Supplemental data for: Jounal of Experimental Botany

Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in *Salvia miltiorrhiza*

Yimian Ma, Lichai Yuan, Bin Wu, Xian’en Li, Shilin Chen, Shanfa Lu

Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, No.151, Malianwa North Road, Haidian District, Beijing 100193, China

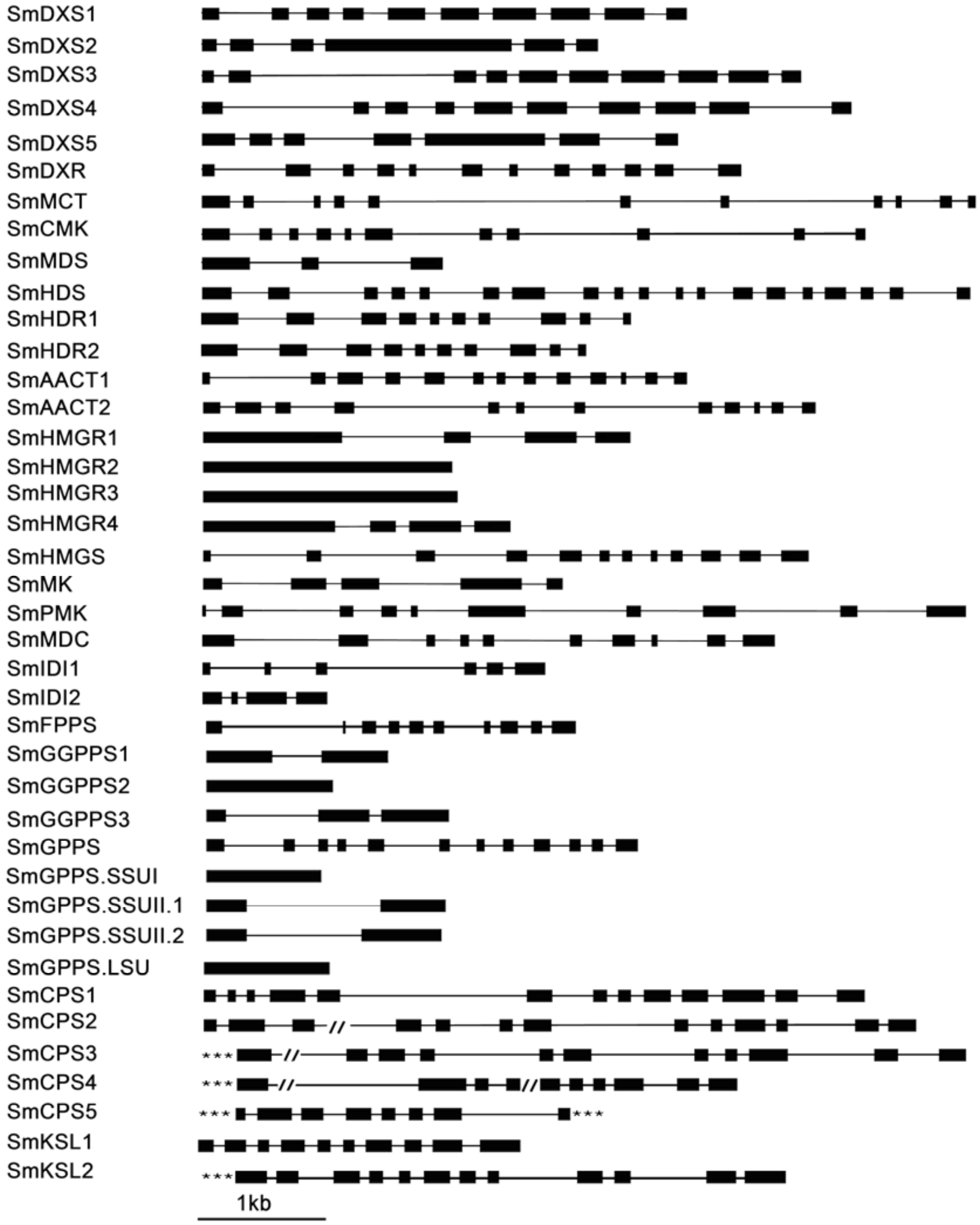
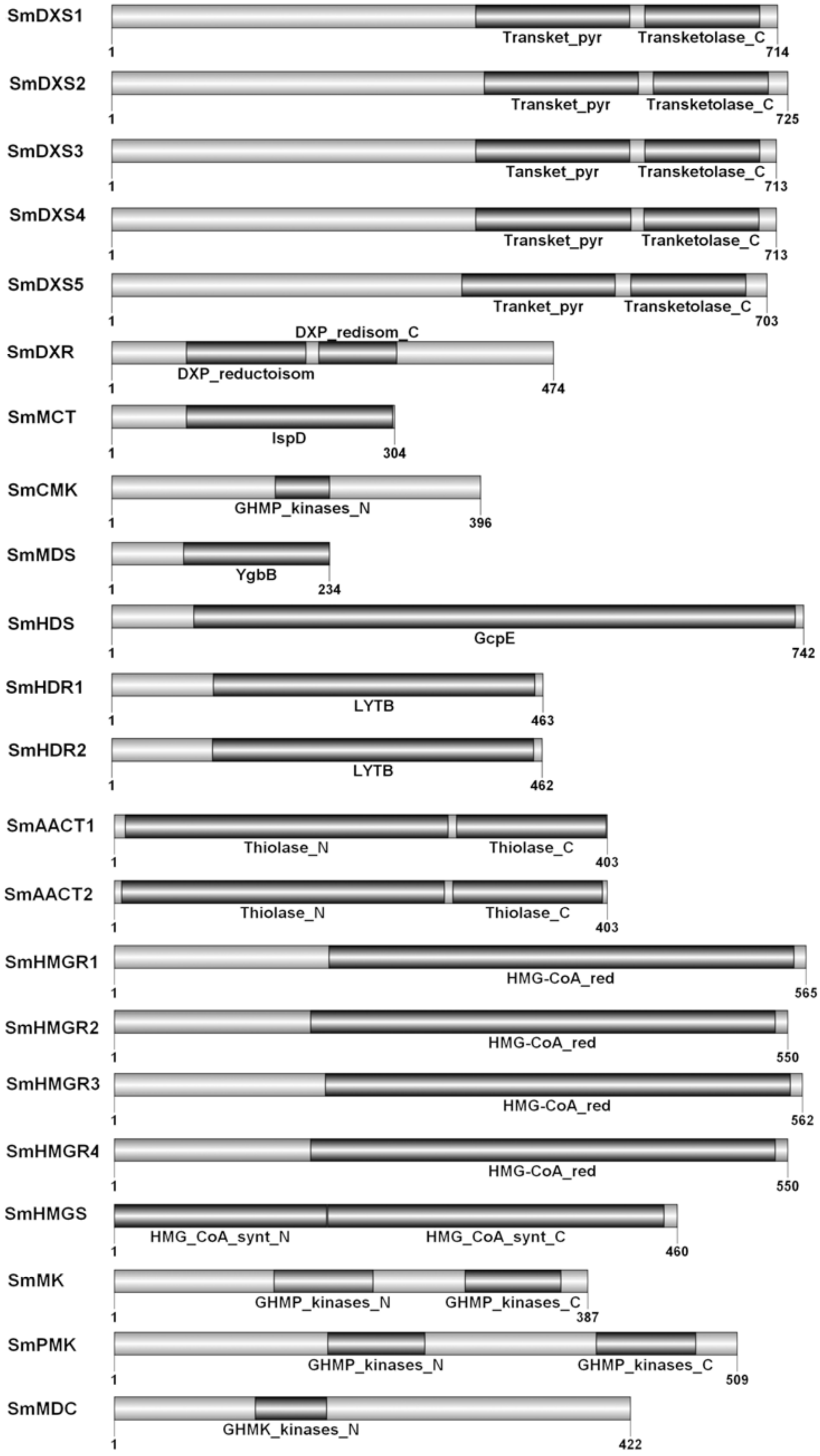


Figure S1. Exon/intron Structures of forty terpenoid biosynthesis-related genes. Heavy black lines represent exons. Thin black lines represent introns. Unknown sequences in the 5' or 3' regions of *SmCPS3*, *SmCPS4*, *SmCPS5* and *SmKSL2* are indicated by ‘***’, whereas unknown sequences in introns are indicated by ‘//’.


SmDXS1
Transket_pyr
Transketolase_C
1
714
SmDXS2
Transket_pyr
Transketolase_C
1
725
SmDXS3
Tansket_pyr
Transketolase_C
1
713
SmDXS4
Transket_pyr
Tranketolase_C
1
713
SmDXS5
Tranket_pyr
Transketolase_C
1
703
DXP_redisom_C
SmDXR
DXP_reductoisom
1
474
SmMCT
IspD
1
304
SmCMK
GHMP_kinases_N
1
396
SmMDS
YgbB
1
234
SmHDS
GcpE
1
742
SmHDR1
LYTB
1
463
SmHDR2
LYTB
1
462
SmAACT1
Thiolase_N
Thiolase_C
1
403
SmAACT2
Thiolase_N
Thiolase_C
1
403
SmHMGR1
HMG-CoA_red
1
565
SmHMGR2
HMG-CoA_red
1
550
SmHMGR3
HMG-CoA_red
1
562
SmHMGR4
HMG-CoA_red
1
550
SmHMGS
HMG_CoA_synt_N
HMG_CoA_synt_C
1
460
SmMK
GHMP_kinases_N
GHMP_kinases_C
1
387
SmPMK
GHMP_kinases_N
GHMP_kinases_C
1
509
SmMDC
GHMK_kinases_N
1
422

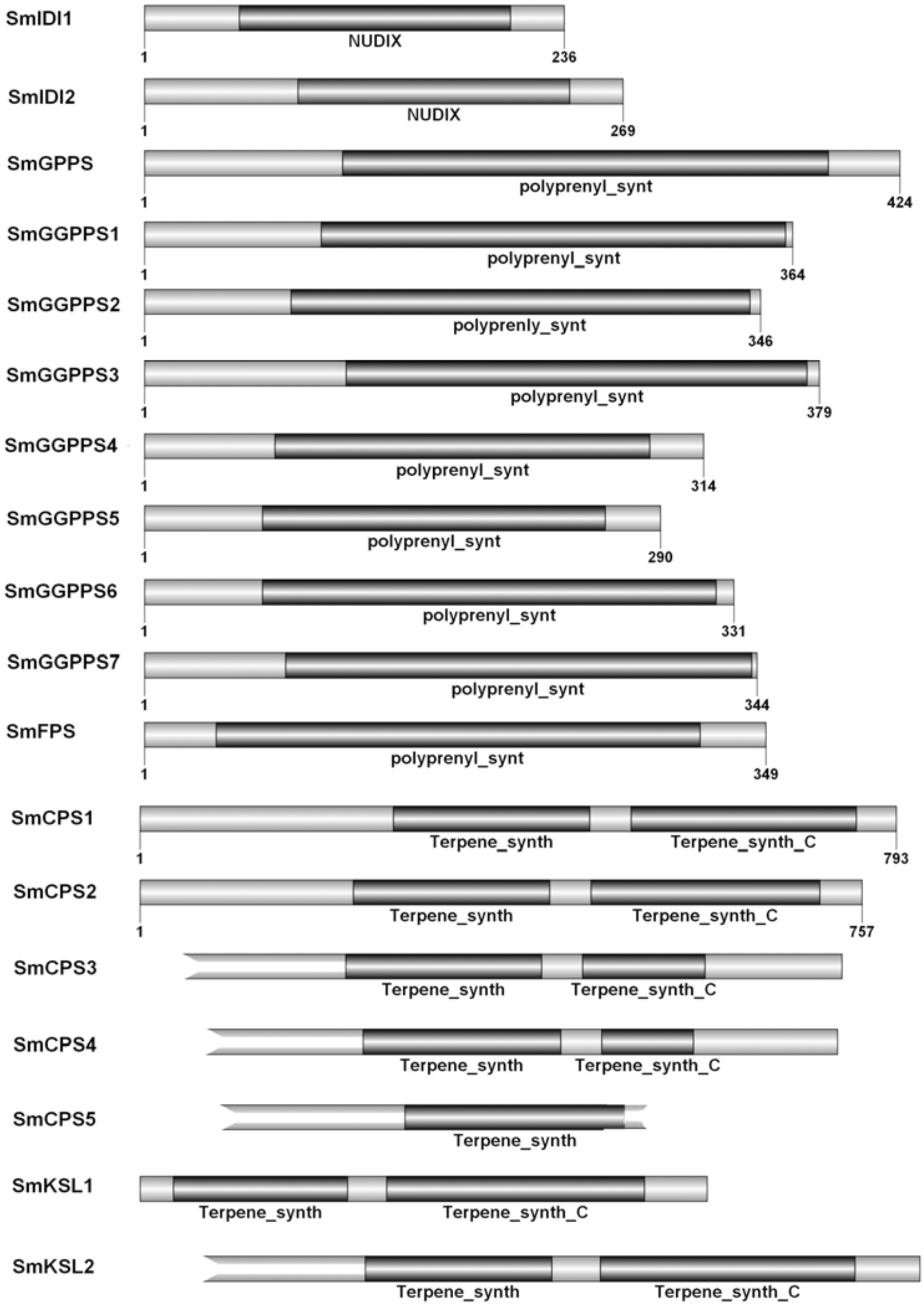
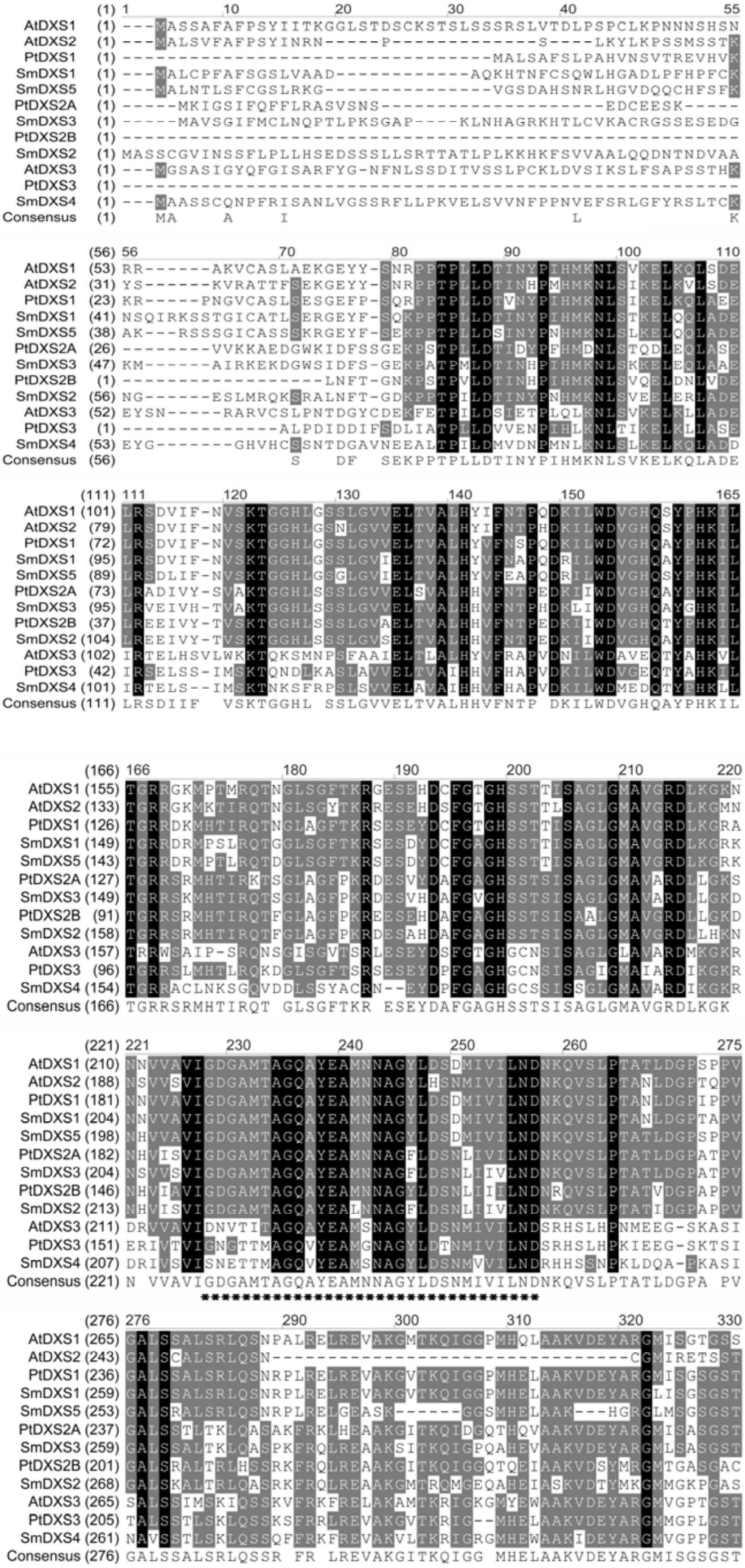
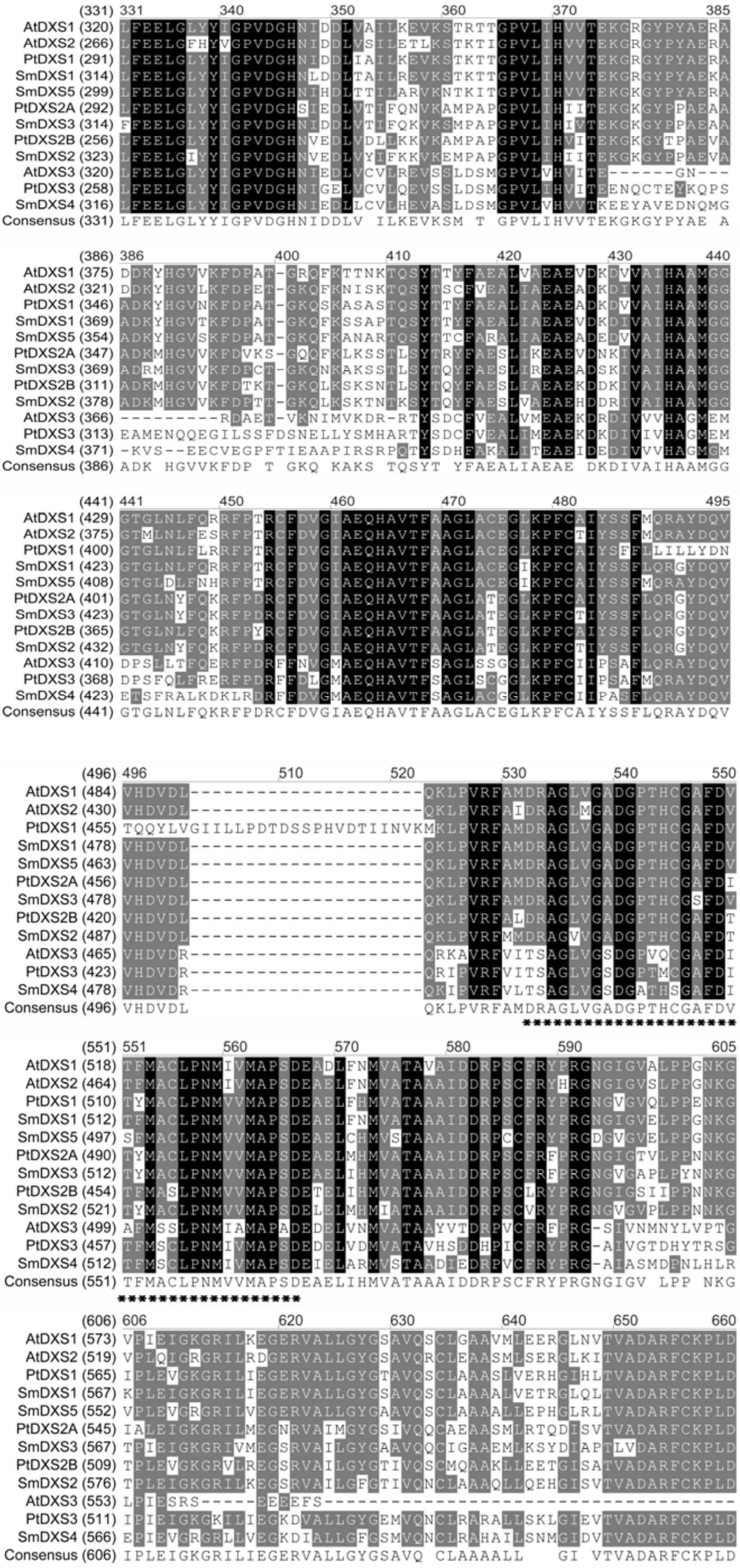


Figure S2. Conserved domains (black boxes) of enzymes involved in terpenoid biosynthesis in *S. miltiorrhiza*. Conserved domains were predicated by searching Pfam HMMs. Names of conserved domains are shown below the black boxes.


(1) 1 10 20 30 40 55
AtDXS1 (1) ---MASSAFAFPSYIITKGGLSTDSCKSTSLSSSRSLVTDLPSPCLKPNNNSHSN
AtDXS2 (1) ---MALSVFAFPSYINRN-----P-------------S----LKYLKPSSMSSTK
PtDXS1 (1) ----------------------------------MALSAFSLPAHVNSVTREVHVK
SmDXS1 (1) ---MALCPFAFSGSLVAAD------------AQKHTNFCSQWLHGADLPFHPFCK
SmDXS5 (1) ---MALNTLSFCGSLRKG---------------VGSDAHSNRLHGVDQQCHFSFK
PtDXS2A (1) -----MKIGSIFQFFLRASVSNS--------------------EDCEESK-----
SmDXS3 (1) -----MAVSGIFMCLNQPTLPKSGAP----KLNHAGRKHTLCVKACRGSSESEDG
PtDXS2B (1) ---
SmDXS2 (1) MASSCGVINSSFLPLLHSEDSSSLLSRTTATLPLKKHKFSVVAALQQDNTNDVAA
AtDXS3 (1) ---MGSASIGYQFGISARFYG-NFNLSSDITVSSLPCKLDVSIKSLFSAPSSTHK
PtDXS3 (1) ---
SmDXS4 (1) ---MAASSCQNPFRISANLVGSSRFLLPKVELSVVNFPPNVEFSRLGFYRSLTCK
Consensus (1) MA A I L K
(56) 56 70 80 90 100 110
AtDXS1 (53) RR------AKVCASLAEKGEYY-SNRPPTPLLDTINYPIHMKNLSVKELKQLSDE
AtDXS2 (31) YS------KVRATTFSEKGEYY-SNRPPTPLLDTINHPMHMKNLSIKELKVLSDE
PtDXS1 (23) KR-----PNGVCASLSESGEFP-SQRPPTPLLDTVNYPIHMKNLSIKELKQLAEE
SmDXS1 (41) NSQIRKSSTGICATLSERGEYF-SQKPPTPLLDTINYPIHMKNLSTKELQQLADE
SmDXS5 (38) AK---RSSSGICASSSKRGEYF-SEKPPTPLLDSINYPNHMKNLSLKELQQLADE
PtDXS2A (26) --------VVKKAEDGWKIDFSSGEKPSTPLLDTIDYPFHMDNLSTQDLEQLASE
SmDXS3 (47) KM------AIRKEKDGWSIDFS-GEKPATPMLDTINHPIHMKNLSKKELEQLAAE
PtDXS2B (1) -------------------LNFT-GNKPSTPVLDTINHPIHMKNLSVQELDNLVDE
SmDXS2 (56) NG------ESLMRQKSRALNFT-GDKPPTPILDTINYPNHMKNLSVEELERLADE
AtDXS3 (52) EYSN-----RARVCSLPNTDGYCDEKFETPILDSIETPLQLKNLSVKELKLLADE
PtDXS3 (1) ---------------ALPDIDDIFSDLIATPLLDVVENPIHLKNLTIKELKLLASE
SmDXS4 (53) EYG-------GHVHCSSNTDGAVNEEALTPILDMVDNPMNLKNLSLKELKQLADD
Consensus (56) S DF SEKPPTPLLDTINYPIHMKNLSVKELKQLADE
(111) 111 120 130 140 150 165
AtDXS1 (101) LRSDVIF-NVSKTGGHLGSSLGVVELTVALHYIFNTPQDKILWDVGHQSYPHKIL
AtDXS2 (79) LRSDVIF-NVSKTGGHLGSNLGVVELTVALHYIFNTPHDKILWDVGHQSYPHKIL
PtDXS1 (72) LRSDVIF-NVSKTGGHLGSSLGVVELTVALHYVFNSPQDKILWDVGHQAYPHKIL
SmDXS1 (95) LRSDVIF-NVSKTGGHLGSSLGVIELTVALHYVFNAPQDRILWDVGHQAYPHKIL
SmDXS5 (89) LRSDLIF-NVSKTGGHLGSGLGVIELTVALHYVFEAPQDRILWDVGHQSYPHKIL
PtDXS2A (73) LRADIVY-SVAKTGGHLSSSLGVVELSVALHHVFNTPEDKIIWDVGHQAYPHKIL
SmDXS3 (95) LRVEIVH-TVAKTGGHLSSSLGVVELTVALHHVFNTPHDKLIWDVGHQAYGHKIL
PtDXS2B (37) LREEIVY-TVSKTGGHLSSSLGVAELTVALHHVFNTPEDKIIWDVGHQTYPHKIL
SmDXS2 (104) LREEIVY-TVSKTGGHLSSSLGVSELTVALHHVFNTPEDKIIWDVGHQAYPHKIL
AtDXS3 (102) IRTELHSVLWKKTQKSMNPSFAAIELTLALHYVFRAPVDNILWDAVEQTYAHKVL
PtDXS3 (42) IRSELSS-IMSKTQNDLKASLAVVELTVAIHHVFHAPVDKILWDVGEQTYAHKIL
SmDXS4 (101) IRTELS--IMSKTNKSFRPSLSVVELAVAIHHVFHAPVDKILWDMEDQTYPHKLL
Consensus (111) LRSDIIF VSKTGGHL SSLGVVELTVALHHVFNTP DKILWDVGHQAYPHKIL
(166) 166 180 190 200 210 220
AtDXS1 (155) TGRRGKMPTMRQTNGLSGFTKRGESEHDCFGTGHSSTTISAGLGMAVGRDLKGKN
AtDXS2 (133) TGRRGKMKTIRQTNGLSGYTKRRESEHDSFGTGHSSTTLSAGLGMAVGRDLKGMN
PtDXS1 (126) TGRRDKMHTIRQTNGLAGFTKRSESEYDCFGTGHSSTTISAGLGMAVGRDLKGRA
SmDXS1 (149) TGRRDRMPSLRQTGGLSGFTKRSESDYDCFGAGHSSTTISAGLGMAVGRDLKGRK
SmDXS5 (143) TGRRDRMPTLRQTDGLSGFTKRSESDYDCFGAGHSSTTISAGLGMAVGRDLKGRK
PtDXS2A (127) TGRRSRMHTIRKTSGLAGFPKRDESVYDAFGAGHSSTSISAGLGMAVARDLLGKS
SmDXS3 (149) TGRRSKMHTIRQTSGLAGFPKRDESVHDAFGVGHSSTSISAGLGMAVARDLLGKD
PtDXS2B (91) TGRRSRMHTIRQTFGLAGFPKREESEHDAFGAGHSSTSISAALGMAVGRDLLGKD
SmDXS2 (158) TGRRSRMHTIRQTFGLAGFPKRDESAHDAFGAGHSSTSISAGLGMAVGRDLLHKN
AtDXS3 (157) TRRWSAIP-SRQNSGISGVTSRLESEYDSFGTGHGCNSISAGLGLAVARDMKGKR
PtDXS3 (96) TGRRSLMHTLRQKDGLSGFTSRSESEYDPFGAGHGCNSISAGIGMAIARDIKGKR
SmDXS4 (154) TGRRACLNKSGQVDDLSSYACRN--EYDPFGAGHGCSSISSGLGMAVARDIKGKR
Consensus (166) TGRRSRMHTIRQT GLSGFTKR ESEYDAFGAGHSSTSISAGLGMAVGRDLKGK
(221) 221 230 240 250 260 275
AtDXS1 (210) NNVVAVIGDGAMTAGQAYEAMNNAGYLDSDMIVILNDNKQVSLPTATLDGPSPPV
AtDXS2 (188) NSVVSVIGDGAMTAGQAYEAMNNAGYLHSNMIVILNDNKQVSLPTANLDGPTQPV
PtDXS1 (181) NNVVAVIGDGAMTAGQAYEAMNNAGYLDSDMIVILNDNKQVSLPTANLDGPIPPV
SmDXS1 (204) NNVVAVIGDGAMTAGQAYEAMNNAGYLDSDMIVILNDNKQVSLPTANLDGPTAPV
SmDXS5 (198) NHVVAVIGDGAMTAGQAYEAMNNAGYLDSDMIVILNDNKQVSLPTATLDGPSPPV
PtDXS2A (182) NHVISVIGDGAMTAGQAYEAMNNAGFLDSNLIVILNDNKQVSLPTATLDGPATPV
SmDXS3 (204) NSVVSVIGDGAMTAGQAYEAMNNAGFLDSNLIIVLNDNKQVSLPTATLDGPATPV
PtDXS2B (146) NHVIAVIGDGAMTAGQAYEAMNNAGYLDSNLIIILNDNRQVSLPTATVDGPAPPV
SmDXS2 (213) NHVISVIGDGAMTAGQAYEALNNAGFLDSNLIIVLNDNKQVSLPTATIDGPAPPV
AtDXS3 (211) DRVVAVIDNVTITAGQAYEAMSNAGYLDSNMIVILNDSRHSLHPNMEEG-SKASI
PtDXS3 (151) ERIVTVIGNGTTMAGQVYEAMGNAGYLDTNMIVILNDSRHSLHPKIEEG-SKTSI
SmDXS4 (207) DRIVSVISNETTMAGQVYEAMSNAGYLDSNMVVILNDRHHSSNPKLDQA-PKASI
Consensus (221) N VVAVIGDGAMTAGQAYEAMNNAGYLDSNMIVILNDNKQVSLPTATLDGPA PV
****
(276) 276 290 300 310 320 330
AtDXS1 (265) GALSSALSRLQSNPALRELREVAKGMTKQIGGPMHQLAAKVDEYARGMISGTGSS
AtDXS2 (243) GALSCALSRLQSN-------------------------------CGMIRETSST
PtDXS1 (236) GALSSALSRLQSNRPLRELREVAKGVTKQIGGPMHELAAKVDEYARGMISGSGST
SmDXS1 (259) GALSSALSRLQSNRPLRELREVAKGVTKQIGGPMHELAAKVDEYARGLISGSGST
SmDXS5 (253) GALSRALSRLQSNRPLRELGEASK------GGSMHELAAK---HGRGLMSGSGST
PtDXS2A (237) GALSSTLTKLQASAKFRKLHEAAKGITKQIDGQTHQVAAKVDEYARGMISASGST
SmDXS3 (259) GALSSALTKLQASPKFRQLREAAKSITKQIGPQAHEVAAKVDEYARGMLSASGST
PtDXS2B (201) GALSRALTRLHSSRKFRQLREAAKGITKQIGGQTQEIAAKVDSYMRGMTGASGAC
SmDXS2 (268) GALSKALTRLQASRKFRQLREAAKGMTRQMGEQAHEIASKVDTYMKGMMGKPGAS
AtDXS3 (265) SALSSIMSKIQSSKVFRKFRELAKAMTKRIGKGMYEWAAKVDEYARGMVGPTGST
PtDXS3 (205) TALSSTLSKLQSSKSFRRLREVAKGVTKRIG--MHELAAKVDEYARGMMGPLGST
SmDXS4 (261) NAVSSTLSKLQSSQFFRKFREVAKVLTKRIGRGMHEWAAKIDEYARGMVGPPGST
Consensus (276) GALSSALSRLQSSR FR LREVAKGITKQIGG MHELAAKVDEYARGMISGSGST



```
          (331) 331      340       350       360       370            385
   AtDXS1 (320) LFEELGLYYIGPVDGHNIDDLVAILKEVKSTRTTGPVLIHVVTEKGRGYPYAERA
   AtDXS2 (266) LFEELGFHYVGPVDGHNIDDLVSILETLKSTKTIGPVLIHVVTEKGRGYPYAERA
   PtDXS1 (291) LFEELGLYYIGPVDGHNIDDLIAILKEVKSTKTTGPVLIHVVTEKGRGYPYAERA
   SmDXS1 (314) LFEELGLYYIGPVDGHNLDDLTAILREVKSTKTTGPVLIHVVTEKGRGYPYAEKA
   SmDXS5 (299) LFEELGLYYIGPVDGHNIHDLTTILARVKNTKITGPVLIHVVTEKGKGYPYAERA
  PtDXS2A (292) LFEELGLYYIGPVDGHSIEDLVTIFQNVKAMPAPGPVLIHIITEKGKGYPPAEAA
   SmDXS3 (314) FFEELGLYYIGPVDGHNIDDLVTIFQKVKSMPAPGPVLIHIVTEKGKGYPPAEAA
  PtDXS2B (256) LFEELGLYYIGPVDGHNVEDLVDLLKKVKAMPAPGPVLIHVITEKGKGYTPAEVA
   SmDXS2 (323) LFEELGIYYIGPVDGHNVEDLVYIFKKVKEMPAPGPVLIHIITEKGKGYPPAEVA
   AtDXS3 (320) LFEELGLYYIGPVDGHNIEDLVCVLREVSSLDSMGPVLVHVITE------GN---
   PtDXS3 (258) LFEELGLYYIGPVDGHNIGELVCVLQEVSSLDSMGPVLIHVITEENQCTEYKQPS
   SmDXS4 (316) LFEELGLYYIGPVDGHNIEDLLCVLHEVASLDSMGPVLVHVVTKEEYAVEDNQMG
Consensus (331) LFEELGLYYIGPVDGHNIDDLV ILKEVKSM T GPVLIHVVTEKGKGYPYAE A

          (386) 386           400       410       420       430      440
   AtDXS1 (375) DDKYHGVVKFDPAT-GRQFKTTNKTQSYTTYFAEALVAEAEVDKDVVAIHAAMGG
   AtDXS2 (321) DDKYHGVLKFDPET-GKQFKNISKTQSYTSCFVEALIAEAEADKDIVAIHAAMGG
   PtDXS1 (346) ADKYHGVNKFDPAT-GKQSKASASTQSYTTYFAEALIAEAEADKDVVAIHAAMGG
   SmDXS1 (369) ADKYHGVTKFDPAT-GKQFKSSAPTQSYTTYFAEALIAEAEVDKDIVAIHAAMGG
   SmDXS5 (354) ADKYHGVSKFDPAT-GKQFKANARTQSYTTCFARALIAEAEADEDVVAIHAAMGG
  PtDXS2A (347) ADKMHGVVKFDVKS-GQQFKLKSSTLSYTRYFAESLIKEAEVDNKIVAIHAAMGG
   SmDXS3 (369) ADRMHGVVKFDPCT-GKQNKAKSSTLSYTQYFAESLIREAEADSKIVAIHAAMGG
  PtDXS2B (311) ADKMHGVVKFDTKT-GKQLKSKSNTLSYTQYFAESLIAEAEKDDKIVAIHAAMGG
   SmDXS2 (378) ADKMHGVVKFDPTT-GKQLKSKTNTKSYTQYFAESLVAEAEHDDRIVAIHAAMGG
   AtDXS3 (366) ---------RDAET-VKNIMVKDR-RTYSDCFVEALVMEAEKDRDIVVVHAGMEM
   PtDXS3 (313) EAMENQQEGILSSFDSNELLYSMHARTYSDCFVEALIMEAEKDKDIVIVHAGMEM
   SmDXS4 (371) -KVS--EECVEGPFTIEAAPIRSRPQTYSDHFAKALITEAEIDEDIVVVHAGMGM
Consensus (386) ADK HGVVKFDP T GKQ KAKS TQSYT YFAEALIAEAE DKDIVAIHAAMGG

          (441) 441      450       460       470       480            495
   AtDXS1 (429) GTGLNLFQRRFPTRCFDVGIAEQHAVTFAAGLACEGLKPFCAIYSSFMQRAYDQV
   AtDXS2 (375) GTMLNLFESRFPTRCFDVGIAEQHAVTFAAGLACEGLKPFCTIYSSFMQRAYDQV
   PtDXS1 (400) GTGLNLFLRRFPTRCFDVGIAEQHAVTFAAGLACEGLKPFCAIYSFFLLILLYDN
   SmDXS1 (423) GTGLNLFQRRFPTRCFDVGIAEQHAVTFAAGLACEGIKPFCAIYSSFLQRGYDQV
   SmDXS5 (408) GTGLDLFNHRFPTRCFDVGIAEQHAVTFAAGLACEGIKPFCAIYSSFMQRAYDQV
  PtDXS2A (401) GTGLNYFQKRFPDRCFDVGIAEQHAVTFAAGLATEGLKPFCAIYSSFLQRGYDQV
   SmDXS3 (423) GTGLNYFQKRFPDRCFDVGIAEQHAVTFAAGLATEGLKPFCTIYSSFLQRGYDQV
  PtDXS2B (365) GTGLNLFQKRFPYRCFDVGIAEQHAVTFAAGLATEGLKPFCAIYSSFLQRGYDQV
   SmDXS2 (432) GTGLNYFQKRFPDRCFDVGIAEQHAVTFAAGLATEGLKPFCTIYSSFLQRGYDQV
   AtDXS3 (410) DPSLLTFQERFPDRFFNVGMAEQHAVTFSAGLSSGGLKPFCIIPSAFLQRAYDQV
   PtDXS3 (368) DPSFQLFRERFPDRFFDLGMAEQHAVTFSAGLSCGGLKPFCIIPSAFMQRAYDQV
   SmDXS4 (423) ETSFRALKDKLRDRFFDVGMAEQHAVTFSAGLACGGLKPFCIIPASFLQRAYDQV
Consensus (441) GTGLNLFQKRFPDRCFDVGIAEQHAVTFAAGLACEGLKPFCAIYSSFLQRAYDQV

          (496) 496           510       520       530       540      550
   AtDXS1 (484) VHDVDL---------------------QKLPVRFAMDRAGLVGADGPTHCGAFDV
   AtDXS2 (430) VHDVDL---------------------QKLPVRFAIDRAGLMGADGPTHCGAFDV
   PtDXS1 (455) TQQYLVGIILLPDTDSSPHVDTIINVKMKLPVRFAMDRAGLVGADGPTHCGAFDV
   SmDXS1 (478) VHDVDL---------------------QKLPVRFAMDRAGLVGADGPTHCGAFDV
   SmDXS5 (463) VHDVDL---------------------QKLPVRFAMDRAGLVGADGPTHCGAFDV
  PtDXS2A (456) VHDVDL---------------------QKLPVRFAMDRAGLVGADGPTHCGAFDI
   SmDXS3 (478) VHDVDL---------------------QKLPVRFAMDRAGLVGADGPTHCGSFDV
  PtDXS2B (420) VHDVDL---------------------QKLPVRFALDRAGLVGADGPTHCGAFDT
   SmDXS2 (487) VHDVDL---------------------QKLPVRFMMDRAGVVGADGPTHCGAFDT
   AtDXS3 (465) VHDVDR---------------------QRKAVRFVITSAGLVGSDGPVQCGAFDI
   PtDXS3 (423) VHDVDR---------------------QRIPVRFVITSAGLVGSDGPTMCGAFDI
   SmDXS4 (478) VHDVDR---------------------QKIPVRFVLTSAGLVGSDGATHSGAFDI
Consensus (496) VHDVDL                     QKLPVRFAMDRAGLVGADGPTHCGAFDV
                                                  **********************

          (551) 551      560       570       580       590            605
   AtDXS1 (518) TFMACLPNMIVMAPSDEADLFNMVATAVAIDDRPSCFRYPRGNGIGVALPPGNKG
   AtDXS2 (464) TFMACLPNMIVMAPSDEAELFNMVATAAAIDDRPSCFRYHRGNGIGVSLPPGNKG
   PtDXS1 (510) TYMACLPNMVVMAPSDEAELFHMVATAAAIDDRPSCFRYPRGNGVGVQLPPENKG
   SmDXS1 (512) TFMACLPNMVVMAPSDEAELFNMVATAAAIDDRPSCFRYPRGNGIGVELPPGNKG
   SmDXS5 (497) SFMACLPNMVVMAPSDEAELCHMVSTAAAIDDRPCCFRYPRGDGVGVELPPGNKG
  PtDXS2A (490) TYMACLPNMVVMAPSDEAELMHMVATAAAIDDRPSCFRFPRGNGIGTVLPPNNKG
   SmDXS3 (512) TYMACLPNMVVMAPSDEAELIHMVATAAAIDDRPSCFRFPRGNGVGAPLPYNNKG
  PtDXS2B (454) TFMASLPNMVVMAPSDETELIHMVATAAAIDDRPSCLRYPRGNGIGSIIPPNNKG
   SmDXS2 (521) TYMACLPNMVVMAPSDELELMHMIATAAAIDDRPSCVRYPRGNGVGVPLPPNNKG
   AtDXS3 (499) AFMSSLPNMIAMAPADEDELVNMVATAAYVTDRPVCFRFPRG-SIVNMNYLVPTG
   PtDXS3 (457) TFMSCLPNMIVMAPSDEDELVDMVATAVHSDDHPICFRYPRG-AIVGTDHYTRSG
   SmDXS4 (512) TFMSCLPNMIVMAPSDEIELARMVSTAADIEDRPVCFRYPRG-AIASMDPNLHLR
Consensus (551) TFMACLPNMVVMAPSDEAELIHMVATAAAIDDRPSCFRYPRGNGIGV LPP NKG
               ******************

          (606) 606           620       630       640       650      660
   AtDXS1 (573) VPIEIGKGRILKEGERVALLGYGSAVQSCLGAAVMLEERGLNVTVADARFCKPLD
   AtDXS2 (519) VPLQIGRGRILRDGERVALLGYGSAVQRCLEAASMLSERGLKITVADARFCKPLD
   PtDXS1 (565) IPLEVGKGRILIEGERVALLGYGTAVQSCLAAASLVERHGIHLTVADARFCKPLD
   SmDXS1 (567) KPLEIGKGRILIEGERVALLGYGSAVQSCLAAAALVETRGLQLTVADARFCKPLD
   SmDXS5 (552) VPLEVGRGRILVEGERVALLGYGAAVQSCLAAAALLEPHGLRLTVADARFCKPLD
  PtDXS2A (545) IALEIGKGRILMEGNRVAIMGYGSIVQQCAEAASMLRTQDISVTVADARFCKPLD
   SmDXS3 (567) TPIEIGKGRIVMEGSRVAILGYGAAVQQCIGAAEMLKSYDIAPTLVDARFCKPLD
  PtDXS2B (509) TPLEVGKGRVLREGSRVAILGYGTIVQSCMQAAKLLEETGISATVADARFCKPLD
   SmDXS2 (576) TPLEIGKGRILKEGSRVAILGFGTIVQNCLAAAQLLQEHGISVTVADARFCKPLD
   AtDXS3 (553) LPIESRS-----EEEEFS-------------------------------------
   PtDXS3 (511) IPIEIGKGKILIEGKDVALLGYGEMVQNCLRARALLSKLGIEVTVADARFCKPLD
   SmDXS4 (566) EPIEVGRGRLLVEGKDIALLGFGSMVQNCLRAHAILSNMGIDVTVADARFCKPLD
Consensus (606) IPLEIGKGRILIEGERVALLGYGSAVQ CLAAAALL    GI VTVADARFCKPLD
```

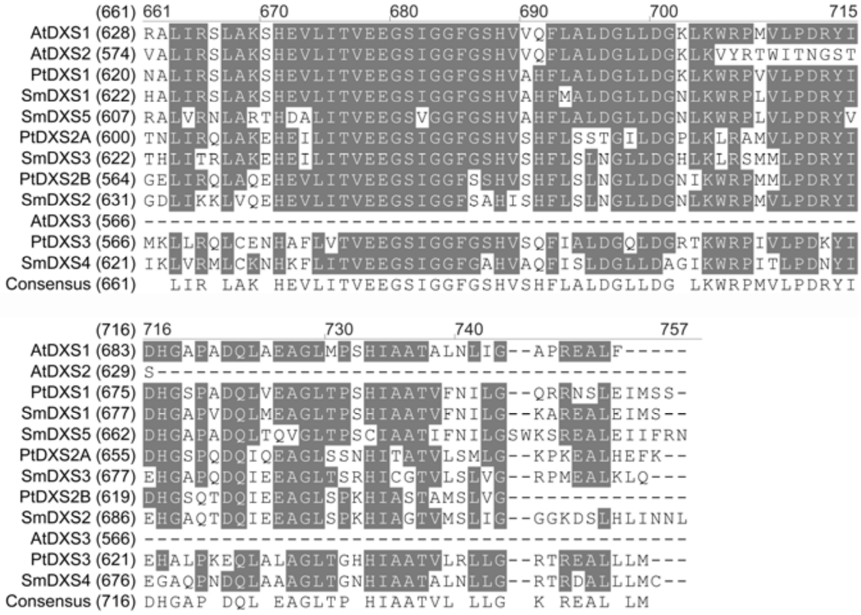
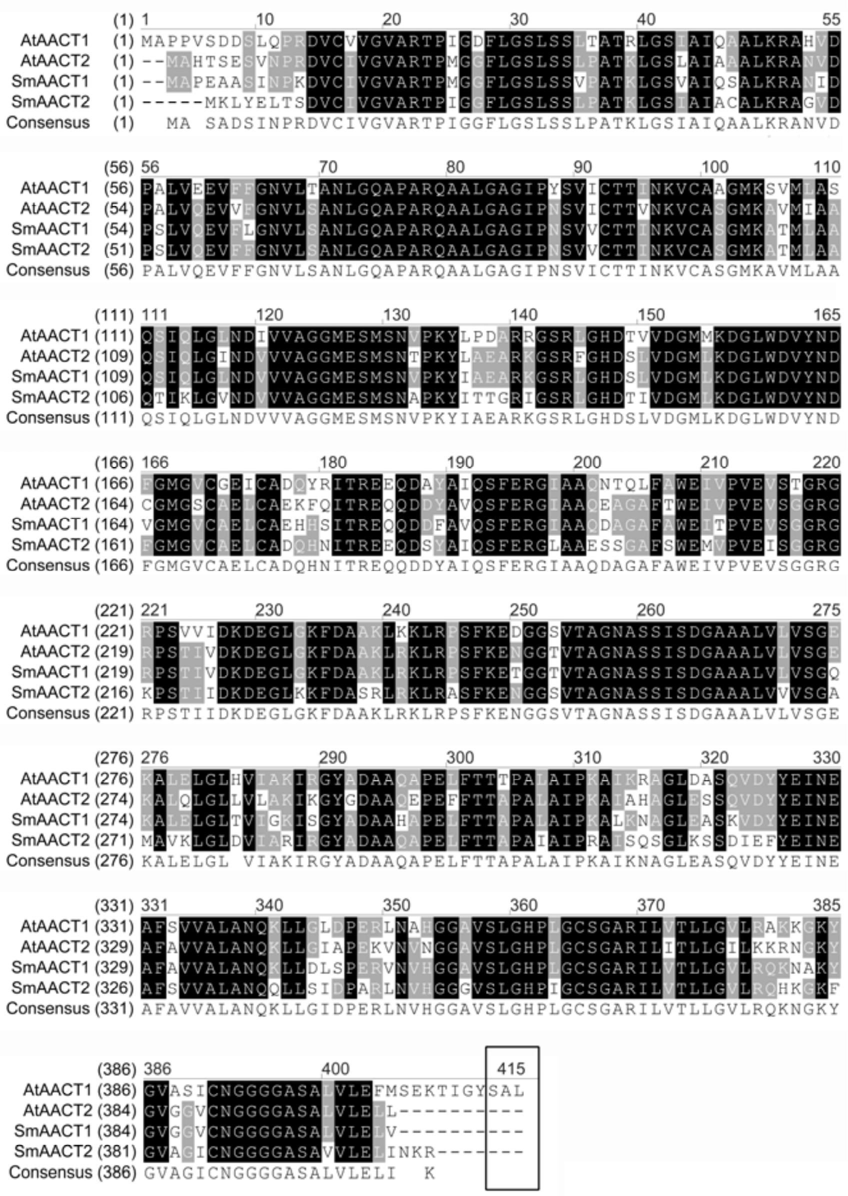
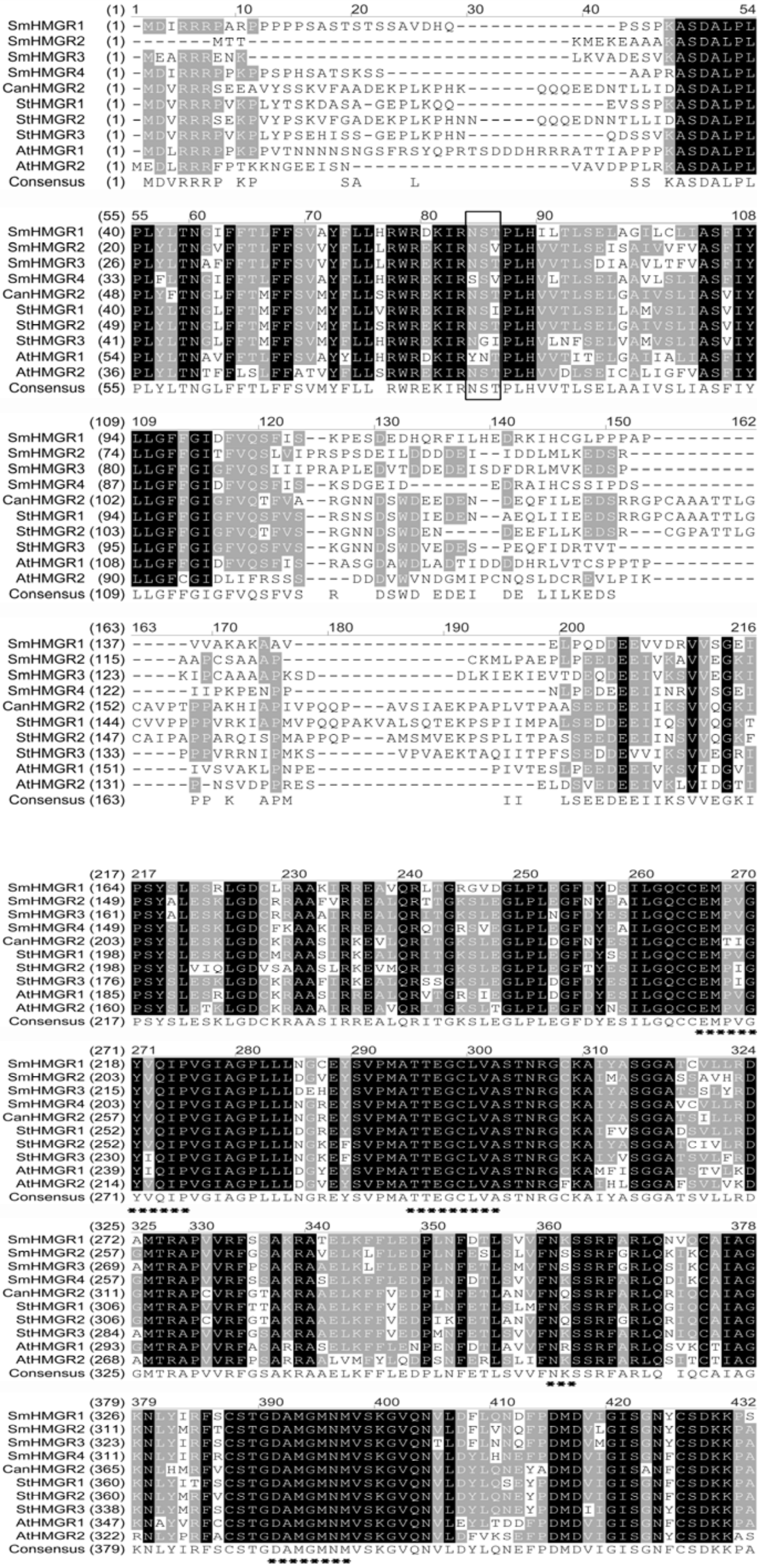


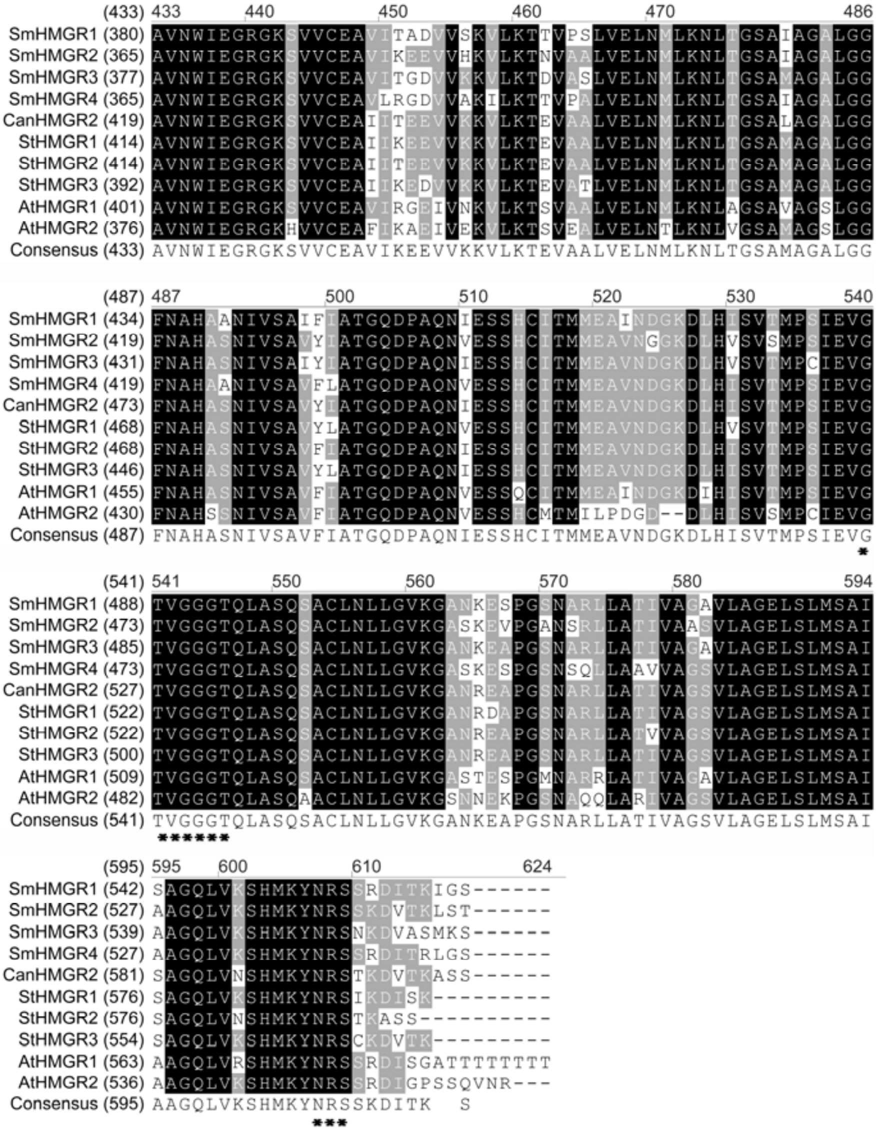
Figure S3. Sequence alignment of DXS proteins from *Salvia miltiorrhiza*, *Arabidopsis thaliana* and *Populus trichocarpa*. DXSs included are *A. thaliana* AtDXS1 (At4g15560), AtDXS2 (At3g21500), AtDXS3 (At 5G11380), *P. trichocarpa* PtDXS1(XP_002312717), PtDXS2A (XP_002303416), PtDXS2B (XP_002331678), PtDXS3 (XP_002308644), *S. miltiorrhiza* SmDXS1 (ACF21004), SmDXS2 (ACQ66107), SmDXS3 (JN831116), SmDXS4 (JN831117), SmDXS5 (JN831118). The consensus thiamine pyrophosphatase - binding motif (GDGAMTAG……VILND) and the pyridine binding domain (DRAGLVGAD……VMAPSD) was indicated by '***'.



```
            (1) 1        10        20        30        40            55
AtAACT1     (1) MAPPVSDDSLQPRDVCVVGVARTPIGDFLGSLSSLTATRLGSIAIQAALKRAHVD
AtAACT2     (1) --MAHTSESVNPRDVCIVGVARTPMGGFLGSLSSLPATKLGSLAIAAALKRANVD
SmAACT1     (1) --MAPEAASINPKDVCIVGVARTPMGGFLGSLSSVPATKLGSVAIQSALKRANID
SmAACT2     (1) -----MKLYELTSDVCIVGVARTPIGGFLGSLSSLPATKLGSIAIACALKRAGVD
Consensus   (1)   MA SADSINPRDVCIVGVARTPIGGFLGSLSSLPATKLGSIAIQAALKRANVD

           (56) 56            70        80        90        100     110
AtAACT1    (56) PALVEEVFFGNVLTANLGQAPARQAALGAGIPYSVICTTINKVCAAGMKSVMLAS
AtAACT2    (54) PALVQEVVFGNVLSANLGQAPARQAALGAGIPNSVICTTVNKVCASGMKAVMIAA
SmAACT1    (54) PSLVQEVFLGNVLSANLGQAPARQAALGAGIPNSVVCTTINKVCASGMKATMLAA
SmAACT2    (51) PSLVQEVFFGNVLSANLGQAPARQAALGAGIPNSVVCTTINKVCASGMKATMLAA
Consensus  (56) PALVQEVFFGNVLSANLGQAPARQAALGAGIPNSVICTTINKVCASGMKAVMLAA

          (111) 111      120       130       140       150          165
AtAACT1   (111) QSIQLGLNDIVVAGGMESMSNVPKYLPDARRGSRLGHDTVVDGMMKDGLWDVYND
AtAACT2   (109) QSIQLGINDVVVAGGMESMSNTPKYLAEARKGSRFGHDSLVDGMLKDGLWDVYND
SmAACT1   (109) QSIQLGLNDVVVAGGMESMSNVPKYIAEARKGSRLGHDSLVDGMLKDGLWDVYND
SmAACT2   (106) QTIKLGVNDVVVAGGMESMSNAPKYITTGRIGSRLGHDTIVDGMLKDGLWDVYND
Consensus (111) QSIQLGLNDVVVAGGMESMSNVPKYIAEARKGSRLGHDSLVDGMLKDGLWDVYND

          (166) 166           180       190       200       210     220
AtAACT1   (166) FGMGVCGEICADQYRITREEQDAYAIQSFERGIAAQNTQLFAWEIVPVEVSTGRG
AtAACT2   (164) CGMGSCAELCAEKFQITREQQDDYAVQSFERGIAAQEAGAFTWEIVPVEVSGGRG
SmAACT1   (164) VGMGVCAELCAEHHSITREQQDDFAVQSFERGIAAQDAGAFAWEITPVEVSGGRG
SmAACT2   (161) FGMGVCAELCADQHNITREEQDSYAIQSFERGLAAESSGAFSWEMVPVEISGGRG
Consensus (166) FGMGVCAELCADQHNITREQQDDYAIQSFERGIAAQDAGAFAWEIVPVEVSGGRG

          (221) 221      230       240       250       260          275
AtAACT1   (221) RPSVVIDKDEGLGKFDAAKLKKLRPSFKEDGGSVTAGNASSISDGAAALVLVSGE
AtAACT2   (219) RPSTIVDKDEGLGKFDAAKLRKLRPSFKENGGTVTAGNASSISDGAAALVLVSGE
SmAACT1   (219) RPSTIVDKDEGLGKFDAAKLRKLRPSFKETGGTVTAGNASSISDGAAALVLVSGQ
SmAACT2   (216) KPSTIIDKDEGLKKFDASRLRKLRASFKENGGSVTAGNASSISDGAAALVVVSGA
Consensus (221) RPSTIIDKDEGLGKFDAAKLRKLRPSFKENGGSVTAGNASSISDGAAALVLVSGE

          (276) 276           290       300       310       320     330
AtAACT1   (276) KALELGLHVIAKIRGYADAAQAPELFTTTPALAIPKAIKRAGLDASQVDYYEINE
AtAACT2   (274) KALQLGLLVLAKIKGYGDAAQEPEFFTTAPALAIPKAIAHAGLESSQVDYYEINE
SmAACT1   (274) KALELGLTVIGKISGYADAAHAPELFTTAPALAIPKALKNAGLEASKVDYYEINE
SmAACT2   (271) MAVKLGLDVIARIRGYADAAQAPELFTTAPAIAIPRAISQSGLKSSDIEFYEINE
Consensus (276) KALELGL VIAKIRGYADAAQAPELFTTAPALAIPKAIKNAGLEASQVDYYEINE

          (331) 331      340       350       360       370          385
AtAACT1   (331) AFSVVALANQKLLGLDPERLNAHGGAVSLGHPLGCSGARILVTLLGVLRAKKGKY
AtAACT2   (329) AFAVVALANQKLLGIAPEKVNVNGGAVSLGHPLGCSGARILITLLGILKKRNGKY
SmAACT1   (329) AFAVVALANQKLLDLSPERVNVHGGAVSLGHPLGCSGARILVTLLGVLRQKNAKY
SmAACT2   (326) AFSVVALANQQLLSIDPARLNVHGGGVSLGHPIGCSGARILVTLLGVLRQHKGKF
Consensus (331) AFAVVALANQKLLGIDPERLNVHGGAVSLGHPLGCSGARILVTLLGVLRQKNGKY

          (386) 386           400          415
AtAACT1   (386) GVASICNGGGGASALVLEFMSEKTIGYSAL
AtAACT2   (384) GVGGVCNGGGGASALVLELL----------
SmAACT1   (384) GVGGVCNGGGGASALVLELV----------
SmAACT2   (381) GVAGICNGGGGASAVVLELINKR-------
Consensus (386) GVAGICNGGGGASALVLELI   K
```

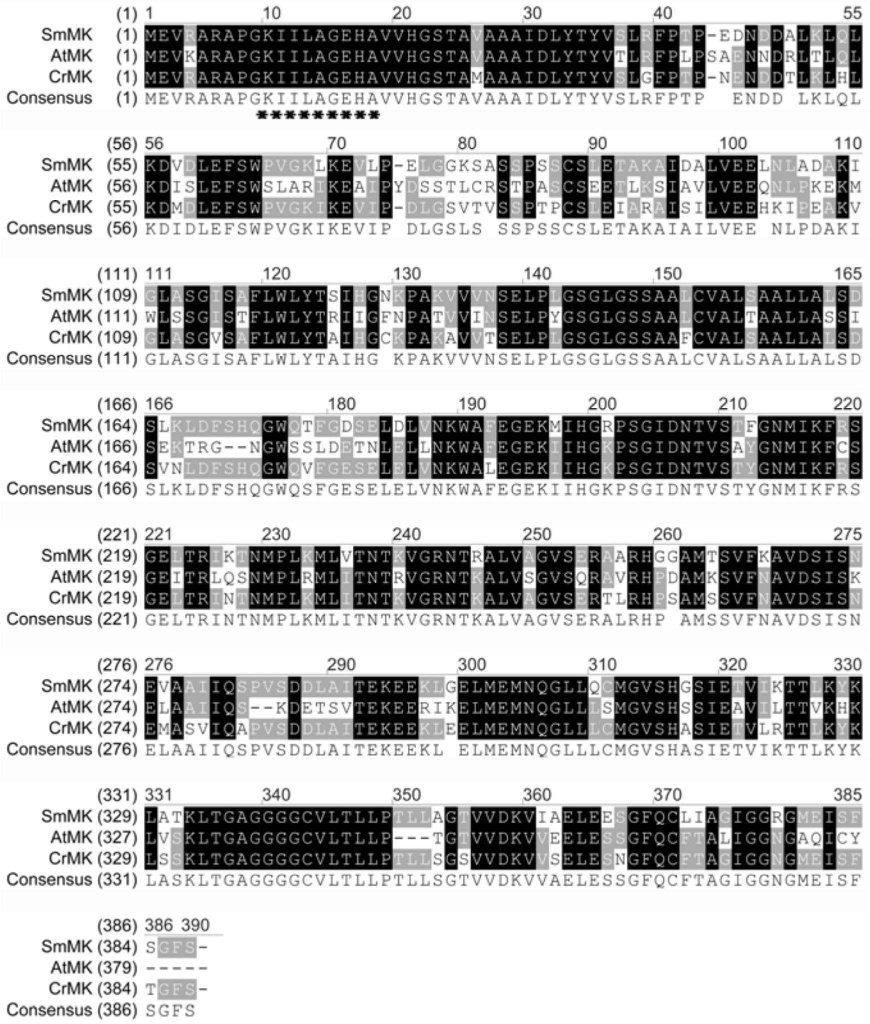
Figure S4. Sequence alignment of AACT proteins from *S. miltiorrhiza* and *A. thaliana*. AACTs included are *A. thaliana* AtAACT1 (At5g47720), AtAACT2 (At5g48230), *S. miltiorrhiza* SmAACT1 (ABV08820), SmAACT2 (JN831101). The peroxisomal targeting signal 1 (PTS1) related motif (SAL) presented in the carboxy terminus of AtAACT1 is boxed.


(1) 1 10 20 30 40 54
SmHMGR1 (1) -MDIRRRPARPPPPPSASTSTSSAVDHQ--------------PSSPKASDALPL
SmHMGR2 (1) -------MTT----------------------------KMEKEAAAKASDALPL
SmHMGR3 (1) -MEARRRENK----------------------------LKVADESVKASDALPL
SmHMGR4 (1) -MDIRRRPPKPPSPHSATSKSS---------------------AAPRASDALPL
CanHMGR2 (1) -MDVRRRSEEAVYSSKVFAADEKPLKPHK------QQQEEDNTLLIDASDALPL
StHMGR1 (1) -MDVRRRPVKPLYTSKDASA-GEPLKQQ-------------EVSSPKASDALPL
StHMGR2 (1) -MDVRRRSEKPVYPSKVFGADEKPLKPHNN-----QQQEDNNTLLIDASDALPL
StHMGR3 (1) -MDVRRRPVKPLYPSEHISS-GEPLKPHN------------QDSSVKASDALPL
AtHMGR1 (1) -MDLRRRPPKPPVTNNNNSNGSFRSYQPRTSDDDHRRRATTIAPPPKASDALPL
AtHMGR2 (1) MEDLRRRFPTKKNGEEISN-------------------VAVDPPLRKASDALPL
Consensus (1) MDVRRRP KP SA L SS KASDALPL
(55) 55 60 70 80 90 108
SmHMGR1 (40) PLYLTNGIFFTLFFSVAYFLLHRWRDKIRNSTPLHILTLSELAGILCLIASFIY
SmHMGR2 (20) PLYLTNGVFFTLFFSVVYFLLLRWREKIRNSVPLHVVTLSEISAIVVFVASFIY
SmHMGR3 (26) PLYLTNAFFFTLFFSVVYFLLLRWREKIRNSTPLHVVTLSDIAAVLTFVASFIY
SmHMGR4 (33) PLFLTNGIFFTLFFSVAYFLLHRWREKIRSSVPLHVLTLSELAAVLSLIASFIY
CanHMGR2 (48) PLYFTNGLFFTMFFSVMYFLLSRWREKIRNSTPLHVVTLSELGAIVSLIASVIY
StHMGR1 (40) PLYLTNGLFFTMFFSVMYFLLVRWREKIRNSIPLHVVTLSELLAMVSLIASVIY
StHMGR2 (49) PLYLTNGLFFTMFFSVMYFLLSRWREKIRNSTPLHVVTLSELGAIVSLIASVIY
StHMGR3 (41) PLYLTNGLFFTMFFSVMYFLLHRWREKIRNGIPLHVLNFSELVAMVSLIASVIY
AtHMGR1 (54) PLYLTNAVFFTLFFSVAYYLLHRWRDKIRYNTPLHVVTITELGAIIALIASFIY
AtHMGR2 (36) PLYLTNTFFLSLFFATVYFLLSRWREKIRNSTPLHVVDLSEICALIGFVASFIY
Consensus (55) PLYLTNGLFFTLFFSVMYFLL RWREKIRNSTPLHVVTLSELAAIVSLIASFIY
(109) 109 120 130 140 150 162
SmHMGR1 (94) LLGFFGIDFVQSFIS--KPESDEDHQRFILHEDRKIHCGLPPPAP---------
SmHMGR2 (74) LLGFFGITFVQSLVIPRSPSDEILDDDDEI--IDDLMLKEDSR-----------
SmHMGR3 (80) LLGFFGIGFVQSIIIPRAPLEDVTDDEDEISDFDRLMVKEDSP-----------
SmHMGR4 (87) LLGFFGIDFVQSFIS--KSDGEID-------EDRAIHCSSIPDS----------
CanHMGR2 (102) LLGFFGIGFVQTFVA--RGNNDSWDEEDEN--DEQFILEEDSRRGPCAAATTLG
StHMGR1 (94) LLGFFGIGFVQSFVS--RSNSDSWDIEDEN--AEQLIIEEDSRRGPCAAATTLG
StHMGR2 (103) LLGFFGIGFVQTFVS--RGNNDSWDEN-----DEEFLLKEDSR---CGPATTLG
StHMGR3 (95) LLGFFGIGFVQSFVS--KGNNDSWDVEDES--PEQFIDRTVT------------
AtHMGR1 (108) LLGFFGIDFVQSFIS--RASGDAWDLADTIDDDDHRLVTCSPPTP---------
AtHMGR2 (90) LLGFCGIDLIFRSSS----DDDVWVNDGMIPCNQSLDCREVLPIK---------
Consensus (109) LLGFFGIGFVQSFVS R DSWD EDEI DE LILKEDS
(163) 163 170 180 190 200 216
SmHMGR1 (137) -----VVAKAKAAV----------------------ELPQDDEEVVDRVVSGEI
SmHMGR2 (115) ----AAPCSAAAP----------------CKMLPAEPLPEEDEEIVKAVVEGKI
SmHMGR3 (123) ----KIPCAAAAPKSD------------DLKIEKIEVTDEQDEEIVKSVVEGKI
SmHMGR4 (122) -----IIPKPENPP----------------------NLPEDEEEIINRVVSGEI
CanHMGR2 (152) CAVPTPPAKHIAPIVPQQP---AVSIAEKPAPLVTPAASEEDEEIIKSVVQGKI
StHMGR1 (144) CVVPPPPVRKIAPMVPQQPAKVALSQTEKPSPIIMPALSEDDEEIIQSVVQGKT
StHMGR2 (147) CAIPAPPARQISPMAPPQP---AMSMVEKPSPLITPASSEEDEEIINSVVQGKF
StHMGR3 (133) ----PPPVRRNIPMKS-------VPVAEKTAQIITPFSSEDDEVVIKSVVEGRI
AtHMGR1 (151) -----IVSVAKLPNPE---------------PIVTESLPEEDEEIVKSVIDGVI
AtHMGR2 (131) -----P-NSVDPPRES-------------------ELDSVEDEEIVKLVIDGTI
Consensus (163) PP K APM II LSEEDEEIIKSVVEGKI
(217) 217 230 240 250 260 270
SmHMGR1 (164) PSYSLESRLGDCLRAAKIRREAVQRLTGRGVDGLPLEGFDYDSILGQCCEMPVG
SmHMGR2 (149) PSYALESKLGDCRRAAFVRREALQRTTGKSLEGLPLEGFNYEAILGQCCEMPVG
SmHMGR3 (161) PSYALESKLGDCRRAAAIRREALQRITGKSLEGLPLNGFDYESILGQCCEMPVG
SmHMGR4 (149) PSYSLESKLGDCFKAAKIRREALQRQTGRSVEGLPLEGFDYEAILGQCCEMPVG
CanHMGR2 (203) PSYSLESKLGDCKRAASIRKEVLQRITGKSLEGLPLDGFNYESILGQCCEMTIG
StHMGR1 (198) PSYSLESKLGDCMRAASIRKEALQRITGKSLEGLPLEGFDYSSILGQCCEMPVG
StHMGR2 (198) PSYSLVIQLGDVSAAASLRKEVMQRITGKSLEGLPLEGFTYESILGQCCEMPIG
StHMGR3 (176) PSYSLESKLGDCKRAAFIRKEALQRSSGKSLEGLPLDGFDYESILGQCCEMPIG
AtHMGR1 (185) PSYSLESRLGDCKRAASIRREALQRVTGRSIEGLPLDGFDYESILGQCCEMPVG
AtHMGR2 (160) PSYSLETKLGDCKRAAAIRREAVQRITGKSLTGLPLEGFDYNSILGQCCEMPVG
Consensus (217) PSYSLESKLGDCKRAASIRREALQRITGKSLEGLPLEGFDYESILGQCCEMPVG
******
(271) 271 280 290 300 310 324
SmHMGR1 (218) YVQIPVGIAGPLLLNGCEYSVPMATTEGCLVASTNRGCKAIYASGGATCVLLRD
SmHMGR2 (203) YVQIPVGIAGPLLLDGVEYSVPMATTEGCLVASTNRGCKAIMASGGASSAVHRD
SmHMGR3 (215) YVQIPVGIAGPLLLDEHEYSVPMATTEGCLVASTNRGCKAIYASGGATSSLYRD
SmHMGR4 (203) YVQIPVGIAGPLLLNGREYSVPMATTEGCLVASTNRGCKAIYASGGAVCVLLRD
CanHMGR2 (257) YVQIPVGIAGPLLLNGREYSVPMATTEGCLVASTNRGCKAIYASGGATSILLRD
StHMGR1 (252) YVQIPVGIAGPLLLDGREYSVPMATTEGCLVASTNRGCKAIFVSGGADSVLLRD
StHMGR2 (252) YVQIPVGIAGPLLLNGKEFSVPMATTEGCLVASTNRGCKAIYASGGATCIVLRD
StHMGR3 (230) YIQIPVGIAGPLLLNGKEFSVPMATTEGCLVASTNRGCKAIYVSGGATSVLFRD
AtHMGR1 (239) YIQIPVGIAGPLLLDGYEYSVPMATTEGCLVASTNRGCKAMFISGGATSTVLKD
AtHMGR2 (214) YVQIPVGIAGPLLLDGVEYSVPMATTEGCLVASTNRGFKAIHLSGGAFSVLVKD
Consensus (271) YVQIPVGIAGPLLLNGREYSVPMATTEGCLVASTNRGCKAIYASGGATSVLLRD
****** *********
(325) 325 330 340 350 360 378
SmHMGR1 (272) AMTRAPVVRFSSAKRATELKFFLEDPLNFDTLSVVFNKSSRFARLQNVQCAIAG
SmHMGR2 (257) GMTRAPVVRFGSAKRAVELKLFLEDPLNFESLSLVFNSSSRFGRLQKIKCAIAG
SmHMGR3 (269) AMTRAPVVRFPSAKRAAELKLFLEDPLNFETLSMVFNSSSRFGRLQSIKCAIAG
SmHMGR4 (257) GMTRAPVVRFSSAKRASELKFFLEDPLNFDTLSVVFNKSSRFARLQDIKCAIAG
CanHMGR2 (311) GMTRAPCVRFGTAKRAAELKFFVEDPINFETLANVFNQSSRFARLQRIQCAIAG
StHMGR1 (306) GMTRAPVVRFTTAKRAAELKFFVEDPLNFETLSLMFNKSSRFARLQGIQCAIAG
StHMGR2 (306) GMTRAPCVRFGTAKRAAELKFFVEDPIKFETLANVFNQSSRFGRLQRIQCAIAG
StHMGR3 (284) AMTRAPVVRFGSAKRAAELKFFVEDPMNFETLSVVFNKSSRFARLQNIQCAIAG
AtHMGR1 (293) GMTRAPVVRFASARRASELKFFLENPENFDTLAVVFNRSSRFARLQSVKCTIAG
AtHMGR2 (268) AMTRAPVVRFPSARRAALVMFYLQDPSNFERLSLIFNKSSRFARLQSITCTIAG
Consensus (325) GMTRAPVVRFGSAKRAAELKFFLEDPLNFETLSVVFNKSSRFARLQ IQCAIAG
***
(379) 379 390 400 410 420 432
SmHMGR1 (326) KNLYIRFSCSTGDAMGMNMVSKGVQNVLDFLQNDFPDMDVIGISGNYCSDKKPS
SmHMGR2 (311) KNLYMRFTCSTGDAMGMNMVSKGVQNVLDFLVNQFPDMDVLGISGNYCSDKKPA
SmHMGR3 (323) KNLYIRFSCSTGDAMGMNMVSKGVQNTLDFLNNQFPDMDVMGISGNYCSDKKPA
SmHMGR4 (311) KNLYIRFRCSTGDAMGMNMVSKGVQNVLDYLHNEFPDMDVIGISGNFCSDKKPA
CanHMGR2 (365) KNLHMRFVCSTGDAMGMNMVSKGVQNVLDYLQNEYADMDVIGISANFCSDKKPA
StHMGR1 (360) KNLYITFSCSTGDAMGMNMVSKGVQNVLDYLQSEYPDMDVIGISGNFCSDKKPA
StHMGR2 (360) KNLYMRFVCSTGDAMGMNMVSKGVQNVLDYLQNEYPDMDVIGISGNFCSDKKPA
StHMGR3 (338) KNLYMRFSCSTGDAMGMNMVSKGVQNVLDYLQNEYPDMDIIGISGNYCSDKKPA
AtHMGR1 (347) KNAYVRFCCSTGDAMGMNMVSKGVQNVLEYLTDDFPDMDVIGISGNFCSDKKPA
AtHMGR2 (322) RNLYPRFACSTGDAMGMNMVSKGVQNVLDFVKSEFPDMDVIGISGNYCSDKKAS
Consensus (379) KNLYIRFSCSTGDAMGMNMVSKGVQNVLDYLQNEFPDMDVIGISGNFCSDKKPA
****



```
          (433) 433    440       450       460       470           486
  SmHMGR1 (380) AVNWIEGRGKSVVCEAVITADVVSKVLKTTVPSLVELNMLKNLTGSAIAGALGG
  SmHMGR2 (365) AVNWIEGRGKSVVCEAVIKEEVVHKVLKTNVAALVELNMLKNLTGSAIAGALGG
  SmHMGR3 (377) AVNWIEGRGKSVVCEAVITGDVVKKVLKTDVASLVELNMLKNLTGSAMAGALGG
  SmHMGR4 (365) AVNWIEGRGKSVVCEAVLRGDVVAKILKTTVPALVELNMLKNLTGSAIAGALGG
 CanHMGR2 (419) AVNWIEGRGKSVVCEAIITEEVVKKVLKTEVAALVELNMLKNLTGSALAGALGG
  StHMGR1 (414) AVNWIEGRGKSVVCEAIIKEEVVKKVLKTEVAALVELNMLKNLTGSAMAGALGG
  StHMGR2 (414) AVNWIEGRGKSVVCEAIITEEVVKKVLKTEVAALVELNMLKNLTGSAMAGALGG
  StHMGR3 (392) AVNWIEGRGKSVVCEAIIKEDVVKKVLKTEVATLVELNMLKNLTGSAMAGALGG
  AtHMGR1 (401) AVNWIEGRGKSVVCEAVIRGEIVNKVLKTSVAALVELNMLKNLAGSAVAGSLGG
  AtHMGR2 (376) AVNWIEGRGKHVVCEAFIKAEIVEKVLKTSVEALVELNTLKNLVGSAMAGSLGG
Consensus (433) AVNWIEGRGKSVVCEAVIKEEVVKKVLKTEVAALVELNMLKNLTGSAMAGALGG

          (487) 487          500       510       520       530     540
  SmHMGR1 (434) FNAHAANIVSAIFIATGQDPAQNIESSHCITMMEAINDGKDLHISVTMPSIEVG
  SmHMGR2 (419) FNAHASNIVSAVYIATGQDPAQNVESSHCITMMEAVNGGKDLHVSVSMPSIEVG
  SmHMGR3 (431) FNAHASNIVSAIYIATGQDPAQNIESSHCITMMEAVNDGKDLHVSVTMPCIEVG
  SmHMGR4 (419) FNAHAANIVSAVFLATGQDPAQNVESSHCITMMEAVNDGKDLHISVTMPSIEVG
 CanHMGR2 (473) FNAHASNIVSAVYIATGQDPAQNIESSHCITMMEAVNDGKDLHISVTMPSIEVG
  StHMGR1 (468) FNAHASNIVSAVYLATGQDPAQNVESSHCITMMEAVNDGKDLHVSVTMPSIEVG
  StHMGR2 (468) FNAHASNIVSAVFIATGQDPAQNIESSHCITMMEAVNDGKDLHISVTMPSIEVG
  StHMGR3 (446) FNAHASNIVSAVYLATGQDPAQNIESSHCITMMEAVNDGKDLHISVTMPSIEVG
  AtHMGR1 (455) FNAHASNIVSAVFIATGQDPAQNVESSQCITMMEAINDGKDIHISVTMPSIEVG
  AtHMGR2 (430) FNAHSSNIVSAVFIATGQDPAQNVESSHCMTMILPDGD--DLHISVSMPCIEVG
Consensus (487) FNAHASNIVSAVFIATGQDPAQNIESSHCITMMEAVNDGKDLHISVTMPSIEVG
                                                                     *

          (541) 541      550       560       570       580         594
  SmHMGR1 (488) TVGGGTQLASQSACLNLLGVKGANKESPGSNARLLATIVAGAVLAGELSLMSAI
  SmHMGR2 (473) TVGGGTQLASQSACLNLLGVKGASKEVPGANSRLLATIVAASVLAGELSLMSAI
  SmHMGR3 (485) TVGGGTQLASQSACLNLLGVKGANKEAPGSNARLLATIVAGAVLAGELSLMSAI
  SmHMGR4 (473) TVGGGTQLASQSACLNLLGVKGASKESPGSNSQLLAAVVAGSVLAGELSLMSAI
 CanHMGR2 (527) TVGGGTQLASQSACLNLLGVKGANREAPGSNARLLATIVAGSVLAGELSLMSAI
  StHMGR1 (522) TVGGGTQLASQSACLNLLGVKGANRDAPGSNARLLATIVAGSVLAGELSLMSAI
  StHMGR2 (522) TVGGGTQLASQSACLNLLGVKGANREAPGSNARLLATVVAGSVLAGELSLMSAI
  StHMGR3 (500) TVGGGTQLASQSACLNLLGVKGANREAPGSNARLLATIVAGSVLAGELSLMSAI
  AtHMGR1 (509) TVGGGTQLASQSACLNLLGVKGASTESPGMNARRLATIVAGAVLAGELSLMSAI
  AtHMGR2 (482) TVGGGTQLASQAACLNLLGVKGSNNEKPGSNAQQLARIVAGSVLAGELSLMSAI
Consensus (541) TVGGGTQLASQSACLNLLGVKGANKEAPGSNARLLATIVAGSVLAGELSLMSAI
                *******

          (595) 595  600       610         624
  SmHMGR1 (542) SAGQLVKSHMKYNRSSRDITKIGS------
  SmHMGR2 (527) AAGQLVKSHMKYNRSSKDVTKLST------
  SmHMGR3 (539) AAGQLVKSHMKYNRSNKDVASMKS------
  SmHMGR4 (527) AAGQLVKSHMKYNRSSRDITRLGS------
 CanHMGR2 (581) SAGQLVNSHMKYNRSTKDVTKASS------
  StHMGR1 (576) SAGQLVKSHMKYNRSIKDISK---------
  StHMGR2 (576) SAGQLVNSHMKYNRSTKASS----------
  StHMGR3 (554) SAGQLVKSHMKYNRSCKDVTK---------
  AtHMGR1 (563) AAGQLVRSHMKYNRSSRDISGATTTTTTTT
  AtHMGR2 (536) AAGQLVKSHMKYNRSSRDIGPSSQVNR---
Consensus (595) AAGQLVKSHMKYNRSSKDITK  S
                            ***
```

Figure S5. Sequence alignment of HMGR proteins from *S. miltiorrhiza* and various other plants. Proteins included are *Arabidopsis thaliana* AtHMGR1 (CAA33139), AtHMGR2 (AAA67317), *Solanum tuberosum* StHMGR1 (AAA93498), StHMGR2 (AAB52551), StHMGR3 (AAB52552), *Capsicum annuum* CanHMGR2 (AAD28179), and four *S. miltiorrhiza* SmHMGRs. The putative HMG-CoA-binding sites (EMPIGYVQIP and TTEGCLVA), NADP(H)-binding sites (DAMGMNM and GTVGGGT) and *N*-linked glycosylation sites (N-X-S/T) are indicated by '***'. The *N*-glycosylation site (NST) is boxed.



```
            (1) 1        10        20        30        40            55
       SmMK (1) MEVRARAPGKIILAGEHAVVHGSTAVAAAIDLYTYVSLRFPTP-EDNDDALKLQL
       AtMK (1) MEVKARAPGKIILAGEHAVVHGSTAVAAAIDLYTYVTLRFPLPSAENNDRLTLQL
       CrMK (1) MEVRARAPGKIILAGEHAVVHGSTAMAAAIDLYTYVSLGFPTP-NENDDTLKLHL
  Consensus (1) MEVRARAPGKIILAGEHAVVHGSTAVAAAIDLYTYVSLRFPTP  ENDD LKLQL
                         *********

           (56) 56            70        80        90        100     110
      SmMK (55) KDVDLEFSWPVGKLKEVLP-ELGGKSASSPSSCSLETAKAIDALVEELNLADAKI
      AtMK (56) KDISLEFSWSLARIKEAIPYDSSTLCRSTPASCSEETLKSIAVLVEEQNLPKEKM
      CrMK (55) KDMDLEFSWPVGKIKEVIP-DLGSVTVSSPTPCSLEIARAISILVEEHKIPEAKV
 Consensus (56) KDIDLEFSWPVGKIKEVIP DLGSLS SSPSSCSLETAKAIAILVEE NLPDAKI

          (111) 111      120       130       140       150          165
     SmMK (109) GLASGISAFLWLYTSIHGNKPAKVVVNSELPLGSGLGSSAALCVALSAALLALSD
     AtMK (111) WLSSGISTFLWLYTRIIGFNPATVVINSELPYGSGLGSSAALCVALTAALLASSI
     CrMK (109) GLASGVSAFLWLYTAIHGCKPAKAVVTSELPLGSGLGSSAAFCVALSAALLALSD
Consensus (111) GLASGISAFLWLYTAIHG KPAKVVVNSELPLGSGLGSSAALCVALSAALLALSD

          (166) 166           180       190       200       210     220
     SmMK (164) SLKLDFSHQGWQTFGDSELDLVNKWAFEGEKMIHGRPSGIDNTVSTFGNMIKFRS
     AtMK (166) SEKTRG--NGWSSLDETNLELLNKWAFEGEKIIHGKPSGIDNTVSAYGNMIKFCS
     CrMK (164) SVNLDFSHQGWQVFGESELELVNKWALEGEKIIHGKPSGIDNTVSTYGNMIKFRS
Consensus (166) SLKLDFSHQGWQSFGESELELVNKWAFEGEKIIHGKPSGIDNTVSTYGNMIKFRS

          (221) 221      230       240       250       260          275
     SmMK (219) GELTRIKTNMPLKMLVTNTKVGRNTRALVAGVSERAARHGGAMTSVFKAVDSISN
     AtMK (219) GEITRLQSNMPLRMLITNTRVGRNTKALVSGVSQRAVRHPDAMKSVFNAVDSISK
     CrMK (219) GELTRINTNMPLKMLITNTKVGRNTKALVAGVSERTLRHPSAMSSVFNAVDSISN
Consensus (221) GELTRINTNMPLKMLITNTKVGRNTKALVAGVSERALRHP AMSSVFNAVDSISN

          (276) 276           290       300       310       320     330
     SmMK (274) EVAAIIQSPVSDDLAITEKEEKLGELMEMNQGLLQCMGVSHGSIETVIKTTLKYK
     AtMK (274) ELAAIIQS--KDETSVTEKEERIKELMEMNQGLLLSMGVSHSSIEAVILTTVKHK
     CrMK (274) EMASVIQAPVSDDLAITEKEEKLEELMEMNQGLLLCMGVSHASIETVLRTTLKYK
Consensus (276) ELAAIIQSPVSDDLAITEKEEKL ELMEMNQGLLLCMGVSHASIETVIKTTLKYK

          (331) 331      340       350       360       370          385
     SmMK (329) LATKLTGAGGGGCVLTLLPTLLAGTVVDKVIAELEESGFQCLIAGIGGRGMEISF
     AtMK (327) LVSKLTGAGGGGCVLTLLP---TGTVVDKVVEELESSGFQCFTALIGGNGAQICY
     CrMK (329) LSSKLTGAGGGGCVLTLLPTLLSGSVVDKVVSELESNGFQCFTAGIGGNGMEISF
Consensus (331) LASKLTGAGGGGCVLTLLPTLLSGTVVDKVVAELESSGFQCFTAGIGGNGMEISF

          (386) 386 390
     SmMK (384) SGFS-
     AtMK (379) -----
     CrMK (384) TGFS-
Consensus (386) SGFS
```

Figure S6. Sequence alignment of MK proteins from *S. miltiorrhiza*, *Arabidopsis thaliana* and *Catharanthus roseus*. Proteins included are *A. thaliana* AtMK (NP_198097), *C. roseus* CrMK (ADR65111), and *S. miltiorrhiza* SmMK (JN831104). The peroxisomal targeting signal 2 (PTS2) related nonapeptide (KIILAGEHA) is indicated by '***'.

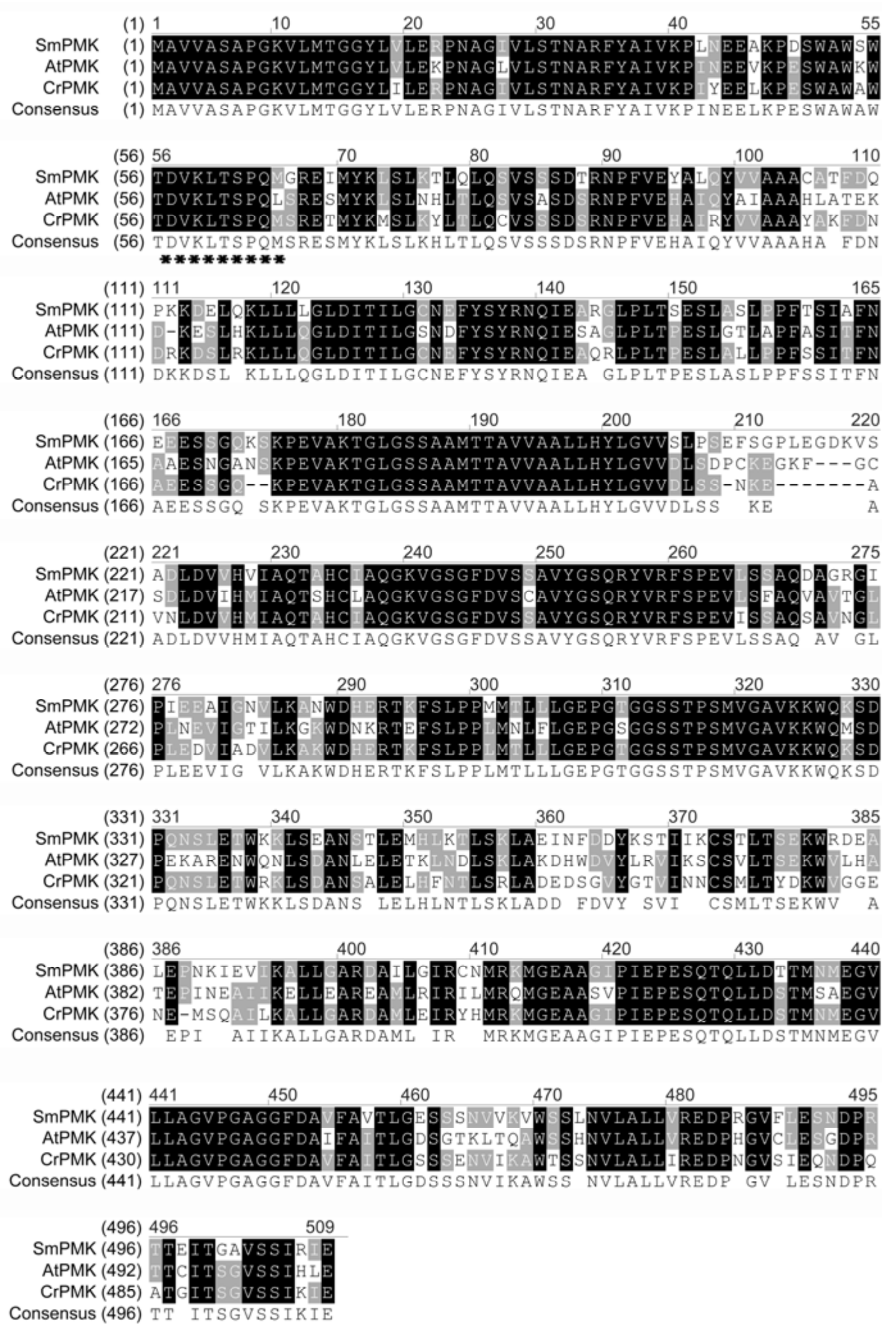


Figure S7. Sequence alignment of PMK proteins from *S. miltiorrhiza*, *Arabidopsis* thaliana and *Catharanthus roseus*. Proteins included are *A. thaliana* AtPMK (NP_174473), *C. roseus* CrPMK (ADR65112), and *S. miltiorrhiza* SmPMK (JN831095). The PTS2 related nonapeptide (DVKLTSPQL/M) is indicated by ‘***’.

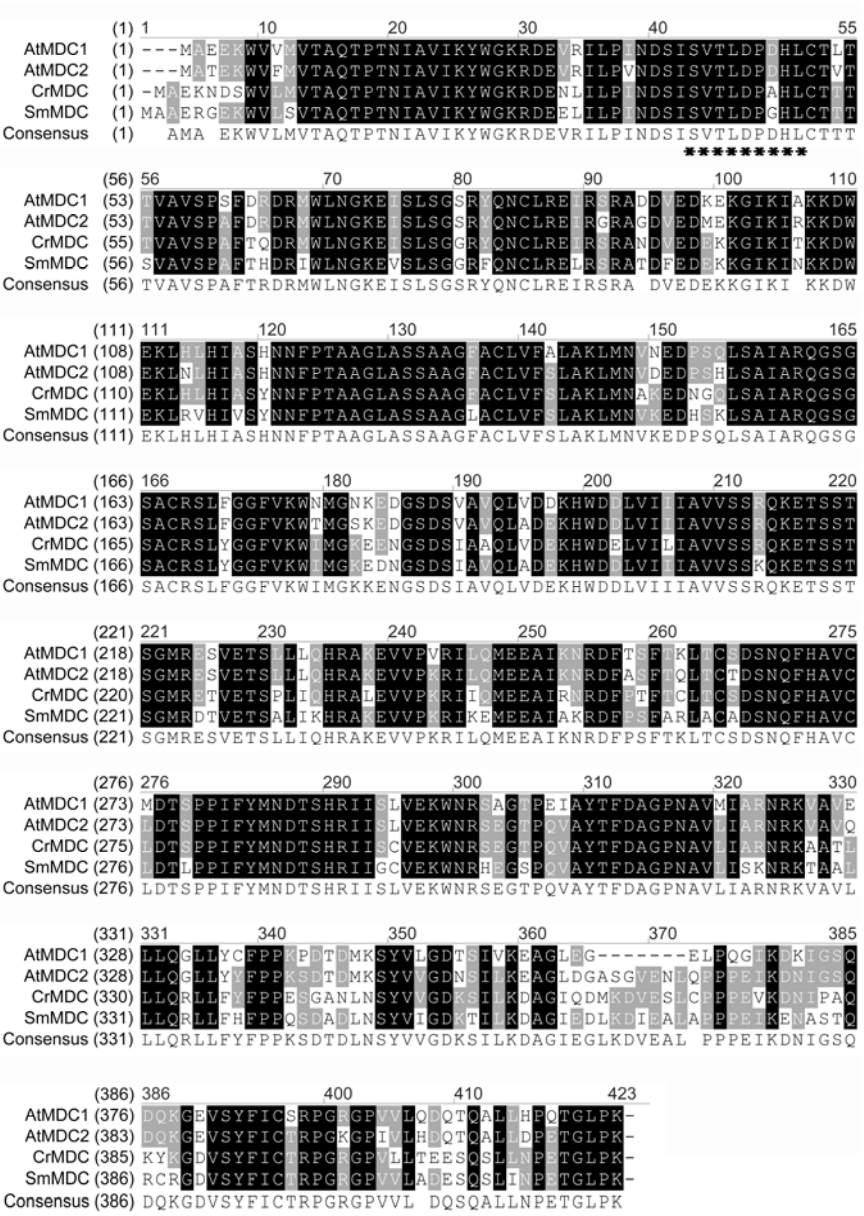


Figure S8. Sequence alignment of MDC proteins from *S. miltiorrhiza*, *Arabidopsis thaliana* and *Catharanthus roseus*. Proteins included are *A. thaliana* AtMDC (NP_174473), *C. roseus* CrMDC (ADR65113), and *S. miltiorrhiza* SmMDC (JN831105). The PTS2 related nonapeptide (SVTLDPXHL) is indicated by ‘***’.

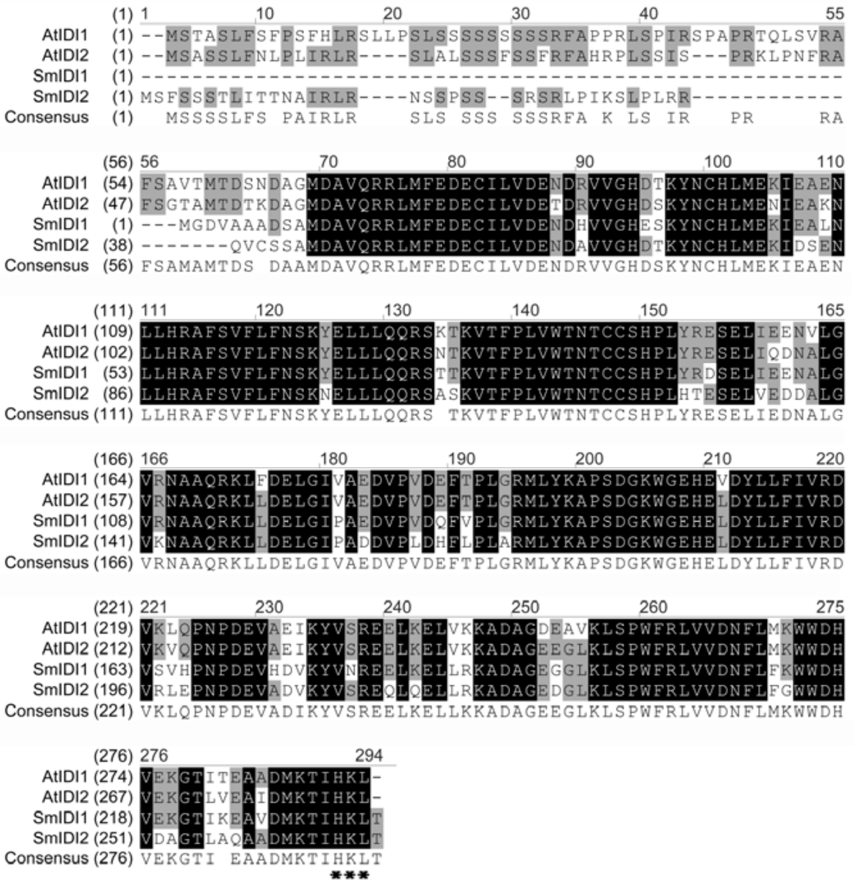
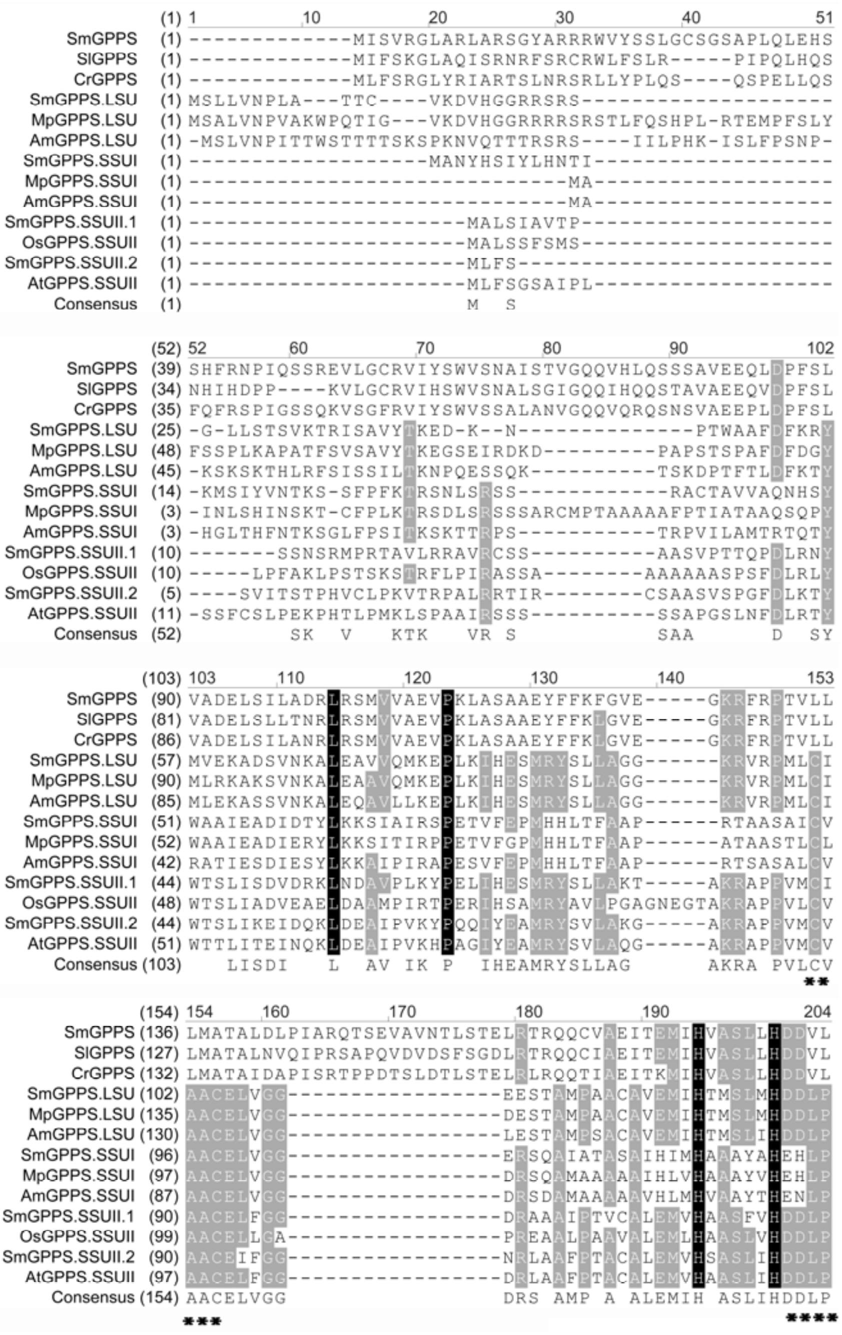
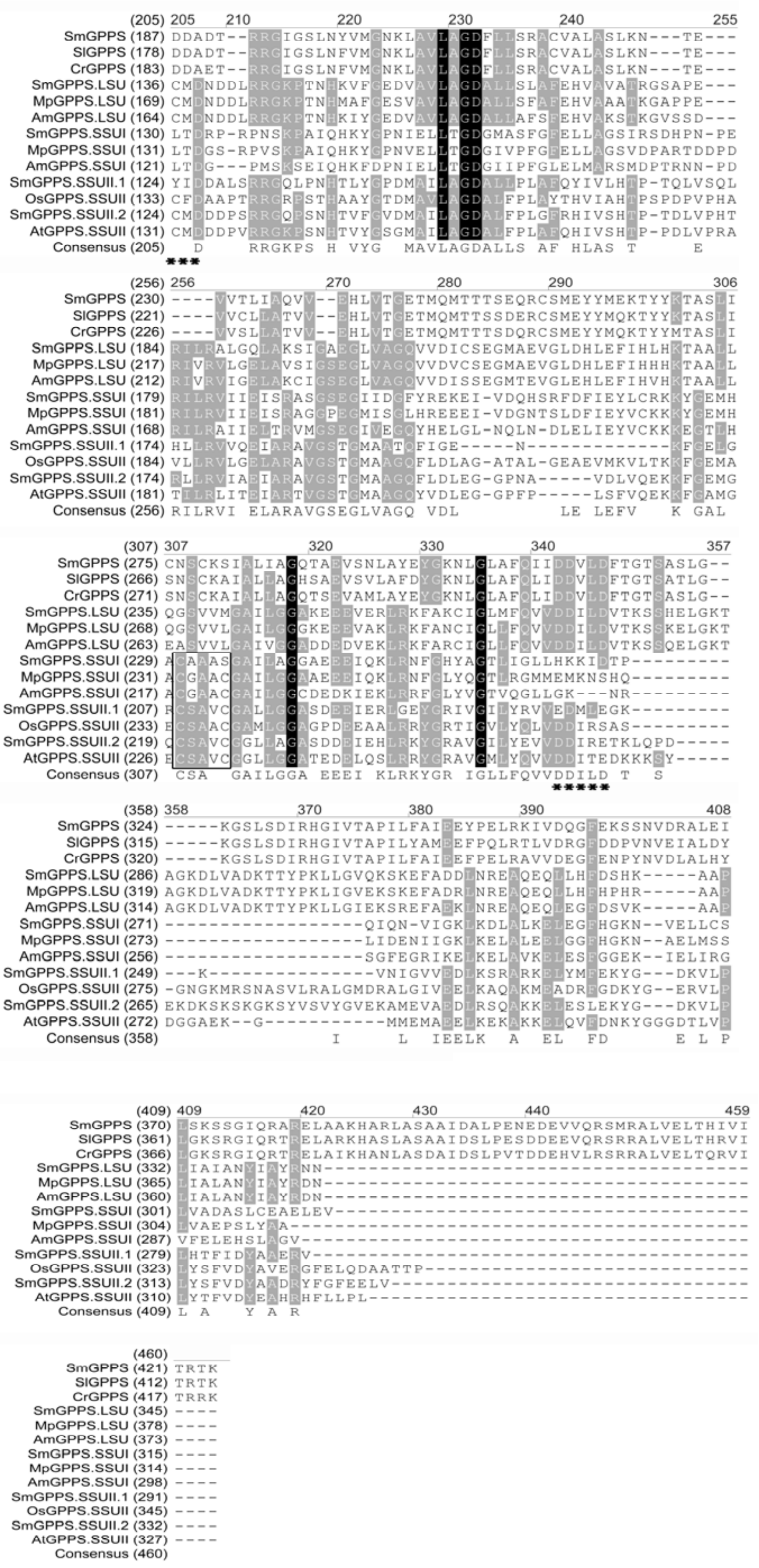


Figure S9. Sequence alignment of IDI proteins from *S. miltiorrhiza* and *Arabidopsis thaliana*. Proteins included are *A. thaliana* AtIDI1 (NP_197148), AtIDI2 (NP_186927), and two *S. miltiorrhiza* SmIDIs.The PTS1 related motif (HKL) is indicated by '***'.



```
                  (1) 1        10        20        30        40         51
SmGPPS            (1) -------------MISVRGLARLARSGYARRRWVYSSLGCSGSAPLQLEHS
SlGPPS            (1) -------------MIFSKGLAQISRNRFSRCRWLFSLR-----PIPQLHQS
CrGPPS            (1) -------------MLFSRGLYRIARTSLNRSRLLYPLQS----QSPELLQS
SmGPPS.LSU        (1) MSLLVNPLA---TTC----VKDVHGGRRSRS--------------------
MpGPPS.LSU        (1) MSALVNPVAKWPQTIG---VKDVHGGRRRRSRSTLFQSHPL-RTEMPFSLY
AmGPPS.LSU        (1) -MSLVNPITTWSTTTTSKSPKNVQTTTRSRS----IILPHK-ISLFPSNP-
SmGPPS.SSUI       (1) -------------------MANYHSIYLHNTI-------------------
MpGPPS.SSUI       (1) ------------------------------MA-------------------
AmGPPS.SSUI       (1) ------------------------------MA-------------------
SmGPPS.SSUII.1    (1) ----------------------MALSIAVTP--------------------
OsGPPS.SSUII      (1) ----------------------MALSSFSMS--------------------
SmGPPS.SSUII.2    (1) ----------------------MLFS-------------------------
AtGPPS.SSUII      (1) ----------------------MLFSGSAIPL-------------------
Consensus         (1)                       M  S

                 (52) 52      60        70        80        90          102
SmGPPS           (39) SHFRNPIQSSREVLGCRVIYSWVSNAISTVGQQVHLQSSSAVEEQLDPFSL
SlGPPS           (34) NHIHDPP----KVLGCRVIHSWVSNALSGIGQQIHQQSTAVAEEQVDPFSL
CrGPPS           (35) FQFRSPIGSSQKVSGFRVIYSWVSSALANVGQQVQRQSNSVAEEPLDPFSL
SmGPPS.LSU       (25) -G-LLSTSVKTRISAVYTKED-K--N--------------PTWAAFDFKRY
MpGPPS.LSU       (48) FSSPLKAPATFSVSAVYTKEGSEIRDKD---------PAPSTSPAFDFDGY
AmGPPS.LSU       (45) -KSKSKTHLRFSISSILTKNPQESSQK----------TSKDPTFTLDFKTY
SmGPPS.SSUI      (14) -KMSIYVNTKS-SFPFKTRSNLSRSS------------RACTAVVAQNHSY
MpGPPS.SSUI       (3) -INLSHINSKT-CFPLKTRSDLSRSSSARCMPTAAAAAFPTIATAAQSQPY
AmGPPS.SSUI       (3) -HGLTHFNTKSGLFPSITKSKTTRPS-----------TRPVILAMTRTQTY
SmGPPS.SSUII.1   (10) -------SSNSRMPRTAVLRRAVRCSS----------AASVPTTQPDLRNY
OsGPPS.SSUII     (10) -----LPFAKLPSTSKSTRFLPIRASSA--------AAAAAASPSFDLRLY
SmGPPS.SSUII.2    (5) ----SVITSTPHVCLPKVTRPALRRTIR--------CSAASVSPGFDLKTY
AtGPPS.SSUII     (11) -SSFCSLPEKPHTLPMKLSPAAIRSSS----------SSAPGSLNFDLRTY
Consensus        (52)         SK  V   KTK   VR S           SAA      D  SY

                (103) 103    110       120       130       140          153
SmGPPS           (90) VADELSILADRLRSMVVAEVPKLASAAEYFFKFGVE-----GKRFRPTVLL
SlGPPS           (81) VADELSLLTNRLRSMVVAEVPKLASAAEYFFKLGVE-----GKRFRPTVLL
CrGPPS           (86) VADELSILANRLRSMVVAEVPKLASAAEYFFKLGVE-----GKRFRPTVLL
SmGPPS.LSU       (57) MVEKADSVNKALEAVVQMKEPLKIHESMRYSLLAGG------KRVRPMLCI
MpGPPS.LSU       (90) MLRKAKSVNKALEAAVQMKEPLKIHESMRYSLLAGG------KRVRPMLCI
AmGPPS.LSU       (85) MLEKASSVNKALEQAVLLKEPLKIHESMRYSLLAGG------KRVRPMLCI
SmGPPS.SSUI      (51) WAAIEADIDTYLKKSIAIRSPETVFEPMHHLTFAAP------RTAASAICV
MpGPPS.SSUI      (52) WAAIEADIERYLKKSITIRPPETVFGPMHHLTFAAP------ATAASTLCL
AmGPPS.SSUI      (42) RATIESDIESYLKKAIPIRAPESVFEPMHHLTFAAP------RTSASALCV
SmGPPS.SSUII.1   (44) WTSLISDVDRKLNDAVPLKYPELIHESMRYSLLAKT-----AKRAPPVMCI
OsGPPS.SSUII     (48) WTSLIADVEAELDAAMPIRTPERIHSAMRYAVLPGAGNEGTAKRAPPVLCV
SmGPPS.SSUII.2   (44) WTSLIKEIDQKLDEAIPVKYPQQIYEAMRYSVLAKG-----AKRAPPVMCV
AtGPPS.SSUII     (51) WTTLITEINQKLDEAIPVKHPAGIYEAMRYSVLAQG-----AKRAPPVMCV
Consensus       (103)    LISDI   L  AV IK P  IHEAMRYSLLAG      AKRA PVLCV
                                                                     **

                (154) 154   160       170       180       190          204
SmGPPS          (136) LMATALDLPIARQTSEVAVNTLSTELRTRQQCVAEITEMIHVASLLHDDVL
SlGPPS          (127) LMATALNVQIPRSAPQVDVDSFSGDLRTRQQCIAEITEMIHVASLLHDDVL
CrGPPS          (132) LMATAIDAPISRTPPDTSLDTLSTELRLRQQTIAEITKMIHVASLLHDDVL
SmGPPS.LSU      (102) AACELVGG-----------------EESTAMPAACAVEMIHTMSLMHDDLP
MpGPPS.LSU      (135) AACELVGG-----------------DESTAMPAACAVEMIHTMSLMHDDLP
AmGPPS.LSU      (130) AACELVGG-----------------LESTAMPSACAVEMIHTMSLIHDDLP
SmGPPS.SSUI      (96) AACELVGG-----------------ERSQAIATASAIHIMHAAAYAHEHLP
MpGPPS.SSUI      (97) AACELVGG-----------------DRSQAMAAAAAIHLVHAAAYVHEHLP
AmGPPS.SSUI      (87) AACELVGG-----------------DRSDAMAAAAAVHLMHVAAYTHENLP
SmGPPS.SSUII.1   (90) AACELFGG-----------------DRAAAIPTVCALEMVHAASFVHDDLP
OsGPPS.SSUII     (99) AACELLGA-----------------PREAALPAAVALEMLHAASLVHDDLP
SmGPPS.SSUII.2   (90) AACEIFGG-----------------NRLAAFPTACALEMVHSASLIHDDLP
AtGPPS.SSUII     (97) AACELFGG-----------------DRLAAFPTACALEMVHAASLIHDDLP
Consensus       (154) AACELVGG                 DRS AMP A ALEMIH ASLIHDDLP
                      ***                                            ****
```



```
               (205) 205  210       220       230       240            255
       SmGPPS (187) DDADT--RRGIGSLNYVMGNKLAVLAGDFLLSRACVALASLKN---TE---
       SlGPPS (178) DDADT--RRGIGSLNFVMGNKLAVLAGDFLLSRACVALASLKN---TE---
       CrGPPS (183) DDAET--RRGIGSLNFVMGNKLAVLAGDFLLSRACVALASLKN---TE---
   SmGPPS.LSU (136) CMDNDDLRRGKPTNHKVFGEDVAVLAGDALLSLAFEHVAVATRGSAPE---
   MpGPPS.LSU (169) CMDNDDLRRGKPTNHMAFGESVAVLAGDALLSFAFEHVAAATKGAPPE---
   AmGPPS.LSU (164) CMDNDDLRRGKPTNHKIYGEDVAVLAGDALLAFSFEHVAKSTKGVSSD---
  SmGPPS.SSUI (130) LTDRP-RPNSKPAIQHKYGPNIELLTGDGMASFGFELLAGSIRSDHPN-PE
  MpGPPS.SSUI (131) LTDGS-RPVSKPAIQHKYGPNVELLTGDGIVPFGFELLAGSVDPARTDDPD
  AmGPPS.SSUI (121) LTDG---PMSKSEIQHKFDPNIELLTGDGIIPFGLELMARSMDPTRNN-PD
SmGPPS.SSUII.1 (124) YIDDALSRRGQLPNHTLYGPDMAILAGDALLPLAFQYIVLHTP-TQLVSQL
  OsGPPS.SSUII (133) CFDAAPTRRGRPSTHAAYGTDMAVLAGDALFPLAYTHVIAHTPSPDPVPHA
SmGPPS.SSUII.2 (124) CMDDDPSRRGQPSNHTVFGVDMAILAGDALFPLGFRHIVSHTP-TDLVPHT
  AtGPPS.SSUII (131) CMDDDPVRRGKPSNHTVYGSGMAILAGDALFPLAFQHIVSHTP-PDLVPRA
     Consensus (205)   D    RRGKPS H VYG  MAVLAGDALLS AF HLAS T     E
                    ***

               (256) 256           270       280       290            306
       SmGPPS (230) ----VVTLIAQVV--EHLVTGETMQMTTTSEQRCSMEYYMEKTYYKTASLI
       SlGPPS (221) ----VVCLLATVV--EHLVTGETMQMTTSSDERCSMEYYMQKTYYKTASLI
       CrGPPS (226) ----VVSLLATVV--EHLVTGETMQMTTTSDQRCSMEYYMQKTYYMTASLI
   SmGPPS.LSU (184) RILRALGQLAKSIGAEGLVAGQVVDICSEGMAEVGLDHLEFIHLHKTAALL
   MpGPPS.LSU (217) RIVRVLGELAVSIGSEGLVAGQVVDVCSEGMAEVGLDHLEFIHHHKTAALL
   AmGPPS.LSU (212) RIVRVIGELAKCIGSEGLVAGQVVDISSEGMTEVGLEHLEFIHVHKTAALL
  SmGPPS.SSUI (179) RILRVIIEISRASGSEGIIDGFYREKEI-VDQHSRFDFIEYLCRKKYGEMH
  MpGPPS.SSUI (181) RILRVIIEISRAGGPEGMISGLHREEEI-VDGNTSLDFIEYVCKKKYGEMH
  AmGPPS.SSUI (168) RILRAIIELTRVMGSEGIVEGQYHELGL-NQLN-DLELIEYVCKKKEGTLH
SmGPPS.SSUII.1 (174) HLLRVVQEIARAVGSTGMAATQFIGE-----N-------------KFGELG
  OsGPPS.SSUII (184) VLLRVLGELARAVGSTGMAAGQFLDLAG-ATAL-GEAEVMKVLTKKFGEMA
SmGPPS.SSUII.2 (174) RLLRVIAEIARAVGSTGMAAGQFLDLEG-GPNA-----VDLVQEKKFGEMG
  AtGPPS.SSUII (181) TILRLITEIARTVGSTGMAAGQYVDLEG-GPFP-----LSFVQEKKFGAMG
     Consensus (256) RILRVI ELARAVGSEGLVAGQ VDL        LE LEFV    K GAL

               (307) 307          320       330       340            357
       SmGPPS (275) CNSCKSIALIAGQTAEVSNLAYEYGKNLGLAFQIIDDVLDFTGTSASLG--
       SlGPPS (266) SNSCKAIALLAGHSAEVSVLAFDYGKNLGLAFQLIDDVLDFTGTSATLG--
       CrGPPS (271) SNSCKAIALLAGQTSEVAMLAYEYGKNLGLAFQLIDDVLDFTGTSASLG--
   SmGPPS.LSU (235) QGSVVMGAILGGAKEEEVERLRKFAKCIGLMFQVVDDILDVTKSSHELGKT
   MpGPPS.LSU (268) QGSVVLGAILGGGKEEEVAKLRKFANCIGLLFQVVDDILDVTKSSKELGKT
   AmGPPS.LSU (263) EASVVLGAIVGGADDEDVEKLRKFARCIGLLFQVVDDILDVTKSSQELGKT
  SmGPPS.SSUI (229) ACAAASGAILAGGAEEEIQKLRNFGHYAGTLIGLLHKKIDTP---------
  MpGPPS.SSUI (231) ACGAACGAILGGAAEEEIQKLRNFGLYQGTLRGMMEMKNSHQ---------
  AmGPPS.SSUI (217) ACGAACGAILGGCDEDKIEKLRRFGLYVGTVQGLLGK---NR---------
SmGPPS.SSUII.1 (207) RCSAVCGALLGGASDEEIERLGEYGRIVGILYRVVEDMLEGK---------
  OsGPPS.SSUII (233) ECSAACGAMLGGAGPDEEAALRRYGRTIGVLYQLVDDIRSAS---------
SmGPPS.SSUII.2 (219) QCSAVCGGLLAGASDDEIEHLRKYGRAVGILYEVVDDIRETKLQPD-----
  AtGPPS.SSUII (226) ECSAVCGGLLGGATEDELQSLRRYGRAVGMLYQVVDDITEDKKKSY-----
     Consensus (307)  CSA  GAILGGA EEEI KLRKYGR IGLLFQVVDDILD T  S
                                                       *****

               (358) 358         370       380       390             408
       SmGPPS (324) -----KGSLSDIRHGIVTAPILFAIEEYPELRKIVDQGFEKSSNVDRALEI
       SlGPPS (315) -----KGSLSDIRHGIVTAPILYAMEEFPQLRTLVDRGFDDPVNVEIALDY
       CrGPPS (320) -----KGSLSDIRHGIVTAPILFAIEEFPELRAVVDEGFENPYNVDLALHY
   SmGPPS.LSU (286) AGKDLVADKTTYPKLLGVQKSKEFADDLNREAQEQLLHFDSHK-----AAP
   MpGPPS.LSU (319) AGKDLVADKTTYPKLIGVEKSKEFADRLNREAQEQLLHFHPHR-----AAP
   AmGPPS.LSU (314) AGKDLVADKTTYPKLLGIEKSREFAEKLNREAQEQLEGFDSVK-----AAP
  SmGPPS.SSUI (271) ------------------QIQN-VIGKLKDLALKELEGFHGKN--VELLCS
  MpGPPS.SSUI (273) ------------------LIDENIIGKLKELALEELGGFHGKN--AELMSS
  AmGPPS.SSUI (256) ------------------SGFEGRIKELKELAVKELESFGGEK--IELIRG
SmGPPS.SSUII.1 (249) ---K---------------VNIGVVEDLKSRARKELYMFEKYG---DKVLP
  OsGPPS.SSUII (275) -GNGKMRSNASVLRALGMDRALGIVEELKAQAKMEADRFGDKYG--ERVLP
SmGPPS.SSUII.2 (265) EKDKSKSKGKSYVSVYGVEKAMEVAEDLRSQAKKELESLEKYG---DKVLP
  AtGPPS.SSUII (272) DGGAEK--G-----------MMEMAEELKEKAKKELQVFDNKYGGGDTLVP
     Consensus (358)                I     L  IEELK  A  EL  FD      E L P

               (409) 409        420       430       440                459
       SmGPPS (370) LSKSSGIQRARELAAKHARLASAAIDALPENEDEVVQRSMRALVELTHIVI
       SlGPPS (361) LGKSRGIQRTRELARKHASLASAAIDSLPESDDEEVQRSRRALVELTHRVI
       CrGPPS (366) LGKSRGIQRTRELAIKHANLASDAIDSLPVTDDEHVLRSRRALVELTQRVI
   SmGPPS.LSU (332) LIAIANYIAYRNN--------------------------------------
   MpGPPS.LSU (365) LIALANYIAYRDN--------------------------------------
   AmGPPS.LSU (360) LIALANYIAYRDN--------------------------------------
  SmGPPS.SSUI (301) LVADASLCEAELEV-------------------------------------
  MpGPPS.SSUI (304) LVAEPSLYAA-----------------------------------------
  AmGPPS.SSUI (287) VFELEHSLAGV----------------------------------------
SmGPPS.SSUII.1 (279) LHTFIDYAAERV---------------------------------------
  OsGPPS.SSUII (323) LYSFVDYAVERGFELQDAATTP-----------------------------
SmGPPS.SSUII.2 (313) LYSFVDYAADRYFGFEELV--------------------------------
  AtGPPS.SSUII (310) LYTFVDYEAHRHFLLPL----------------------------------
     Consensus (409) L A   Y A R

               (460)
       SmGPPS (421) TRTK
       SlGPPS (412) TRTK
       CrGPPS (417) TRRK
   SmGPPS.LSU (345) ----
   MpGPPS.LSU (378) ----
   AmGPPS.LSU (373) ----
  SmGPPS.SSUI (315) ----
  MpGPPS.SSUI (314) ----
  AmGPPS.SSUI (298) ----
SmGPPS.SSUII.1 (291) ----
  OsGPPS.SSUII (345) ----
SmGPPS.SSUII.2 (332) ----
  AtGPPS.SSUII (327) ----
     Consensus (460)
```

Figure S10. Sequence alignment of GPPS proteins from *S. miltiorrhiza* and various other plants. GPPS proteins inlucled are *Solanum lycopersicum* SlGPPS (ABB88703), *Catharanthus roseus* CrGPPS (ACC77966), *Mentha x piperita* MpGPPS.LSU (AF182828), *Antirrhinum majus* AmGPPS.LSU (AAS82860), MpGPPS.SSUI (AF182827), AmGPPS.SSUI (AAS82859), rice OsGPPS.SSUII (EAY87007), *Arabidopsis thaliana* AtGPPS.SSUII (At4g38460), and five *S. miltiorrhiza* GPPSs. The FARM (the first aspartate-rich motif, DDLPCMD), SARM (the second aspartate-rich motif, DDXXD) and CXXXC motifs are indicated by '***'.

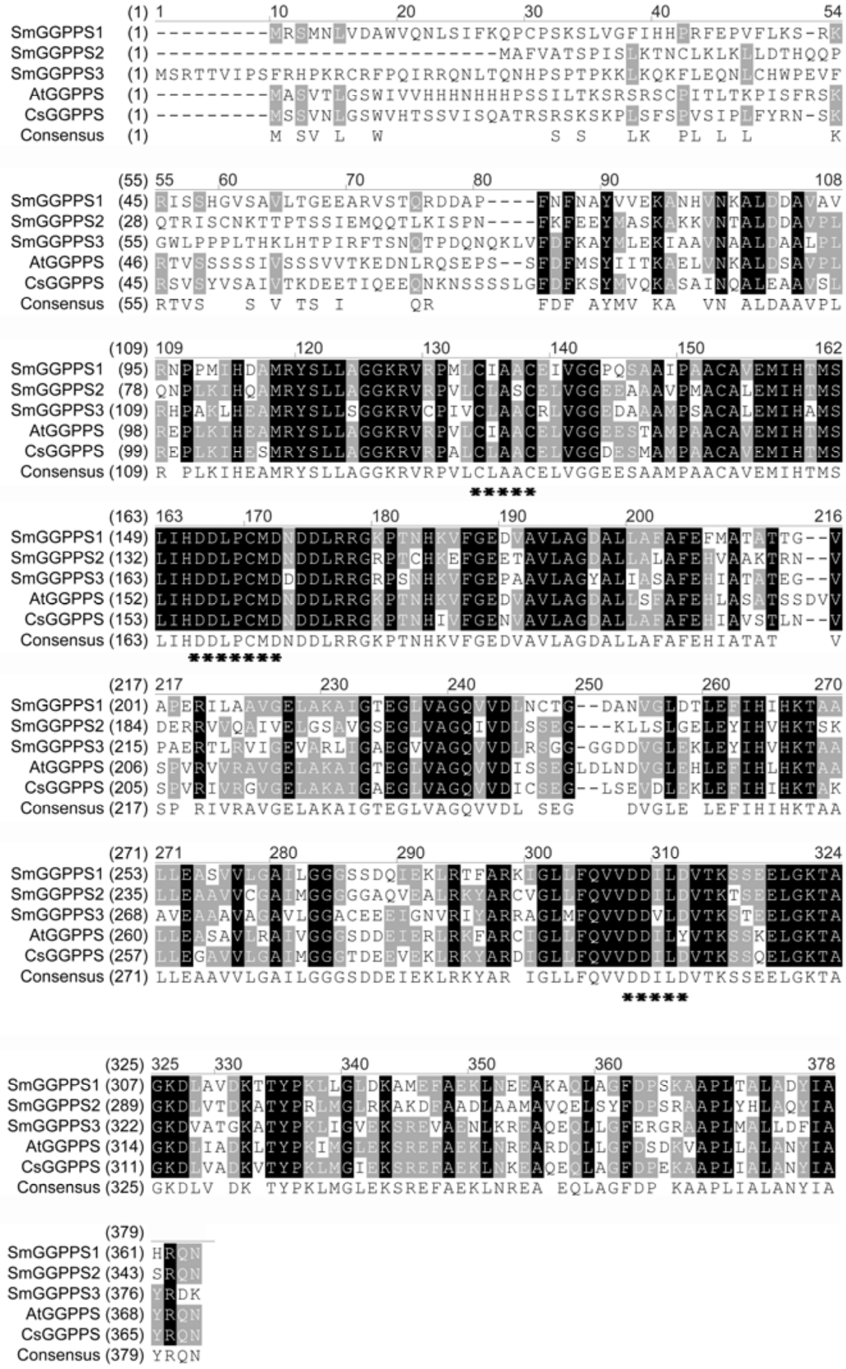


Figure S11. Sequence alignment of GGPPS proteins from *S. miltiorrhiza* and various other plants. GGPPS proteins included are *A. thaliana* AtGGPPS (AAM65107), *Croton sublyratus* CsGGPPS (BAA86284), and three *S. miltiorrhiza* SmGGPPSs. The FARM (the first aspartate-rich motif, DDLPCMD), SARM (the second aspartate-rich motif, DDXXD) and CXXXC motifs are indicated by '***'.

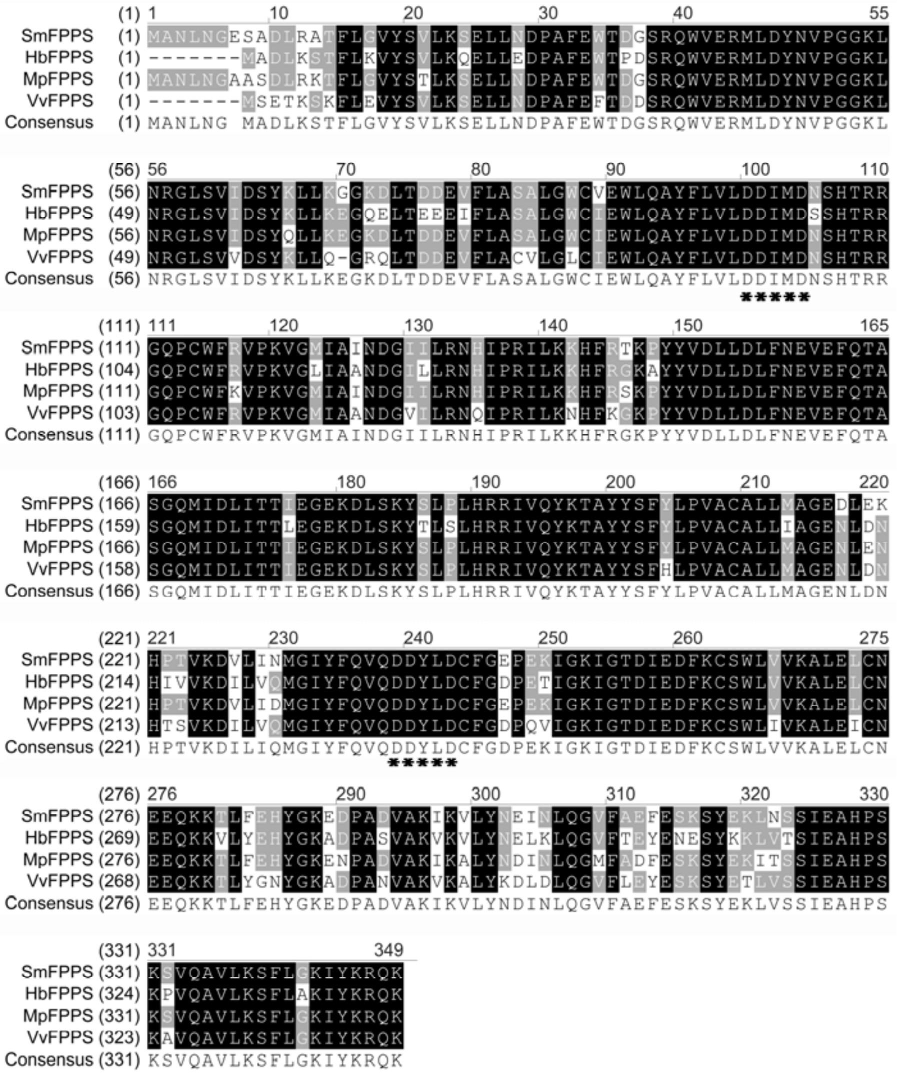


Figure S12. Sequence alignment of FPPS proteins from *S. miltiorrhiza* and various other plants. FPPS proteins included are *Hevea brasiliensis* HbFPPS (AAM98379), *Mentha x piperita* MpFPPS (AF384040), *Vitis vinifera* VvFPPS (AAX76910), and the *S. miltiorrhiza* SmFPPS. The FARM (the first aspartate-rich motif, DDIMD) and SARM motifs (the second aspartate-rich motif, DDXXD) are indicated by '***'.

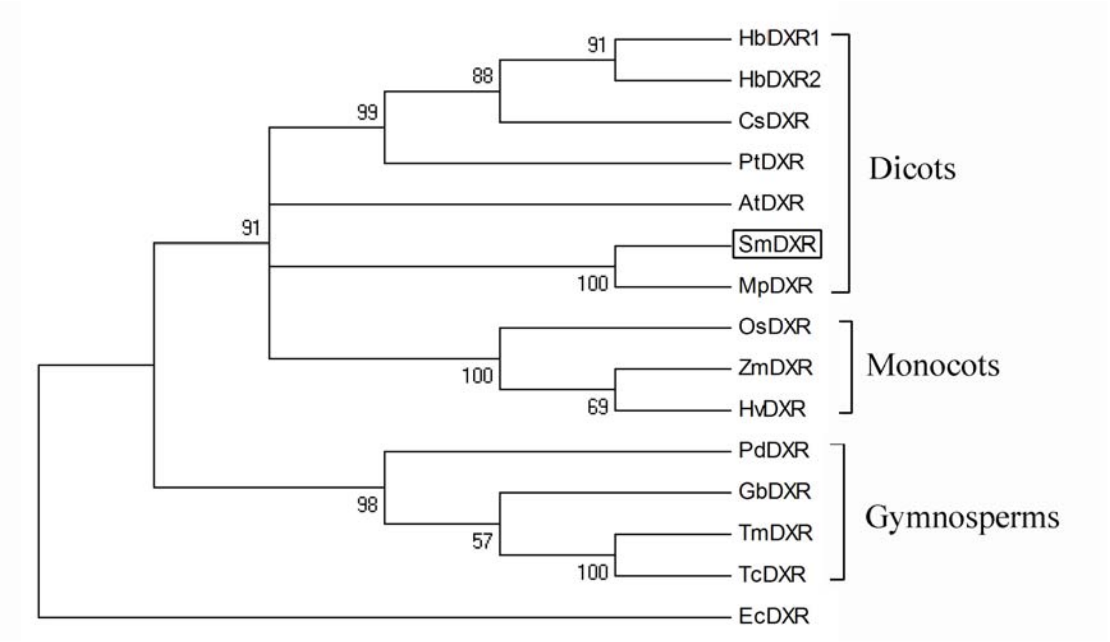


Figure S13. Phylogenetic relationships of DXRs form *S. miltiorrhiza* and various other plants. DXR proteins included are *S. miltiorrhiza* SmDXR (ABJ80680), *Hevea brasiliensis* HbDXR1 (AAS94121), HbDXR2 (ABF18928), *A. thaliana* AtDXR (AED97658), *Populus trichocarpa* PtDXR (XP_002318048), *Zea mays* ZmDXR (CAC03581), *Mentha* × *piperita* MpDXR (AF116825), rice OsDXR (AF367205), *H. vulgare* HvDXR (CAE47438), *C. stellatopilosus* CsDXR (ABO38177), *P. densiflora* PdDXR (ACC54558), *G. biloba* GbDXR (AAR95700), *Taxus x media* TmDXR (AAU87836), and *T. cuspidate* TcDXR (AAT47184). *E. coli* EcDXR (NP_414715) was used as an out group.

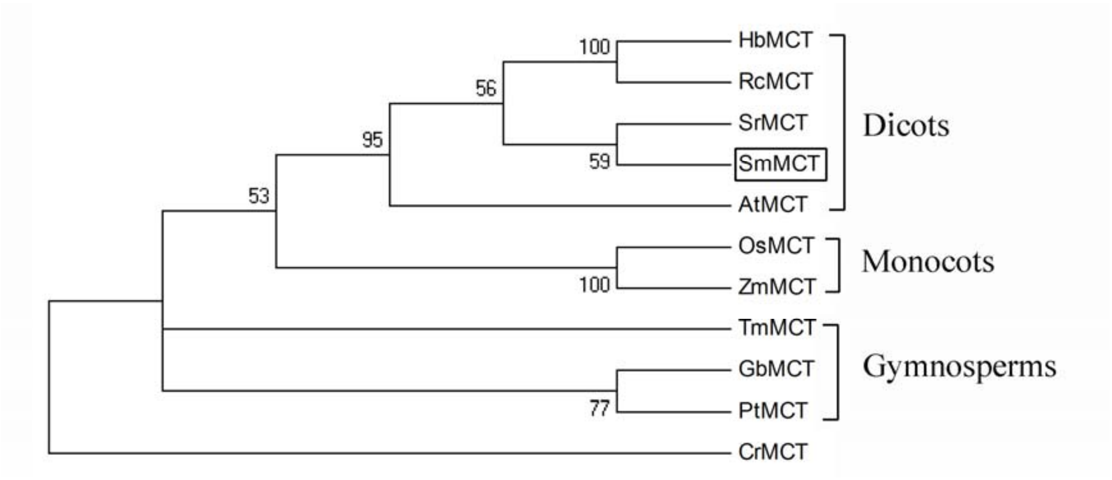


Figure S14. Phylogenetic relationships of MCTs form *S. miltiorrhiza* and various other plants. MCT proteins included are *S. rebaudiana* SrMCT (ABB88837), *S. miltiorrhiza* SmMCT (JN831096), *A. thaliana* AtMCT (NP_565286), *H. brasiliensis* HbMCT (BAF98291), *R. communis* RcMCT (XP_002519366), rice OsMCT (BAD82245), *Z. mays* ZmMCT (NP_001140440), *Taxus x media* TmMCT (ABU48538), *G. biloba* GbMCT (AAZ80386), and *P. sitchensis* PtMCT (ACN40114). *E. coli* EcMCT (XP_001698942) was used as an out group.

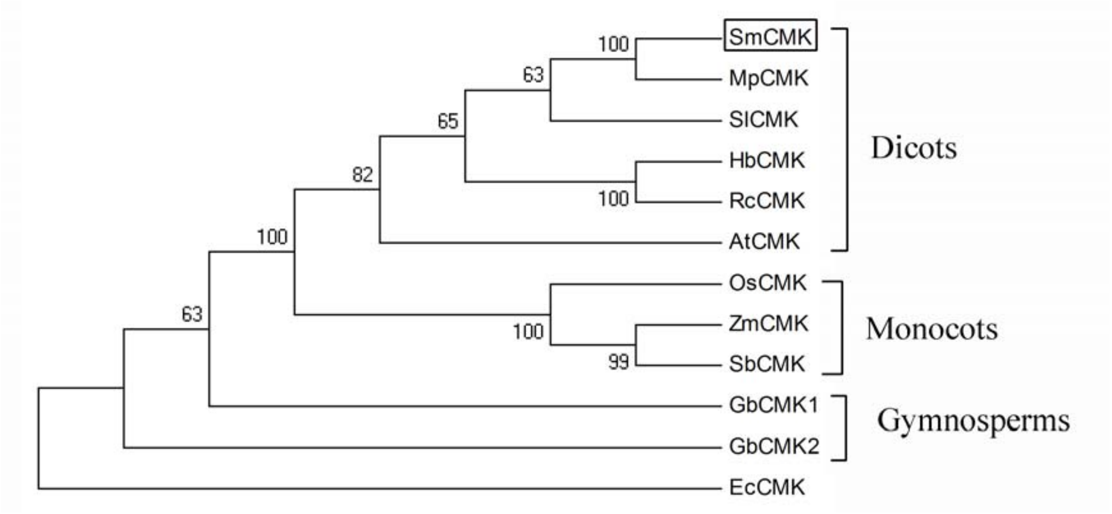


Figure S15. Phylogenetic relationships of CMKs form *S. miltiorrhiza* and various other plants. CMK proteins included are *G. biloba* GbCMK1 (AAZ80384), GbCMK2 (DQ102359), *A. thaliana* AtCMK (BT000208), *L. esculentum* LeCMK (AAB49936), *S. miltiorrhiza* SmCMK (ABP96842), rice OsCMK (NP_915832), *Z. mays* ZmCMK (NP_001149138), *S. bicolor* SbCMK (XP_002456497), *M. piperita* MpCMK (CAB65292), *H. brasiliensis* HbCMK (BAF98293), and *R. communis* RbCMK (XP_002523216). *E. coli* EcCMK (NP_415726) was used as an out group.

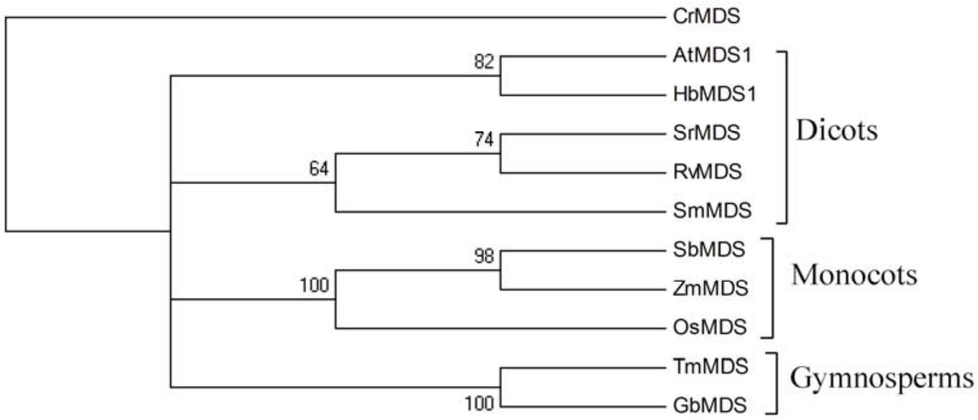


Figure S16. Phylogenetic relationships of MDSs form *S. miltiorrhiza* and various other plants. MDS proteins included are SrMDS (*S. rebaudiana*, ABG23395), RvMDS (*R. verticillata*, ABV89583), HbMDS1 (*H. brasiliensis*, BAF98294), AtMDS1 (*A. thaliana*, NP_850971), OsMDS (*O.sativa*, BAD29384), SbMDS (*S. bicolor*, XP_002452756), ZmMDS (*Z,mays*, NP_001150687), GbMDS (*G. biloba*, ABL77405), TmMDS (*Taxus x media*, ABU48537), and *S. miltiorrhiza* SmMDS (JN831097). CrMDS (*C. reinhardtii*, XP_001690985) was used as an out group.

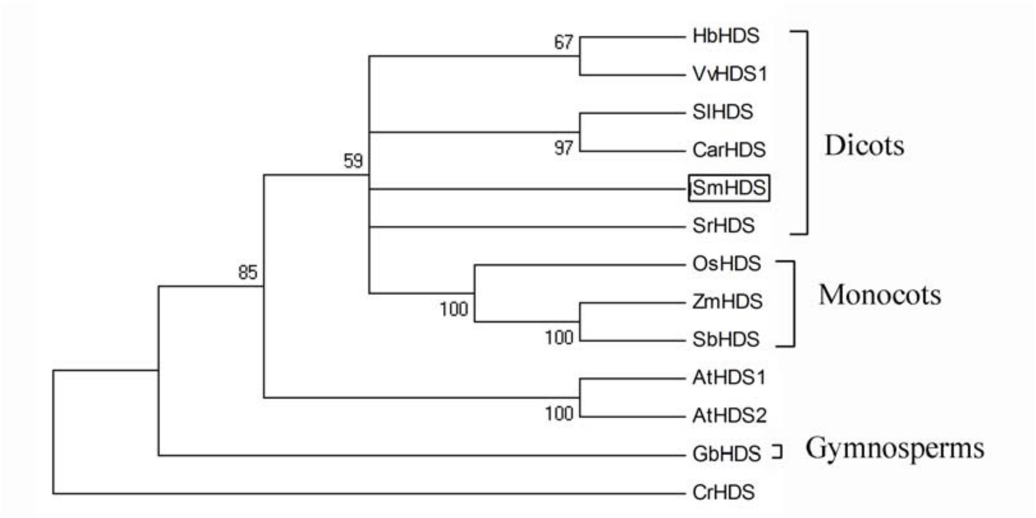


Figure S17. Phylogenetic relationships of HDSs form *S. miltiorrhiza* and various other plants. HDS proteins included are AtHDS1 (*A. thaliana*, NP_200868.2), HbHDS (*H. brasiliensis*, BAF98296), AtHDS2/GcpE protein (BAB09833), CarHDS (*C. roseus*, AAO24774), GbHDS (*G. biloba*, ABB78087), VvHDS1 (*V. vinifera*, XP_002285130), SlHDS (*S. lycopersicum*, AF435086), SrHDS (*S. rebaudiana*, ABG75916.2), OsHDS (*O. sativa*, AAO72576), ZmHDS (*Z. mays*, ACL53652), SbHDS (*S. bicolor*, XP_002454137), and *S. miltiorrhiza* SmHDS (JN831098). CrHDS (*C. reinhardtii*, XP_001690937) was used as an out group.

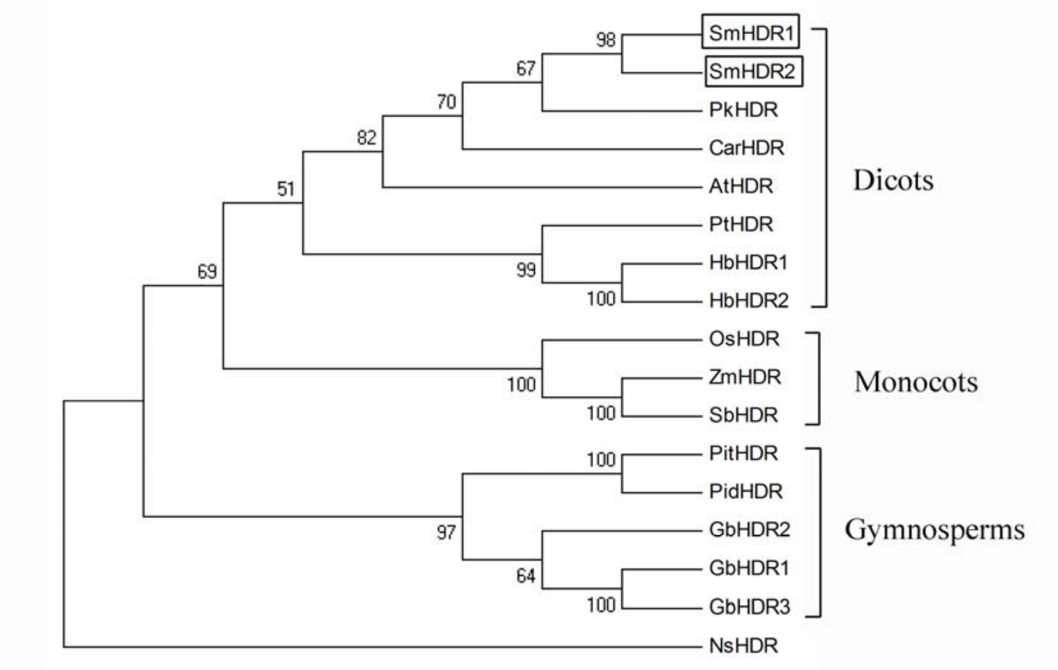


Figure S18. Phylogenetic relationships of HDRs form *S. miltiorrhiza* and various other plants. HDR proteins included are PtHDR (*P. trichocarpa*, XP_002305413), HbHDR1 (*H. brasiliensis*, ACG55683), HbHDR2 (*H. brasiliensis*, BAF98297), AtHDR (*A. thaliana*, NP_567965), CarHDR (*C. roseus*, ABI30631), PkHDR (*P. kurrooa*, ABM89226), PitHDR (*P. taeda*, ABO26588), PidHDR (*P. densiflora*, ACC54561), GbHDR1 (*G. biloba*, ABC84344), GbHDR2 (*G. biloba*, ABB78089), GbHDR3 (*G. biloba*, ABB78090), OsHDR (*O. sativa*, NP_001051167), ZmHDR (*Z. mays*, NP_001169300), SbHDR (*S. bicolor*, XP_002463933), and two *S. miltiorrhiza* SmHDRs. NsHDR (*Nostoc sp. PCC 7120*, NP_485028) was used as an out group.

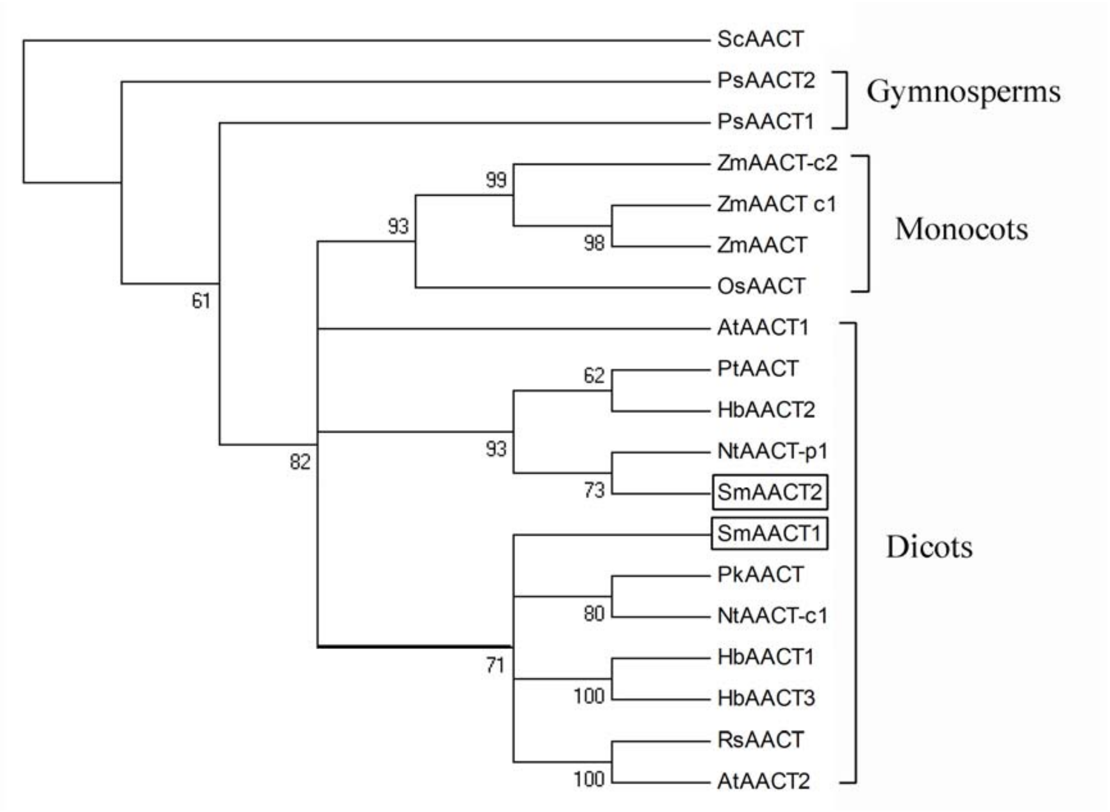


Figure S19. Phylogenetic relationships of AACTs form *S. miltiorrhiza* and various other plants. AACT proteins included are ZmAACT (*Z. mays*, ACF85124), ZmAACT-c1 (ACG34735), ZmAACT-c2 (NP_001148667), OsAACT (*O. sativa*, NP_001041797), AtAACT1 (*A. thaliana*, AT5G47720), AtAACT2 (AT5G48230), PtAACT (*P. sitchensis*, XP_002308755), HbAACT1 (*H. brasiliensis*, BAF98276), HbAACT2 (AF429383), HbAACT3 (BAF98277), RsAACT (*R. sativus*, CAA55006), PkAACT (*P. kurrooa*, ABC74567), PsAACT1 (*P. sitchensis*, ACN40771), PsAACT2 (ACN41149.1), NtAACT-c1 (*N. tabacum*, AAU95618), NtAACT-p1 (AAU95619), and two *S. miltiorrhiza* SmAACTs. ScAACT (*S. cerevisiae*, P41338.3) was used as an out group.

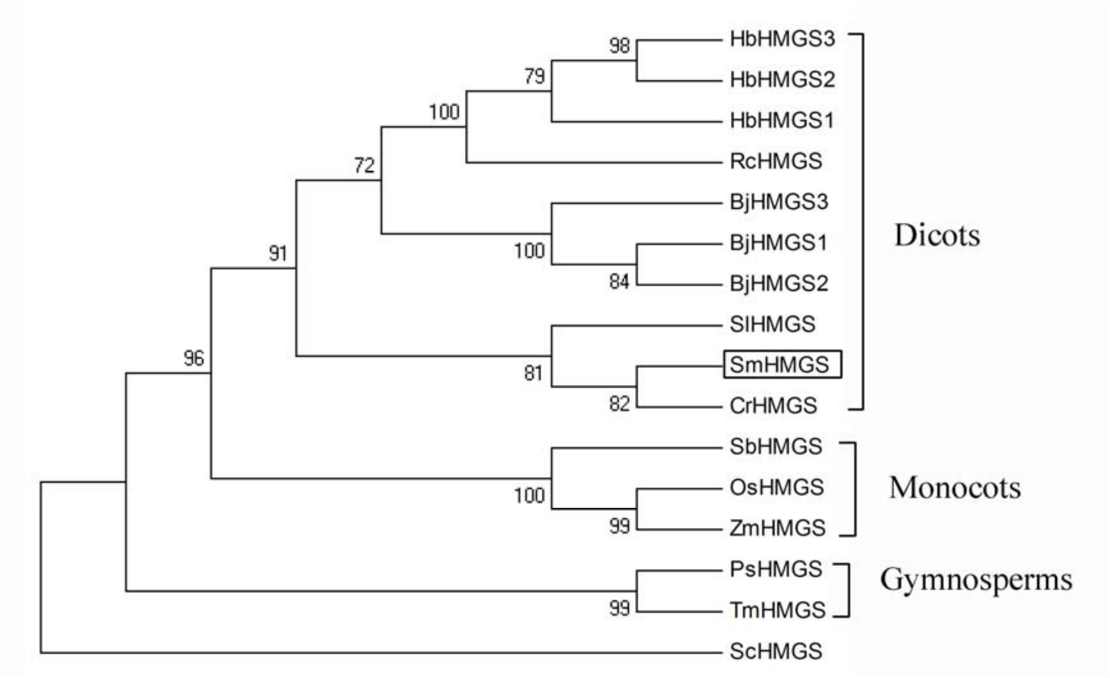


Figure S20. Phylogenetic relationships of HMGSs form *S. miltiorrhiza* and various other plants. HMGS proteins included are BjHMGS1 (*B. juncea*, AF188639), BjHMGS2 (AF188640), BJHMGS3 (AF148847_1), OsHMGS (*O. sativa*, Os08g0544900), SlHMGS (*S. lycopersicum*, ABX55778), HbHMGS1 (*H. brasiliensis*, AF429389), HbHMGS2 (AAS46245.1), HbHMGS3 (BAF98279), TmHMGS (*Taxus x media*, AAT73206), CrHMGS (*C. roseus*, AEC13715), PsHMGS (*P. sylvestris*, CAA65250), SmHMGS (*S. miltiorrhiza*, ACV65039), RcHMGS (*R. communis*, XP_002509692), SbHMGS (*S. bicolor*, XP_002468628), ZmHMGS (*Z. mays*, ACG33137). ScHMGS (*S. cerevisiae*, P54839) was used as an out group.

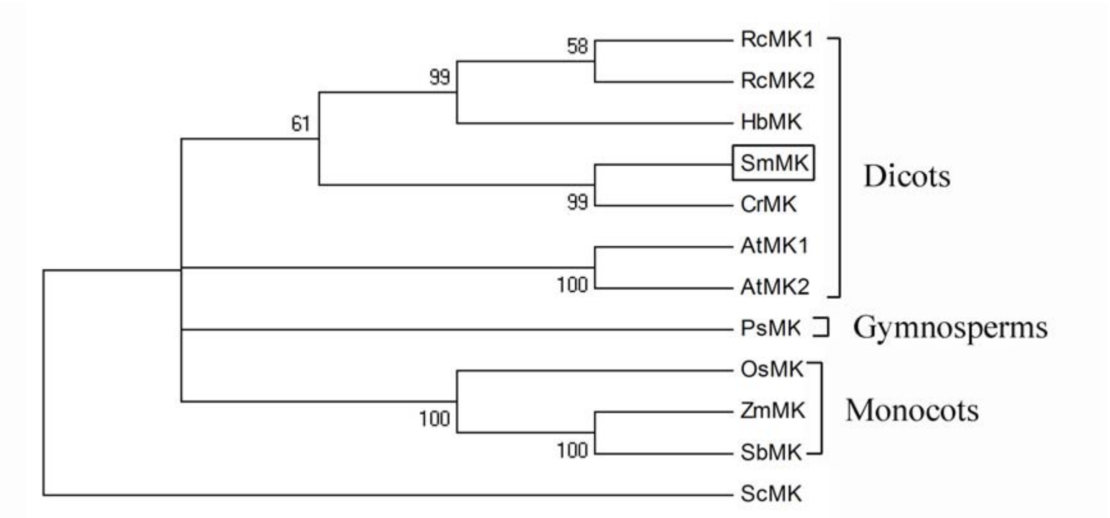


Figure S21. Phylogenetic relationships of MKs form *S. miltiorrhiza* and various other plants. MK proteins included are OsMK (*O. sativa*, Os10g0329300), AtMK1 (*A. thaliana*, NP_198097), AtMK2 (L77688), HbMK (*H. brasiliensis*, AF429384), ZmMK (*Z. mays*, ACG46416), PsMK (*P. sitchensis*, ACN40231), SbMK (*S. bicolor,* XP_002453181), RcMK1 (*R. communis*, XP_002512365), RcMK2 (XP_002529655), CrMK (*C. roseus*, ADR65111), and SmMK (*S. miltiorrhiza*, JN831104). ScMK (*S. cerevisiae*, P24521) was used as an outgroup.

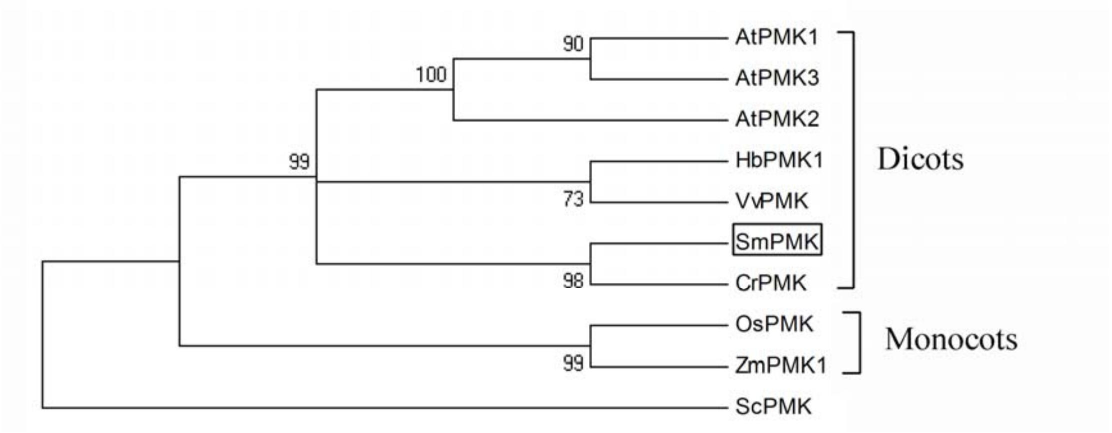


Figure S22. Phylogenetic relationships of PMKs form *S. miltiorrhiza* and various other plants. PMK proteins included are OsPMK (*O. sativa*, Os03g0253100), AtPMK1 (*A. thaliana*, NP_174473), AtPMK2 (NP_001185124), AtPMK3 (BAD43274), HbPMK1 (*H. brasiliensis*, AF429385), ZmPMK1 (Z. mays, ACG35008), VvPMK (V. vinifera, XP_002275808), CrPMK (C. roseus, ADR65112), and SmPMK (*S. miltiorrhiza*, JN831095). ScPMK (*S. cerevisiae* RM11-1a, EDV11711) was used as an out group.

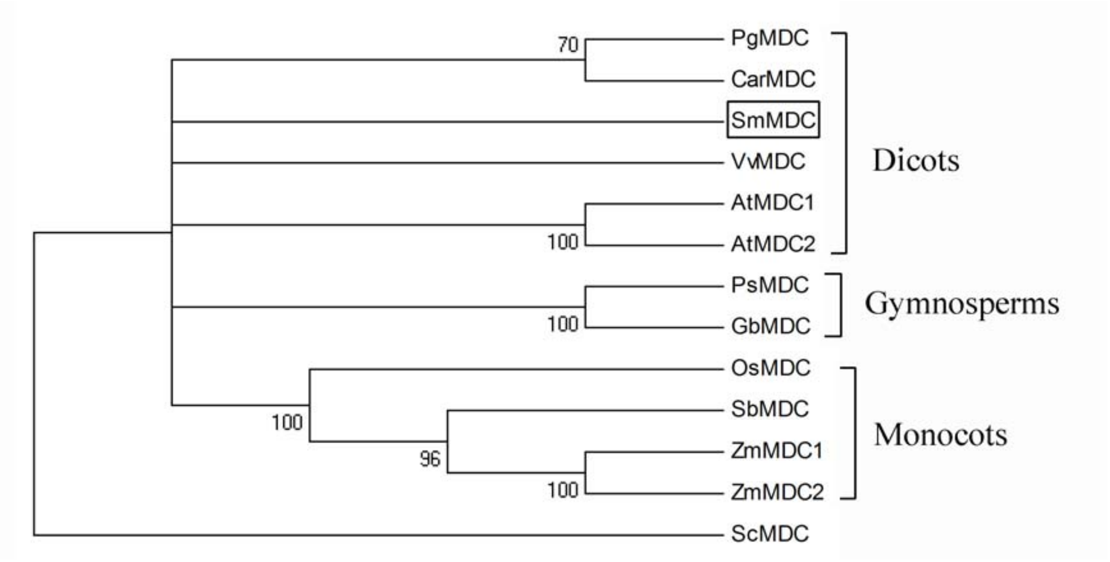


Figure S23. Phylogenetic relationships of MDCs form *S. miltiorrhiza* and various other plants. MDC proteins included are VvMDC (*V. vinifera*, CAN82519), AtMDC1 (*A. thaliana*, NP_181404), AtMDC2 (NP_566995), ZmMDC1 (*Z. mays*, ACF86239), ZmMDC2 (NP_001149256), PsMDC (*P. sitchensis,* ACN41090), SbMDC (*S. bicolor*, XP_002452980), PgMDC (P. ginseng, ADI80345), CrMDC (*C. roseus*, ADR65113), OsMDC (*O.sativa*, BAD27942), GbMDC (*G. biloba*, AAV32433), and SmMDC (*S. miltiorrhiza*, JN831105). ScMDC (*S. cerevisiae*, P32377.2) was used as an out group.

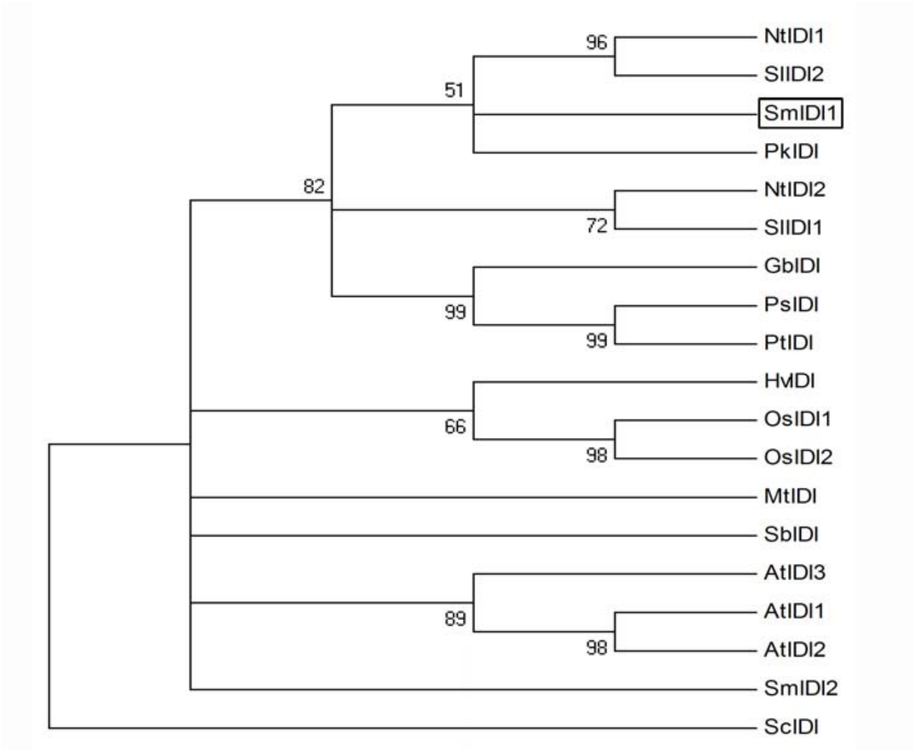


Figure S24. Phylogenetic relationships of IDIs form *S. miltiorrhiza* and various other plants. IDI proteins included are AtIDI1 (*A. thaliana*, AF188066), AtIDI2 (NP_186927), AtIDI3 (AF188067), GbIDI (*G. biloba*, ACU56979), MtIDI (*M. truncatula*, ACJ84381), NtIDI1 (N. tabacum, BAB40973), NtIDI2 (BAB40974), OsIDI1(O. sativa, ABR26078), OsIDI2 (AF188065), PkIDI (*P. kurrooa*, ABO14800), PsIDI (*P. sitchensis*, ACN41037), PtIDI (*P. taeda*, ACU56978), SbIDI (*S. bicolor*, XP_002439804), SlIDI1 (*S. lycopersicum*, ABX55779), SlIDI2 (ACS34993), HvIDI (H. vulgare, BAJ86792), SmIDI1 (*S. miltiorrhiza*, ABV08818) and SmIDI2 (JN831106). ScIDI (*S. cerevisiae*, AAB68245) was used as an out group.

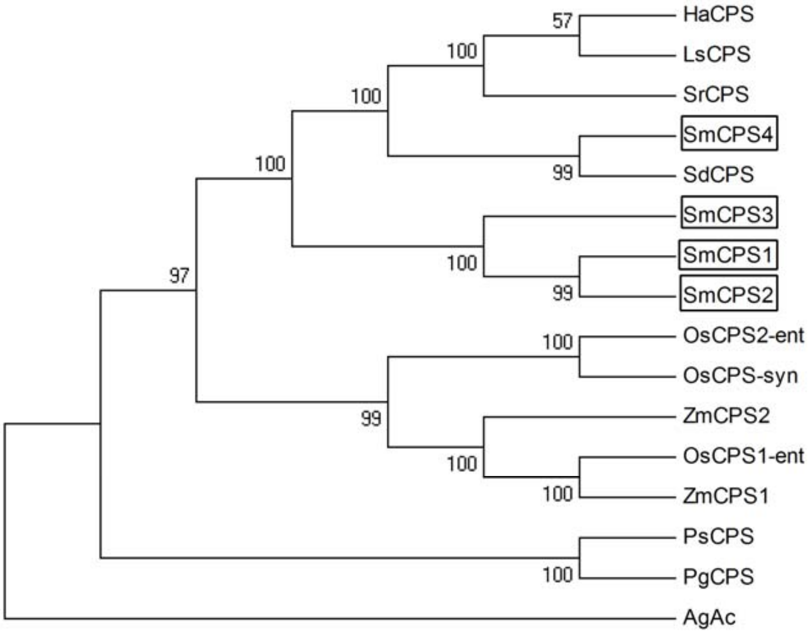


Figure S25. Phylogenetic relationships of CPSs form *S. miltiorrhiza* and various other plants. CPS proteins included are HaCPS (*H. annuus*, CBL42915), SrCPS (*S. rebaudiana*, AAB87091), LsCPS (*L. sativa*, BAB12440), SdCPS (*S. dulcis*, BAB03594), PsCPS (*P. sitchensis*, ADB55709), PgCPS (*P. glauca*, ADB55707), OsCPS1-ent (*O. sativa*, BAD42449.2), OsCPS2-ent/OsCyc2 (BAD42452), OsCPS-syn/OsCyc1 (BAD42451), ZmCPS1 (*Z. mays*, AAA73960), ZmCPS2 AAT70083), SmCPS1 (*S. miltiorrhiza*, ABV57835), SmCPS2 (JN831114), SmCPS3 (JN831115), and SmCPS4 (JN831120). AgAc (*A. grandis*, AAB05407) was used as an outgroup. Sequences in upstream of the partially conserved motif SPYDTAWVAL were trimmed.

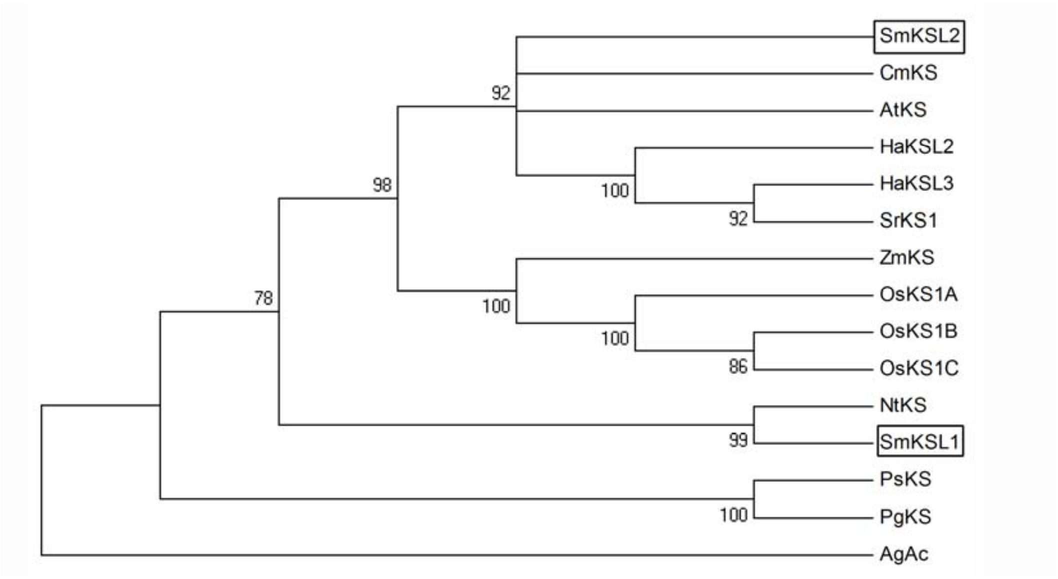
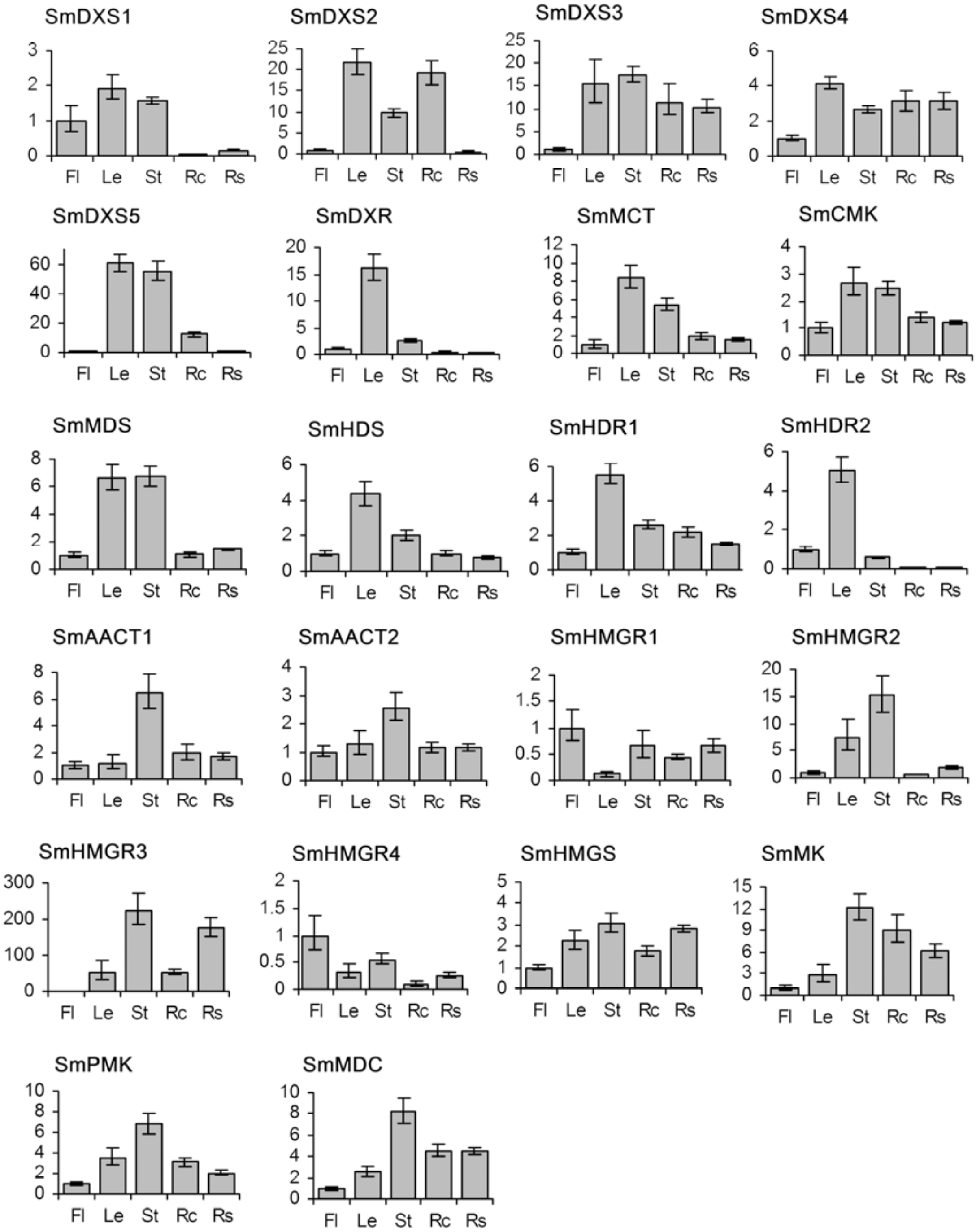


Figure S26. Phylogenetic relationships of KS and KSLs form *S. miltiorrhiza* and various other plants. Proteins included are NtKS (*N. tabacum*, AAS98912), SmKS1 (*S. miltiorrhiza*, ABV08817), SmKSL2 (JN831119), HaKSL2 (*H. annuus*, CBL42916), HaKSL3 (CBL42917), SrKS1 (*S. rebaudiana*, AF097310), CmKS (*C. mollissima*, AEF32083), AtKS (*A. thaliana*, AAC39443), PsKS (*P. sitchensis*, ADB55710), PgKs (*P. glauca*, ADB55708), OsKS1A (*O. sativa*, AAQ72559), OsKS1B (AAQ72560), OsKS1C (AAQ72561.1), and ZmKS (*Z. mays*, AF105149). AgAc (*A. grandis*, AAB05407) was used as an out group. About 200 amino acids in the N-terminal of most KS proteins were trimmed because SmKSL1 lacks those sequences.


SmDXS1
3
2
1
0
SmDXS2
25
20
15
10
5
0
SmDXS3
25
20
15
10
5
0
SmDXS4
6
4
2
0
SmDXS5
60
40
20
0
SmDXR
20
15
10
5
0
SmMCT
12
10
8
6
4
2
0
SmCMK
4
3
2
1
0
SmMDS
8
6
4
2
0
SmHDS
6
4
2
0
SmHDR1
6
4
2
0
SmHDR2
6
4
2
0
SmAACT1
8
6
4
2
0
SmAACT2
4
3
2
1
0
SmHMGR1
2
1.5
1
0.5
0
SmHMGR2
20
15
10
5
0
SmHMGR3
300
200
100
0
SmHMGR4
2
1.5
1
0.5
0
SmHMGS
5
4
3
2
1
0
SmMK
15
12
9
6
3
0
SmPMK
10
8
6
4
2
0
SmMDC
10
8
6
4
2
0
Fl
Le
St
Rc
Rs

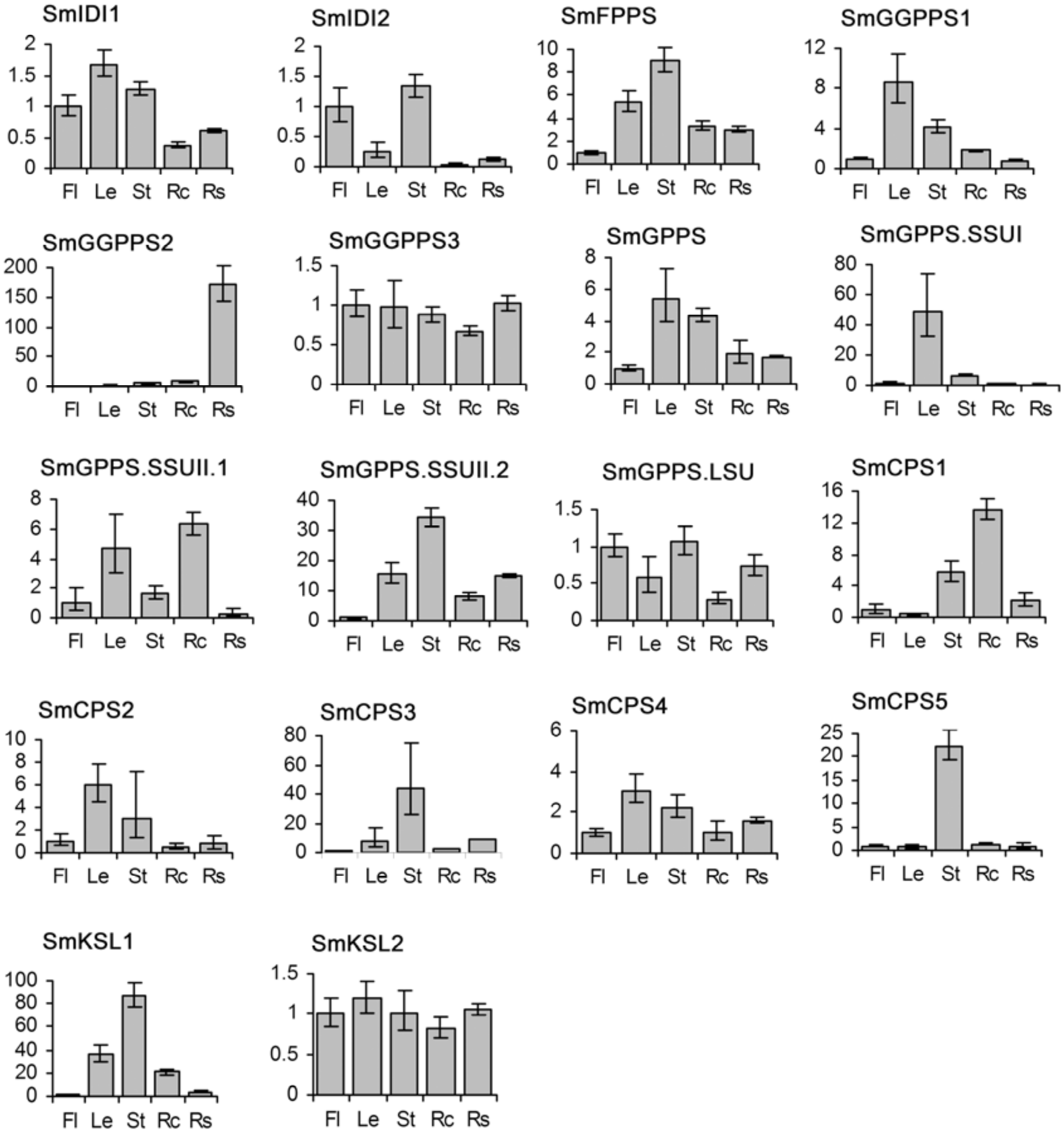
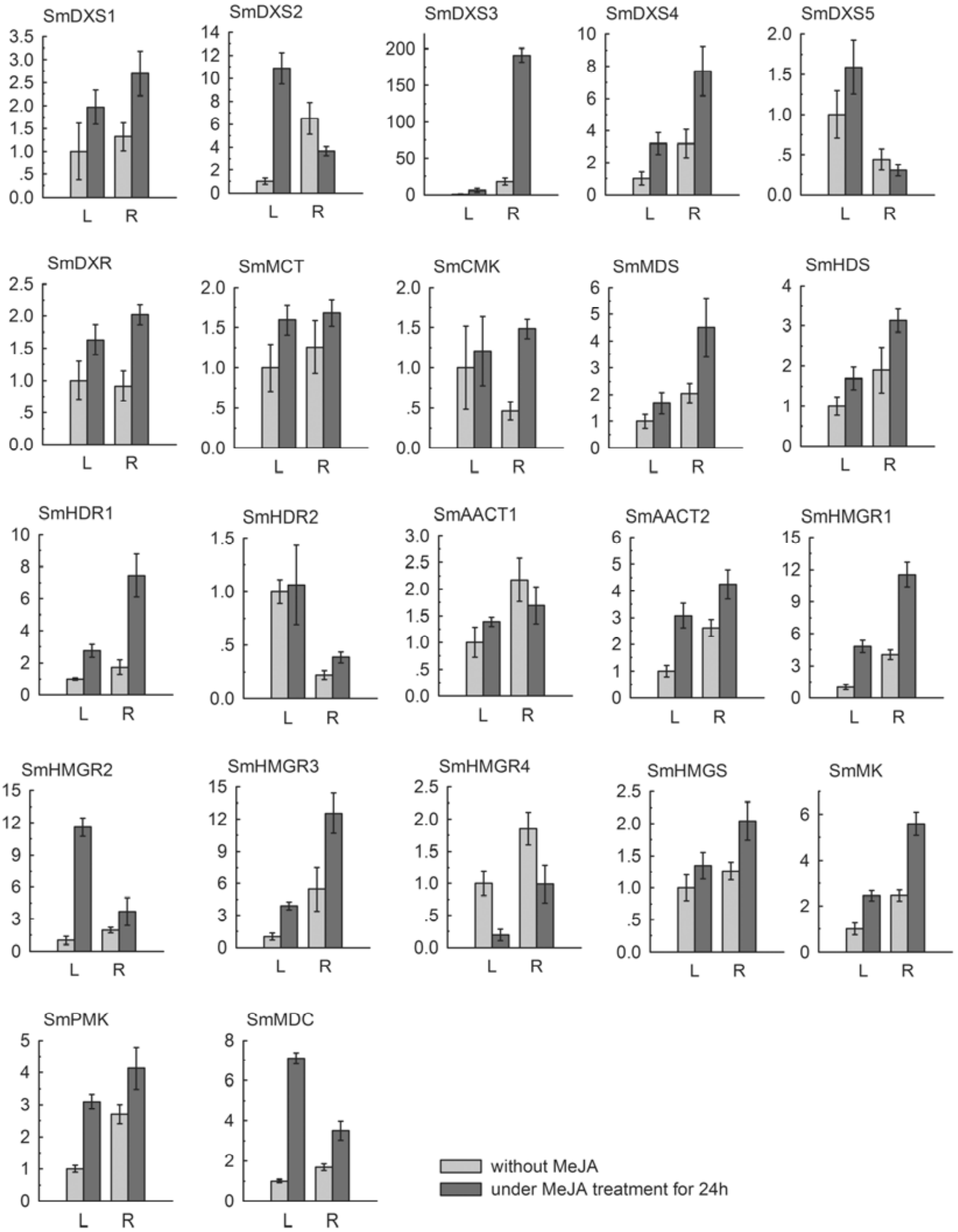


Figure S27. Expression patterns of forty terpenoid biosynthesis-related genes in various tissues of *S. miltiorrhiza* plants. Fold changes of transcript levels in flowers (Fl), leaves (Le), stems (St), root cortices (Rc) and root steles (Rs) of *S. miltiorrhiza* plants grown in soil are shown. Transcript levels in flowers were arbitrarily set to 1.


SmDXS1
SmDXS2
SmDXS3
SmDXS4
SmDXS5
SmDXR
SmMCT
SmCMK
SmMDS
SmHDS
SmHDR1
SmHDR2
SmAACT1
SmAACT2
SmHMGR1
SmHMGR2
SmHMGR3
SmHMGR4
SmHMGS
SmMK
SmPMK
SmMDC
L
R
without MeJA
under MeJA treatment for 24h

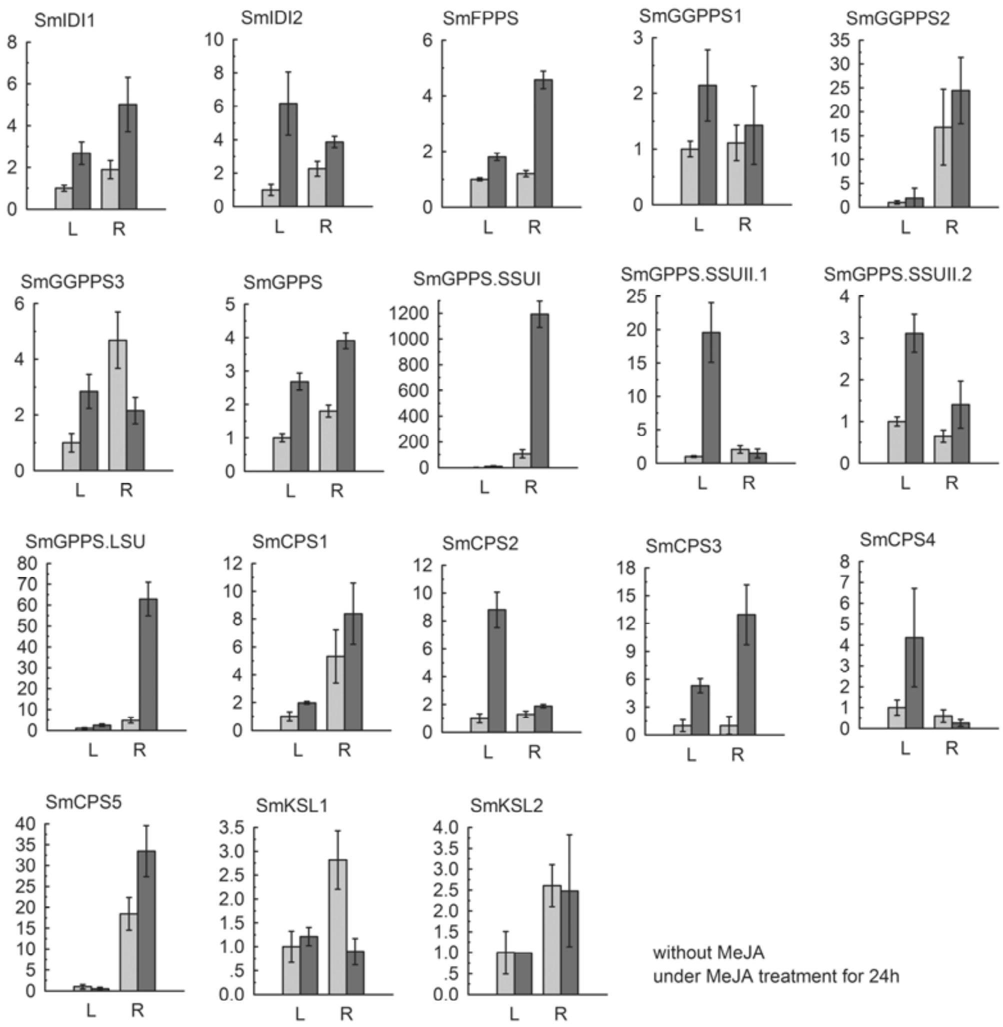


Figure S28. Expression patterns of forty pterpenoid biosynthesis-related genes in *S. miltiorrhiza* plantlets treated with MeJA. Fold changes of transcript levels in leaves (L) and roots (R) of *S. miltiorrhiza* plantlets treated with MeJA for 0 and 24 hours. Transcript levels in leaves of plantlets without treatment were arbitrarily set to 1.

Table S1. Primers used for quantitative real-time RT-PCR.

Gene name	Primer name	Sequence (5' to 3')
SmUBQ	UBQ-F	agatgggcggacacttgctgatta
	UBQ-R	actctccacctccaaagtgatggt
SmDXS1	DXS1-F	catgctcttatacggagcttggcca
	DXS1-R	ggcgtgagtcctgcttccatcagt
SmDXS2	DXS2-F	cagcgcccatatttctcatttcttat
	DXS2-R	gaagactgtcttttcccccaccaat
SmDXS3	DXS3-F	ggctccattgggggatttggttca
	DXS3-R	cactgaagtttgagagcttccatag
SmDXS4	DXS4-F	tctctcgacggccttctcgatgct
	DXS4-R	catcaagagaagagcgtcgcgggt
SmDXS5	DXS5-F	gttggtggctttggatcacacgtt
	DXS5-R	tctctaaagcttctctggatttccaa
SmDXR	DXR-F	tggaggcaccatgaccggagttct
	DXR-R	gggcccactggtcgtagtggatga
SmMCT	MCT-F	acctgaaatttgcattgcctgggaa
	MCT-R	tccttgattgtagccttagcaggaa
SmCMK	CMK-F	ccaagagtggtcgggtgagattgg
	CMK-R	ggggcatgcctcttgtggcttaat
SmMDS	MDS-F	ttggccacggatttgaccttcatc
	MDS-R	tgccctatatctgggagcccaagc
SmHDS	HDS-F	cgtggtgacgaaccccaagaagag
	HDS-R	gccagcaccgataaccaagtcgtc
SmHDR1	HDR1-F	ctcagaattgcgtggaattccatca
	HDR1-R	ttaagccgattgcaatacttctgctt
SmHDR2	HDR2-F	atcgtattgggttgacagcgagaaa
	HDR2-R	tatgccgtttgcaatacttcatcgc
SmAACT1	AACT1-F	aaagctctggagctcgggcttaca
	AACT1-R	tccaaaccagcattcttgagtgcct
SmAACT2	AACT2-F	gcaatggcagttaaacttggattgga
	AACT2-R	gagattgctctcggtatcgcaatg
SmHMGR1	HMGR1-F	caataaggaggctcccggatccaa
	HMGR1-R	acttcatgctcgcaacatccttatt
SmHMGR2	HMGR2-F	ctcaacctgctgggagtcaaggga
	HMGR2-R	ttgagagcttagtaacatctttggaa
SmHMGR3	HMGR3-F	caataaggaggctcccggatccaa
	HMGR3-R	acttcatgctcgcaacatccttatt
SmHMGR4	HMGR4-F	tagcatcacaatcggcttgcctca
	HMGR4-R	ctcgaccggttgtacttcatgtga
SmHMGS	HMGS-F	aggactgcagccttcttgcaccag

	HMGS-R	ttcgcaatggccttcttggcatag
SmMK	MK-F	ttcgtctgctgctctttgtgttgc
	MK-R	agatggcctcccatggatcatctt
SmPMK	PMK-F	tgcatgtcatcgctcaaactgctc
	PMK-R	tcaattggaatccctcgaccagca
SmMDC	MDC-F	agtgcttgccgcagcttgtatggt
	MDC-R	ctggtttcaaccgtgtcacgcatt
SmIDI1	IDI1-F	gcatccaaatccagacgaagtgca
	IDI1-R	tttctcgacgtggtcccaccactt
SmIDI2	IDI2-F	gtgaagtacgtgagccgcgagcaa
	IDI2-R	tttcatgtcagccgcttgggcga
SmGPPS	GPPS-F	atacaaggcggggtattggctcttt
	GPPS-R	tgcatggtctcaccagtgaccaga
SmGGPPS1	GGPPS1-F	ctgcattgttagaggcatctgtagttt
	GGPPS1-R	agcttgggatacgtggtcttgtcg
SmGGPPS2	GGPPS2-F	ccagattgtggacttgtcgagcga
	GGPPS2-R	caacacacctggcgtacttcctcaa
SmGGPPS3	GGPPS3-F	gcgtgcgaggaggagatcgggaat
	GGPPS3-R	cctggacttctccactccgatgagt
SmGPPS.SSUI	GPPS.SSUI-F	tcacgagcatctccctctgaccga
	GPPS.SSUI-R	gctggcccgactgatctctattata
SmGPPS.SSUII.1	GPPS.SSUII.1-F	cgaggccaactacccaaccacacat
	GPPS.SSUII.1-R	tctcctgcacgactcgaaggaggt
SmGPPS.SSUII.2	GPPS.SSUII.2-F	tgacttgccgtgcatggatgatgat
	GPPS.SSUII.2-R	cccactgcccttgctatctcagca
SmGPPS.LSU	GPPS.LSU-F	gttcgccaagtgcatcggcctca
	GPPS.LSU-R	agccttgtgagaatcgaagtgaagg
SmFPPS	FPPS-F	gcgggtgaggacctggagaaacat
	FPPS-R	cagggcctttacaaccagccaagaa
SmCPS1	CPS1-F	ccacatcgccttcagggaagaaat
	CPS1-R	tttatgctcgatttcgctgcgatct
SmCPS2	CPS2-F	ggtctcatcgccttcaacgaagat
	CPS2-R	tccttatcctttatgctcccatcca
SmCPS3	CPS3-F	ggagatgccaattcgaacatcaga
	CPS3-R	tcaaatatagttgcggcggccaaa
SmCPS4	CPS4-F	cggctgccttgggctacaacaata
	CPS4-R	tccctggtgacctcctccttccca
SmCPS5	CPS5-F	tagaagatgcagctactttctctgct
	CPS5-R	catcatcttcaccgccgtactgtt
SmKSL1	KSL1-F	tggaaacagtgtgacccttctgct
	KSL1-R	gcttgcatacaaataacacccaatcct
SmKSL2	KSL2-F	ttagttttggagggcaagaagagtgt
	KSL2-R	ctcctgtttggtcgttgagaagaata

Table S2. Conserved motifs of PTs in various plant species.

Different types of PTs	**The first CXXXC motif**	**FARM motif ($\mathbf{DDX_{2\text{-}4D}}$)**	**The second CXXXC motif**	**SARM motif (DDXXD)**
SmGPPS,SlGPPS,CrGPPS (subunits of homodimeric GPPSs)		DDVLD		DDVLD
SmGPPS.LSU, MpGPPS.LSU, AmGPPS.LSU (large subunits of heteromeric GPPSs)	CIAAC	DDLPCMD		DDILD
SmGPPS.SSUI, MpGPPS.SSUI, AmGPPS.SSUI (small subunits of heteromeric GPPSs, typeI)	CV/LAAC		CGAAC (except for SmGPPS.SSUI)	
SmGPPS.SSUII.1, OsGPPS.SSUII.1, SmGPPS.SSUII.2, AtGPPS.SSUII.2 (small subunits of heteromeric GPPSs, type II)	CV/IAAC	DDLPXXD	CSAV/AC	
SmGGPPS1, SmGGPPS2, SmGGPPS3, AtGGPPS, CsGGPPS	CLAAC	DDLPCMD		DDXXD
SmFPPS, HbFPPS, MpFPPS, VvFPPS		DDIMD		DDYLD