

Enzyme	Gene family size		
	<i>Zea mays</i> GenBank ID	<i>Oryza sativa</i> Gramene ID	<i>Sorghum bicolor</i> Phytozome ID
DXS1	AY946271	LOC_Os05g33840	Sb09g020140
DXS2	AY946270	LOC_Os07g09190	Sb02g005380
DXS3	DR967452	LOC_Os06g05100	Sb10g002960
DXR	AJ297566	LOC_Os01g01710	Sb03g008650
HDS	AY562489	LOC_Os02g39160	Sb04g025290
HDR	DR789385	LOC_OS03g52170	Sb01g009140
IPPI1	AF330034	LOC_Os07g36190	Sb02g035700
IPPI2	DN215017	LOC_Os05g34180	Sb09g020370
IPPI3	CO530856	NA	NA
GGPPS1	EF417573	LOC_Os07g39270	Sb02g037510
GGPPS2	EF417574	NA	NA
GGPPS3	EF417575	LOC_Os01g14630	Sb03g009380
PSY1	AY324431	LOC_Os06g51290	Sb10g031020
PSY2	AY325302	LOC_Os12g43130	Sb08g022310
PSY3	DQ372936	LOC_Os09g38320	Sb02g032370
PDS	L39266	LOC_Os03g08570	Sb06g030030
ZDS	AF047490	LOC_Os07g10490	Sb02g006100
CRTISO1	DR812825	LOC_Os11g36440	Sb05g022240
CRTISO2	AC183901	NA	NA
LCYB	AY206862	LOC_Os02g09750	Sb04g006120
LCYE	EB674039	LOC_Os01g39960	Sb03g026020
ZEP1	DR820114	LOC_Os04g37619	Sb06g018220
ZEP2	CO532283	NA	NA

Table S1. Gene family sizes in three species of the Poaceae. Enzymes highlighted in blue or red are encoded by a maize gene family member showing positive or negative correlation, respectively, between transcripts and total carotenoid content in maize endosperm. Accession numbers are obtained as stated: *Zea mays*, GenBank; *Oryza sativa*, Gramene; *Sorghum bicolor*, Phytozome

Gene	Primer No.	Sequence	Gen Bank Acc #
DXS1	1206	CAGAACAGGGAGGCTCTTG	AY946271
	1207	GCGACACTTTCATCCAAAGA	
DXS2	1212	GCTGAACTACTTCCAGAAGCG	AY946270
	1213	CTGCAGGAACGACGAGTAGA	
DXS3	1194	CTGTGCCAGCAGCATAGTTT	DR967452
	1195	TTCGACCGTCAAGTAGACCA	
DXR	1164	TCCAATGTCACGCTTCTAGC	AJ297566
	1165	TGGCGAGCAACTTCTATGAC	
HDS	1176	CAGTCCAGGGTAGGAAGGAA	AY562489
	1177	TAGGGCTACCGTAGGAGCAA	
HDR	1218	GCCCAACTCGTTACATCCTT	DR789385
	1219	TGAGATCAAGCGTCAGGAAA	
IPPI1	1159A	TGAGCTTATCCAGGAGAATA	AF330034
	1163A	ATGGTGCCTTGTAAGCATT	
IPPI2	1333	TCAGGCTGGTGGTGGACAACCTT	DN215017
	1335	TGATGGCTAAAACAAAGGCA	
IPPI3	1333	TCAGGCTGGTGGTGGACAACCTT	CO530856
	1334	AGCACCAAACAGCAGGACAA	
GGPPS1	1475	CGGAGGAGTTGCTCTCTGAT	EF417573
	1476	AGTCACAACAGGTCCGACAA	
GGPPS2	1477	CAGAGAAGTTGGTCTCTGAC	EF417574
	1478	AGATCGACAGTGAGCAAGAA	
GGPPS3	1426	ATCCATGTGCACAAGACTGC	EF417575
	1427	CACGTCCAGAACGTCGTC	
PSY1	505	CATCTTCAAAGGGGTCGTC	AY324431
	254	CAGGATCTGCCTGTACAACA	
PSY2	503	TCACCCATCTCGACTCTGCTA	AY325302
	676	GATGTGATCTACGGATGGTTCAT	
PSY3	996	GCCAGAGCCTTCTTCAGGCAGG	DQ372936
	997	GTCTTCGGAACGTAGGCCCTC	
PDS	933	GAAATCATCGATGCAACTATGGAA	L39266
	934	CTTCGATAGGTGACCTTTGGA	
ZDS	281	GTGTGGTAAAGATCGGACAA	AF047490
	917	AGAGAGTTGCTCCTTCCAT	
CrtISO1	1325	GGTTGTCAAATGGGAAGGAG	DR812825
	1326	GAATGATGGTGCCTTGACAT	
LCYB	1095	CATCGTAAGGTTCTCGACA	AY206862
	1096	ATGCCGAAGCAGAAGAATC	
LCYE	1037	TTT ACG TGC AAA TGC AGT CAA	EB674039
	1038	TGA CTC TGA AGC TAG AGA AAG	
ZEP1	1664	TTGGTGGCAACAGCTCAAAA	DR820114
	1665	CCTTCACTTGTTGCGATGAG	
ZEP2	1400	TTGGTGGCAATAGCTCAAAG	CO532283
	1401	GCTTCACTTGTTGCGAAGAG	
Actin	1134	CGATTGAGCATGGCATTGTCA	J01238
	1135	CCCCTAGCGTACAACGAA	

Table S2. Primers used in the study

Relative transcript levels of DXS3, DXR, HDR, GGPPS1, PSY1, CrtISO1 in diverse maize lines						
20 DAP						
inbred	DXS3	DXR	HDR	GGPPS1	PSY1	CrtISO1
A619	1.37 (0.12)	7.02 (0.97)	2.89 (0.42)	6.52 (1.26)	2.69 (0.52)	0.86 (0.01)
B73	2.03 (0.21)	11.14 (3.86)	4.89 (0.64)	2.73 (0.83)	1.53 (0.05)	1.88 (0.08)
B37	3.59 (0.44)	7.47 (0.85)	3.17 (0.86)	12.05 (1.05)	3.74 (0.52)	0.7 (0.03)
CI7	1.80 (0.60)	3.32 (0.93)	3.58 (0.37)	6.98 (2.00)	1.71 (0.11)	0.41 (0.08)
C131A	5.14 (0.34)	3.21 (0.35)	2.47 (0.69)	0.81 (0.12)	2.15 (0.49)	2.8 (0.35)
DE3	4.16 (0.19)	8.23 (0.46)	4.26 (0.65)	2.18 (0.48)	3.04 (0.06)	0.42 (0.04)
KUI2007	4.80 (0.14)	20.93 (4.59)	3.39 (0.69)	4.86 (0.43)	3.65 (0.17)	0.21 (0.02)
NC300	2.55 (0.53)	4.26 (0.59)	2.9 (0.94)	2.51 (0.52)	1.01 (0.31)	2.08 (0.54)
SD44	6.43 (0.38)	10.96 (2.53)	4.36 (0.63)	5.15 (0.40)	1.98 (0.35)	1.77 (0.30)
TZI18	4.51 (0.18)	3.56 (0.37)	2.84 (0.43)	4.38 (0.52)	4.47 (0.96)	0.25 (0.05)
25 DAP						
inbred	DXS3	DXR	HDR	GGPPS1	PSY1	CrtISO1
A619	6.56 (0.21)	19.43 (0.95)	4.42 (0.72)	2.69 (0.46)	9.88 (1.76)	1.02 (0.34)
B73	4.79 (0.19)	14.87 (0.88)	6.84 (0.19)	5.76 (0.41)	1.15 (0.25)	3.87 (0.88)
B37	5.43 (0.66)	7.57 (1.33)	13.11 (1.81)	7.55 (1.08)	16.21 (0.29)	1.01 (0.29)
CI7	2.73 (0.21)	4.58 (0.46)	10.62 (1.88)	6.81 (1.40)	1.52 (0.42)	0.93 (0.27)
C131A	4.51 (0.28)	3.49 (0.11)	4.52 (0.14)	1.32 (0.35)	0.74 (0.23)	5.26 (0.06)
DE3	4.80 (0.19)	16.84 (0.88)	5.81 (1.29)	6.86 (0.97)	2.03 (0.74)	1.55 (0.01)
KUI2007	7.80 (0.11)	27.68 (2.66)	10.16 (1.48)	11.48 (0.40)	1.20 (0.24)	3.75 (0.29)
NC300	2.60 (0.50)	5.69 (0.38)	4.7 (0.02)	2.42 (0.40)	0.94 (0.19)	2.17 (0.12)
SD44	2.16 (0.11)	7.30 (0.11)	6.52 (1.18)	3.37 (0.50)	2.37 (0.46)	1.46 (0.12)
TZI18	11.72 (0.48)	24.43 (3.30)	8.11 (1.13)	5.69 (0.10)	3.98 (0.40)	0.59 (0.13)

Table S3. Transcript levels in developing endosperm from diverse maize lines. Transcripts were measured quantitatively from endosperm collected at the indicated DAP. Parentheses show SD, standard deviation. Quantitative RT-PCR was performed using gene specific primers (Table S2) and normalized to actin, as previously described (Li et al., 2008a). Values are expressed as the mean of three RT-PCR replicates +/- standard deviation.

	Relative transcript levels of <i>ZEP1</i> and <i>ZEP2</i> (SD)			
	20 DAP		25 DAP	
	<i>ZEP 1</i>	<i>ZEP2</i>	<i>ZEP1</i>	<i>ZEP2</i>
A619	0.67 (0.10)	1.30 (0.29)	0.32 (0.01)	0.53 (0.07)
B73	0.40 (0.03)	1.45 (0.24)	0.09 (0.03)	0.94 (0.22)
B37	0.51 (0.21)	0.88 (0.22)	0.22 (0.03)	0.27 (0.01)
CI7	0.75 (0.00)	1.12 (0.01)	0.09 (0.02)	0.18 (0.05)
C131A	0.94 (0.37)	1.06 (0.37)	0.46 (0.05)	0.89 (0.09)
DE3	0.40 (0.06)	0.96 (0.17)	0.02 (0.01)	0.62 (0.00)
KUI2007	0.04 (0.02)	0.05 (0.00)	0.14 (0.00)	0.56 (0.04)
NC300	0.66 (0.26)	1.41 (0.21)	0.84 (0.05)	1.11 (0.07)
SD44	0.63 (0.16)	1.15 (0.14)	0.79 (0.19)	0.97 (0.25)
TZI18	0.34 (0.08)	0.41(0.13)	0.07 (0.01)	0.16 (0.02)

Table S4. Transcript levels of *ZEP* genes in developing endosperm from diverse maize lines. Transcripts were measured quantitatively from endosperm collected at the indicated DAP. Parentheses show SD, standard deviation. Quantitative RT-PCR was performed using gene specific primers (Table S2) and normalized to actin, as previously described (Li et al., 2008a). Values are expressed as the mean of three RT-PCR replicates +/- standard deviation.



```

OsDXS3 : MGGRYLHSPVANLVCEESTFVVRSTSCCLHSLCLCRCKTSLNKLPCRTLVFNMPSSRRFNRRVAAALPD---VDD---FFWEKDPPIPLDTIDAPIHLKLNLSKSELKQLADEVRSEIAFTMSRKCQPCN-TGHSVVELLTAIH
SbDXS3 : MDTAFLRPLARNLVYDEFAVLHPTSYPFHTLRLYRCNPMYSRPLLTSPASPSRGLTQRMAALPD---VDD---FFWEKDPPIPLDTIDAPIHLKLNLSKSELKQLADEVRSEIAFTMSRKCQPCG-PGRSVVELLTAIH
ZmDXS3 : MDTAFLRPLARNLVYDEFAVLHPTSYPFHTLRLYRCNPMYSRPLLTSPASPSRGLTQRMAALPD---VDD---FFWEKDPPIPLDTIDAPIHLKLNLSKSELKQLADEVRSEIAFTMSRKCQPCG-ADRSVVELLTAIH
OsDXS2 : MALQASS-PSMFRAPITNTNASCRRKLVQVRAASAAAAAN----GGDGVMMRK-----EAS---GAWKIDYS-GEKBPPLLDVTVYVPHMKNLSTPELEQLAAELRAEIVVHTVSKTGGHLS-SSLGVELLVALH
SbDXS2 : MALQASS-PSAFRAVATANASCRRQFQVRAQVAGSSSSSVGADGAGKMMFKAEPAAAAATSSGSSGPKIDFTSGBKBPATPLLDVTVYVPLMKNLSTSELEQLAAELRAEIVVHTVSKTGGHLS-SSLGVELLVALH
ZmDXS2 : MALQASSPSAFRAVATANASCRRQFQVRAQVAGSSSSSISGADGG-KMMVSK-EPAAAAATSS---GWKIDFS-GEKBPPLLDVTVYVPLMKNLSTSELEQLAAELRAEIVVHTVSKTGGHLS-SSLGVELLVALH
OsDXS1 : -----MALTTFSISRGGFVGLPQ-EGHFAPAABELSHLKLQS-RPHKAR-RSSSSISASLS---TEREAAYHSQREPPPLLDVTVYVPHMKNLSELKQLADELRSDVIFHVSKTGGHLG-SSLGVELLVALH
SbDXS1 : -----MALWTFESVPRG-FLAVFAHDSHFASAVEQLHGHLKQARPLRPRRHRPACVSASLS---PERE-AEYYSQREPPPLLDVTVYVPHMKNLSELKQLADELRSDVIFHVSKTGGHLG-SSLGVELLVALH
ZmDXS1 : -----MALSTFSVPRG-FLGVPAQ-DSHFASAVE-LHVNKLQARPINLKPRRRACVSASLS---SERE-AEYYSQREPPPLLDVTVYVPHMKNLSELKQLADELRSDVIFHVSKTGGHLG-SSLGVELLVALH

OsDXS3 : YVFNAPMDKILWDACQHTYAHKILTGRRSLFHTIKRKNLSGFTSRSESEYDPPFCAGHGCNLSLACLGMAVARDLGRKRNRIIVTIVSNWTTMAGQVYEAAMGHAGFLDSNMVVLNDSRHHTLTPKADSQ-SKMSINALSSA
SbDXS3 : YVFNAPMDKILWDACQHAYAHKILTGRRSQFHTIKRKNLSGFTSRSESEYDPPFCAGHGCNLSLACLGMAVARDINGRKNRIIVTIVSNWTTMAGQVYEAAMGHAGFLDSNMVVLNDSCHTLLPKADGR-EKMSVNAEASSA
ZmDXS3 : YVFNAPMDKILWDACQHAYAHKILTGRRSLFHTIKRKNLSGFTSRSESEYDPPFCAGHGCNLSLACLGMAVARDINGRKNRIIVTIVSNWTTMAGQVYEAAMGHAGFLDSNMVVLNDSCHTLLPKADGG-AKMSVNAEASSA
OsDXS2 : HVFTPEBDKI IWDVGHQAYPHKILTGRSRMHTI RQTSGLAGPKRDESAHDAFGVGHSSSTISAAALGMVARDLLGRKNHVTIVSICDGAMTAGQAYEAMNNSGYLDSNMIVVLNDNKQVSLPTATLDGSPVPVCAALSSA
SbDXS2 : HVFTPEBDKI IWDVGHQAYPHKILTGRSRMHTI RQTSGLAGPKRDESAHDAFGVGHSSSTISAAALGMVARDLLGRKNHVTIVSICDGAMTAGQAYEAMNNSGYLDSNMIVVLNDNKQVSLPTATLDGSPVPVCAALSSA
ZmDXS2 : HVFTPEBDKI IWDVGHQAYPHKILTGRSRMHTI RQTSGLAGPKRDESAHDAFGVGHSSSTISAAALGMVARDLLGRKNHVTIVSICDGAMTAGQAYEAMNNSGYLDSNMIVVLNDNKQVSLPTATLDGSPVPVCAALSSA
OsDXS1 : YVFNAPQDRILWDVGHQSYPHKILTGRDKMPTMRCQTNGLAGTKRAESEYDSFGTGHSSSTISAAALGMVARDLKGKNNVAVICDGAMTAGQAYEAMNAGYLDSDMIVLNDNKQVSLPTATLDGSPVPVCAALSSA
SbDXS1 : YVFNAPQDRILWDVGHQSYPHKILTGRDKMPTMRCQTNGLAGTKRAESEYDSFGTGHSSSTISAAALGMVARDLKGKNNVAVICDGAMTAGQAYEAMNAGYLDSDMIVLNDNKQVSLPTATLDGSPVPVCAALSSA
ZmDXS1 : YVFNAPQDRILWDVGHQSYPHKILTGRDKMPTMRCQTNGLAGTKRAESEYDSFGTGHSSSTISAAALGMVARDLKGKNNVAVICDGAMTAGQAYEAMNAGYLDSDMIVLNDNKQVSLPTATLDGSPVPVCAALSSA

OsDXS3 : LSKVQSSKGRFRFREAAGLAKWFCCKGMHEFAAKIDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
SbDXS3 : LSKIQSSKGRFRFREAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITGTESDTDNIRSEITASEECPNSNSNDLLKPLETG-LRSTYNDCEVEA
ZmDXS3 : LSKIQSSKGRFRFREAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITGTENDTCGIGSEINAVEECPNSNSPDLKFLGTG-LRSTYNDCEVEA
OsDXS2 : LTKLQSSTKLRRRREAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
SbDXS2 : LTKLQSSTKLRRRREAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
ZmDXS2 : LTKLQSSTKLRRRREAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
OsDXS1 : LSKLQSSRPLRELEVAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
SbDXS1 : LSKLQSSRPLRELEVAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES
ZmDXS1 : LSKLQSSRPLRELEVAAGLAKWFCCKGMHEFAAKVDEYARGMIGPHCATLFEELGLYYIGPIDGNNDLLICVWKEVSTLSDTGPVLVHVITNEKDSGGEFNSEITPDEECPDSSQDILKFLENG-LRSTYNDCEVES

OsDXS3 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
SbDXS3 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
ZmDXS3 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
OsDXS2 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
SbDXS2 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
ZmDXS2 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
OsDXS1 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
SbDXS1 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID
ZmDXS1 : LIAEAEQDSKIVAIHAAMCGGTGLNYFLRRFPPRCFDVGLAEQHAVFAAGLACEGLKPPCAIMSSFLQRGYDQVVDVLDLQKLPVRFAMDRAGLVGADGPTHCCAFDVAAYMACLPNMVMAPADEABELCHMVATAAAID

OsDXS3 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
SbDXS3 : DRPFCFRYPKRGAVGTSG-TLYYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
ZmDXS3 : DRPFCFRYPKRGAVGTSG-SVTYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
OsDXS2 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
SbDXS2 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
ZmDXS2 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
OsDXS1 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
SbDXS1 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP
ZmDXS1 : DRPFCFRYPKRGAVGTSG-TLAYGNPLEIGKGEILVBC-KEIAFLGYGVVQRCILARSILSNFGIQATVANARFCKPLDIDLIRLTCQCHSFLITVEEGTVGGFGSHVSOFTSLDGLLDGKIKWRPVLVLDPRYIDHGASP

OsDXS3 : TEQIDMAGLTAHHIAATALTLLCRHRDAILLMK---- (722)
SbDXS3 : AEQIDMAGLTAHHIAATALTLLCRHRDAILLMK---- (722)
ZmDXS3 : AEQIDMAGLTAHHIAATALTLLCRHRDAILLMK---- (722)
OsDXS2 : VDQIEEAGLTPRHIAATVLSLLCRPLEAOLK---- (713)
SbDXS2 : QDQIEEAGLTPRHIAATVLSLLCRPLEAOLK---- (727)
ZmDXS2 : QDQIEEAGLTPRHIAATVLSLLCRPLEAOLK---- (723)
OsDXS1 : ADQIAEAGLTPSHIAATVFNVLQAREALAMTVPNA (720)
SbDXS1 : ADQIAEAGLTPSHIAATVFNVLQAREALAMTVPNA (722)
ZmDXS1 : ADQIAEAGLTPSHIAATVFNVLQAREALAMTVPNA (719)

```

Figure S1 Multiple-alignment of DXS proteins. Amino acid sequences across available grass species (*Zea mays*, *Zm*; *Oryza sativa*, *Os*; *Sorghum bicolor*, *Sb*) were used to show the sequence signatures of transketolase and DXS enzymes (Rodriguez-Concepcion et al., 2003). These include thiamine di-phosphate binding domain (underlined red) and two transketolase motifs (underlined blue), with HIS residue for activity (blue triangle). Total number of aminoacids were marked at the end of each sequence in blue. The variant N-terminal amino acid sequence includes the transit peptide region. For sequence accessions, see Table S1.

```

ZmGGPPS1 : ...MAMSFHFHPLAASRVHFSPLPFAAAAAGAPSPSSVAIAAHQHGRRRFSATVATAAASAAATTFDFKAYMGERAV
ZmGGPPS2 : .....MAFHFHPLIVASRVHFSPLPFAAATGAPSPSSVAIAAHYHHGRRRFSATVATAT.....ITTFNFKAYMERAV
ZmGGPPS3 : MNKLASCFLOHGAPHTQIFKSYHVQRSPSLQLENRSVSMTRHRAADRAARCTIVDVAVDSG...TSFDFESYLSAKAR

ZmGGPPS1 : AVNRALDAAVPAGPEPPAALHDMRYALLAGGKVRPALCLAACAVVGGPEPWAMPAAAAVEMVHTMSLVHDDLPCMDDD
ZmGGPPS2 : AVNRALDAAI PAGGLPAALHDMRYALLAGGKVRPALCLAACAVVGGPEPWAMPAAAAVEMVHTMSLVHDDLPCMDDD
ZmGGPPS3 : AVHDALDLTLQGLRCEVLSSESMRYSVLAGGKRLRPLAIAACELVGGTAAAVPVACAVEMIHTASLIHDDMPCMDDDA
                                     (I)

ZmGGPPS1 : LRRGKPTCHVYVGEPTAVLAGDALLSLSFHHMASVGSYPPDVDEKHPARVVRAIGELARCIGSEGLVAGQVVDLEMTGT
ZmGGPPS2 : LRRGKPTCHVYVGEPTAVLAGDALLSLSFHHMASVGSYPPAVDPEKHPARVVRAIGELARCIGSEGLVAGQVVDLEMTGT
ZmGGPPS3 : LRRGRPSNHVAFGEPTALLAGDALLALAFEHVARG.....SAGAGVPADRALRAVVELGSVAGVGGIAGOVADMASEGA
                                     (II)                                     (III)                                     (IV)

ZmGGPPS1 : .SETVPLEERLEYIHLHKTAALLEASVVI GAIIGGGTDEQIEERLKYARSIGLLFQVVDLDVTKSSEELGKTAGKDLAS
ZmGGPPS2 : .SOTVPLELLEYIHLHKTAALLEASVVI GAIIGGGTDEQIEERLKYARSIGLLFQVVDLDVTKSSEELGKTAGKDLAS
ZmGGPPS3 : PSGSVSLGLEYIHVHKTARLVEAAAVSGAVVGGGDEGEVEVRRYAHFLGLLGQVVDLDVDTGTSEQLGKTAGKDVA
                                     (V)                                     (VI)

ZmGGPPS1 : DKITYPKLLGLLEKSRFAEELSDAVTEQLACFDKEKAAPLHLANYIAHRQN
ZmGGPPS2 : DKITYPKLLGLAKSRFAEKLVS DATTEQLACFDEEKAAPLHLANYIAHRQN
ZmGGPPS3 : GKATYPRLMGLKGARAYMGELAKAEALDGLDAARTAPLRHLARFMAHROH
                                     (VII)

```

Figure S2. GGPPS proteins of maize. Blue lines indicate seven conserved domains among GGPPS gene family proteins that are highly conserved among prenyl diphosphate synthases with two Asp-rich motifs for substrate binding (domains II and VI) (Okada et al., 2000). Accession numbers: ZmGGPPS1, EF417573; ZmGGPPS2, EF417574; ZmGGPPS3, EF417575.

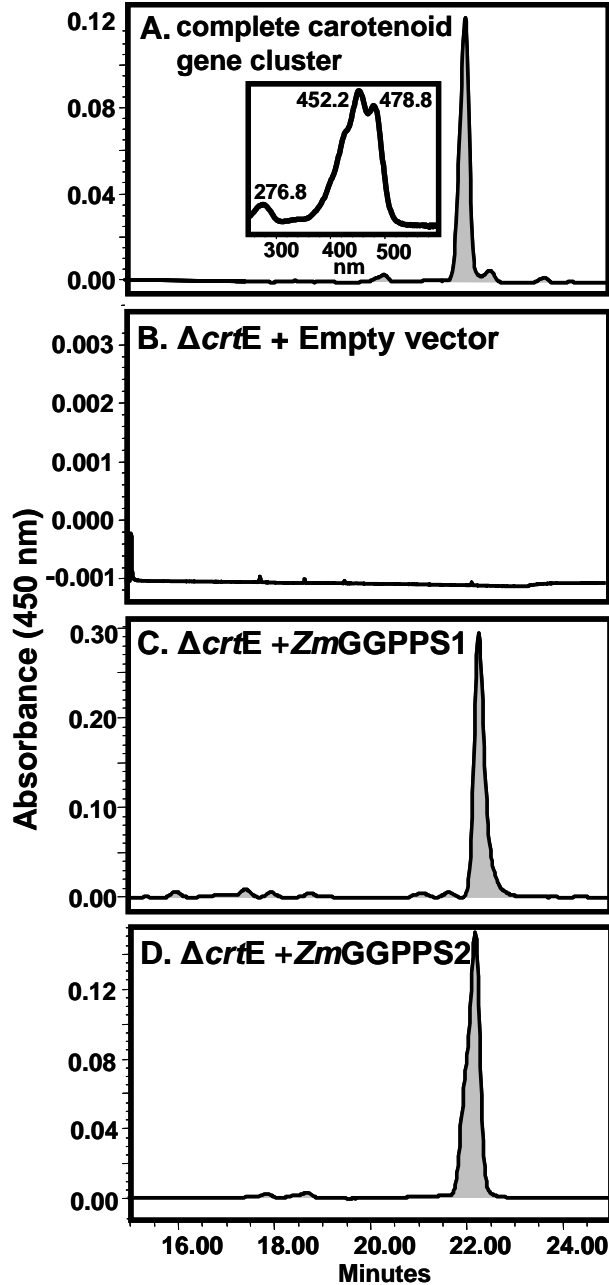


Figure S3. Functional complementation of GGPPS1 and GGPPS2 from maize. *E. coli* cells were transformed with A, pACCAR25 (for zeaxanthin accumulation); B, pACCAR25 $\Delta crtE$ (missing bacterial GGPPS)+ pET23a (empty vector); C, pACCAR25 $\Delta crtE$ + maize GGPPS1; D, pACCAR25 $\Delta crtE$ + maize GGPPS2. Chromatograms show HPLC separation of extracted pigments; inset in A shows spectral fine structure for the pathway end product, zeaxanthin diglucoside.

```

AtCrtISO : MDLCFQNPVKCGDR[F]SALNTSTYYKLGTSNLGFNGFVLENR[KKKKKLP]MVTVK[SVSSV]WASTVQGTKRD[GG]---[ESL]
OsCrtISO : ----MPL[LLLLAAARPCAPL]LAPSAAGAVVGR[ESTARPLGR]TRRGAC[AAAAAAVAAE]KTVVKAEE[E]EEGG-----[GE]
SbCrtISO : -----MPPLAAR[VHAPILLAP]---VAP-HEATRLCASS[CFRGGVGGFRRGALASEKPPQAVAVA]KAGGEV[GGGGEGGP]
ZmCrtISO1 : -----MPPLAAR[VHAPILLVP]---AASPR[FEATRLCAAS]CFRG-VGGFRRGALASEKPP[VVAVA]KAGGE[GGGGQGGP]
ZmCrtISO2 : -----MPPLAAR[VYAPLN]LAPAVPAASL[FEATRLCAAS]CFRN-VGGFRRGAL[A*]EKAAAVAVA[KAGGE]GGGG[DEGP]

AtCrtISO : YDAIVIGSGIGGLVAATQLA[VKEARVLVLEKYLIPGGSSGFY]ERDGYTFDVGSSVMFGFSDKGNLNLITQAL[KAVGRKME]
OsCrtISO : YDAIVVSGIGGMVAATQLA[AKGARVLVLEKYVIPGSSGYRRDGF]TFDVGSSVMFGFSDKGNLNLITQAL[EAVGHKME]
SbCrtISO : YDAIVIGSGIGGLVAATQLA[AKGARVLVLEKYLIPGGSSGYRRDGF]TFDVGSSVMFGFSDKGNLNLITQAL[EAVGRKME]
ZmCrtISO1 : YDAIVIGSGIGGLVAATQLA[AKGARVLVLEKYLIPGGSSGYRRDGF]TFDVGSSVMFGFSDKGNLNLITQAL[EAVGRKME]
ZmCrtISO2 : YDAIVIGSGIGGLVAATQLA[AKGARVLVLEKYLIPGGSSGYRRDGF]TFDVGSSVMFGFSDKGNLNLITQAL[EAVGRKMR]

AtCrtISO : VIPDPSTVHFHLP[NNLSVRIHREYD]DFIAEL[ITSKFPHEKEGILKFY]GDCW[KIFNSLNSLELKSLEEP]IYLF[Q]OFFOKP[LE]
OsCrtISO : VIPDPSTVHFHLP[GDLSVLVHREYD]DFVTE[LVNKKFPHEKEGILKFY]GTCW[KIFNSLNSLELKSLEEP]LYLF[Q]OFFOKP[LE]
SbCrtISO : VLPDPSTVHFHLP[GDLSVLVHRKYE]DFINEL[ISKFPHEKEGILKFY]GTCW[KIFNSLNSLELKSLEEP]LYLF[Q]OFFOKP[LE]
ZmCrtISO1 : VLPDPSTVHFHLP[GDLSVLVHRKYE]DFINEL[ISKFPHEKEGILKFY]GTCW[KIFNSLNSLELKSLEEP]LYLF[Q]OFFOKP[LE]
ZmCrtISO2 : VLPDPSTVHFHLP[GDLSVLVHRKYE]DFINEL[ISKFPHEKEGILKFY]GTCW[KIFNSLNSLELKSLEEP]LYLF[Q]OFFOKP[RE]

AtCrtISO : CLTLAYYLPQNAG[AIARKYIKD]Q[ELLSFIDAECFIVSTVNALQTPMINAS]MVLCDRH[YGGINYPVGGVGGI]AKSLA[E]GLV
OsCrtISO : CLTLAYYLPQNAG[DIARKFIK]D[QELLSFIDAECFIVSTVNALQTPMINAS]MVLCDRH[FGGINYPVGGVGGI]AMSLA[DG]LV
SbCrtISO : CLTLAYYLPQNAG[DIARKFIK]D[QELLSFIDAECFIVSTVNALQTPMINAS]MVLCDRH[FGGINYPVGGVGGI]IASLA[DG]LV
ZmCrtISO1 : CLTLAYYLPQNAG[DIARKFIK]D[QELLSFIDAECFIVSTVNALQTPMINAS]MVLCDRH[FGGINYPVGGVGGI]IALSLA[DG]LV
ZmCrtISO2 : CLTLAYYLPQNAG[DIARKFIK]D[QELLSFIDAECFIVSTVNALQTPMINAS]MVLCDRH[FGGINYPVGGVGGI]IASLA[DG]LV

AtCrtISO : DQ[GSEI]QYKANVKS[ITLDHGKAVGVRL]ADGREFFAKT[II]SNATRWDTFGKLLK[E]KLPKEE[ENFQKV]YVKAPSFLSIHMG
OsCrtISO : D[KGSEI]RYKANVTNVILENGKAVGVRLSNGKEFFAKT[VISNATRWDTFGKLLKVE]ELPEE[KNFQKN]YVKAPSFLSIHMG
SbCrtISO : E[KGSEI]RYKANVTNVILENGKAVGVRLSNGKEFFAKT[VISNATRWDTFGKLLKE]ELPEE[KNFQKN]YVKAPSFLSIHMG
ZmCrtISO1 : E[KGSEI]RYKANVTNVILENGKAVGVRLSNGKEFFAKT[VISNATRWDTFGKLLKE]ELPEE[KNFQKN]YVKAPSFLSIHMG
ZmCrtISO2 : E[KGSEI]RYKANVTNVILENGKAVGVRLSNGKEFFAKT[VISNATRWDTFGKLVK-]ELPEE[KNFQKN]YVKAPSFLSIHMG

AtCrtISO : VKAEVLEP[PD]DCHHFVLEDDW[KNLE]E[YGSI]FLSIPTVLD[SLAPD]GRHILHIFTT[SSIE]DWEGL[PP]KEYE[A]KKE[D]VAAR
OsCrtISO : VKASVLPADT[D]CHHFVLEDDW[ANLE]K[YGSI]FLSIPTVLD[SLAPR]GHILHIFTTAGIE[DWEGL]SRKDYE[K]KKE[V]VATE
SbCrtISO : VKASVLPAGT[D]CHHFVLEDNWN[KNLE]K[YGSI]FLSIPTVLD[SLAPR]GHILHVFTTAGIE[DWEGL]SRKEYE[E]KKE[V]VANE
ZmCrtISO1 : VKASVLPAGT[D]CHHFVLEDNWN[KNLE]K[YGSI]FLSIPTVLD[SLAPR]GHILHIFTTAGIE[DWEGL]SRKEYE[E]KKE[V]VANE
ZmCrtISO2 : VKASVLPAGT[D]CHHFVLEDNWN[KNLE]K[YGSI]FLSIPTVLD[SLAPR]GHILHIFTTAGIE[DWEGL]SRKEYE[E]KKE[V]VANE

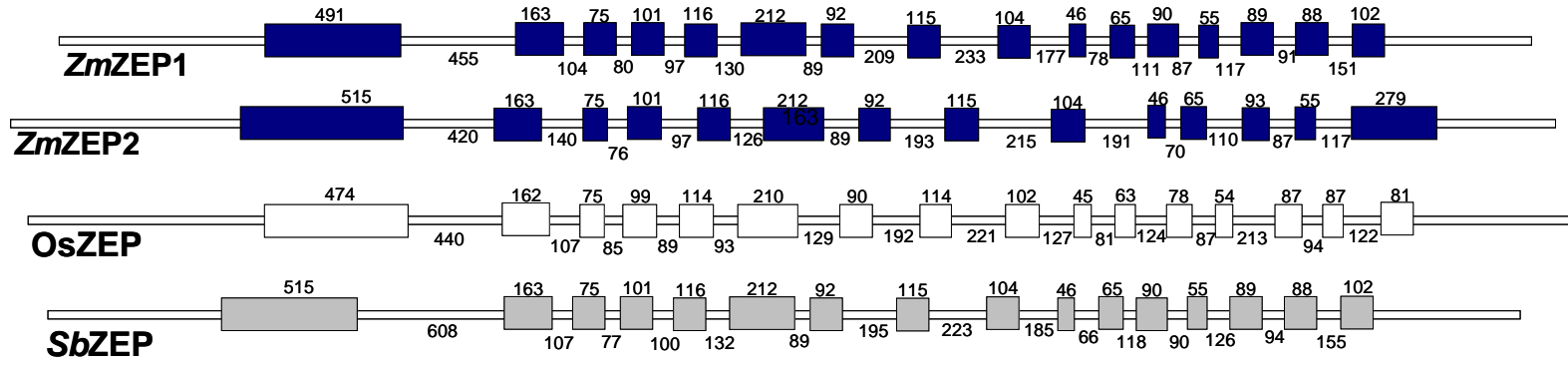
AtCrtISO : IIR[ORLEK]KLFPG[LS]SITF[KEVGT]PRTHRR[FLARN]DGTYGMP[PRG]TPKGLL[MPFN]T[AT]DGLYCVGDSCFP[QGQ]VIAVA
OsCrtISO : IIR[ORLEK]KLFPG[LD]SIVL[KEVGS]PKTHRR[FLARN]DGTYGMP[PRG]KPKGLLAMPFN[TT]SIDGLYCVGDSCFP[QGQ]VIAVA
SbCrtISO : IIR[ORLEK]KLFPG[LD]SIVL[KEVGS]PKTHRR[FLARN]DGTYGMP[PRG]KPKGLLAMPFN[TT]SIDGLYCVGDSCFP[QGQ]VIAVA
ZmCrtISO1 : IIR[ORLEK]KLFPG[LD]SIVL[KEVGS]PKTHRR[FLARN]DGTYGMP[PRG]KPKGLLAMPFN[TT]SIDGLYCVGDSCFP[QGQ]VIAVA
ZmCrtISO2 : IIR[ORLEK]KLFPG[LD]SIVL[KEVGS]PKTHRR[FLARN]DGTYGMP[PRG]KPKGLLAMPFN[TT]SIDGLYCVGDSCFP[QGQ]VIAVA

AtCrtISO : FSGVMCAHRVAADIGLEKKS[RVLDV]GLLGLL[GW]LRTLA (595)
OsCrtISO : FSGIMCAHRVAADIGLEQRS[PVLD]AGLLGLL[RW]LRTLA (586)
SbCrtISO : FSGIMCAHRVAADIGLEOKS[PVLD]AGLLGLL[RW]LRTLA (588)
ZmCrtISO1 : FSGIMCAHRVAADIGLEOKS[PALD]AGLLGLL[RW]LRTLA (587)
ZmCrtISO2 : FSGIMCAHRVAADIGLEOKS[PVLD]AGLLGLL[RW]LRTLA (588)

```

Figure S4. Multiple-alignment of CrtISO protein sequences. Amino acid sequences across available grass species (*Zea mays*, *Zm*; *Oryza sativa*, *Os*; *Sorghum bicolor*, *Sb*) and *Arabidopsis thaliana* (*At*) were used. Total number of amino acids were marked at the end of each sequence in blue. Accession numbers: **AtCrtISO**, At1g06820; **OsCrtISO**, LOC_Os11g36440; **SbCrtISO**, Sb05g022240; **ZmCrtISO1**, AC218991; **ZmCrtISO2**, AC183901.

A



B

```

OsZEP : -----MAPLPSSS--TRATSLVAIPGASCPHGQAQRLLAFR---PPRRGSQWRGLCVSRGRHATVAMAMPAAQAAG--RRARVLVAGGGIGGLVFALAARRKGFVVLVLERDMSAVRGEGRYRGPQLQSN
ZmZEP1 : MPTVLSTRRAASEETISLPSRS--SHSRITASPLHEVSRGN--RALRLLAALRSPSAPRPLRLRRPGLPPARAGLAAAAMPAPET----KARLLVAGGGIGGLVFALAARRKGFVVLVLERDMSAVRGEGRYRGPQLQSN
SbZEP : MPATLSTRAP--PRATSLPSSR---ASQATLLRHVSRRN--RALRLLALP--SPSAPRRRSPGLLPPARAGLVATAAMPAPPEP---KARVLVAGGGIGGLVFALAARRKGFVVLVLERDMSAVRGEGRYRGPQLQSN
ZmZEP2 : ---MLSTTWAS--EKATPLPSSP---ASQATRLRHVSSN--RTLRLVLP--SPSAPRRR--RPGLPARAGLVATAAMPAPPEP---KARVLVAGGGIGGLVFALAARRKGFVVLVLERDMSAVRGEGRYRGPQLQSN

OsZEP : ALAVLEAVDAGAADQVMDAGCITGNRVNGIVDGVSGSWYIKFDFTFPAERGLPVTRVISRMTLQQILARAVGDDAIIINGSHVVDFIDDGSKVTAILEDGRKFEGLLVGADGIWSKVRKTLFGHSDATYSGYTCYTGIA
ZmZEP1 : ALAALEAVDAAADEVMDAGCVTGDVRVNGIVDGVSGSWYIKFDFTFPAERGLPVTRVISRMTLQQILARAVGDDAIIINGSHVVDFIDDGSKVTAILEDGRKFEGLLVGADGIWSKVRKTLFGHSDATYSGYTCYTGIA
SbZEP : ALAALEAVDAAADEIMDAGCVTGDVRVNGIVDGVSGSWYIKFDFTFPAERGLPVTRVISRMTLQQILARAVGDDAIIINGSHVVDFIDDGSKVTAILEDGRKFEGLLVGADGIWSKVRKTLFGHSDATYSGYTCYTGIA
ZmZEP2 : ALAALEAVDAAADEIMDSGCVTGDVRVNGIVDGVSGSWYIKFDFTFPAERGLPVTRVISRMTLQQILARAVGDDAIIINGSHVVDFIDDGSKVTAILEDGRKFEGLLVGADGIWSKVRKTLFGHSDATYSGYTCYTGIA

OsZEP : DFVPPDIDTVGYRVFLGHKQYFVSSDVGAGKMQWYAFHKEPAGGIDPENGKKRLLLEIFDNGWCDNVVDLINATDEEAILRRDIYDRPPIINWGKGRVTLTGDSVHAMQPNLQGGCMAIEDGYQLAVELEKSWQESAKSG
ZmZEP1 : DFVPPDIDTVGYRVFLGHKQYFVSSDVGAGKMQWYAFHKEPAGGIDPENGKKRLLLEIFDNGWCDNVVDLINATDEEAILRRDIYDRPPIINWGKGRVTLTGDSVHAMQPNLQGGCMAIEDGYQLAVELENAWQESVKTG
SbZEP : DFVPPDIDTVGYRVFLGHKQYFVSSDVGAGKMQWYAFHNEPAGGIDPENGKKRLLLEIFDNGWCDNVVDLINATDEEAVLRRDIYDRPPIINWGKGRVTLTGDSVHAMQPNLQGGCMAIEDGYQLAVELENAWQESVKTG
ZmZEP2 : DFVPPDIDTVGYRVFLGHKQYFVSSDVGAGKMQWYAFHNEPAGGIDPENGKKRLLLEIFDNGWCDNVVDLINATDEEAVLRRDIYDRPPIINWNRKGRVTLTGDSVHAMQPNLQGGCMAIEDGYQLAVELENAWQESVKTG

OsZEP : TPMDIVSSLRRYEKERRLRVVIHGLARMAAIMATTYREYLVGLGPLSFLTKLRIPHPGRVGGRRFFIKYGMPLMISWVLGGNSKLEGRPLSCLSDKANDQLRRWFEDDDALEQAMGGEWYLFATSEANNNSLQOPIR
ZmZEP1 : TPMDIVSSLRRYEKERRVRVVIHGLARMAAIMATTYREYLVGLGPLSFLTKLRIPHPGRVGGRRFFIKYGMPLMISWVLGGNSKLEGRPLSCLSDKANDQLRRWFEDDDALEQAMGGEWYLFATSEANNNSLQOPIR
SbZEP : TPMDIVSSLRRYEKERRLRVVIHGLARMAAIMATTYREYLVGLGPLSFLTKLRIPHPGRVGGRRFFIKYGMPLMISWVLGGNSKLEGRPLSCLSDKANDQLRRWFEDDDALEQAMGGEWYLFATSEANNNSLQOPIR
ZmZEP2 : TPMDIVSSLRRYEKERRLRVVIHGLARMAAIMATTYREYLVGLGPLSFLTKLRIPHPGRVGGRRFFIKYGMPLMISWVLGGNSKLEGRPLSCLSDKANDQLRRWFEDDDALEQAMGGEWYLFATSEANNNSLQOPIR

OsZEP : LLRDEQRSLSVGSRSDANDSASSLSLPEPQISERHATITCKNKAFYLDLHGEHGTWITDNEGRYRVPNPFVRFHPSDVEIFGSDKKAMFRVKVLTLPYESARSCKQOGLQOQEVLQAA
ZmZEP1 : LLRDEQRSLSVGSRSDANDSASSLSLPEPQISERHATITCKNKAFYLDLHGEHGTWITDNEGRYRVPNPFVRFHPSDVEIFGSDKKAMFRVKVLTLPYESARSCKQOGLQOQEVLQAA
SbZEP : LLRDEQRSLSVGSRSDANDSASSLSLPEPQISERHATITCKNKAFYLDLHGEHGTWITDNEGRYRVPNPFVRFHPSDVEIFGSDKKAMFRVKVLTLPYESARSCKQOGLQOQEVLQAA
ZmZEP2 : LLRDEQRSLSVGSRSDANDSASSLSLPEPQISERHATITCKNKAFYLDLHGEHGTWITDNEGRYRVPNPFVRFHPSDVEIFGSDKKAMFRVKVLTLPYESARSCKQOGLQOQEVLQAA

```

Figure S5. ZEP gene family structure. (A) Gene organizational structure showing paralogs and orthologs of the ZEP gene family in *Zea mays* (*ZmZEP1*, AZM5_13315; *ZmZEP2*, AZM5_13312); *Oryza sativa* (*OsZEP*, LOC_Os04g37619) and *Sorghum bicolor* (*SbZEP*, Sb06g018220) species representing two subfamilies of the Poaceae (grasses). Boxes and lines indicate exons and introns respectively, sizes for which are in base pairs (B) Multiple-alignment of ZEP paralogs in maize and orthologs from *Oryza sativa* (*Os*) and *Sorghum bicolor* (*Sb*).