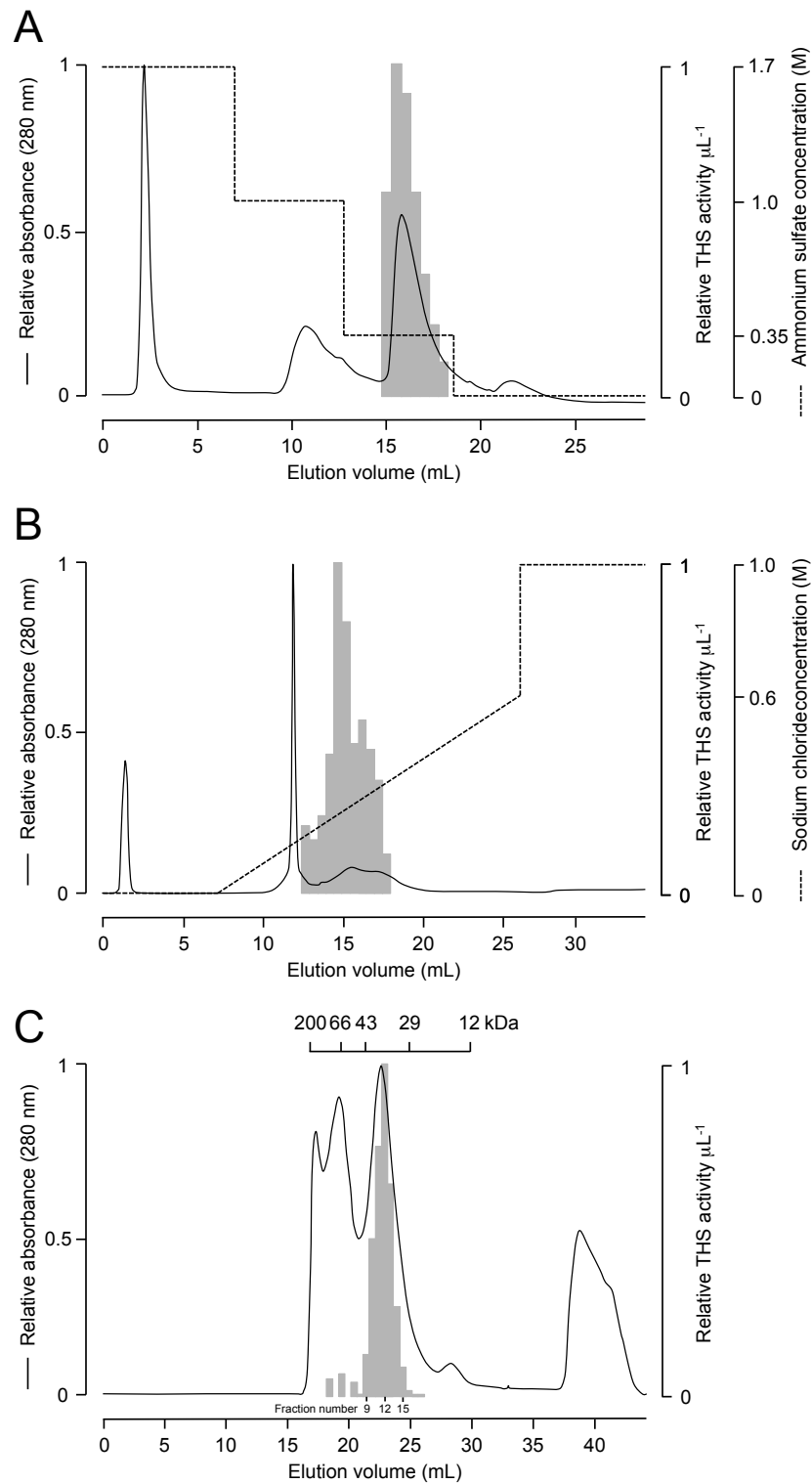


Supplementary Figure 3. High-resolution mass spectral fragmentation analysis of alkaloid byproduct (m/z 330). (A) High-resolution MS² collision-induced dissociation (CID) spectrum of the m/z 330-byproduct. Fragmentation at 35% normalized collision energy (NCE) was conducted in the linear ion trap portion of an LTQ-Orbitrap XL instrument followed by full-scan FTMS detection in the Orbitrap MS (m/z 90-340). The spectrum represents an average of 161 individual scans captured over 5 minutes of continuous sample infusion

continued on next page

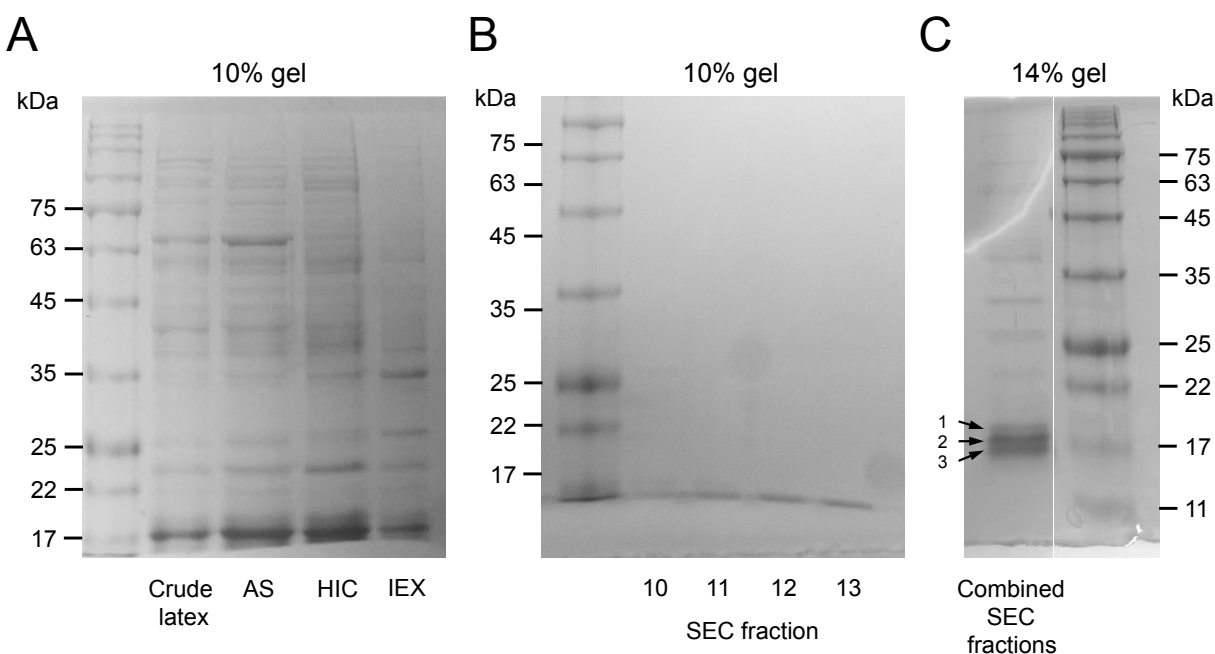
continued from previous page

(5 $\mu\text{L}/\text{min}$). Ionization was performed by ESI at room temperature. Mass error was < 2 ppm across all MS^n datasets, allowing reliable prediction of elemental formula (Supplementary Table 1). The expanded region (m/z 300-340) highlights the occurrence of a fragment with m/z 312, possibly a dehydration product of the parent ion. **(B)** MS^n analysis reveals similarities and differences between (7*S*)-salutaridinol and the m/z 330-byproduct. MS^n on the m/z 330-byproduct yielded ions shown in black. Although some byproduct ions were common with those obtained from similar (7*S*)-salutaridinol analysis (*a, f, g*), others (*c, d, e, h*) were not. Conversely, ions shown in red were unique to (7*S*)-salutaridinol MS^n (*i, j, k, l, m, n*) and were not observed in the m/z 330-byproduct spectra. An important fragmentation route for (7*S*)-salutaridinol is amine loss through cleavage of the piperidine ring ($-\text{CH}_2\text{CHNHCH}_3$), which was not evidenced in the case of m/z 330-byproduct. Instead the byproduct decomposed to fragment *a* by loss of a methylamine ($-\text{CH}_3\text{NH}_2$), leading to the formation of other observed ions. The presence of fragment *b* in m/z 330-byproduct MS^2 spectra could result from a loss of water, a fragmentation route common to ^{14}C -hydroxylated opiates. Fragment *b* could not be isolated in MS^3 spectra, thus, its degradation to fragment *c* is theoretical (dotted arrow) [Pennanen, K., Kotiaho, T., Huikko, K. & Kostinen, R., *J. Mass Spectrom.* 36, 791–797 (2001)]. The experiment was repeated independently three times with similar results.



Supplementary Figure 4. Partial purification of THS using multi-resin chromatography.

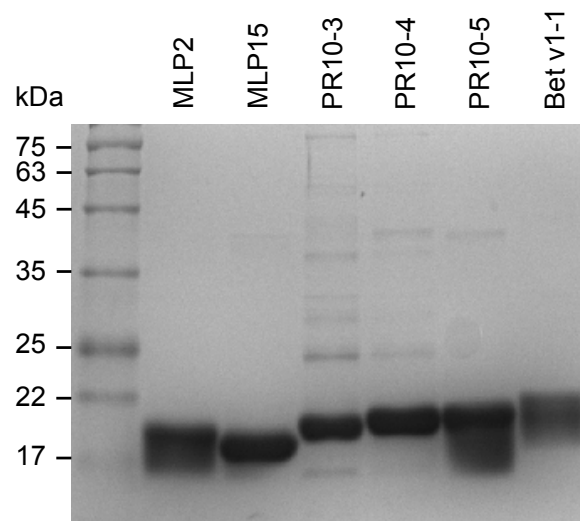
Elution profiles of protein fractionated in sequential order using (A) hydrophobic interaction chromatography (HIC) with a butyl-S sepharose column, (B) ion-exchange chromatography (IEX) with a HiTrap Q column, and (C) size-exclusion chromatography (SEC) using two 75 HR 10/30 column connected in series. Shaded portions of the elution profile indicating fractions that showed THS activity, were pooled, analyzed by LC-MS/MS, and subjected to the subsequent chromatography step. Assays were repeated twice with similar results.



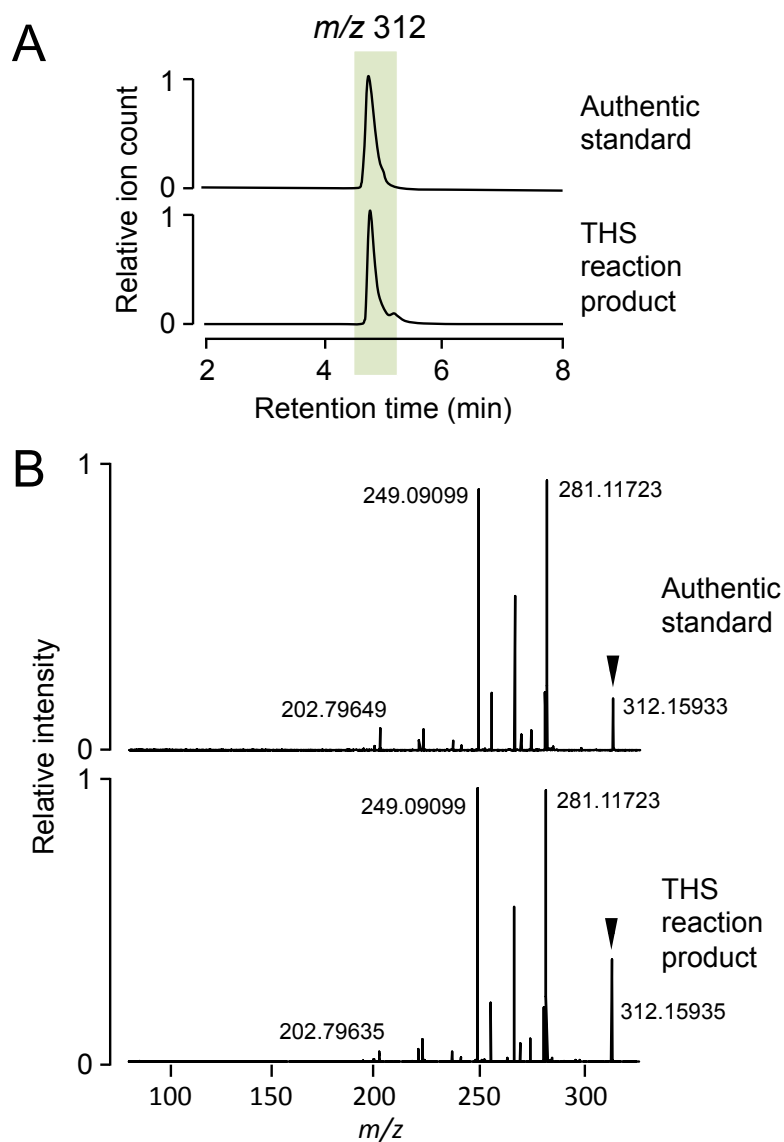
Supplementary Figure 5. SDS-PAGE analysis of latex protein chromatography fractions displaying THS activity. Approximately 20 μ g protein was loaded in each lane, and resolution was achieved using 10% (w/v) or 14% (w/v) acrylamide gels, as indicated. **(A)** Protein extracted directly from latex following a low-speed spin (5000 g) to remove insoluble material (i.e. crude latex) compared with THS-active protein enriched by ammonium precipitation (AS), hydrophobic interaction chromatography (HIC), and ion exchange chromatography (IEX). **(B)** Protein in fractions 10 through 13 obtained by size exclusion chromatography (SEC) showed THS activity. **(C)** THS-active fractions obtained by SEC were pooled and resolved by SDS-PAGE on a 14% (w/v) gel. Arrows indicate three prominent bands at ~17 kDa, which were excised from the gel and submitted for LC-MS/MS analysis. SDS-PAGE was repeated twice with similar results.

Betv1-1	MAP-LGVSGLVGKLSTELVDCDAEKYYNMYKHGEDVKKAVPHLCVDVKI	49
PR10-3	MAH-HGVSGLVGKIVTELEVNCNADEFYKILKRDEDVPRAVSDLFPPVKI	49
PR10-4	MAHPHPISGLVGKLVTELEVNCDAKYYKIFKHEDVKKAVPHMYTSVKV	50
PR10-5	MAH-HGVSGLVGKLVTELEVHCNADAYYKIFKHQEDVKKAMPHLYTGKVV	49
MLP-2	MAHAHGISGLVGKLITESEVNCNADKFFYQMFKHDENITNIIPHIYTSFKV	50
MLP-15	MAHQHTISGLVGKLITESEVNCNADKYYQIFKHEDLPSAIPHIYTSVKA	50
Betv1-1	ISGDPTSSGCIKEWNVNIDGKTIRSVEETHDDETKTLRHRVFEGLVMKD	99
PR10-3	AKGDGLVSGCIKEWDCVLDGKAMSGKEETTHNDETRTLRHRELEGLMKD	99
PR10-4	VEGHGITSGLVKEWGYLLEGKELIVKETTTTYTDETRTIHHSVAVGCHMTKI	100
PR10-5	ISGDATRSGCIKEWNYILEGKALIAVEETHDDETRTLTHRITGDLTKD	99
MPL-2	VEGDGLISGCTKEWGYLSEKARIVKEQTTFDDETRTIHHCAGDMMND	100
MLP-15	VEGHGTTSGCVKEWCYILEGKPLTVKEKTTYNDETRITNHNGIEGMMTD	100
Betv1-1	YKKFDTIMVVPKPKDCNGCVVTRSIYEKTNENSPTPFDYLFQCHQAIED	149
PR10-3	YKKFDSIIIEVNPKNCHGHSIVTWISIEYEKMNEDSPAPFAYIASFHQNVVE	149
PR10-4	YKKFDATLVVNPKPSCHGSTVSWTIDYEKINEDSPVPPIFYLAFFHKLIED	150
PR10-5	YKKFVKIIVEVNPKNCHGHSIVTVSLVYEKMNEGSPTPFNYLFQVHQTIIVG	149
MLP-2	YKKFVLTLLVNPKAHGQGSTVKWIIDYEKINEDSPVPEAYLSLCIKITEG	150
MLP-15	YKKFVATLLVVKPKANGQGSIVTWIVDYEKINEDSPVPFDYLAFFQQTIED	150
Betv1-1	MNKYLIRDSESN	160
PR10-3	VDSHLCLESE	158
PR10-4	LNSHLCASD	159
PR10-5	LNSHICAS	157
MLP-2	LNSHIYASE	159
MLP-15	LNSHLCASD	159

Supplementary Figure 6. Amino acid alignment of six THS candidates from opium poppy latex identified by LC-MS/MS. Highlighted residues are conserved in other candidates in comparison with Betv1-1. Abbreviations: MLP, major latex protein; PR, pathogenesis-related.

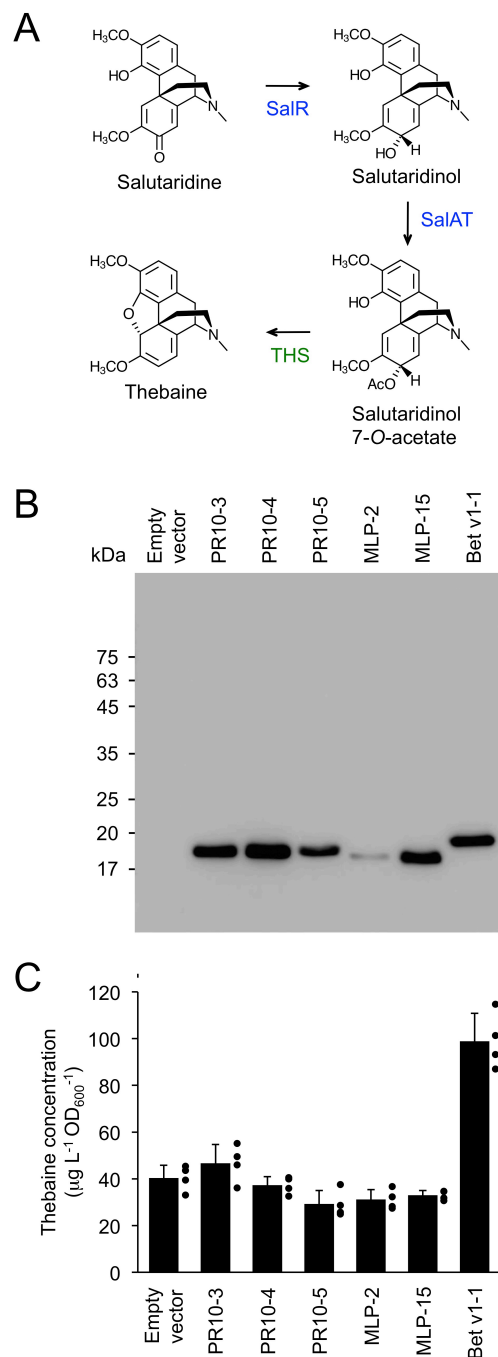


Supplementary Figure 7. SDS-PAGE of candidate proteins. SDS-PAGE of candidate proteins. His₆-tagged, codon-optimized THS genes were expressed in *Escherichia coli* and purified by affinity chromatography. Protein purification was repeated twice with similar results.

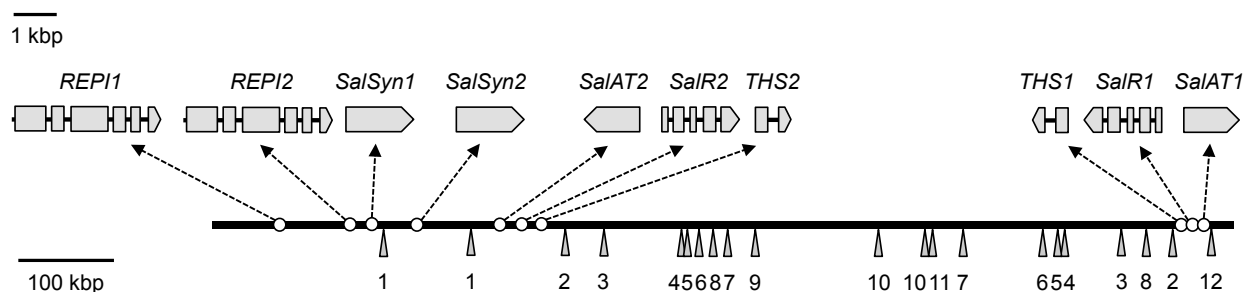


Supplementary Figure 8. Identification of the THS reaction product by LC-MS/MS.

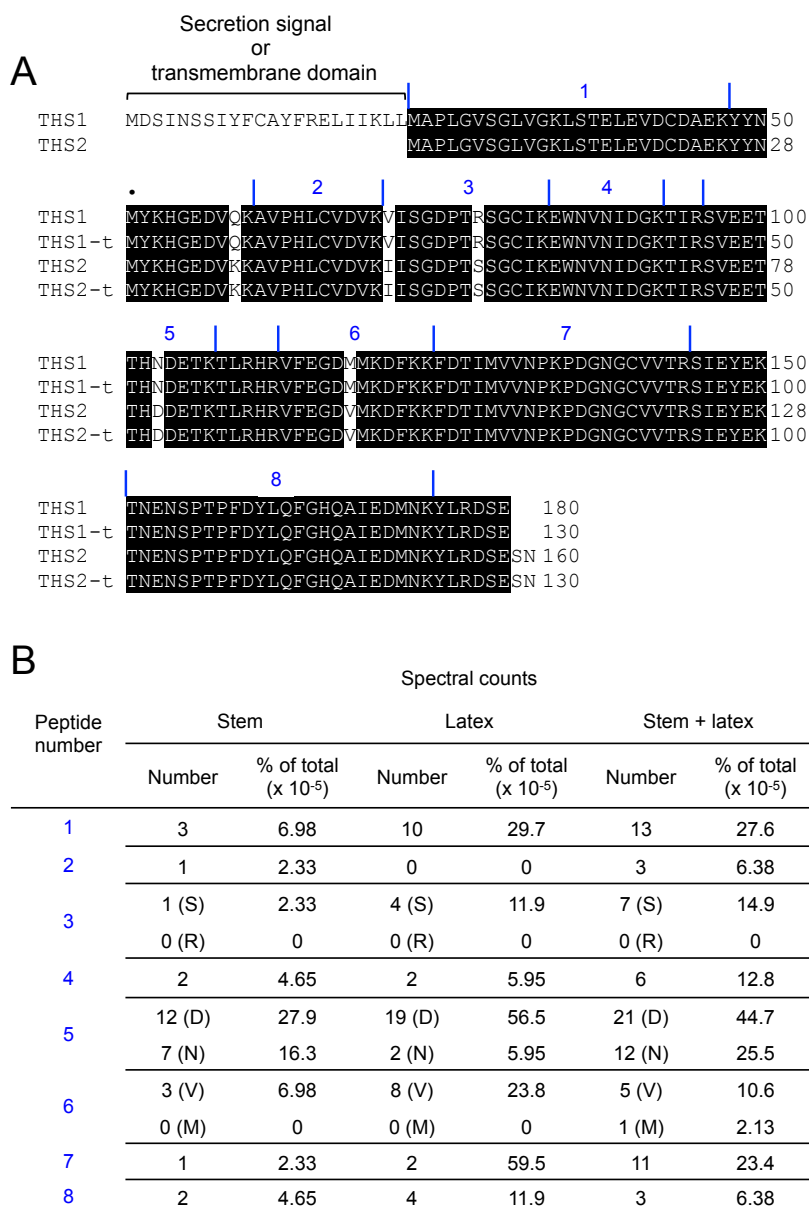
Extracted ion chromatogram at m/z 312 showing peaks for an authentic thebaine standard and the THS reaction product (**A**). The MS² collision-induced dissociation (CID) spectra of the authentic thebaine standard and the THS reaction product (**B**). Arrowheads indicate the parent ion. Exact masses for only a representative selection of MS² CID fragments are shown. A complete list of corresponding high-resolution MS² CID fragments is provided in Supplementary Table 4. Experiments were repeated independently three times with similar results.



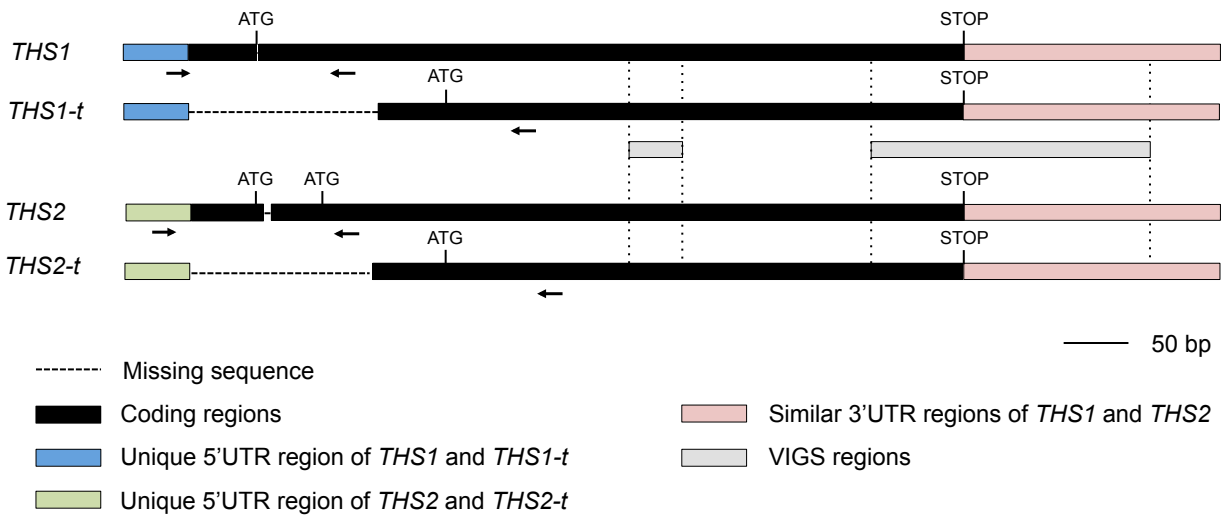
Supplementary Figure 9. *In vivo* assay of yeast expressing six candidate genes identifies Bet v1-1 as THS. (A) Schematic representation of the metabolic pathway used for assays in yeast. *Saccharomyces cerevisiae* strain CEN.PK102-5B with chromosomally integrated copies of genes encoding salutaridine reductase (SalR) and salutaridine 7-*O*-acetyltransferase (SalAT) (Sc-1; Supplementary Table 7) was transformed with either the pEV-1 empty vector or a construct harboring one of six THS candidate genes. Yeast fermentation was performed with the addition of 100 μM salutaridine, which was converted *in vivo* to salutaridinol 7-*O*-acetate by SalR and SalAT. THS activity (green) was identified by an increase in thebaine concentration in the culture medium. Panels (B) and (C) show immunoblot detection of HA-tagged protein, and thebaine content in the culture medium, respectively, for yeast 24 h after the addition of salutaridine. Only the Bet v1-1 candidate increased thebaine accumulation. The immunoblot (B) was repeated twice with similar results. Bars (C) represent the mean $\pm$ standard deviation obtained using four independently transformed yeast strains (black dots).



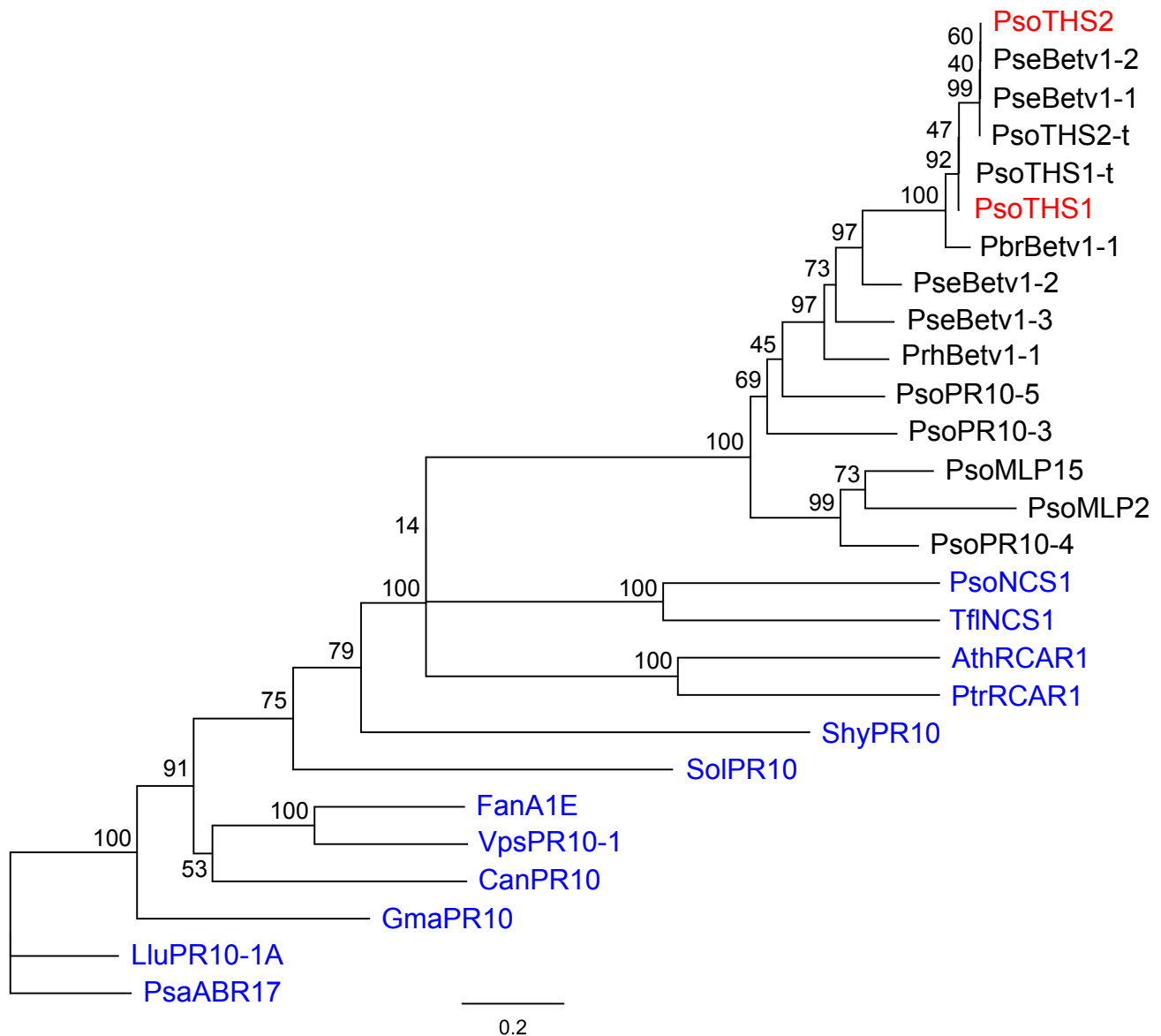
Supplementary Figure 10. Schematic representation showing the clustering of five genes involved in the biosynthesis of thebaine from (*S*)-reticuline. A draft opium poppy genome sequence revealed a scaffold of ~1 megabase pairs (Mb; contiguous black line) that included duplicate gene copies of reticuline epimerase (REPI), salutaridine synthase (SalSyn), salutaridine reductase (SalR), salutaridinol 7-*O*-acetyltransferase (SalAT), and thebaine synthase (THS). Open grey boxes and black lines represent the exons and introns, respectively, found in these five genes. Arrowheads indicate the location of other putative genes. Predicted translation products are as follows: 1, *O*-methyltransferase; 2, purine permease; 3, methylstylophine 14-hydroxylase; 4, B2 protein; 5, starch synthase; 6, IQ domain 1; 7, F-box protein; 8, transposon; 9, disease resistance protein; 10, E3 ubiquitin-protein ligase; 11, V-type proton ATPase; 12, SalAT-like enzyme.



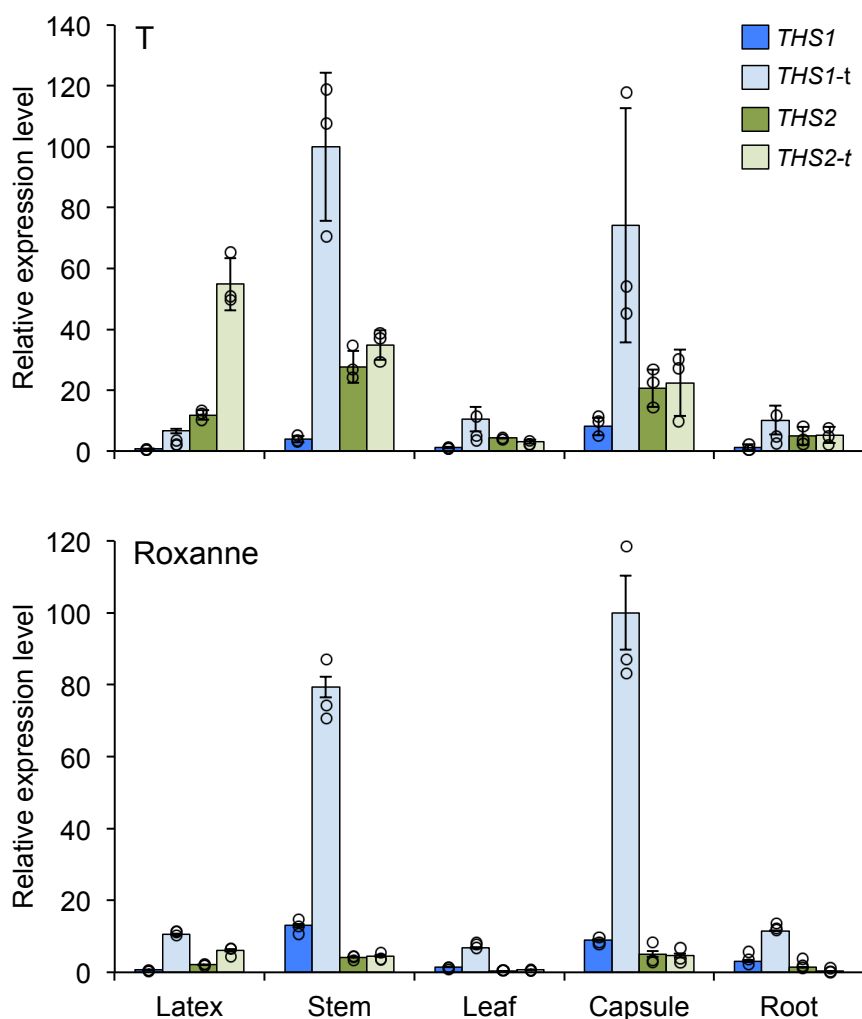
Supplementary Figure 11. Amino acid alignment and proteomics analysis of different THS isoforms. (A) Amino acid alignment of four THS isoforms, showing peptides detected by proteomics (numbered in blue). Residues strictly conserved across all candidates are highlighted in black. The unique 22 N-terminal amino acids of THS1 were predicted to comprise either a signal peptide or transmembrane region using two different programs (Phobius, <http://phobius.sbc.su.se>; SignalP 4.1, <http://www.cbs.dtu.dk/services/SignalP/>). The black dot marks the first Met residue common to all isoforms, and the presumed beginning of ‘short’ isoforms THS1-t and THS2-t. (B) Spectral counts obtained by LC-MS/MS analysis of opium poppy stem and/or latex tissues of detected peptides corresponding to THS1 and THS2 full-length or truncated isoforms. Three proteome databases for (i) latex, (ii) stem minus latex, and (iii) stem plus latex were screened for THS-specific peptides. The number of each peptide was recorded and expressed as a percentage of total peptides in each database. Bracketed letters designate isoform-specific residues. For example, peptide number 5 contained either aspartate (D) or asparagine (N), with 12 and 7 spectral counts respectively in stem, suggesting predominance of THS2/THS2-t over THS1/THS1-t isoforms in this tissue.



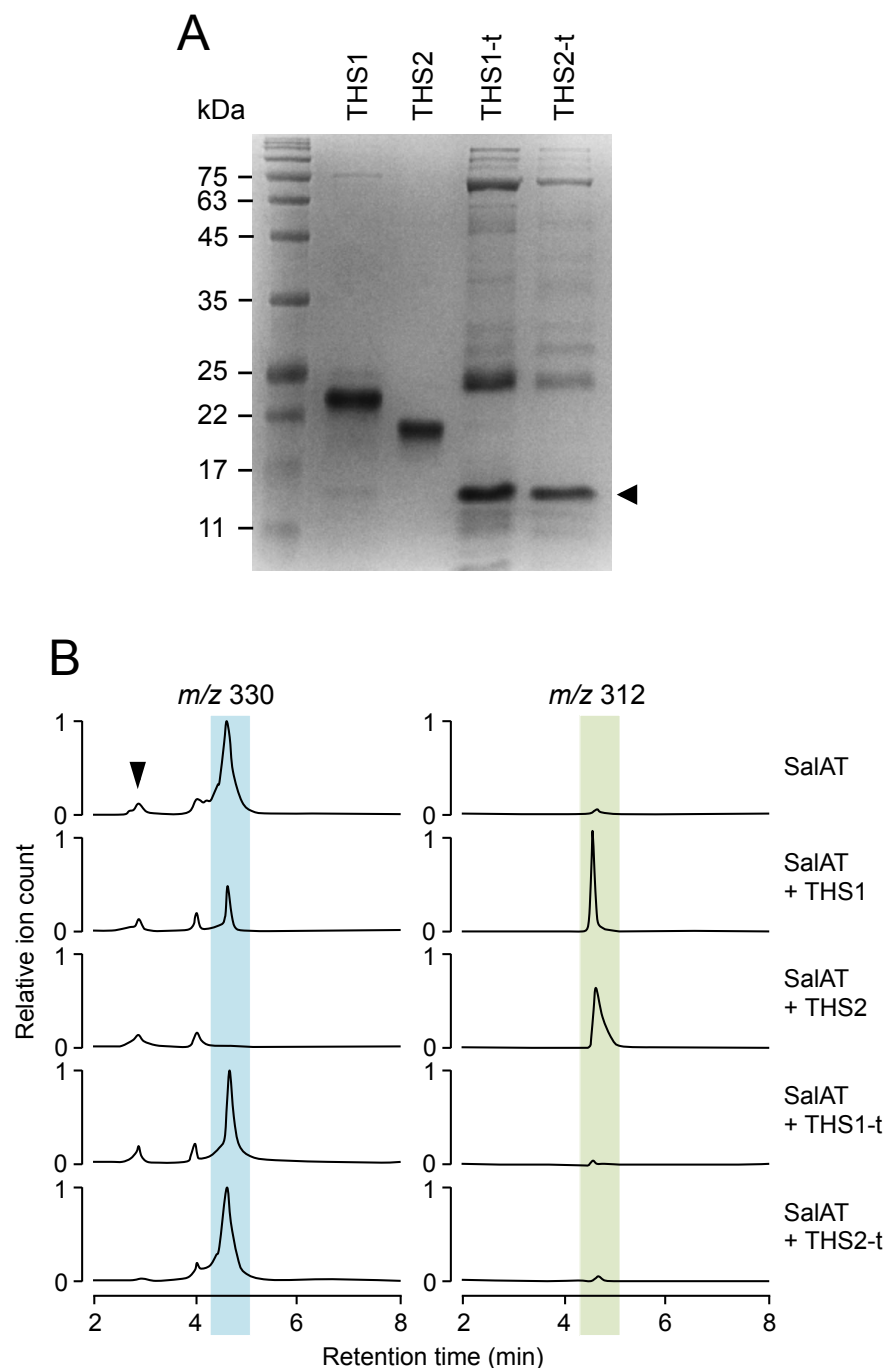
Supplementary Figure 12. Schematic representation of cDNA encoding four THS isoforms, showing regions targeted for isoform-specific amplification by RT-qPCR, and gene suppression by VIGS. Forward and reverse qPCR primers (black arrows) were designed to distinguish *THS1* and *THS2* full-length transcripts, and corresponding transcripts (*THS1-t* and *THS2-t*) with an internal deletion. Regions used to construct the pTHS-V vector for VIGS analysis (grey boxes) were conserved in all isoforms to ensure co-suppression of all THS transcripts. ATG designates potential start codons for each THS isoform. Note that in *THS2*, the first ATG is followed by four-nucleotide deletion resulting in a frame-shift mutation; thus, the subsequent ATG is the more likely start codon.



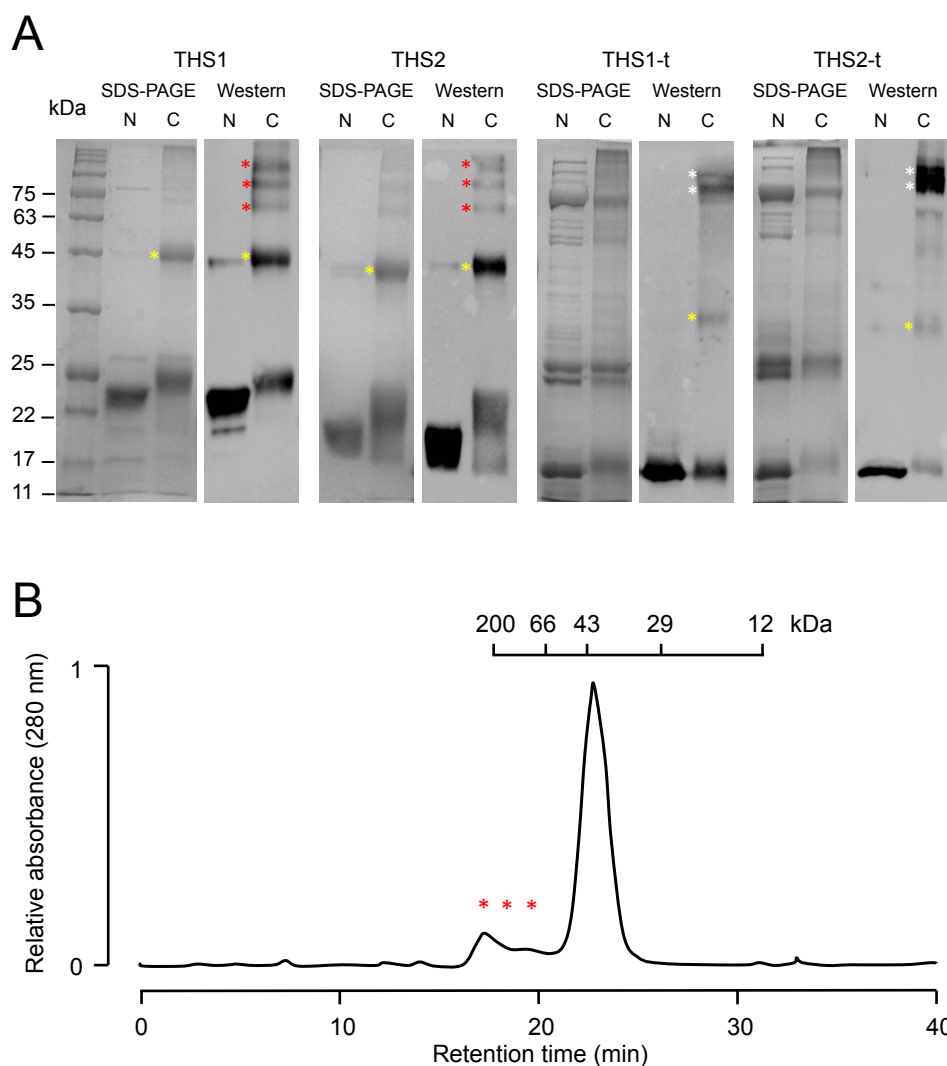
Supplementary Figure 13. Phylogenetic tree showing the relationships among THS variants from opium poppy and members of the Bet v1, PR10 and major latex protein (MLP) superfamily. Characterized THS paralogs are in red. Uncharacterized orthologs of THS were selected from *Papaver bracteatum* (Pbr), *P. rhoeas* (Prh), and *P. setigerum* (Pse) by blasting the 1000 Plant (1KP) transcriptome database (<https://db.cngb.org/blast4onekp/>) with the *P. somniferum* (Pso) THS2 sequence (Supplementary Note 3). Characterized PR10 proteins in blue are from *Papaver somniferum* (PsoNCS2, AY860500), *Thalictrum flavum* (TfINCS, Q67A25), *Arabidopsis thaliana* (AthRCAR1, OAP12964), *Populus trichocarpa* (PtrRCAR1, XP_002324762), *Saccharum hybrid* (ShyPR10, AMQ67074), *Spinacia oleraceae* (SolPR10, AGE12484), *Fragaria x ananassa* FanA1E, CAJ85646), *Vitis pseudoreticulata* (VpsPR1-1, ABC86747), *Capsicum annum* (CanPR10, AAF63519), *Glycine max* (GmaPR10, ADC31789), *Lupinus luteus* (LluPR10-1A, P52778), *Pisium sativum* (PsaABR17, Q06931) and *Thalictrum flavum* (TfINCS, Q67A25). An alignment was constructed using the CLUSTALW feature of Geneious (<http://www.geneious.com>) with a cost matrix of BLOSUM. The alignment formed the basis for phylogenetic tree construction using PHYMYL, employing the 'LG' substitution model [Le, S. & Gascuel, O., *Mol. Biol. Evol.* 25, 1307–1320 (2008)] and 100 iterations to yield bootstrap values. The scale bar refers to the number of substitutions per site.



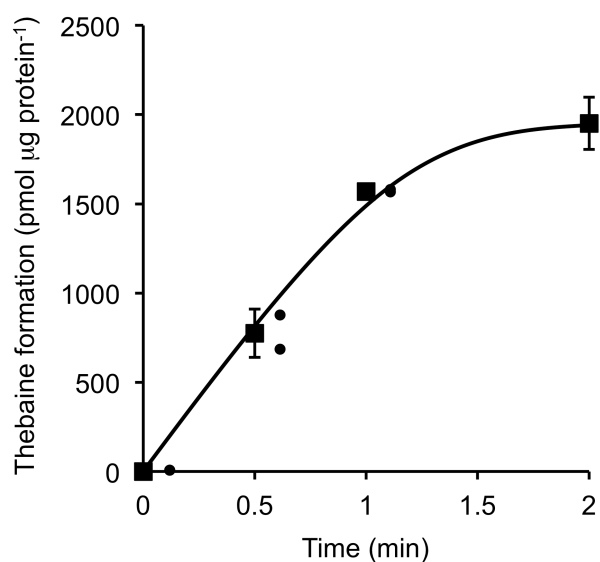
Supplementary Figure 14. Relative transcript abundance of four THS isoforms in different organs of two opium poppy chemotypes. The T chemotype accumulates high levels of thebaine and the related alkaloid oripavine [Desgagné-Penix, I., Farrow, S.C., Cram, D., Nowak, J. & Facchini, P.J., *Plant Mol. Biol.* 79, 295–313 (2012)], at the expense of morphine and codeine, and was used as the source of latex for THS purification. The Roxanne chemotype accumulates moderate to high levels of morphine and codeine and was the source of DNA used to the draft genome assembly. qRT-PCR was performed using primers listed in Supplementary Table 5, targeting regions unique to each THS isoform as shown in Supplementary Figure 12. Bars represent the mean $\pm$ standard deviation obtained using samples collected from three independent plant samples (open dots).



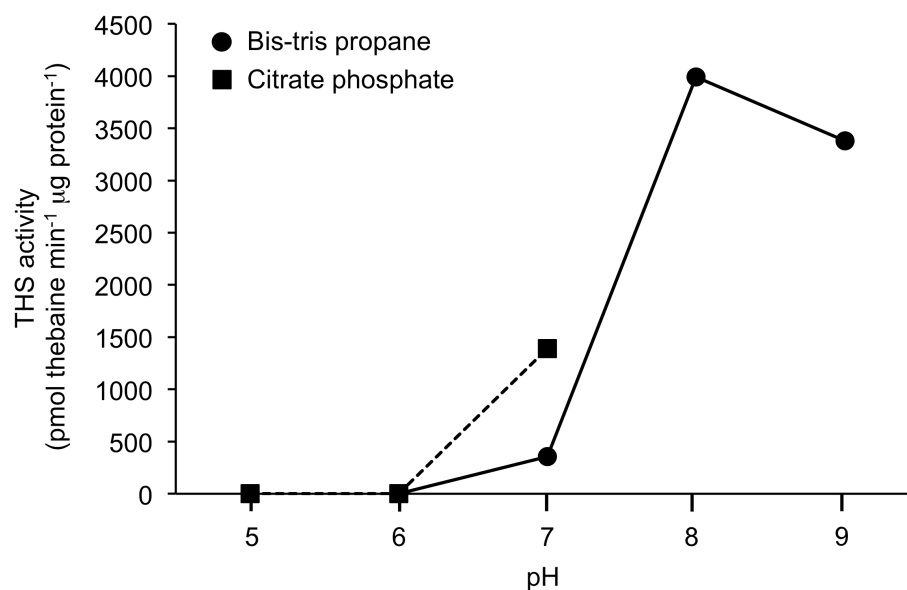
Supplementary Figure 15. SalAT-coupled assay conducted in the presence of different THS isoforms shows activity of THS1 and THS2, but not N-terminally truncated THS1-t and THS2-t isoforms. (A) SDS-PAGE of the THS isoforms used in the assays. His₆-tagged, codon-optimized THS genes were expressed in *Escherichia coli* and partially purified by affinity chromatography. (B) Coupled assays consisted of (7*S*)-salutaridinol substrate, purified recombinant SalAT enzyme, and one of four recombinant THS isoforms. Panels represent LC-MS extracted ion chromatographs (EIC) of either *m/z* 330 or *m/z* 312. The retention time (*R*_t) of the *m/z* 330-byproduct was ~4.8 minutes (blue highlight) in contrast with the (7*S*)-salutaridinol substrate, which eluted at *R*_t 3.1 minutes (arrow). The substrate, which was mostly consumed (arrow) by thebaine formation, exhibited the same exact mass and predicted elemental formula as the byproduct (Supplementary Table 1). Green highlight shows thebaine (*m/z* 312) eluting at *R*_t 4.8 minutes. Experiments were repeated independently three times with similar results.



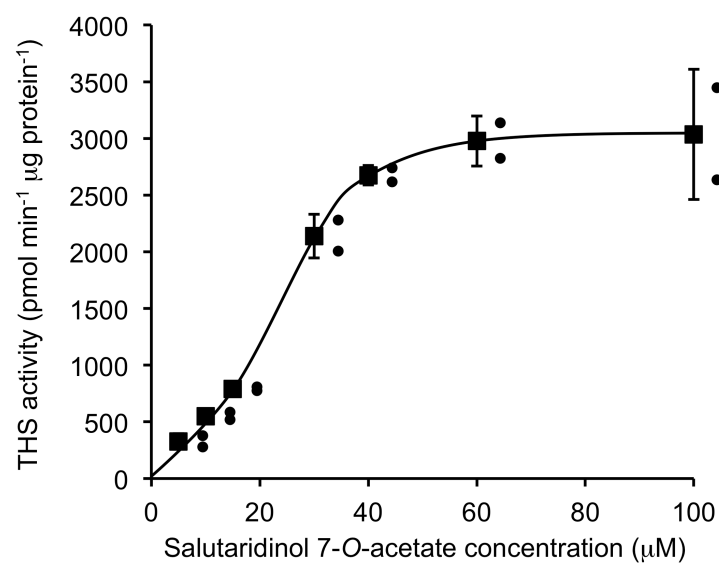
Supplementary Figure 16. Homodimerization of THS proteins. (A) Chemical cross-linking of recombinant, purified THS isoforms using BS3 (bis[sulfosuccinimidyl]suberate) followed by SDS-PAGE and immunoblot analyses, supporting the occurrence of THS1 and THS2 as homodimers (~40 kDa; yellow asterisk) and, to a lesser extent, higher-order oligomers (red asterisks). In contrast, protein bands corresponding to potential homodimers represented reduced bands in the case of the N-terminally truncated isoforms THS1-t and THS2-t. Abbreviations: N, native; C, cross-linked. (B) Fractionation of pure, recombinant THS2 protein by size exclusion chromatography (SEC) showed most of the active protein as a homodimer, with a small proportion of potential oligomers (red asterisks). Experiments were repeated independently three times with similar results.



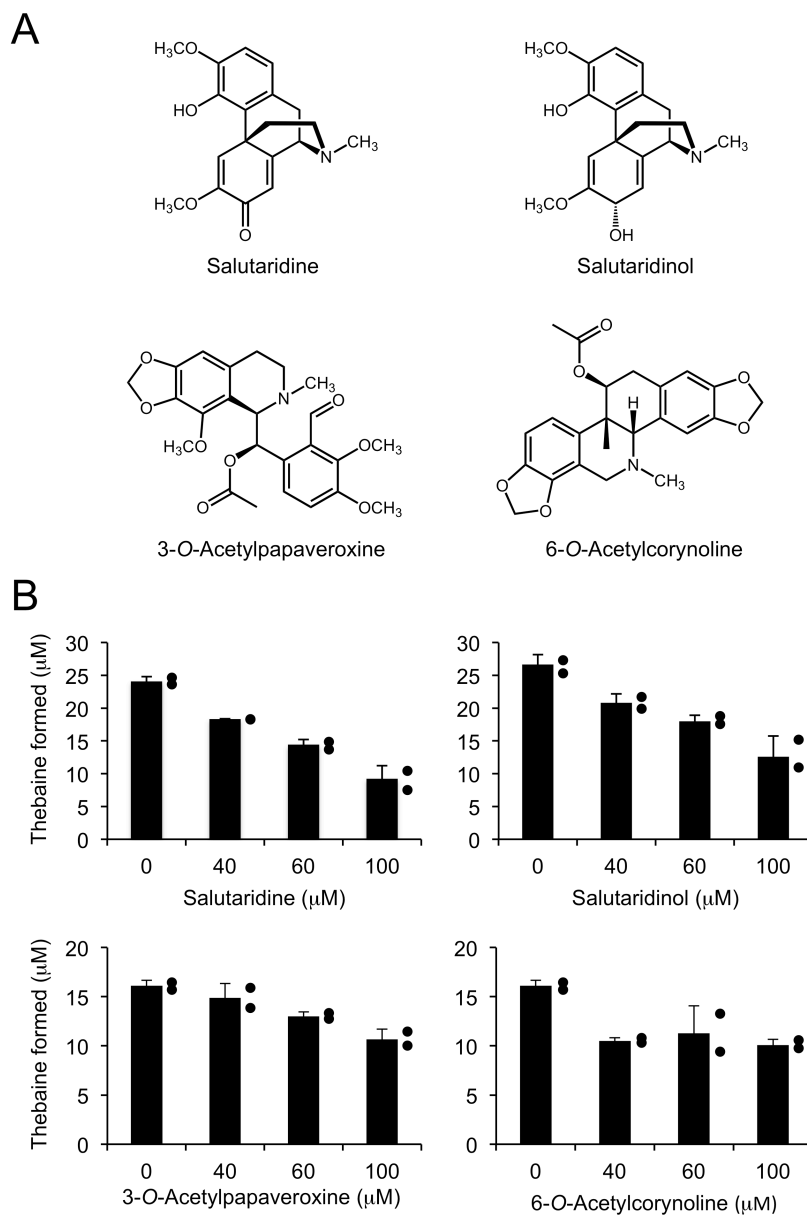
Supplementary Figure 17. Time dependence of thebaine formation from (7*S*)-salutaridinol 7-*O*-acetate. Standard, direct assays were performed using 40 μ M of (7*S*)-salutaridinol 7-*O*-acetate and 0.4 μ g of purified recombinant THS2. A short incubation time (30 sec) was essential to limit degradation of (7*S*)-salutaridinol 7-*O*-acetate to the *m/z* 330-byproduct and maintain linearity in the direct THS assay. Squares represent the mean $\pm$ standard difference of two independent assays (black dots).



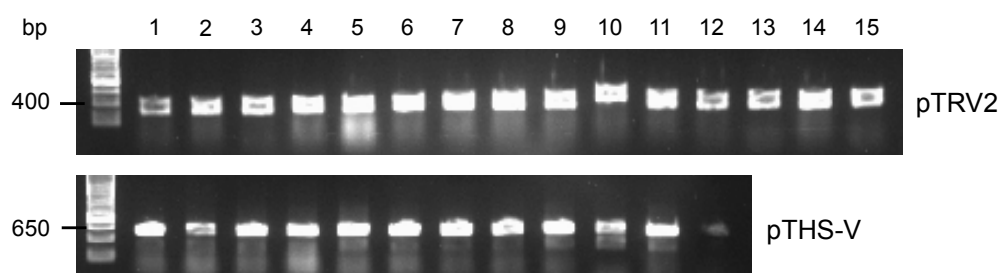
Supplementary Figure 18. Dependence of THS2 activity on pH. Assays were performed for 30 sec using 40 μM salutaridinol 7-*O*-acetate substrate, 0.6 μg of purified recombinant THS2, and pH-appropriate buffers as indicated. Experiments were repeated independently three times with similar results.



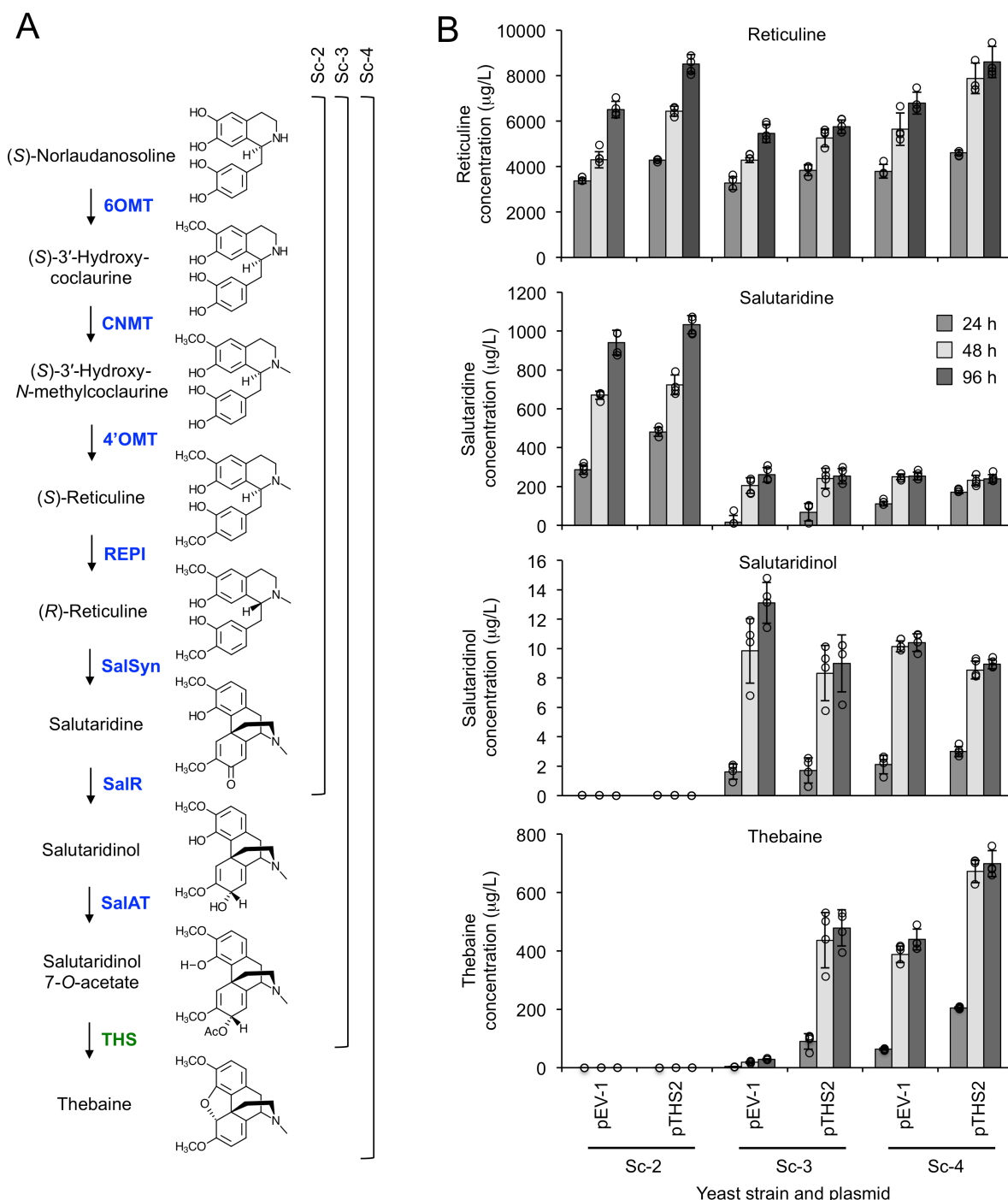
Supplementary Figure 19. Sigmoidal dependence of THS2 activity on substrate concentration reveals allosteric enzyme kinetic behavior. Standard, direct assays were performed for 30 sec using variable concentrations of (7*S*)-salutaridinol 7-*O*-acetate and 0.4 μg of purified recombinant THS2. Squares represent the mean ± standard difference of two independent assays (black dots).



Supplementary Figure 20. Substrate-competition assays for THS2. (A) Compounds tested as alternative substrates or competitive inhibitors of salutaridinol 7-*O*-acetate. None of the compounds were accepted by THS as alternative substrates. (B) Dependence of thebaine formation on the addition of increasing concentrations of potential inhibitors using the direct THS assay with (7*S*)-salutaridinol 7-*O*-acetate as the substrate. Bars represent the mean $\pm$ standard difference of two independent assays (black dots).



Supplementary Figure 21. Tobacco rattle virus (TRV) infection in opium poppy plants infiltrated with pTRV2 vectors for viral-induced gene silencing (VIGS). Following infiltration of seedlings with *Agrobacterium tumefaciens* harboring control (pTRV2) or *THS*-silencing (pTHS-V) constructs, stems were harvested at anthesis to isolate RNA for RT-PCR. Of ~30 plants subjected to infiltration with *A. tumefaciens* harboring each vector, 15 and 12 individuals tested positive for infection. The RT-PCR primers (TRV2-MCS_F and TRV2-MCS_R) targeted a region surrounding the multiple-cloning site of the pTRV2-based vectors, with predicted amplicon sizes of 400 bp (empty vector) and 650 bp (pTHS-V), respectively (Supplementary Table 5).



Supplementary Figure 22. Impact of THS2 on the yield of thebaine and the accumulation of selected pathway intermediates in yeast engineered with a multi-step, opiate production pathway. (A) Three yeast strains were built by genomic integration of BIA biosynthetic genes, enabling production from (*R,S*)-norlaudanoline of salutaridine (Sc-2), (7*S*)-salutaridinol 7-*O*-acetate and low levels of thebaine (Sc-3), and high levels of thebaine (Sc-4). (B) Levels of reticuline, salutaridine, salutaridinol, and thebaine in strains Sc-2, Sc-3, and Sc-4 transformed with either pEV-1 or pTHS2 after 24, 48 and 96 h of fermentation. Bars (B) represent the mean $\pm$ standard deviation obtained from four independently transformed yeast lines (open dots).

Supplementary Table 1. High-resolution fragmentation analysis of m/z 330-hyproduct generated by degradation of (7*S*)-salutaridinol 7-*O*-acetate.

MS² 330 →

Empirical m/z	Relative abundance (%)	Theoretical m/z	Error (ppm)	Composition [M+H] ⁺ or [M] ⁺	Product ion
281.11725	100	281.11722	-0.1	C ₁₈ H ₁₇ O ₃	[MH-H ₂ O-CH ₃ NH ₂] ⁺
299.12782	61	299.12779	-0.1	C ₁₈ H ₁₉ O ₄	[MH-CH ₃ NH ₂] ⁺
249.09117	40	249.09101	-0.6	C ₁₇ H ₁₃ O ₂	[MH-H ₂ O-CH ₃ NH ₂ -CH ₃ OH] ⁺
271.13309	22	271.13287	-0.8	C ₁₇ H ₁₉ O ₃	[MH-CH ₃ NH ₂ -CO] ⁺
267.10185	9	267.10157	-1.0	C ₁₇ H ₁₅ O ₃	[MH-CH ₃ NH ₂ -CH ₃ OH] ⁺
221.09624	9	221.09609	-0.7	C ₁₆ H ₁₃ O	[MH-H ₂ O-CH ₃ NH ₂ -CH ₃ OH-CO] ⁺
266.09402	4	266.09375	-1.0	C ₁₇ H ₁₄ O ₃	-
330.17005	4	330.16998	-0.2	C ₁₉ H ₂₄ NO ₄	MH ⁺
237.09127	2	237.09101	-1.1	C ₁₆ H ₁₃ O ₂	-
239.10692	2	239.10666	-1.1	C ₁₆ H ₁₅ O ₂	[MH-CH ₃ NH ₂ -CH ₃ OH-CO] ⁺
256.10963	2	256.10940	-0.9	C ₁₆ H ₁₆ O ₃	-
312.15968	0.4	312.15942	-0.8	C ₁₉ H ₂₂ NO ₃	[MH-H ₂ O] ⁺

MS³ 330 → 299 →

Empirical m/z	Relative abundance (%)	Theoretical m/z	Error (ppm)	Composition [M+H] ⁺ or [M] ⁺	Product ion
281.11730	100	281.11722	-0.3	C ₁₈ H ₁₇ O ₃	[MH-CH ₃ NH ₂ -H ₂ O] ⁺
249.09113	32	249.09101	-0.5	C ₁₇ H ₁₃ O ₂	[MH-CH ₃ NH ₂ -H ₂ O-CH ₃ OH] ⁺
299.12790	23	299.12779	-0.4	C ₁₈ H ₁₉ O ₄	[MH-CH ₃ NH ₂] ⁺
221.09617	21	221.09609	-0.4	C ₁₆ H ₁₃ O	[MH-CH ₃ NH ₂ -H ₂ O-CH ₃ OH-CO] ⁺
252.07822	9	252.07810	-0.5	C ₁₆ H ₁₂ O ₃	-
266.09396	7	266.09375	-0.8	C ₁₇ H ₁₄ O ₃	-
267.10179	6	267.10157	-0.2	C ₁₇ H ₁₅ O ₃	[MH-CH ₃ NH ₂ -CH ₃ OH] ⁺
271.13305	6	271.13287	-0.7	C ₁₇ H ₁₉ O ₃	[MH-CH ₃ NH ₂ -CO] ⁺

MS³ 330 → 281 →

Empirical m/z	Relative abundance (%)	Theoretical m/z	Error (ppm)	Composition [M+H] ⁺ or [M] ⁺	Product ion
281.11722	100	281.11722	0	C ₁₈ H ₁₇ O ₃	[MH-H ₂ O-CH ₃ NH ₂] ⁺
249.09104	94	249.09101	-0.1	C ₁₇ H ₁₃ O ₂	[MH-H ₂ O-CH ₃ NH ₂ -CH ₃ OH] ⁺
266.09383	23	266.09375	-0.3	C ₁₇ H ₁₄ O ₃	-
221.09618	4	221.09609	-0.4	C ₁₆ H ₁₃ O	[MH-H ₂ O-CH ₃ NH ₂ -CH ₃ OH-CO] ⁺
218.07271	4	218.07262	-0.4	C ₁₆ H ₁₀ O	-
234.06759	4	234.06753	-0.3	C ₁₆ H ₁₀ O ₂	-

MS³ 330 → 249 →

Empirical m/z	Relative abundance (%)	Theoretical m/z	Error (ppm)	Composition [M+H] ⁺ or [M] ⁺	Product ion
249.09105	100	249.09101	-0.2	C ₁₇ H ₁₃ O ₂	-
221.09614	45	221.09609	-0.2	C ₁₆ H ₁₃ O	-
234.06760	22	234.06753	-0.3	C ₁₆ H ₁₀ O ₂	-
218.07270	12	218.07262	-0.4	C ₁₆ H ₁₀ O	-
217.06484	8	217.06479	-0.2	C ₁₆ H ₉ O	-

Supplementary Table 2. Partial purification of THS from opium poppy latex tissue.

Specific activity was determined using a SalAT-coupled assay. Abbreviations: AS, ammonium sulfate; HIC, hydrophobic interaction chromatography; IEX, ion exchange chromatography; SEC, size-exclusion chromatography.

Purification step	Total protein (mg)	Specific activity (m/z 312 ion counts μg^{-1} protein)	Fold purification	Yield (%)
AS	48	1254087	1	100
HIC	16	1410214	1.1	37.4
IEX	4.2	5045200	4.0	35.2
SEC	0.1	25457966	20.2	4.2

Supplementary Table 3. Six THS candidates identified by LC-MS/MS proteomics analysis of partially purified proteins following size exclusion chromatography. The six candidates were the major proteins detected in three ~17-kDa bands on a SDS-PAGE 14% polyacrylamide gel. Accession numbers correspond to transcripts derived from a latex transcriptome that matched peptides identified by LC-MS/MS. Spectral counts for each peptide indicate approximate abundance of each candidate in each of the original three gel bands.

Accession Number	Molecular weight (kDa)	LC-MS/MS spectral count			Annotation
		Band 1	Band 2	Band 3	
c15408_g1_i1	18	39	22	17	Bet v1-1
c25055_g1_i1	18	58	34	37	PR10-3
c38417_g1_i1	17	32	75	59	PR10-5
c50593_g1_i1	18	18	29	24	PR10-4
c37788_g2_i4	16	8	15	55	MLP-2
c27108_g1_i1	18	16	42	92	MLP15

Supplementary Table 4. Identification of the THS reaction product by LC-MS/MS.
The collision-induced dissociation spectra of an authentic thebaine standard and the enzymatic reaction product are shown.

Authentic standard		THS reaction product	
<i>m/z</i>	Relative abundance	<i>m/z</i>	Relative abundance
281.11723	100	281.11723	99
249.09099	97	249.09099	100
266.09374	57	266.09373	57
312.15933	19	312.15935	37
281.14100	24	281.14100	24
280.13322	21	280.13320	20
255.10159	21	255.10158	22
223.07533	8	223.07532	8
269.11722	6	269.11724	7
241.08590	2	241.08593	2
237.09104	4	237.09093	4
221.09606	4	221.09604	5
202.79649	8	202.79635	4

Supplementary Table 5. Sequence-specific primers used for *THS* gene or *THS* transcript amplification from gDNA or cDNA. qRT-PCR primer sets targeted regions specific for each *THS* isoform (Supplementary Figure 9). cDNA/genomic primers were used to determine gene and transcript structure and relative abundance. VIGS primers were used to confirm the tobacco rattle virus infection of plants using either the pTRV2 or pTHS-V vectors. The synthetic VIGS fragment was for *THS*-silencing and consists of two fusion regions (Supplementary Figure 9).

qRT-PCR primers	Sequence	Target region
qTHS2F	GAATCCCCGGCCAACCTATC	<i>THS2</i> -specific
qTHS2R	ACCTGAAACACCGAGAGGAGC	<i>THS2</i> -specific
qTHS2SF	TAACAGCTAATCCTGAACAGTTTAGTCG	<i>THS2s</i> -specific
qTHS2SR	CGCAAAGATGAGGAACTGCCT	<i>THS2s</i> -specific
qTHS1F	GCAACATGCCAAAGTTAGAACAGC	<i>THS1</i> -specific
qTHS1R	GAAACACCGAGAGGAGCCATC	<i>THS1</i> -specific
qTHS1SF	GCAACATGCCAAAGTCTAGTCG	<i>THS1s</i> -specific
qTHS1sR	AACCTGAACGGGTTGGGTCT	<i>THS1s</i> -specific

cDNA/genomic PCR primers	Sequence	Target region
BV1_F1	TGGGCTCATATGTGACACTGCTTAACCAC	<i>THS2</i> 5' UTR
BV1_F2	TTACATCAATTCTTTATAGCAACATGCCAAAG	<i>THS1</i> 5' UTR
BV1_F3	ACAACGGGACTATACTACTAATACCGAATC	<i>THS2</i> 5' UTR (nested with F1)
BV1_R1	TATGTACGTGCATAGGTACACGTACACC	<i>THS1</i> and <i>THS2</i> (conserved 3' UTR)
BV1_F4	CGATAATGGTAGTCAATCCAAAGCC	<i>THS1/2</i> forward
BV1_R2	GGCCTGATGGCCGAATTGTA	<i>THS1/2</i> reverse

VIGS primers	Sequence	Target region
TRV2-MSC_F	GGTCAAGGTACGTAGTAGAG	Vector MCS
TRV2-MSC_R	CGAGAATGTCAATCTCGTAGG	Vector MCS

Synthetic VIGS fragment		
AATGTTAACATTGATGGTAAGACGATTGCTCGGTACAATTCGGCCATCAGGCCATTGAGGACATG AACAAGTACTTACGCGATTCTGAAAGTAACTGATTGATGCCATTTATACATGAACCCATTCCGGTGTT GGTGTACGTGTACCTATGCACGTACATATTTTATATTTATCAGTTGCAACTTGTGTGTATTTTATCGT TTCAGTGACTCCAATATTGTGTCAGTGACCCGGTGTATGGGTGTGATT		

Supplementary Table 6. Kinetic constants for THS2. All data was acquired using the direct THS assay. Values represent mean $\pm$ standard deviation of three replicates.

Substrate	$V_{\max}$ ($\text{pmol min}^{-1} \mu\text{g protein}^{-1}$)	η	K_A (μM)	K' (μM)	R^2
Salutaridinol 7-O-acetate	3216 ± 183	2.316 ± 0.352	21.97 ± 2.036	1282 ± 1277	0.972

Supplementary Table 7. Engineered yeast (*Saccharomyces cerevisiae*) strains used for alkaloid fermentation.

Strain Name	Genotype
CEN.PK	<i>Mata</i> ; <i>ura3-52</i> ; <i>trp1-289</i> ; <i>leu2-3,112</i> ; <i>his3Δ1</i> ; <i>MAL2-8C</i> ; <i>SUC2</i>
Sc-1	CEN.PK102-5B P_{GAL10} -PsomSalR- T_{CYC1} P_{GAL1} -PsomSalAT- T_{ADH1}
Sc-2	CEN.PK2-1C $T_{MET13}::T_{CBS1}$ P_{GAL1} -Pbra6OMT- T_{ADH1} $T_{LYS2}::T_{CST1}$ P_{GAL10} -Psom4OMT- T_{PMA1} P_{GAL1} -CjapCNMT- T_{PGK1} $T_{LYS4}::T_{CHS2}$ P_{GAL10} -PsoREPI-2- T_{ADH1} $T_{MET17}::T_{TUB1}$ P_{GAL10} -PbrSalSyn- T_{ADE3} P_{GAL1} -PbrCPR1- T_{PDR5}
Sc-3	Sc2 $T_{MET16}::T_{CHS2}$ P_{GAL10} -PsomSalR- T_{HPC2} P_{TEA1} -PsomSalAT- T_{ROX3}
Sc-4	Sc3 $T_{NCL1}::T_{STE5}$ P_{GAL1} -PsomBetv1-1- T_{ADH1}

Supplementary Table 8. MRM transitions of alkaloid standards.

Compound Name	Parent (<i>m/z</i>)	Daughter (<i>m/z</i>)	Cone (V)	Collision (V)
Reticuline	330	192	38	20
Salutaridine	328	58	48	26
Salutaridinol	330	239	40	24
Thebaine	312	58	22	16
<i>m/z</i> 330-byproduct	330	249	40	20
Thebaine- <i>N</i> -(<i>methyl</i> -d3)	315	61	30	18

Supplementary Note 1. Sequences of integration cassettes for construction of yeast strains.

Name	Sequence
$T_{MET13}::T_{CBS1} P_{GAL1}$ -Pbra6OMT-T _{ADH1}	CATTCTATCATCACTATAAACTCTCAACCTCAAGTCAACGGCATTAGGTCCAATGACAAAATTCATGG TTGGGGACCCAAGGATGGTTACGTTTACCAGAAGCAATATTTGGAATTTATGTTGCCCAAGACTAAG TTGCCCAAGTTGATTGACACCTTGA AAAACAATGAGTCTTGACCTACTTCGCCATCGACTCTCAAG GTGACCTGCTAAGTAATCATCCAGACAACCTCCAAGTCCAACGCTGTGACTTGGGGTATTTTCCCGG CAGAGAAATTTCTCAACCTACCATTGTGCGAGAAAATTCGTTCTTAGCGTGGAAGGAGGAGTTCTAT CATATCTTGAATGAATGGAACTAAACATGAATAAATACGATAAACCCGCATAGTGCCCAATTCATTCA GTCCTTGATTGACGATTACTGCTTGGTCAATATTGTTGACAATGACTACATTTCTCCTGACGACCAAA TTCATTCTATTCTTCTAAGTCTATAGTACAAAAGAAACGATGCCTTTTTGTCCGTTATCCCTTTGTTTT ATTTACAATAAGTACGTAAAAACAATTATTTCTTATTTTTGTTTATAATGTTACAAATCTTTTTATCCCC ATTCTTCCCAATTTTTTTTTCTTTTCTTTTCTTAAATTAGGGCATGAAGCCCATTTGATTGTTAA TCACTGTCACTGTGCTACTATCACTACTGTCTTCTTCAAGATCGGAATAATTATTATTATTTATGATT TCATCCTTTTCAAAAATTTCAAGGTTTTCTTTTCTTGTAGTTGGCGCAGTCGAGTAGTTGCCACTT AGTTTTCTTTTGTCTATTGTTGTTATTTTTCTTCAAAATATCTCTCATAAAATTTGCAATAGTCGCTT CATCTAAATATCCAAATCAATTTCAATTTCACTTTCATTGGAAGTTAGTATTTCACTCTCGGGTAATG AATTTTTAATGATACGAATCATCCCATTTCTAAATCTAGAGCTCGAGGACGGATTAGAAGCCG CCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTCCTCGTGCGTCTCCTGTTCCACCGGTGCGG TTCCTGAAACGCAGATGTGCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATATGACTCTT TTATGGTTATGAAGAGGAAAAATTTGGCAGTAACCTGGCCCCACAAACCTTCAATGAACGAATCAAA TTAACAACCATAGGATGATAATGCGATTAGTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAG CGATGATTTTTGATCTATTAACAGATATATAAATGCAAAAACCTGCATAACCACTTTAACTAATACTTTC AACATTTTTCGGTTTGATTACTTCTTATTCAAAATGTAATAAAAGTATCAACAAAAATTTGTAATAACC TCTATACTTTAACGTCAAAATGGAGACTGTCTCTAAGATTGACCAACAGAATCAGGCTAAGATTTGGA AACAGATTTACGGTTTCGCTGAGTCTTTAGTTTTGAAATGCGCTGTCCAATTAGAAATAGCTGAGACT TTACATAACAATGTCAAACCTATGTCATTGTGAGAATTAGCTTCAAAATTACCAGTCCCACAGCCAGT CAATGAGGATAGGTTGTTGAGAATAATGAGGTATTTAGTTCACATGGAGTTGTTCAAAATAGATGCAA CTACACAAAAGTATTCTATTGGCACCACCTGCTAAGTATTTGTTGAGGGGTTGGGAGAAATCTATGGT TGACTCTATATTGTGCATTAACGATAAGGATTTCTTGGCTCCATGGCATCACTTGGGAGATGGATTGA CAGGTAACGTGTGACGCTTTTGAAGGGCTTTAGGTAATCTATATGGGTCTATATGTCTGAGAATCCT GAAAAGAATCAGTTGTTCAATGCAGCTATGGCTTGCATACAAGATTAGTTACATCTGCATTGGCAAA TGAGTGTAAAGTCTATATTTTCAAGATGGAATTAAGACTTTAGTTGACGTTGGAGGTGGAAGTGAAGT GCTGTTAAGACTATATCTAAGGCTTTTCCAGACATTAAGTGTACAATTTACGATTTACCACACGTTATA GCTGACTCTCCTGAAATACCAAAACATTACAAAAATATCTGGAGATATGTTTAAAGTCTATACCATCAGC TGACGCTATATTTATGAAATGTATTTTACATGACTGGAACGATGAGGAATGTATTCAGATTTTGAAAA GGTGTAAAGGACATTGGCTTAAAGACGGTAAAGTTATAATTGTCGATGTCGTTATAGATATGGATTTC AACTCATCCATACGCTAAGATTAGGTTGACATTGGACTTGGACATGATGTTGAACACTGGAGGTAAG GAAAGGACTAAGGAAGAGTGGAAAGATTTGTTTACGCTGCTGGTTTCGCTTCACATAAGGTTACAC AAATATCTGCTGTTCAAGTCTGTCATTGAGGCTTATCCATACTAAATCGATCGGAATTTCTTATGATTTA TGATTTTTTATTATAAAGTTATAAAAAAATAAGTGATATACAAATTTTAAAGTACCTTTAGGTTTT AAAACGAAAATTTCTTATTCTTGAGTAACCTTTTCTGTAGGTGACGTTGCTTTCTCAGGTATAGCATG AGGTCGCTCTTATTGACCACACCTCTACCGGATGCCGAGCAAATGCCTGCAAATCGCTCCCCATTT CACCCAATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGATTTTATGTCCTCA GAGGACAACACCTGTTGTAATCGTTCTCCACACGGATCCACAGCCTAGCCTTCAGTTGGGCTCTAT CTTCATCGTCATTGATGCTACTAGCCCCCTACCTGAGCTTCAAGACGTTATATCGCTTTATGT ATCATGATCTTATCTTGAGATATGAATACATAAATATATTTACTCAAGTGATACGTGCATCCTTTTTTT ACGGCAGCATTTTTTTTTTCAACTCTGATCGCCCCCTCCTGCAGAATTCGGATCCGGTACCATGGTAA CACAGACAACGATAACAGTTCTTTAACAGAATTGTCAACCCCCCTCTTTGATTACATTCAACATTA CATTGCATTTTTTTTTTTTTATTCAATTGACTCATGTTCTTTCTTTTTTACCAGTTTGTCTCTTT TCCATTGTTTTACCATTCTTTTTTAACTAGCATTTCACATGTGTTTGGTTAACACCCCTTCTTTTTTT CAGGAAAATCCTTTCAATTTCTCTCATACTTTCAACAAAGTTTTTAAAGGTACTTTAAATAGTTCAAC ACCCTCTTCTTCAATTTATTTATCTCTTCAATTTCAACATACTCGAAAAGGAAGAACTAAAAGTAC TTACATTTTACATGTATGTATACCTATATATATATATATACTCTTATAGATATATTTACAAATTA AGGAAAAATAATAAAATAACCTCCCTGTCAAGTTAAAACACGGCCCCATCACTTATAAATAGTCT CGACTCGTG
$T_{LYS2}::T_{CST1}$ P_{GAL10^-} Psom4OMT-T _{PMA1} P_{GAL1} -CjapCNMT- T _{PGK1}	GCTACGCTTTGAATCCTCCCAAAGAAAATGAATTGGCCGTTGCTCAAGTAACGGGTCAACCAAGAA TATTATTCAAAGACTACTTGATACCTTTACACGATTATGGTTACGATGTCGAAATCGAAAGCTATTCTA AATGGAAGAAATCATTGGAGGCGTCTGTTATTGACAGGAATGAAGAAAATGCGTTGATCCTTTGCT ACACATGGTCTTAGACAACCTTACCTGAAAGTACCAAAGCTCCGGAAGTACGATAGGAACGCCGT GGCATCTTTAAGAAAGACACCGCATGGACAGGTGTTGATTGGTCTAATGGAATAGGTGTTACTCCA GAAGAGGTTGGTATATATATTGCAATTTTAAACAAGGTTGGATTTTACCTCCCAACTCATAATGA CAAACTTCTCTACCTAGTATCGAATTGACTCAAGCGCAAATCAGTTTGGTTGCTAGTGGTCTGGA GCTAGAGGTAGCTCAGCAGCTGCATAGAGTTATTAGAGATTGATTCATAGAGTCCGAATATTTTTTT TGCAATATTCCTTTATTTTTACCTCATCTCACCCTCTCTTTATTTGCTCAAGGTACCCTGCCTTCA TGTTTTTACGTTTTTATAATCGACTACTTTATATCTATAACGAATTTTAAATTTCTCGATTTACTGTG TTTCTACCATTTTGCGAATTCGTGCCCCATCATGACTTATTTTATTTAAATTTTACATATCAAAGGACA CTGGTTTACAAACATAAAAAATGATTGTTGAGAATGCCAATTCGCCCCACTTACTACTTTTCATGTGTTA AATAGTACGCATAAACACGAGATCTGACTTCGATTGTTTCTCGTCTTCGTCAAACATCGGTTTCCAA

	TTAACCATTTC AATTAGCTGAGGAACCAATGGAAC TCGTTCTTAGCCAATTCGAAGAATTC TAAATC CTTCTCCAACAATTTTGGTCTATGCGGTGAAAACTACCAACGCTCTAGAGCTCGAGGAATAAATTA GTGTGTGTGCATTATATATATTAATAAATTAAGAATTAGACTAAATAAAGTGTTC TAAAAAATATTA GTTGAAATGTGCGTGTTGTGAATTGTGCTCTATTAGAATAATTATGACTTGTGTGCGTTCATATTTTA AAATAGGAAATAACCAAGAAAGAAAAAGTACCATCCAGAGAAACCAATTATATCAAATCAAATAAAC AACCAGCTTCGGTGTGTGTGTGTGTGAAGCTAAGAGTTGATGCCATTTAATCTAAAAATTTTAAGG TGTGTGTGTGGATAAAATATTAGAATGACAATTCGAGATGAAATTTTAAGCAAAC TCTAGTAGGAAAT AAGCGGCTTATTCTGTGGCTCCTAATTCCTTTTAGTGATCAGTCCCATTTGATAAAAAAATTA TAAAAATTAGAAAAATTAACCAAGAAAAATCAAGTTGATTAAATGTGACAAAAATTATGATTAAATGC TACTTCAACAGGAATCGATTTCATGGAAAAGCCTCAATGACAGACTGAATTGCTCTTATTGGCCTAATC TTATGTGATCTGAAACCTGCTCTCTTCAACAAATTCCTCCAGTCATCAGCAGTCTTTCTCTACCTCC AGTATTAACTAACATGTCAATGTCCAATATCAACCTAGTCTTAGTTAATTCGTGATTTGATTCCTCTTC TAAAGCGACATCGACAATTATAACCTTACCTGTTTCCTTTGGTATAGCCTCTTTCGATTTCTTCAATAT ATTGACACACTCTTCGTCAAGTCCAATCGTGCAATATCAACTTCAACAAAATTGCTTGTGCAGAGGAA CTGATTTGAACATGTCACCAGGAACCTTTCTATATTAGGCAAGTCATGTGAATTTGCAACGACATGA GGCAAGTCATATAAAGTGCAATTTTATATGAGGAAATGCTTCGTAATTTGATTTGACCGTTACCGTTA CCTCCACCTATATCGACCAATGAATCCAAACCTTGAATGTATCCCTACAGTCTTCTATCAATGTTTTT GTCAATAATCTTGTTCCTCCAGCCATACCCTCATTGAACAAC TGAAGACTGATCAGGATTTCTTCTAA ATACTTCCATATATCCATACCATACCTTTTTCAAAGCTGTAGTTGAACCGTTACCTAAACCTTCTTT CATAAAGTGCCATGAAACCATAAAGTCTTTCTGAGTCATACCTAAATCATTGGGACCATTGATCTCT CAGCGTCTCTTAACAACAAAGTTCCAACAGGTTTCAAAGAGTAAACCTTTTCGACACCTCATTGCAA GTCTCTTGCTCAATTATATCAAATGGACCAAGTATCTGACCATCCTATATAAATAGTCAGATGAAAC GTTAGTTATTGGTAATTTTGCAGCTAATTGTGCCAATGTAATTCACCGTCAATTATTCTTAATAATGTC AGCTATACCTATTTTCGACAGCGCATCTTAAAACTAAAGAGTCAGCGAAACCGTAAATTTATTTCCACA ACTGAGCCTGATCCTTAATTGAAACCTCTTGTGTAGCAGCAGCAGGTTTTCGATTAAGATCCCAT GTTGAATTCGAATTTTCAAATTTCTTACTTTTTTTTTTGGATGGACGCAAAGAAGTTTAATAATCATATT ACATGGCATTACCACCATATACATATCCATATACATATCCATATCTAATCTTACTTATATGTTGTGGAA ATGTAAGAGAGCCCCATTATCTTAGCCTAAAAAAACCTTCTCTTTGGAACCTTCAGTAAACGCTTAAC CTCATTGCTATATTGAAGTACGGATTAGAAGCCGCGAGCGGGTGACAGCCTCCAGGAAGAC TCTCCTCCGTGCGTCTCGTCTTACCAGGTGCGGTTCTGAAACGCAGATGTGCCTCGCGCCGCAC TGCTCCGAACAATAAAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAATTGGCAGTAAC CTGGCCCCACAACCTTCAAATGAACGAATCAAATTAACAACCATAGGATGATAATGCGATTAGTTTT TTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTAAACAGATATATAATG CAAAAAC TGCATAACCACTTTAACTAATCTTTCAACATTTTCGGTTTGTATTACTTCTTTATTCAATGT AATAAAAGTATCAACAAAAAATTGTTAATATACCTCTATACTTTAACGTCAAATGGCAGTGGAAGCTAA ACAAACTAAGAAGGCAGCGATCGTGGAAC TACTGAAGCAGTTGGAATTTGGGTCTAGTTCATACGA CGATATTAAGCAGTTGATACGGCGTGAAC TGGCAAGAAGGTTGCAGTGGGGTTACAAGCCAACTTA CGAAGAACAGATTGCTGAAATACAAAAC TTTGACTCACTCTTTACGTCAGATGAAGATCGTACTGAA GTAGAAACTTTGGACTCTCAGTTGTACGAGATCCCTATTGAATTTCTGAAGATCATGAACGGTTCAAA CCTAAAGGGCTCTTGCTGTTACTTCAAGGAGGACTCAACTACACTAGACGAAGCAGAAATCGCCATG CTAGACCTGTACTGTGAACGTGCTCAGATCCAGGATGGCCAAAGTGATTGGACTTGGGTTGTGGT CAGGGAGCTTTAACCTTGCATGTGGCTCAGAAGTACAAGAAGTGTGCTGTCACTGTGTTACAAACT CTGTGAGTCAAAAGGAGTACATCGAGGAAGAAAGCCGAAGAAGGAAC TTAACCTGGAAGTAA AGCTGGCGGATATTACAAC TCAAGATGGCAGAACTTACGACAGAATACTGGTGATCGAACTCTT CGAGCACATGAAGAACTACGAAC TGTATTACGTAAGATATCGGAATGGATTTCTAAGGACGGCTTG TTGTTCTAGAACATATCTGTACATAAGACTTTGCTTACCCTATGAACCATTAGACGACGATGATTG GTTACAGAGTACGTGTTCCCGCTGGTACTATGATAATACCAAGTGCAAGTTTCTCTGTACTTCC AGGATGACGTTTCTGTGGTAAATCATTGGACTCTTAGTGGCAAGCACTTCTCTCGCACTAACGAAGA GTGGTTAAAGAGACTGGACGCTAACCTTGATGTAATAAAGCCCATGTTTGAACCTCTGATGGGCAAC GAGGAAGAAGCGGTCAAGCTGATAAACTACTGGCGTGGTTTCTGTTGAGTGGCATGGAATGTTT GGTTACAACAACGGAGAAGAGTGGATGGCTAGTCACGTTCTGTTCAAAAAGAAAGTAAATCGGATATTG AATTGAATTGAAAACGATAGATCAATTTTTTTCTTTTCTTTTCCCCATCCTTACGCTAAAATAATAGT TTATTTTATTTTTGAATATTTTTTATTTATATACGTATATATAGACTATTATTTATCTTTAATGATTATT AAGATTTTTATTAATAAAAAAATTCGCTCCTCTTTAATGCCTTTATGCAGTTTTTTTTTCCCATTGATA TTTCTATGTTCGGGTTCAGCGTATTTTAAGTTTAATAACTCGAAAAATTCGCGTTCGTTAAAGCTTTCC AGAAGGATATTATTTGAAATAAACCGTGTGTGTGAAGCTTGAAGCCTTTTTGCGCTGCCAATATTCT TATCCATCTATTGTACTCTTAGATCCAGTATAGTGATTCTTCTGCTCCAAGCTCATCCCACCTTGCA ACAAAAAAGTCTAATCTTCTGCAATAATTTCCATCCTTGGCATTGAGACATATTTGGTCAATCG GTTTTAATTTTGTCTTCCCTGCGAATTTCGGTACCATTGGGTTGAGCATTCGATGATA TGTCATGTACAATAATTAATATGAATTAGGAGAAAGACTTAGCTTCTTTTTCGGGTGATGTCACTTA AAAAC TCCGAGAATAATATATAAAGAGAATAAAATATTAGTTATTGAATAAGAAGTGAATACAGCT GGCGTTAGTCTGCTAATGGCAGCTTCATCTTGGTTTATTGTAGCATGAATCATATTTGCCTTTTTTTC CTGTAATTCATGATTCTTGCTTCTATACTATCCTCAATGCAAAACCTTGTGATCTTCCAGGTCGATA CTGACCAATTCTATGAAC TCTATCAACCTTTGCGATTCAACACTAGGTTCCACCATTGGGTCTAAAA TGAATACTTGCGAAGCTTCACAAAGATTCAAAGCAACACCCGCCCTTTAAACTGACCAAGAAAC CTCGCATTGAATGTTGTTTCATGAAATACTTGATGGTTTCATCTCTTTCGCTCGGTGACATACTACCCT GA
T_{LYS4}:: T_{CHS2} P_{GAL10} -PsoREPI-2-T_{ADH1}	ATATATCCTGGTAAGTACACTTATCAGGATGATGTGCCTAAAGAAAAAGTGGCGCAAGTTTGTATGG AAAATTATGATGCCGAGTTTCAAGCAAGGTTTCATCCAGGTGATATAGTGGTCAGTGGGTTCAATTT CGGTACCGGTTCTCCAGGGAACAAGCGGCCACCGCCTTATTGGCTAAAGGTATCAACTTAGTTGT

TTACAGGATCTTTTGGTAATATTTTTCAAGAACTCCATTAACAATGCTCTTCTGACCTTGAAATCCC
AGCATTAAATCAAAAAATTACGTGAGAAATATCAAGGTGCTCCAAAAGAACTTACAAGAAGAACTGGTT
GGTTTTTGAATGGGATGTAGCTGATGCTAAAGTGGTCGTTACCGAAGGTTCTTTGGACGGCCCTGT
GATCTTGGAGCAAAAAGTGGGTGAGCTAGGTAAAGAACCTACAGGAGATCATAGTAAAGGGTGGCCT
AGAGGGATGGGTCAAATCACAGCTTAATGCCAGTATTTTTCAGCTAATTTCTCGTCATTCCCTCTTT
TTTCTCATTCTGTTTTAGAAATGTACAGGTTATGACAACCATTACCTTCTAGGATATTTTGACATTTCTC
TCTTAGACTTTTTATATATAATATACATGTTTCAATACCAATATAATGTAGCATAACATAATCTATCTT
TTTACTCATAGACCTTTACTTTTCTCTCTAAAAGCCGTGTCGCAGTAACCTGGTGACGCTAAAGTTCTC
CGGCGTTTTTTTATGTATGAGCTGGGCATCAAGAGCTTTTGAGATTTTCCAAGTAGTAACTCATCTTT
CTGAGTGTGCTATCAAATACATACTAAGGAGAATAAACTCTTGTTATTACGTATTCTTCATCCTTATGG
GTAGAGAGCGCACTGTTTTAGTACATTTTCTAGACGTCGAAACGTAGAGCAATTGTGCGATAAAAAAA
AAAAAAGTAAGAAGATATATGAATAGGAGCTGTCGCTAGAACTAGTAATCTAGAGCTCGAGGGGTG
GATCCGGGGTTTTTTCTCCTTGACGTTAAAGTATAGAGGTATATTAACAATTTTTTGTGATACTTTTA
TTACATTTGAATAAGAAGTAATACAAACCGAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGTTT
TTGCATTTATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCAGAAAATAAGGC
TAAAAAATAATCGCATTATCATCTTATGTTGTTGTTGATTGCTTCAAGTTTGTGGGGTGGGCTG
CAGGTTACTGCCAATTTTCTCTCTCATAAACCAAAAAGCTAGTATTGTAGAACTTTTATTGTTCCGAG
CAGTGGCGCGCAGGACATCTGCGTTTCAGGAACGCGACCGGTGAAGACGAGGACGCGACGGAG
GAGAGTCTTCTTCGGAGGGCTGTCAACCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAAT
GAGCAGTTAAGCGTATTACTGAAAGTTCCAAAGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCT
TTACATTTCCACAACATATAAGTAAGATTAGATATGGATATGTATATGGATATGTATGTGGTAATG
CCATGTAATATGATTATTAACCTTCTTTGCGTCCATCCAAAAAAGTAAGAATTTTTGAAAATTCGA
ATTC AACCTCACTAAAGGGCGGCCGCAATGAGGAAGACATTTTTGAATAACTACTCATCATCACC
TGCATCATCCACAAAAACAGCGGTACTTTCTCATCAGCGGCAGCAGTCTGTGTCATTGCCAATTTCC
GGTCTCCTCATATTTTTCATGAATAAAAACGGCTTAATTCATGTAACCTTGGAAATGTGGTGAATA
ATACGGTCCGATTTTCAGTTTCCCAACAGGTAGCCATAGAACTCTCGTTGTGAGCAGTTGGGAGATG
GTAAAAGAGTGTTTTACTGGCAACAATGACACTGCTTCTCAAACCGTCTATCCCGTTAGCTTTTAA
GACTATATTCTATGCATGCGGTGGCATAGACTCATACGGTCTTTCGAGTGTACCTTATGGAAAATATT
GGAGGGAGCTTCAAAAGGTCTGTGTGCATAACCTCCTGTCTAATCAACAACCTCAAGTTCAGACA
CTTGATAATTTCTCAAGTCGATACGTCTTTCAATAAGCTGTATGAGTTATGCAAAAACCTCTGAAGACA
ACCAGGGAACTATCCTACTACTACTACCGCAGCTGGCATGGTGAGAATCGATGATTGGCTCGCCG
AATTATCGTTCAATGTGATAGGAAGAATAGTCTGCGGATTCCAATCAGGCCCTAAGACAGGTGCTCC
AAGCAGGGTGGAACAATTTAAAGAAGCAATTAATGAAGCATCTTATTTTATGTCGACATCTCCAGTGT
CATGAATGTTCCAATGCTAGGGTGGATTGACCAATTGACAGGTCTTACGAGAATATGAAGCACTG
CGGAAAGAAATTAGACTTGGTGGTCGAGAGCATAATTAATGATCATCGTCAAAAGAGACGATTCTCT
AGAATCAAAGGAGGAGATGAGAAGGACGATGAACAAGATGACTTCATCGACATTTGTTTGTCAATAA
TGGAGCAACCACAGCTTCTGGCAACAATAATCCTTCTCAGATACCTATCAAATCTATTGTCTCTGGAC
ATGATAGGTGGGGCACTGACACCACAAAACCTGACCACCATCTGGACCTTTCTGCTGCTGTGGAAC
AACCCCATGTGTTGGACAAGGC AAAACAAGAAGTGATGCACACTTTGAAACCAAAAGGAGATCA
ACAAATGATGCAGCAGCAGCCGTGGTGGATTTTGTATGATATTCGTAACCTTGTCTACATCCAGGCAA
TCATCAAAGAATCAATGCGGTTGTATCCAGCCAGCCCCGTGGTGGAGCGCATGAGCGGCGAAGATT
GTGTGGTGGTGGGTTTTCATGTACAGCAGGGACGAGATTATGGGCTAACGTTATGGAAGATGCAAC
GAGACCTAAAGTATGGGATGATCCATTGGTGTTCGACCAGACAGATTTTGTAGCGATGAACAGAA
GATGGTTGATGTAAGGGGTCAAATTATGAGCTGTTACCATTGGAGCCGGTTCGACGTGTATGTCCA
GGTGTATCCTTCTCTTTGGATCTAATGCAACTGGTACTGACTCGTCTTATTCTCGAGTTTGAAATGAA
GTCTCCTAGCGGAAAAGTGGACATGACAGCAACAGGATTAATGAGTTACAAGGTGATCCCCCT
TGACATTTCTGCTCACCCATCGTCGCATAAAGCCGTGTGTGCAGTCAGCAGCGTCTGAGAGAGACAT
GGAGAGTAGTGGTGTACCAATCACTCTGGGCTCGGGCAAGGTGATGCCTGTTCTGGGCATGG
GAACATTTGAGAAAGTTGGTAAAGGGTCCGAAAGAGAGAGGTTGGCGATTTTAAAGCGATAGAGG
TGGGTTACAGATACTTCGATACAGCTGCTGCATACGAACTGAAGAGGTTCTTGAGAAGCTATTGC
TGAAAGCACTTCAACTTGGCCTAGTCAAATCTCGAGATGAACCTTTTTCATCAGTTCCATGCTCTGGTGCA
CTGATGCTCACGCTGATCGTGTCTCTCGCTCTTCAGAAATTCGCTGAGGAATCTTAAATTGGAGTA
TGTGGATCTATATATGTTACCCTTCCCGGCAAGCTTGAAGCCTGGGAAGATAACGATGGACATACCA
GAGGAAGATATTTGTGCGATGGACTACAGGTCTGTATGGGCAGCCATGGAAGAGTGTCAAAACCTT
GGCTTCACTAAATCAATCGGTGTTAGCAATTTCTCCTGCAAAAAGCTTCAGGAATTGATGGCGACCG
CCAACATCCCTCCAGCTGTGAATCAAGTGGAGATGAGCCCGCTTTCCAACAAAAGAAGCTGAGAG
AGTATTGCAACGCAAATAATATATTAGTCAGTGCAATCTCTGTACTGGGATCAAACGGAACCCCATG
GGGCTCCAATGCAGTTTTGGGTTCTGAGGTGCTTAAGAAAATTGCTATGGCCAAAGGAAAATCTGTT
GCTCAGGTTAGTATGAGATGGGTTTACGAGCAAGGCGCAGTCTTGTGTTGTTGTTGTTGTTGTTGTTG
GAGAGATTGAGGGAAAACCTGAACATATTTGACTGGGAACCTACTAAGGAAGACCATGAAAAGATCG
GTGAGATTCCACAGTGCAGAATCTTGAGTGCTTATTTTTTGGTCTCACCTAATGGACCTTCAAATCT
CAAGAAGAGTTGTGGGATGATGAAGCTCTACTAGTATCGATGGATTACAAGGATGACGACGATAAGA
TCTGAAGATCTGAGCTCTTAATTAACAATTTCTGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTG
TCGGCTTGTCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGCTCAAATCGTTGGTAGATACGT
TGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAATAAGTTATAAAAAA
ATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTTAAACGAAAATTTCTTATTTCTTGAGTAACTCTT
TCCTGTAGGTGAGGTTGCTTTCTCAGGTATAGCATGAGGTGCGCTCCCTGCAGAAATTCGGATCCGGTA
CCATTGGTGTGCATAAATATCTAGAAAAATCGACCTTTTCTTTATTTATTTATTTATTTATTTA
CGGTTATTCACGTATATTTTGTAAATCTATTGATACACAAGTTCTTATTTGTAATTTTTCTACTTCTTTT
CATTAATTTCAAAGATGCTCTTGTGACGTCCAAACAGTGTCTGACCTAAAGCCCGTGGGAAAAA

	CAACGTAATTTCTACATAAAAAAGTCAAGTAGGCGCTCGAAAGAGTTTCAGAAAAATACCCAGCTGA CCATTAACTGAAGTAGCAAACATCTATGGATAAATATACTGCTTTGATTACGATGAGAATTTTTCAAC TCTTACATTGAATGTCTCGAGATACCCTAAAAGTTTGGCATATTGGGAGAAGTTACTGAACTATATAG TAAAGGCTTCGGCACCAATATGCAAGTCCACTGAGCCTCAATTATTAAGTTAATACGATGTACTTAT TCTTCTATGCTGAATGAGTTCCCTAT
$T_{MET17}::T_{TUB1}$ P_{GAL10}-PbrSalSyn- T_{ADE3} P_{GAL1}-PbrCPR1- T_{PDR5}	TTTGCCTCTTTCTTGCTACTACAAGGTGTTGAAACATTATCTTTGAGAGCTGAAAGACACGGTGA TGCATTGAAGTTAGCCAAATGGTTAGAACAAATCCCCATACGTATCTTGGGTTTCATACCCCTGGTTT CATCTCATTCTCATCATGAAAAATGCTAAGAAGTATCTATCTAACGGTTTCGGTGGTGTCTTATCTTT GGTGTAAAAGACTTACCAAAATGCCGACAAGGAACTGACCCATTCAAACCTTCTGGTGCTCAAGTTG TTGACAATTTAAAGCTTGCCCTCTAACTTGGCCAAATGTTGGTGATGCCAAGACCTTAGTCATTGCTCCA TACTTCACTACCCACAAACAATTAATGACAAAGAAAAGTTGGCATCTGGTGTTACTAAAGACTTGAT CCGTGTTTCTGTTGGTATCGAATTTATTGATGACATTATTGCAGACTTTTCAGCAATCATTGCAAAACAG TCTTCGACGGTCAAAGCCCTTAAGGTTCACTCTCGCCCCAACCTCCTTATCCTTTCTATATTTCAAG TATATTGTAAAAAATTTCCCTACTGTTATATCTCTTGGTGATATCTTATCTATAATCTACTTA TTATATTTCTCTAATATCTATGTACAAAAATAAAAAAACCATTCTCCTTCTCTAACCTGCAATTTG ATGAAAGCTTAGTAAGTCTCATCGTAGTTGATGGGCCAAGACATTTGACTTGTTCATCAGATTTCT TACATCACCCATTCCATGATTATTTACGGGATTTTATTAGTAATTGCATTTTGTATCCCTGATTTCT TTTCTCCCGCTTGCCGTGCAATTGAAAAAATGGAATGAGGACATCTTCCATTGATATCCAG GACAAAGCAAGAGTAGAAGTAACTATTAGCAAGGAGAATTCCAAAAATAGTGAACCTCAAGGTAAAG AGATCTATAGTTCAATATCTGATTTTCTGCTTTTCAAACCTTACTAGAGCTCGAGGGGTATATATT ATCCAAGAATTCTATCGTGAGTCCAATACCGTTTGTGACTTATCACACATTTCTAGCTTACTACTTT TTATATTATTCACCTCAGAGATGCAATGGAACGCTTTTCTTTGTGCTTATGTATATCTTCGTTATTT TATTTCTGTTCCGCCCAAATATTGTGCTAGCCAGTGTTGCTTGTCTTGGTATAGTATCATTTAGATTG CTAAAAGTTTCATACATTTATTGTTCTCATTAATGCGTCTCCCGCTTCGTTTGGTAATTAGGTTTCT CCTCCTTAACGGAATCGGAAATTTTCGACGTTAAGGAACTGGGGTTAAATTAAGAAAAAGGGTCG CGGAACGTTAAATCCCTGTGTAAAAAATATAAATCATTATTAGTCCCAATAGTGTTCTGATTATC GTTATTATTATTATACATCCATCTTTAAAAAATCTTTCGATTGTCTTTATTAATTTCTATATAATTAA GTTGTCTTACTGTCCAAAGGTTTGTACGATCGAGCATCTCGAGGAGTAATTTTGGCTTCAAGTGGG TTTTTCATGTAGAGAATGAAGCTAACAGCTTCTAAGATCAGGCAATTTCCCTCAGAAACACAATC CCATTTGAATTCGTTGACCATACTTGAAGAGCAAAACCTACCATCTGCTTGGCTAATTTCTACACCTC CACAAATTCATACACGACCAAAATGGTATCAATGAACCTTCCATTGTGTTGATATCACCAAAACGT CCATCACTATTAACATCCTTTAAGAATCTCTCTGGTATGAATTTATACGGTGCAGGGGAAAACGTTAGG GTCGTGATGAATAGCATACAGATTACCATTTTGTACCTTTAACGATTCTCCTCCCTTGAATG TAGTATCTTTTACTGCTACATGAGGGACCGTAGTGGTGCGATGGGTTTCTGCTCAATTTCTTCTTTC ATCACGGCTTGTAGATATGGCAGCTCTTCAAATCTGCAACTTTCACCTGGTCTTGGCCACCCGTCA GGTTGTTTATTTTCGCGGTAAAGTTTCTTGAATTTCTGCTTGGCGACTAAAAAGGAGAGTGCCCAA ACTGCCGTTGCTGCAGTACTATCTACTCCCAAACCAAAATTTCAAGTACCATTGAGTAAATAATTTCT ATCACTGTAATCTTGAGATGTAAGAAAAATGTAAGAAAAAATTTGCAGGAAGATTAGAGACGATATGTG GTTGAATCAAATTATATACTCTGTCTCTTGTGGCGTACATATCCCGAATATTTTCTCTGGCTTGGTA AATATTTAGCAATTTTGAAGCATCAGCTAACTTGCCAACCCACTTATGCGTACCATCTCATCGAGG GCGTCTTTCACACCAATAACGAGATCCTCATTTGAAAAGTCTTGACCAATAAATCTACTCAGTAA TCGAATAGACGCTTCTTGGCATGATCAAGAGGTTTAAAAATTCATTCTGCTCTGACGCTTCTTCTT GCATGGATTTGATGAGATTTTGCATGTCTTTTCTGTAACGAGATAGAGACATAACATTTGATGGA TTTATACAACCTTTGAAGACCTTTCTTTAAATGATGCCAAAATGGACCAGAATCACAAATCAGCAATATTC TTCCCATTTGCTGTGATGATTTTAAAGATATCAGGTACATCTCTAGCGGAATAATCAGCAGATTTATTA ACAAGAACTTCCCGAGCGTTGTGATATCGTAAATGACGATCATTGGAACCCAGCTTCCAATCCAAA TTGTGAATGCACCTCCAATAAATTTGGATAAGTTTGCCAAAGCAACCTGTGAAGGGAACACCTCCCTCC CAACTGATGAAGGTTTCAATTATGGGTAGTGTGTTTGGACCTTCTGGCCATTCCGTCGTATTTGCAG ATTGATAAAAAGAAAACATGAACCTTACCAATGCAAAATCTATTATTACAGTGATGCTATCATCCAAA AATCATTCTCCTCTATATTAATCGGAGCCATTATAGTTTTTCTCCTTGACGTTAAAGTATAGAGGTAT ATTAACAATTTTTTTGTTGATACTTTTATGACATTTGACATTTGAATAAGAAGTAATAAATGTTGAAA GTATTAGTTAAAGTGGTTATGCAGCTTTTCCATTTATATATCTGTTAATAGATCAAAAAATCATCGCTTC GCTGATTAATTACCCAGAAATAAGGCTAAAAAATCAATCGCATTATCATCCTATGTTTGTAAATTTGA TTCGTTAATTTGAAGGTTTGTGGGGCCAGGTTACTGCCAATTTTCTCTTTCATAACCATAAAAGCTA GTATTGTAGAATCTTTATTGTTTCGGAGCAGTGCGGCGGAGGCACATCTGCTTTTCAGGAACGCGA CCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCGGTCCGAGGGCTGTGCGCCGCTCGGCGGC TTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAGTTTCAAAGAGAAGGTT TTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGGATATGT ATATGGTGGTAATGCCATGTAATATGATTATTAACCTTTTTCGCTCCATCCAAAAAATGAAGAAT TTTTGAAAATTCAAATATAAATGTTGATCATGGGTTCTAACAACTTGGCCAACTATCAGTAATCCGATGT AGGTATTTCTATTGGTTCCGAATACATTTCTGATCCAATTTTCTTATGGTTACCACTGTTGCTTCTAT GTTGATTGGTTTCGGTTTCTTCGTCTACATGAAATCCTCCTCTTCTCAATCTAAGCCAATTGAAACTTA CAAACCAATCATCGATAAAGAAGAAGAAGAAATGAAGTTGATCCAGGTAAAAATCAAGTTGACCATTT TCTTCGGTACTCAAACCGGTACTGCTGAAGGTTTTCGCTAAGGCTTTGGCTGAAGAAATTAAGGCTAA ATACAAGAAGGCTGTCGTCAAGGTTGTTGACTTAGATGACTACGCCGCTGAAGATGACCAATACGAA GAAAAGTTAAAGAAGGAATCTTTGGTCTTCTTCATGGTTGCTACTTACGGTGACGGTGAACCAACCG ATAACGCTGCTAGATTTTATAAGTGGTTACCCCAAGAACACGAAAGAGGTGACTGGTTGCAACAATT GACTTACGGTGTTTTCGGTTTATAGTAATCGTCAATACGAGCATTTCAACAAGATGCGCGCTGACGCTT GATGAACAATTTGGTAAGCAAGGTGCTAAAAAATCGTCCAAGTCGGTTTGGGTGATGACGACCAA TGCATCGAAGACGACTTCACTGCTTGGAGAGAGTTATTGTGGACTGAATTAGACCAATTGTTAAAGG

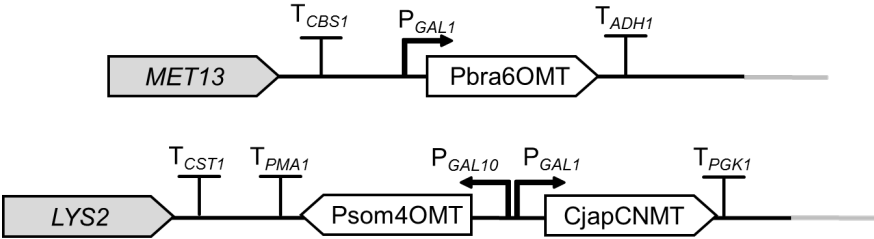
	ATGAGGATGCTGCCCCATCTGTGCTACTCCATACATCGCTACTGTCCCTGAGTACCGTGTGCTTAT TCACGACACTACTGTCGCCCTCCTTGGAAGAAAAGCACATCAACACTGTCAACGGTGATGTTGCTTTC GACATTTTGCACCCATGTCGTGCTATTGTGCTCAACAAAAGAAATTGCATAAGCCTAAATCTGATC GTTCTTGCACTTCACTTGGAATTCGACATTTTGGGTTCTTCTTTGACTTATGAAACTGGTGATCAGTT GGTGTCTATGCCGAAAACGTGACGAAACCGTTGAAGAGGCCGGTAAATTGTTGGGTCAACCATTG GATTTATTGTTCTCTATCCATACTGACAAGGAGGATGGTTCTCCACAAGGTTCTTCCCTTGGCACCACC TTTTCCAGGTCCATGTACTTTGAGATCTGCCTTGGCTAGATACGCCGACTTGTGAAACCCACCTAGA AAGGCTTCTTTGATTGCCTTGTCCGCTCACGCTTCCGTTCCATCTGAAGCCGAAAGATTGCGTTTTT GTCTTCTCCATTGGGTAAGAACGAATACTCTAAGTGGGTTGTGCGTTCCCAAAATATCTTTGTTGAAA TCATGGCCGAATTCCCATCCGCCAAGCCACCATTGGGTGTTTTCTTTGCTGCTGTTGCTCCAAGATT GCCACCAAGATACTACTCCATCTCCTCTTCCCCAAAGTTCGCTCCATCTAGAATCCATGTTACTTGTG CTTTAGTCTATGGTGAATCTCCAACCGGTAGAGTTCATAGAGGTGTCTGTTCTACTTGGATGAAACA CGCCGTTCCACAAGACTCTTGGGCTCCAATTTTCTGTCGTACCTCTAAGTTCAAGTTCAGCTGAC CCATCTACTCCAATCATCATGGTCCGTCCAGGTACCGGTTTGGCTCCATTTAGAGGTTTCTTACAAG AAAGAATGGCCTTGAAAGAAAACGGTGCCCAATTGGGTCCAGCTGTTTTATTCTTCGGTTGTAGAAA CAGAGACATGGACTTCATTACGAAGACGAATTGAACAAGTTCGTTGAACAAGGTTATTCATCCGAA TTGGTCATCGCCTTCTCCAGAGAAGGTGAAAAGGAATACGTCCAACACAAGATGATGGAAAAG GCCACCGATGTTTGAACGTTATTTCTGGTGACGGTTTCTGTATGTTTGGTGATGCTAAGGGTA TGGCTAGAGATGTCCACCGTACTTTACATACCATTGCTCAAGAACAAGGTCCAATGGAATCTTCTGC TGCCGAAGCCGCTGTCAAGAAATTGCAAGTTGAAGGTAGATACTTAAGAGATGCTGGTAATAGAAT TTTGAATTTGGTTAAGAAAAGAAACTTACCAAGATGGACTTTTTTAATAAACATACATCACTACAT ATAGGTGCGTAATAATAAGTTTTTATTTTTTTTTCTTAATTCAGCGAGCTTTCTACATTTCAATTTCT TGAATTTACCAGCTCTGATAAATCAAAGTTCAATGTCCGAAAGAATTCCGAGAATTTTATTCCTAGGT TATCTCATCGGTACTTTTTTTTATAGGCATCAGCATTTAGTGTGAAAACATATTTCTTAACTGTGCC TTACGGTTAAAAAGAAAGGACGCAATGGATATAGTTAAATATAAGATATATTATCAATCACTGTCC GAATGGCGAACCGTCTTCTCATATACGGCCTTATTTTGTGGGTAAGTATAATCGGATCTTTTGCATTA GATAGGAACAAAACGTGCTCAAAATGCGAAGATTGGTCTACATGATACCAGTGTGATCAGCACTGGAA GTACAACGAACGTGCCTGCAGAATTCGGATCCGGTACCATGGCTCAGCCAAACGTGATCTAAATTA GTATAAACCTACTTTCTCTCACAAGTACTATCTTTTATAAACGAACTTTATTGAAATCAATCCTT TTTTTCCCTTGTTACATGTCGTGACTCGTACTTTGAACCTAAATTGTTCTAACATCAAAGAACAGTGTT AATTCGCAGTCGAGAAGAAAAATATGGTGAACAAGACTCATCTACTTCATGAGACTACTTTACGCCCTC CTATAAAGCTGTCACACTGGATAAATTTATTGTAGGACCAAGTTACAAAAGAGGATGATGGAGGTTTC TTTACAATAAAGAAGCACATGTGTTAACGTTTTTAGTATTGCTTGTATGTAATCAGGAAAACCT CGCGGGATTGGTTGGATGCTACTTTCCATACATAAATATTATAGATCTAAAAAGCCAAATACAAAG TAAAGATTAGTAAAGCTGTTGGAATTCATCGTTGATAAAAATGTTAGTTATTAAATATAAAA
T_{MET16}:: T_{CHS2} P_{GAL10}-P_{somSalR}- T_{HPC2} P_{TEA1}-P_{somSalAT}- T_{ROX3}	TACCAGCCTAAAAATCAAACCATTCACGTATATAAGCCGGATGGATGTGAATCGGAGGCAGATTTTG CCTCGAAATACGGGGATTCTTATGGGAGAAAGATGATGACAAGTACGATTATCTGGCCAAAGTGGA ACCTGCACATCGTGCCTACAAAGAGCTACATATAAGTGCTGTGTTTACTGGTGAAGAAAATCACAA GGTTCTGCCCGCTCCCAACTGTCGATTATTGAAATAGACGAACCTTAATGGAATCTTAAAAATAAATCC ATTGATCAATTGGACGTTTCGAGCAGGTTAAACAGTATATAGATGCAAACATGTACCATACAACGAAC TTTTGGACCTTGGATATAGATCCATTGGTGATTACCATTCCACACAACCCGTCAGGAAGGTGAAGA TGAGAGAGCAGGAAGATGGAAGGGCAAGGCTAAAGCTGAATGTGGTATCATAGAGCTTCTAGATT CGCTCAATTCCTAAAGCAAGATGCTTAATGCCAGTATTTTTCAGCTAATTTCTCGTCATTCCCTCTTTT TTCTCATTCTGTTTTAGAAATGTACAGGTTATGACAACCATTACCTTCTAGGATATTTTGACATTTCTCT CTTAGACTTTTTATATATAATATACATGTTTCAATACCAAATAAATGTAGCATAACATAATCTATCTTT TTACTCATAGACCTTTACTTTCTCTCTAAAAGCCGTGTGCGAGTAACTGGTGACGCTAAAGTTCTCC GGCGTTTTTTTATGTATGAGCTGGGCATCAAGAGCTTTTGAGATTTTCCAAGTATGACCTCATCTTTCT TGAGTGTGCTATCAAATACATACTAAGGAGAATAAACTCTTGTATTACGTATTCTTCATCCTTATGGG TAGAGAGCGCACTGTTTTAGTACATTTTCTAGACGTCGAAACGTAGAGCAATTGTGATAAAACAAAA AAAAAGTAAGAAGATATATGAATAGGAGCTGTGCTAGAACTAGTAATCTAGAGCTCGAGGATCGCT TCGCTGATTAATTACCCGAGAAATAAGGCTAAAAAACTAATCGCATTATCATCCTTATGTTGTTAATTT GATTGCTTAATTTGAAGGTTTGTGGGGCCAGGTTACTGCCAATTTTTCTCTTTCATAACCATAAAAGC TAGTATTGTAGAATCTTTATTGTTTCGAGCAGTGCGGCGCGAGGCACATCTGCGTTTCAGGAACGC GACCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCCGTGCGAGGGCTGTGCGCCCGCTCGGCG GCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAAGTTCCAAAGAGAAGG TTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGGATAT GTATATGGTGGTAATGCCATGTAATAGATTATTAACCTTCTTTCGCTCCATCCAAAAAAAAGTAAG AATTTTTGAAAATTCAATATAAATGCCTGAAACTTGTCCAAACACTGTGACTAAGCGCAGATGCGCTG TGGTGACTGGAGGCAACAAGGGTATAGGTTTCGAGATCTGCAACACAGCTCAAGTAACGGCATCA TGGTTGTTTTGACATGTAGGGACGTAACTAAAGGTCACGAAGCTGTTGAAAAGCTAAAAACAGTAA CCACGAAAACGTAGTGTTCACCAAGTTGGACGTAACGGACCCCATAGCGACTATGTCATCCTTAGCA GATTTTCATCAAGACTCATTTTCGGCAAGTTAGACATTTCTGGTCAACAACGCTGGAGTGGCAGGCTTCA GTGTGGACGCAGATAGATTCAAAGCGATGATTTCCGACATCGGGGAAGATTGAGAAGAACTGGTTA AAATATATGAAAAGCCTGAGGCTCAGGAACTAATGTCCGAAACTTACGAGTTGGCGGAAGAGTGCT AAAGATAAACTACAACGGTGTAAGTCTGTCACTGAGGTCTAATACCACTGCTACAGCTGTCTGAT TCGCCCAGGATCGTTAATGTTTCTAGCTCTACAGGGTCACTAAAGTACGTTTCTAACGAAACAGCAC TTGAAATACTAGGCGATGGCGATGCTCTCACTGAAGAAAGGATAGATATGGTAGTGAATATGCTACT AAAGGACTTCAAAGAAAATCTAATTTGAGACTAAGCGCTGGCCTTCTTTTCGGTTCAGCTTACACCACT TCTAAGGCTTGTCTCAATGCCTACACCCGCTCCTGGCAAACAAGATTCCAAAGTTCCAGGTAAACT GTGTTTGTCTGGCTTGGTAAGACGGAGATGAACTACGGAATCGGAACTACACTGCTGAAGAAG

	GTGCAGAACATGTGGTACGCATCGCACTATTCCCCGATGACGGGCCATCTGGCTTTTTCTACGATTG CTCTGAACTATCTGCTTTCTAAATCGATGCAATTGTGTGAGAGATTTTCCAAAATCCCTGAACGACATA AAAAATGAAATAAAGTGTGCCTATATCTGAAAGTATATAGTCAGTTTATATGAAGATTTTTCCGAGCG AATGCTTTTTCTATCCATGTATCTATGTCCGATAATGTGGTGAAGAAAAATAGATAATTTTCACTTTAG AGAATCTATAAACTTGGCTATAAAACCAGAAACGAGTATGGCATGTTCAATATCTTGGTGAAATGAA ATAGATAAGAACCTCATTTTATACCAGTAAAAATAGTCATTTAGAGATTCGTTTTATCCACAATTTCCCTT ATACTTCGGAAGTTTATCAGGTTTCGACTGACTTCAAATTAGAATCAATCTTCGAACAAACAATCTTCA GTAGATCTTCATCCCATGTTTTTTTTTAGCACAATGAAAAATTTCAAGTAAGTTAGCAGTAATAATATAT CACTCTTCTTGGCATCTGCCGTAATGAATCTTGGATAGTCTGCAGACCTCGAGCGTATAGTTGGTC CCTTGGGCATGTTAGAAGGAACACTTTACGCGTCCGCTCCTCTGCATGACGAATTATAACTACTCGC AGGAAGAGTTTCATTACGCGACGGTTCCAAGAGCGTGAAAGAATTGCTTCTGCGCATAAGTCTTCTTC CATAATAGATTTTTATATCATTTTTAGAACATAAAATTTCTCGCCAAGAGTTTTTTAAGTAAAAAGCAGAT TATCCAGTAACATGGGTTCCGCTTTTTGAGCGAATACTATGAACGAGTTTTGCCGCTGCTCCCCATG ACTCATTAGATTCTTTGATGACAAAAGATCACTGGCAGAAGTCTGTTTCGCTTCTTTTCTTTTTTTTTT TTTTTTTTTTTTTTTCGAGCAGATTGCAATCAGCGATGAGACAAAAATGGAAGTAGGTATGGGCTG ACGATTGCCTAAACGGTTATCAATTGGTTACTATAAACGCCCTCTATTATATGAATACCTATTATTG AATCCCTTCATTAATATTTAAGAAGACAGAGGCTATATATCATCTGGATCAAAAACTGCAGGGCTAG CAAAAAATGGCGACTATGTACAGTGCTGCTGTAGAAGTAATCTCTAAGGAACTATAAAACCAACTA CTCCAACACCAAGTCAACTGAAGAACCTCAATCTATCTTTGCTGGACCAAGTGTTCCTCTATATTAC TACGTGCCCATCATCCTGTTCTACCCAGCAACAGCAGCAAACTCTACTGGCAGTTCTAACCATCACG ACGATTTTGATCTGCTAAAAAGTAGTTTGTCAAAGACTCTGGTACACTTCTACCTTATGGCAGGCAG GATGATAGATAACATCTTAGTTGATTGTACGACCAAGGAATCAACTTCTACAAGGTGAAGATAAGG GGTAAGATGTGCGAATTTATGTCACAACCAGACGTGCCTCTGTCTCAATTACTGCCCAAGTGAAGTGG TTTCAGCTTCTGTCCCAAAGGAAGCACTGGTAATCGTCCAAGTAAATATGTTGACTGTGGCGGCAC TGCTATCTGTAGTTCTGTCTCGCACAAAGATCGAGATCGGGCAACTATGTTACATTATTAGTCTG TGGGCTAGTACTACTAAAACTTCTAGGTCTGGAGGATCGACTGCGGCAGTGACAGACCAAAAGTTG ATCCCGTCTTTTCGACTCAGCTTCGCTGTTCCACCATCTGAAAGGCTAACTTCTCCAAGTGGCATGT CTGAAATACCGTTCAGTAGCACTCCAGAGGACACAGAAGATGATAAGACCGTGTCAAAGCGGTTCCG TCTTTGACTTTGCAAAGATAACCTCTGTACAGAAAAAGCTACAAGTTCTGATGCACGCAATTACAAG AGTAGGAGACAGACTAGAGTGGAAGTTGTAACAAGTCTAATATGGAAGTCTGTATGAAAAGTACAC CAGCTGGTTTCTACCAGTTGTACATCAGCTGTCAACCTAAGAAAGAAGATGGACCCACCACTTCA GGACGTTAGTTTGGAACTTGTCTGTGACTGTGTACGCGTTCCTTCCCGCAACTACTACAACCTACG ACTAACGCTGTGAACAAGACGATTAACCTACTTCTCGGAATCTCAAGTGGTCTTACACGAGTTGC ACGATTTTCATCGCGCAATGAGGTGAGAAATAGACAAGGTTAAGGGCGATAAAGGCTCTCTGGAAA AAGTTATACAGAACTTCGCTTCTGGTCATGACGCAAGTATTAAGAAGATTAACGACGTGGAAGTAATT AATTTCTGGATTCTTCGTGGTGTGCTATGGGTCTGTACGAAATAGACTTCGGCTGGGGCAAACCAA TCTGGGTAAGTGTGGACCCAAATCAAGCCAAACAAGAACTGCTTCTTCATGAACGACACTAAGTG TGGTGAAGGCATAGAAGTGTGGGCTCTTTCCCTGGAAGATGATATGGCAAAATTTGAACCTACACTTG TCTGAAATACTAGAATTGATCTAAATCGATCTTTACGCCATCGCATATCAGCCCCACTATAAATAAAG AGCAAGTGTTTCATGGAAGTGGCACGTTTGTGCTTCGGCGGTAAGTACATGCAGTATATATGTTAAC TTTTTTGAATATAGAATATTAATTAACACGGAATAGGCTATTTTGAGACAGGCATGCTTTTTCAAACA TATCAAGCCATTTAAAAATTTTGAACCAAGTAGGCTCAACGAGAATACAATTTGGCCTTGCCAAATTTT TTTGCCATCTCCTGCTTATGGTTGAAACCCCAAATATGTGGTAAGTTCCAACCTCTAGTTCCATTTCCG AGAAGAACTTCCAACGAAACTGTTCCAAGTGCCCTCAATGATACACTTTCTGCCTTTAATCTCTT TGACGTTTGCTTTTAGTATAATCGTAGTATCTACAGGAATGTCCTCAAAAACTCTAGCGACAGCCTT GCAGTTACACCTCTTTTACTAGGCAACAAAGGAAGCCACAATAGCACAATGCTCCTGCAGAATTCCG GATCCGGTACCATGGATAGATAGATAGATACGATATTAACCATATGTATGTGACTAATTTATTTCCTTA ATAACACCAATGATTACAACCTTTCTAAAGCTGGCGGAGAATTGCGCTGTACGAGAAAAGAGCGAAA ACAGAGGAATATTCAACTAAGAACCAAACCTGCGATAAAGAGGATTGAAAGGAAAAACGAAAGAAAA GGTAACTGACAAATATATACATTAACCGATGGGTAATTTGAGTTTCTATAAAAAACCAAGCTACCA CCAGGGTTTATCAATGCTCGCATACTTAGGGATAACCTTCAAAAGACAACAATTTAAAGCAATGAAAT CCTTGTTAAATCTTTGAAATTCATCGCTAGAAATATGAACCTTCCAACAAAACCTGAGGTTGGAGGCTC AGTTAAACTAAATGCATTGCCAACTACATGAGATCTACGCAGATCAAAAAATAGGTGCGTGGACTC TGGACATGCAAGATTTGTTCTAAGTGATTTAGGTTGTGTA
T_{NCL1}::T_{STE5} P_{GAL1}-PsomBetv1- 1-T_{ADH1}	ATTATTTATTCTGGTGTAATTTGTTTGTCTCTCAAAGAAGTGATATCGAATGTTTCATGGAGAATCCAA AGTGAATCATTGCCAATAATGAAACACCATATGAAATCTAATAGAATTGTTGAAGCTAATTTAGAGAT GTTAAACACTTGTTAATCGAATCTTTCCCTAACTTTGACGACATTCGTTTGAAGAATCGATAATG ATTTTGTGAAAAGATGACAAAATTAAGCTCTGGTTGCGCCTTTATTGATGTGTCAAGAAATGACCCCT GCCAAAGAAAACCTTATTCTTGCTGTGTGAAAGGCAACAAGTGATCAATTTGATGGTTTGAAGA AGATACTCATGAGCTATTATATAGGATCTTTGGTATTGATGCGAATGCCAAGGCTACTCCAAGCGCT GAAGAAAAAGAAAAAGAAAAAGAAACGACTGAATCTCCCGCAGAACTACTACAGGAACATCAACAG AAGCTCATAGCGCTGCTAATTGAAGTATACACTAAATTTTATGCAATAATAAAAAAGAAAGCATCCCGC CAAACGTTTCGTAACATATTTGTTACATAGTTTGATTCCGCTGAATTTGAAGTGGACGCAAGTTCTTC TTAGCTTTCAAGTTTCATAGTAAATAGACATTACCCAAAGATCTCTTTCCGTCTTGTGCTTGCATGA AAACCATTGACTTGCCGCTCACACCTAACTTTTTAGTGAAATGGATCTTATCTGTGTTACTTGAACA TTAGAGTTATTCAAGTCAGCTTCCGGAGCAGAAGGTGGGTCAATATTATTATCTTGTCTACCATAATC TTGCGTATATTGCTTTTCATTCCCTTTCTTCTGTAAGCGATTGCTAGAAAGGGCAGCCCTTGTCCAA ACCAAAGCGTAAAGCATCGCGAATCTCGTCTTCTACAGATCGCCATCCCTCCCATAAACTCGTGG GAGAGCGTGGGATCCATAGTTTATTTCTGTAAGTACGATCTCTCTTCTAGAGCTCGAGGTTTGCG GCCGCCCTTAGTGAGGGTTGAATTCGAATTTTCAAAAAATCTTACTTTTTTTTTTGATGGACGCAA

	GAAGTTTAATAATCATATTACATGGCATTACCACCATATACATATCCATATACATATCCATATCTAATC TTACTTATATGTTGTGGAAATGTAAAGAGCCCCATTATCTTAGCCTAAAAAACCTTCTCTTTGGAAC TTCAGTAATACGCTTAAGTCTCATTGCTATATTGAAGTACGGATTAGAAGCCGCCGAGCGGGTGAC AGCCCTCCGAAGGAAGACTCTCCTCCGTGCGTCTCGTTCACCGGTGCGGTTCTGAAACGCAG ATGTGCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATACTAGCTTTTATGGTTATGAAG AGGAAAAATTGGCAGTAACCTGGCCCCACAAACCTTCAAATGAACGAATCAAATTAACAACCATAGG ATGATAATGCGATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATC TATTAACAGATATATAAATGCAAAAACGTCATAACCACTTTAACTAATACTTTCAACATTTTCGGTTTG TATTACTTCTTATTCAAATGTAATAAAAGTATCAACAAAAAATTGTTAATATACCTCTATACTTTAACGT CAAGGAGAAAAAACCCCGGATCCACCAATGGCACCATTAGGTGTTTCTGGTTTGGTTGGTAAATTGT CAACTGAATTGGAAGTTGATTGTGATGCTGAAAAGTACTACAACATGTACAACATGGTGAAGATGTT AAGAAAGCAGTTCCACATTTGTGTGTTGATGTTAAGATCATCTCTGGTGACCAACTTCTTCAGGTTG TATCAAGGAATGGAACGTTAACATCGATGGTAAACAATCAGATCAGTTGAAGAACTACACATGAT GATGAACTAAGACATTGAGACATAGAGTTTTCGAAGGTGACGTTATGAAGGATTTCAAGAAATTCG ATACTATTATGGTTGTTAATCCAAAACAGATGGTAATGGTTGTGTTGTACAGATCTATCGAATAC GAAAAGACTAACGAAACTCACCAACACCATTCGATTACTTACAATTTGGTCATCAAGCTATTGAAGA TATGAATAAGTATTTGAGAGATTCTGAATCAAATTAAGCTCTTAATTAACAATTCTTCGCCAGAGGTTT GGTCAAGTCTCCAATCAAGGTTGTCGGCTTGCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAG GGTCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTT ATTATTAATAAGTTATAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAACGAA AATTCCTTATTCTTGAGTAACCTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTCGCT CCCTGCAGAATTCGGATCCGGTACCATGGTAGTTATCATGTTTCTTATCCCCTCCACTGTAAAGTAAA TGTATATTATTAAGGCAAAAATAAGAATATAATATTCAGTAATGAGCCTTTTTTATTCTGGTCTTGT TACTTCTTTTTTACTTCATTTTTTTTTTTTTTTTGGATAGTAAATAATGCACCTTTGACTAGATCTTGAG TATTTAGTTGAATTGAAATTAAGATACAGTGTATAAAAAATTATACAAAAAATAAGGAAACTTGATAT TGAAAATAGGATTGCCAATACATGGATTGAGACAGAGAATTTTTATCTTCATATCCAGATATTCGTA GGAATAACAAAGTTTTAGTGACCCAAGGTATAAATTGCGAAAGACCTGCGGAGTTGGCGGCGAACA CTGACTTTTTGGGCATCAACAAAGGAATCAACTACAACCTTTGATAGCTCTATCCAAATCATAAGAAGA CACAAATTCAGATAATCTCATTTTAGCAAAAGATTCTGCAATTCCTAGA
--	--

Blue: Homologous recombination sites
Green: Promoters
Orange: CDS
Red: Terminators

Example schematic of integration cassettes



Supplementary Note 2. Nucleotide sequences of pEV-1 and pTHS used to transform engineered strains of *Saccharomyces cerevisiae* in fermentation study.

>pEV-1 (circular vector sequence)

```
TTTCACACCGCATAGATCCGTCGAGTTCAAGAGAAAAAAGAAAAAGCAAAAAGAAAAAGGAAAGCGC
GCCTCGTTTCAGAATGACACGTATAGAATGATGCATTACCTTGTCTATCTTCAGTATCATACTGTTTCGTATAC
ATACTTACTGACATTCATAGGTATACATATATACACATGTATATATATCGTATGCTGCAGCTTTAAATAAT
CGGTGTCACTACATAAGAACACCTTTGGTGGAGGGAACATCGTTGGTACCATTGGGCGAGGTGGCTTCTCT
TATGGCAACCGCAAGAGCCTTGAACGCACTCTCACTACGGTGATGATCATTCTTGCCTCGCAGACAATCAA
CGTGGAGGGTAATTCTGCTAGCCTCTGCAAAGCTTTCAAGAAAATGCGGGATCATCTCGCAAGAGAGATCT
CCTACTTTTCTCCCTTTGCAAACCAAGTTCGACAACCTGCGTACGGCCTGTTTCGAAAGATCTACCACCGCTCT
GGAAAGTGCCTCATCCAAAGGCGCAAATCCTGATCCAAACCTTTTTACTCCACGCACGGCCCCCTAGGGCCT
CTTTAAAAGCTTGACCGAGAGCAATCCCGCAGTCTTTCAGTGGTGTGATGGTCGTCTATGTGTAAGTCACCA
ATGCACCTCAACGATTAGCAGACCAGCCGGAATGCTTGGCCAGAGCATGTATCATATGGTCCAGAAACCCCTAT
ACCTGTGTGGACGTTAATCACTTGCATTGTGTGGCCTGTTCTGCTACTGCTTCTGCCTCTTTTTTCTGGGA
AGATCGAGTGCTCTATCGCTAGGGGACCACCCTTTAAAGAGATCGCAATCTGAATCTTGGTTTTCATTTGTA
ATACGCTTTTACTAGGGCTTTTCTGCTCTGTCTATCTTTGCCTTCGTTTATCTTGCCTGCTCATTTTTTTAGTAT
ATTCTTTCGAAGAAATCACATTACTTTTATATAATGTATAATTCATTATGTGATAATGCCAATCGCTAAGAAA
AAAAAAGAGTCATCCGCTAGGTGGAAAAAAAAAATGAAAATCATTACCGAGGCATAAAAAAATATAGAGT
GTACTAGAGGAGGCCAAGAGTAATAGAAAAAGAAAATTGCGGGAAGGACTGTGTTATGACTTCCCTGACT
AATGCCGTGTTCAAACGATACCTGGCAGTGACTCCTAGCGCTCACCAAGCTCTTAAAACGGGAATTTATGG
TGCACCTCTCAGTACACGCGCCAGATCTGTTTAGCTTGCCTCGTCCCCGCCGGGTACCCGGCCAGCGACAT
GGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTGCGCCCGT
ACATTTAGCCCATACATCCCCATGTATAATCATTTCGATCCATACATTTTGATGGCCGCACGGCGCGAAGC
AAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCC
CCACGCCGCGCCCCCTGTAGAGAAATATAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTTCATATACTTC
CTTTTAAAATCTTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAAGA
CTCACGTTTTCGAGGCCGCGATTAAATTCGAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGAT
AATGTGCGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTTCTGAA
ACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTTA
TGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCC
GGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTACAGGTGAAAATATTGTTGATGCGCTGGCAGT
GTTCTGCGCCGGTTGCATTTCGATTCCTGTTTGTAAATTGTCCTTTTAAACAGCGATCGCGTATTTTCGTCTCG
CTCAGGCGCAATCACGAATGAATAACGGTTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGG
CCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGG
TGATTTCTCACTTGATAACCTTATTTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCG
GAATCGCAGACCGATAACCAGGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTTCTCCTTCATTACAG
AAACGGCTTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATGCAGTTTCATTGTGATGCTCGA
TGAGTTTTTCTAATCAGTACTGACAATAAAAAAGTATCTTGTTTTTCAAGAACTTGTCATTTGTATAGTTTTT
TTATTTTGTAGTTGTTTCTATTTTTAATCAAATGATTGCGTGATTTATATTTTTTTTTTCGCTTCGACATCATCT
GCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATA
CTGCTGTGCGATTGATACTAACGCCGCCATCCAGTGTCGAATTCGCCATTACAGGCTGCGCAACTGTTGGGA
AGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGAATTGGAGCGACCTCATGCTATACCTGAGAA
AGCAACCTGACCTACAGGAAAGAGTTACTCAAGAATAAGAATTTTCGTTTTTAAACCTAAGAGTCACTTTTA
AAATTTGTATACACTTATTTTTTTTTTATAACTTATTTAATAATAAAAATCATAAATCATAAGAAATTCGCTT
ATTTAGAAGTGTCAACAACGTATCTACCAACGATTTGACCCTTTTCCATCTTTTCGTAAATTTCTGGCAAG
GTAGACAAGCCGACAACCTTGATTGGAGACTTGACCAACCTCTGGCGAAGAATTGTTAATTAAGAGCTCA
GATCTTTTTCGCGCCGCCCTTTAGTGAGGGTTGAATTCGAATTTTCAAAAATTTCTTACTTTTTTTTTTGATG
GACGCAAAGAAGTTTAATAATCATATTACATGGCATTACCACCATATACATATCCATATACATATCCATAT
CTAATCTTACTTATATGTTGTGGAATGTAAAGAGCCCCATTATCTTAGCCTAAAAAACCTTCTCTTTGG
AACTTTCAGTAATACGCTTAACTGCTCATTGCTATATGAAGTACGGATTAGAAGCCGCCGAGCGGGTGAC
AGCCCTCCGAAGGAAGACTCTCCTCCGTGCGTCTCGTCTTACCAGGTGCGGTTCTGAAACGCAGATGTG
CCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAAT
TGGCAGTAACCTGGCCCCACAAACCTTCAAATGAACGAATCAAATTAACAACCATAGGATGATAATGCGAT
TAGTTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTTGATCTATTAACAGATATAT
AAATGCAAAAACCTGCATAACCACTTTAACTAATACTTTCAACATTTTTCGGTTTGTATTACTTCTTATTCAA
ATGTAATAAAAGTATCAACAAAAAATTTGTTAATATACCTCTATACTTTAACGTCAAGGAGAAAAAACCCG
GATCCACCAGCTTGGTACCGCGGCTAGCTAAGATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGA
```

AGTCTAGGTCCCTATTTATTTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTTC
TTTTTTTTTCTGTACAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGA
CGCTCGAAGATCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTGCGTATTGGGCGCTC
TTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA
AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCGTTTCAGCCCCAGCGCTGCGCCTTATCCGGTAACATATCGTCTTGAGTCCA
ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
GGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTG
CGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
GTAGCGGTGGTTTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGT
AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTCA
TCCATAGTTGCCTGACTCCCCGCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC
TGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGG
CCGAGCGCAGAAGTGGTCCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA
GTAAGTAGTTTCGCCAGTTAATAGTTTTCGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTC
GTCGTTTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT
GCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTGTCAGAAGTAAGTTGGCCGCGAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGA
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG
ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTC
TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACCTGATCTTCAGCATC
TTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGT
CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG
AAAAGTGCCACCTGAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATT
TTTCAAACAAAGAATCTGAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACG
AGAATCTGTGCTTCATTTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTTCAAACAAAGAATCT
GAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTTTACCAACAAAGAATCTATACTTCTT
TTTTGTTCTACAAAAATGCATCCCGAGAGCGCTATTTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTC
CTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTTGCAGTGTAGGTCCGTTAAGGTTAGAAGAAGGC
TACTTTGGTGTCTATTTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTTTACTGATTACTAGCG
AAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATTATATTCTATACCGATGTGGATTGCGCATA
CTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTCATTGGTCAGAAAATTATGAACGGTTTTCTTCTATT
TTGTCTCTATATACTACGTATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTT
CTTACTACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAAACATAAAAAATGTAGAGGTCGAGTTTGA
TGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGA
GATACTTTTTGAGCAATGTTTGTGGAAGCGGTATTTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCCTT
TTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGTTTTTCAAAGCGCTCTGAAGTTCTTATACTTTT
TAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGC
GAGCTGCGCACATACAGCTCACTGTTACGTCGCACCTATATCTGCGTGTTGCCTGTATATATATATACAT
GAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAG
GTAGTCTAGTACCTCCTGTGATATTATCCCATTCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCT
TTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCTCATCCTTCAATGCTATCATTTCTCTTTGAT
ATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTT
TCGTCTCGCGCGTTTTCGGTGATGACGGTGAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTT
GTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGC
TGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATA

>pTHS2 (circular vector sequence)

TATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCC
GCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAG
CTGCATGTGTACAGAGTTTTTACCCTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT

TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGTATGATCCAATATCAAAGGAAATGATAGCATTG
AAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATAGAACAGCTAAAGGGTAGTGCTGAAGGAAGCAT
ACGATACCCCGCATGGAATGGGATAATATCACAGGAGGTACTAGACTACCTTTTCATCCTACATAAATAGAC
GCATATAAGTACGCATTTAAGCATAAACACGCACATATGCCGTTCTTCTCATGTATATATATATACAGGCAA
CACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCT
CGTTTTTCGGAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCAGAGC
GCTTTTGAAAACCAAAGCGCTCTGAAGACGCACCTTTCAAAAAACCAAACGCACCGGACTGTAACGAGC
TACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGC
TATATCCCTATATAACCTACCCATCCACCTTTTCGCTCCTTGAACCTGCATCTAACTCGACCTCTACATTT
TTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAATTGTAGTAAGAAGTATTTCATAGAGTGAATCGA
AAACAATACGAAAATGTAAACATTTCTATACGTAGTATATAGAGACAAAATAGAAGAAACCGTTCATAAT
TTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTTCACAAAGTATGCGCAATCCACATCGGTAT
AGAATATAATCGGGGATGCCTTTATCTTGAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGG
GAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGA
CCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAGTAATCTAAGAT
GCTTTGTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAAGAAAGTATAGATTCTTTGTTG
GTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTA
GCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTTCG
CGTTGCATTTCTGTTCTGTAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCAT
TTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTGAGTGGCACTTTTCGGGGAATGTGCGCGGAACCC
CTATTTGTTTATTTTTCTAAATACATTCAAAATATGTATCCGCTCATGAGACAATAACCTGATAAAGCTT
CAATAATATTTAAAAAGGAAGATATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTTCGCGC
ATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG
CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGCT
TTTCCAATGATGAGCACTTTTAAAGTTCGTCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGA
GCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATC
TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAAC
TTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAC
TCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG
TAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTA
ATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTAT
TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGC
CCTCCCGTATCTGAGTTATCTACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT
GAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTACAGCAAGTTTACTCATATATACCTTTAGATTGA
TTTTAAACTTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCC
CTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCGCTAGAAAAGATCAAAGGATCTTCTTGAGATCCT
TTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGCCGGA
TCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCACCTACATACCTCGCTCTGCTAATC
CTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC
GGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACA
CCGAAGTGAATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG
TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCT
TTATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA
GCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTTACGGTTTCTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCG
CCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGC
CTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGATCTTCGAGCGTCCCAAAACCTTCTCAAGCAAG
GTTTTTCAGTATAATGTTACATGCGTACACGCGTCTGTACAGAAAAAAGAAAAATTTGAAATATAAATAA
CGTTCTTAATACTAATACTATAAAAAAATAAATAGGGACCTAGACTTCAGGTTGTCTAACTCCTTCCT
TTTCGGTTAGAGCGGATCTTAGCTAGCCGCGGTACCAAGCTGGTGGATCCGGGGTTTTTTCTCCTTGACGT
TAAAGTATAGAGGTATATTAACAATTTTTTTGTTGATACTTTTATTACATTTGAATAAGAAGTAATACAAAC
CGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGTTTTTGCATTTATATATCTGTTAATAGATCAAA
AATCATCGCTTCGCTGATTAATTACCCAGAAATAAGGCTAAAAAACTAATCGCATTATCATCCTATGGTT
GTTAATTTGATTGTTTCATTTGAAGGTTTTGTGGGGCCAGGTTACTGCCAATTTTTCTCTTCATAACCATA
AAAGCTAGTATTGTAGAATCTTTATTGTTTCGGAGCAGTGCGGCGCGAGGCACATCTGCGTTTCAGGAACGC
GACCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGTACCCGCTCGGCGGCTTCTA
ATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAGTTCCAAAGAGAAGGTTTTTTTAGG

CTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGGATATGTATATGGATAT
GTATATGGTGGTAATGCCATGTAATATGATTATTAACTTCTTTGCGTCCATCCAAAAAAGTAAGAAT
TTTTGAAAATTCGAATTC AACCTCACTAAAGGGCGGCCGCATGTATCCATACGATGTTCCAGATTATGCT
GCACCATTAGGTGTTTCTGGTTTGGTTGGTAAATTGTCAACTGAATTGGAAGTTGATTGTGATGCTGAAAA
GTACTACAACATGTACAAACATGGTGAAGATGTTAAGAAAGCAGTTCCACATTTGTGTGTTGATGTTAAGA
TCATCTCTGGTGACCCAACTTCTTCAGGTTGTATCAAGGAATGGAACGTTAACATCGATGGTAAAAACAATC
AGATCAGTTGAAGAACTACACATGATGATGAACTAAGACATTGAGACATAGAGTTTTTCGAAGGTGACGT
TATGAAGGATTTCAAGAAATTCGATACTATTATGGTTGTTAATCCAAAACCAGATGGTAATGGTTGTGTTG
TTACAAGATCTATCGAATACGAAAAGACTAACGAAAACCTCACCAACACCATTTCGATTACTTACAATTTGGT
CATCAAGCTATTGAAGATATGAATAAGTATTTGAGAGATTCTGAATCAAATTAAGAGCTCTTAATTAACAA
TTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTGCGCTTGTCTACCTTGCCAGAAATTTACGA
AAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGAT
TTATGATTTTTTATTATTAAATAAGTTATAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTT
TTAAAACGAAAATTTCTTATTCTTGAGTAACCTCTTCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGA
GGTCGCTCCAATTCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATTCGACACTGGATGGCGGCGTTAGTATCGAATCGACAGCAGTATAGCGACCAGCATTCA
CATACGATTGACGCATGATATTACTTTCTGCGCACTTAACTTCGCATCTGGGCAGATGATGTGCGAGGCGAA
AAAAATATAAATCACGCTAACATTTGATTAAAAATAGAACAACCTACAATATAAAAAAATATAACAATGAC
AAGTTCTTGAAAACAAGAATCTTTTTATTGTCTAGTACTGATTAGAAAACTCATCGAGCATCAAATGAAAC
TGCAATTTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAGCCGTTTCTGTAATGAGGAGAAAA
CTCACCAGGCGAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAA
TACAACCTATTAATTTCCCTCGTCAAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAA
TCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTC
ATCAAAATCACTCGCATCAACCAACCGTTATTTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGAT
CGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACA
ATATTTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTGCCGGGGATCGCAGTGGTGAG
TAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCGGTCAGCCAGT
TTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTTCAGAAACAACCTCTGGC
GCATCGGGCTTCCCATACAATCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCCATTTATA
CCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCCTCGAAACGTGAGTCTTTTCTTACCCATGG
TTGTTTATGTTTCGGATGTGATGTGAGAACTGTATCCTAGCAAGATTTTAAAAGGAAGTATATGAAAGAAGA
ACCTCAGTGGCAAATCCTAACCTTTTATATTTCTCTACAGGGGCGCGGCGTGGGGACAATTC AACCGCTCT
GTGAGGGGAGCGTTTTCCCTGCTCGCAGGTCTGCAGCGAGGAGCCGTAATTTTTGCTTCGCGCCGTGCGGCC
ATCAAAATGATGGATGCAAAATGATTATACATCGGGGATGTATGGGCTAAATGTACGGGCGCAGTCAATC
ATGCCCCTGAGCTGCGCACGTCAAGACTGTCAAGGAGGGTATTCTGGGCCTCCATGTCTGCTGGCCGGGTGA
CCCGGCGGGGACGAGGCAAGCTAAACAGATCTGGCGCGTGTACTGAGAGTGCACCATAAATTTCCGTTTTTA
AGAGCTTGGTGAGCGCTAGGAGTCACTGCCAGGTATCGTTTTGAACACGGCATTAGTCAGGGAAGTCATAAC
ACAGTCCTTTCCCGCAATTTTCTTTTTCTATTACTCTTGGCCTCCTCTAGTACACTCTATATTTTTTTATG
CCTCGGTAATGATTTTTCATTTTTTTTTTTTCCACCTAGCGGATGACTCTTTTTTTTTTCTTAGCGATTGGCAT
TATCACATAATGAATTATACATTATATAAAGTAATGTGATTTCTTCGAAGAATATACTAAAAAATGAGCAG
GCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGAAACCAAGAT
TCAGATTGCGATCTCTTTAAAGGGTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAG
AAGCAGTAGCAGAACAGGCCACACAATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCAT
ATGATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGTGACTTACACATAGA
CGACCATCACACCACTGAAGACTGCGGGATGCTCTCGGTCAAGCTTTTAAAGAGGCCCTAGGGGCCGTGC
GTGGAGTAAAAAGGTTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTTCCAGAGCGGTGGTAGATCTT
TCGAACAGGCCGTACGCAGTTGTGCAACTTGGTTTTGCAAAGGGAGAAAGTAGGAGATCTCTCTTGCAGAT
GATCCCGCATTTTCTTGAAAGCTTTGCAGAGGCTAGCAGAATTACCCTCCACGTTGATTGTCTGCGAGGCA
AGAATGATCATCACCGTAGTGAGAGTGC GTTCAAGGCTCTTGGCGTTGCCATAAGAGAAGCCACCTCGCCC
AATGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATGTAGTGACACCGATTATTTAAAGCTGCAGC
ATACGATATATATACATGTGTATATATGTATACCTATGAATGTGAGTAAGTATGTATACGAACAGTATGAT
ACTGAAGATGACAAGGTAATGCATCATTCTATACGTGTCATTCTGAACGAGGCGCGCTTTTCTTTTTTCTT
TTTGCTTTTTTCTTTTTTTTTTCTTCTTGAACCTCGACGGATCTATGCGGTGTGAAA

Supplementary Note 3. Sequences of PR10-type proteins used for phylogenetic analysis.

>PsoTHS1

MDSINSSIIYFCAYFRELI IKLLMAPPGVSGLVGKLSTELEVNCDAEKYYNMYKHGEDVQKAVPHLCVDVKVISG
DPTRSGCIKEWNVNIDGKTIRSVEETHNDETKTLRHRVFEGDMMKDFKKFDTIMVVNPKPDGNGCVVTRSIEY
EKTNENSPTPFDY LQFGHQAIEDMNKYLRDSE

>PsoTHS1-t

MYKHGEDVQKAVPHLCVDVKVISGDPTSGCIKEWNVNIDGKTIRSVEETHNDETKTLRHRVFEGDMMKDFKK
FDTIMVVNPKPDGNGCVVTRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKYLRDSE

>PsoTHS2

MAPLGVSGLVGKLSTELEVDCDAEKYYNMYKHGEDVKKAVPHLCVDVKIISGDPTSSGCIKEWNVNIDGKTIRS
VEETHHDETKTLRHRVFEGDVMKDFKKFDTIMVVNPKPDGNGCVVTRSIEYEKTNENSPTPFDY LQFGHQAIE
DMNKYLRDSESN

>PsoTHS2-t

MYKHGEDVKKAVPHLCVDVKIISGDPTSSGCIKEWNVNIDGKTIRSVEETHHDETKTLRHRVFEGDVMKDFKK
FDTIMVVNPKPDGNGCVVTRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKYLRDSESN

>PseBetv1-1

MDSINSSIIYFCAYFRELI IKLLMAPLGVSGLVGKLSTELEVDCDAEKYYNMYKHGEDVKKAVPHLCVDVKIISG
DPTSSGCIKEWNVNIDGKTIRSVEETHHDETKTLRHRVFEGDVMKDFKKFDTIMVVNPKPDGNGCVVTRSIEY
EKTNENSPTPFDY LQFGHQAIEDMNKYLRDSESN

>PseBetv1-2

MAPLGVSGLVGKLSTELEVDCDAEKYYNMYKHGEDVKKAVPHLCVDVKIISGDPTSSGCIKEWNVNIDGKTIRS
VEETHHDETKTLRHRVFEGDVMKDFKKFDTIMVVNPKPDGNGCVVTRSIEYEKTNENSPTPFDY LQFGHQAIE
DMNKYLRDSESN

>PseBetv1-3

MACHGVSGLVGKLITELEVNCDAKYYKIYKRPEDVRKAIPHLCTSIKLINGDASRSGCIKEWNFILEGKMIRT
VEETHNDETRTLHHRIFEGDLMKDYYKKFDSII EVNPKPNGNCIVKRSIVYEKINKDSPTPFAYVPFCHQAIE
DMNKHLCGSE

>PbrBetv1-1

MAPRGVSGLVGKLSTELDVNCDAEKYYNMYKNGEDVQKAVPHLCMDVKVISGDATRSGCIKEWNVNIDGKTIRS
VEETHNDESKTLRHRVFEGDMMKDYYKKFDTIMEVKPKPDGNGCVVTRSIEYEKVNENSPTPFDY LQFGHQAME
DMNKYLRDSE

>PbrBetv1-2

MEVSCDADKYYKIYKHAEDVQKAIPHLCIDVKVINGDATRSGCIKEWNFIIEGKMVRSVEETHNDETKTIHHR
VFEGDLMKDYYKKFDSVIQVNPKPNGSVVTRSIAYEKINEDSPTPFAYILFSHQAIEDMNKYLCDSE

>PrhBetv1-1

MAHHHSISGLVGKLVTELEVNCDAKYYTIFKHQEDVPNAVPHLYSDVKVINGHAIHSGCIKEWNFIVEGKTIC
AVEETHNDDTRTLHHRFEGDLMKDYYKKFDSII EVNPKPNGNCVVTWSIAYEKINEDSPVPFAYIPFCHQVIE
DMNIHLCDTE

>PsoMLP2

MAHAHGISGLVGKLITESEVNCNADKFYQMFKH DENITNIIPHIYTSFKVVEGDGLISGCTKEWGYLSEGKARI
VKEQTTFDDETRTIHHCAGDMMNDYKKFVLT LVNPKAHGQGSTVKWIIDYEKINEDSPVPFAYLSLCIKIT
EGLNSHIYASE

>PsoMLP15

MAHQHTISGLVGKLITeseVNCNADKYYQIFKHHEDLPSAIPHIYTSVKAVEGHGTTSGCVKEWCYILEGKPLT
VKEKTTYNDETRTINHNGIEGGMMTDYKKFVATLVVKPKANGQGSIVTWIVDYEKINEDSPVPFDYLAFFQQNI
EDLNSHLCASD

>PsoPR10-3

MAHHGVSGLVGKIVTELEVNCNADEFYKILKRDEDVPRAVSDLFPPVKIAKGDGLVSGCIKEWDCVLDGKAMSG
KEETHHNDETRTRLRHRELEGDLMDYKKFDSIIIEVNPKNPGHGSIVTWSIEYEKMNEDSPAPFAYLASFHQNVV
EVDShLCLSE

>PsoPR10-4

MAHPHPISGLVGKLVTELEVNCADADKYYKIFKHHEDVPKAVPHMYTSVKVVEGHGITSVCVKEWGYLLEGKELI
VKETTTYTDETRTIHHSAGVGHMTKIYKKFDATLVVNPKNPSGHGSTVSWTIDYEKINEDSPVPIPYLAFFHKLI
EDLNSHLCASD

>PsoPR10-5

MAHHGVSGLVGKLVTELEVHCNADAYYKIFKHQEDVPKAMPHLYTGKKVISGDATRSGCIKEWNYILEGKALIA
VEETHHDDETRTRLTHRITGGDLTKDYKKFVKIVEVNPKNPGHGSIVTVSLVEKMNEGSPTPFNYLQFVHQTVI
GLNSHICAS

>PsoNCS2

MSKLITTEPLKSMAEVISNYAMKQQSVSERNIPKKQSLRKEITYETEVTQTSADSIWNVYSSPDIPRLLRDVLL
PGVFEKLDVIAGNGGVGTVLDIAFPLGAVPRRYKEKFVKINHEKRLKEVVMIEGGYLDMGCTFYMDRIHIFEKT
PNSCVIESSIIEVKEEYAGKMAKLITTEPLESMAEVISGYVLKKRLQVFGFEIKPKLRFNLLLCLIIICLVIAG
GMFVAGVPL

>Tf1NCS

MMKMEVVVFVFLMLLGTINCQKLILTGRPFLLHHQGIINQVSTVTKVIHHELEVAASADDIWTVYSWPGLAKHLPD
LLPGAFAFEKLEIIGDGGVGTILDMTFVPGEFPHEYKEKFILVDNEHRLKKVQMIIEGGYLDLGVTYYMDTIHVPT
GKDSCVIKSSTEYHVKPEFVKIVEPLITTGPLAAMADAISKLVLEHKSXNSDEIEAAIITV

>AthRCAR1

MMDGVEGGTAMYGGLETVQYVVRTHHQHLCRENQCTSALVKHIKAPLHLVWSLVRRFDQPQKYKPFVSRCTVIGD
PEIGSLREVNKSGLPATTSTERLELLDDEEHILGIKIIGDHRLLKNYSSILTVHPEIIEGRAGTMVIESFVVD
VPQGNTKDETCYFVEALIRCNLKSLADVSERLASQDITQ

>PtrRCAR1

MDTNQAPPPQGLTQEEYMELKPLIDTYHKFEPAPNTCTSLITQRIDAPAHVVPFVRRFDNPQKYKHFIKSCNM
SAGDGGVGSVREVAVVSGLPASTSIERLEILDDENHILSFRVVGGEHRLNNYKSVTSVNEFNKEGKVYAIVLES
YIVDIPGGNTGEDTKMFVDTVVKLNLQKLAVVAIASLHGHE

>ShyPR10

MASANSWTLEIASPVAAPRLFRAAVLDWHTLAPKLASHVVASAHPVEGDGGVGSIRQFNFTSVMPTFTTKERLD
FLDLDKCECKQTLIEGGGIGVAIETATSHIKVEPAAGGGSVAKVETTFKPLPGVEVKDEITKAKESLTAIFKGA
EAYLVANPDAYN

>SolPR10

MGVYTFTVIDEASSVAPARLFEALCLDNHNVFPKVVPFIVKSVDVFEVDTTSGVCVKQMNFAADGGPYKYMKTKI
NELDVDKYYAKYTNFEGDVLNVLECVVYESSIESTGSGSHYKLVGHFHTKGDVVYTLEDAKMGLQSMQVSYKA
VEEYLSNNPEIYA

>FanA1E

MGVYTYENEFTSDIPAPKLFKAFVLDADNLIPIKAPQAVKCAEILEGDGGPGTIKKITFGEGSHYGYVKHKIHS
IDKVNHTYSYSLIEGDALSENIEKIDYETKLVSAHPGGTIKTTSKYHTKGDVEIKEEHVKAGKEKAAHLFKLI
EGYLDKHPSEYN

>VpsPR10-1

MGVFTYESEVTSSVPPAKMFKAAILDSDNLIPKIRPQDIKSVEILQGEGGPGTIKKIHFGEGRKFKSMTHRIDA
IDKENFTFSYTVIDGDVLTSGIESISHELKVVSSPDGGCIYKNTKKYHTKAGVEISEEHVKGKGKEESLALFKAI
EAYIIAHPDAY

>CanPR10

MGAYTFTDKSTASVAPSRLFKALVIDFNNLVSKLAPDVKSIENVEGDGGAGTIKKMTFVEGGPIKYMKHKIHVI
DEKNLVTKYSLIESDVTENKAESVDYDGKFEASADGGSVCTTVTVYNTKGDYVVTEEEHNVHKEKANDLLKAIE
AYLLANPSVYV

>GmaPR10

MGVVTTQIYDTPAAVPPTLRFKAMTLDFHNLFPKLVDSIHSIVFTQGNGGPGTIKKITTIEGDKTKYVLHRVDAI
DEANFVYNFSITEGTALADTLEKVSFESQLVEAPNGGSIRKVSVOFFTKGDATLSEEELTANKAKIQGLVKLVE
GYLLANPDY

>LluPR10-1A

MGIFAFENEQSSTVAPAKLYKALTKDSDEIVPKVIEPIQSVEIVEGNGGPGTIKKIIAHDGHTSFVLHKLDAI
DEANLTYNYSIIGGEGLDSELEKISYESKILPGPDGGSIGKINVKFHTKGDVLSETVRDQAKFKGLGLFKAI
EYVLAHPDY

>PsaABR17

MGVFVFDDEYVSTVAPPKLYKALAKDADEIVPKVIEAQQVEIIEGNGGPGTIKKLSILEDGKTNYVLHKLDAV
DEANFGYNYSLVGGPGLHESLEKVAFETIILAGSDGGSIVKISVKYHTKGAALSDAVRDETKAKGTGLIKAIE
GYVLANPGY